UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

LEANDRO MARCIO MOREIRA

Análise estrutural e funcional do genoma de
Xanthomonas axonopodis* pv. *citri

São Paulo
Data do Depósito na SPG:
12/09/2006

LEANDRO MARCIO MOREIRA

**Análise estrutural e funcional do genoma de
Xanthomonas axonopodis pv. *citri***

Tese apresentada ao Instituto de Química da
Universidade de São Paulo para obtenção do
Título de Doutor em Bioquímica

Orientadora: Profa. Dra. Aline Maria da Silva

São Paulo
2006

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL 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

Moreira, Leandro Marcio.

Análise estrutural e funcional do genoma de *Xanthomonas axonopodis* pv. *citri*. / Leandro Marcio Moreira; orientadora Aline Maria da Silva.

São Paulo, 2006.

170 f.

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

1.Genômica comparativa. 2.Cancrose. 3.Compostos aromáticos. 4.Fitopatógeno. 5.Microarranjos de DNA. 6.Osmoproteção.

Leandro Marcio Moreira

Análise estrutural e funcional do genoma de *Xanthomonas axonopodis* pv. *citri*

Tese apresentada ao Instituto de Química da
Universidade de São Paulo para obtenção do
Título de Doutor em Bioquímica

Aprovado em: **12 de Setembro de 2006**

Banca Examinadora

Prof. Dr. João Paulo Kitajima

Instituição: Alellyx Applied Genomics

Especialidade: Bioinformata

Prof. Dra. Cláudia Monteiro-Vitorello

Instituição: Laboratório Nacional de Computação Científica - LNCC

Especialidade: Genômica

Prof. Dra. Regina Lúcia Baldini

Instituição: Instituto de Química da Universidade de São Paulo

Especialidade: Bioquímica e biologia molecular de microorganismos

Prof. Dr. Shaker Chuck Farah

Instituição: Instituto de Química da Universidade de São Paulo

Especialidade: Bioquímica de proteínas

Prof. Dra. Aline Maria da Silva

Instituição: Instituto de Química da Universidade de São Paulo

Especialidade: Bioquímica e biologia molecular de microorganismos

Suplentes

Prof. Dr. Gonçalo A. Guimarães Pereira

Instituição: Instituto de biologia - UNICAMP

Prof. Dr. Julio Rodrigues Neto

Instituição: Instituto Biológico - Campinas

Prof. Dr. Ricardo Harakava

Instituição: Instituto Biológico - SP

A minha esposa, Edmara Rocha dos Santos Moreira, que tanto contribuiu, fora do laboratório, para que este trabalho fosse concluído.

Eu amo você.

AGRADECIMENTOS

Agradeço a minha esposa, **Edmara Rocha dos Santos Moreira**, pelo apoio incondicional durante os momentos mais difíceis, e não foram poucos, ocorridos durante o período em que me dediquei a este trabalho. Ao meu pai, **José Osmar Moreira**, e minha falecida mãe, **Maria P. C. Moreira**, pelo esforço para que um dia eu pudesse chegar nesta etapa da vida.

Agradeço à **Profa. Dra. Aline Maria da Silva** por ter me "adotado" no momento em que eu mais precisava de apoio e sobretudo pela orientação e pela paciência, a qual ela teve de sobra, mesmo que muitas vezes eu a tenha "tirado do sério".

Agradeço à **Profa. Dra. Ana Cláudia Rases da Silva**, minha orientadora no Mestrado e nos primeiros dois anos do Doutorado por ter me oferecido a oportunidade de trabalhar na área da genômica de fitopatógenos.

Aos **Profs. Drs. Julio César de Oliveira, Jesus Aparecido Ferro e Luiz Roberto Furlan** pelo apoio e sugestões durante o desenvolvimento deste projeto.

Agradecimento inestimável ao amigo e professor **Dr. João Carlos Setúbal** pela orientação informal ao longo deste projeto e pelos ensinamentos em gestão de pessoal.

Um agradecimento especial aos amigos **Robson Francisco de Souza, Stefano Pashalidis, Jean Marcel, Alex Willian, Marcelo Luiz de Laia, Paulo Zaini, Paulo Paiva, José Roberto, Denise Yamamoto, Adriana Matsukuma, Júlio César Levano, Andréa Fogaça, Patrícia Pessoa, Cássia Docena, Daniela Gonzalez, Alexandre Sanchez, Luciano Digiampietri** e tantos outros aqui não nomeados.

Aos técnicos e funcionários do **Instituto de Química da USP**, do **Laboratório de Bioinformática do Instituto de Computação da UNICAMP** e do **Laboratório de Biologia Molecular da UNESP Jaboticabal** pela enorme ajuda durante todos estes anos.

Ao **sabiá** (*Turdus rufiventris*) que durante todas as manhãs ficava a cantar de frente ao bloco zero dando forças para que o dia fosse sempre mais animado, e aos **sapos** (*Bufo* spp) e **grilos** (*Grillus campestris*) que me acompanhavam na saída do Instituto durante as noites chuvosas e abafadas.

À **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)** pela concessão da bolsa de Doutorado e à **Universidade de São Paulo** em nome do Instituto de Química pela oportunidade.

Finalmente, agradeço a **Deus** por mais esta conquista.

*"O que afeta diretamente uma pessoa, afeta a todos indiretamente."
"Enfrentaremos a força física com a nossa força moral."
"Tenho visto demasiado ódio para querer odiar."
"Eu tentei ser direito e caminhar ao lado do próximo."
"Não permita que ninguém o faça descer tão baixo a ponto de você sentir ódio."*

Martin Luther King

*"Nunca permita que um problema a ser resolvido se torne
mais importante do que uma pessoa a ser amada."*

Barbara Johnson

LISTA DE ABREVIATURAS E SIGLAS

ABC	Motivo de ligação de ATP, do inglês <i>ATP Binding Cassete</i>
Avr	Genes de avirulência
BLAST	<i>Basic Local Alignment Search Tool</i>
CDS	Sequências Codificadoras, do inglês <i>Coding Sequences</i>
COG	<i>Clusters of Orthologous Groups</i>
CVC	Clorose Variegada dos Citros
DF	Moléculas difusíveis, do inglês <i>Diffusible Factor</i>
DPO	Degradação de Poli e Oligossacarídeos
DSF	Moléculas sinalizadoras, do inglês <i>Diffusible Signal Factor</i>
ECA	<i>Erwinia carotovora subsp. atroseptica</i>
EDPCV	Enzimas de Degradação de Parede Celular Vegetal
EPS	Polissacarídeos extracelulares, do inglês <i>Extracellular Polysaccharides</i>
GC %	Porcentagem (%) de nucleotídeos G e C
GSP	Via secretória geral, do inglês <i>General Secretory Pathway</i>
HR	Resposta de hiperssensibilidade, do inglês <i>Hypersensitive Response</i>
Hrp	Genes envolvidos com respostas de interação planta-patógeno, do inglês <i>Hypersensitive Response and Pathogenicity</i>
IP	Ilhas de Patogenicidade
ITL	Ilhas de Transferência Lateral
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
Kpb	Milhares de pares de bases
LPS	Lipopolissacarídeos, do inglês <i>Lipopolysaccharides</i>
LXX	<i>Leifsonia xyli subsp. xyli</i>
MCP	<i>Methyl-accepting Chemotactic Protein</i> , envolvida com quimiotaxia
MFS	Principal superfamília de transportadores, do inglês <i>Major Facilitator Superfamily of transporters</i>
Mpb	Milhões de pares de bases
NCBI	<i>National Center for Biotechnology Information</i>

PAMGO	Ontologia de genes de microorganismos associados a plantas, do inglês <i>Plant-Associated Microbe Gene Ontology</i>
pb	Pares de b ases
PCR	Reação em cadeia da polimerase, do inglês <i>Polymerase Chain Reaction</i>
PD	Doença de Pierce, do inglês <i>Pierce's Disease</i>
PFAM	<i>Protein FAM</i>ily
PInDel	Regiões prováveis de inserção e deleção, do inglês <i>Putative Insertion and Deletion Regions</i>
POY	<i>Phytoplasma asteris</i> cepa OY
PSP	<i>Pseudomonas syringae</i> pv. <i>phaseolicola</i> cepa 1448A
PST	<i>Pseudomonas syringae</i> pv. <i>tomato</i> cepa DC3000
PTS	Sistema de fosfo-transferência, do inglês <i>Phosphotransferase System</i>
RM	Sistema de R estrição e M odificação de DNA
RT-qPCR	PCR quantitativo precedido de transcrição reversa
SB	S íntese de B iotina
SPI-7	Ilha de patogenicidade determinada em <i>Salmonella enterica</i>
SS-I	Sistema S ecretório tipo I
SS-II	Sistema S ecretório tipo II
SS-III	Sistema S ecretório tipo III
SS-IV	Sistema S ecretório tipo IV
Swiss-Prot	<i>Swiss-Prot Protein knowledgebase</i>
TIGR	<i>The Institute of Genome Research</i>
UFC	Unidades F ormadoras de C olônias
UV	Radiação u ltravioleta
XAA-B	<i>Xanthomonas axonopodis</i> pv. <i>aurantifolii</i> cepa B
XAA-C	<i>Xanthomonas axonopodis</i> pv. <i>aurantifolii</i> cepa C
XAC/XAC-306	<i>Xanthomonas axonopodis</i> pv. <i>citri</i> cepa 306
XAC-A*	<i>Xanthomonas axonopodis</i> pv. <i>citri</i> variante A*
XAC-W	<i>Xanthomonas axonopodis</i> pv. <i>citri</i> variante W
XAM	<i>Xanthomonas axonopodis</i> pv. <i>manihots</i>
XAP	<i>Xanthomonas axonopodis</i> pv. <i>phaseoli</i>
XCC	<i>Xanthomonas campestris</i> pv. <i>campestris</i> cepa não especificada

XCC-8004	<i>Xanthomonas campestris</i> pv. <i>campestris</i> cepa 8004
XCC-ATCC	<i>Xanthomonas campestris</i> pv. <i>campestris</i> cepa ATCC33913
XCC-B100	<i>Xanthomonas campestris</i> pv. <i>campestris</i> cepa B100
XCP	<i>Xanthomonas campestris</i> pv. <i>passiflorae</i>
Xcs	Segundo agrupamento de genes que codificam o SS-II
XCV	<i>Xanthomonas campestris</i> pv. <i>vesicatoria</i>
XFA	<i>Xylella fastidiosa</i>
XF-Ann-1	<i>Xylella fastidiosa</i> pv. <i>oleander</i> cepa Ann-1
XF-CVC	<i>Xylella fastidiosa</i> cepa 9a5c, agente causal da CVC
XF-Dixon	<i>Xylella fastidiosa</i> pv. <i>almond</i> cepa Dixon
XF-PD	<i>Xylella fastidiosa</i> cepa temecula, agente causal da PD
XM	<i>Xanthomonas malvacearum</i>
XOO	<i>Xanthomonas oryzae</i> pv. <i>oryzae</i> cepa KAC10331
Xps	Primeiro agrupamento de genes que codificam o SS-II
XT	<i>Xanthomonas translucens</i>

RESUMO

Moreira, L.M. **Análise estrutural e funcional do genoma de *Xanthomonas axonopodis* pv. *citri***. 2006. 170p. Tese (Doutorado) - Programa de Pós-Graduação em Bioquímica. Instituto de Química. Universidade de São Paulo.

O cancro cítrico é uma doença que afeta diversas espécies de *Citrus*, cujo agente causal é *Xanthomonas axonopodis* pv *citri* (XAC). O genoma desta fitobactéria consiste de um cromossomo de ~5 Mpb e dois plasmídeos, que juntos codificam 4313 CDS (seqüências codificadoras), das quais 2710 apresentam similaridade com proteínas conhecidas. Neste trabalho realizamos uma análise comparativa detalhada do genoma de XAC com genomas de três fitopatógenos, *Xanthomonas campestris campestris*, *Xylella fastidiosa* 9a5c e *Xylella fastidiosa* temecula. Com esta análise identificamos genes espécie e gênero-específicos, potencialmente relevantes para adaptação aos seus respectivos nichos ou hospedeiros, além de ilhas de inserção e deleção genômica putativas. Também identificamos vias metabólicas relacionadas com osmoproteção/osmorregulação e com degradação de compostos aromáticos em XAC, que possivelmente são determinantes na eficácia de sua interação com o hospedeiro. Analisamos o nível de expressão de 9 CDS após crescimento de XAC em diferentes concentrações de glicose e verificamos que este açúcar modula positivamente a expressão de CDS relacionadas à síntese de goma e ao sistema de osmoproteção. Além disso, descrevemos a construção de microarranjos de DNA representando 2760 CDS de XAC, constituindo-se uma nova ferramenta para estudos de genômica comparativa e expressão gênica deste fitopatógeno.

Palavras chave: genômica comparativa, cancrose, compostos aromáticos, fitopatógeno, microarranjos de DNA, osmoproteção

ABSTRACT

Moreira, L.M. **Structural and functional analyses of *Xanthomonas axonopodis* pv. *citri* genome**. 2006. 170p. PhD Thesis. Biochemistry Graduate Program. Instituto de Química. Universidade de São Paulo.

Xanthomonas axonopodis pv *citri* (XAC) is the bacterial pathogen that causes citrus canker disease in several species of *Citrus* plants. XAC genome consists of a main chromosome of ~5 Mpb and two plasmids that together encode 4313 CDS (coding sequences). Approximately 63% of the CDS have assigned biological functions. In this work, we present a detailed genomic comparison between the genomes of XAC and of three other phytopathogens, *X. campestris* *campestris*, *Xylella fastidiosa* 9a5c and *X. fastidiosa* Temecula. Based on this analysis, we identified species and genus-specific genes that might be relevant for adaptation to their niches and hosts. We mapped putative insertion/deletion regions in the XAC genome possibly related to gene gains and losses during the divergence of the four bacterial lineages. We have identified the metabolic pathways related to osmoprotection/osmoregulation and aromatic compound degradation important for XAC efficient host colonization and interaction. Expression levels of 9 CDS were analyzed after XAC growth under different glucose concentrations revealing that this sugar upregulates the expression of CDS related to gum synthesis and to osmoregulation. In addition, we describe here the construction of a DNA microarray representing 2760 CDS of XAC as a new tool for comparative genomic and gene expression studies in this phytopathogen.

Keywords: aromatic compounds, citrus canker, comparative genomic, DNA microarrays, osmoprotection, phytopatogen

SUMÁRIO

1. INTRODUÇÃO.....	16
1.1. A IMPORTÂNCIA DA CITRICULTURA BRASILEIRA	16
1.2. O CANCRO CÍTRICO (CANCROSE)	17
1.3. GENÔMICA DE BACTÉRIAS FITOPATOGÊNICAS	24
1.3.1. <i>Erwinia carotovora</i> subsp. <i>atroseptica</i> (ECA)	24
1.3.2. <i>Leifsonia xyli</i> subsp. <i>xyli</i> (LXX)	25
1.3.3. <i>Phytoplasma asteris</i> cepa OY (POY).....	25
1.3.4. <i>Pseudomonas</i>	26
1.3.5. <i>Xanthomonas</i>	27
1.3.6. <i>Xylella fastidiosa</i>	29
1.3.7. <i>Análise comparativa dos genomas de algumas fitobactérias</i>	29
2. OBJETIVOS	34
3. METODOLOGIAS	36
3.1. METODOLOGIAS PARA ANÁLISE IN SILICO DE GENOMAS	36
3.1.1. <i>Agrupamento de genes em famílias</i>	36
3.1.2. <i>Análise filogenética</i>	37
3.1.3. <i>Classificação das integrases e das ilhas prováveis de inserção/</i> <i>deleção (PinDels)</i>	38
3.1.4. <i>Análise de vias de metabólicas</i>	38
3.2. CULTIVO E MANIPULAÇÃO DE <i>XANTHOMONAS AXONOPODIS</i> PV. CITRI	39
3.2.1. <i>Cultivo de XAC</i>	39
3.2.2. <i>Inoculação de XAC em folhas de citros</i>	40
3.3. ISOLAMENTO DE RNA TOTAL DE XAC E ENSAIOS DE RT-QPCR.....	40
3.4. CONSTRUÇÃO DOS MICROARRANJOS DE DNA DE XAC (XACARRAY)	42
3.4.1. <i>Seleção dos clones da biblioteca genômica e amplificação dos</i> <i>insertos por PCR</i>	42
3.4.2. <i>Preparação dos microarranjos de DNA</i>	44
3.4.3. <i>Preparação de DNA marcado e hibridação dos microarranjos</i>	44

3.4.4. *Deteção, quantificação e normalização dos sinais de fluorescência* 45

4. RESULTADOS E DISCUSSÃO.....46

4.1. ANÁLISE COMPARATIVA DOS GENOMAS COMPLETOS DAS BACTÉRIAS XAC, XCC, XF-CVC E XF-PD.....46

4.1.1. *Comparação dos quatro genomas de acordo com as categorias funcionais de anotação*.....49

4.1.2. *Comparação dos quatro fitopatógenos com base na presença ou ausência de genes*.....54

4.1.2.1. Genes mais representados e exclusivos em cada uma das espécies/cepas analisadas54

Genes exclusivos de XAC.....61

Genes exclusivos de XCC63

Genes exclusivos de XF-CVC.....67

Genes exclusivos de XF-PD67

4.1.2.2. Genes mais representados em cada um dos gêneros69

4.1.2.2.1. Genes mais representados no genoma das duas bactérias do gênero *Xanthomonas*69

Sistema Secretório Tipo II (SS-II)69

Enzimas de degradação de parede celular vegetal (EDPCV).....73

Genes relacionados com aquisição de açúcares.....73

Genes relacionados com captação e metabolismo de ferro75

Sistema secretório tipo III (SS-III)75

Quimiotaxia e genes relacionados com síntese de flagelo78

Genes com funções regulatórias85

4.1.2.2.2 Genes mais representados no genoma de bactérias do gênero *Xylella*.....86

Pili86

Colicinas-V89

Sistema de Restrição e Modificação de DNA93

4.1.2.3. Genes mais representados exclusivamente nos genomas de XAC e XF-CVC.....94

Ilha de patogenicidade SPI-7	94
4.1.2.4. Genes mais representados na família Xanthomonadaceae.....	97
4.1.3. <i>Determinação de PlnDels nos quatro genomas comparados</i>	101
4.2. ANÁLISE DO SISTEMA DE OSMOPROTEÇÃO E OSMORREGULAÇÃO E AS VIAS DE DEGRADAÇÃO DE COMPOSTOS AROMÁTICOS CODIFICADOS NO GENOMA DE XAC...	109
4.2.1. <i>Análise do sistema de osmoproteção e osmorregulação</i>	109
4.2.2 <i>Análise das vias de degradação de compostos aromáticos</i>	119
4.2.2.1. Implicação da via de degradação do β -cetoadipato na virulência de XAC.....	127
4.3. CONSTRUÇÃO DOS MICROARRANJOS DE DNA DE XAC (XACARRAY)	129
4.3.1. <i>Validação da qualidade do XACarray</i>	135
4.4. <i>Análise do perfil de expressão de um conjunto seletivo de CDS em XAC cultivada em diferentes concentrações de glicose</i>	138
<i>nuoA</i> (XAC2704)	140
<i>pykA</i> (XAC3345)	140
<i>kdpA</i> (XAC0756)	140
<i>rpfC</i> (XAC1878)	140
<i>rpfB</i> (XAC1880) e <i>rpfF</i> (XAC1879).....	141
<i>gumK</i> (XAC2576) e <i>gumD</i> (XAC2583).....	141
5. CONSIDERAÇÕES FINAIS	143
<u>6. REFERÊNCIAS BIBLIOGRÁFICAS</u>	148
LISTA DE ANEXOS ver CD-ROM como material suplementar.....	170

1. INTRODUÇÃO

1.1. A importância da citricultura brasileira

O Brasil é o maior produtor de laranjas, sendo responsável atualmente por 27% da produção mundial. Entretanto, deve-se ao suco de laranja concentrado o real ganho de divisas oriundas de exportações pela captação de mais de US\$ 1,5 bilhões/ano, além de proporcionar a geração de 400 mil empregos diretos e cerca de outros 3 milhões de empregos indiretos (FUNDECITRUS, 2006). Atualmente, a área total cultivada no território brasileiro é superior a 1 milhão de hectares, dos quais cerca de 80% estão localizados na região sudeste do País, de tal maneira que 98% da produção frutífera nacional é proveniente do Estado de São Paulo, predominantemente da região de Campinas, São Carlos, São José do Rio Preto e Barretos. É importante destacar que nos últimos anos o nordeste brasileiro, mais precisamente os Estados da Bahia e Sergipe, está emergindo como uma nova e promissora área de produção cítrica, sobretudo em função da ausência de doenças e pragas encontradas na região sudeste, o que tenderá a destacar ainda mais o País no exigente mercado externo (NEVES *et al.*, 2001, FUNDECITRUS, 2006).

Mesmo com estes significativos ganhos anuais, a produção cítrica sofre perdas consideráveis em consequência de diversas pragas e doenças. Assim, existe grande interesse na obtenção de espécies e variedades de plantas mais resistentes e tolerantes, bem como no aprimoramento de medidas profiláticas e de erradicação destas enfermidades. Dentre as doenças que afetam a citricultura, merecem destaque o cancro cítrico, a clorose variegada dos citros (CVC), a pinta preta, o *greening*, a morte súbita, a leprose e a tristeza dos citros (ROSSETI *et al.*, 1990, LEITE, 2003, LOCALI *et al.*, 2003, ROMÁN *et al.*, 2004, FUNDECITRUS, 2006).

1.2. O cancro cítrico (cancrose)

O cancro cítrico ou cancrose é uma doença que afeta a maioria das espécies do gênero *Citrus*, as quais exibem diferentes respostas de resistência ao patógeno (GOTTWALD *et al.*, 2002). Em plantas compatíveis os sintomas são observados em folhas, frutos ou ramos como mostrado na Figura 1. Nas folhas, as lesões inicialmente são pequenas pústulas, puntiformes e esponjosas, circundadas ou não por um halo amarelo (clorótico) em ambos os lados da folha (Fig. 1A, B, C) (BRUNINGS & GABRIEL, 2003). Estas lesões espalham-se rapidamente (2 a 4 semanas) pelo tecido, aumentando de diâmetro e tornando-se pardacentas ou escuras, com aspecto corticoso e saliente (Fig. 1C). A fotossíntese é prejudicada em função da destruição da área foliar que muitas vezes sofre desfolha. Em frutos e ramos os sintomas são semelhantes aos observados em folhas, porém, a parte corticosa é bem pronunciada e as lesões podem ou não estar envolvidas por um halo clorótico (Fig. 1D, E). A formação clorótica das folhas está associada a um aumento na produção de etileno, como consequência da infecção bacteriana, o que leva a ativação da cascata de degradação de clorofila e indução da síntese de β -caroteno no tecido vegetal (DOMINGUES *et al.*, 2001). Esta alteração metabólica também é observada no fruto, cuja casca passa a apresentar rupturas que resultam na exposição do mesocarpo, comprometendo seu aspecto e tornando-o inadequado para o consumo (Fig 1F, G) (BRUNINGS & GABRIEL, 2003).

Três formas de cancrose já foram descritas até o momento e nomeadas como cancrose A, B e C. A cancrose A, considerada a mais relevante, é originária da Ásia e está atualmente disseminada por quase todos os continentes, acometendo todas as variedades e espécies de citros. Nos Estados Unidos da América, a cancrose A surgiu por volta de 1910, na região norte da Flórida (SCHUBERT & SUN, 1996). No Brasil, o primeiro relato da doença data de 1957 no município de Presidente Prudente, extremo oeste do Estado de São Paulo (BITANCOURT, 1957). O agente causal da cancrose A é a bactéria *Xanthomonas axonopodis* pv *citri* (**XAC**) (BITANCOURT, 1957), cujo genoma foi

completamente seqüenciado por um consórcio de grupos de pesquisa paulistas (DA SILVA *et al.*, 2002). A Tabela 1 resume as características principais do genoma de XAC, que além do cromossomo principal de aproximadamente 5Mpb, é composto por dois plasmídeos, sendo um de aproximadamente 64 Kpb e outro de 33 Kpb. No cromossomo principal foram anotadas 4.313 CDS (Seqüências codificadoras), das quais 2.710 (63 %) têm similaridade como proteínas de função conhecida em outros organismos. A complexidade do genoma de XAC será detalhada e discutida na seção **4.1** desta Tese (Resultados), comparativamente aos genomas de outros fitopatógenos pertencentes a mesma família.

Recentemente outras duas variantes da cancrose A foram descritas. A primeira foi detectada no sudeste da Ásia em 1998, infectando *Citrus aurantifolia*. O patógeno foi isolado e nomeado como *X. axonopodis* pv. *citri* variante A* (**XAC-A***) com base em testes sorológicos que o distinguiram de XAC (VERNIÈRE *et al.*, 1998). A segunda foi isolada no sul da Flórida (EUA) em 2003 infectando *Citrus aurantifolia* (limão mexicano ou galego) e *Citrus macrophylla* (alemow). O patógeno foi nomeado como *X. axonopodis* pv. *citri* variante A^w (**XAC-A^w**) (SUN *et al.*, 2004).

A cancrose B ou falsa cancrose foi identificada originalmente na Argentina em 1923 e descrita como uma cancrose restrita a *Citrus limon* (limão siciliano) (Fig. 1G) (CIVEROLO, 1984). Entretanto, a cancrose B também acomete *Citrus sinensis* (laranja doce) e *Citrus paradisi* (pomelo) (CIVEROLO, 1984). O agente causal foi isolado e classificado como *X. axonopodis* pv. *aurantifolii* tipo B (**XAA-B**). O tempo de aparecimento de sintomas na cancrose B é quatro a cinco vezes maior do que observado para cancrose A, e por isto recebeu a denominação de falsa cancrose. A cancrose C foi descrita inicialmente em 1963 no Brasil e restringe sua patogênese a plantas de *Citrus aurantifolia* (limão mexicano ou galego), mesmo em pomares constituídos de outras variedades de citros (NAMEKATA, 1971). O agente causal foi então identificado e nomeado como *Xanthomonas axonopodis* pv. *aurantifolii* tipo C (**XAA-C**).

Em todos os casos de cancrose já descritos, a disseminação e transmissão da doença geralmente ocorrem por intermédio de exudados de lesões, principalmente em dias com umidade relativa elevada ou chuvosos. As gotas de água dispersam *X. axonopodis* a curtas distâncias ao colidirem diretamente com as lesões. Na presença de ventos, a dispersão do patógeno atinge distâncias maiores. Também são agentes potenciais de dispersão de *X. axonopodis*, os animais, o maquinário e utensílios agrícolas, além das pessoas que circulam nos pomares (GOTTWALD *et al.*, 2002). Em contato com aberturas naturais do tecido vegetal, como estômatos, lenticelas, hidatódios, poros de água ou mesmo lesões teciduais, *X. axonopodis* passa a colonizar o mesófilo foliar provocando os sintomas clássicos da doença em hospedeiros compatíveis, ou uma resposta de hipersensibilidade (HR) em hospedeiros incompatíveis (HAMMOND-KOSACK & JONES, 1996).

Até o momento, não há nenhuma forma de cura descrita para a cancrose. Portanto, a eliminação do agente causal e erradicação das plantas contaminadas é a melhor maneira de limitar a propagação da doença. Exemplos bem sucedidos de erradicação ocorreram em vários países tais como África do Sul, Austrália, Ilhas Fiji, Moçambique, Nova Zelândia e Estados Unidos, onde a doença está praticamente dizimada e novos focos são imediatamente controlados (KOIZUMI, 1985, SCHUBERT & SUN, 1996). Na tentativa de erradicar a doença, países como Brasil, Uruguai e Argentina também adotaram um rigoroso programa de prevenção e controle da cancrose, que inclui a utilização de quebra-ventos, plantio de mudas certificadas e de variedades mais tolerantes em regiões selecionadas, pulverizações com agro-químicos, quando necessário, além de constante vigilância e inspeção dos pomares (LEITE JUNIOR & MOHAN, 1990). É importante destacar que XAC tende a usufruir de aberturas vegetais para sua propagação (Fig. 1H), assim, o combate e a prevenção à larva minadora dos citros (*Phyllocnistis citrella*), uma das pragas que provoca grandes danos aos tecidos vegetais, é de fundamental importância no controle do cancro cítrico.

Além do cancro cítrico, bactérias do gênero *Xanthomonas* são responsáveis por doenças em várias outras plantas. Infecções induzidas por *Xanthomonas* já foram descritas em pelo menos 124 plantas monocotiledôneas e 268 plantas dicotiledôneas causando sintomas distintos (LEYNS *et al.*, 1984). Dentre estes sintomas merecem destaque especial as formações necróticas ou hiperplásicas em diferentes tecidos vegetais. As doenças provocam ainda alterações no parênquima e em tecidos vasculares que comprometem dramaticamente o crescimento e o desenvolvimento da planta, consequentemente sua sobrevivência (SWINGS & CIVEROLO, 1993). Entre as doenças causadas por outras espécies de *Xanthomonas* se destacam a podridão negra das crucíferas (*X. campestris* pv. *campestris*; XCC), a mancha bacteriana em pimenteiras e tomateiros (*X. campestris* pv. *vesicatoria*; XCV), a estria bacteriana do arroz (*X. oryzae* pv. *oryzae*; XOO), o crestamento bacteriano do feijoeiro (*X. axonopodis* pv. *phaseoli*; XAP), a bacteriose da mandioca (*X. axonopodis* pv. *manihotis*; XAM), a mancha-angular dos algodoeiros (*X. malvacearum*; XM), a mancha oleosa do maracujazeiro (*X. campestris* pv. *passiflorae*; XCP) e a bacteriose de cereais, como o trigo, e de gramíneas (*X. translucens* group; XT).

As *Xanthomonas* são proteobactérias gram-negativas da divisão gama e pertencentes à família das Xanthomonadaceae. Morfologicamente apresentam formato bacilar com diâmetro variando entre 0.2-0.6 µm e comprimento entre 0.8-2.9 µm, geralmente com um único flagelo na posição polar (monotríquia) e um genoma em torno 5 Mpb, que pode ou não estar associado com plasmídeos (SWINGS & CIVEROLO, 1993). Estas bactérias apresentam um pigmento amarelado denominado *xanthomonadina*, que é freqüentemente utilizado como fator de classificação taxonômica e de identificação de espécies (JENKINS & STARR, 1985). Este pigmento é formado por uma estrutura de aril-polienos, que por sua vez está inserido na parede bacteriana, conferindo a estes organismos resistência contra danos provocados por radiação UV, portanto, servindo como um fator de foto-proteção (POPLAWSKY *et al.*, 2000). Graças a este pigmento, bactérias deste gênero podem sobreviver como organismos epífitos, embora

sejam classicamente relatados como patógenos vegetais (HIRANO & UPPER, 1983).

Como discutido acima, bactérias do gênero *Xanthomonas* apresentam uma singular importância na fitopatologia, por serem capazes de infectar e provocar danos em plantas de elevada importância sócio-econômica, como é o caso do arroz, algodão, feijão, batata, mandioca trigo e cereais. Ainda que o modo de transmissão e infecção das diferentes espécies de *Xanthomonas* seja similar, a sintomatologia gerada em cada hospedeiro é decorrente da espécie bacteriana. Há espécies sistêmicas, como é o caso de XCC ao induzir podridão negra das crucíferas, outras causam infecções locais com disseminação célula-a-célula, como é o caso de XAC em plantas de citros. Assim, aprofundar os conhecimentos sobre a biologia de *Xanthomonas* é fundamental para uma melhor compreensão de seus mecanismos de patogenicidade e de interação com seus hospedeiros.

Figura 1: Sintomatologia da cancriose em citros. **A)** Lesões corticosas na face inferior e na região inter-nervuras da folha (região abaxial); **B)** Detalhes de lesões corticosas na face superior da folha (região adaxial); **C)** Lesões foliares rodeadas por aros cloróticos; **D)** Detalhe de lesões salientes em ramos do caule; **E)** Lesões causadas pelo cancro cítrico em frutos de laranja pêra; **F)** Aglutinação da lesão e conseqüente ruptura da casca do fruto; **G)** Frutos de limão siciliano infectados com cancriose B; **H)** Lesões em forma de galerias produzidas pela larva minadora dos citros (*Phyllocnistis citrella*) (à direita) e conseqüente infecção e formação de cancriose por XAC na região previamente lesionada do tecido vegetal (à esquerda). Fotos A, B, C, D, E, F e H retiradas do sítio da Fundecitrus (www.fundecitrus.com.br) e foto G retirada do sítio da "The Phytopathological Society" (www.aspnet.org).

Tabela 1: Características gerais do genoma de XAC^a

Características	Cromossomo	pXac64^b	pXac33^b
Tamanho em pb	5.175.554	64.920	33.699
Conteúdo de GC (%)	64.7	61.4	61.9
Seqüências codificadoras (CDS) (%)	85.59	89.26	81.59
Total de CDS anotadas	4.313	73	42
CDS com funções determinadas	2.710	39	21
CDS hipotéticas conservadas	1.272	7	7
CDS hipotéticas	331	27	14
<i>Operon</i> de rRNAs	2	-	-
RNA transportadores (tRNAs)	54	-	-
Elementos de inserção	87	10	11

^aDados adaptados de da Silva e colaboradores (DA SILVA *et al.*, 2002). ^bpXac64 e pXac33, plasmídeos de XAC.

1.3. Genômica de bactérias fitopatogênicas

Nos últimos sete anos, o seqüenciamento e anotação de genomas de várias fitobactérias foram total ou parcialmente concluídos. Até junho de 2006, 17 espécies ou subespécies/patovares destas bactérias tiveram seus genomas completamente seqüenciados enquanto que os genomas de outras 25 estão em fase de seqüenciamento e/ou anotação. A Tabela 2, lista estas 42 bactérias, as doenças que causam ou suas aplicações bem como seus principais hospedeiros, que em sua maioria representam culturas agrícolas de grande interesse sócio-econômico. Do total de genomas de fitobactérias seqüenciados, 80% são de proteobactérias, das quais 67% são da divisão gama. Destacamos a seguir algumas informações derivadas da análise do genoma de algumas das espécies listadas na Tabela 2, como descrito em Setubal e colaboradores (SETUBAL *et al.*, 2005) ¹

1.3.1. *Erwinia carotovora* subsp. *atroseptica* (ECA)

ECA é uma bactéria do grupo das enterobacteriáceas cujos membros incluem patógenos humanos muito bem estudados, tais como *E.coli*, *Salmonella enterica* e *Yersinia pestis*, que também tiveram seus genomas seqüenciados (PERNA *et al.*, 1998, MCCLELLAND *et al.*, 2001, PARKHILL *et al.*, 2001a, PARKHILL *et al.*, 2001b, PERNA *et al.*, 2001, DENG *et al.*, 2002). O seqüenciamento do genoma de ECA revelou que aproximadamente 1500 genes (33% do total do genoma de 5Mpb) são específicos da cepa fitopatogênica com relação às cepas enteropatogênicas, sugerindo extrema especialização do genoma em relação ao hospedeiro vegetal (*Solanum tuberosum*) (BELL *et al.*, 2004). Este grupo de genes exclusivos de ECA compreende genes classicamente relacionados à patogenicidade, muitos dos quais presentes em

¹ Setubal, J. C., Moreira, L. M. & da Silva, A. C. (2005). Bacterial phytopathogens and genome science. *Curr Opin Microbiol.* 8, 595-600. (ANEXO 1).

ilhas de patogenicidade (IP). Mutações em alguns dos genes que compõem duas destas ilhas, por inserção de um elemento de transposição, alterou a virulência de ECA (BELL *et al.*, 2004). Uma destas ilhas contém os genes que codificam para o sistema secretório tipo IV (SS-IV) ao passo que a outra contém genes envolvidos com a síntese de fitotoxinas (poliquetídeos).

1.3.2. *Leifsonia xyli* subsp. *xyli* (LXX)

Leifsonia xyli subsp. *xyli* foi a primeira fitobactéria gram-positiva que teve seu genoma de 2.6Mpb completamente seqüenciado (MONTEIRO-VITORELLO *et al.*, 2004). Esta bactéria habita o xilema de *Saccharum officinarum* (cana-de-açúcar) (DAVIS *et al.*, 1980) e é responsável pela doença conhecida como raquitismo-da-soqueira da cana-de-açúcar, que afeta cerca de 20% da cultura canavieira (CARNEIRO-JUNIOR *et al.*, 2004). É interessante observar que aproximadamente 13% do genoma de LXX é composto por pseudogenes. Além disso, o genoma possui muitos genes codificadores de transportadores do tipo ABC, relacionados com a aquisição de açúcares e outros nutrientes (MONTEIRO-VITORELLO *et al.*, 2004). Por outro lado, o número de genes relacionados à patogenicidade é bastante reduzido se comparado a outros fitopatógenos, porém inclui genes que codificam para hemolisinas, hemaglutininas e proteínas envolvidas com adesão celular, das quais algumas se encontram em IP.

1.3.3. *Phytoplasma asteris* cepa OY (POY)

O genoma de POY é, até o momento, o menor genoma de fitobactéria descrito (OSHIMA *et al.*, 2004). Seu genoma tem aproximadamente 870 Kpb, tamanho seis vezes menor que o genoma de bactérias do gênero *Xanthomonas* e três vezes menor do que bactérias do gênero *Xylella*. É interessante destacar que POY possui menos funções metabólicas que o próprio *Mycoplasma* spp, outrora relatado como sendo o organismo que apresentava o chamado “genoma mínimo” (HANCOCK, 1996, RAZIN, 1997, WEGRZYN, 2001). Em POY foi descrita a ausência de todas as subunidades que compõe o complexo de ATP-

sintase (subunidades F⁰ e F¹), estando a síntese de ATP limitada à via glicolítica. Por esta e outras perdas gênicas, POY é considerado um endosimbionte intracelular obrigatório. Entretanto, POY apresenta pelo menos 27 genes que codificam proteínas relacionadas ao transporte de íons, ATP e outras moléculas, alguns dos quais representados em alto número de cópias como resultado de eventos de duplicação gênica, compensando a escassez de vias importantes.

1.3.4. *Pseudomonas*

O gênero *Pseudomonas* compreende uma grande variedade de espécies patogênicas para animais e vegetais (WIDMER *et al.*, 1998). Embora várias espécies de *Pseudomonas* tenham sido seqüenciadas, *Pseudomonas syringae* pv. *tomato* cepa DC3000 (PST) foi a primeira espécie fitopatogênica deste gênero a ter seu genoma completamente seqüenciado (BUELL *et al.*, 2003). Em seu genoma de 6 Mpb foram identificados vários agrupamentos gênicos que codificam para proteínas efetoras do sistema secretório do tipo III (SS-III), classicamente relacionado com patogenicidade e virulência. Além dos 31 genes efetores previamente caracterizados experimentalmente, a análise *in silico* do genoma revelou a existência de 19 novos candidatos, muitos dos quais também presentes em bactérias de outros gêneros, tais como *Xanthomonas* e *Ralstonia*. Este trabalho aumentou significativamente a lista de genes candidatos a efetores de virulência previamente caracterizados, que somam atualmente mais de 200 genes (ALFANO & COLLMER, 2004) (<http://pseudomonas-syringae.org>). A análise revelou ainda outros 298 genes potencialmente relacionados à patogenicidade em PST, dos quais 96 são exclusivos de PST em relação a *P. aeruginosa*, a espécie melhor estudada deste gênero.

Outra espécie do gênero *Pseudomonas* que teve seu genoma seqüenciado foi *Pseudomonas syringae* pv. *phaseolicola* cepa 1448A (PSP), o qual também possui 6 Mpb (JOARDAR *et al.*, 2005). Estruturalmente o genoma de PSP é muito semelhante a PST. Do total de genes preditos como envolvidos com virulência no genoma de PSP, cerca de 240 (81%) estão presentes em PST, incluindo genes que codificam para efetores do tipo Hop (LINDEBERG *et*

al., 2005), para proteínas que compreendem os sistemas secretórios, para enzimas de degradação de parede celular vegetal e para proteínas envolvidas com adesão celular, enquanto que o restante parece denotar especificidade pelo hospedeiro, como é o caso dos genes que participam da síntese de fitotoxina coronatina, uma toxina bacteriana exclusiva desta espécie. Além disto, entre os genes exclusivos de PSP há um agrupamento de genes que codificam para SS-III que é homólogo ao encontrado em *Rhizobium*, *Photorhabdus*, *Aeromonas* spp. ou mesmo em *P. aeruginosa*. PSP ainda apresenta genes envolvidos com a síntese de *phaseolotoxins*, toxinas específicas desta espécie, e oito regiões envolvidas com síntese não-ribossomal de peptídeos, cujos genes apresentam domínios de poliquetídio sintase.

1.3.5. *Xanthomonas*

Até o momento já foram completamente seqüenciados os genomas de sete espécies ou patovares de *Xanthomonas*, sendo que o genoma de mais seis deverão ser completados em breve (DA SILVA *et al.*, 2002, VORHOLTER *et al.*, 2003, LEE *et al.*, 2005, QIAN *et al.*, 2005, THIEME *et al.*, 2005). A Tabela 3 resume as principais características de oito destes genomas. Análise detalhada da seqüência genômica de *Xanthomonas oryzae* pv. *oryzae* cepa KAC10331 (XOO) que infecta exclusivamente *Oryzae sativa*, uma monocotiledônea de elevada importância econômica (LEE *et al.*, 2005), revelou um total de 80% de similaridade com outras duas espécies previamente seqüenciadas, XAC e XCC-ATCC (DA SILVA *et al.*, 2002). Entretanto, o alinhamento dos cromossomos destas três bactérias revelou a ocorrência de rearranjos genômicos relevantes. Esta quantidade de rearranjos pode ser decorrente do fato de XOO apresentar pelo menos o dobro do número de elementos de transposição previamente descritos nas outras duas espécies. Do total de genes presentes em XOO, 245 são exclusivos desta espécie e, portanto, são candidatos potenciais para o estudo de interação entre XOO e seu hospedeiro específico.

Recentemente, o genoma de *Xanthomonas campestris* pv. *campestris* cepa 8004 (XCC-8004) foi também sequenciado (QIAN *et al.*, 2005). A

comparação deste genoma com outra cepa da mesma espécie XCC-ATCC (DA SILVA *et al.*, 2002) revelou que rearranjos estruturais observados são bem maiores entre estas duas cepas do que o que observado na comparação com uma espécie relacionada, por exemplo XOO. Este trabalho revelou ainda que 108 genes são exclusivos de XCC-8004 e outros 64 exclusivos de XCC-ATCC. A função de alguns destes genes anotados foi então avaliada através da análise fenotípica de 16.512 mutantes de XCC-8004 obtidos por inserção aleatória de um elemento de transposição. Entre os mutantes que apresentaram fenótipos alterados de patogenicidade quando inoculados em plantas de *Brassica oleracea*, muitos deles continham inserções em regiões exclusivas de XCC-8004 com relação a XCC-ATCC, sugerindo a existência de ilhas de patogenicidade específicas dentro de variantes de mesma espécie.

Xanthomonas campestris pv. *vesicatoria* (XCV) é a espécie de *Xanthomonas* cujos genes de virulência foram muito bem caracterizados funcionalmente, como é o caso dos genes da família *avrBs* (RONALD & STASKAWICZ, 1988, FOUTS *et al.*, 2003). O genoma de XCV é mais colinear com o genoma de XAC do que com qualquer outro membro deste gênero, cujos genomas já foram sequenciados não seguindo os padrões de distribuição filogenética por uso dos genes ribossomais 16S (THIEME *et al.*, 2005). Enquanto XCC não possui plasmídeos, em XCV foram descritos quatro, tendo o maior deles aproximadamente 185 Kpb. Este plasmídeo contém uma cópia adicional do SS-IV homólogo ao encontrado em *Legionella* spp. Adicionalmente, em XCV foram descritos um número maior de genes **MCP** (**Methyl-Accepting Chemotactic Protein**) posicionados em série, totalizando 14, contra os 10 previamente descritos em XAC ou XCC-ATCC (DA SILVA *et al.*, 2002, MOREIRA *et al.*, 2004, MOREIRA *et al.*, 2005), além de 68 elementos de transposição, alguns dos quais flanqueando regiões com nítida variação de composição de nucleotídeos, evidenciando possíveis ilhas de transferência lateral (ITL).

1.3.6. *Xylella fastidiosa*

O genoma da bactéria *Xylella fastidiosa* cepa 9a5c (XF-CVC), agente causal da clorose variegada dos citros (CVC) foi o primeiro genoma de um fitopatógeno a ser completamente elucidado (SIMPSON *et al.*, 2000). Dois anos depois, foi relatada a seqüência incompleta de genomas de *Xylella fastidiosa* pv. *almond* cepa Dixon (XF-Dixon) e *Xylella fastidiosa* pv. *oleander* cepa Ann-1 (XF-Ann-1), agentes causais da escaldadura da amendoeira (*Prunus dulcis*) e da espirradeira (*Nerium oleander*), respectivamente (BHATTACHARYYA *et al.*, 2002a, BHATTACHARYYA *et al.*, 2002b). Em 2003, foi finalizado o seqüenciamento completo do genoma de outra cepa, *Xylella fastidiosa* cepa temecula (XF-PD) (VAN SLUYS *et al.*, 2003), agente causal da doença de Pierce que ocorre em videiras. A Tabela 3 resume as principais características dos genomas destas quatro cepas de *Xylella fastidiosa*. XF-CVC apresenta um cromossomo de 2,67 Mpb e dois plasmídeos, com um total de 2848 seqüências codificadoras (CDS), sendo que ~50% destas apresentam uma função putativa (SIMPSON *et al.*, 2000, VAN SLUYS *et al.*, 2003). XF-PD, por sua vez, apresenta um cromossomo de 2,51 Mpb, mas apenas um plasmídeo, com um total de 2066 CDS, sendo que para 68% foi atribuída uma função putativa (VAN SLUYS *et al.*, 2003). É importante destacar que a diferença observada nos números de CDS entre XF-CVC e XF-PD decorre da diferença de critérios utilizados durante o processo de anotação, que resultou na eliminação de um conjunto de cerca de 500 CDS com tamanho reduzido.

1.3.7. Análise comparativa de alguns dos genomas de fitobactérias

A análise dos genomas dos fitopatógenos descritos acima (Tabela 2) permite classificá-los em dois grandes grupos: o grupo composto pelas bactérias POY, LXX e *Xylella fastidiosa*, cujos genomas apresentam tamanho menor o que pode refletir a ocorrência de redução progressiva ao longo da sua evolução, e o grupo composto pelas *Xanthomonas*, *Pseudomonas* e ECA, que aparentemente sofreram expansão de seus genomas ao longo de sua evolução.

Entre os genomas que compõem o primeiro grupo, observa-se 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 18% do genoma de POY resulte de duplicação gênica (OSHIMA *et al.*, 2004) ou que seqüências completas de fagos tenham se inserido no genoma de XF-CVC (SIMPSON *et al.*, 2000).

Embora, os genomas de *Xylella fastidiosa* e LXX não apresentem depleção equivalente àquela observada para POY, estas duas bactérias carecem de genes e vias importantes, fato este que provavelmente reflete ou é consequência do crescimento destas bactérias exclusivamente no xilema de seus respectivos hospedeiros. Além disso, POY, XF-CVC e LXX dependem de um vetor animal para sua propagação, o que também pode estar relacionado às perdas gênicas, também observadas em *Wigglesworthia glossinidia brevipalpis* (AKMAN *et al.*, 2002) e *Blochmannia floridanus* (DEGNAN *et al.*, 2005), bactérias que infectam insetos.

Em resumo, a disponibilidade das seqüências genômicas destes diversos fitopatógenos (Tabela 2) têm possibilitado estudos comparativos *in silico*, facilitando a identificação de genes relacionados a mecanismos de adaptação, de patogenicidade e de virulência (BHATTACHARYYA *et al.*, 2002a, BHATTACHARYYA *et al.*, 2002b, VAN SLUYS *et al.*, 2002, VAN SLUYS *et al.*, 2003, MOREIRA *et al.*, 2004, MOREIRA *et al.*, 2005). Recentemente, uma análise comparativa dos genomas de PST, XAC, XCC-ATCC, *A. tumefaciens*, *R. solanacearum* e XF-CVC revelou a existência de variação na composição de bases dos genes supostamente super-expressos nestas bactérias (FU *et al.*, 2005). A partir desta observação, foi elaborado um modelo capaz de determinar outros genes com perfil similar de composição, os chamados genes PHX, ou genes com perfil de serem altamente expressos, do inglês ***Predicted Highly Expressed***. Interessantemente, muitos dos genes PHX compreendem genes relacionados a patogenicidade, genes que codificam para proteínas de membrana e genes relacionados a biossíntese do flagelo. Outro exemplo de comparação *in silico* de genomas de fitopatógenos é o estudo realizado por

STUDHOLME e colaboradores (2005), que identificaram e analisaram domínios proteicos característicos que estão relacionados a mecanismos de interação entre planta e patógeno. Vale destacar que está em desenvolvimento um banco de dados com informações ontológicas, denominado PAMGO, do inglês *Plant-Associated Microbe Gene Ontology* (<http://pamgo.vbi.vt.edu>), cujo objetivo central é expandir o consórcio nomeado GOC ou *Gene Ontology Consortium* (<http://www.geneontology.org>), incluindo termos e relações relativas aos microorganismos associados a plantas.

Ademais, o conhecimento do genoma destes patógenos tem possibilitado a utilização de diversas tecnologias para análise funcional genomas tais como os microarranjos de DNA (DE SOUZA *et al.*, 2003, NUNES *et al.*, 2003, DE SOUZA *et al.*, 2004, KOIDE *et al.*, 2004, ASTUA-MONGE *et al.*, 2005, DE SOUZA *et al.*, 2005, PASHALIDIS *et al.*, 2005), ensaios de interação entre proteínas no sistema de duplo-híbrido (ALEGRIA *et al.*, 2004, KABISCH *et al.*, 2005, MACCHERONI *et al.*, 2005) e análise de proteomas (SMOLKA *et al.*, 2003, KAZEMI-POUR *et al.*, 2004, NOEL-GEORIS *et al.*, 2004, WATT *et al.*, 2005).

Tabela 2: Lista das fitobactérias com genomas completa ou parcialmente seqüenciados.

Bactéria	Grupo	Doença ou Aplicação	Hospedeiro Principal	Situação	Referência ^a
<i>Agrobacterium radiobacter</i> K84	Alpha	Biocontrole para galha-da-coroa	Plantas dicotiledôneas em geral	Em andamento	AGRO
<i>Agrobacterium tumefaciens</i> C58	Alpha	Galha-da-coroa	Plantas dicotiledôneas em geral	Completo	(GOODNER <i>et al.</i> , 2001, WOOD <i>et al.</i> , 2001a)
<i>Agrobacterium vitis</i> S4	Alpha	Galha-da-coroa	Videiras	Em andamento	AGRO
<i>Burkholderia ambifaria</i> AMMD	Beta	Biocontrole para plantas de ervilha	Plantas de ervilha	Em andamento	JGI
<i>Burkholderia cepacia</i>	Beta	Podridão bacteriana da escama	Plantas de cebola	Em andamento	JGI
<i>Burkholderia glumae</i>	Beta	Prodrisão da semente e panicle blight	Plantas de arroz	Em andamento	GGL
<i>Ralstonia metallidurans</i>	Beta	Biorremediação de metais	Não patogênica	Em andamento	JGI
<i>Ralstonia solanacearum</i> 1609	Beta	Murcha bacteriana	Batateiros	Em andamento	GENOSCOPE
<i>Ralstonia solanacearum</i> GM11000	Beta	Murcha bacteriana	Batateiros/Bananeiras/Tomateiros	Completo	(SALANOUBAT <i>et al.</i> , 2002)
<i>Ralstonia solanacearum</i> MolK2	Beta	Murcha bacteriana	Bananeiras	Em andamento	GENOSCOPE
<i>Ralstonia solanacearum</i> race 3 biovar 2	Beta	Murcha bacteriana	Batateiro/Bananeiras/Tomateiros	Em andamento	(GABRIEL <i>et al.</i> , 2006)
<i>Aster yellows witches' broom phytoplasma</i> AY-WB	Firmicutes	<i>Aster yellows witches' broom</i>	Plantas de milho	Em andamento	OSU
<i>Phytoplasma asteris</i> cepa OY	Firmicutes	Amarelão das cebolas	Cebola	Completo	(OSHIMA <i>et al.</i> , 2004)
<i>Spiroplasma kunkelii</i>	Firmicutes	Enfezamento do milho	Plantas de Milho	Em andamento	USDA/UO
<i>Spiroplasma citri</i>	Firmicutes	<i>Stubborn</i> dos citros	Plantas de citros	Em andamento	GENAMICS
<i>Western X phytoplasma</i>	Firmicutes	<i>Western X disease</i>	Pessegueiras e cerejeiras	Em andamento	IG
<i>Erwinia amylovora</i> Ea273	Gamma	Fire Blight	Macieira/Pereira	Em andamento	SANGER
<i>Erwinia carotovora atroseptica</i>	Gamma	Talo oco dos batateiros	Batateiro	Completo	(BELL <i>et al.</i> , 2004)
<i>Erwinia chrysanthemi</i> 3937	Gamma	Podridão da raiz em diversas plantas	Diversas plantas	Em andamento	WISC/TIGR
<i>Pseudomonas syringae</i> pv. <i>phaseolicola</i> 1448A	Gamma	Podridão do feijoeiro	Feijoeiro	Completo	(JOURNET <i>et al.</i> , 2001)
<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	Gamma	Queima bacteriana do feijão	Feijoeiro	Completo	(FEIL <i>et al.</i> , 2005)
<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000	Gamma	Mancha bacteriana pequena	Tomateiros/ <i>Arabidopsis</i>	Completo	(BUELL <i>et al.</i> , 2003)
<i>Xanthomonas albilineans</i> GPE PC73	Gamma	Escaldadura das folhas	Cana-de-açúcar	Em andamento	GENOSCOPE
<i>Xanthomonas axonopodis aurantifolii</i> B	Gamma	Cancrose B	Plantas de citros	Em andamento	ONSA
<i>Xanthomonas axonopodis aurantifolii</i> C	Gamma	Cancrose C	Plantas de citros	Em andamento	ONSA
<i>Xanthomonas axonopodis citri</i>	Gamma	Cancrose A	Plantas de citros	Completo	(DA SILVA <i>et al.</i> , 2002)
<i>Xanthomonas campestris campestris</i> 8004	Gamma	Podridão negra das crucíferas	<i>Brassica</i> e <i>Arabidopsis</i>	Completo	(QIAN <i>et al.</i> , 2005)
<i>Xanthomonas campestris campestris</i> ATCC 33913	Gamma	Podridão negra das crucíferas	<i>Brassica</i> e <i>Arabidopsis</i>	Completo	(DA SILVA <i>et al.</i> , 2002)
<i>Xanthomonas campestris campestris</i> B100	Gamma	Podridão negra das crucíferas	<i>Brassica</i> e <i>Arabidopsis</i>	Em andamento	UB (VORHOLTER <i>et al.</i> , 2003)
<i>Xanthomonas campestris</i> pv. <i>vesicatoria</i>	Gamma	Mancha bacteriana	Pimenteiro e Tomateiro	Completo	(THIEME <i>et al.</i> , 2005)
<i>Xanthomonas oryzae oryzae</i>	Gamma	Estria bacteriana do arroz	Plantas de arroz	Completo	(LEE <i>et al.</i> , 2005)
<i>Xanthomonas oryzae oryzicola</i>	Gamma	Doença do traço no arroz	Plantas de arroz	Completo	(ZHAO <i>et al.</i> , 2005)
<i>Xanthomonas oryzae oryzae</i> MAFF 311018	Gamma	Estria bacteriana do arroz I	Plantas de Arroz	Completo	NIAS
<i>Xanthomonas smithii</i> subsp. <i>citri</i>	Gamma	Cancro cítrico	Plantas de citros	Em andamento	NCBI
<i>Xylella fastidiosa</i> 9a5c	Gamma	Clorose variegada dos citros	Plantas de citros	Completo	(SIMPSON <i>et al.</i> , 2000)
<i>Xylella fastidiosa</i> Ann 1	Gamma	Escaldadura da espirradeira	Espirradeira	Em andamento	JGI
<i>Xylella fastidiosa</i> Dixon	Gamma	Escaldadura da amendoeira	Amendoeira	Em andamento	JGI
<i>Xylella fastidiosa</i> Temecula	Gamma	Doença de Pierce	Videiras	Completo	(VAN SLUYS <i>et al.</i> , 2003)
<i>Clavibacter michiganensis</i> subsp. <i>sepedonicus</i>	High GC Gram +	Cancro bacteriano do batateiro	Batateiros	Em andamento	SANGER
<i>Clavibacter michiganensis</i> subsp. <i>michiganensis</i>	High GC Gram +	Cancro bacteriano do tomateiro	Tomateiros	Em andamento	UB
<i>Leifsonia xyli xyli</i>	High GC Gram +	Raquitismo da soqueira	Cana-de-açúcar	Completo	(MONTEIRO-VITORELLO <i>et al.</i> , 2004)

Tabela 2: ^a As referências indicam os artigos publicados que descreveram os respectivos genomas ou as organizações responsáveis pelo seqüenciamento, cujos sítios estão indicados a seguir:

AGRO: A public web resource for the Agrobacterium research community: <http://www.agrobacterium.org>;

JGI: Joint Genome Institute: <http://www.jgi.doe.gov>;

GGL: Göttingen Genomics Laboratory: http://www.g2l.bio.uni-goettingen.de/projects/f_projects.html

GENOSCOPE: Centre National de Séquençage – France: http://www.genoscope.cns.fr/externe/English/Projets/Projet_Y/organisme_Y.html.

OSU: Iowa State University Department of Plant Pathology: <http://www.public.iastate.edu/~ajbog>

USDA: United States Department of Agricultural <http://www.ba.ars.usda.gov/mppl/research/kunkelii.html>;

UO: University of Oklahoma: Advanced Center for Genome Technology: <http://www.genome.ou.edu/spiro.html>;

GENAMICS: Genamics:research From Your Desktop:

<http://genamics.com/cgi-bin/genamics/genomes/genomesearch.cgi?field=ID&query=1658>

IG: Integrated Genomic: <http://www.integratedgenomics.com/>;

SANGER: Sanger Institute – Bacterial genomes: <http://www.sanger.ac.uk/Projects/Microbes>;

WISC: Erwinia chrysanthemi 3937 Genome Project - Wisconsin University: <http://www.ahabs.wisc.edu/~pernalab/erwinia/index.html>;

TIGR: The Institute for Genomic Research: <http://www.tigr.org>;

ONSA: Organization for Nucleotide Sequencing and Analysis – FAPESP: <http://genoma4.iq.usp.br/xanthomonas>;

UB: Universität Bielefeld: <http://www.genetik.uni-bielefeld.de/>;

NIAS: National Institute of Agrobiological Sciences: <http://microbe.dna.affrc.go.jp/Xanthomonas/>

NCBI: National Center for Biotechnology Information: <http://www.ncbi.nlm.nih.gov/>

Tabela 3: Características gerais dos 12 genomas da família *Xanthomonadaceae* já seqüenciados.

Composição	Genomas											
	XF 9a5c	XF Dixon	XF Ann-1	XF PD	XAC 306	XCC ATCC	XCC B100	XCC 8004	XOO	XCV	XAA ^c B	XAA ^c C
Tamanho (Mpb) (C+P*) ^a	2.73	2.43	2.63	2.52	5.3	5.1	nd	5.2	5.0	5.4	4.3	4.3
Conteúdo de GC (%) (C+P*)	53	52	52	52	65	65	nd	65	64	60	63	63
Seqüências codificadoras (C+P*)	2848	2681	2870	2066 ^b	4428	4182	3076	4273	4637	4726	4225	4247
Com função putativa	1314	1593	1713	1362	2770	2708	2992	2671	3340	3080	nd	nd
Hipotéticos	1534	1088	1157	704	1658	1474	84	1602	1297	1646	nd	nd
Operon de rRNAs	2	2	2	2	2	2	2	2	2	2	2	2
RNA transportadores (rRNA)	49	47	48	49	54	53	19	53	54	56	52	52
Plasmídeo	2	0	0	1	2	0	nd	0	0	4	2	2
Referência Bibliográfica	(1)	(2,3)	(2,3)	(4)	(5)	(5)	(6)	(7)	(8)	(9)	(10)	(10)

^aC+P*, resultado das informações relativas ao cromossomo principal e plasmídeos. ^bValor baseado em uma nova anotação que eliminou alguns genes previamente anotados no genoma de XF-CVC. nd, informações não disponíveis no referido artigo ou *web site* oficial. ^cDados não publicados. Referências Bibliográficas: (1) - (SIMPSON *et al.*, 2000); (2,3) - (BHATTACHARYYA *et al.*, 2002a, BHATTACHARYYA *et al.*, 2002b); (4) - (VAN SLUYS *et al.*, 2003); (5) - (DA SILVA *et al.*, 2002); (6) - (VORHOLTER *et al.*, 2003); (7) - (QIAN *et al.*, 2005); (8) - (LEE *et al.*, 2005); (9) - (THIEME *et al.*, 2005); (10) – (MOREIRA *et al.*, manuscrito em preparação).

2. OBJETIVOS

Como apresentado na Introdução, bactérias do gênero *Xanthomonas* constituem fitopatógenos importantes, responsáveis por doenças em diversas plantas de interesse sócio-econômico. Entre estas doenças está o cancro cítrico, cujo agente causal é *Xanthomonas axonopodis* pv. *citri* (**XAC**). Nesta tese de doutorado tivemos como meta principal aprofundar os conhecimentos relativos à biologia e aos mecanismos de patogenicidade desta bactéria. Para tal, propusemos os quatro objetivos descritos a seguir:

1. Realizar uma análise comparativa dos genomas completos das bactérias XAC, XCC, XF-CVC e XF-PD.

Tendo em vista que XAC e *Xylella fastidiosa* cepa 9a5c (XF-CVC) são fitopatógenos de um mesmo hospedeiro (*Citrus sinensis*) propuzemos a realização de uma análise comparativa de seus genomas já completamente seqüenciados (SIMPSON *et al.*, 2000, DA SILVA *et al.*, 2002). Este objetivo foi expandido pela incorporação de dois outros genomas, *Xanthomonas campestris* pv. *campestris* (XCC) (DA SILVA *et al.*, 2002) e *Xylella fastidiosa* cepa temecula (XF-PD) (VAN SLUYS *et al.*, 2003). A comparação entre quatro genomas de bactérias de uma mesma família e de dois gêneros distintos, até o momento, não havia sido realizada. Os resultados desta parte do trabalho estão detalhados na seção 4.1.

2. Analisar o sistema de osmoproteção e osmoregulação e vias de degradação de compostos aromáticos codificados no genoma de XAC.

Os sistemas relacionados aos processos de osmoproteção e osmoregulação não haviam sido estudados em bactérias do gênero *Xanthomonas*. Assim, propusemos a identificação dos componentes que compõem este sistema a partir da análise do genoma de XAC. Paralelamente, analisamos os componentes das possíveis vias de degradação de compostos

aromáticos, uma possível resposta adaptativa ao processo de interação da bactéria com seu hospedeiro. Estas vias também não haviam sido analisadas em bactérias do gênero *Xanthomonas*. Resultados desta parte do trabalho estão descritos na seção **4.2**.

3. Planejar, construir e validar microarranjos de DNA em lâminas de vidro, a partir de um conjunto selecionado de clones oriundos do projeto de seqüenciamento do genoma de XAC (XACarray).

Para o seqüenciamento de XAC foram geradas bibliotecas de clones genômicos, representando 98% do genoma com uma cobertura de sete vezes (DA SILVA *et al.*, 2002). Tendo em vista a disponibilidade destes clones, propusemos sua utilização para a construção de microarranjos de DNA em lâminas de vidro (XACarray), alternativamente à amplificação de regiões genômicas de interesse utilizando-se pares distintos de oligonucleotídeos. Estes microarranjos de DNA representativos de uma fração significativa do genoma de XAC estariam disponíveis para análise da expressão gênica de XAC submetida a crescimento em diferentes condições. Detalhes sobre a preparação e validação do XACarray estão apresentados na seção **4.3**.

4. Verificar o perfil de expressão de alguns genes de XAC em resposta ao seu crescimento em diferentes concentrações de glicose

Paralelamente à construção e validação do **XACarray**, propusemos uma análise do perfil de expressão dos genes relacionados com crescimento de XAC em diferentes concentrações de glicose. Estudos anteriores relataram que este açúcar interfere positivamente na produção de goma xanthana em *Xanthomonas* (VOJNOV *et al.*, 2001). Assim, pretendemos avaliar o perfil de expressão de genes relacionados à biossíntese do polissacarídeo extracelular (EPS) bem como a sua regulação (genes *rpf*) através da metodologia de RT-qPCR. Os resultados deste estudo estão apresentados na seção **4.4**.

3. METODOLOGIAS

3.1. Metodologias para análise in silico de genomas

3.1.1. Agrupamento de genes em famílias

O agrupamento de genes em famílias foi feito em três etapas utilizando o programa BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL *et al.*, 1990) para comparação entre todas as seqüências de aminoácidos deduzidas a partir dos genes ou CDS anotados nos genomas das quatro bactérias (XAC, XCC, XF-CVC e XF-PD). Inicialmente selecionamos todos os pares de genes que apresentaram o melhor resultado de comparação bidirecional (BBH, do inglês *Bi-directional Best Hit*), cujos *e-values* foram $\leq 10^{-20}$. Em seguida, uma seqüência (*g*) era adicionada a uma família de tamanho *n* já existente, se seu primeiro resultado de BLAST (*k*) apresentasse similaridade com algum gene desta família. Os parâmetros usados para esta determinação foram $k=0,8n$, sendo 0,8 um valor definido empiricamente. Por último, as famílias f_1 e f_2 foram agrupadas, obedecendo ao critério de que duas famílias só seriam unidas se cada uma tivesse pelo menos 80% de seus genes com alto grau de similaridade com a outra família. Estes critérios são similares aos usados por (DIGIAMPIETRI *et al.*, 2003).

Classificamos um gene (*g*) como exclusivo de um determinado genoma (*x*), se a seqüência que o codifica é única na análise comparativa ou se ela pertence a uma família gênica que não inclui membros de outros genomas (família de parálogos). Esta definição foi também aplicada para declarar genes ou famílias gênicas exclusivas de dois ou mais genomas em relação aos genomas definidos como referência. O grupo de genes ou famílias compartilhadas em dois ou mais genomas (ortólogos) compreende todos os genes que apresentam pelo menos um exemplar em cada genoma de referência. O conjunto de genes parálogos de um genoma *x* inclui genes de uma família com pelo menos dois genes do mesmo genoma.

Para automatização destas análises foram utilizados programas e *plugins* escritos pelo Prof. Dr. João Carlos Setúbal e pelo aluno de Doutorado Luciano A. Digiampietri da UNICAMP. Em alguns casos, famílias gênicas foram definidas após checagem manual da qualidade dos alinhamentos.

3.1.2. Análise filogenética

A seleção dos genes homólogos para análise filogenética foi realizada através de diferentes metodologias para genes isolados ou concatenados. Os genes isolados foram comparados com seqüências do banco de dados UNIPROT (APWEILER *et al.*, 2000, APWEILER *et al.*, 2004) usando o programa BLAST (ALTSCHUL *et al.*, 1990), e as seqüências com diferentes graus de identidade foram então selecionadas com base no *e-value* de 10^{-5} e uma cobertura mínima de 60%. Alinhamentos de seqüências concatenadas foram construídos para prováveis *operons*, selecionando-se para cada gene de um agrupamento gênico bem definido de uma certa bactéria, geralmente XAC e XF-CVC, as seqüências homólogas obtidas dos bancos de dados KEGG (KANEHISA & GOTO, 2000) e COG (TATUSOV *et al.*, 2000). Alguns genes e organismos foram excluídos da análise comparativa por não apresentarem genes importantes para os alinhamentos finais.

Uma vez definido o grupo de seqüências homólogas para cada gene, estas foram submetidas a um alinhamento múltiplo usando o programa CLUSTAL-W versão 1.74 (THOMPSON *et al.*, 1994), com parâmetros padrão. As regiões conservadas foram selecionadas usando o programa GBLOCKS (CASTRESANA, 2000). No caso de *operons* ou agrupamentos gênicos, as seqüências foram concatenadas nesta etapa. Todos os alinhamentos foram então analisados com o programa PROTML (ADACHI & HASEGAWA, 1996) para inferências de máxima verossimilhança na determinação da filogenia, a partir do modelo de Poisson como matriz de substituição de aminoácidos.

3.1.3. Classificação das integrases e das ilhas prováveis de inserção/deleção (PinDels)

Inserções de fagos e de transposases são conhecidos mediadores de rearranjos genômicos e de eventos de transferência genética lateral. Por esta razão definimos como regiões PinDels (*Putative Insertion and Deletion regions*) as regiões flanqueadas por integrases de fagos e ricas em genes hipotéticos ou relacionados a fagos. Também consideramos como PinDels, as regiões não flanqueadas por integrases, mas com evidente variação na composição de nucleotídeos GC e na frequência de códons, desde que também fossem ricas em genes hipotéticos ou relacionados a fagos. As análises da variação no conteúdo GC e na frequência de códons foi realizada de acordo com a metodologia de Karlin (KARLIN, 1998, KARLIN, 2001).

A análise acima foi complementada pela classificação das integrases segundo suas similaridades de seqüências e sítios de integração. De acordo com a similaridade de seqüências, foram definidos os grupos I, II, III e IV além de um grupo adicional formado pelos genes *xerC* e *xerD* que codificam para recombinases específicas. Os grupos A, B e C de integrases foram definidos de acordo com os diferentes sítios de integração. As integrases do grupo A sempre flanqueiam, a montante ou jusante, genes que codificam tRNAs; integrases do grupo B sempre flanqueiam obrigatoriamente PinDels putativas, mas nunca genes de tRNAs, e as integrases do grupo C não flanqueiam tRNAs e se localizam no interior de uma PinDel.

A determinação do provável término de replicação foi baseada no valor máximo de assimetria na frequência de bases acumulada nas duas fitas do DNA cromossomal (LOBRY, 1996, LOBRY & SUEOKA, 2002).

3.1.4. Análise de vias metabólicas

Para definição e análise de vias metabólicas utilizamos dados da literatura pertinente e informações extraídas diretamente dos bancos de genes e genomas disponibilizados nos sítios dos projetos genoma de XAC

(<http://genoma4.iq.usp.br/>) e XF-CVC (<http://aeg.lbi.ic.unicamp.br/xf/>) e de bancos de dados públicos tais como:

- NCBI (*National Center for Biotechnology Information*, www.ncbi.nlm.nih.gov)
- KEGG (*Kyoto Encyclopedia of Genes and Genomes*, www.genome.jp/kegg/)
- COG (*Clusters of Orthologous Groups*, www.ncbi.nlm.nih.gov/COG/)
- Swiss-Prot (www.ebi.ac.uk/swissprot/)
- PFAM (www.sanger.ac.uk/Software/Pfam/)

Para pesquisa de similaridade de seqüências de proteínas utilizamos a ferramenta BLAST (ALTSCHUL *et al.*, 1990) e para o alinhamento de seqüências de aminoácidos utilizamos o programa ClustalW (THOMPSON *et al.*, 1994, CHENNA *et al.*, 2003).

3.2. Cultivo e manipulação de *Xanthomonas axonopodis* pv. *citri*

3.2.1. Cultivo de XAC

As cepas selvagem de XAC isolado 306 ou mutantes (selecionados do banco de mutantes gerado por inserção de transposon) foram extraídas a partir de folhas infectadas de *Citrus limonia* (limão cravo) e mantidas em glicerol 30% a -80°C, seguindo-se semeadura em meio sólido NA (3 g de extrato de carne, 5 g de peptona, 5 g de cloreto de sódio, 15 g de ágar) e crescimento a 28°C por 3 dias. O crescimento das cepas mutantes foi sempre realizado em meio de cultura contendo 100 µg/mL de canamicina. Após pré-cultivo em meio sólido, as bactérias foram transferidas para 20 mL de meio líquido NA ou meio XDM₂ (LEMOS *et al.*, 2003) contendo 50 mM de glicose e incubadas a 28°C sob agitação de 220 rpm por 24h. Em seguida, 1 mL desta cultura foi transferida para 100 mL de meio NA ou XDM₂, seguindo-se nas mesmas condições anteriores. Em alguns experimentos a concentração de glicose no meio XDM₂ foi de 10 mM ou 250 mM. Alternativamente, as cepas selvagem ou mutantes foram crescidas em meio sólido TSA (5 g de Triptona, 5 g de Sacarose, 1 g/L de Glutamato de sódio e 7,5 g de Ágar) por 72 horas a 28°C e em seguida transferidas para 50 mL de meio TSA líquido. Após cultivo a 28°C por 24h, sob

agitação de 200 rpm, 1 mL da cultura foi transferido 50 mL de meio de cultura TSA líquido fresco, seguindo-se incubação nas mesmas condições.

3.2.2. Inoculação de XAC em folhas de citros

Após crescimento em meio TSA como descrito acima, a suspensão bacteriana das cepas selvagem ou mutante foi ajustada para $DO_{600nm} = 0,01$ (aproximadamente 10^4 UFC/mL) e 50 uL desta suspensão foram infiltrados em folhas de *Citrus limonia*. A quantificação do número de bactérias por área foi determinada retirando-se três discos de 1 cm^2 de cada folha inoculada, nos tempos 0, 2, 4, 6, 8 e 10 dias após a inoculação. Os discos foliares foram macerados em 1 mL de água estéril, com a ajuda de um pistilo, e foram feitas diluições seriadas de 10^{-1} até 10^{-7} . Dez μL de cada uma das diluições seriadas foram semeadas em meio de cultura TSA sólido contendo canamicina, 100 $\mu\text{g/mL}$ (exceto para o selvagem). As placas foram mantidas a 28°C por dois dias, quando as colônias isoladas foram contadas para cada uma das diluições. Alternativamente, cultivos das cepas selvagem ou mutantes obtidos em meio TSA, como descrito acima, tiveram a concentração ajustada para $DO_{600nm} = 0,3$ (10^8 UFC/mL) antes de serem infiltrados na parte abaxial da mesma folha. Após três dias da inoculação, as folhas foram fotografadas com uma câmera Sony DSC-S75/S85 (4,1 megapixels de resolução) para documentação dos sintomas apresentados. Os experimentos de infecção de plantas foram realizados no laboratório dos Profs. Drs. Julio César de Oliveira e Jesus A. Ferro (UNESP/Jaboticabal), o qual é adequado a manipulação segura de XAC.

3.3. Isolamento de RNA total de XAC e ensaios de RT-qPCR

RNA total de XAC foi purificado pelo método do TRIzol™ (Invitrogen) de acordo com instruções do fabricante, em geral a partir de 50 mL de cultura de XAC. O RNA total isolado foi tratado com RQ1 DNase livre de RNAase (Promega) para remoção de DNA genômico e 5 μg de RNA total livre de DNA foi utilizado na síntese da primeira fita de cDNA utilizando-se 200 U da enzima transcriptase reversa Superscript II™ (Invitrogen) e 500 μg de oligonucleotídeos

nonaméricos aleatórios. O cDNA obtido foi diluído com água livre de nuclease para 35 µg/µL e armazenado a -20°C até o uso.

PCR quantitativo (qPCR) foi realizado no equipamento ABI PRISM 5700 *Sequence Detection System*TM (Applied Biosystems) utilizando-se parâmetros padrão. A mistura para qPCR incluía 10 µL de Platinum[®] SYBR[®] Green qPCR SuperMix UDG (Invitrogen), 800 nM de cada oligonucleotídeo e 180 ng de cDNA molde. Para confirmar a geração de produtos específicos, a PCR foi imediatamente seguida por análise da curva de fusão do produto de PCR, de acordo com as recomendações do fabricante.

Os oligonucleotídeos específicos para cada CDS foram planejados com o programa PRIMER EXPRESS 2.0 (Applied Biosystems). As seqüências dos oligonucleotídeos utilizados estão apresentadas na Tabela 4. A CDS XAC2704 (*NADH-ubiquinone oxidoreductase, NQO7 subunit*) foi utilizada como controle para normalizar a quantidade de RNA total por amostra. A razão de expressão em cada amostra, comparada a condição de referência, foi calculada pelo método $2^{-\Delta\Delta CT}$, como descrito em (LIVAK & SCHMITTGEN, 2001) a partir de dois experimentos independentes. Nós consideramos uma CDS como diferencialmente expressa em relação a condição de referência, quando a razão obtida descontada de dois desvios padrões das médias das triplicatas é diferente da obtida para a mesma CDS na condição de referência.

Tabela 4: Oligonucleotídeos planejados para as reações de RT-qPCR

ID ^a	Gene	Produto	Oligonucleotídeo F (5' - 3')	Oligonucleotídeo R (5' - 3')
XAC0756	kdpA	<i>Potassium-transporting ATPase, A chain</i>	GCCACACAAGCGTCGATTG	ACTTCCGGCTCCGAAATCTC
XAC1878	rpfC	<i>RpfC protein</i>	GGCCCGATAGCGAACATG	TGCGGCAACAACACATACG
XAC1879	rpfF	<i>RpfF protein</i>	TGGGCGCATATTCCTTCCT	GCGTCCTTCGAGAATGATTTTC
XAC1880	rpfB	<i>RpfB protein</i>	CACTTCGCTGCTTACCTGCTT	GGCAGTTGGGCATCATCAG
XAC2576	gumK	<i>GumK protein</i>	GCCAGCCGTGACAATGTCTA	GACGGGTCGCCGAGTTG
XAC2583	gumD	<i>GumD protein</i>	GGCGAGACCCCGAACT	GACCAACGGCGGATGTAGTC
XAC2704	nuoA	<i>NADH-ubiquinone oxidoreductase, NQO7 subunit</i>	TTTGCCGAGTCTGCTGTTTCT	GGAATCGACCGACCAACATC
XAC2975	ptsK	<i>HPK kinase/phosphatase</i>	CGATCAGCAAGAACCAGTCATG	AGATCCACAGCGGCGTATTG
XAC3345	pykA	<i>Pyruvate kinase type II</i>	CCGGGCCTGCAGAAGAA	ATCTGCGTGGCGGTGATC

^a ID – identificação do gene de acordo com DA SILVA e colaboradores (2002).

3.4. Construção dos microarranjos de DNA de XAC (XACArray)

3.4.1. Seleção dos clones da biblioteca genômica e amplificação dos insertos por PCR

Os clones que compunham a biblioteca genômica de XAC foram gerados usando-se a metodologia de *shotgun* em pUC18 com insertos apresentando entre 1 e 4 Kpb de tamanho. Os insertos dos clones selecionados foram amplificados por PCR após extração do DNA pela técnica de *boiling prep*. As reações foram realizadas com uma desnaturação inicial a 95°C por 5 min, seguida de 40 ciclos de desnaturação a 95°C por 45 s, anelamento a 50°C por 30 s e extensão a 72°C por 1 min. Os produtos da amplificação/reamplificação foram purificados em placas de filtração Multiscreen da Millipore (cat. # MAFB NOB 50) e eluídos em 50 µL de Tris 10 mM pH8,0, diretamente em placas de 96 poços, as quais foram combinadas manualmente em placas de 384 poços, que é o formato de placa para carregar o *spotter*.

Em todos os passos os produtos de PCR foram avaliados por eletroforese em gel de agarose 1% corado com brometo de etídio (AUSUBEL *et al.*, 1990). O tamanho e quantidade dos *amplicons* foram estimados pela comparação com um padrão (fragmentos de DNA) em quantidade e tamanho conhecidos. Este padrão com fragmentos de 1697 pb, 689 pb e 297 pb foi gerado pela digestão do plasmídeo pUC19 com as enzimas de restrição *HindIII*, *NdeI* e *ScaI*. A Figura 2 mostra eletroforese em gel de agarose 1% de produtos de PCR para alguns clones obtidos após a etapa de purificação, para exemplificar a quantidade e pureza dos produtos de PCR utilizados na confecção dos microarranjos.

Figura 2: Eletroforese em gel de agarose 1% dos produtos de PCR para alguns clones da biblioteca de shotgun após a etapa de purificação. Na primeira canaleta de cada gel foi aplicado marcador de tamanho correspondente a 2683 pb (plasmídeo pUC19 linearizado), 1697 pb, 689 pb e 297 pb, de cima para baixo.

3.4.2. Preparação dos microarranjos de DNA

Aos produtos de PCR purificados, acondicionados nas placas de 384 poços, adicionamos igual volume de DMSO para que a concentração final fosse entre 200-400 fmol/ μ L de DNA. Em seguida o DNA foi depositado em lâminas espelhadas tipo 7*, utilizando-se o *Generation III Microarrays Spotter*TM (Amersham Biosciences/GE Healthcare) no Laboratório de *Microarrays* do Departamento de Bioquímica do IQ-USP. Este *spotter* permite a deposição de até 4608 amostras de DNA organizadas em 12 subconjuntos (*subarrays*) de 384 pontos (12x32). O conjunto de 4608 é depositado em duplicata nas duas metades longitudinais da lâmina, configurando o que chamamos de *set A* e *set B*. Após a deposição do DNA, as lâminas foram submetidas a 50mJ de luz UV, fixação covalente do DNA, e em seguida armazenadas em dessecador (umidade relativa ~5%).

3.4.3. Preparação de DNA marcado e hibridação dos microarranjos

DNA dupla fita marcado com fluoróforos foi gerado pela incorporação de Cy3-dCTP ou Cy5-dCTP (Amersham Biosciences/GE Healthcare) durante a polimerização a partir de 2 μ g de DNA genômico de XAC fragmentado ~14 vezes em seringa de insulina. A reação foi realizada com DNA-polimerase Klenow Fragment (Gibco) de alta concentração (40 U/ μ L) e 500 μ g de iniciadores não-números aleatórios (KOIDE *et al.*, 2004)². Em seguida, o DNA marcado foi purificado em placas de filtração Multiscreen (Millipore) e o fluoróforo total incorporado foi quantificado medindo-se absorvância em 550 nm (Cy3) e em 650 nm (Cy5).

Utilizamos quantidades equivalentes de DNA fluorescente nas hibridizações dos microarranjos, as quais foram realizadas em tampão de hibridação (Amersham Biosciences/GE Healthcare) contendo 50% formamida

² 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. & Gomes, S. L. (2004). *DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of Xylella fastidiosa delineate genes important for bacterial virulence*. **J. Bacteriol.** 186, 5442-9. (ANEXO 2).

por ~16 horas a 42°C. Após algumas lavagens com SSC/SDS (1x/0,2% e 0,1x/0,2%), as lâminas foram secas sob gás nitrogênio e analisadas no *Generation III Scanner* (Amersham Biosciences/GE Healthcare) para obtenção das imagens que refletem a intensidade de fluorescências dos pontos do microarranjo.

3.4.4. Detecção, quantificação e normalização dos sinais de fluorescência

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 4 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. Em seguida, para cada ponto foi calculada a razão das intensidades de sinal dos canais (ICy3 e ICy5) e 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 $M = \log_2(ICy5 / ICy3)$ e $S = \frac{1}{2} * \log_2(ICy5 + ICy3)$), para cada lâmina. Neste espaço, as razões foram normalizadas usando um algoritmo de regressão linear com peso local (Lowess), implementado usando o pacote R (<http://www.r-project.org/>) para corrigir influências sistemáticas nos dados, originadas de pequenas diferenças na marcação e/ou eficiência de detecção entre os corantes fluorescentes (KOIDE *et al.*, 2004).

4. RESULTADOS E DISCUSSÃO

4.1. Análise comparativa dos genomas completos das bactérias XAC, XCC, XF-CVC e XF-PD

Os dados que descrevemos a seguir resultam da análise computacional de seqüências genômicas de quatro fitobactérias da família Xanthomonadaceae, duas do gênero *Xanthomonas* (XAC e XCC) e duas do gênero *Xylella* (XF-CVC e XF-PD). Vale lembrar que genomas completos de XAC e XCC já haviam sido comparados entre si (DA SILVA *et al.*, 2002). Posteriormente estes dois genomas foram novamente comparados em uma análise com outros seis genomas de bactérias associadas a plantas, incluindo XF-CVC (Van Sluys *et al.*, 2002). A análise comparativa entre os genomas de XF-CVC e XF-PD bem como de XF-CVC com XF-Ann1 e XF-Dixon também já foram reportadas (VAN SLUYS *et al.*, 2003).

As características gerais dos genomas das quatro bactérias (XAC, XCC, XF-CVC e XF-PD) estão sumarizadas na Tabela 5. Inicialmente fizemos uma avaliação da complexidade destes organismos considerando seus potenciais metabólicos, a partir de uma análise da composição gênica baseada nas categorias funcionais de anotação. Paralelamente, as análises computacionais revelaram genes exclusivos de cada uma destas bactérias, os quais poderiam estar envolvidos com adaptação aos seus hospedeiros. Estes genes exclusivos foram então classificados em famílias, com base em suas similaridades de seqüências. Os 3636 genes exclusivos dos genomas de ambas as cepas de *Xanthomonas* foram agrupados em 1470 famílias gênicas, enquanto que os 1026 genes exclusivos dos genomas de ambas as cepas de *Xylella* foram agrupados em 375 famílias gênicas distintas. Em seguida, identificamos os genes, ou conjuntos de genes, associados a uma mesma função em cada um dos gêneros ou entre gêneros e que eram mais representados. Com esta abordagem foi identificada uma IP putativa, nomeada como SPI-7 a qual é exclusiva de XAC e XF-CVC, mas previamente descrita em *Salmonella enterica* (PICKARD *et al.*, 2003). Nas duas bactérias do gênero *Xanthomonas* foi possível determinar não apenas genes exclusivos, mas também regiões exclusivas, as

quais foram minuciosamente analisadas. Entre os genes mais representados do gênero *Xanthomonas* se destacaram os que codificam para enzimas de degradação de parede celular vegetal, proteases, receptores de ferro e sideróforos, sistemas secretórios tipo II e III, além de genes envolvidos com fotoproteção e com síntese e regulação de flagelo. Por outro lado, genes envolvidos com síntese e secreção de proteínas do tipo colicinas e diferentes cópias de genes relacionados à síntese de *pili* estão mais representados nos genomas das cepas de *Xylella*. Finalmente, mapeamos em cada genoma as prováveis ilhas de transferência lateral com base na presença de regiões classificadas como **PInDels** (**P**utative **I**nsertion and **D**eletion regions), observando uma correlação entre a sequência da integrase e possíveis sítios de integração.

A seguir descrevemos os resultados da análise comparativa entre estes quatro genomas (MOREIRA *et al.*, 2004, MOREIRA *et al.*, 2005)³.

³ **Moreira, L. M.**, de Souza, R. F., Almeida, N. F., Jr., Setubal, J. C., Oliveira, J. C., Furlan, L. R., Ferro, J. A. & da Silva, A. C. (2004). *Comparative genomics analyses of citrus-associated bacteria*. **Annu Rev Phytopathol.** 42, 163-84. (**ANEXO 3**)
Moreira, L. M., de Souza, R. F., Digiampietri, L. A., Da Silva, A. C. & Setubal, J. C. (2005). *Comparative analyses of Xanthomonas and Xylella complete genomes*. **Omics.** 9, 43-76. (**ANEXO 4**)

Tabela 5: Características gerais e genômicas de XF-CVC, XF-PD, XAC e XCC.

	XF-CVC	XF-PD	XAC	XCC
CARACTERÍSTICAS GERAIS				
Hospedeiros compatíveis	<i>Citrus spp</i>	Videiras	<i>Citrus spp</i>	<i>Brassica spp</i>
Doença	Clorose Variegada dos Citrus (CVC)	Doença de Pierce (PD)	Cancro cítrico (CC)	Podridão negra das crucíferas (BR)
Tecido infectado	Xilema	Xilema	Mesófilo	Mesófilo e tecidos vasculares
Sintomas	Áreas cloróticas na superfície superior da folha geralmente seguidas de lesões em tom marrom. Os frutos apresentam redução de tamanho e consistência pétrea.	Lesões foliares geralmente acompanhadas de halos cloróticos, com dano aos pecíolos e redução no tamanho do fruto.	Lesões cancróticas seguidas de abscisão foliar e de frutos, geralmente culminando para o declínio do vegetal.	Clorose nas bordas foliares associada com escurecimento dos tecidos vasculares, extensiva morte celular e necrose.
Forma de transmissão - Vetor (es)	Insetos sugadores (cigarrinhas).	Insetos sugadores (cigarrinhas).	Animais, condições ambientais e utensílios agrícolas.	Animais, condições ambientais e utensílios agrícolas.
CARACTERÍSTICAS GENÔMICAS				
Seqüenciamento do genoma	Completo	Completo	Completo	Completo
Referência	(SIMPSON <i>et al.</i> , 2000) http://aeg.lbi.ic.unicamp.br/xf/	(VAN SLUYS <i>et al.</i> , 2003) http://aeg.lbi.ic.unicamp.br/world/xfpd/	(DA SILVA <i>et al.</i> , 2002) http://genoma4.iq.usp.br	(DA SILVA <i>et al.</i> , 2002) http://genoma4.iq.usp.br
Tamanho cromossomo principal	2,679,305 pb	2,519,802 pb	5,175,554 pb	5,076,187 pb
Plasmídeo (CDS) ^a	pXF51 (64) pXF1.3 (2) 2848	pXFPD1.3 (2) 2066	pXAC33 (42) pXAC64 (73) 4313	---- 4182
Número de CDS (cromossomo + plasmídeo)				
Com funções preditas (%)	1314 (58.42%)	1408 (68.15%)	2770 (64.22%)	2708 (64.75%)
Conteúdo médio de GC (%)	52.7 %	51.8 %	64.7 %	65.0 %
Operons de rRNA	2	2	2	2
tRNA (AA) ^b	49 (20)	49 (20)	54 (20)	53 (20)
Proteínas de fago (%)	82 (3.64%)	139 (6.72%)	35 (0.81%)	43 (1.03%)
Transposases (%)	7 (0.31%)	7 (0.33%)	84 (1.94%)	108 (2.58%)
ANÁLISE COMPARATIVA				
Genes exclusivos (%)	283 (12.58%)	104 (5.98%)	665 (15.41%)	555 (13.27%)
Com funções determinadas (%)	87 (30.74%)	74 (71.15%)	207 (31.12%)	215 (38.73%)
Intersecções exclusivas	XF-PD 1026; XAC 47; XCC 8	XF-CVC 1026; XAC 2; XCC 38	XCC 3636; XF-CVC 47; XF-PD 2	XAC 3636; XF-CVC 8; XF-PD 38

^aCDS, número de seqüências codificadoras; ^b AA, número de aminoácidos com tRNA especificado.

4.1.1. Comparação dos quatro genomas de acordo com as categorias funcionais de anotação

Usando como referencial o número de genes nas diferentes categorias funcionais de anotação, os genomas de XAC, XCC, XF-CVC e XF-PD foram comparados como mostrado na Tabela 6. Em quatro das nove categorias funcionais a diferença entre os números de genes dos genomas de membros do gênero *Xanthomonas* é superior ao número total de genes de mesma categoria em membros do gênero *Xylella*, em particular nas categorias de genes de patogenicidade, virulência e adaptação, processos celulares e metabolismo intermediário. Com uma análise mais detalhada das subcategorias envolvidas no metabolismo intermediário, especialmente metabolismo energético, observamos que as diferenças são decorrentes da ausência total de algumas vias metabólicas e ausências específicas de genes fundamentais em outras, tornando-as incompletas nas cepas de *Xylella* (Tabela 7). As bactérias do gênero *Xanthomonas* apresentam via glicolítica, gliconeogênese, ciclo de Krebs, via das pentoses fosfato, biossíntese e degradação de ácidos graxos e via do glioxilato completas (Figura 3A), enquanto que em membros do gênero *Xylella* estão completamente ausentes a via do glioxilato e a via de degradação de ácidos graxos, além de estarem incompletas a gliconeogênese, a via das pentoses e o ciclo da uréia (Figura 3B). Esta complexidade no metabolismo energético de *Xanthomonas* coincide com o fato deste organismo apresentar duas cópias do SS-II, uma série de EDPCV, responsáveis pelo aumento na demanda de açúcares, e um número elevado de transportadores capazes de internalizar estes compostos, como será melhor discutido adiante. Uma outra divergência metabólica interessante entre estes genomas é a presença de dois conjuntos gênicos que codificam para o complexo citocromo oxidase em *Xanthomonas* e apenas um em membros do gênero *Xylella*. Ambas possuem o complexo citocromo O oxidase (bo3) codificado pelos genes *cyoABCDE*, mas apenas em *Xanthomonas* foram encontrados os genes *cydAB* capazes de codificar as proteínas responsáveis pela formação do complexo citocromo d

oxidase (bd). É importante citar que existe um terceiro complexo citocromo oxidase nomeado como cbb3 presente e descrito em outros organismos. Entretanto, nenhum dos quatro membros de ambos os gêneros apresentam genes que codificam para este complexo, o que, teoricamente, os tornariam incapazes de sobreviverem em condições de anaerobiose estrita (INGLEDEW & POOLE, 1984). As diferenças metabólicas, em detrimento da presença de complexos citocromo oxidase diferentes, estão relacionadas com parâmetros cinéticos, de tal maneira que o complexo bo3, mesmo tendo baixa afinidade pelo oxigênio, teria uma alta velocidade máxima (V_{max}) e, portanto, uma maior eficiência em gerar um gradiente de prótons para a síntese de ATP, do que em relação ao complexo bd (SVENSSON-EK *et al.*, 1996). Assim sendo, XAC teria uma maior facilidade de modular sua cadeia respiratória de acordo com a oferta de oxigênio do meio. Em contrapartida, XF-CVC estaria restrita a ambientes com condições de aerobiose bem controladas, como por exemplo, o xilema.

Devido ao seu potencial metabólico, é previsível que membros do gênero *Xanthomonas* também possuam uma complexa gama de genes relacionados à proteção contra danos oxidativos. Para se precaver destes danos, estas bactérias apresentam em sua composição cerca de 40 genes que codificam para oxidoredutases, contra apenas 6 descritos nos genomas de *Xylella*.

É interessante também observar que, embora a diferença no número de genes envolvidos com elementos genéticos móveis entre os genomas seja baixa, há uma diferença muito grande no tipo de elemento presente em cada organismo. Em *Xanthomonas* é notória a presença de elementos de transposição (DA SILVA *et al.*, 2002), enquanto que em *Xylella* fica evidenciada a presença de genes que codificam para proteínas de fagos (SIMPSON *et al.*, 2000).

Tabela 6: Distribuição das CDS dos quatro fitopatógenos por categoria funcional.

Categorias funcionais gerais	Número de CDS			
	XAC ^a	XCC ^a	XF-CVC ^b	XF-PD ^c
I - Metabolismo Intermediário	730	694	226	207
II - Biossíntese de pequenas moléculas	353	354	216	239
III - Metabolismo de macromoléculas	497	511	356	329
IV - Estrutura celular	223	197	128	136
V - Processos celulares	359	380	118	130
VI - Elementos genéticos móveis	149	173	135	161
VII - Patogenicidade, virulência e adaptação	296	285	128	156
VIII - Genes hipotéticos e hipotéticos conservados	1696	1467	1517	658
IX - Genes com categorias indefinidas	135	121	24	48
Total	4438	4182	2848	2066

^a - (DA SILVA *et al.*, 2002) ; ^b - (SIMPSON *et al.*, 2000); ^c - (VAN SLUYS *et al.*, 2003).

Tabela 7: Distribuição das CDS dos quatro fitopatógenos por subcategorias funcionais do metabolismo intermediário.

Metabolismo (Vias principais)	Situação da via ou número de CDS			
	XAC ^a	XCC ^a	XF-CVC ^b	XF-PD ^c
Glicólise	Completa	Completa	Completa	Completa
Gliconeogênese	Completa	Completa	Incompleta	Incompleta
Ciclo de Krebs	Completa	Completa	Completa	Completa
Via das Pentoses	Completa	Completa	Incompleta	Incompleta
Via do Glioxilato	Completa	Completa	Ausente	Ausente
Ciclo da Uréia	Completa	Completa	Incompleta	Incompleta
Biossíntese de ácidos graxos	Completa	Completa	Completa	Completa
Metabolismo de aminoácidos	Completa	Completa	Completa	Completa
Degradação de ácidos graxos	Completa	Completa	Ausente	Ausente
Degradação de poli e oligossacarídeos	44 CDS	43 CDS	4 CDS	5 CDS
Degradação de pequenas moléculas	93 CDS	88 CDS	33 CDS	19 CDS
Funções regulatórias	296 CDS	280 CDS	77 CDS	64 CDS

^a - (DA SILVA *et al.*, 2002) ; ^b - (SIMPSON *et al.*, 2000) ; ^c - (VAN SLUYS *et al.*, 2003).

Figura 3A: Perfil do metabolismo energético e de pequenas moléculas em bactérias do gênero *Xanthomonas*.

Figura 3B: Perfil do metabolismo energético e de pequenas moléculas em bactérias do gênero *Xylella*.

4.1.2. Comparação dos quatro fitopatógenos com base na presença ou ausência de genes

Embora muitos genes exclusivos de XAC, XCC, XF-CVC e XF-PD tenham sido descritos anteriormente (SIMPSON *et al.*, 2000, DA SILVA *et al.*, 2002, VAN SLUYS *et al.*, 2003), uma análise comparativa mais detalhada revelou novos genes com esta característica e que possivelmente estão envolvidos com os fenômenos de adaptação, patogenicidade, virulência ou mesmo especificidade a ao hospedeiro. Estes genes foram classificados em quatro grupos: **1)** genes mais representados⁴ e exclusivos em cada uma das espécies/cepas analisadas; **2)** genes mais representados por gênero; **3)** genes mais representados e exclusivos nos genomas de XAC e XF-CVC e **4)** genes mais representados na família Xanthomonadaceae. Além disso, estes quatro grupos foram subdivididos em: a) genes agrupados, muitas vezes formando uma única unidade transcricional (*operon*), ou b) genes dispersos pelo genoma. A seguir detalhamos cada um destes grupos de genes nos genomas comparados. A Figura 4 sumariza as informações mais importantes resultantes da análise comparativa dos quatro genomas, sendo que a figura foi dividida em linhas e colunas identificadas por números e letras respectivamente, visando facilitar a localização de um dado processo fisiológico ou via metabólica descrita no texto.

4.1.2.1. Genes mais representados e exclusivos em cada uma das espécies/cepas analisadas

Utilizando as metodologias para comparação de genomas completos descritas na seção 3.1, determinamos as famílias e os genes exclusivos em cada uma das espécies/cepas. Os 3636 genes exclusivos dos genomas de ambas as cepas de *Xanthomonas* foram agrupados em 1470 famílias gênicas, enquanto que os 1026 genes exclusivos dos genomas de ambas as cepas de

⁴ Consideramos genes mais representados aqueles presentes em maior número de cópias, seja via comparação intra-gêneros ou inter-gêneros.

Xylella foram agrupados em 375 famílias gênicas distintas. A Tabela 8 relaciona o número de genes e famílias gênicas exclusivos.

A maioria dos genes exclusivos das cepas de *Xylella* (607) foi classificada como hipotética e não será discutida. Por outro lado, em *Xanthomonas* a maioria dos genes exclusivos deste gênero (2046) apresenta uma função provável. A Tabela 9 relaciona as funções atribuídas às famílias e genes exclusivos das duas *Xanthomonas* em comparação com as duas cepas de *Xylella*.

É importante destacar que 425 genes, dos quais 232 exclusivos do genoma de XAC, bem como 162 genes, dos quais 88 exclusivos do genoma de XCC estão inseridos em ITL putativas. Estas ilhas foram determinadas por apresentarem variação na frequência de códons e na composição de nucleotídeos, além de possuírem elementos que classicamente favorecem sua identificação, tais como seqüências plasmidiais, integrases, transposases ou proteínas de fagos flanqueando a suposta ilha (Tabela 10).

Figura 4: Esquema comparativo dos processos biológicos relacionados ao estilo de vida de XF-CVC (A), XF-PD (B), XAC (C) e XCC (D). As linhas pretas na posição vertical e horizontal separam os processos biológicos de cada bactéria em seus respectivos ambientes de vida. No plano horizontal é possível comparar bactérias de mesmo gênero, cepas de *Xylella* acima e cepas de *Xanthomonas* abaixo. No plano vertical podemos comparar as bactérias em seus distintos hospedeiros. XF-CVC e XAC são capazes de infectar um mesmo hospedeiro, mas em diferentes tecidos, xilema e mesófilo foliar respectivamente. O círculo central limita a informação relacionada ao metabolismo energético das bactérias. Vias ausentes nas cepas de *Xylella* estão em linhas vermelhas. A linha que delimita o círculo central foi utilizada para representar o cromossomo principal circular, destacando possíveis informações genéticas exclusivas destas bactérias, por exemplo, a ilha SPI-7 e os genes envolvidos com síntese de LPS. O significado dos símbolos está indicado abaixo. As linhas e colunas indicadas por números e letras, respectivamente, posicionam os diferentes processos e vias metabólicas referidas no texto. Esquema retirado de MOREIRA e colaboradores (MOREIRA *et al.*, 2005).

Tabela 8: Distribuição de genes e famílias gênicas exclusivas nos quatro genomas comparados.

		XAC		XCC		XF-CVC		XF-PD	
XAC	P	198	208	2046		16		2	
	H	453	457	1590		31		0	
XCC	P	752		188	215	2		6	
	H	718		335	340	6		2	
XF-CVC	P	8		1		74	82	419	
	H	15		3		165	179	607	
XF-PD	P	1		2		147		74	82
	H	0		1		228		30	36

As interseções em fundo cinza destacam os números de famílias de genes exclusivos, enquanto que em fundo branco destacam os números de genes exclusivos. (P) – função putativa atribuída. (H) – genes ou famílias sem função atribuída (hipotéticos ou hipotéticos conservados). Células da tabela que apresentam contorno preto destacam genes exclusivos encontrados em cada um dos genomas analisados.

Tabela 9: Famílias de genes exclusivos das duas bactérias do gênero *Xanthomonas* organizados de acordo com a função predita na anotação de seus genomas.

Produto	XAC ^a	XCC ^b	#Famílias ^c	Maior número de parálogos ^d
Proteínas ribossomais do tipo 50S	6	6	6	----
ABC transportadores	6	6	6	----
Acetiltransferases	8	8	8	----
Proteínas envolvidas com transferência de grupos acil	21	20	13	683 (5)
Proteínas de quimiotaxia	35	36	12	1029 (19)
Proteínas hipotéticas conservadas	775	784	713	5393/1329 (9)
Citocromos	17	19	13	132/133 (3)
Desidrogenases	7	6	5	2112 (3)
Proteínas flagelares	30	30	30	----
Glutathione S-transferases	12	9	9	1254/1743/4689 (2)
Proteínas envolvidas com transferência de grupos glicosil	8	8	8	----
Proteínas envolvidas com goma	4	4	4	----
Hpa/Hrc/Hrp	27	26	23	1104/11,20/7,17/409 (2)*
Hidrolases	12	12	11	4994 (2)
Proteínas hipotéticas	22	10	5	51/5392 (7)
Proteínas de membrana interna	4	6	4	5210 (3)
Proteínas integrais de membrana	5	5	5	----
Integrases	4	2	2	5236 (3)
Transposases	85	70	5	161/4727 (21), 2670 (17)
Proteínas de membrana	7	6	6	5727 (2)
MFS transportadores	11	12	9	1884/1902/2097 (2)
Proteínas envolvidas com metabolismo e transporte de molibdênio	9	9	8	1656 (2)
Proteínas de membrana externa	22	17	10	2025 (9)
Oxidoredutases	14	16	11	1266 (3)
Proteínas relacionadas com fagos	15	14	14	4912 (2)
Degradação de compostos aromáticos	4	4	4	----
Receptores TonB	12	14	6	5486/5151 (4), 2396 (3)
Reguladores de transcrição	56	55	49	1952 (4)
Duplos componentes regulatórios	14	14	10	4958/432 (2)
Proteínas do sistema secretório tipo II	8	8	8	----
Proteína de degradação de vanilato	4	4	3	3488 (2)
Proteínas virB	8	10	6	5446 (5)
Total	1272	1250	1026	----

Os grupos de genes foram definidos de acordo com a participação do gene em uma determinada via metabólica e/ou função predita. As linhas em negrito/fundo cinza indicam sistemas detalhados no texto (item 3.1.2.2) ^a Número de genes em XAC; ^b Número de genes em XCC; ^c – Número de famílias gênicas; ^d – Identificador da família gênica com maior número de parálogos (número de cópias de genes parálogos entre parênteses). Famílias com mesmo número de genes estão separadas por barras.

Tabela 10: Determinação dos agrupamentos gênicos exclusivos (ITL putativas) do genoma de XAC e XCC.

Ilha	Número da CDS de XAC			PinDel	XAC CHP / HP	Transposases HXAC/UXAC	Inserção de fagos (int) HXAC/UXAC	Composição	CDS com problema ^b HXAC/UXAC
	Inicial	Final	T (U)					tRNA ^a XAC	
XaUC01	0843	0860	18 (13)	----	-- / 3	-- / --	-- / --	--	2 / --
XaUC02	0918	0924	07 (04)	----	-- / --	-- / --	-- / --	--	-- / 2
XaUC03	1101	1108	08 (05)	2	-- / 2	2 / --	(1) / 1	S	-- / --
XaUC04	1489	1511	23 (13)	3	3 / 7	2 / --	(2) / 1	tm	1 / --
XaUC05	1809	1818	10 (05)	----	-- / 2	-- / --	-- / --	R	-- / --
XaUC06	2174	2204	31 (15)	5	4 / 6	2 / --	(1) / 2	--	-- / 1
XaUC07	2214	2234	31 (13)	6	2 / 3	2 / --	(1) / 2	--	-- / --
XaUC08	2269	2286	18 (13)	7	4 / 7	-- / --	(1) / 1	--	-- / --
XaUC09	2417	2445	29 (21)	8	5 / 6	5 / 1	1+1# / 3#	A	-- / 1
XaUC10	2901	2904	04 (04)	10	3 / --	-- / --	-- / (1)	--	-- / --
XaUC11	3018	3025	08 (08)	----	-- / 8	-- / --	-- / --	--	-- / --
XaUC12	3221	3234	11 (09)	----	-- / 4	3 / 1	-- / --	--	-- / 3
XaUC13	3251	3299	48 (23)	11	2 / 18	3 / --	-- / (2)+1#	G	1 / 2
XaUC14	3503	3531	29 (08)	----	1 / 5	2 / --	-- / --	--	4 / --
XaUC15	3702	3732	31 (08)	----	-- / --	-- / --	-- / --	--	1 / 1
XaUC16	3763	3786	24 (15)	12	6 / 7	2 / --	(1) / --	M	2 / 1
XaUC17	3932	3989	58 (28)	13	10 / 11	3 / 2	-- / (1)	A	1 / --
XaUC18	4112	4148	37 (27)	----	16 / 5	1 / 1	-- / --	--	1 / 2
Total	----	----	425 (232)	10	56 / 94	27 / 5	8(7)+1# / 11(4)+4#	7	13 / 13
Ilha	Número da CDS de XCC			PinDel	XCC CHP / HP	Transposases HXCC/UXCC	Inserção de fagos (int) HXCC/UXCC	Composição	CDS com problema ^b HXCC/UXCC
	Inicial	Final	T (U)					tRNA ^a XCC	
XcUC01	0319	0347	29 (20)	----	5 / 8	-- / --	-- / --	--	-- / --
XcUC02	0524	0545	22 (08)	----	3 / 1	3 / --	-- / --	--	-- / --
XcUC03	0599	0618	20 (10)	----	-- / 1	4 / 1	-- / --	--	1 / 1
XcUC04	0735	0751	17 (08)	----	3 / 4	-- / --	-- / --	--	-- / --
XcUC05	1307	1321	13 (06)	----	3 / 3	-- / --	-- / --	--	-- / --
XcUC06	1446	1463	18 (09)	1	4 / 3	4 / --	2(1) / 1	--	-- / --
XcUC07	2091	2113	22 (16)	4	3 / 2	3 / 1	-- / (1)	GGLC	-- / --
XcUC08	2413	2429	17 (07)	----	1 / 2	-- / --	-- / --	F	-- / --
XcUC09	4048	4051	04 (04)	----	1 / --	-- / --	-- / --	--	-- / --
Total	---	---	162 (88)	2	23 / 24	14 / 2	2(1) / 2(1)	2	1 / 1

XaUC_n, *Xanthomonas axonopodis* **Unique Cluster** + identificador *n*; **XcUC_n**, *Xanthomonas campestris* **Unique Cluster** + identificador *n*; **PinDel**, **Putative insertion or Deletion**; T, número total de genes; U – número de genes exclusivos;; CHP/HP, genes hipotéticos conservados e hipotéticos, respectivamente; H_/U_, seqüências homólogas em outros organismos ou exclusiva do organismo referido, respectivamente; **HXAC** ou **HXCC**, homólogos; **UXAC** ou **UXCC**, exclusivos; **Int**, integrases. #, genes relacionados a plasmídeos. ^a Estão indicados os respectivos aminoácidos dos tRNAs identificados; ^b CDS que apresentam mutações pontuais alteração no canal de leitura ou ambas alterações.

Genes exclusivos de XAC – Como indicado na Tabela 10, identificamos 18 agrupamentos gênicos exclusivos de XAC com características de ITLs. Estes agrupamentos foram nomeados seguindo-se a nomenclatura **XaUC n** , onde: **Xa** refere-se ao gênero e espécie do organismo, *Xanthomonas axonopodis*; **UC** indica agrupamento exclusivo, do inglês *Unique Cluster*; e **n** a classificação numérica de identificação de 1 a 18. A região **XaUC1**, que compreende os genes XAC0855-XAC0860, inclui um grupo de genes com função de transportadores ABC, codificados pelos genes *oppDCB*, previamente descrito como exclusivos de XAC (DA SILVA *et al.*, 2002) (Figura 4, posição C8). A região **XaUC2** apresenta em sua composição quatro genes envolvidos com trans-hidrogenação de nucleotídeos de pirimidinas, codificados pelos genes XAC0918-XAC0924. Dois destes genes, denominados de *pntA* codificam proteínas responsáveis pela conversão de NAD em NADP. Entretanto, estes genes apresentam uma mutação pontual que altera a fase de leitura e possivelmente codificam proteínas inativas. A região **XaUC6** é composta por 31 genes. Um deles (XAC2197) codifica para uma proteína ligadora de cálcio do tipo hemolisina. Flanqueando este gene estão *hlyD* e *hlyB* (XAC2201-02), que codificam para o SS-I responsável pela secreção desta hemolisina. A região **XaUC9** apresenta em sua composição 29 CDS, das quais 21 se encontram em uma sub-região flanqueada por um tRNA^{Ala}, além de seis genes que codificam para transposases, uma proteína de fago e quatro outras relacionadas com plasmídeos. Dentro do grupo dos 21 genes se destacam dois genes que participam do sistema RM-II, codificado pelos genes *Xaml* (XAC2437) e sua respectiva metiltransferase (XAC2436) (DA SILVA *et al.*, 2002) (Figura 4, posição F12 e F13). A região **XaUC12** apresenta as CDS que codificam as denominadas *cointegrate resolution proteins*, cujos genes XAC3227-XAC3229 estão flanqueados por transposases. Duas destas proteínas, XAC3227 e XAC3229, são homólogas às integrases XACb0009 e XACb0010 respectivamente, encontradas no mega plasmídeo de XAC e, outrora descritos no agrupamento de genes que codificam para o SS-III de XCV (NOEL *et al.*, 2003). A sequência do gene XAC3228 é muito similar à região N-terminal codificada pelo gene XAC3227, e

possivelmente compreende a uma cópia truncada e inativada deste gene. A região **XaUC15** apresenta um grupo de genes que codificam proteínas do sistema HMS, cujas subunidades são codificadas pelos genes *hmsRFK* (XAC1811-XAC1813) (PENDRAK & PERRY, 1991). Em *Yersinia pestis* estes genes estão inseridos no *locus pmg* (de pigmentação), cuja função seria a de internalizar e estocar grupos heme exógenos (Figura 4, posição C7). Mutantes destes genes apresentam uma significativa redução em sua patogenicidade e em sua taxa de crescimento (PENDRAK & PERRY, 1991, SCHUBERT *et al.*, 1998).

As outras 12 regiões exclusivas **XaUC3-5,7-8,10-11,13-14,16-18** são quase que totalmente compostas por genes hipotéticos, geralmente acompanhados por genes codificadores de proteínas de fago ou elementos de transposição, e todas flanqueadas por tRNA (Tabela 10). Recentemente foram delimitadas outras cinco regiões únicas, supostamente envolvidas com transferência lateral em XAC, evidenciadas pela presença de GC e códons atípicos (LIMA *et al.*, 2005).

Entre os genes exclusivos, dispersos pelo genoma, destacamos o gene *kdgT* (XAC0337) que codifica um transportador denominado *2-keto-3-deoxy-D-gluconate*, cuja função está diretamente relacionada com internalização de subprodutos da degradação de pectina e que está flanqueado por três outros genes que codificam para transposases. É importante observar que, embora este gene esteja ausente em XCC, os genes flanqueadores são sintênicos o que mostra uma curiosa perda ou ganho de um gene exclusivo entre os dois genomas. XAC ainda possui dois genes envolvidos com metabolismo de fitoeno (XAC2744 and XAC3594), um importante precursor da biossíntese de carotenóides que poderia auxiliar as xanthomonadinas nos processos de fotoproteção durante a fase epífita. Três cópias de genes para peptidases encontrados no genoma de XAC apresentam similaridade com enzimas como *pseudomonapepsin* and *xanthomonapepsin carboxypeptidases* envolvidas com internalização e degradação de oligopeptídeos (DA SILVA *et al.*, 2002) (Figura 4, posição C8).

Genes exclusivos de XCC – Da mesma forma como descrito em XAC, as regiões exclusivas de XCC foram nomeadas como **XcUCn**, neste caso Xc se referindo a *Xanthomonas campestris* (Tabela 10). A região **XcUC1**, composta pelas seqüências codificadoras XCC0319-XCC0347, inclui genes envolvidos com a biossíntese de xanthomonadina. Dos 29 genes que compõe a região, 20 são exclusivos de XCC, de tal maneira que 6 genes posicionados entre XCC0332 e XCC0342 apresentam alta similaridade de seqüência com a região amino terminal do *pigH* de XOO, caracterizando uma possível duplicação gênica em série.

A região **XcUC3** apresenta 10 genes exclusivos (XCC0599–XCC0618), os quais estão relacionados com síntese de antígeno-O e LPS, como previamente descrito (VORHOLTER *et al.*, 2001) (Figura 5). De acordo com estes autores, mutação induzida em genes que compõem esta região resulta em uma redução na virulência de XCC. É importante destacar que esta região exclusiva se localiza à jusante dos genes *xanAB*, *rmcABCD* e *ispJI*, todos direta ou indiretamente relacionados com biossíntese de LPS (STEINMANN *et al.*, 1997, VORHOLTER *et al.*, 2001) (Figura 6). Uma comparação da região onde se encontra **XcUC3** contra o genoma de XAC, revelou que na mesma posição há uma região exclusiva de XAC, denominada XAC5 (LIMA *et al.*, 2005), a qual contém genes com funções relacionadas com biossíntese de LPS, mas cuja composição é distinta de XCC (Figura 6). Como esquematizado na Figura 5 as vias de síntese da unidade estrutural de O-antígeno de XCC e XAC são distintas, refletindo a ausência de vários genes em XAC.

XCC ainda apresenta em sua composição sete genes que podem estar envolvidos com assimilação e conversão de nitritos e nitratos em amônia (DA SILVA *et al.*, 2002) (Figura 4, posição I7 até K7). Este agrupamento gênico está inserido na região do término de replicação e posicionado a jusante de outras três regiões exclusivas dela, duas destas XCC2 e XCC3 descrita por outros autores (LIMA *et al.*, 2005), e XcUC7 descrita por nós. As regiões **XcUC7** e **XcUC9** juntas apresentam em sua composição quatro cópias do gene que codifica para proteínas similares a *avirulence/effector proteins*,

reconhecidamente importantes para a interação entre planta e patógeno e responsável pela colonização do organismo em plantas compatíveis (LIM & KUNKEL, 2004, THARA *et al.*, 2004). A região **XcUC7** ainda se destaca pela presença de várias transposases seguidas de uma tanase (Figura 4, posição I8 até K8). Esta proteína é importante para prevenir danos induzidos pela produção de taninos (pela planta) durante um processo de infecção (BHAT *et al.*, 1998). A região determinada como uma putativa região de inserção e deleção (PInDel 4) de XCC está inserida na região exclusiva XcUC7, e se caracteriza como a região de XCC com o maior desvio no conteúdo de GC e de códons (MOREIRA *et al.*, 2005), como será descrito na seção **4.1.3**.

Outro agrupamento exclusivo de XCC apresenta em sua composição genes que codificam para três enzimas que participam do sistema RM-III (Figura 4, posição G12 e G13). Este sistema RM é composto por duas helicases (XCC1067-69) e uma DNA metilase (XCC1068), e sua função está relacionada a clivagem de moléculas de DNA exógeno por intermédio de regiões curtas de reconhecimento. Entre os genes exclusivos dispersos pelo genoma de XCC, merecem destaque apenas uma proteína *avr* (*avrXccA2*) codificada pelo gene XCC2396; um gene que codifica um transportador de ácido málico, e um gene que codifica para uma bacteriocina nomeada como nisina (XCC3409).

Figura 5: Esquema da via de síntese de O-antígeno em XCC e XAC. A numeração de 1 a 18 representa os genes numerados na Figura 6. As reações circundadas por linhas contínuas (em rosa e amarelo) são exclusivas de XCC em relação aos outros três genomas (XAC, XF-CVC e XF-PD). As vias circundadas por linhas tracejadas são exclusivas das cepas de *Xanthomonas* em relação às duas cepas de *Xylella* utilizadas na comparação genômica. Esquema adaptado de VORHOLTER e colaboradores (2001)

Figura 6: Análise da região que compreende genes envolvidos com a síntese de LPS em XCC e XAC. O esquema de XCC (modelo) foi adaptado de VORHOLTER e colaboradores (2001), sendo que as setas em rosa representam genes relacionados com a síntese de O-antígeno. Os nomes dos genes estão indicados acima das setas que apontam a orientação de transcrição. Os números identificadores dos genes para XCC e XAC estão indicados de acordo com DA SILVA e colaboradores (2002). A linha preta horizontal, posicionada acima dos nomes dos genes, corresponde a uma referência numérica dos genes destacados na Figura 5. As linhas verticais mais claras indicam genes homólogos. XcUC3 corresponde uma das regiões exclusivas do genoma de XCC (Tabela 9) e a região XAC5 corresponde a uma região exclusiva previamente descrita (LIMA *et al.*, 2005). Setas brancas indicam genes sem função descrita para a síntese de O-antígeno. Setas pretas indicam a presença de elementos de transposição e X representa alteração na fase de leitura de um gene.

Genes exclusivos de XF-CVC – Em XF-CVC foram encontradas cinco grandes regiões de genes exclusivos no cromossomo principal, todas tendo no mínimo oito genes em sua composição, sendo que destes, 95% foram classificados como genes hipotéticos. Todas estas regiões estão flanqueadas por genes codificadores de proteínas de fago. O genoma de XF-CVC codifica 20 proteínas relacionadas a transferência de DNA por conjugação. Onze destes estão presentes no plasmídeo pXF51 e os demais (9 genes) no cromossomo (XF2048-XF2079) (MARQUES *et al.*, 2001). O gene que codifica para *daunorubicin C-13 ketoreductase* é exclusivo de XF-CVC e está disperso no cromossomo principal (XF1741). Esta enzima está envolvida na produção de antibiótico em *Streptomyces peucetius* (JIANG & HUTCHINSON, 2006). Finalmente, os genes que codificam para uma proteína de competência (XF0063), envolvida com a formação de *pili* tipo I, uma proteína associada à aderência de *pili* tipo IV (XF0479) e uma proteína envolvida com RM-I (XF2741) (Figura 4, posição F4 e F5) foram identificados como exclusivos do genoma de XF-CVC. No plasmídeo pXF51 dois genes exclusivos merecem destaque: um que codifica para uma *nickase protein* que se caracteriza como uma endonuclease sítio específica, similar ao gene *virD2* do plasmídeo de *Agrobacterium tumefaciens*, e uma *plasmid maintenance protein* responsável pela manutenção do plasmídeo durante a divisão celular. Dados recentes demonstraram que o gene *virD2* é altamente induzido em condições de estresse (KOIDE *et al.*, 2006).

Genes exclusivos de XF-PD – Dentre os quatro genomas que analisamos, XF-PD é o que apresenta o menor número de genes exclusivos (Tabela 8), contidos predominantemente em quatro regiões de destaque, PD0577-79, PD0906-951, PD1340-46 e PD1667-69. A primeira região é composta por três genes, envolvidos com sistema de RM (Figura 4, posição G4 e G5). A segunda região apresenta em sua composição 30 genes únicos, dos quais 16 foram classificadas como sendo codificadores de proteínas de fago e o restante como proteínas hipotéticas. A terceira região é composta por quatro genes: o primeiro anotado como *proteic killer suppression protein*, envolvida na regulação da

atividade de toxinas mediada por plasmídeos específicos; uma proteína associada à virulência nomeada como *vapI* e relatada como tendo uma importante função em plasmídeos integrativos; um gene hipotético; e o gene que codifica para a proteína *HicA*, homólogo em *Haemophilus influenzae* (MHLANGA-MUTANGADURA *et al.*, 1998, READ *et al.*, 2000). A quarta região apresenta em sua composição um gene que codifica uma endolisina de fago, e proteínas do sistema RM-II, a enzima de restrição *nspV* e sua respectiva metilase codificadas pelos genes PD1667-PD1668 (Figura 4, posição G4 e G5). Apenas dois genes exclusivos dispersos no genoma de XF-PD foram encontrados. Ambos envolvidos com formação de parede celular bacteriana, uma *UDP-N-acetylmuramate-L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase* (PD0223), capaz de catalisar o último passo da via de reaproveitamento de mureína e *UDP-N-acetylglucosamine N-acetylmuramyl (pentapeptide) pyrophosphorylundeca-prenol N-acetylglucosamine transferase* (PD0688), que catalisa o último passo da biossíntese de peptídeoglicano.

4.1.2.2. Genes mais representados em cada um dos gêneros

4.1.2.2.1. Genes mais representados no genoma das duas bactérias do gênero *Xanthomonas*

Sistema Secretório Tipo II (SS-II)

O denominado SS-II compreende um complexo protéico inserido nas membranas de bactérias e é responsável pela secreção de proteínas e outras moléculas. Um conjunto entre 12 a 14 genes altamente conservados é responsável pela síntese do aparato (Figura 7A). Este sistema é também referido como via geral de secreção ou GSP, do inglês **General Secretory Pathway** (BUELL *et al.*, 2003). Nas espécies de *Xanthomonas* analisadas, dois agrupamentos gênicos que codificam para o aparato do SS-II foram encontrados, um codificado pelo conjunto de genes XAC3534-3544/XCC0660-0670, nomeado **xps**, e o outro codificado pelo conjunto de genes XAC0694-0705/XCC3415-3426, nomeado **xcs** (DA SILVA *et al.*, 2002) (Figura 7B e Figura 4, posição C11 e J11). Em contrapartida, nas cepas de *Xylella* foram observados apenas um destes agrupamentos, codificado pelos genes XF1517-1527/PD0732-0742 (SIMPSON *et al.*, 2000) homólogo ao agrupamento **xps** de *Xanthomonas* (Figura 7B e Figura 4, posição C3 e J3).

Os genes que compõe o agrupamento **xcs** de *Xanthomonas* apresentam alta identidade (variando entre 72 e 88%) tanto quanto a sequência de nucleotídeos como de aminoácidos com os genes que compõe o mesmo agrupamento em *Caulobacter crescentus* (dados não apresentados). Pelo menos outros sete genes que circundam esta cópia do SS-II também exibiram alta similaridade com *Caulobacter crescentus*, sugerindo que esta região seja uma possível ilha de transferência lateral (ITL) composta por 17 genes (Figura 7B). Em concordância com esta hipótese, esta região apresenta conteúdo de GC similar ao encontrado em *Caulobacter crescentus* e sutilmente diferenciado da média do genoma de *Xanthomonas*. Além disto, esta possível ITL está flanqueada por conjuntos de genes que, em *Xylella*, apresentam-se concatenados. Assim sendo, dois possíveis eventos distintos poderiam explicar esta composição gênica: *Xylella* pode ter perdido o conjunto **xcs** ou *Xanthomonas* adquiriu este sistema após divergência entre as bactérias do gênero *Xylella* e *Xanthomonas*. Como mostrado na Figura 8, a análise

filogenética dos genes que compõem estes agrupamentos sugere que a divergência entre os dois agrupamentos antecede o ancestral da família *Xanthomonadaceae*, implicando que ambos os agrupamentos existiam antes da divergência dos gêneros e, portanto, provavelmente o agrupamento xcs foi perdido do genoma de *Xylella*.

Figura 7: Agrupamentos gênicos que codificam para o SS-II em *Xylella* e *Xanthomonas*.
A) Esquema do complexo protéico do SS-II (adaptado de <http://www.genome.jp/kegg>). **B)** Organização dos genes relacionados ao SS-II em *Xylella* e *Xanthomonas*. Estão representados os dois agrupamentos gênicos xps e xcs, para XAC e XF-CVC, representando aqui exemplares de seus respectivos gêneros, visto que a organização de XF-PD é idêntica a de XF-CVC e de XCC é idêntica a de XAC (omitidos no esquema). As cores dos genes seguem a nomenclatura de anotação do genoma de ambos os organismos. Em rosa estão os genes categorizados como tendo a função relacionada com patogenicidade, virulência e adaptação. Os genes delimitados pela caixa e nomeados de D a N compreendem as subunidades do SS-II. #, indica genes com alta identidade com *Caulobacter crescentus*. Genes dispostos entre as linhas verticais pontilhadas representam a possível ITL perdida em *Xylella*.

Figura 8: Árvore filogenética não enraizada gerada a partir das seqüências concatenadas dos genes que codificam para as subunidades D, E, F, G, H e K do SS-II de *Xanthomonas*, *Xylella* e outras bactérias. A seleção dos genes homólogos foi feita com base nos resultados de Blast, usando como referência 60% de cobertura e $e\text{-value} \leq 10^{-10}$. A árvore foi gerada usando máxima verossimilhança e assumindo o modelo de evolução de Poisson. Os valores de bootstrap estão presentes nos nós. As linhas coloridas dispostas à direita da filogenia representam o grupo a que pertencem a maioria das seqüências. A análise foi realizada com seqüências dos genes das subunidades D, E, F, G, H e K de XAC; XCC; XF-CVC (XF-9a5c); XF-PD (XF-temecula); **CAUCR**, *Caulobacter crescentus*; **BRAJA**, *Bradyrhizobium japonicum*; **CHRVO**, *Chromobacterium violaceum*; **ECOL11**, *Escherichia coli* CFT073; **ECOL12**, *Escherichia coli* K12; **RHILO**, *Mesorhizobium loti*; **PSEPK**, *Pseudomonas putida* cepa KT2440; **PSESM**, *Pseudomonas syringae* pv. tomato; **RALSO**, *Ralstonia solanacearum*; **PSEAE**, *Pseudomonas aeruginosa*; **SHEON**, *Shewanella oneidensis*; **VIBCH**, *Vibrio cholerae*; **VIBPA**, *Vibrio parahaemolyticus*; **VIBVU**, *Vibrio vulnificus*; **VIBVY**, *Vibrio vulnificus* cepa YJ016; **YERPE1**, *Yersinia pestis* CO92; **YERPE2**, *Yersinia pestis* KIM; **LEPIN**, *Leptospira interrogans*. O número posicionado ao final de cada identificador de espécie determina o número da cópia deste sistema nos diferentes organismos. A filogenia foi gerada por **Robson Francisco de Souza (IQ-USP)**.

Enzimas de degradação de parede celular vegetal (EDPCV)

Muitas enzimas são secretadas por uma bactéria ao meio externo, geralmente por intermédio do SS-II, descrito anteriormente. No caso dos fitopatógenos, ao entrarem em contato com o tecido vegetal estes organismos passam a liberar EDPCV na tentativa de colonizar um determinado tecido para posterior dispersão dentro da planta (BUELL *et al.*, 2003, DOI & KOSUGI, 2004, JAYASINGHE *et al.*, 2004).

Xanthomonas apresenta uma complexa e extensiva maquinaria de degradação de parede celular vegetal (DA SILVA *et al.*, 2002), quando comparada com *Xylella* (SIMPSON *et al.*, 2000). Esta maquinaria em *Xanthomonas* compreende 6 cópias de genes com função de degradação de pectina e 12 outros genes com função de degradação de celulose e hemicelulose, contra apenas 1 e 4 genes para a degradação dos respectivos compostos em *Xylella*. Este número de EDPCV pode ter uma correlação direta com a cópia adicional do SS-II e ambos podem estar associados com encharcamento e necrose intensa que decorrem da infecção por *Xanthomonas*, mas não por *Xylella*. Além disso, *Xylella* por depender exclusivamente das cigarrinhas para sua transferência entre hospedeiros não necessita, pelo menos durante a infecção, da ativação destas enzimas.

Genes relacionados com aquisição de açúcares

Degradação de parede celular vegetal geralmente tende a produzir uma somatória muito grande de carboidratos, essenciais ao metabolismo e crescimento de organismos bacterianos. Para adquirirem estes compostos do meio, as bactérias geralmente se utilizam do sistema de fosfo-transferência de açúcares ou PTS. Este sistema é composto por duas proteínas ativadas por gasto de energia (*energy-coupling proteins*), uma proteína que contém um resíduo de histidina capaz de ser fosforilado (*histidine-containing phosphoprotein* - HPr), e vários receptores específicos para cada tipo de açúcar (REIZER *et al.*, 1992, REIZER *et al.*, 1999).

Tanto as duas cepas de *Xanthomonas* quanto as de *Xylella* analisadas apresentam o PTS específico para aquisição de glicose incompletos

(XAC2974-2979/XCC2804-2809 e XF1408-1401/PD0631-0635) (DA SILVA *et al.*, 2002, MEIDANIS *et al.*, 2002). Em contrapartida o PTS específico para frutose está completo em *Xanthomonas*, mas ausente em *Xylella*, sugerindo que *Xylella* seja incapaz de internalizar este composto. É interessante observar que o primeiro gene a montante do agrupamento de genes que codificam para o PTS de frutose em *Xanthomonas* é o gene *rpfN* (XAC2504/XCC2373), também ausente em *Xylella*, e que por sua vez está envolvido com a regulação dos fatores de patogenicidade (genes *rpf*). Este gene foi descrito como responsável pela codificação de uma proteína formadora de poro (porina) sensível a glicose, que seria capaz de se estruturar na membrana externa, assim como a proteína *oprB* de *Pseudomonas aeruginosa* (WYLIE *et al.*, 1993, WYLIE & WROBEC, 1995), internalizando glicose na ausência do PTS específico.

Analisando outros genes e complexos envolvidos com internalização de açúcares, nota-se que as duas cepas de *Xanthomonas*, mas não *Xylella*, teriam a capacidade de internalizar oligogalacturonídeos, galacturonatos ou mesmo 2-ceto-deoxigluconatos (KDG) por transportadores específicos nomeados *exuT* e *kdgT* (exclusivo de XAC), codificados respectivamente pelos genes XAC4255/XCC3243 e XAC0337. Uma vez internalizados, estes compostos poderiam ser convertidos em piruvato ou 3-fosfoglicerídeos pelos genes *kduIDKA* (XAC0168-171/XCC0151-0154) (RODIONOV *et al.*, 2004), servindo assim como uma fonte de carbono alternativa, quase sempre decorrente da degradação de subprodutos vegetais, essencialmente parede celular.

Os três sistemas que detalhamos, sistema secretório tipo II, enzimas de degradação de parede celular vegetal e genes relacionados com aquisição de açúcares, em associação ao completo metabolismo energético de *Xanthomonas* são, provavelmente, muito relevantes para a sobrevivência de organismos deste gênero durante a adaptação no interior do vegetal. Este metabolismo mais dinâmico de *Xanthomonas* certamente favorece sua maior eficácia de infecção em relação à *Xylella*.

Genes relacionados com captação e metabolismo de ferro

O ferro é um dos metais mais importantes para o metabolismo de bactérias, pois está diretamente correlacionado com a síntese de citocromos, aminoácidos aromáticos e pirimidinas, além de participar como cofator de uma série de enzimas de função fundamental à sobrevivência das bactérias (EARHART, 1996, CROSA *et al.*, 2004). Para captar e internalizar este metal, geralmente as bactérias utilizam dois mecanismos. O primeiro está envolvido com a captação direta de Fe^{2+} do meio, por intermédio de receptores específicos. O segundo através do seqüestro de Fe^{3+} por intermédio de moléculas orgânicas conhecidas como sideróforos, capazes de ligar Fe^{3+} e internalizar o metal através de receptores específicos (CROSA *et al.*, 2004). Genes classicamente envolvidos com a síntese destes sideróforos não foram ainda identificados em nenhuma das cepas das bactérias aqui analisadas, embora WIGGERICH e colaboradores (1997, 2000), e mais recentemente SILVA-STENICO e colaboradores (2005), tenham demonstrado que XAC e XF-CVC, respectivamente, seriam capazes de produzir alguma espécie de composto quelante com função de sideróforos. Assim, é provável que algumas das CDS hipotéticas destes organismos possam estar relacionadas à síntese destes compostos.

Embora o número de genes envolvidos com metabolismo de ferro seja praticamente o mesmo, a diferença no número de genes que atuam como receptores é bastante significativa em *Xanthomonas*, uma média de 60 genes candidatos contra apenas 4 nos genomas das cepas de *Xylella*. Isso sugere um elevado poder de aquisição deste metal em bactérias do gênero *Xanthomonas*.

Sistema secretório tipo III (SS-III)

O SS-III corresponde uma maquinaria de secreção presente e homóloga entre patógenos animais e vegetais, cuja função básica é a de exportar ao citossol da célula hospedeira proteínas efetoras, que geralmente estão envolvidas com a modulação da sua função celular (BONAS *et al.*, 1991, GALAN & COLLMER, 1999). *Xanthomonas* apresenta uma cópia dos genes que participam da síntese e regulação deste sistema no seu cromossomo

principal, composto por 24 genes em XAC e 23 em XCC (XAC0393-0417/XCC1217-1241), nomeados de **hrp** (*hypersensitive response pathogenesis*) (DA SILVA *et al.*, 2002), enquanto que em *Xylella* nenhum gene relacionado com este sistema foi encontrado.

Do ponto de vista estrutural, em XAC este agrupamento gênico está flanqueado por outros dois agrupamentos com funções não correlacionadas. A jusante está um grupo de genes envolvidos com a síntese de biotina (SB) e a montante um grupo de genes envolvidos com degradação de poli e oligossacarídeos (DPO) (Figura 9). Os onze genes que flanqueiam os três agrupamentos gênicos citados acima (SB-HRP-DPO) são encontrados em XF-CVC, praticamente concatenados, com ausência dos agrupamentos *hrp* e DPO. Entretanto, o SB se encontra fragmentado em três regiões distintas do genoma, um outro dado que reforça a hipótese de rearranjo e perda gênica em XF-CVC. A perda dos agrupamentos *hrp* e DPO, mais uma vez, parece ser um reflexo do mecanismo de propagação de XF-CVC que é dependente do inseto. Já em XCC, embora os mesmos genes estruturais do SS-III estejam presentes, nota-se que o agrupamento Hrp está inserido num outro sítio não colinear ao genoma de XAC. A este agrupamento gênico tem sido atribuída a capacidade de se integrar em diferentes regiões, em alguns casos sendo até identificado como uma ilha de transferência lateral putativa (HANSEN-WESTER & HENSEL, 2001, MUDGETT, 2005).

Figura 9: Representação esquemática da região genômica de XAC que codifica para os genes do SS-III. A figura explicita as possíveis perdas dos agrupamentos Hrp (SS-III) e DPO (degradação de poli e oligossacarídeos), e fragmentação do agrupamento SB (Síntese de Biotina) nos genomas das cepas de *Xylella*, representado por XF-CVC, com relação a *Xanthomonas*, representado por XAC. Os triângulos pretos denotam a presença de um tRNA com a identificação do aminoácido que transporta (C – cisteína e R – arginina).

Quimiotaxia e genes relacionados com síntese de flagelo

As bactérias do gênero *Xanthomonas* são classicamente reconhecidas como bactérias que apresentam um único flagelo (monotríquias) na posição polar (HU *et al.*, 2005). Coincidente com esta observação morfológica, os genomas de XAC e XCC codificam todos os genes relacionados à biossíntese e regulação do flagelo (DA SILVA *et al.*, 2002). Estes genes estão organizados em seis regiões genômicas nomeadas como regiões I a VI (Figura 10) nos genomas de XAC e XCC. Três destas regiões estão agrupadas em cerca de 130 Kpb e se posicionam perto do término de replicação (regiões II, III e IV) (Figura 10 e Figura 11A). A primeira das três regiões concatenadas contém as seqüências codificadoras entre XAC1888-1909/XCC1866-1891, onde se destacam os genes relacionados ao processo quimiotático, tais como *tsr* (Figuras 11A e 11B). A segunda região está contida entre as seqüências codificadoras XAC1330-1957/XCC1903-1927, e apresenta genes que participam da estrutura do motor flagelar e da sua regulação, tais como genes *che*, *flh* e *fli* (Figuras 11A e 11B). Finalmente, a terceira região está contida entre as seqüências codificadoras XAC1973-1989/XCC1939-1935, e codifica genes relacionados a estrutura do filamento flagelar (Figuras 11A e 11B). As outras três regiões restantes são menores e estão distribuídas pelo genoma. São compostas pelos genes *cheBR* (XAC1280-1281/XCC1183-1184), região I, *cheAWRB* (XAC2865-2870/XCC2700-2705), região V, e *motAB* (XAC3693-3694/XCC3650-3658), região VI (Figura 10).

É importante ressaltar que entre a região II e a região III do genoma de XAC existem 20 outros genes com funções não correlacionadas, dos quais 10 classificam-se como hipotéticos e 5 outros ligados a elementos genéticos móveis, cuja média do conteúdo de GC é discrepante da média do genoma de XAC, indicando um possível evento de inserção. Da mesma forma, entre as regiões III e IV também existem 17 genes, porém, todos neste caso, apresentam uma provável função putativa determinada por anotação, sem variação na média do conteúdo de GC (Figura 10). Curiosamente, embora XCC tenha os agrupamentos I e III-VI sintênicos aos de XAC, os agrupamentos de genes que não apresentam função correlacionada (entre II e III e III e IV) são diferentes. Isto sugere a ocorrência de pressão seletiva para a manutenção dos

genes estruturais do flagelo devido à importância deste aparato para a sobrevivência destes organismos.

Uma característica peculiar da região II é que nela estão contidos 10 cópias do gene *tsr* concatenados em XAC, sendo que XCC apresenta 9 em disposição parecida (Figura 10 e Figura 11). Estes genes codificam as proteínas relacionadas com captação e transferência de sinais quimiotáticos ao interior da bactéria (pertencentes a família de genes **MCP**), via cascata de sinalização mediada pelos genes *che* (STOCK & SURETTE, 1996). Análise filogenética destes genes revelou a ocorrência de uma possível duplicação gênica em série (Figura 12). Estas análises incluíram todos os genes *mcp* de XAC e XCC, inclusive os genes *tsr*, envolvendo algumas outras bactérias como referência. Observamos que os genes *tsr* em série se apresentam compactados em um único clado, com valores de *bootstrap* bastante altos (72 a 100), sugerindo fortemente que tais genes se duplicaram antes da divergência das espécies (Figura 12). Paralelamente a esta análise filogenética, comparamos os domínios que compõem os genes *tsr* de XAC e XCC. Os resultados indicam uma extrema especialização destes *tsr* (Figura 13). Usando a ferramenta Pfam para determinar domínios, verificamos que, como esperado, o domínio *mcp* (carboxi-terminal) seria conservado em todos eles, bem como o domínio *Hamp* (região central), ambos envolvidos com a cascata de sinalização intracelular. Entretanto, alguns aspectos peculiares destas seqüências foram observados: **i)** Há dois domínios entre os domínios *Hamp* e *mcp* nos genes *tsr* de *Xanthomonas*. Estes domínios não apresentam características funcionais bem definidas, por isso se agrupam entre os genes que compõem o chamado Pfam-B, com códigos identificadores 2069 e 1759. É interessante observar que ambos os domínios parecem ser bem específicos aos genomas de *Xanthomonas*, uma vez que apenas 27 e 30 seqüências homólogas a estes respectivos domínios (**N-hits**) foram encontrados, sendo 25 e 23 destes pertencentes aos genes *mcp* de XAC e XCC, respectivamente (**X-hits**); **ii)** Os domínios sensoriais, que ficam expostos no periplasma bacteriano, apresentam hipervariabilidade de seqüências, diferentemente do domínio de sinalização, o que sugere diferentes especificidades a distintos compostos. Da mesma forma que nos domínios 2069 e 1759, muitos domínios desta região também parecem ser exclusivos dos genes *tsr* de *Xanthomonas*; **iii)** A filogenia

coincide em grande parte com a análise da composição de domínios, uma vez que tal distribuição tanto na árvore filogenética quanto na análise esquemática, agrupam os mesmo genes (Figuras 12 e Figura 13).

Nos genomas das cepas de *Xylella* que analisamos, nenhum gene com função de síntese e regulação de flagelo foi encontrado, como esperado. A ausência de flagelo correlaciona-se bem com o estilo de vida de *Xylella*, a qual tem sua infectividade restrita a um vetor específico que é capaz de inoculá-la, dispensando assim o complexo aparato de infecção existente em *Xanthomonas*. As cepas de *Xylella* possuem apenas um gene denominado *chpA* a jusante do agrupamento gênico que participa da síntese de síntese *pili* (*pilJIG*) cuja função está provavelmente relacionada a ativação e movimento do tipo *twitching motility* (ver discussão adiante no item 4.1.2.2.2).

Figura 10: Análise estrutural da distribuição de genes e regiões genômicas relacionados com síntese ou regulação do flagelo de XAC e XCC. Um esquema representativo do genoma de XAC e XCC está mostrado à esquerda e direita, respectivamente. Os números indicam tamanho em Mpb. Os nomes acima de cada uma das seis regiões (I - VI) indicam o nome dos respectivos genes. Os padrões das setas indicam as atividades e funções descritas na legenda.

Figura 12: Árvore filogenética não enraizada para os genes *tsr* existentes nos genomas de XAC e XCC. Os genes de *Xanthomonas* foram nomeados de acordo com o identificador numérico de anotação do gene, enquanto que para outras bactérias foi utilizada a nomenclatura do campo GN do programa UNIPROT (APWEILER *et al.*, 2004). A seleção dos genes homólogos foi feita com base nos resultados de Blast, usando como referência o gene XCC1869 com pelo menos 60% de cobertura e $e\text{-value} \leq 10^{-10}$. A árvore foi gerada usando máxima verossimilhança e assumindo o modelo de evolução de Poisson. Os valores de bootstrap estão presentes nos nós. As linhas coloridas dispostas à direita da filogenia representam uma forma esquemática de distribuição dos genes *tsr*, explicitados em análises de domínios na Figura 13. Os códigos das espécies seguem a legenda abaixo: **AGRT5** – *Agrobacterium tumefaciens* cepa C58/ATCC33970; **BRAJA** – *Bradyrhizobium japonicum*; **CHRVO** - *Chromobacterium violaceum*; **GEOSL** - *Geobacter sulfurreducens*; **HALCU** - *Halobacterium cutirubrum*; **HALN1** - *Halobacterium* sp. cepa NRC-1/ATCC 700922/JCM11081; **HALSA** - *Halobacterium salinarium*; **METMA** - *Methanosarcina mazei*; **NITEU** - *Nitrosomonas europaea*; **PSEAE** - *Pseudomonas aeruginosa*; **RHILE** - *Rhizobium leguminosarum*; **RHOPA** - *Rhodopseudomonas palustris*; **VIBCH** - *Vibrio cholerae*; e **WOLSU** - *Wolinella succinogenes*. A filogenia foi gerada por **Robson Francisco de Souza (IQ-USP)**.

Figura 13: Análise da composição de domínios encontrados nos genes *tsr* dos genomas de XAC e XCC. N_gene – determina o número de genes *MCP*, não especificamente *tsr*, encontrados no genoma de ambos os organismos. Os números na escala horizontal indicam o total de aminoácidos das seqüências. Seta preta indica a posição final da maioria das proteínas. Na mesma escala, os fragmentos de linha na cor vermelha determinam as regiões usadas na filogenia detalhada na Figura 12. Observe que a região escolhida não abrange o domínio hipervariável dos genes *tsr* (domínio sensor), o que torna o resultado da filogenia independente da análise de domínios; Gene_id – denota a nomenclatura de anotação dos genes *tsr* de XAC e XCC. As linhas verticais coloridas, posicionadas ao lado esquerdo da figura, representam uma distribuição esquemática dos genes *tsr* na filogenia da Figura 12. A região consenso, posicionada abaixo da figura principal, procura agrupar os principais domínios encontrados em cada região das seqüências dos genes *tsr*, detalhando o possível nome do domínio mais representativo. A estrutura básica apresentada no banco de dados do Interpro despreza os domínios intermediários entre os domínios *MCP* e *Hamp*, denominando o todo como domínio de sinalização. Diferentemente dos genes *tsr* de outros organismos (*E.coli* como referência), os genes *tsr* de *Xanthomonas* apresentam apenas um domínio transmembrana (TM), posicionado entre os domínios hipervariável e Hamp. Em outras bactérias o domínio TM adicional fica posicionado anteriormente ao domínio hipervariável, o que dá ao mesmo um aspecto de alça em forma de grampo (Figura 11B). As cores simbolizam domínios do PFAM-B indicados na tabela à direita. Nesta tabela o código determina o número identificador do domínio (data de referência abril de 2004); N-hits, número de vezes que este domínio aparece nos banco de dados do Pfam; X-hits, número de vezes que este domínio encontra-se presente nos genes *tsr* de *Xanthomonas*. Observe que muitos destes domínios são exclusivos de *Xanthomonas*. Alinhamento gerado por Robson Francisco de Souza (IQ-USP).

Genes com funções regulatórias

Todos os tópicos abordados anteriormente demonstram a maior complexidade de *Xanthomonas* quando comparada com *Xylella*, o que se relaciona diretamente aos seus estilos de vida distintos. Coincidente com estas observações está o fato de que *Xanthomonas* apresenta um número bem superior de genes relacionados com funções regulatórias. São 296 e 285 genes anotados como tendo esta função respectivamente nos genomas de XAC e XCC, ou seja, aproximadamente 8% dos genes presentes em seus genomas. Este total de genes equivale a quatro vezes mais do que o número de genes com mesma função anotados em *Xylella fastidiosa* (~77 genes, 4.2%) (SIMPSON *et al.*, 2000, DA SILVA *et al.*, 2002). Na tentativa de facilitar o estudo dos genes regulatórios, em *Xanthomonas* os mesmos foram divididos em quatro grupos: I) Sistemas de dois componentes, compostos pelo gene que codifica o sensor e pelo gene que codifica o regulador de resposta. Em alguns casos os genes são híbridos, codificando ambas as funções. Neste grupo foram anotados 69 genes em XAC e 59 genes em XCC contra apenas 22 em XF-CVC e 19 em XF-PD; II) Ativadores e repressores de transcrição gênica. Nesta categoria são 164 genes de XAC e 160 genes de XCC contra apenas 41 genes de XF-CVC e 29 de XF-PD; III) Quinases e fosfatases, 32 genes anotados em XAC e 28 em XCC contra apenas 3 genes em XF-CVC e 4 em XF-PD; e IV) Fatores sigma e outros componentes regulatórios, totalizando 31 genes em XAC e 33 em XCC contra apenas 11 em XF-CVC e 12 em XF-PD.

Observamos em *Xanthomonas* a existência do sistema de dois componentes denominado *nodVW*, que em outros organismos está envolvido com regulação dos genes envolvido com nodulação (LOH *et al.*, 1997, LOH *et al.*, 2002). Também destacamos o sistema denominado *ntrBC* que classicamente está relacionado com regulação da expressão de genes envolvidos com fixação de nitrogênio (MACFARLANE & MERRICK, 1985, MARTINEZ-ARGUDO *et al.*, 2001). Entretanto nenhum destes dois fenômenos (nodulação e fixação de nitrogênio) foi descrito, até o momento, para bactérias do gênero *Xanthomonas*. Assim, seria interessante realizar mutagênese direcionada a estes genes, para averiguar suas reais funções nestas bactérias, já que não se sabe a real função destes genes nestes organismos.

4.1.2.2.2 Genes mais representados no genoma de bactérias do gênero *Xylella*

Utilizando a mesma estratégia de análise do genoma de bactérias do gênero *Xanthomonas*, descrevemos a seguir estão descritos genes e complexos gênicos que foram classificados como estando mais representados no genoma de bactérias do gênero *Xylella*.

Pili

Os *pili* compreendem estruturas celulares bacterianas muito importantes, pois são utilizados como ferramentas de adesão e colonização. Estas estruturas têm sido classificadas em dois tipos: *pili* do tipo I (fímbrias do tipo I) (CAPITANI *et al.*, 2006) e *pili* do tipo IV (WALL & KAISER, 1999). Os dados a seguir se referem aos *pili* do tipo IV, cuja estrutura se caracteriza por filamentos localizados na posição polar das bactérias e que apresentam a função de promoverem movimentos denominados de *twitching motility* (ALM & MATTICK, 1997, MATTICK, 2002). Tanto *Xanthomonas* quanto *Xylella* apresentam genes que codificam para *pili* tipo IV. Entretanto, cepas de *Xylella* apresentam um número maior de genes para esta função. Análise para identificar genes homólogos nos genomas revelou dois grupos de genes com esta função em *Xylella*, ambos contendo o gene *pilE*, que por sua vez, codifica para a subunidade protéica que é polimerizada para a formação da estrutura filamentosa propriamente dita. O primeiro agrupamento é composto pelos genes *pilB-pilC-pilE-pilD* tendo uma cópia em cada genoma, ao passo que o segundo agrupamento é formado pelos genes *pilE -pilY1-pilX-pilW-pilN-fimT* com duas cópias em *Xylella* e apenas uma em *Xanthomonas*. Ademais, outros genes anotados como tendo função correlacionada estão distribuídos aleatoriamente no genoma de ambas as bactérias, e podem estar diretamente envolvidos com a síntese e regulação da atividade deste aparato.

Estudos filogenéticos para ambos os agrupamentos gênicos sugere a presença de duas cópias do agrupamento *pilE-fimT* no ancestral das *Xanthomonadacea*, sendo que uma cópia foi perdida em XAC e a outra em XCC, mas ambas foram mantidas nas cepas de *Xylella* (Figura 14). A manutenção de mais cópias do agrupamento dos genes envolvidos com

síntese de *pili* tipo IV em *Xylella* parece refletir a importância da motilidade para colonização do hospedeiro. Recentemente, foi demonstrado que *Xylella fastidiosa* utiliza o processo de *twitching motility* para se movimentar contra o fluxo de seiva do xilema. A análise de mutantes que não possuem *pili* do tipo IV confirmou o envolvimento desta estrutura na motilidade de *Xylella* e na sua capacidade de estabelecer uma colonização eficiente (MENG *et al.*, 2005).

Figura 14: Árvore filogenética não enraizada para os genes que compõem o agrupamento *pilB-pilD* em associação a uma árvore filogenética gerada nas mesmas condições para os genes *fimT* individualizados ou agrupados na região *pilE-fimT*. Para gerar as filogenias, as seqüências de aminoácidos deduzidas para cada um dos genes foram previamente alinhadas utilizando-se Clustal-W com padrão de máxima verossimilhança e modelo de Poisson de matriz de substituição de aminoácidos. Os valores de bootstrap estão indicados nos nós. Os genes candidatos foram selecionados pela presença da região amino-terminal conservada do gene de pilina, mesmo proveniente de genes inteiros ou fragmentados. Algumas seqüências truncadas na região amino-terminal foram excluídas da análise. Todas as cópias dos genes *pilE* encontrados nos genomas de XF-CVC (XF-9a5c) e XF-PD (XF-temecula) são membros de agrupamentos gênicos que *pilB-pilD* nas quatro bactérias. Note também que este mesmo gene é duplicado no interior do agrupamento *pilB-pilD* em todas as linhagens. A filogenia foi gerada por **Robson Francisco de Souza (IQ-USP)**.

Colicinas-V

As colicinas-V (Col-V) constituem um dos diversos grupos de peptídeos antimicrobianos produzidos por bactérias, as quais apresentam atividade contra espécies bacterianas relacionadas, geralmente para controlar o crescimento de espécies que habitam o mesmo ambiente, diminuindo a competição por nutrientes (RILEY & WERTZ, 2002a, RILEY & WERTZ, 2002b). Em *E. coli*, o operon Col-V é composto por quatro genes plasmidiais: *cvaC*, *cvaA*, *cvaB* e *cvi* (GILSON *et al.*, 1987). O gene *cvaC* codifica a colicina propriamente dita é sintetizada como um precursor de 103 aminoácidos, que por sua vez é processado pela remoção de 15 aminoácidos de sua extremidade aminoterminal. Já as proteínas codificadas pelos genes *cvaA* e *cvaB* estão envolvidas com a secreção da Col-V, que em conjunto com TolC compõem um transportador ABC (FATH *et al.*, 1994). A proteína codificada pelo gene *cvi* é responsável pela proteção da bactéria produtora da Col-V contra os efeitos deletérios desta toxina, encontrando-se ancorada na membrana interna. Além disso, um gene cromossômico, denominado *cvpA*, também é necessário para a síntese e secreção da Col-V (ZHANG *et al.*, 1995). No cromossomo de *E. coli*, o gene *cvpA* está inserido no operon *purF* (FATH *et al.*, 1989, GILSON *et al.*, 1990).

A anotação inicial do genoma de XF-CVC revelou, a existência de dois genes *cvaC* (XF0262 e XF0263) codificadores de peptídeos similares a Col-V bem como de genes relacionados a sua biossíntese e transporte, tais como *cvaA*, *cvaB* e *cvpA* (SIMPSON *et al.*, 2000). Uma análise mais detalhada do genoma de XF-CVC nos permitiu identificar uma cópia adicional de *cvaC* (XF0264) bem como propor que a CDS XF0261, então classificada como hipotética, codifica uma provável proteína de imunidade similar a *cvi* descrita no operon Col-V de *E. coli*. O provável gene *cvi* de XF-CVC, é vizinho das três cópias dos genes *cvaC*, mas aparentemente não compõe uma mesma unidade transcricional (PASHALIDIS *et al.*, 2005)⁵. A demonstração da funcionalidade

⁵ 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. & Da Silva, A. C. (2005). *Whole-genome expression profiling of Xylella fastidiosa in response to growth on glucose*. **Omic**s. 9, 77-90. (**ANEXO 5**).

do sistema de síntese e produção de bacteriocinas da família Col-V em XF-CVC ainda aguarda experimentação adicional. Entretanto, já foi descrito que membros desse sistema, incluindo os genes *cvaC*, são expressos em diferentes condições experimentais (DE SOUZA *et al.*, 2003, PASHALIDIS *et al.*, 2005).

Realizamos uma análise comparativa dos genomas de *Xanthomonas* e de *Xylella* com o objetivo de investigar a organização dos genes relacionados à produção de Col-V. Como esperado, o genoma de XF-PD codifica todos os genes deste sistema descritos em XF-CVC (Tabela 11). Entretanto, observamos que XAC e XCC não codificam nenhum gene com similaridade à *cvaC*, responsável pela síntese de colicina V propriamente dita. Os genomas de XAC e XCC também não codificam o gene *cvaA*, ainda que XAC possua uma cópia do gene *cvaB* (Figura 15A). Curiosamente, observamos nesta região do genoma de XAC a inserção de duas cópias das transposases *ISxac3* (Figura 15A). Por outro lado, vale ressaltar que ambas as cepas de *Xylella* e de *Xanthomonas* exibem perfeita sintonia de seu *operon purF* quando comparadas ao *E. coli* (Figura 15B). Como já mencionado este *operon* contém o gene *cvpA*, que regula a expressão dos genes relacionados a síntese e secreção de Col-V. Assim, concluímos que em XAC e XCC, ainda que tenham conservado alguns dos componentes relacionados a produção de Col-V, não produzem bacteriocinas deste grupo.

Tabela 11: Genes relacionados a regulação, biossíntese, secreção e imunidade do sistema Col-V identificado no genoma de *Xylella*.

Nome do gene	Função ^a	CDS ^b	
		XF-CVC	XF-PD
<i>cvpA</i>	Gene regulatório cromossomal. Regula positivamente a expressão dos genes do plasmídeo pColV de <i>E.coli</i>	1948	0852
<i>cvaC</i>	Codifica o polipeptídeo precursor de colicina V	0262	0215
		0263	0216
		0264	0217
<i>cvi</i>	Codifica a proteína de auto-imunidade	0261 ^c	0214 ^c
<i>cvaA</i>	Codifica para uma proteína associada às membranas interna e externa. Está envolvida com a transferência de colicina V do periplasma para o meio externo. É análoga à proteína <i>HlyD</i> do sistema secretório tipo I.	1216	0496
<i>cvaB</i>	Codifica para um transportador ABC, responsável pela transferência de colicina V do citoplasma para o periplasma, simultaneamente à clivagem do peptídeo sinal do precursor da proteína <i>cvaC</i> . É análoga a <i>HlyB</i> do sistema secretório tipo I.	1220	0499
<i>tolC</i>	Codifica para uma proteína acessória posicionada na membrana externa e que apresenta a função de auxiliar no transporte mediado por <i>cvaA</i> .	2586	1964

^a Informações sobre a função destes genes foram extraídas das referências citadas no texto

^b Identificação das CDS anotadas nos genomas de XF-CVC e XF-PD

^c CDS similar a *cvi* de *E.coli* (PASHALIDIS *et al.*, 2005)

A **B**

Figura 15: Sintenia dos genes envolvidos com regulação da biossíntese e secreção de colicina V. Em amarelo estão destacados os genes envolvidos diretamente com regulação (**B**) e transporte (**A**) de colicina V. Cinza, genes hipotéticos. Preto, elementos genéticos móveis. Branco, genes com outras funções não relacionadas. **A** – Esquema de sintenia dos genes envolvidos na secreção de colicina V, usando como referência duas cepas de *Xylella*. Note a ausência do gene *cvaA* em XAC, possivelmente induzida pela inserção de dois elementos de transposição a montante do gene *cvaB* (setas em cinza escuro). Já em XCC o que se observa é a ausência total dos genes envolvidos com esta função (Tabela 11). **B** – Sintenia do operon *purF* que contém o gene *cvpA*, regulador da síntese de colicina V em *E.coli*.

Sistema de Restrição e Modificação de DNA

Os sistemas de restrição e modificação (RM) de DNA em bactérias são compostos por endonucleases, metilases e desmetilases de DNA, sendo responsáveis por diversos processos celulares, mas principalmente por limitar a entrada de DNA exógeno. Estes sistemas são classificados em três tipos: RM-I, RM-II e RM-III (WILSON & MURRAY, 1991). Análise da composição genômica de XF-CVC revelou a presença de quatro cópias do sistema RM-I, três das quais inseridas em uma região cromossômica de aproximadamente 18 Kpb. Cada uma das cópias do sistema RM-I apresenta três genes em sua composição, nomeados de *hsdSRM*, que codificam respectivamente para um determinante de especificidade, uma endonuclease e uma DNA metilase. A quarta cópia do sistema RM-I de XF-CVC apresenta em sua composição dois pseudogenes flanqueados por um gene que codifica para um tRNA^{Phe}. Embora XF-CVC apresente múltiplas cópias do sistema RM-I, nenhuma cópia dos outros dois tipos de sistemas RM foi encontrada em seu genoma (SIMPSON *et al.*, 2000). Por outro lado, XF-PD possui, além das quatro cópias homólogas do sistema RM-I de XF-CVC, uma cópia adicional do sistema RM-II, codificado pelos genes *nspV* e sua respectiva metilase (PD1607-1608), a qual também está flanqueada por um gene que codifica para um tRNA^{Asn} (VAN SLUYS *et al.*, 2003). Esta região em XF-CVC, é constituída de uma ilha genômica putativa, flanqueada por uma integrase e que foi denominada anteriormente como ilha G11 (NUNES *et al.*, 2003).

Diferentemente de *Xylella*, XAC apresenta apenas uma cópia do sistema RM-I, codificada pelos genes XAC2898-2900, e uma cópia do sistema RM-II, codificada pelos genes XAC2437-2438. O significado destas diferenças biológicas ainda não é conhecido. Entretanto, especulamos que possa haver uma correlação entre o número de cópias e o tipo de sistema de RM com a organização estrutural dos genomas destas bactérias, caracterizando distintas proteções de seus materiais genéticos.

4.1.2.3. Genes mais representados exclusivamente nos genomas de XAC e XF-CVC

Ilha de patogenicidade SPI-7

Dentro do grupo dos 1495 genes ortólogos, descritos entre os genomas de XAC e XF-CVC, o total de 23 se destaca como sendo exclusivos destes genomas em relação à XCC e XF-PD. Dezesesseis destes merecem destaque especial, dos quais 14 foram anotados como sendo hipotéticos (Tabela 12). Entretanto, quando analisados em maior detalhe, estes genes apresentam alto grau de similaridade e mantêm uma considerável sintenia com os genes presentes na ilha de patogenicidade nomeada como SPI-7 de *Salmonella enterica* (PICKARD *et al.*, 2003) (Figura 4, posição E8 e E6). O mais interessante é que esta sintenia também é observada em outros organismos, muitos deles bactérias de solo, como *P.fluorescens*, *B.fulgorum* ou *R.metallidurans*, caracterizando uma possível ilha de transferência lateral relativamente freqüente.

Em ambos os genomas estes genes estão agrupados em uma única região genômica que compreende a PInDel 8 de XF-CVC e a PInDel 4 de XAC (ver definição e classificação de PInDels na seção 4.1.3). Observamos que em *Xylella* há uma inserção de 37 genes que compreendem a PInDel 7 (XF1718-XF1754) a montante de SPI-7, estando esta inserção flanqueada por um tRNA^{Gly}. Em XAC há um agrupamento adicional de quatorze genes no interior da ilha SPI-7, com destaque especial para uma integrase (XAC2220), que talvez explique a inserção destes genes, mediada por fago (Figura 16). Um dado que reforça a idéia de aquisição da PInDel 7 de XF-CVC por transferência lateral é o fato da média de conteúdo de GC desta região ser de 68% contra apenas 53% da média do genoma. Além disso, vale ressaltar que a média do conteúdo de GC da ilha SPI-7 é de 65% em ambos os genomas, indicando a ocorrência de inserção recente em ambos os organismos. Estas observações permitem especular que esta ilha possivelmente teria alguma importância na interação destes fitopatógenos com seus hospedeiros específicos.

Tabela 12: Pares de genes homólogos em XAC e XF-CVC que compõe a ilha de patogenicidade putativa SPI-7.

Gene	CDS em XAC	Posição inicial	Posição final	Tamanho (pb)	CDS em XF-CVC	Posição inicial	Posição final	Tamanho (pb)	Produto
1	XAC2205	2589160	2590005	845	XF1785	1702552	1703427	875	Proteína relacionada com o particionamento do cromossomo
2	XAC2206	2590213	2591895	1682	XF1784	1701566	1702327	761	Proteína hipotética
3	XAC2207	2591892	2592452	560	XF1782	1700095	1700655	560	Proteína hipotética
4	XAC2208	2592457	2593680	1223	XF1781	1698842	1700098	1256	Proteína hipotética
5	XAC2209	2593977	2594825	848	XF1780	1697908	1698657	749	Proteína hipotética
6	XAC2210	2595000	2595359	359	XF1779	1697378	1697911	533	Proteína hipotética
7	XAC2211	2595501	2595887	386	XF1778	1696852	1697304	452	Proteína de ligação ao DNA fita simples
8	XAC2213	2598564	2600027	1463	XF1774	1690011	1694029	4018	Metiltransferase de DNA
9	XAC2217	2603332	2603745	413	XF1772	1688664	1989074	410	Proteína hipotética
10	XAC2218	2603765	2604157	392	XF1771	1688250	1688642	392	Proteína hipotética
11	XAC2221	2606053	2606418	365	XF1764	1680412	1680765	353	Proteína hipotética
12	XAC2236	2619587	2620810	1223	XF1761	1677741	1678868	1127	Proteína hipotética
13	XAC2237	2620870	2621571	701	XF1760	1677021	1677683	662	Proteína hipotética
14	XAC2238	2621664	2622008	344	XF1759	1676603	1676926	323	Proteína relacionada a plasmídeo
15	XAC2239	2622116	2622523	407	XF1758	1676095	1676502	407	Proteína hipotética
16	XAC2240	2622538	2622798	260	XF1757	1675818	1676078	260	Proteína hipotética

Figura 16: Análise da região que compõe a PinDel 4 de XAC e PinDel 8 de XF-CVC, homóloga à ilha SPI-7. Os genes homólogos detalhados na Tabela 12 e que definem a ilha SPI-7 estão indicados por linhas verticais pontilhadas, assim como descrito para a ilha SPI-7 de *S. enterica* (PICKARD *et al.*, 2003). PinDel 4 de XAC está sombreada em verde claro e a PinDel 8 de XF-CVC está sombreada em rosa claro. A região entre os genes XF1718 e XF1754 corresponde à PinDel 7 de XF-CVC (sombreado em azul), flanqueada a montante e a jusante por integrases (retângulos pretos). Retângulos cinza representam genes hipotéticos da ilha SPI-7. Retângulos amarelos representam genes não relacionados com a ilha SPI-7. SS-I, Sistema secretório do tipo I. As setas em preto indicam tRNAs. * determinam genes com conteúdo de GC mais elevado.

4.1.2.4. Genes mais representados na família Xanthomonadaceae

Entre os genes exclusivos da família Xanthomonadaceae, estão aqueles envolvidos com a síntese de DSF, do inglês *Diffusible Signal Factor*, codificado pelos genes *rpf* (reguladores de fatores de patogenicidade) e envolvidos com a sinalização entre bactérias (TANG *et al.*, 1991, BARBER *et al.*, 1997). Em relação ao agrupamento de genes *rpf*, observamos que XAC difere de XCC por não possuir os genes *rpfH* e *rpfI*. A inserção de dois elementos de transposição entre os genes *recJ* e *prfB* pode ter sido o fator que desencadeou a perda do gene *rpfI* (DOW *et al.*, 2000). Nesta mesma região existe um gene *wapA* que codifica para uma proteína associada à membrana pertencente à família *rhs* de proteínas e contém nove repetições do domínio de mesmo nome (APWEILER *et al.*, 2000, APWEILER *et al.*, 2004). Este domínio *rhs* já foi descrito em *E.coli* como sendo uma região de alta frequência de inserção de elementos de transposição, bem como de eventos de rearranjos e duplicações gênicas (LIN *et al.*, 1984, MINET & CHIQUET-EHRISMANN, 2000), o que poderia justificar a instabilidade estrutural desta região genômica em XAC. O agrupamento de genes *rpf* em ambas as cepas de *Xylella* não possui os genes *rpfI*, *rpfH* e *rpfD* e apresenta ainda reorganização entre os genes *rpfF* e *rpfB* (DA SILVA *et al.*, 2002, MOREIRA *et al.*, 2004).

É importante destacar que os genes *rpf* ainda participam da regulação dos genes de goma constituída pelo polissacarídeo extracelular, o qual tem sido associado a patogenicidade de bactérias (VOJNOV *et al.*, 2001). Entre as cepas de *Xanthomonas*, os genes relacionados à produção da goma apresentam alta identidade, chegando a 98% em pares homólogos. Entretanto, menor identidade destes genes em relação aos genes correspondentes em *Xylella* (65 e 83% de identidade). Além disso, o genoma de *Xylella* não possui os genes *gumGIL*, responsáveis pela polimerização e modificação de resíduos de manose do EPS (SIMPSON *et al.*, 2000). A ausência destes genes pode implicar na produção de uma goma com características físicoquímicas distintas (DA SILVA *et al.*, 2001).

O agrupamento de genes que codificam para xanthomonadinas, pigmentos amarelados que conferem fotoproteção (STARR *et al.*, 1977, POPLAWSKY *et al.*, 2000), compreende um outro exemplo de genes família-específico, ainda que sejam somente sintetizados por bactérias do gênero *Xanthomonas*. Estes pigmentos apresentam como característica básica a insolubilidade e, *in vivo*, encontram-se alojados na membrana externa (GOEL *et al.*, 2002). Esta proteção adicional confere a estas bactérias maior capacidade de sobreviverem como organismos epífitos durante seu ciclo de propagação e infecção. Além disso, em laboratório, as xanthomonadinas são costumeiramente utilizadas como fatores de classificação fenotípica e taxonômica de diversas espécies (GREBE & STOCK, 1998). Quimicamente, são compostos formados por mono ou dibromoaril polienos e são subprodutos da esterificação de glicerolfosforil sorbitol ou glicerofosfato (STARR *et al.*, 1977).

A região que contém os genes envolvidos com a síntese deste composto foi primeiramente descrita em XOO (GOEL *et al.*, 2001, GOEL *et al.*, 2002) ocupando 20 Kpb do genoma deste organismo codificando um total de 14 CDS, as quais foram nomeadas em números crescentes, sendo os genes 03, 06, 07 e 10 considerados essenciais para a sua síntese (Figura 17). XCC e XAC apresentam este agrupamento de genes idênticos ao descrito em XOO (Figura 4, posição E14 e I14), exceto pelo fato dos dois primeiros genes estarem localizados em outras regiões no genoma de ambas as *Xanthomonas* (Figura 17). Em XF-CVC se nota apenas a presença vestigial deste agrupamento gênico, já que a sequência homóloga ao gene 03 de XOO foi identificada entre os genes XF0777 e 0778 com uma identidade de 67%, caracterizando a formação de um pseudogene. O mesmo acontece com a região entre os genes XF0777 e 0776 que apresentou homologia com o gene 05 de XOO (Figura 17).

Denominamos a região codificada entre os genes 03 e 12 como agrupamento maior (*large cluster*), já que identificamos um novo agrupamento de genes que possivelmente tem relação com a síntese deste pigmento, e que o denominamos “pequeno agrupamento” (*small cluster*) (Figura 17). Esta provável correlação deve-se a função e localização destes genes, que estão relacionados

com síntese de ácidos graxos e fosfatídicos, fundamentais para a síntese de membrana celular. Um dado peculiar que destacamos na Figura 17 se refere às diferenças estruturais entre os dois putativos agrupamentos envolvidos com a síntese de xanthomonadina. Os genes XAC4094, XAC4095 e XCC4006, em particular, representam cópias da região N-terminal dos genes homólogos XAC4093, XCC4005 e CDS 4 de XOO. O gene 14 de XOO é homólogo a um dos genes inseridos na região **XcUC1**. O pequeno agrupamento em ambas as cepas de *Xylella* está flanqueado por uma região de instabilidade, XFP4 em XF-CVC e PInDel 6 em XF-PD. O ponto de fragmentação de ambos os agrupamentos nas cepas de *Xylella*, correspondem a regiões com seqüências repetitivas no DNA das bactérias, homóloga à mesma região descrita em XOO (GOEL *et al.*, 2002).

Figura 17: Análise comparativa da composição de genes envolvidos com biossíntese de xanthomonadina em cinco bactérias da família Xanthomonadaceae. O esquema dos genes envolvidos na síntese de xanthomonadina está baseado no modelo proposto para XOO (GOEL *et al.*, 2002). Em rosa estão representados os genes que compõem o grande agrupamento gênico, relacionado à biossíntese de xanthomonadina, enquanto que em amarelo estão dispostos o pequeno agrupamento dos genes, potencialmente relacionados a esta função. As barras horizontais delimitam estas regiões. Note que nas cepas de *Xylella* são observadas várias rupturas da estrutura sintênica dos genes que participam da síntese xanthomonadina em *Xanthomonas* (triângulos vermelhos). A filogenia esquematizada ao lado da figura é baseada na filogenia de espécies, tomando como referencial a sequência do gene ribossomal 16S. A simbologia usada está descrita abaixo do esquema. Os números dos genes indicados no esquema do agrupamento de XOO compreendem aos genes: 1, Proteína hipotética conservada; 2, proteína hipotética; 3, proteína hipotética; 4, proteína de membrana; 5, Proteína transmembranar putativa; 6, proteína hipotética; 7, proteína hipotética; 8, *Lysophosphatidic acid acyltransferase*; 9, proteína hipotética; 10, proteína carreadora de grupos acil; 11, halogenase; 12, dipeptidil peptidase; 13, proteína hipotética; 14, proteína hipotética.

4.1.3. Determinação de PInDels nos quatro genomas comparados

Com o objetivo de detectar rearranjos não descritos anteriormente nos genomas de XAC, XCC, XF-CVC e XF-PD, comparamos estes quatro genomas e focamos a busca por regiões que caracterizassem inserções ou deleções gênicas, as quais denominamos **PInDels**. As Figuras 18 e 19 esquematizam os resultados obtidos a partir das análises estruturais, do conteúdo de GC e da variação de códons ao longo dos quatro genomas, subdivididos dois a dois pelo gênero.

Nossas análises indicam a ocorrência de modificações por rearranjos e ocorrências de 11 PInDels no genoma de XF-CVC e 10 no genoma de XF-PD. Em XF-CVC já haviam sido descritas seqüências de quatro fagos completos integrados no seu genoma (SIMPSON *et al.*, 2000), que podem ter sido adquiridos de bactérias que habitam ou co-habitam o solo, ou mesmo diretamente por sifofagos (BHATTACHARYYA *et al.*, 2002a, BHATTACHARYYA *et al.*, 2002b). Além dos genes que codificam para proteínas de fagos dispostas nas regiões que compreendem estas quatro completas inserções, outras 22 CDSs que codificam proteínas para a mesma função foram encontrados dispersas, particularmente, nas 11 PInDels encontradas no genoma de XF-CVC, (Figura 18 e Tabela 13). Oito destas 11 regiões já haviam sido descritas (NUNES *et al.*, 2003), e estão indicadas na Tabela 13. As PInDels adicionais que identificamos foram designadas de PInDel 6, 9 e 10 respectivamente. É importante observar que a PInDel 6 corresponde a uma região repetitiva invertida, composta basicamente por genes hipotéticos e de fago, sugerindo um remota inserção seguida de perda parcial.

Nossas análises também buscaram a presença de integrases nos genomas analisados. Estas integrases correspondem a proteínas de bacteriófagos e são geralmente responsáveis pela integração do material genético exógeno no DNA hospedeiro (BENNETT, 2004). Todos os quatro fagos completos integrados no genoma de XF-CVC apresentam estas integrases, e para completar, outras integrases foram encontradas flanqueando seis das 11 PInDels que identificamos (Tabela 13).

As 15 integrases que identificamos foram divididas em 4 grupos de acordo com a metodologia descrita no item 3.1. Observamos que todas as integrases classificadas como pertencentes ao grupo II estão flanqueando um fago ou uma PInDel (Tabela 13). Além disso, a maioria das integrases do grupo I e praticamente todas as integrases do grupo IV se apresentam ao lado de um gene que codifica para um tRNA (Tabela 13). Um outro dado importante foi a observação de que seis integrases, numeradas como 3, 4, 7, 13, 16 e 17, estão posicionadas em XF-CVC exatamente na borda das regiões caracterizadas como rearranjos genômicos, em relação ao genoma de XF-PD. Em um trabalho anterior, apenas 4 delas haviam sido posicionadas (VAN SLUYS *et al.*, 2003). Portanto, similaridade de seqüências e sitio de integração de integrases estão correlacionados, demonstrando possivelmente que todos os rearranjos entre as cepas de *Xylella* (CVC e PD) podem ter sido decorrentes da inserção de fagos.

Embora nenhuma seqüência de bacteriófago completo tenha sido encontrada no genoma de XF-PD, um total de 139 CDSs de fago foram observadas dispostas em 10 PinDels, oito das quais previamente descritas (VAN SLUYS *et al.*, 2003). Análise posicional destas PInDels mostra que a PInDel 2 está localizada no término de replicação (Figura 18) e que provavelmente a PInDel 6 corresponde a uma fusão de duas distintas inserções de fago (VAN SLUYS *et al.*, 2003). A correlação entre as PInDels de XF-CVC e XF-PD está detalhada na Tabela 13.

Em *Xanthomonas* o número de genes que codificam para proteínas de fagos é bem menor do que o descrito em *Xylella* (DA SILVA *et al.*, 2002). Nenhuma inserção de fago havia sido determinada no genoma de XAC, enquanto que uma inserção de fago LΦ foi encontrada no genoma de XCC. Usando a metodologia para determinação de regiões de deleção e inserção gênica, identificamos 13 PinDels no genoma de XAC e apenas 7 no genoma de XCC (Figura 19). A grande maioria destas PInDels estão flanqueadas por integrases de fago, com exceção das PinDels 3 e 7 de XCC que foram determinadas exclusivamente com base na variação de composição de GC e de códons usados. As PinDels numeradas de 1, 2, 5, 6, 7, 8, 9, 10, 12 e 13 de XAC e as PinDels numeradas de 3, 4, 5 e 6 em XCC são exclusivas destas bactérias,

e provavelmente foram adquiridas ou perdidas após o evento de divergência das espécies. A PInDel 3 de XAC é homóloga à PInDel 1 de XCC, assim como as PinDels 4 e 11 de XAC são homólogas às PinDels 2 e 6 de XCC, respectivamente.

A análise da composição de PinDels das cepas de *Xanthomonas* revelou outros dados interessantes. Por exemplo, uma das cópias do SS-II denominada como xcs, apresenta-se flanqueada por dois agrupamentos gênicos conservados, denominados como A e B (Figura 19). O mesmo tipo de estrutura foi observado por genes que flanqueiam o SS-III (T3SS) de *Xanthomonas*, e neste caso estas regiões flanqueadoras foram denominadas de C e D. Outra característica peculiar sobre as duas cópias do SS-II e a cópia do SS-III presente em ambas as cepas de *Xanthomonas*, é que embora seus posicionamentos sejam distintos, as cópias do SS-II são flanqueadas pelas recombinases *xerC* e *xerD*. Estas enzimas estão envolvidas com rearranjos clássicos do DNA bacteriano (HENDRICKS *et al.*, 2000) (Figura 19).

Finalmente, o agrupamento gênico nomeado de E, que compreende os genes *fimT*, *uvrB* e um tRNA^{val}, flanqueia os genes que codificam para o SS-IV (T4SS) em XAC (Figura 19). O agrupamento E se encontra em XF-CVC entre os rearranjos II e III e a montante do rearranjo II de XF-PD (Figura 18).

As integrases de XAC foram classificadas em três grupos baseados na similaridade de seqüências, sendo algumas destas integrases homólogas às descritas em *Xylella* (Tabela 13). Embora não tenhamos conseguido determinar um padrão de similaridade de seqüências de integrases com seus respectivos sítios de integração, como observado em XF-CVC, em XAC 11 das 18 integrases descritas flaqueiam as PInDels que identificamos. O mesmo pode ser observado com 9 das 11 descritas no genoma de XCC. Assim, embora não tenhamos determinado possíveis seqüências alvos de integração destas proteínas de fago, é provável que a presença de integrases sinalize possíveis eventos por aquisição gênica. É importante destacar que estas regiões são em geral seguidas de um número muito grande de genes hipotéticos com variação nas suas composições de nucleotídeos quando comparado com a média global do genoma.

**Figura 18 – Estrutura e composição gênica dos genomas das bactérias do gênero *Xylella*.
Ver legenda a seguir**

Figura 18: Os esquemas representativos dos genomas das cepas de XF-CVC e XF-PD foram gerados com base nas análises de composição estrutural, conteúdo de GC e variação de códons, como indicado. A linha vertical pontilhada indica o provável término de replicação. A simbologia utilizada está descrita abaixo. Os números arábicos menores posicionados acima das setas pretas correspondem a: 1- Complexo NADH desidrogenase (genes *nuo*); 2 – Proteína ligadora de cálcio do tipo hemolisina; 3 – Agrupamento de genes envolvidos com peptideoglicanos e sáculos mureínicos; 4 – genes *xylA*, *nahA*, *dadA* e *alr*; 5 – Proteína secretada do tipo hemaglutinina; 6 – Agrupamento de genes envolvidos com metabolismo de arginina e prolina; 7 - Proteína ligadora de cálcio do tipo hemolisina; 8 – Agrupamentos de genes que codificam para proteínas ribossomais; 9 - Proteína secretada do tipo hemaglutinina; 10 – Genes de goma; 11 – Bacteriocina; 12 – RNA polimerase e agrupamento de proteínas ribossomais; 13 a 15 – Sistema de restrição e modificação agrupamentos 1, 2 e 3 respectivamente; 16 - Proteína secretada do tipo hemaglutinina.

Fagos		Genes/agrupamentos gênicos conservados
PlnDel (int) (GC)		<i>Operon</i> RNA ribossomal (16-5-23S)
Integrase do grupo A		Agrupamento de genes de flagelo + MCP
Integrase do grupo B	1- 18	Número das integrases
Integrase do grupo C	1-13	Classificação das PlnDels
Rearranjos genômicos	TxSS	Sistemas secretórios TxSS (x = 1, 2, 3 ou 4)
Regiões destacadas	SPI-7	Ilha genômica sintênica em <i>Salmonella</i>
Fluxo de transcrição de um <i>operon</i>		Agrupamento envolvido com assimilação de nitratos
Direção de transcrição do tRNA-aa		Região repetitiva e invertida
Zoom de uma região destacada	!	Término de Replicação

**Figura 19 – Estrutura e composição gênica dos genomas das bactérias do gênero *Xanthomonas*.
Ver legenda a seguir**

Figura 19: Os esquemas representativos dos genomas das cepas de XF-CVC e XF-PD foram gerados com base nas análises de composição estrutural, conteúdo de GC e variação de códons, como indicado. A linha vertical pontilhada indica o provável término de replicação. A simbologia utilizada está descrita abaixo. Os números arábicos menores posicionados acima das setas pretas correspondem a: **1** – proteínas de fago e transposases; **2** - Transposases; **3** – Degradação de compostos aromáticos; **4** – Agrupamento gênico envolvido com metabolismo de biotina; **5** - Genes envolvidos com degradação de poli e oligossacarídeos; **6** - Agrupamento de genes envolvidos com peptideoglicanos e sáculos mureínicos; **7** – Agrupamentos de genes que codificam para proteínas ribossomais; **8** – HMS, hemaglutinina e hemolisina; **9** – Agrupamento de genes envolvidos com metabolismo de histidina; **10** – Genes *rpf*; **11** – Genes *tsr*; **12** - Proteína ligadora de cálcio do tipo hemolisina; **13** – Agrupamento de genes envolvidos com metabolismo de arginina e prolina; **14** – Genes de goma; **15** - Complexo NADH desidrogenase (genes *nuo*); **16** – Agrupamento de genes envolvidos com metabolismo de cobalamina; **17** – Xanthomonadina.

Fagos		Genes/agrupamentos gênicos conservados
PlnDel (int) (GC)		<i>Operon</i> RNA ribossomal (16-5-23S)
Integrase do grupo A		Agrupamento de genes de flagelo + MCP
Integrase do grupo B	1- 18	Número das integrases
Integrase do grupo C	1-13	Classificação das PlnDels
Rearranjos genômicos	TxSS	Sistemas secretórios TxSS (x = 1, 2, 3 ou 4)
Regiões destacadas	SPI-7	Ilha genômica sintênica em <i>Salmonella</i>
Fluxo de transcrição de um <i>operon</i>		Agrupamento envolvido com assimilação de nitratos
Direção de transcrição do tRNA-aa		Região repetitiva e invertida
Zoom de uma região destacada	!	Término de Replicação

Tabela 13 – Classificação das integrases e determinação das PInDels de XF-CVC e XF-PD.

Número Integrase	Genoma	Número CDS	Grupo de Integrase			AA	Loc.	Inserção		Início	Fim	Rearranjos		Homólogos próximos		Ilhas já ^b descritas
			Inserção	Blast	Fita			Nome	Loc.			Número	Loc.	XAC	XCC	
1	XF-CVC	XF0480	A	IV	←	N	Up	PInDel 1	Montante	XF0480	XF0558	----	----	XAC2628	XCC3012	pGI2
2	XF-CVC	XF0631	C	III	→	----	----	PInDel 3	Interna	XF0623	XF0648	----	----	XAC2222/2183	----	GI3
3	XF-CVC	XF0678	A	I	←	V	Up	XFP1	Montante	XF0678	XF0733	I	Jusante	----	----	----
4	XF-CVC	XF0968	A	I	←	V	Up	----	----	XF0968	----	II	Jusante	----	----	----
5	XF-CVC	XF1425	<i>xerD</i>	Recomb	←	----	----	----	----	XF1425	----	----	----	XAC3551	XCC0654	----
6	XF-CVC	XF1483	<i>xerC</i>	Recomb	→	----	----	----	----	XF1483	----	----	----	XAC0636	XCC3497	----
7	XF-CVC	XF1555/56	B	II	←	----	----	XFP3	Montante	XF1555	XF1596	III	Jusante	----	----	----
8	XF-CVC	XF1642	B	II	←	----	----	XFP4	Montante	XF1642	XF1711	----	----	----	----	----
9	XF-CVC	XF1718	A	III	→	G	Up	PInDel 7	Montante	XF1718	XF1754	----	----	----	XCC2110	pI1
10	XF-CVC	XF1754	B	III	→	----	----	PInDel 7/8	Ju/Montante	XF1718	XF1754	----	----	----	XCC2110?	GI5
11	XF-CVC	XF1789	A	I	←	T	Down	PInDel 8	Jusante	XF1754	XF1793	----	----	----	----	GI5
12	XF-CVC	XF2028	B	Singlet ^a	←	----	----	PInDel 9	Montante	XF2028	XF2081	----	----	XACb0061	XCC1630	----
13	XF-CVC	XF2131/32	B	II	→	----	----	PInDel 11	Jusante	XF2108	XF2132	IV	Jusante	----	----	pGI6
14	XF-CVC	XF2288	B	II	←	----	----	PInDel 12	Montante	XF2288	XF2309	----	----	----	----	----
15	XF-CVC	XF2309	A	I	→	G	Down	PInDel 12	Jusante	XF2288	XF2309	----	----	----	----	----
16	XF-CVC	XF2478	B	I	→	----	----	XFP2	Montante	XF2478	XF2530	V	Jusante	----	----	----
17	XF-CVC	XF2530	B	II	→	----	----	XFP2	Jusante	XF2478	XF2530	V	Jusante	----	----	----
18	XF-CVC	XF2761	A	IV	←	PRHK	Down	PInDel 13	Montante	XF2761	XF2773	----	----	XAC2628	XCC3012	----
8	XF-PD	PD0384	B	II	→	----	----	PInDel 1	Jusante	PD0363	PD0384	----	----	----	----	GI2
5	XF-PD	PD0652	<i>xerD</i>	Recomb	←	----	----	----	----	----	----	----	----	XAC3551	XCC0654	----
6	XF-PD	PD0700	<i>xerC</i>	Recomb	→	----	----	----	----	----	----	----	----	XAC0636	XCC3497	----
8	XF-PD	PD0764	B	II	←	----	----	PInDel 2	Jusante	PD0764	PD0772	III	Jusante	----	----	GI1
12	XF-PD	PD0789	C	Singlet	→	----	----	DePo #	Montante	----	----	----	----	XACb0061	XCC1630	----
9	XF-PD	PD0990	B	III	←	----	----	PInDel 5	Montante	PD0990	PD1019	----	----	XAC2183	----	GI2
13	XF-PD	PD1019	B	II	→	----	----	PInDel 5	Jusante	PD0990	PD1019	----	----	----	----	GI2
1	XF-PD	PD1075	A	IV	←	T	Up	PInDel 6	Montante	PD1075	PD1139	----	----	XAC2628	XCC3012	GI4
11	XF-PD	PD1078	A	I	→	G	Down	PInDel 6	Montante	PD1075	PD1139	----	----	----	----	GI4
8	XF-PD	PD1139	B	II	→	----	----	PInDel 6	Jusante	PD1075	PD1139	----	----	----	----	GI4
8	XF-PD	PD1196	B	II	→	----	----	PInDel 7	Jusante	PD1166	PD1196	----	----	----	----	pI1
14	XF-PD	PD1320 *	C	II	←	----	----	----	----	PD1320	PD1323	----	----	----	----	----
16	XF-PD	PD1495	A	I	→	V	Down	----	----	----	----	V	Jusante	----	----	----
1	XF-PD	PD1605 *	A	IV	→	N	Down	PInDel 8	Jusante	PD1592	PD1605	----	----	XAC2628	XCC3012	GI5
16	XF-PD	PD1732 *	A	I	→	V	Down	PInDel 9	Jusante	PD1714	PD1732	I	Jusante	----	----	----

^a Singlet – denota integrases de seqüências únicas e que não puderam ser classificadas em nenhum dos grupos destacados; # DPO, integrase flanqueando o agrupamento de genes envolvidos com Degradação de Poli e Oligossacarídeos DPO. Fita: → (+) e ← (-). Loc, localização. ^b Ilhas genômicas foram descritas anteriormente (NUNES *et al.*, 2003), com os seguintes códigos: GI, ilhas genômicas; pGI, ilhas genômicas putativas; pl, regiões de integração de plasmídeos. *, genes que apresetam problemas do tipo mutação pontual ou alteração no canal de leitura.

4.2. Análise do sistema de osmoproteção e osmorregulação e as vias de degradação de compostos aromáticos codificados no genoma de XAC

4.2.1. Análise do sistema de osmoproteção e osmorregulação

Com o objetivo de investigar a complexidade do sistema de osmorregulação e osmoproteção em *Xanthomonas*, realizamos uma análise detalhada dos genes que estão hipoteticamente relacionados a estes processos. A maquinaria celular de osmoproteção possibilita a manutenção adequada do turgor celular em resposta a alterações na composição do meio externo (ROESSER & MULLER, 2001). Geralmente esta resposta está associada a mecanismos moleculares que controlam o transporte ou síntese de alguns compostos capazes de balancear este turgor (NORRIS & MANNERS, 1993). Um destes mecanismos se baseia no efluxo de água do meio intracelular, acarretando redução na pressão interna, por intermédio de transportadores específicos denominados aquaporinas (FISCHBARG *et al.*, 1990, CALAMITA *et al.*, 1995, CALAMITA *et al.*, 1997, CALAMITA, 2000). Outro mecanismo modelo usa o influxo de íons ao interior da célula proporcionando a redução do efluxo de água, evento este comumente observado em bactérias halofíticas e halotolerantes (VENTOSA *et al.*, 1998a, VENTOSA *et al.*, 1998b, PFLUGER & MULLER, 2004). As bactérias ainda são capazes de transportar ou sintetizar moléculas capazes de aumentar sua osmolaridade interna. Moléculas osmoprotetoras são altamente solúveis e incapazes de induzir estresse celular, tendo a capacidade de serem posteriormente metabolizadas e aproveitadas pela célula como fonte alternativa de carbono ou nitrogênio (DOMINGUE, 1991). Os solutos usados por microorganismos como osmoprotetores compreendem um repertório ubíquo de compostos, tais como açúcares (trealose), amino ácidos livres (prolina e glutamato), polióis (glicerol and glucosilglicerol), derivados de aminoácidos (betaina e ectoina), aminas quaternárias e seus análogos sulfonados (glicinabetaina, carnitina and dimetilsulfoniopropionato), ésteres de sulfato (colina-O-sulfato) e di-aminoácidos *N*-acetilados e pequenos peptídeos (*N*-acetilornitina and *N*-acetilglutaminilglutamina) (GARCIA-PEREZ & BURG,

1991, KEMPF & BREMER, 1998, WOOD *et al.*, 2001b). A Figura 20 mostra que XAC é capaz de transportar ou sintetizar vários destes compostos osmoprotetores, perfil que verificamos a partir dos dados de anotação de seu genoma (DA SILVA *et al.*, 2002).

Ver legenda a seguir

Figura 20: Genes e sistemas envolvidos com osmoproteção encontrados no genoma de XAC. Os quadrados em cinza escuro mostram as moléculas possivelmente envolvidas com osmoproteção em XAC. Os transportadores indicados na cor branca estão ausentes no genoma de XAC (*Trk*, *ProU* e *AqzP*). *MscL* (*large-conductance mechanosensitive channel*), envolvido com transdução física de estresse na membrana celular em resposta a um gradiente eletroquímico. *OmpC* e *OmpF* (*outer membrane proteins*), proteínas de membrana externa envolvidas com internalização de moléculas específicas. *ProP*, transportador de prolina/betaina. *ProU*, transportador de glicina-betaina e L-prolina. *BetT*, transportador de alta afinidade por colina. *GltT*, proteína simporte de transporte de glutamato. *FruA* + *PtsKHI*, PTS de frutose. *KdgT*, sistema transportador de 2-ceto-3-deoxi-D-gluconato. *KdgIDKA*, genes envolvidos com metabolismo e conversão de 2-ceto-3-deoxi-D-gluconato em piruvato. *ExuT*, transportador de exuronato. *FucP*, transportador de glicose, galactose e fucose. *GlupP*, transportador de glicose e galactose. Demais genes indicados e relacionados ao influxo de K^+ estão descritos na Tabela 14.

O processo de adaptação de bactérias a um ambiente de alta osmolaridade (*upshock*) ocorre em duas fases (KEMPF & BREMER, 1998). Na primeira fase se verifica intensa internalização de íons K^+ mediada por transportadores específicos, que respondem a alterações no volume celular. Esta aquisição se caracteriza como sendo a resposta mais rápida de estímulo em condições de *upshock* relatados em microorganismos, ficando aquém apenas do acúmulo de prolina (CSONKA & EPSTEIN, 1996). O potássio, uma vez internalizado coordena a síntese de outros osmoprotetores, caracterizando-se, portanto, como um mensageiro secundário da resposta de adaptação em condições de estresse osmótico (EPSTEIN *et al.*, 1993). Em *E.coli*, um organismo modelo de estudo para maquinaria de osmorregulação, três sistemas distintos estão envolvidos com a internalização de K^+ (*kup*, *trk* e *kdp*), enquanto que apenas um sistema está relacionado com o seu efluxo (*kef*) (Figura 21). Dentre estes sistemas, apenas o codificado pelos genes *trk* está ausente em XAC. A função de cada gene, os respectivos genes homólogos em XAC, juntamente com os valores de identidade e similaridade estão apresentados detalhadamente na Tabela 14.

Kup e Trk são constitutivamente sintetizados em *E.coli* ao passo que Kdp (K^+ ***dependent growth***) é influenciado pelas concentrações de K^+ do meio, representando, portanto, o sistema de maior afinidade por este cátion ($K_M = 2 \mu M$) (SILVER, 1996). O sistema Kdp é formado por três proteínas estruturais de membrana, codificadas pelos genes *kdpABC*, além de duas outras que compõem um sistema de dois componentes, com *kdpD* sendo a proteína sensora e *kdpE* o regulador de resposta (SILVER, 1996). Em XAC estes genes estão contidos em uma só região (XAC0756-0760) formando duas unidades transcricionais com um promotor localizado a montante do gene *kdpA* e outro a montante do gene *kdpD* (Figura 22A). Em *E.coli* este sistema é expresso apenas em resposta a baixa concentração de K^+ , sendo regulado pelos produtos dos genes *kdpDE* (WALDERHAUG *et al.*, 1992). O sistema *Kup*, previamente denominado *TrkD*, corresponde ao sistema constitutivo de menor capacidade de captação K^+ e é codificado por um único gene cuja proteína é composta por 12

domínios transmembrana com um motivo adicional de ligação ao DNA (SCHLEYER & BAKKER, 1993). Diferentemente dos outros osmoprotetores, K⁺ em excesso no citosol pode afetar importantes interações entre proteínas ou mesmo entre proteínas e DNA (STUMPE *et al.*, 1996), causando alterações no metabolismo e na coordenação do ciclo celular.

O sistema *kdp* pode estar diretamente relacionado à adaptação de XAC. Recentemente foi isolado um mutante do gene *kdpA* que exibiu um fenótipo aparente de menor virulência em *Citrus limonia* (J.C.F de OLIVEIRA e M.L de LAIA, comunicação pessoal). Este mutante selecionado de uma biblioteca gerada por inserção aleatória de transposon apresenta o gene *kdpA* interrompido próximo a região que codificadora do aminoácido 287, entre o quarto e quinto domínio transmembranar (Figura 22B). A proteína codificada pelo gene *kdpA* é responsável pelo reconhecimento do K⁺ (BUURMAN *et al.*, 1995). Assim, o fenótipo de redução de hiperplasia e diminuição do processo de *watersoaking* observados quando o mutante *kdpA*⁻ é inoculado em folhas de *Citrus limonia* parece estar correlacionado com sua menor taxa de crescimento *in vitro*, reforçando a importância deste gene para os fenômenos adaptativos de XAC (J.C.F de OLIVEIRA e M.L de LAIA, comunicação pessoal). O fato da análise deste mutante estar ainda em fase inicial leva à necessidade de experimentos adicionais para demonstrar a importância relativa do sistema *kdp* no processo infeccioso/adaptativo por XAC.

Embora os sistemas de captação de potássio sejam importantes para a sobrevivência de bactérias, os genomas de bactérias do gênero *Xylella* codificam apenas os genes *kup* (MEIDANIS *et al.*, 2002) (Figura 21). Esta peculiar ausência dos demais sistemas pode ser talvez justificada pelo fato desta bactéria habitar um ambiente de baixa osmolaridade como o xilema. Além disso, a presença de um único sistema de captação de potássio pode estar correlacionado ao crescimento fastidioso exibido por esta espécie bacteriana.

A segunda fase de adaptação a condições de *upshock* ocorre após a internalização de potássio e é mediada pela síntese de glutamato no citosol bacteriano. Em organismos como *E.coli* e *Salmonella typhimurium*, esta síntese

se inicia aproximadamente em um minuto após *upshock* osmótico mediado pelo influxo de K^+ sendo totalmente dependente deste processo para manter a concentração de K^+ durante o estado estacionário (MCLAGGAN *et al.*, 1994, YAN *et al.*, 1996). A análise do genoma de XAC mostra que esta bactéria apresenta a via de síntese de glutamato completa (DA SILVA *et al.*, 2002) (Figura 20). Assim, as análises e resultados iniciais discutidos acima evidenciam que XAC apresenta uma complexa maquinaria de osmorregulação e osmoproteção capaz de favorecer sua adaptação em diferentes condições fisiológicas (Figura 20), e esta flexibilidade pode até ser correlacionada com a sua excelente capacidade de sobrevivência tanto como organismo epífita, fora do vegetal, como agente patogênico.

Figure 22: Sistemas envolvidos com osmoregulação mediada por potássio, encontrados no genoma de XAC. Os sistemas homólogos de *E. coli* verificados em XAC estão coloridos em rosa claro. Em branco se destacam os sistemas ausentes em XAC. O sistema kup indicado em amarelo é encontrado em XAC e XF-CVC. Esquema adaptado do modelo proposto para *E. coli* (SILVER, 1996). A função de cada um dos genes está detalhado na Tabela 14. GS-X, Glutathione-S-transferase.

Tabela 14: Genes envolvidos com osmorregulação mediada por influxo de potássio.

Gene	Produtoa	ECO Gene_id	Id / sim (%)c	XAC Gene_id	Funçãoa	Localização Celulare	Fluxo de K+	Ref
<i>Kup</i> (<i>TrkD</i>)	Sistema de transporte de baixa afinidade TC 2.A.72	b3747 (P63183)	48 / 65	XAC1484	Responsável pelo sistema de transporte com baixa afinidade por potássio para o interior da célula, com o provável co-transporte de prótons.	Proteína integral de membrana, ligada à membrana interna.	Influxo	(ZAKHARYAN & TRCHOUNIAN, 2001)
<i>TrkH</i>	Proteínas que participam do transporte de potássio mediado pelo sistema <i>Trk</i>	b3849 (P21166)	---	---	Sistema de transporte de baixa afinidade por potássio. <i>TrkH</i> e <i>trkG</i> interagem com <i>trkA</i> (liga NAD ⁺ e NADH.) e requerem <i>trkE</i> para ter atividade. <i>TrkE</i> ainda apresenta função incerta, mas é requerido por <i>trkG</i> e <i>trkH</i> , o mesmo ocorre com <i>trkE</i>	Proteína integral de membrana, ligada à membrana interna.	Influxo	(SCHLOSSER <i>et al.</i> , 1995)
<i>TrkG</i>		b1363 (P23849)	---	---		Proteína integral de membrana, ligada à membrana interna.	Influxo	(SCHLOSSER <i>et al.</i> , 1995)
<i>TrkE</i>		---	---	---			Influxo	(SCHLOSSER <i>et al.</i> , 1993)
<i>TrkA</i>		b3290 (P23868)	---	---		Membrana interna, com sítio de ligação para <i>trkG</i> e <i>trkH</i>	Influxo	(BOSSEMEYER <i>et al.</i> , 1989, SCHLOSSER <i>et al.</i> , 1993)
<i>KdpA</i>	Cadeia A, B e C, respectivamente, com função de ATPase que atuam no transporte ativo de potássio EC 3.6.3.12 (A, B e C)	b0698 (P03959)	35 / 54	XAC0756	A e B são componentes de um sistema de transporte ativo com alta afinidade por potássio. A hidrólise do ATP está associada a uma troca de hidrogênio por potássio. Atividade catalítica ATP + H ₂ O + K ⁺ (fora) = ADP + fosfato + K ⁺ (dentro). A subunidade C parece estar envolvida com a união das subunidades A e B na formação do complexo.	Proteínas integrais de membrana ligadas à membrana interna.	Influxo	(SIEBERS & ALTENDORF, 1988, ALTENDORF <i>et al.</i> , 1998)
<i>KdpB</i>		b0697 (P03960)	59 / 75	XAC0757			Influxo	(SIEBERS & ALTENDORF, 1988, ALTENDORF <i>et al.</i> , 1998)
<i>KdpC</i>		b0696 (P03961)	42 / 54	XAC0758			Influxo	(SIEBERS & ALTENDORF, 1988, ALTENDORF <i>et al.</i> , 1998)
<i>KdpD</i>	Proteína sensora EC 2.7.3.-	b0695 (P21865)	36 / 53	XAC0759	Membro do sistema de dois componentes <i>kdpD/kdpE</i> envolvidos na regulação do operon <i>KDP</i> . Pode atuar como uma proteína quinase de membrana fosforilando <i>kdpE</i> em resposta a um sinal do ambiente.	Proteína integral de membrana, ligada à membrana interna.	Influxo	(SIEBERS & ALTENDORF, 1988, ALTENDORF <i>et al.</i> , 1998)
<i>KdpE</i>	Proteína regulatória	b0694 (P21866)	49 / 68	XAC0760	Membro do sistema de dois componentes <i>kdpD/kdpE</i> envolvidos na regulação do operon <i>KDP</i> .	Citoplasma	Influxo	(SIEBERS & ALTENDORF, 1988, ALTENDORF <i>et al.</i> , 1998)
<i>KefB</i>	Sistema de proteínas regulados por glutatona em associação com a saída de potássio	b3350 (P45522)	41 / 60	XAC2291	Sistema de transporte que facilita a saída de potássio, possivelmente associada à entrada de prótons.	Proteína integral de membrana, ligada à membrana interna.	Efluxo	(BAKKER <i>et al.</i> , 1987, MUNRO <i>et al.</i> , 1991)
<i>KefC</i>	Antiporter K(+)/H(+)	b0047 (P03819)	43 / 61	XAC4080			Efluxo	(BAKKER <i>et al.</i> , 1987, MUNRO <i>et al.</i> , 1991)

^a A função de cada um dos genes foi obtida a partir dos artigos de referência indicadas na tabela ^b Produtos e funções obtidas a partir do UniProtKB/Swiss-Prot; ^{c,d} Sequências obtidas do sequenciamento genômico de XAC e XF-CVC; ^e informações obtidas do UniProtKB/Swiss-Prot protein family; TC – Número de classificação dos transportadores (IUPAC- IUBMB); EC – Número de classificação de enzimas (IUPAC- IUBMB).

Figura 22: Análise do regulon que codifica proteínas envolvidas com internalização de K^+ em XAC. (A) Os dois promotores relacionados com a transcrição dos genes estão indicados por uma pequena seta preta a montante de *kdpDE* (em salmão) e a montante de *kdpABC* (em amarelo). **(B)** – Esquema da inserção do transposon (em azul) no gene *kdpA* no mutante *kdpA⁻*. A inserção rompeu a alça externa da proteína que se localiza entre os domínios transmembrânicos 4 e 5, próximo ao resíduo 287 da proteína. A determinação exata do ponto de inserção foi obtida por seqüenciamento das pontas do transposon, usando oligonucleotídeos específicos (setas pretas). ME, meio externo; EP, espaço periplasmático; MI, meio interno.

4.2.2 Análise das vias de degradação de compostos aromáticos

Entre os compostos aromáticos mais freqüentes na biomassa terrestre, está a lignina, um complexo polímero aromático cuja degradação é realizada por um número limitado de microorganismos que expressam enzimas específicas como as *laccases* e peroxidases de lignina, entre outras (ORTH *et al.*, 1993, MARTINEZ *et al.*, 2005). Da mesma maneira, apenas um limitado grupo de organismos, é capaz de degradar outras classes de compostos aromáticos encontrados no ambiente, tais como os hidrocarbonetos aromáticos policíclicos (YANG *et al.*, 1994).

A degradação aeróbica do anel aromático, em geral, ocorre em duas fases. Na primeira, ocorre modificações no anel aromático tais como mono ou di-oxigenações. Na segunda, ocorre a clivagem do anel, seguindo-se de reações para geração de intermediários do ciclo do ácido tricarboxílico. A clivagem do anel é catalisada por dioxigenases e pode ocorrer de acordo com três rotas, denominadas de clivagem orto (via do β -cetoadipato), meta ou gentisato, dependendo da ligação que é clivada (HARWOOD & PARALES, 1996).

A rota da meta-clivagem é importante na degradação de substratos catecolíticos metilados (por exemplo: xileno e tolueno), que por serem utilizados em diversos processos industriais, são contaminantes ambientais relevantes (BARBIERI *et al.*, 1993, KIM *et al.*, 2002). A rota da meta-clivagem tem sido estudada em bactérias utilizadas em processos de biodegradação, tais como *Pseudomonas putida*, sendo que os genes que codificam as enzimas desta via são plasmidiais (DIAZ *et al.*, 2001, JIMENEZ *et al.*, 2002). A via do β -cetoadipato, por outro lado, é de ocorrência freqüente em diversos grupos taxonômicos de bactérias e fungos, estando os genes das enzimas desta via geralmente codificados no cromossomo principal. Esta via tem particular importância na degradação de compostos aromáticos produzidos por plantas e na geração de substratos essenciais para a síntese de aminoácidos aromáticos (HARWOOD & PARALES, 1996)

Com objetivo de investigarmos com mais detalhe um dos prováveis processos importantes na interação de XAC com seu hospedeiro vegetal, passamos a analisar as possíveis vias de degradação de compostos aromáticos codificadas no genoma de XAC, utilizando como referência as vias descritas para *P. putida* e *E.coli*. A Figura 23 mostra as rotas de degradação de três compostos aromáticos pela via do β -cetoadipato que deciframos a partir da análise do genoma de XAC. Os genes relacionados a cada passo da vias de degradação estão listados na Tabela 15. A organização destes genes no genoma de XAC foi comparada com outras bactérias (Figura 24). É interessante observar que a maioria dos genes relacionados a estas vias de degradação apresentam-se agrupados (XAC0349-XAC0373), exceto dois deles (*pcaI* e *pcaJ*) que se localizam inseridos em uma região distinta do genoma. Apesar do agrupamento em XAC ser interrompido por alguns genes não-relacionados a estas vias, observamos certa sintonia comparativamente aos agrupamentos contendo genes análogos em *A.tumefaciens*, *C.crescentus*, *R. metallidurans* e *P.fluorescens*. Vale ressaltar que o agrupamento de *Xanthomonas* é mais diversificado quanto a composição de genes comparativamente a estas quatro bactérias (Figura 24).

A comparação cuidadosa dos genomas de *Xanthomonas* e *Xylella*, indicou claramente a ausência das vias descritas na Figura 23 no genoma de *Xylella fastidiosa* (Figura 25). Observamos, contudo, que alguns poucos genes que flanqueiam ou estão inseridos no agrupamento gênico relacionado a via do β -cetoadipato de XAC são conservados no genoma de *Xylella*, tais como algumas CDS hipotéticas e CDS relacionadas a síntese de glicerol (Figura 25).

Além das vias de degradação mostradas na Figura 23, também identificamos genes que potencialmente estão relacionados a outras vias do metabolismo de compostos aromáticos. Estas vias estão sumarizadas na Figura 26. A via que denominados como I está relacionada a síntese de ubiquinol a partir de hidroxibenzaldeído ou hidroxibenzoato. A via II, por sua vez, relaciona-se a síntese de triptofano a partir de dihidroxibenzoato. A via III, corresponde à via explicitada na Figura 23 capaz de converter os dois substratos da via I mais

vanilato em intermediários do ciclo do ácido tricarboxílico. A Via IV envolve a degradação dos aminoácidos aromáticos fenilalanina e tirosina em acetoacetato e piruvato. Finalmente, a via V está envolvida com tolerância ao tolueno.

As vias I, II e V estão codificadas nos genomas de *Xanthomonas* (XAC e XCC) e *Xylella* (XF-CVC e XF-PD), entretanto as vias III e IV estão ausentes em *Xylella*, caracterizando uma possível maior capacidade adaptativa de *Xanthomonas* mediante a presença de compostos aromáticos oriundas do hospedeiro vegetal.

Figura 23: Degradação dos compostos 4-hidroxibenzaldeído, 4-hidroxibenzoato e vanilato pela via do cetoadipato em *Xanthomonas*. Detalhes de cada um dos genes envolvidos nesta via, em azul, estão mostrados na Tabela 15. Note que oxoadipil-coA pode resultar em dois produtos, acetil-coA e succinil-coA, após catálise da enzima codificada pelo gene *pcaF*. Os quadrados em azul e vermelho destacam as modificações bioquímicas promovidas pela enzima *pcaF*.

Tabela 15: Genes de XAC envolvidos com degradação de 4-hidroxibenzaldeído, 4-hidroxibenzoato e vanilato.

CDS	Nome do gene	Produto ^a	Posição Inicial ^b	Posição Final ^c	Tamanho (aa) ^d
XAC0354	xyIC	<i>Benzaldehyde dehydrogenase II</i>	421742	423220	494
XAC0311	vanA	<i>Vanillate O-demethylase oxygenase subunit</i>	373056	372007	351
XAC0363	vanA	<i>Vanillate O-demethylase oxygenase subunit</i>	434670	433597	359
XAC0310	vanB	<i>Vanillate O-demethylase oxidoreductase</i>	371986	371030	320
XAC0356	pobA	<i>P-hydroxybenzoate hydroxylase</i>	424556	425737	395
XAC0368	pcaG	<i>Protocatechuate 3,4-dioxygenase alpha chain</i>	438595	439155	188
XAC0367	pcaH	<i>Protocatechuate 3,4-dioxygenase beta chain</i>	437856	438587	245
XAC0878	pcaH	<i>Protocatechuate 3,4-dioxygenase beta chain</i>	1039714	1040607	299
XAC0369	pcaB	<i>3-carboxy-cis,cis-muconate cycloisomerase</i>	439511	440860	451
XAC0371	pcaC	<i>4-carboxymuconolactone decarboxylase</i>	441683	442093	138
XAC1886	pcaD	<i>Beta-ketoadipate enol-lactone hydrolase</i>	2191610	2192452	282
XAC3577	pcaI	<i>lpsJ protein</i>	4243328	4242699	211
XAC3578	pcaJ	<i>lpsJ protein</i>	4244056	4243328	244
XAC0366	pcaF	<i>Beta-ketoadipyl CoA thiolase</i>	436568	437773	403
XAC0364	gctA	<i>Glutaconate CoA transferase subunit A</i>	434923	435783	288
XAC0365	gctB	<i>Glutaconate CoA transferase subunit B</i>	435783	436568	263

^a O nome dos produtos foi extraído diretamente da anotação do genoma de XAC (<http://genoma4.iq.usp.br/>)

^{b,c} Coordenadas de localização dos genes no genoma de XAC.

^d Número de aminoácidos (aa) do produto codificado pela CDS.

Figura 24: Organização dos genes relacionados a via do β -cetoadipato em XAC e em outras bactérias. Setas em: Laranja, genes envolvidos com catabolismo de β -cetoadipato; Vermelho, Genes com funções regulatórias; Azul, genes envolvidos com transporte de compostos aromáticos; Amarelo, Gene codificador da enzima *Protochatechuete dioxxygenase*; Rosa, Gene codificador da enzima *Hydroxybenzoate hydroxylase*; Verde, Genes relacionados com outras funções de degradação de compostos aromáticos; Branco, genes com funções não relacionadas com as vias de degradação de compostos aromáticos. Barras paralelas em diagonal determinam fragmentação de sentença e espaçamento para outra região do genoma. Os números e as barras posicionadas abaixo dos genes determinam um sub-agrupamento gênico conservado nas bactérias.

Figura 25: Comparação dos genomas de *Xanthomonas* e *Xylella* nas regiões que flanqueiam o agrupamento de genes relacionados à degradação de compostos aromáticos. Os genes relacionados com a síntese de glicerol em *Xanthomonas* (XAC0358-0361) e *Xylella* (XF2266-68, PD1302-04) estão representados por setas pretas. Genes hipotéticos em *Xanthomonas* e conservados em *Xylella* estão representados por setas brancas.

Figura 26: Vias envolvidas no metabolismo de compostos aromáticos em *Xanthomonas* e *Xylella*. Os números romanos denotam e classificam as vias encontradas de maneira comparativa entre as bactérias do gênero *Xanthomonas* (XAC e XCC) e *Xylella* (XF-CVC e XF-PD). Observe que em *Xylella* não foram encontradas as vias III e IV. --- indica a ausência do(s) gene(s) para uma determinada via. Cada um dos quadrados em branco detalha um composto aromático que pode iniciar ou finalizar uma via metabólica. () indicam transportadores de compostos aromáticos específicos. Note que todos os genes envolvidos na via de degradação III, estão contidos em um único agrupamento gênico detalhado na Figura 24.

4.2.2.1. Implicação da via de degradação do β -cetoadipato na virulência de XAC

Ao examinarmos o banco de mutantes de XAC, obtido por inserção aleatória de um elemento de transposição, encontramos um mutante que apresentava o gene *pobA* interrompido pelo transposon. Como já mencionado, o gene *pobA* codifica a enzima p-hidroxybenzoato hidroxilase que converte 4-hidroxibenzoato em protocatecuato (Figura 23 e Tabela 15).

Os clones mutantes do banco investigado foram selecionados por apresentarem fenótipo de virulência alterado em citros. Assim, a mutação do gene *pobA* poderia estar diretamente relacionada a observação de que este mutante, quando inoculado em folhas de *Citrus limonia*, não induz a hiperplasia do tecido como verificado com a cepa selvagem, gerando inclusive uma resposta semelhante à resposta de hipersensibilidade (J.C.F de OLIVEIRA e M.L de LAIA, comunicação pessoal).

Uma hipótese que explicaria este fenótipo do mutante *pobA*⁻ seria a incapacidade de metabolizar 4-hidroxibenzoato, resultando em um acúmulo deste composto, causando um efeito tóxico para o patógeno. Uma investigação preliminar desta hipótese consistiu na avaliação do crescimento da cepa selvagem e mutante *pobA*⁻ na presença de diferentes concentrações de 4-hidroxibenzoato. Resultados preliminares (dados não apresentados) sugeriu que a cepa selvagem é mais tolerante ao 4-hidroxibenzoato do que o mutante *pobA*⁻. Observamos que enquanto a cepa selvagem exibe algum crescimento em 60 ou 90 mM deste composto, o mutante apresenta crescimento apenas na concentração de 30 vezes menor (3 mM). Esta observação é um indício de que o composto 4-hidroxibenzoato afeta o crescimento do mutante *pobA*⁻. Entretanto este efeito é revertido após transferência das bactérias para meio de cultura livre deste composto aromático, demonstrando que o efeito do 4-hidroxibenzoato é bacteriostático para o mutante *pobA*⁻.

Considerando os resultados sobre o fenótipo exibido pelo mutante *pobA*⁻, sugerimos que a via de degradação de compostos aromáticos deve

desempenhar um papel importante na infecção de XAC em hospedeiros compatíveis. Provavelmente compostos aromáticos produzidos pelo hospedeiro durante respostas à infecção, tais como 4-hidroxibenzaldeído, 4-hidroxibenzoato, benzoato, vanilato, protocatecuato, são metabolizados por XAC como fontes de carbono alternativas, de modo similar ao observado em outras bactérias (DISPENSA *et al.*, 1992, EGLAND *et al.*, 1997, GIBSON *et al.*, 1997, BREESE & FUCHS, 1998, DIAZ *et al.*, 2001, GESCHER *et al.*, 2005)

4.3. Construção dos microarranjos de DNA de XAC (XACArray)

Entre as tecnologias que permitem a avaliação do transcrito, os microarranjos de DNA são considerados uma alternativa excelente, se não, a melhor. Esta tecnologia possibilita a análise quantitativa dos níveis de expressão de milhares de genes simultaneamente, permitindo, com relativa facilidade, a avaliação global da expressão gênica em organismos, tecidos ou células (CONWAY & SCHOOLNIK, 2003). Além disso, os microarranjos de DNA são uma ferramenta potente na utilização da enorme quantidade de informações geradas através dos projetos de seqüenciamento de genomas (PANDA *et al.*, 2003).

Em bactérias, a metodologia de microarranjos de DNA tem sido amplamente utilizada para a detecção de alterações na expressão gênica em resposta as mais variadas condições de cultivo, uma vez que a regulação ao nível da transcrição desempenha um papel fundamental em procariotos (LUCCHINI *et al.*, 2001). Outra aplicação freqüente dos microarranjos é na análise comparativa de genomas microbianos (SCHOOLNIK, 2002). Após a conclusão do seqüenciamento completo do genoma de XAC (DA SILVA *et al.*, 2002), iniciamos o planejamento e construção de um microarranjo de DNA de XAC. O fluxograma das etapas realizadas está apresentado na Figura 27 e será detalhado a seguir. Primeiramente selecionamos um conjunto de clones nas bibliotecas genômicas utilizadas no seqüenciamento, o qual fosse representativo das CDS anotadas no genoma de XAC.

A biblioteca genômica de *shotgun* seqüenciada totalizava 46462 clones com insertos variando entre 400 e 5200 bp (Figura 28A), os quais estavam acondicionados em 484 placas de 96 poços, armazenadas a -80°C. Os critérios que definimos para a seleção dos clones a partir desta biblioteca permitiram a identificação de 4421 clones representativos de 3084 CDSs. Como pode ser observado na Figura 28A, a distribuição dos tamanhos dos insertos é distinta dos tamanhos dos genes anotados. Consideramos um clone como representativo, quando este continha a maior cobertura de um gene de

interesse, mesmo que no inserto existissem outras seqüências. Este critério impossibilitou a identificação de clones representativos para 1229 CDS. A Figura 28B mostra a distribuição do número de genes e clones com seus tamanhos em pb. Clones representativos foram selecionados de acordo com a sobreposição do gene no inserto. A Figura 29 apresenta um esquema dos quatro tipos de clones encontrados: (A) – a CDS se apresenta completamente representada no inserto e está flanqueada por regiões intergênicas ou mesmo fragmentos de genes vizinhos; (B) - o inserto se caracteriza como sendo um fragmento da CDS e, portanto, não há regiões que representam outros genes ou mesmo regiões intergênicas; (C) - parte da CDS está sobreposta à extremidade esquerda do inserto; (D) - parte da CDS está sobreposta à extremidade direita. Nos dois últimos casos, a região do inserto não sobreposta à CDS pode também compreender regiões intergênicas ou outras seqüências codificantes. O total de CDS classificadas em um destes tipos está detalhado na Tabela 16, que amostra também o tamanho médio de sobreposição e não sobreposição gene-inserto. Os clones que consideramos mais adequados correspondem ao tipo B, por não apresentarem regiões não relacionadas a CDS representada. Além disso, é importante salientar que do total de pares de bases não sobrepostas nos outros três tipos de clones selecionados corresponde predominantemente a regiões intergênicas, com tamanho médio de 200 bp. O tipo de clone que representa uma dada CDS é considerado para efeito da análise dos dados.

Após a seleção *in silico*, os 4421 clones foram rearranjados em 47 placas de 96 poços, seguindo-se a amplificação dos insertos por PCR, utilizando-se um par de oligonucleotídeos universais. Verificamos que cerca de aproximadamente 5% dos clones selecionados não eram viáveis e 10% apresentaram inserto com tamanho diferente do predito (dados não mostrados), sugerindo prováveis erros de endereçamento dos clones. Além disso, outros 15% deram resultado negativo na amplificação. Por estes motivos, foi necessário o re-seqüenciamento dos clones viáveis. Assim, 2653 clones rearranjados em 27 placas de 96 poços foram re-seqüenciados em uma das suas extremidades. As seqüências foram comparadas com o genoma de XAC e os clones tiveram sua identificação

corrigida quando necessário. Infelizmente, ao final de todas as análises aproximadamente 40% dos clones inicialmente selecionados não puderam ser aproveitados. Embora o número de clones tenha sido severamente reduzido, passamos a dispor de um conjunto de clones bastante confiável quanto à sua identidade. Para evitar erro no reordenamento dos produtos de PCR previamente obtidos, fizemos novas reações de amplificação, sendo que apenas 1% dos clones deu resultado negativo de amplificação, resultando ao final na obtenção de 2639 produtos de PCR.

A este conjunto de 2639 produtos de PCR, somamos 121 oriundos de amplificações de seqüências derivadas de DNA genômico com oligonucleotídeos específicos, por representarem genes de interesse especial. Assim, 2760 produtos de PCR⁶ foram então purificados, rearranjados em 8 placas de 384 poços e utilizados na confecção dos microarranjos como descrito no item material e métodos.

⁶ A lista completa dos genes (*amplicons*) fixados no XACarray se apresenta incluso a esta tese como material ANEXO 6.

Figura 27: Fluxograma de etapas executadas para construção dos microarranjos de XAC (XACarray).

Figura 28: Distribuição do número de clones ou CDS em função de seus respectivos tamanhos em pb. (A) – Distribuição dos 46462 clones de *shotgun* e das 4313 CDS anotadas no genoma de XAC. Note que os genes (CDS) que apresentavam entre 0 e 200 bp de comprimento e uma pequena fração dos genes até 400 bp não foram encontrados na biblioteca de *shotgun* segundo os critérios de seleção descritos no texto. **(B)** – Distribuição dos 4421 clones selecionados e das 3084 CDS representadas por estes clones.

Figura 29: Esquema dos tipos de clones selecionados a partir da biblioteca genômica de XAC. A região em branco representa a sequência do vetor de clonagem pUC19. A região delimitada por linhas verticais pontilhadas delimita o inserto, sendo o cinza a região de não sobreposição à CDS de interesse, e preto a correspondente sobreposição. Regiões serrilhadas representam a ruptura da sequência dos genes de interesse. Setas representam os oligonucleotídeos utilizados para a reação de PCR.

Tabela 16: Análise da seleção dos clones agrupados de acordo com os modelos de sobreposições da Figura 29.

Tipo de clone	Total de clones selecionados	Tamanho médio de sobreposição gene-inserto	Tamanho médio de não sobreposição gene-inserto
A	2055	635	835
B	463	1121	0
C	827	822	418
D	1124	726	481
Total	4421	830.5 ^a	578 ^b

^{a,b} Antes de comparar o tamanho da região não sobreposta com a região sobreposta (CDS de interesse), devemos desprezar uma média de 200 pb dos valores não sobrepostos dos tipos C e D e 400 pb do tipo A, representando o tamanho médio das regiões intergênicas. Quanto maior a relação sobreposição - não sobreposição, melhor é o clone.

4.3.1. Validação da qualidade do XACarray

Para uma validação inicial da qualidade do **XACarray** realizamos a hibridação com DNA genômico de XAC. O DNA de XAC foi fragmentado por sonicação, marcado com nucleotídeos fluorescentes, e em seguida utilizado na hibridação dos microarranjos, como descrito no item material e métodos. A Figura 30 mostra a imagem, após varredura no *scanner* em um dos lados da lâmina do XACarray, hibridado com DNA genômico de XAC. A análise das imagens revelou que obtivemos sinal detectável acima do ruído (*background*) para 86% das CDS representadas no XACarray. Vale ressaltar que o XACarray contém 3072 elementos em duplicata dos quais 2760 representam CDS de XAC e o restante corresponde aos controles descritos no material e métodos.

A Tabela 17 resume as características dos 2670 *amplicons* depositados no XACarray, quanto ao tipo de sobreposição no inserto do clone original e tamanho médio dos *amplicons*. Para apenas 292 CDS temos a situação considerada ideal, ou seja, o amplicon representa apenas uma CDS e não inclui outras regiões genômicas. A Tabela 17 mostra ainda que 63% (1675/2639) CDS representadas no XACarray tem função predita com base na anotação do genoma completo. Analisamos também a distribuição das CDS representadas no XACarray de acordo com sua categorização funcional. A Tabela 18 mostra que 100% das categorias funcionais de anotação identificadas no genoma de XAC estão representadas no XACarray.

Figura 30: Resultado da hibridação de amostras de DNA de XAC marcadas com os fluoróforos Cy3 e Cy5. O subarray 7 foi destacado da lâmina com intuito de detalhar a qualidade de hibridação. A imagem deste subarray foi decomposta nas imagens monocromáticas resultantes da varredura a laser em 650nm (Cy5) e 550nm (Cy3).

Tabela 17: Análise das CDS depositadas no XACarray

Tipo de Sobreposição ^a	Análise Física			Função predita ^b		Sinal na lâmina ^c	
	Número de Clones	Tamanho médio do inserto (bp)	Porcentagem do gene em sobreposição ao inserto (%)	Número de genes com função putativa - I	Número de genes sem função putativa - II	I	II
A	1184	1487	56,8	631	553	574	485
B	292	1121	100	235	57	215	52
C	490	1211	69,9	343	147	323	123
D	673	1209	70,5	466	207	436	184
Total	2639	1324	82,1	1675	964	1521	844

^a A classificação de sobreposição gene-inserto segue descrita na Tabela 16.

^b Função predita com base na anotação do genoma de XAC (DA SILVA *et al.*, 2002)

^c Sinal nos microarranjos após hibridação com DNA genômico. Note que mais de 86% das CDS estão representadas no XACarray (2365/2760 deram sinal de hibridação acima do *background*).

Tabela 18: Categorias funcionais de anotação representadas no XACarray

a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
IA.1	44	27	61	IC.7	18	10	56	IID.6	4	4	100	IIIB.3	14	9	64	VA.7	175	104	59
IA.2	93	47	50	IC.8	8	7	88	IID.7	2	2	100	IIIB.4	41	25	61	VB	27	14	52
IA.3	13	11	85	ID.1	69	39	57	IID.8	7	2	29	IIIB.5	28	18	64	VC	77	41	53
IB.1	4	1	25	ID.2	164	88	50	IID.9	7	4	57	IIIB.6	11	5	45	VD	6	2	33
IB.2	2	1	50	ID.3	32	23	72	IID.10	15	11	73	IIIC.1	48	24	50	VIA	43	28	65
IB.3	3	2	67	ID.4	31	18	58	IID.11	13	8	62	IIIC.2	17	12	71	VIB	42	24	57
IB.4	2	2	100	IIA.1	18	8	44	IID.12	15	9	60	IIIC.3	92	65	71	VIC	112	65	58
IB.5	0	0	0	IIA.2	50	31	62	IID.13	0	0	0	IIID.1	20	12	60	VIIA	10	3	30
IB.6	4	3	75	IIA.3	11	8	73	IID.14	8	5	63	IIID.2	25	18	72	VII B	27	13	48
IB.7	3	1	33	IIA.4	22	15	68	IID.15	3	2	67	IIID.3	12	6	50	VII C	126	75	60
IB.8	3	2	67	IIA.5	10	6	60	IID.16	3	0	0	IVA.1	51	26	51	VII D	20	16	80
IB.9	9	7	78	IIB.1	22	12	55	IID.17	20	9	45	IVA.2	43	27	63	VII E	15	13	87
IB.10	27	15	56	IIB.2	11	6	55	IIF	50	31	62	IV.B	31	17	55	VII F	14	11	79
IB.11	6	3	50	IIB.3	6	2	33	II F	3	1	33	IV.C	54	35	65	VII G	24	16	67
IB.12	9	6	67	IIB.4	12	9	75	IIA.1	35	19	54	IV.D	25	19	76	VII H	70	47	67
IC.1	30	21	70	IIC	3	1	33	IIIA.2	13	7	54	VA.1	23	15	65	VIII.A	958	512	52
IC.2	10	8	80	IID.1	13	6	46	IIIA.3	24	17	71	VA.2	12	6	50	VIII.B	372	180	49
IC.3	116	66	58	IID.2	8	2	25	IIIA.4	47	30	64	VA.3	21	7	33	VIII.C	328	168	50
IC.4	14	9	64	IID.3	2	2	100	IIIA.5	16	10	63	VA.4	34	22	65	IX	107	63	59
IC.5	6	4	67	IID.4	11	7	64	IIIB.1	61	1	2	VA.5	2	2	100	Total	4501	2516	56
IC.6	7	4	57	IID.5	3	3	100	IIIB.2	54	32	59	VA.6	15	7	47				

Os dados a respeito de categorias funcionais foram extraídos diretamente do banco de dados de XAC (<http://genoma4.iq.usp.br>). a, indica a categoria funcional de anotação; b, número de genes por categoria; c, número de genes fixados na lâmina por categoria e; d, percentual de representatividade de cada categoria no **XACarray**.

4.4. Análise do perfil de expressão de um conjunto seletivo de CDS em XAC cultivada em diferentes concentrações de glicose

Paralelamente à construção do XACarray descrito anteriormente, optamos por comparar os níveis de expressão de um conjunto de 10 CDS (Figura 31) por RT-qPCR em XAC cultivada em meio XDM₂ contendo 10 mM, 50 mM e 250 mM de glicose. Tais genes foram selecionados após análise da literatura pertinente para identificarmos alguns genes reconhecidamente modulados por variações na concentração de açúcar do meio. Também selecionamos outros genes que tínhamos interesse em avaliar os níveis de expressão nestas condições. Foram realizadas reações de RT-qPCR em triplicata a partir de RNA total preparado de XAC crescida por 22 horas em XDM₂ contendo 10 mM, 50 mM e 250 mM de glicose. Os resultados obtidos estão mostrados na Figura 31. A seguir discutimos o resultado obtido para cada CDS.

ID ^a	Gene	Produto
XAC0756	kdpA	Potassium-transporting ATPase, A chain
XAC1878	rpfC	RpfC protein
XAC1879	rpfF	RpfF protein
XAC1880	rpfB	RpfB protein
XAC2576	gumK	GumK protein
XAC2583	gumD	GumD protein
XAC2704	nuoA	NADH-ubiquinone oxidoreductase, NQO7 subunit
XAC2975	ptsK	HPr kinase/phosphatase
XAC3345	pykA	Pyruvate kinase type II

Figura 31: Análise dos níveis de expressão por RT-qPCR de alguns genes de XAC selecionados. Genes que apresentam níveis de expressão na escala 1 (linha pontilhada) da ordenada, compreendem genes que não sofreram modificações no processo transcrricional quando variada a condição fisiológica de crescimento, como é o caso do gene XAC2704 (*nuoA*) previamente escolhido como um putativo gene *housekeeping* de XAC, e agora validado como tendo esta característica em diferentes concentrações de glicose. ^a ID – identificação do gene de acordo com da Silva e colaboradores (DA SILVA *et al.*, 2002).

***nuoA* (XAC2704)** - O gene *nuoA* codifica para uma *NADH-ubiquinone oxidoreductase*, *NQO7 subunit* foi selecionado como um provável gene *housekeeping*, e, como esperadonão apresentou alterações em seu nível de expressão, nas condições testadas.

***pykA* (XAC3345)** - Este gene codifica a enzima *pyruvate kinase type II* responsável pelo último passo da via glicolítica e apresentou níveis de expressão proporcionais à concentração de açúcar do meio. Isto provavelmente reflete um maior deslocamento de glicose internalizada para a via glicolítica o que coincide com uma maior taxa de crescimento de XAC em 250 mM de glicose (dados não mostrados).

***kdpA* (XAC0756)** - O aumento da expressão do gene *kdpA* parece refletir a real condição fisiológica em que a célula se encontra, ou seja, o aumento da concentração de glicose no meio parece ter gerado um desbalanço no controle osmótico, ocasionando um aumento na expressão deste gene como discutido anteriormente (seção 4.2). *kdpA* é um transportador de potássio, com atividade aumentada em condições de estresse osmótico (LAIMINS *et al.*, 1978, TREUNER-LANGE *et al.*, 1997). Este potássio transportado com maior eficiência acaba sendo internalizado e acumulado no interior da célula de modo a manter um equilíbrio osmótico decorrente das concentrações de solutos no ambiente (glicose). Como apresentado na seção 4.2, o mutante deste gene apresentou virulência reduzida.

***rpfC* (XAC1878)** – Este gene codifica uma proteína transmembranar envolvida com internalização e cascata de sinalização intracelular mediada por DSF em eventos de *quorum sensing*. *RpfC* não apresentou variação do seu nível de expressão em nenhuma condição analisada, o que poderia justificar sua função no perfil metabólico do sistema de genes *rpf*. Por codificar uma proteína de membrana responsável por captar o estímulo externo e transferir uma

mensagem via cascata de sinalização intracelular, *rpfC* parece não sofrer regulação pela variação na concentração de açúcar.

***rpfB* (XAC1880) e *rpfF* (XAC1879)** codificam respectivamente para uma acil-coA ligase e uma proteína cuja função não foi ainda totalmente determinada, mas também regulada em decorrência da concentração de DSF do meio. Ambas apresentaram alteração nas suas expressões diferencial nas variações de concentração de açúcar testadas, porém não parecem ser co-reguladas. Isso sugere que em XAC não existe um *operon* clássico (Figura 32A) entre os genes *rpfB* e *rpfF* como descrito previamente para XCC (SLATER *et al.*, 2000, VOJNOV *et al.*, 2001), e que possivelmente existem promotores independentes. Na busca de promotores, identificamos, na região intergênica entre *rpfB* e *rpfF*, uma provável seqüência formadora de alça de término de transcrição, como mostrado na Figura 32, o que poderia justificar esta variação nos níveis de transcritos dos dois genes.

***gumK* (XAC2576) e *gumD* (XAC2583)** – estes genes codificam para glicosiltransferases que atuam em passos diferentes da biossíntese de goma xanthana. Observamos que o nível de seus transcritos apresentam variação, aumentando em ambas concentrações de glicose, 50 e 250 mM em relação a 10 mM quando analisados por RT-qPCR. Resultados prévios descritos na literatura apontam para um provável promotor inativo a montante do gene *gumK*, restando um único promotor a montante do gene *gumB*, capaz de transcrever todo um *operon* composto por 12 genes (KATZEN *et al.*, 1996). Assim sendo, esperávamos obter resultados semelhantes para os dois genes nas diferentes concentrações de glicose. A discrepância de nossos resultados e o que seria esperado pode ter relação com a região promotora destes genes, que seriam independentes no caso de XAC.

Figura 32: Análise da região que codifica para os genes *rpf* nos genomas de XCC e XAC. O esquema em (A) mostra a organização dos genes *rpf* nos genoms de XAC e XCC. As setas finas apontam a orientação de transcrição. A provável alça de terminação de transcrição está indicada entre os genes *rpfF* e *rpfB*. (B), análise da região onde foi encontrada a seqüência do provável terminador de transcrição, detalhado em (C).

5. CONSIDERAÇÕES FINAIS

Desde a publicação do primeiro artigo descrevendo a sequência completa e a anotação do genoma da bactéria *Xylella fastidiosa*, há pouco mais de seis anos (SIMPSON *et al.*, 2000), um crescente número de genomas completa ou parcialmente seqüenciados de outras bactérias fitopatogênicas vem sendo reportados (SETUBAL *et al.*, 2005) Tabela 2. A análise comparativa *in silico* destes genomas tem se revelado como uma abordagem extremamente útil na identificação de genes comuns ou exclusivos que potencialmente constituem um repertório de candidatos envolvidos na interação planta-patógeno e/ou em mecanismos comuns ou específicos de patogenicidade (BHATTACHARYYA *et al.*, 2002a, BHATTACHARYYA *et al.*, 2002b, VAN SLUYS *et al.*, 2002, VAN SLUYS *et al.*, 2003, VORHOLTER *et al.*, 2003, MOREIRA *et al.*, 2004, WOOD *et al.*, 2004, MOREIRA *et al.*, 2005, THIEME *et al.*, 2005).

A comparação do genoma de *Xanthomonas axonopodis* pv *citri* com os genomas de três outros fitopatógenos (XCC, XF-CVC e XF-PD) que descrevemos neste trabalho revelou que o conjunto de genes destas bactérias reflete diretamente a adaptação a seus respectivos ambientes de vida bem como aos seus mecanismos de patogenicidade. Por exemplo, os agrupamentos gênicos relacionados à síntese de goma (genes *rpf*) e à síntese de xanthomonadinas (genes *pig*), embora presentes em todas as quatro bactérias comparadas, apresentam variações na sua organização e composição, o que parece ter associação direta com os mecanismos de patogenicidade de cada uma delas. Nós sugerimos que a ruptura do agrupamento gênico entre os genes *rpfF* e *rpfB* poderia estar propiciando síntese reduzida de DSF nas cepas de *Xylella*, em comparação a *Xanthomonas* (MOREIRA *et al.*, 2004, MOREIRA *et al.*, 2005). De fato, foi demonstrado que XF-CVC produz quatorze vezes menos DSF que XAC (SCARPARI *et al.*, 2003). Assim, é bastante provável que a deficiência na síntese de DSF afete proporcionalmente a indução do fenômeno de *quorum sensing* e a retro-inibição de outros genes *rpf* (*rpfCHG*) que participam da regulação da síntese de goma, interferindo na formação do biofilme e na colonização eficiente de

plantas de *Citrus* spp. Além disso, diferentemente do que se observa em *Xanthomonas* (VOJNOV *et al.*, 2001), recentemente demonstramos que em *Xylella* a indução de genes *rpf* não é modulada pela variação na concentração de açúcar (glicose) do meio (PASHALIDIS *et al.*, 2005), em concordância com o fato desta bactéria habitar o xilema, um ambiente que provavelmente carece desta fonte de carbono.

Em XF-CVC, o agrupamento gênico relacionado à síntese de xanthomonadina sofreu modificações recentes em sua estrutura, o que é evidenciado pela presença de pseudogenes em sua composição. Especulamos que este complexo gênico tende a desaparecer ou inativar-se completamente em *Xylella*, dado que as bactérias deste gênero não apresentam a fase epífita de sobrevivência como observado em *Xanthomonas*, onde pigmentos como xanthomonadina teriam um papel fotoprotetor importante (POPLAWSKY *et al.*, 2000, GOEL *et al.*, 2001). Ao lado das variações na composição e organização de genes descritos acima, observamos que, comparativamente a *Xanthomonas*, as bactérias do gênero *Xylella* possuem menor número de genes envolvidos com degradação de parede celular vegetal e completa ausência dos clássicos sistemas de fosfotransferência, envolvidos no transporte de açúcares (sistema PTS). Esta maquinaria deficitária na captação de fontes de carbono é agravada pelo metabolismo energético incompleto verificado em *Xylella fastidiosa* (MOREIRA *et al.*, 2004), fatos que certamente tem relação com o crescimento fastidioso exibido por esta espécie bacteriana.

Dois outros importantes exemplos de diferenças na composição gênica entre bactérias do gênero *Xanthomonas* e *Xylella* são a presença de uma segunda cópia do SS-II, nomeada como *xcs*, e a cópia completa do SS-III em *Xanthomonas*. Em contraposição aos sistemas citados nos parágrafos anteriores, estes sistemas estão completamente ausentes no genoma de *Xylella*. Estas diferenças podem refletir possíveis eventos de perda ao longo da evolução, no caso das cepas de *Xylella*, ou ganho, no caso das cepas de *Xanthomonas*, após a divergência das espécies. Entretanto, no caso do SS-II, a hipótese de perda para o *operon xcs* em *Xylella* é reforçada por dados filogenéticos (Figura 8).

A análise comparativa que realizamos revelou que os genomas das duas cepas de *Xanthomonas* (XAC e XCC) apresentam expansões, que possivelmente estão relacionadas com adaptação destes organismos aos seus respectivos nichos. Por exemplo, a família de genes *tsr* (genes que codificam receptores de sinais quimiotáticos) apresenta enorme expansão no genoma de XAC e XCC e provavelmente ocorreu antes da separação das espécies (MORALES *et al.*, 2004, MOREIRA *et al.*, 2004). Este mesmo fenômeno já havia sido descrito para genes *tsr* de *Vibrio cholerae* (HEIDELBERG *et al.*, 2000), os quais seriam de extrema importância para a sobrevivência deste organismo no interior do hospedeiro humano.

Por outro lado, alguns sistemas estão mais representados ou completos no genoma das cepas de *Xylella* em relação a *Xanthomonas*. Este é o caso dos genes envolvidos com regulação, síntese e secreção de toxinas do tipo colicina V, que em *Xanthomonas* se apresentam ausentes. A completa perda dos genes *cvaA* e *cvaB* em XCC, representa, sob o ponto de vista evolutivo, um estado derivativo, já que XAC representa um estado intermediário de depleção por apresentar o gene *cvaB* na sua composição, sendo as cepas de *Xylella* classificadas como ancestrais em relação a atividade e funcionalidade destes genes. As toxinas similares a colicina-V produzidas pela *Xylella* possivelmente lhe conferem vantagens para sobreviver em ambientes ricos em microorganismos endofíticos, como a planta ou aparelho bucal do inseto (ARAUJO *et al.*, 2002, LACAVA *et al.*, 2004, PASHALIDIS *et al.*, 2005).

Outro aspecto que nossa análise comparativa detalhou foi a presença de duas cópias estruturais das subunidades protéicas que codificam para *pili* do tipo IV nas cepas de *Xylella* relativamente a *Xanthomonas*. Estas estruturas são responsáveis pela motilidade de *Xylella*, como recentemente demonstrado (MENG *et al.*, 2005). Por outro lado, vale ressaltar que *Xanthomonas* tem flagelo, ausente em *Xylella*, que acaba tendo uma função tão fundamental quanto o *pili* em *Xylella*, já que necessita desta estrutura para dar início ao processo infeccioso e de propagação no interior do vegetal.

Do ponto de vista da estrutura genômica, todos os quatro organismos apresentam importantes rearranjos genômicos. No caso das duas espécies de *Xanthomonas* foram descritos três rearranjos específicos, quando comparadas entre si. Um destes rearranjos compreende uma inversão localizada no término de replicação destes organismos; e os outros dois compreendem uma translocação recíproca com inversão das regiões supostamente posicionadas nas proximidades de duas e dez horas, tomando o cromossomo no formato circular (DA SILVA *et al.*, 2002). Nos genomas das duas cepas de *Xylella*, em contrapartida, foi observado um conjunto maior de rearranjos quando comparadas entre si, incluindo uma translocação do término de replicação (VAN SLUYS *et al.*, 2003). Muito provavelmente o aspecto mais importante destes rearranjos genômicos em *Xylella*, tem relação com a presença de proteínas de fago, mais especificamente integrases, localizadas justamente nas bordas destes rearranjos ou mesmo flanqueando putativas regiões de inserção e deleção gênica (PInDels), que identificamos neste trabalho. Em contraposição a estes dados, nenhuma relação pode ser observada entre a presença e o sítio de inserção das integrases de fago com rearranjos genômicos nas cepas de *Xanthomonas*. Acompanhando as integrases de fago observamos a presença predominante de outros genes que codificam proteínas de fago e, principalmente, genes hipotéticos nos genomas dos quatro fitopatógenos, o que é mais evidente nas cepas de *Xylella*. Muitas destas regiões, compostas basicamente por estes genes, formam regiões únicas e exclusivas nos genomas destes organismos, algumas das quais com extrema variação no conteúdo de CG e de códons, reforçando ainda mais a suposição de aquisição lateral destas regiões.

A análise detalhada do genoma de XAC possibilitou ainda a identificação de uma complexa rede de genes envolvidos com osmoproteção e osmorregulação, as quais não haviam sido previamente descritas em bactérias deste gênero. Verificamos também que o gene *kdpA* (gene que codifica para um transportador de potássio) tem sua expressão aumentada em resposta ao aumento na concentração de glicose, o que pode refletir a ativação de uma resposta osmoprotetora. Estas observações embasam a realização de experimentos para

caracterizar bioquimicamente esta rede e determinar a sua importância na interação da bactéria com seu hospedeiro. Entre estes experimentos está uma análise completa do fenótipo exibido por uma linhagem mutante de XAC que possui o gene *kdpA* interrompido por um transposon, a qual apresentou uma aparente redução na virulência em citros.

Outro aspecto que identificamos no genoma de XAC foi a existência de vias relacionadas a degradação de compostos aromáticos, que da mesma forma que as vias de osmorregulação, são descritas pela primeira vez em bactérias do gênero *Xanthomonas*. XAC apresenta pelo menos cinco vias completas envolvidas com degradação destes compostos, enquanto que passo que bactérias do gênero *Xylella* apresentam apenas três. A descoberta da via do β -cetoadipato íntegra em XAC e num único agrupamento gênico é uma forte evidência de sua funcionalidade. Especulamos que XAC teria a capacidade de burlar os mecanismos de defesa da planta, mediante resposta sistêmica após infecção, destoxificando os compostos aromáticos por ela produzidos, utilizando-os num segundo momento como uma fonte alternativa de carbono, já que a via leva a síntese de compostos intermediários do ciclo do ácido tricarboxílico. O mutante de XAC que tem uma inserção por transposon no gene *pobA* (p-hidroxibenzoato hidroxilase) apresentou virulência reduzida na infecção em citros, o que reforça nossa especulação

Finalmente, neste trabalho descrevemos o planejamento e construção de microarranjos de DNA representando 2760 CDS codificadas no genoma de XAC, classificadas em todas as categorias e subcategorias funcionais de anotação. Híbridações deste microarranjo com DNA genômico de XAC mostraram que 89% de todos os *amplicons* fixados na lâmina (2365/2639) apresentaram sinal de hibridação superior à intensidade de fluorescência do ruído (*background*). Ainda que experimentos de hibridação com RNA não tenham sido efetivamente realizados, estamos confiantes que estes microarranjos de DNA constituem-se uma nova ferramenta para estudos de genômica comparativa e expressão gênica de XAC.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ADACHI, J. & HASEGAWA, M. (1996). **MOLPHY version 2.3: Programs for Molecular Phylogenetics Based on Maximum Likelihood**. Edited by J. Adachi. Tokyo: The Institute of Statistical Mathematics.
- AKMAN, L., YAMASHITA, A., WATANABE, H., OSHIMA, K., SHIBA, T., HATTORI, M. & AKSOY, S. (2002). **Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia***. *Nat Genet* 32, 402-7.
- ALEGRIA, M. C., DOCENA, C., KHATER, L., RAMOS, C. H., DA SILVA, A. C. & FARAH, C. S. (2004). **New protein-protein interactions identified for the regulatory and structural components and substrates of the type III Secretion system of the phytopathogen *Xanthomonas axonopodis* Pathovar *citri***. *J Bacteriol* 186, 6186-97.
- ALFANO, J. R. & COLLMER, A. (2004). **Type III secretion system effector proteins: double agents in bacterial disease and plant defense**. *Annu Rev Phytopathol* 42, 385-414.
- ALM, R. A. & MATTICK, J. S. (1997). **Genes involved in the biogenesis and function of type-4 fimbriae in *Pseudomonas aeruginosa***. *Gene* 192, 89-98.
- ALTENDORF, K., GASSEL, M., PUPPE, W., MOLLENKAMP, T., ZEECK, A., BODDIEN, C., FENDLER, K., BAMBERG, E. & DROSE, S. (1998). **Structure and function of the Kdp-ATPase of *Escherichia coli***. *Acta Physiol Scand Suppl* 643, 137-46.
- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. (1990). **Basic local alignment search tool**. *J Mol Biol* 215, 403-10.
- APWEILER, R., ATTWOOD, T. K., BAIROCH, A., BATEMAN, A., BIRNEY, E., BISWAS, M., BUCHER, P., CERUTTI, L., CORPET, F., CRONING, M. D., DURBIN, R., FALQUET, L., FLEISCHMANN, W., GOUZY, J., HERMJAKOB, H., HULO, N., JONASSEN, I., KAHN, D., KANAPIN, A., KARAVIDOPOULOU, Y., LOPEZ, R., MARX, B., MULDER, N. J., OINN, T. M., PAGNI, M., SERVANT, F., SIGRIST, C. J. & ZDOBNOV, E. M. (2000). **InterPro--an integrated documentation resource for protein families, domains and functional sites**. *Bioinformatics* 16, 1145-50.
- APWEILER, R., BAIROCH, A., WU, C. H., BARKER, W. C., BOECKMANN, B., FERRO, S., GASTEIGER, E., HUANG, H., LOPEZ, R., MAGRANE, M., MARTIN, M. J., NATALE, D. A., O'DONOVAN, C., REDASCHI, N. & YEH, L. S. (2004). **UniProt: the Universal Protein knowledgebase**. *Nucleic Acids Res* 32, D115-9.
- ARAUJO, W. L., MARCON, J., MACCHERONI, W., JR., VAN ELSAS, J. D., VAN VUURDE, J. W. & AZEVEDO, J. L. (2002). **Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants**. *Appl Environ Microbiol* 68, 4906-14.

- ASTUA-MONGE, G., FREITAS-ASTUA, J., BACOCINA, G., RONCOLETTA, J., CARVALHO, S. A. & MACHADO, M. A. (2005). **Expression profiling of virulence and pathogenicity genes of *Xanthomonas axonopodis* pv. *citri***. *J Bacteriol* 187, 1201-5.
- AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STRUHL, K. (1990). **Current protocols in Molecular Biology**, 3th edn. New York:: Greene Publishing Associates and Wiley-Interscience
- BAKKER, E. P., BOOTH, I. R., DINNBIER, U., EPSTEIN, W. & GAJEWSKA, A. (1987). **Evidence for multiple K⁺ export systems in *Escherichia coli***. *J Bacteriol* 169, 3743-9.
- BARBER, C. E., TANG, J. L., FENG, J. X., PAN, M. Q., WILSON, T. J., SLATER, H., DOW, J. M., WILLIAMS, P. & 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.
- BARBIERI, P., PALLADINO, L., DI GENNARO, P. & GALLI, E. (1993). **Alternative pathways for o-xylene or m-xylene and p-xylene degradation in a *Pseudomonas stutzeri* strain**. *Biodegradation* 4, 71-80.
- BELL, K. S., SEBAIHIA, M., PRITCHARD, L., HOLDEN, M. T., HYMAN, L. J., HOLEVA, M. C., THOMSON, N. R., BENTLEY, S. D., CHURCHER, L. J., MUNGALL, K., ATKIN, R., BASON, N., BROOKS, K., CHILLINGWORTH, T., CLARK, K., DOGGETT, J., FRASER, A., HANCE, Z., HAUSER, H., JAGELS, K., MOULE, S., NORBERTCZAK, H., ORMOND, D., PRICE, C., QUAIL, M. A., SANDERS, M., WALKER, D., WHITEHEAD, S., SALMOND, G. P., BIRCH, P. R., PARKHILL, J. & TOTH, I. K. (2004). **Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors**. *Proc Natl Acad Sci U S A* 101, 11105-10.
- BENNETT, P. M. (2004). **Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement**. *Methods Mol Biol* 266, 71-113.
- BHAT, T. K., SINGH, B. & SHARMA, O. P. (1998). **Microbial degradation of tannins--a current perspective**. *Biodegradation* 9, 343-57.
- BHATTACHARYYA, A., STILWAGEN, S., IVANOVA, N., D'SOUZA, M., BERNAL, A., LYKIDIS, A., KAPATRAL, V., ANDERSON, I., LARSEN, N., LOS, T., REZNIK, G., SELKOV, E., JR., WALUNAS, T. L., FEIL, H., FEIL, W. S., PURCELL, A., LASSEZ, J. L., HAWKINS, T. L., HASELKORN, R., OVERBEEK, R., PREDKI, P. F. & KYRPIDES, N. C. (2002a). **Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains**. *Proc Natl Acad Sci U S A* 99, 12403-8.
- BHATTACHARYYA, A., STILWAGEN, S., REZNIK, G., FEIL, H., FEIL, W. S., ANDERSON, I., BERNAL, A., D'SOUZA, M., IVANOVA, N., KAPATRAL, V., LARSEN, N., LOS, T., LYKIDIS, A., SELKOV, E., JR., WALUNAS, T. L., PURCELL, A., EDWARDS, R. A., HAWKINS, T., HASELKORN, R., OVERBEEK, R., KYRPIDES, N. C. & PREDKI, P. F. (2002b). **Draft sequencing and**

comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res* 12, 1556-63.

BITANCOURT, A. A. (1957). **O Cancro Cítrico.** . *Biológico* 23, 110-111.

BONAS, U., SCHULTE, R., FENSELAU, S., MINSAVAGE, G. V., STASKAWICZ, B. J. & STALL, R. E. (1991). **Isolation of a gene cluster from *Xanthomonas campestris* pv. vesicatoria that determines pathogenicity and the hypersensitive response on pepper and tomato.** *Mol. Plant Microbe Interact.* 4, 81–88.

BOSSEMEYER, D., BORCHARD, A., DOSCH, D. C., HELMER, G. C., EPSTEIN, W., BOOTH, I. R. & BAKKER, E. P. (1989). **K⁺-transport protein TrkA of *Escherichia coli* is a peripheral membrane protein that requires other trk gene products for attachment to the cytoplasmic membrane.** *J Biol Chem* 264, 16403-10.

BREESE, K. & FUCHS, G. (1998). **4-Hydroxybenzoyl-CoA reductase (dehydroxylating) from the denitrifying bacterium *Thauera aromatica*--prosthetic groups, electron donor, and genes of a member of the molybdenum-flavin-iron-sulfur proteins.** *Eur J Biochem* 251, 916-23.

BRUNINGS, A. M. & GABRIEL, D. W. (2003). ***Xanthomonas citri*: breaking the surface.** *Molecular Plant Pathology* 4, 141-157.

BUELL, C. R., JOARDAR, V., LINDEBERG, M., SELENGUT, J., PAULSEN, I. T., GWINN, M. L., DODSON, R. J., DEBOY, R. T., DURKIN, A. S., KOLONAY, J. F., MADUPU, R., DAUGHERTY, S., BRINKAC, L., BEANAN, M. J., HAFT, D. H., NELSON, W. C., DAVIDSEN, T., ZAFAR, N., ZHOU, L., LIU, J., YUAN, Q., KHOURI, H., FEDOROVA, N., TRAN, B., RUSSELL, D., BERRY, K., UTTERBACK, T., VAN AKEN, S. E., FELDBLYUM, T. V., D'ASCENZO, M., DENG, W. L., RAMOS, A. R., ALFANO, J. R., CARTINHO, S., CHATTERJEE, A. K., DELANEY, T. P., LAZAROWITZ, S. G., MARTIN, G. B., SCHNEIDER, D. J., TANG, X., BENDER, C. L., WHITE, O., FRASER, C. M. & COLLMER, A. (2003). **The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000.** *Proc Natl Acad Sci U S A* 100, 10181-6.

BUURMAN, E. T., KIM, K. T. & EPSTEIN, W. (1995). **Genetic evidence for two sequentially occupied K⁺ binding sites in the Kdp transport ATPase.** *J Biol Chem* 270, 6678-85.

CALAMITA, G. (2000). **The *Escherichia coli* aquaporin-Z water channel.** *Mol Microbiol* 37, 254-62.

CALAMITA, G., BISHAI, W. R., PRESTON, G. M., GUGGINO, W. B. & AGRE, P. (1995). **Molecular cloning and characterization of AqpZ, a water channel from *Escherichia coli*.** *J Biol Chem* 270, 29063-6.

CALAMITA, G., KEMPF, B., RUDD, K. E., BONHIVERS, M., KNEIP, S., BISHAI, W. R., BREMER, E. & AGRE, P. (1997). **The aquaporin-Z water channel gene of *Escherichia coli*: structure, organization and phylogeny.** *Biol Cell* 89, 321-9.

- CAPITANI, G., EIDAM, O., GLOCKSHUBER, R. & GRUTTER, M. G. (2006). **Structural and functional insights into the assembly of type 1 pili from *Escherichia coli*.** *Microbes Infect.*
- CARNEIRO-JUNIOR, J. B., SILVEIRA, S. F., SOUZA-FILHO, G. A., OLIVARES, F. L. & GIGLIOTI, E. A. (2004). **Policlonal antiserum specificity against *Leifsonia xyli* subsp. *xyli*.** *Fitopatol. bras.* 29.
- CASTRESANA, J. (2000). **Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.** *Mol Biol Evol* 17, 540-52.
- CHENNA, R., SUGAWARA, H., KOIKE, T., LOPEZ, R., GIBSON, T. J., HIGGINS, D. G. & THOMPSON, J. D. (2003). **Multiple sequence alignment with the Clustal series of programs.** *Nucleic Acids Res* 31, 3497-500.
- CIVEROLO, E. L. (1984). **Bacterial canker disease of citrus.** *J. Rio Grande Valley Hortic. Soc.* 37, 127-145.
- CONWAY, T. & SCHOOLNIK, G. K. (2003). **Microarray expression profiling: capturing a genome-wide portrait of the transcriptome.** *Mol Microbiol* 47, 879-89.
- CROSA, J. H., MEY, A. R. & PAYNE, S. M. (2004). **Iron Transport in Bacteria**, pp. 532. Washington, D.C: ASM Press.
- CSONKA, L. N. & EPSTEIN, W. (1996). **Osmoregulation.** In *Escherichia coli and Salmonella: Cellular and Molecular biology*, Second edn, pp. 1210-1223. Edited by F. C. Neidhart, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: ASM Press.
- DA SILVA, A. C., FERRO, J. A., REINACH, F. C., FARAH, C. S., FURLAN, L. R., QUAGGIO, R. B., MONTEIRO-VITORELLO, C. B., VAN SLUYS, M. A., ALMEIDA, N. F., ALVES, L. M., DO AMARAL, A. M., BERTOLINI, M. C., CAMARGO, L. E., CAMAROTTE, G., CANNAVAN, F., CARDOZO, J., CHAMBERGO, F., CIAPINA, L. P., CICARELLI, R. M., COUTINHO, L. L., CURSINO-SANTOS, J. R., EL-DORRY, H., FARIA, J. B., FERREIRA, A. J., FERREIRA, R. C., FERRO, M. I., FORMIGHIERI, E. F., FRANCO, M. C., GREGGIO, C. C., GRUBER, A., KATSUYAMA, A. M., KISHI, L. T., LEITE, R. P., LEMOS, E. G., LEMOS, M. V., LOCALI, E. C., MACHADO, M. A., MADEIRA, A. M., MARTINEZ-ROSSI, N. M., MARTINS, E. C., MEIDANIS, J., MENCK, C. F., MIYAKI, C. Y., MOON, D. H., MOREIRA, L. M., NOVO, M. T., OKURA, V. K., OLIVEIRA, M. C., OLIVEIRA, V. R., PEREIRA, H. A., ROSSI, A., SENA, J. A., SILVA, C., DE SOUZA, R. F., SPINOLA, L. A., TAKITA, M. A., TAMURA, R. E., TEIXEIRA, E. C., TEZZA, R. I., TRINDADE DOS SANTOS, M., TRUFFI, D., TSAI, S. M., WHITE, F. F., SETUBAL, J. C. & KITAJIMA, J. P. (2002). **Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities.** *Nature* 417, 459-63.
- DA SILVA, F. R., VETTORE, A. L., KEMPER, E. L., LEITE, A. & ARRUDA, P. (2001). **Fastidious gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity.** *FEMS Microbiol Lett* 203, 165-71.

- DAVIS, M. J., GILLASPIE, A. G., HARRIS, R. W. & LAWSON, R. H. (1980). **Ratoon stunting disease of sugarcane: isolation of the causal bacterium.** . *Science* 210, 1365-1367.
- 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. & 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.
- 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. & MACHADO, M. A. (2004). **Gene expression profile of the plant pathogen *Xylella fastidiosa* during biofilm formation in vitro.** *FEMS Microbiol Lett* 237, 341-53.
- DE SOUZA, A. A., TAKITA, M. A., PEREIRA, E. O., COLETTA-FILHO, H. D. & MACHADO, M. A. (2005). **Expression of Pathogenicity-Related genes of *Xylella fastidiosa* in vitro and in planta.** *Current Microbiology* 50, 1-6 In Press.
- DEGNAN, P. H., LAZARUS, A. B. & WERNEGREN, J. J. (2005). **Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects.** *Genome Res* 15, 1023-33.
- DENG, W., BURLAND, V., PLUNKETT, G., 3RD, BOUTIN, A., MAYHEW, G. F., LISS, P., PERNA, N. T., ROSE, D. J., MAU, B., ZHOU, S., SCHWARTZ, D. C., FETHERSTON, J. D., LINDLER, L. E., BRUBAKER, R. R., PLANO, G. V., STRALEY, S. C., MCDONOUGH, K. A., NILLES, M. L., MATSON, J. S., BLATTNER, F. R. & PERRY, R. D. (2002). **Genome sequence of *Yersinia pestis* KIM.** *J Bacteriol* 184, 4601-11.
- DIAZ, E., FERRANDEZ, A., PRIETO, M. A. & GARCIA, J. L. (2001). **Biodegradation of aromatic compounds by *Escherichia coli*.** *Microbiol Mol Biol Rev* 65, 523-69, table of contents.
- DIGIAMPIETRI, L. A., MEDEIROS, C. B. & SETUBAL, J. C. (2003). **A data model for comparative genomics.** *Rev. Tecnol. Informat.* 3, 35-40.
- DISPENSA, M., THOMAS, C. T., KIM, M. K., PERROTTA, J. A., GIBSON, J. & HARWOOD, C. S. (1992). **Anaerobic growth of *Rhodopseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators.** *J Bacteriol* 174, 5803-13.
- DOI, R. H. & KOSUGI, A. (2004). **Cellulosomes: plant-cell-wall-degrading enzyme complexes.** *Nat Rev Microbiol* 2, 541-51.
- DOMINGUE, G. J. (1991). **Bacterial osmoprotection.** *J Lab Clin Med* 118, 203-4.
- DOMINGUES, M. C. S., ONO, E. O. & RODRIGUES, J. D. (2001). **Indução do amadurecimento de frutos cítricos em pós-colheita com aplicação de ethephon.** *Revista Brasileira de Fruticultura* 23, 555-558.

- DOW, J. M., FENG, J. X., BARBER, C. E., TANG, J. L. & DANIELS, M. J. (2000). **Novel genes involved in the regulation of pathogenicity factor production within the rpf gene cluster of *Xanthomonas campestris***. *Microbiology* 146 (Pt 4), 885-91.
- EARHART, C. F. (1996). **Uptake and metabolism of iron and molybdenum**. . In ***Escherichia coli and Salmonella: Cellular and Molecular Biology***. . Edited by F. C. Neidhardt, R. I. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikopp, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: : ASM Press.
- EGLAND, P. G., PELLETIER, D. A., DISPENSA, M., GIBSON, J. & HARWOOD, C. S. (1997). **A cluster of bacterial genes for anaerobic benzene ring biodegradation**. *Proc Natl Acad Sci U S A* 94, 6484-9.
- EPSTEIN, W., BUURMAN, E., MCLAGGAN, D. & NAPRSTEK, J. (1993). **Multiple mechanisms, roles and controls of K⁺ transport in *Escherichia coli***. *Biochem Soc Trans* 21, 1006-10.
- FATH, M. J., MAHANTY, H. K. & KOLTER, R. (1989). **Characterization of a purF operon mutation which affects colicin V production**. *J Bacteriol* 171, 3158-61.
- FATH, M. J., ZHANG, L. H., RUSH, J. & KOLTER, R. (1994). **Purification and characterization of colicin V from *Escherichia coli* culture supernatants**. *Biochemistry* 33, 6911-7.
- FEIL, H., FEIL, W. S., CHAIN, P., LARIMER, F., DIBARTOLO, G., COPELAND, A., LYKIDIS, A., TRONG, S., NOLAN, M., GOLTSMAN, E., THIEL, J., Malfatti, S., LOPER, J. E., LAPIDUS, A., DETTER, J. C., LAND, M., RICHARDSON, P. M., KYRPIDES, N. C., IVANOVA, N. & LINDOW, S. E. (2005). **Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000**. *Proc Natl Acad Sci U S A* 102, 11064-9.
- FISCHBARG, J., KUANG, K. Y., VERA, J. C., ARANT, S., SILVERSTEIN, S. C., LOIKE, J. & ROSEN, O. M. (1990). **Glucose transporters serve as water channels**. *Proc Natl Acad Sci U S A* 87, 3244-7.
- FOUTS, D. E., BADEL, J. L., RAMOS, A. R., RAPP, R. A. & COLLMER, A. (2003). **A *pseudomonas syringae* pv. *tomato* DC3000 Hrp (Type III secretion) deletion mutant expressing the Hrp system of bean pathogen *P. syringae* pv. *syringae* 61 retains normal host specificity for tomato**. *Mol Plant Microbe Interact* 16, 43-52.
- FU, Q. S., LI, F. & CHEN, L. L. (2005). **Gene expression analysis of six GC-rich Gram-negative phytopathogens**. *Biochem Biophys Res Commun* 332, 380-7.
- FUNDECITRUS (2006). **Fundo de Defesa da Citricultura**. São Paulo.
- GABRIEL, D. W., ALLEN, C., SCHELL, M., DENNY, T. P., GREENBERG, J. T., DUAN, Y. P., FLORES-CRUZ, Z., HUANG, Q., CLIFFORD, J. M., PRESTING, G., GONZALEZ, E. T., REDDY, J., ELPHINSTONE, J., SWANSON, J., YAO, J., MULHOLLAND, V., LIU, L., FARMERIE, W., PATNAIKUNI, M., BALOGH, B.,

- NORMAN, D., ALVAREZ, A., CASTILLO, J. A., JONES, J., SADDLER, G., WALUNAS, T., ZHUKOV, A. & MIKHAILOVA, N. (2006). **Identification of open reading frames unique to a select agent: *Ralstonia solanacearum* race 3 biovar 2.** *Mol Plant Microbe Interact* 19, 69-79.
- GALAN, J. E. & COLLMER, A. (1999). **Type III secretion machines: bacterial devices for protein delivery into host cells.** *Science* 284, 1322-8.
- GARCIA-PEREZ, A. & BURG, M. B. (1991). **Role of organic osmolytes in adaptation of renal cells to high osmolality.** *J Membr Biol* 119, 1-13.
- GESCHER, J., EISENREICH, W., WORTH, J., BACHER, A. & FUCHS, G. (2005). **Aerobic benzoyl-CoA catabolic pathway in *Azoarcus evansii*: studies on the non-oxygenolytic ring cleavage enzyme.** *Mol Microbiol* 56, 1586-600.
- GIBSON, J., DISPENSA, M. & HARWOOD, C. S. (1997). **4-hydroxybenzoyl coenzyme A reductase (dehydroxylating) is required for anaerobic degradation of 4-hydroxybenzoate by *Rhodopseudomonas palustris* and shares features with molybdenum-containing hydroxylases.** *J Bacteriol* 179, 634-42.
- GILSON, L., MAHANTY, H. K. & KOLTER, R. (1987). **Four plasmid genes are required for colicin V synthesis, export, and immunity.** *J Bacteriol* 169, 2466-70.
- GILSON, L., MAHANTY, H. K. & KOLTER, R. (1990). **Genetic analysis of an MDR-like export system: the secretion of colicin V.** *Embo J* 9, 3875-84.
- GOEL, A. K., RAJAGOPAL, L., NAGESH, N. & SONTI, R. V. (2002). **Genetic locus encoding functions involved in biosynthesis and outer membrane localization of xanthomonadin in *Xanthomonas oryzae* pv. *oryzae*.** *J Bacteriol* 184, 3539-48.
- GOEL, A. K., RAJAGOPAL, L. & SONTI, R. V. (2001). **Pigment and virulence deficiencies associated with mutations in the *aroE* gene of *Xanthomonas oryzae* pv. *oryzae*.** *Appl Environ Microbiol* 67, 245-50.
- GONCALVES, E. R. & ROSATO, Y. B. (2002). **Phylogenetic analysis of *Xanthomonas* species based upon 16S-23S rDNA intergenic spacer sequences.** *Int J Syst Evol Microbiol* 52, 355-61.
- GOODNER, B., HINKLE, G., GATTUNG, S., MILLER, N., BLANCHARD, M., QUROLLO, B., GOLDMAN, B. S., CAO, Y., ASKENAZI, M., HALLING, C., MULLIN, L., HOUMIEL, K., GORDON, J., VAUDIN, M., IARTCHOUK, O., EPP, A., LIU, F., WOLLAM, C., ALLINGER, M., DOUGHTY, D., SCOTT, C., LAPPAS, C., MARKELZ, B., FLANAGAN, C., CROWELL, C., GURSON, J., LOMO, C., SEAR, C., STRUB, G., CIELO, C. & SLATER, S. (2001). **Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58.** *Science* 294, 2323-8.
- GOTTWALD, T. R., GRAHAM, J. H. & SCHUBERT, T. S. (2002). **Citrus Canker: The Pathogen and Its Impact.** *APS Net: Peer-Reviewed by Plant Health Progress*

- GREBE, T. W. & STOCK, J. (1998). **Bacterial chemotaxis: the five sensors of a bacterium.** *Curr Biol* 8, R154-7.
- HAMMOND-KOSACK, K. E. & JONES, J. D. (1996). **Resistance gene-dependent plant defense responses.** *Plant Cell* 8, 1773-91.
- HANCOCK, J. M. (1996). **Simple sequences in a "minimal" genome.** *Nat Genet* 14, 14-5.
- HANSEN-WESTER, I. & HENSEL, M. (2001). **Salmonella pathogenicity islands encoding type III secretion systems.** *Microbes Infect* 3, 549-59.
- HARWOOD, C. S. & PARALES, R. E. (1996). **The beta-ketoadipate pathway and the biology of self-identity.** *Annu Rev Microbiol* 50, 553-90.
- HEIDELBERG, J. F., EISEN, J. A., NELSON, W. C., CLAYTON, R. A., GWINN, M. L., DODSON, R. J., HAFT, D. H., HICKEY, E. K., PETERSON, J. D., UMayAM, L., GILL, S. R., NELSON, K. E., READ, T. D., TETTELIN, H., RICHARDSON, D., ERMOLAEVA, M. D., VAMATHEVAN, J., BASS, S., QIN, H., DRAGOI, I., SELLERS, P., MCDONALD, L., UTTERBACK, T., FLEISHMANN, R. D., NIERMAN, W. C., WHITE, O., SALZBERG, S. L., SMITH, H. O., COLWELL, R. R., MEKALANOS, J. J., VENTER, J. C. & FRASER, C. M. (2000). **DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.** *Nature* 406, 477-83.
- HENDRICKS, E. C., SZERLONG, H., HILL, T. & KUEMPEL, P. (2000). **Cell division, guillotining of dimer chromosomes and SOS induction in resolution mutants (dif, xerC and xerD) of *Escherichia coli*.** *Mol Microbiol* 36, 973-81.
- HIRANO, S. S. & UPPER, C. D. (1983). **Ecology and epidemiology of foliar bacterial plant pathogens.** *Annual Review of Phytopathology* 21, 243-269.
- HU, R. M., YANG, T. C., YANG, S. H. & TSENG, Y. H. (2005). **Deduction of upstream sequences of *Xanthomonas campestris* flagellar genes responding to transcription activation by FleQ.** *Biochem Biophys Res Commun* 335, 1035-43.
- INGLEDEW, W. J. & POOLE, R. K. (1984). **The respiratory chains of *Escherichia coli*.** *Microbiol Rev* 48, 222-71.
- JAYASINGHE, C. K., WIJAYARATNE, S. C. & FERNANDO, T. H. (2004). **Characterization of cell wall degrading enzymes of *Thanatephorus cucumeris*.** *Mycopathologia* 157, 73-9.
- JENKINS, C. L. & STARR, M. P. (1985). **Formation of halogenated aryl-polyene (xanthomonadin) pigments by the type and other yellow-pigmented strains of *Xanthomonas maltophilia*.** *Ann Inst Pasteur Microbiol* 136B, 257-64.
- JIANG, H. & HUTCHINSON, C. R. (2006). **Feedback regulation of doxorubicin biosynthesis in *Streptomyces peucetius*.** *Res Microbiol*.

- JIMENEZ, J. I., MINAMBRES, B., GARCIA, J. L. & DIAZ, E. (2002). **Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440.** *Environ Microbiol* 4, 824-41.
- JOARDAR, V., LINDEBERG, M., JACKSON, R. W., SELENGUT, J., DODSON, R., BRINKAC, L. M., DAUGHERTY, S. C., DEBOY, R., DURKIN, A. S., GIGLIO, M. G., MADUPU, R., NELSON, W. C., ROSOVITZ, M. J., SULLIVAN, S., CRABTREE, J., CREASY, T., DAVIDSEN, T., HAFT, D. H., ZAFAR, N., ZHOU, L., HALPIN, R., HOLLEY, T., KHOURI, H., FELDBLYUM, T., WHITE, O., FRASER, C. M., CHATTERJEE, A. K., CARTINHOOR, S., SCHNEIDER, D. J., MANSFIELD, J., COLLMER, A. & BUELL, C. R. (2005). **Whole-genome sequence analysis of *Pseudomonas syringae* pv. phaseolicola 1448A reveals divergence among pathovars in genes involved in virulence and transposition.** *J Bacteriol* 187, 6488-98.
- JOURNET, L., BOUVERET, E., RIGAL, A., LLOUBES, R., LAZDUNSKI, C. & BENEDETTI, H. (2001). **Import of colicins across the outer membrane of *Escherichia coli* involves multiple protein interactions in the periplasm.** *Mol Microbiol* 42, 331-44.
- KABISCH, U., LANDGRAF, A., KRAUSE, J., BONAS, U. & BOCH, J. (2005). **Type III secretion chaperones ShcS1 and ShcO1 from *Pseudomonas syringae* pv. tomato DC3000 bind more than one effector.** *Microbiology* 151, 269-80.
- KANEHISA, M. & GOTO, S. (2000). **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 28, 27-30.
- KARLIN, S. (1998). **Global dinucleotide signatures and analysis of genomic heterogeneity.** *Curr Opin Microbiol* 1, 598-610.
- KARLIN, S. (2001). **Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes.** *Trends Microbiol* 9, 335-43.
- KATZEN, F., BECKER, A., ZORREGUIETA, A., PUHLER, A. & IELPI, L. (1996). **Promoter analysis of the *Xanthomonas campestris* pv. *campestris* gum operon directing biosynthesis of the xanthan polysaccharide.** *J Bacteriol* 178, 4313-8.
- KAZEMI-POUR, N., CONDEMIN, G. & HUGOUVIEUX-COTTE-PATTAT, N. (2004). **The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*.** *Proteomics* 4, 3177-86.
- KEMPF, B. & BREMER, E. (1998). **Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments.** *Arch Microbiol* 170, 319-30.
- KIM, D., KIM, Y. S., KIM, S. K., KIM, S. W., ZYLSTRA, G. J., KIM, Y. M. & KIM, E. (2002). **Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17.** *Appl Environ Microbiol* 68, 3270-8.

- KOIDE, T., VENCIO, R. Z. & GOMES, S. L. (2006). **Global gene expression analysis of the heat shock response in the phytopathogen *Xylella fastidiosa*.** *J Bacteriol* 188, 5821-30.
- 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. & GOMES, S. L. (2004). **DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence.** *J Bacteriol* 186, 5442-9.
- KOIZUMI, M. (1985). **Citrus canker: the world situation** In *Citrus canker: An international perspective*, First edn, pp. 2-7. Edited by L. W. Timmer. Florida: Lake Alfred.
- LACAVA, P. T., ARAUJO, W. L., MARCON, J., MACCHERONI, W., JR. & AZEVEDO, J. L. (2004). **Interaction between endophytic bacteria from citrus plants and the phytopathogenic bacteria *Xylella fastidiosa*, causal agent of citrus-variegated chlorosis.** *Lett Appl Microbiol* 39, 55-9.
- LAIMINS, L. A., RHOADS, D. B., ALTENDORF, K. & EPSTEIN, W. (1978). **Identification of the structural proteins of an ATP-driven potassium transport system in *Escherichia coli*.** *Proc Natl Acad Sci U S A* 75, 3216-9.
- LEE, B. M., PARK, Y. J., PARK, D. S., KANG, H. W., KIM, J. G., SONG, E. S., PARK, I. C., YOON, U. H., HAHN, J. H., KOO, B. S., LEE, G. B., KIM, H., PARK, H. S., YOON, K. O., KIM, J. H., JUNG, C. H., KOH, N. H., SEO, J. S. & GO, S. J. (2005). **The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice.** *Nucleic Acids Res* 33, 577-86.
- LEITE, J., R.P. (2003). **Surviving with citrus canker in Brazil.** . In *Proceedings of the 9th Congress of the International Society of Citriculture*, pp. 890-896. Orlando, FL: International Society of Citriculture.
- LEITE JUNIOR, R. P. & MOHAN, S. K. (1990). **Integrated management of the citrus bacterial canker disease caused by *Xanthomonas campestris* pv. *citri* in the state of Paraná, Brazil.** *Crop Protection* 9, 3-7.
- LEMOS, E. G., ALVES, L. M. & CAMPANHARO, J. C. (2003). **Genomics-based design of defined growth media for the plant pathogen *Xylella fastidiosa*.** *FEMS Microbiol Lett* 219, 39-45.
- LEYNS, F., DECLEENE, M., SWINGS, J. G. & DELEY, J. (1984). **The host range of the genus *Xanthomonas*.** *Botanical Review* 50, 308-356.
- LIM, M. T. & KUNKEL, B. N. (2004). **The *Pseudomonas syringae* type III effector AvrRpt2 promotes virulence independently of RIN4, a predicted virulence target in *Arabidopsis thaliana*.** *Plant J* 40, 790-8.
- LIMA, W. C., VAN SLUYS, M. A. & MENCK, C. F. (2005). **Non-gamma-proteobacteria gene islands contribute to the *Xanthomonas* genome.** *Omics* 9, 160-72.

- LIN, R. J., CAPAGE, M. & HILL, C. W. (1984). **A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome.** *J Mol Biol* 177, 1-18.
- LINDEBERG, M., STAVRINIDES, J., CHANG, J. H., ALFANO, J. R., COLLMER, A., DANGL, J. L., GREENBERG, J. T., MANSFIELD, J. W. & GUTTMAN, D. S. (2005). **Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen Pseudomonas syringae.** *Mol Plant Microbe Interact* 18, 275-82.
- LIVAK, K. J. & 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.
- LOBRY, J. R. (1996). **Asymmetric substitution patterns in the two DNA strands of bacteria.** *Mol Biol Evol* 13, 660-5.
- LOBRY, J. R. & SUEOKA, N. (2002). **Asymmetric directional mutation pressures in bacteria.** *Genome Biol* 3, RESEARCH0058.
- LOCALI, E. C., FREITAS-ASTUA, J., DE SOUZA, A. A., TAKITA, M. A., ASTUA-MONGE, G., ANTONIOLI, R., KITAJIMA, E. W. & MACHADO, M. A. (2003). **Development of a molecular tool for the diagnosis of leprosis, a major threat to citrus production in the Americas.** *Plant Disease* 87, 1317-1321.
- LOH, J., GARCIA, M. & STACEY, G. (1997). **NodV and NodW, a second flavonoid recognition system regulating nod gene expression in Bradyrhizobium japonicum.** *J Bacteriol* 179, 3013-20.
- LOH, J., LOHAR, D. P., ANDERSEN, B. & STACEY, G. (2002). **A two-component regulator mediates population-density-dependent expression of the Bradyrhizobium japonicum nodulation genes.** *J Bacteriol* 184, 1759-66.
- LUCCHINI, S., THOMPSON, A. & HINTON, J. C. (2001). **Microarrays for microbiologists.** *Microbiology* 147, 1403-14.
- MACCHERONI, W., ALEGRIA, M. C., GREGGIO, C. C., PIAZZA, J. P., KAMLA, R. F., ZACHARIAS, P. R., BAR-JOSEPH, M., KITAJIMA, E. W., ASSUMPCAO, L. C., CAMAROTTE, G., CARDOZO, J., CASAGRANDE, E. C., FERRARI, F., FRANCO, S. F., GIACHETTO, P. F., GIRASOL, A., JORDAO, H., JR., SILVA, V. H., SOUZA, L. C., AGUILAR-VILDOSO, C. I., ZANCA, A. S., ARRUDA, P., KITAJIMA, J. P., REINACH, F. C., FERRO, J. A. & DA SILVA, A. C. (2005). **Identification and genomic characterization of a new virus (Tymoviridae family) associated with citrus sudden death disease.** *J Virol* 79, 3028-37.
- MACFARLANE, S. A. & MERRICK, M. (1985). **The nucleotide sequence of the nitrogen regulation gene ntrB and the glnA-ntrBC intergenic region of Klebsiella pneumoniae.** *Nucleic Acids Res* 13, 7591-606.

- MARQUES, M. V., DA SILVA, A. M. & GOMES, S. L. (2001). **Genetic organization of plasmid pXF51 from the plant pathogen *Xylella fastidiosa***. *Plasmid* 45, 184-99.
- MARTINEZ-ARGUDO, I., MARTIN-NIETO, J., SALINAS, P., MALDONADO, R., DRUMMOND, M. & CONTRERAS, A. (2001). **Two-hybrid analysis of domain interactions involving NtrB and NtrC two-component regulators**. *Mol Microbiol* 40, 169-78.
- MARTINEZ, A. T., SPERANZA, M., RUIZ-DUENAS, F. J., FERREIRA, P., CAMARERO, S., GUILLEN, F., MARTINEZ, M. J., GUTIERREZ, A. & DEL RIO, J. C. (2005). **Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin**. *Int Microbiol* 8, 195-204.
- MATTICK, J. S. (2002). **Type IV pili and twitching motility**. *Annu Rev Microbiol* 56, 289-314.
- MCCLELLAND, M., SANDERSON, K. E., SPIETH, J., CLIFTON, S. W., LATREILLE, P., COURTNEY, L., PORWOLLIK, S., ALI, J., DANTE, M., DU, F., HOU, S., LAYMAN, D., LEONARD, S., NGUYEN, C., SCOTT, K., HOLMES, A., GREWAL, N., MULVANEY, E., RYAN, E., SUN, H., FLOREA, L., MILLER, W., STONEKING, T., NHAN, M., WATERSTON, R. & WILSON, R. K. (2001). **Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2**. *Nature* 413, 852-6.
- MCLAGGAN, D., NAPRSTEK, J., BUURMAN, E. T. & EPSTEIN, W. (1994). **Interdependence of K⁺ and glutamate accumulation during osmotic adaptation of *Escherichia coli***. *J Biol Chem* 269, 1911-7.
- MEIDANIS, J., BRAGA, M. D. & VERJOVSKI-ALMEIDA, S. (2002). **Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa***. *Microbiol Mol Biol Rev* 66, 272-99.
- MENG, Y., LI, Y., GALVANI, C. D., HAO, G., TURNER, J. N., BURR, T. J. & HOCH, H. C. (2005). **Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility**. *J Bacteriol* 187, 5560-7.
- MHLANGA-MUTANGADURA, T., MORLIN, G., SMITH, A. L., EISENSTARK, A. & GOLOMB, M. (1998). **Evolution of the major pilus gene cluster of *Haemophilus influenzae***. *J Bacteriol* 180, 4693-703.
- MINET, A. D. & CHIQUET-EHRISMANN, R. (2000). **Phylogenetic analysis of teneurin genes and comparison to the rearrangement hot spot elements of *E. coli***. *Gene* 257, 87-97.
- MONTEIRO-VITORELLO, C. B., CAMARGO, L. E., VAN SLUYS, M. A., KITAJIMA, J. P., TRUFFI, D., DO AMARAL, A. M., HARAKAVA, R., DE OLIVEIRA, J. C., WOOD, D., DE OLIVEIRA, M. C., MIYAKI, C., TAKITA, M. A., DA SILVA, A. C., FURLAN, L. R., CARRARO, D. M., CAMAROTTE, G., ALMEIDA, N. F., JR., CARRER, H., COUTINHO, L. L., EL-DORRY, H. A., FERRO, M. I., GAGLIARDI, P. R., GIGLIOTI, E., GOLDMAN, M. H., GOLDMAN, G. H., KIMURA, E. T., FERRO, E. S., KURAMAE, E. E., LEMOS, E. G., LEMOS, M. V., MAURO, S. M., MACHADO, M.

- A., MARINO, C. L., MENCK, C. F., NUNES, L. R., OLIVEIRA, R. C., PEREIRA, G. G., SIQUEIRA, W., DE SOUZA, A. A., TSAI, S. M., ZANCA, A. S., SIMPSON, A. J., BRUMBLEY, S. M. & SETUBAL, J. C. (2004). **The genome sequence of the gram-positive sugarcane pathogen *Leifsonia xyli* subsp. *xyli***. *Mol Plant Microbe Interact* 17, 827-36.
- MORALES, P., FELIU, I., FERNANDEZ-GARCIA, E. & NUNEZ, M. (2004). **Volatile compounds produced in cheese by Enterobacteriaceae strains of dairy origin**. *J Food Prot* 67, 567-73.
- MOREIRA, L. M., DE SOUZA, R. F., ALMEIDA, N. F., JR., SETUBAL, J. C., OLIVEIRA, J. C., FURLAN, L. R., FERRO, J. A. & DA SILVA, A. C. (2004). **Comparative genomics analyses of citrus-associated bacteria**. *Annu Rev Phytopathol* 42, 163-84.
- MOREIRA, L. M., DE SOUZA, R. F., DIGIAMPIETRI, L. A., DA SILVA, A. C. & SETUBAL, J. C. (2005). **Comparative analyses of *Xanthomonas* and *Xylella* complete genomes**. *Omics* 9, 43-76.
- MUDGETT, M. B. (2005). **New insights to the function of phytopathogenic bacterial type III effectors in plants**. *Annu Rev Plant Biol* 56, 509-31.
- MUNRO, A. W., RITCHIE, G. Y., LAMB, A. J., DOUGLAS, R. M. & BOOTH, I. R. (1991). **The cloning and DNA sequence of the gene for the glutathione-regulated potassium-efflux system KefC of *Escherichia coli***. *Mol Microbiol* 5, 607-16.
- NAMEKATA, T. (1971). **Estudo comparativo entre *Xanthomonas citri* [Hasse] Dow., agente causal do cancro cítrico e *Xanthomonas citri* [Hasse] Dow., N. F. SP. *Aurantifolii*, agente causal da cancriose do limoeiro 'Galego'**. In *Plant Pathology*. Piracicaba: Escola Superior Luiz de Queiroz.
- NEVES, E. M., DAYOUB, M., DRAGONE, D. S. & NEVES, M. F. (2001). **Brazilian Citrus industry: economic and financial effects, 1996-2000**. *Rev. Bras. Frutic.* 23, 432-436.
- NOEL-GEORIS, I., VALLAEYS, T., CHAUVAUX, R., MONCHY, S., FALMAGNE, P., MERGEAY, M. & WATTIEZ, R. (2004). **Global analysis of the *Ralstonia metallidurans* proteome: prelude for the large-scale study of heavy metal response**. *Proteomics* 4, 151-79.
- NOEL, L., THIEME, F., GABLER, J., BUTTNER, D. & BONAS, U. (2003). **XopC and XopJ, two novel type III effector proteins from *Xanthomonas campestris* pv. *vesicatoria***. *J Bacteriol* 185, 7092-102.
- NORRIS, V. & MANNERS, B. (1993). **Deformations in the cytoplasmic membrane of *Escherichia coli* direct the synthesis of peptidoglycan. The hernia model**. *Biophys J* 64, 1691-700.
- 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. & 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.
- ORTH, A. B., ROYSE, D. J. & TIEN, M. (1993). **Ubiquity of lignin-degrading peroxidases among various wood-degrading fungi.** *Appl Environ Microbiol* 59, 4017-23.
- OSHIMA, K., KAKIZAWA, S., NISHIGAWA, H., JUNG, H. Y., WEI, W., SUZUKI, S., ARASHIDA, R., NAKATA, D., MIYATA, S., UGAKI, M. & NAMBA, S. (2004). **Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma.** *Nat Genet* 36, 27-9.
- PANDA, S., SATO, T. K., HAMPTON, G. M. & HOGENESCH, J. B. (2003). **An array of insights: application of DNA chip technology in the study of cell biology.** *Trends Cell Biol* 13, 151-6.
- PARKHILL, J., DOUGAN, G., JAMES, K. D., THOMSON, N. R., PICKARD, D., WAIN, J., CHURCHER, C., MUNGALL, K. L., BENTLEY, S. D., HOLDEN, M. T., SEBAIHIA, M., BAKER, S., BASHAM, D., BROOKS, K., CHILLINGWORTH, T., CONNERTON, P., CRONIN, A., DAVIS, P., DAVIES, R. M., DOWD, L., WHITE, N., FARRAR, J., FELTWELL, T., HAMLIN, N., HAQUE, A., HIEN, T. T., HOLROYD, S., JAGELS, K., KROGH, A., LARSEN, T. S., LEATHER, S., MOULE, S., O'GAORA, P., PARRY, C., QUAIL, M., RUTHERFORD, K., SIMMONDS, M., SKELTON, J., STEVENS, K., WHITEHEAD, S. & BARRELL, B. G. (2001a). **Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.** *Nature* 413, 848-52.
- PARKHILL, J., WREN, B. W., THOMSON, N. R., TITBALL, R. W., HOLDEN, M. T., PRENTICE, M. B., SEBAIHIA, M., JAMES, K. D., CHURCHER, C., MUNGALL, K. L., BAKER, S., BASHAM, D., BENTLEY, S. D., BROOKS, K., CERDENO-TARRAGA, A. M., CHILLINGWORTH, T., CRONIN, A., DAVIES, R. M., DAVIS, P., DOUGAN, G., FELTWELL, T., HAMLIN, N., HOLROYD, S., JAGELS, K., KARLYSHEV, A. V., LEATHER, S., MOULE, S., OYSTON, P. C., QUAIL, M., RUTHERFORD, K., SIMMONDS, M., SKELTON, J., STEVENS, K., WHITEHEAD, S. & BARRELL, B. G. (2001b). **Genome sequence of Yersinia pestis, the causative agent of plague.** *Nature* 413, 523-7.
- 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. & DA SILVA, A. C. (2005). **Whole-genome expression profiling of Xylella fastidiosa in response to growth on glucose.** *Omics* 9, 77-90.
- PENDRAK, M. L. & PERRY, R. D. (1991). **Characterization of a hemin-storage locus of Yersinia pestis.** *Biol Met* 4, 41-7.
- PERNA, N. T., MAYHEW, G. F., POSFAI, G., ELLIOTT, S., DONNENBERG, M. S., KAPER, J. B. & BLATTNER, F. R. (1998). **Molecular evolution of a pathogenicity island from enterohemorrhagic Escherichia coli O157:H7.** *Infect Immun* 66, 3810-7.

- PERNA, N. T., PLUNKETT, G., 3RD, BURLAND, V., MAU, B., GLASNER, J. D., ROSE, D. J., MAYHEW, G. F., EVANS, P. S., GREGOR, J., KIRKPATRICK, H. A., POSFAI, G., HACKETT, J., KLINK, S., BOUTIN, A., SHAO, Y., MILLER, L., GROTEBECK, E. J., DAVIS, N. W., LIM, A., DIMALANTA, E. T., POTAMOUSIS, K. D., APODACA, J., ANANTHARAMAN, T. S., LIN, J., YEN, G., SCHWARTZ, D. C., WELCH, R. A. & BLATTNER, F. R. (2001). **Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.** *Nature* 409, 529-33.
- PFLUGER, K. & MULLER, V. (2004). **Transport of compatible solutes in extremophiles.** *J Bioenerg Biomembr* 36, 17-24.
- PICKARD, D., WAIN, J., BAKER, S., LINE, A., CHOCHAN, S., FOOKES, M., BARRON, A., GAORA, P. O., CHABALGOITY, J. A., THANKY, N., SCHOLLES, C., THOMSON, N., QUAIL, M., PARKHILL, J. & DOUGAN, G. (2003). **Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7.** *J Bacteriol* 185, 5055-65.
- POPLAWSKY, A. R., URBAN, S. C. & CHUN, W. (2000). **Biological role of xanthomonadin pigments in *Xanthomonas campestris* pv. *campestris*.** *Appl Environ Microbiol* 66, 5123-7.
- QIAN, W., JIA, Y., REN, S. X., HE, Y. Q., FENG, J. X., LU, L. F., SUN, Q., YING, G., TANG, D. J., TANG, H., WU, W., HAO, P., WANG, L., JIANG, B. L., ZENG, S., GU, W. Y., LU, G., RONG, L., TIAN, Y., YAO, Z., FU, G., CHEN, B., FANG, R., QIANG, B., CHEN, Z., ZHAO, G. P., TANG, J. L. & HE, C. (2005). **Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*.** *Genome Res* 15, 757-67.
- RAZIN, S. (1997). **The minimal cellular genome of mycoplasma.** *Indian J Biochem Biophys* 34, 124-30.
- READ, T. D., SATOLA, S. W. & FARLEY, M. M. (2000). **Nucleotide sequence analysis of hypervariable junctions of *Haemophilus influenzae* pilus gene clusters.** *Infect Immun* 68, 6896-902.
- REIZER, J., BACHEM, S., REIZER, A., ARNAUD, M., SAIER, M. H., JR. & STULKE, J. (1999). **Novel phosphotransferase system genes revealed by genome analysis - the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*.** *Microbiology* 145 (Pt 12), 3419-29.
- REIZER, J., SUTRINA, S. L., WU, L. F., DEUTSCHER, J., REDDY, P. & SAIER, M. H., JR. (1992). **Functional interactions between proteins of the phosphoenolpyruvate:sugar phosphotransferase systems of *Bacillus subtilis* and *Escherichia coli*.** *J Biol Chem* 267, 9158-69.
- RILEY, M. A. & WERTZ, J. E. (2002a). **Bacteriocin diversity: ecological and evolutionary perspectives.** *Biochimie* 84, 357-64.
- RILEY, M. A. & WERTZ, J. E. (2002b). **Bacteriocins: evolution, ecology, and application.** *Annu Rev Microbiol* 56, 117-37.

- RODIONOV, D. A., GELFAND, M. S. & HUGOUVIEUX-COTTE-PATTAT, N. (2004). **Comparative genomics of the KdgR regulon in *Erwinia chrysanthemi* 3937 and other gamma-proteobacteria.** *Microbiology* 150, 3571-90.
- ROESSER, M. & MULLER, V. (2001). **Osmoadaptation in bacteria and archaea: common principles and differences.** *Environ Microbiol* 3, 743-54.
- ROMÁN, M. P., CAMBRA, M., JUÁREZ, J., MORENO, P., DURAN-VILA, N., TANAKA, F. O. A., ALVES, E., KITAJIMA, E. W., YAMAMOTO, P. T., BASSANEZI, R. B., TEIXEIRA, D. C., JESUS JUNIOR, W. C., AYRES, A. J., GIMENES-FERNANDES, N., RABENSTEIN, F., GIROTTO, L. F. & BOVÉ, J. M. (2004). **Sudden death of citrus in Brazil: a graft-transmissible bud union disease.** *Plant Disease* 88, 453-467.
- RONALD, P. C. & STASKAWICZ, B. J. (1988). **The avirulence gene *avrBs1* from *Xanthomonas campestris* pv. *vesicatoria* encodes a 50-kD protein.** *Mol Plant Microbe Interact* 1, 191-8.
- ROSSETI, V., GARNIER, M. B., BERETTA, M. J. G., TEIXEIRA, A. R. R., QUAGGIO, J. A., BATTAGLIA, O. C., GOMES, M. P., DE NEGRI, J. D. & BOVÉ, J. M. (1990). **Resultados preliminares de estudos sobre uma nova anormalidade dos citros observada nos estados de São Paulo e Minas Gerais.** *Summa Phytopathol.* 16, 13.
- SALANOUBAT, M., GENIN, S., ARTIGUENAVE, F., GOUZY, J., MANGENOT, S., ARLAT, M., BILLAULT, A., BROTTIER, P., CAMUS, J. C., CATTOLICO, L., CHANDLER, M., CHOISNE, N., CLAUDEL-RENARD, C., CUNNAC, S., DEMANGE, N., GASPIN, C., LAVIE, M., MOISAN, A., ROBERT, C., SAURIN, W., SCHIEX, T., SIGUIER, P., THEBAULT, P., WHALEN, M., WINCKER, P., LEVY, M., WEISSENBAACH, J. & BOUCHER, C. A. (2002). **Genome sequence of the plant pathogen *Ralstonia solanacearum*.** *Nature* 415, 497-502.
- SCARPARI, L. M., LAMBAIS, M. R., SILVA, D. S., CARRARO, D. M. & 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.
- SCHLEYER, M. & BAKKER, E. P. (1993). **Nucleotide sequence and 3'-end deletion studies indicate that the K(+)-uptake protein *kup* from *Escherichia coli* is composed of a hydrophobic core linked to a large and partially essential hydrophilic C terminus.** *J Bacteriol* 175, 6925-31.
- SCHLOSSER, A., HAMANN, A., BOSSEMEYER, D., SCHNEIDER, E. & BAKKER, E. P. (1993). **NAD⁺ binding to the *Escherichia coli* K(+)-uptake protein *TrkA* and sequence similarity between *TrkA* and domains of a family of dehydrogenases suggest a role for NAD⁺ in bacterial transport.** *Mol Microbiol* 9, 533-43.
- SCHLOSSER, A., MELDORF, M., STUMPE, S., BAKKER, E. P. & EPSTEIN, W. (1995). ***TrkH* and its homolog, *TrkG*, determine the specificity and kinetics of cation transport by the *Trk* system of *Escherichia coli*.** *J Bacteriol* 177, 1908-10.

- SCHOOLNIK, G. K. (2002). **Functional and comparative genomics of pathogenic bacteria.** *Curr Opin Microbiol* 5, 20-6.
- SCHUBERT, S., RAKIN, A., KARCH, H., CARNIEL, E. & HEESEMANN, J. (1998). **Prevalence of the "high-pathogenicity island" of Yersinia species among Escherichia coli strains that are pathogenic to humans.** *Infect Immun* 66, 480-5.
- SCHUBERT, T. S. & SUN, X. (1996). **Bacterial Citrus Canker.** *Plant Pathology Circular: Edited by Industry DoACSDoP* 377, 1-6.
- SETUBAL, J. C., MOREIRA, L. M. & DA SILVA, A. C. (2005). **Bacterial phytopathogens and genome science.** *Curr Opin Microbiol* 8, 595-600.
- SIEBERS, A. & ALTENDORF, K. (1988). **The K⁺-translocating Kdp-ATPase from Escherichia coli. Purification, enzymatic properties and production of complex- and subunit-specific antisera.** *Eur J Biochem* 178, 131-40.
- SILVA-STENICO, M. E., PACHECO, F. T., RODRIGUES, J. L., CARRILHO, E. & TSAI, S. M. (2005). **Growth and siderophore production of Xylella fastidiosa under iron-limited conditions.** *Microbiol Res* 160, 429-36.
- SILVER, S. (1996). **Transport of inorganic cations.** In *Escherichia coli and Salmonella: Cellular and Molecular biology.*, Second edn, pp. 1091-1102. Edited by F. C. Neidhart, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: ASM Press.
- SIMPSON, A. J., REINACH, F. C., ARRUDA, P., ABREU, F. A., ACENCIO, M., ALVARENGA, R., ALVES, L. M., ARAYA, J. E., BAIA, G. S., BAPTISTA, C. S., BARROS, M. H., BONACCORSI, E. D., BORDIN, S., BOVE, J. M., BRIONES, M. R., BUENO, M. R., CAMARGO, A. A., CAMARGO, L. E., CARRARO, M., D., CARRER, H., COLAUTO, N. B., COLOMBO, C., COSTA, F. F., COSTA, M. C., COSTA-NETO, C. M., COUTINHO, L. L., CRISTOFANI, M., DIAS-NETO, E., DOCENA, C., EL-DORRY, H., FACINCANI, A. P., FERREIRA, A. J., FERREIRA, V. C., FERRO, J. A., FRAGA, J. S., FRANCA, S. C., FRANCO, M. C., FROHME, M., FURLAN, L. R., GARNIER, M., GOLDMAN, G. H., GOLDMAN, M. H., GOMES, S. L., GRUBER, A., HO, P. L., HOHEISEL, J. D., JUNQUEIRA, M. L., KEMPER, E. L., KITAJIMA, J. P., KRIEGER, J. E., KURAMAE, E. E., LAIGRET, F., LAMBAIS, M. R., LEITE, L. C., LEMOS, E. G., LEMOS, M. V., LOPES, S. A., LOPES, C. R., MACHADO, J. A., MACHADO, M. A., MADEIRA, A. M., MADEIRA, H. M., MARINO, C. L., MARQUES, M. V., MARTINS, E. A., MARTINS, E. M., MATSUKUMA, A. Y., MENCK, C. F., MIRACCA, E. C., MIYAKI, C. Y., MONTERIRO-VITORELLO, C. B., MOON, D. H., NAGAI, M. A., NASCIMENTO, A. L., NETTO, L. E., NHANI, A., JR., NOBREGA, F. G., NUNES, L. R., OLIVEIRA, M. A., DE OLIVEIRA, M. C., DE OLIVEIRA, R. C., PALMIERI, D. A., PARIS, A., PEIXOTO, B. R., PEREIRA, G. A., PEREIRA, H. A., JR., PESQUERO, J. B., QUAGGIO, R. B., ROBERTO, P. G., RODRIGUES, V., DE, M. R. A. J., DE ROSA, V. E., JR., DE SA, R. G., SANTELLI, R. V., SAWASAKI, H. E., DA SILVA, A. C., DA SILVA, A. M., DA SILVA, F. R., DA SILVA, W. A., JR., *et al.* (2000). **The**

genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *Nature* 406, 151-7.

- SLATER, H., ALVAREZ-MORALES, A., BARBER, C. E., DANIELS, M. J. & DOW, J. M. (2000). **A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*.** *Mol Microbiol* 38, 986-1003.
- SMOLKA, M. B., MARTINS, D., WINCK, F. V., SANTORO, C. E., CASTELLARI, R. R., FERRARI, F., BRUM, I. J., GALEMBECK, E., DELLA COLETTA FILHO, H., MACHADO, M. A., MARANGONI, S. & NOVELLO, J. C. (2003). **Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution.** *Proteomics* 3, 224-37.
- STARR, M. P., JENKINS, L. B., BUSSEY, L. B. & ANDREWES, A. G. (1977). **Chemotaxonomic significance of the xanthomonadins, novel brominated aryl-polyene pigments produced by bacteria of the genus *Xanthomonas*.** *Arch. Microbiol.* 113, 1-9.
- STEINMANN, D., KOPLIN, R., PUHLER, A. & NIEHAUS, K. (1997). ***Xanthomonas campestris* pv. *campestris* *lpsI* and *lpsJ* genes encoding putative proteins with sequence similarity to the alpha- and beta-subunits of 3-oxoacid CoA-transferases are involved in LPS biosynthesis.** *Arch Microbiol* 168, 441-7.
- STOCK, J. B. & SURETTE, M. G. (1996). **Chemotaxis.** In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, pp. 1103-29. Edited by F. C. Neidhardt, R. I. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikopp, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: ASP Press.
- STUDHOLME, D. J., DOWNIE, J. A. & PRESTON, G. M. (2005). **Protein domains and architectural innovation in plant-associated Proteobacteria.** *BMC Genomics* 6, 17.
- STUMPE, S., SCHLOSSER, A., SCHLEYER, M. & BAKKER, E. P. (1996). **K⁺ circulation across prokaryotic cell membrane: K⁺ uptake systems.** In *Transport process in eukaryotic and prokaryotic organisms*, pp. 473 - 499. Edited by W. N. Konings, H. R. Kaback & J. S. Lolkema. Amsterdam: Elsevier.
- SUN, X., STALL, R. E., JONES, J. B., CUBERO, J., GOTTWALD, T. R., GRAHAM, J. H., DIXON, W. N., SCHUBERT, T. S., CHALOUX, P. H. & STROMBERG, V. K. (2004). **Detection and Characterization of a new strain of citrus canker bacteria from Key/Mexican lime and alemow in south Florida.** *Plant Disease* 88, 1179-1188.
- SVENSSON-EK, M., THOMAS, J. W., GENNIS, R. B., NILSSON, T. & BRZEZINSKI, P. (1996). **Kinetics of electron and proton transfer during the reaction of wild type and helix VI mutants of cytochrome bo3 with oxygen.** *Biochemistry* 35, 13673-80.

- SWINGS, J. G. & CIVEROLO, E. L. (1993). **Xanthomonas**, First edn, pp. 399. London.
- TANG, J. L., LIU, Y. N., BARBER, C. E., DOW, J. M., WOOTTON, J. C. & DANIELS, M. J. (1991). **Genetic and molecular analysis of a cluster of rpf genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in Xanthomonas campestris pathovar campestris**. *Mol Gen Genet* 226, 409-17.
- TATUSOV, R. L., GALPERIN, M. Y., NATALE, D. A. & KOONIN, E. V. (2000). **The COG database: a tool for genome-scale analysis of protein functions and evolution**. *Nucleic Acids Res* 28, 33-6.
- THARA, V. K., SEILANIANITZ, A. R., DENG, Y., DONG, Y., YANG, Y., TANG, X. & ZHOU, J. M. (2004). **Tobacco genes induced by the bacterial effector protein AvrPto**. *Mol Plant Microbe Interact* 17, 1139-45.
- THIEME, F., KOEBNIK, R., BEKEL, T., BERGER, C., BOCH, J., BUTTNER, D., CALDANA, C., GAIGALAT, L., GOESMANN, A., KAY, S., KIRCHNER, O., LANZ, C., LINKE, B., MCHARDY, A. C., MEYER, F., MITTENHUBER, G., NIES, D. H., NIESBACH-KLOSGEN, U., PATSCHKOWSKI, T., RUCKERT, C., RUPP, O., SCHNEIKER, S., SCHUSTER, S. C., VORHOLTER, F. J., WEBER, E., PUHLER, A., BONAS, U., BARTELS, D. & KAISER, O. (2005). **Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium Xanthomonas campestris pv. vesicatoria revealed by the complete genome sequence**. *J Bacteriol* 187, 7254-66.
- THOMPSON, J. D., HIGGINS, D. G. & 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.
- TREUNER-LANGE, A., KUHN, A. & DURRE, P. (1997). **The kdp system of Clostridium acetobutylicum: cloning, sequencing, and transcriptional regulation in response to potassium concentration**. *J Bacteriol* 179, 4501-12.
- VAN SLUYS, M. A., DE OLIVEIRA, M. C., MONTEIRO-VITORELLO, C. B., MIYAKI, C. Y., FURLAN, L. R., CAMARGO, L. E., DA SILVA, A. C., MOON, D. H., TAKITA, M. A., LEMOS, E. G., MACHADO, M. A., FERRO, M. I., DA SILVA, F. R., GOLDMAN, M. H., GOLDMAN, G. H., LEMOS, M. V., EL-DORRY, H., TSAI, S. M., CARRER, H., CARRARO, D. M., DE OLIVEIRA, R. C., NUNES, L. R., SIQUEIRA, W. J., COUTINHO, L. L., KIMURA, E. T., FERRO, E. S., HARAKAVA, R., KURAMAE, E. E., MARINO, C. L., GIGLIOTI, E., ABREU, I. L., ALVES, L. M., DO AMARAL, A. M., BAIA, G. S., BLANCO, S. R., BRITO, M. S., CANNAVAN, F. S., CELESTINO, A. V., DA CUNHA, A. F., FENILLE, R. C., FERRO, J. A., FORMIGHIERI, E. F., KISHI, L. T., LEONI, S. G., OLIVEIRA, A. R., ROSA, V. E., JR., SASSAKI, F. T., SENA, J. A., DE SOUZA, A. A., TRUFFI, D., TSUKUMO, F., YANAI, G. M., ZAROS, L. G., CIVEROLO, E. L., SIMPSON, A. J., ALMEIDA, N. F., JR., SETUBAL, J. C. & KITAJIMA, J. P. (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.

- 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. & SIMPSON, A. J. (2002). **Comparative genomic analysis of plant-associated bacteria.** *Annu Rev Phytopathol* 40, 169-89.
- VENTOSA, A., MARQUEZ, M. C., GARABITO, M. J. & ARAHAL, D. R. (1998a). **Moderately halophilic gram-positive bacterial diversity in hypersaline environments.** *Extremophiles* 2, 297-304.
- VENTOSA, A., NIETO, J. J. & OREN, A. (1998b). **Biology of moderately halophilic aerobic bacteria.** *Microbiol Mol Biol Rev* 62, 504-44.
- VERNIÈRE, C., HARTUNG, J. S., PRUVOST, O. P., CIVEROLO, E. L., ALVAREZ, A. M., MAESTRI, P. & LUISETTI, J. (1998). **Characterization of phenotypically distinct strains of *Xanthomonas axonopodis* pv. *citri* from Southwest Asia.** *European Journal of Plant pathology* 104, 477-487.
- VOJNOV, A. A., SLATER, H., DANIELS, M. J. & DOW, J. M. (2001). **Expression of the gum operon directing xanthan biosynthesis in *Xanthomonas campestris* and its regulation in planta.** *Mol Plant Microbe Interact* 14, 768-74.
- VORHOLTER, F. J., NIEHAUS, K. & PUHLER, A. (2001). **Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core.** *Mol Genet Genomics* 266, 79-95.
- VORHOLTER, F. J., THIAS, T., MEYER, F., BEKEL, T., KAISER, O., PUHLER, A. & NIEHAUS, K. (2003). **Comparison of two *Xanthomonas campestris* pathovar *campestris* genomes revealed differences in their gene composition.** *J Biotechnol* 106, 193-202.
- WALDERHAUG, M. O., POLAREK, J. W., VOELKNER, P., DANIEL, J. M., HESSE, J. E., ALTENDORF, K. & EPSTEIN, W. (1992). **KdpD and KdpE, proteins that control expression of the kdpABC operon, are members of the two-component sensor-effector class of regulators.** *J Bacteriol* 174, 2152-9.
- WALL, D. & KAISER, D. (1999). **Type IV pili and cell motility.** *Mol Microbiol* 32, 1-10.
- WATT, S. A., WILKE, A., PATSCHKOWSKI, T. & NIEHAUS, K. (2005). **Comprehensive analysis of the extracellular proteins from *Xanthomonas campestris* pv. *campestris* B100.** *Proteomics* 5, 153-67.
- WEGRZYN, G. (2001). **The minimal genome paradox.** *J Appl Genet* 42, 385-92.
- WIDMER, F., SEIDLER, R. J., GILLEVET, P. M., WATRUD, L. S. & DI GIOVANNI, G. D. (1998). **A highly selective PCR protocol for detecting 16S rRNA genes of the genus *Pseudomonas* (sensu stricto) in environmental samples.** *Appl Environ Microbiol* 64, 2545-53.
- WIGGERICH, H. G., KLAUKE, B., KOPLIN, R., PRIEFER, U. B. & PUHLER, A. (1997). **Unusual structure of the tonB-exb DNA region of *Xanthomonas campestris***

pv. campestris: tonB, exbB, and exbD1 are essential for ferric iron uptake, but exbD2 is not. *J Bacteriol* 179, 7103-10.

WIGGERICH, H. G. & PUHLER, A. (2000). **The exbD2 gene as well as the iron-uptake genes tonB, exbB and exbD1 of Xanthomonas campestris pv. campestris are essential for the induction of a hypersensitive response on pepper (Capsicum annuum).** *Microbiology* 146 (Pt 5), 1053-60.

WILSON, G. G. & MURRAY, N. E. (1991). **Restriction and modification systems.** *Annu Rev Genet* 25, 585-627.

WOOD, D. W., SETUBAL, J. C., KAUL, R., MONKS, D. E., KITAJIMA, J. P., OKURA, V. K., ZHOU, Y., CHEN, L., WOOD, G. E., ALMEIDA, N. F., JR., WOO, L., CHEN, Y., PAULSEN, I. T., EISEN, J. A., KARP, P. D., BOVEE, D., SR., CHAPMAN, P., CLENDENNING, J., DEATHERAGE, G., GILLET, W., GRANT, C., KUTYAVIN, T., LEVY, R., LI, M. J., MCCLELLAND, E., PALMIERI, A., RAYMOND, C., ROUSE, G., SAENPHIMMACHAK, C., WU, Z., ROMERO, P., GORDON, D., ZHANG, S., YOO, H., TAO, Y., BIDDLE, P., JUNG, M., KRESPAN, W., PERRY, M., GORDON-KAMM, B., LIAO, L., KIM, S., HENDRICK, C., ZHAO, Z. Y., DOLAN, M., CHUMLEY, F., TINGEY, S. V., TOMB, J. F., GORDON, M. P., OLSON, M. V. & NESTER, E. W. (2001a). **The genome of the natural genetic engineer Agrobacterium tumefaciens C58.** *Science* 294, 2317-23.

WOOD, D. W., SETUBAL, J. C. & NESTER, E. W. (2004). **Genome sequence analysis of prokaryotic plant pathogens.** In *Plant microbiology*, pp. 223-241. Edited by A. Gillings MaH. London: BIOS Scientific Publishers (Oxford, UK) in cooperation with Marcel Dekker.

WOOD, J. M., BREMER, E., CSONKA, L. N., KRAEMER, R., POOLMAN, B., VAN DER HEIDE, T. & SMITH, L. T. (2001b). **Osmosensing and osmoregulatory compatible solute accumulation by bacteria.** *Comp Biochem Physiol A Mol Integr Physiol* 130, 437-60.

WYLIE, J. L., BERNEGGER-EGLI, C., O'NEIL, J. D. & WOROBEK, E. A. (1993). **Biophysical characterization of OprB, a glucose-inducible porin of Pseudomonas aeruginosa.** *J Bioenerg Biomembr* 25, 547-56.

WYLIE, J. L. & WOROBEK, E. A. (1995). **The OprB porin plays a central role in carbohydrate uptake in Pseudomonas aeruginosa.** *J Bacteriol* 177, 3021-6.

YAN, D., IKEDA, T. P., SHAUGER, A. E. & KUSTU, S. (1996). **Glutamate is required to maintain the steady-state potassium pool in Salmonella typhimurium.** *Proc Natl Acad Sci U S A* 93, 6527-31.

YANG, Y., CHEN, R. F. & SHIARIS, M. P. (1994). **Metabolism of naphthalene, fluorene, and phenanthrene: preliminary characterization of a cloned gene cluster from Pseudomonas putida NCIB 9816.** *J Bacteriol* 176, 2158-64.

ZAKHARYAN, E. & TRCHOUNIAN, A. (2001). **K⁺ influx by Kup in Escherichia coli is accompanied by a decrease in H⁺ efflux.** *FEMS Microbiol Lett* 204, 61-4.

- ZHANG, L. H., FATH, M. J., MAHANTY, H. K., TAI, P. C. & KOLTER, R. (1995). **Genetic analysis of the colicin V secretion pathway.** *Genetics* 141, 25-32.
- ZHAO, B., LIN, X., POLAND, J., TRICK, H., LEACH, J. & HULBERT, S. (2005). **A maize resistance gene functions against bacterial streak disease in rice.** *Proc Natl Acad Sci U S A* 102, 15383-8.

ANEXOS

Anexos inclusos no CD-ROM suplementar a esta Tese

- **ANEXO 1: ARTIGO PUBLICADO**

Setubal JC, Moreira LM, da Silva ACR. *Bacterial Phytopathogens and Genomic Science*. **Curr Opin Microbiol**. 2005; 8(5): 595 - 600.

- **ANEXO 2: ARTIGO PUBLICADO**

Koide T, Zaini PA, Moreira LM, Vencio RZ, Matsukuma AY, Durham AM, Teixeira DC, El-Dorry H, Monteiro PB, da Silva AC, Verjovski-Almeida S, da Silva AM, Gomes SL. *DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of Xylella fastidiosa delineate genes important for bacterial virulence*. **J Bacteriol**. 2004 Aug;186(16):5442-9.

- **ANEXO 3: ARTIGO PUBLICADO**

Moreira LM, de Souza RF, Almeida NF Jr, Setubal JC, Oliveira JC, Furlan LR, Ferro JA, da Silva AC. *Comparative genomics analyses of citrus-associated bacteria*. **Annu Rev Phytopathol**. 2004; 42:163-84.

- **ANEXO 4: ARTIGO PUBLICADO**

Moreira LM, De Souza RF, Digiampietri LA, Da Silva AC, Setubal JC. *Comparative analyses of Xanthomonas and Xylella complete genomes*. **OMICS**. 2005 Spring;9(1):43-76.

- **ANEXO 5: ARTIGO PUBLICADO**

Pashalidis S, Moreira LM, Zaini PA, Campanharo JC, Alves LM, Ciapina LP, Vencio RZ, Lemos EG, Da Silva AM, Da Silva AC. *Whole-genome expression profiling of Xylella fastidiosa in response to growth on glucose*. **OMICS**. 2005 Spring;9(1):77-90.

- **ANEXO 6: LISTA DE GENES PRESENTES NO XACARRAY**

- **ANEXO 7: SÚMULA CURRICULAR**

ANEXO 1

**TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA**

ARTIGO PUBLICADO

Bacterial Phytopathogens and Genomic Science

Setubal JC, Moreira LM, da Silva ACR

Curr Opin Microbiol. 2005; 8(5): 595 - 600.

PMID: 16125997 - PubMed - indexed for MEDLINE.

ELSEVIER

Bacterial phytopathogens and genome science

João C Setubal¹, Leandro M Moreira² and Ana CR da Silva³

There are now fourteen completed genomes of bacterial phytopathogens, all of which have been generated in the past six years. These genomes come from a phylogenetically diverse set of organisms, and range in size from 870 kb to more than 6 Mb. The publication of these annotated genomes has significantly helped our understanding of bacterial plant disease. These genomes have also provided important information about bacterial evolution. Examples of recently completed genomes include: *Pseudomonas syringae* pv *tomato*, which is notable for its large repertoire of effector proteins; *Leifsonia xyli* subsp. *xyli*, the first Gram-positive bacterial genome to be sequenced; and *Phytoplasma asteris*, the small genome that lacks important functions previously thought to be essential in a bacterium.

Addresses

¹ Virginia Bioinformatics Institute and Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060-0477, USA

² Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Lineu Prestes 748, Cx. Postal 05509-900, São Paulo, SP, Brazil

³ Alellyx Applied Genomics, Rua James C. Maxwell 320, Techno Park, 13067-850, Campinas, SP, Brazil

Corresponding author: da Silva, Ana CR (ana.rasera@alellyx.com.br)

Current Opinion in Microbiology 2005, **8**:595–600

This review comes from a themed issue on
Genomics
Edited by Gerhard Gottschalk and Stephan C
Schuster
Available online 25th August 2005

1369-5274/\$ – see front matter
© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.mib.2005.08.015

Introduction

It has been five years since Simpson *et al.* [1] published the genome of *Xylella fastidiosa* 9a5c — the first plant pathogen (phytopathogen) to be sequenced. This article aims to briefly review the status of genomics of bacterial plant pathogens since that landmark achievement.

The current status of bacterial plant pathogen genome projects are presented in Table 1. This shows that there are now 14 completely sequenced genomes of phytopathogens, and that the sequencing of 19 others is in progress or has already been partially completed. In terms of phylogenetic diversity, the proteobacteria predominate, corresponding to 79% of entries in the table; of those, bacteria in the gamma subdivision are the most

common (54% of the phytopathogens). It appears that the majority of bacterial diseases of important crops have been covered by the projects listed. Table 1 also shows that the stage of strain sequencing has been reached for phytopathogens: there are eight species for which more than one sequence has been generated, with *X. fastidiosa* contributing four sequences to the table. This means that fine-grained comparative genomics has become as important for the analysis of bacterial plant pathogens as it has for genomics in general.

In this review we highlight recently completed genome projects. We also review post-genomic work that uses the results of genome projects as a basis. These efforts show that genome sequencing of bacterial plant pathogens is providing valuable insights into genome science, in addition to the potential impact that it can bring to the control of agriculturally important diseases.

Recently completed genomes

Van Sluys *et al.* [2] have published an article that reviews the results of plant-associated bacterial genome projects performed before mid-2002; that work was complemented shortly afterwards by Wood *et al.* [3], who focused on plant pathogens. Since then, six new projects have been completed and their genomes published. In this section, we will review these results.

Pseudomonas syringae pv. *tomato* DC3000

Pseudomonas is a bacterial genus that contains a wide range of animal and plant pathogens. *Pseudomonas syringae* pv. *tomato* DC3000 (PTO) was the first *Pseudomonas* phytopathogen to be sequenced [4]. A major result of this project was the identification of several clusters of genes that encode type III secretion system effector proteins (31 proteins confirmed and 19 predicted effectors). Most of these proteins are essential for pathogenicity. Orthologous genes are present in *Xanthomonas* and *Ralstonia* species, but many of them have not yet been confirmed to be expressed and secreted [5^{*}]. Subsequent work on *P. syringae* strains has increased the list of predicted effector-coding genes to more than 200 (<http://pseudomonas-syringae.org>). Buell *et al.* [4] presented a detailed analysis of additional genes in PTO that are related to pathogenicity — a total of 298 genes were discovered. Of these, 96 are unique to PTO with respect to *Pseudomonas aeruginosa* and *Pseudomonas putida*; thus, they are presumed to play a role specific to plant–PTO interactions.

Xanthomonas oryzae pv. *oryzae* KAC10331 strain

Xanthomonas oryzae pv. *oryzae* KAC10331 strain (XOO) was the first *Xanthomonas* genome to be sequenced from

Glossary

Niche restriction: A sub-population becomes segregated in an ecological niche (such as the xylem) in which evolutionary pressures are different to those of the previous environment. The genome then responds to these different evolutionary conditions. Among the possible responses are gene loss and genome shrinkage.

Putative highly-expressed genes: These are genes in a genome that have an amino acid composition close to that of known highly-expressed genes, such as ribosomal proteins.

Transposon insertional mutant library: A population of bacteria artificially generated in laboratory that has incorporated a transposon insertion. These insertions cause disruptions of genes and probable mutant phenotypes. It is a routine approach that is used to find the function of unknown genes.

a species that infects exclusively monocotyledonous plants, in this case rice [6]. This genome was found to be similar to the two previously sequenced *Xanthomonas* genomes (*Xanthomonas axonopodis* pv. *citri* strain 306 [XAC] and *Xanthomonas campestris* pv. *campestris* strain ATCC [XCC-ATCC]; [7]) in terms of gene content (more than 80% of XOO's genes have orthologs in the other two species), but whole-chromosome alignments showed substantial rearrangements compared to both XAC and XCC-ATCC. The fact that XOO contains more than twice the number of transposable elements than either XAC or

Table 1

Status of bacterial plant pathogen genome projects (based on Wood *et al.* [3]).

Organism	Phylo. group ^a	Disease	Major hosts	Status	Reference
<i>Agrobacterium tumefaciens</i> C58	Alpha	Crown gall	Dicotyledonous plants	Complete	[49,50]
<i>Agrobacterium vitis</i> S4	Alpha	Crown gall	Grape	In progress	AGRO
<i>Agrobacterium radiobacter</i> K84	Alpha	Biocontrol for crown gall	Biocontrol	In progress	AGRO
<i>Burkholderia cepacia</i>	Beta	Sour skin	Onion	In progress	JGI
<i>Ralstonia solanacearum</i> GM11000	Beta	Bacterial wilt	Potato, banana, tomato	Complete	[51]
<i>Ralstonia solanacearum</i> MolK2	Beta	Bacterial wilt	Banana	In progress	GENOSCOPE
<i>Ralstonia solanacearum</i> 1609	Beta	Bacterial wilt	Potato	In progress	GENOSCOPE
<i>Ralstonia metallidurans</i>	Beta	Potential for bioremediation of metals	Non-pathogenic	In progress	JGI
<i>Erwinia amylovora</i> Ea273	Gamma	Fire Blight	Apple, pear	In progress	SANGER
<i>Erwinia carotovora atroseptica</i>	Gamma	Potato blackleg	Potato	Complete	[9**]
<i>Erwinia chrysanthemi</i> 3937	Gamma	Bulb soft rot	Many crops	In progress	WISC/TIGR
<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000	Gamma	Bacterial spot	Tomato, <i>Arabidopsis</i>	Complete	[4]
<i>Pseudomonas syringae</i> pv. <i>phaseolicola</i> 1448A	Gamma	Plant rot	Bean	Complete	PPI
<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	Gamma	Bacterial brown spot of bean	Bean	Complete	JGI
<i>Xanthomonas axonopodis citri</i>	Gamma	Citrus Canker	Citrus	Complete	[7]
<i>Xanthomonas axonopodis aurantifolii</i> B	Gamma	Cancrosis B	Citrus	In progress	ONSA
<i>Xanthomonas axonopodis aurantifolii</i> C	Gamma	Cancrosis C	Citrus	In progress	ONSA
<i>Xanthomonas campestris campestris</i> ATCC 33913	Gamma	Black rot of cruciferous	Brassica	Complete	[7]
<i>Xanthomonas campestris campestris</i> 8004	Gamma	Black rot of cruciferous	Brassica	Complete	[8**]
<i>Xanthomonas campestris campestris</i> B100	Gamma	Black rot of cruciferous	Brassica	In progress	UB
<i>Xanthomonas oryzae oryzae</i>	Gamma	Bacterial blight	Rice	Complete	[6]
<i>Xanthomonas oryzae oryzicola</i>	Gamma	Bacterial leaf streak	Rice	In progress	IOWAST
<i>Xylella fastidiosa</i> 9a5c	Gamma	Citrus variegated chlorosis	Citrus	Complete	[1]
<i>Xylella fastidiosa Temecula</i>	Gamma	Pierce's Disease	Grape	Complete	[16]
<i>Xylella fastidiosa Ann 1</i>	Gamma	Oleander leaf scorch	Oleander	In progress	JGI
<i>Xylella fastidiosa Dixon</i>	Gamma	Almond leaf scorch	Almond	In progress	JGI
<i>Clavibacter michiganensis</i> subsp. <i>michiganensis</i>	High GC Gram +	Mich canker of tomato	Tomato	In progress	UB
<i>Clavibacter michiganensis</i> subsp. <i>sepedonicus</i>	High GC Gram +	Bacterial ring rot of potato	Potato	In progress	SANGER
<i>Leifsonia xyli xyli</i>	High GC Gram +	Ratoon stunting	Sugar cane	Complete	[16]
<i>Spiroplasma kunkelii</i>	Firmicutes	Corn stunt disease	Corn	In progress	USDA/UO
Onion yellows phytoplasma OY-M	Firmicutes	Onion yellows disease	Onion	Complete	[11**]
Aster yellows witches' broom phytoplasma AY-WB	Firmicutes	Aster yellows witches' broom	Aster, maize	In progress	OSU
Western X phytoplasma	Firmicutes	Western X disease	Peach, cherry, nut	In progress	IG

Web sites are as follows: AGRO, www.agrobacterium.org; AY-WB, <http://www.oardc.ohio-state.edu/phytoplasma>; CNHGC, <http://www.chgc.sh.cn/>; GENOSCOPE, http://www.genoscope.cns.fr/externe/English/Projets/Projet_Y/organisme_Y.html; IG, <http://www.integratedgenomics.com/>; IOWAST, JGI (Joint Genome Institute), www.jgi.doe.gov; <http://www.public.iastate.edu/~ajbog>; ONSA, <http://genoma4.iq.usp.br/xanthomonas>; OSU, <http://www.oardc.ohio-state.edu/phytoplasma/>; OU, <http://www.genome.ou.edu/spiro.html>; PPI, www.pseudomonas-syringae.org; SANGER, www.sanger.ac.uk/Projects/Microbes; TIGR, <http://www.tigr.org>; UB, <http://www.genetik.uni-bielefeld.de/>; USDA, <http://www.ba.ars.usda.gov/mppl/research/kunkelii.html>; WISC, <http://www.ahabs.wisc.edu/~pernalab/erwinia/index.html>.

^a Alpha, beta, and gamma refer to subdivisions of the Proteobacteria.

XCC-ATCC (which already contain quite a few of these) suggests that these elements are pervasive in the *Xanthomonas* genus and that they play a major role in the rearrangements. XOO was found to have 245 unique genes with respect to XAC and XCC-ATCC; among these were found a restriction-modification gene, a TonB-dependent siderophore receptor, genes for toxin production (e.g. *mlrB* and *rtx*), and a type III secretion system effector.

***Xanthomonas campestris* pv. *campestris* 8004**

This is a good example of how the study of closely related genomes can yield important insights. *Xanthomonas campestris* pv. *campestris* 8004 (XCC-8004) and XCC-ATCC [7] are two strains of the same species, but their pathogenicity properties in at least two hosts are remarkably different [8^{••}]. As expected, their genome sequences are very similar. As in the case of XOO, several chromosomal rearrangements were observed, with respect to both XAC and XCC-ATCC. In fact, surprisingly, XAC and XCC-ATCC have fewer rearrangements between them than XCC-ATCC has with respect to XCC-8004. Qian *et al.* [8^{••}] identified 108 and 62 putative genes specific to XCC-8004 and XCC-ATCC, respectively. The authors went one step beyond genome sequencing and performed a screening of a high-density transposon insertional mutant library with 16 512 clones against the host *Brassica oleracea* (see glossary). This effort confirmed previous predictions of pathogenicity-related genes and revealed new ones: in particular, some were found in a chromosomal region specific of XCC-8004 with respect to XCC-ATCC. This shows a direct correlation between genome dynamics and XCC virulence [8^{••}], an important result.

Erwinia carotovora* subsp. *atroseptica

Erwinia carotovora subsp. *atroseptica* (ECA) is a member of the Enterobacteriaceae, which contains several well-studied human pathogens, such as *E. coli*. The authors report that about 1500 (33%) of its predicted genes were specific to it compared to other sequenced enterobacteria, which suggests that significant genomic specialization has occurred [9^{••}]. This set was found to have a high proportion of pathogenicity-related genes, many of which are present within genomic islands. Two groups of these genes, one of which was related to type IV secretion and the other to polyketide phytotoxin synthesis, were then confirmed to be virulence determinants by an insertional mutant study.

Leifsonia xyli* subsp. *xyli

The genome of *Leifsonia xyli* subsp. *xyli* (LXX) is notable for being the first Gram-positive plant pathogen to be sequenced [10^{••}]. This study suggested that LXX is in a process of genome shrinkage, possibly as a result of niche (the sugar cane xylem) restriction (see glossary). Strong evidence of this comes from the relatively large fraction of pseudogenes (13%), which is significantly higher than in other plant pathogens sequenced to date. Other clues

include the apparent lack of a complete set of genes for oxidative phosphorylation and glucogenesis as well as for cobalamin, biotin and thiamin synthesis. The sugar cane xylem is nutritionally poor, but the number of ATP-binding cassette transporter-coding genes and those involved with sugar intake was found to be comparable to that of free-living soil organisms. This could be evidence that LXX has a free-living ancestor. The number of pathogenicity-related genes is relatively small when compared to other bacterial phytopathogens (105, of which 20 are pseudogenes). As is the case in many other bacterial pathogens, several pathogenicity-related genes (that code for proteins such as cellulase, pectinase, wilt-inducing protein, lysozyme and desaturase) are probably the result of lateral transfer.

***Phytoplasma asteris* OY strain**

Of bacterial plant pathogens sequenced to date, *Phytoplasma asteris* OY strain (PA-OY) has the smallest genome, of only 870 kbp [11^{••}]. It encodes fewer metabolic functions than *Mycoplasmas*, which has long been thought to be close to the 'minimal bacterial genome'. This is the first time that a prokaryotic genome has been found to lack all eight ATP-synthase subunits. It might import host ATP by a mechanism that has yet to be discovered or it might be particularly dependent on the glycolysis pathway. PA-OY has 27 genes that encode different transporter functions; some of these are duplicated, which suggests that PA-OY depends heavily on metabolites from the host cell. Indeed, this dependence could explain some of the disease symptoms. Only two possible virulence factors have been identified: a glucanase and a hemolysin-like protein.

Post-genome analyses

Thanks to projects such as those described above, there is now enough genomic information with regard to phytopathogens to allow relatively large-scale *in silico* comparative studies. We now summarize two recent studies of this kind.

Fu *et al.* [12[•]] looked for variations in base-pair composition to model putative highly-expressed genes among six bacterial phytopathogens (see glossary) (PTO, XAC, XCC-ATCC, *Agrobacterium tumefaciens*, *Ralstonia solanacearum* and *X. fastidiosa* 9a5c). Among these, many genes were found that are implicated in pathogenicity, such as a flagellin-coding gene and some that code for outer membrane proteins (e.g. a molybdate-responsive regulator and a peptidoglycan-associated lipoprotein). One interesting finding by Fu *et al.* [12[•]] is that xanthan gum genes in *X. fastidiosa* 9a5c should be highly expressed, whereas apparently they are not so in *Xanthomonas* species. This finding might be related to a hypothesis put forward by Moreira *et al.* [13[•]] and Pashalidis *et al.* [14], that gum genes in *X. fastidiosa* 9a5c are regulated by a mechanism other than sugar concentration levels.

The other study, by Studholme *et al.* [15^{••}], looked at protein domains in plant-associated proteobacteria. They showed that certain well-known pathogenicity-related genes contain domains that are specific to plant-associated bacteria; the same method yielded domains of unknown function that are specific to the same group, therefore proposing a valuable hypothesis to test. They also discovered domains that appear to be over-represented in plant-associated proteobacterial genomes and are therefore likely to play important roles in bacterial–host interactions.

The availability of complete genomic sequences has opened up a wide research front that is creating a rapidly increasing amount of literature. Several comparative genomics studies have targeted genes involved in mechanisms of adaptation, pathogenicity or virulence related to a specific pathogen or host [2,4,6,7,8^{••},9^{••},13[•],16,17–19,20[•],21–23]. Several significant large-scale functional genomics analyses have been performed. Examples of these include: transcription analysis of a determined physiological condition or investigation of gene composition in strains or pathovars of a species based on DNA microarray analysis [24–30]; knockout of genes by a transposon insertion that induces phenotype variation or alteration of a specific metabolic pathway [8^{••},31,32,33]; proteome analysis [34–37]; and analysis of protein–protein interactions using a two-hybrid system as a model [38[•],39,40]. Other relevant works include protein structure determination [41,42], analysis of genes that might be involved in vertical or horizontal gene transfer [43–46], and *in silico* analyses that determined improved medium culture and new metabolic pathways based only on genomic information [47,48].

Conclusions

The stories told by the six recently sequenced bacterial plant pathogens emphasize themes that are now becoming familiar in genomics. These themes include genome rearrangements and lateral transfer events as powerful factors in shaping evolution and survival; niche restrictions that lead to genome decay (LXX) and to extreme genome shrinkage (PA-OY); and development of a gene repertoire that is specific to plant hosts, most notably in PTO. In conclusion, quite apart from the potential economic relevance of these genomic sequencing efforts of agricultural pests, it is clear that the newly sequenced genomes are making significant contributions to genome science in general.

Owing to these genome projects, we now have extensive catalogs of pathogenicity-related genes that belong to most crop-relevant bacterial species. As was done for ECA and XCC-8004, the genome information can be further leveraged by functional studies based on large-scale insertional mutants. However, much remains to be learnt. In common with all bacterial genomes sequenced to date, bacterial plant pathogen genomes still have

sizable fractions of genes for which no putative function can be assigned using computational methods. In the most recently completed genome (XCC-8004) and in PA-OY, large fractions of genes have no functional assignment (37% and 34%, respectively). This is significant as these two genomes are of quite different size. Many of these unassigned genes are members of homologous families that are shared by several distinct species. Targeted studies or new computational approaches are required to shed light on the function of these genes. The array of new experimental data generated by functional studies, integrated in a meaningful way with genomic data and *in silico* cross-genome studies, should enable researchers to achieve substantial progress in understanding more precisely the mechanisms that cause the various diseases. The next five years will probably see many new genome sequences, both because sequencing costs have been reduced and because plant pathogens are now recognized as potential bioterrorism weapons.

One initiative that might have a large impact on computational analyses of bacterial plant pathogen genomes is the Plant-Associated Microbe Gene Ontology Interest Group (<http://pamgo.vbi.vt.edu>). The goals of this group are to expand the Gene Ontology (<http://www.geneontology.org>) to include terms and relationships that reflect knowledge about microbial genes relevant to their associations with plants, and to use these new terms to create annotated reference microbial gene sets to be made available to the community. The hope is that this will lead to new discoveries that have been hampered in the past by the lack of standards regarding microbe–host interaction terms. Among the microbial genomes that will be used to create the reference gene sets are three bacterial plant pathogens: *Erwinia carotovora chrysanthemi*, PTO and *Agrobacterium tumefaciens* C58.

Acknowledgements

We thank Boris Vinatzer for critical reading of this manuscript. This work was supported in part by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; PhD fellowship to LMM).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Simpson AJ, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, Alves LM, Araya JE, Baia GS, Baptista CS *et al.*: **The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis.** *Nature* 2000, **406**:151–157.
2. Van Sluys MA, Monteiro-Vitorello CB, Camargo LE, Menck CF, Da Silva AC, Ferro JA, Oliveira MC, Setubal JC, Kitajima JP, Simpson AJ: **Comparative genomic analysis of plant-associated bacteria.** *Annu Rev Phytopathol* 2002, **40**:169–189.
3. Wood DW, Setubal JC, Nester EW: **Genome sequence analysis of prokaryotic plant pathogens.** In *Plant Microbiology*. Edited by

- Gillings M and Holmes A. BIOS Scientific Publishers in cooperation with Marcel Dekker; 2004:223-241.
4. Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, Gwinn ML, Dodson RJ, Deboy RT, Durkin AS, Kolonay JF *et al.*: **The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000.** *Proc Natl Acad Sci USA* 2003, **100**:10181-10186.
 5. Chang JH, Urbach JM, Law TF, Arnold LW, Hu A, Gombor S, Grant SR, Ausubel FM, Dangi JL: **A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*.** *Proc Natl Acad Sci USA* 2005, **102**:2549-2554.
- An interesting study that uses a smart approach to screen interactions between type III secretion system effectors in *Pseudomonas*.
6. Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS *et al.*: **The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice.** *Nucleic Acids Res* 2005, **33**:577-586.
 7. da Silva AC, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Van Sluys MA, Almeida NF, Alves LM *et al.*: **Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities.** *Nature* 2002, **417**:459-463.
 8. Qian W, Jia Y, Ren SX, He YQ, Feng JX, Lu LF, Sun Q, Ying G, Tang DJ, Tang H *et al.*: **Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*.** *Genome Res* 2005, **15**:757-767.
- This study shows that the type four secretion system is also important for pathogenicity, in addition to other secretion systems.
9. Bell KS, Sebaihia M, Pritchard L, Holden MT, Hyman LJ, Holeva MC, Thomson NR, Bentley SD, Churcher LJ, Mungall K *et al.*: **Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors.** *Proc Natl Acad Sci USA* 2004, **101**:11105-11110.
- The authors combine genomic sequencing and insertion mutants to reveal the importance of type four secretion system proteins and phytoalexins in pathogenicity of *E. carotovora*.
10. Monteiro-Vitorello CB, Camargo LE, Van Sluys MA, Kitajima JP, Truffi D, do Amaral AM, Harakava R, de Oliveira JC, Wood D, de Oliveira MC *et al.*: **The genome sequence of the Gram-positive sugarcane pathogen *Leifsonia xyli* subsp. *xyli*.** *Mol Plant Microbe Interact* 2004, **17**:827-836.
- This paper, together with [11*], shows two interesting examples of bacterial adaptation.
11. Oshima K, Kakizawa S, Nishigawa H, Jung HY, Wei W, Suzuki S, Arashida R, Nakata D, Miyata S, Ugaki M *et al.*: **Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma.** *Nat Genet* 2004, **36**:27-29.
- See annotation to [10*].
12. Fu QS, Li F, Chen LL: **Gene expression analysis of six GC-rich Gram-negative phytopathogens.** *Biochem Biophys Res Commun* 2005, **332**:380-387.
- The authors use gene sequence analysis to infer patterns of gene expression for whole genomes, and derive interesting results.
13. Moreira LM, de Souza RF, Almeida NF Jr, Setubal JC, Oliveira JC, Furlan LR, Ferro JA, da Silva AC: **Comparative genomics analyses of citrus-associated bacteria.** *Annu Rev Phytopathol* 2004, **42**:163-184.
- This paper, together with [20*], reveals important differences between *Xanthomonadaceae* genomes that might be related to pathogenicity and lifestyle.
14. Pashalidis S, Moreira LM, Zaini PA, Campanharo JC, Alves LM, Ciapina LP, Vencio RZ, Lemos EG, Da Silva AM, Da Silva AC: **Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose.** *OMICS* 2005, **9**:77-90.
 15. Studholme DJ, Downie JA, Preston GM: **Protein domains and architectural innovation in plant-associated Proteobacteria.** *BMC Genomics* 2005, **6**:17.
- The authors were able to identify protein domains that are specific to plant associated bacteria. Some of these domains are found in distantly-related bacteria.
16. Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LE, da Silva AC, Moon DH, Takita MA, Lemos EG *et al.*: **Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*.** *J Bacteriol* 2003, **185**:1018-1026.
 17. Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, Bernal A, D'Souza M, Ivanova N, Kapatral V *et al.*: **Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights.** *Genome Res* 2002, **12**:1556-1563.
 18. Bhattacharyya A, Stilwagen S, Ivanova N, D'Souza M, Bernal A, Lykidis A, Kapatral V, Anderson I, Larsen N, Los T *et al.*: **Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains.** *Proc Natl Acad Sci USA* 2002, **99**:12403-12408.
 19. Vorholter FJ, Thias T, Meyer F, Bekel T, Kaiser O, Puhler A, Niehaus K: **Comparison of two *Xanthomonas campestris* pathovar *campestris* genomes revealed differences in their gene composition.** *J Biotechnol* 2003, **106**:193-202.
 20. Moreira LM, de Souza RF, Digiampietri LA, da Silva ACR, Setubal JC: **Comparative analyses of *Xanthomonas* and *Xylella* complete genomes.** *OMICS* 2005, **1**:43-76.
- See annotation to [13*].
21. Rodionov DA, Gelfand MS, Hugouvieux-Cotte-Pattat N: **Comparative genomics of the KdgR regulon in *Erwinia chrysanthemi* 3937 and other gamma-proteobacteria.** *Microbiology* 2004, **150**:3571-3590.
 22. Zhao Y, Ma Z, Sundin GW: **Comparative genomic analysis of the pPT23A plasmid family of *Pseudomonas syringae*.** *J Bacteriol* 2005, **187**:2113-2126.
 23. Buttner D, Noel L, Thieme F, Bonas U: **Genomic approaches in *Xanthomonas campestris* pv. *vesicatoria* allow fishing for virulence genes.** *J Biotechnol* 2003, **106**:203-214.
 24. Astua-Monge G, Freitas-Astua J, Bacocina G, Roncoletta J, Carvalho SA, Machado MA: **Expression profiling of virulence and pathogenicity genes of *Xanthomonas axonopodis* pv. *citri*.** *J Bacteriol* 2005, **187**:1201-1205.
 25. de Souza AA, Takita MA, Coletta-Filho HD, Caldana C, Goldman GH, Yanai GM, Muto NH, de Oliveira RC, Nunes LR, Machado MA: **Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity.** *Mol Plant Microbe Interact* 2003, **16**:867-875.
 26. de Souza AA, Takita MA, Coletta-Filho HD, Caldana C, Yanai GM, Muto NH, de Oliveira RC, Nunes LR, Machado MA: **Gene expression profile of the plant pathogen *Xylella fastidiosa* during biofilm formation *in vitro*.** *FEMS Microbiol Lett* 2004, **237**:341-353.
 27. Koide T, Zaini PA, Moreira LM, Vencio RZ, Matsukuma AY, Durham AM, Teixeira DC, El-Dorry H, Monteiro PB, da Silva AC *et al.*: **DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence.** *J Bacteriol* 2004, **186**:5442-5449.
 28. Pashalidis S, Moreira LM, Zaini PA, Campanharo JC, Alves LMC, Ciapina LP, Vencio RZN, Lemos EGM, da Silva AM, da Silva ACR: **Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose.** *OMICS* 2005, **1**:77-90.
 29. de Souza AA, Takita MA, Pereira EO, Coletta-Filho HD, Machado MA: **Expression of pathogenicity-related genes of *Xylella fastidiosa* *in vitro* and *in planta*.** *Curr Microbiol* 2005, **50**:223-228.
 30. Nunes LR, Rosato YB, Muto NH, Yanai GM, da Silva VS, Leite DB, Gonçalves ER, de Souza AA, Coletta-Filho HD, Machado MA *et al.*: **Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements.** *Genome Res* 2003, **13**:570-578.
 31. Koide T, da Silva Neto JF, Gomes SL, Marques MV: **Insertional transposon mutagenesis in the *Xylella fastidiosa* Citrus Variegated Chlorosis strain with transposome.** *Curr Microbiol* 2004, **48**:247-250.

32. da Silva Neto JF, Koide T, Gomes SL, Marques MV: **Site-directed gene disruption in *Xylella fastidiosa***. *FEMS Microbiol Lett* 2002, **210**:105-110.
 33. Brumbley SM, Petrasovits LA, Murphy RM, Nagel RJ, Candy JM, Hermann SR: **Establishment of a functional genomics platform for *Leifsonia xyli* subsp. *xyli***. *Mol Plant Microbe Interact* 2004, **17**:175-183.
 34. Watt SA, Wilke A, Patschkowski T, Niehaus K: **Comprehensive analysis of the extracellular proteins from *Xanthomonas campestris* pv. *campestris* B100**. *Proteomics* 2005, **5**:153-167.
 35. Smolka MB, Martins D, Winck FV, Santoro CE, Castellari RR, Ferrari F, Brum IJ, Galembeck E, Della Coletta Filho H, Machado MA *et al.*: **Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution**. *Proteomics* 2003, **3**:224-237.
 36. Kazemi-Pour N, Condemine G, Hugouvieux-Cotte-Pattat N: **The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi***. *Proteomics* 2004, **4**:3177-3186.
 37. Noel-Georis I, Vallaeys T, Chauvaux R, Monchy S, Falmagne P, Mergeay M, Wattiez R: **Global analysis of the *Ralstonia metallidurans* proteome: prelude for the large-scale study of heavy metal response**. *Proteomics* 2004, **4**:151-179.
 38. Alegria MC, Docena C, Khater L, Ramos CH, da Silva AC, Farah CS: **New protein-protein interactions identified for the regulatory and structural components and substrates of the type III secretion system of the phytopathogen *Xanthomonas axonopodis* Pathovar *citri***. *J Bacteriol* 2004, **186**:6186-6197.
- An interesting paper in which the authors nicely integrate genomic data and yeast two-hybrid experiments to reveal important interactions between proteins of the type III secretion system.
39. Alegria MC, Souza DP, Andrade MO, Docena C, Khater L, Ramos CH, da Silva AC, Farah CS: **Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. *citri***. *J Bacteriol* 2005, **187**:2315-2325.
 40. Kabisch U, Landgraf A, Krause J, Bonas U, Boch J: **Type III secretion chaperones ShcS1 and ShcO1 from *Pseudomonas syringae* pv. *tomato* DC3000 bind more than one effector**. *Microbiology* 2005, **151**:269-280.
 41. Arcuri HA, Canduri F, Pereira JH, da Silveira NJ, Camera Junior JC, de Oliveira JS, Basso LA, Palma MS, Santos DS, de Azevedo Junior WF: **Molecular models for shikimate pathway enzymes of *Xylella fastidiosa***. *Biochem Biophys Res Commun* 2004, **320**:979-991.
 42. Osiro D, Muniz JR, Coleta Filho HD, de Sousa AA, Machado MA, Garratt RC, Colnago LA: **Fatty acid synthesis in *Xylella fastidiosa*: correlations between genome studies, ¹³C NMR data, and molecular models**. *Biochem Biophys Res Commun* 2004, **323**:987-995.
 43. Chen J, Civerolo EL, Jarret RL, Van Sluys MA, de Oliveira MC: **Genetic discovery in *Xylella fastidiosa* through sequence analysis of selected randomly amplified polymorphic DNAs**. *Curr Microbiol* 2005, **50**:78-83.
 44. Patil PB, Sonti RV: **Variation suggestive of horizontal gene transfer at a lipopolysaccharide (lps) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the bacterial leaf blight pathogen of rice**. *BMC Microbiol* 2004, **4**:40.
 45. Sarkar SF, Guttman DS: **Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen**. *Appl Environ Microbiol* 2004, **70**:1999-2012.
 46. Lima WC, Paquola ACM, Van Sluys M-A, Menck CFM: **Non-gamma proteobacteria gene islands contribution to *Xanthomonas* genome**. *OMICS Submitted* 2005, **9**:160-172.
 47. Lemos EG, Alves LM, Campanharo JC: **Genomics-based design of defined growth media for the plant pathogen *Xylella fastidiosa***. *FEMS Microbiol Lett* 2003, **219**:39-45.
 48. Etchegaray A, Silva-Stenico ME, Moon DH, Tsai SM: **In silico analysis of nonribosomal peptide synthetases of *Xanthomonas axonopodis* pv. *citri*: identification of putative siderophore and lipopeptide biosynthetic genes**. *Microbiol Res* 2004, **159**:425-437.
 49. Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B, Goldman BS, Cao Y, Askenazi M, Halling C *et al.*: **Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58**. *Science* 2001, **294**:2323-2328.
 50. Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF Jr *et al.*: **The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58**. *Science* 2001, **294**:2317-2323.
 51. Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Ariat M, Billault A, Brottier P, Camus JC, Cattolico L *et al.*: **Genome sequence of the plant pathogen *Ralstonia solanacearum***. *Nature* 2002, **415**:497-502.

Elsevier joins major health information initiative

Elsevier has joined with scientific publishers and leading voluntary health organizations to create patientINFORM, a groundbreaking initiative to help patients and caregivers close a crucial information gap. patientINFORM is a free online service dedicated to disseminating medical research and is scheduled to launch in 2005.

Elsevier will provide the voluntary health organizations with increased online access to our peer-reviewed biomedical journals immediately upon publication, together with content from back issues. The voluntary health organizations will integrate the information into materials for patients and link to the full text of selected research articles on their websites.

patientINFORM has been created to allow patients seeking the latest information about treatment options online access to the most up-to-date, reliable research available for specific diseases.

'Not only will patientINFORM connect patients and their caregivers with the latest research, it will help them to put it into context. By making it easier to understand research findings, patientINFORM will empower patients to have a more productive dialogue with their physicians and make well-informed decisions about care', said Harmon Eyre, M.D., national chief medical officer of the American Cancer Society.

For more information, visit www.patientinform.org

ANEXO 2

TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA

ARTIGO PUBLICADO

DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineate genes important for bacterial virulence.

Koide T, Zaini PA, Moreira LM, Vencio RZ, Matsukuma AY, Durham AM, Teixeira DC, El-Dorry H, Monteiro PB, da Silva AC, Verjovski-Almeida S, da Silva AM, Gomes SL.

J Bacteriol. 2004 Aug;186(16):5442-9.

PMID: 15292146 -PubMed - indexed for MEDLINE.

DNA Microarray-Based Genome Comparison of a Pathogenic and a Nonpathogenic Strain of *Xylella fastidiosa* Delineates Genes Important for Bacterial Virulence†

Tie Koide,^{1‡} Paulo A. Zaini,^{1‡} Leandro M. Moreira,¹ Ricardo Z. N. Vêncio,² Adriana Y. Matsukuma,¹ Alan M. Durham,³ Diva C. Teixeira,⁴ Hamza El-Dorry,¹ Patrícia B. Monteiro,^{4§} Ana Claudia R. da Silva,¹ Sergio Verjovski-Almeida,¹ Aline M. da Silva,^{1*} and Suely L. Gomes^{1*}

Departamento de Bioquímica, Instituto de Química,¹ and Departamento de Estatística² and Ciência da Computação,³ Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, and Fundo de Defesa da Citricultura, Araraquara,⁴ Brasil

Received 20 February 2004/Accepted 22 March 2004

Xylella fastidiosa is a phytopathogenic bacterium that causes serious diseases in a wide range of economically important crops. Despite extensive comparative analyses of genome sequences of *Xylella* pathogenic strains from different plant hosts, nonpathogenic strains have not been studied. In this report, we show that *X. fastidiosa* strain J1a12, associated with citrus variegated chlorosis (CVC), is nonpathogenic when injected into citrus and tobacco plants. Furthermore, a DNA microarray-based comparison of J1a12 with 9a5c, a CVC strain that is highly pathogenic and had its genome completely sequenced, revealed that 14 coding sequences of strain 9a5c are absent or highly divergent in strain J1a12. Among them, we found an arginase and a fimbrial adhesin precursor of type III pilus, which were confirmed to be absent in the nonpathogenic strain by PCR and DNA sequencing. The absence of arginase can be correlated to the inability of J1a12 to multiply in host plants. This enzyme has been recently shown to act as a bacterial survival mechanism by down-regulating host nitric oxide production. The lack of the adhesin precursor gene is in accordance with the less aggregated phenotype observed for J1a12 cells growing in vitro. Thus, the absence of both genes can be associated with the failure of the J1a12 strain to establish and spread in citrus and tobacco plants. These results provide the first detailed comparison between a nonpathogenic strain and a pathogenic strain of *X. fastidiosa*, constituting an important step towards understanding the molecular basis of the disease.

Xylella fastidiosa is a gram-negative bacterium, limited to the plant xylem vessels, which is responsible for worldwide economic losses due to diseases caused in a variety of plants of agricultural relevance. This bacterium is transmitted to new host plants during xylem sap feeding by insect vectors and spreads from the site of infection to colonize the xylem, a water transport network of vessels composed of lignified dead cells. Bacterial cells attach to the vessel wall, forming biofilm-like colonies that, depending on the size, can occlude the xylem vessels, blocking water transport and causing water stress symptoms (22, 35).

Different strains of *X. fastidiosa* have been reported to infect a wide range of plants, including grapevines and citrus, almond, and pear trees, among others (26). In the United States for instance, Pierce's disease prevents profitable viticulture if leafhopper vectors are present at high densities (1, 14). In Brazil, citrus variegated chlorosis (CVC) is responsible for

great financial losses to the citrus agroindustry, being detected in one-third of the citrus trees. Orange production quickly decreases in orchards affected by CVC, as fruits become hardened and of no commercial value. Interestingly, within the majority of host plants, *X. fastidiosa* behaves as a harmless endophyte (27).

Several *X. fastidiosa* strains have had their genomes completely or partially sequenced, and genome comparative analysis with different pathogenic strains of *X. fastidiosa* pointed to common candidate virulence determinants as well as strain-specific genomic signatures (4, 24, 32, 37). However, no information is available about the genome composition of nonpathogenic *Xylella* strains, which would contribute to more direct insights on pathogenicity mechanisms.

Genome-wide comparison between pathogenic and nonpathogenic strains within a species is a useful strategy for identifying candidate genes important for virulence. DNA microarray-based genome composition analysis is a good alternative to full genome sequencing and has been used in comparative studies to analyze various bacterial pathogens including *Mycobacterium tuberculosis* (3), *Helicobacter pylori* (29), *Pseudomonas aeruginosa* (38), *Bacillus anthracis* (28), *Yersinia pestis*, and *Yersinia pseudotuberculosis* (12).

In this report, we show that *X. fastidiosa* strain J1a12, which was isolated from citrus and is suitable for genetic transformation (6, 23), elicits few or no CVC symptoms when inoculated into citrus and tobacco plants. Furthermore, a DNA micro-

* Corresponding author. Mailing address: Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-900 São Paulo, SP, Brasil. Phone, fax, and e-mail for Suely L. Gomes: 5511 3091 3826, 5511 3815 5579, sulgomes@iq.usp.br. Phone fax, and e-mail for Aline M. da Silva: 5511 3091 2182, 5511 3815 5579, almsilva@iq.usp.br.

† Supplemental material for this article may be found at <http://jb.asm.org/>.

‡ These authors contributed equally to this work.

§ Present address: Allex Applied Genomics, Campinas, SP, Brasil.

TABLE 1. Evaluation of *C. sinensis* and *N. tabacum* plants inoculated with *X. fastidiosa* CVC strains 9a5c and J1a12

Strain	Test	No. of infected plants/total no. of plants at time (mo) postinoculation				
		<i>C. sinensis</i>			<i>N. tabacum</i>	
		5	8	15	1.5	3
9a5c	PCR ^a	8/14	14/14	13/13 ^b	5/5	5/5
	Symptom evaluation	0/14	3/14	10/13 ^b	5/5	5/5
J1a12	PCR	0/14	1/14	0/14	0/5	0/5
	Symptom evaluation	0/14	0/14	0/14	0/5	0/5

^a PCR experiments were performed on samples drawn from the plant xylem by using primers designed to identify *Xylella* citrus isolates (25).

^b One of the plants died 15 months after inoculation.

array-based genome composition analysis was performed by comparing J1a12 with strain 9a5c, which produces typical CVC symptoms (20) but is resistant to transformation with DNA in vitro (23), a drawback for its genetic manipulation. Our microarray data revealed that the great majority of the coding sequences (CDS) are highly conserved on both strains. However, 14 CDS were shown to be absent or highly divergent in the nonpathogenic strain. Expression profiling of both strains, PCR and reverse transcription (RT)-PCR with CDS-specific primers, and DNA sequence analysis were used to validate the genomic differences observed.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and pathogenicity tests. Triply cloned *X. fastidiosa* strains 9a5c (20) and J1a12 (23) isolated from CVC symptomatic *Citrus sinensis* (L.) Osbeck trees (sweet orange) were grown in periwinkle wilt broth medium (7) at 28°C in the dark with 100 rpm rotatory agitation. As shown in Table 1, strain J1a12 is nonpathogenic, despite its isolation from a citrus plant with CVC symptoms. This is possibly due to a mixed population of bacteria infecting the plant (see "Final remarks" below). A culture started from a single colony was weekly passaged through serial transfers at a 1/100 dilution in periwinkle wilt medium. For citrus plant experiments, 9a5c and J1a12 strains with 9 and 24 weekly passages, respectively, were used; for tobacco plant experiments, 11 weekly passages were used for both strains. Mechanical inoculation of plants was performed essentially as described in reference 20 for *C. sinensis* and as described in reference 21 for *Nicotiana tabacum* (accession clelandii). Further details are found in the supplemental material. The detection of *X. fastidiosa* in host plants followed the procedure described in reference 25, which consists of PCR experiments with a pair of primers (CVC-1 and 272-int) designed to identify *Xylella* strains isolated from citrus plants.

Microarray construction and hybridization. A 6,152-element DNA microarray was printed containing unique internal fragments of 2,692 CDS spotted at least in duplicate, representing 94.5% of all of the 2,848 CDS annotated by Simpson et al. (32). DNA fragments ranging from 200 to 1,000 bp were PCR amplified with CDS-specific primers (18- to 23-mers) designed with PRIMER3 (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) and based on the annotated genome sequence of *X. fastidiosa* strain 9a5c (<http://aeg.lbi.ic.unicamp.br/xf/>). A full list of primers, PCR product sizes, and their nucleotide sequences are available at the project site (<http://verjo19.iq.usp.br/xylella/microarray/>). The arrays were hybridized with DNA fragments from strain 9a5c in combination with itself, J1a12, or grapevine strain Temecula (kindly provided by Marie-Anne Van Sluys, University of São Paulo) labeled separately with either Cy3- or Cy5-dCTP analogs (see the supplemental material). Expression profiling studies were carried out by labeling total RNA with the Cy-Scribe post labeling kit (Amersham Biosciences) according to the manufacturer's instructions.

Data acquisition and normalization. Microarray slides were scanned by using a Generation III DNA scanner (Amersham Biosciences), and fluorescence intensity values (I_{Cy3} and I_{Cy5}) from each spot were extracted by using Array-Vision, version 6.0, software (Imaging Research, Inc.). Raw fluorescence intensity and normalized data are available at the project site (<http://verjo19.iq.usp.br/xylella/microarray/>). Data normalization was carried out by LOWESS fitting on

an M versus A plot (39), where M is the ratio of fluorescence intensities of the two measurements for each spot [defined as $M = \log_2(I_{Cy5}/I_{Cy3})$] and A is the geometric mean of the fluorescence intensities [defined as $A = 1/2 \times \log_2(I_{Cy5} \times I_{Cy3})$]. The normalization script is available at the project site.

CDS classification process. To determine hybridization noise and to estimate dynamic cutoff values for classifying a CDS as equally present in both strains (9a5c and J1a12) we used the hybridization data collected from three independent homotypic experiments (9a5c versus 9a5c). For this kind of experiment, also called self-self hybridization, the microarrays were cohybridized with strain 9a5c DNA separately labeled with either Cy3- or Cy5-dCTP analogs. As verified in the M versus A plot, there is a dependence of the hybridization intensity log ratio of each spot (M) with the mean log intensity of each spot (A). Thus, we have determined a cutoff value for each interval in the A axis by using kernel density estimators. We chose kernel density estimators (31) instead of the normal probability density function (16) because we experimentally derived the null distribution as the result of homotypic experiments and verified that it does not present a Gaussian behavior (further information is available at the project site [<http://verjo19.iq.usp.br/xylella/microarray/>]). The density distribution was integrated around the mode peak until 0.995 probability was reached, defining intensity-dependent noise threshold cutoff values (credibility intervals) based on experimental data from 9a5c versus 9a5c homotypic experiments, thus setting an interval where the hybridization ratio is considered to be 1:1 (e.g., $-2.5 < M < 1.8$, for the lowest accepted intensities at $A = 2$). These credibility intervals were subsequently used in the analysis of replicas of 9a5c versus J1a12 hybridization experiments to nonparametrically estimate the null distribution of the statistical test H_0 : CDS is present in both J1a12 and 9a5c strains. Spots outside the credibility intervals present strong evidence against a 1:1 ratio. Using these criteria, four categories were defined for the CDS in the J1a12 genome based on its orthologous 9a5c counterpart: (i) equally present in both strains, (ii) divergent in strain J1a12, (iii) highly divergent or absent in J1a12, and (iv) higher copy number in J1a12. Category i includes all of the CDS for which $\geq 60\%$ of the replicas were inside the credibility intervals.

CDS presenting negative log ratio values outside the credibility intervals could be classified as divergent in strain J1a12 (category ii) or highly divergent or absent in J1a12 (category iii). To distinguish between these categories, we performed four control experiments where the microarrays were cohybridized with DNA from the sequenced grapevine strain Temecula (37) and strain 9a5c labeled with either Cy3- or Cy5-dCTP analogs. Next, we derived a correspondence between the hybridization intensity log ratio for each CDS amplicon and its respective nucleotide sequence identity in both strains. Amplicon sequences exhibiting nucleotide identities smaller than 20% were defined as category iii, highly divergent or absent. Their respective log ratio M was taken as the cutoff threshold (M_{cutoff}) of category iii. The log ratio value with the least false callings was an M_{cutoff} value of -1.7 . At this threshold, 19 false positives and 23 false negatives were observed (0.76 and 0.92%, respectively). Category iii includes CDS with an M value of < -1.7 and a P value smaller than 0.05 in a t test against the null hypothesis H_0 : $M \geq -1.7$. Sequencing some amplicons from J1a12 that were outside the credibility intervals and checking the divergence between 9a5c and J1a12 sequences further supported the adequacy of the threshold. The remaining CDS that have a negative log ratio value, with an M value of ≥ -1.7 , were considered divergent (category ii). Those CDS with positive log ratio values were considered to be present at higher copy numbers in J1a12 (category iv). CDS with low reproducibility (less than 60% of replicas in a single category) were excluded from the analysis.

Validation of microarray data. PCR and RT-PCR experiments were performed with CDS-specific primers to further investigate the status of CDS in the genome of strain J1a12 classified by using the microarray data. The reactions were carried out with genomic DNA or cDNA from strain 9a5c or J1a12 with 35 cycles of amplification. A sample of CDS presenting log ratios outside the credibility intervals, as determined by homotypic hybridization experiments (9a5c versus 9a5c), was chosen to perform the validation. Among the 64 CDS with negative log ratios (categories ii and iii), the following 33 CDS were randomly chosen for PCR validation: XF0077, XF0496, XF0497, XF0500, XF0663, XF0684, XF0696, XF0890, XF1250, XF1306, XF1581, XF1588, XF1589, XF1646, XF1663, XF1664, XF1686, XF1708, XF1709, XF1860, XF1863, XF1874, XF1877, XF1878, XF1884, XF1968, XF2193, XF2195, XF2307, XF2722, XF2768, XF2772, and XFb0001. In addition, the following CDS with positive log ratios (category iv) were also tested by PCR: XF0513, XF0514, XF0515, XF0516, XF0518, XF0519, XF0521, XF1933, XF1934, XF1935, XF1936, and XF1937. A 4- μ l sample of each reaction mixture was electrophoresed in agarose gels, and DNA was stained with ethidium bromide. The amplicons were then classified by visual inspection as absent, same copy number, or more abundant in strain J1a12 in relation to strain 9a5c. In addition, DNA sequence determination was carried out

for a few CDS. For that, PCR products obtained with primers based on neighboring CDS were cloned in pGEM-TEasy vector (Promega) and dideoxy sequencing reactions were performed with 100 ng of plasmid DNA in Big Dye Terminator sequencing reactions (Applied Biosystems) according to the manufacturer's instructions. Sequencing reactions were carried out with either CDS-specific or T7 promoter primers.

RESULTS AND DISCUSSION

***X. fastidiosa* CVC strain J1a12 is nonpathogenic.** In planta pathogenicity tests were carried out with strains 9a5c and J1a12, and results are presented in Table 1. Inoculation of citrus plants with strain 9a5c resulted in multiplication of *X. fastidiosa* in all plants analyzed 8 months after infection. Plants were also evaluated visually for characteristic CVC symptoms, and positive symptoms were observed in the leaves of 77% of the citrus plants 15 months after inoculation. In contrast, no symptoms were observed in the leaves of plants inoculated with strain J1a12 up to 15 months after infection.

Similar results were observed with tobacco plants as the experimental host. As shown in Table 1, none of the plants inoculated with J1a12 presented symptoms or were colonized by the bacteria. In contrast, all tobacco plants inoculated with 9a5c presented the lesions characteristic of *X. fastidiosa* CVC infection, as previously described (21).

These results indicate that strain J1a12 shows a nonpathogenic phenotype, in contrast to the highly pathogenic behavior of strain 9a5c. In addition, Table 1 shows that plant colonization by strain J1a12 is very inefficient, as no bacteria were detected in the plant xylem by PCR experiments with a pair of primers (CVC-1 and 272-int) which are specific for *Xylella* isolates from citrus plants (25). This pair of primers amplifies a genomic region of approximately 500 bp, from chromosome position 1051239 to 1051745 of the 9a5c genome, encompassing 195 bp of CDS XF1100 and an intergenic region upstream of this CDS. It is important to stress that these primers can amplify the correct DNA fragment when directly tested in J1a12 in vitro cultures (6, 23).

Genotyping by DNA microarray analysis. To investigate whether the differences in phenotype observed between strains 9a5c and J1a12 could be associated with differences at the DNA level, we have constructed a DNA microarray encompassing 2,692 CDS, which represents 94.5% of all CDS described in the reference strain 9a5c (32). Total DNAs isolated from strains 9a5c and J1a12 were separately labeled with either Cy3- or Cy5-dCTP fluorescent analogs, and competitive microarray hybridizations were performed. Raw and normalized hybridization data are available at the project site (<http://verjo19.iq.usp.br/xylella/microarray/>). An initial screening revealed that 292 CDS presented either low signal intensity or poor reproducibility and were excluded from further analysis. As detailed in Material and Methods, the remaining 2,400 CDS were classified into four categories according to the normalized hybridization fluorescence intensity ratios of J1a12 over 9a5c DNA samples determined for each CDS. Among the 2,400 CDS, approximately 96% were found to have a log ratio (M) around 0 and were classified as equally present in both strains (category i). One example is shown in Fig. 1A. This figure shows an M versus A plot, i.e., normalized intensities log ratios (M) versus the average of log intensities (A) of all the replicas for a given CDS. This kind of graph shows the dependence of

the ratio on the overall intensity of each spot, indicating that, for genes with low hybridization intensity signals (A values below 2), the observed ratios have a higher intrinsic dispersion, as determined by homotypic hybridizations. As a result, different cutoff values for M were used for different ranges of intensity (A) when classifying a CDS as equally present in both strains. M versus A plots showing the reproducibility of the data for each CDS are available at the project site.

Fifty CDS were found to have a hybridization intensity log ratio of $-1.7 < M < -0.3$ and were classified as divergent in strain J1a12 (category ii, an example is shown in Fig. 1B). Table 2 lists only the CDS with hybridization intensity log ratios between -0.5 and -1.7 . For the complete list of CDS in this category see Table S1 in the supplemental material. Fourteen CDS were found to have the log ratios of < -1.7 and were classified as absent or highly divergent in strain J1a12 (Table 3). A typical example is shown in Fig. 1C. Within this group, three CDS, namely XF0077, XF1250, and XF1646, were especially interesting due to their putative involvement in pathogenesis and will be discussed in more detail later. Most of the CDS in this category belong to the previously defined flexible gene pool of *Xylella*, which includes integrated prophages and genomic islands (4, 24). For instance, CDS XF1860, XF1874, XF1878, and XF1884 belong to genomic island 4 (24). This region has a different GC content and altered codon bias, which are common features of laterally transferred elements (15). However, CDS XF0077, XF1250, XF1646, XF1707, and XF1708 are not mapped within any genomic island and do not show altered GC content or codon bias.

In addition, 40 CDS (see Table S2 in the supplemental material) were found to have a log ratio of > 0 and were considered as possibly presenting a higher number of copies in J1a12 (an example is shown in Fig. 1D). The majority of the CDS in this category are of unknown function or have phage-related functions. Among them, there are 2 groups of contiguous CDS (XF0512 to XF0523 and XF1932 to XF1937). The first set is within genomic island 1 (24). The other group includes genes that belong to different functional categories such as DNA metabolism and transport, suggesting them as possible newborn paralogs in the J1a12 genome.

Despite all the information obtained with DNA microarray genotyping, it is necessary to stress that frameshifts and point mutations cannot be identified by this method. In addition, our DNA microarray experiments will not detect genes present exclusively in the unsequenced strain J1a12, impairing detection of genes that would eventually attenuate pathogenicity.

Validation of CDS classification. To further validate the distinction between CDS potentially absent or highly divergent and those classified as divergent, a sample of 33 CDS (listed in Materials and Methods) were analyzed by PCR with CDS-specific primers. All CDS tested were PCR negative for strain J1a12, including those with log hybridization intensity ratios between -0.5 and -1.7 .

Figure 2A shows an example of a CDS (XF0077, encoding a fimbrial adhesin precursor) classified as absent or highly divergent by microarray analysis. PCR and RT-PCR experiments with CDS-specific primers corroborated its classification in this category. In addition, a PCR experiment with a pair of specific primers flanking XF0077 encompassing CDS XF0075 to XF0079 produced a smaller amplicon in strain J1a12 than in

FIG. 1. *Xylella* strain J1a12 CDS classification based on DNA microarray hybridization ratios. Examples of CDS classified in each of the four categories are shown. (A) XF1621, classified as equally present in both 9a5c and J1a12 strains, category i; (B) XF0262, classified as divergent in J1a12, category ii; (C) XF0496, classified as absent or highly divergent in strain J1a12, category iii; (D) XF1937, classified as higher copy number in J1a12, category iv. Orange dots represent the results for all CDS in the microarray from three homotypic control experiments (9a5c labeled with Cy5 versus 9a5c labeled with Cy3). Graphs show the fluorescence intensity ratio $M = \log_2(I_{Cy5}/I_{Cy3})$ versus the fluorescence intensity mean $A = 1/2 \times \log_2(I_{Cy5} \times I_{Cy3})$. Green dots represent the hybridization data (9a5c labeled with Cy3 versus J1a12 labeled with Cy5 or vice versa) from multiple replicas for the indicated CDS, where the fluorescence intensity ratio $M = \log_2(I_{J1a12}/I_{9a5c})$. Similar graphs for each of the CDS are available at the project site (<http://verjo19.iq.usp.br/cagexylella/private/>).

strain 9a5c (Fig. 2A). The exact position of the deleted region was determined by DNA sequence analysis, as described in the next section.

Data for XF1968, which encodes a putative methyltransferase classified as divergent by microarray analysis (mean log ratio of -1.4), is shown in Fig. 2B. The PCR and RT-PCR results with CDS-specific primers were negative, suggesting the absence of the CDS. However, PCR experiments with a pair of specific primers flanking XF1968 (encompassing XF1967 to XF1969) showed amplicons with the same size in both strains (Fig. 2B). Nucleotide sequencing of both amplicons confirmed that the methyltransferase encoded by XF1968 is divergent between 9a5c and J1a12 (Table 4).

Three other examples of CDS classified as divergent according to microarray experiments and found to be PCR negative are also depicted in Table 4. Sequence analysis of these CDS

cloned from strain J1a12 has confirmed their divergence, as they present nucleotide identities between 92 and 55% when compared to strain 9a5c (Table 4). These results show that PCR validation can be misleading, given that the primers used were based on the genome sequence of strain 9a5c. A few mismatches in the regions where the primers should anneal in the J1a12 DNA template can give PCR-negative results, leading to a wrong conclusion about the presence or absence of a gene.

RNA expression studies by microarray hybridization comparing strains J1a12 and 9a5c were performed as an additional validation of CDS classification. RNA hybridization data are available at the project site (<http://verjo19.iq.usp.br/cagexylella/private/>). Hybridization signals were found to be at the background level for the 14 CDS classified as absent or highly divergent in strain J1a12. On the other hand, all of these CDS showed significant expression levels in strain 9a5c.

TABLE 2. CDS divergent in strain J1a12

Gene ^a	Product
XF0075	Hypothetical protein
XF0078	Fimbrial adhesin precursor
XF0157	Hypothetical protein
XF0262	Colicin V precursor
XF0263	Colicin V precursor
XF0500	Phage-related repressor protein
XF0501	Conserved hypothetical protein
XF0626	Hypothetical protein
XF0663	Hypothetical protein
XF0665	Hypothetical protein
XF0666	Hypothetical protein
XF0668	Hemolysin-type calcium binding protein
XF0684	Phage-related protein
XF0696	Phage-related repressor protein
XF1057	Hypothetical protein
XF1306	Hypothetical protein
XF1588	Hypothetical protein
XF1589	Plasmid stabilization protein
XF1590	Plasmid stabilization protein
XF1609	Glucose/galactose transporter
XF1664	Hypothetical protein
XF1746	Alcohol dehydrogenase
XF1756	Hypothetical protein
XF1758	Hypothetical protein
XF1851	Serine protease
XF1859	Hypothetical protein
XF1862	Conserved hypothetical protein
XF1873	Conserved hypothetical protein
XF1877	Hypothetical protein
XF1883	Hypothetical protein
XF1968	Methyltransferase
XF2193	Hypothetical protein
XF2194	Hypothetical protein
XF2195	Hypothetical protein
XF2217	Imidazoleglycerolphosphate dehydratase/histidinol-phosphate phosphatase/bifunctional enzyme
XF2307	Hypothetical protein
XF2406	Hypothetical protein
XF2407	Bacteriocin
XF2542	Fimbrial protein
XF2722	Type I restriction-modification system specificity determinant
XF2726	Type I restriction-modification system specificity determinant
XF2768	Hypothetical protein
XF2770	Hypothetical protein
XF2772	Hypothetical protein

^a Only 44 CDS with log hybridization intensity ratios between -1.7 and -0.5 are shown in this table.

The expression levels of CDS that were classified as equally present in both strains (category i, see Materials and Methods) were studied under standard bacterial growth conditions. This class of CDS was chosen to eliminate possible hybridization artifacts due to sequence divergence. We found that among the 2,296 CDS in this category, which presented detectable hybridization intensity values, approximately 97% exhibited comparable RNA expression levels on both strains, i.e., differences in expression were smaller than twofold, with P values smaller than 0.05 in a t test. However, about 1% of the CDS presented a higher RNA expression level (twofold or more) in strain 9a5c and about 2% had higher expression in strain J1a12 (see Tables S3 and S4 in the supplemental material). No obvious

correlation could be made between the CDS presenting differential expression and the phenotypes of each strain.

Functional characteristics of genes absent or highly divergent in strain J1a12. Among the 14 CDS classified as absent or highly divergent (Table 3), 10 have no similarity to known genes and no function could be assigned. Therefore, they will not be further discussed, although their involvement in pathogenesis cannot be excluded.

The *X. fastidiosa* strain 9a5c genome encodes three fimbrial adhesin subunits from type III pilus (XF0077, XF0078, and XF0080). Our microarray data classified XF0077 as absent or highly divergent, and we have confirmed its deletion in strain J1a12 by DNA sequence analysis. The deleted region, encompassing 1,050 nucleotides (Fig. 2A), extends from position 76845 to 77895 (numbers from the strain 9a5c main chromosome). XF0078 was classified as divergent in J1a12, and DNA sequence analysis has shown a 92.2% nucleotide identity with its ortholog in 9a5c (Table 4). The third fimbrial adhesin paralog (XF0080) was classified as equally present in both strains (category i). The three paralogs from 9a5c are similar to *mrkD* from *Klebsiella pneumoniae* and share high similarity with each other (XF0077 and XF0078 share 70% amino acid sequence identity and both display around 60% amino acid identity to XF0080). As reported for the adhesins of *K. pneumoniae* (30), the N-terminal region exhibits a greater degree of variability, probably conferring on strain 9a5c the ability to adhere to different bacterial and/or host cell components or even producing an extracellular matrix with greater bonding capacity. In *K. pneumoniae*, *mrkD* null mutants are fimbriate but non-adhesive (34) and the *mrkD* gene product is not required for biofilm formation (18). Proteomic and mass spectrometric analyses of whole-cell lysates and extracellular components have demonstrated that structural and adhesive subunits of fimbriae are ubiquitous in cultures of *Xylella* strain 9a5c (33). Despite the presence of two CDS encoding the adhesion subunit precursors XF0078 and XF0080 in J1a12, we have observed that this strain displays a much less aggregated phenotype in vitro than 9a5c cells, as shown in Fig. 3. One possible explanation for this phenotype is that the presence of the

TABLE 3. CDS absent or highly divergent in strain J1a12

Gene ^a	Product	M^b
XF0077	Fimbrial adhesion precursor	-2.77
XF0496	Conserved hypothetical protein	-2.31
XF0497	Conserved hypothetical protein	-2.08
XF0667	Hypothetical protein	-1.99
XF1250	Arginine deaminase	-1.96
XF1646	UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase	-2.01
XF1707	Hypothetical protein	-2.30
XF1708	Conserved hypothetical protein	-2.22
XF1860	Hypothetical protein	-2.79
XF1874	Hypothetical protein	-2.13
XF1878	Hypothetical protein	-3.11
XF1884	Hypothetical protein	-2.21
XFb0001	Replication protein	-3.55
XFb0002	Hypothetical protein	-4.18

^a CDS shown here have a P value smaller than 0.05 in a t test for the null hypothesis $H_0: M \geq -1.7$.

^b DNA hybridization intensity ratios [$M = \log_2(I_{J1a12}/I_{9a5c})$] were calculated. For each gene, the values shown are the means of the results from at least eight replicates.

FIG. 2. Validation of CDS classification. CDS-specific primers for XF0077 (A) and XF1968 (B) were employed to perform PCR amplifications with DNA from strains 9a5c and J1a12 (left panels) or to perform RT-PCR amplifications with total RNA (central panels). PCRs were also carried out with DNA of both strains and primers based on the sequence of the CDS flanking XF0077 (amplicon XF0075 to XF0079) or XF1968 (amplicon XF1967 to XF1969) (right panels). The sizes of the amplicons are shown in base pairs.

adhesin encoded by XF0077 is important for the adhesion of *X. fastidiosa* cells.

Interestingly, another CDS confirmed to be deleted from strain J1a12 is XF1250, which encodes a putative arginase. The deleted region of 963 nucleotides extends from position 1203200 to 1204163 (numbers from the strain 9a5c main chromosome). This enzyme catalyzes the first step of arginine degradation in the urea cycle, which may thus be incomplete in strain J1a12. Besides being a substrate for arginases, arginine is also a substrate for nitric oxide (NO) synthase, which converts L-arginine to L-citrulline, releasing NO. In *H. pylori*, it has recently been shown that the arginase (*rocF*) encoded by this bacterium inhibits NO production by macrophages at physiologic concentrations of L-arginine. *H. pylori rocF* mutants cocultured with macrophages were killed due to the restoration of normal levels of NO. These results indicate that bacterial arginase down-regulates NO production, acting as a survival mechanism that contributes to the successful infection of the host (10). Thus, it is tempting to speculate that the absence of arginase in *X. fastidiosa* strain J1a12 is linked to its reduced growth in planta and its incapacity to colonize the xylem vessels; J1a12 cells would not be able to inhibit NO production by the plant, analogous to the *rocF* mutant of *H. pylori* (10). In fact, it is known that in plants, NO has an immune protective role, mediating the plant defense against pathogens and serving as a signal in hormonal responses. Indeed, two NO-synthesizing enzymes have recently been found in plants, one of which is pathogen inducible (5, 11).

XF1646, classified as absent or highly divergent in J1a12, was annotated as a putative UDP-3-*O*-(*R*-3-hydroxymyristoyl)-glucosamine *N*-acyltransferase, which is similar to the *lpxD* gene from *Rickettsia rickettsii*. This enzyme catalyzes the third step in lipid A biosynthesis, a constituent of lipopolysaccharides from the outer membrane. In *Escherichia coli*, an *lpxD* mutant had its susceptibility to various antibiotics increased as high as 512-fold, indicating alterations in the outer membrane permeability barrier (36). An altered outer membrane structure due to the incomplete lipid A biosynthesis may render J1a12 more

TABLE 4. Nucleotide and amino acid identity of selected divergent CDS of strain J1a12 compared to strain 9a5c

Gene	Product	% Nucleotide identity ^a	% Amino acid identity ^a
XF0078	Fimbrial adhesin precursor	92.2	86.4
XF1968	Methyltransferase	78.8	75.5
XF2542	Fimbrial protein	73.6	65.8
XF2726	Type I restriction-modification system specificity determinant	54.9	40.4

^a Comparison was carried out with the complete sequence of each CDS.

sensitive to antimicrobial agents eventually produced by the plant host. The increased permeability of its outer membrane may also explain why J1a12 is amenable to DNA transformation, whereas 9a5c is not (6, 17, 23).

The two highest DNA hybridization intensity ratios obtained when comparing 9a5c versus J1a12 were derived from the two CDS encoded by the mini plasmid pXF1.3 (XFb0001 and XFb0002). This result reflects the presence of multiple copies of this plasmid in strain 9a5c (32) and its absence in strain J1a12, as previously reported (6).

Final remarks. Among the 14 genes classified as absent or highly divergent, three CDS encoding a fimbrial adhesin precursor, an arginase, and a UDP-3-*O*-(*R*-3-hydroxymyristoyl)-glucosamine *N*-acyltransferase are conspicuously absent in the nonpathogenic strain J1a12. Due to their putative role in bacterial survival in infected hosts, they emerge here as important players in *Xylella* pathogenicity. The observation that several other genes are missing in J1a12 gives support to the hypothesis that bacterial pathogenesis is a multifactorial process and that each of these factors may contribute somewhat quantitatively to the development of disease. In fact, inactivation of a single fimbrial adhesin gene (PD0058) or a single fimbriae protein gene (PD0062) in *X. fastidiosa* grapevine strain Temecula was not sufficient to decrease bacterial pathogenicity, causing only a slight reduction in the bacterial population (9).

Recent microarray expression studies comparing *X. fastidiosa* 9a5c cells freshly isolated from citrus with bacteria attenuated after several passages in axenic culture have shown that most genes found to be induced in the freshly isolated condition were associated with adhesion and with possible adaptation to the host environment (8). However, the set of genes observed in that study is different from the genes found to be absent in the nonpathogenic strain analyzed in the present report, reinforcing the multifactorial hypothesis of bacterial pathogenesis. Furthermore, our results obtained with J1a12 are independent of the number of passages in culture, since the microarray experiments were performed with DNA from bacterial cells obtained after either 14 or 24 passages and no differences were observed (data not shown).

Our plant colonization assays showing that strain J1a12 is unable to induce CVC symptoms or even sustain itself in host plants raise the question of how strain J1a12 was originally isolated from a citrus tree. A recent report about the diversity of the endophytic bacterial community in citrus trees (2) can shed light upon this intriguing question. Possibly, the presence of strain J1a12 in symptomatic plants is dependent on other microorganisms and/or other *X. fastidiosa* pathogenic strains

FIG. 3. *X. fastidiosa* strain J1a12 has lost the ability to aggregate. Optical microscopy of *X. fastidiosa* strains 9a5c (A) and J1a12 (B). Magnification, $\times 1,000$.

eventually present in the biofilms formed by the aggregated cells and clogging the xylem vessels of infected plant hosts (13, 19).

Among the CDS found to be absent in J1a12 and proposed here to play a role in disease development, four CDS do have orthologs in the three other pathogenic *Xylella* strains that have been sequenced (see Table S5 in the supplemental material). Thus, despite the diversity of hosts, geographical location, and disease symptoms, different *Xylella* strains may present similar mechanisms of pathogenesis. Our results suggest that common strategies could be undertaken to control the diseases that are caused by different *X. fastidiosa* strains infecting various host plants. Given the importance of the set of genes found here, further functional characterization is warranted. Towards this end, we are currently trying to complement strain J1a12 with the CDS shown to be absent from its genome. Different from 9a5c, the nonpathogenic J1a12 strain is amenable to DNA transformation with plasmid vectors and the transposome system (6, 17, 23). Thus, we believe that with complementation studies it will be possible to evaluate the role of these CDS in bacterial virulence and the ability to colonize xylem vessels.

ACKNOWLEDGMENTS

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

We are greatly indebted to Hugo A. Armelin for coordinating the Cooperation for Analysis of Gene Expression (CAGE) Project and for strongly supporting this work. We thank João Carlos Setubal and João Kitajima for providing information about reannotation of the *Xylella* genomic sequence and Apuã C. M. Paquola, Milton Y. Nishiyama, Jr., and Abimael A. Machado for providing bioinformatic tools. We thank Jesus Ferro for coordinating our PCR amplification data bank and for the *Xylella* cosmid library. We also thank Sanvai R. P. Rocha, Mateus de Almeida Santos, and Anelise G. Mariano for help in pathogenicity tests and Helder Nakaya and Ari J. S. Ferreira for valuable help in the beginning of this work.

A.M.d.S., H.E.-D., S.L.G., and S.V.-A. were partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, (CNPq). T.K., L.M.M., R.Z.N.V., and P.A.Z. are FAPESP doctoral fellows.

REFERENCES

- Almeida, R. P. P., and A. H. Purcell. 2003. Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera: Cicadellidae). *J. Econ. Entomol.* **96**:264–271.
- Araujo, W. L., J. Marcon, W. Maccheroni, Jr., J. D. Van Elsland, J. W. Van Vuurde, and J. L. Azevedo. 2002. Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants. *Appl. Environ. Microbiol.* **68**:4906–4914.
- Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane, and P. M. Small. 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284**:1520–1523.
- Bhattacharyya, A., S. Stilwagen, G. Reznik, H. Feil, W. S. Feil, I. Anderson, A. Bernal, M. D'Souza, N. Ivanova, V. Kapatral, N. Larsen, T. Los, A. Lykidis, E. Selkov, Jr., T. L. Walunas, A. Purcell, R. A. Edwards, T. Hawkins, R. Haselkorn, R. Overbeek, N. C. Kyrpides, and P. F. Predki. 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res.* **12**:1556–1563.
- Chandok, M. R., A. J. Ytterberg, K. J. van Wijk, and D. F. Klessig. 2003. The pathogen-inducible nitric oxide synthase (iNOS) in plants is a variant of the P protein of the glycine decarboxylase complex. *Cell* **113**:469–482.
- da Silva Neto, J. F., T. Koide, S. L. Gomes, and M. V. Marques. 2002. Site-directed gene disruption in *Xylella fastidiosa*. *FEMS Microbiol. Lett.* **210**:105–110.
- Davis, M. J., W. J. French, and N. W. Schaad. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. *Curr. Microbiol.* **6**:309–314.
- de Souza, A. A., M. A. Takita, H. D. Coletta, C. Caldana, G. H. Goldman, G. M. Yanai, N. H. Muto, R. C. de Oliveira, L. R. Nunes, and M. A. Machado. 2003. Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity. *Mol. Plant-Microbe Interact.* **16**:867–875.
- Feil, H., W. S. Feil, J. C. Detter, A. H. Purcell, and S. E. Lindow. 2003. Site-directed disruption of the *fimA* and *fimF* fimbrial genes of *Xylella fastidiosa*. *Phytopathology* **93**:675–682.
- Gobert, A. P., D. J. McGee, M. Akhtar, G. L. Mendz, J. C. Newton, Y. Cheng, H. L. Mobley, and K. T. Wilson. 2001. *Helicobacter pylori* arginase inhibits nitric oxide production by eukaryotic cells: a strategy for bacterial survival. *Proc. Natl. Acad. Sci. USA* **98**:13844–13849.
- Guo, F. Q., M. Okamoto, and N. M. Crawford. 2003. Identification of a plant nitric oxide synthase gene involved in hormonal signaling. *Science* **302**:100–103.
- Hinchliffe, S. J., K. E. Isherwood, R. A. Stabler, M. B. Prentice, A. Rakin, R. A. Nichols, P. C. F. Oyston, J. Hinds, R. W. Titball, and B. W. Wren. 2003. Application of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and *Yersinia pseudotuberculosis*. *Genome Res.* **13**:2018–2029.
- Hopkins, D. L. 1989. *Xylella fastidiosa*-xylem-limited bacterial pathogen of plants. *Annu. Rev. Phytopathol.* **27**:271–290.
- Hopkins, D. L., and A. H. Purcell. 2002. *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. *Plant Dis.* **86**:1056–1066.
- Karlin, S. 2001. Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol.* **9**:335–343.
- Kim, C. C., E. A. Joyce, K. Chan, and S. Falkow. 2002. Improved analytical methods for microarray-based genome-composition analysis. *Genome Biol.* **3**:RESEARCH0065.
- Koide, T., J. F. da Silva Neto, S. L. Gomes, and M. V. Marques. 2004. Insertional transposon mutagenesis in the *Xylella fastidiosa* citrus variegated chlorosis strain with transposome. *Curr. Microbiol.* **48**:247–250.
- Langstraat, J., M. Bohse, and S. Clegg. 2001. Type 3 fimbrial shaft (MrkA) of *Klebsiella pneumoniae*, but not the fimbrial adhesin (MrkD), facilitates biofilm formation. *Infect. Immun.* **69**:5805–5812.
- Leite, B., M. L. Ishida, E. Alves, H. Carrer, S. F. Pascholati, and E. W.

- Kitajima. 2002. Genomics and X-ray microanalysis indicate that Ca^{2+} and thiols mediate the aggregation and adhesion of *Xylella fastidiosa*. *Braz. J. Med. Biol. Res.* **35**:645–650.
20. Li, W. B., L. Zreik, N. G. Fernandes, V. S. Miranda, D. C. Teixeira, A. J. Ayres, M. Garnier, and J. M. Bove. 1999. A triply cloned strain of *Xylella fastidiosa* multiplies and induces symptoms of citrus variegated chlorosis in sweet orange. *Curr. Microbiol.* **39**:106–108.
 21. Lopes, S. A., D. M. Ribeiro, P. G. Roberto, S. C. Franca, and J. M. Santos. 2000. *Nicotiana tabacum* as an experimental host for the study of plant-*Xylella fastidiosa* interactions. *Plant Dis.* **84**:827–830.
 22. Machado, M. A., A. A. Souza, H. D. Coletta-Filho, E. E. Kuramae, and M. A. Takita. 2001. Genome and pathogenicity of *Xylella fastidiosa*. *Mol. Biol. Today* 2:33–43.
 23. Monteiro, P. B., D. C. Teixeira, R. R. Palma, M. Garnier, J. M. Bove, and J. Renaudin. 2001. Stable transformation of the *Xylella fastidiosa* citrus variegated chlorosis strain with *oriC* plasmids. *Appl. Environ. Microbiol.* **67**:2263–2269.
 24. Nunes, L. R., Y. B. Rosato, N. H. Muto, G. M. Yanai, V. S. da Silva, D. B. Leite, E. R. Gonçalves, A. A. de Souza, H. D. Coletta-Filho, M. A. Machado, S. A. Lopes, and R. C. de Oliveira. 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res.* **13**:570–578.
 25. Pooler, M. R., and J. S. Hartung. 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. *Curr. Microbiol.* **31**:377–381.
 26. Purcell, A. H., and D. L. Hopkins. 1996. Fastidious xylem-limited bacterial plant pathogens. *Annu. Rev. Phytopathol.* **34**:131–151.
 27. Purcell, A. H., and S. R. Saunders. 1999. Fate of Pierce's disease strains of *Xylella fastidiosa* in common riparian plants in California. *Plant Dis.* **83**:825–830.
 28. Read, T. D., S. N. Peterson, N. Tourasse, L. W. Baillie, I. T. Paulsen, K. E. Nelson, H. Tettelin, D. E. Fouts, J. A. Eisen, S. R. Gill, E. K. Holtzapple, O. A. Okstad, E. Helgason, J. Rilstone, M. Wu, J. F. Kolonay, M. J. Beanan, R. J. Dodson, L. M. Brinkac, M. Gwinn, R. T. DeBoy, R. Madpu, S. C. Daugherty, A. S. Durkin, D. H. Haft, W. C. Nelson, J. D. Peterson, M. Pop, H. M. Khouri, D. Radune, J. L. Benton, Y. Mahamoud, L. X. Jiang, I. R. Hance, J. F. Weidman, K. J. Berry, R. D. Plaut, A. M. Wolf, K. L. Watkins, W. C. Nierman, A. Hazen, R. Cline, C. Redmond, J. E. Thwaite, O. White, S. L. Salzberg, B. Thomason, A. M. Friedlander, T. M. Koehler, P. C. Hanna, A. B. Kolsto, and C. M. Fraser. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**:81–86.
 29. Salama, N., K. Guillemin, T. K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow. 2000. A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc. Natl. Acad. Sci. USA* **97**:14668–14673.
 30. Sebgathi, T. A. S., T. K. Korhonen, D. B. Hornick, and S. Clegg. 1998. Characterization of the type 3 fimbrial adhesins of *Klebsiella* strains. *Infect. Immun.* **66**:2887–2894.
 31. Silverman, B. W. 1986. Density estimation. Chapman and Hall, London, England.
 32. Simpson, A. J., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bove, M. R. Briones, M. R. Bueno, A. A. Camargo, L. E. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. Costa, C. M. Costa-Neto, L. L. Coutinho, M. Cristofani, E. Dias-Neto, C. Docena, H. El-Dorry, A. P. Facincani, A. J. Ferreira, V. C. Ferreira, J. A. Ferro, J. S. Fraga, S. C. Franca, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. H. Goldman, S. L. Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. Leite, E. G. Lemos, M. V. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, M. A. Machado, A. M. Madeira, H. M. Madeira, C. L. Marino, M. V. Marques, E. A. Martins, E. M. Martins, A. Y. Matsukuma, C. F. Menck, E. C. Miracca, C. Y. Miyaki, C. B. Monteiro-Vitorello, D. H. Moon, M. A. Nagai, A. L. Nascimento, L. E. Netto, A. Nhani, Jr., F. G. Nobrega, L. R. Nunes, M. A. Oliveira, M. C. de Oliveira, R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. Pereira, H. A. Pereira, Jr., J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, A. J. de M. Rosa, V. E. de Rosa, Jr., R. G. de Sa, R. V. Santelli, H. E. Sawasaki, A. C. da Silva, A. M. da Silva, F. R. da Silva, W. A. da Silva, Jr., J. F. da Silveira, M. L. Silvestri, W. J. Siqueira, A. A. de Souza, A. P. de Souza, M. F. Terezi, D. Truffi, S. M. Tsai, M. H. Tshako, H. Vallada, M. A. Van Sluys, S. Verjovski-Almeida, A. L. Vettore, M. A. Zago, M. Zatz, J. Meidanis, and J. C. Setubal. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature* **406**:151–157.
 33. Smolka, M. B., D. Martins, F. V. Winck, C. E. Santoro, R. R. Castellari, F. Ferrari, I. J. Brum, E. Galembeck, H. Della Coletta Filho, M. A. Machado, S. Marangoni, and J. C. Novello. 2003. Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution. *Proteomics* **3**:224–237.
 34. Tarkkanen, A. M., R. Virkola, S. Clegg, and T. K. Korhonen. 1997. Binding of the type 3 fimbriae of *Klebsiella pneumoniae* to human endothelial and urinary bladder cells. *Infect. Immun.* **65**:1546–1549.
 35. Tyson, G. E., B. J. Stojanovic, R. F. Kuklinski, T. J. Divittorio, and M. L. Sullivan. 1985. Scanning electron-microscopy of Pierce's disease bacterium in petiolar xylem of grape leaves. *Phytopathology* **75**:264–269.
 36. Vaara, M., and M. Nurminen. 1999. Outer membrane permeability barrier in *Escherichia coli* mutants that are defective in the late acyltransferases of lipid A biosynthesis. *Antimicrob. Agents Chemother.* **43**:1459–1462.
 37. Van Sluys, M. A., M. C. de Oliveira, C. B. Monteiro-Vitorello, C. Y. Miyaki, L. R. Furlan, L. E. Camargo, A. C. da Silva, D. H. Moon, M. A. Takita, E. G. Lemos, M. A. Machado, M. I. Ferro, F. R. da Silva, M. H. Goldman, G. H. Goldman, M. V. Lemos, H. El-Dorry, S. M. Tsai, H. Carrer, D. M. Carraro, R. C. de Oliveira, L. R. Nunes, W. J. Siqueira, L. L. Coutinho, E. T. Kimura, E. S. Ferro, R. Harakava, E. E. Kuramae, C. L. Marino, E. Gigliotti, I. L. Abreu, L. M. Alves, A. M. do Amaral, G. S. Baia, S. R. Blanco, M. S. Brito, F. S. Cannavan, A. V. Celestino, A. F. da Cunha, R. C. Fenille, J. A. Ferro, E. F. Formighieri, L. T. Kishi, S. G. Leoni, A. R. Oliveira, V. E. Rosa, Jr., F. T. Sasaki, J. A. Sena, A. A. de Souza, D. Truffi, F. Tsukumo, G. M. Yanai, L. G. Zaro, E. L. Civerolo, A. J. Simpson, N. F. Almeida, Jr., J. C. Setubal, and J. P. Kitajima. 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. *J. Bacteriol.* **185**:1018–1026.
 38. Wolfgang, M. C., B. R. Kulasekara, X. Y. Liang, D. Boyd, K. Wu, Q. Yang, C. G. Miyada, and S. Lory. 2003. Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **100**:8484–8489.
 39. Yang, Y. H., S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**:e15.

ANEXO 3

**TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA**

ARTIGO PUBLICADO

Comparative genomics analyses of citrus-associated bacteria.

Moreira LM, de Souza RF, Almeida NF Jr, Setubal JC, Oliveira JC,
Furlan LR, Ferro JA, da Silva AC.

Annu Rev Phytopathol. 2004; 42:163-84.

PMID: 15283664 - PubMed - indexed for MEDLINE.

COMPARATIVE GENOMICS ANALYSES OF CITRUS-ASSOCIATED BACTERIA

Leandro M. Moreira,¹ Robson F. de Souza,¹
Nalvo F. Almeida Jr,² João C. Setubal,^{3*}
Julio Cezar F. Oliveira,⁵ Luiz R. Furlan,⁴
Jesus A. Ferro,^{5,6} and Ana C.R. da Silva^{1,6}

¹Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil; email: lmorei@iq.usp.br, rfsouza@citri.iq.usp.br, ana.rasera@alellyx.com.br; ²Departamento de Computação e Estatística, Universidade Federal do Mato Grosso do Sul, Campo Grande, MS, Brazil, email: nalvojr@det.ufms.br; ³Instituto de Computação, Universidade Estadual de Campinas, Campinas, Brazil, email: setubal@ic.unicamp.br; ⁴Departamento de Melhoramento e Nutrição Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista, Fazenda Lageado, Botucatu, SP, Brazil, email: cedral@fca.unesp.br; ⁵Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP, Brazil, email: jesus@fcav.unesp.br; jcyler@fcav.unesp.br; ⁶Alellyx Applied Genomics, Techno Park 13067–850, Campinas, SP, Brazil.

Key Words *Xylella fastidiosa*, *Xanthomonas axonopodis* pv. *citri*, comparative genomic, phytopathogens, CVC and citrus canker

■ **Abstract** *Xylella fastidiosa* 9a5c (XF-9a5c) and *Xanthomonas axonopodis* pv. *citri* (XAC) are bacteria that infect citrus plants. Sequencing of the genomes of these strains is complete and comparative analyses are now under way with the genomes of other bacteria of the same genera. In this review, we present an overview of this comparative genomic work. We also present a detailed genomic comparison between XF-9a5c and XAC. Based on this analysis, genes and operons were identified that might be relevant for adaptation to citrus. XAC has two copies of a type II secretion system, a large number of cell wall-degrading enzymes and sugar transporters, a complete energy metabolism, a whole set of avirulence genes associated with a type III secretion system, and a complete flagellar and chemotactic system. By contrast, XF-9a5c possesses more genes involved with type IV pili biosynthesis than does XAC, contains genes encoding for production of colicins, and has 4 copies of Type I restriction/modification system while XAC has only one.

*Current address: Virginia Bioinformatics Institute and Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24060-0477; email: setubal@vbi.vt.edu

INTRODUCTION

Xanthomonas axonopodis pv. *citri* (XAC) and *Xylella fastidiosa* strain 9a5c (XF-9a5c) are two gram-negative proteobacteria of the family *Xanthomonadaceae* that infect the same host, sweet orange (*Citrus sinensis*). Both cause great losses to fruit production and its subsidiaries and are therefore of major concern to the citrus industry. XAC, the causal agent of citrus canker, causes the formation of cankers associated with water-soaked lesions on leaves. These cankers are surrounded by chlorotic haloes and surface-penetrating necrotic lesions on fruit. Characteristics of the disease include early fruit abscission and general tree decline (23). The bacterium can survive and multiply outside the host as an epiphyte, but once it contacts stomates, hydathodes, water pores, or lesions of plant tissues, it can gain access to and colonize the mesophyll, and cause the canker symptoms. XF-9a5c, the causal agent of citrus variegated chlorosis (CVC), induces the formation of chlorotic areas on the upper side of leaves; the infections result in small-sized fruits with a hard consistency (10, 32, 50). CVC is transmitted by leafhoppers (34) and infects the xylem, causing vessel occlusion by cellular agglutination, a process that seems to be critical for the development of symptoms.

The genome sequences of these two phytopathogens have been published (14, 54). Subsequent comparative genomic analyses using the XF-9a5c sequences have also been published. Van Sluys and collaborators (61) compared eight fully sequenced genomes of plant-associated bacteria, including XF-9a5c and XAC. Bhattacharyya and collaborators (6, 7) compared three different strains of *Xylella fastidiosa* (XF-9a5c and two others that infect almond and oleander, respectively). Van Sluys and collaborators (60) published the sequence of the grapevine strain of *X. fastidiosa* and compared it to XF-9a5c. In this review, we present an overview of the comparative work within the genera *Xylella* and *Xanthomonas*. We also present a detailed genomic comparison between XF-9a5c and XAC. Although these two pathogens belong to the same family and infect the same host, they are adapted to different environments within the host. We identified differences that we propose may be responsible for the different processes of pathogenesis and adaptation to their respective environments.

The genome of XAC is considerably more complex than that of XF-9a5c. The overall size of the XAC genome is larger than XF-9a5c (XAC: 5,175,554 bp; XF-9a5c: 2,679,305 bp) and the gene content of the XAC genome is more diverse than that of XF-9a5c. Based on comparative analysis, several systems might have contributed to the particular adaptation of XAC. It has two divergent copies of a type II secretion system, a larger number of cell wall-degrading enzymes and sugar transporters (relative to XF-9a5c), a complete energy metabolism with a rich machinery of oxide reduction, a large set of putative genes related to avirulence effector proteins that are associated with a type III secretion system, and a complete flagellar and chemotactic system. By contrast, XF-9a5c possesses more genes involved with type IV pili biosynthesis, production of colicins, and a restriction/modification system that may reflect adaptation to permit survival in the plant

xylem and the insect foregut, and is associated with different microbe communities within both of these contexts.

GENOMIC COMPARISON BETWEEN *Xylella fastidiosa* 9a5c (XF-9a5c) AND OTHER STRAINS OF *Xylella fastidiosa*

Previous studies using techniques as distinct as 16S-23S rRNA analysis, DNA homology, structural protein analysis, DNA-DNA hybridization, RFLPs, and RAPD-PCR have shown a certain degree of divergence between the genomes from different strains of *X. fastidiosa* (3, 5, 11, 13, 25, 28, 37, 44). However, these studies could not determine the degree of divergence with precision because their measurements were based only on the overall number of polymorphisms or single mutations, which do not always reflect modification at the protein level. Since publication of the complete genome sequence of XF-9a5c in July 2000 (54), other strains have been sequenced, opening the way for comparative analysis of these closely related genomes. These studies should help in identifying divergences that may explain the adaptation of *X. fastidiosa* to both its insect and plant host, and to pathogenicity in the plant host. This information may identify targets that can be used for disease control.

Bhattacharyya and collaborators (6, 7) compared the XF-9a5c genome sequence with the incomplete genome sequences of *X. fastidiosa* pv. *almond* (XF-Dixon) and *X. fastidiosa* pv. *oleander* (XF-Ann-1). The latter two bacteria are the causal agents of almond and oleander leaf scorch, respectively. These studies identified 1705 conserved protein clusters which exhibit a similarity averaging at least 82%. These authors speculated that a putative phage insertion, located next to the hypothetical replication terminus, may account for strain specificity to citrus (7). The XF-Dixon genome contains 133 ORFs that are specific to it with respect to the other two. Of these, 78 are hypothetical and 55 were found to have similarity outside the genus *Xylella*. In this last group, there are genes involved in conjugal transfer and type II restriction and modification systems. In the XF-Ann-1 genome, there are 188 specific ORFs, most of which are involved with phage-related proteins or plasmid maintenance. In XF-9a5c, there are 389 specific ORFs, some of which are involved in transcription regulation, transporters, and antibiotic resistance (6).

In 2003, Van Sluys and collaborators (60) completed sequencing the genome of the PD strain of *X. fastidiosa* (XF-PD), the causal agent of Pierce's disease (PD) in grapevine, and compared it with XF-9a5c. XF-PD contains a small (1.3-Kb) plasmid that was also reported in XF-9a5c, and its chromosome is 150 kb smaller than that of XF-9a5c. Approximately 98% of all genes present in XF-PD are found in XF-9a5c, with a high average identity (98%), which suggests similar metabolic functions during plant and insect colonization and pathogenesis. Among the genes that exhibit a higher degree of divergence are those that may be related to plant and insect host interactions such as fimbriellins, hemagglutinins, colicins, hemolysins, toxins, drug resistance, and DNA restriction modification enzymes.

A total of 29 genes in XF-PD and 16 genes in XF-CVC are pseudogenes (with frameshift or stop codon in frame) that represent approximately 1.4% and 0.71% of all genes in their respective genomes. Surprisingly, one of the Xf-9a5c pseudogenes encodes a polygalacturonase precursor that could facilitate migration between vessels; the ortholog in XF-PD is intact. This difference could have several implications if the polygalacturonases serves as a pathogenicity factor: it could suggest that XF-PD might be more aggressive than XF-9a5c, or it could reflect differences in the xylem vessel structure between citrus and grapevine.

XF-PD has 51 strain-specific genes (2.47%) and XF-9a5c has 152 (6.78%), in both cases approximately half of these are hypothetical genes, and a significant proportion are associated with mobile genetic elements. Three major genomic rearrangements have been noted, most with a putative phage-integrase at one border, suggesting that they might be phage mediated. Interestingly, if the prophage regions are excluded from the sequence and rearrangements are reoriented, both genomes are very similar and colinear. This indicates that the structural divergence between the two genomes might have been mediated by lateral gene transfer, mostly by phage vectors. In addition to these rearrangements, each genome has a specific genomic island, named giCVC for XF-9a5c and giPD for XF-PD. The giCVC island was described by Bhattacharya as a citrus-specific island (7). However, van Sluys and collaborators have shown that it is also present in different strains of *Xylella* from South America (60).

GENOMIC COMPARISON BETWEEN *Xanthomonas axonopodis* pv. *citri* (XAC) AND *Xanthomonas campestris* pv. *campestris* (XCC)

The complete sequencing and comparison of the genomes of *Xanthomonas axonopodis* pv. *citri* (XAC) and *Xanthomonas campestris* pv. *campestris* (XCC) have shown a high degree of identity between the two genomes (14). However, there are groups of strain-specific genes that may be associated with each species distinct pathology and specificity to different hosts (Table 1). Both genomes are rich in transposable elements but these elements appear not to be responsible for structural differences since the two genomes are almost colinear. There are 109 genes corresponding to mobile elements in XCC and 108 in XAC, and these are grouped into 27 different types of elements. Both genomes contain an extensive repertoire of cell wall-degrading enzymes divided among cellulolytic, pectinolytic, and hemicellulolytic activities. The greater number of genes involved with these functions in XCC may account for its potential to infect plants systemically. On the other hand, XAC accomplishes disease by a high level of local cell proliferation and subsequent necrosis, without moving systemically in the plant. One possible explanation for the difference in pathogenicity styles is the fact that the number of genes capable of causing a massive degeneration of host tissue, the genes encoding CWDE, is smaller in XAC than in XCC. Other important differences found between these two genomes are related to putative avirulence/effector-protein coding genes,

TABLE 1 General comparison between *Xac* and *Xcc*

Features that are similar and present in both genomes	Complete intermediary, DNA, and small molecule metabolism	
	Xanthomonadin production	
	Flagella structural genes	
	Type IV fimbriae genes	
	Adhesin-related genes	
	Exopolysaccharide operon	
Features that are different but are present in both genomes	Four types of secretory systems	
	Mobile genetic elements	
	<i>Avr</i> genes	
	O-antigen biosynthesis genes	
	Number of proteases and polysaccharide-degrading enzymes	
	Leucin-rich proteins	
Features that are present only in <i>Xcc</i> or <i>Xac</i>	<i>HpaA</i> , <i>hpaI</i> , <i>hrpF</i> , <i>hpaP</i> and <i>hrpW</i> genes	
	<i>Xcc</i>	<i>Xac</i>
	Nitrate assimilation	Absent
	Tannase and 1,6 beta glucan synthetase	Absent
	<i>RpfI</i> and <i>rpfH</i>	Absent
	<i>PilE</i> ortholog in type IV fimbriae	Absent
	Macrolide/streptomycin-related operon	Absent
	Absent	Transport of peptides as source of nitrogen
	Absent	Presence of RTX-like toxins
	Absent	Syringomicin biosynthesis genes
	Absent	<i>FimA</i> ortholog in type IV fimbriae
	Absent	Extra copy of <i>Vir</i> operon
	Absent	Two plasmids and <i>pthA</i> genes
	Absent	Type II restriction system

which are crucial to the development of the disease, and O-antigen biosynthesis, which may be related to host range (Table 1). XCC presents a greater diversity of effectors, even though it lacks members of the *avrBs3/pthA* gene family that are found in four copies in the XAC genome, and which are associated with plant cell proliferation in citrus canker (17).

GENOMIC COMPARISON BETWEEN *Xanthomonas axonopodis* pv. *citri* (XAC) AND *Xylella fastidiosa* 9a5c (XF-9a5c)

We compared the predicted protein sets of XAC and Xf-9a5c by running the BLASTP program comparing every predicted protein sequence in one genome against all predicted protein sequences in the other genome and vice versa.

We declare that a given predicted protein is present in both genomes when one is the first hit for the other and vice versa, the e-value in both cases is at most 10^{-5} , and the alignments include at least 60% of each sequence. Using this methodology we found that XAC and Xf-9a5c share 1495 proteins. This number represents almost 74% of the XF-9a5c genes and 38% of the XAC genes. The discrepancy in percentages reflects the difference in genome size. Similar numbers are obtained if we compare XF-9a5c with XCC or XF-PD with XAC (39). A group of 16 genes (Table 2), most of them hypothetical, are present in both XF-9a5c and XAC, but not in any other strains of *Xanthomonas* or *Xylella* sequenced to date (XCC, XF-PD, XF-Dixon or XF-Ann-1). In XAC, this group is located inside the 133-Kbp region, which has been described as specific for XAC (14). Interestingly, the 133-Kbp region has synteny with the SP1-7 island in *Salmonella enterica* (42), which also has synteny with the genome of several soil bacteria such as *P. fluorescens*, *B. fulgorum*, or *R. metallidurans* and is assumed to be a mobile element. In XF-9a5c, the strain-specific group is located inside the giCVC island described by van Sluys and collaborators as being specific to South American *Xylella*, regardless of its host (citrus, coffee, periwinkle, or hibiscus) (60).

Although the two bacteria have many genes in common, a series of XAC genes are absent or reduced in numbers in XF-9a5c. Among these overrepresented genes in XAC are 12 groups of genes related to virulence, cellular growth, and adaptation. There are genes involved in the following: (a) the type II secretion system, (b) cell wall-degrading enzymes, (c) sugar transporters, (d) energy and general metabolism components, (e) oxide-reduction enzymes, (f) iron acquisition, (g) type III secretion system components, (h) flagella and the chemotactic system, (i) regulators of pathogenicity factors genes, (j) xanthomonadin biosynthesis, (k) gum genes, and (l) transcriptional factors (39). Despite its much smaller size, XF-9a5c has genes that are either not found in XAC or are present in more copies in XF-9a5c. These genes are related to type IV pili biosynthesis, colicin production, and particular restriction/modification systems (39). We suggest that the differences between XAC and XF-9a5c may reflect their different styles and location of pathogenicity and the fact that Xf also colonizes insects. Figure 1 compares the gene content and lifestyle of XAC and XF-9a5c.

Genes Overrepresented in *Xanthomonas axonopodis* pv. *citri*

TYPE II SECRETION SYSTEM (T2SS) Gram-negative bacteria require a protein complex that can break through the physical obstacle of the cellular membranes in order to discharge plant cell wall-degrading enzymes into the environment. This is called the type II secretion system (T2SS), or general secretory pathway (GSP) (52). A cluster of 12 to 14 genes with a high degree of conservation among different organisms codes for this system. XAC has two clusters of genes that putatively encode this system: XAC3534-3544 and XAC0694-0705 (14) whereas XF-9a5c has only one cluster: XF-1517-1527 (54). Although it is not known if both are functional, the presence of two clusters in XAC might be related to the larger number of cell

TABLE 2 Pairs of XF-9a5c/XAC-specific genes^a

XAC gene ID	Start position	End position	XF-9a5c gene ID	Start position	End position	Gene product
XAC2205	2589,160	2590,005	XF1785	17,02552	17,03427	Chromosome partitioning related protein
XAC2206	2590,213	2591,895	XF1784	17,01566	17,02327	Hypothetical protein
XAC2207	2591,892	2592,452	XF1782	17,00095	17,00655	Hypothetical protein
XAC2208	2592,457	2593,680	XF1781	16,98842	17,00098	Hypothetical protein
XAC2209	2593,977	2594,825	XF1780	16,97908	16,98657	Hypothetical protein
XAC2210	2595,000	2595,359	XF1779	16,97378	16,97911	Hypothetical protein
XAC2211	2595,501	2595,887	XF1778	16,96852	16,97304	Single-stranded DNA binding protein
XAC2213	2598,564	2600,027	XF1774	16,90011	16,94029	DNA methyltransferase
XAC2217	2603,332	2603,745	XF1772	16,88664	16,89074	Hypothetical protein
XAC2218	2603,765	2604,157	XF1771	16,88250	16,88642	Hypothetical protein
XAC2221	2606,053	2606,418	XF1764	16,80412	16,80765	Hypothetical protein
XAC2236	2619,587	2620,810	XF1761	16,77741	16,78868	Hypothetical protein
XAC2237	2620,870	2621,571	XF1760	16,77021	16,77683	Hypothetical protein
XAC2238	2621,664	2622,008	XF1759	16,76603	16,76926	Plasmid-related protein
XAC2239	2622,116	2622,523	XF1758	16,76095	16,76502	Hypothetical protein
XAC2240	2622,538	2622,798	XF1757	16,75818	16,76078	Hypothetical protein

^aSpecific in relation to XCC, XF-PD, XF-Ann-1, XF-Dixon genome.

Figure 1 (See legend on next page.)

Figure 1 Comparative view of the biological processes involved in the lifestyle of *X. axonopodis* pv. *citri* (A) and *X. fastidiosa* 9a5c (B). The categories with differences between the two bacteria are divided in three groups (adaptation, metabolism, and motility) and are numbered as follows: 1, regulators of pathogenicity factor genes; 2, xanthomonadins; 3, xanthan gum biosynthesis genes; 4, cell wall-degrading enzymes; 5, type II secretion system; 6, uptake of sugars; 7, type III secretion system; 8, oxide-reductases; 9, colicins; 10, restriction and modification systems; 11, energy and general metabolism; 11a, biosynthesis of fatty acids; 11b, degradation of fatty acids; 12, oxidative phosphorylation (cytochromes); 13, flagellum and chemotactic system; 14, pili.

Figure 2 Schematic alignments of the genomic region corresponding to T2SS of *X. axonopodis* pv. *citri* (XAC-xps and XAC-xcs) and *X. fastidiosa* 9a5c (XF-xps). A, alignment of the syntenic arrangement between XF-xps and XAC-xps. B, alignment of the flanking region of XAC-xcs on the genome of *X. fastidiosa*. XAC-xcs genes are more similar to T2SS of *C. crescentus* than to XAC-xps. (#), genes that also have high similarity to genes of *C. crescentus*. XpsD–N, and XcsC–N, components of T2SS.

wall-degrading enzymes encoded by XAC than XF-9a5c, and the mode of infection and dissemination in the plant, as described below. One cluster in XAC has been named xps and is similar to the XF-9a5c xps cluster (Figure 2A) (14). The other cluster has much greater similarity to a cluster in *Caulobacter crescentus* (CC) (40), and has been called xcs (14). Interestingly, seven genes that surround xcs have strong similarity to other CC genes and are located in the same relative position (Figure 2B), forming an “island” composed of 17 genes (39). This island is flanked on both sides by clusters of genes that are linked in the XF-9a5c genome (Figure 2B). There are two possible interpretations for this observation: either XF-9a5c lost the xcs cluster or XAC gained it by lateral transfer after divergence from XF-9a5c. A phylogenetic analysis of these clusters (data not shown) suggests that the divergence between the two operons is ancient and dates back at least to ancestral *Xanthomonadaceae*, implying that both clusters existed before the divergence of *Xylella* and *Xanthomonas*.

Cell Wall–Degrading Enzymes (CWDE)

Several hydrolytic enzymes are secreted by xanthomonads. Once they come into contact with plant tissues, these enzymes first expose the wall layer and then release masses of fibrillar material to facilitate colonization and dispersal to new tissues (52). XAC has extensive and diverse machinery for CWDE (14) compared with XF-9a5c (54). XAC has 6 copies of genes for pectinolytic enzymes and 12 copies of genes for cellulolytic and hemicellulolytic enzymes, whereas XF-9a5c encodes only 1 pectinolytic enzyme and 4 cellulolytic enzymes. The larger number of CWDE and the presence of two copies of the T2SS, which may be involved in

secretion of the CWDE, may account for the necrosis phenomenon that follows water soaking and hyperplasia in infection by *Xanthomonas*, which is not observed in infection by XF-9a5c. XCC has even more copies of genes encoding CWDE than does XAC, and, if all are translated and transcribed, this could account for the fact that XCC causes much more tissue necrosis than does XAC (14). XF-9a5c, which relies on an insect vector to gain access to the plant xylem vessels, and only would require CWDE to move between vessel elements, may not require the extensive complement of CWDE found in the *Xanthomonas* species.

Uptake of Sugars

Cell wall degradation might produce a set of sugars, providing a critical energy resource for bacterial metabolism. For uptake of these compounds from the environment, bacteria use a sugar phosphotransferase system (PTS), usually composed of two general energy-coupling proteins, a histidine-containing phosphoprotein (HPr), and several sugar-specific enzymes (49). Both XAC and XF-9a5c have an incomplete operon coding the PTS system specific to glucose (XAC2974-2979 and XF1408-1401) (38). By contrast, the PTS specific to fructose is present in XAC and absent in XF-9a5c, suggesting that PTS is incapable of processing this sugar in XF-9a5c. Surprisingly, the first gene downstream of the PTS fructose cluster in XAC is *rpfN*, which is involved in regulating pathogenicity factors but is absent in XF-9a5c. This gene was described as a possible outer-membrane porin or glucose-sensitive porin with great similarity to the *oprB* gene of *Pseudomonas aeruginosa*. Its function therefore is related to carbohydrate uptake, mainly fructose (65).

Besides the systems described above, it is possible that only XAC has the capability to internalize oligogalacturonides, galacturonates, or 2-keto-deoxygluconate (KDG) by specific transporters coded by *exuT* and *kdgT* (XAC4255 and XAC0337, respectively) while Xf-9a5c does not have this system. In the cell, these compounds could be converted to pyruvate and 3-phosphoglyceride using enzymes encoded by *kduIDKA* genes (26).

Energy and General Metabolism

XAC has genes encoding a complete energy metabolism including glycolysis, glyconeogenesis, Krebs cycle, pentose phosphate pathway (oxidative and nonoxidative phases), biosynthesis, and degradation of fatty acids (β -oxidation) (14). In contrast, some genes that code for important enzymes of these pathways are absent in XF-9a5c, for example, missing genes are those encoding fructose-1,6-biphosphate and phosphoenolpyruvate carboxykinase from the gluconeogenesis pathway; malate synthase and isocitrate lyase from the glyoxilate pathway; glucose-1-dehydrogenase and transaldolase B from the pentose phosphate pathway, and the entire pathway of β -oxidation (54). Therefore, the presence of two T2SS, a large number of CWDE, and a possible fructose PTS system associated with a complete energy metabolism may be related to XAC's adaptation to obtain and metabolize sugars during its infection of the mesophyll. On the other

hand, XF-9a5c, as a xylem-limited bacterium, has only a single copy of T2SS, a small number of CWDE, and, therefore, limited abilities to acquire energy from its hosts. These differences in pathogen gene complements could be related to differences in the nutrient content of the xylem relative to the intercellular spaces of the mesophyll.

Another metabolic divergence observed between these genomes is the presence of two types of terminal oxidases in XAC but only one in XF-9a5c. Both bacteria have the cytochrome O oxidase (bo_3) coded by genes *cyoABCDE* (XAC1258-61, 1736, and XF1387-90, 1360). The second system of XAC, called cytochrome d oxidase or complex bd, is coded by genes *cydAB* (Xac2336-37). Both systems are able to reduce O_2 in the presence of ubiquinol-8 (27). Note that both bacteria lack the *cbb₃* complex and therefore are incapable of living in an anaerobic environment. A comparative analysis on the presence or absence of different terminal oxidases in other organisms has shown a great divergence in the composition and presence of these complexes. However, XF-9a5c is unique as a phytopathogen in having only one type of oxidase complex (bo_3), and as the only organism lacking the bd complex (39). The biochemical differences caused by these terminal oxidases are related to kinetic parameters since the bo_3 complex has a higher V_{max} and is more efficient in creating a proton electrochemical gradient than the bd complex, although it has a lower affinity for oxygen. Expression of the bo_3 complex is therefore higher under aerobic conditions whereas the bd complex is synthesized and accumulated under low oxygen concentrations (2, 43). Therefore, the presence of two complexes on XAC indicates a possible modulation of the respiratory chain in adaptation to differences in environment. On the other hand, XF-9a5c is restricted to the specific aerobic conditions of the xylem, and that restriction may also be related to the fastidious growth of this organism (7).

Bacteria require protection against the production of reactive oxygen species (ROS) by the plant in response to the secretion of cell wall-degrading enzymes and *avr/hrp* genes (described below), or from injuries induced by photo exposure during their epiphytic phase. To prevent damage from ROS, XAC contains a greater number of genes involved with oxide-reduction (41 genes versus 6 genes on XF-9a5c's genome). Once again, the number of genes reflects a possible adaptation to different environments and lifestyle of these organisms (39).

Iron Uptake and Metabolism

Iron uptake in bacteria is important because of its use in metabolism, such as synthesis of cytochromes, aromatic amino acids, and pyrimidines, and its role as a cofactor in enzymatic reactions (18). To enable uptake of iron, bacteria use two pathways: (a) by taking oxidized Fe^{2+} from the environment, which is not easy because iron is usually found in its reduced form Fe^{3+} ; or (b) by chelating Fe^{3+} through the use of siderophores, which transform Fe^{3+} to Fe^{2+} . Genes that encode the well-known siderophores were not found in either genome, although Wiggerich and collaborators showed that *Xanthomonas* is capable of producing these compounds (64). We hypothesize that since the siderophores synthesized

by these bacteria are not characterized, the enzymes involved in their synthesis are not known, and the genes would show up as hypothetical genes. Even though the genes related to metabolism of iron are very similar in number and sequence (14, 54), XAC has a larger number of genes responsible for ferric uptake, with 63 genes, compared with only 4 genes in XF-9a5c. This difference might translate to a higher efficiency of uptake of iron by XAC, and may be related to its host-tissue specificity and growth during colonization.

Type III Secretion System (T3SS)

The type III secretion system (T3SS) machinery is highly conserved among animal and plant bacteria, especially the structural genes. The function of this system is associated with the export of effector proteins into the cytosol of the host cell, and is responsible for modulating host cellular function (8, 20). XAC has a copy of this gene cluster, composed of 24 genes (XAC0393-0417) (14). XF-9a5c does not have a T3SS (54). In XAC there are four genes upstream of the T3SS cluster that are involved in biotin synthesis (BioSy), and downstream of the cluster there are six genes involved in poly- and oligosaccharide degradation (PoDe) (Figure 3). The genes flanking this entire region (BioSy, Hrp, and PoDe) are also found in XF-9a5c, but with an 11-gene gap (16.3 Kbp), indicating a probable rearrangement (Figure 3) (39). The HRP and PoDe clusters are absent in XF-9a5c, and the BioSy gene cluster is fragmented into three parts; no avirulence effector genes similar to

Figure 3 Schematic representation of the region around the T3SS on *X. axonopodis* pv. *citri* (XAC) and the corresponding regions on the genome of *X. fastidiosa* 9a5c (XF). On the XAC genome the cluster that encodes enzymes involved with biotin syntheses (BioSy), T3SS genes (HRPs), and with polysaccharide degradation (PoDe) are together. In XF, the HRP and PoDe clusters are absent and the flanking regions of the entire cluster (black bar) are separated by 16.3 Kbp that contain the genes corresponding to XAC2499-2496 and XAC1596/95. The BioSy cluster (gray bar) is fragmented and is located far apart. R and C represent Arg and Cys tRNAs, respectively.

those found in XAC were found (54). These findings indicate that XF-9a5c does not interact with the citrus cells in the same way as does XAC.

Flagellum and Chemotatic Systems

The flagellum is a cellular organelle with a filamentous structure that is responsible for bacterial motility, as well as an export apparatus to transfer proteins involved in synthesis of the filament. XAC contains a complete set of genes involved with synthesis and regulation of flagella (14), but no genes related to this function were found in XF-9a5c (54). In XAC the flagellar apparatus is divided into three clusters of genes in a region of approximately 126 Kbp. Between clusters 1 and 2 is an interval of 20 genes, of which 10 are hypothetical genes and 5 are related to mobile elements, suggesting that these clusters once formed a single cluster (39). Between clusters 2 and 3 are 17 genes, all with putatively assigned functions. Interestingly, cluster 1 has 10 tandem copies of the *tsr* gene. The product of this gene, methyl chemotatic protein (MCP), is involved in the capture and transfer of chemotatic signals to the bacteria in response to the presence of serine, alanine, or glycine (56). Multiple alignments of these sequences show a high degree of identity, in fact, higher than that found in other MCPs from the XAC genome. A phylogenetic analysis (data not shown) of all *mcp* and *tsr* genes from XAC and other bacteria indicated that although all copies of *tsr* from this region belong to a single clade, *mcp* genes from outside this region also cluster together with high bootstrap support values. This suggests either that some copies were duplicated and later translocated to distant regions or that not all copies of *tsr* were duplicated in a tandem fashion. The inclusion of *tsr* genes from XCC shows that the duplication of *tsr* genes occurred at least in the ancestor of today's *Xanthomonas*. Therefore, these genes are a real duplication event, probably related to adaptation and indicating an important role for metabolism of certain amino acids in this organism. The absence of flagellum in the genome of XF-9a5c can be related to specialization to an insect vector or, recalling the absence of T3SS, to adaptation to the environment (xylem), where the flux of fluid inside the plant (sap) interferes with motility.

Regulator of Pathogenicity Factor (RPF) Genes

Quorum sensing in gram-negative bacteria is normally mediated by N-acyl derivatives of homoserine lactones (acyl HSLs) that work as signaling molecules to control the expression of various physiological functions (63). However, in *Xanthomonas* there is another system mediated by a low-molecular-weight signaling molecule called a diffusible signal factor (DSF) (4). The synthesis of DSF molecules is mediated by a regulator of pathogenicity factor genes (*rpf*), which is also responsible for controlling expression of genes for xanthan gum biosynthesis and extracellular enzymes (59). The *rpf* gene cluster is composed of 9 genes, *rpfABFCHGDIE*; the genera *Xanthomonas* and *Xylella* are the only organisms known to date to have this cluster (16). Compared with the XAC genomic arrangement, XF-9a5c shows rearrangements in *rpf* clusters, with a 780-Kbp gap

between *rpjB* and *rpjF*, splitting the cluster into two parts (16). The proximity of *rpjB* to *rpjF* in *Xanthomonas* might facilitate the synthesis of DSF, because of the presence of a promoter upstream of *rpjB* that combines expression of B and F genes, and an independent promoter upstream of *rpjF* that enhances expression of the F gene (55). Because there is a split between *rpjB* and F in XF-9a5c, this rearrangement may cause a fall in *rpjF* levels and consequently in a cascade of cellular events that decreases DSF synthesis compared to that in XAC. In agreement with this hypothesis, Scarpari and collaborators (53) have demonstrated DSF production in XF-9a5c at a level 14-fold lower than that in XCC. However the presence of a functional RpfF protein seems to be very important to colonization of insect (39b).

Xanthomonadins

Xanthomonadins are yellow pigments derived from brominated aryl-polyenes that are found in the outer membrane of xanthomonads. They confer protection against photobiological damage during the bacterium's epiphytic phase (22, 46, 47). The xanthomonadin gene cluster, previously identified in *X. campestris* pv. *campestris*, consists of seven transcriptional units (pigABCDEFG) scattered within an 18.6-Kbp region in the main chromosome (45). One of its transcriptional units, pigB, has been described as a regulatory gene(s) involved in the biosynthesis of a diffusible factor, DF, that is required for production of xanthan gum and xanthomonadins (45, 46). The function of DF is very similar to that of DSF produced by *rpj* genes, regarding the quorum sensing mechanism and cell-to-cell signaling. However, the results from cross-complementation studies of pigB, *rpjF*, and *rpjB* mutants suggest that they act independently to regulate biosynthesis of pigment and extracellular enzymes, respectively (46).

Recently, a 20-Kbp region was partially sequenced and characterized in *Xanthomonas oryzae* pv. *oryzae* (XOO) as a xanthomonadin gene cluster. There are 14 genes in this region (Figure 4), of which genes 3, 6, 7, and 10 are required for xanthomonadin production (21).

Figure 4 Schematic alignments of the cluster of xanthomonadins in the genomes of *Xanthomonas oryzae* pv. *oryzae*, *X. axonopodis* pv. *citri*, and *X. fastidiosa*. Gray arrows represent putative genes. White arrows represent intergenic regions of the *Xylella* genome identified by blastX to have amino acid similarity to protein sequence of the corresponding genes in *X. o.* pv. *oryzae*. Numbers above the arrows as defined by Goel and collaborators (21).

As expected, XAC and XCC contain a genomic region that corresponds to the cluster mapped by Goel in XOO (14, 21). Although there is no experimental evidence for xanthomonadins and DF production in XF-9a5c, its genome has a cluster of genes that are similar to those present in the three species of *Xanthomonas* (*axonopodis*, *oryzae*, and *campestris*), indicating a vestigial presence of this cluster in XF-9a5c (Figure 4). In this region, genes XF0771, 0774, 0775, and 0777 are putative orthologs of genes 10, 07, 06, and 04, which Goel and collaborators characterized as being essential to xanthomonadin synthesis and transport to the outer membrane in XOO (21). Gene 03, also described by these authors as important in the production of xanthomonadins, has 67% identity with 141 amino acids coded by the intergenic region between gene XF0777 and 0778. The loss and degeneration of the other genes involved in xanthomonadin synthesis might be a consequence of *Xylella*'s adaptation to the xylem and, at the same time, to a specific vector for infection, thereby avoiding the need for adaptation to an epiphytic phase.

Xanthan Gum Biosynthesis Genes

The *Xanthomonadaceae* family produces extracellular polysaccharides (EPS), also known as xanthan gum. This substance has a high commercial value and is involved in water-soaking formation during the colonization of tissue by *Xanthomonas* (51). The biosynthesis of xanthan gum in members of the genus *Xanthomonas* is coded by 12 genes that make up a single cluster of 16 Kbp in which each gene has a function related to the synthesis or export of xanthan gum (9, 24, 29, 30). Expression of these genes is regulated by RPF (58, 59) and by sugar concentration in the environment (62), and mutations of these genes may decrease pathogenicity in vivo (12, 30). The xanthan gum gene cluster in different strains of *Xanthomonas* is well conserved, with 98% identity (14). However, when compared with XF-9a5c genes, similarity decreases to values between 65% and 83%; the genes *gumGIL*, which are involved in elongation and modification of the polymer (15), are absent in XF-9a5c. These differences may confer different physicochemical properties on XF-9a5c gum and may denote an adaptative advantage for life in the xylem or in the vector.

Transcriptional Factors and Functional Regulators

The complexity of XAC's lifestyle in relation to that of XF-9a5c is illustrated by comparing the number of regulatory genes that they each have. XAC has 296 regulatory genes (7.6% of the whole genome), four times more than XF-9a5c (77 genes, 4.2% of genome) (14, 54). These genes have been classified into four groups: (a) two-component systems with 69 genes in XAC and 22 genes in XF-9a5c; (b) activators and repressors with 164 genes in XAC and 41 genes in XF-9a5c; (c) kinases and phosphatases with 32 genes in XAC and 3 genes in XF-9a5c; and (d) sigma factors and other regulatory components with 31 genes in XAC and 11 genes in XF-9a5c. Within these groups, some genes warrant special attention

in XAC. First, a set of two-component systems with similarity to *nodVW* that is involved in the regulation of nodulation genes (33), and second, a gene with similarity to *ntrBC* that is involved with nitrogen assimilation (35, 36). Note that although XAC has genes with similarity to regulators of nodulation and nitrogen assimilation, there is no proof of the occurrence of either phenotype.

Genes Overrepresented in the Genome of *Xylella fastidiosa* 9a5c

TYPE IV PILI SYSTEM The pili complex is a very important mechanism utilized by bacteria during the process of adhesion and colonization of host tissue. The pili complex is divided into classes, of which the best studied is the type IV pili, with a filamentous structure and polar position responsible for twitching motility on adhesion surfaces (1, 57). Both XF-9a5c and XAC have genes that encode type IV-A pili (14, 54). However, XF-9a5c has more genes related to this structure than does XAC.

The identification of homologous genes involved in pili biosynthesis between XF-9a5c and XAC revealed two groups of genes that were identified as putative operons and that include the pilin gene (*pilE*): (a) cluster *pilB-pilC-pilE-pilD*, with one copy in each genome; and (b) cluster *pilE-pilY1-pilX-pilW-pilV-fimT*, with two copies present in the XF-9a5c genome and one copy in the XAC genome. Other copies of pilin genes and some pieces of the two clusters are dispersed around the genomes of XF-9a5c and XAC. Phylogenetic studies for each cluster suggest the presence of two copies of the *pilE-fimT* operon in the ancestor of *Xanthomonadaceae*, one copy being lost in XAC and the other in XCC, but both still present in XF-9a5c (data not shown).

Additional evidence for the distinct histories for the pilin gene clusters *pilB-pilD* and *pilE-fimT* comes from a phylogenetic analysis of the pilin gene. This analysis has shown that the pilin genes of these two clusters are isolated into distinct monophyletic groups (data not shown). Moreover, the isolated copies of pilin found in the genome of XF-9a5c belong to the same group as the copy of pilin from the *pilB-pilD* gene cluster. All the evidence points to recent duplications of the pilin gene in XF-9a5c and suggests different cellular roles for the two operons containing the pilin gene.

This expansion of genes involved in synthesis and regulation of pili in XF-9a5c may reflect their importance to attachment to the insect and/or colonization and survival in the citrus xylem. The capacity to form aggregates may contribute to obstruction of vessels, which reduces the level of water in higher parts of the plant. Formation of these aggregates may be facilitated by pili, which is consistent with their higher complexity in XF-9a5c.

Colicins

Colicins are produced by a large number of bacteria and promote bacterial antagonism, reducing the competition for nutrients in the environment (31). The complete complement of genes for the synthesis of type V colicins was found in XF-9a5c (14, 54); however, two important genes for this process are not found

in XAC. The genes that code for synthesis of this compound in *Escherichia coli* are inside the plasmid pColV-K30 and are regulated by *cvpA-purF* genes (19). In both phytopathogens, we found that the genes *cvpA-purF* show a high degree of conservation (XAC1029-32 and XF1946-49). However, in XAC the genes *cvaC* and *cvaA* are absent; these genes are, respectively, a precursor of colicin synthesis, present in three tandem copies on the genome of XF-9a5c, and a protein of family MFP (membrane fusion protein) involved in secretion of this compound. In the case of the *cvaA* gene, it is disrupted by a transposon insertion in XAC (14). The presence of colicins in XF-9a5c and its absence in XAC correlate with their lifestyle in as much as XF-9a5c lives in the xylem in competition with an endophytic community whereas XAC alone colonizes the mesophyll.

Restriction and Modification System (RM)

Restriction and modification systems (RM) are responsible for cellular protection and maintenance of genetic material against invasion of exogenous DNA. Three models of RM, types I, II, and III, are described in the literature (48). XF-9a5c has four copies of a cluster with three or four genes related to RM-I. Three of these clusters occur near the replication terminus. The fourth cluster has two pseudogenes, flanked upstream by Phe-tRNA. XAC has only one copy of RM-I (Xac2898-900) and one other copy of RM-II (XAC2437-38). Although XF-9a5c has many copies of RM-I, no gene has been found that is involved with RM-II, although there are two RM-II genes in XF-PD, the enzyme *nspV* and its methylase (XF-PD1607-08), flanked by a tRNA-N (60). This same region in XF-9a5c presents a putative insertion of a genomic island, described by Nunes and collaborators (41) as GI1, flanked by an integrase.

Interestingly, XF-9a5c has more regions on the genome that are related to bacteriophages (around 7% of the genome) than does XAC, including a total of 82 genes that encode phage-related proteins. However, XF-9a5c has only seven genes related to transposable elements. In contrast, XAC has a large number of transposable elements (84 genes) and only 35 genes related to phage insertions. Divergences in the type of exogenous elements present in the genome, such as phages or transposons, might be the result of the types of restriction/modification systems found in these bacteria. Furthermore, the differences in the number and type of RM might be related to the environment in which they live; because XF-9a5c lives in the xylem, it might be more exposed than XAC to exogenous DNA and need a greater number of a particular type of RM system genes.

CONCLUDING REMARKS

The genera *Xanthomonas* and *Xylella* belong to the same family and are made up of several species. XAC and XF-9a5c are two examples of this family that invade the same host, sweet orange, but live in different environments and cause different types of pathology. In our comparative analysis, for each bacterium we

found gene clusters that may well be important to colonization and survival of these organisms in their particular environment. Interestingly, we were not able to identify a common set of genes that could be related to life in the same host, perhaps because these strains live in distinct parts of the plant. However, most of the conclusions we draw here can be adapted to the other bacteria of these families for which the genomes have been sequenced, such as XCC (14) or XF-PD (60). They possess similar systems and live in the same habitat as XAC and XF-9a5c. This observation indicates that adaptation to a specific host such as citrus is due to slight differences in sequence composition or in the number of genes.

The Annual Review of Phytopathology is online at <http://phyto.annualreviews.org>

LITERATURE CITED

1. Alm RA, Mattick JS. 1997. Genes involved in the biogenesis and function of type-4 fimbriae in *Pseudomonas aeruginosa*. *Gene* 192:89–98
2. Anraku Y, Gennis RB. 1987. The aerobic respiratory chain of *Escherichia coli*. *Trends Biochem. Sci.* 12:262–66
3. Banks D, Albibi R, Chen J, Lamikanra O, Jarret RL, Smith BJ. 1999. Specific detection of *Xylella fastidiosa* Pierce's disease strains. *Curr. Microbiol.* 39:85–88
4. Barber CE, Tang JL, Feng JX, Pan MQ, Wilson TJ, et al. 1997. A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol. Microbiol.* 24:555–66
5. Barry T, Collieran GGM, Dunican LK, Gannon F. 1991. The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. *PCR Methods Appl.* 1:233–36
6. Bhattacharyya A, Stilwagen S, Ivanova N, D'Souza M, Bernal A, et al. 2002. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. *Proc. Natl. Acad. Sci. USA* 99: 12403–8
7. Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, et al. 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res.* 12:1556–63
8. Bonas U, Schulte R, Fenselau S, Minsavage GV, Staskawicz BJ, Stall RE. 1991. Isolation of a gene cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. *Mol. Plant Microbe Interact.* 4:81–88
9. Capage MA, Doherty DH, Betlach MR, Vanderslice RW. 1987. *Int. Patent No. WO87/05938*
10. Chang CJ, Garnier M, Zreik L, Rossetti V, Bové JM. 1993. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. *Curr. Microbiol.* 27:137–42
11. Chang CJ, Schaad NW. 1982. Electrophoretic protein profiles of total cell envelopes of xylem-limited plant pathogenic rickettsia-like bacteria (RLB). *Phytopathology* 72:935–36
12. Chou FL, Chou HC, Lin YS, Yang BY, Lin NT, et al. 1997. The *Xanthomonas campestris* gumD gene required for synthesis of xanthan gum is involved in normal pigmentation and virulence in causing black rot. *Biochem. Biophys. Res. Commun.* 233:265–9
13. da Costa PI, Franco CF, Miranda VS, Teixeira DC, Hartung JS. 2000. Strains of *Xylella fastidiosa* rapidly distinguished by arbitrarily primed-PCR. *Curr. Microbiol.* 40:279–82

14. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, et al. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417:459–63
15. da Silva FR, Vettore AL, Kemper EL, Leite A, Arruda P. 2001. Fastidious gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *FEMS Microbiol. Lett.* 203:165–71
16. Dow JM, Daniels MJ. 2000. *Xylella* genomics and bacterial pathogenicity to plants. *Yeast* 17:263–71
17. Duan YP, Castaneda A, Zhao G, Erdos G, Gabriel DW. 1999. Expression of a single, host-specific, bacterial pathogenicity gene in plant cells elicits division, enlargement and cell death. *Mol. Plant Microbe Interact.* 12:556–60
18. Earhart CF. 1996. Uptake and metabolism of iron and molybdenum. See Ref. 39a, pp. 1075–90
19. Fath MJ, Mahanty HK, Kolter R. 1989. Characterization of a *purF* operon mutation which affects colicin V production. *J. Bacteriol.* 171:3158–61
20. Galan JE, Collmer A. 1999. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* 284:1322–28
21. Goel AK, Rajagopal L, Nagesh N, Sonti RV. 2002. Genetic locus encoding functions involved in biosynthesis and outer membrane localization of xanthomonadin in *Xanthomonas oryzae* pv. *oryzae*. *J. Bacteriol.* 184:3539–48
22. Goel AK, Rajagopal L, Sonti RV. 2001. Pigment and virulence deficiencies associated with mutations in the *aroE* gene of *Xanthomonas oryzae* pv. *oryzae*. *Appl. Environ. Microbiol.* 67:245–50
23. Gottwald TR, Graham JH. 2000. Bacterial diseases. In *Compendium of Citrus Diseases*, ed. LW Timmer, SM Garnsey, JH Graham. St. Paul: APS Press
24. Harding NE, Cleary JM, Ielpi L. 1995. Genetics and biochemistry of xanthan gum production by *Xanthomonas campestris*. In *Food Biotechnology Microorganisms*, ed. YH Hui, G Khachatourians, pp. 495–514. New York: VCH Publ.
25. Henderson M, Purcell AH, Chen D, Smart C, Guilhabert M, Kirkpatrick B. 2001. Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. *Appl. Environ. Microbiol.* 67:895–903
26. Hugouvieux-Cotte-Pattat N, Condemine G, Nasser W, Reverchon S. 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annu. Rev. Microbiol.* 50:213–57
27. Ingledew WJ, Poole RK. 1984. The respiratory chains of *Escherichia coli*. *Microbiol. Rev.* 48:222–71
28. Kamper SM, French WJ, DeKloet SR. 1985. Genetic relationships of some fastidious xylem-limited bacteria. *Int. J. Syst. Bacteriol.* 35:185–88
29. Katzen F, Becker A, Zorreguieta A, Puhler A, Ielpi L. 1996. Promoter analysis of the *Xanthomonas campestris* pv. *campestris* gum operon directing biosynthesis of the xanthan polysaccharide. *J. Bacteriol.* 178:4313–18
30. Katzen F, Ferreira DU, Oddo CG, Ielmini MV, Becker A, et al. 1998. *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J. Bacteriol.* 180:1607–17
31. Lazdunski CJ, Bouveret E, Rigal A, Journet L, Llobes R, Benedetti H. 1998. Colicin import into *Escherichia coli* cells. *J. Bacteriol.* 180:4993–5002
32. Lee RF, Beretta MJG, Hartung JH, Hooker ME, Derrick KS. 1993. Citrus variegated chlorosis: confirmation of a *Xylella fastidiosa* as the causal agent. *Summa Phytopathol.* 19:123–25
33. Loh J, Garcia M, Stacey G. 1997. NodV and NodW, a second flavonoid recognition system regulating nod gene expression in *Bradyrhizobium japonicum*. *J. Bacteriol.* 179:3013–20
34. Lopes JRS, Beretta MJG, Harakava R, Almeida RPP, Krugner R, Garcia A Jr. 1996. Confirmação da transmissão por cigarrinhas do agente causal da clorose

- variada dos citros, *Xylella fastidiosa*. *Fitopatol. Bras.* 1:343
35. MacFarlane SA, Merrick M. 1985. The nucleotide sequence of the nitrogen regulation gene *ntfB* and the *glnA-ntfBC* intergenic region of *Klebsiella pneumoniae*. *Nucleic Acids Res.* 13:7591–606
36. Martinez-Argudo I, Martin-Nieto J, Salinas P, Maldonado R, Drummond M, Contreras A. 2001. Two-hybrid analysis of domain interactions involving *NtrB* and *NtrC* two-component regulators. *Mol. Microbiol.* 40:169–78
37. Mehta A, Rosato YB. 2001. Phylogenetic relationships of *Xylella fastidiosa* strains from different hosts, based on 16S rDNA and 16S-23S intergenic spacer sequences. *Int. J. Syst. Evol. Microbiol.* 51:311–18
38. Meidanis J, Braga MD, Verjovski-Almeida S. 2002. Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa*. *Microbiol. Mol. Biol. Rev.* 66:272–99
39. Moreira LM, da Silva ACR. 2002. *Análise comparativa entre os genomas dos fitopatógenos Xylella fastidiosa e Xanthomonas axonopodis pv citri*. MSc. thesis. Univ. Sao Paulo, Chem. Inst.
- 39a. Neidhardt FC, Curtiss RI, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikopp WS, Riley M, Schaechter M, Umbarger HE, eds. 1996. *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Washington, DC: ASM Press
- 39b. Newman K, Almeida RPP, Purcell AH, Lindow SE. 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. *Proc. Natl. Acad. Sci. USA* 101:1737–42
40. Nierman WC, Feldblyum TV, Laub MT, Paulsen IT, Nelson KE, et al. 2001. Complete genome sequence of *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. USA* 98: 4136–41
41. Nunes LR, Rosato YB, Muto NH, Yanai GM, da Silva VS, et al. 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res.* 13:570–78
42. Pickard D, Wain J, Baker S, Line A, Chohan S, et al. 2003. Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7. *J. Bacteriol.* 185:5055–65
43. Poole RK, Ingledew WJ. 1987. Pathways of electrons to oxygen. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. FC Neidhardt, JL Ingraham, KB Low, B Magasanik, M Schaechter & HE Umbarger, pp. 170–200. Washington, DC: ASM Press
44. Pooler MR, Hartung JS. 1995. Genetic relationships among strains of *Xylella fastidiosa* from RAPD-PCR data. *Curr. Microbiol.* 31:134–7
45. Poplawsky AR, Chun W. 1997. *pigB* determines a diffusible factor needed for extracellular polysaccharide slime and xanthomonadin production in *Xanthomonas campestris* pv. *campestris*. *J. Bacteriol.* 179:439–44
46. Poplawsky AR, Chun W. 1998. *Xanthomonas campestris* pv. *campestris* requires a functional *pigB* for epiphytic survival and host infection. *Mol. Plant Microbe Interact.* 11:466–75
47. Poplawsky AR, Urban SC, Chun W. 2000. Biological role of xanthomonadin pigments in *Xanthomonas campestris* pv. *campestris*. *Appl. Environ. Microbiol.* 66: 5123–27
48. Redaschi N, Bickle TA. 1996. DNA restriction and modification systems. See Ref. 39a, pp. 773–81
49. Reizer J, Bachem S, Reizer A, Arnaud M, Saier MH Jr, Stulke J. 1999. Novel phosphotransferase system genes revealed by genome analysis—the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*. *Microbiology* 145:3419–29
50. Rossetti V, Garnier M, Bove JM, Beretta MJG, Teixeira ARR, et al. 1990. Présence de bactéries dans le xylème d'orangers

- atteints de chlorose variéguee, une nouvelle maladie des agrumes au Brésil. *C. R. Acad. Paris* 310:345–49
51. Rudolph K. 1993. Infection of the plant by *Xanthomonas*. In *Xanthomonas*, ed. JG Swings, EL Civerolo, pp. 193–264. London: Chapman & Hall
52. Russel M. 1998. Macromolecular assembly and secretion across the bacterial cell envelope: type II protein secretion systems. *J. Mol. Biol.* 279:485–99
53. Scarpari LM, Lambais MR, Silva DS, Carraro DM, 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
54. Simpson AJ, Reinach FC, Arruda P, Abreu FA, Acencio M, et al. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature* 406:151–57
55. Slater H, Alvarez-Morales A, Barber CE, Daniels MJ, Dow JM. 2000. A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol. Microbiol.* 38: 986–1003
56. Stock JB, Surette MG. 1996. Chemotaxis. See Ref. 39a, pp. 1103–29
57. Strom MS, Lory S. 1993. Structure-function and biogenesis of the type IV pili. *Annu. Rev. Microbiol.* 47:565–96
58. Tang JL, Feng JX, Li QQ, Wen HX, Zhou DL, et al. 1996. Cloning and characterization of the *rpfC* gene of *Xanthomonas oryzae* pv. *oryzae*: involvement in exopolysaccharide production and virulence to rice. *Mol. Plant Microbe Interact.* 9:664–66
59. Tang JL, Liu YN, Barber CE, Dow JM, Wootton JC, Daniels MJ. 1991. Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. *Mol. Gen. Genet.* 226:409–17
60. Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, 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
61. Van Sluys MA, Monteiro-Vitorello CB, Camargo LE, Menck CF, da Silva AC, et al. 2002. Comparative genomic analysis of plant-associated bacteria. *Annu. Rev. Phytopathol.* 40:169–89
62. Vojnov AA, Slater H, Daniels MJ, Dow JM. 2001. Expression of the gum operon directing xanthan biosynthesis in *Xanthomonas campestris* and its regulation in planta. *Mol. Plant Microbe Interact.* 14: 768–74
63. Whitehead NA, Barnard AM, Slater H, Simpson NJ, Salmond GP. 2001. Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol. Rev.* 25:365–404
64. Wiggerich HG, Klauke B, Koplin R, Priefer UB, Puhler A. 1997. Unusual structure of the *tonB-exb* DNA region of *Xanthomonas campestris* pv. *campestris*: *tonB*, *exbB*, and *exbD1* are essential for ferric iron uptake, but *exbD2* is not. *J. Bacteriol.* 179:7103–10
65. Wylie JL, Worobec EA. 1995. The *OprB* porin plays a central role in carbohydrate uptake in *Pseudomonas aeruginosa*. *J. Bacteriol.* 177:3021–26

ANEXO 4

**TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA**

ARTIGO PUBLICADO

Comparative analyses of Xanthomonas and Xylella complete genomes.

Moreira LM, De Souza RF, Digiampietri LA, Da Silva AC, Setubal JC.
OMICS. 2005 Spring;9(1):43-76.

PMID: 15805778 - PubMed - indexed for MEDLINE.

Comparative Analyses of *Xanthomonas* and *Xylella* Complete Genomes

*LEANDRO M. MOREIRA,¹ *ROBSON F. DE SOUZA,¹
LUCIANO A. DIGIAMPIETRI,² ANA C.R. DA SILVA,³ and JOÃO C. SETUBAL^{2,4}

ABSTRACT

Computational analyses of four bacterial genomes of the *Xanthomonadaceae* family reveal new unique genes that may be involved in adaptation, pathogenicity, and host specificity. The *Xanthomonas* genus presents 3636 unique genes distributed in 1470 families, while *Xylella* genus presents 1026 unique genes distributed in 375 families. Among *Xanthomonas*-specific genes, we highlight a large number of cell wall degrading enzymes, proteases, and iron receptors, a set of energy metabolism genes, second copy of the type II secretion system, type III secretion system, flagella and chemotactic machinery, and the xanthomonadin synthesis gene cluster. Important genes unique to the *Xylella* genus are an additional copy of a type IV pili gene cluster and the complete machinery of colicin V synthesis and secretion. Intersections of gene sets from both genera reveal a cluster of genes homologous to *Salmonella*'s SPI-7 island in *Xanthomonas axonopodis* pv *citri* and *Xylella fastidiosa* 9a5c, which might be involved in host specificity. Each genome also presents important unique genes, such as an HMS cluster, the *kdgT* gene, and O-antigen in *Xanthomonas axonopodis* pv *citri*; a number of *avrBS* genes and a distinct O-antigen in *Xanthomonas campestris* pv *campestris*, a type I restriction-modification system and a nickase gene in *Xylella fastidiosa* 9a5c, and a type II restriction-modification system and two genes related to peptidoglycan biosynthesis in *Xylella fastidiosa* temecula 1. All these differences imply a considerable number of gene gains and losses during the divergence of the four lineages, and are associated with structural genome modifications that may have a direct relation with the mode of transmission, adaptation to specific environments and pathogenicity of each organism.

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

²Instituto de Computação, Universidade Estadual de Campinas, Campinas, Brazil.

³Alellyx Applied Genomics, Campinas, Brazil.

⁴Virginia Bioinformatics Institute and Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

*These authors contributed equally to this work.

INTRODUCTION

XYLELLA AND *XANTHOMONAS* ARE TWO GENERA in the gamma subdivision of the proteobacteria that belong to the same family, the *Xanthomonadaceae*. All known species in these genera are phytopathogens. In this paper, we compare four species in this family, the only ones for which complete genomes are publicly available.

In the *Xylella* genus, two genomes have been published: *Xylella fastidiosa* 9a5c (Xf-9a5c), the causal agent of citrus variegated chlorosis (CVC) (Simpson et al., 2000), and *Xylella fastidiosa* Temecula 1 (Xf-temecula), the causal agent of Pierce's disease, which affects grapevine (Van Sluys et al., 2003). Two additional species have been partially sequenced: *Xylella fastidiosa* Ann-1 (infects oleander) and *Xylella fastidiosa* Dixon (infects almond) (Bhattacharyya et al., 2002a,b). In the *Xanthomonas* genus, for two species there are complete genomes publicly available: *Xanthomonas axonopodis* pv *citri* (XAC) and *Xanthomonas campestris* pv *campestris* (XCC) (da Silva et al., 2002).

Some previous comparative analyses involving these species have been done. Bhattacharyya et al. (2002a,b) compared Xf-9a5c with the *X. fastidiosa* Ann-1 and Dixon strains. Van Sluys et al. (2003) compared Xf-9a5c and Xf-temecula. Da Silva et al. (2002) and Lima et al. (2005) compared the two *Xanthomonas* species. Van Sluys et al. (2002) compared the genomes of eight plant-associated bacteria, including Xf-9a5c, XAC, and XCC. Finally, Moreira et al. (2004), using the common citrus host as a reference, compared Xf-9a5c and XAC. These analyses have shown that the genomes of all four organisms are closely related, while highlighting significant differences. In this paper, we present the first analysis of all four completely sequenced representatives of the *Xanthomonadaceae*. We expand on previous analyses and report several new results. These results shed new light on the pathogenicity mechanisms used by these organisms and bring understanding of the evolution of these related pathogens to a new level.

General features of Xylella

Xylella fastidiosa is a fastidious and xylem-limited bacteria that causes a great variety of diseases in distinct plant hosts (over 100 plant species), such as Pierce's disease of grape, periwinkle wilt, phony peach disease, citrus variegated chlorosis, and almond, plum, elm, maple, oak, and sycamore leaf scorch (Henderson et al., 2001; Hopkins, 1989; Purcell, 1980; Wells et al., 1987). Although these strains are generally classified as a single species, some genetic studies suggested the existence of multiple species (Schaad et al., 2004). *Xylella* cells are approximately 0.3–0.5 μm in diameter and 3–5 μm in length, with a single chromosome of approximately 2.6 Mbp, and may contain plasmids (Holt, 1994; Wells et al., 1981). Transmission of *Xylella* involves the obligatory participation of specific insect vectors (leafhoppers), entering the oral apparatus of insects while they feed on phloem sap from plants (Lopes et al., 1996). Inside the xylem, these bacteria cause vessel occlusion through cellular agglutination, a process that seems critical for the development of symptoms, which include chlorotic areas on the upper side of leaves, reduction of fruit size and hardened consistence (Chang, 1993; Lee, 1993; Rossetti, 1990). The two isolates we analyze here, Xf-9a5c and Xf-temecula, are responsible for reduction of citrus and wine production in São Paulo state (Brazil) and California (USA), respectively. Xf-9a5c has a single chromosome of 2.67 Mbp and two plasmids, with a total of 2249 coding sequences (CDSs), 58.4% of which have been assigned putative functions (Simpson et al., 2000; Van Sluys et al., 2002). Xf-temecula carries one chromosome of 2.51 Mbp and one plasmid, with 2066 predicted CDSs, 68.1% of which were assigned putative functions (Van Sluys et al., 2003) (Table 1).

General features of Xanthomonas

Infections induced by *Xanthomonas* spp. have been described for 124 monocotyledons and 268 dicotyledons, generally inducing tissue necrosis, vascular or parenchymal diseases, and damage of leaves and fruits that induce plant decline (Long and Staskawicz, 1993). *Xanthomonas* cells are rod shaped, with a diameter of 0.2–0.6 μm and a length of 0.8–2.9 μm , with a single chromosome of slightly more than 5 Mbp that may be associated with plasmids (Swings and Civerolo, 1993).

XAC is the causal agent of citrus canker, a disease characterized by the formation of canker associated with water-soaked lesions on leaves, generally surrounded by chlorotic haloes (Gottwald and Graham, 2000).

TABLE 1. GENERAL FEATURES OF THE FOUR *XANTHOMONADACEAE* GENOMES ANALYZED

Features	<i>Xf-9a5c</i>	<i>Xf-temecula</i>	XAC	XCC
General features				
Compatible host	Citrus plant	Grapevine plants	Citrus plant	<i>Brassica</i> sp
Disease	Citrus variegated chlorosis (CVC)	Pierce's disease (PD)	Citrus canker (CC)	Black rot of cruciferous (BR)
Tissues infected	Xylem	Xylem	Mesophyll	Mesophyll and vessels
Symptoms	Conspicuous variegations on older leaves; chlorotic areas on the upper side; corresponding light brown lesions; reduced and hardened fruits	Scorched leaves that detach; bare petioles attached to the canes; wilted, shriveled and raisin-like fruit	Canker lesions; abscission of fruit and leaves; general tree decline	Marginal leaf chlorosis; darkening of vascular tissue; extensive wilting and necrosis
Vector(s)	Sharpshooter leafhoppers	Sharpshooter leafhoppers		
Importance	CVC is considered the most important citrus disease in the State of São Paulo and is responsible for large annual losses.	PD is a major threat to the viability of the California wine industry	Animals, ambient conditions and utensils CC causes major economic losses to the citrus industry and is a nuisance to people with ornamental citrus trees.	Animals, ambient conditions and utensils BR is the most serious disease of crucifer crops world wide when environmental conditions (high temperature and humidity) are favorable
Genome features				
Genome status	Complete	Complete	Complete	Complete
Website	(Simpson et al., 2000) www.lbi.ic.unicamp.br	(Van Sluys et al., 2003) www.lbi.ic.unicamp.br	(da Silva et al., 2002) http://genoma4.iq.usp.br	(da Silva et al., 2002) http://genoma4.iq.usp.br
Genome size	2,679,305 bp	2,519,802 bp	5,175,554 bp	5,076,187 bp
Plasmid(s) (CDS)	pXF51 (64) pXF1.3 (2)	pXFPD1.3 (2)	pXAC33 (42) pXAC64 (73)	—
CDS number	2249	2066	4313	4182
With assigned function (%)	1314 (58.42%)	1408 (68.15%)	2770 (64.22%)	2708 (64.75%)
GC %	52.7%	51.8%	64.7%	65.0%
Ribosomal RNA operons	2	2	2	2
Transfer RNA (AA)	49 (20)	49 (20)	54 (20)	53 (20)
Phage proteins (%)	82 (3.64%)	139 (6.72%)	35 (0.81%)	43 (1.03%)
Transposases (%)	7 (0.31%)	7 (0.33%)	84 (1.94%)	108 (2.58%)
Comparative features				
Unique genes (%)	283 (12.58%)	104 (5.98%)	665 (15.41%)	555 (13.27%)
With classified function (%)	87 (30.74%)	74 (71.15%)	207 (31.12%)	215 (38.73%)
Unique Intersection	XF-temecula 1026; XAC 47; XCC 8	XF-9A5C 1026; XAC 2; XCC 38	XCC 3636; XF-9A5C 47; XF-temecula 2	XAC 3636; XF-9A5C 8; XF-temecula 38
Plant-associated genome Comparisons	(Bhattacharyya et al., 2002a; Bhattacharyya et al., 2002b; Koide et al., 2004; Moreira et al. 2004;	(Nunes et al., 2003 Van Sluys et al., 2003)	(da Silva et al., 2002; Lima et al., 2004; Moreira et al. 2004; Van Sluys et al., 2002)	(da Silva et al., 2002; Lima et al., 2004)
Involving Xanthomonadaceae genomes	Nunes et al., 2003; Van Sluys et al., 2003; Van Sluys et al., 2002)			
CDSs, coding sequences; AA, amino acids.				

The surface penetrating necrotic lesions on fruit lead to abscission of fruit and general tree decline, causing great losses in citrus production (Gottwald and Graham, 2000). These bacteria can survive and multiply outside the host as epiphytic organisms, but in contact with stomates, hydathodes, water pores or lesions of plant tissues, they may colonize the mesophyll producing the classical symptoms of the disease in compatible hosts. Dissemination and transmission generally occur through bacteria exudates derived from lesions under wet weather and by splash dispersal at short range, windblown rain at medium to long range, and human-assisted movement at all ranges (Graham et al., 2004). XAC contains one chromosome of 5.17 Mbp and two plasmids with a total of 4313 CDSs, 64.7% of which have been assigned a putative function (da Silva et al., 2002) (Table 1).

XCC is the causal agent of black rot in cruciferous plants (*Brassica* sp) and also infects weeds, including *Arabidopsis thaliana*. XCC propagates from host to host through the same mechanisms as XAC, but causes a systemic infection (Schaad and Alvarez, 1993). In compatible hosts, XCC induces marginal leaf chlorosis, darkening of vascular tissue, extensive wilting, and necrosis. Its cells contain a single chromosome of 5.07 Mbp with 4182 CDSs, 65% of them with assigned putative functions (da Silva et al., 2002) (Table 1).

MATERIALS AND METHODS

Gene clusterization

We clustered genes into a set of gene families, using results from an all-against-all BLAST analysis (Altschul et al., 1997) for all proteins coded by the four genomes. The clustering was performed in three steps: first, we selected all pairs of bidirectional best hits with e-values of $\leq 10^{-20}$, creating mutually exclusive families of two members each. In the second step, a singleton gene g was added to an already existing family of size n if at least the first k BLAST hits of g belonged to that family ($k = 0.8n$; the value 0.8 was empirically derived). The last step joined families. Two families f_1 and f_2 were joined if each had a sufficiently high number of genes with a sufficiently high number of hits in the other family. “Sufficiently high” threshold was chosen to be 80%. These criteria were derived from similar ones used by Digiampietri et al. (2003).

We call a gene g unique (or specific) to genome x if it is a singleton or if it belongs to a gene family that does not include members from other genomes (a paralogous family). This definition is extended in the obvious way to declare genes unique to two genomes with respect to the other reference genomes. The set of genes shared by two or more genomes (orthologous families) are all genes in the families that have at least one gene from each reference genome. The set of paralogous genes from genome x is comprised of the genes in families that contain at least two genes from genome x .

Phylogenetic analysis

Selection of homologous genes for phylogenetic analysis was performed through different methodologies for isolated and concatenated genes. Isolated genes were compared to sequences in the UNIPROT database (Apweiler et al., 2004) using BLAST (Altschul et al., 1997), and valid matches were selected based on varying degrees of sequence identity, alignment coverage and e-value. Concatenated alignments were built for putative operons by selecting, for each gene in a well defined seed cluster from a certain organism (usually XAC or XF-9a5c), all homologous sequences from the KEGG (Kanehisa et al., 2004) and COG (Tatusov et al., 2001) databases, and later selecting all homologous CDSs in different genomes that were separated by no more than ten CDSs from other homologous genes of the same seed cluster. Some genes and organisms were then dropped in order to avoid missing genes in the final alignment.

Once a set of homologous sequences was chosen for each gene, selected sequences were aligned using CLUSTALW, version 1.74 (Thompson et al., 1994), with default parameters, and conserved regions were selected using the program GBLOCKS (Castresana, 2000). When dealing with clusters, alignments were concatenated at this point. All alignments were then analyzed by the PROTML program (Adachi and Hasegawa, 1996) for maximum likelihood inference of phylogenies, using the Poisson model of amino acid substitution.

Integrase classification and putative insertion/deletion islands (PinDels)

Phage insertions and transposases are known mediators of rearrangements and horizontal gene transfer events, and we searched putative insertion/deletion events (PinDels) associated with phages. We defined PinDels as regions flanked by phage integrase genes and rich in hypothetical and phage-related genes. In a second step, we also accepted, as PinDels, regions without flanking integrases but with significant GC and codon biases, and also rich in hypothetical and phage-related genes. The GC and codon biases were computed following Karlin's methodology (Karlin, 2001).

As a complement to the analysis above, integrases were classified based on two criteria: (i) sequence similarity and (ii) integration site class (Table 2). Based on sequence similarity, four distinct paralogous groups were formed for integrases (I, II, III, and IV), and an independent group was defined for recombinases xerC and xerD. Three integration site classes were defined: (A) integrases flanked by t-RNA genes, (B) integrases flanking PinDels but not t-RNAs, and (C) integrases with no flanking t-RNAs or located inside a PinDel.

Determination of putative terminus of replication was based on maximum value for the cumulative strand bias (Lobry, 1996).

RESULTS*Structural comparisons and PinDels of the four genomes*

Figure 1 illustrates the rearrangements of relevant gene clusters among the genomes of the four *Xanthomonadaceae*. Graphs below each genome map represent variations in genome composition, measured by codon and GC biases.

In Xf-9a5c, four complete phages were found (Simpson et al., 2000), which could have been acquired from soil-inhabiting bacteria or directly by siphophage (Bhattacharyya et al., 2002a; Bhattacharyya et al., 2002b). In addition, 22 CDSs coding for phage proteins were found dispersed on the Xf-9a5c genome and in 11 PinDels (Fig. 1). Nunes et al. (2003) described eight of these Xf-9a5c PinDels as regions found only in citrus-related *Xylella* strains (in parenthesis the designation given by Nunes et al.): PinDel-1 (pGI2), PinDel-2 (GI1), PinDel-3 (GI3), PinDel-4 and PinDel-5 (GI2), PinDel-6 (GI4), PinDel-7 (pI1), PinDel-8 (GI5), and PinDel-11 (pGI6). Nunes et al. (2003) described a strong correlation of these regions with unique genes of *Xylella*. We detected two new PinDels, which we called PinDel-9 and PinDel-10. Interestingly, PinDel-6 is an inverted repeat, that is, it is composed by hypothetical and phage-related proteins that are duplicated and in reverse order. Xf-9a5c integrases flank all four phages and six PinDels (Table 2). All integrases of group II flank a phage or a PinDel (B-II). The majority of group I and all group IV Xf-9a5c integrases are flanked by tRNA genes (A-I and A-IV in Table 2). Also, six Xf-9a5c integrases (3, 4, 7, 13, 16, and 17) are positioned at the border of regions that characterize rearrangements when compared to Xf-temecula, and not just four, as described by Van Sluys et al. (2003). Therefore, sequence similarity and integration site classes are correlated, and all rearrangements between the *Xylella* strains could have been induced by phage insertion.

Although no complete phage insertion was found in Xf-temecula, 139 CDS's are grouped in 10 PinDels, 8 of which were previously described (Van Sluys et al., 2003). Xf-temecula PinDel-2 is located very close to the replication terminus (Fig. 1). Previous analysis has suggested that Xf-temecula PinDel-6 could have originated from a fusion of two different phages (Van Sluys, et al., 2003).

In *Xanthomonas* the number of genes coding for phage proteins is much smaller, with 35 CDSs in XAC and 43 CDSs in XCC (Moreira et al., 2004). No complete phage is present in XAC and a single phage (Lf) is found in the XCC genome. We identified thirteen PinDels in XAC and seven in XCC, all flanked by phage integrases, except PinDel-3 and PinDel-7 from XCC, which were identified based on anomalous GC content. PinDels 1, 2, 5, 6, 7, 8, 9, 10, 12, and 13 in XAC and 3, 4, 5, and 6 in XCC might have been inserted or lost after divergence of the two strains. PinDel-3 of XAC is homologous to PinDel-1 of XCC and the same holds for PinDel-4 and PinDel-11 of XAC and PinDel-2 and PinDel-6 of XCC, respectively. While *Xanthomonas* presents a smaller number of integrases and phage-related proteins than *Xylella*, it has many genes coding for transposases, which are not so abundant in *Xylella*.

TABLE 2. XYLELLA PHAGE INTEGRASES CLASSIFICATION

Integrase number		Genome	Integrase Group				Insertion			Rearrangements			Nearest homologues		
ORF number	Insertion		Blast	Strand	AA	Location	Name	Location	Start	End	Number	Location	XAC	XCC	
1	XF0480	XF-9a5c	A	IV	←	N	Up	PinDel-1	Up	XF0480	XF0558	—	—	XAC2682	XCC3012
2	XF0631	XF-9a5c	C	III	→	—	—	PinDel-3	Inside	XF0623	XF0648	—	—	XAC2222/XAC2183	—
3	XF0678	XF-9a5c	A	I	←	V	Up	XFP1	Up	XF0678	XF0733	I	Down	—	—
4	XF0968	XF-9a5c	A	I	←	V	Up	—	—	XF0968	—	II	Down	—	—
5	XF-9a5c	XF1425	<i>xerD</i>	Recomb	←	—	—	—	—	XF1425	—	—	—	XAC3551	XCC0654
6	XF1483	XF-9a5c	<i>xerC</i>	Recomb	→	—	—	—	—	XF1483	—	—	—	XAC0636	XCC3497
7	XF1555/XF1556	XF-9a5c	B	II	←	—	—	XFP3	Up	XF1555	XF1596	III	Down	—	—
8	XF1642	XF-9a5c	B	II	←	—	—	XFP4	Up	XF1642	XF1711	—	—	—	—
9	XF1718	XF-9a5c	A	III	→	G	Up	PinDel-7	Up	XF1718	XF1754	—	—	—	XCC2110
10	XF1754	XF-9a5c	B	III	→	—	—	PinDel-7/8	Dwn/Up	XF1718	XF1754	—	—	—	XCC2110?
11	XF1789	XF-9a5c	A	I	←	T	Down	PinDel-8	Down	XF1754	XF1793	—	—	—	—
12	XF2028	XF-9a5c	B	Singlet	←	—	—	PinDel-9	Up	XF2028	XF2081	—	—	XACb0061	XCC1630
13	XF2131/XF2132	XF-9a5c	B	II	→	—	—	PinDel-11	Down	XF2108	XF2132	IV	Down	—	—
14	XF2288	XF-9a5c	B	II	←	—	—	PinDel-12	Up	XF2288	XF2309	—	—	—	—
15	XF-9a5c	XF2309	A	I	→	G	Down	PinDel-12	Down	XF2288	XF2309	—	—	—	—
16	XF2478	XF-9a5c	B	I	→	—	—	XFP2	Up	XF2478	XF2530	V	Down	—	—
17	XF2530	XF-9a5c	B	II	→	—	—	XFP2	Down	XF2478	XF2530	V	Down	—	—
18	XF2761	XF-9a5c	A	IV	←	PRHK	Down	PinDel-13	Up	XF2761	XF2773	—	—	XAC2628	XCC3012
8	PD0384	XF-temecula	B	II	→	—	—	PinDel-1	Down	PD0363	PD0384	—	—	—	—
5	PD0652	XF-temecula	<i>xerD</i>	Recomb	←	—	—	—	—	—	—	—	—	XAC3551	XCC0654
6	PD0700	XF-temecula	<i>xerC</i>	Recomb	→	—	—	—	—	—	—	—	—	XAC0636	XCC3497
8	PD0764	XF-temecula	B	II	←	—	—	PinDel-2	Down	PD0764	PD0772	III	Down	—	—
12	PD0789	XF-temecula	C	Singlet	→	—	—	DePo #	Up	—	—	—	—	XACb0061	XCC1630
9	PD0990	XF-temecula	B	III	←	—	—	PinDel-5	Up	PD0990	PD1019	—	—	XAC2183	—
13	PD1019	XF-temecula	B	II	→	—	—	PinDel-5	Down	PD0990	PD1019	—	—	—	—
1	PD1075	XF-temecula	A	IV	←	T	Up	PinDel-6	Up	PD1075	PD1139	—	—	XAC2628	XCC3012
11	PD1078	XF-temecula	A	I	→	G	Down	PinDel-6	Up	PD1075	PD1139	—	—	—	—
8	PD1139	XF-temecula	B	II	→	—	—	PinDel-6	Down	PD1075	PD1139	—	—	—	—
8	PD1196	XF-temecula	B	II	→	—	—	PinDel-6	Down	PD1166	PD1196	—	—	—	—
14	PD1320*	XF-temecula	C	II	←	—	—	—	—	PD1320	PD1323	—	—	—	—
16	PD1495	XF-temecula	A	I	→	V	Down	—	—	—	—	V	Down	—	—
1	PD1605*	XF-temecula	A	IV	→	N	Down	PinDel-8	Down	PD1592	PD1605	—	—	XAC2628	XCC3012
16	PD1732*	XF-temecula	A	I	→	V	Down	PinDel-9	Down	PD1714	PD1732	I	Down	—	—

Integrase number 7 and 13 of Xf-9a5c are fragmented into two genes. Up, upstream; Down, downstream. *Gene with frameshift or point mutation. #Although DePo is not a PinDel, it is discussed in the text and therefor3e was highlighted. Recomb, recombinases; Xf-temecula integrase number was defined using the number of homologous integrase genes in Xf-9a5c.

The location of one copy of the type II secretion system (T2SS) genes, a cluster known as *xcs* (Fig. 1), is interesting. Although this system is absent from *Xylella*, the regions in *Xylella* homologous to the *xcs* flanking regions from XAC (A and B in Fig. 1) are joined together (Moreira et al., 2004). The pattern is the same as the one found for the flanking clusters (C and D) of genes in the type III secretion system (T3SS), which are also neighbors in *Xylella* (Moreira et al., 2004). Although the copies of T2SS *xps* and *xcs* are located at different regions in the XAC genome, these clusters belong to homologous translocated regions containing recombinases *xerC* and *xerD* (Fig. 1), also observed in Xf-9a5c. These two proteins are important to separate concatenated chromosomes during replication and have been involved in genome rearrangements (Hendricks et al., 2000). Region E (*fimT*, *uvrB* and a tRNA-Val gene) flanks genes in the Type IV secretion system in XAC, and is found between rearrangements II and III in Xf-9a5c, and upstream to rearrangement II in Xf-temecula.

For XAC, we have classified the integrases in three groups of sequence similarity and identified some that are homologous to the *Xylella* integrases (Table 3). Although we have not found any correlation between sequence similarity group and integration site class in XAC, eleven of eighteen integrases flank PinDels. In XCC nine of twelve integrases flank PinDels.

Analysis based on presence/absence of genes

Although many unique genes were described in previous work (da Silva et al., 2002; Moreira et al., 2004; Simpson et al., 2000; Van Sluys et al., 2002; 2003), a careful and detailed analyses of the four *Xanthomonadaceae* revealed new unique genes and systems, which seem to be involved in the processes of adaptation, colonization, and pathogenicity. These genes were classified into four distinct major groups: (i) genes unique to each strain; (ii) genes unique to each genus; (iii) genes unique to the intersection of strains of different genera; and (iv) genes unique to *Xanthomonas* and *Xylella* genera, that is, genes not found in organisms that do not belong to one of these genera. In the presentation that follows, these categories were divided into two subclasses: genes arranged in tandem, that is, close enough to be considered members of a gene cluster, and scattered unique genes.

Unique genes in each genome

Table 4 gives the number of unique genes in each genome with respect to the others as well as unique genes in combinations of genomes. The composition of these gene sets is analyzed next.

XAC-specific genes. There are 18 clusters of unique genes in XAC, which we call **XaUC n** , where the first **Xa** stands for *Xanthomonas axonopodis*, **UC** highlights that this is a unique cluster and n is a number from 1 to 18 (Table 5). XaUC1 (XAC0855-XAC0860) includes an ABC transporter permease complex, previously described as unique by da Silva et al. (2002), composed by the subunits *oppD*, *oppC*, and *oppB*, and an associated oligopeptide binding protein *oppA* (Fig. 2, square 8C). The cointegrate resolution protein subunits (SST) form a gene cluster (XAC3227-XAC3229) inside the region XaUC12, with a Tn5044 transposase gene (XAC3226) located upstream of the cluster and another transposase downstream, after three hypothetical genes. We note that XAC3227 and XAC3229 are homologous to integrase and cointegrase genes found in the pXAC64 plasmid (XACb0009 and XACb0010, respectively) and in a T3SS cluster from *X. campestris* pv. *vesicatoria* genome (Noel et al., 2003). The entire sequence of XAC3228, on the other hand, is highly similar to the N-terminal region of XAC3227, and thus seems to be a truncated copy of this gene. HMS proteins (Pendrak and Perry, 1991) form a single cluster comprising the subunits R, F, and H (XAC1811-XAC1813), known as pigmentation locus *pmg* in *Yersinia pestis* (Yp) (Fig. 2, square 7C). This cluster is inserted inside the XaUC15 region and is involved in uptake and storage of exogenous hemin groups (Pendrak and Perry, 1991; Schubert et al., 1998). XAC also has a XamI adenine-specific methyltransferase (XAC2436) and a type II XamI restriction enzyme (XAC2437) (da Silva et al., 2002; Gomez et al., 1997) (Fig. 2, squares 12F and 13F). These genes are located inside region XaUC9, which is composed by 29 CDSs, including 21 unique genes flanked by a tRNA-A, and six transposases, one phage and four plasmid-related proteins. The richness of plasmid-related and transposase genes suggests the involvement of plasmids in the acquisition of this region. The cluster formed by four pyridine nucleotide transhy-

FIG. 1. Structure and composition of the four genomes. Each genome is depicted based on three parameters: structural composition, GC content variation, and codon bias variation. The key to symbols used are given below the *Xylella* section. Gene clusters related to pathogenicity or adaptation are depicted by numbered black arrows just above the GC content graph; the key to numbers is given below the *Xanthomonas* section. The vertical dashed line represents our estimate for the location of the terminus of replication.

Xylella strains (XF-9a5c as model): 1 - NADH dehydrogenase complex (*nuo* genes); 2 - hemolysin-type calcium binding protein; 3 - Murein sacculus, peptidoglycan gene cluster; 4 - *xyI4*, *nahA*, *dad4*, *atr*; 5 - hemagglutinin-like secreted protein; 6 - arginine/proline metabolism; 7 - hemolysin-type calcium binding protein; 8 - ribosomal proteins; 9 - hemagglutinin-like secreted protein; 10 - gum genes; 11 - bacteriocin; 12 - RNA polymerase + ribosomal proteins; 13 to 15 - type I restriction-modification system; 16 - hemagglutinin-like secreted protein. **Xanthomonas strains** (XAC as model): 1 - phage-related proteins + transposase; 2 - transposase; 3 - phenolic compounds degradation; 4 - biotin metabolism cluster; 5 - degradation of polysaccharides cluster; 6 - Murein sacculus, peptidoglycan gene cluster; 7 - ribosomal proteins; 8 - HMS + haemagglutinin + hemolysin; 9 - histidine metabolism; 10 - *ppf* genes; 11 - *tr* genes; 12 - hemolysin-type calcium binding protein; 13 - arginine/proline metabolism; 14 - *gum* genes; 15 - NADH dehydrogenase complex (*nuo* genes); 16 - cobalamin metabolism; 17 - xanthomonadins small and large cluster. * Xanthomonadins cluster unique from XCC.

TABLE 3. XAC AND XCC UNIQUE GENE CLUSTERS

Island	XAC ORF number			PinDel	XAC CHP/HP	Transposases HXAC/UXAC	Phage proteins (int) HXAC/UXAC	tRNA XAC	ORF with problem HXAC/UXAC
	Initial	Final	T (U)						
XaUC01	0843	0860	18 (13)	—	—/3	—/—	—/—	—	2/—
XaUC02	0918	0924	07 (04)	—	—/—	—/—	—/—	—	—/2
XaUC03	1101	1108	08 (05)	2	—/2	2/—	(1)/1	S	—/—
XaUC04	1489	1511	23 (13)	3	3/7	2/—	(2)/1	tm	1/—
XaUC05	1809	1818	10 (05)	—	—/2	—/—	—/—	R	—/—
XaUC06	2174	2204	31 (15)	5	4/6	2/—	(1)/2	—	—/1
XaUC07	2214	2234	31 (13)	6	2/3	2/—	(1)/2	—	—/—
XaUC08	2269	2286	18 (13)	7	4/7	—/—	(1)/1	—	—/—
XaUC09	2417	2245	29 (21)	8	5/6	5/1	1 + 1#3#	A	—/1
XaUC10	2901	2904	04 (04)	10	3/—	—/—	—/(1)	—	—/—
XaUC11	3018	3025	08 (08)	—	—/8	—/—	—/—	—	—/—
XaUC12	3221	3234	11 (09)	—	—/4	3/1	—/—	—	—/3
XaUC13	3251	3299	48 (23)	11	2/18	3/—	—/(2) + 1#	G	1/2
XaUC14	3503	3531	29 (08)	—	1/5	2/—	—/—	—	4/—
XaUC15	3702	3732	31 (08)	—	—/—	—/—	—/—	—	1/1
XaUC16	3763	3786	24 (15)	12	6/7	2/—	(1)/—	M	2/1
XaUC17	3932	3989	58 (28)	13	10/11	3/2	—/(1)	A	1/—
XaUC18	4112	4148	37 (27)	—	16/5	1/1	—/—	—	1/2
Total	—	—	425 (232)	10	56/94	27/5	8(7) + 1#11(4) + 4#	7	13/13

Island	XCC ORF number			PinDel	XCC CHP/HP	Transposases HXCC/UXCC	Phage proteins (int) HXCC/UXCC	tRNA XCC	ORF with problem HXCC/UXCC
	Initial	Final	T (U)						
XcUC01	0319	0347	29 (20)	—	5/8	—/—	—/—	—	—/—
XcUC02	0524	0545	22 (08)	—	3/1	3/—	—/—	—	—/—
XcUC03	0599	0618	20 (10)	—	—/1	4/1	—/—	—	1/1
XcUC04	0735	0751	17 (08)	—	3/4	—/—	—/—	—	—/—
XcUC05	1307	1321	13 (06)	—	3/3	—/—	—/—	—	—/—
XcUC06	1446	1463	18 (09)	1	4/3	4/	2(1)/1	—	—/—
XcUC07	2091	2113	22 (16)	4	3/2	3/1	/(1)	GGLC	—/—
XcUC08	2413	2429	17 (07)	—	1/2	—/—	—/—	F	—/—
XcUC09	4048	4051	04 (04)	—	1/—	—/—	—/—	—	—/—
Total	—	—	162 (88)	2	23/24	14/2	2(1)/2(1)	2	1/1

XaUC_n, *Xanthomonas axonopodis* Unique Cluster number *n*; **XcUC_n**, *Xanthomonas campestris* Unique Cluster number *n*; T, Total; U, Unique; **PinDel**, Putative insertion or Deletion; CHP, Conserved Hypothetical Proteins; HP, Hypothetical Proteins; **HXAC** or **HXCC**, homologous; **UXAC** or **UXCC**, Unique; **Int**, integrases. #Plasmid-related protein.

drogenases (XAC0918-XAC0924) was named XaUC2 and has two *pntA* subunits (both with frameshifts) that promote the reversal conversion of NAD in NADP, and two *pntB* subunits that promote the transhydrogenation of NADH and NADPH. XaUC6 is formed by 31 genes and has a unique hemolysin-type calcium binding protein, placed side by side with *hlyD* and *hlyB* genes, coding for the T1SS machinery. Other eleven **XaUC_n** regions are clusters of unique genes composed almost entirely of hypothetical proteins, always accompanied by phage proteins or transposases and usually flanked by tRNAs (Table 3).

Lima et al. (2005), based on atypical GC content and codon bias, described five other unique regions in XAC. Although no genes in these five clusters presented relevant functions, some isolated genes have known

TABLE 4. NUMBER OF UNIQUE GENES AND FAMILIES FOR ALL INTERSECTIONS OF THE FOUR GENOMES

		XAC		XCC		Xf-9a5c		Xf-temecula	
XAC	P	198	208	2046		16		2	
	H	453	457	1590		31		0	
XCC	P	752		188	215	2		6	
	H	718		335	340	6		2	
Xf-9a5c	P	8		1		74	82	419	
	H	15		3		165	179	607	
Xf-temecula	P	1		2		147		74	82
	H	0		1		228		30	36

Gray background cells show the number of unique gene families per genome or genome intersection. White background cells show the number of unique genes. P, genes with a putative function; H, hypothetical and conserved hypothetical genes.

functions, like an antirestriction protein that inhibits both restriction and modification by each of the four type I restriction systems in *Escherichia coli* (Delver et al., 1991), two drug resistance genes (*matE*) that mediate resistance against specific drugs, two genes related to pectin degradation, an important step in plant tissue colonization and pathogenicity, and two components of the type V autotransporter secretion system involved in translocation of specific proteins across the outer membrane via a transmembrane pore.

Among scattered unique genes of XAC, we highlight the gene that codes for 2-keto-3-deoxy-D-gluconate transport system, named *kdgT* (XAC0337), involved in uptake and transport of degraded pectin into the cell (Fig. 2, square 9C), which is placed near two copies of ISXac3 transposases. Although XCC does not have a copy of *kdgT*, genes flanking *kdgT* in XAC show a conserved order in XCC. Two genes involved in phytoene metabolism (XAC2744 and XAC3594), an important precursor of carotenoid biosynthesis, were also found in XAC and may be components of a photoprotection system that is not mediated by xanthomonadins. Three copies of XAC peptidases have homology to pseudomonapepsin and xanthomonapepsin carboxypeptidases, involved in oligopeptide absorption and degradation (da Silva et al., 2002) (Fig. 2, square 8C). XAC also contains one hrp-associated gene, *hpaF* (XAC0391), which is part of T3SS and is not found in XCC.

XCC-specific genes. As was done above for XAC, we named genomic regions in XCC enriched with unique genes as **XcUC_n** (Table 5). XcUC1 (XCC0319–XCC0347) includes a cluster of genes that code for proteins involved in xanthomonadin biosynthesis. This region has a total of 29 genes, 20 of which are found only in XCC, among the four bacteria. Six genes between XCC0332 and XCC0342 align to the amino terminal region of the *Xanthomonas oryzae* pv. *oryzae* *pigH* gene, with high similarity, suggesting a tandem duplication event. XcUC3 has 10 unique genes involved in biosynthesis of LPS O-antigen and the LPS core (XCC0599–XCC0618), previously described by Vorholter et al. (2001). According to these authors, knock-out of genes in this region may impact on XCC pathogenic interactions with the host. Interestingly, this cluster is positioned downstream of the *xanAB*, *rmcABCD* and *ispJI* gene clusters, all related to LPS biosynthesis (Steinmann et al., 1997; Vorholter et al., 2001). Comparing XcUC3 with the genome of XAC, we note that a unique XAC gene cluster, named as XAC5 by Lima et al. (2005), is placed at a locus homologous to XcUC3 and has two ABC transporters (*rmd* and *gmd*) in common (Fig. 3a). Other genes that are inserted into XAC5 have functions that might be related to LPS biosynthesis. Therefore, a detailed functional analysis of this cluster in XAC may reveal new genes involved in LPS biosynthesis (Fig. 3b).

XCC has a cluster of seven genes that may be required to assimilate and convert nitrate and nitrite into ammonium (da Silva et al., 2002) (Fig. 2, squares 7I to 7K). This cluster is placed at the end of replication and upstream of other three XCC unique regions (XCC2, XCC3, and XcUC7) (da Silva et al., 2002). XcUC7 and XcUC9 have a total of four copies of avirulence/effector protein-coding genes, which are very important to plant-host interaction and mediate bacteria colonization of compatible plants. Another important unique gene cluster is composed by three enzymes of the type III restriction-modification system (Fig. 2,

TABLE 5. *XANTHOMONAS* UNIQUE GENE FAMILIES, ORGANIZED BY PUTATIVE FUNCTION

<i>Product</i>	<i>XAC</i> ^a	<i>XCC</i> ^b	<i>No. of families</i> ^c	<i>Largest</i> ^d
50S Ribosomal proteins	6	6	6	—
ABC transporters	6	6	6	—
Acetyltransferases	8	8	8	—
Acyl proteins	21	20	13	683 (5)
Chemotaxis proteins	35	36	12	1029 (19)
Conserved Hypothetical proteins	775	784	713	5393/1329 (9)
Cytochromes	17	19	13	132/133 (3)
Dehydrogenase	7	6	5	2112 (3)
Flagellar proteins	30	30	30	—
Glutathione S-transferase	12	9	9	1254/1743/4689 (2)
Glycosyl proteins	8	8	8	—
Gum proteins	4	4	4	—
Hpa/Hrc/Hrp	27	26	23	1104/1120/717/409 (2)
Hydrolase	12	12	11	4994 (2)
Hypothetical	22	10	5	51/5392 (7)
Inner membrane protein	4	6	4	5210 (3)
Integral membrane protein	5	5	5	—
Integrase	4	2	2	5236 (3)
Transposase	85	70	5	161/4727 (21), 2670 (17)
Membrane protein	7	6	6	5727 (2)
MFS transporter	11	12	9	1884/1902/2097 (2)
Molybdenum/molybdopterin proteins	9	9	8	1656 (2)
Outer membrane protein	22	17	10	2025 (9)
Oxidoreductase	14	16	11	1266 (3)
Phage-related protein	15	14	14	4912 (2)
Protocatechuate degradation	4	4	4	—
TonB receptors/like proteins	12	14	6	5486/5151 (4), 2396 (3)
Transcriptional regulator	56	55	49	1952 (4)
Two-component system	14	14	10	4958/432 (2)
Type II secretion system protein	8	8	8	—
Vanillate degradation	4	4	3	3488 (2)
VirB proteins	8	10	6	5446 (5)
Total	1272	1250	1026	—

^aNumber of *X. axonopodis* pv. *citri* genes.

^bNumber of *X. campestris* pv. *campestris* genes.

^cNumber of orthologous/paralogous gene families.

^dGene family id for this function with largest number of paralogous genes (the number of genes is given in parenthesis immediately after the gene family id; when more than one family has the same number of genes, the families are separated by slashes).

The grouping was based either on the predicted protein participation in metabolic pathways and/or the predicted protein molecular function. Lines in bold type highlight systems discussed in the text.

squares 12G and 13G), which are inserted into the conserved region number 9 of XCC and XAC genomes, described previously by Lima et al. (2005). The three genes of this restriction-modification system are the only CDSs inside this region that are absent from XAC. This system is composed of a helicase subunit (XCC1067), whose gene is located upstream of a possible DNA methylase (XCC1068), and another helicase (XCC1069), which may be involved in endonucleolytic cleavage of DNA, recognition of specific short DNA sequences and cleavage of sites distant from the recognition sequence.

Previous analyses involving XCC and XAC revealed other five regions unique to XCC (Lima et al., 2005). The most interesting among these clusters is XcUC7 (region 4 of Lima et al.), which includes several transposases and a gene that codes for a tannase protein (Fig. 2, square 8I to 8K). This protein is

important to prevent damage induced by plant tannins produced as a response to infection (da Silva et al., 2002), and is located near some transposases and phage-related genes. XCC PinDel-4 is located inside XcUC7, and is one of the regions in XCC with largest deviations in GC content and codon bias (Fig. 1).

Among the scattered unique genes of XCC an *avr* protein, named *avrXccA2* (XCC2396), is important to host-plant interactions (Table 5). Other genes include a putative malic acid transporter (XCC3392), a gene coding for a protein related to nisin (bacteriocin) resistance (XCC3409), and a virulence associated protein (XCC4197).

Xf-9a5c-specific genes. *Xf-9a5c* has five large chromosomal unique regions with more than eight genes and at least 95% of hypothetical or conserved hypothetical genes, always flanked by phage proteins. *Xf-9a5c* has 20 genes that code for conjugal transfer proteins located in two unique regions. Eleven are in the pXF51 plasmid and nine are in the chromosome (XF2048-XF2079).

The daunorubicin C-13 ketoreductase is a unique *Xf-9a5c* gene involved in antibiotic production in *Streptomyces peucetius*. There is also a nickase protein similar to a *virD2* protein from an *Agrobacterium tumefaciens* plasmid, which is a site-specific endonuclease, and a plasmid maintenance protein responsible for stable maintenance of the plasmid during cell division, both located at the pXF51 plasmid. Finally, a type IV pilin involved in competence, an adherence associated pilin and a type I restriction-modification system specificity determinant protein, involved in restriction of exogenous DNA sequences, are also genes unique to *Xf-9a5c* (Fig. 2, squares 4F and 5F).

Xf-temecula-specific genes. Among the genomes compared, *Xf-temecula* presents the smallest number of unique genes (Table 4). *Xf-temecula* has four unique regions that deserve attention. Three genes, including a site-specific DNA-methyltransferase involved in restriction and modification of DNA sequences, compose the first region (Fig. 2, squares 4G and 5G). The second region (PD0906–PD0951), is the largest region unique to *Xf-temecula* with 30 unique CDSs, including 16 phage-related proteins and 14 hypothetical proteins. The third region comprises four genes, a proteic killer suppression protein, involved in regulation of toxin activity mediated by specific plasmids, a virulence-associated protein (*vapI*), related to integrative plasmids, and a *HicA*-related protein that in *Haemophilus influenzae* is inserted into the *Hif*-contiguous pilus cluster. The fourth region has a phage-related endolysin, a type II restriction enzyme (*nspV*), and its respective methylase (PD1667-PD1668), which are involved in site-specific restriction and modification (Fig. 2, squares 4G and 5G).

Xf-temecula has two unique genes that are involved in cell wall formation: a UDP-*N*-acetylmuramate-L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase, which catalyzes the last step in the murein tripeptide recycling pathway, and a UDP-*N*-acetylglucosamine-*N*-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol *N*-acetylglucosamine transferase, that catalyzes the last step of peptidoglycan biosynthesis.

Genes specific for several genomes

Genes unique to Xanthomonas with respect to Xylella. Among the *Xanthomonas* genes arranged in clusters, we have found genes coding for chemotaxis proteins, flagellar structural units, copies of type II and III secretory systems and iron receptors. Both *Xanthomonas* possess a complete set of genes involved in synthesis and regulation of flagella (Fig. 2, squares 15E to 15H, and 16E to 16H) (da Silva et al., 2002; Moreira et al., 2004). Ten XAC and nine XCC copies of a methyl chemotactic protein gene (*mcp*), called *tsr*, are located in a tandem duplication, comprising about 15 Kb. The *tsr* gene is involved in capture and transference of chemotactic signals from the environment to the bacterial cytoplasm, and responds to presence of the amino acids serine, alanine or glycine (Stock and Surette, 1994; Stock et al., 1994; Stock et al., 1992). Phylogenetic analysis of *mcp* and *tsr* genes from XAC, XCC and some other bacteria have shown that, although all copies of *tsr* in this region belong to a single clade, some *mcp* genes from outside this region also cluster together with high bootstrap support values in both genomes. Therefore, these genes must have been duplicated before divergence of XAC and XCC lineages (Fig. 4) (Moreira et al., 2004).

The type II secretory system (T2SS) is responsible for the export of exoenzymes essential to host colonization in plant pathogens (Housby et al., 1998; Lee et al., 2004; Lee et al., 2000; Thomas et al., 1997). In *Xanthomonas*, two copies of this system were described (Fig. 2, square 11C and 11J), while for *Xylella* strains just the copy known as xps was found (Fig. 2, square 3C and 3J) (Moreira et al., 2004). Both T2SS clusters of *Xanthomonas* genomes are located in the largest genomic rearrangements between XAC and XCC genomes (da Silva et al., 2002), near to genes coding for XerC and XerD recombinases (Fig. 1). A phylogeny for the two copies of T2SS (Fig. 5) suggests that divergence between these two operons is ancient, dating at least to the ancestor of *Xanthomonadaceae* (Moreira et al., 2004).

The type III secretory system (T3SS) is a molecular device of gram-negative bacteria specialized in delivery of effectors proteins across the membrane barrier of compatible hosts (Collmer et al., 2000). Both *Xanthomonas* strains have genes that code for T3SS, known as *hrp* genes (Fig. 2, square 11C and 11J) (da Silva et al., 2002). There are nine XAC and thirteen XCC *hrp* genes, nine XAC/XCC *hrp*-conserved genes (*hrc*), six XAC and five XCC *hrp*-associated genes (*hpa*) and one unknown gene unique to XAC. Copies of T3SS in XAC and XCC are located at non-equivalent positions, that is, they are flanked by non-homologous regions at clearly different genome loci (Fig. 1). *Xylella* strains do not have a copy of this system (Simpson et al., 2000; Van Sluys et al., 2003) (Fig. 2); they also lack a cluster that codes for poly- and oligosaccharide (DePo) degradation, which flanks the T3SS cluster on both *Xanthomonas* (Fig. 1) (Moreira et al., 2004).

Besides the complexes and genes cited above, the two *Xanthomonas* present some other unique gene clusters that may be related to adaptation and pathogenesis. A unique cluster related to uptake and metabolism of carbon sources codes for a C4-dicarboxylate transport system and two tandem copies of 4-hydroxy-2-oxoglutarate-aldolase/2-dehydro-3-deoxyphosphogluconate aldolase, related to degradation of D-galactonate, might provide an alternative route for carbon uptake. A complete PTS fructose system, related to uptake and metabolism of fructose, is present, including the subunits IIC and IIB, which are absent in *Xylella*, and may easy entrance of sugars into the cell (da Silva et al., 2002) (Fig. 2, squares 9C, 9D, 9I and 9J).

Two systems related to cytochrome biosynthesis are also unique to *Xanthomonas*: a cluster related to cobalamin and pyrroloquinoline quinone (PQQ) biosynthesis, and three clusters coding for cytochrome D ubiquinol oxidase I (*cydA*) and II (*cydB*) subunits. Cytochrome D ubiquinol oxidase was implicated in maintaining low intracellular oxygen concentrations as a requirement for eventual nitrogen fixation in nitrogen-fixing bacteria (Kelly et al., 1990). The *Xanthomonas* species also have two unique clusters possibly involved in osmotic stress: a potassium-transporting ATPase cluster, a cluster involved in putrescine binding, and a transport protein required for the periplasmic transport of putrescine. There are also three copies of microcystin dependent proteins, in tandem, which are related to inhibition of protein phosphatases (Sivonen et al., 1992). Microcystin is a nonribosomally produced cyclic heptapeptide found in toxic strains of *Microcystis*, *Anabaena*, *Nostoc*, and *Oscillatoria* (Sivonen et al., 1992).

FIG. 2. Comparative view of the Xanthomonadaceae biological processes. Comparative view of the biological processes involved in the lifestyle of Xf-9a5c (A), Xf-temecula (B), XAC (C), and XCC (D). Horizontal and vertical black lines separate each organism's biological processes according to the environments in which they live. Using the horizontal axis, it is possible to compare same genus organisms. The central circle represents the main chromosome and within this circle is shown the central metabolism found in these bacteria. The central circle border alludes to the main chromosome DNA molecule, and some unique systems were plotted on it, such as the SPI-7 island (Fig. 7) and LPS synthesis genes (Fig. 3).

Unique *Xanthomonas* genes were also found scattered in both genomes. Most of these genes have a regulatory function and may be directly involved in regulation of pathogenicity and adaptation. Among them, we highlight a large number of genes involved directly or indirectly with sugar metabolism: a copy of fructose-1,6-bisphosphatase, an important enzyme in the gluconeogenesis pathway; *kdgA* gene (KDPG and KHG aldolase), which codes a key enzyme in the Entner-Doudoroff pathway (Fig. 2, squares 9D and 9J); two xylose repressor genes involved in regulation of the *xylBAFGHR* operon; a *tldD* gene that suppresses the inhibitory activity of the carbon storage regulator (*csrA*), possibly affecting glycogen biosynthesis, gluconeogenesis, cell size and surface properties; two copies in XAC and a copy in XCC of 2-keto-3-deoxygluconate kinase, that catalyzes phosphorylation of the first common intermediate in the D-glucuronate and D-galacturonate catabolic pathways. Besides genes for cell-wall degradation, sugar uptake and metabolism, the two *Xanthomonas* genomes also have two subunits of gum genes that are absent in *Xylella* (da Silva et al., 2001).

The two *Xanthomonas* genomes have unique genes related to the uptake of nitrogen, such as an ABC transporter amino acid permease, which eases uptake and processing of nitrogen, associated to a nitrogen regulatory IIA protein and a fumarate and nitrate reduction regulatory protein, which may have a role in linking carbon and nitrogen assimilation. Both genomes also have two copies of a sulfate permease that may provide alternative routes of energy production based on sulfur metabolism.

Both *Xanthomonas* present a large number of genes related to adaptation after the plant response to infection. Eight and six copies of chloroacetaldehyde dehydrogenase were found on XAC and XCC respectively. This gene is involved in detoxification of chloroacetaldehyde to chloroacetic acid, and therefore is important to adaptation and colonization of mesophyll. Genes related to vanilate, protocatechuate, hydroxybenzoate and benzoate degradation, all phenolic compounds, also exist and are directly involved in defense against plant chemical defenses produced during systemic acquired resistance (SAR) (Fig. 2, squares 14B to 14E and 14H to 14K).

The two *Xanthomonas* genomes studied here have a large number of genes coding for efflux and influx pumps, and genes related to detoxification and antibiotic resistance. Among the unique genes that mediate the interchange of compounds between the cytosol and the environment, we highlight nine copies in XAC and six copies in XCC of outer membrane proteins homologous to an efflux system protein called *nodT*, a component of multidrug efflux pump; four copies of a multidrug resistance efflux pump (*fusE*), responsible for export of toxins; a gene that codes for a toxin secretion ABC transporter ATP-binding protein; genes associated with iron storage and detoxification (bacterioferritin and bacterioferritin-associated ferredoxin); and genes related to the synthesis of choline, betaine and taurine, compounds involved in osmotic control.

Both *Xanthomonas* present a large number of genes that code for different iron receptors, with 63 genes in XAC and 64 genes in XCC (Fig. 2, squares 8C and 8J). Among these genes, there are receptors for specific siderophores, such as: pyoverdine, enterobactin and pseudobactin; and for other compounds that also chelate iron, such as hemin, cobalamin and citrate (Braun and Braun, 2002; Buyer and Leong, 1986; Cor-

←
FIG. 4. Unrooted phylogenetic tree inferred for *tsr/mcp* genes from XAC, XCC, and other bacteria. *Xanthomonas* genes are identified by their ordered locus names (XAC#### or XCC####), as available in the GN field from UniProt. These genes form a monophyletic group at the top, and most *tsr* genes in this group are present in orthologous pairs and arranged in tandem in both *Xanthomonas*, thus implying a lineage specific expansion of *tsr* genes before the divergence of XAC and XCC. Selection of homologs was made requiring all sequences to be BLAST hits of XCC1869 with at least 60% coverage and e-value of $\leq 10^{-10}$. Phylogeny inference was performed using maximum likelihood and assuming a Poisson model of evolution. Numbers close to internal branches indicate bootstrap support values. Sequences from other bacteria are identified by their SwissProt accession numbers (four alphanumeric + underscore + five letters) or by a TrEMBL ID and SwissProt species code. Species codes are *Agrobacterium tumefaciens* (strain C58/ATCC 33970) (**AGRT5**); *Bradyrhizobium japonicum* (**BRAJA**); *Chromobacterium violaceum* (**CHRVO**); *Geobacter sulfurreducens* (**GEOSL**); *Halobacterium cutirubrum* (**HALCU**); *Halobacterium sp.* (strain NRC-1/ATCC 700922 / JCM 11081) (**HALN1**); *Halobacterium salinarum* (**HALSA**); *Methanosarcina mazei* (**METMA**); *Nitrosomonas europaea* (**NITEU**); *Pseudomonas aeruginosa* (**PSEAE**); *Rhizobium leguminosarum* (**RHILE**); *Rhodopseudomonas palustris* (**RHOPA**); *Vibrio cholerae* (**VIBCH**) and *Wolinella succinogenes* (**WOLSU**).

nelis and Matthijs, 2002; Khalil-Rizvi et al., 1997; Koster et al., 1995). The *Xylella* genomes have significantly fewer of these receptors. Six genes for iron receptors are found in Xf-temecula and four genes were described in Xf-9a5c (Fig. 2, squares 6C and 6J).

Genes unique to Xylella with respect to *Xanthomonas*. Although *Xylella* has fewer unique genes than *Xanthomonas* (Table 4), some important systems are found among them. Type IV pili are filamentous structures placed at bacterial polar regions and are responsible for movements called twitching motility on adhesion surfaces (Alm and Mattick, 1997; Strom and Lory, 1993). Both *Xylella* and *Xanthomonas* have genes that code for components of this structure, but *Xylella* shows a higher number of copies of such genes, including several copies of the structural subunit gene *pilE* and, most importantly, two copies of the cluster *pilE-fimT* (Moreira et al., 2004) (Fig. 2, squares 2F and 2G). Phylogenetic analysis suggests the presence and divergence of two copies of the *pilE-fimT* cluster in the ancestor of *Xanthomonadaceae*, both copies being preserved in *Xylella* (Fig. 6), while one copy was lost in XAC and the other in XCC, thus implying independent losses after divergence of the two *Xanthomonas* lineages (Fig. 2, squares 15D and 15I). The *pilE-fimT* clusters in XAC (XAC0974-XAC0984) and XCC (XCC2488-XCC2495) have low %GC (54%) when compared to the genome average (65%) and to GC content in its flanking regions, which are rich in phage and transposase insertions.

Among the several unique genes found scattered in *Xylella* genomes, those involved in biosynthesis and secretion of colicin and the gene that codes for arginine deaminase deserve special attention. Colicins are small proteins that promote bacterial antagonism during nutritional limitation (Lazdunski et al., 1998). *Xylella* strains have copies of all genes necessary for colicin biosynthesis, but both *Xanthomonas* strains have no copy of the *cvaC* and *cvaA* genes and the *cvaB* gene is present only in XAC (Fig. 2, squares 2C, 2D, 2I, 2J). The absence of *cvaA* gene in XAC is paralleled by the insertion of two transposases that might have mediated rupture and depletion of the *cvaAB* cluster. In XCC, both colicin genes are absent and there are no transposases. The genes *cvaABC* are regulated by the *cvpA-purF* gene cluster, which was first identified in *Escherichia coli* plasmid pColV-K30 (Fath et al., 1989; Waters and Crosa, 1991). All four phytopathogens have copies of the *cvpA-purF* cluster, showing a high degree of similarity to *E. coli* genes, but located in the chromosome (Fath et al., 1989; Lazdunski et al., 1998; Waters and Crosa, 1991). A recent *in vitro* approach revealed that the colicin genes are functional and are expressed under high sugar concentration (Pashalidis et al., 2005).

Arginine deaminase (*rocF*) was formerly described as a gene that may define the pathogenic or non-pathogenic phenotype for strains of *Xylella fastidiosa* (Koide et al., 2004). Curiously, this gene is not found in *Xanthomonas* and is located inside a region that presents other seven unique genes (XF1232-XF1255 and PD0508-PD0521), most of them hypothetical or conserved hypothetical proteins. This gene may be important against NO produced by the plant.

FIG. 5. Unrooted phylogenetic tree for the concatenated alignment of subunits D, E, F, G, H, and K of the T2SS gene cluster. The *xps* clusters found in XAC and XCC are homologous to the T2SS from *Xylella* (Xf-9a5c and Xf-temecula) and belong to a well-supported clade together with other plant-associated bacteria, while *xcs* is more related to secretory systems found in enterobacteria. *Leptospira* (LEPIN) was used to orient the tree in order to separate the systems found in entero- and plant-associated bacteria. Selection of homologous sequences was made based on KEGG and COG orthologous groups. Phylogenies were inferred using maximum likelihood and a Poisson model for aminoacid substitution. Numbers close to internal branches indicate bootstrap support values. Codes for species are XAC—*Xanthomonas axonopodis* (pv. citri); XCC—*Xanthomonas campestris* (pv. campestris); XF-9a5c—*Xylella fastidiosa*; XF-temecula—*Xylella fastidiosa* (strain Temecula1/ATCC 700964); BRAJA—*Bradyrhizobium japonicum*; CAUCR—*Caulobacter crescentus*; CHRVO—*Chromobacterium violaceum*; ECOLI1—*Escherichia coli* CFT073; ECOLI2—*Escherichia coli* K-12 MG1655; LEPIN—*Leptospira interrogans*; RHIL0—*Mesorhizobium loti*; PSEAE—*Pseudomonas aeruginosa*; PSEPK—*Pseudomonas putida* (strain KT2440); PSESM—*Pseudomonas syringae* (pv. tomato); RALSO—*Ralstonia solanacearum*; SHEON—*Shewanella oneidensis*; VIBCH—*Vibrio cholerae*; VIBPA—*Vibrio parahaemolyticus*; VIBVU—*Vibrio vulnificus*; VIBVY—*Vibrio vulnificus* (strain YJ016); YERPE1—*Yersinia pestis* CO92; YERPE2—*Yersinia pestis* KIM. Ending numbers or letters after subscripts highlight different copies of the system in a single organism.

Unique genes in XAC and Xf-9a5c/Xf-temecula with respect to XCC. Eighteen of these genes are located inside PinDel-8 of Xf-9a5c and PinDel-4 of XAC, comprising a region that is homologous to the SPI-7 pathogenicity island from *Salmonella thiphymurium* (Fig. 2, squares 8E and 6E) (Moreira et al., 2004; Pickard et al., 2003). Detailed analysis of the SPI-7 region in both genomes suggests probable lateral transfer events in each lineage inside this region: insertion of 37 genes comprising PinDel-7 in Xf-9a5c (XF1718-XF1754), which is located upstream of SPI-7 in Xf-9a5c and is flanked by a tRNA-G, and an added cluster of fourteen genes, including an integrase (XAC2220), inside the SPI-7 region in XAC (Fig. 7). Our analysis has shown that, in Xf-9a5c, 13 genes in PinDel-7 and 14 genes in SPI-7 show atypically high GC content (68%) when compared to the Xf-9a5c genome average (53%). Their homologs in the region SPI-7 from XAC have GC content very close to the average (65%) for XAC genes. In spite of this, and given the absence of SPI-7 both in XCC and Xf-temecula, plus the other evidence presented above, we speculate that the presence of SPI-7 in XAC and Xf-9a5c resulted from two separate lateral transfer events; this is consistent with Pickard et al.'s hypothesis (Pickard et al., 2003) that SPI-7 may be a mobile element.

Other CDSs in the intersection of XAC and Xf-9a5c gene sets include three conserved hypothetical proteins: a transcriptional regulator, a plasmid stabilization protein, and a cytochrome B561-like gene. A single gene was classified as unique to the intersection between the XAC and Xf-temecula gene sets: CDP-diacylglycerol-glycerol-3-phosphate-3-phosphatidyltransferase, which catalyzes a committed step to the synthesis of the acidic phospholipids, and is located near transposases and phage-related genes in Xf-temecula.

Genes unique to XCC and Xf-9a5c/Xf-temecula with respect to XAC. Four genes are unique to XCC and Xf-9a5c gene sets: three conserved hypothetical proteins and a phage-related protein. Only three genes are unique to XCC and Xf-temecula: two phage-related proteins, which are duplicated in XCC, and one hypothetical protein.

Xanthomonadaceae unique genes. These were selected based on published results. Among the most important of the genes unique to the *Xanthomonadaceae* are those responsible for the synthesis of DFS (Diffusible Signal Factor) signaling molecules, the *rpf* cluster (regulation of pathogenicity factors) (Barber et al., 1997; Tang et al., 1991). Inside the *rpf* operon, XAC differs from XCC because it does not have the *rpfH* and *rpfI* genes. The insertion of two transposases between the genes *recJ* and *prfB* may have replaced *rpfI* (Dow et al., 2000), accompanied by the insertion of a gene coding for a wall-associated protein (*wapA*) of the family *rhs*, composed by nine copies of the *rhs* repeat (Apweiler et al., 2000). The *rhs* domain can act as a recombination hot spot (Minet and Chiquet-Ehrismann, 2000) and was related to duplications and rearrangements in *E. coli* (Lin et al., 1984). *Rpf* clusters in both *Xylella* strains lack copies of the *rpfI*, *rpfH* and *rpfD* genes, in addition showing a rupture of this cluster between genes *rpfF* and *rpfB* (da Silva et al., 2001; Moreira et al., 2004). *Rpf* genes also regulate expression of gum genes, responsible for the synthesis and exporting of the xanthan gum, an extracellular polysaccharide important to pathogenesis (Vojnov et al., 2001). Gum genes from *Xanthomonas* strains show a high degree of similarity, with 98% of identity for pairs of orthologs, while in *Xylella*, identities between genes from different pathovars decrease to values between 65 and 83%, with absence of the genes *gumGIL*, responsible for joining and modification of mannos residues (da Silva et al., 2001).

FIG. 6. Unrooted phylogenetic tree for the genes *pilE* and *fimT* found in XAC, XCC, Xf-9a5c, and Xf-temecula, based on their encoded aminoacid sequences. Horizontal lines separate copies of *pilE* located at the clusters *pilBD* or *pilE_fimT* and the copies of *fimT*. All unique copies of *pilE* found in Xf-9a5c and Xf-temecula are members of a cluster that includes all copies of *pilE* from the *pilB-pilD* cluster in the four genomes. Note that *pilE* is also duplicated inside the cluster *pilBD* in all lineages. This phylogeny was built using maximum likelihood and a Poisson model of aminoacid substitution. ORFs were selected based on the presence of a conserved amino-terminal pilin signature, and some truncated sequences were later discarded. Although the resulting alignment, after selection of conserved regions, was too short (40 aa), we inspected phylogenies based on different selection of homologs (using the procedures described in Methods for isolated genes) and observed that this phylogeny reproduces all important features obtained when using longer alignments.

FIG. 7. Analysis of the region comprising PinDel-6 of XAC and PinDel-5 of XF-9a5c, which corresponds to island SPI-7 described for *Salmonella typhimurium* (Pickard et al., 2003). The region between genes XF1718 and XF1754 corresponds to PinDel-4 in XF-9a5c. Gray rectangles, genes homologous to genes in the SPI-7 island; white rectangles, genes with function not related to this analysis; black rectangles, phage integrases; black arrows, tRNA genes. Genes with atypically high GC content in Xf-9a5c.

The xanthomonadins cluster is another group of genes specific to the *Xanthomonadaceae* (Starr et al., 1977). These genes code for proteins involved in the synthesis of pigments useful for protection against photobiological damage (Jenkins and Starr, 1985; Poplawsky and Chun, 1997; Poplawsky et al., 2000; Rajagopal et al., 1997). XCC and XAC xanthomonadin clusters have the same structure described for *Xanthomonas oryzae* pv. *oryzae* (XOO) (Goel et al., 2002) (Fig. 2, squares 14E and 14I), except for the first two genes, located away from the xanthomonadin operon in both genomes. Upstream to XAC and XCC genes homologous to genes 3-12 in XOO (referred to as *large cluster* in Fig. 8), a new cluster of genes, which we call the *small cluster*, may be associated with xanthomonadins biosynthesis, because genes in this relatively conserved cluster are related to fatty acids and phosphatidic acid biosynthesis. An interesting feature shown in Figure 8 is the extensive modifications of the region containing the two clusters in XOO. The entire small cluster was either lost or translocated in XOO, and the same happened to isolated genes around and inside the large cluster. In particular, the genes XAC4094, XAC4095, and XCC4006 are truncated copies of the N-terminal region of the homologous XAC4093, XCC4005 and the CDS 4 from XOO. The gene flanking the xanthomonadins cluster in XOO, CDS14, is homologous to a gene located in the XcUC1 region, close to the tandem duplication of xanthomonadins-related genes inside this region.

In both *Xylella* genomes, the small cluster is located far downstream from the large cluster. The small cluster is flanked by phage XFP4 in Xf-9a5c and by PinDel-6 in Xf-temecula. The breakpoint for the separation of the two clusters in both *Xylella* genomes corresponds to the location of the repetitive non-coding DNA sequences in XOO. Three CDSs in both Xf-9a5c clusters have frameshifts and therefore are considered to be pseudogenes, suggesting degeneration of this system (Moreira et al., 2004). Homologs of these pseudogenes in Xf-temecula show no frameshifts.

Finally, although genes *tonB*, *exbB*, and *exbD*, responsible for iron uptake, form a cluster in many bacteria and a second copy of *exbD* (*exbD2*) is common, the *Xanthomonadaceae* are unusual for having all these genes in a single cluster, with *exbD2* located downstream to *tonB-exbB-exbD*. Knock-out of *exbD2* in XCC induces hypersensitive response from the host (Wiggerich and Puhler, 2000), but differently from the other genes in this cluster, it is not essential for penetration of phage Φ L7 and shows no impact on iron uptake (Hung et al., 2003; Wiggerich et al., 1997).

DISCUSSION

Differences in gene content between *Xanthomonas* and *Xylella* complete genome sequences have been characterized previously (Van Sluys, M.A., et al., 2003; da Silva, et al., 2002; Moreira, LM, et al., 2004). The data presented here complements such analyses with an extensive description of important genome rearrangements and a comparison of the putative functions of all genes leading to variations in gene content among the four strains analyzed. As presented above and explored in previous works, all four bacteria studied here share some characteristic systems that are commonly used to identify members of their family, and were shown to have important roles in their survival and adaptation. For instance, the gum operon, the extended xanthomonadin cluster, and the *rpf* genes, although present in all four bacteria, show important variations in gene order and gene content among these organisms, and these variations are likely to be a factor in each strain's phenotype (da Silva et al., 2002).

Following a trend seen in other shared systems, copies of the xanthomonadin cluster in *Xylella* strains are highly fragmented when compared to the structure seen in *Xanthomonas*. In Xf-9a5c, the fragmentation of the cluster is paralleled by the presence of three pseudogenes inside the large and small clusters. Although Xf-temecula has intact copies of the Xf-9a5c pseudogenes, experimental data has shown that both *Xylella* strains do not synthesize the xanthomonadin pigment (Almeida et al., 2004; Almeida and Purcell, 2003; Chang and Donaldson, 2000; Leite et al., 2004). Considering the changes in the xanthomonadin clusters from XAC and XCC to Xf-temecula and Xf-9a5c, the trend represented in Figure 8 that we wish to highlight is an apparent progressive depletion of the xanthomonadin cluster along the evolution of *Xylella*.

A similar fragmentation is observed for the *rpf* gene cluster in *Xylella* (Moreira et al., 2004). It was suggested that the fragmented structure of the *rpf* cluster in *Xylella* might induce a reduction of DSF synthesis, thus leading to reduced quorum sensing signaling and no feedback for regulation of expression of the

FIG. 8. Analysis of genes predicted to be involved with xanthomonadin synthesis in XCC, XAC, XF-9a5c, and XF-temecula, using *Xanthomonas oryzae* pv *oryzae* (XOO) as model, following Goel et al. (2002). The dendrogram on the left shows the 16S phylogeny. Although XOO and XCC are closer phylogenetically, the gene organization of XCC is more similar to that of XAC. This suggests a rearrangement or deletion of the "small cluster" in XOO. A transposase and a repetitive element present upstream of the large cluster in XOO are evidence in favor of this hypothesis. It can be seen in this comparison that the *Xylella* genomes appear to have been subject to many rearrangements in this region, consistent with data shown in Figure 1. This analysis suggests that XF-9a5c phage 4 is related to PinDel-6 in XF-temecula; note also that XOO gene 14 has an ortholog in XcUCI.

rpf genes themselves (Moreira et al., 2004). We note that the hypothesis of Moreira et al., (2004) for the consequences of *rpf* fragmentation is compatible with the relative concentration of DSF in *Xylella* and *Xanthomonas* cultures (Scarpari et al., 2003). Such loss of feedback control over *rpf* genes probably reduces *Xylella*'s virulence, because it lowers the level of chemical signals that activate gum synthesis. With a reduced DSF-mediated quorum sensing signaling to *rpf* and living in a low sugar milieu (the xylem), it is likely that *Xylella* either possesses another system for gum synthesis regulation or express this system constitutively, differently from *Xanthomonas*, which are regulated by sugar concentration (Vojnov et al., 2001). At any rate, such structural differences in the *rpf* operon point to the relevance that changes in gene order might have over the phenotype of bacteria. The presence of a complete PTS system only in *Xanthomonas*, associated to the presence of clusters involved in degradation of plant cell-wall and sugar storage, suggests that *Xanthomonas* might use degraded cell walls as carbon sources, another function probably absent from *Xylella* cells, due to the degradation of the PTS system and a reduced number of cell-wall degrading enzymes (CWDEs) in *Xylella*. The presence of genes encoding the enzyme 2-keto-3-deoxygluconate kinase, along with genes for CWDEs, suggests that *Xanthomonas* may use the CWDEs to obtain energy from host cell-wall degradation.

Some systems were found to be complete in *Xylella* and reduced in *Xanthomonas*, probably due to the existence of negative selection against the loss of such functions, which are probably necessary for adaptation into the xylem. This is likely to be the case for the presence of a complete operon for colicin production in *Xylella*, which is not found in *Xanthomonas*. The lack of both *cvaA* and *cvaB* genes in XCC is probably a derived state, XAC representing an intermediate depletion state, while *Xylella* probably corresponds to the ancestral set of colicin genes, all being preserved because synthesis of these compounds is essential to avoid competition with other endophytic bacteria of the plant or resident at the insect foregut, as proposed by Pashalidis et al. (2005). In fact, it has been demonstrated that Xf-9a5c lives in an environment in citrus plants that is rich in endophytic bacteria (Araujo et al., 2002; Lacava et al., 2004), and therefore the presence of a complete and functional colicin V biosynthesis and secretion system may enhance its chances of survival in the xylem.

Two other important examples of differences in gene content between *Xanthomonas* and *Xylella* are the presence in *Xanthomonas* of a second copy of the type II secretion systems (T2SS), known as *xcs*, and of a type III secretion system (T3SS). These systems differ from the systems discussed above for being entirely absent from *Xylella* genomes, instead of partially depleted. Therefore, although one can argue, for the systems discussed above, that a partially degraded copy of a cluster of genes is enough evidence to support a hypothesis of vertical heritage followed by gene loss, the full absence of both secretion systems from *Xylella* might well be explained by loss of the cluster in the *Xylella* lineage or acquisition by *Xanthomonas* after divergence, through lateral gene transfer. In the case of *xcs*, we believe this system was lost from a common ancestor of Xf-temecula and Xf-9a5c. Two pieces of evidence support this hypothesis: (i) the regions flanking this cluster in *Xanthomonas* are homologous to neighboring regions in *Xylella* (Fig. 1, regions A and B) (Moreira et al., 2004); (ii) the phylogeny shown in Figure 5 indicates that copies of the T2SS distributed among several proteobacteria radiate in a pattern that can be easily reconciled with the species tree, by postulating just a few events of gene duplication and loss (Page and Charleston, 1997). In particular, the copy of *xcs* present in XAC and XCC is placed at the base of the branch that leads to copies of T2SS present in *Pseudomonas* and several enterobacteria, thus reproducing the location of *Xanthomonas* in a species tree (Lerat et al., 2003). On the other hand, the *Caulobacter* T2SS copy is placed in an unexpected location, well inside a γ -proteobacteria group, although this organism belongs to the alpha subdivision of Proteobacteria. Even more suggestive is the grouping of *xps*, the other T2SS cluster found in *Xanthomonas* and *Xylella*, with plant-associated bacteria of the α -proteobacteria subdivision (*Bradyrhizobium japonicum* and *Mesorhizobium loti* in Fig. 5), which suggests that this cluster, instead of *xcs*, has originated from lateral transfer from plant-associated bacteria. We note however that branch lengths in the *xps* subtree leading to *B. japonicum* and *M. loti* are long, thus implicating early divergence of these genes, maybe dating to the divergence among these organisms' lineages. Whatever the origins of *xps* and *xcs*, their phylogeny and sequence alignment show that their divergence is very old, probably dating to the ancestor of all proteobacteria. This suggests that these systems have specialized to perform different physiological roles, and therefore *Xylella* may lack some functionality associated with these systems.

For T3SS, a single BLAST run is able to show how easy it is to find such a system among members of the *Xanthomonas* genus. Sequence and phylogenetic analyses have shown T3SS clusters from members of the β -proteobacteria subdivision to be the closest relatives of the T3SS from XAC and XCC (data not shown). Such a relationship is in agreement with the species tree, but contradicts the fact that the XCC copy of T3SS possesses much lower GC composition and high codon bias when compared to the rest of XCC genes, a feature common to laterally transferred genes, although not unique to them. Loci homologous to regions C and D in *Xylella* (Fig. 1), which flank *Xanthomonas* T3SS, are located near the replication terminus in both *Xylella* strains, a pattern that suggests deletion of this system from *Xylella*. So we cannot, with known evidence, draw conclusions on what is the origin of T3SS found *Xanthomonas*.

Lineage specific expansion and diversification is another evolutionary process involved in variations in gene content between the two genera. This is the mechanism behind the much higher number of chemotactic genes in *Xanthomonas*, and was especially important in the evolution of *tsr* genes. This family of genes has undergone fast expansion in the *Xanthomonas* lineage after separation of the *Xylella* branch (Moreira et al., 2004) (Fig. 4). It is possible that a mechanism of dosage compensation has prompted the very first duplication events of the *tsr* gene in the ancestor of XAC and XCC, as proposed by Heidelberg et al. (2000) in relation to the genes involved in chemotaxis in *Vibrio cholerae*.

The presence of several copies of the structural subunit of type IV pili in *Xylella* may also have involved a similar mechanism of dosage compensation, followed by divergence of some copies (Fig. 6). The presence of two distinct clusters of pili type IV biosynthesis in *Xylella*, as described by Moreira et al. (2004), associated with a complete cluster of chemotaxis genes related to twitching motility, may have a relation to the lifestyle of *Xylella*. Each pili cluster may have a specific function such as, for example, fixation into the xylem vessel and fixation into the insect foregut. In addition, the presence of additional copies of genes for fimbrial adhesins may augment the expression of proteins needed for colonization and cell agglomeration into the plant. A recent analysis of genome composition of pathogenic (CVC) and nonpathogenic *Xylella fastidiosa* isolates showed that the absence of fimbrial adhesins can determine the pathogenic phenotype of an isolate (Koide et al., 2004).

Another mechanism involved in determining variations in gene content among the four bacteria is the process of lateral gene transfer, already referred to when discussing the secretory systems in *Xanthomonas*. Several phage- and plasmid-related genes are present in all four genomes and most certainly were acquired directly from other organisms and phages through conjugation or transfection followed by recombination. We note also that many of the unique genes described for the four genomes that had assigned functions were restriction enzymes, known to be transmitted horizontally (Gelfand and Koonin, 1997; Sekizaki et al., 2001; Sharp et al., 1992). The many transposases seen in *Xanthomonas* might be related to plasmids and therefore be related to the transposition of genes from plasmids into the genome. An important example of a gene likely to have been acquired laterally is the copy of *kdgT*, found only in XAC, and placed near two ISxac3 transposases. This gene is related to pathogenicity, since it is involved in the transport of degraded pectin products in the bacterial cell, where these products may be used as carbon and energy sources.

As a final comment on our detailed characterization of unique gene functions among the two *Xanthomonadaceae* genera analyzed here, some unique genes related to differences in nitrogen metabolism for XCC and XAC and NO resistance in *Xylella* are worth mentioning. The presence of a unique cluster that codes for nitrate assimilation in XCC may have a horizontal origin and may be coupled with the *Xanthomonas* unique cytochrome genes *cydAB*, which have been implicated in nitrogen fixation (da Silva et al., 2002), thus characterizing a unique form of nitrogen uptake among the bacteria analyzed here. In association with the *opp* genes located in the XaUC1 region, the pseudomonapepsin and xanthomonapepsin carboxypeptidases may be part of an oligopeptide absorption and degradation pathway that may be linked to XAC's ability to grow in different environments (da Silva et al., 2002). Since *opp* genes code for components of an ABC permease complex that is associated to an oligopeptide binding protein (*oppA*), it is possible that these genes may help entry of small oligopeptide products for use as a nitrogen source.

In *Xylella* strains, the unique *rocF* gene is homologous to *Helicobacter pylori*'s arginase *rocF*. In *Helicobacter pylori*, this gene inhibits NO production by macrophages at physiological concentrations of arginine. *RocF* expression is aborted at normal levels of NO, indicating that this gene down-regulates NO production, thus acting as a survival mechanism and contributing to the success of infection (Gobert et al.,

2001). Gobert et al. (2002) showed that macrophage apoptosis is induced by activation of the arginase II gene of *H. pylori*. Based on such observations, Koide et al. (2004) proposed that absence of this gene in *Xylella fastidiosa* J1a12, a non-pathogenic strain, is linked to reduced growth in the plant and incapacity to colonize the xylem vessels, due to inability to inhibit NO production by the plant host. On the other hand, the hypothesis of Koide et al. (2004) for *Xylella fastidiosa* J1a12 cannot be extended to *Xanthomonas*, since these bacteria lack *rocF* but grow very well inside the plant and, in the case of *Xanthomonas campestris*, are able to colonize vascular tissues. This suggests that either *Xanthomonas* has other NO detoxification systems or that it has alternative routes to use arginine produced by the plant.

PinDels and phages. We have discussed above evolutionary events that may have lead to the observed variations in gene content among XAC, XCC, Xf-9a5c, and Xf-temecula. Two other important processes were characterized regarding the evolution of these genomes: the rearrangements observed between species in each genus and the distribution of putative insertion/deletion islands, referred to as PinDels.

All four genomes present important rearrangements. The *Xanthomonas* genomes have three important rearrangements with respect to each other: an inversion at the terminus of replication and a reciprocal translocation with inversion of the translocated regions (da Silva et al., 2002). The *Xylella* genomes, on the other hand, exhibit a more varied set of rearrangements with respect to each other, including a translocation of the replication terminus (Van Sluys et al., 2003) (Fig. 1). Probably the most important aspect of the *Xylella* genome rearrangements is the correlation between the borders of shuffled regions and the presence of phages or PinDels near such borders, which indicate a possible involvement of phage integrases in the rearrangements observed in these two lineages, as previously suggested by Van Sluys et al. (2003); no similar connection can be made between phage integrase distribution and rearrangements in *Xanthomonas*.

All strains analyzed show a high number of phage-related genes, many of them located inside unique gene islands rich in hypothetical genes. Note however that *Xylella* strains possess much more phage-related and complete integrated phages than *Xanthomonas*, and that *Xanthomonas* integrases appear to have less activity than in *Xylella*, since the higher number of rearrangements in this lineage seems related to phage integrases. Also, most phage-related genes in *Xanthomonas* are located inside unique regions that seem to be remains of degraded phages. Which mechanism is behind the abundance of complete phages integrated into the genomes of the two *Xyellae*? Multiple alignments of all integrases found in both genomes lead to classification of these integrases into groups based on sequence similarity and integration site class, as described in Materials and Methods. In Xf-9a5c, integrases of group A may be related to insertion on the 3' end of tRNAs (Campbell, 2003; Williams, 2002), whereas the integrases of group B may determine prophage termination. Although no experimental work has been done to verify the relation between the pili machinery in Xf-9a5c and this organism's capacity to acquire phage insertions, pili biosynthesis may be an important mediator of the entrance of phages into *Xylella*, as suggested by Bradley (1972, 1974) for *Pseudomonas aeruginosa*. Also, besides mediating the uptake of iron and other compounds, such as cobalamin and heme groups, the tonB and tolA complexes may be involved in interactions with lytic bacteriophages and the bacteria cell host (Braun et al., 1976), and facilitate entrance of phages into the genome of *Xylella*, thus providing another mechanism through which phage invasion may be facilitated in *Xylella*. The abundance of phage-related genes in *Xylella*, as well as the greater abundance of transposases in *Xanthomonas*, may be related to distinct restriction and modification systems in each bacterium. Interestingly, although our initial approach to identify probable insertions was based on phage integrases and site integration, this approach identified nearly the same regions obtained from an analysis based on GC content and codon bias, as shown in Figure 1.

Although molecular mechanisms that promote a higher rate of phage integration into the genome of *Xylella* might be involved in the presence of a higher number of complete phages and phage-related genes, a population genetics explanation could as well account for such differences. With nearly half the genome size of *Xanthomonas*, it is possible that the *Xylella* genome is evolving under a regime of accumulation of deleterious mutations (Crow and Kimura, 1970), both in the form of degraded gene systems and pseudogenes as well as integrated phages. Together with its limited environmental range, restricted to the xylem, slow growth and fastidious phenotype, *Xylella* presents limitations to its propagation derived from the dependence on an insect vector, while *Xanthomonas* can and usually does spread easily from contaminated

to noncontaminated hosts, being able to survive under adverse conditions over leaf surfaces and on the ground (Lopez et al., 1999). As *Xanthomonas* also have a much smaller doubling time both in plants and the laboratory, we believe that the effective population size for *Xanthomonas* strains is much greater than those for *Xylella* strains. If real, the reduction of effective population size for *Xylella* might be regarded as a mechanism that enables fixation of both gene deletion events, leading to a higher number of depleted systems and pseudogenes, and the integration of phages into its genome. Both gene deletion and phage insertion might be considered deleterious events to bacteria and are expected to be removed from sufficiently large populations by means of natural selection. On smaller populations, though, these mutations might become fixed by random genetic drift (Woolfit and Bromham, 2003).

As a final note, a link between two of the main threads of this paper, namely variation in gene content among phylogenetically close genomes and the observed correlation between phage-related genes and regions rich in unique genes, is provided by recent analyses that show that phage-bacteria interactions are important mechanisms for new gene creation in bacterial hosts (Daubin and Ochman, 2004).

ACKNOWLEDGMENTS

This work was supported in part by Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP (Ph.D. fellowship to L.M.M.), by Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq (Ph.D. fellowship to R.F.S.), and by Coordenação para Aperfeiçoamento de Pessoal de Ensino Superior, CAPES (Ph.D. fellowship to L.A.D.), all of them Brazilian funding agencies.

REFERENCES

- ADACHI, J., and HASEGAWA, M. (1996). *MOLPHY Version 2.3: Programs for Molecular Phylogenetics Based on Maximum Likelihood*. J. Adachi, ed. (Institute of Statistical Mathematics, Tokyo).
- ALMRA, and MATTICK, J.S. (1997). Genes involved in the biogenesis and function of type-4 fimbriae in *Pseudomonas aeruginosa*. *Gene* **192**, 89–98.
- ALMEIDA, R.P., MANN, R., and PURCELL, A.H. (2004). *Xylella fastidiosa* cultivation on a minimal solid defined medium. *Curr Microbiol* **48**, 368–372.
- ALMEIDA, R.P., and PURCELL, A.H. (2003). Biological traits of *Xylella fastidiosa* strains from grapes and almonds. *Appl Environ Microbiol* **69**, 7447–7452.
- ALTSCHUL, S.F., MADDEN, T.L., SCHAFER, A.A., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- APWEILER, R., ATTWOOD, T.K., BAIROCH, A., et al. (2000). InterPro: an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16**, 1145–1150.
- APWEILER, R., BAIROCH, A., WU, C.H., et al. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, D115–D119.
- ARAUJO, W.L., MARCON, J., MACCHERONI, W., Jr., et al. (2002). Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants. *Appl Environ Microbiol* **68**, 4906–4914.
- BARBER, C.E., TANG, J.L., FENG, J.X., et al. (1997). A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol Microbiol* **24**, 555–566.
- BHATTACHARYYA, A., STILWAGEN, S., IVANOVA, N., et al. (2002a). Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. *Proc Natl Acad Sci USA* **99**, 12403–12408.
- BHATTACHARYYA, A., STILWAGEN, S., REZNIK, G., et al. (2002b). Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res* **12**, 1556–1563.
- BRADLEY, D.E. (1972). Shortening of *Pseudomonas aeruginosa* pili after RNA-phage adsorption. *J Gen Microbiol* **72**, 303–319.
- BRADLEY, D.E. (1974). The adsorption of *Pseudomonas aeruginosa* pilus-dependent bacteriophages to a host mutant with nonretractile pili. *Virology* **58**, 149–163.
- BRAUN, V., and BRAUN, M. (2002). Iron transport and signaling in *Escherichia coli*. *FEBS Lett* **529**, 78–85.
- BRAUN, V., HANCOCK, R.E., HANTKE, K., et al. (1976). Functional organization of the outer membrane of *Escherichia coli*: phage and colicin receptors as components of iron uptake systems. *J Supramol Struct* **5**, 37–58.

- BUYER, J.S., and LEONG, J. (1986). Iron transport-mediated antagonism between plant growth-promoting and plant-deleterious *Pseudomonas* strains. *J Biol Chem* **261**, 791–794.
- CAMPBELL, A. (2003). Prophage insertion sites. *Res Microbiol* **154**, 277–282.
- CASTRESANA, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**, 540–552.
- CHANG, C.J., and DONALDSON, R.C. (2000). Nutritional requirements of *Xylella fastidiosa*, which causes Pierce's disease in grapes. *Can J Microbiol* **46**, 291–293.
- CHANG, C.J., GARNER, M., ZREIK, L., et al. (1993). Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and identification as a strain of *Xylella fastidiosa*. *Curr Microbiol* **27**, 137–142.
- COLLMER, A., BADEL, J.L., CHARKOWSKI, A.O., et al. (2000). *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proc Natl Acad Sci USA* **97**, 8770–8777.
- CORNELIS, P., and MATTHIJS, S. (2002). Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. *Environ Microbiol* **4**, 787–798.
- CROW, J.F., and KIMURA, M. (1970). *An Introduction to Population Genetics Theory* (Harper and Row, New York).
- DA SILVA, A.C., FERRO, J.A., REINACH, F.C., et al. (2002). Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**, 459–463.
- DA SILVA, F.R., VETTORE, A.L., KEMPER, E.L., et al. (2001). Fastidious gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *FEMS Microbiol Lett* **203**, 165–171.
- DAUBIN, V., and OCHMAN, H. (2004). Start-up entities in the origin of new genes. *Curr Opin Genet Dev* **14**, 616–619.
- DELVER, E.P., KOTOVA, V.U., ZAVILGELSKY, G.B., et al. (1991). Nucleotide sequence of the gene (ard) encoding the antirestriction protein of plasmid collb-P9. *J Bacteriol* **173**, 5887–5892.
- DIGIAMPIETRI, L.A., MEDEIROS, C.B., and SETUBAL, J.C. (2003). A data model for comparative genomics. *Rev Tecnol Inform* **3**, 35–40.
- DOW, J.M., FENG, J.X., BARBER, C.E., et al. (2000). Novel genes involved in the regulation of pathogenicity factor production within the *rpf* gene cluster of *Xanthomonas campestris*. *Microbiology* **146**, 885–891.
- FATH, M.J., MAHANTY, H.K., and KOLTER, R. (1989). Characterization of a *purF* operon mutation which affects colicin V production. *J Bacteriol* **171**, 3158–3161.
- GELFAND, M.S., and KOONIN, E.V. (1997). Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes. *Nucleic Acids Res* **25**, 2430–2439.
- GOBERT, A.P., CHENG, Y., WANG, J.Y., et al. (2002). *Helicobacter pylori* induces macrophage apoptosis by activation of arginase II. *J Immunol* **168**, 4692–4700.
- GOBERT, A.P., MCGEE, D.J., AKHTAR, M., et al. (2001). *Helicobacter pylori* arginase inhibits nitric oxide production by eukaryotic cells: a strategy for bacterial survival. *Proc Natl Acad Sci USA* **98**, 13844–13849.
- GOEL, A.K., RAJAGOPAL, L., NAGESH, N., et al. (2002). Genetic locus encoding functions involved in biosynthesis and outer membrane localization of xanthomonadin in *Xanthomonas oryzae* pv. *oryzae*. *J Bacteriol* **184**, 3539–3548.
- GOMEZ, P., RIBAS-APARICIO, R.M., PELAEZ, A.I., et al. (1997). Isolation and nucleotide sequence of the gene encoding the XamI DNA methyltransferase of *Xanthomonas campestris* pv. *amaranthicola*. *Biochim Biophys Acta* **1351**, 261–266.
- GOTTWALD, T., and GRAHAM, J. (2000). Bacteria diseases. In *Compendium of Citrus Diseases*, L.W. Timmer, S.M. Garnsey, and J.H. Graham, eds. (APS Press, St. Paul).
- GRAHAM, J.H., GOTTWALD, T.R., CUBERO, J., et al. (2004). *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. *Mol Plant Pathol* **5**, 1–15.
- HEIDELBERG, J.F., EISEN, J.A., NELSON, W.C., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477–483.
- HENDRICKS, E.C., SZERLONG, H., HILL, T., et al. (2000). Cell division, guillotining of dimer chromosomes and SOS induction in resolution mutants (*dif*, *xerC* and *xerD*) of *Escherichia coli*. *Mol Microbiol* **36**, 973–981.
- HENDSON, M., PURCELL, A.H., CHEN, D., et al. (2001). Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. *Appl Environ Microbiol* **67**, 895–903.
- HOLT, J.G. (1994). *Xylella*. *Bergey's Manual of Determinative Bacteriology*, 9th ed. J.G. Holt, ed. (Williams and Wilkins, London), pp. 100–115.
- HOPKINS, D.L. (1989). *Xylella fastidiosa*: xylem-limited bacterial pathogen of plants. *Annu Rev Phytopathol* **27**, 271–290.
- HOUSBY, J.N., THOMAS, J.D., WHARAM, S.D., et al. (1998). Conditional mutations in OutE and OutL block exoenzyme secretion across the *Erwinia carotovora* outer membrane. *FEMS Microbiol Lett* **165**, 91–102.
- HUNG, C.H., YANG, C.F., YANG, C.Y., et al. (2003). Involvement of *tonB*-*exbD1D2* operon in infection of *Xanthomonas campestris* phage phi L7. *Biochem and Biophys Res Commun* **302**, 878–884.

- JENKINS, C.L., and STARR, M.P. (1985). Formation of halogenated aryl-polyene (xanthomonadin) pigments by the type and other yellow-pigmented strains of *Xanthomonas maltophilia*. *Ann l'Institut Pasteur Microbiol* **136B**, 257–264.
- KANEHISA, M., GOTO, S., KAWASHIMA, S., et al. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res* **32**, D277–D280.
- KARLIN, S. (2001). Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol* **9**, 335–343.
- KELLY, M.J., POOLE, R.K., YATES, M.G., et al. (1990). Cloning and mutagenesis of genes encoding the cytochrome bd terminal oxidase complex in *Azotobacter vinelandii*: mutants deficient in the cytochrome d complex are unable to fix nitrogen in air. *J Bacteriol* **172**, 6010–6019.
- KHALIL-RIZVI, S., TOTH, S.I., VAN DER HELM, D., et al. (1997). Structures and characteristics of novel siderophores from plant deleterious *Pseudomonas fluorescens* A225 and *Pseudomonas putida* ATCC 39167. *Biochemistry* **36**, 4163–4171.
- KOIDE, T., ZAINI, P.A., MOREIRA, L.M., et al. (2004). DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. *J Bacteriol* **186**, 5442–5449.
- KOSTER, M., OVAA, W., BITTER, W., et al. (1995). Multiple outer membrane receptors for uptake of ferric pseudobactins in *Pseudomonas putida* WCS358. *Mol Gen Genet* **248**, 735–743.
- LACAVA, P.T., ARAUJO, W.L., MARCON, J., et al. (2004). Interaction between endophytic bacteria from citrus plants and the phytopathogenic bacteria *Xylella fastidiosa*, causal agent of citrus-variegated chlorosis. *Lett Appl Microbiol* **39**, 55–59.
- LAZDUNSKI, C.J., BOUVERET, E., RIGAL, A., et al. (1998). Colicin import into *Escherichia coli* cells. *J Bacteriol* **180**, 4993–5002.
- LEE, H.M., CHEN, J.R., LEE, H.L., et al. (2004). Functional dissection of the XpsN (GspC) protein of the *Xanthomonas campestris* pv. *campestris* type II secretion machinery. *J Bacteriol* **186**, 2946–2955.
- LEE, H.M., WANG, K.C., LIU, Y.L., et al. (2000). Association of the cytoplasmic membrane protein XpsN with the outer membrane protein XpsD in the type II protein secretion apparatus of *Xanthomonas campestris* pv. *campestris*. *J Bacteriol* **182**, 1549–1557.
- LEE, R.F.E.A., BERETTA, M.J.G., HARTUNG, J.H., et al. (1993). Citrus variegated chlorosis: confirmation of a *Xylella fastidiosa* as the causal agent. *Summa Phytopathol* **19,2**, 123–125.
- LEITE, B., ANDERSEN, P.C., and ISHIDA, M.L. (2004). Colony aggregation and biofilm formation in xylem chemistry-based media for *Xylella fastidiosa*. *FEMS Microbiol Lett* **230**, 283–290.
- LERAT, E., DAUBIN, V., and MORAN, N.A. (2003). From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-Proteobacteria. *PLoS Biol* **1**, E19.
- LIMA, W.C., PAQUOLA, A.C.M., VAN SLUYS, M.-A., et al. (2005). Non-gamma proteobacteria gene islands contribution to *Xanthomonas* genome. *OMICS* (this issue).
- LIN, R.J., CAPAGE, M., and HILL, C.W. (1984). A repetitive DNA sequence, rhs, responsible for duplications within the *Escherichia coli* K-12 chromosome. *J Mol Biol* **177**, 1–18.
- LOBRY, J.R. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol* **13**, 660–665.
- LONG, S.R., and STASKAWICZ, B.J. (1993). Prokaryotic plant parasites. *Cell* **73**, 921–935.
- LOPES, J.R.S., BERETTA, M.J.G., HAKAKA, R., et al. (1996). Confirmação da transmissão por cigarrinhas do agente causal da clorose variegada dos citros, *Xylella fastidiosa*. *Fitopatol Brasil* **21**(supplement), 343.
- LOPEZ, N.I., HAEDO, A.S., and MENDEZ, B.S. (1999). Evaluation of *Xanthomonas campestris* survival in a soil microcosm system. *Int Microbiol* **2**, 111–114.
- MINET, A.D., and CHIQUET-EHRISMANN, R. (2000). Phylogenetic analysis of teneurin genes and comparison to the rearrangement hot spot elements of *E. coli*. *Gene* **257**, 87–97.
- MOREIRA, L.M., DE SOUZA, R.F., ALMEIDA, N.F., JR., et al. (2004). Comparative genomics analyses of citrus-associated bacteria. *Annu Rev Phytopathol* **42**, 163–184.
- NOEL, L., THIEME, F., GABLER, J., et al. (2003). XopC and XopJ, two novel type III effector proteins from *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol* **185**, 7092–7102.
- NUNES, L.R., ROSATO, Y.B., MUTO, N.H., et al. (2003). Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res* **13**, 570–578.
- PAGE, R.D., and CHARLESTON, M.A. (1997). From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. *Mol Phylogenet Evol* **7**, 231–240.
- PASHALIDIS, S., MOREIRA, L.M., ZAINI, P.A., et al. (2005). Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose. *OMICS* (this issue).
- PENDRAK, M.L., and PERRY, R.D. (1991). Characterization of a hemin-storage locus of *Yersinia pestis*. *Biol Met* **4**, 41–47.

- PICKARD, D., WAIN, J., BAKER, S., et al. (2003). Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7. *J Bacteriol* **185**, 5055–5065.
- POPLAWSKY, A.R., and CHUN, W. (1997). *pigB* determines a diffusible factor needed for extracellular polysaccharide slime and xanthomonadin production in *Xanthomonas campestris* pv. *campestris*. *J Bacteriol* **179**, 439–444.
- POPLAWSKY, A.R., URBAN, S.C., and CHUN, W. (2000). Biological role of xanthomonadin pigments in *Xanthomonas campestris* pv. *campestris*. *Appl Environ Microbiol* **66**, 5123–5127.
- PURCELL, A.H. (1980). Almond leaf scorch: leafhopper and spittle bug vectors. *J Econ Entomol* **73**, 834–838.
- RAJAGOPAL, L., SUNDARI, C.S., BALASUBRAMANIAN, D., et al. (1997). The bacterial pigment xanthomonadin offers protection against photodamage. *FEBS Lett* **415**, 125–128.
- ROSSETTI, V., GARNIER, M., BOVÉ, J.M., et al. (1990). Presence de bacteries dans le xyleme d'oranges atteints de chlorose varieguee, une nouvelle maladie des agrumes au Bresil. *CRAcad Paris* **310**, 345–349.
- SCARPARI, L.M., LAMBAIS, M.R., SILVA, D.S., et al. (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.
- SCHAAD, N.W., and ALVAREZ, A. (1993). *Xanthomonas campestris* pv. *campestris*: cause of black rot of crucifers. In *Xanthomonas*, 1st ed. J.G. Swings and E.L. Civerolo, eds. London: Chapman & Hall, pp. 51–53.
- SCHAAD, N.W., POSTNIKOVA, E., LACY, G., et al. (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.
- SCHUBERT, S., RAKIN, A., KARCH, H., et al. (1998). Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect Immun* **66**, 480–485.
- SEKIZAKI, T., OTANI, Y., OSAKI, M., et al. (2001). Evidence for horizontal transfer of *Ssu*DATII restriction-modification genes to the *Streptococcus suis* genome. *J Bacteriol* **183**, 500–511.
- SHARP, P.M., KELLEHER, J.E., DANIEL, A.S., et al. (1992). Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc Natl Acad Sci USA* **89**, 9836–9840.
- SIMPSON, A.J., REINACH, F.C., ARRUDA, P., et al. (2000). The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *Nature* **406**, 151–157.
- SIVONEN, K., NAMIKOSHI, M., EVANS, W.R., et al. (1992). Three new microcystins, cyclic heptapeptide hepatotoxins, from *Nostoc* sp. strain 152. *Chem Res Toxicol* **5**, 464–469.
- STARR, M.P., JENKINS, C.L., BUSSEY, L.B., et al. (1977). Chemotaxonomic significance of the xanthomonadins, novel brominated aryl-polyene pigments produced by bacteria of the genus *Xanthomonas*. *Arch Microbiol* **113**, 1–9.
- STEINMANN, D., KOPLIN, R., PUHLER, A., et al. (1997). *Xanthomonas campestris* pv. *campestris* *lpsI* and *lpsJ* genes encoding putative proteins with sequence similarity to the alpha- and beta-subunits of 3-oxoacid CoA-transferases are involved in LPS biosynthesis. *Arch Microbiol* **168**, 441–447.
- STOCK, J., and SURETTE, M. (1994). Bacterial chemotaxis. The motor connection. *Curr Biol* **4**, 143–144.
- STOCK, J., SURETTE, M., and PARK, P. (1994). Chemosensing and signal transduction in bacteria. *Curr Opin Neurobiol* **4**, 474–480.
- STOCK, J.B., SURETTE, M.G., MCCLEARY, W.R., et al. (1992). Signal transduction in bacterial chemotaxis. *J Biol Chem* **267**, 19753–19756.
- STROM, M.S., and LORY, S. (1993). Structure-function and biogenesis of the type IV pili. *Annu Rev Microbiol* **47**, 565–596.
- SWINGS, J.G., and CIVEROLO, E.L. (1993). *Xanthomonas*, 1st ed. (Chapman & Hall, London).
- TANG, J.L., LIU, Y.N., BARBER, C.E., et al. (1991). Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. *Mol Gen Genet* **226**, 409–417.
- TATUSOV, R.L., NATALE, D.A., GARKAVTSEV, I.V., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* **29**, 22–28.
- THOMAS, J.D., REEVES, P.J., and SALMOND, G.P. (1997). The general secretion pathway of *Erwinia carotovora* subsp. *carotovora*: analysis of the membrane topology of OutC and OutF. *Microbiology* **143**, 713–720.
- THOMPSON, J.D., HIGGINS, D.G., and 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**, 4673–4680.
- VAN SLUYS, M.A., DE OLIVEIRA, M.C., MONTEIRO-VITORELLO, C.B., 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**, 1018–1026.
- VAN SLUYS, M.A., MONTEIRO-VITORELLO, C.B., CAMARGO, L.E., et al. (2002). Comparative genomic analysis of plant-associated bacteria. *Annu Rev Phytopathol* **40**, 169–189.

- VOJNOV, A.A., SLATER, H., NEWMAN, M.A., et al. (2001). Regulation of the synthesis of cyclic glucan in *Xanthomonas campestris* by a diffusible signal molecule. *Arch Microbiol* **176**, 415–420.
- VORHOLTER, F.J., NIEHAUS, K., and PUHLER, A. (2001). Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. *Mol Genet Genomics* **266**, 79–95.
- WATERS, V.L., and CROSA, J.H. (1991). Colicin V virulence plasmids. *Microbiol Rev* **55**, 437–450.
- WELLS, J.M., RAJU, B.C., HUNG, H.Y., et al. (1987). *Xylella fastidiosa* gen. nov., sp. nov.: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *Int J Syst Bacteriol* **37**, 136–143.
- WELLS, J.M., RAJU, B.C., THOMPSON, J.M., et al. (1981). Etiology of phony peach and plum leaf scald diseases. *Phytopathology* **71**, 1156–1161.
- WIGGERICH, H.G., KLAUKE, B., KOPLIN, R., et al. (1997). Unusual structure of the tonB-exb DNA region of *Xanthomonas campestris* pv. *campestris*: tonB, exbB, and exbD1 are essential for ferric iron uptake, but exbD2 is not. *J Bacteriol* **179**, 7103–7110.
- WIGGERICH, H.G., and PUHLER, A. (2000). The exbD2 gene as well as the iron-uptake genes tonB, exbB and exbD1 of *Xanthomonas campestris* pv. *campestris* are essential for the induction of a hypersensitive response on pepper (*Capsicum annuum*). *Microbiology* **146**, 1053–1060.
- WILLIAMS, K.P. (2002). Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* **30**, 866–875.
- WOOLFIT, M., and BROMHAM, L. (2003). Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. *Mol Biol Evol* **20**, 1545–1555.

Address reprint requests to:

Dr. João C. Setubal

Virginia Bioinformatics Institute and Department of Computer Science

Virginia Polytechnic Institute and State University

Bioinformatics I

Blacksburg, VA 24060–0477

E-mail: setubal@vbi.vt.edu

NOTE ADDED IN PROOF

After this paper had been written, Lee et al. (2005) published a new *Xanthomonas* genome. [Lee, B.M., Park, Y.J., Park, D.S., Kang, H.W., Kim, J.G., Song, E.S., Park, I.C., Yoon, U.H., Hahn, J.H., Koo, B.S., Lee G.B., Kim, H., Park, H.S., Yoon, K.O., Kim, J.H., Jung, C.H., Koh, N.H., Seo, J.S. and Go, S.J. (2005). The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res* **33**, 577–586.]

ANEXO 5

TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA

ARTIGO PUBLICADO

Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose.

Pashalidis S, Moreira LM, Zaini PA, Campanharo JC, Alves LM, Ciapina LP, Vencio RZ, Lemos EG, Da Silva AM, Da Silva AC.

OMICS. 2005 Spring;9(1):77-90.

PMID: 15805779 - PubMed - indexed for MEDLINE.

Whole-Genome Expression Profiling of *Xylella fastidiosa* in Response to Growth on Glucose

***STEFANO PASHALIDIS,¹ *LEANDRO M. MOREIRA,¹ *PAULO A. ZAINI,¹
JOÃO C. CAMPANHARO,² LÚCIA M.C. ALVES,² LUCIANE P. CIAPINA,²
RICARDO Z.N. VÊNCIO,³ ELIANA G.M. LEMOS,² ALINE M. DA SILVA,¹
and ANA C.R. DA SILVA¹**

ABSTRACT

Xylella fastidiosa is the etiologic agent of diseases in a wide range of economically important crops including citrus variegated chlorosis, a major threat to the Brazilian citrus industry. The genomes of several strains of this phytopathogen have been completely sequenced enabling large-scale functional studies. In this work we used whole-genome DNA microarrays to investigate the transcription profile of *X. fastidiosa* grown in defined media with different glucose concentrations. Our analysis revealed that while transcripts related to fastidial gum production were unaffected, colicin-V-like and fimbria precursors were induced in high glucose medium. Based on these results, we suggest a model for colicin-defense mechanism in *X. fastidiosa*.

INTRODUCTION

XYLELLA FASTIDIOSA IS A GRAM-NEGATIVE XYLEM-LIMITED BACTERIUM which is the etiologic agent of economically important plant diseases, such as Pierce's disease of grapevines (PD) and citrus variegated chlorosis (CVC) (Tyson et al., 1985; Chang et al., 1993; Hopkins and Purcell, 2002). The latter is a major concern to the citrus industry, being responsible for annual losses over \$100 million. CVC is considered one of the most devastating citrus diseases and according to Fundecitrus (www.fundecitrus.com.br), of the 180 million productive trees, 70 million will be affected during the next five years.

Upon transmission from infected plants by xylem sap-feeding sharpshooter leafhoppers, *X. fastidiosa* spreads systemically throughout the plant and attaches to xylem vessel walls. Depending on size, these biofilm-like colonies eventually occlude the vascular system, leading to the water-stress symptoms typical of CVC (Roberto et al., 1996; McElrone et al., 2001). Furthermore, it has been proposed that the fastidial exopolysaccharide and fimbriae are involved in the bacterial attachment and survival in the plant as well as in the insect vector (da Silva et al., 2001; Feil et al., 2003; Newman et al., 2004). Thus, the formation

¹Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, Brasil.

²Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, Universidade Estadual Paulista, Jaboticabal, Brasil.

³BIOINFO-USP-Núcleo de Pesquisas em Bioinformática, Universidade de São Paulo, São Paulo, Brasil.

*These authors contributed equally to this work.

of biofilm is currently accepted as a crucial factor in the pathogenicity of *X. fastidiosa* (Simpson et al., 2000; Machado et al., 2001; Newman et al., 2003; Osiro et al., 2004). Virulence mechanism might also include toxins, antibiotics and ion sequestration systems (Dow and Daniels, 2000; Simpson et al., 2000), but these remain poorly characterized.

Several *X. fastidiosa* strains have had their genomes completely or partially sequenced and comparative genome analysis of different strains pointed to common virulence determinants as well as strain-specific genomic signatures (Simpson et al., 2000; Bhattacharyya et al., 2002; Nunes et al., 2003; Van Sluys et al., 2003; Koide et al., 2004). These genome sequence data paved the way for high-throughput functional studies such as transcriptome and proteome analyses in an effort to improve our understanding of physiology and pathogenicity of this organism (de Souza et al., 2003; Nunes et al., 2003; Smolka et al., 2003; de Souza et al., 2004; Koide et al., 2004).

In this work we used full-genome DNA microarrays to investigate the transcription profile of *X. fastidiosa* grown in defined media with different glucose concentrations. Our analysis revealed that while transcripts related to fastidian gum production were unaffected, colicin-V-like and fimbria precursors were induced in high glucose medium. Based on these results we suggest a model for colicin-defense mechanism in *X. fastidiosa*.

MATERIALS AND METHODS

Bacterial strain and growth conditions

Triply cloned *X. fastidiosa* strain 9a5c (Li et al., 1999) isolated from citrus variegated chlorosis (CVC) symptomatic *Citrus sinensis* (L.) Osbeck trees (sweet orange) was grown on solid XDM₂ medium (Lemos et al., 2003) at 28°C in the dark. Cells with less than 20 passages were used. After 6 days (late log-phase) of growth, cells were collected and washed in liquid XDM₂ medium without glucose. A 1 mL aliquot of the cell suspension was then plated in XDM₂ with defined glucose concentrations (1, 50, and 250 mM) and grown for 6 more days for total RNA isolation.

To monitor the growth at different concentrations of glucose, the cultures were prepared as described above until the washing step, and then transferred to liquid XDM₂ medium with defined glucose concentrations. Samples were taken every 4 days during 12 days. Growth was monitored by measuring the optical density of resuspended cells and by total protein content. For this, samples were resuspended in 1.0 mL of sonication buffer (10mM Tris-HCl and 5mM MgCl₂, pH 7.0) by vortexing. A 50- μ L aliquot was removed and diluted in 950 μ L of water for measurement of optical density at 600 nm. The remaining cells were sonicated (Sonifier Branson, model 250 at 85W) in an ice water-bath for 5 min, and the extract was clarified by centrifugation at $10,600 \times g$ for 10 min at 4°C and total protein was estimated by the Hartree assay (Hartree, 1972).

Microarray construction, fluorescent labeling, and hybridization

A 6152-element DNA microarray containing unique internal fragments of 2692 CDS spotted at least in duplicate, representing 94.5% of all the 2848 CDS annotated by Simpson et al. (2000) was constructed as previously described (Koide et al., 2004). Expression profiling studies were carried out labeling 20 μ g of total RNA with indocarbocyanine or indodicarbocyanine (Cy3 or Cy5) using Cy-Scribe Post Labeling kit (GE Healthcare) according to manufacturer's instructions. Total RNA from 6-day cultures was extracted with TRIzol (Invitrogen), treated with RQ1 DNase I (Promega). Complete removal of genomic DNA was evaluated by PCR.

Fluorescent cDNAs were combined and denatured by heating to 95°C for 2 min and quickly chilled on ice. The targets were then applied to the microarray slide and covered with a 24 $\times$ 60 mm coverslip (Corning). The hybridization proceeded for 16 h at 42°C. The slides were then washed in 1 $\times$ SSC + 0.2% SDS for 10 min at 55°C, twice in 0.1 $\times$ SSC + 0.2% SDS also for 10 min at 55°C, 0.1 $\times$ SSC for 1 min at room temperature and finally in deionized water for 1 min at room temperature and dried with N₂ gas.

Data acquisition, ratio normalization, and statistical analysis

Microarray slides were scanned using a Generation III DNA scanner (GE Healthcare) and fluorescence intensity values from each spot were extracted using ArrayVision 7.0 software (Imaging Research Inc.). We used the usual “reference design”, with the 50 mM glucose growth condition as the reference, and prepared three biological replicates for each growth condition (Churchill, 2002). Each microarray slide has at least two technical replicates measured for both channels and, therefore, each CDS has at least $3 \times 2 \times 2 = 12$ intensity measures. A CDS is defined as “detectable” if the majority of its intensity measurements are brighter than the 90% percentile of its surrounding background intensity. In microarray context, a given CDS intensity measurement cannot be detected significantly above its local background due to a series of reasons ranging from limitations on dynamic range or high level of local background hybridization, to real lack of transcriptional activity. On the other hand, a given CDS successfully detected can have its intensity value due to unspecific hybridization or artifacts. We deemed detected CDS as being transcriptionally active but we do not call the not detected CDS as not transcribed acknowledging the limitations above. Since our hybridization stringency is relatively high and the array was constructed with specific primers (Koide et al., 2004), we believe that our “detectable thus expressed” approximation is reasonable.

The $\log_2$ -ratio results are calculated only for those CDS that could be detected in both co-hybridized conditions. CDS that could be observed only in one condition relative to other are treated separately since are differentially expressed but do not have measurable expression ratio. Data normalization was carried out essentially as previously described by Koide et al. (2004) using LOWESS fitting on M versus S plot, where M is the $\log_2$ -ratio of background subtracted fluorescence intensities (I_{CY3} and I_{CY5}), defined as $M = \log_2(I_{CY5}/I_{CY3})$, and S is the logarithm of the average intensity, defined as $S = \log_2(I_{CY5}/2 + I_{CY3}/2)$ (see supplemental Figs. S1.1–S1.23).

To determine hybridization noise and to estimate dynamic cutoff values for classifying a CDS according to its relative expression levels we used the hybridization data collected from two duplicated biological replicates of self-self hybridization (independent harvest of *X. fastidiosa* grown at 50 mM glucose). The intensity-dependent noise-threshold cutoff values were obtained according to the rationale introduced by Koide et al. (2004), with the improvements described in <http://blasto.iq.usp.br/~rvencio/HTself>. We used a sliding window of size 1.0, step 0.2 and defined 0.995 credibility intervals. A CDS was considered as consistently differentially expressed if the majority of its M replicated measurements were outside of the 0.995 curves defined by the self-self hybridizations. Using these criteria, up- or down-regulated categories were defined for samples grown at 1 mM and at 250 mM glucose (see supplemental Fig. S2). The complete data set is publicly available under the MIAME guidelines (Brazma et al., 2001) at supplemental web site (<http://verjo19.iq.usp.br/xylella/microarray/glucose>).

Real-time quantitative RT-PCR

cDNA was generated from 5 μ g of total RNA using 200 U of SuperscriptII reverse transcriptase (Invitrogen) and 500 ng of random nonamer primers, according to the manufacture’s instructions. cDNA was diluted with nuclease free water to 35 ng/ μ L and stored at -20°C before use. Real-time quantitative PCR was performed using an ABI PRISM 5700 Sequence Detection System (Applied Biosystems) using default parameters. The PCR mixture included 10 μ L of Platinum® SYBR® Green qPCR SuperMix UDG (Invitrogen), 800 nM forward and reverse primers and 180ng template cDNA. In order to confirm the generation of specific PCR products, the PCR was immediately followed by melting curve analysis of the RT-PCR product according to the manufacturer’s recommendations (Applied Biosystems). Primers were purchased from Bio-Synthesis, Inc. Primers for amplification of XF0305 or XF2157 were used as endogenous controls to normalize the amount of total RNA per sample. The fold-change of each gene was calculated by using the $2^{-\Delta\Delta\text{CT}}$ method, as described by Livak and Schmittgen (2001) from three independent experiments. We considered a CDS differentially expressed relative to 50mM glucose standard condition if the mean minus one standard deviation on obtained fold-change exceeded the 1.5-fold value cutoff. PCR primers were designed with PRIMER EXPRESS 2.0 software (Applied Biosystems), and their sequences are given in Table 1.

TABLE 1. GENES AND PRIMERS USED IN REAL-TIME QUANTITATIVE PCR

Gene ID ^a	Gene name	Primers
XF0262	<i>cvaC</i>	F 5'GGCGGTATTGCTGGTGCTAT R 5'AGCGAAGGTGCCGTTAAAGA
XF0287	<i>rpfB</i>	F 5'CGGGCTTCGACAAACTTGA R 5'GACGGAGCGCTGGATCAC
XF0305	<i>nuoA</i>	F 5'TTCATCGTGCCTTGGACTCA R 5'CAGCGCTCCCTTCTTCCATA
XF1115	<i>rpfF</i>	F 5'GCTTGGAGGAGGCTTCGAA R 5'CACCCCTTCCTCAGCTACGA
XF2361	<i>gumK</i>	F 5'GCATGCATCTTTCGGTATTGC R 5'CCAAAGCGTGTCCGATAAGAG
XF2157	<i>dnaQ</i>	F 5'GGTGCCGAAGTATTATTCACA R 5'CAACCGCGATAACTCGTAATCAA

^aSimpson et al. (2000).

DNA and protein sequence analysis

DNA and protein sequence similarity searches were done with BLAST tools (Altschul et al., 1997) at the NCBI databases and at the *X. fastidiosa* 9a5c genome sequencing consortium database (<http://aeg.lbi.ic.unicamp.br/xf/>). Sequence alignment was performed with CLUSTAL W software (Thompson et al., 1994). Protein secondary structure prediction was done with NNPREDICT software (Kneller et al., 1990).

RESULTS AND DISCUSSION

Analysis of *X. fastidiosa* growth at different concentrations of glucose

The recent development of defined media for *X. fastidiosa* (Campanharo et al., 2003; Lemos et al., 2003; Almeida et al., 2004; Leite et al., 2004) allowed us to examine the growth kinetics and global gene expression under different glucose concentrations. As shown in Figure 1, growth in XDM₂ containing 1 mM and 250 mM glucose is strikingly different compared to standard XDM₂, which has 50 mM glucose. After 8 days of growth, cell density reached OD₆₀₀ ~1.2 at the standard conditions and was 3–4-fold lower at 1 mM and 250 mM glucose. It can also be observed that cultures grown at low glucose reach the stationary phase within 4 days of growth while this stage is only reached after 8 days at the standard conditions. Interestingly, even after 12 days the stationary phase was not reached for cultures grown at 250 mM glucose. Their log phase was delayed for ~8 days, possibly when glucose declined to optimal levels. In all of the conditions the cellular density correlated well with total protein content. This correlation suggests that *X. fastidiosa* does not alter synthesis of fastidian gum (exopolysaccharide) in response to glucose variations.

Global analysis of *X. fastidiosa* gene expression in response to glucose concentration

We next monitored global transcription profiles of *X. fastidiosa* grown in XDM₂ with 1, 50, and 250 mM glucose. For this we used DNA microarrays representing 94.5% of the 2848 previously annotated coding sequences (CDS) (Simpson et al., 2000; Koide et al., 2004). The microarray slides were competitively hybridized with fluorescently labeled cDNA prepared from RNA samples of 6 day-old cultures. As detailed in Materials and Methods, we considered a CDS as expressed if its intensity measurements were consistently and significantly above the local background. As represented in the diagrams of Figure 2, the analysis of hybridization data show that 2143 CDS were considered expressed at 50mM glucose. On the other hand, only 554 and 1711 CDS were expressed by *X. fastidiosa* cells grown at 1 mM and 250 mM, respectively. These results show that when *X. fastidiosa* was grown in 1 mM glucose XDM₂ an enormous set of

FIG. 1. Growth curves of *Xylella fastidiosa* 9a5c in XDM₂ medium with different concentrations of glucose. Closed symbols represent optical density readings at 600 nm. Open symbols represent total protein content. Squares, circles and triangles represent cultures grown at 1, 50, and 250 mM of glucose, respectively.

transcripts were down-regulated compared to the standard medium. This effect was also observed in cells growing in 250 mM glucose, though to a lesser extent. CDS expressed in each condition are listed in supplemental Table S1 and represented on interactive transcriptome maps that summarize all our hybridization results at the project site (<http://verjo19.iq.usp.br/xylella/microarray/glucose>). The transcriptome maps provide an overview of genome activity in each hybridization condition, enriched with additional information on each CDS. The project site also offers raw and normalized data and additional information on data collection, processing and statistical analysis.

In addition, the gene expression profiles revealed that 211 out of 551 sequences (see supplemental Table S2), that had been considered invalid CDS in the re-annotation of *X. fastidiosa* strain 9a5c genome (Van Sluys et al., 2002), were transcriptionally active in the standard XDM₂ medium. The fact that we have detected transcriptional activity of these predicted CDS demonstrates the usefulness of DNA microarrays as a complementary tool to validate genome annotations. Given 187 of these expressed CDS are conserved in *X. fastidiosa* strain Temecula, responsible for Pierce's disease in grapevines, we intend to reannotate these CDS as conserved hypotheticals.

FIG. 2. Gene expression changes in three glucose concentrations. The number of CDSs expressed in cells grown at 50 versus 1 mM (A), and 50 versus 250 mM glucose (B). White numbers correspond to CDS detected in both conditions, for which differential expression analysis was performed. The number of up- and down-regulated CDS in each condition are indicated below the circles. Undetectable CDS were excluded from further analysis.

Functional characteristics of differentially expressed genes in response to growth on glucose

As mentioned above, transcriptional activity was drastically reduced in *X. fastidiosa* cells grown at 1 mM glucose given that a total of 1589 CDS were not expressed. Moreover, cells grown at 250 mM glucose also reduced the overall number of expressed genes considering 452 CDS were not expressed. Even though these large sets of CDS (supplemental Table S1) should be assigned as differentially expressed relative to growth at 50mM, their precise hybridization ratios were not be calculated. Therefore, the biological implication of these observations will not be addressed in this work. We chose to analyze differentially expressed genes exclusively with genes expressed in both conditions and for these precise ratios were calculated (Fig. 2). When comparing 50 mM versus 1 mM, 14 and 8 out of 554 CDS were found to be up- and down-regulated, respectively (Tables 2 and S3). When 50 mM versus 250 mM were compared, 39 and 22 out of 1691 CDS were up- and down-regulated, respectively (Tables 3 and S3).

Among the 14 CDS up-regulated in 1 mM glucose listed in Table 2, we highlight one that encodes a sugar ABC transporter ATP-binding protein (XF1067) and another that encodes an ATP-dependent Clp protease proteolytic subunit (XF1187). This protease assembles into a disk-like structure with a central cavity, resembling the structure of eukaryotic proteasomes (Kessel et al., 1995). Interestingly, Clp-dependent proteolysis has been recently linked to general bulk protein breakdown at the transition from growing to non-growing phases in *B. subtilis* (Kock et al., 2004).

Table 3 presents the 39 CDS found to be up-regulated upon growth in 250 mM glucose. Twenty-six of them were annotated as hypothetical or conserved hypothetical. Eleven of these reside between CDS XF1648 and XF1693 that comprises a region denoted as Xfp4 (for phage number 4). Although the region harboring CDS XF1648-XF1657 is currently not annotated as belonging to this phage, we suspect this phage region starts at CDS XF1642 and not at XF1658 as currently proposed (Simpson et al., 2000). XF1642 encodes a phage integrase whose genomic location is indicative of an island or horizontally transferred region in *X. fastidiosa* 9a5c (Moreira et al., 2004). Expression of phage sequences have been shown to be under the control of nutrient availability in a microarray comparison of *X. fastidiosa* grown in rich and nutrient-limited media (Nunes et al., 2003).

Other up-regulated genes in this condition include XF0395, XF1189, XF2234 and XF2625. The first encodes bacterioferritin, an intracellular reservoir of ferric iron, the second an ATP-dependent serine endopeptidase (Lon), and the other two encode heat shock proteins. This observation suggests *X. fastidiosa* might be undergoing some kind of stress due to growth on high glucose concentration. In *Pseudomonas aeruginosa*, bacterioferritin has been proposed to be involved indirectly in the resistance to redox stress

TABLE 2. UP-REGULATED GENES AT 1 mM GLUCOSE

<i>Gene ID^a</i>	<i>Product</i>	<i>Gene name</i>	<i>Ratio</i>
XF0094	Cell division protein	<i>ftsJ</i>	1.6
XF0388	Hypothetical protein	—	2.0
XF0561	Hypothetical protein	—	2.0
XF1067	Sugar ABC transporter ATP-binding protein	DR2153	1.9
XF1187	ATP-dependent Clp protease proteolytic subunit	<i>clpP</i>	2.1
XF1189	ATP-dependent serine proteinase La	<i>lopA</i>	1.8
XF1530	Subunit C of alkyl hydroperoxide reductase	<i>ahpC</i>	1.9
XF1554	Fumarate hydratase	<i>fumC</i>	1.9
XF1693	Hypothetical protein	—	2.1
XF1796	Bifunctional transcriptional repressor of the biotin operon/ biotin acetyl-CoA-carboxylase synthetase	<i>birA</i>	1.4
XF1810	Conserved hypothetical protein	HI0961	2.0
XF2377	Hypothetical protein	—	2.0
XFa0034	Conserved hypothetical protein	STMD1.84	1.8
XFb0001	Replication protein	—	2.1

^aSimpson et al. (2000).

WHOLE-GENOME EXPRESSION PROFILING OF *XYLELLA FASTIDIOSA*

TABLE 3. UP-REGULATED GENES AT 250 mM GLUCOSE

<i>Gene ID</i> ^a	<i>Product</i>	<i>Gene name</i>	<i>Ratio</i>
XF0032	PilY1 gene product	<i>pilY1</i>	2.1
XF0261	Colicin V immunity protein ^b	<i>cvi</i>	2.3
XF0262	Colicin V precursor	<i>cvaC</i>	2.5
XF0264	Colicin V precursor ^b	<i>cvaC</i>	2.4
XF0395	Bacterioferritin	<i>bfr</i>	2.3
XF0667	Hypothetical protein	—	1.9
XF0700	Hypothetical protein	—	2.3
XF0703	Hypothetical protein	—	2.0
XF1189	ATP-dependent serine proteinase La	<i>lopA</i>	1.7
XF1217	Hypothetical protein	—	3.0
XF1224	PilY1 gene product	<i>pilY1</i>	2.2
XF1287	Hypothetical protein	—	2.0
XF1550	Oxoglutarate dehydrogenase	<i>odhA</i>	1.5
XF1556	Hypothetical protein	—	1.8
XF1648	Hypothetical protein	—	3.6
XF1649	Conserved hypothetical protein	b2360	3.1
XF1650	Hypothetical protein	—	2.3
XF1655	Hypothetical protein	—	2.6
XF1659	Hypothetical protein	—	3.9
XF1661	Hypothetical protein	—	2.8
XF1662	Hypothetical protein	—	2.0
XF1663	Phage-related protein	—	1.9
XF1673	Hypothetical protein	—	2.4
XF1683	Hypothetical protein	—	2.0
XF1693	Hypothetical protein	—	2.4
XF1754	Conserved hypothetical protein	<i>orf1</i>	2.1
XF1948	Colicin V production protein	<i>cvpA</i>	2.0
XF2005	Hypothetical protein	—	1.8
XF2173	Hypothetical protein	—	2.0
XF2234	Low molecular weight heat shock protein	<i>hspA</i>	2.7
XF2377	Hypothetical protein	—	2.0
XF2380	Hypothetical protein	—	2.0
XF2510	Hypothetical protein	—	2.8
XF2550	Outer membrane hemolysin activator protein	<i>hecB</i>	1.6
XF2625	Heat shock protein	<i>htpX</i>	2.5
XFa0024	Hypothetical protein	—	1.8
XFa0033	Hypothetical protein	—	1.6
XFa0045	Conserved hypothetical protein	—	1.6
XFb0001	Replication protein	—	2.1

^aSimpson et al. (2000).

^bCurrently annotated as hypothetical proteins.

(Ma et al., 1999) and the Lon protease homologue of *Brucella abortus* is essential for survival in a variety of stresses (Robertson et al., 2000).

When cultivated at 250 mM glucose *X. fastidiosa* also up-regulates two CDS (XF0032 and XF1224) that encode PilY1, a transmembrane protein involved in export and correct assembly of the Type-IV pili (Alm et al., 1996). Type IV pili are multifunctional devices at the bacterial surface that can act as virulence factors because of pilus-based motility or formation of biofilms (Nudleman and Kaiser, 2004). Fimbriae- and pili-mediated attachment of *X. fastidiosa* to xylem or cibarium has been suggested as essential to successful colonization (Purcell and Hopkins, 1996; Simpson et al., 2000). Other virulence determinants induced on high glucose include colicin-V relates genes, which will be further discussed below.

Surprisingly, based on our microarray data, the *rpf* and *gum* genes were not induced by the glucose increment used in this study. This was further confirmed by real-time quantitative PCR performed with selected genes (Fig. 3). Our initial premise was that fastidian gum production would be increased as a function of glucose concentration, as observed for *Xanthomonas campestris* pv. *campestris* (Vojnov et al., 2001) and *Xanthomonas axonopodis* pv. *citri* (our unpublished results). These results are specially intriguing since RpfB and RpfF are 70% and 44% identical, respectively, between *X. fastidiosa* and *X. axonopodis* pv. *citri*. These genes are essential to the synthesis of a diffusible signal factor (DSF) (Barber et al., 1997) which in turn activates a transduction pathway that leads to transcriptional regulation of genes required for pathogenic traits, such as EPS and exoenzyme synthesis (Slater et al., 2000). The expression of *rpf* and *gum* genes in *X. fastidiosa* has been recently shown to be affected by cell density *in vitro*, being significantly induced at high density conditions (Scarpari et al., 2003). Taken together, these observations lead us to conclude that cell density but not nutrient availability triggers fastidian gum production.

X. fastidiosa expresses colicin-V-like toxins and putative anti-toxin

X. fastidiosa 9a5c genome encodes two colicin-like precursor proteins encoded by *cvaC* genes XF0262 and XF0263 (Simpson et al., 2000). Colicin-V is an antibacterial polypeptide toxin produced by *E. coli*,

^a according to Simpson *et al.*, 2000.

FIG. 3. Gene expression of selected CDS in response to glucose variations. Real-time quantitative polymerase chain reactions using templates derived from total RNA samples of *X. fastidiosa* cultures grown at 1 mM glucose (white bars) and 250 mM glucose (gray bars) compared to the standard growth condition (50 mM glucose). Standard error bars from three independent experiments are shown. XF0305 (NADH-ubiquinone oxidoreductase, NQO7 subunit) was used as the normalizer of total RNA input.

WHOLE-GENOME EXPRESSION PROFILING OF *XYLELLA FASTIDIOSA*

which acts against closely related sensitive bacteria (Havarstein et al., 1994). Its synthesis is regulated by CvpA and activation and secretion depend on two other proteins, CvaB and CvaA. The first is an ATP-dependent ABC transporter inserted in the inner membrane of the cell, which transports colicin from the cytosol to the periplasm after the cleavage of a signal peptide, and the second is a transport protein inserted in the outer membrane that transports colicin from the periplasm to the external medium, usually with assistance of TolC (Waters and Crosa, 1991; Skvirsky et al., 1995; Zhang et al., 1995). In *X. fastidiosa* 9a5c *cvaA* and *cvaB* are dispersed in the chromosome (Simpson et al., 2000), different from *E. coli*, where these genes are organized in operon present in plasmid pColV (Gilson et al., 1987). Processed colicin-V is lethal against bacteria that have a specific receptor (Cir) but do not express a protein (Civ) that confers immunity against it (Zhang et al., 1995).

Our microarray results indicated that *cvaC* XF0262 and the neighboring XF0264 are up-regulated in cells grown at 250 mM glucose (Table 3). Induction of XF0262 was also verified by real-time quantitative PCR (Fig. 3). XF0264 is currently annotated as hypothetical, despite its strong aminoacid similarity to the other two *cvaC* found in *X. fastidiosa* 9a5c genome (Fig. 4A). The tandem organization of the *cvaC* genes in this strain suggests a genomic expansion as this gene is found as single-copy in other strains of *X. fastidiosa* with genomes completely (*X. fastidiosa*-Temecula) or partially sequenced (*X. fastidiosa*-Ann1 and *X. fastidiosa*-Dixon) (Bhattacharyya et al., 2002; Van Sluys et al., 2003). *cvpA* (XF1948) was also up-regulated in this condition, while *cvaA*, *cvaB* and *tolC* were not.

A

```

XF0263 MRELTSIEMN NVSGGDLATR IEASIVFGVS AFFAGSIWGG TRGGDGGGIL GVGSIAGQVG
XF0264 MRELTSIEMN NVSGGDFATR LEASIVGGIA AFFAGSIWGG TRGGDGGGVL GIGSIGQGVG
XF0262 MRELTLITEID NVSGADLGSR LSAAIVGGVA AFFAGSIWGG TRGGDGGGIL GVGSIAGQVG
      ***** *: : *****.*: : * :.*: ** *: : ***** * *****: * *:***.*****

XF0263 MVYGGIVGGI GGLIAGFVLD KNVTYNYAVG FYNSLFNGTFTK (283nt)
XF0264 MVYGGIAGAI GGAIGCFVLD KDVTYSYTTG FMNSLFSGNFAK (227nt)
XF0262 MVYGGIAGAI GGAIAGFVLD KDVIYSYTNQ FMSSIFNGTFAK (205nt)
      *****.*.* ** *. ******:* *.*: * *.*:.*.*.*.*

```

B

```

-----HHHHHHHHHH--HHHEHHHHHHH-H-----HHHHEHHHHHHHEEEEE-----
Q841V5 MDRKRTKLELLFAFIINATAIYIALAIYDCVFREKDFLSMHTFCFSALMSAICYFVGDNV
XF0261 MRSLFGQVFNTNPSDVNEDTVSASLAG-FWFKLSSRIERVLGSLGRQFAGQGY
      -----E-----HHH-HHH.HHHHH--HHHHHHHHH-HHHH-----

-----
Q841V5 YSISDKIKRRSYENSSDSK
XF0261 LWERTDSAQNFELN
      EEE-----

```

FIG. 4. (A) Comparison of *X. fastidiosa* 9a5c CvaC proteins. The deduced aminoacid sequences of XF0262, XF0263 and XF0264 were aligned using CLUSTALW. Identical and conserved aminoacids are denoted with (*) and (:), respectively. Sequences encoded by spotted amplicons are underlined. (B) Pairwise alignment of candidate *X. fastidiosa* Cvi with *E. coli* Cvi. The deduced amino acid sequences of XF0261 and *E. coli* cvi (UniProt accession number Q841V5) were aligned. Bold and underlined letters represent identical and conserved aminoacids, respectively. Secondary structures were predicted using NNPREDICT. H and E represent helices and β -strands, respectively.

Although the colicin-V secretory machinery is complete in *X. fastidiosa* 9a5c, paradoxically no gene encoding a putative immunity protein (*cvi*) was annotated based on Blast searches (Simpson et al., 2000). We then reasoned if a *cvi* analog could be among hypothetical CDS co-regulated with *cvaC*. Among these, the gene that fulfilled the requirements for a *X. fastidiosa* *cvi* was the hypothetical CDS XF0261, exactly upstream of the *cvaC* genes. XF0261 encodes a protein with similar size to *E. coli* *cvi* (66 and 74 aminoacids, respectively) and despite their low sequence similarity shown by pair-wise alignment, their predicted secondary structure turned out to be quite similar (Fig. 4B). Both are rich in helices in the central region and have a small fraction of β -sheets at the carboxi terminus. It has been proposed that the helix-rich structure allows *cvi* insertion in the inner membrane in *E. coli* (Waters and Crosa, 1991). Further experimentation to demonstrate the functionality of this putative *X. fastidiosa* Cvi is warranted. If indeed XF0261 is a *cvi* analog, we can postulate that *X. fastidiosa* is capable of synthesizing, secreting and defending itself against colicin-V. A model for this response is presented in Figure 5. Colicin secretion would allow efficient competition within the xylem and/or insect foregut. In fact, it has been demonstrated that *X. fastidiosa* lives in an endophytic environment in citrus plants (Araujo et al., 2001; Lacava et al., 2004).

X. axonopodis pv *citri*, a member of the same family as *X. fastidiosa*, also infects orange trees but seems incapable of producing colicin-V since *cvaA* and *cvaC* are absent. In *X. campestris* pv *campestris*, besides *cvaA* and *cvaC*, *cvaB* was not found (da Silva et al., 2002). A possible explanation for this is the different life style and habitat of *Xanthomonas*. By living in the mesophyl, an environment rich in nutrients and carbon sources, and by being able to intensively degrade plant tissues, they might not need to antagonize other bacteria in order to reduce local competition, differently from *Xylella*.

FIG. 5. Model of regulation, synthesis and immunity against colicin-V produced by *X. fastidiosa* 9a5c. Presence of the XF0261 protein inhibits colicin-V toxicity against *X. fastidiosa* 9a5c but allows K^+ efflux, cell membrane imbalance and death of susceptible bacteria. IM, PS, and OM indicate inner membrane, periplasm and outer membrane, respectively.

ACKNOWLEDGMENTS

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). We are greatly indebted to Dr Hugo A. Armelin for coordinating the Cooperation for Analysis of Gene Expression (CAGE) Project. A.M.D.S. and S.P. were partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, (CNPq). L.M.M., R.Z.N.V. and P.A.Z. are FAPESP doctoral fellows.

SUPPLEMENTAL METHODS

DNA microarray construction

According to Koide et al. (2004). See: <http://verjo19.iq.usp.br/xylella/microarray/glucose>. PCR primers were designed to amplify unique internal fragments of 200-1000 bp of each predicted CDS described in the annotated genome sequence of *X. fastidiosa* strain 9a5c (<http://aeg.lbi.ic.unicamp.br/xf>). Primers (18–23mers) with equivalent predicted melting temperature were designed with the use of a perl program that ran PRIMER3 (www-genome.wi.mit.edu/genome_software/other/primer3.html) for the complete CDS list, automatically testing many parameter settings and also guaranteeing that primers hybridized only to a single genome location. Oligonucleotides were synthesized by MWG and Operon Technologies. Genomic or cosmid DNA, obtained in the *X. fastidiosa* genome sequencing project (Simpson et al., 2000), were used as template in the first round of PCR amplification, and 200-fold-diluted PCR products were used as templates for PCR reamplification to increase product concentration when necessary.

The reactions were done in 96-well plates. The mixture in each well contained 100 ng of DNA, 0.5 U of Biolase Taq polymerase (Bioline), 0.2 mM of each dNTP (Invitrogen), 1.5 mM MgCl₂ and the primers at 0.5 μ M, in a total volume of 100 μ L. A 5-min denaturing step at 95°C was applied, followed by 40 cycles of 95°C for 45 sec, 50°C for 30 sec, 72°C for 1 min and a final step at 72°C for 10 min. 4 μ L of each PCR reaction were checked for product size and concentration by electrophoresis in 1.2% agarose gels. The amplicons were then purified with 96-well MultiScreen purification plates (Millipore), and an equal volume of dimethyl sulfoxide was added to the purified products (~100 ng/ μ L final concentration). Generation III DNA spotter (Amersham Biosciences) was used to array the samples onto coated type-7 glass slides (Amersham Biosciences). This spotter arrays two technical replicas of each sample, one in each longitudinal half of the slides. Thus, a 6152-element array was printed, representing 2692 CDS spotted at least in duplicate. After deposition, the spotted DNA samples were crosslinked to the coated slides by applying 50 mJ of UV light and the slides were stored desiccated at ~10% relative humidity at room temperature until use.

Hybridization conditions

Labeled DNA or cDNA fragments from both strains were combined in the hybridization mixture containing 50% formamide and hybridization buffer (Amersham Biosciences) in a final volume of 54 μ L. The mixture was heated to 92°C for 2 min, cooled on ice, and applied to the microarray. A cover slip was used to spread the solution throughout the microarray and the slide was then placed in a 50-mL Falcon tube that was sealed and horizontally positioned for hybridization in a 42°C water bath for 16 h. After hybridization, slides were washed at 55°C for 10 min in 1 $\times$ SSC buffer containing 0.2% SDS and twice for 10 min in 0.1 $\times$ SSC buffer containing 0.2% SDS, followed by 1 min at room temperature in 0.1 $\times$ SSC and a quick rinse in ddH₂O. After drying with N₂ in a clean room, the slides were ready for scanning.

Data acquisition and analysis

A Generation III DNA scanner (Amersham Biosciences) was used to acquire monochromatic images of 10 μ m/pixel from the microarray slides, corresponding to channels Cy3 (532-nm laser and PMT at 700 V) and Cy5 (633-nm laser and PMT at 750 V). Images were analyzed with ArrayVision 7.0 software (Imaging Research Inc.). The median-trimmed mean density (MTMdens) of signal intensity was the measure chosen for signal quantification. This measure removes pixels with signals below or above 4 MADs (median

of absolute deviation) of the mean signal intensity of all pixels within a spot, such as those representing dust particles. The median of the background intensity was calculated for a frame of $24,000 \mu\text{m}^2$ around each spot of $180 \mu\text{m}$ in diameter, and the value was subtracted from the spot's MTMdens value. This represents the raw data used in the normalization procedure.

Normalization procedure

Several imbalance errors affect the true ratio measure. We have assumed that all imbalance due to enzyme efficiency, wavelength detection, dye and brightness can be approximated by multiplicative factors that are contained into just one normalization constant that depends non-linearly on signal intensities. In order to normalize the Cy3 and Cy5 background-subtracted signal intensities (I_{Cy3} and I_{Cy5}), we have used the hypothesis that the great majority of genes are not differentially expressed and therefore the predominant ratio must be one. This is a reasonable hypothesis given that our microarray contains fragments of all CDS of the 9a5c strain genome. We have performed the LOWESS fitting on M vs S plot in order to obtain, locally and non-parametrically, the normalization constant and thus normalized ratios, following Yang et al. (2002). As defined in our work, $S = \log_2((I_{\text{Cy5}} + I_{\text{Cy3}})/2)$ and $M = \log_2(I_{\text{Cy5}}/I_{\text{Cy3}})$. Lowess normalization eliminated the dye bias of the ratios.

Intensity-dependent ratio cutoff level

We have performed homotypical (or self-self) hybridizations (9a5c strain independently grown in 50 mM of glucose labeled with either Cy3-dCTP or Cy5-dCTP and hybridized simultaneously to the same microarray) in order to derive intrinsic experimental variability of the 1:1 ratio (noise) and to set an upper and lower limit for this noise. With this approach, we have detected a clear dependence between ratio, estimated by M, and average foreground intensity of each CDS, estimated by S. Fold change of CDS with intensities above but close to background strongly varied inside $[0;\infty]$, or in logarithm scale, inside $[-;\infty]$ when the expected result is ratio = 1 or $\log_2(\text{ratio}) = 0$. We used three independent experiments (three slides, six images, two biological replicas) as samplings from experimental error around 1:1 bi-dimensional probability density distribution. We conditioned this distribution in arbitrary S intervals to make it one-dimensional and estimated the density distribution using Kernel Density Estimators (Silverman, 1986). Finally, we integrated this density around mode peak until 0.995 probability was reached, to determine an intensity-dependent ratio cutoff level. These cutoffs levels were subsequently used in the analysis of replicas of hybridization experiments; spots outside these credibility intervals present strong evidence against 1:1 ratio. The method used is an improvement of the rationale introduced by Koide et al. (2004) and is available online at <http://blasto.iq.usp.br/~rvencio/HTself>.

REFERENCES

- ALM, R.A., HALLINAN, J.P., WATSON, A.A. et al. (1996). Fimbrial biogenesis genes of *Pseudomonas aeruginosa*: pilW and pilX increase the similarity of type 4 fimbriae to the GSP protein-secretion systems and pilY1 encodes a gonococcal PilC homologue. *Mol Microbiol* **22**, 161–173.
- ALMEIDA, R.P., MANN, R., and PURCELL, A.H. (2004). *Xylella fastidiosa* cultivation on a minimal solid defined medium. *Curr Microbiol* **48**, 368–372.
- ALTSCHUL, S.F., MADDEN, T.L., SCHAFER, A.A., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- ARAUJO, W.L., MACCHERONI, W., JR., AGUILAR-VILDOSO, C.I., et al. (2001). Variability and interactions between endophytic bacteria and fungi isolated from leaf tissues of citrus rootstocks. *Can J Microbiol* **47**, 229–236.
- BARBER, C.E., TANG, J.L., FENG, J.X., et al. (1997). A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol Microbiol* **24**, 555–566.
- BHATTACHARYA, A., STILWAGEN, S., REZNIK, G., et al. (2002). Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res* **12**, 1556–1563.
- BRAZMA, A., HINGAMP, P., QUACKENBUSH, J., et al. (2001). Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nat Genet* **29**, 365–371.

- CAMPANHARO, J.C., LEMOS, M.V., and LEMOS, E.G. (2003). Growth optimization procedures for the phytopathogen *Xylella fastidiosa*. *Curr Microbiol* **46**, 99–102.
- CHANG, C.J., GARNIER, M., ZREIK, L., et al. (1993). Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. *Curr Microbiol* **27**, 137–142.
- CHURCHILL, G.A. (2002). Fundamentals of experimental design for cDNA microarrays. *Nat Genet (Suppl)*: **32**, 490–495.
- DA SILVA, A.C., FERRO, J.A., REINACH, F.C., et al. (2002). Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**, 459–463.
- DA SILVA, F.R., VETTORE, A.L., KEMPER, E.L., et al. (2001). Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *Fems Microbiol Lett* **203**, 165–171.
- DE SOUZA, A.A., TAKITA, M.A., COLETTA-FILHO, H.D., et al. (2003). Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity. *Mol Plant Microbe Interact* **16**, 867–875.
- DE SOUZA, A.A., TAKITA, M.A., COLETTA-FILHO, H.D., et al. (2004). Gene expression profile of the plant pathogen *Xylella fastidiosa* during biofilm formation in vitro. *FEMS Microbiol Lett* **237**, 341–353.
- DOW, J.M., and DANIELS, M.J. (2000). *Xylella* genomics and bacterial pathogenicity to plants. *Yeast* **17**, 263–271.
- FEIL, H., FEIL, W.S., DETTER, J.C., et al. (2003). Site-directed disruption of the fimA and fimF fimbrial genes of *Xylella fastidiosa*. *Phytopathology* **93**, 675–682.
- GILSON, L., MAHANTY, H.K., and KOLTER, R. (1987). Four plasmid genes are required for colicin V synthesis, export, and immunity. *J Bacteriol* **169**, 2466–2470.
- HARTREE, E.F. (1972). Determination of protein: a modification of the Lowry method that gives a linear photometric response. *Anal Biochem* **48**, 422–427.
- HOPKINS, D.L., and PURCELL, A.H. (2002). *Xylella fastidiosa*: Cause of Pierce's disease of grapevine and other emergent diseases. *Plant Dis* **86**, 1056–1066.
- KESSEL, M., MAURIZI, M.R., KIM, B., et al. (1995). Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26 S proteasome. *J Mol Biol* **250**, 587–594.
- KNELLER, D.G., COHEN, F.E., and LANGRIDGE, R. (1990). Improvements in protein secondary structure prediction by an enhanced neural network. *J Mol Biol* **214**, 171–182.
- KOCK, H., GERTH, U., and HECKER, M. (2004). The ClpP peptidase is the major determinant of bulk protein turnover in *Bacillus subtilis*. *J Bacteriol* **186**, 5856–5864.
- KOIDE, T., ZAINI, P.A., MOREIRA, L.M., et al. (2004). DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. *J Bacteriol* **186**, 5442–5449.
- LACAVA, P.T., ARAUJO, W.L., MARCON, J., et al. (2004). Interaction between endophytic bacteria from citrus plants and the phytopathogenic bacteria *Xylella fastidiosa*, causal agent of citrus-variegated chlorosis. *Lett Appl Microbiol* **39**, 55–59.
- LEITE, B., ANDERSEN, P.C., and ISHIDA, M.L. (2004). Colony aggregation and biofilm formation in xylem chemistry-based media for *Xylella fastidiosa*. *FEMS Microbiol Lett* **230**, 283–290.
- LEMONS, E.G., ALVES, L.M., and CAMPANHARO, J.C. (2003). Genomics-based design of defined growth media for the plant pathogen *Xylella fastidiosa*. *FEMS Microbiol Lett* **219**, 39–45. [Erratum in *FEMS Microbiol Lett* (2004) **231**, 299.]
- LI, W.B., ZREIK, L., FERNANDES, N.G., et al. (1999). A triply cloned strain of *Xylella fastidiosa* multiplies and induces symptoms of citrus variegated chlorosis in sweet orange. *Curr Microbiol* **39**, 106–108.
- LIVAK, K.J., and 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**, 402–408.
- MA, J.F., OCHSNER, U.A., KLOTZ, M.G., et al. (1999). Bacterioferritin A modulates catalase A (KatA) activity and resistance to hydrogen peroxide in *Pseudomonas aeruginosa*. *J Bacteriol* **181**, 3730–3742.
- MACHADO, M.A., SOUZA, A.A., COLETTA-FILHO, H.D., et al. (2001). Genome and pathogenicity of *Xylella fastidiosa*. *Mol. Biol. Today* **2**, 33–43.
- MCELDRONE, A.J., SHERALD, J.L., and FORSETH, I.N. (2001). Effects of water stress on symptomatology and growth of *Parthenocissus quinquefolia* infected by *Xylella fastidiosa*. *Plant Dis* **85**, 1160–1164.
- MOREIRA, L.M., DE SOUZA, R.F., ALMEIDA, N.F., JR., et al. (2004). Comparative genomics analyses of citrus-associated bacteria. *Annu Rev Phytopathol* **42**, 163–184.
- NEWMAN, K.L., ALMEIDA, R.P., PURCELL, A.H., et al. (2003) Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. *Appl Environ Microbiol* **69**, 7319–7327.
- NEWMAN, K.L., ALMEIDA, R.P., PURCELL, A.H., et al. (2004). Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. *Proc Natl Acad Sci USA* **101**, 1737–1742.

- NUDLEMAN, E., and KAISER, D. (2004). Pulling together with type IV pili. *J Mol Microbiol Biotechnol* **7**, 52–62.
- NUNES, L.R., ROSATO, Y.B., MUTO, N.H., et al. (2003). Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res* **13**, 570–578.
- OSIRO, D., COLNAGO, L.A., OTOBONI, A.M., et al. (2004). A kinetic model for *Xylella fastidiosa* adhesion, biofilm formation, and virulence. *FEMS Microbiol Lett* **236**, 313–318.
- PURCELL, A.H., and HOPKINS, D.L. (1996). Fastidious xylem-limited bacterial plant pathogens. *Annu Rev Phytopathol* **34**, 131–151.
- ROBERTO, S.R., COUTINHO, A., DE LIMA, J.E.O., et al. (1996). Transmissão de *Xylella fastidiosa* pelas cigarrinhas *Dilobopterus costalimai*, *Acrogonia terminalis* e *Oncometopia facialis* em citros. *Fitopatol. Bras.* **21**, 517–518.
- ROBERTSON, G.T., KOVACH, M.E., ALLEN, C.A., et al. (2000). The *Brucella abortus* Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol Microbiol* **35**, 577–588.
- SCARPARI, L.M., LAMBAIS, M.R., SILVA, D.S., et al. (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.
- SILVERMAN, B.W. (1986). *Density Estimation* (Chapman and Hall, London).
- SIMPSON, A.J., REINACH, F.C., ARRUDA, P., et al. (2000). The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *Nature* **406**, 151–157.
- SKVIRSKY, R.C., REGINALD, S., and SHEN, X. (1995). Topology analysis of the colicin V export protein CvaA in *Escherichia coli*. *J Bacteriol.* **177**, 6153–6159.
- SLATER, H., ALVAREZ-MORALES, A., BARBER, C.E., et al. (2000). A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol Microbiol* **38**, 986–1003.
- SMOLKA, M.B., MARTINS, D., WINCK, F.V., et al. (2003). Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution. *Proteomics* **3**, 224–237.
- THOMPSON, J.D., HIGGINS, D.G., and 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**, 4673–4680.
- TYSON, G.E., STOJANOVIC, B.J., KUKLINSKI, R.F., et al. (1985). Scanning electron microscopy of Pierce's disease bacterium in petiolar xylem of grape leaves. *Phytopathology* **75**, 264–269.
- VAN SLUYS, M.A., DE OLIVEIRA, M.C., MONTEIRO-VITORELLO, C.B., 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**, 1018–1026.
- VAN SLUYS, M.A., MONTEIRO-VITORELLO, C.B., CAMARGO, L.E., et al. (2002). Comparative genomic analysis of plant-associated bacteria. *Annu Rev Phytopathol* **40**, 169–189.
- VOJNOV, A.A., SLATER, H., DANIELS, M.J., et al. (2001). Expression of the gum operon directing xanthan biosynthesis in *Xanthomonas campestris* and its regulation in planta. *Mol Plant Microbe Interact* **14**, 768–774.
- WATERS, V.L., and CROSA, J.H. (1991). Colicin V virulence plasmids. *Microbiol Rev* **55**, 437–450.
- YANG, Y.H., DUDOIT, S., LUU, P., et al. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**, e15.
- ZHANG, L.H., FATH, M.J., MAHANTY, H.K., (1995). Genetic analysis of the colicin V secretion pathway. *Genetics* **141**, 25–32.

Address reprint requests to:
 Dr. Ana C.R. da Silva
 Alellyx Applied Genomics
 Rua James C. Maxwell 320
 Techno Park, 13067-850
 Campinas, SP, Brasil

E-mail: ana.rasera@alellyx.com.br

ANEXO 6

**TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA**

LISTA DE GENES PRESENTES NO XACARRAY

LISTA DE GENES E CLONES AMPLIFICADOS E FIXADOS NO XACARRAY

Os identificadores das colunas estão apresentados abaixo:

A - Linha ocupada em cada *subarray*, correspondente a cada uma das placas de 384 poços

(total de 9 placas)

B - *Subarray* indicado na lâmina

(total de 12 *subarrays* seguindo numeração crescente de cima para baixo)

C - Coluna ocupada em cada *subarray*, correspondente a cada posição linha x coluna das placas de 384 poços

(total de 32 colunas)

D - Posição ocupada pelo *amplicon* na placa de 384 poços referenciada

(01A01: número da placa 01, posição A01)

E - Identificador do gene no genoma de XAC

(de acordo com da Silva e colaboradores (2002), <http://genoma4.iq.usp.br/xanthomonas>)

F - Posição ocupada pelo *amplicon* na placa de 96 poços

(cada conjunto de 4 placas de 96 poços forma 1 placa de 384 poços)

G - Código identificador do clone que compõe a biblioteca de *shotgun* de XAC

(usado para a geração do *amplicon*)

H - Categoria funcional de anotação ao qual pertence o gene selecionado dentro do *amplicon*

(<http://genoma4.iq.usp.br/xanthomonas>)

I - Produto codificado pelo gene identificado na coluna E

SC - Score card, NC, no clone (ausência de clone para o gene)

A	B	C	D	E	F	G	H	I
01	01	01	01A01	XAC2172	01A01	A0JJ0101D04	I.C.1	<i>NADH dehydrogenase</i>
01	02	01	01A02	XAC1282	02A01	A0EC1088H01	I.D.1	<i>two-component system, sensor protein</i>
01	03	01	01A03	XAC1997	01A02	A0AC0115B12	VIII.A	<i>conserved hypothetical protein</i>
01	04	01	01A04	XAC0652	02A02	A0EC0311G07	I.C.2	<i>alcohol dehydrogenase class III</i>
01	05	01	01A05	XAC0913	01A03	A0AC0113A06	VIII.A	<i>conserved hypothetical protein</i>
01	06	01	01A06	XAC0326	02A03	A0JJ0702A07	I.D.1	<i>two-component system, sensor protein</i>
01	07	01	01A07	XAC0202	01A04	A0AC0113A10	VIII.C	<i>Xanthomonas conserved hypothetical</i>
01	08	01	01A08	XAC2405	02A04	A0AM1084A08	III.A.4	<i>DNA mismatch repair protein MutL</i>
01	09	01	01A09	XAC3649	01A05	A0CE0105B09	I.C.8	<i>ATP synthase, beta chain</i>
01	10	01	01A10	XAC1477	02A05	A0JE1044A10	III.A.1	<i>replicative DNA helicase</i>
01	11	01	01A11	XAC4119	01A06	A0AC0113B06	VIII.A	<i>conserved hypothetical protein</i>
01	12	01	01A12	XAC3312	02A06	A0JJ1004A08	I.A.1	<i>glycosyl hydrolase</i>
02	01	01	02A01	XAC3233	05A01	A0QH1645G09	VI.C	<i>transposase</i>
02	02	01	02A02	XAC3814	06A01	A0QR5311D08	VII.C	<i>multidrug efflux protein</i>
02	03	01	02A03	XACb0043	05A02	A0AM1594G01	VIII.B	<i>hypothetical protein</i>
02	04	01	02A04	XAC2939	06A02	A0CE5420H08	IX	<i>acetyltransferase</i>
02	05	01	02A05	XAC1328	05A03	A0JJ1626B12	VIII.A	<i>conserved hypothetical protein</i>
02	06	01	02A06	XAC2905	06A03	A0CE5420H09	III.A.2	<i>single-stranded DNA binding protein</i>
02	07	01	02A07	XAC2710	05A04	A0UV5307G04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	08	01	02A08	XAC4230	06A04	A0CE5422A09	I.A.1	<i>xylosidase/arabinosidase</i>
02	09	01	02A09	XAC1230	05A05	A0UV5307D10	VIII.A	<i>conserved hypothetical protein</i>
02	10	01	02A10	XAC3303	06A05	A0QR5529G08	III.A.4	<i>DNA mismatch repair protein</i>
02	11	01	02A11	XAC2005	05A06	A0QR5206F01	II.B.3	<i>thioredoxin reductase</i>
02	12	01	02A12	XAC4062	06A06	A0QR5529A02	V.A.7	<i>TonB-dependent receptor</i>
03	01	01	03A01	XAC3925	09A01	A0QH6307G05	VIII.A	<i>conserved hypothetical protein</i>
03	02	01	03A02	XAC2309	10A01	A0QR6377C08	V.A.3	<i>ABC transporter sugar permease</i>
03	03	01	03A03	XAC0405	09A02	A0CE6342H06	VII.B	<i>HrcV protein</i>
03	04	01	03A04	XAC2344	10A02	A0QR6367C11	VIII.A	<i>conserved hypothetical protein</i>
03	05	01	03A05	XAC1396	09A03	A0CE6341D04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>

03	06	01	03A06	XAC0687	10A03	A0QR6377E07	VIII.A	<i>conserved hypothetical protein</i>
03	07	01	03A07	XAC2113	09A04	A0CE6344A06	VIII.A	<i>conserved hypothetical protein</i>
03	08	01	03A08	XAC3358	10A04	A0QR6375H05	V.A.2	<i>molybdate-binding periplasmic protein; permease</i>
03	09	01	03A09	XAC3879	09A05	A0CE6341E05	I.C.3	<i>cytochrome C oxidase assembly factor</i>
03	10	01	03A10	XAC1391	10A05	A0AC6405C02	V.A.1	<i>S-methylmethionine permease</i>
03	11	01	03A11	XAC3298	09A06	A0CE6341E08	VI.A	<i>integrase</i>
03	12	01	03A12	XAC2191	10A06	A0QR6389D01	I.D.1	<i>two-component system, regulatory protein</i>
04	01	01	04A01	XAC1996	13A01	A0UE6612E05	V.C	<i>chemotaxis protein</i>
04	02	01	04A02	XAC1983	14A01	A0UV6711B02	V.C	<i>flagellar biosynthesis, hook protein</i>
04	03	01	04A03	XAC2672	13A02	A0QR6613G03	IV.A.2	<i>Oar protein</i>
04	04	01	04A04	XAC3977	14A02	A0UV6710D04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	05	01	04A05	XAC2772	13A03	A0UV6708A07	IV.A.2	<i>outer membrane protein</i>
04	06	01	04A06	XAC3425	14A03	A0UV6709G11	I.C.3	<i>cytochrome C6</i>
04	07	01	04A07	XAC3683	13A04	A0UV6708B12	I.D.3	<i>sensor histidine kinase</i>
04	08	01	04A08	XAC4236	14A04	A0UV6709C11	IX	<i>ring canal kelch - like protein</i>
04	09	01	04A09	XAC1772	13A05	A0UV6708A12	VIII.A	<i>conserved hypothetical protein</i>
04	10	01	04A10	XAC0548	14A05	A0UV6711C07	II.C	<i>gluconolactonase precursor</i>
04	11	01	04A11	XAC0512	13A06	A0UV6707B09	VIII.B	<i>hypothetical protein</i>
04	12	01	04A12	XAC0617	14A06	A0UV6710G03	VIII.B	<i>hypothetical protein</i>
05	01	01	05A01	XAC1207	17A01	A0UV6733D02	I.D.2	<i>transcriptional regulator araC/xylS family</i>
05	02	01	05A02	XAC4060	18A01	A0CE6740G01	V.A.7	<i>heavy metal transporter</i>
05	03	01	05A03	XAC3395	17A02	A0UV6733G12	II.B.1	<i>guanylate kinase</i>
05	04	01	05A04	XAC2949	18A02	A0CE6742C10	IX	<i>calcium-binding protein</i>
05	05	01	05A05	XAC1226	17A03	A0UV6736C11	V.B	<i>cell division inhibitor</i>
05	06	01	05A06	XAC1540	18A03	A0CE6742H08	VIII.B	<i>hypothetical protein</i>
05	07	01	05A07	XAC0717	17A04	A0UV6736H01	III.B.4	<i>tRNA nucleotidyltransferase</i>
05	08	01	05A08	XAC2954	18A04	A0CE6743G07	VIII.A	<i>conserved hypothetical protein</i>
05	09	01	05A09	XAC1857	17A05	A0UV6736H03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	10	01	05A10	XAC3870	18A05	A0CE6740E06	II.D.2	<i>dihydroneopterin aldolase</i>
05	11	01	05A11	XAC0372	17A06	A0QR6737F10	III.D.2	<i>hydrolase</i>
05	12	01	05A12	XAC0775	18A06	A0CE6743H10	IV.B	<i>UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase</i>
06	01	01	06A01	XAC3238	21A01	A0QH6810F04	I.D.1	<i>two-component system, regulatory protein</i>
06	02	01	06A02	XAC0085	22A01	A0UV6818F09	VIII.A	<i>conserved hypothetical protein</i>
06	03	01	06A03	XAC3987	21A02	A0QH6810F05	III.C.3	<i>leucine aminopeptidase</i>
06	04	01	06A04	XAC1120	22A02	A0UV6819G07	VIII.A	<i>conserved hypothetical protein</i>
06	05	01	06A05	XAC1571	21A03	A0UT6814E08	III.B.6	<i>ribonuclease T</i>
06	06	01	06A06	XAC3219	22A03	A0UV6819G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	07	01	06A07	XAC1330	21A04	A0UT6815B05	VIII.A	<i>conserved hypothetical protein</i>
06	08	01	06A08	XAC4188	22A04	A0UV6821D06	IX	<i>RTS beta protein</i>
06	09	01	06A09	XAC4235	21A05	A0UT6816D03	VIII.B	<i>hypothetical protein</i>
06	10	01	06A10	XAC3473	22A05	A0UV6821D09	I.D.3	<i>sensor histidine kinase</i>
06	11	01	06A11	XAC0469	21A06	A0UT6817B01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	12	01	06A12	XAC0479	22A06	A0AC6825E12	VIII.A	<i>conserved hypothetical protein</i>
07	01	01	07A01	XAC2424	25A01	A0CE9725H09	VI.C	<i>ISxcd1 transposase</i>
07	02	01	07A02	XAC1206	26A01	A0UE9829C11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	01	07A03	XAC1972	25A02	A0CE9728H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	04	01	07A04	XAC0651	26A02	A0UV9803D04	IV.C	<i>surface antigen gene</i>
07	05	01	07A05	XAC2620	25A03	A0QH9737H04	VII.H	<i>VirB9 protein</i>
07	06	01	07A06	XAC3064	26A03	A0UV9805C07	VIII.A	<i>conserved hypothetical protein</i>
07	07	01	07A07	XAC0633	25A04	A0QR9739H11	VIII.A	<i>conserved hypothetical protein</i>
07	08	01	07A08	XAC0896	26A04	A0QR9814B08	I.D.1	<i>two-component system, regulatory protein</i>
07	09	01	07A09	XAC3069	25A05	A0QH9738F03	VIII.A	<i>conserved hypothetical protein</i>
07	10	01	07A10	XAC1354	26A05	A0QR9815E05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	11	01	07A11	XAC0190	25A06	A0QR9776E12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>

07	12	01	07A12	XAC0762	26A06	A0QR9814G03	I.C.2	<i>D-lactate dehydrogenase</i>
08	01	01	08A01	XAC1090	29A01	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	02	01	08A02	XAC1265	30A01	genomic_DNA	VII.B	<i>HrpG protein</i>
08	03	01	08A03	XAC3245	29A02	genomic_DNA	IX	<i>RhsD protein</i>
08	04	01	08A04	XAC1088	30A02	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	05	01	08A05	XAC0674	29A03	genomic_DNA	IX	<i>pre-B cell enhancing factor related protein</i>
08	06	01	08A06	XAC0095	30A03	genomic_DNA	VII.B	<i>HrcS protein</i>
08	07	01	08A07	XAC2605	29A04	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	08	01	08A08	XAC1088	30A04	genomic_DNA	VII.B	<i>HrcN protein</i>
08	09	01	08A09	XAC2618	29A05	genomic_DNA	VII.H	<i>VirB11 protein</i>
08	10	01	08A10	XAC1879	30A05	genomic_DNA	VII.H	<i>RpfF protein</i>
08	11	01	08A11	XAC2610	29A06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	01	08A12	no_clone	30A06	no_clone	NC	<i>no_gene</i>
09	01	01	09A01	Score_card	33A01	Score_card	SC	<i>Score_card</i>
09	02	01	09E17	Score_card	33C09	Score_card	SC	<i>Score_card</i>
09	03	01	09K09	Score_card	33F05	Score_card	SC	<i>Score_card</i>
09	04	01	09A02	Score_card	34A01	Score_card	SC	<i>Score_card</i>
09	05	01	09E18	Score_card	34C09	Score_card	SC	<i>Score_card</i>
09	06	01	09K10	Score_card	34F05	Score_card	SC	<i>Score_card</i>
09	07	01	09B02	Score_card	35A01	Score_card	SC	<i>Score_card</i>
09	08	01	09F18	Score_card	35C09	Score_card	SC	<i>Score_card</i>
09	09	01	09L10	Score_card	35F05	Score_card	SC	<i>Score_card</i>
09	10	01	09B01	Score_card	36A01	Score_card	SC	<i>Score_card</i>
09	11	01	09F17	Score_card	36C09	Score_card	SC	<i>Score_card</i>
09	12	01	09L09	Score_card	36F05	Score_card	SC	<i>Score_card</i>
01	01	10	01J01	XAC3712	04E01	A0JJ1577F08	III.C.3	<i>metallopeptidase</i>
01	02	10	01J02	XAC3486	03E01	A0JE1208C11	II.E	<i>acetoacetyl-CoA reductase</i>
01	03	10	01J03	XAC2069	04E02	A0JJ1546A01	I.C.5	<i>6-phosphogluconolactonase</i>
01	04	10	01J04	XAC4250	03E02	A0RN1193G11	I.A.1	<i>beta-galactosidase</i>
01	05	10	01J05	XAC3148	04E03	A0UV1501F08	V.A.4	<i>potassium uptake protein</i>
01	06	10	01J06	XAC0223	03E03	A0JJ1371F07	VIII.A	<i>conserved hypothetical protein</i>
01	07	10	01J07	XAC2614	04E04	A0AM1489E07	VII.H	<i>VirB4 protein</i>
01	08	10	01J08	XAC2068	03E04	A0JJ1389E04	I.B.2	<i>6-phosphogluconate dehydratase</i>
01	09	10	01J09	XAC4162	04E05	A0AM1490C05	VII.C	<i>cation efflux system protein</i>
01	10	10	01J10	XAC1925	03E05	A0AM1363D10	I.D.2	<i>transcriptional regulator</i>
01	11	10	01J11	XAC2504	04E06	A0UV1502G04	VII.H	<i>regulator of pathogenicity factors</i>
01	12	10	01J12	XAC0553	03E06	A0AC1335A10	VIII.A	<i>conserved hypothetical protein</i>
02	01	10	02J01	XAC3708	08E01	A0QR6002D08	I.D.2	<i>transcriptional regulator</i>
02	02	10	02J02	XAC0193	07E01	A0QR5702C06	VIII.A	<i>conserved hypothetical protein</i>
02	03	10	02J03	XAC2623	08E02	A0QR6204E07	VII.H	<i>VirD4 protein</i>
02	04	10	02J04	XAC0194	07E02	A0QR5701E10	VIII.A	<i>conserved hypothetical protein</i>
02	05	10	02J05	XAC2452	08E03	A0QR6002E01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	06	10	02J06	XAC4030	07E03	A0QR5702D01	VII.C	<i>catalase</i>
02	07	10	02J07	XAC1867	08E04	A0QR6106H05	VIII.A	<i>conserved hypothetical protein</i>
02	08	10	02J08	XAC3799	07E04	A0QR5704B09	IX	<i>Sun protein</i>
02	09	10	02J09	XAC3525	08E05	A0QR6204F07	VIII.A	<i>conserved hypothetical protein</i>
02	10	10	02J10	XAC2220	07E05	A0QR5704E08	VIII.B	<i>hypothetical protein</i>
02	11	10	02J11	XAC3883	08E06	A0UV6207H05	VIII.A	<i>conserved hypothetical protein</i>
02	12	10	02J12	XAC1063	07E06	A0QR5704E09	VI.A	<i>phage-related lysozyme</i>
03	01	10	03J01	XAC1807	12E01	A0UV6485F05	VIII.A	<i>conserved hypothetical protein</i>
03	02	10	03J02	XAC3662	11E01	A0EC6427H05	VIII.A	<i>conserved hypothetical protein</i>
03	03	10	03J03	XAC0013	12E02	A0CE6504G10	VIII.A	<i>conserved hypothetical protein</i>
03	04	10	03J04	XAC2885	11E02	A0UV6435G05	III.D.2	<i>phospholipase A1</i>
03	05	10	03J05	XAC0287	12E03	A0QR6481F04	I.C.3	<i>quinone oxidoreductase</i>
03	06	10	03J06	XAC3961	11E03	A0UV6434F04	I.D.2	<i>transcriptional regulator tetR family</i>

03	07	10	03J07	XAC2004	12E04	A0UV6512A06	VIII.A	<i>conserved hypothetical protein</i>
03	08	10	03J08	XAC2445	11E04	A0UV6436A10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	09	10	03J09	XAC1780	12E05	A0UV6485C08	III.D.1	<i>N-acetylmuramoyl-L-alanine amidase</i>
03	10	10	03J10	XAC1413	11E05	A0AC6446H11	IV.A.2	<i>outer membrane antigen</i>
03	11	10	03J11	XAC1340	12E06	A0QR6490F04	III.A.1	<i>helicase-related protein</i>
03	12	10	03J12	XAC0577	11E06	A0UV6455E06	III.A.5	<i>cytosine methyltransferase</i>
04	01	10	04J01	XAC2124	16E01	A0QR6730H12	VIII.A	<i>conserved hypothetical protein</i>
04	02	10	04J02	XAC2990	15E01	A0QR6724G05	VIII.B	<i>hypothetical protein</i>
04	03	10	04J03	XAC0694	16E02	A0QR6730D07	VII.H	<i>type II secretion system protein C</i>
04	04	10	04J04	XAC2960	15E02	A0QR6725B05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	05	10	04J05	XAC2141	16E03	A0QR6730C04	I.D.1	<i>two-component system, regulatory protein</i>
04	06	10	04J06	XAC0473	15E03	A0QR6725B06	IV.A.1	<i>membrane protein</i>
04	07	10	04J07	XAC3423	16E04	A0UV6733A12	II.A.5	<i>histidinol-phosphate aminotransferase</i>
04	08	10	04J08	XAC0938	15E04	A0QR6730A07	II.D.10	<i>thioredoxin</i>
04	09	10	04J09	XAC2975	16E05	A0UV6733B04	V.A.3	<i>HPr kinase/phosphatase</i>
04	10	10	04J10	XAC2408	15E05	A0QR6728E02	VIII.A	<i>conserved hypothetical protein</i>
04	11	10	04J11	XAC0593	16E06	A0QH6732E10	VIII.A	<i>conserved hypothetical protein</i>
04	12	10	04J12	XAC4032	15E06	A0QR6728E11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	01	10	05J01	XAC2959	20E01	A0UV6766H12	II.B.1	<i>phosphoribosylformylglycinamide cyclo-ligase</i>
05	02	10	05J02	XAC1480	19E01	A0UE6759F11	I.D.2	<i>transcriptional regulator</i>
05	03	10	05J03	XAC0907	20E02	A0QR6768A02	VII.C	<i>alkyl hydroperoxide reductase subu</i>
05	04	10	05J04	XAC1919	19E02	A0UV6761B10	VIII.B	<i>hypothetical protein</i>
05	05	10	05J05	XAC4029	20E03	A0UV6766G10	VII.C	<i>catalase [precursor]</i>
05	06	10	05J06	XAC2913	19E03	A0UV6761D09	VIII.A	<i>conserved hypothetical protein</i>
05	07	10	05J07	XAC2030	20E04	A0QR6768A07	III.A.4	<i>exodeoxyribonuclease III</i>
05	08	10	05J08	XAC2996	19E04	A0UE6759A08	VIII.A	<i>conserved hypothetical protein</i>
05	09	10	05J09	XAC0150	20E05	A0QR6801A02	VIII.B	<i>hypothetical protein</i>
05	10	10	05J10	XAC0126	19E05	A0UV6763B05	V.A.7	<i>iron transporter</i>
05	11	10	05J11	XAC1101	20E06	A0QR6802H12	III.C.2	<i>heat shock protein G homolog</i>
05	12	10	05J12	XAC3252	19E06	A0UV6763B11	III.B.3	<i>ribosomal protein S6 modification protein</i>
06	01	10	06J01	XAC1173	24E01	A0UV6844E10	I.D.2	<i>transcriptional regulator uid family</i>
06	02	10	06J02	XAC0603	23E01	A0AC6825C11	VIII.A	<i>conserved hypothetical protein</i>
06	03	10	06J03	XAC0065	24E02	A0QR6840H03	IX	<i>microcystin dependent protein</i>
06	04	10	06J04	XAC1484	23E02	A0AC6824C10	I.C.3	<i>short chain dehydrogenase</i>
06	05	10	06J05	XAC1200	24E03	A0QR6840H04	III.C.3	<i>prolyl oligopeptidase family protein</i>
06	06	10	06J06	XAC1245	23E03	A0CE6828B09	VIII.A	<i>conserved hypothetical protein</i>
06	07	10	06J07	XAC1292	24E04	A0QR9705B04	III.B.2	<i>30S ribosomal protein S16</i>
06	08	10	06J08	XAC2352	23E04	A0CE6830B04	II.A.1	<i>ornithine carbamoyltransferase</i>
06	09	10	06J09	XAC3316	24E05	A0UV9712H02	III.B.4	<i>tRNA/rRNA methyltransferase</i>
06	10	10	06J10	XAC4192	23E05	A0CE6831A06	VIII.A	<i>conserved hypothetical protein</i>
06	11	10	06J11	XAC1034	24E06	A0UV6845D06	III.C.3	<i>peptidyl-Asp metalloendopeptidase</i>
06	12	10	06J12	XAC4082	23E06	A0QH6836E02	VIII.A	<i>conserved hypothetical protein</i>
07	01	10	07J01	XAC3211	28E01	A0UV9900B12	VII.G	<i>trehalose-6-phosphate synthase</i>
07	02	10	07J02	XAC1842	27E01	A0CE9840B04	V.A.1	<i>cationic amino acid transporter</i>
07	03	10	07J03	XAC2851	28E02	A0UV9905A07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	04	10	07J04	XAC1718	27E02	A0QH9865C01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	05	10	07J05	XAC4233	28E03	A0UV9905A10	VII.C	<i>bleomycin resistance protein</i>
07	06	10	07J06	XAC3938	27E03	A0UE9858A03	VI.C	<i>ISxac3 transposase</i>
07	07	10	07J07	XAC3415	28E04	A0AC9896F12	II.D.8	<i>thiamin-phosphate pyrophosphorylase</i>
07	08	10	07J08	XAC0038	27E04	A0QR9851A02	VIII.A	<i>conserved hypothetical protein</i>
07	09	10	07J09	XAC2366	28E05	A0UV9901F11	I.A.2	<i>ethanolamine ammonia-lyase light chain</i>
07	10	10	07J10	XAC2753	27E05	A0UV9872C09	III.D.3	<i>lipoprotein</i>
07	11	10	07J11	XAC1899	28E06	A0UT9891E08	V.C	<i>chemotaxis protein</i>
07	12	10	07J12	XAC0838	27E06	A0UV9871G08	VIII.A	<i>conserved hypothetical protein</i>
08	01	10	08J01	no_clone	32E01	no_clone	NC	<i>no_gene</i>

08	02	10	08J02	no_clone	31E01	no_clone	NC	no_gene
08	03	10	08J03	no_clone	32E02	no_clone	NC	no_gene
08	04	10	08J04	no_clone	31E02	no_clone	NC	no_gene
08	05	10	08J05	no_clone	32E03	no_clone	NC	no_gene
08	06	10	08J06	no_clone	31E03	no_clone	NC	no_gene
08	07	10	08J07	no_clone	32E04	no_clone	NC	no_gene
08	08	10	08J08	no_clone	31E04	no_clone	NC	no_gene
08	09	10	08J09	no_clone	32E05	no_clone	NC	no_gene
08	10	10	08J10	no_clone	31E05	no_clone	NC	no_gene
08	11	10	08J11	no_clone	32E06	no_clone	NC	no_gene
08	12	10	08J12	no_clone	31E06	no_clone	NC	no_gene
09	01	10	09A19	Score_card	33A10	Score_card	SC	Score_card
09	02	10	09G11	Score_card	33D06	Score_card	SC	Score_card
09	03	10	09M03	Score_card	33G02	Score_card	SC	Score_card
09	04	10	09A20	Score_card	34A10	Score_card	SC	Score_card
09	05	10	09G12	Score_card	34D06	Score_card	SC	Score_card
09	06	10	09M04	Score_card	34G02	Score_card	SC	Score_card
09	07	10	09B20	Score_card	35A10	Score_card	SC	Score_card
09	08	10	09H12	Score_card	35D06	Score_card	SC	Score_card
09	09	10	09N04	Score_card	35G02	Score_card	SC	Score_card
09	10	10	09B19	Score_card	36A10	Score_card	SC	Score_card
09	11	10	09H11	Score_card	36D06	Score_card	SC	Score_card
09	12	10	09N03	Score_card	36G02	Score_card	SC	Score_card
01	01	11	01K01	XAC2787	01F01	A0AC0113C03	VIII.B	hypothetical protein
01	02	11	01K02	XAC2332	02F01	A0JJ1065G04	II.A.2	homoserine O-acetyltransferase
01	03	11	01K03	XAC3034	01F02	A0JJ0102F06	VIII.A	conserved hypothetical protein
01	04	11	01K04	XAC3869	02F02	A0JE1079C03	I.A.1	beta-glucosidase
01	05	11	01K05	XAC3455	01F03	A0JJ0101D12	II.A.2	2-isopropylmalate synthase
01	06	11	01K06	XAC1944	02F03	A0JE1041F05	V.C	flagellar biosynthetic protein
01	07	11	01K07	XAC3289	01F04	A0JJ0101E01	VIII.C	Xanthomonas conserved hypothetical protein
01	08	11	01K08	XAC1140	02F04	A0JJ1003G11	VIII.A	conserved hypothetical protein
01	09	11	01K09	XAC4016	01F05	A0AC0115F09	VIII.A	conserved hypothetical protein
01	10	11	01K10	XAC0083	02F05	A0EC0204C09	I.C.3	short chain dehydrogenase
01	11	11	01K11	XAC1636	01F06	A0JJ0102G11	I.A.2	formylglutamate amidohydrolase
01	12	11	01K12	XAC0286	02F06	A0JJ1085D10	VII.A	avirulence protein
02	01	11	02K01	XAC2021	05F01	A0RN1634F09	IX	GTP-binding protein
02	02	11	02K02	XAC3167	06F01	A0QR5407C11	VIII.A	conserved hypothetical protein
02	03	11	02K03	XAC1168	05F02	A0QR5103F10	VIII.A	conserved hypothetical protein
02	04	11	02K04	XAC0944	06F02	A0UV5523D08	III.C.1	peptide chain release factor 1
02	05	11	02K05	XAC3024	05F03	A0QR5206B05	VIII.B	hypothetical protein
02	06	11	02K06	XAC3824	06F03	A0QR5520F08	I.D.4	RNA polymerase sigma-32 factor
02	07	11	02K07	XAC3080	05F04	A0QR5205C09	I.A.2	ribokinase
02	08	11	02K08	XAC1796	06F04	A0UV5523H05	I.A.1	mannan endo-1,4-beta-mannosidase
02	09	11	02K09	XAC2044	05F05	A0UV5307A02	III.A.4	7,8-dihydro-8-oxoguanine-triphosphatase
02	10	11	02K10	XAC4342	06F05	A0QR5521D08	VII.G	toluene tolerance protein
02	11	11	02K11	XAC3389	05F06	A0UV5307A09	III.B.2	50S ribosomal protein L31
02	12	11	02K12	XAC0766	06F06	A0UV5525H03	VIII.A	conserved hypothetical protein
03	01	11	03K01	XAC2126	09F01	A0EC6321E11	VIII.A	conserved hypothetical protein
03	02	11	03K02	XAC3283	10F01	A0AC6357D06	VI.C	ISxac2 transposase
03	03	11	03K03	XAC0153	09F02	A0QH6326F12	VI.C	ISxac3 transposase
03	04	11	03K04	XAC0743	10F02	A0AC6358D03	II.A.3	serine hydroxymethyltransferase
03	05	11	03K05	XAC3817	09F03	A0QR6334A05	VIII.A	conserved hypothetical protein
03	06	11	03K06	XAC3179	10F03	A0UV6362C12	V.A.7	transport protein
03	07	11	03K07	XAC2076	09F04	A0QH6327E01	I.C.7	succinate dehydrogenase, membrane anchor subunit
03	08	11	03K08	XAC2158	10F04	A0QR6390H12	I.D.3	histidine kinase/response regulator hybrid protein

03	09	11	03K09	XAC1977	09F05	A0QH6326A08	V.C	<i>flagellar protein</i>
03	10	11	03K10	XAC3063	10F05	A0AC6408C01	VIII.A	<i>conserved hypothetical protein</i>
03	11	11	03K11	XAC0009	09F06	A0QH6327F01	VII.C	<i>biopolymer transport ExbB protein</i>
03	12	11	03K12	XAC1419	10F06	A0QH6382C03	I.B.8	<i>uridylyate kinase</i>
04	01	11	04K01	XACa0041	13F01	A0QR6704E11	VI.B	<i>partition protein A</i>
04	02	11	04K02	XAC0199	14F01	A0UV6709F06	VIII.A	<i>conserved hypothetical protein</i>
04	03	11	04K03	XAC1132	13F02	A0QR6705C03	III.A.1	<i>DNA polymerase III, delta' subunit</i>
04	04	11	04K04	XAC0674	14F02	A0UV6711H03	IX	<i>pre-B cell enhancing factor related protein</i>
04	05	11	04K05	XAC2263	13F03	A0UV6707B04	VIII.B	<i>hypothetical protein</i>
04	06	11	04K06	XAC0149	14F03	A0UV6711E03	VIII.B	<i>hypothetical protein</i>
04	07	11	04K07	XAC2756	13F04	A0QR6705E04	II.E	<i>acyl-CoA thioester hydrolase</i>
04	08	11	04K08	XAC4142	14F04	A0UV6710D09	VIII.A	<i>conserved hypothetical protein</i>
04	09	11	04K09	XAC0877	13F05	A0QR6705D03	I.D.2	<i>transcriptional regulator gntR family</i>
04	10	11	04K10	XAC2127	14F05	A0UV6709H05	VIII.B	<i>hypothetical protein</i>
04	11	11	04K11	XAC0807	13F06	A0QR6705D04	I.D.2	<i>transcriptional regulator tetR/acrR family</i>
04	12	11	04K12	XAC0880	14F06	A0UV6709G01	I.D.2	<i>transcriptional regulator</i>
05	01	11	05K01	XAC3826	17F01	A0UV6735H06	IX	<i>response regulator protein</i>
05	02	11	05K02	XAC4033	18F01	A0CE6740D02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	03	11	05K03	XAC0501	17F02	A0UV6735E05	VIII.A	<i>conserved hypothetical protein</i>
05	04	11	05K04	XAC0935	18F02	A0CE6740D11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	05	11	05K05	XAC0668	17F03	A0UV6736B07	II.D.3	<i>lipoic acid synthetase</i>
05	06	11	05K06	XAC0259	18F03	A0CE6743A02	VIII.A	<i>conserved hypothetical protein</i>
05	07	11	05K07	XAC2713	17F04	A0UV6736B08	I.C.3	<i>oxidoreductase</i>
05	08	11	05K08	XAC3800	18F04	A0CE6743C01	III.B.4	<i>10-Formyltetrahydrofolate:L-methionyl-tRNA(fMet) N-formyltransferase</i>
05	09	11	05K09	XAC1163	17F05	A0UV6736E11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	10	11	05K10	XAC3859	18F05	A0CE6743F06	III.C.3	<i>D-alanyl-D-alanine dipeptidase</i>
05	11	11	05K11	XAC4111	17F06	A0UV6739F07	VIII.A	<i>conserved hypothetical protein</i>
05	12	11	05K12	XAC2151	18F06	A0CE6746F11	VII.F	<i>YapH protein</i>
06	01	11	06K01	XAC3648	21F01	A0QR6804H10	I.C.8	<i>ATP synthase, epsilon chain</i>
06	02	11	06K02	XAC1403	22F01	A0UV6818F10	VIII.B	<i>hypothetical protein</i>
06	03	11	06K03	XAC3647	21F02	A0UT6817D11	II.A.4	<i>chorismate mutase/prephenate dehydratase</i>
06	04	11	06K04	XAC1954	22F02	A0AC6822C03	V.C	<i>flagellar protein</i>
06	05	11	06K05	XAC2669	21F03	A0UV6818C11	IV.A.2	<i>pre-pilin like leader sequence</i>
06	06	11	06K06	XAC0815	22F03	A0UV6820D05	IX	<i>methyltransferase</i>
06	07	11	06K07	XAC3462	21F04	A0UT6814F08	III.C.1	<i>L-isoaspartate protein carboxymethyltransferase</i>
06	08	11	06K08	XAC1259	22F04	A0UV6819F02	I.C.1	<i>cytochrome O ubiquinol oxidase, subunit I</i>
06	09	11	06K09	XAC0064	21F05	A0QH6813G09	IX	<i>acetyltransferase</i>
06	10	11	06K10	XAC1237	22F05	A0AC6824C12	II.B.1	<i>phosphoribosylglycinamide formyltransferase 2</i>
06	11	11	06K11	XAC4260	21F06	A0UT6814G02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	12	11	06K12	XAC4051	22F06	A0AC6824A06	II.D.17	<i>tryptophan halogenase</i>
07	01	11	07K01	XAC2934	25F01	A0UV9747G02	VIII.A	<i>conserved hypothetical protein</i>
07	02	11	07K02	XAC0225	26F01	A0UV9820H12	I.D.1	<i>two-component system, sensor protein</i>
07	03	11	07K03	XAC0873	25F02	A0CE9723F08	IX	<i>VisC protein</i>
07	04	11	07K04	XAC1998	26F02	A0UV9822A04	III.B.4	<i>tRNA methyltransferase</i>
07	05	11	07K05	XAC3010	25F03	A0QH9730E12	II.A.4	<i>shikimate kinase</i>
07	06	11	07K06	XAC2883	26F03	A0UE9829H01	IX	<i>hydrolase</i>
07	07	11	07K07	XAC0270	25F04	A0CE9724C12	VIII.A	<i>conserved hypothetical protein</i>
07	08	11	07K08	XAC1894	26F04	A0UV9804A04	V.C	<i>chemotaxis protein</i>
07	09	11	07K09	XAC3607	25F05	A0QR9753D06	IV.A.2	<i>type II secretion system protein-like protein</i>
07	10	11	07K10	XAC1094	26F05	A0UE9829C03	IV.C	<i>saccharide biosynthesis regulatory protein</i>
07	11	11	07K11	XAC3027	25F06	A0QR9776F05	V.A.7	<i>MFS transporter</i>
07	12	11	07K12	XAC2434	26F06	A0UE9829C06	VIII.A	<i>conserved hypothetical protein</i>
08	01	11	08K01	XAC1495	29F01	genomic_DNA	VII.H	<i>virulence regulator</i>
08	02	11	08K02	XAC0330	30F01	genomic_DNA	VII.B	<i>HpaP protein</i>

08	03	11	08K03	XAC2555	29F02	genomic_DNA	I.D.3	<i>sensor histidine kinase</i>
08	04	11	08K04	XAC0095	30F02	genomic_DNA	VII.B	<i>HrpF protein</i>
08	05	11	08K05	XAC3273	29F03	genomic_DNA	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
08	06	11	08K06	XAC0095	30F03	genomic_DNA	VII.B	<i>HrpB2 protein</i>
08	07	11	08K07	XACb0035	29F04	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	08	11	08K08	XAC1877	30F04	genomic_DNA	VII.H	<i>response regulator</i>
08	09	11	08K09	XAC2617	29F05	genomic_DNA	VII.H	<i>VirB1 protein</i>
08	10	11	08K10	XAC1874	30F05	genomic_DNA	VII.H	<i>response regulator</i>
08	11	11	08K11	XACb0026	29F06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	11	08K12	no_clone	30F06	no_clone	NC	<i>no_gene</i>
09	01	11	09A21	Score_card	33A11	Score_card	SC	<i>Score_card</i>
09	02	11	09G13	Score_card	33D07	Score_card	SC	<i>Score_card</i>
09	03	11	09M05	Score_card	33G03	Score_card	SC	<i>Score_card</i>
09	04	11	09A22	Score_card	34A11	Score_card	SC	<i>Score_card</i>
09	05	11	09G14	Score_card	34D07	Score_card	SC	<i>Score_card</i>
09	06	11	09M06	Score_card	34G03	Score_card	SC	<i>Score_card</i>
09	07	11	09B22	Score_card	35A11	Score_card	SC	<i>Score_card</i>
09	08	11	09H14	Score_card	35D07	Score_card	SC	<i>Score_card</i>
09	09	11	09N06	Score_card	35G03	Score_card	SC	<i>Score_card</i>
09	10	11	09B21	Score_card	36A11	Score_card	SC	<i>Score_card</i>
09	11	11	09H13	Score_card	36D07	Score_card	SC	<i>Score_card</i>
09	12	11	09N05	Score_card	36G03	Score_card	SC	<i>Score_card</i>
01	01	12	01L01	XAC3487	04F01	A0JJ1442E11	I.D.2	<i>transcriptional regulator</i>
01	02	12	01L02	XAC3655	03F01	A0JE1211A02	I.C.8	<i>ATP synthase, A chain</i>
01	03	12	01L03	XAC3315	04F02	A0JJ1553D03	IX	<i>carboxylesterase</i>
01	04	12	01L04	XAC2516	03F02	A0JE1208D07	I.A.2	<i>L-lysine 6-aminotransferase</i>
01	05	12	01L05	XAC2341	04F03	A0UT1521D03	VII.C	<i>glutaryl-7-ACA acylase precursor</i>
01	06	12	01L06	XAC2254	03F03	A0AM1402D10	VIII.B	<i>hypothetical protein</i>
01	07	12	01L07	XAC2980	04F04	A0UV1502C04	V.A.4	<i>Mg++ transporter</i>
01	08	12	01L08	XAC1958	03F04	A0AM1402E11	VIII.A	<i>conserved hypothetical protein</i>
01	09	12	01L09	XAC0904	04F05	A0UV1502E01	VIII.A	<i>conserved hypothetical protein</i>
01	10	12	01L10	XAC3011	03F05	A0QH1368B09	II.A.4	<i>3-dehydroquinase synthase</i>
01	11	12	01L11	XAC2122	04F06	A0JJ1537D08	I.C.3	<i>dehydrogenase</i>
01	12	12	01L12	XAC2600	03F06	A0UV1356E10	V.A.7	<i>TonB-dependent receptor</i>
02	01	12	02L01	XAC1969	08F01	A0QR6106E06	I.D.4	<i>RNA polymerase sigma-54 factor</i>
02	02	12	02L02	XAC3323	07F01	A0QR5702D10	IX	<i>acidic amino acid rich protein</i>
02	03	12	02L03	XAC0488	08F02	A0UV6205H02	III.B.2	<i>30S ribosomal protein S9</i>
02	04	12	02L04	XAC3565	07F02	A0QR5702C10	III.D.2	<i>phosphatidylserine synthase</i>
02	05	12	02L05	XAC2808	08F03	A0QR6106H04	VIII.A	<i>conserved hypothetical protein</i>
02	06	12	02L06	XAC0443	07F03	A0QR5704B06	I.C.6	<i>dihydrolipoamide acyltransferase</i>
02	07	12	02L07	XAC0457	08F04	A0UV6206E05	VIII.A	<i>conserved hypothetical protein</i>
02	08	12	02L08	XAC2079	07F04	A0QR5704G12	VIII.A	<i>conserved hypothetical protein</i>
02	09	12	02L09	XAC2523	08F05	A0UV6206F11	II.D.10	<i>gamma-glutamyltranspeptidase</i>
02	10	12	02L10	XAC0888	07F05	A0QR5704F07	V.D	<i>glucose-fructose oxidoreductase</i>
02	11	12	02L11	XAC2880	08F06	A0QR6303E06	III.A.2	<i>pirin</i>
02	12	12	02L12	XAC3440	07F06	A0QR5704H04	I.B.9	<i>H+ translocating pyrophosphate synthase</i>
03	01	12	03L01	XAC2691	12F01	A0UV6486F10	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO14 subunit</i>
03	02	12	03L02	XAC0991	11F01	A0AC6430D09	III.B.2	<i>50S ribosomal protein L15</i>
03	03	12	03L03	XAC1169	12F02	A0CE6507B08	VIII.B	<i>hypothetical protein</i>
03	04	12	03L04	XAC0583	11F02	A0AC6439C06	I.C.3	<i>oxidoreductase</i>
03	05	12	03L05	XAC1527	12F03	A0UV6485H08	VII.C	<i>RND efflux membrane fusion protein</i>
03	06	12	03L06	XAC4283	11F03	A0UV6435H03	I.D.3	<i>sensor histidine kinase</i>
03	07	12	03L07	XAC3545	12F04	A0UV6485C05	III.C.3	<i>protease</i>
03	08	12	03L08	XAC3393	11F04	A0AC6446A01	I.D.4	<i>pentaphosphate guanosine-3'-pyrophosphohydrolase</i>
03	09	12	03L09	XAC3146	12F05	A0UV6486B07	VIII.A	<i>conserved hypothetical protein</i>

03	10	12	03L10	XAC0472	11F05	A0AC6447F05	I.B.6	<i>D-ribulose-5-phosphate 3-epimerase</i>
03	11	12	03L11	XAC3015	12F06	A0AC6495A07	III.A.3	<i>RebB protein</i>
03	12	12	03L12	XAC2833	11F06	A0UV6458D01	III.C.3	<i>extracellular serine protease</i>
04	01	12	04L01	XAC0328	16F01	A0QH6731B06	VII.C	<i>multidrug efflux transporter</i>
04	02	12	04L02	XAC2046	15F01	A0UV6712C03	III.D.2	<i>CDP-diacylglycerol--serine o-phosphase, cytochrome C1 subunit</i>
04	03	12	04L03	XAC3693	16F02	A0QR6730E11	V.C	<i>chemotaxis protein</i>
04	04	12	04L04	XAC1470	15F02	A0UV6712C04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	05	12	04L05	XAC3031	16F03	A0QR6730D08	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
04	06	12	04L06	XAC2007	15F03	A0QR6725E09	VIII.A	<i>conserved hypothetical protein</i>
04	07	12	04L07	XAC3045	16F04	A0QR6730C05	VIII.A	<i>conserved hypothetical protein</i>
04	08	12	04L08	XAC1489	15F04	A0QR6725H12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	09	12	04L09	XAC0040	16F05	A0QR6730C07	VIII.B	<i>hypothetical protein</i>
04	10	12	04L10	XAC0875	15F05	A0QR6728H07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	11	12	04L11	XAC0917	16F06	A0UV6733A03	I.D.2	<i>transcriptional regulator</i>
04	12	12	04L12	XAC3546	15F06	A0QR6728H08	VII.F	<i>outer membrane protein</i>
05	01	12	05L01	XAC4308	20F01	A0QR6768D08	V.A.7	<i>dicarboxylate transport protein</i>
05	02	12	05L02	XAC1688	19F01	A0UE6759H09	VIII.A	<i>conserved hypothetical protein</i>
05	03	12	05L03	XAC2816	20F02	A0QR6770C02	VIII.A	<i>conserved hypothetical protein</i>
05	04	12	05L04	XAC3734	19F02	A0UV6761H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	05	12	05L05	XAC3659	20F03	A0QR6768E09	I.C.6	<i>dihydrolipoamide dehydrogenase</i>
05	06	12	05L06	XAC3558	19F03	A0UV6761H11	III.A.1	<i>DNA polymerase III holoenzyme chi subunit</i>
05	07	12	05L07	XAC4094	20F04	A0QR6768F01	IV.A.1	<i>membrane protein</i>
05	08	12	05L08	XAC3654	19F04	A0UE6759C03	I.C.8	<i>ATP synthase, C chain</i>
05	09	12	05L09	XAC3755	20F05	A0QR6802A08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	10	12	05L10	XAC4107	19F05	A0UV6763G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	11	12	05L11	XAC2508	20F06	A0QR6803A04	VI.C	<i>transposase</i>
05	12	12	05L12	XAC0385	19F06	A0UV6764C05	II.D.1	<i>biotin biosynthesis protein</i>
06	01	12	06L01	XAC1743	24F01	A0QR6839D09	I.D.2	<i>carbon storage regulator</i>
06	02	12	06L02	XAC0260	23F01	A0AC6825E09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	03	12	06L03	XAC0770	24F02	A0UV6843B06	VIII.A	<i>conserved hypothetical protein</i>
06	04	12	06L04	XAC2278	23F02	A0AC6824F02	VIII.A	<i>conserved hypothetical protein</i>
06	05	12	06L05	XAC2112	24F03	A0UV6843B12	III.B.5	<i>pseudouridylate synthase</i>
06	06	12	06L06	XAC2951	23F03	A0CE6830E04	V.A.5	<i>DNA transport competence protein</i>
06	07	12	06L07	XAC1599	24F04	A0QR9705F04	III.A.4	<i>exodeoxyribonuclease I</i>
06	08	12	06L08	XAC2663	23F04	A0CE6830E06	VI.C	<i>transposase</i>
06	09	12	06L09	XAC1263	24F05	A0UV6845D04	III.A.4	<i>DNA repair protein</i>
06	10	12	06L10	XAC2105	23F05	A0CE6831E07	IV.A.1	<i>polysaccharide biosynthetic protein</i>
06	11	12	06L11	XAC1831	24F06	A0UV6845F04	II.A.5	<i>imidazoleglycerolphosphate dehydratase/ histidinol-phosphate phosphatase bifunctional enzyme</i>
06	12	12	06L12	XAC1824	23F06	A0CE6829B06	VIII.A	<i>conserved hypothetical protein</i>
07	01	12	07L01	XAC2952	28F01	A0UV9900H03	II.A.2	<i>succinyl-diaminopimelate desuccinylase</i>
07	02	12	07L02	XAC1843	27F01	A0CE9842A09	VIII.A	<i>conserved hypothetical protein</i>
07	03	12	07L03	XAC1776	28F02	A0UT9891B09	I.A.2	<i>xylose isomerase</i>
07	04	12	07L04	XAC2622	27F02	A0UV9871D06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	05	12	07L05	XAC0347	28F03	A0UT9894B03	III.C.3	<i>proteinase inhibitor</i>
07	06	12	07L06	XAC1456	27F03	A0QH9865D07	III.C.3	<i>peptidyl-dipeptidase</i>
07	07	12	07L07	XAC1521	28F04	A0UV9900F05	III.C.2	<i>heat shock protein GrpE</i>
07	08	12	07L08	XAC2633	27F04	A0UE9854H08	VI.C	<i>ISxac3 transposase</i>
07	09	12	07L09	XAC3431	28F05	A0UV9903F03	I.C.3	<i>short chain dehydrogenase</i>
07	10	12	07L10	XAC0110	27F05	A0UT9877B11	V.A.1	<i>proline/betaine transporter</i>
07	11	12	07L11	XAC3658	28F06	A0AC9895E08	VIII.B	<i>hypothetical protein</i>
07	12	12	07L12	XAC0831	27F06	A0UT9876A06	I.C.3	<i>oxidoreductase</i>
08	01	12	08L01	no_clone	32F01	no_clone	NC	<i>no_gene</i>
08	02	12	08L02	no_clone	31F01	no_clone	NC	<i>no_gene</i>

08	03	12	08L03	no_clone	32F02	no_clone	NC	no_gene
08	04	12	08L04	no_clone	31F02	no_clone	NC	no_gene
08	05	12	08L05	no_clone	32F03	no_clone	NC	no_gene
08	06	12	08L06	no_clone	31F03	no_clone	NC	no_gene
08	07	12	08L07	no_clone	32F04	no_clone	NC	no_gene
08	08	12	08L08	no_clone	31F04	no_clone	NC	no_gene
08	09	12	08L09	no_clone	32F05	no_clone	NC	no_gene
08	10	12	08L10	no_clone	31F05	no_clone	NC	no_gene
08	11	12	08L11	no_clone	32F06	no_clone	NC	no_gene
08	12	12	08L12	no_clone	31F06	no_clone	NC	no_gene
09	01	12	09A23	Score_card	33A12	Score_card	SC	Score_card
09	02	12	09G15	Score_card	33D08	Score_card	SC	Score_card
09	03	12	09M07	Score_card	33G04	Score_card	SC	Score_card
09	04	12	09A24	Score_card	34A12	Score_card	SC	Score_card
09	05	12	09G16	Score_card	34D08	Score_card	SC	Score_card
09	06	12	09M08	Score_card	34G04	Score_card	SC	Score_card
09	07	12	09B24	Score_card	35A12	Score_card	SC	Score_card
09	08	12	09H16	Score_card	35D08	Score_card	SC	Score_card
09	09	12	09N08	Score_card	35G04	Score_card	SC	Score_card
09	10	12	09B23	Score_card	36A12	Score_card	SC	Score_card
09	11	12	09H15	Score_card	36D08	Score_card	SC	Score_card
09	12	12	09N07	Score_card	36G04	Score_card	SC	Score_card
01	01	13	01M01	XAC0670	01G01	A0AC0113G03	VIII.A	conserved hypothetical protein
01	02	13	01M02	XAC2916	02G01	A0JE1075C02	II.B.2	aspartate carbamoyltransferase
01	03	13	01M03	XAC0915	01G02	A0JJ0104E10	VIII.A	conserved hypothetical protein
01	04	13	01M04	XAC3085	02G02	A0EC1095B04	VIII.C	Xanthomonas conserved hypothetical protein
01	05	13	01M05	XAC0919	01G03	A0JJ0102D01	I.B.10	pyridine nucleotide transhydrogenase
01	06	13	01M06	XAC1499	02G03	A0AM1050F07	I.D.2	transcriptional regulator
01	07	13	01M07	XAC0878	01G04	A0JJ0102D03	I.A.2	protocatechuate 3,4-dioxygenase beta chain
01	08	13	01M08	XAC3643	02G04	A0JJ1015G05	I.D.3	histidine kinase/response regulator hybrid protein
01	09	13	01M09	XAC3344	01G05	A0JJ0102D12	I.C.4	fructose-bisphosphate aldolase
01	10	13	01M10	XAC3952	02G05	A0JJ0705F03	VIII.A	conserved hypothetical protein
01	11	13	01M11	XAC1377	01G06	A0CE0105A02	I.C.3	dehydrogenase
01	12	13	01M12	XAC0779	02G06	A0JJ1101C04	IV.B	UDP-N-acetylglucosamine--N-acetylmuramyl- (pentapeptide) pyrophosphoryl-undecaprenol
02	01	13	02M01	XAC2873	05G01	A0QH1645H07	VIII.C	Xanthomonas conserved hypothetical protein
02	02	13	02M02	XAC3214	06G01	A0QR5407F12	VIII.A	conserved hypothetical protein
02	03	13	02M03	XAC1968	05G02	A0AM1594G07	IX	response regulator
02	04	13	02M04	XAC2933	06G02	A0UV5526E01	VIII.A	conserved hypothetical protein
02	05	13	02M05	XAC2783	05G03	A0QR5206H01	II.D.10	thioredoxin
02	06	13	02M06	XAC1859	06G03	A0UV5526G05	VIII.A	conserved hypothetical protein
02	07	13	02M07	XAC1525	05G04	A0QR5206B07	II.A.4	chorismate mutase/prephenate dehydrogenase
02	08	13	02M08	XAC0773	06G04	A0UV5525C11	V.B	cell division protein
02	09	13	02M09	XAC1936	05G05	A0UV5307E02	V.C	flagellar protein
02	10	13	02M10	XAC1191	06G05	A0UV5524D12	VIII.C	Xanthomonas conserved hypothetical protein
02	11	13	02M11	XAC4255	05G06	A0UV5307E10	V.A.3	hexuronate transporter
02	12	13	02M12	XAC4228	06G06	A0CE5420A10	IV.C	sialic acid-specific 9-O-acetyltransferase
03	01	13	03M01	XAC0725	09G01	A0EC6322D06	IX	GTP-binding protein
03	02	13	03M02	XAC1081	10G01	A0AC6360H09	III.A.2	histone-like protein
03	03	13	03M03	XAC4067	09G02	A0QR6330A02	V.A.2	permease
03	04	13	03M04	XAC1267	10G02	A0AC6361A07	VII.G	Hsp90xo protein
03	05	13	03M05	XAC1068	09G03	A0CE6340D07	VI.A	phage-related tail protein
03	06	13	03M06	XAC2022	10G03	A0UV6363H11	II.D.4	molybdopterin biosynthesis
03	07	13	03M07	XAC0480	09G04	A0QR6330B12	II.A.4	anthranilate synthase component II
03	08	13	03M08	XAC1005	10G04	A0QR6391G07	III.C.1	peptidyl-prolyl cis-trans isomerase

03	09	13	03M09	XAC1544	09G05	A0QR6334C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	10	13	03M10	XAC3294	10G05	A0QH6382A07	VIII.B	<i>hypothetical protein</i>
03	11	13	03M11	XAC2570	09G06	A0CE6341H04	IX	<i>GumP protein</i>
03	12	13	03M12	XAC3248	10G06	A0QR6389D09	VI.C	<i>ISxac3 transposase</i>
04	01	13	04M01	XAC3951	13G01	A0QR6601F08	VIII.B	<i>hypothetical protein</i>
04	02	13	04M02	XAC2641	14G01	A0UV6709G10	VI.A	<i>phage-related capsid packaging protein</i>
04	03	13	04M03	XAC2992	13G02	A0QR6705F10	III.C.3	<i>endoproteinase Arg-C</i>
04	04	13	04M04	XAC2513	14G02	A0UV6709C04	VII.H	<i>queuine tRNA-ribosyltransferase</i>
04	05	13	04M05	XAC0041	13G03	A0UV6707D05	IV.C	<i>mannosyltransferase B</i>
04	06	13	04M06	XAC1700	14G03	A0UV6711G04	III.D.1	<i>hexosyltransferase</i>
04	07	13	04M07	XAC0147	13G04	A0QR6705H01	VIII.A	<i>conserved hypothetical protein</i>
04	08	13	04M08	XAC1089	14G04	A0UV6710F11	III.B.6	<i>ribonuclease H</i>
04	09	13	04M09	XAC4050	13G05	A0QR6705E08	IX	<i>Pass1-related protein</i>
04	10	13	04M10	XAC0653	14G05	A0UV6710B09	V.A.7	<i>TonB-dependent receptor</i>
04	11	13	04M11	XACb0003	13G06	A0QR6705H11	VIII.B	<i>hypothetical protein</i>
04	12	13	04M12	XAC0187	14G06	A0UV6710C02	IV.B	<i>HipA protein</i>
05	01	13	05M01	XAC0822	17G01	A0UV6733D04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	02	13	05M02	XAC3354	18G01	A0CE6740G03	IV.A.2	<i>outer membrane protein W</i>
05	03	13	05M03	XAC2409	17G02	A0UV6735H09	V.A.7	<i>iron-sulfur cluster-binding protein</i>
05	04	13	05M04	XAC1918	18G02	A0CE6740G08	VII.C	<i>hemolysin related protein</i>
05	05	13	05M05	XAC1632	17G03	A0UV6736D02	VIII.B	<i>hypothetical protein</i>
05	06	13	05M06	XAC0918	18G03	A0CE6743B10	I.B.10	<i>pyridine nucleotide transhydrogenase</i>
05	07	13	05M07	XAC3348	17G04	A0UV6736D07	VIII.A	<i>conserved hypothetical protein</i>
05	08	13	05M08	XAC1657	18G04	A0CE6743D06	VIII.B	<i>hypothetical protein</i>
05	09	13	05M09	XAC3950	17G05	A0UV6736H08	VIII.A	<i>conserved hypothetical protein</i>
05	10	13	05M10	XAC1849	18G05	A0CE6743H09	III.C.1	<i>elongation factor P</i>
05	11	13	05M11	XAC1107	17G06	A0UV6739H05	VI.A	<i>integrase</i>
05	12	13	05M12	XAC1759	18G06	A0QR6749A01	I.D.2	<i>glycine cleavage system transcriptional repressor</i>
06	01	13	06M01	XAC3093	21G01	A0QR6805D02	VIII.B	<i>hypothetical protein</i>
06	02	13	06M02	XAC3873	22G01	A0UV6819A06	VIII.A	<i>conserved hypothetical protein</i>
06	03	13	06M03	XAC1822	21G02	A0UV6818B09	VIII.A	<i>conserved hypothetical protein</i>
06	04	13	06M04	XAC3017	22G02	A0UV6819B10	III.A.3	<i>RebB protein</i>
06	05	13	06M05	XAC0483	21G03	A0QH6813E03	VII.H	<i>CAP-like protein</i>
06	06	13	06M06	XAC0533	22G03	A0UV6820G04	II.A.4	<i>catabolic dehydroquinase</i>
06	07	13	06M07	XACb0037	21G04	A0UT6815B09	VII.H	<i>VirB11 protein</i>
06	08	13	06M08	XAC3900	22G04	A0UV6820D09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	13	06M09	XAC1964	21G05	A0UT6814G01	II.E	<i>3-oxoacyl-[ACP] synthase</i>
06	10	13	06M10	XAC1111	22G05	A0AC6824F03	III.A.3	<i>recombination protein RecR</i>
06	11	13	06M11	XAC3720	21G06	A0UT6815C04	VIII.A	<i>conserved hypothetical protein</i>
06	12	13	06M12	XAC1766	22G06	A0AC6825D09	I.A.2	<i>4-hydroxy-2-oxoglutarate aldolase/2-deydro-3-deoxyphosphogluconate aldolase</i>
07	01	13	07M01	XAC3075	25G01	A0CE9723E05	I.A.1	<i>beta-mannosidase</i>
07	02	13	07M02	XAC3898	26G01	A0UE9829D08	VIII.A	<i>conserved hypothetical protein</i>
07	03	13	07M03	XAC1765	25G02	A0QH9730E11	I.A.2	<i>4-hydroxy-2-oxoglutarate aldolase/2-deydro-3-deoxyphosphogluconate aldolase</i>
07	04	13	07M04	XAC2232	26G02	A0UE9829F01	I.D.2	<i>repressor binding protein</i>
07	05	13	07M05	XAC2863	25G03	A0QR9732F12	VIII.B	<i>hypothetical protein</i>
07	06	13	07M06	XAC4035	26G03	A0CE9837F11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	07	13	07M07	XAC2784	25G04	A0QH9730G11	VIII.A	<i>conserved hypothetical protein</i>
07	08	13	07M08	XAC3339	26G04	A0UV9805D06	I.D.2	<i>transcriptional regulator lysR family</i>
07	09	13	07M09	XAC4154	25G05	A0QR9756C01	I.D.2	<i>transcriptional regulator</i>
07	10	13	07M10	XAC3149	26G05	A0UV9804A05	III.A.3	<i>holliday junction binding protein, DNA helicase</i>
07	11	13	07M11	XAC1112	25G06	A0UV9751H04	I.A.2	<i>histidine triad protein homolog (HIT-like protein)</i>
07	12	13	07M12	XAC3402	26G06	A0UV9804D09	VIII.A	<i>conserved hypothetical protein</i>
08	01	13	08M01	XAC3136	29G01	genomic_DNA	I.D.1	<i>two-component system, sensor protein</i>

08	02	13	08M02	XAC0330	30G01	genomic_DNA	VII.B	<i>HrpE protein</i>
08	03	13	08M03	XAC2231	29G02	genomic_DNA	IX	<i>B-cell mitogen related protein</i>
08	04	13	08M04	XAC1088	30G02	genomic_DNA	VII.B	<i>HrpD6 protein</i>
08	05	13	08M05	XAC2614	29G03	genomic_DNA	VII.H	<i>VirB4 protein</i>
08	06	13	08M06	XAC0279	30G03	genomic_DNA	VII.B	<i>HrpB4 protein</i>
08	07	13	08M07	XACb0049	29G04	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	08	13	08M08	XAC2493	30G04	genomic_DNA	I.D.1	<i>two-component system, regulatory protein</i>
08	09	13	08M09	XACb0041	29G05	genomic_DNA	VII.H	<i>VirB6 protein</i>
08	10	13	08M10	XAC1882	30G05	genomic_DNA	VII.H	<i>aconitase</i>
08	11	13	08M11	XACb0027	29G06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	13	08M12	no_clone	30G06	no_clone	NC	<i>no_gene</i>
09	01	13	09C01	Score_card	33B01	Score_card	SC	<i>Score_card</i>
09	02	13	09G17	Score_card	33D09	Score_card	SC	<i>Score_card</i>
09	03	13	09M09	Score_card	33G05	Score_card	SC	<i>Score_card</i>
09	04	13	09C02	Score_card	34B01	Score_card	SC	<i>Score_card</i>
09	05	13	09G18	Score_card	34D09	Score_card	SC	<i>Score_card</i>
09	06	13	09M10	Score_card	34G05	Score_card	SC	<i>Score_card</i>
09	07	13	09D02	Score_card	35B01	Score_card	SC	<i>Score_card</i>
09	08	13	09H18	Score_card	35D09	Score_card	SC	<i>Score_card</i>
09	09	13	09N10	Score_card	35G05	Score_card	SC	<i>Score_card</i>
09	10	13	09D01	Score_card	36B01	Score_card	SC	<i>Score_card</i>
09	11	13	09H17	Score_card	36D09	Score_card	SC	<i>Score_card</i>
09	12	13	09N09	Score_card	36G05	Score_card	SC	<i>Score_card</i>
01	01	14	01N01	XAC4106	04G01	A0RN1457G01	III.C.3	<i>dipeptidyl peptidase</i>
01	02	14	01N02	XAC2722	03G01	A0JJ1110G03	IV.D	<i>FimV protein</i>
01	03	14	01N03	XAC3498	04G02	A0JJ1577H11	V.A.4	<i>outer membrane receptor for ferric iron uptake</i>
01	04	14	01N04	XAC0360	03G02	A0JE1211B06	I.A.2	<i>glycerol-3-phosphate dehydrogenase</i>
01	05	14	01N05	XAC0860	04G03	A0JJ1536D10	V.A.1	<i>ABC transporter ATP-binding protein</i>
01	06	14	01N06	XAC1736	03G03	A0JJ1304C12	III.C.3	<i>GTP-binding protein</i>
01	07	14	01N07	XAC1080	04G04	A0UV1533G01	III.C.3	<i>ATP-dependent serine proteinase La</i>
01	08	14	01N08	XAC4265	03G04	A0AM1405E08	VIII.B	<i>hypothetical protein</i>
01	09	14	01N09	XAC1276	04G05	A0JJ1536G05	V.A.7	<i>TonB-dependent receptor</i>
01	10	14	01N10	XAC0923	03G05	A0JJ1382C08	I.B.10	<i>pyridine nucleotide transhydrogenase, subunit alpha</i>
01	11	14	01N11	XAC0594	04G06	A0JJ1541F09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	12	14	01N12	XAC1279	03G06	A0AM1363E05	I.D.1	<i>two-component system, regulatory protein</i>
02	01	14	02N01	XAC2185	08G01	A0QR6202D01	V.A.7	<i>ferrichrome-iron receptor</i>
02	02	14	02N02	XAC2792	07G01	A0QR5702G11	VIII.A	<i>conserved hypothetical protein</i>
02	03	14	02N03	XAC3517	08G02	A0QR5904H11	VIII.A	<i>conserved hypothetical protein</i>
02	04	14	02N04	XAC2746	07G02	A0QR5702E03	III.C.3	<i>metallopeptidase</i>
02	05	14	02N05	XAC4161	08G03	A0QR6203D08	VII.C	<i>cation efflux system protein</i>
02	06	14	02N06	XAC0077	07G03	A0JE5637E10	VIII.A	<i>conserved hypothetical protein</i>
02	07	14	02N07	XAC0573	08G04	A0QR6002C11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	08	14	02N08	XAC0957	07G04	A0QR5902F09	III.C.1	<i>elongation factor Tu</i>
02	09	14	02N09	XAC0638	08G05	A0QR6002C12	III.C.3	<i>ATP-dependent HslUV protease ATP-binding subunit HslU</i>
02	10	14	02N10	XAC0388	07G05	A0QR5902B06	II.D.1	<i>biotin synthase</i>
02	11	14	02N11	XAC1942	08G06	A0QR6304H09	V.C	<i>flagellar biosynthesis</i>
02	12	14	02N12	XAC2826	07G06	A0QR5902E10	I.C.2	<i>alcohol dehydrogenase</i>
03	01	14	03N01	XAC2070	12G01	A0QR6490H11	I.C.4	<i>glucose kinase</i>
03	02	14	03N02	XACa0034	11G01	A0UV6433D05	VI.C	<i>Tn5045 transposase</i>
03	03	14	03N03	XAC3144	12G02	A0QR6484B08	VII.C	<i>TolR protein</i>
03	04	14	03N04	XAC3291	11G02	A0AC6441G04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	05	14	03N05	XAC3319	12G03	A0UV6487E06	VIII.B	<i>hypothetical protein</i>
03	06	14	03N06	XAC2893	11G03	A0AC6439C07	I.C.3	<i>oxidoreductase</i>
03	07	14	03N07	XAC1220	12G04	A0UV6485H09	VIII.A	<i>conserved hypothetical protein</i>

03	08	14	03N08	XAC3340	11G04	A0QR6413A02	II.D.12	<i>siroheme synthase</i>
03	09	14	03N09	XAC3587	12G05	A0AC6494F09	I.C.3	<i>electron transfer flavoprotein alpha subunit</i>
03	10	14	03N10	XAC0297	11G05	A0UV6455E01	VIII.A	<i>conserved hypothetical protein</i>
03	11	14	03N11	XAC3973	12G06	A0CE6506D09	V.B	<i>cell division inhibitor</i>
03	12	14	03N12	XAC0778	11G06	A0UE6464G09	V.B	<i>cell division protein</i>
04	01	14	04N01	XAC1104	16G01	A0QH6731D06	VI.B	<i>plasmid mobilization protein</i>
04	02	14	04N02	XAC1096	15G01	A0UV6712D12	VIII.A	<i>conserved hypothetical protein</i>
04	03	14	04N03	XAC1826	16G02	A0QR6730G09	III.B.4	<i>histidyl-tRNA synthetase</i>
04	04	14	04N04	XAC0730	15G02	A0UV6712E01	I.D.1	<i>two-component system, regulatory protein</i>
04	05	14	04N05	XAC0930	16G03	A0QR6730F02	III.C.3	<i>extracellular protease</i>
04	06	14	04N06	XAC0558	15G03	A0QR6727H11	V.A.7	<i>iron utilization protein</i>
04	07	14	04N07	XAC3912	16G04	A0QR6730D09	IV.A.1	<i>phosphomannomutase</i>
04	08	14	04N08	XAC0959	15G04	A0QR6727H12	V.A.6	<i>preprotein translocase subunit</i>
04	09	14	04N09	XAC0136	16G05	A0QR6730E03	I.D.1	<i>two-component system, regulatory protein</i>
04	10	14	04N10	XAC2087	15G05	A0QR6729D02	V.A.7	<i>ABC transporter ATP-binding protein</i>
04	11	14	04N11	XAC0477	16G06	A0QR6730C08	II.A.3	<i>threonine aldolase</i>
04	12	14	04N12	XAC0189	15G06	A0QR6729E09	I.C.3	<i>indolepyruvate ferredoxin oxidoreductase chain alpha</i>
05	01	14	05N01	XAC3707	20G01	A0QR6770B09	VIII.A	<i>conserved hypothetical protein</i>
05	02	14	05N02	XAC3060	19G01	A0UV6761B08	I.B.10	<i>glycine cleavage H protein</i>
05	03	14	05N03	XACb0057	20G02	A0QR6801F04	VIII.A	<i>conserved hypothetical protein</i>
05	04	14	05N04	XAC2085	19G02	A0UE6759A01	V.A.7	<i>biopolymer transport protein</i>
05	05	14	05N05	XAC3735	20G03	A0QR6770C03	I.C.3	<i>cyanide insensitive terminal oxidase</i>
05	06	14	05N06	XAC0982	19G03	A0UE6759A02	III.B.2	<i>50S ribosomal protein L14</i>
05	07	14	05N07	XAC0707	20G04	A0UV6766A07	I.A.1	<i>beta-galactosidase (truncated)</i>
05	08	14	05N08	XAC4195	19G04	A0UE6759E06	IV.C	<i>NdvB protein</i>
05	09	14	05N09	XAC0975	20G05	A0QR6802H01	III.B.2	<i>50S ribosomal protein L2</i>
05	10	14	05N10	XAC1205	19G05	A0UV6764E06	VIII.A	<i>conserved hypothetical protein</i>
05	11	14	05N11	XACb0032	20G06	A0QR6803F06	VIII.A	<i>conserved hypothetical protein</i>
05	12	14	05N12	XAC0979	19G06	A0UV6764E11	III.B.2	<i>50S ribosomal protein L16</i>
06	01	14	06N01	XAC0146	24G01	A0QR6840H02	VIII.B	<i>hypothetical protein</i>
06	02	14	06N02	XAC3854	23G01	A0AC6827A05	I.A.2	<i>hydrolase, haloacid dehalogenase-like family</i>
06	03	14	06N03	XAC1433	24G02	A0UV6843F01	II.A.2	<i>asparagine synthase B</i>
06	04	14	06N04	XAC0135	23G02	A0AC6825A10	I.D.1	<i>two-component system, sensor protein</i>
06	05	14	06N05	XAC0023	24G03	A0UV6843F05	III.C.3	<i>carboxyl-terminal protease</i>
06	06	14	06N06	XAC0292	23G03	A0CE6831A04	VIII.A	<i>conserved hypothetical protein</i>
06	07	14	06N07	XAC3705	24G04	A0UV9710G08	VIII.A	<i>conserved hypothetical protein</i>
06	08	14	06N08	XAC3191	23G04	A0CE6831A05	II.D.14	<i>cob(I)alamin adenosyltransferase</i>
06	09	14	06N09	XAC2728	24G05	A0UV6845E12	II.E	<i>phosphatidylserine decarboxylase</i>
06	10	14	06N10	XAC2512	23G05	A0QH6836G12	V.A.6	<i>preprotein translocase YajC subunit</i>
06	11	14	06N11	XAC3129	24G06	A0UV6845G11	III.B.5	<i>pseudouridylate synthase</i>
06	12	14	06N12	XAC0344	23G06	A0QH6835D01	VI.A	<i>phage-related integrase</i>
07	01	14	07N01	XAC3147	28G01	A0UV9902A03	III.A.3	<i>holliday junction binding protein, DNA helicase</i>
07	02	14	07N02	XAC2027	27G01	A0UT9843F09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	14	07N03	XAC1342	28G02	A0UT9894B02	III.B.5	<i>mRNA 3'-end processing factor</i>
07	04	14	07N04	XAC1353	27G02	A0CE9840H04	VIII.A	<i>conserved hypothetical protein</i>
07	05	14	07N05	XAC2206	28G03	A0UV9900C08	VIII.A	<i>conserved hypothetical protein</i>
07	06	14	07N06	XAC1742	27G03	A0UV9871D12	III.B.4	<i>alanyl-tRNA synthetase</i>
07	07	14	07N07	XAC0048	28G04	A0UV9901E10	VIII.B	<i>hypothetical protein</i>
07	08	14	07N08	XAC3794	27G04	A0UE9858C06	VIII.A	<i>conserved hypothetical protein</i>
07	09	14	07N09	XAC0516	28G05	A0UV9904E04	VIII.A	<i>conserved hypothetical protein</i>
07	10	14	07N10	XAC1384	27G05	A0UV9879B07	I.C.3	<i>ferredoxin II</i>
07	11	14	07N11	XAC0829	28G06	A0UV9905F12	V.A.7	<i>ABC transporter substrate binding protein</i>
07	12	14	07N12	XAC1556	27G06	A0UT9877E01	V.A.3	<i>glucose/galactose transporter</i>
08	01	14	08N01	no_clone	32G01	no_clone	NC	<i>no_gene</i>
08	02	14	08N02	no_clone	31G01	no_clone	NC	<i>no_gene</i>

08	03	14	08N03	no_clone	32G02	no_clone	NC	<i>no_gene</i>
08	04	14	08N04	no_clone	31G02	no_clone	NC	<i>no_gene</i>
08	05	14	08N05	no_clone	32G03	no_clone	NC	<i>no_gene</i>
08	06	14	08N06	no_clone	31G03	no_clone	NC	<i>no_gene</i>
08	07	14	08N07	no_clone	32G04	no_clone	NC	<i>no_gene</i>
08	08	14	08N08	no_clone	31G04	no_clone	NC	<i>no_gene</i>
08	09	14	08N09	no_clone	32G05	no_clone	NC	<i>no_gene</i>
08	10	14	08N10	no_clone	31G05	no_clone	NC	<i>no_gene</i>
08	11	14	08N11	no_clone	32G06	no_clone	NC	<i>no_gene</i>
08	12	14	08N12	no_clone	31G06	no_clone	NC	<i>no_gene</i>
09	01	14	09C03	Score_card	33B02	Score_card	SC	<i>Score_card</i>
09	02	14	09G19	Score_card	33D10	Score_card	SC	<i>Score_card</i>
09	03	14	09M11	Score_card	33G06	Score_card	SC	<i>Score_card</i>
09	04	14	09C04	Score_card	34B02	Score_card	SC	<i>Score_card</i>
09	05	14	09G20	Score_card	34D10	Score_card	SC	<i>Score_card</i>
09	06	14	09M12	Score_card	34G06	Score_card	SC	<i>Score_card</i>
09	07	14	09D04	Score_card	35B02	Score_card	SC	<i>Score_card</i>
09	08	14	09H20	Score_card	35D10	Score_card	SC	<i>Score_card</i>
09	09	14	09N12	Score_card	35G06	Score_card	SC	<i>Score_card</i>
09	10	14	09D03	Score_card	36B02	Score_card	SC	<i>Score_card</i>
09	11	14	09H19	Score_card	36D10	Score_card	SC	<i>Score_card</i>
09	12	14	09N11	Score_card	36G06	Score_card	SC	<i>Score_card</i>
01	01	15	01O01	XAC2555	01H01	A0AC0115A08	I.D.3	<i>sensor histidine kinase</i>
01	02	15	01O02	XAC3141	02H01	A0EC1090B10	IV.B	<i>outer membrane protein P6 precursor</i>
01	03	15	01O03	XAC2275	01H02	A0CE0105A12	VIII.A	<i>conserved hypothetical protein</i>
01	04	15	01O04	XAC2908	02H02	A0JJ0133H04	IV.B	<i>UDP-N-acetylmuramoylalanine--D-glutamate ligase</i>
01	05	15	01O05	XAC2793	01H03	A0JJ0104E12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	06	15	01O06	XAC3020	02H03	A0JJ1070B04	VIII.B	<i>hypothetical protein</i>
01	07	15	01O07	XACb0035	01H04	A0JJ0104F07	VIII.B	<i>hypothetical protein</i>
01	08	15	01O08	XAC2227	02H04	A0JJ1035G03	VIII.A	<i>conserved hypothetical protein</i>
01	09	15	01O09	XAC1180	01H05	A0CE0105B11	I.C.3	<i>oxidoreductase</i>
01	10	15	01O10	XAC1586	02H05	A0AR0913E08	III.A.4	<i>MutT/nudix family protein</i>
01	11	15	01O11	XAC2380	01H06	A0CE0105C02	III.C.1	<i>elongation factor P</i>
01	12	15	01O12	XAC3113	02H06	A0JJ1110H07	I.D.4	<i>ATP:GTP 3'-pyrophosphotranferase</i>
02	01	15	02O01	XAC0529	05H01	A0QR5106D10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	02	15	02O02	XAC1554	06H01	A0QR5407H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	03	15	02O03	XAC0669	05H02	A0UV1613G04	III.C.3	<i>tail-specific protease</i>
02	04	15	02O04	XAC0380	06H02	A0QR5408H07	VIII.A	<i>conserved hypothetical protein</i>
02	05	15	02O05	XAC3378	05H03	A0UV5307C05	VIII.A	<i>conserved hypothetical protein</i>
02	06	15	02O06	XAC0447	06H03	A0QR5529G06	VII.F	<i>nuclease</i>
02	07	15	02O07	XAC0981	05H04	A0QR5302A08	III.B.2	<i>30S ribosomal protein S17</i>
02	08	15	02O08	XAC3253	06H04	A0UV5526H02	III.B.3	<i>ribosomal protein S6 modification protein</i>
02	09	15	02O09	XAC1439	05H05	A0QR5205H12	VII.C	<i>thiopurine methyltransferase</i>
02	10	15	02O10	XAC1190	06H05	A0UV5525G03	VIII.A	<i>conserved hypothetical protein</i>
02	11	15	02O11	XAC4073	05H06	A0CE5309D04	VIII.A	<i>conserved hypothetical protein</i>
02	12	15	02O12	XAC0222	06H06	A0UV5502C01	III.D.2	<i>glycerol-3-phosphate dehydrogenase</i>
03	01	15	03O01	XAC0741	09H01	A0AC6336B05	V.A.7	<i>ABC transporter ATP-binding protein</i>
03	02	15	03O02	XAC1369	10H01	A0UV6363F10	VIII.A	<i>conserved hypothetical protein</i>
03	03	15	03O03	XAC0519	09H02	A0AC6337H03	II.E	<i>CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase-related protein</i>
03	04	15	03O04	XAC0291	10H02	A0QR6367D10	IV.A.2	<i>Oar protein</i>
03	05	15	03O05	XAC3931	09H03	A0CE6341D10	VII.F	<i>competence related protein</i>
03	06	15	03O06	XAC1102	10H03	A0QR6366C08	VI.C	<i>ISxac3 transposase</i>
03	07	15	03O07	XAC3255	09H04	A0CE6340D09	VIII.B	<i>hypothetical protein</i>
03	08	15	03O08	XAC2475	10H04	A0UE6399B02	VII.C	<i>transport protein</i>

03	09	15	03O09	XAC2742	09H05	A0AC6337B06	V.A.7	<i>TonB-dependent receptor</i>
03	10	15	03O10	XAC3536	10H05	A0QR6388H06	VII.H	<i>general secretion pathway protein M</i>
03	11	15	03O11	XAC4186	09H06	A0CE6345D12	I.C.3	<i>oxidoreductase</i>
03	12	15	03O12	XAC2227	10H06	A0QR6391E10	VIII.A	<i>conserved hypothetical protein</i>
04	01	15	04O01	XAC1435	13H01	A0UE6610E07	V.A.7	<i>iron receptor</i>
04	02	15	04O02	XAC0970	14H01	A0UV6710B01	III.C.1	<i>elongation factor Tu</i>
04	03	15	04O03	XAC3088	13H02	A0UV6707A05	VIII.B	<i>hypothetical protein</i>
04	04	15	04O04	XAC2881	14H02	A0UV6709F07	VII.G	<i>carbon starvation protein A</i>
04	05	15	04O05	XAC1203	13H03	A0UV6707G02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	06	15	04O06	XAC3739	14H03	A0UV6711H08	VIII.A	<i>conserved hypothetical protein</i>
04	07	15	04O07	XAC0481	13H04	A0UV6707B07	II.A.4	<i>indole-3-glycerol phosphate synthase</i>
04	08	15	04O08	XAC1266	14H04	A0UV6711A02	VII.B	<i>HrpX protein</i>
04	09	15	04O09	XAC0210	13H05	A0QR6705H04	VII.C	<i>superoxide dismutase</i>
04	10	15	04O10	XAC2879	14H05	A0UV6710D10	VIII.A	<i>conserved hypothetical protein</i>
04	11	15	04O11	XAC2680	13H06	A0UV6707B10	VIII.A	<i>conserved hypothetical protein</i>
04	12	15	04O12	XAC0885	14H06	A0UV6710D12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	01	15	05O01	XAC2552	17H01	A0UV6733F01	V.B	<i>cell division protein</i>
05	02	15	05O02	XAC3581	18H01	A0CE6742A11	I.B.11	<i>UDP-glucose dehydrogenase</i>
05	03	15	05O03	XAC2125	17H02	A0UV6736B03	IV.C	<i>glycosyl transferase related protein</i>
05	04	15	05O04	XAC0178	18H02	A0CE6742A12	VIII.A	<i>conserved hypothetical protein</i>
05	05	15	05O05	XAC0459	17H03	A0UV6736E08	VII.G	<i>PhaE protein</i>
05	06	15	05O06	XAC2900	18H03	A0CE6743D05	III.A.5	<i>type I restriction-modification system DNA methylase</i>
05	07	15	05O07	XAC0999	17H04	A0UV6736E09	IV.A.2	<i>colicin I receptor</i>
05	08	15	05O08	XAC3459	18H04	A0CE6743G09	I.D.2	<i>transcriptional regulator lysR family</i>
05	09	15	05O09	XAC0500	17H05	A0QR6737B06	VIII.B	<i>hypothetical protein</i>
05	10	15	05O10	XAC2607	18H05	A0CE6746F08	VII.H	<i>VirB6 protein</i>
05	11	15	05O11	XAC1607	17H06	A0UV6736D09	VIII.A	<i>conserved hypothetical protein</i>
05	12	15	05O12	XAC3641	18H06	A0CE6746C05	V.A.7	<i>ABC transporter permease</i>
06	01	15	06O01	XAC3261	21H01	A0QH6810D04	VIII.B	<i>hypothetical protein</i>
06	02	15	06O02	XAC3502	22H01	A0UV6819D06	VIII.A	<i>conserved hypothetical protein</i>
06	03	15	06O03	XAC4371	21H02	A0QH6813E01	IX	<i>polysaccharide deacetylase</i>
06	04	15	06O04	XAC4006	22H02	A0UV6819D11	III.B.4	<i>tryptophanyl-tRNA synthetase</i>
06	05	15	06O05	XAC0753	21H03	A0UT6814F07	VIII.A	<i>conserved hypothetical protein</i>
06	06	15	06O06	XAC2950	22H03	A0UV6821A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	07	15	06O07	XAC1316	21H04	A0UT6815F10	I.A.2	<i>3-hydroxyisobutirate dehydrogenase</i>
06	08	15	06O08	XAC3264	22H04	A0UV6820G11	VIII.B	<i>hypothetical protein</i>
06	09	15	06O09	XAC0170	21H05	A0UT6815F11	IX	<i>sugar-phosphate isomerase</i>
06	10	15	06O10	XAC3032	22H05	A0AC6825A11	VIII.A	<i>conserved hypothetical protein</i>
06	11	15	06O11	XAC4241	21H06	A0UT6815G11	II.D.12	<i>glutamate-1-semialdehyde aminotransferase (aminomutase)</i>
06	12	15	06O12	XAC0713	22H06	A0AC6825F07	I.D.2	<i>transcriptional regulator lacI family</i>
07	01	15	07O01	XAC3933	25H01	A0CE9726D08	VIII.A	<i>conserved hypothetical protein</i>
07	02	15	07O02	XAC1209	26H01	A0AC9835A01	VIII.B	<i>hypothetical protein</i>
07	03	15	07O03	XAC1052	25H02	A0QR9732E07	VI.C	<i>ISxac3 transposase</i>
07	04	15	07O04	XAC0834	26H02	A0UV9803E05	I.D.1	<i>two-component system, regulatory protein</i>
07	05	15	07O05	XAC2491	25H03	A0QH9738D01	VIII.A	<i>conserved hypothetical protein</i>
07	06	15	07O06	XAC0561	26H03	A0UV9805C11	II.E	<i>delta subunit of malonate decarboxylase</i>
07	07	15	07O07	XAC1296	25H04	A0QR9732G07	VIII.A	<i>conserved hypothetical protein</i>
07	08	15	07O08	XAC2131	26H04	A0QR9814C03	VI.C	<i>ISxac3 transposase</i>
07	09	15	07O09	XAC0905	25H05	A0QR9757D08	I.D.2	<i>oxidative stress transcriptional regulator</i>
07	10	15	07O10	XAC3802	26H05	A0UV9805F05	VIII.A	<i>conserved hypothetical protein</i>
07	11	15	07O11	XAC0562	25H06	A0QR9756D01	II.D.1	<i>beta subunit of malonate decarboxylase</i>
07	12	15	07O12	XAC4202	26H06	A0UV9805G09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	01	15	08O01	XAC2126	29H01	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	02	15	08O02	XAC0028	30H01	genomic_DNA	VII.H	<i>HpaF protein</i>

08	03	15	08O03	XAC1088	29H02	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	04	15	08O04	XAC0028	30H02	genomic_DNA	VII.B	<i>HrpD5 protein</i>
08	05	15	08O05	XAC2623	29H03	genomic_DNA	VII.H	<i>VirD4 protein</i>
08	06	15	08O06	XAC0330	30H03	genomic_DNA	VII.B	<i>HrpB5 protein</i>
08	07	15	08O07	XAC2616	29H04	genomic_DNA	VII.H	<i>VirB2 protein</i>
08	08	15	08O08	XAC2504	30H04	genomic_DNA	VII.H	<i>regulator of pathogenicity factors</i>
08	09	15	08O09	XAC2625	29H05	genomic_DNA	III.A.4	<i>excinuclease ABC subunit B</i>
08	10	15	08O10	no_clone	30H05	no_clone	NC	<i>no_gene</i>
08	11	15	08O11	XACb0028	29H06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	15	08O12	no_clone	30H06	no_clone	NC	<i>no_gene</i>
09	01	15	09C05	Score_card	33B03	Score_card	SC	<i>Score_card</i>
09	02	15	09G21	Score_card	33D11	Score_card	SC	<i>Score_card</i>
09	03	15	09M13	Score_card	33G07	Score_card	SC	<i>Score_card</i>
09	04	15	09C06	Score_card	34B03	Score_card	SC	<i>Score_card</i>
09	05	15	09G22	Score_card	34D11	Score_card	SC	<i>Score_card</i>
09	06	15	09M14	Score_card	34G07	Score_card	SC	<i>Score_card</i>
09	07	15	09D06	Score_card	35B03	Score_card	SC	<i>Score_card</i>
09	08	15	09H22	Score_card	35D11	Score_card	SC	<i>Score_card</i>
09	09	15	09N14	Score_card	35G07	Score_card	SC	<i>Score_card</i>
09	10	15	09D05	Score_card	36B03	Score_card	SC	<i>Score_card</i>
09	11	15	09H21	Score_card	36D11	Score_card	SC	<i>Score_card</i>
09	12	15	09N13	Score_card	36G07	Score_card	SC	<i>Score_card</i>
01	01	16	01P01	XAC0747	04H01	A0JE1466A02	VIII.A	<i>conserved hypothetical protein</i>
01	02	16	01P02	XAC0442	03H01	A0AM1113C08	III.B.5	<i>ATP-dependent RNA helicase</i>
01	03	16	01P03	XAC1882	04H02	A0AM1427B07	VII.H	<i>aconitase</i>
01	04	16	01P04	XAC0358	03H02	A0RN1228F09	I.B.10	<i>glycerol kinase</i>
01	05	16	01P05	XAC2077	04H03	A0AM1540G12	I.C.7	<i>succinate dehydrogenase, flavoprotein subunit</i>
01	06	16	01P06	XAC1685	03H03	A0JJ1312D09	I.C.3	<i>cytochrome C</i>
01	07	16	01P07	XAC3124	04H04	A0JJ1536F06	III.A.3	<i>DNA helicase</i>
01	08	16	01P08	XAC4108	03H04	A0JJ1304G04	VIII.A	<i>conserved hypothetical protein</i>
01	09	16	01P09	XAC1166	04H05	A0RN1548D04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	10	16	01P10	XAC1448	03H05	A0AM1402G03	I.A.1	<i>beta-glucosidase</i>
01	11	16	01P11	XAC4328	04H06	A0RN1548D09	VI.C	<i>ISxac1 transposase</i>
01	12	16	01P12	XAC2926	03H06	A0QH1368D04	II.A.1	<i>pyrroline-5-carboxylate reductase</i>
02	01	16	02P01	XAC3851	08H01	A0QR6204C10	VIII.A	<i>conserved hypothetical protein</i>
02	02	16	02P02	XAC2011	07H01	A0QR5704A11	VIII.A	<i>conserved hypothetical protein</i>
02	03	16	02P03	XAC1469	08H02	A0QR6001D02	VIII.A	<i>conserved hypothetical protein</i>
02	04	16	02P04	XAC0196	07H02	A0QR5703C03	VIII.A	<i>conserved hypothetical protein</i>
02	05	16	02P05	XAC2506	08H03	A0QR6204E12	VIII.B	<i>hypothetical protein</i>
02	06	16	02P06	XAC3843	07H03	A0QR5701B02	VIII.A	<i>conserved hypothetical protein</i>
02	07	16	02P07	XAC2993	08H04	A0QR6002E06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	08	16	02P08	XAC1074	07H04	A0QR5902H06	I.D.3	<i>sensor histidine kinase</i>
02	09	16	02P09	XAC3058	08H05	A0QR6203F09	III.A.2	<i>histone H1</i>
02	10	16	02P10	XAC2611	07H05	A0QR5902E05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	11	16	02P11	XAC1519	08H06	A0QH6307G12	III.A.3	<i>recombination protein N</i>
02	12	16	02P12	XAC2786	07H06	A0QR5902G02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	01	16	03P01	XAC1113	12H01	A0UV6511H05	IV.A.2	<i>outer membrane protein Slp</i>
03	02	16	03P02	XAC0768	11H01	A0QR6443E11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	03	16	03P03	XAC2657	12H02	A0UV6485G11	VIII.A	<i>conserved hypothetical protein</i>
03	04	16	03P04	XAC3970	11H02	A0QR6443G07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	05	16	03P05	XAC0858	12H03	A0QR6489D01	V.A.7	<i>ABC transporter permease</i>
03	06	16	03P06	XAC2393	11H03	A0QR6443G08	III.D.2	<i>carboxylesterase</i>
03	07	16	03P07	XAC2596	12H04	A0UV6487F02	III.D.1	<i>cyclomaltodextrin glucanotransferase (CGTase)</i>
03	08	16	03P08	XAC1966	11H04	A0QR6417F05	IV.C	<i>nucleotide sugar transaminase</i>
03	09	16	03P09	XAC3501	12H05	A0CE6507E11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>

03	10	16	03P10	XAC1719	11H05	A0UV6457A11	I.C.4	<i>enolase</i>
03	11	16	03P11	XAC0262	12H06	A0CE6507H12	III.C.3	<i>dipeptidyl aminopeptidase</i>
03	12	16	03P12	XAC3337	11H06	A0EC6474G10	VIII.B	<i>hypothetical protein</i>
04	01	16	04P01	XAC2823	16H01	A0QH6731H04	III.A.4	<i>6-O-methylguanine-DNA methyltransferase</i>
04	02	16	04P02	XAC2453	15H01	A0UV6712F08	I.D.2	<i>stringent starvation protein B</i>
04	03	16	04P03	XAC3362	16H02	A0QH6731E05	VII.C	<i>TonB-like protein</i>
04	04	16	04P04	XAC1192	15H02	A0UV6712F09	VII.C	<i>tetracenomycin polyketide synthesis protein</i>
04	05	16	04P05	XAC2737	16H03	A0QR6730H02	VIII.A	<i>conserved hypothetical protein</i>
04	06	16	04P06	XAC0341	15H03	A0QR6728C07	VIII.A	<i>conserved hypothetical protein</i>
04	07	16	04P07	XAC0849	16H04	A0QR6730F05	V.A.7	<i>sulfonate binding protein</i>
04	08	16	04P08	XAC4217	15H04	A0QR6728C09	IV.A.1	<i>SEC-independent protein translocase</i>
04	09	16	04P09	XAC1307	16H05	A0QR6730F07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	10	16	04P10	XAC4210	15H05	A0QR6729H03	III.B.4	<i>glycyl-tRNA synthetase beta chain</i>
04	11	16	04P11	XAC0438	16H06	A0QR6730H07	V.A.7	<i>component of multidrug efflux system</i>
04	12	16	04P12	XAC1631	15H06	A0QR6729G01	III.A.1	<i>DNA gyrase subunit A</i>
05	01	16	05P01	XAC1959	20H01	A0QR6801D12	VIII.A	<i>conserved hypothetical protein</i>
05	02	16	05P02	XAC3368	19H01	A0UV6761D07	VIII.A	<i>conserved hypothetical protein</i>
05	03	16	05P03	XAC3293	20H02	A0QR6802C03	VIII.B	<i>hypothetical protein</i>
05	04	16	05P04	XAC3806	19H02	A0UE6759C01	VIII.A	<i>conserved hypothetical protein</i>
05	05	16	05P05	XAC2683	20H03	A0QR6772H12	III.B.6	<i>polynucleotide phosphorylase</i>
05	06	16	05P06	XAC3309	19H03	A0UE6759C02	III.C.3	<i>aminopeptidase</i>
05	07	16	05P07	XAC3669	20H04	A0UV6766C05	V.A.7	<i>ABC transporter ATP-binding protein</i>
05	08	16	05P08	XAC0171	19H04	A0UE6759H05	I.A.1	<i>rhamnogalacturonan acetyltransferase</i>
05	09	16	05P09	XACa0042	20H05	A0UV6766A11	VI.B	<i>KfrA protein</i>
05	10	16	05P10	XAC2502	19H05	A0UV6762F01	I.C.4	<i>1-phosphofructokinase (fructose 1-phosphate kinase)</i>
05	11	16	05P11	XAC2429	20H06	A0QR6804A06	VIII.A	<i>conserved hypothetical protein</i>
05	12	16	05P12	XAC2197	19H06	A0UV6765E04	VII.C	<i>hemolysin- type calcium binding protein</i>
06	01	16	06P01	XAC2187	24H01	A0UV6843A10	VIII.A	<i>conserved hypothetical protein</i>
06	02	16	06P02	XAC2432	23H01	A0AC6827D07	VI.C	<i>transposase</i>
06	03	16	06P03	XAC0554	24H02	A0UV6843H07	IX	<i>nitroreductase</i>
06	04	16	06P04	XAC2148	23H02	A0AC6825C12	V.A.7	<i>outer membrane efflux protein</i>
06	05	16	06P05	XAC2927	24H03	A0UV6843H08	III.A.2	<i>histone-like protein</i>
06	06	16	06P06	XAC1179	23H03	A0QH6835F09	VIII.A	<i>conserved hypothetical protein</i>
06	07	16	06P07	XAC0315	24H04	A0UV9712G03	VIII.B	<i>hypothetical protein</i>
06	08	16	06P08	XAC2497	23H04	A0QH6836G10	I.D.2	<i>transcriptional regulator tetR/acrR family</i>
06	09	16	06P09	XAC0133	24H05	A0UV6846B06	I.C.2	<i>L-lactate dehydrogenase</i>
06	10	16	06P10	XAC4244	23H05	A0CE6829F03	III.D.1	<i>xylulose kinase</i>
06	11	16	06P11	XAC2299	24H06	A0UV6846E06	II.B.2	<i>cytidylate kinase</i>
06	12	16	06P12	XAC2647	23H06	A0QH6836C04	VI.A	<i>phage-related tail protein</i>
07	01	16	07P01	XAC1891	28H01	A0UV9905A05	V.C	<i>chemotaxis protein</i>
07	02	16	07P02	XAC0275	27H01	A0UT9846C06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	16	07P03	XAC1553	28H02	A0AC9896B08	VIII.A	<i>conserved hypothetical protein</i>
07	04	16	07P04	XAC2406	27H02	A0CE9842C02	IV.B	<i>N-acetylmuramoyl-L-alanine amidase</i>
07	05	16	07P05	XAC0948	28H03	A0UV9903D07	I.B.10	<i>4-diphosphocytidyl-2-C-methyl-D-erythritol kinase</i>
07	06	16	07P06	XAC2735	27H03	A0CE9841F02	VIII.A	<i>conserved hypothetical protein</i>
07	07	16	07P07	XAC1158	28H04	A0UV9905D05	II.B.1	<i>adenylosuccinate synthetase</i>
07	08	16	07P08	XAC0545	27H04	A0QR9859F12	II.A.4	<i>phospho-2-dehydro-3-deoxyheptonate aldolase, phe-sensitive</i>
07	09	16	07P09	XAC1767	28H05	A0UV9905E10	I.D.2	<i>galactose-binding protein regulator</i>
07	10	16	07P10	XAC1461	27H05	A0UV9880H08	VII.C	<i>glutathione S-transferase</i>
07	11	16	07P11	no_clone	28H06	no_clone	NC	<i>no_gene</i>
07	12	16	07P12	XAC3196	27H06	A0UV9880D09	V.A.7	<i>ABC transporter ATP-binding subunit</i>
08	01	16	08P01	no_clone	32H01	no_clone	NC	<i>no_gene</i>
08	02	16	08P02	no_clone	31H01	no_clone	NC	<i>no_gene</i>
08	03	16	08P03	no_clone	32H02	no_clone	NC	<i>no_gene</i>

08	04	16	08P04	no_clone	31H02	no_clone	NC	<i>no_gene</i>
08	05	16	08P05	no_clone	32H03	no_clone	NC	<i>no_gene</i>
08	06	16	08P06	no_clone	31H03	no_clone	NC	<i>no_gene</i>
08	07	16	08P07	no_clone	32H04	no_clone	NC	<i>no_gene</i>
08	08	16	08P08	no_clone	31H04	no_clone	NC	<i>no_gene</i>
08	09	16	08P09	no_clone	32H05	no_clone	NC	<i>no_gene</i>
08	10	16	08P10	no_clone	31H05	no_clone	NC	<i>no_gene</i>
08	11	16	08P11	no_clone	32H06	no_clone	NC	<i>no_gene</i>
08	12	16	08P12	no_clone	31H06	no_clone	NC	<i>no_gene</i>
09	01	16	09C07	Score_card	33B04	Score_card	SC	<i>Score_card</i>
09	02	16	09G23	Score_card	33D12	Score_card	SC	<i>Score_card</i>
09	03	16	09M15	Score_card	33G08	Score_card	SC	<i>Score_card</i>
09	04	16	09C08	Score_card	34B04	Score_card	SC	<i>Score_card</i>
09	05	16	09G24	Score_card	34D12	Score_card	SC	<i>Score_card</i>
09	06	16	09M16	Score_card	34G08	Score_card	SC	<i>Score_card</i>
09	07	16	09D08	Score_card	35B04	Score_card	SC	<i>Score_card</i>
09	08	16	09H24	Score_card	35D12	Score_card	SC	<i>Score_card</i>
09	09	16	09N16	Score_card	35G08	Score_card	SC	<i>Score_card</i>
09	10	16	09D07	Score_card	36B04	Score_card	SC	<i>Score_card</i>
09	11	16	09H23	Score_card	36D12	Score_card	SC	<i>Score_card</i>
09	12	16	09N15	Score_card	36G08	Score_card	SC	<i>Score_card</i>
01	01	17	01A13	XAC4037	01A07	A0AC0113B07	III.A.5	<i>endonuclease</i>
01	02	17	01A14	XAC1723	02A07	A0AM1113F06	IX	<i>hydrogenase subunit</i>
01	03	17	01A15	XAC3273	01A08	A0AC0113B08	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
01	04	17	01A16	XAC2167	02A08	A0AM1113H12	I.D.1	<i>two-component system, sensor protein</i>
01	05	17	01A17	XACb0011	01A09	A0JJ0104E03	VII.A	<i>avirulence protein</i>
01	06	17	01A18	XAC0111	02A09	A0JE1144B02	VIII.C	<i>Xanthomonas conserved hypotheticalprotein</i>
01	07	17	01A19	XAC4148	01A10	A0EC0204F05	I.D.2	<i>transcriptional regulator</i>
01	08	17	01A20	XAC1653	02A10	A0JJ1166F08	III.B.4	<i>seryl-tRNA synthetase</i>
01	09	17	01A21	XAC3539	01A11	A0JJ1072B03	VII.H	<i>general secretion pathway protein J</i>
01	10	17	01A22	XAC0016	02A11	A0JE1145F05	VIII.B	<i>hypothetical protein</i>
01	11	17	01A23	XAC4307	01A12	A0JJ1072E02	VIII.A	<i>conserved hypothetical protein</i>
01	12	17	01A24	XAC1701	02A12	A0JJ1160H11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	01	17	02A13	XAC2294	05A07	A0QR5205A01	IV.C	<i>lipopolysaccharide core biosynthesis protein</i>
02	02	17	02A14	XAC3816	06A07	A0QR5529E02	VIII.A	<i>conserved hypothetical protein</i>
02	03	17	02A15	XAC3280	05A08	A0QR5323A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	04	17	02A16	XAC2768	06A08	A0QR5529E06	VIII.A	<i>conserved hypothetical protein</i>
02	05	17	02A17	XAC3769	05A09	A0QR5327H09	III.A.5	<i>endonuclease precursor</i>
02	06	17	02A18	XAC3909	06A09	A0UV5526C09	III.D.1	<i>dolichol-phosphate mannosyltransferase</i>
02	07	17	02A19	XAC0195	05A10	A0CE5318C03	II.E	<i>cardiolipin synthase</i>
02	08	17	02A20	XAC3626	06A10	A0QR5703C09	VIII.B	<i>hypothetical protein</i>
02	09	17	02A21	XAC0866	05A11	A0QR5322C07	VII.C	<i>organic solvent tolerance precursor</i>
02	10	17	02A22	XAC1199	06A11	A0JE5637B08	III.A.1	<i>DNA polymerase III, alpha chain</i>
02	11	17	02A23	XAC2245	05A12	A0QR5407C08	VIII.B	<i>hypothetical protein</i>
02	12	17	02A24	XAC1674	06A12	A0JE5637B11	I.C.3	<i>C-type cytochrome biogenesis protein</i>
03	01	17	03A13	XAC2917	09A07	A0UV6348D11	VIII.A	<i>conserved hypothetical protein</i>
03	02	17	03A14	XAC2205	10A07	A0QR6392E07	V.B	<i>chromosome partitioning related protein</i>
03	03	17	03A15	XAC1323	09A08	A0QH6326C11	III.C.1	<i>signal peptidase I</i>
03	04	17	03A16	XAC2422	10A08	A0QR6389E11	VI.B	<i>plasmid-related protein</i>
03	05	17	03A17	XAC0977	09A09	A0UV6362F06	III.B.2	<i>50S ribosomal protein L22</i>
03	06	17	03A18	XAC3575	10A09	A0QR6392H04	I.C.3	<i>flavoprotein-ubiquinone oxidoreductase</i>
03	07	17	03A19	XAC3118	09A10	A0QR6376D07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	08	17	03A20	XAC3136	10A10	A0AC6407D05	I.D.1	<i>two-component system, sensor protein</i>
03	09	17	03A21	XAC1573	09A11	A0QR6369A06	I.D.2	<i>phosphate regulon transcriptional regulator</i>
03	10	17	03A22	XAC2013	10A11	A0AC6405C01	I.A.3	<i>3-hydroxyacyl-CoA dehydrogenase</i>

03	11	17	03A23	XAC1057	09A12	A0QR6377B03	VIII.B	<i>hypothetical protein</i>
03	12	17	03A24	XAC4302	10A12	A0UV6433B10	II.D.2	<i>GTP cyclohydrolase I</i>
04	01	17	04A13	XAC3804	13A07	A0UV6708B02	IX	<i>Smg protein</i>
04	02	17	04A14	XAC0799	14A07	A0UV6710G07	VIII.A	<i>conserved hypothetical protein</i>
04	03	17	04A15	XAC0701	13A08	A0UV6707H02	VII.H	<i>type II secretion system protein J</i>
04	04	17	04A16	XAC1085	14A08	A0QH6713A01	III.C.1	<i>peptidyl-prolyl cis-trans isomerase</i>
04	05	17	04A17	XAC2995	13A09	A0UV6708D03	II.D.17	<i>tryptophan halogenase</i>
04	06	17	04A18	XAC1133	14A09	A0QR6723A03	IV.D	<i>type IV fimbriae assembly protein</i>
04	07	17	04A19	XAC1463	13A10	A0UV6708D05	III.D.2	<i>phospholipase</i>
04	08	17	04A20	XAC3244	14A10	A0UE6718B01	VIII.A	<i>conserved hypothetical protein</i>
04	09	17	04A21	XAC2997	13A11	A0UV6711F06	VIII.A	<i>conserved hypothetical protein</i>
04	10	17	04A22	XAC4128	14A11	A0QH6713B12	I.D.3	<i>extracytoplasmic sigma factor</i>
04	11	17	04A23	XAC3331	13A12	A0UV6711D05	I.B.12	<i>NADPH-sulfite reductase, iron-sulfur protein</i>
04	12	17	04A24	XAC0143	14A12	A0QH6713C11	VII.D	<i>2-keto-3-deoxygluconate kinase</i>
05	01	17	05A13	XAC1385	17A07	A0UV6736F02	V.B	<i>conserved hypothetical protein</i>
05	02	17	05A14	XAC1252	18A07	A0QR6749A05	VII.H	<i>virulence factor</i>
05	03	17	05A15	XAC3039	17A08	A0UV6736C01	II.A.2	<i>cystathionine gamma-synthase</i>
05	04	17	05A16	XAC3272	18A08	A0QR6748G06	VIII.A	<i>conserved hypothetical protein</i>
05	05	17	05A17	XAC0424	17A09	A0UV6739C11	VIII.A	<i>conserved hypothetical protein</i>
05	06	17	05A18	XAC4329	18A09	A0QR6755A05	VIII.A	<i>conserved hypothetical protein</i>
05	07	17	05A19	XAC3752	17A10	A0UV6739G05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	08	17	05A20	XAC0646	18A10	A0QR6752B04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	09	17	05A21	XAC1658	17A11	A0CE6740B08	VIII.B	<i>hypothetical protein</i>
05	10	17	05A22	XAC2921	18A11	A0UE6759B01	VIII.A	<i>conserved hypothetical protein</i>
05	11	17	05A23	XAC1803	17A12	A0CE6740F05	IV.A.1	<i>integral membrane protein</i>
05	12	17	05A24	XAC2888	18A12	A0UE6758B09	VIII.A	<i>conserved hypothetical protein</i>
06	01	17	06A13	XAC1793	21A07	A0UT6816C03	VII.D	<i>glucan 1,4-beta-glucosidase</i>
06	02	17	06A14	XAC1402	22A07	A0AC6827B01	VIII.A	<i>conserved hypothetical protein</i>
06	03	17	06A15	XAC3976	21A08	A0UT6815C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	04	17	06A16	XAC1438	22A08	A0AC6825B03	V.A.4	<i>bacterioferritin</i>
06	05	17	06A17	XAC0960	21A09	A0UT6815D06	III.B.5	<i>transcription antitermination factor</i>
06	06	17	06A18	XAC2228	22A09	A0AC6823F04	I.D.2	<i>transcriptional regulator tetR family</i>
06	07	17	06A19	XAC1692	21A10	A0UV6819H06	IV.C	<i>lipopolysaccharide biosynthesis protein</i>
06	08	17	06A20	XAC4362	22A10	A0CE6828A09	III.C.3	<i>dipeptidase</i>
06	09	17	06A21	XAC1270	21A11	A0UV6820E10	I.D.2	<i>negative regulator of sigma-B</i>
06	10	17	06A22	XAC0740	22A11	A0AC6825H08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	11	17	06A23	XAC3448	21A12	A0UV6819G03	V.A.7	<i>TonB-dependent receptor</i>
06	12	17	06A24	XAC2674	22A12	A0AC6824E01	VIII.A	<i>conserved hypothetical protein</i>
07	01	17	07A13	XAC2233	25A07	A0QR9757H03	I.C.3	<i>short chain dehydrogenase</i>
07	02	17	07A14	XAC2896	26A07	A0QR9816A01	I.C.3	<i>alcohol dehydrogenase</i>
07	03	17	07A15	XAC3751	25A08	A0QR9767E08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	04	17	07A16	XAC3240	26A08	A0UT9845G08	IV.D	<i>fimbrillin</i>
07	05	17	07A17	XAC0814	25A09	A0QR9759B03	VIII.A	<i>conserved hypothetical protein</i>
07	06	17	07A18	XAC1283	26A09	A0CE9842E10	I.D.1	<i>two-component system, sensor protein</i>
07	07	17	07A19	XAC4305	25A10	A0QR9767F09	VII.C	<i>fusaric acid resistance protein</i>
07	08	17	07A20	XAC2442	26A10	A0CE9842A03	VIII.B	<i>hypothetical protein</i>
07	09	17	07A21	XAC0464	25A11	A0QH9769H04	VIII.A	<i>conserved hypothetical protein</i>
07	10	17	07A22	XAC4021	26A11	A0UV9870E01	VIII.A	<i>conserved hypothetical protein</i>
07	11	17	07A23	XAC0045	25A12	A0QH9770D03	VIII.B	<i>hypothetical protein</i>
07	12	17	07A24	XAC0006	26A12	A0QR9860E10	VIII.A	<i>conserved hypothetical protein</i>
08	01	17	08A13	XACb0029	29A07	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	02	17	08A14	no_clone	30A07	no_clone	NC	<i>no_gene</i>
08	03	17	08A15	XAC1271	29A08	genomic_DNA	I.D.2	<i>sigma-B negative effector</i>
08	04	17	08A16	no_clone	30A08	no_clone	NC	<i>no_gene</i>
08	05	17	08A17	XAC2620	29A09	genomic_DNA	VII.H	<i>VirB9 protein</i>

08	06	17	08A18	no_clone	30A09	no_clone	NC	no_gene
08	07	17	08A19	XAC2801	29A10	genomic_DNA	VIII.A	conserved hypothetical protein
08	08	17	08A20	no_clone	30A10	no_clone	NC	no_gene
08	09	17	08A21	XAC1380	29A11	genomic_DNA	I.D.4	RNA polymerase sigma factor
08	10	17	08A22	no_clone	30A11	no_clone	NC	no_gene
08	11	17	08A23	no_clone	29A12	no_clone	NC	no_gene
08	12	17	08A24	no_clone	30A12	no_clone	NC	no_gene
09	01	17	09C09	Score_card	33B05	Score_card	SC	Score_card
09	02	17	09I01	Score_card	33E01	Score_card	SC	Score_card
09	03	17	09M17	Score_card	33G09	Score_card	SC	Score_card
09	04	17	09C10	Score_card	34B05	Score_card	SC	Score_card
09	05	17	09I02	Score_card	34E01	Score_card	SC	Score_card
09	06	17	09M18	Score_card	34G09	Score_card	SC	Score_card
09	07	17	09D10	Score_card	35B05	Score_card	SC	Score_card
09	08	17	09J02	Score_card	35E01	Score_card	SC	Score_card
09	09	17	09N18	Score_card	35G09	Score_card	SC	Score_card
09	10	17	09D09	Score_card	36B05	Score_card	SC	Score_card
09	11	17	09J01	Score_card	36E01	Score_card	SC	Score_card
09	12	17	09N17	Score_card	36G09	Score_card	SC	Score_card
01	01	18	01B13	XAC2636	04A07	A0JJ1569A11	VIII.B	hypothetical protein
01	02	18	01B14	XAC1502	03A07	A0JE1376A05	VIII.B	hypothetical protein
01	03	18	01B15	XAC2700	04A08	A0RN1551E05	I.C.1	NADH-ubiquinone oxidoreductase, NQO2 subunit
01	04	18	01B16	XAC2945	03A08	A0UV1356F01	VIII.A	conserved hypothetical protein
01	05	18	01B17	XAC4368	04A09	A0RN1551F10	V.A.4	TonB-dependent receptor
01	06	18	01B18	XAC0796	03A09	A0JJ1342A04	VIII.B	hypothetical protein
01	07	18	01B19	XAC0059	04A10	A0UV1651F08	II.A.2	asparagine synthetase like protein
01	08	18	01B20	XAC3742	03A10	A0JE1331H08	IV.C	UDP-galactopyranose mutase
01	09	18	01B21	XAC2701	04A11	A0JJ1591E02	I.C.1	NADH-ubiquinone oxidoreductase, NQO4 subunit
01	10	18	01B22	XAC3543	03A11	A0AM1404D09	VII.H	general secretion pathway protein F
01	11	18	01B23	XAC2193	04A12	A0RN1627F02	V.A.7	TonB-dependent receptor
01	12	18	01B24	XAC2356	03A12	A0JE1380B10	VII.C	drug:proton antiporter
02	01	18	02B13	XAC2251	08A07	A0QH6309B11	VIII.B	hypothetical protein
02	02	18	02B14	XAC4179	07A07	A0QR5902H10	II.E	acetyl coenzyme A synthetase
02	03	18	02B15	XAC2966	08A08	A0QH6313H04	VIII.A	conserved hypothetical protein
02	04	18	02B16	XAC0418	07A08	A0QR5902F03	VIII.C	Xanthomonas conserved hypothetical protein
02	05	18	02B17	XAC3815	08A09	A0UV6208H01	VIII.A	conserved hypothetical protein
02	06	18	02B18	XAC2490	07A09	A0QR5902C08	VIII.C	Xanthomonas conserved hypothetical protein
02	07	18	02B19	XAC4254	08A10	A0UV6210E03	I.A.1	xylanase
02	08	18	02B20	XAC4005	07A10	A0QR5704G03	VII.C	beta-lactamase related protein
02	09	18	02B21	XAC3965	08A11	A0QR6304B05	VIII.A	conserved hypothetical protein
02	10	18	02B22	XAC2217	07A11	A0QR5704F02	VIII.A	conserved hypothetical protein
02	11	18	02B23	XAC3284	08A12	A0QH6309B02	VI.C	ISxac2 transposase
02	12	18	02B24	XAC2567	07A12	A0QR5904C05	VIII.C	Xanthomonas conserved hypothetical protein
03	01	18	03B13	XAC3555	12A07	A0UV6514C04	VIII.A	conserved hypothetical protein
03	02	18	03B14	XAC2795	11A07	A0EC6476E08	VIII.A	conserved hypothetical protein
03	03	18	03B15	XAC4044	12A08	A0CE6508A11	VIII.A	conserved hypothetical protein
03	04	18	03B16	XAC3251	11A08	A0EC6476F02	VIII.B	hypothetical protein
03	05	18	03B17	XAC3249	12A09	A0QR6704E03	I.D.1	two-component system, sensor protein
03	06	18	03B18	XAC3073	11A09	A0CE6477A05	VIII.A	conserved hypothetical protein
03	07	18	03B19	XAC0396	12A10	A0QR6613B03	VII.B	HpaB protein
03	08	18	03B20	XAC1738	11A10	A0EC6474C05	II.D.11	ubiquinone biosynthesis protein
03	09	18	03B21	XAC0232	12A11	A0QR6614F03	VIII.A	conserved hypothetical protein
03	10	18	03B22	XAC2437	11A11	A0EC6475E05	III.A.5	type II restriction enzyme XamI
03	11	18	03B23	XAC1617	12A12	A0QR6614F04	VIII.C	Xanthomonas conserved hypothetical protein
03	12	18	03B24	XAC0074	11A12	A0EC6474F02	V.A.7	TonB-dependent receptor

04	01	18	04B13	XAC1364	16A07	A0QH6731A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	02	18	04B14	XAC1649	15A07	A0QR6729H05	II.A.4	<i>P-protein</i>
04	03	18	04B15	XAC0697	16A08	A0UV6735E08	VII.H	<i>type II secretion system protein F</i>
04	04	18	04B16	XAC2576	15A08	A0QR6730B08	VII.E	<i>GumK protein</i>
04	05	18	04B17	XAC1629	16A09	A0UV6733C04	VIII.A	<i>conserved hypothetical protein</i>
04	06	18	04B18	XAC0656	15A09	A0QR6730B09	IV.B	<i>rod shape-determining protein</i>
04	07	18	04B19	XAC0468	16A10	A0UV6736A06	VIII.A	<i>conserved hypothetical protein</i>
04	08	18	04B20	XAC3682	15A10	A0QR6730A03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	09	18	04B21	XAC3102	16A11	A0UV6735C05	IV.D	<i>pilus protein</i>
04	10	18	04B22	XAC0695	15A11	A0QR6730B11	VII.H	<i>type II secretion system protein D</i>
04	11	18	04B23	XAC2805	16A12	A0UV6733C12	V.A.4	<i>cation:proton antiporter</i>
04	12	18	04B24	XAC3795	15A12	A0UV6733A05	VIII.A	<i>conserved hypothetical protein</i>
05	01	18	05B13	XAC2658	20A07	A0QR6804B10	VIII.B	<i>hypothetical protein</i>
05	02	18	05B14	XAC0556	19A07	A0UV6762D04	VIII.B	<i>hypothetical protein</i>
05	03	18	05B15	XAC2134	20A08	A0QR6805E06	VIII.A	<i>conserved hypothetical protein</i>
05	04	18	05B16	XAC0470	19A08	A0UV6762F05	II.B.1	<i>phosphoribosylaminoimidazole-succinocarboxamide synthase</i>
05	05	18	05B17	XAC3276	20A09	A0QR6805E12	VIII.B	<i>hypothetical protein</i>
05	06	18	05B18	XAC3006	19A09	A0UV6763D05	V.A.7	<i>potassium channel related protein</i>
05	07	18	05B19	XAC2565	20A10	A0QR6803C07	I.B.10	<i>deoxyxylulose-5-phosphate synthase</i>
05	08	18	05B20	XAC1425	19A10	A0UV6764F12	IV.A.2	<i>outer membrane usher protein FasD</i>
05	09	18	05B21	XAC4057	20A11	A0QR6803D01	VIII.A	<i>conserved hypothetical protein</i>
05	10	18	05B22	XAC3372	19A11	A0UV6764B08	I.B.6	<i>transketolase 1</i>
05	11	18	05B23	XAC3022	20A12	A0QH6810E09	VIII.B	<i>hypothetical protein</i>
05	12	18	05B24	XAC1370	19A12	A0QR6801B04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	01	18	06B13	XAC0800	24A07	A0UT9701H12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	02	18	06B14	XAC3170	23A07	A0QH6836E04	II.D.1	<i>cytochrome P-450 hydroxylase</i>
06	03	18	06B15	XAC0132	24A08	A0QR9705D11	VIII.A	<i>conserved hypothetical protein</i>
06	04	18	06B16	XAC0546	23A08	A0QH6835G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	05	18	06B17	XAC2500	24A09	A0UT9704H04	I.D.2	<i>transcriptional regulator lacI family</i>
06	06	18	06B18	XAC0164	23A09	A0CE6831F11	V.A.7	<i>C4-dicarboxylate transport protein</i>
06	07	18	06B19	XAC2907	24A10	A0UV9709A11	VIII.A	<i>conserved hypothetical protein</i>
06	08	18	06B20	XAC1510	23A10	A0QH6835H03	VI.A	<i>phage-related integrase</i>
06	09	18	06B21	XAC0664	24A11	A0CE9727H07	IV.A.1	<i>penicillin-binding protein 6</i>
06	10	18	06B22	XAC3230	23A11	A0UV6844D03	VIII.B	<i>hypothetical protein</i>
06	11	18	06B23	XAC4001	24A12	A0CE9725C03	IV.A.1	<i>integral membrane protein</i>
06	12	18	06B24	XAC0406	23A12	A0QR6839E11	VII.B	<i>HrcU protein</i>
07	01	18	07B13	no_clone	28A07	no_clone	NC	<i>no_gene</i>
07	02	18	07B14	XAC2582	27A07	A0UV9881F05	VII.E	<i>GumE protein</i>
07	03	18	07B15	no_clone	28A08	no_clone	NC	<i>no_gene</i>
07	04	18	07B16	XAC2732	27A08	A0UV9881F06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	05	18	07B17	no_clone	28A09	no_clone	NC	<i>no_gene</i>
07	06	18	07B18	XAC2239	27A09	A0UV9880E09	VIII.A	<i>conserved hypothetical protein</i>
07	07	18	07B19	no_clone	28A10	no_clone	NC	<i>no_gene</i>
07	08	18	07B20	XAC1695	27A10	A0UV9880F08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	09	18	07B21	no_clone	28A11	no_clone	NC	<i>no_gene</i>
07	10	18	07B22	XAC1832	27A11	A0UT9878C05	II.A.5	<i>amidotransferase</i>
07	11	18	07B23	no_clone	28A12	no_clone	NC	<i>no_gene</i>
07	12	18	07B24	XAC0350	27A12	A0UT9876H09	VIII.A	<i>conserved hypothetical protein</i>
08	01	18	08B13	no_clone	32A07	no_clone	NC	<i>no_gene</i>
08	02	18	08B14	no_clone	31A07	no_clone	NC	<i>no_gene</i>
08	03	18	08B15	no_clone	32A08	no_clone	NC	<i>no_gene</i>
08	04	18	08B16	no_clone	31A08	no_clone	NC	<i>no_gene</i>
08	05	18	08B17	no_clone	32A09	no_clone	NC	<i>no_gene</i>
08	06	18	08B18	no_clone	31A09	no_clone	NC	<i>no_gene</i>

08	07	18	08B19	no_clone	32A10	no_clone	NC	<i>no_gene</i>
08	08	18	08B20	no_clone	31A10	no_clone	NC	<i>no_gene</i>
08	09	18	08B21	no_clone	32A11	no_clone	NC	<i>no_gene</i>
08	10	18	08B22	no_clone	31A11	no_clone	NC	<i>no_gene</i>
08	11	18	08B23	no_clone	32A12	no_clone	NC	<i>no_gene</i>
08	12	18	08B24	no_clone	31A12	no_clone	NC	<i>no_gene</i>
09	01	18	09C11	Score_card	33B06	Score_card	SC	<i>Score_card</i>
09	02	18	09I03	Score_card	33E02	Score_card	SC	<i>Score_card</i>
09	03	18	09M19	Score_card	33G10	Score_card	SC	<i>Score_card</i>
09	04	18	09C12	Score_card	34B06	Score_card	SC	<i>Score_card</i>
09	05	18	09I04	Score_card	34E02	Score_card	SC	<i>Score_card</i>
09	06	18	09M20	Score_card	34G10	Score_card	SC	<i>Score_card</i>
09	07	18	09D12	Score_card	35B06	Score_card	SC	<i>Score_card</i>
09	08	18	09J04	Score_card	35E02	Score_card	SC	<i>Score_card</i>
09	09	18	09N20	Score_card	35G10	Score_card	SC	<i>Score_card</i>
09	10	18	09D11	Score_card	36B06	Score_card	SC	<i>Score_card</i>
09	11	18	09J03	Score_card	36E02	Score_card	SC	<i>Score_card</i>
09	12	18	09N19	Score_card	36G10	Score_card	SC	<i>Score_card</i>
01	01	19	01C13	XAC2412	01B07	A0AC0113D09	VIII.A	<i>conserved hypothetical protein</i>
01	02	19	01C14	XAC3250	02B07	A0JE1138D03	I.D.1	<i>two-component system, regulatory protein</i>
01	03	19	01C15	XAC0521	01B08	A0AC0113F06	I.B.8	<i>phosphatidate cytidyltransferase</i>
01	04	19	01C16	XAC1458	02B08	A0JE1142A02	I.C.3	<i>ferredoxin-NADP reductase</i>
01	05	19	01C17	XAC3573	01B09	A0CE0105A06	VIII.A	<i>conserved hypothetical protein</i>
01	06	19	01C18	XAC1939	02B09	A0JJ1165H10	I.D.4	<i>GGDEF family protein</i>
01	07	19	01C19	XAC2301	01B10	A0JJ0705F07	VIII.A	<i>conserved hypothetical protein</i>
01	08	19	01C20	XAC3477	02B10	A0RN1196E12	I.C.3	<i>rhizopine catabolism protein mocA</i>
01	09	19	01C21	XAC3223	01B11	A0AR1086E11	VI.C	<i>ISxac3 transposase</i>
01	10	19	01C22	XAC1078	02B11	A0JJ1159G08	III.C.3	<i>ATP-dependent Clp protease proteolytic subunit</i>
01	11	19	01C23	XAC3426	01B12	A0AR1086H06	II.D.9	<i>flavin monoamine oxidase-related protein</i>
01	12	19	01C24	XAC3904	02B12	A0JJ1167E03	VIII.A	<i>conserved hypothetical protein</i>
02	01	19	02C13	XAC3844	05B07	A0QR5206A02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	02	19	02C14	XAC3188	06B07	A0RN5611D07	II.D.14	<i>cobyric acid synthase</i>
02	03	19	02C15	XACb0034	05B08	A0QR5407G02	VIII.B	<i>hypothetical protein</i>
02	04	19	02C16	XAC3049	06B08	A0JE5637A05	VIII.A	<i>conserved hypothetical protein</i>
02	05	19	02C17	XAC2920	05B09	A0QR5407H11	VIII.A	<i>conserved hypothetical protein</i>
02	06	19	02C18	XAC3497	06B09	A0JE5637A08	VIII.B	<i>hypothetical protein</i>
02	07	19	02C19	XAC2876	05B10	A0CE5319H04	VIII.B	<i>hypothetical protein</i>
02	08	19	02C20	XAC1228	06B10	A0QR5704C05	I.D.1	<i>two-component system, sensor protein</i>
02	09	19	02C21	XAC0382	05B11	A0QR5323F11	II.A.2	<i>aspartyl/asparaginyl beta-hydroxylase</i>
02	10	19	02C22	XAC3098	06B11	A0JE5637F02	IV.D	<i>PilL protein</i>
02	11	19	02C23	XAC1379	05B12	A0QR5407F06	VIII.A	<i>conserved hypothetical protein</i>
02	12	19	02C24	XAC4187	06B12	A0JE5637F05	I.A.2	<i>2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/ 5-carboxymethyl-2-oxo-hex-3-ene-1,7-dioatedecarboxylase</i>
03	01	19	03C13	XAC2525	09B07	A0EC6322C08	VIII.A	<i>conserved hypothetical protein</i>
03	02	19	03C14	XAC3270	10B07	A0UE6397D04	VIII.B	<i>hypothetical protein</i>
03	03	19	03C15	XAC2464	09B08	A0AC6336B04	VIII.A	<i>conserved hypothetical protein</i>
03	04	19	03C16	XAC0897	10B08	A0QR6391F05	I.D.1	<i>two-component system, sensor protein</i>
03	05	19	03C17	XAC0498	09B09	A0UV6364F04	VIII.A	<i>conserved hypothetical protein</i>
03	06	19	03C18	XAC1877	10B09	A0UE6398E11	VII.H	<i>response regulator</i>
03	07	19	03C19	XAC2201	09B10	A0UE6378F07	VII.C	<i>hemolysin secretion protein D</i>
03	08	19	03C20	XAC4294	10B10	A0QR6409B04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	09	19	03C21	XAC2159	09B11	A0UV6349C05	II.D.12	<i>siroheme synthase</i>
03	10	19	03C22	XAC2640	10B11	A0AC6407E11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	11	19	03C23	XAC2224	09B12	A0UE6379C03	VI.C	<i>ISxac3 transposase</i>
03	12	19	03C24	XAC1523	10B12	A0UV6434C01	III.C.2	<i>DnaJ protein</i>

04	01	19	04C13	XAC0327	13B07	A0UV6708C10	VII.C	<i>acriflavin resistance protein</i>
04	02	19	04C14	XAC3159	14B07	A0UV6711A05	I.B.9	<i>phospholipase C</i>
04	03	19	04C15	XAC1077	13B08	A0UV6708B04	III.C.2	<i>peptidyl-prolyl cis-trans isomerase</i>
04	04	19	04C16	XAC2723	14B08	A0UE6718E03	II.A.2	<i>aspartate semialdehyde dehydrogenase</i>
04	05	19	04C17	XAC1801	13B09	A0QR6705C01	V.A.7	<i>Prop transport protein</i>
04	06	19	04C18	XAC1269	14B09	A0QR6725A03	I.D.2	<i>positive regulator of sigma-B</i>
04	07	19	04C19	XAC3220	13B10	A0UV6708F05	II.D.7	<i>NH3-dependent NAD synthetase</i>
04	08	19	04C20	XAC0788	14B10	A0UE6718F01	V.A.6	<i>preprotein translocase SecA subunit</i>
04	09	19	04C21	XAC2618	13B11	A0UV6711G09	VII.H	<i>VirB11 protein</i>
04	10	19	04C22	XAC4365	14B11	A0QH6713H05	V.A.7	<i>export protein</i>
04	11	19	04C23	XAC2846	13B12	A0UV6711F10	I.D.2	<i>transcriptional regulator fur family</i>
04	12	19	04C24	XAC4223	14B12	A0QH6713H07	VIII.A	<i>conserved hypothetical protein</i>
05	01	19	05C13	XAC2779	17B07	A0QR6737C04	III.A.1	<i>DNA polymerase III, delta subunit</i>
05	02	19	05C14	XAC3612	18B07	A0QR6749B04	IX	<i>peptidase</i>
05	03	19	05C15	XAC1579	17B08	A0UV6736D10	IV.A.2	<i>polyphosphate-selective porin O</i>
05	04	19	05C16	XAC2724	18B08	A0QR6752G06	II.A.3	<i>2-hydroxyacid dehydrogenase</i>
05	05	19	05C17	XAC3956	17B09	A0UV6739F12	III.D.3	<i>outer membrane lipoprotein Blc</i>
05	06	19	05C18	XAC2859	18B09	A0QR6749F10	VIII.B	<i>hypothetical protein</i>
05	07	19	05C19	XAC2365	17B10	A0CE6740B06	I.A.2	<i>ethanolamine ammonia-lyase large subunit</i>
05	08	19	05C20	XAC2279	18B10	A0QR6752G03	VIII.A	<i>conserved hypothetical protein</i>
05	09	19	05C21	XAC2553	17B11	A0CE6740B12	III.A.4	<i>A/G-specific adenine glycosylase</i>
05	10	19	05C22	XAC1517	18B11	A0UE6759C05	I.D.2	<i>ferric uptake regulator</i>
05	11	19	05C23	XAC2621	17B12	A0CE6742C01	VII.H	<i>VirB8 protein</i>
05	12	19	05C24	XAC4219	18B12	A0UE6759D02	VIII.A	<i>conserved hypothetical protein</i>
06	01	19	06C13	XAC2035	21B07	A0UT6816E09	I.C.3	<i>non-heme chloroperoxidase</i>
06	02	19	06C14	XAC3591	22B07	A0AC6827F01	I.C.3	<i>short chain dehydrogenase</i>
06	03	19	06C15	XAC3019	21B08	A0UT6815H01	VIII.B	<i>hypothetical protein</i>
06	04	19	06C16	XAC0124	22B08	A0AC6825D11	I.B.3	<i>fructose-1,6-bisphosphatase</i>
06	05	19	06C17	XAC2480	21B09	A0UT6815H11	II.A.1	<i>glutamine synthase</i>
06	06	19	06C18	XAC3371	22B09	A0AC6824D07	III.C.3	<i>proline imino-peptidase</i>
06	07	19	06C19	XAC1770	21B10	A0UV6820B11	VII.D	<i>cellulase</i>
06	08	19	06C20	XAC1549	22B10	A0AC6823B06	V.A.7	<i>ABC transporter vitamin B12 uptakepermease</i>
06	09	19	06C21	XAC1445	21B11	A0UV6820H09	VII.C	<i>multidrug resistance efflux pump</i>
06	10	19	06C22	XAC1161	22B11	A0AC6827C09	VIII.A	<i>conserved hypothetical protein</i>
06	11	19	06C23	XAC0941	21B12	A0UV6820A02	I.D.2	<i>transcriptional regulator</i>
06	12	19	06C24	XAC2421	22B12	A0AC6825C07	VIII.B	<i>hypothetical protein</i>
07	01	19	07C13	XAC0618	25B07	A0QR9767D08	VII.G	<i>periplasmic glucan biosynthesis protein</i>
07	02	19	07C14	XAC3301	26B07	A0UV9819D11	VIII.A	<i>conserved hypothetical protein</i>
07	03	19	07C15	XAC0601	25B08	A0QH9769B02	VIII.B	<i>hypothetical protein</i>
07	04	19	07C16	XAC0550	26B08	A0UT9846G07	I.D.4	<i>glutamine synthetase adenyllyltransferase</i>
07	05	19	07C17	XAC3864	25B09	A0AC9763E09	V.A.4	<i>cationic amino acid transporter</i>
07	06	19	07C18	XAC3921	26B09	A0UT9845G11	IX	<i>glucosyltransferase</i>
07	07	19	07C19	XAC1542	25B10	A0QH9769G01	I.C.7	<i>fumarate hydratase</i>
07	08	19	07C20	XAC0403	26B10	A0CE9842G03	VII.B	<i>HrcQ protein</i>
07	09	19	07C21	XAC0764	25B11	A0QR9776E04	VIII.A	<i>conserved hypothetical protein</i>
07	10	19	07C22	XAC1298	26B11	A0CE9839F02	I.D.2	<i>transcriptional regulator</i>
07	11	19	07C23	XAC0720	25B12	A0QR9776E07	V.A.7	<i>high-affinity choline transport</i>
07	12	19	07C24	XAC3694	26B12	A0QH9865A02	V.C	<i>chemotaxis MotB protein</i>
08	01	19	08C13	XACb0022	29B07	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	02	19	08C14	no_clone	30B07	no_clone	NC	<i>no_gene</i>
08	03	19	08C15	XAC2606	29B08	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	04	19	08C16	no_clone	30B08	no_clone	NC	<i>no_gene</i>
08	05	19	08C17	XACb0041	29B09	genomic_DNA	VII.H	<i>VirB6 protein</i>
08	06	19	08C18	no_clone	30B09	no_clone	NC	<i>no_gene</i>
08	07	19	08C19	XAC2555	29B10	genomic_DNA	I.D.3	<i>sensor histidine kinase</i>

08	08	19	08C20	no_clone	30B10	no_clone	NC	no_gene
08	09	19	08C21	XAC4128	29B11	genomic_DNA	I.D.3	extracytoplasmic sigma factor
08	10	19	08C22	no_clone	30B11	no_clone	NC	no_gene
08	11	19	08C23	no_clone	29B12	no_clone	NC	no_gene
08	12	19	08C24	no_clone	30B12	no_clone	NC	no_gene
09	01	19	09C13	Score_card	33B07	Score_card	SC	Score_card
09	02	19	09I05	Score_card	33E03	Score_card	SC	Score_card
09	03	19	09M21	Score_card	33G11	Score_card	SC	Score_card
09	04	19	09C14	Score_card	34B07	Score_card	SC	Score_card
09	05	19	09I06	Score_card	34E03	Score_card	SC	Score_card
09	06	19	09M22	Score_card	34G11	Score_card	SC	Score_card
09	07	19	09D14	Score_card	35B07	Score_card	SC	Score_card
09	08	19	09J06	Score_card	35E03	Score_card	SC	Score_card
09	09	19	09N22	Score_card	35G11	Score_card	SC	Score_card
09	10	19	09D13	Score_card	36B07	Score_card	SC	Score_card
09	11	19	09J05	Score_card	36E03	Score_card	SC	Score_card
09	12	19	09N21	Score_card	36G11	Score_card	SC	Score_card
01	01	02	01B01	XAC0560	04A01	A0UT1518G10	II.E	alpha subunit of malonate decarboxylase
01	02	02	01B02	XAC0811	03A01	A0JE1149B03	V.A.7	TonB-dependent receptor
01	03	02	01B03	XAC3508	04A02	A0AM1489A01	VIII.A	conserved hypothetical protein
01	04	02	01B04	XAC0478	03A02	A0JJ1132B08	II.A.4	anthranilate synthase component II
01	05	02	01B05	XAC0043	04A03	A0JJ1443G01	IV.C	UDP-glucose lipid carrier transferase
01	06	02	01B06	XAC4012	03A03	A0JJ1311A05	VIII.C	Xanthomonas conserved hypothetical protein
01	07	02	01B07	XAC0096	04A04	A0JJ1546D10	VIII.C	Xanthomonas conserved hypothetical protein
01	08	02	01B08	XAC4327	03A04	A0JE1321A12	I.A.2	urea amidolyase
01	09	02	01B09	XAC1615	04A05	A0JJ1541D02	I.D.2	transcriptional regulator protein Pai2
01	10	02	01B10	XAC1363	03A05	A0AC1314F05	V.A.7	MFS transporter
01	11	02	01B11	XAC1663	04A06	A0JJ1437G11	VIII.B	hypothetical protein
01	12	02	01B12	XAC2664	03A06	A0AM1405F09	IV.A.2	PilE protein
02	01	02	02B01	XAC1717	08A01	A0QR5904A01	IV.C	2-dehydro-3-deoxyphosphooctonate aldolase
02	02	02	02B02	XAC4196	07A01	A0JE5637C08	V.A.7	cation symporter
02	03	02	02B03	XAC3023	08A02	A0UV6205D02	VIII.B	hypothetical protein
02	04	02	02B04	XAC4193	07A02	A0QR5704D01	I.D.1	two-component system, sensor protein
02	05	02	02B05	XAC0700	08A03	A0QR6002A03	VII.H	type II secretion system protein I
02	06	02	02B06	XAC3038	07A03	A0QR5704D11	II.A.2	homoserine dehydrogenase
02	07	02	02B07	XAC2619	08A04	A0QR6001A01	VII.H	VirB10 protein
02	08	02	02B08	XAC2815	07A04	A0QR5701F02	VIII.C	Xanthomonas conserved hypothetical protein
02	09	02	02B09	XAC3180	08A05	A0QR6002G02	V.A.7	iron transporter
02	10	02	02B10	XAC1208	07A05	A0QR5903B01	VIII.C	Xanthomonas conserved hypothetical protein
02	11	02	02B11	XAC2264	08A06	A0UV6205C06	VIII.B	hypothetical protein
02	12	02	02B12	XAC2498	07A06	A0QR5902H09	VII.C	multidrug resistance protein
03	01	02	03B01	XAC0112	12A01	A0EC6476C04	VIII.A	conserved hypothetical protein
03	02	02	03B02	XAC1394	11A01	A0AC6440G07	VIII.A	conserved hypothetical protein
03	03	02	03B03	XAC1181	12A02	A0UV6485F06	VIII.A	conserved hypothetical protein
03	04	02	03B04	XAC2410	11A02	A0QR6414A07	III.A.4	exodeoxyribonuclease VII large subunit
03	05	02	03B05	XAC0810	12A03	A0QR6489A09	VIII.C	Xanthomonas conserved hypothetical protein
03	06	02	03B06	XAC3784	11A03	A0QH6412H03	VIII.B	hypothetical protein
03	07	02	03B07	XAC4227	12A04	A0QR6491H08	I.A.1	alpha-glucuronidase
03	08	02	03B08	XAC1350	11A04	A0QH6412H04	I.A.2	leucine dehydrogenase
03	09	02	03B09	XAC3726	12A05	A0QR6489E03	VIII.A	conserved hypothetical protein
03	10	02	03B10	XAC3518	11A05	A0EC6427B05	III.D.1	cellulose synthase
03	11	02	03B11	XAC3623	12A06	A0UV6514G03	II.E	beta-hydroxydecanoyl-ACP dehydratase
03	12	02	03B12	XAC1526	11A06	A0QR6460G08	VII.C	outer membrane component of multidrug efflux pump
04	01	02	04B01	XAC1963	16A01	A0QR6730C02	II.E	3-oxoacyl-[ACP] reductase
04	02	02	04B02	XAC1235	15A01	A0UV6712H03	VIII.C	Xanthomonas conserved hypothetical protein

04	03	02	04B03	XAC2531	16A02	A0QH6732C04	V.A.7	<i>TonB-dependent receptor</i>
04	04	02	04B04	XAC1339	15A02	A0QH6713E05	VIII.A	<i>conserved hypothetical protein</i>
04	05	02	04B05	XAC3716	16A03	A0QH6731H06	VIII.B	<i>hypothetical protein</i>
04	06	02	04B06	XAC2864	15A03	A0QH6713F05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	07	02	04B07	XAC2715	16A04	A0QH6731H12	II.E	<i>acetyl-coenzyme A carboxylase carboxyl transferase</i>
04	08	02	04B08	XAC3398	15A04	A0QR6728D12	VIII.A	<i>conserved hypothetical protein</i>
04	09	02	04B09	XAC3364	16A05	A0QR6730H03	VIII.A	<i>conserved hypothetical protein</i>
04	10	02	04B10	XAC3310	15A05	A0QR6729A10	I.D.2	<i>transcriptional regulator lacI family</i>
04	11	02	04B11	XAC1136	16A06	A0QR6730H04	I.D.2	<i>propionate catabolism regulatory protein</i>
04	12	02	04B12	XAC1721	15A06	A0QR6730B05	II.E	<i>4-diphosphocytidyl-2C-methyl-D-erythritol synthase</i>
05	01	02	05B01	XAC0728	20A01	A0QR6801C12	II.D.10	<i>glutamate-cysteine ligase precursor</i>
05	02	02	05B02	XAC2798	19A01	A0UV6762B01	VIII.A	<i>conserved hypothetical protein</i>
05	03	02	05B03	XAC1345	20A02	A0QR6802E09	I.D.3	<i>sensor histidine kinase</i>
05	04	02	05B04	XAC3120	19A02	A0UV6761H03	I.C.4	<i>glucose kinase</i>
05	05	02	05B05	XAC0449	20A03	A0QR6802F08	V.A.1	<i>di-tripeptide transporter</i>
05	06	02	05B06	XAC2572	19A03	A0UE6759E02	IX	<i>GumN protein</i>
05	07	02	05B07	XAC4359	20A04	A0QR6802C08	I.D.2	<i>sugar diacyl regulator</i>
05	08	02	05B08	XACb0045	19A04	A0UE6759E05	VII.H	<i>VirB4 protein</i>
05	09	02	05B09	XAC3685	20A05	A0UV6766F06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	10	02	05B10	XAC3551	19A05	A0UV6761A08	III.A.3	<i>integrase/recombinase XerD</i>
05	11	02	05B11	XAC3468	20A06	A0UV6766D03	VIII.A	<i>conserved hypothetical protein</i>
05	12	02	05B12	XAC2322	19A06	A0UV6762G02	II.A.2	<i>succinyl-diaminopimelate transaminase</i>
06	01	02	06B01	XAC1662	24A01	A0UV6843G12	VI.A	<i>phage-related protein</i>
06	02	02	06B02	XAC3370	23A01	A0AC6823C04	V.A.4	<i>outer membrane receptor for ferric iron uptake</i>
06	03	02	06B03	XAC3427	24A02	A0UV6843H04	V.A.7	<i>TonB-dependent receptor</i>
06	04	02	06B04	XAC2709	23A02	A0AC6827F10	IV.A.2	<i>cyanoglobin</i>
06	05	02	06B05	XAC3634	24A03	A0UV6844G04	VIII.A	<i>conserved hypothetical protein</i>
06	06	02	06B06	XAC3780	23A03	A0AC6825E10	V.A.7	<i>chloride channel</i>
06	07	02	06B07	XAC1337	24A04	A0UV6844G05	VII.G	<i>cold shock protein</i>
06	08	02	06B08	XAC0547	23A04	A0QH6836D01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	02	06B09	XAC1812	24A05	A0UV6845E06	IX	<i>HmsF protein</i>
06	10	02	06B10	XAC0057	23A05	A0CE6829A09	V.A.7	<i>transport protein</i>
06	11	02	06B11	XAC3807	24A06	A0UV6846E01	III.A.1	<i>DNA topoisomerase I</i>
06	12	02	06B12	XAC2580	23A06	A0CE6831A10	VII.E	<i>GumG protein</i>
07	01	02	07B01	XAC2889	28A01	A0UV9882E10	VI.C	<i>ISxsc2 transposase</i>
07	02	02	07B02	XAC3035	27A01	A0QR9859C02	VIII.A	<i>conserved hypothetical protein</i>
07	03	02	07B03	XAC4243	28A02	A0UT9891A10	I.A.2	<i>L-fucose-phosphate aldolase</i>
07	04	02	07B04	XAC2836	27A02	A0UT9848A06	V.A.7	<i>MFS transporter</i>
07	05	02	07B05	XAC4173	28A03	A0UV9900C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	06	02	07B06	XAC3401	27A03	A0UT9846D11	VIII.A	<i>conserved hypothetical protein</i>
07	07	02	07B07	XAC4181	28A04	A0UV9904C05	VIII.A	<i>conserved hypothetical protein</i>
07	08	02	07B08	XAC3195	27A04	A0CE9842D07	III.C.3	<i>ATP-dependent Clp protease subunit</i>
07	09	02	07B09	XAC4118	28A05	A0UT9891E01	VIII.A	<i>conserved hypothetical protein</i>
07	10	02	07B10	XAC2998	27A05	A0QR9864E12	V.A.7	<i>TonB-dependent receptor</i>
07	11	02	07B11	XAC3476	28A06	A0AC9895B11	I.D.2	<i>transcriptional regulator</i>
07	12	02	07B12	XAC2748	27A06	A0UV9882F03	IV.A.1	<i>integral membrane protein</i>
08	01	02	08B01	no_clone	32A01	no_clone	NC	<i>no_gene</i>
08	02	02	08B02	no_clone	31A01	no_clone	NC	<i>no_gene</i>
08	03	02	08B03	no_clone	32A02	no_clone	NC	<i>no_gene</i>
08	04	02	08B04	no_clone	31A02	no_clone	NC	<i>no_gene</i>
08	05	02	08B05	no_clone	32A03	no_clone	NC	<i>no_gene</i>
08	06	02	08B06	no_clone	31A03	no_clone	NC	<i>no_gene</i>
08	07	02	08B07	no_clone	32A04	no_clone	NC	<i>no_gene</i>
08	08	02	08B08	no_clone	31A04	no_clone	NC	<i>no_gene</i>
08	09	02	08B09	no_clone	32A05	no_clone	NC	<i>no_gene</i>

08	10	02	08B10	no_clone	31A05	no_clone	NC	<i>no_gene</i>
08	11	02	08B11	no_clone	32A06	no_clone	NC	<i>no_gene</i>
08	12	02	08B12	no_clone	31A06	no_clone	NC	<i>no_gene</i>
09	01	02	09A03	Score_card	33A02	Score_card	SC	<i>Score_card</i>
09	02	02	09E19	Score_card	33C10	Score_card	SC	<i>Score_card</i>
09	03	02	09K11	Score_card	33F06	Score_card	SC	<i>Score_card</i>
09	04	02	09A04	Score_card	34A02	Score_card	SC	<i>Score_card</i>
09	05	02	09E20	Score_card	34C10	Score_card	SC	<i>Score_card</i>
09	06	02	09K12	Score_card	34F06	Score_card	SC	<i>Score_card</i>
09	07	02	09B04	Score_card	35A02	Score_card	SC	<i>Score_card</i>
09	08	02	09F20	Score_card	35C10	Score_card	SC	<i>Score_card</i>
09	09	02	09L12	Score_card	35F06	Score_card	SC	<i>Score_card</i>
09	10	02	09B03	Score_card	36A02	Score_card	SC	<i>Score_card</i>
09	11	02	09F19	Score_card	36C10	Score_card	SC	<i>Score_card</i>
09	12	02	09L11	Score_card	36F06	Score_card	SC	<i>Score_card</i>
01	01	20	01D13	XAC1621	04B07	A0JJ1439G07	III.B.2	<i>30S ribosomal protein S18</i>
01	02	20	01D14	XAC0826	03B07	A0JJ1382F06	VIII.A	<i>conserved hypothetical protein</i>
01	03	20	01D15	XAC2423	04B08	A0JJ1571E03	VI.C	<i>IS1478 transposase</i>
01	04	20	01D16	XAC0311	03B08	A0AM1363H10	I.A.2	<i>vanillate O-demethylase oxygenasesubunit</i>
01	05	20	01D17	XAC2196	04B09	A0QH6836E03	VIII.A	<i>conserved hypothetical protein</i>
01	06	20	01D18	XAC1214	03B09	A0UV1356H05	I.A.2	<i>glycine decarboxylase</i>
01	07	20	01D19	XAC1708	04B10	A0QR5104B11	IV.C	<i>ExoD protein</i>
01	08	20	01D20	XAC1754	03B10	A0JJ1343C04	VIII.A	<i>conserved hypothetical protein</i>
01	09	20	01D21	XAC2759	04B11	A0JJ1599F12	I.B.9	<i>alkaline phosphatase</i>
01	10	20	01D22	XAC1620	03B11	A0AM1424A11	III.B.2	<i>30S ribosomal protein S6</i>
01	11	20	01D23	XAC4105	04B12	A0RN1632B02	VII.C	<i>AMP-ligase</i>
01	12	20	01D24	XAC2981	03B12	A0AM1424E09	VIII.A	<i>conserved hypothetical protein</i>
02	01	20	02D13	XAC0411	08B07	A0CE6315H09	VII.B	<i>HrpB5 protein</i>
02	02	20	02D14	XAC3103	07B07	A0QR5903G01	II.D.10	<i>glutathione synthetase</i>
02	03	20	02D15	XAC0605	08B08	A0EC6316B08	VIII.A	<i>conserved hypothetical protein</i>
02	04	20	02D16	XAC0692	07B08	A0QR5902G07	VIII.A	<i>conserved hypothetical protein</i>
02	05	20	02D17	XAC0280	08B09	A0UV6210E02	IX	<i>ATPase</i>
02	06	20	02D18	XAC1897	07B09	A0QR5902F04	V.C	<i>chemotaxis protein</i>
02	07	20	02D19	XAC3884	08B10	A0QH6306C09	I.C.1	<i>cytochrome C oxidase, polypeptide III</i>
02	08	20	02D20	XAC0612	07B10	A0QR5902D04	VII.D	<i>cellulase</i>
02	09	20	02D21	XAC0780	08B11	A0QH6308G12	IV.B	<i>UDP-N-acetylmuramate--alanine ligase</i>
02	10	20	02D22	XAC3550	07B11	A0QR5704G06	III.C.1	<i>disulfide isomerase</i>
02	11	20	02D23	XAC0432	08B12	A0QH6309F03	VIII.A	<i>conserved hypothetical protein</i>
02	12	20	02D24	XAC2374	07B12	A0QR5904E12	VII.D	<i>polygalacturonase</i>
03	01	20	03D13	XAC1875	12B07	A0UV6514H09	III.C.1	<i>peptide chain release factor 2</i>
03	02	20	03D14	XAC2606	11B07	A0CE6480B03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	03	20	03D15	XAC0052	12B08	A0UV6511H02	VIII.B	<i>hypothetical protein</i>
03	04	20	03D16	XAC1490	11B08	A0CE6480B05	VIII.A	<i>conserved hypothetical protein</i>
03	05	20	03D17	XAC3385	12B09	A0QR6606F05	IV.D	<i>fimbrial assembly membrane protein</i>
03	06	20	03D18	XAC3216	11B09	A0CE6480C08	VIII.A	<i>conserved hypothetical protein</i>
03	07	20	03D19	XAC2413	12B10	A0QR6614A05	III.B.4	<i>ribonuclease D</i>
03	08	20	03D20	XAC3317	11B10	A0CE6478A11	III.B.3	<i>acetyltransferase</i>
03	09	20	03D21	XAC0355	12B11	A0QR6701H06	I.D.2	<i>PobR regulator</i>
03	10	20	03D22	XAC1713	11B11	A0CE6478E03	III.C.3	<i>carboxypeptidase-related protein</i>
03	11	20	03D23	XAC0007	12B12	A0QR6702C11	VIII.A	<i>conserved hypothetical protein</i>
03	12	20	03D24	XAC3566	11B12	A0EC6475E09	VIII.A	<i>conserved hypothetical protein</i>
04	01	20	04D13	XAC2938	16B07	A0QH6731G06	II.A.2	<i>cysteine desulfurase</i>
04	02	20	04D14	XAC1775	15B07	A0QR6730B07	I.A.2	<i>D-xylulokinase</i>
04	03	20	04D15	XAC3989	16B08	A0UV6736A01	I.D.4	<i>ECF sigma factor</i>
04	04	20	04D16	XAC0658	15B08	A0QR6725F07	IV.B	<i>rod shape-determining protein</i>

04	05	20	04D17	XACb0046	16B09	A0UV6733D10	VII.H	<i>VirB3 protein</i>
04	06	20	04D18	XAC4048	15B09	A0QR6725G09	V.A.7	<i>TonB-dependent receptor</i>
04	07	20	04D19	XAC1881	16B10	A0UV6733C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	08	20	04D20	XAC4248	15B10	A0QR6730B10	I.A.2	<i>gluconolactonase precursor</i>
04	09	20	04D21	XAC3125	16B11	A0UV6736A07	VIII.A	<i>conserved hypothetical protein</i>
04	10	20	04D22	XAC0251	15B11	A0QR6730C11	I.D.2	<i>transcriptional regulator tetR family</i>
04	11	20	04D23	XAC1474	16B12	A0UV6733E09	VII.C	<i>glutathione S-transferase</i>
04	12	20	04D24	XAC0339	15B12	A0QR6730C01	I.C.3	<i>oxidoreductase</i>
05	01	20	05D13	XAC3744	20B07	A0QR6804D07	V.A.7	<i>ATP binding transporter 1</i>
05	02	20	05D14	XAC1534	19B07	A0UV6762G05	I.C.7	<i>dihydrolipoamide S-succinyltransferase</i>
05	03	20	05D15	XAC1664	20B08	A0QH6810A05	VI.A	<i>phage-related integrase</i>
05	04	20	05D16	XAC0623	19B08	A0UV6762G07	VIII.A	<i>conserved hypothetical protein</i>
05	05	20	05D17	XAC2727	20B09	A0QH6810A11	VIII.A	<i>conserved hypothetical protein</i>
05	06	20	05D18	XAC0238	19B09	A0UV6763G05	II.E	<i>NAD(P)H steroid dehydrogenase</i>
05	07	20	05D19	XAC0319	20B10	A0QR6804C09	IX	<i>chloroperoxidase</i>
05	08	20	05D20	XAC3333	19B10	A0UV6765A07	VIII.A	<i>conserved hypothetical protein</i>
05	09	20	05D21	XAC2985	20B11	A0QR6803H02	V.A.1	<i>amino acid transporter</i>
05	10	20	05D22	XAC1128	19B11	A0UV6765D02	II.E	<i>acyl carrier protein</i>
05	11	20	05D23	XAC1275	20B12	A0QH6811A01	I.A.1	<i>xylosidase/arabinosidase</i>
05	12	20	05D24	XAC4267	19B12	A0QR6802B03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	01	20	06D13	XAC1016	24B07	A0QR9706E11	II.A.3	<i>2-amino-3-ketobutyrate CoA ligase</i>
06	02	20	06D14	XAC3262	23B07	A0QR6838A01	VIII.B	<i>hypothetical protein</i>
06	03	20	06D15	XAC4206	24B08	A0UV9710E01	VIII.B	<i>hypothetical protein</i>
06	04	20	06D16	XAC0504	23B08	A0QH6836C08	IV.A.1	<i>transmembrane protein</i>
06	05	20	06D17	XAC0087	24B09	A0QR9705E01	VIII.B	<i>hypothetical protein</i>
06	06	20	06D18	XAC1609	23B09	A0QH6835H02	VIII.B	<i>hypothetical protein</i>
06	07	20	06D19	XAC3454	24B10	A0UV9710F07	I.A.2	<i>threonine dehydratase catabolic</i>
06	08	20	06D20	XAC1783	23B10	A0QH6836C12	III.B.5	<i>polynucleotide adenyltransferase</i>
06	09	20	06D21	XAC2744	24B11	A0QH9738G07	II.D.17	<i>phytoene dehydrogenase</i>
06	10	20	06D22	XAC3365	23B11	A0QR6839C01	VIII.A	<i>conserved hypothetical protein</i>
06	11	20	06D23	XAC3206	24B12	A0QR9732C03	VIII.A	<i>conserved hypothetical protein</i>
06	12	20	06D24	XAC4345	23B12	A0QR6840C01	IX	<i>enterocin biosynthesis related protein</i>
07	01	20	07D13	no_clone	28B07	no_clone	NC	<i>no_gene</i>
07	02	20	07D14	XAC0530	27B07	A0UV9884D05	II.E	<i>biotin carboxylase subunit of acetyl CoA carboxylase</i>
07	03	20	07D15	no_clone	28B08	no_clone	NC	<i>no_gene</i>
07	04	20	07D16	XAC0972	27B08	A0UV9882H06	III.B.2	<i>50S ribosomal protein L3</i>
07	05	20	07D17	no_clone	28B09	no_clone	NC	<i>no_gene</i>
07	06	20	07D18	XAC3457	27B09	A0UV9881F07	II.A.2	<i>3-isopropylmalate dehydratase small subunit</i>
07	07	20	07D19	no_clone	28B10	no_clone	NC	<i>no_gene</i>
07	08	20	07D20	XAC1336	27B10	A0UV9881G05	II.B.4	<i>purine nucleoside phosphorylase</i>
07	09	20	07D21	no_clone	28B11	no_clone	NC	<i>no_gene</i>
07	10	20	07D22	XAC0619	27B11	A0UV9880G04	IX	<i>carboxylesterase</i>
07	11	20	07D23	no_clone	28B12	no_clone	NC	<i>no_gene</i>
07	12	20	07D24	XAC3065	27B12	A0UT9878E04	VIII.A	<i>conserved hypothetical protein</i>
08	01	20	08D13	no_clone	32B07	no_clone	NC	<i>no_gene</i>
08	02	20	08D14	no_clone	31B07	no_clone	NC	<i>no_gene</i>
08	03	20	08D15	no_clone	32B08	no_clone	NC	<i>no_gene</i>
08	04	20	08D16	no_clone	31B08	no_clone	NC	<i>no_gene</i>
08	05	20	08D17	no_clone	32B09	no_clone	NC	<i>no_gene</i>
08	06	20	08D18	no_clone	31B09	no_clone	NC	<i>no_gene</i>
08	07	20	08D19	no_clone	32B10	no_clone	NC	<i>no_gene</i>
08	08	20	08D20	no_clone	31B10	no_clone	NC	<i>no_gene</i>
08	09	20	08D21	no_clone	32B11	no_clone	NC	<i>no_gene</i>
08	10	20	08D22	no_clone	31B11	no_clone	NC	<i>no_gene</i>
08	11	20	08D23	no_clone	32B12	no_clone	NC	<i>no_gene</i>

08	12	20	08D24	no_clone	31B12	no_clone	NC	<i>no_gene</i>
09	01	20	09C15	Score_card	33B08	Score_card	SC	<i>Score_card</i>
09	02	20	09I07	Score_card	33E04	Score_card	SC	<i>Score_card</i>
09	03	20	09M23	Score_card	33G12	Score_card	SC	<i>Score_card</i>
09	04	20	09C16	Score_card	34B08	Score_card	SC	<i>Score_card</i>
09	05	20	09I08	Score_card	34E04	Score_card	SC	<i>Score_card</i>
09	06	20	09M24	Score_card	34G12	Score_card	SC	<i>Score_card</i>
09	07	20	09D16	Score_card	35B08	Score_card	SC	<i>Score_card</i>
09	08	20	09J08	Score_card	35E04	Score_card	SC	<i>Score_card</i>
09	09	20	09N24	Score_card	35G12	Score_card	SC	<i>Score_card</i>
09	10	20	09D15	Score_card	36B08	Score_card	SC	<i>Score_card</i>
09	11	20	09J07	Score_card	36E04	Score_card	SC	<i>Score_card</i>
09	12	20	09N23	Score_card	36G12	Score_card	SC	<i>Score_card</i>
01	01	21	01E13	XAC2937	01C07	A0AC0113H06	V.A.7	<i>ABC transporter permease</i>
01	02	21	01E14	XAC3450	02C07	A0JJ1169E10	VII.C	<i>gamma-glutamyltranspeptidase</i>
01	03	21	01E15	XACa0028	01C08	A0AC0115A01	VI.B	<i>plasmid stable inheritance protein I</i>
01	04	21	01E16	XAC1004	02C08	A0RN1171D08	V.B	<i>GTP-binding elongation factor protein</i>
01	05	21	01E17	XAC1234	01C09	A0CE0105C09	VIII.A	<i>conserved hypothetical protein</i>
01	06	21	01E18	XAC4300	02C09	A0JJ1178E07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	07	21	01E19	XAC3507	01C10	A0JJ1001G02	VII.D	<i>cellulase S (truncated)</i>
01	08	21	01E20	XAC2547	02C10	A0JE1206C10	II.A.2	<i>dihydrodipicolinate synthetase</i>
01	09	21	01E21	XAC0393	01C11	A0EC1093D06	VII.B	<i>HpaF protein</i>
01	10	21	01E22	XAC1712	02C11	A0JJ1167A10	VIII.A	<i>conserved hypothetical protein</i>
01	11	21	01E23	XAC4197	01C12	A0JJ0133G11	I.A.2	<i>gluconokinase</i>
01	12	21	01E24	XAC0759	02C12	A0JJ1197D07	I.D.1	<i>two-component system, sensor protein</i>
02	01	21	02E13	XAC4086	05C07	A0UV5307B04	II.E	<i>beta-ketoacyl-[ACP] synthase II</i>
02	02	21	02E14	XAC2351	06C07	A0CE5420D02	II.A.1	<i>argininosuccinate synthase</i>
02	03	21	02E15	XAC3860	05C08	A0QR5407H10	IV.B	<i>N-acetylmuramoyl-L-alanine amidase</i>
02	04	21	02E16	XAC1872	06C08	A0CE5420G07	VI.C	<i>transposase</i>
02	05	21	02E17	XAC2181	05C09	A0CE5309H01	VIII.A	<i>conserved hypothetical protein</i>
02	06	21	02E18	XAC1901	06C09	A0JE5637A11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	07	21	02E19	XAC0758	05C10	A0QR5323E10	V.A.7	<i>potassium-transporting ATPase, C chain</i>
02	08	21	02E20	XAC0484	06C10	A0JE5637B03	II.F	<i>S-adenosyl methionine decarboxylase proenzyme</i>
02	09	21	02E21	XACb0072	05C11	A0QR5407H04	VI.B	<i>resolvase</i>
02	10	21	02E22	XACa0011	06C11	A0QR5701C07	VIII.B	<i>hypothetical protein</i>
02	11	21	02E23	XAC2174	05C12	A0QR5317B09	VI.C	<i>ISxac3 transposase</i>
02	12	21	02E24	XAC4145	06C12	A0QR5701A01	VIII.A	<i>conserved hypothetical protein</i>
03	01	21	03E13	XAC1739	09C07	A0QH6326B08	I.D.2	<i>LexA repressor</i>
03	02	21	03E14	XAC2575	10C07	A0QR6404E10	VII.E	<i>GumL protein</i>
03	03	21	03E15	XAC3674	09C08	A0CE6341C02	VIII.A	<i>conserved hypothetical protein</i>
03	04	21	03E16	XAC0513	10C08	A0QR6392E10	II.B.1	<i>bifunctional purine biosynthesis protein</i>
03	05	21	03E17	XAC3288	09C09	A0QR6366E02	VIII.A	<i>conserved hypothetical protein</i>
03	06	21	03E18	XAC0348	10C09	A0QR6403B06	IX	<i>transferase</i>
03	07	21	03E19	XAC0882	09C10	A0UV6349B08	I.C.3	<i>aldehyde dehydrogenase</i>
03	08	21	03E20	XAC3114	10C10	A0QR6410F09	II.D.11	<i>pyrroloquinoline quinone biosynthesis protein G</i>
03	09	21	03E21	XAC3479	09C11	A0UV6351E06	VIII.A	<i>conserved hypothetical protein</i>
03	10	21	03E22	XAC1669	10C11	A0QR6409B06	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
03	11	21	03E23	XAC0787	09C12	A0UV6349G11	III.C.3	<i>peptidase</i>
03	12	21	03E24	XAC0847	10C12	A0UV6435A01	V.A.7	<i>ABC transporter ATP-binding protein</i>
04	01	21	04E13	XAC0233	13C07	A0UV6708E11	II.E	<i>3-oxoacyl-[ACP] synthase III</i>
04	02	21	04E14	XAC0018	14C07	A0UV6711D01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	03	21	04E15	XAC0926	13C08	A0UV6708E12	VIII.A	<i>conserved hypothetical protein</i>
04	04	21	04E16	XAC0599	14C08	A0UE6718F11	VIII.B	<i>hypothetical protein</i>
04	05	21	04E17	XAC3830	13C09	A0QR6705D10	II.D.10	<i>thioredoxin</i>
04	06	21	04E18	XAC0179	14C09	A0QR6725B12	V.A.7	<i>ABC transporter ATP-binding protein</i>

04	07	21	04E19	XAC3579	13C10	A0UV6709B03	VII.E	<i>phosphoglucomutase/phosphomannomutase</i>
04	08	21	04E20	XAC4045	14C10	A0UE6718G10	VIII.B	<i>hypothetical protein</i>
04	09	21	04E21	XAC3929	13C11	A0UV6709B06	I.D.2	<i>nitrogen regulatory protein P-II</i>
04	10	21	04E22	XAC3053	14C11	A0UE6718C01	VIII.A	<i>conserved hypothetical protein</i>
04	11	21	04E23	XAC4183	13C12	A0UV6709B07	I.A.1	<i>xylosidase</i>
04	12	21	04E24	XACa0009	14C12	A0UE6718C05	VI.C	<i>ISxac3 transposase</i>
05	01	21	05E13	XAC0690	17C07	A0UV6738A04	V.A.7	<i>TonB-dependent receptor</i>
05	02	21	05E14	XAC1529	18C07	A0CE6746D03	I.D.2	<i>hydrogen peroxide-inducible genes activator</i>
05	03	21	05E15	XACb0064	17C08	A0UV6736F09	VIII.B	<i>hypothetical protein</i>
05	04	21	05E16	XAC3636	18C08	A0QR6750A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	05	21	05E17	XAC3718	17C09	A0UV6736C04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	06	21	05E18	XAC1301	18C09	A0QR6752A06	VII.C	<i>catalase/peroxidase</i>
05	07	21	05E19	XAC0942	17C10	A0QR6737D11	II.D.4	<i>molybdopterin biosynthesis protein B</i>
05	08	21	05E20	XAC1418	18C10	A0QR6749H10	III.C.1	<i>ribosome recycling factor</i>
05	09	21	05E21	XAC3935	17C11	A0CE6740F04	VI.C	<i>IS1389 transposase</i>
05	10	21	05E22	XACb0002	18C11	A0UE6759E09	VIII.B	<i>hypothetical protein</i>
05	11	21	05E23	XAC2698	17C12	A0CE6742F03	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO3 subunit</i>
05	12	21	05E24	XAC1196	18C12	A0UE6759H07	I.D.2	<i>LexA repressor</i>
06	01	21	06E13	XAC3762	21C07	A0UT6817B06	VIII.A	<i>conserved hypothetical protein</i>
06	02	21	06E14	XAC2284	22C07	A0AC6827H05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	03	21	06E15	XAC1714	21C08	A0UT6817D06	III.A.1	<i>topoisomerase IV subunit B</i>
06	04	21	06E16	XAC4023	22C08	A0AC6825G01	I.D.1	<i>two-component system, regulatory protein</i>
06	05	21	06E17	XAC2316	21C09	A0UT6817C02	VIII.A	<i>conserved hypothetical protein</i>
06	06	21	06E18	XAC3171	22C09	A0AC6824F11	V.A.4	<i>cation:proton antiporter</i>
06	07	21	06E19	XAC2092	21C10	A0UV6821B04	III.A.4	<i>excinuclease ABC subunit C</i>
06	08	21	06E20	XAC4285	22C10	A0AC6823F05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	21	06E21	XAC1351	21C11	A0UV6821B05	VIII.A	<i>conserved hypothetical protein</i>
06	10	21	06E22	XAC0141	22C11	A0AC6827F08	VIII.A	<i>conserved hypothetical protein</i>
06	11	21	06E23	XAC1492	21C12	A0UV6820C01	VIII.A	<i>conserved hypothetical protein</i>
06	12	21	06E24	XAC2535	22C12	A0AC6825E06	V.A.7	<i>TonB-dependent receptor</i>
07	01	21	07E13	XAC1758	25C07	A0QR9768H06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	02	21	07E14	XAC3918	26C07	A0UV9822F10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	21	07E15	XAC0448	25C08	A0QR9776G05	II.D.17	<i>tryptophan 2,3-dioxygenase</i>
07	04	21	07E16	XAC4214	26C08	A0UT9849C09	II.B.1	<i>GMP synthase</i>
07	05	21	07E17	XAC0928	25C09	A0QR9776G09	III.C.3	<i>extracellular protease</i>
07	06	21	07E18	XAC4304	26C09	A0UT9846G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	07	21	07E19	XAC2302	25C10	A0QR9775D08	VIII.A	<i>conserved hypothetical protein</i>
07	08	21	07E20	XAC3192	26C10	A0UT9845H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	09	21	07E21	XAC4049	25C11	A0UT9800D12	IV.A.2	<i>SapC protein</i>
07	10	21	07E22	XAC0774	26C11	A0CE9842A04	IV.B	<i>penicillin-binding protein 3</i>
07	11	21	07E23	XAC3598	25C12	A0UT9800E08	IV.C	<i>O-antigen biosynthesis protein (truncated)</i>
07	12	21	07E24	XAC2381	26C12	A0UV9870E02	VIII.A	<i>conserved hypothetical protein</i>
08	01	21	08E13	XACb0023	29C07	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	02	21	08E14	no_clone	30C07	no_clone	NC	<i>no_gene</i>
08	03	21	08E15	XAC2611	29C08	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	04	21	08E16	no_clone	30C08	no_clone	NC	<i>no_gene</i>
08	05	21	08E17	XACb0015	29C09	genomic_DNA	VII.A	<i>avirulence protein</i>
08	06	21	08E18	no_clone	30C09	no_clone	NC	<i>no_gene</i>
08	07	21	08E19	XAC3723	29C10	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	08	21	08E20	no_clone	30C10	no_clone	NC	<i>no_gene</i>
08	09	21	08E21	no_clone	29C11	no_clone	NC	<i>no_gene</i>
08	10	21	08E22	no_clone	30C11	no_clone	NC	<i>no_gene</i>
08	11	21	08E23	no_clone	29C12	no_clone	NC	<i>no_gene</i>
08	12	21	08E24	no_clone	30C12	no_clone	NC	<i>no_gene</i>
09	01	21	09C17	Score_card	33B09	Score_card	SC	<i>Score_card</i>

09	02	21	09I09	Score_card	33E05	Score_card	SC	Score_card
09	03	21	09O01	Score_card	33H01	Score_card	SC	Score_card
09	04	21	09C18	Score_card	34B09	Score_card	SC	Score_card
09	05	21	09I10	Score_card	34E05	Score_card	SC	Score_card
09	06	21	09O02	Score_card	34H01	Score_card	SC	Score_card
09	07	21	09D18	Score_card	35B09	Score_card	SC	Score_card
09	08	21	09J10	Score_card	35E05	Score_card	SC	Score_card
09	09	21	09P02	Score_card	35H01	Score_card	SC	Score_card
09	10	21	09D17	Score_card	36B09	Score_card	SC	Score_card
09	11	21	09J09	Score_card	36E05	Score_card	SC	Score_card
09	12	21	09P01	Score_card	36H01	Score_card	SC	Score_card
01	01	22	01F13	XAC1639	04C07	A0AM1452B11	I.A.2	<i>atrazine chlorohydrolase</i>
01	02	22	01F14	XAC4323	03C07	A0AM1403G12	VI.C	<i>ISxac3 transposase</i>
01	03	22	01F15	XAC4336	04C08	A0JJ1440C09	III.A.3	<i>exodeoxyribonuclease V beta chain</i>
01	04	22	01F16	XAC3903	03C08	A0AC1383E07	II.B.2	<i>orotate phosphoribosyl transferase</i>
01	05	22	01F17	XAC0840	04C09	A0RN1587E01	VIII.A	<i>conserved hypothetical protein</i>
01	06	22	01F18	XAC2496	03C09	A0QR1369H04	VIII.A	<i>conserved hypothetical protein</i>
01	07	22	01F19	XAC3116	04C10	A0QR5111B02	II.D.11	<i>PqqC/D protein</i>
01	08	22	01F20	XAC4314	03C10	A0UV1358B01	VI.B	<i>plasmid stability protein</i>
01	09	22	01F21	XAC3849	04C11	A0RN1627D03	VII.C	<i>acriflavin resistance protein</i>
01	10	22	01F22	XAC0685	03C11	A0JJ1311A02	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
01	11	22	01F23	XAC3560	04C12	A0QH1645B04	V.A.7	<i>TonB-dependent receptor</i>
01	12	22	01F24	XAC4194	03C12	A0AM1424F01	VIII.A	<i>conserved hypothetical protein</i>
02	01	22	02F13	XAC1431	08C07	A0UT6211B06	VIII.A	<i>conserved hypothetical protein</i>
02	02	22	02F14	XAC1333	07C07	A0QR5904A09	VIII.A	<i>conserved hypothetical protein</i>
02	03	22	02F15	XAC2871	08C08	A0UV6208G11	III.D.2	<i>cardiolipin synthase</i>
02	04	22	02F16	XAC2646	07C08	A0QR5902H11	VI.A	<i>phage-related capsid completion protein</i>
02	05	22	02F17	XAC3056	08C09	A0UT6211G01	VIII.A	<i>conserved hypothetical protein</i>
02	06	22	02F18	XAC1449	07C09	A0QR5902G10	VIII.A	<i>conserved hypothetical protein</i>
02	07	22	02F19	XAC2190	08C10	A0QH6308F06	VIII.A	<i>conserved hypothetical protein</i>
02	08	22	02F20	XAC2563	07C10	A0QR5902F05	I.A.3	<i>acyl-CoA dehydrogenase</i>
02	09	22	02F21	XAC0578	08C11	A0QH6309D10	VI.C	<i>ISxac3 transposase</i>
02	10	22	02F22	XAC2202	07C11	A0QR5902B02	VII.C	<i>hemolysin secretion protein B</i>
02	11	22	02F23	XAC1047	08C12	A0CE6315A06	VIII.A	<i>conserved hypothetical protein</i>
02	12	22	02F24	XAC1009	07C12	A0QR5904G10	VIII.B	<i>hypothetical protein</i>
03	01	22	03F13	XAC3308	12C07	A0QR6483H04	V.A.7	<i>large-conductance mechanosensitive channel</i>
03	02	22	03F14	XAC3619	11C07	A0AC6449D12	VIII.A	<i>conserved hypothetical protein</i>
03	03	22	03F15	XAC3701	12C08	A0UV6514C09	V.A.4	<i>Na⁺:H⁺ antiporter</i>
03	04	22	03F16	XAC1121	11C08	A0AC6450A07	VIII.A	<i>conserved hypothetical protein</i>
03	05	22	03F17	XAC3932	12C09	A0QR6613A10	VI.C	<i>integrase/recombinase</i>
03	06	22	03F18	XAC0997	11C09	A0AC6450B03	III.B.2	<i>50S ribosomal protein L17</i>
03	07	22	03F19	XAC0230	12C10	A0QR6701H05	III.B.4	<i>tRNA/rRNA methyltransferase</i>
03	08	22	03F20	XAC1689	11C10	A0CE6480F01	VIII.A	<i>conserved hypothetical protein</i>
03	09	22	03F21	XAC0244	12C11	A0QR6702G10	III.B.5	<i>pseudouridylate synthase</i>
03	10	22	03F22	XAC0853	11C11	A0CE6480F04	VIII.B	<i>hypothetical protein</i>
03	11	22	03F23	XAC2522	12C12	A0QR6702G12	VII.D	<i>cellulase</i>
03	12	22	03F24	XAC1761	11C12	A0CE6478F04	VIII.A	<i>conserved hypothetical protein</i>
04	01	22	04F13	XAC2708	16C07	A0UV6733C02	I.C.3	<i>dehydrogenase</i>
04	02	22	04F14	XAC3975	15C07	A0QR6727F12	I.D.1	<i>two-component system, sensor protein</i>
04	03	22	04F15	XAC1572	16C08	A0UV6733C03	VIII.B	<i>hypothetical protein</i>
04	04	22	04F16	XAC0320	15C08	A0QR6728B03	I.D.2	<i>SIR2-like regulatory protein</i>
04	05	22	04F17	XAC4077	16C09	A0UV6733F10	VIII.A	<i>conserved hypothetical protein</i>
04	06	22	04F18	XAC3601	15C09	A0QR6727G10	V.A.7	<i>ABC transporter permease</i>
04	07	22	04F19	XAC0677	16C10	A0UV6733G02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	08	22	04F20	XAC2373	15C10	A0QR6727H10	VII.D	<i>pectate lyase (degenerated)</i>

04	09	22	04F21	XAC2086	16C11	A0UV6733E07	V.A.7	<i>biopolymer transport protein</i>
04	10	22	04F22	XAC3741	15C11	A0QR6730E05	VIII.A	<i>conserved hypothetical protein</i>
04	11	22	04F23	XAC1558	16C12	A0UV6733G09	VIII.A	<i>conserved hypothetical protein</i>
04	12	22	04F24	XAC1481	15C12	A0QR6730E06	I.C.3	<i>dehydrogenase</i>
05	01	22	05F13	XAC3407	20C07	A0QR6804H11	VIII.A	<i>conserved hypothetical protein</i>
05	02	22	05F14	XAC2551	19C07	A0UV6763F01	I.D.2	<i>transcriptional regulator</i>
05	03	22	05F15	XAC2558	20C08	A0QH6810D11	III.A.4	<i>excinuclease ABC subunit C homolog</i>
05	04	22	05F16	XAC3676	19C08	A0UV6763A04	I.C.3	<i>oxidoreductase</i>
05	05	22	05F17	XAC0727	20C09	A0QH6810E01	I.D.2	<i>transcriptional regulator</i>
05	06	22	05F18	XAC2203	19C09	A0UV6764B05	VIII.B	<i>hypothetical protein</i>
05	07	22	05F19	XAC3711	20C10	A0QR6804G05	VIII.A	<i>conserved hypothetical protein</i>
05	08	22	05F20	XAC2435	19C10	A0UV6765G02	VI.B	<i>plasmid-related protein</i>
05	09	22	05F21	XAC1594	20C11	A0QR6804B03	VIII.A	<i>conserved hypothetical protein</i>
05	10	22	05F22	XAC2149	19C11	A0UV6765G03	I.D.1	<i>nodulation protein</i>
05	11	22	05F23	XAC1817	20C12	A0QR6803D12	VIII.B	<i>hypothetical protein</i>
05	12	22	05F24	XAC1176	19C12	A0QR6802E02	IV.C	<i>glycosyl hydrolase</i>
06	01	22	06F13	XAC2974	24C07	A0UV9709G05	I.D.4	<i>nitrogen regulatory IIA protein</i>
06	02	22	06F14	XAC3941	23C07	A0CE6829B07	III.A.4	<i>DNA helicase</i>
06	03	22	06F15	XAC2237	24C08	A0UV6845D10	VIII.A	<i>conserved hypothetical protein</i>
06	04	22	06F16	XAC1498	23C08	A0QH6836E07	VI.A	<i>integrase</i>
06	05	22	06F17	XAC3703	24C09	A0UV9709A04	VIII.A	<i>conserved hypothetical protein</i>
06	06	22	06F18	XAC2438	23C09	A0QH6836C10	VIII.B	<i>hypothetical protein</i>
06	07	22	06F19	XAC1426	24C10	A0UV6845E04	III.C.2	<i>pili assembly chaperone</i>
06	08	22	06F20	XAC3322	23C10	A0QH6836F08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	22	06F21	XAC4199	24C11	A0UV9746F02	I.C.3	<i>polyvinylalcohol dehydrogenase</i>
06	10	22	06F22	XAC0314	23C11	A0QR6840F09	VIII.A	<i>conserved hypothetical protein</i>
06	11	22	06F23	XAC1503	24C12	A0QH9737C01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	12	22	06F24	XAC1399	23C12	A0QR6840G03	IX	<i>sensor protein</i>
07	01	22	07F13	no_clone	28C07	no_clone	NC	<i>no_gene</i>
07	02	22	07F14	XAC0283	27C07	A0UE9886C11	IX	<i>hydrolase</i>
07	03	22	07F15	no_clone	28C08	no_clone	NC	<i>no_gene</i>
07	04	22	07F16	XAC3119	27C08	A0UV9884D08	VIII.A	<i>conserved hypothetical protein</i>
07	05	22	07F17	no_clone	28C09	no_clone	NC	<i>no_gene</i>
07	06	22	07F18	XAC3380	27C09	A0UV9884F02	VIII.B	<i>hypothetical protein</i>
07	07	22	07F19	no_clone	28C10	no_clone	NC	<i>no_gene</i>
07	08	22	07F20	XAC1752	27C10	A0UV9884F11	VIII.A	<i>conserved hypothetical protein</i>
07	09	22	07F21	no_clone	28C11	no_clone	NC	<i>no_gene</i>
07	10	22	07F22	XAC0641	27C11	A0UV9882C07	VII.C	<i>multidrug resistance efflux pump</i>
07	11	22	07F23	no_clone	28C12	no_clone	NC	<i>no_gene</i>
07	12	22	07F24	XAC4156	27C12	A0UV9880G11	I.C.3	<i>FldA protein</i>
08	01	22	08F13	no_clone	32C07	no_clone	NC	<i>no_gene</i>
08	02	22	08F14	no_clone	31C07	no_clone	NC	<i>no_gene</i>
08	03	22	08F15	no_clone	32C08	no_clone	NC	<i>no_gene</i>
08	04	22	08F16	no_clone	31C08	no_clone	NC	<i>no_gene</i>
08	05	22	08F17	no_clone	32C09	no_clone	NC	<i>no_gene</i>
08	06	22	08F18	no_clone	31C09	no_clone	NC	<i>no_gene</i>
08	07	22	08F19	no_clone	32C10	no_clone	NC	<i>no_gene</i>
08	08	22	08F20	no_clone	31C10	no_clone	NC	<i>no_gene</i>
08	09	22	08F21	no_clone	32C11	no_clone	NC	<i>no_gene</i>
08	10	22	08F22	no_clone	31C11	no_clone	NC	<i>no_gene</i>
08	11	22	08F23	no_clone	32C12	no_clone	NC	<i>no_gene</i>
08	12	22	08F24	no_clone	31C12	no_clone	NC	<i>no_gene</i>
09	01	22	09C19	Score_card	33B10	Score_card	SC	<i>Score_card</i>
09	02	22	09I11	Score_card	33E06	Score_card	SC	<i>Score_card</i>
09	03	22	09O03	Score_card	33H02	Score_card	SC	<i>Score_card</i>

09	04	22	09C20	Score_card	34B10	Score_card	SC	Score_card
09	05	22	09I12	Score_card	34E06	Score_card	SC	Score_card
09	06	22	09O04	Score_card	34H02	Score_card	SC	Score_card
09	07	22	09D20	Score_card	35B10	Score_card	SC	Score_card
09	08	22	09J12	Score_card	35E06	Score_card	SC	Score_card
09	09	22	09P04	Score_card	35H02	Score_card	SC	Score_card
09	10	22	09D19	Score_card	36B10	Score_card	SC	Score_card
09	11	22	09J11	Score_card	36E06	Score_card	SC	Score_card
09	12	22	09P03	Score_card	36H02	Score_card	SC	Score_card
01	01	23	01G13	XAC3044	01D07	A0AC0115G04	VIII.A	conserved hypothetical protein
01	02	23	01G14	XAC3390	02D07	A0RN1194B11	II.B.4	inosine-uridine preferring nucleoside hydrolase
01	03	23	01G15	XAC4022	01D08	A0AC0115B07	I.D.1	two-component system, sensor protein
01	04	23	01G16	XAC2468	02D08	A0JJ1203B12	V.A.4	magnesium and cobalt transport protein
01	05	23	01G17	XAC2238	01D09	A0AC0113F11	VI.B	plasmid-related protein
01	06	23	01G18	XAC1204	02D09	A0RN1195F02	III.C.3	alanyl dipeptidyl peptidase
01	07	23	01G19	XAC1240	01D10	A0JJ1004A12	VIII.A	conserved hypothetical protein
01	08	23	01G20	XAC3429	02D10	A0JE1209E08	II.A.1	acetylornithine aminotransferase
01	09	23	01G21	XAC1596	01D11	A0AC0115H07	II.B.4	5'-nucleotidase
01	10	23	01G22	XAC2687	02D11	A0JJ1197C07	III.C.1	protein chain initiation factor IF-2
01	11	23	01G23	XAC4231	01D12	A0JJ0707A01	VII.D	glucan 1,4-beta-glucosidase
01	12	23	01G24	XAC4358	02D12	A0JE1210H03	VIII.A	conserved hypothetical protein
02	01	23	02G13	XAC2470	05D07	A0UV5307F09	V.A.1	putrescine transport protein; permease
02	02	23	02G14	XAC1393	06D07	A0UV5502C11	VIII.A	conserved hypothetical protein
02	03	23	02G15	XAC0300	05D08	A0QR5408C07	II.A.2	serine-pyruvate aminotransferase
02	04	23	02G16	XAC4018	06D08	A0UV5502C12	I.D.2	bifunctional transcriptional repressor of the biotin operon/biotin acetyl-CoA-carboxylase synthetase
02	05	23	02G17	XAC4166	05D09	A0UT5316B08	I.B.9	alkaline phosphatase D
02	06	23	02G18	XAC3029	06D09	A0JE5637E11	I.D.3	histidine kinase/response regulator hybrid protein
02	07	23	02G19	XAC0245	05D10	A0QR5407B10	VIII.A	conserved hypothetical protein
02	08	23	02G20	XAC1007	06D10	A0JE5637E12	VII.C	glutathione S-transferase
02	09	23	02G21	XAC0891	05D11	A0QR5311C09	VIII.B	hypothetical protein
02	10	23	02G22	XAC4319	06D11	A0QR5701E04	VIII.C	Xanthomonas conserved hypothetical protein
02	11	23	02G23	XAC4160	05D12	A0QR5322D03	VII.C	cation efflux system protein
02	12	23	02G24	XAC0835	06D12	A0QR5701E06	I.D.1	two-component system, sensor protein
03	01	23	03G13	XAC1841	09D07	A0QR6330F10	V.A.1	cationic amino acid transporter
03	02	23	03G14	XAC1434	10D07	A0AC6406B12	VIII.A	conserved hypothetical protein
03	03	23	03G15	XAC4198	09D08	A0CE6344H02	VIII.B	hypothetical protein
03	04	23	03G16	XAC2638	10D08	A0AC6406G08	VIII.B	hypothetical protein
03	05	23	03G17	XAC0901	09D09	A0QR6376B01	VIII.A	conserved hypothetical protein
03	06	23	03G18	XAC1900	10D09	A0AC6407C10	V.C	chemotaxis protein
03	07	23	03G19	XAC1870	09D10	A0UV6350G04	VIII.B	hypothetical protein
03	08	23	03G20	XAC0990	10D10	A0QH6383E03	III.B.2	50S ribosomal protein L30
03	09	23	03G21	XAC1065	09D11	A0AC6356G07	VIII.B	hypothetical protein
03	10	23	03G22	XAC2581	10D11	A0QH6411H07	VII.E	GumF protein
03	11	23	03G23	XAC3374	09D12	A0UV6351E09	VIII.A	conserved hypothetical protein
03	12	23	03G24	XAC2813	10D12	A0AC6440G05	III.B.5	ATP-dependent RNA helicase
04	01	23	04G13	XAC3709	13D07	A0UV6708G05	I.D.2	tryptophan repressor binding protein
04	02	23	04G14	XAC1152	14D07	A0UV6711F01	III.C.2	curved DNA binding protein
04	03	23	04G15	XAC0029	13D08	A0UV6709A09	VII.D	cellulase
04	04	23	04G16	XAC2693	14D08	A0QR6725B08	I.C.1	NADH-ubiquinone oxidoreductase, NQO12 subunit
04	05	23	04G17	XAC3681	13D09	A0QR6705F08	I.A.2	L-sorbose dehydrogenase
04	06	23	04G18	XAC2835	14D09	A0UV6712B07	I.C.3	oxidoreductase
04	07	23	04G19	XAC3541	13D10	A0UV6709B05	VII.H	general secretion pathway protein H
04	08	23	04G20	XAC1577	14D10	A0QR6723A07	V.A.2	ABC transporter phosphate binding
04	09	23	04G21	XAC0208	13D11	A0UV6709D04	I.D.1	two-component system, regulatory protein

04	10	23	04G22	XAC1854	14D11	A0UE6718F03	V.A.4	<i>ferrous iron transport protein</i>
04	11	23	04G23	XAC1157	13D12	A0UV6709D05	VIII.A	<i>conserved hypothetical protein</i>
04	12	23	04G24	XAC1853	14D12	A0UE6718F04	I.A.3	<i>enoyl-CoA hydratase</i>
05	01	23	05G13	XAC2047	17D07	A0UV6738H03	IX	<i>PHA synthase subunit</i>
05	02	23	05G14	XAC0585	18D07	A0QR6748F06	VIII.A	<i>conserved hypothetical protein</i>
05	03	23	05G15	XAC0145	17D08	A0UV6736H12	VIII.A	<i>conserved hypothetical protein</i>
05	04	23	05G16	XAC4190	18D08	A0QR6752D12	V.A.3	<i>fucose permease</i>
05	05	23	05G17	XAC4139	17D09	A0UV6736E01	VIII.B	<i>hypothetical protein</i>
05	06	23	05G18	XAC2487	18D09	A0QR6755B11	I.C.2	<i>formate dehydrogenase b chain</i>
05	07	23	05G19	XAC3613	17D10	A0UV6738F07	V.A.7	<i>TonB-dependent receptor</i>
05	08	23	05G20	XAC1144	18D10	A0QR6752C06	III.B.5	<i>inosine-uridine preferring nucleoside hydrolase</i>
05	09	23	05G21	XAC3632	17D11	A0CE6740H08	I.B.10	<i>lactoylglutathione lyase</i>
05	10	23	05G22	XAC2774	18D11	A0UV6761A09	V.A.7	<i>TonB like protein</i>
05	11	23	05G23	XAC3789	17D12	A0CE6743B01	VIII.A	<i>conserved hypothetical protein</i>
05	12	23	05G24	XAC0182	18D12	A0UV6762A12	V.A.7	<i>ABC transporter ATP-binding protein</i>
06	01	23	06G13	XAC2252	21D07	A0UT6817D05	VIII.B	<i>hypothetical protein</i>
06	02	23	06G14	XAC0969	22D07	A0CE6828D06	III.C.1	<i>elongation factor G</i>
06	03	23	06G15	XAC2912	21D08	A0UT6817F07	VIII.A	<i>conserved hypothetical protein</i>
06	04	23	06G16	XAC0816	22D08	A0AC6827B04	VIII.A	<i>conserved hypothetical protein</i>
06	05	23	06G17	XAC0343	21D09	A0UT6817D07	VI.C	<i>ISxac3 transposase</i>
06	06	23	06G18	XAC1122	22D09	A0AC6825B08	III.B.2	<i>50S ribosomal protein L32</i>
06	07	23	06G19	XACb0056	21D10	A0UV6821D10	VI.B	<i>replication protein A</i>
06	08	23	06G20	XAC0054	22D10	A0AC6824C06	I.B.11	<i>UDP-glucose epimerase (degenerated)</i>
06	09	23	06G21	XAC1613	21D11	A0UV6821E01	VIII.B	<i>hypothetical protein</i>
06	10	23	06G22	XAC3599	22D11	A0CE6828B01	VIII.B	<i>hypothetical protein</i>
06	11	23	06G23	XAC1601	21D12	A0UV6820E11	II.D.17	<i>kynureninase</i>
06	12	23	06G24	XAC4370	22D12	A0AC6827A04	VII.C	<i>thiophene and furan oxidation protein</i>
07	01	23	07G13	XAC1955	25D07	A0UV9750F07	V.C	<i>flagellar protein</i>
07	02	23	07G14	XAC3347	26D07	A0UE9830F02	I.C.4	<i>phosphoglycerate kinase</i>
07	03	23	07G15	XAC1806	25D08	A0UV9750G03	VIII.A	<i>conserved hypothetical protein</i>
07	04	23	07G16	XAC2515	26D08	A0UE9858E07	I.D.2	<i>transcriptional regulator asnC/lrp family</i>
07	05	23	07G17	XAC1557	25D09	A0UV9750G10	I.A.2	<i>fructokinase</i>
07	06	23	07G18	XAC2807	26D09	A0UE9858E12	III.A.5	<i>S-adenosylmethionine:2-demethylmenaquinone methyltransferase</i>
07	07	23	07G19	XAC1906	25D10	A0UT9800D01	V.C	<i>chemotaxis protein</i>
07	08	23	07G20	XAC1660	26D10	A0UT9846G09	VI.C	<i>ISxac3 transposase</i>
07	09	23	07G21	XAC3042	25D11	A0QR9753D04	III.C.1	<i>peptide chain release factor 3</i>
07	10	23	07G22	XAC3512	26D11	A0CE9842H05	VII.C	<i>arsenate reductase</i>
07	11	23	07G23	XAC2152	25D12	A0UV9804D10	VIII.A	<i>conserved hypothetical protein</i>
07	12	23	07G24	XAC3603	26D12	A0CE9839F11	II.A.3	<i>cystathionine beta-synthase</i>
08	01	23	08G13	XACb0021	29D07	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	02	23	08G14	no_clone	30D07	no_clone	NC	<i>no_gene</i>
08	03	23	08G15	XAC2620	29D08	genomic_DNA	VII.H	<i>VirB9 protein</i>
08	04	23	08G16	no_clone	30D08	no_clone	NC	<i>no_gene</i>
08	05	23	08G17	XACb0019	29D09	genomic_DNA	VI.B	<i>partition protein A</i>
08	06	23	08G18	no_clone	30D09	no_clone	NC	<i>no_gene</i>
08	07	23	08G19	XAC1269	29D10	genomic_DNA	I.D.2	<i>positive regulator of sigma-B</i>
08	08	23	08G20	no_clone	30D10	no_clone	NC	<i>no_gene</i>
08	09	23	08G21	no_clone	29D11	no_clone	NC	<i>no_gene</i>
08	10	23	08G22	no_clone	30D11	no_clone	NC	<i>no_gene</i>
08	11	23	08G23	no_clone	29D12	no_clone	NC	<i>no_gene</i>
08	12	23	08G24	no_clone	30D12	no_clone	NC	<i>no_gene</i>
09	01	23	09C21	Score_card	33B11	Score_card	SC	<i>Score_card</i>
09	02	23	09I13	Score_card	33E07	Score_card	SC	<i>Score_card</i>
09	03	23	09O05	Score_card	33H03	Score_card	SC	<i>Score_card</i>

09	04	23	09C22	Score_card	34B11	Score_card	SC	Score_card
09	05	23	09I14	Score_card	34E07	Score_card	SC	Score_card
09	06	23	09O06	Score_card	34H03	Score_card	SC	Score_card
09	07	23	09D22	Score_card	35B11	Score_card	SC	Score_card
09	08	23	09J14	Score_card	35E07	Score_card	SC	Score_card
09	09	23	09P06	Score_card	35H03	Score_card	SC	Score_card
09	10	23	09D21	Score_card	36B11	Score_card	SC	Score_card
09	11	23	09J13	Score_card	36E07	Score_card	SC	Score_card
09	12	23	09P05	Score_card	36H03	Score_card	SC	Score_card
01	01	24	01H13	XAC0663	04D07	A0RN1497B05	III.D.3	<i>rare lipoprotein A</i>
01	02	24	01H14	XAC0102	03D07	A0RN1406D06	VIII.A	<i>conserved hypothetical protein</i>
01	03	24	01H15	XAC3177	04D08	A0AM1452F12	VIII.A	<i>conserved hypothetical protein</i>
01	04	24	01H16	XAC3404	03D08	A0JJ1392B10	VIII.B	<i>hypothetical protein</i>
01	05	24	01H17	XAC0987	04D09	A0AM1595G04	III.B.2	<i>50S ribosomal protein L6</i>
01	06	24	01H18	XAC1725	03D09	A0JE1377G06	VII.G	<i>survival protein</i>
01	07	24	01H19	XAC0460	04D10	A0JJ1591D03	VII.G	<i>PhaD protein</i>
01	08	24	01H20	XAC2401	03D10	A0AM1365A08	II.E	<i>acetoacetyl-CoA reductase</i>
01	09	24	01H21	XAC0444	04D11	A0RN1631E12	VIII.A	<i>conserved hypothetical protein</i>
01	10	24	01H22	XAC3495	03D11	A0AM1318C01	VIII.A	<i>conserved hypothetical protein</i>
01	11	24	01H23	XAC1504	04D12	A0QR5101G05	VI.C	<i>ISxcd1 transposase</i>
01	12	24	01H24	XAC2505	03D12	A0JJ1441A12	VIII.B	<i>hypothetical protein</i>
02	01	24	02H13	XAC0274	08D07	A0QH6215E06	III.A.4	<i>nuclease</i>
02	02	24	02H14	XAC1830	07D07	A0QR5904F10	II.A.5	<i>histidinol-phosphate aminotransferase</i>
02	03	24	02H15	XAC1256	08D08	A0UV6210A11	VII.C	<i>penicillin tolerance protein</i>
02	04	24	02H16	XAC2017	07D08	A0QR5903G03	IV.D	<i>fimbrial biogenesis protein</i>
02	05	24	02H17	XAC3438	08D09	A0QR6302H01	I.C.4	<i>6-phosphofructokinase</i>
02	06	24	02H18	XAC3198	07D09	A0QR5902H12	V.A.7	<i>nitrate transport protein</i>
02	07	24	02H19	XAC4245	08D10	A0QH6309D04	VIII.B	<i>hypothetical protein</i>
02	08	24	02H20	XAC0188	07D10	A0QR5902G11	I.D.2	<i>transcriptional regulator</i>
02	09	24	02H21	XAC2269	08D11	A0EC6316F07	VIII.B	<i>hypothetical protein</i>
02	10	24	02H22	XAC4159	07D11	A0QR5902D09	III.B.2	<i>50S ribosomal protein L28</i>
02	11	24	02H23	XAC0263	08D12	A0EC6321D09	II.E	<i>biotin carboxylase</i>
02	12	24	02H24	XAC1818	07D12	A0QR5704F05	VII.F	<i>hemagglutinin/hemolysin-related protein</i>
03	01	24	03H13	XAC3324	12D07	A0UV6485D07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	02	24	03H14	XAC2175	11D07	A0UV6456C10	VI.C	<i>ISxac3 transposase</i>
03	03	24	03H15	XAC2684	12D08	A0UV6514H12	III.B.2	<i>30S ribosomal protein S15</i>
03	04	24	03H16	XAC4031	11D08	A0UV6456E02	III.A.4	<i>ATP-dependent helicase</i>
03	05	24	03H17	XAC1195	12D09	A0QR6614E08	VIII.B	<i>hypothetical protein</i>
03	06	24	03H18	XAC2385	11D09	A0QR6453G04	I.D.2	<i>extragenic suppressor protein SuhB</i>
03	07	24	03H19	XAC1165	12D10	A0QR6702G06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	08	24	03H20	XAC2312	11D10	A0UV6456F01	VIII.A	<i>conserved hypothetical protein</i>
03	09	24	03H21	XAC1605	12D11	A0QR6704B07	VIII.A	<i>conserved hypothetical protein</i>
03	10	24	03H22	XAC0820	11D11	A0AC6450H03	VIII.B	<i>hypothetical protein</i>
03	11	24	03H23	XAC3231	12D12	A0UE6610B09	VIII.B	<i>hypothetical protein</i>
03	12	24	03H24	XAC0093	11D12	A0CE6480F05	VI.C	<i>ISxac1 transposase</i>
04	01	24	04H13	XAC1745	16D07	A0UV6733D08	VIII.B	<i>hypothetical protein</i>
04	02	24	04H14	XAC2862	15D07	A0QR6728A12	VIII.B	<i>hypothetical protein</i>
04	03	24	04H15	XAC1597	16D08	A0UV6733D09	VIII.A	<i>conserved hypothetical protein</i>
04	04	24	04H16	XAC2313	15D08	A0QR6728F03	I.D.2	<i>transcriptional regulator lacI family</i>
04	05	24	04H17	XAC1008	16D09	A0UV6733H05	VIII.A	<i>conserved hypothetical protein</i>
04	06	24	04H18	XAC0167	15D09	A0QR6728D09	VIII.B	<i>hypothetical protein</i>
04	07	24	04H19	XAC4054	16D10	A0UV6733H12	V.A.4	<i>sodium ABC transporter ATP-binding protein</i>
04	08	24	04H20	XAC1352	15D10	A0QR6728B07	VIII.A	<i>conserved hypothetical protein</i>
04	09	24	04H21	XAC0389	16D11	A0UV6733G07	VII.F	<i>competence protein F</i>
04	10	24	04H22	XAC1193	15D11	A0QR6730H08	VIII.A	<i>conserved hypothetical protein</i>

04	11	24	04H23	XAC0420	16D12	A0UV6734C11	VIII.A	<i>conserved hypothetical protein</i>
04	12	24	04H24	XAC2455	15D12	A0QR6730G02	I.C.1	<i>ubiquinol cytochrome C oxidoreductase, cytochrome C1 subunit</i>
05	01	24	05H13	XAC0476	20D07	A0QH6810D05	II.A.4	<i>anthranilate synthase component I</i>
05	02	24	05H14	XAC0211	19D07	A0UV6763H05	I.B.10	<i>lactoylglutathione lyase</i>
05	03	24	05H15	XAC0625	20D08	A0QH6810G03	VIII.A	<i>conserved hypothetical protein</i>
05	04	24	05H16	XAC3329	19D08	A0UV6764B02	I.B.12	<i>ATP sulfurylase, small subunit</i>
05	05	24	05H17	XAC3995	20D09	A0QH6810G09	VII.C	<i>acriflavin resistance protein</i>
05	06	24	05H18	XAC0273	19D09	A0UV6764D05	I.C.3	<i>cytochrome C5</i>
05	07	24	05H19	XAC1800	20D10	A0QR6805C04	III.D.2	<i>phosphatidylglycerophosphatase B-related protein</i>
05	08	24	05H20	XAC2539	19D10	A0UV6762E09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	09	24	05H21	XAC3490	20D11	A0QR6804D02	III.D.1	<i>amylsucrase or alpha amylase</i>
05	10	24	05H22	XAC1551	19D11	A0UV6765G04	I.B.11	<i>UDP-glucose dehydrogenase</i>
05	11	24	05H23	XAC3505	20D12	A0QR6804D04	I.A.1	<i>rhamnogalacturonase B</i>
05	12	24	05H24	XAC3287	19D12	A0UV6766B05	VIII.B	<i>hypothetical protein</i>
06	01	24	06H13	XAC4150	24D07	A0UV6845F05	IV.C	<i>nodulation protein</i>
06	02	24	06H14	XAC1452	23D07	A0CE6829F10	VIII.B	<i>hypothetical protein</i>
06	03	24	06H15	XAC1311	24D08	A0UV6845F08	I.D.2	<i>transcriptional regulator</i>
06	04	24	06H16	XAC4085	23D08	A0QR6839A11	VIII.A	<i>conserved hypothetical protein</i>
06	05	24	06H17	XAC2055	24D09	A0UV9711G01	I.D.1	<i>two-component system, regulatory protein</i>
06	06	24	06H18	XAC3217	23D09	A0QR6839B05	III.B.3	<i>ribosomal large subunit pseudouridine synthase D</i>
06	07	24	06H19	XAC1062	24D10	A0UV6846A04	VIII.B	<i>hypothetical protein</i>
06	08	24	06H20	XAC2556	23D10	A0QR6839B09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	24	06H21	XAC4091	24D11	A0CE9723A11	VIII.B	<i>hypothetical protein</i>
06	10	24	06H22	XAC0748	23D11	A0UV6843A02	II.D.9	<i>riboflavin synthase alpha chain</i>
06	11	24	06H23	XAC4279	24D12	A0QR9739A10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	12	24	06H24	XAC3117	23D12	A0UV6843C10	II.D.11	<i>PqqE protein</i>
07	01	24	07H13	no_clone	28D07	no_clone	NC	<i>no_gene</i>
07	02	24	07H14	XAC4263	27D07	A0UV9872G02	VIII.B	<i>hypothetical protein</i>
07	03	24	07H15	no_clone	28D08	no_clone	NC	<i>no_gene</i>
07	04	24	07H16	XAC4316	27D08	A0UE9889C04	VIII.A	<i>conserved hypothetical protein</i>
07	05	24	07H17	no_clone	28D09	no_clone	NC	<i>no_gene</i>
07	06	24	07H18	XAC1315	27D09	A0UE9886D10	I.A.3	<i>enoyl-CoA hydratase</i>
07	07	24	07H19	no_clone	28D10	no_clone	NC	<i>no_gene</i>
07	08	24	07H20	XAC2796	27D10	A0UE9886G07	VIII.A	<i>conserved hypothetical protein</i>
07	09	24	07H21	no_clone	28D11	no_clone	NC	<i>no_gene</i>
07	10	24	07H22	XAC0252	27D11	A0UV9883B05	VIII.A	<i>conserved hypothetical protein</i>
07	11	24	07H23	no_clone	28D12	no_clone	NC	<i>no_gene</i>
07	12	24	07H24	XAC2311	27D12	A0UE9890D05	VIII.A	<i>conserved hypothetical protein</i>
08	01	24	08H13	no_clone	32D07	no_clone	NC	<i>no_gene</i>
08	02	24	08H14	no_clone	31D07	no_clone	NC	<i>no_gene</i>
08	03	24	08H15	no_clone	32D08	no_clone	NC	<i>no_gene</i>
08	04	24	08H16	no_clone	31D08	no_clone	NC	<i>no_gene</i>
08	05	24	08H17	no_clone	32D09	no_clone	NC	<i>no_gene</i>
08	06	24	08H18	no_clone	31D09	no_clone	NC	<i>no_gene</i>
08	07	24	08H19	no_clone	32D10	no_clone	NC	<i>no_gene</i>
08	08	24	08H20	no_clone	31D10	no_clone	NC	<i>no_gene</i>
08	09	24	08H21	no_clone	32D11	no_clone	NC	<i>no_gene</i>
08	10	24	08H22	no_clone	31D11	no_clone	NC	<i>no_gene</i>
08	11	24	08H23	no_clone	32D12	no_clone	NC	<i>no_gene</i>
08	12	24	08H24	no_clone	31D12	no_clone	NC	<i>no_gene</i>
09	01	24	09C23	Score_card	33B12	Score_card	SC	<i>Score_card</i>
09	02	24	09I15	Score_card	33E08	Score_card	SC	<i>Score_card</i>
09	03	24	09O07	Score_card	33H04	Score_card	SC	<i>Score_card</i>
09	04	24	09C24	Score_card	34B12	Score_card	SC	<i>Score_card</i>

09	05	24	09I16	Score_card	34E08	Score_card	SC	Score_card
09	06	24	09O08	Score_card	34H04	Score_card	SC	Score_card
09	07	24	09D24	Score_card	35B12	Score_card	SC	Score_card
09	08	24	09J16	Score_card	35E08	Score_card	SC	Score_card
09	09	24	09P08	Score_card	35H04	Score_card	SC	Score_card
09	10	24	09D23	Score_card	36B12	Score_card	SC	Score_card
09	11	24	09J15	Score_card	36E08	Score_card	SC	Score_card
09	12	24	09P07	Score_card	36H04	Score_card	SC	Score_card
01	01	25	01I13	XAC3408	01E07	A0JJ0101F10	VIII.A	conserved hypothetical protein
01	02	25	01I14	XAC1243	02E07	A0JE1208E03	III.D.2	acyl-CoA thioesterase II
01	03	25	01I15	XAC0177	01E08	A0AC0115D10	VIII.A	conserved hypothetical protein
01	04	25	01I16	XAC2361	02E08	A0JE1208H07	III.C.3	peptidase
01	05	25	01I17	XAC0899	01E09	A0AC0115A07	VIII.A	conserved hypothetical protein
01	06	25	01I18	XAC2481	02E09	A0JJ1205A02	I.C.3	oxidoreductase
01	07	25	01I19	XAC0401	01E10	A0JJ1016G09	VII.B	HrcS protein
01	08	25	01I20	XAC1386	02E10	A0JE1211H05	III.B.4	methionyl-tRNA synthetase
01	09	25	01I21	XAC1726	01E11	A0JJ0209E09	III.C.1	L-isoaspartate protein carboxymethyltransferase type II
01	10	25	01I22	XAC3458	02E11	A0JE1210D03	II.A.2	3-isopropylmalate dehydratase large subunit
01	11	25	01I23	XAC0155	01E12	A0JJ1003E03	III.D.1	trehalose synthase
01	12	25	01I24	XAC2853	02E12	A0RN1227A01	III.C.3	cysteine protease
02	01	25	02I13	XAC2375	05E07	A0CE5309E12	VIII.A	conserved hypothetical protein
02	02	25	02I14	XAC2625	06E07	A0UV5503H06	III.A.4	excinuclease ABC subunit B
02	03	25	02I15	XAC2242	05E08	A0CE5309G05	VI.B	plasmid-related protein
02	04	25	02I16	XAC3089	06E08	A0UE5505E07	VIII.A	conserved hypothetical protein
02	05	25	02I17	XAC4028	05E09	A0CE5319F01	IV.A.1	ankyrin like protein
02	06	25	02I18	XAC2443	06E09	A0JE5637H06	VIII.A	conserved hypothetical protein
02	07	25	02I19	XAC3051	05E10	A0QR5407F02	VIII.A	conserved hypothetical protein
02	08	25	02I20	XAC3689	06E10	A0JE5637H08	I.D.2	leucine responsive regulatory protein
02	09	25	02I21	XAC0330	05E11	A0QR5317B07	I.D.2	conditioned medium factor
02	10	25	02I22	XACb0060	06E11	A0QR5702B12	VII.H	virulence plasmid protein
02	11	25	02I23	XAC1787	05E12	A0QR5324A02	II.D.5	aspartate 1-decarboxylase precursor
02	12	25	02I24	XAC2736	06E12	A0QR5702A03	I.A.2	carboxymethylenebutenolidase
03	01	25	03I13	XAC3453	09E07	A0CE6344H01	II.A.2	acetolactate synthase isozyme II, small subunit
03	02	25	03I14	XAC0538	10E07	A0AC6408F04	VIII.B	hypothetical protein
03	03	25	03I15	XAC3770	09E08	A0UV6348F07	VIII.A	conserved hypothetical protein
03	04	25	03I16	XAC0201	10E08	A0AC6408F07	I.C.2	alcohol dehydrogenase
03	05	25	03I17	XAC4275	09E09	A0UE6378D11	II.D.17	tryptophan halogenase
03	06	25	03I18	XAC0173	10E09	A0UT6386G04	VIII.C	Xanthomonas conserved hypothetical protein
03	07	25	03I19	XAC1790	09E10	A0UV6351D08	VIII.A	conserved hypothetical protein
03	08	25	03I20	XAC3957	10E10	A0UT6387C05	VIII.C	Xanthomonas conserved hypothetical protein
03	09	25	03I21	XAC1949	09E11	A0AC6359E01	V.C	flagellar protein
03	10	25	03I22	XAC3511	10E11	A0QR6417F12	VIII.B	hypothetical protein
03	11	25	03I23	XAC0699	09E12	A0AC6357C10	VII.H	type II secretion system protein H
03	12	25	03I24	XAC0631	10E12	A0EC6426B09	III.C.3	protease II (oligopeptidase B)
04	01	25	04I13	XAC2617	13E07	A0UV6709A03	VII.H	VirB1 protein
04	02	25	04I14	XAC2327	14E07	A0UV6711G08	I.C.3	C-type cytochrome biogenesis protein
04	03	25	04I15	XAC3122	13E08	A0QR6705B12	III.A.1	ATP-dependent RNA helicase
04	04	25	04I16	XAC4220	14E08	A0UV6712B04	II.D.12	ferrochelataase
04	05	25	04I17	XAC1124	13E09	A0QR6706H12	VIII.A	conserved hypothetical protein
04	06	25	04I18	XAC3993	14E09	A0UV6712D02	I.D.1	two-component system, regulatory protein
04	07	25	04I19	XAC3825	13E10	A0UV6709G02	III.A.4	uracil-DNA glycosylase
04	08	25	04I20	XAC1915	14E10	A0QR6725C09	VIII.C	Xanthomonas conserved hypothetical protein
04	09	25	04I21	XAC3998	13E11	A0UV6709G03	VIII.A	conserved hypothetical protein
04	10	25	04I22	XAC2372	14E11	A0UE6718H01	VI.C	IS1479 transposase
04	11	25	04I23	XAC1934	13E12	A0UV6709G07	V.B	flagellar biosynthesis switch protein

04	12	25	04I24	XAC3531	14E12	A0UE6718H06	V.A.5	<i>PnuC protein</i>
05	01	25	05I13	XAC2045	17E07	A0UV6739B01	I.C.3	<i>D-beta-hydroxybutyrate dehydrogenase</i>
05	02	25	05I14	XAC2023	18E07	A0QR6749B11	II.D.4	<i>molybdopterin biosynthesis protein</i>
05	03	25	05I15	XAC2108	17E08	A0QR6737C11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	04	25	05I16	XAC0567	18E08	A0QR6749E10	V.A.3	<i>dicarboxylate carrier protein</i>
05	05	25	05I17	XAC3653	17E09	A0UV6736G02	I.C.8	<i>ATP synthase, B chain</i>
05	06	25	05I18	XAC1003	18E09	A0QR6749G10	VIII.A	<i>conserved hypothetical protein</i>
05	07	25	05I19	XAC2915	17E10	A0UV6739A04	V.D	<i>osmotically inducible protein</i>
05	08	25	05I20	XAC0642	18E10	A0QR6752G05	V.A.7	<i>MFS transporter</i>
05	09	25	05I21	XAC1160	17E11	A0CE6742E07	I.C.3	<i>oxidoreductase</i>
05	10	25	05I22	XAC0679	18E11	A0UV6761C11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	11	25	05I23	XAC3874	17E12	A0CE6743C04	VIII.B	<i>hypothetical protein</i>
05	12	25	05I24	XAC3946	18E12	A0UE6757F04	VIII.B	<i>hypothetical protein</i>
06	01	25	06I13	XAC0165	21E07	A0UT6817H07	I.A.1	<i>xylosidase/arabinosidase</i>
06	02	25	06I14	XAC3292	22E07	A0AC6823A05	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
06	03	25	06I15	XAC0536	21E08	A0UV6818E05	VII.H	<i>virulence regulating protein</i>
06	04	25	06I16	XAC1360	22E08	A0AC6827F03	VIII.A	<i>conserved hypothetical protein</i>
06	05	25	06I17	XAC1310	21E09	A0UT6817F08	V.A.7	<i>TonB-dependent receptor</i>
06	06	25	06I18	XAC0368	22E09	A0AC6825E03	I.A.2	<i>protocatechuate 3,4-dioxygenase alpha chain</i>
06	07	25	06I19	XAC3651	21E10	A0UV6818F05	I.C.8	<i>ATP synthase, alpha chain</i>
06	08	25	06I20	XAC2840	22E10	A0AC6824D11	VIII.A	<i>conserved hypothetical protein</i>
06	09	25	06I21	XAC2875	21E11	A0UV6821F08	III.A.4	<i>endonuclease V</i>
06	10	25	06I22	XAC2970	22E11	A0CE6828E06	VIII.A	<i>conserved hypothetical protein</i>
06	11	25	06I23	XAC3363	21E12	A0UV6820H11	I.D.2	<i>transcriptional regulator blal family</i>
06	12	25	06I24	XAC4075	22E12	A0AC6827C10	II.B.3	<i>ribonucleoside-diphosphate reductase alpha chain</i>
07	01	25	07I13	XAC2319	25E07	A0QR9756F03	VIII.A	<i>conserved hypothetical protein</i>
07	02	25	07I14	XAC3352	26E07	A0CE9838C10	I.C.4	<i>glyceraldehyde-3-phosphate dehydrogenase</i>
07	03	25	07I15	XAC0792	25E08	A0QR9753C01	VIII.A	<i>conserved hypothetical protein</i>
07	04	25	07I16	XAC3067	26E08	A0QR9864F01	III.A.4	<i>ADP compounds hydrolase</i>
07	05	25	07I17	XAC1258	25E09	A0QR9756A02	I.C.1	<i>cytochrome O ubiquinol oxidase, subunit II</i>
07	06	25	07I18	XAC1184	26E09	A0QR9860B10	VIII.A	<i>conserved hypothetical protein</i>
07	07	25	07I19	XAC1710	25E10	A0UV9750H08	VIII.A	<i>conserved hypothetical protein</i>
07	08	25	07I20	XAC1020	26E10	A0UT9849H09	V.A.2	<i>sulfate ABC transporter ATP-binding protein</i>
07	09	25	07I21	XAC2568	25E11	A0QR9756B10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	10	25	07I22	XAC0333	26E11	A0UT9846A01	I.D.2	<i>transcriptional regulator metE/metH family</i>
07	11	25	07I23	XAC2559	25E12	A0QR9814G07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	12	25	07I24	XAC3713	26E12	A0CE9842A06	III.C.3	<i>peptidase</i>
08	01	25	08I13	XACb0020	29E07	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	02	25	08I14	no_clone	30E07	no_clone	NC	<i>no_gene</i>
08	03	25	08I15	XAC2606	29E08	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	04	25	08I16	no_clone	30E08	no_clone	NC	<i>no_gene</i>
08	05	25	08I17	XAC1089	29E09	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	06	25	08I18	no_clone	30E09	no_clone	NC	<i>no_gene</i>
08	07	25	08I19	XAC0922	29E10	genomic_DNA	I.D.4	<i>ECF sigma factor</i>
08	08	25	08I20	no_clone	30E10	no_clone	NC	<i>no_gene</i>
08	09	25	08I21	no_clone	29E11	no_clone	NC	<i>no_gene</i>
08	10	25	08I22	no_clone	30E11	no_clone	NC	<i>no_gene</i>
08	11	25	08I23	no_clone	29E12	no_clone	NC	<i>no_gene</i>
08	12	25	08I24	no_clone	30E12	no_clone	NC	<i>no_gene</i>
09	01	25	09E01	Score_card	33C01	Score_card	SC	<i>Score_card</i>
09	02	25	09I17	Score_card	33E09	Score_card	SC	<i>Score_card</i>
09	03	25	09O09	Score_card	33H05	Score_card	SC	<i>Score_card</i>
09	04	25	09E02	Score_card	34C01	Score_card	SC	<i>Score_card</i>
09	05	25	09I18	Score_card	34E09	Score_card	SC	<i>Score_card</i>
09	06	25	09O10	Score_card	34H05	Score_card	SC	<i>Score_card</i>

09	07	25	09F02	Score_card	35C01	Score_card	SC	Score_card
09	08	25	09J18	Score_card	35E09	Score_card	SC	Score_card
09	09	25	09P10	Score_card	35H05	Score_card	SC	Score_card
09	10	25	09F01	Score_card	36C01	Score_card	SC	Score_card
09	11	25	09J17	Score_card	36E09	Score_card	SC	Score_card
09	12	25	09P09	Score_card	36H05	Score_card	SC	Score_card
01	01	26	01J13	XAC0974	04E07	A0UV1503C09	III.B.2	50S ribosomal protein L23
01	02	26	01J14	XAC3561	03E07	A0AM1309B03	IV.B	soluble lytic murein transglycosylase
01	03	26	01J15	XAC1945	04E08	A0RN1498A05	V.C	flagellar protein
01	04	26	01J16	XAC2911	03E08	A0AM1404C04	II.A.2	bifunctional diaminopimelate decarboxylase/aspartate
01	05	26	01J17	XAC1010	04E09	A0AM1616D10	I.A.2	2,4-dienoyl-CoA reductase
01	06	26	01J18	XAC2820	03E09	A0AC1384C06	VIII.C	Xanthomonas conserved hypothetical protein
01	07	26	01J19	XAC3670	04E10	A0AM1617C01	VIII.A	conserved hypothetical protein
01	08	26	01J20	XAC3547	03E10	A0QR1369H12	III.C.3	serine protease
01	09	26	01J21	XAC2903	04E11	A0AR1644F08	VIII.A	conserved hypothetical protein
01	10	26	01J22	XAC1289	03E11	A0JE1333C11	V.A.6	signal recognition particle protein
01	11	26	01J23	XAC0757	04E12	A0AM1592A02	V.A.7	potassium-transporting ATPase, B chain
01	12	26	01J24	XAC0544	03E12	A0RN1457A07	VIII.A	conserved hypothetical protein
02	01	26	02J13	XAC0857	08E07	A0QR6303E10	V.A.4	ABC transporter permease
02	02	26	02J14	XAC4373	07E07	A0QR5704E10	III.B.4	ribonuclease P, protein component
02	03	26	02J15	XAC1454	08E08	A0QH6306H08	VIII.C	Xanthomonas conserved hypothetical protein
02	04	26	02J16	XAC1580	07E08	A0QR5904E06	I.B.10	carbonic anhydrase
02	05	26	02J17	XAC3360	08E09	A0QR6303H12	V.A.2	ATP-binding component of molybdate transport
02	06	26	02J18	XAC0375	07E09	A0QR5903C07	III.D.2	lipase
02	07	26	02J19	XAC0861	08E10	A0QH6313D01	II.B.4	diadenosine tetraphosphatase
02	08	26	02J20	XAC0033	07E10	A0QR5903A03	II.A.1	glutamate synthase, alpha subunit
02	09	26	02J21	XAC2902	08E11	A0UV6209D03	VIII.A	conserved hypothetical protein
02	10	26	02J22	XAC1703	07E11	A0QR5902F06	VIII.C	Xanthomonas conserved hypothetical protein
02	11	26	02J23	XAC2439	08E12	A0UV6209E03	VIII.A	conserved hypothetical protein
02	12	26	02J24	XAC0392	07E12	A0QR5704G08	VIII.B	hypothetical protein
03	01	26	03J13	XAC3373	12E07	A0UV6486D07	V.A.1	proton glutamate symport protein
03	02	26	03J14	XAC0440	11E07	A0UV6458D02	I.C.3	oxidoreductase
03	03	26	03J15	XAC1106	12E08	A0UV6515D03	VIII.B	hypothetical protein
03	04	26	03J16	XAC3145	11E08	A0UV6458D05	VII.C	TolQ protein
03	05	26	03J17	XAC1304	12E09	A0QR6702B05	VIII.B	hypothetical protein
03	06	26	03J18	XAC0012	11E09	A0UV6456E10	II.D.6	pyridoxal phosphate biosynthetic protein
03	07	26	03J19	XAC2655	12E10	A0QR6704E07	VI.A	phage-related baseplate assembly protein
03	08	26	03J20	XAC2081	11E10	A0UV6458E12	III.D.3	lipoprotein releasing system transmembrane protein
03	09	26	03J21	XAC3139	12E11	A0UV6515H02	IX	radical activating enzyme
03	10	26	03J22	XAC1066	11E11	A0UV6456H07	VI.A	phage-related protein
03	11	26	03J23	XAC1746	12E12	A0QR6701H12	V.C	chemotaxis protein
03	12	26	03J24	XAC3247	11E12	A0UV6456H08	VI.C	ISxsc3 transposase
04	01	26	04J13	XAC3460	16E07	A0UV6733H02	VIII.A	conserved hypothetical protein
04	02	26	04J14	XAC0323	15E07	A0QR6728D05	VIII.C	Xanthomonas conserved hypothetical protein
04	03	26	04J15	XAC1465	16E08	A0UV6733H03	VII.G	major cold shock protein
04	04	26	04J16	XACa0005	15E08	A0QR6729B12	VI.C	ISxsc2 transposase
04	05	26	04J17	XAC4225	16E09	A0UV6734F12	I.A.2	xylose isomerase
04	06	26	04J18	XAC0716	15E09	A0QR6728F09	V.A.7	TonB-dependent receptor
04	07	26	04J19	XAC3578	16E10	A0UV6734C08	IV.A.1	lpsJ protein
04	08	26	04J20	XAC4002	15E10	A0QR6729C08	VIII.A	conserved hypothetical protein
04	09	26	04J21	XAC4071	16E11	A0UV6734C10	VIII.B	hypothetical protein
04	10	26	04J22	XAC3911	15E11	A0QH6731D04	VIII.C	Xanthomonas conserved hypothetical protein
04	11	26	04J23	XAC1216	16E12	A0UV6734G05	VIII.C	Xanthomonas conserved hypothetical protein
04	12	26	04J24	XAC2772	15E12	A0QH6731G11	IV.A.2	outer membrane protein
05	01	26	05J13	XAC3684	20E07	A0QH6810G02	VIII.C	Xanthomonas conserved hypothetical protein

05	02	26	05J14	XAC0217	19E07	A0UV6764C10	IV.C	<i>glycosyltransferase</i>
05	03	26	05J15	XAC0951	20E08	A0QR6803C01	III.B.2	<i>50S ribosomal protein L25</i>
05	04	26	05J16	XAC1224	19E08	A0UV6765C04	V.B	<i>cell division topological specificity factor</i>
05	05	26	05J17	XAC1026	20E09	A0QR6803C02	VIII.A	<i>conserved hypothetical protein</i>
05	06	26	05J18	XAC0421	19E09	A0UV6765C11	IV.B	<i>phosphoglycerol transferase I</i>
05	07	26	05J19	XAC1756	20E10	A0QR6805F05	VIII.A	<i>conserved hypothetical protein</i>
05	08	26	05J20	XACb0062	19E10	A0UV6762F12	VI.C	<i>ISxac3 transposase</i>
05	09	26	05J21	XAC0036	20E11	A0QR6804H02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	10	26	05J22	XAC3388	19E11	A0UV6766B01	I.C.7	<i>citrate synthase</i>
05	11	26	05J23	XAC1691	20E12	A0QR6804H07	IV.B	<i>aminotransferase</i>
05	12	26	05J24	XAC1667	19E12	A0UV6766D08	I.C.3	<i>oxidoreductase</i>
06	01	26	06J13	XAC0061	24E07	A0UV6845G12	VIII.A	<i>conserved hypothetical protein</i>
06	02	26	06J14	XAC0950	23E07	A0CE6830D06	II.B.1	<i>phosphoribosyl pyrophosphate synthetase</i>
06	03	26	06J15	XAC3818	24E08	A0UV6845H07	III.A.1	<i>primosomal protein N'</i>
06	04	26	06J16	XAC2362	23E08	A0CE6829G05	VIII.A	<i>conserved hypothetical protein</i>
06	05	26	06J17	XAC1464	24E09	A0UV6845G02	VIII.A	<i>conserved hypothetical protein</i>
06	06	26	06J18	XAC0575	23E09	A0CE6830E03	I.A.1	<i>arabinogalactan endo-1,4-beta-galactosidase</i>
06	07	26	06J19	XAC3420	24E10	A0QR9705A06	II.D.12	<i>glutamate-1-semialdehyde 2,1-aminomutase</i>
06	08	26	06J20	XAC3439	23E10	A0QR6839E02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	26	06J21	XAC2155	24E11	A0CE9724F02	VIII.A	<i>conserved hypothetical protein</i>
06	10	26	06J22	XAC0940	23E11	A0UV6843C09	VIII.A	<i>conserved hypothetical protein</i>
06	11	26	06J23	XAC2223	24E12	A0UT9742F03	VIII.A	<i>conserved hypothetical protein</i>
06	12	26	06J24	XAC3234	23E12	A0UV6844D07	VIII.B	<i>hypothetical protein</i>
07	01	26	07J13	no_clone	28E07	no_clone	NC	<i>no_gene</i>
07	02	26	07J14	XAC2315	27E07	A0QR9874F02	VIII.A	<i>conserved hypothetical protein</i>
07	03	26	07J15	no_clone	28E08	no_clone	NC	<i>no_gene</i>
07	04	26	07J16	XAC3396	27E08	A0UV9872A02	VIII.A	<i>conserved hypothetical protein</i>
07	05	26	07J17	no_clone	28E09	no_clone	NC	<i>no_gene</i>
07	06	26	07J18	XAC1916	27E09	A0UE9889E10	VI.C	<i>ISxac1 transposase</i>
07	07	26	07J19	no_clone	28E10	no_clone	NC	<i>no_gene</i>
07	08	26	07J20	XAC0035	27E10	A0UE9889F07	VIII.A	<i>conserved hypothetical protein</i>
07	09	26	07J21	no_clone	28E11	no_clone	NC	<i>no_gene</i>
07	10	26	07J22	XAC2964	27E11	A0UE9886H11	VIII.A	<i>conserved hypothetical protein</i>
07	11	26	07J23	no_clone	28E12	no_clone	NC	<i>no_gene</i>
07	12	26	07J24	XAC3811	27E12	A0UV9872C07	I.B.10	<i>tropinone reductase</i>
08	01	26	08J13	no_clone	32E07	no_clone	NC	<i>no_gene</i>
08	02	26	08J14	no_clone	31E07	no_clone	NC	<i>no_gene</i>
08	03	26	08J15	no_clone	32E08	no_clone	NC	<i>no_gene</i>
08	04	26	08J16	no_clone	31E08	no_clone	NC	<i>no_gene</i>
08	05	26	08J17	no_clone	32E09	no_clone	NC	<i>no_gene</i>
08	06	26	08J18	no_clone	31E09	no_clone	NC	<i>no_gene</i>
08	07	26	08J19	no_clone	32E10	no_clone	NC	<i>no_gene</i>
08	08	26	08J20	no_clone	31E10	no_clone	NC	<i>no_gene</i>
08	09	26	08J21	no_clone	32E11	no_clone	NC	<i>no_gene</i>
08	10	26	08J22	no_clone	31E11	no_clone	NC	<i>no_gene</i>
08	11	26	08J23	no_clone	32E12	no_clone	NC	<i>no_gene</i>
08	12	26	08J24	no_clone	31E12	no_clone	NC	<i>no_gene</i>
09	01	26	09E03	Score_card	33C02	Score_card	SC	<i>Score_card</i>
09	02	26	09I19	Score_card	33E10	Score_card	SC	<i>Score_card</i>
09	03	26	09O11	Score_card	33H06	Score_card	SC	<i>Score_card</i>
09	04	26	09E04	Score_card	34C02	Score_card	SC	<i>Score_card</i>
09	05	26	09I20	Score_card	34E10	Score_card	SC	<i>Score_card</i>
09	06	26	09O12	Score_card	34H06	Score_card	SC	<i>Score_card</i>
09	07	26	09F04	Score_card	35C02	Score_card	SC	<i>Score_card</i>
09	08	26	09J20	Score_card	35E10	Score_card	SC	<i>Score_card</i>

09	09	26	09P12	Score_card	35H06	Score_card	SC	Score_card
09	10	26	09F03	Score_card	36C02	Score_card	SC	Score_card
09	11	26	09J19	Score_card	36E10	Score_card	SC	Score_card
09	12	26	09P11	Score_card	36H06	Score_card	SC	Score_card
01	01	27	01K13	XAC2718	01F07	A0JJ0102H10	I.D.2	<i>transcriptional regulator</i>
01	02	27	01K14	XAC1851	02F07	A0JE1211F08	I.A.2	<i>hydroxymethylglutaryl-CoA lyase</i>
01	03	27	01K15	XAC3054	01F08	A0AC0115G05	I.A.3	<i>acyl-CoA dehydrogenase</i>
01	04	27	01K16	XAC1825	02F08	A0JE1211G05	I.D.2	<i>fumarate and nitrate reduction regulatory protein</i>
01	05	27	01K17	XAC3418	01F09	A0AC0115E05	IV.A.2	<i>Oar protein</i>
01	06	27	01K18	XAC1097	02F09	A0JE1211G09	II.D.4	<i>molybdenum cofactor biosynthesis protein A</i>
01	07	27	01K19	XAC0682	01F10	A0JJ1038G03	VIII.A	<i>conserved hypothetical protein</i>
01	08	27	01K20	XAC0349	02F10	A0AR1103D07	V.A.7	<i>MFS transporter</i>
01	09	27	01K21	XAC3948	01F11	A0JJ1002D08	VIII.A	<i>conserved hypothetical protein</i>
01	10	27	01K22	XAC2589	02F11	A0AM1112D12	III.B.4	<i>phenylalanyl-tRNA synthetase betachain</i>
01	11	27	01K23	XAC3691	01F12	A0AR1008E04	VIII.A	<i>conserved hypothetical protein</i>
01	12	27	01K24	XAC2210	02F12	A0JJ1110B05	VIII.A	<i>conserved hypothetical protein</i>
02	01	27	02K13	XAC1648	05F07	A0UT5315A07	II.A.3	<i>phosphoserine aminotransferase</i>
02	02	27	02K14	XAC2074	06F07	A0QR5517C12	VIII.A	<i>conserved hypothetical protein</i>
02	03	27	02K15	XAC0394	05F08	A0UT5316A09	VII.B	<i>HrpF protein</i>
02	04	27	02K16	XAC2305	06F08	A0QR5517E11	VI.B	<i>pheromone shutdown protein</i>
02	05	27	02K17	XAC3383	05F09	A0QR5321F07	IV.D	<i>fimbrial assembly membrane protein</i>
02	06	27	02K18	XAC1013	06F09	A0QR5701F12	VIII.A	<i>conserved hypothetical protein</i>
02	07	27	02K19	XAC0266	05F10	A0QR5408D08	I.D.2	<i>transcriptional regulator acrR family</i>
02	08	27	02K20	XAC3013	06F10	A0QR5701C06	II.D.12	<i>uroporphyrinogen decarboxylase</i>
02	09	27	02K21	XAC1071	05F11	A0CE5318C08	VI.C	<i>ISxac3 transposase</i>
02	10	27	02K22	XAC3836	06F11	A0QR5702D06	VIII.A	<i>conserved hypothetical protein</i>
02	11	27	02K23	XAC1883	05F12	A0QR5407A09	VIII.A	<i>conserved hypothetical protein</i>
02	12	27	02K24	XAC4146	06F12	A0QR5702C01	VIII.A	<i>conserved hypothetical protein</i>
03	01	27	03K13	XAC1324	09F07	A0CE6345G11	VIII.A	<i>conserved hypothetical protein</i>
03	02	27	03K14	XAC3661	10F07	A0QH6382D04	I.C.6	<i>dihydrolipoamide acetyltransferase</i>
03	03	27	03K15	XAC1427	09F08	A0UV6350F11	VII.F	<i>protein U</i>
03	04	27	03K16	XAC4369	10F08	A0UT6386F11	I.B.9	<i>phosphatase precursor</i>
03	05	27	03K17	XAC1153	09F09	A0UV6351D03	VIII.B	<i>hypothetical protein</i>
03	06	27	03K18	XAC1070	10F09	A0QR6389C01	VI.C	<i>ISxac3 transposase</i>
03	07	27	03K19	XAC1970	09F10	A0AC6359C10	IX	<i>response regulator</i>
03	08	27	03K20	XAC1487	10F10	A0QR6390F12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	09	27	03K21	XAC3696	09F11	A0UV6363B08	VIII.A	<i>conserved hypothetical protein</i>
03	10	27	03K22	XAC2639	10F11	A0QR6418G06	III.A.5	<i>site-specific DNA-methyltransferase</i>
03	11	27	03K23	XAC0373	09F12	A0AC6360F08	I.D.2	<i>transcriptional regulator</i>
03	12	27	03K24	XAC3954	10F12	A0AC6430D04	VIII.B	<i>hypothetical protein</i>
04	01	27	04K13	XAC3057	13F07	A0QR6705B06	VII.C	<i>beta-lactamase</i>
04	02	27	04K14	XAC2444	14F07	A0UV6712A10	VIII.A	<i>conserved hypothetical protein</i>
04	03	27	04K15	XAC2965	13F08	A0QR6705D09	IV.C	<i>UDP-N-acetylglucosamine 1-carboxyvinyltransferase</i>
04	04	27	04K16	XAC3728	14F08	A0UV6712E06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	05	27	04K17	XAC3650	13F09	A0UV6707F08	I.C.8	<i>ATP synthase, gamma chain</i>
04	06	27	04K18	XAC3589	14F09	A0UV6712E09	IV.A.1	<i>integral membrane protein</i>
04	07	27	04K19	XAC2050	13F10	A0UV6709H07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	08	27	04K20	XAC1873	14F10	A0UV6712B08	IV.A.1	<i>inner membrane protein</i>
04	09	27	04K21	XAC2057	13F11	A0UV6709H11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	10	27	04K22	XAC1884	14F11	A0QR6723A08	VIII.A	<i>conserved hypothetical protein</i>
04	11	27	04K23	XAC2033	13F12	A0UV6710A10	II.D.4	<i>molybdopterin guanine dinucleotide synthase</i>
04	12	27	04K24	XAC1552	14F12	A0QR6725B03	VIII.A	<i>conserved hypothetical protein</i>
05	01	27	05K13	XAC3176	17F07	A0UV6739C10	V.A.4	<i>outer membrane receptor; citrate-dependent iron transporter</i>
05	02	27	05K14	XAC2222	18F07	A0CE6746E01	VI.A	<i>phage-related integrase</i>

05	03	27	05K15	XAC3630	17F08	A0UV6738A05	VII.C	<i>copper resistance protein A precursor</i>
05	04	27	05K16	XACa0032	18F08	A0QR6752G10	VIII.A	<i>conserved hypothetical protein</i>
05	05	27	05K17	XAC3100	17F09	A0UV6739A02	IV.D	<i>pilus biogenesis protein</i>
05	06	27	05K18	XAC0242	18F09	A0QR6752A12	II.D.11	<i>ubiquinone biosynthesis protein</i>
05	07	27	05K19	XAC2009	17F10	A0UV6739D01	VIII.A	<i>conserved hypothetical protein</i>
05	08	27	05K20	XAC2182	18F10	A0QR6755D05	VIII.A	<i>conserved hypothetical protein</i>
05	09	27	05K21	XAC2179	17F11	A0CE6743A12	III.A.4	<i>RadC family protein</i>
05	10	27	05K22	XAC2256	18F11	A0UV6761G01	VIII.A	<i>conserved hypothetical protein</i>
05	11	27	05K23	XAC0067	17F12	A0CE6743E05	IX	<i>microcystin dependent protein</i>
05	12	27	05K24	XAC2536	18F12	A0UE6758C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	01	27	06K13	XACa0027	21F07	A0UV6818E02	VI.B	<i>plasmid stable inheritance protein K</i>
06	02	27	06K14	XAC3673	22F07	A0AC6823F01	I.D.3	<i>histidine kinase</i>
06	03	27	06K15	XAC1038	21F08	A0QH6813C11	IV.C	<i>glycosyl transferase</i>
06	04	27	06K16	XAC4211	22F08	A0AC6827H07	III.B.4	<i>glycyl-tRNA synthetase alpha chain</i>
06	05	27	06K17	XAC0408	21F09	A0UV6818E12	VII.B	<i>HrpB2 protein</i>
06	06	27	06K18	XAC1694	22F09	A0AC6825G10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	07	27	06K19	XAC0555	21F10	A0UV6818H09	VIII.A	<i>conserved hypothetical protein</i>
06	08	27	06K20	XAC1381	22F10	A0AC6824G06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	27	06K21	XAC0098	21F11	A0UV6821H09	VIII.B	<i>hypothetical protein</i>
06	10	27	06K22	XAC2583	22F11	A0AC6823C03	VII.E	<i>GumD protein</i>
06	11	27	06K23	XAC1869	21F12	A0UV6821B07	VIII.B	<i>hypothetical protein</i>
06	12	27	06K24	XAC2865	22F12	A0AC6827F09	V.C	<i>chemotaxis histidine protein kinase</i>
07	01	27	07K13	XAC0518	25F07	A0QR9759B02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	02	27	07K14	XAC3213	26F07	A0CE9839A04	V.C	<i>chemotaxis protein</i>
07	03	27	07K15	XAC3831	25F08	A0QR9753H08	III.B.5	<i>transcription termination factor Rho</i>
07	04	27	07K16	XAC1212	26F08	A0UV9870B09	I.C.3	<i>oxidoreductase</i>
07	05	27	07K17	XAC0172	25F09	A0QR9756G08	VIII.A	<i>conserved hypothetical protein</i>
07	06	27	07K18	XACb0036	26F09	A0QR9864G07	VII.H	<i>VirB1 protein</i>
07	07	27	07K19	XAC2716	25F10	A0QR9756A07	II.A.4	<i>tryptophan synthase alpha chain</i>
07	08	27	07K20	XAC3184	26F10	A0UE9854A06	II.D.14	<i>cobalamin synthase</i>
07	09	27	07K21	XAC1988	25F11	A0QR9757C01	V.C	<i>flagellar protein</i>
07	10	27	07K22	XAC2944	26F11	A0UT9846H03	VIII.A	<i>conserved hypothetical protein</i>
07	11	27	07K23	XAC1099	25F12	A0QR9817A01	II.D.4	<i>molybdopterin-converting factor chain 1</i>
07	12	27	07K24	XAC1110	26F12	A0UT9843A02	VIII.A	<i>conserved hypothetical protein</i>
08	01	27	08K13	XAC3788	29F07	genomic_DNA	I.D.4	<i>RNA polymerase sigma-70 factor</i>
08	02	27	08K14	no_clone	30F07	no_clone	NC	<i>no_gene</i>
08	03	27	08K15	XAC2610	29F08	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	04	27	08K16	no_clone	30F08	no_clone	NC	<i>no_gene</i>
08	05	27	08K17	XAC0677	29F09	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	06	27	08K18	no_clone	30F09	no_clone	NC	<i>no_gene</i>
08	07	27	08K19	XAC1989	29F10	genomic_DNA	V.C	<i>flagellar protein</i>
08	08	27	08K20	no_clone	30F10	no_clone	NC	<i>no_gene</i>
08	09	27	08K21	no_clone	29F11	no_clone	NC	<i>no_gene</i>
08	10	27	08K22	no_clone	30F11	no_clone	NC	<i>no_gene</i>
08	11	27	08K23	no_clone	29F12	no_clone	NC	<i>no_gene</i>
08	12	27	08K24	no_clone	30F12	no_clone	NC	<i>no_gene</i>
09	01	27	09E05	Score_card	33C03	Score_card	SC	<i>Score_card</i>
09	02	27	09I21	Score_card	33E11	Score_card	SC	<i>Score_card</i>
09	03	27	09O13	Score_card	33H07	Score_card	SC	<i>Score_card</i>
09	04	27	09E06	Score_card	34C03	Score_card	SC	<i>Score_card</i>
09	05	27	09I22	Score_card	34E11	Score_card	SC	<i>Score_card</i>
09	06	27	09O14	Score_card	34H07	Score_card	SC	<i>Score_card</i>
09	07	27	09F06	Score_card	35C03	Score_card	SC	<i>Score_card</i>
09	08	27	09J22	Score_card	35E11	Score_card	SC	<i>Score_card</i>
09	09	27	09P14	Score_card	35H07	Score_card	SC	<i>Score_card</i>

09	10	27	09F05	Score_card	36C03	Score_card	SC	Score_card
09	11	27	09J21	Score_card	36E11	Score_card	SC	Score_card
09	12	27	09P13	Score_card	36H07	Score_card	SC	Score_card
01	01	28	01L13	XAC0870	04F07	A0UV1534F03	VIII.A	<i>conserved hypothetical protein</i>
01	02	28	01L14	XAC1889	03F07	A0AC1317A07	V.C	<i>chemotaxis protein</i>
01	03	28	01L15	XAC3764	04F08	A0UT1515C10	VI.C	<i>ISxac2 transposase</i>
01	04	28	01L16	XAC3523	03F08	A0AM1309C07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	05	28	01L17	XAC0450	04F09	A0RN1627B01	VIII.A	<i>conserved hypothetical protein</i>
01	06	28	01L18	XAC0791	03F09	A0AM1393B06	II.A.2	<i>5,10-methylenetetrahydrofolate reductase</i>
01	07	28	01L19	XAC2200	04F10	A0RN1627C07	VIII.B	<i>hypothetical protein</i>
01	08	28	01L20	XAC0672	03F10	A0JE1379F06	VIII.A	<i>conserved hypothetical protein</i>
01	09	28	01L21	XAC3822	04F11	A0JJ1656C06	I.C.1	<i>NADH dehydrogenase</i>
01	10	28	01L22	XAC2599	03F11	A0JJ1343E05	I.A.1	<i>alpha-glucosidase</i>
01	11	28	01L23	XAC2212	04F12	A0CE1622F02	III.A.1	<i>DNA topoisomerase III</i>
01	12	28	01L24	XAC3158	03F12	A0JJ1465G12	V.A.7	<i>TonB-dependent receptor</i>
02	01	28	02L13	XAC2369	08F07	A0QH6306H03	VII.G	<i>general stress protein</i>
02	02	28	02L14	XACb0059	07F07	A0QR5704F12	VII.H	<i>virulence associated protein</i>
02	03	28	02L15	XAC2130	08F08	A0QH6308B05	VIII.A	<i>conserved hypothetical protein</i>
02	04	28	02L16	XAC3422	07F08	A0QR5704E12	I.C.3	<i>electron transfer protein azurin I</i>
02	05	28	02L17	XAC1948	08F09	A0QH6308F02	V.C	<i>flagellar protein</i>
02	06	28	02L18	XAC2824	07F09	A0QR5903G08	II.B.4	<i>phosphodiesterase-nucleotide pyrophosphatase</i>
02	07	28	02L19	XAC4352	08F10	A0CE6314E07	VII.C	<i>glutathione S-transferase</i>
02	08	28	02L20	XAC0576	07F10	A0QR5903D03	I.C.6	<i>pyruvate dehydrogenase</i>
02	09	28	02L21	XAC2486	08F11	A0UV6210H06	I.C.2	<i>formate dehydrogenase a chain</i>
02	10	28	02L22	XAC0684	07F11	A0QR5902H03	I.D.1	<i>two-component system, regulatory protein</i>
02	11	28	02L23	XAC1472	08F12	A0UV6210H10	I.A.2	<i>glutaryl-CoA dehydrogenase</i>
02	12	28	02L24	XAC0042	07F12	A0QR5902D11	IV.C	<i>glycosyltransferase</i>
03	01	28	03L13	XAC1885	12F07	A0UV6488D09	I.C.7	<i>aconitate hydratase 2</i>
03	02	28	03L14	XAC3314	11F07	A0QR6462H04	VIII.A	<i>conserved hypothetical protein</i>
03	03	28	03L15	XAC2394	12F08	A0QR6613A01	VII.C	<i>glutathione S-transferase</i>
03	04	28	03L16	XAC4089	11F08	A0UE6463C09	II.D.17	<i>halogenase</i>
03	05	28	03L17	XAC2901	12F09	A0QR6704A09	VIII.A	<i>conserved hypothetical protein</i>
03	06	28	03L18	XAC1404	11F09	A0UV6458E06	VIII.A	<i>conserved hypothetical protein</i>
03	07	28	03L19	XAC1922	12F10	A0UV6515G01	VIII.B	<i>hypothetical protein</i>
03	08	28	03L20	XAC3243	11F10	A0QR6460F05	IV.D	<i>type IV pre-pilin leader peptidase</i>
03	09	28	03L21	XAC0976	12F11	A0UE6610B04	III.B.2	<i>30S ribosomal protein S19</i>
03	10	28	03L22	XAC3155	11F11	A0QR6460F11	VIII.A	<i>conserved hypothetical protein</i>
03	11	28	03L23	XACa0029	12F12	A0QR6702H02	VI.C	<i>transposase</i>
03	12	28	03L24	XAC2349	11F12	A0UV6458G12	II.A.1	<i>acetylornithine deacetylase</i>
04	01	28	04L13	XAC1779	16F07	A0UV6734B07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	02	28	04L14	XAC2231	15F07	A0QR6728H12	IX	<i>B-cell mitogen related protein</i>
04	03	28	04L15	XAC2089	16F08	A0UV6734F10	IV.C	<i>3-deoxy-manno-octulosonate cytidyltransferase</i>
04	04	28	04L16	XAC2433	15F08	A0QR6729D08	VI.B	<i>resolvase</i>
04	05	28	04L17	XAC0908	16F09	A0UV6734G12	II.D.12	<i>protoporphyrinogen oxidase</i>
04	06	28	04L18	XAC4274	15F09	A0QR6729C07	VII.F	<i>OmpA-related protein</i>
04	07	28	04L19	XAC0925	16F10	A0UV6734E07	VIII.A	<i>conserved hypothetical protein</i>
04	08	28	04L20	XAC4068	15F10	A0QR6729E04	II.B.2	<i>2-dehydropantoate 2-reductase</i>
04	09	28	04L21	XAC2471	16F11	A0UV6734G03	V.A.1	<i>polyamine transport protein</i>
04	10	28	04L22	XAC1171	15F11	A0QH6731G10	I.D.3	<i>serine/threonine kinase</i>
04	11	28	04L23	XAC3239	16F12	A0UV6734H04	IV.D	<i>pilus biogenesis protein</i>
04	12	28	04L24	XAC0015	15F12	A0QH6732D11	VIII.B	<i>hypothetical protein</i>
05	01	28	05L13	XAC2000	20F07	A0QH6811D05	VIII.A	<i>conserved hypothetical protein</i>
05	02	28	05L14	XAC1705	19F07	A0UV6764H04	V.A.7	<i>MFS transporter</i>
05	03	28	05L15	XAC4216	20F08	A0QR6803G02	V.A.6	<i>sec-independent protein translocase</i>
05	04	28	05L16	XAC2731	19F08	A0UV6762E02	VIII.A	<i>conserved hypothetical protein</i>

05	05	28	05L17	XAC3680	20F09	A0QH6810A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	06	28	05L18	XAC3966	19F09	A0UV6765F01	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	07	28	05L19	XAC3736	20F10	A0QH6810B01	I.C.3	<i>cyanide insensitive terminal oxidase</i>
05	08	28	05L20	XAC2494	19F10	A0UV6762H07	VII.C	<i>drug resistance translocase</i>
05	09	28	05L21	XAC2963	20F11	A0QR6805C08	VIII.A	<i>conserved hypothetical protein</i>
05	10	28	05L22	XAC0364	19F11	A0UV6766G04	I.C.2	<i>glutaconate CoA transferase subunit A</i>
05	11	28	05L23	XAC1819	20F12	A0QR6805C10	I.D.2	<i>tryptophan-rich sensory protein</i>
05	12	28	05L24	XAC2743	19F12	A0UV6766G06	IV.A.2	<i>Oar protein</i>
06	01	28	06L13	XAC3766	24F07	A0UV6846C05	VIII.B	<i>hypothetical protein</i>
06	02	28	06L14	XAC2857	23F07	A0CE6831B10	VIII.A	<i>conserved hypothetical protein</i>
06	03	28	06L15	XAC1423	24F08	A0UV6846C06	III.C.2	<i>pili assembly chaperone</i>
06	04	28	06L16	XAC1602	23F08	A0CE6830E02	VIII.A	<i>conserved hypothetical protein</i>
06	05	28	06L17	XAC1154	24F09	A0UV6846C09	I.D.2	<i>regulatory protein pilH family</i>
06	06	28	06L18	XAC0228	23F09	A0CE6830H03	VIII.B	<i>hypothetical protein</i>
06	07	28	06L19	XAC4055	24F10	A0QR9705E12	III.C.3	<i>cysteine proteinase</i>
06	08	28	06L20	XAC4066	23F10	A0QR6840F06	I.A.2	<i>phenol hydroxylase</i>
06	09	28	06L21	XAC3692	24F11	A0UV9748E09	VIII.A	<i>conserved hypothetical protein</i>
06	10	28	06L22	XAC2133	23F11	A0UV6844A02	IX	<i>oxidoreductase</i>
06	11	28	06L23	XAC1380	24F12	A0UV9747D03	I.D.4	<i>RNA polymerase sigma factor</i>
06	12	28	06L24	XAC1666	23F12	A0QR6839F02	V.C	<i>chemotaxis protein</i>
07	01	28	07L13	no_clone	28F07	no_clone	NC	<i>no_gene</i>
07	02	28	07L14	XAC2012	27F07	A0UT9876A10	I.A.3	<i>3-ketoacyl-CoA thiolase</i>
07	03	28	07L15	no_clone	28F08	no_clone	NC	<i>no_gene</i>
07	04	28	07L16	XAC2660	27F08	A0UT9875A08	VI.A	<i>phage-related baseplate assembly protein</i>
07	05	28	07L17	no_clone	28F09	no_clone	NC	<i>no_gene</i>
07	06	28	07L18	XAC1836	27F09	A0UT9875C05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	07	28	07L19	no_clone	28F10	no_clone	NC	<i>no_gene</i>
07	08	28	07L20	XACb0040	27F10	A0UV9872A07	VII.H	<i>VirB8 protein</i>
07	09	28	07L21	no_clone	28F11	no_clone	NC	<i>no_gene</i>
07	10	28	07L22	XAC2712	27F11	A0UE9890C09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	11	28	07L23	no_clone	28F12	no_clone	NC	<i>no_gene</i>
07	12	28	07L24	XAC0502	27F12	A0UT9876A01	VI.C	<i>ISxac1 transposase</i>
08	01	28	08L13	no_clone	32F07	no_clone	NC	<i>no_gene</i>
08	02	28	08L14	no_clone	31F07	no_clone	NC	<i>no_gene</i>
08	03	28	08L15	no_clone	32F08	no_clone	NC	<i>no_gene</i>
08	04	28	08L16	no_clone	31F08	no_clone	NC	<i>no_gene</i>
08	05	28	08L17	no_clone	32F09	no_clone	NC	<i>no_gene</i>
08	06	28	08L18	no_clone	31F09	no_clone	NC	<i>no_gene</i>
08	07	28	08L19	no_clone	32F10	no_clone	NC	<i>no_gene</i>
08	08	28	08L20	no_clone	31F10	no_clone	NC	<i>no_gene</i>
08	09	28	08L21	no_clone	32F11	no_clone	NC	<i>no_gene</i>
08	10	28	08L22	no_clone	31F11	no_clone	NC	<i>no_gene</i>
08	11	28	08L23	no_clone	32F12	no_clone	NC	<i>no_gene</i>
08	12	28	08L24	no_clone	31F12	no_clone	NC	<i>no_gene</i>
09	01	28	09E07	Score_card	33C04	Score_card	SC	<i>Score_card</i>
09	02	28	09I23	Score_card	33E12	Score_card	SC	<i>Score_card</i>
09	03	28	09O15	Score_card	33H08	Score_card	SC	<i>Score_card</i>
09	04	28	09E08	Score_card	34C04	Score_card	SC	<i>Score_card</i>
09	05	28	09I24	Score_card	34E12	Score_card	SC	<i>Score_card</i>
09	06	28	09O16	Score_card	34H08	Score_card	SC	<i>Score_card</i>
09	07	28	09F08	Score_card	35C04	Score_card	SC	<i>Score_card</i>
09	08	28	09J24	Score_card	35E12	Score_card	SC	<i>Score_card</i>
09	09	28	09P16	Score_card	35H08	Score_card	SC	<i>Score_card</i>
09	10	28	09F07	Score_card	36C04	Score_card	SC	<i>Score_card</i>
09	11	28	09J23	Score_card	36E12	Score_card	SC	<i>Score_card</i>

09	12	28	09P15	Score_card	36H08	Score_card	SC	Score_card
01	01	29	01M13	XAC3160	01G07	A0CE0105A05	I.B.9	<i>phospholipase C</i>
01	02	29	01M14	XAC1255	02G07	A0JJ1101F08	III.C.1	<i>lipoprotein signal peptidase</i>
01	03	29	01M15	XAC2739	01G08	A0JJ0101G03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	04	29	01M16	XAC4137	02G08	A0AR1102A01	VI.C	<i>ISxac1 transposase</i>
01	05	29	01M17	XAC2008	01G09	A0AC0115H01	III.C.2	<i>outer-membrane lipoproteins carrier protein precursor</i>
01	06	29	01M18	XAC3723	02G09	A0AR1102A02	VIII.B	<i>hypothetical protein</i>
01	07	29	01M19	XAC3530	01G10	A0JE1044B04	V.A.7	<i>iron receptor</i>
01	08	29	01M20	XAC1268	02G10	A0AM1111G06	VIII.A	<i>conserved hypothetical protein</i>
01	09	29	01M21	XAC1524	01G11	A0JJ1004G06	II.D.6	<i>pyridoxine kinase</i>
01	10	29	01M22	XAC2828	02G11	A0JE1129A05	VIII.A	<i>conserved hypothetical protein</i>
01	11	29	01M23	XAC2383	01G12	A0JJ1065F04	IX	<i>phosphate-binding protein</i>
01	12	29	01M24	XAC2891	02G12	A0AM1112H04	VIII.A	<i>conserved hypothetical protein</i>
02	01	29	02M13	XAC4122	05G07	A0CE5319B12	VIII.A	<i>conserved hypothetical protein</i>
02	02	29	02M14	XAC1061	06G07	A0QR5520B10	VIII.A	<i>conserved hypothetical protein</i>
02	03	29	02M15	XAC3140	05G08	A0QR5317H04	VIII.A	<i>conserved hypothetical protein</i>
02	04	29	02M16	XAC2785	06G08	A0QR5520C04	VIII.B	<i>hypothetical protein</i>
02	05	29	02M17	XAC1711	05G09	A0CE5310F06	V.A.7	<i>transport protein</i>
02	06	29	02M18	XAC3305	06G09	A0QR5702B01	VIII.A	<i>conserved hypothetical protein</i>
02	07	29	02M19	XAC1018	05G10	A0QR5317A08	V.A.2	<i>ABC transporter sulfate permease</i>
02	08	29	02M20	XAC2080	06G10	A0QR5703A09	VIII.A	<i>conserved hypothetical protein</i>
02	09	29	02M21	XAC0864	05G11	A0QR5323G04	II.D.6	<i>pyridoxal phosphate biosynthetic protein</i>
02	10	29	02M22	XAC2066	06G11	A0QR5702G05	V.A.7	<i>transport protein</i>
02	11	29	02M23	XAC3610	05G12	A0QR5407C09	III.B.5	<i>ATP-dependent RNA helicase</i>
02	12	29	02M24	XAC1568	06G12	A0QR5702D09	VIII.A	<i>conserved hypothetical protein</i>
03	01	29	03M13	XAC3529	09G07	A0UV6348E01	V.A.7	<i>iron receptor</i>
03	02	29	03M14	XAC2024	10G07	A0UT6386F04	V.A.7	<i>TonB-dependent receptor</i>
03	03	29	03M15	XAC1971	09G08	A0UV6351C12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	04	29	03M16	XAC0564	10G08	A0QR6389B09	II.D.1	<i>malonate decarboxylase</i>
03	05	29	03M17	XAC3443	09G09	A0AC6359C09	I.D.2	<i>response regulator</i>
03	06	29	03M18	XAC1408	10G09	A0QR6393B10	IV.C	<i>lipid A disaccharide synthase</i>
03	07	29	03M19	XAC2488	09G10	A0UV6364F08	V.A.7	<i>integral membrane transporter</i>
03	08	29	03M20	XAC1001	10G10	A0QR6393C05	IV.A.1	<i>membrane protein</i>
03	09	29	03M21	XAC3906	09G11	A0QR6369G12	V.B	<i>chromosome partitioning protein</i>
03	10	29	03M22	XAC2822	10G11	A0AC6440D03	III.A.4	<i>DNA methylation and regulatory protein</i>
03	11	29	03M23	XAC3036	09G12	A0UV6363B10	I.A.2	<i>L-serine dehydratase</i>
03	12	29	03M24	XAC2961	10G12	A0UV6433C09	II.B.1	<i>5'-phosphoribosylglycinamide transformylase</i>
04	01	29	04M13	XAC4019	13G07	A0QR6705D05	VIII.A	<i>conserved hypothetical protein</i>
04	02	29	04M14	XAC2287	14G07	A0UV6712B02	II.B.1	<i>glutamine amidotransferase</i>
04	03	29	04M15	XAC4261	13G08	A0QR6705F02	VIII.B	<i>hypothetical protein</i>
04	04	29	04M16	XAC3104	14G08	A0UE6718A11	VII.C	<i>TonB protein</i>
04	05	29	04M17	XAC0160	13G09	A0UV6708A05	I.A.1	<i>xylanase</i>
04	06	29	04M18	XAC2605	14G09	A0QH6713B11	VIII.A	<i>conserved hypothetical protein</i>
04	07	29	04M19	XAC1608	13G10	A0UV6710G08	VIII.B	<i>hypothetical protein</i>
04	08	29	04M20	XAC4372	14G10	A0UV6712D07	IV.A.1	<i>60kDa inner-membrane protein</i>
04	09	29	04M21	XAC1432	13G11	A0UV6710E06	II.A.2	<i>succinyl-diaminopimelate desuccinylase</i>
04	10	29	04M22	XAC2530	14G11	A0UV6712D08	VIII.A	<i>conserved hypothetical protein</i>
04	11	29	04M23	XAC3099	13G12	A0UV6710C12	IV.D	<i>pilus biogenesis protein</i>
04	12	29	04M24	XAC2440	14G12	A0UV6712C01	VI.B	<i>plasmid mobilization protein</i>
05	01	29	05M13	XAC0356	17G07	A0UV6739F09	I.A.2	<i>P-hydroxybenzoate hydroxylase</i>
05	02	29	05M14	XAC2189	18G07	A0QR6749B12	VIII.B	<i>hypothetical protein</i>
05	03	29	05M15	XAC4176	17G08	A0UV6739A01	V.A.7	<i>solute:Na⁺ symporter</i>
05	04	29	05M16	XAC3838	18G08	A0QR6752A05	VIII.A	<i>conserved hypothetical protein</i>
05	05	29	05M17	XAC3071	17G09	A0UV6739C12	V.A.7	<i>TonB-dependent receptor</i>
05	06	29	05M18	XAC2886	18G09	A0QR6755C06	VIII.A	<i>conserved hypothetical protein</i>

05	07	29	05M19	XAC1611	17G10	A0UV6739F02	VIII.B	<i>hypothetical protein</i>
05	08	29	05M20	XAC1483	18G10	A0QR6755E05	VII.C	<i>RND multidrug efflux transporter MexF</i>
05	09	29	05M21	XAC4133	17G11	A0CE6743E02	VIII.A	<i>conserved hypothetical protein</i>
05	10	29	05M22	XAC2290	18G11	A0UV6762A10	VIII.A	<i>conserved hypothetical protein</i>
05	11	29	05M23	XAC0289	17G12	A0CE6743F09	VIII.A	<i>conserved hypothetical protein</i>
05	12	29	05M24	XAC4221	18G12	A0UE6759D03	IX	<i>hydrolase</i>
06	01	29	06M13	XAC0226	21G07	A0QH6813C04	I.D.1	<i>two-component system, regulatory protein</i>
06	02	29	06M14	XAC3254	22G07	A0AC6824D06	I.A.1	<i>glycogen debranching enzyme</i>
06	03	29	06M15	XAC4061	21G08	A0UT6814D02	VIII.A	<i>conserved hypothetical protein</i>
06	04	29	06M16	XAC2146	22G08	A0CE6828D10	VIII.B	<i>hypothetical protein</i>
06	05	29	06M17	XAC2128	21G09	A0UV6818H08	I.A.2	<i>2-keto-gluconate dehydrogenase</i>
06	06	29	06M18	XAC1295	22G09	A0AC6827B11	III.B.2	<i>50S ribosomal protein L19</i>
06	07	29	06M19	XAC4053	21G10	A0UV6819C04	V.A.4	<i>ABC transporter sodium permease</i>
06	08	29	06M20	XAC2775	22G10	A0AC6825C03	VIII.A	<i>conserved hypothetical protein</i>
06	09	29	06M21	XAC1251	21G11	A0UV6818F06	III.B.2	<i>30S ribosomal protein S20</i>
06	10	29	06M22	XAC0797	22G11	A0AC6823F12	I.A.2	<i>D-galactose 1-dehydrogenase</i>
06	11	29	06M23	XAC1303	21G12	A0UV6821E02	III.A.4	<i>DNA mismatch repair protein</i>
06	12	29	06M24	XAC3664	22G12	A0CE6828B06	IV.A.2	<i>outer membrane protein</i>
07	01	29	07M13	XAC0221	25G07	A0AC9761F03	III.C.2	<i>protein-export protein</i>
07	02	29	07M14	XAC2956	26G07	A0CE9841G04	III.A.1	<i>replication related protein</i>
07	03	29	07M15	XAC2818	25G08	A0QR9756F07	IV.A.1	<i>inner membrane protein</i>
07	04	29	07M16	XAC0243	26G08	A0CE9839A10	VIII.A	<i>conserved hypothetical protein</i>
07	05	29	07M17	XAC0131	25G09	A0QR9758E02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	06	29	07M18	XAC4087	26G09	A0UV9870C07	VIII.A	<i>conserved hypothetical protein</i>
07	07	29	07M19	XAC2267	25G10	A0QR9759C08	VIII.B	<i>hypothetical protein</i>
07	08	29	07M20	XAC1036	26G10	A0UE9858G06	VIII.A	<i>conserved hypothetical protein</i>
07	09	29	07M21	XAC0370	25G11	A0QR9758G03	I.A.2	<i>b-ketoadipate enol-lactone hydrolase</i>
07	10	29	07M22	XAC2336	26G11	A0UT9850A04	I.C.1	<i>cytochrome D ubiquinol oxidase subunit I</i>
07	11	29	07M23	XAC2924	25G12	A0UV9822G07	IV.D	<i>twitching motility protein</i>
07	12	29	07M24	XAC3083	26G12	A0UT9846C04	VIII.A	<i>conserved hypothetical protein</i>
08	01	29	08M13	XAC3824	29G07	genomic_DNA	I.D.4	<i>RNA polymerase sigma-32 factor</i>
08	02	29	08M14	no_clone	30G07	no_clone	NC	<i>no_gene</i>
08	03	29	08M15	XAC2614	29G08	genomic_DNA	VII.H	<i>VirB4 protein</i>
08	04	29	08M16	no_clone	30G08	no_clone	NC	<i>no_gene</i>
08	05	29	08M17	XAC2047	29G09	genomic_DNA	IX	<i>PHA synthase subunit</i>
08	06	29	08M18	no_clone	30G09	no_clone	NC	<i>no_gene</i>
08	07	29	08M19	XAC4129	29G10	genomic_DNA	I.D.4	<i>ECF sigma factor</i>
08	08	29	08M20	no_clone	30G10	no_clone	NC	<i>no_gene</i>
08	09	29	08M21	no_clone	29G11	no_clone	NC	<i>no_gene</i>
08	10	29	08M22	no_clone	30G11	no_clone	NC	<i>no_gene</i>
08	11	29	08M23	no_clone	29G12	no_clone	NC	<i>no_gene</i>
08	12	29	08M24	no_clone	30G12	no_clone	NC	<i>no_gene</i>
09	01	29	09E09	Score_card	33C05	Score_card	SC	<i>Score_card</i>
09	02	29	09K01	Score_card	33F01	Score_card	SC	<i>Score_card</i>
09	03	29	09O17	Score_card	33H09	Score_card	SC	<i>Score_card</i>
09	04	29	09E10	Score_card	34C05	Score_card	SC	<i>Score_card</i>
09	05	29	09K02	Score_card	34F01	Score_card	SC	<i>Score_card</i>
09	06	29	09O18	Score_card	34H09	Score_card	SC	<i>Score_card</i>
09	07	29	09F10	Score_card	35C05	Score_card	SC	<i>Score_card</i>
09	08	29	09L02	Score_card	35F01	Score_card	SC	<i>Score_card</i>
09	09	29	09P18	Score_card	35H09	Score_card	SC	<i>Score_card</i>
09	10	29	09F09	Score_card	36C05	Score_card	SC	<i>Score_card</i>
09	11	29	09L01	Score_card	36F01	Score_card	SC	<i>Score_card</i>
09	12	29	09P17	Score_card	36H09	Score_card	SC	<i>Score_card</i>
01	01	03	01C01	XAC3556	01B01	A0JJ0101G09	III.C.3	<i>aminopeptidase A/I</i>

01	02	03	01C02	XAC3467	02B01	A0EC1093E08	IX	<i>glycosyltransferase</i>
01	03	03	01C03	XAC1281	01B02	A0AC0115E12	V.C	<i>chemotaxis protein</i>
01	04	03	01C04	XAC2450	02B02	A0JJ0901E01	III.A.1	<i>ATP-dependent DNA helicase</i>
01	05	03	01C05	XAC2170	01B03	A0AC0113C04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	06	03	01C06	XAC3719	02B03	A0AM0909B02	I.D.2	<i>transcriptional regulator protein</i>
01	07	03	01C07	XAC2831	01B04	A0AC0113C05	III.C.3	<i>extracellular serine protease</i>
01	08	03	01C08	XAC3109	02B04	A0EC1091D09	IV.A.1	<i>penicillin-binding protein 1B</i>
01	09	03	01C09	XAC1297	01B05	A0AC0113B03	VIII.A	<i>conserved hypothetical protein</i>
01	10	03	01C10	XAC0497	02B05	A0JJ1056C05	VIII.A	<i>conserved hypothetical protein</i>
01	11	03	01C11	XAC4058	01B06	A0AC0113D07	I.A.1	<i>beta-xylosidase</i>
01	12	03	01C12	XAC4354	02B06	A0JJ1037C10	V.A.1	<i>amino acid transporter</i>
02	01	03	02C01	XAC1349	05B01	A0QR5106A11	III.C.3	<i>serine protease</i>
02	02	03	02C02	XAC2968	06B01	A0QR5317D10	VIII.A	<i>conserved hypothetical protein</i>
02	03	03	02C03	XAC0610	05B02	A0CE1624D01	I.D.3	<i>histidine kinase/response regulator hybrid protein</i>
02	04	03	02C04	XAC0144	06B02	A0UV5502G02	V.A.7	<i>TonB-dependent receptor</i>
02	05	03	02C05	XAC0248	05B03	A0RN1630E07	II.A.2	<i>asparaginase</i>
02	06	03	02C06	XAC3221	06B03	A0UV5502H11	VI.C	<i>ISxac3 transposase</i>
02	07	03	02C07	XAC3052	05B04	A0QR5205G06	I.D.2	<i>transcriptional regulator</i>
02	08	03	02C08	XAC3203	06B04	A0QR5508C08	II.D.10	<i>glutathione transferase</i>
02	09	03	02C09	XAC3185	05B05	A0QR5205H07	VIII.A	<i>conserved hypothetical protein</i>
02	10	03	02C10	XAC2462	06B05	A0QR5408H09	IX	<i>ATP-binding protein</i>
02	11	03	02C11	XAC2914	05B06	A0UV5307E08	VIII.A	<i>conserved hypothetical protein</i>
02	12	03	02C12	XAC1217	06B06	A0QR5530C03	III.C.3	<i>dipeptidyl carboxypeptidase</i>
03	01	03	03C01	XAC0676	09B01	A0QH6309B08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	02	03	03C02	XAC1720	10B01	A0UE6379E04	VIII.A	<i>conserved hypothetical protein</i>
03	03	03	03C03	XAC2978	09B02	A0CE6345B10	V.A.3	<i>phosphotransferase system HPr enzyme</i>
03	04	03	03C04	XAC0268	10B02	A0QR6375E10	VIII.A	<i>conserved hypothetical protein</i>
03	05	03	03C05	XAC2281	09B03	A0CE6345C01	III.A.4	<i>RadC family protein</i>
03	06	03	03C06	XAC4253	10B03	A0QH6380C04	VIII.A	<i>conserved hypothetical protein</i>
03	07	03	03C07	XAC2839	09B04	A0CE6345D04	I.D.2	<i>transcriptional regulator</i>
03	08	03	03C08	XAC4131	10B04	A0QR6377E09	VIII.A	<i>conserved hypothetical protein</i>
03	09	03	03C09	XAC3178	09B05	A0CE6344C10	VIII.A	<i>conserved hypothetical protein</i>
03	10	03	03C10	XAC2578	10B05	A0QH6381H05	VII.E	<i>GumI protein</i>
03	11	03	03C11	XAC2661	09B06	A0CE6344D09	VI.C	<i>ISxac3 transposase</i>
03	12	03	03C12	XAC1985	10B06	A0QR6392D03	V.C	<i>flagellar biosynthesis, cell-proximal portion of basal-body rod</i>
04	01	03	04C01	XAC1105	13B01	A0QR6614D04	VIII.B	<i>hypothetical protein</i>
04	02	03	04C02	XAC1482	14B01	A0UV6711D08	VII.C	<i>RND multidrug efflux membrane fusion protein</i>
04	03	03	04C03	XAC3875	13B02	A0QR6614E04	III.B.4	<i>ribonuclease BN</i>
04	04	03	04C04	XAC3724	14B02	A0UV6710F02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	05	03	04C05	XAC3086	13B03	A0UV6708B09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	06	03	04C06	XAC2431	14B03	A0UV6710B03	VI.C	<i>transposase</i>
04	07	03	04C07	XAC3449	13B04	A0UV6708D12	V.C	<i>chemotaxis protein</i>
04	08	03	04C08	XAC0852	14B04	A0UV6709D10	VII.C	<i>TonB-dependent receptor</i>
04	09	03	04C09	XAC3259	13B05	A0UV6708C03	VIII.B	<i>hypothetical protein</i>
04	10	03	04C10	XAC1914	14B05	A0UV6711G06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	11	03	04C11	XAC0910	13B06	A0UV6708B01	VIII.A	<i>conserved hypothetical protein</i>
04	12	03	04C12	XAC3381	14B06	A0UV6711A03	IV.D	<i>fimbrial assembly protein</i>
05	01	03	05C01	XAC2135	17B01	A0UV6733E11	VIII.B	<i>hypothetical protein</i>
05	02	03	05C02	XAC2071	18B01	A0CE6742C08	I.C.5	<i>glucose-6-phosphate 1-dehydrogenase</i>
05	03	03	05C03	XAC1451	17B02	A0UV6734A09	VIII.A	<i>conserved hypothetical protein</i>
05	04	03	05C04	XAC3863	18B02	A0CE6742G04	VIII.A	<i>conserved hypothetical protein</i>
05	05	03	05C05	XAC3996	17B03	A0QR6737A06	V.A.7	<i>ABC transporter ATP-binding protein</i>
05	06	03	05C06	XAC3810	18B03	A0CE6743D02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	07	03	05C07	XAC2666	17B04	A0QR6737A09	IV.D	<i>PilX protein</i>
05	08	03	05C08	XAC1302	18B04	A0CE6740E03	VIII.A	<i>conserved hypothetical protein</i>

05	09	03	05C09	XAC1844	17B05	A0UV6738G09	II.A.3	<i>D-3-phosphoglycerate dehydrogenase</i>
05	10	03	05C10	XAC1467	18B05	A0CE6742E05	II.D.4	<i>molybdopterin biosynthesis protein</i>
05	11	03	05C11	XAC3678	17B06	A0UV6738H02	III.B.3	<i>ribosomal small subunit pseudouridylate synthase</i>
05	12	03	05C12	XAC3983	18B06	A0CE6746A11	VIII.A	<i>conserved hypothetical protein</i>
06	01	03	06C01	XAC3657	21B01	A0QH6811G03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	02	03	06C02	XAC1479	22B01	A0UV6819D05	IV.A.2	<i>OmpA family protein</i>
06	03	03	06C03	XAC2935	21B02	A0QH6811D03	V.A.7	<i>ABC transporter permease</i>
06	04	03	06C04	XAC1242	22B02	A0UV6820C12	VII.H	<i>pathogenicity-related protein</i>
06	05	03	06C05	XAC3888	21B03	A0UT6815A09	I.C.1	<i>cytochrome C oxidase, polypeptide II</i>
06	06	03	06C06	XAC0636	22B03	A0UV6820D02	III.A.3	<i>site-specific recombinase</i>
06	07	03	06C07	XAC0305	21B04	A0UT6815E09	II.D.10	<i>gamma-glutamyltranspeptidase</i>
06	08	03	06C08	XAC3428	22B04	A0UV6821G12	IX	<i>hydrolase</i>
06	09	03	06C09	XAC1069	21B05	A0UT6817A11	VIII.B	<i>hypothetical protein</i>
06	10	03	06C10	XAC0678	22B05	A0UV6821H03	VIII.A	<i>conserved hypothetical protein</i>
06	11	03	06C11	XAC0750	21B06	A0UT6817E08	II.D.9	<i>6,7-dimethyl-8-ribityllumazine synthase</i>
06	12	03	06C12	XAC2053	22B06	A0AC6827A11	I.D.2	<i>transcription-related protein</i>
07	01	03	07C01	XAC3278	25B01	A0CE9728G11	VIII.B	<i>hypothetical protein</i>
07	02	03	07C02	XAC0342	26B01	A0AC9832C06	VI.C	<i>ISxac3 transposase</i>
07	03	03	07C03	XAC1707	25B02	A0QR9739G04	VII.G	<i>general stress protein</i>
07	04	03	07C04	XAC4165	26B02	A0UV9805C01	III.A.1	<i>glucose inhibited division protein B</i>
07	05	03	07C05	XAC0741	25B03	A0QR9739G05	V.A.7	<i>ABC transporter ATP-binding protein</i>
07	06	03	07C06	XAC1347	26B03	A0QR9815A07	IV.A.2	<i>outer membrane protein</i>
07	07	03	07C07	XAC3775	25B04	A0UV9746D12	VIII.B	<i>hypothetical protein</i>
07	08	03	07C08	XAC2473	26B04	A0UV9819H06	VII.C	<i>outer membrane protein OprN precursor</i>
07	09	03	07C09	XAC1509	25B05	A0UT9742C09	VIII.A	<i>conserved hypothetical protein</i>
07	10	03	07C10	XAC1314	26B05	A0UV9819A08	I.A.3	<i>enoyl-CoA hydratase</i>
07	11	03	07C11	XAC2388	25B06	A0UV9750E01	VIII.A	<i>conserved hypothetical protein</i>
07	12	03	07C12	XAC2648	26B06	A0QR9815H11	VI.A	<i>phage-related protein</i>
08	01	03	08C01	XAC2150	29B01	genomic_DNA	I.D.1	<i>nodulation protein</i>
08	02	03	08C02	XAC1266	30B01	genomic_DNA	VII.B	<i>HrpX protein</i>
08	03	03	08C03	XAC1282	29B02	genomic_DNA	I.D.1	<i>two-component system, sensor protein</i>
08	04	03	08C04	XAC0095	30B02	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	05	03	08C05	XAC3285	29B03	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	06	03	08C06	XAC0279	30B03	genomic_DNA	VII.B	<i>HrcQ protein</i>
08	07	03	08C07	XAC2613	29B04	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	08	03	08C08	XAC0028	30B04	genomic_DNA	VII.B	<i>HrpB7 protein</i>
08	09	03	08C09	XACb0042	29B05	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	10	03	08C10	XAC1864	30B05	genomic_DNA	VII.H	<i>regulatory protein</i>
08	11	03	08C11	XAC2609	29B06	genomic_DNA	IV.B	<i>carboxypeptidase</i>
08	12	03	08C12	no_clone	30B06	no_clone	NC	<i>no_gene</i>
09	01	03	09A05	Score_card	33A03	Score_card	SC	<i>Score_card</i>
09	02	03	09E21	Score_card	33C11	Score_card	SC	<i>Score_card</i>
09	03	03	09K13	Score_card	33F07	Score_card	SC	<i>Score_card</i>
09	04	03	09A06	Score_card	34A03	Score_card	SC	<i>Score_card</i>
09	05	03	09E22	Score_card	34C11	Score_card	SC	<i>Score_card</i>
09	06	03	09K14	Score_card	34F07	Score_card	SC	<i>Score_card</i>
09	07	03	09B06	Score_card	35A03	Score_card	SC	<i>Score_card</i>
09	08	03	09F22	Score_card	35C11	Score_card	SC	<i>Score_card</i>
09	09	03	09L14	Score_card	35F07	Score_card	SC	<i>Score_card</i>
09	10	03	09B05	Score_card	36A03	Score_card	SC	<i>Score_card</i>
09	11	03	09F21	Score_card	36C11	Score_card	SC	<i>Score_card</i>
09	12	03	09L13	Score_card	36F07	Score_card	SC	<i>Score_card</i>
01	01	30	01N13	XAC1313	04G07	A0JJ1537D09	I.A.3	<i>acyl-CoA dehydrogenase</i>
01	02	30	01N14	XAC3576	03G07	A0JJ1329D09	VIII.A	<i>conserved hypothetical protein</i>
01	03	30	01N15	XAC3962	04G08	A0UV1534F12	VIII.B	<i>hypothetical protein</i>

01	04	30	01N16	XAC2041	03G08	A0AC1317A11	I.B.3	<i>phosphoenolpyruvate synthase</i>
01	05	30	01N17	XAC4256	04G09	A0RN1630G06	V.A.7	<i>TonB-dependent receptor</i>
01	06	30	01N18	XAC1250	03G09	A0JJ1421C01	IX	<i>GTP-binding protein</i>
01	07	30	01N19	XAC0265	04G10	A0UV1654C03	I.A.3	<i>acyl-CoA dehydrogenase</i>
01	08	30	01N20	XACb0067	03G10	A0AC1387F05	VI.C	<i>Tn5045 transposase</i>
01	09	30	01N21	XAC0817	04G11	A0QR5104G09	VIII.B	<i>hypothetical protein</i>
01	10	30	01N22	XAC1555	03G11	A0UV1358B10	I.D.2	<i>transcriptional regulator</i>
01	11	30	01N23	XAC1680	04G12	A0RN1628B05	III.C.3	<i>serine protease</i>
01	12	30	01N24	XAC2720	03G12	A0AM1488G08	III.B.4	<i>tRNA pseudouridine synthase A</i>
02	01	30	02N13	XAC3016	08G07	A0QH6308B03	III.A.3	<i>RebA protein</i>
02	02	30	02N14	XAC2218	07G07	A0QR5704H06	VIII.A	<i>conserved hypothetical protein</i>
02	03	30	02N15	XAC0028	08G08	A0CE6314A10	VII.D	<i>cellulase</i>
02	04	30	02N16	XAC0492	07G08	A0QR5704G02	I.C.3	<i>bacterioferritin-associated ferredoxin</i>
02	05	30	02N17	XAC1387	08G09	A0QH6309D01	VIII.A	<i>conserved hypothetical protein</i>
02	06	30	02N18	XAC3105	07G09	A0QR5904B07	I.D.4	<i>dinitrogenase reductase activation glycohydrolase</i>
02	07	30	02N19	XAC3332	08G10	A0UV6210G07	I.B.12	<i>3'-phosphoadenosine 5'-phosphosulfate reductase</i>
02	08	30	02N20	XAC0099	07G10	A0QR5904C02	VIII.B	<i>hypothetical protein</i>
02	09	30	02N21	XAC2528	08G11	A0QR6304D02	III.C.2	<i>heat shock protein G</i>
02	10	30	02N22	XAC2604	07G11	A0QR5903A07	VI.C	<i>ISxac4 transposase</i>
02	11	30	02N23	XAC3871	08G12	A0QR6303E04	III.C.3	<i>O-sialoglycoprotein endopeptidase</i>
02	12	30	02N24	XAC4129	07G12	A0QR5902H04	I.D.4	<i>ECF sigma factor</i>
03	01	30	03N13	XAC4017	12G07	A0QR6490H06	I.D.2	<i>transcriptional regulator</i>
03	02	30	03N14	XAC1924	11G07	A0UE6465C06	VI.C	<i>transposase</i>
03	03	30	03N15	XAC0756	12G08	A0QR6702B04	V.A.7	<i>potassium-transporting ATPase, A chain</i>
03	04	30	03N16	XAC2358	11G08	A0UE6465F02	IX	<i>DnaK supressor</i>
03	05	30	03N17	XAC2788	12G09	A0QR6704E05	VIII.A	<i>conserved hypothetical protein</i>
03	06	30	03N18	XAC3417	11G09	A0QR6460D02	VIII.B	<i>hypothetical protein</i>
03	07	30	03N19	XAC3128	12G10	A0UE6612A02	VIII.B	<i>hypothetical protein</i>
03	08	30	03N20	XAC0345	11G10	A0EC6472A01	II.A.2	<i>dihydroxy-acid dehydratase</i>
03	09	30	03N21	XAC0116	12G11	A0QR6613D02	VIII.A	<i>conserved hypothetical protein</i>
03	10	30	03N22	XAC3411	11G11	A0UE6464B05	I.B.6	<i>ribose-5-phosphate isomerase A</i>
03	11	30	03N23	XAC3256	12G12	A0QR6704D08	VII.H	<i>virulence regulator</i>
03	12	30	03N24	XAC3823	11G12	A0UE6464E04	IV.A.1	<i>conserved hypothetical protein</i>
04	01	30	04N13	XAC1305	16G07	A0UV6734G09	IV.A.2	<i>wall-associated protein</i>
04	02	30	04N14	XAC0362	15G07	A0QR6729B08	I.A.2	<i>phenoxybenzoate dioxygenase beta subunit</i>
04	03	30	04N15	XAC1941	16G08	A0UV6735E09	V.C	<i>flagellar biosynthetic protein</i>
04	04	30	04N16	XAC1850	15G08	A0QR6729F01	I.A.3	<i>3-hydroxyacyl-CoA dehydrogenase type II</i>
04	05	30	04N17	XAC4226	16G09	A0UV6735B12	I.D.2	<i>sal operon transcriptional repressor</i>
04	06	30	04N18	XAC1030	15G09	A0QR6729D10	VIII.A	<i>conserved hypothetical protein</i>
04	07	30	04N19	XAC1606	16G10	A0UV6734G02	VIII.B	<i>hypothetical protein</i>
04	08	30	04N20	XAC2018	15G10	A0QR6729F03	VIII.A	<i>conserved hypothetical protein</i>
04	09	30	04N21	XAC1055	16G11	A0UV6734H03	VIII.B	<i>hypothetical protein</i>
04	10	30	04N22	XAC1893	15G11	A0QH6732B10	V.C	<i>chemotaxis protein</i>
04	11	30	04N23	XAC3094	16G12	A0UV6735E02	VIII.B	<i>hypothetical protein</i>
04	12	30	04N24	XAC1182	15G12	A0QH6732F07	II.D.10	<i>thioredoxin reductase</i>
05	01	30	05N13	XAC3585	20G07	A0QR6803G01	IV.C	<i>dTDP-glucose 4,6-dehydratase</i>
05	02	30	05N14	XAC3538	19G07	A0UV6765B09	VII.H	<i>general secretion pathway protein K</i>
05	03	30	05N15	XAC0783	20G08	A0QR6804C07	V.B	<i>cell division protein</i>
05	04	30	05N16	XAC3972	19G08	A0UV6762G08	VIII.A	<i>conserved hypothetical protein</i>
05	05	30	05N17	XAC1975	20G09	A0QH6810E06	V.C	<i>flagellar protein</i>
05	06	30	05N18	XAC0257	19G09	A0UV6762F09	I.B.4	<i>isocitrate lyase</i>
05	07	30	05N19	XAC4072	20G10	A0QH6810E08	II.D.10	<i>thioredoxin</i>
05	08	30	05N20	XAC2321	19G10	A0UV6763D12	IX	<i>hydrolase</i>
05	09	30	05N21	XAC1058	20G11	A0QR6805F07	VIII.B	<i>hypothetical protein</i>
05	10	30	05N22	XAC2755	19G11	A0QR6770B02	VIII.A	<i>conserved hypothetical protein</i>

05	11	30	05N23	XAC3406	20G12	A0QR6805G03	VIII.A	<i>conserved hypothetical protein</i>
05	12	30	05N24	XAC1507	19G12	A0QR6768G03	VI.B	<i>plasmid mobilization protein</i>
06	01	30	06N13	XAC3936	24G07	A0UV6846E09	VI.C	<i>IS1389 transposase</i>
06	02	30	06N14	XAC3914	23G07	A0CE6831F09	III.A.1	<i>DNA/pantothenate metabolism flavoprotein</i>
06	03	30	06N15	XAC3679	24G08	A0UV6846F03	III.B.3	<i>ribosomal RNA small subunit methyltransferase C</i>
06	04	30	06N16	XAC3743	23G08	A0CE6830H01	VIII.A	<i>conserved hypothetical protein</i>
06	05	30	06N17	XAC4200	24G09	A0UV6846F06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	06	30	06N18	XAC3865	23G09	A0CE6831D06	VIII.A	<i>conserved hypothetical protein</i>
06	07	30	06N19	XAC2117	24G10	A0UV9709D02	VIII.A	<i>conserved hypothetical protein</i>
06	08	30	06N20	XAC2285	23G10	A0UV6843F08	VI.A	<i>phage-related protein</i>
06	09	30	06N21	XAC1254	24G11	A0QR9717B01	III.B.4	<i>isoleucyl-tRNA synthetase</i>
06	10	30	06N22	XAC4222	23G11	A0UV6844G10	VIII.A	<i>conserved hypothetical protein</i>
06	11	30	06N23	XAC1075	24G12	A0UV9748F06	I.D.3	<i>sensor histidine kinase</i>
06	12	30	06N24	XAC2594	23G12	A0QR6840D03	III.B.4	<i>threonyl-tRNA synthetase</i>
07	01	30	07N13	no_clone	28G07	no_clone	NC	<i>no_gene</i>
07	02	30	07N14	XAC3437	27G07	A0UT9877E06	II.B.1	<i>adenylate kinase</i>
07	03	30	07N15	no_clone	28G08	no_clone	NC	<i>no_gene</i>
07	04	30	07N16	XAC2277	27G08	A0UT9876C09	I.A.2	<i>3-hydroxyisobutirate dehydrogenase</i>
07	05	30	07N17	no_clone	28G09	no_clone	NC	<i>no_gene</i>
07	06	30	07N18	XAC0539	27G09	A0UT9876D05	I.C.3	<i>oxidoreductase</i>
07	07	30	07N19	no_clone	28G10	no_clone	NC	<i>no_gene</i>
07	08	30	07N20	XAC0992	27G10	A0UT9875G10	V.A.6	<i>preprotein translocase SecY subunit</i>
07	09	30	07N21	no_clone	28G11	no_clone	NC	<i>no_gene</i>
07	10	30	07N22	XAC0592	27G11	A0UV9872B12	VIII.A	<i>conserved hypothetical protein</i>
07	11	30	07N23	no_clone	28G12	no_clone	NC	<i>no_gene</i>
07	12	30	07N24	XAC1437	27G12	A0UT9878G03	VII.C	<i>penicillin acylase II</i>
08	01	30	08N13	no_clone	32G07	no_clone	NC	<i>no_gene</i>
08	02	30	08N14	no_clone	31G07	no_clone	NC	<i>no_gene</i>
08	03	30	08N15	no_clone	32G08	no_clone	NC	<i>no_gene</i>
08	04	30	08N16	no_clone	31G08	no_clone	NC	<i>no_gene</i>
08	05	30	08N17	no_clone	32G09	no_clone	NC	<i>no_gene</i>
08	06	30	08N18	no_clone	31G09	no_clone	NC	<i>no_gene</i>
08	07	30	08N19	no_clone	32G10	no_clone	NC	<i>no_gene</i>
08	08	30	08N20	no_clone	31G10	no_clone	NC	<i>no_gene</i>
08	09	30	08N21	no_clone	32G11	no_clone	NC	<i>no_gene</i>
08	10	30	08N22	no_clone	31G11	no_clone	NC	<i>no_gene</i>
08	11	30	08N23	no_clone	32G12	no_clone	NC	<i>no_gene</i>
08	12	30	08N24	no_clone	31G12	no_clone	NC	<i>no_gene</i>
09	01	30	09E11	Score_card	33C06	Score_card	SC	<i>Score_card</i>
09	02	30	09K03	Score_card	33F02	Score_card	SC	<i>Score_card</i>
09	03	30	09O19	Score_card	33H10	Score_card	SC	<i>Score_card</i>
09	04	30	09E12	Score_card	34C06	Score_card	SC	<i>Score_card</i>
09	05	30	09K04	Score_card	34F02	Score_card	SC	<i>Score_card</i>
09	06	30	09O20	Score_card	34H10	Score_card	SC	<i>Score_card</i>
09	07	30	09F12	Score_card	35C06	Score_card	SC	<i>Score_card</i>
09	08	30	09L04	Score_card	35F02	Score_card	SC	<i>Score_card</i>
09	09	30	09P20	Score_card	35H10	Score_card	SC	<i>Score_card</i>
09	10	30	09F11	Score_card	36C06	Score_card	SC	<i>Score_card</i>
09	11	30	09L03	Score_card	36F02	Score_card	SC	<i>Score_card</i>
09	12	30	09P19	Score_card	36H10	Score_card	SC	<i>Score_card</i>
01	01	31	01O13	XAC3868	01H07	A0CE0105C06	I.C.3	<i>dehydrogenase</i>
01	02	31	01O14	XAC4121	02H07	A0AM1111C09	VIII.A	<i>conserved hypothetical protein</i>
01	03	31	01O15	XAC2456	01H08	A0JJ0102E05	I.C.1	<i>ubiquinol cytochrome C oxidoreductase, cytochrome B subunit</i>
01	04	31	01O16	XAC3405	02H08	A0RN1116D05	III.C.3	<i>aminopeptidase P</i>
01	05	31	01O17	XAC3204	01H09	A0AC0115H02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>

01	06	31	01O18	XAC2261	02H09	A0AM1111F12	VI.B	<i>plasmid-related protein</i>
01	07	31	01O19	XAC4273	01H10	A0AR1060F11	VII.F	<i>OmpA-related protein</i>
01	08	31	01O20	XAC1185	02H10	A0JJ1118F10	VIII.A	<i>conserved hypothetical protein</i>
01	09	31	01O21	XAC2048	01H11	A0JE1039A04	II.E	<i>poly (3-hydroxybutyric acid) synthase</i>
01	10	31	01O22	XAC0527	02H11	A0JE1147G09	VIII.B	<i>hypothetical protein</i>
01	11	31	01O23	XAC0439	01H12	A0JJ1073F04	V.A.7	<i>cation efflux system protein</i>
01	12	31	01O24	XAC1633	02H12	A0JJ1131A01	I.C.5	<i>glucose dehydrogenase</i>
02	01	31	02O13	XAC0587	05H07	A0QR5321D02	VIII.A	<i>conserved hypothetical protein</i>
02	02	31	02O14	XAC0114	06H07	A0UV5526B08	VIII.A	<i>conserved hypothetical protein</i>
02	03	31	02O15	XAC0192	05H08	A0QR5321E04	V.B	<i>partition protein</i>
02	04	31	02O16	XAC0213	06H08	A0UV5523B02	I.A.3	<i>acyl-CoA thiolase</i>
02	05	31	02O17	XAC4296	05H09	A0UT5316D03	IX	<i>transglycolase/epimerase</i>
02	06	31	02O18	XAC0532	06H09	A0QR5702F05	II.E	<i>biotin carboxyl carrier protein of acetyl-CoA carboxylase</i>
02	07	31	02O19	XAC0121	05H10	A0CE5319H08	III.C.3	<i>TldD protein</i>
02	08	31	02O20	XAC2695	06H10	A0QR5704C06	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO10 subunit</i>
02	09	31	02O21	XAC0510	05H11	A0QR5407A08	VIII.A	<i>conserved hypothetical protein</i>
02	10	31	02O22	XAC1740	06H11	A0QR5704A08	III.A.3	<i>RecA protein</i>
02	11	31	02O23	XAC2514	05H12	A0QR5407F08	III.B.4	<i>S-adenosylmethionine:tRNA ribosyltransferase-isomerase</i>
02	12	31	02O24	XAC1271	06H12	A0QR5704A09	I.D.2	<i>sigma-B negative effector</i>
03	01	31	03O13	XAC2598	09H07	A0EC6323B03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	02	31	03O14	XAC0425	10H07	A0QR6389B08	III.D.1	<i>glycogen synthase</i>
03	03	31	03O15	XAC0659	09H08	A0AC6358F03	IV.B	<i>penicillin-binding protein 2</i>
03	04	31	03O16	XAC3403	10H08	A0QR6391F08	III.C.3	<i>proline dipeptidase</i>
03	05	31	03O17	XAC2297	09H09	A0QR6374G07	III.A.3	<i>integration host factor, beta subunit</i>
03	06	31	03O18	XAC1039	10H09	A0UE6398F12	I.B.9	<i>exopolyphosphatase</i>
03	07	31	03O19	XAC1973	09H10	A0QR6366G09	V.C	<i>flagellar protein</i>
03	08	31	03O20	XAC2946	10H10	A0UE6399A01	VIII.A	<i>conserved hypothetical protein</i>
03	09	31	03O21	XAC0267	09H11	A0QR6375D12	VIII.A	<i>conserved hypothetical protein</i>
03	10	31	03O22	XAC3557	10H11	A0QH6412B10	VIII.B	<i>hypothetical protein</i>
03	11	31	03O23	XAC2741	09H12	A0QR6375E06	VIII.A	<i>conserved hypothetical protein</i>
03	12	31	03O24	XAC3521	10H12	A0QR6437A01	II.D.7	<i>nicotinate phosphoribosyltransferase</i>
04	01	31	04O13	XAC3645	13H07	A0QR6705E12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	02	31	04O14	XAC3375	14H07	A0UV6712F12	VIII.A	<i>conserved hypothetical protein</i>
04	03	31	04O15	XAC3260	13H08	A0QR6706B01	VI.B	<i>plasmid mobilization protein</i>
04	04	31	04O16	XAC1928	14H08	A0UE6718E11	VIII.B	<i>hypothetical protein</i>
04	05	31	04O17	XAC0795	13H09	A0UV6708B07	III.C.3	<i>protease</i>
04	06	31	04O18	XAC1879	14H09	A0QH6713H04	VII.H	<i>RpfF protein</i>
04	07	31	04O19	XAC3605	13H10	A0UV6711A09	IV.A.2	<i>outer membrane protein</i>
04	08	31	04O20	XAC2585	14H10	A0UV6712G11	VII.E	<i>GumB protein</i>
04	09	31	04O21	XAC1246	13H11	A0UV6710H02	VIII.A	<i>conserved hypothetical protein</i>
04	10	31	04O22	XAC3913	14H11	A0UV6712G12	II.B.4	<i>dUTPase</i>
04	11	31	04O23	XAC3747	13H12	A0UV6710E07	I.C.3	<i>alcohol dehydrogenase (Zn-dependent)</i>
04	12	31	04O24	XAC1684	14H12	A0UV6712F06	I.C.3	<i>cytochrome C2</i>
05	01	31	05O13	XACa0012	17H07	A0CE6740A04	VIII.B	<i>hypothetical protein</i>
05	02	31	05O14	XAC1786	18H07	A0CE6746F05	II.D.5	<i>pantoate-beta-alanine ligase</i>
05	03	31	05O15	XAC3033	17H08	A0UV6739B05	I.C.3	<i>cytochrome B561</i>
05	04	31	05O16	XAC3798	18H08	A0QR6752G11	VIII.A	<i>conserved hypothetical protein</i>
05	05	31	05O17	XAC4213	17H09	A0UV6739E10	II.A.4	<i>shikimate kinase</i>
05	06	31	05O18	XAC2025	18H09	A0QR6749G11	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	07	31	05O19	XAC2331	17H10	A0UV6739G08	I.C.3	<i>C-type cytochrome biogenesis protein</i>
05	08	31	05O20	XAC2733	18H10	A0UE6757A02	IV.A.1	<i>integral membrane protein</i>
05	09	31	05O21	XAC1861	17H11	A0CE6740C06	II.B.2	<i>carbamoyl-phosphate synthase small chain</i>
05	10	31	05O22	XAC1747	18H11	A0UE6757E06	VIII.A	<i>conserved hypothetical protein</i>
05	11	31	05O23	XAC4020	17H12	A0CE6740D01	VIII.A	<i>conserved hypothetical protein</i>
05	12	31	05O24	XAC3618	18H12	A0UE6759F06	VIII.A	<i>conserved hypothetical protein</i>

06	01	31	06O13	XAC3584	21H07	A0UT6814G10	IV.A.1	<i>glucose-1-phosphate thymidyltransferase</i>
06	02	31	06O14	XAC1729	22H07	A0AC6824F10	VIII.A	<i>conserved hypothetical protein</i>
06	03	31	06O15	XAC0254	21H08	A0UT6815A01	V.A.4	<i>Na⁺/H⁺-exchanging protein</i>
06	04	31	06O16	XAC0081	22H08	A0AC6823A06	VIII.A	<i>conserved hypothetical protein</i>
06	05	31	06O17	XAC4344	21H09	A0UV6819F06	III.D.3	<i>lipoprotein</i>
06	06	31	06O18	XACb0054	22H09	A0AC6827F05	VI.B	<i>partition protein A</i>
06	07	31	06O19	XAC3394	21H10	A0UV6819F08	III.B.5	<i>RNA polymerase omega subunit</i>
06	08	31	06O20	XAC1006	22H10	A0AC6825E05	I.C.7	<i>malate dehydrogenase</i>
06	09	31	06O21	XAC3861	21H11	A0UV6819C06	VIII.A	<i>conserved hypothetical protein</i>
06	10	31	06O22	XAC0183	22H11	A0AC6824C07	V.A.1	<i>ABC transporter amino acid permease</i>
06	11	31	06O23	XAC2860	21H12	A0UV6821G02	VIII.B	<i>hypothetical protein</i>
06	12	31	06O24	XAC2609	22H12	A0CE6828E07	IV.B	<i>carboxypeptidase</i>
07	01	31	07O13	XAC2270	25H07	A0AC9763E08	VIII.B	<i>hypothetical protein</i>
07	02	31	07O14	XAC0890	26H07	A0CE9842E01	VIII.A	<i>conserved hypothetical protein</i>
07	03	31	07O15	XAC0572	25H08	A0QR9758C10	IX	<i>lcfG protein</i>
07	04	31	07O16	XAC0895	26H08	A0CE9841H03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	05	31	07O17	XAC0441	25H09	A0AC9764C06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	06	31	07O18	XAC1401	26H09	A0CE9839E07	VIII.A	<i>conserved hypothetical protein</i>
07	07	31	07O19	XAC0586	25H10	A0QR9767H07	VIII.A	<i>conserved hypothetical protein</i>
07	08	31	07O20	XAC1037	26H10	A0QR9860E02	VIII.A	<i>conserved hypothetical protein</i>
07	09	31	07O21	XAC1376	25H11	A0QR9759D12	VIII.A	<i>conserved hypothetical protein</i>
07	10	31	07O22	XAC2652	26H11	A0UE9854D02	VI.A	<i>phage-related tail protein</i>
07	11	31	07O23	XAC0427	25H12	A0QR9825E08	I.A.1	<i>maltooligosyltrehalose trehalohydrolase</i>
07	12	31	07O24	XAC1024	26H12	A0UT9850B07	III.D.2	<i>non-hemolytic phospholipase C</i>
08	01	31	08O13	XAC1682	29H07	genomic_DNA	I.D.4	<i>RNA polymerase sigma-E factor</i>
08	02	31	08O14	no_clone	30H07	no_clone	NC	<i>no_gene</i>
08	03	31	08O15	XAC2618	29H08	genomic_DNA	VII.H	<i>VirB11 protein</i>
08	04	31	08O16	no_clone	30H08	no_clone	NC	<i>no_gene</i>
08	05	31	08O17	XAC1742	29H09	genomic_DNA	III.B.4	<i>alanyl-tRNA synthetase</i>
08	06	31	08O18	no_clone	30H09	no_clone	NC	<i>no_gene</i>
08	07	31	08O19	XAC0570	29H10	genomic_DNA	I.D.4	<i>anti-sigma F factor antagonist</i>
08	08	31	08O20	no_clone	30H10	no_clone	NC	<i>no_gene</i>
08	09	31	08O21	no_clone	29H11	no_clone	NC	<i>no_gene</i>
08	10	31	08O22	no_clone	30H11	no_clone	NC	<i>no_gene</i>
08	11	31	08O23	no_clone	29H12	no_clone	NC	<i>no_gene</i>
08	12	31	08O24	no_clone	30H12	no_clone	NC	<i>no_gene</i>
09	01	31	09E13	Score_card	33C07	Score_card	SC	<i>Score_card</i>
09	02	31	09K05	Score_card	33F03	Score_card	SC	<i>Score_card</i>
09	03	31	09O21	Score_card	33H11	Score_card	SC	<i>Score_card</i>
09	04	31	09E14	Score_card	34C07	Score_card	SC	<i>Score_card</i>
09	05	31	09K06	Score_card	34F03	Score_card	SC	<i>Score_card</i>
09	06	31	09O22	Score_card	34H11	Score_card	SC	<i>Score_card</i>
09	07	31	09F14	Score_card	35C07	Score_card	SC	<i>Score_card</i>
09	08	31	09L06	Score_card	35F03	Score_card	SC	<i>Score_card</i>
09	09	31	09P22	Score_card	35H11	Score_card	SC	<i>Score_card</i>
09	10	31	09F13	Score_card	36C07	Score_card	SC	<i>Score_card</i>
09	11	31	09L05	Score_card	36F03	Score_card	SC	<i>Score_card</i>
09	12	31	09P21	Score_card	36H11	Score_card	SC	<i>Score_card</i>
01	01	32	01P13	XAC3757	04H07	A0AM1542C08	VIII.A	<i>conserved hypothetical protein</i>
01	02	32	01P14	XAC1627	03H07	A0AC1335D07	III.A.1	<i>DNA ligase</i>
01	03	32	01P15	XAC2324	04H08	A0AM1540C10	V.A.6	<i>ABC transporter heme permease</i>
01	04	32	01P16	XAC0205	03H08	A0JE1331G02	I.D.2	<i>nitrogen regulatory protein P-II</i>
01	05	32	01P17	XAC2858	04H09	A0JE1637F04	III.A.4	<i>transcription-repair coupling factor</i>
01	06	32	01P18	XAC0288	03H09	A0AC1317G08	I.C.3	<i>oxidoreductase</i>
01	07	32	01P19	XAC2554	04H10	A0QR5104D09	VIII.A	<i>conserved hypothetical protein</i>

01	08	32	01P20	XAC0437	03H10	A0AM1393H03	I.D.2	<i>transcriptional regulator tetR/acrR family</i>
01	09	32	01P21	XAC3152	04H11	A0JJ1599G11	VIII.A	<i>conserved hypothetical protein</i>
01	10	32	01P22	XAC3189	03H11	A0AM1365C09	II.D.14	<i>cobalamin biosynthetic protein</i>
01	11	32	01P23	XAC0422	04H12	A0RN1632F03	V.A.7	<i>ABC transporter substrate binding protein</i>
01	12	32	01P24	XAC2654	03H12	A0RN1498F11	VIII.A	<i>conserved hypothetical protein</i>
02	01	32	02P13	XAC1382	08H07	A0QH6313B06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	02	32	02P14	XAC3700	07H07	A0QR5902C04	V.A.1	<i>ABC transporter ATP-binding protein</i>
02	03	32	02P15	XAC1321	08H08	A0EC6316B12	III.C.3	<i>periplasmic protease</i>
02	04	32	02P16	XAC1329	07H08	A0QR5902A01	III.B.3	<i>RNA methyltransferase</i>
02	05	32	02P17	XAC1518	08H09	A0EC6316C07	VIII.A	<i>conserved hypothetical protein</i>
02	06	32	02P18	XAC1974	07H09	A0QR5904G07	V.C	<i>flagellar protein</i>
02	07	32	02P19	XAC2637	08H10	A0QR6303D04	VIII.B	<i>hypothetical protein</i>
02	08	32	02P20	XAC1728	07H10	A0QR5904G08	III.D.3	<i>lipoprotein</i>
02	09	32	02P21	XAC4109	08H11	A0QH6306E03	II.D.12	<i>coproporphyrinogen III oxidase, aerobic</i>
02	10	32	02P22	XAC3982	07H11	A0QR5903D09	VIII.A	<i>conserved hypothetical protein</i>
02	11	32	02P23	XAC1929	08H12	A0QR6304F02	VI.C	<i>ISxac1 transposase</i>
02	12	32	02P24	XAC0647	07H12	A0QR5903A09	VIII.A	<i>conserved hypothetical protein</i>
03	01	32	03P13	XAC1286	12H07	A0AC6496B06	I.A.1	<i>alpha-L-arabinosidase</i>
03	02	32	03P14	XAC1768	11H07	A0EC6475B01	V.A.7	<i>TonB-dependent receptor</i>
03	03	32	03P15	XACb0030	12H08	A0QR6702F05	VI.B	<i>TrwB protein</i>
03	04	32	03P16	XAC2958	11H08	A0EC6473A09	VIII.A	<i>conserved hypothetical protein</i>
03	05	32	03P17	XAC3095	12H09	A0QR6603B06	III.B.3	<i>ribosomal protein alanine acetyltransferase</i>
03	06	32	03P18	XAC3277	11H09	A0UE6463D09	VIII.B	<i>hypothetical protein</i>
03	07	32	03P19	XAC4052	12H10	A0QR6613C01	VII.C	<i>TonB-like protein</i>
03	08	32	03P20	XAC2734	11H10	A0EC6474F01	III.B.5	<i>transcription elongation factor and transcript cleavage</i>
03	09	32	03P21	XAC0063	12H11	A0QR6614C06	II.D.17	<i>aryl sulfotransferase</i>
03	10	32	03P22	XAC2671	11H11	A0EC6473D05	VIII.A	<i>conserved hypothetical protein</i>
03	11	32	03P23	XAC1792	12H12	A0QR6704E10	I.A.2	<i>alkaline phosphatase</i>
03	12	32	03P24	XAC3092	11H12	A0EC6473G01	III.D.1	<i>asparaginase</i>
04	01	32	04P13	XAC3506	16H07	A0UV6734H11	VII.D	<i>cellulase S (truncated)</i>
04	02	32	04P14	XAC2758	15H07	A0QR6729E12	V.A.1	<i>glutamate symport protein</i>
04	03	32	04P15	XAC1938	16H08	A0UV6736A05	I.D.4	<i>GGDEF family protein</i>
04	04	32	04P16	XAC0745	15H08	A0QR6729G04	IV.C	<i>acetyltransferase</i>
04	05	32	04P17	XACb0038	16H09	A0UV6735F02	VII.H	<i>VirB10 protein</i>
04	06	32	04P18	XAC1501	15H09	A0QR6729F02	VIII.B	<i>hypothetical protein</i>
04	07	32	04P19	XAC2371	16H10	A0UV6735A05	VI.C	<i>IS1479 transposase</i>
04	08	32	04P20	XACb0041	15H10	A0QR6730A05	VII.H	<i>VirB6 protein</i>
04	09	32	04P21	XAC4065	16H11	A0UV6736A08	V.A.7	<i>ABC transporter ATP-binding protein</i>
04	10	32	04P22	XAC0691	15H11	A0QH6732D05	VIII.A	<i>conserved hypothetical protein</i>
04	11	32	04P23	XAC2214	16H12	A0UV6736A10	VI.A	<i>phage-related protein</i>
04	12	32	04P24	XAC0549	15H12	A0UV6733A07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
05	01	32	05P13	XAC4232	20H07	A0QR6804A07	I.A.2	<i>mannitol dehydrogenase</i>
05	02	32	05P14	XAC1863	19H07	A0UV6765E06	III.B.5	<i>transcriptional elongation factor</i>
05	03	32	05P15	XAC4204	20H08	A0QR6804E08	VIII.A	<i>conserved hypothetical protein</i>
05	04	32	05P16	XAC3959	19H08	A0UV6763A08	VIII.A	<i>conserved hypothetical protein</i>
05	05	32	05P17	XAC1911	20H09	A0QH6811B11	VIII.B	<i>hypothetical protein</i>
05	06	32	05P18	XAC3199	19H09	A0UV6764B07	I.C.3	<i>oxidoreductase</i>
05	07	32	05P19	XAC3553	20H10	A0QH6810H12	VIII.B	<i>hypothetical protein</i>
05	08	32	05P20	XAC2136	19H10	A0UV6763G07	I.C.3	<i>oxidoreductase</i>
05	09	32	05P21	XAC0329	20H11	A0QH6810B12	IV.A.2	<i>outer membrane protein</i>
05	10	32	05P22	XAC4180	19H11	A0QR6770F02	I.D.1	<i>two-component system, regulatory protein</i>
05	11	32	05P23	XAC0310	20H12	A0QH6810C02	I.A.2	<i>vanillate O-demethylase oxidoreductase</i>
05	12	32	05P24	XAC4313	19H12	A0QR6770B03	VIII.A	<i>conserved hypothetical protein</i>
06	01	32	06P13	XAC0722	24H07	A0UT9703A10	III.C.1	<i>thiol:disulfide interchange protein</i>
06	02	32	06P14	XAC3281	23H07	A0QH6835E02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>

06	03	32	06P15	XAC0212	24H08	A0UT9703B08	VIII.A	<i>conserved hypothetical protein</i>
06	04	32	06P16	XAC1932	23H08	A0CE6831C02	V.C	<i>chemotaxis protein</i>
06	05	32	06P17	XAC1878	24H09	A0QR9705E09	VII.H	<i>RpfC protein</i>
06	06	32	06P18	XAC2923	23H09	A0QH6835F06	IV.D	<i>twitching motility protein</i>
06	07	32	06P19	XAC3985	24H10	A0CE9724E02	V.A.7	<i>transport protein</i>
06	08	32	06P20	XAC0886	23H10	A0UV6843H12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	09	32	06P21	XAC0034	24H11	A0CE9723B02	VIII.A	<i>conserved hypothetical protein</i>
06	10	32	06P22	XAC0197	23H11	A0QR6839C05	IV.C	<i>acetyltransferase</i>
06	11	32	06P23	XAC4360	24H12	A0QR9717C04	I.A.2	<i>glycerate kinase</i>
06	12	32	06P24	XAC0458	23H12	A0UV6843C11	VII.G	<i>PhaF protein</i>
07	01	32	07P13	no_clone	28H07	no_clone	NC	<i>no_gene</i>
07	02	32	07P14	XAC1374	27H07	A0UV9880E06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	32	07P15	no_clone	28H08	no_clone	NC	<i>no_gene</i>
07	04	32	07P16	XAC3878	27H08	A0UV9879H12	IX	<i>disulphide-isomerase</i>
07	05	32	07P17	no_clone	28H09	no_clone	NC	<i>no_gene</i>
07	06	32	07P18	XAC3161	27H09	A0UV9880A05	VIII.B	<i>hypothetical protein</i>
07	07	32	07P19	no_clone	28H10	no_clone	NC	<i>no_gene</i>
07	08	32	07P20	XAC1561	27H10	A0UT9876G07	I.D.2	<i>transcriptional regulator</i>
07	09	32	07P21	no_clone	28H11	no_clone	NC	<i>no_gene</i>
07	10	32	07P22	XAC3628	27H11	A0UT9875H06	II.A.3	<i>cysteine synthase</i>
07	11	32	07P23	no_clone	28H12	no_clone	NC	<i>no_gene</i>
07	12	32	07P24	XAC2103	27H12	A0UV9880B10	III.A.3	<i>DNA recombinase</i>
08	01	32	08P13	no_clone	32H07	no_clone	NC	<i>no_gene</i>
08	02	32	08P14	no_clone	31H07	no_clone	NC	<i>no_gene</i>
08	03	32	08P15	no_clone	32H08	no_clone	NC	<i>no_gene</i>
08	04	32	08P16	no_clone	31H08	no_clone	NC	<i>no_gene</i>
08	05	32	08P17	no_clone	32H09	no_clone	NC	<i>no_gene</i>
08	06	32	08P18	no_clone	31H09	no_clone	NC	<i>no_gene</i>
08	07	32	08P19	no_clone	32H10	no_clone	NC	<i>no_gene</i>
08	08	32	08P20	no_clone	31H10	no_clone	NC	<i>no_gene</i>
08	09	32	08P21	no_clone	32H11	no_clone	NC	<i>no_gene</i>
08	10	32	08P22	no_clone	31H11	no_clone	NC	<i>no_gene</i>
08	11	32	08P23	no_clone	32H12	no_clone	NC	<i>no_gene</i>
08	12	32	08P24	no_clone	31H12	no_clone	NC	<i>no_gene</i>
09	01	32	09E15	Score_card	33C08	Score_card	SC	<i>Score_card</i>
09	02	32	09K07	Score_card	33F04	Score_card	SC	<i>Score_card</i>
09	03	32	09O23	Score_card	33H12	Score_card	SC	<i>Score_card</i>
09	04	32	09E16	Score_card	34C08	Score_card	SC	<i>Score_card</i>
09	05	32	09K08	Score_card	34F04	Score_card	SC	<i>Score_card</i>
09	06	32	09O24	Score_card	34H12	Score_card	SC	<i>Score_card</i>
09	07	32	09F16	Score_card	35C08	Score_card	SC	<i>Score_card</i>
09	08	32	09L08	Score_card	35F04	Score_card	SC	<i>Score_card</i>
09	09	32	09P24	Score_card	35H12	Score_card	SC	<i>Score_card</i>
09	10	32	09F15	Score_card	36C08	Score_card	SC	<i>Score_card</i>
09	11	32	09L07	Score_card	36F04	Score_card	SC	<i>Score_card</i>
09	12	32	09P23	Score_card	36H12	Score_card	SC	<i>Score_card</i>
01	01	04	01D01	XAC1730	04B01	A0UV1535A05	VIII.A	<i>conserved hypothetical protein</i>
01	02	04	01D02	XAC1773	03B01	A0AM1161C01	I.A.1	<i>alpha-xylosidase</i>
01	03	04	01D03	XAC3181	04B02	A0UT1520D03	II.A.2	<i>diaminopimelate decarboxylase</i>
01	04	04	01D04	XAC2065	03B02	A0JE1149F10	V.A.7	<i>transport protein</i>
01	05	04	01D05	XACb0031	04B03	A0RN1459D02	VI.B	<i>TrwC protein</i>
01	06	04	01D06	XAC3808	03B03	A0JE1333F09	VIII.A	<i>conserved hypothetical protein</i>
01	07	04	01D07	XAC2121	04B04	A0AR1557H09	VII.C	<i>O-methyltransferase</i>
01	08	04	01D08	XAC3237	03B04	A0JE1334D10	I.D.1	<i>two-component system, sensor protein</i>
01	09	04	01D09	XAC0812	04B05	A0RN1548A04	I.B.10	<i>phosphoanhydride phosphohydrolase</i>

01	10	04	01D10	XAC0526	03B05	A0RN1324B11	III.B.3	<i>ribosomal protein L11 methyltransferase</i>
01	11	04	01D11	XAC0026	04B06	A0RN1447G12	VIII.A	<i>conserved hypothetical protein</i>
01	12	04	01D12	XAC1785	03B06	A0RN1306A09	II.D.5	<i>3-methyl-2-oxobutanoate hydroxymethyltransferase</i>
02	01	04	02D01	XAC3391	08B01	A0QR5904C11	III.A.3	<i>ATP-dependent DNA helicase</i>
02	02	04	02D02	XAC2184	07B01	A0JE5637F07	VIII.B	<i>hypothetical protein</i>
02	03	04	02D03	XAC3960	08B02	A0QR5904H09	I.C.3	<i>oxidoreductase</i>
02	04	04	02D04	XAC3143	07B02	A0JE5637C09	VII.C	<i>TolA protein</i>
02	05	04	02D05	XAC1856	08B03	A0QR6002C04	VIII.A	<i>conserved hypothetical protein</i>
02	06	04	02D06	XAC1335	07B03	A0JE5637D08	II.B.4	<i>hypoxanthine-guanine phosphoribosyltransferase</i>
02	07	04	02D07	XAC0491	08B04	A0QR6002A09	I.B.7	<i>probable (di)nucleoside polyphosphate hydrolase</i>
02	08	04	02D08	XAC2789	07B04	A0QR5702A11	III.C.1	<i>peptidyl-prolyl cis-trans isomerase</i>
02	09	04	02D09	XAC0635	08B05	A0QR6102H01	VIII.A	<i>conserved hypothetical protein</i>
02	10	04	02D10	XAC2612	07B05	A0QR5903E03	VII.H	<i>VirB6 protein</i>
02	11	04	02D11	XAC3813	08B06	A0UV6206H02	III.C.3	<i>protease IV</i>
02	12	04	02D12	XAC3441	07B06	A0QR5904A08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	01	04	03D01	XAC0634	12B01	A0CE6480F07	II.A.2	<i>diaminopimelate epimerase</i>
03	02	04	03D02	XAC3269	11B01	A0QR6443E06	III.A.4	<i>RadC family protein</i>
03	03	04	03D03	XAC1512	12B02	A0UV6487C09	III.C.3	<i>serine peptidase</i>
03	04	04	03D04	XAC1032	11B02	A0AC6430G09	II.B.1	<i>amidophosphoribosyltransferase</i>
03	05	04	03D05	XAC2051	12B03	A0CE6504G12	I.C.3	<i>oxidoreductase</i>
03	06	04	03D06	XAC2694	11B03	A0QR6418E03	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO11 subunit</i>
03	07	04	03D07	XAC2520	12B04	A0CE6503F05	V.A.7	<i>TonB-dependent receptor</i>
03	08	04	03D08	XAC3275	11B04	A0QR6418E09	VIII.B	<i>hypothetical protein</i>
03	09	04	03D09	XAC2114	12B05	A0CE6505F06	VIII.A	<i>conserved hypothetical protein</i>
03	10	04	03D10	XAC1677	11B05	A0AC6429E07	I.C.3	<i>C-type cytochrome biogenesis protein</i>
03	11	04	03D11	XAC1618	12B06	A0QR6482E05	III.B.4	<i>asparaginyl-tRNA synthetase</i>
03	12	04	03D12	XAC3617	11B06	A0EC6473G06	VIII.A	<i>conserved hypothetical protein</i>
04	01	04	04D01	XAC2407	16B01	A0QR6730D02	VIII.A	<i>conserved hypothetical protein</i>
04	02	04	04D02	XAC2529	15B01	A0UE6718C07	IX	<i>RhsD protein</i>
04	03	04	04D03	XAC4127	16B02	A0QH6732E02	I.D.3	<i>serine/threonine kinase</i>
04	04	04	04D04	XAC2562	15B02	A0UE6718D02	VIII.A	<i>conserved hypothetical protein</i>
04	05	04	04D05	XAC3792	16B03	A0QH6732C05	II.D.9	<i>riboflavin biosynthesis protein</i>
04	06	04	04D06	XAC3902	15B03	A0UE6718A09	III.A.4	<i>exodeoxyribonuclease III</i>
04	07	04	04D07	XAC2469	16B04	A0QH6732C06	I.B.10	<i>succinate-semialdehyde dehydrogenase</i>
04	08	04	04D08	XAC1550	15B04	A0QR6728G09	III.C.1	<i>FKBP-type peptidyl-prolyl cis-trans isomerase (rotamase)</i>
04	09	04	04D09	XAC0709	16B05	A0QH6732A04	III.D.1	<i>N-acetylglucosaminidase</i>
04	10	04	04D10	XAC4027	15B05	A0QR6729E07	VIII.B	<i>hypothetical protein</i>
04	11	04	04D11	XAC1409	16B06	A0QH6731G05	IV.C	<i>UDP-N-acetylglucosamine acyltransferase</i>
04	12	04	04D12	XAC2670	15B06	A0QR6727F06	IX	<i>alginate biosynthesis protein</i>
05	01	04	05D01	XAC0423	20B01	A0QR6802B09	VIII.A	<i>conserved hypothetical protein</i>
05	02	04	05D02	XAC4299	19B01	A0QR6755F11	VIII.A	<i>conserved hypothetical protein</i>
05	03	04	05D03	XAC0588	20B02	A0UV6765H03	VIII.B	<i>hypothetical protein</i>
05	04	04	05D04	XAC1727	19B02	A0UE6757G05	VIII.A	<i>conserved hypothetical protein</i>
05	05	04	05D05	XAC1847	20B03	A0UV6765H07	VIII.A	<i>conserved hypothetical protein</i>
05	06	04	05D06	XAC4101	19B03	A0UE6759G07	II.E	<i>acyl carrier protein</i>
05	07	04	05D07	XAC4039	20B04	A0UV6766A02	VIII.A	<i>conserved hypothetical protein</i>
05	08	04	05D08	XAC4189	19B04	A0UE6759H04	VIII.A	<i>conserved hypothetical protein</i>
05	09	04	05D09	XAC2628	20B05	A0UV6766H06	VI.A	<i>phage-related integrase</i>
05	10	04	05D10	XAC3702	19B05	A0UV6761F09	VIII.B	<i>hypothetical protein</i>
05	11	04	05D11	XAC2249	20B06	A0QR6768C03	VIII.B	<i>hypothetical protein</i>
05	12	04	05D12	XAC2984	19B06	A0UV6764C03	III.C.3	<i>peptidase</i>
06	01	04	06D01	XAC3922	24B01	A0QR6840G09	II.D.15	<i>ATP-dependent serine activating enzyme</i>
06	02	04	06D02	XAC1808	23B01	A0AC6823G01	I.C.3	<i>aldehyde dehydrogenase</i>
06	03	04	06D03	XAC0667	24B02	A0UV6844B10	II.D.3	<i>lipoate biosynthesis protein B</i>
06	04	04	06D04	XAC3583	23B02	A0CE6828B08	IV.C	<i>dTDP-4-dehydrorhamnose 3,5-epimerase</i>

06	05	04	06D05	XAC0017	24B03	A0QR6839E01	VIII.A	<i>conserved hypothetical protein</i>
06	06	04	06D06	XAC2244	23B03	A0AC6827A10	VIII.B	<i>hypothetical protein</i>
06	07	04	06D07	XAC3607	24B04	A0UV6845C08	IV.A.2	<i>type II secretion system protein-like protein</i>
06	08	04	06D08	XAC1860	23B04	A0QH6836F09	II.A.2	<i>dihydrodipicolinate reductase</i>
06	09	04	06D09	XAC4009	24B05	A0UV6846B01	I.A.2	<i>arginase</i>
06	10	04	06D10	XAC3891	23B05	A0CE6829E08	VIII.B	<i>hypothetical protein</i>
06	11	04	06D11	XAC0983	24B06	A0UT9703H01	III.B.2	<i>50S ribosomal protein L24</i>
06	12	04	06D12	XAC3274	23B06	A0QH6835B05	I.D.2	<i>single-domain response regulator</i>
07	01	04	07D01	XAC0285	28B01	A0UV9883F02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	02	04	07D02	XAC2919	27B01	A0QR9863F07	III.A.4	<i>DNA-3-methyladenine glycosylase I</i>
07	03	04	07D03	XAC4339	28B02	A0AC9896A06	VII.G	<i>toluene tolerance protein</i>
07	04	04	07D04	XAC1000	27B02	A0UE9857B11	II.A.4	<i>family II 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthase</i>
07	05	04	07D05	XAC1511	28B03	A0UV9901D05	III.B.1	<i>tmRNA</i>
07	06	04	07D06	XAC0681	27B03	A0UT9848F09	I.C.3	<i>oxidoreductase</i>
07	07	04	07D07	XAC4095	28B04	A0UV9905C04	VIII.A	<i>conserved hypothetical protein</i>
07	08	04	07D08	XAC3205	27B04	A0UT9844G06	VIII.A	<i>conserved hypothetical protein</i>
07	09	04	07D09	XAC0903	28B05	A0UT9894D12	I.D.2	<i>regulator of nucleoside diphosphate kinase</i>
07	10	04	07D10	XAC2699	27B05	A0UV9869E01	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO1 subunit</i>
07	11	04	07D11	XAC1053	28B06	A0UV9900G05	VI.C	<i>ISxac3 transposase</i>
07	12	04	07D12	XAC0846	27B06	A0UV9883G09	I.A.2	<i>FMNH2-dependent monooxygenase</i>
08	01	04	08D01	no_clone	32B01	no_clone	NC	<i>no_gene</i>
08	02	04	08D02	no_clone	31B01	no_clone	NC	<i>no_gene</i>
08	03	04	08D03	no_clone	32B02	no_clone	NC	<i>no_gene</i>
08	04	04	08D04	no_clone	31B02	no_clone	NC	<i>no_gene</i>
08	05	04	08D05	no_clone	32B03	no_clone	NC	<i>no_gene</i>
08	06	04	08D06	no_clone	31B03	no_clone	NC	<i>no_gene</i>
08	07	04	08D07	no_clone	32B04	no_clone	NC	<i>no_gene</i>
08	08	04	08D08	no_clone	31B04	no_clone	NC	<i>no_gene</i>
08	09	04	08D09	no_clone	32B05	no_clone	NC	<i>no_gene</i>
08	10	04	08D10	no_clone	31B05	no_clone	NC	<i>no_gene</i>
08	11	04	08D11	no_clone	32B06	no_clone	NC	<i>no_gene</i>
08	12	04	08D12	no_clone	31B06	no_clone	NC	<i>no_gene</i>
09	01	04	09A07	Score_card	33A04	Score_card	SC	<i>Score_card</i>
09	02	04	09E23	Score_card	33C12	Score_card	SC	<i>Score_card</i>
09	03	04	09K15	Score_card	33F08	Score_card	SC	<i>Score_card</i>
09	04	04	09A08	Score_card	34A04	Score_card	SC	<i>Score_card</i>
09	05	04	09E24	Score_card	34C12	Score_card	SC	<i>Score_card</i>
09	06	04	09K16	Score_card	34F08	Score_card	SC	<i>Score_card</i>
09	07	04	09B08	Score_card	35A04	Score_card	SC	<i>Score_card</i>
09	08	04	09F24	Score_card	35C12	Score_card	SC	<i>Score_card</i>
09	09	04	09L16	Score_card	35F08	Score_card	SC	<i>Score_card</i>
09	10	04	09B07	Score_card	36A04	Score_card	SC	<i>Score_card</i>
09	11	04	09F23	Score_card	36C12	Score_card	SC	<i>Score_card</i>
09	12	04	09L15	Score_card	36F08	Score_card	SC	<i>Score_card</i>
01	01	05	01E01	XAC3336	01C01	A0JJ0104E07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	02	05	01E02	XAC0056	02C01	A0JJ0709F09	VII.E	<i>polysaccharide export protein</i>
01	03	05	01E03	XAC1668	01C02	A0JJ0101B03	I.D.2	<i>transcriptional regulator for cryptic hemolysin</i>
01	04	05	01E04	XAC3074	02C02	A0JE1034C03	IV.A.2	<i>beta-hexosaminidase</i>
01	05	05	01E05	XAC1655	01C03	A0AC0113G09	I.D.2	<i>transcriptional regulator</i>
01	06	05	01E06	XAC2259	02C03	A0JJ1003G10	VIII.A	<i>conserved hypothetical protein</i>
01	07	05	01E07	XAC0833	01C04	A0AC0113H01	II.E	<i>acyl-CoA thioesterase I</i>
01	08	05	01E08	XAC1231	02C04	A0EC1095H07	VIII.A	<i>conserved hypothetical protein</i>
01	09	05	01E09	XAC1064	01C05	A0AC0113D01	VI.A	<i>phage-related DNA maturase</i>
01	10	05	01E10	XAC2436	02C05	A0AM1084H03	III.A.5	<i>XamI DNA methyltransferase</i>

01	11	05	01E11	XAC3492	01C06	A0AC0113H05	VIII.A	<i>conserved hypothetical protein</i>
01	12	05	01E12	XAC3574	02C06	A0JE1044B02	III.A.4	<i>DNA repair system specific for alkylated DNA</i>
02	01	05	02E01	XAC3480	05C01	A0AM1594B03	I.D.2	<i>transcriptional regulator luxR/uhpA family</i>
02	02	05	02E02	XAC2804	06C01	A0QR5320G09	I.D.1	<i>two-component system, sensor protein</i>
02	03	05	02E03	XAC0249	05C02	A0RN1630A06	III.C.3	<i>peptidyl-dipeptidase</i>
02	04	05	02E04	XAC1213	06C02	A0QR5508A07	III.D.2	<i>carboxylesterase, type B</i>
02	05	05	02E05	XAC0078	05C03	A0JE1637C05	III.B.5	<i>ATP-dependent RNA helicase</i>
02	06	05	02E06	XAC3241	06C03	A0QR5508B06	IV.D	<i>fimbrillin</i>
02	07	05	02E07	XAC0419	05C04	A0QR5206B06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	08	05	02E08	XAC2093	06C04	A0QR5512D11	III.D.2	<i>CDP-diacylglycerol-glycerol-3- phosphate 3-phosphatidyltransferase</i>
02	09	05	02E09	XAC3112	05C05	A0UV5307D11	VIII.A	<i>conserved hypothetical protein</i>
02	10	05	02E10	XAC2026	06C05	A0UV5503C08	VIII.A	<i>conserved hypothetical protein</i>
02	11	05	02E11	XAC2019	05C06	A0QR5205E09	VIII.A	<i>conserved hypothetical protein</i>
02	12	05	02E12	XAC0986	06C06	A0UV5502B11	III.B.2	<i>30S ribosomal protein S8</i>
03	01	05	03E01	XAC3890	09C01	A0QH6309G08	I.A.2	<i>bifunctional PutA protein</i>
03	02	05	03E02	XAC2987	10C01	A0UV6349H02	III.C.3	<i>proline imino-peptidase</i>
03	03	05	03E03	XAC1771	09C02	A0QR6346E11	I.A.2	<i>sialic acid-specific 9-O-acetylesterase</i>
03	04	05	03E04	XAC3496	10C02	A0UV6350A12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	05	05	03E05	XAC3321	09C03	A0QR6347A10	VI.C	<i>ISxac3 transposase</i>
03	06	05	03E06	XAC3667	10C03	A0UV6351B09	IV.A.2	<i>outer membrane protein</i>
03	07	05	03E07	XAC2650	09C04	A0QR6347B11	VI.A	<i>phage-related lytic enzyme</i>
03	08	05	03E08	XAC0818	10C04	A0QH6380E01	I.A.2	<i>ribokinase</i>
03	09	05	03E09	XAC2326	09C05	A0CE6345D07	VIII.A	<i>conserved hypothetical protein</i>
03	10	05	03E10	XAC3207	10C05	A0QR6389C08	V.A.4	<i>ferric enterobactin receptor</i>
03	11	05	03E11	XAC3840	09C06	A0CE6345D09	VIII.A	<i>conserved hypothetical protein</i>
03	12	05	03E12	XAC2983	10C06	A0UE6397B12	I.C.1	<i>quinol oxidase, subunit I</i>
04	01	05	04E01	XAC3857	13C01	A0QR6702A05	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	02	05	04E02	XAC0317	14C01	A0UV6711G02	V.A.7	<i>MFS transporter</i>
04	03	05	04E03	XAC0808	13C02	A0QR6702E10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	04	05	04E04	XAC0931	14C02	A0UV6710H08	I.D.2	<i>transcriptional regulator</i>
04	05	05	04E05	XAC1372	13C03	A0UV6708F07	VIII.A	<i>conserved hypothetical protein</i>
04	06	05	04E06	XAC2199	14C03	A0UV6710D06	VIII.B	<i>hypothetical protein</i>
04	07	05	04E07	XAC4047	13C04	A0UV6708F09	VII.C	<i>glutathione S-transferase</i>
04	08	05	04E08	XAC0744	14C04	A0UV6709F09	VIII.A	<i>conserved hypothetical protein</i>
04	09	05	04E09	XAC1170	13C05	A0UV6708E04	VIII.B	<i>hypothetical protein</i>
04	10	05	04E10	XAC3245	14C05	A0UV6711H09	IX	<i>RhsD protein</i>
04	11	05	04E11	XAC3193	13C06	A0UV6708C04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	12	05	04E12	XAC0485	14C06	A0UV6711H10	III.C.2	<i>SugE protein</i>
05	01	05	05E01	XAC3500	17C01	A0UV6733G11	II.D.10	<i>glutaredoxin-like protein</i>
05	02	05	05E02	XACb0053	18C01	A0CE6742F12	VI.B	<i>partition protein B</i>
05	03	05	05E03	XAC3110	17C02	A0UV6734D11	IX	<i>glycosyltransferase</i>
05	04	05	05E04	XAC0784	18C02	A0CE6743B07	V.B	<i>cell division protein</i>
05	05	05	05E05	XAC3343	17C03	A0UV6739A05	VIII.A	<i>conserved hypothetical protein</i>
05	06	05	05E06	XAC3850	18C03	A0CE6743G05	VII.C	<i>acriflavin resistance protein</i>
05	07	05	05E07	XAC0898	17C04	A0UV6738G03	I.D.1	<i>two-component system, regulatory protein</i>
05	08	05	05E08	XAC2268	18C04	A0CE6740H03	VIII.B	<i>hypothetical protein</i>
05	09	05	05E09	XAC0609	17C05	A0UV6739A08	III.C.3	<i>zinc protease</i>
05	10	05	05E10	XAC2300	18C05	A0CE6743A09	III.B.2	<i>50S ribosomal protein L36</i>
05	11	05	05E11	XAC3463	17C06	A0UV6739A12	VII.C	<i>TolC protein</i>
05	12	05	05E12	XAC2318	18C06	A0CE6746F10	III.B.5	<i>pseudouridylate synthase</i>
06	01	05	06E01	XAC2246	21C01	A0QR6803F03	VIII.B	<i>hypothetical protein</i>
06	02	05	06E02	XAC2485	22C01	A0UV6820E12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	03	05	06E03	XAC1319	21C02	A0UT6815A06	I.D.4	<i>RNA polymerase sigma-H factor</i>
06	04	05	06E04	XAC3106	22C02	A0UV6821C09	VIII.A	<i>conserved hypothetical protein</i>

06	05	05	06E05	XAC0070	21C03	A0UT6815D11	IV.A.1	<i>ankyrin-like protein</i>
06	06	05	06E06	XAC3424	22C03	A0UV6821A10	IX	<i>TdcF protein</i>
06	07	05	06E07	XAC2877	21C04	A0UT6817A04	III.A.2	<i>pirin-related protein</i>
06	08	05	06E08	XAC0900	22C04	A0AC6822D08	III.C.1	<i>peptide methionine sulfoxide reductase</i>
06	09	05	06E09	XAC1591	21C05	A0UT6817E06	VIII.A	<i>conserved hypothetical protein</i>
06	10	05	06E10	XAC0929	22C05	A0AC6822D09	III.C.3	<i>extracellular protease</i>
06	11	05	06E11	XAC2685	21C06	A0UT6817H04	III.B.4	<i>tRNA pseudouridine synthase B</i>
06	12	05	06E12	XAC2861	22C06	A0AC6827E01	VIII.B	<i>hypothetical protein</i>
07	01	05	07E01	XAC0968	25C01	A0QR9732D01	III.B.2	<i>30S ribosomal protein S7</i>
07	02	05	07E02	XAC2291	26C01	A0QR9812C06	V.A.7	<i>transport protein</i>
07	03	05	07E03	XAC3218	25C02	A0UV9746C10	VII.F	<i>competence lipoprotein</i>
07	04	05	07E04	XAC3627	26C02	A0QR9812E12	III.C.3	<i>oligopeptidase A</i>
07	05	05	07E05	XAC1344	25C03	A0UV9746D08	VIII.A	<i>conserved hypothetical protein</i>
07	06	05	07E06	XAC1103	26C03	A0QR9818G06	VI.C	<i>ISxac3 transposase</i>
07	07	05	07E07	XAC0520	25C04	A0UV9748A06	IV.C	<i>acyltransferase</i>
07	08	05	07E08	XAC0334	26C04	A0UV9822C06	I.C.3	<i>NADH-dependent FMN reductase</i>
07	09	05	07E09	XAC2579	25C05	A0UV9746E08	VII.E	<i>GumH protein</i>
07	10	05	07E10	XAC2209	26C05	A0UV9820B04	VIII.A	<i>conserved hypothetical protein</i>
07	11	05	07E11	XAC0374	25C06	A0QR9756C08	VIII.A	<i>conserved hypothetical protein</i>
07	12	05	07E12	XAC3833	26C06	A0UV9819D07	I.D.2	<i>B-lactamase regulatory protein</i>
08	01	05	08E01	XAC3285	29C01	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	02	05	08E02	XAC1994	30C01	genomic_DNA	VII.B	<i>HrpX related protein</i>
08	03	05	08E03	XAC4263	29C02	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	04	05	08E04	XAC0279	30C02	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	05	05	08E05	XAC2731	29C03	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	06	05	08E06	XAC1088	30C03	genomic_DNA	VII.B	<i>HrcV protein</i>
08	07	05	08E07	XACb0037	29C04	genomic_DNA	VII.H	<i>VirB11 protein</i>
08	08	05	08E08	XAC0095	30C04	genomic_DNA	VII.B	<i>Hpa2 protein</i>
08	09	05	08E09	XACb0034	29C05	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	10	05	08E10	XAC1093	30C05	genomic_DNA	I.D.2	<i>conditioned medium factor</i>
08	11	05	08E11	XACb0039	29C06	genomic_DNA	VII.H	<i>VirB9 protein</i>
08	12	05	08E12	no_clone	30C06	no_clone	NC	<i>no_gene</i>
09	01	05	09A09	Score_card	33A05	Score_card	SC	<i>Score_card</i>
09	02	05	09G01	Score_card	33D01	Score_card	SC	<i>Score_card</i>
09	03	05	09K17	Score_card	33F09	Score_card	SC	<i>Score_card</i>
09	04	05	09A10	Score_card	34A05	Score_card	SC	<i>Score_card</i>
09	05	05	09G02	Score_card	34D01	Score_card	SC	<i>Score_card</i>
09	06	05	09K18	Score_card	34F09	Score_card	SC	<i>Score_card</i>
09	07	05	09B10	Score_card	35A05	Score_card	SC	<i>Score_card</i>
09	08	05	09H02	Score_card	35D01	Score_card	SC	<i>Score_card</i>
09	09	05	09L18	Score_card	35F09	Score_card	SC	<i>Score_card</i>
09	10	05	09B09	Score_card	36A05	Score_card	SC	<i>Score_card</i>
09	11	05	09H01	Score_card	36D01	Score_card	SC	<i>Score_card</i>
09	12	05	09L17	Score_card	36F09	Score_card	SC	<i>Score_card</i>
01	01	06	01F01	XAC1616	04C01	A0AM1540D02	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	02	06	01F02	XAC2001	03C01	A0JJ1169A07	III.C.3	<i>ATP-dependent Clp protease subunit</i>
01	03	06	01F03	XAC2342	04C02	A0JJ1536A12	II.A.1	<i>gamma-glutamyl phosphate reductase</i>
01	04	06	01F04	XAC3168	03C02	A0AM1161H10	V.A.4	<i>ferric enterobactin receptor</i>
01	05	06	01F05	XAC3446	04C03	A0EC1468G03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	06	06	01F06	XAC0671	03C03	A0UV1358D08	I.D.2	<i>transcriptional regulator</i>
01	07	06	01F07	XAC4297	04C04	A0JJ1578A09	VIII.A	<i>conserved hypothetical protein</i>
01	08	06	01F08	XAC3330	03C04	A0AM1365H02	I.B.12	<i>NADPH-sulfite reductase, flavoprotein subunit</i>
01	09	06	01F09	XAC1876	04C05	A0JJ1434B09	III.B.4	<i>lysyl-tRNA synthetase heat inducible</i>
01	10	06	01F10	XAC2078	03C05	A0JE1334E12	I.C.7	<i>succinate dehydrogenase iron-sulfur protein</i>
01	11	06	01F11	XAC4324	04C06	A0EC1462C07	VIII.A	<i>conserved hypothetical protein</i>

01	12	06	01F12	XAC3548	03C06	A0AC1317A05	VII.F	<i>outer membrane protein</i>
02	01	06	02F01	XAC4364	08C01	A0QR5904F03	I.C.3	<i>oxidoreductase</i>
02	02	06	02F02	XAC0978	07C01	A0QR5701A03	III.B.2	<i>30S ribosomal protein S3</i>
02	03	06	02F03	XAC2166	08C02	A0QR6002D09	I.D.2	<i>transcriptional regulator</i>
02	04	06	02F04	XAC1829	07C02	A0JE5637G06	II.A.5	<i>histidinol dehydrogenase</i>
02	05	06	02F05	XAC0407	08C03	A0QR6002G09	VII.B	<i>HrpB1 protein</i>
02	06	06	02F06	XAC2489	07C03	A0JE5637G08	II.A.2	<i>omega-amino acid-pyruvate aminotransferase</i>
02	07	06	02F07	XAC2391	08C04	A0QR6002C07	II.B.4	<i>adenine phosphoribosyltransferase</i>
02	08	06	02F08	XAC3048	07C04	A0QR5702E11	VII.G	<i>heat shock protein</i>
02	09	06	02F09	XAC0387	08C05	A0QR6107C02	II.D.1	<i>8-amino-7-oxononanoate synthase</i>
02	10	06	02F10	XAC4330	07C05	A0QR5904D07	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	11	06	02F11	XAC0206	08C06	A0QR6002D03	V.A.4	<i>ammonium transporter</i>
02	12	06	02F12	XAC1357	07C06	A0QR5904D08	VII.G	<i>heat shock protein</i>
03	01	06	03F01	XAC2778	12C01	A0CE6480G11	II.B.4	<i>nicotinate-nucleotide adenyllyltransferase</i>
03	02	06	03F02	XAC1898	11C01	A0QR6418C10	VIII.B	<i>hypothetical protein</i>
03	03	06	03F03	XAC1457	12C02	A0UV6488G11	VII.C	<i>glutathione peroxidase-like protein</i>
03	04	06	03F04	XAC4040	11C02	A0UV6433E07	II.D.12	<i>delta-aminolevulinic acid dehydratase</i>
03	05	06	03F05	XAC3464	12C03	A0CE6507B09	IV.C	<i>3-deoxy-D-manno-octulosonic acid transferase</i>
03	06	06	03F06	XAC0843	11C03	A0AC6429C08	VIII.B	<i>hypothetical protein</i>
03	07	06	03F07	XACa0010	12C04	A0CE6505A02	VI.C	<i>ISxac3 transposase</i>
03	08	06	03F08	XAC1334	11C04	A0AC6429D09	I.B.1	<i>N-acetyl-beta-glucosaminidase</i>
03	09	06	03F09	XAC2541	12C05	A0UV6512F12	III.C.3	<i>peptidase</i>
03	10	06	03F10	XAC4171	11C05	A0UV6436A12	III.A.4	<i>exodeoxyribonuclease III</i>
03	11	06	03F11	XAC3009	12C06	A0UV6486C09	II.D.6	<i>pyridoxamine 5'-phosphate oxidase</i>
03	12	06	03F12	XAC1532	11C06	A0EC6474F07	VIII.A	<i>conserved hypothetical protein</i>
04	01	06	04F01	XAC4096	16C01	A0QR6730E10	II.E	<i>fatty acyl CoA synthetase</i>
04	02	06	04F02	XAC1201	15C01	A0UE6718F05	VIII.A	<i>conserved hypothetical protein</i>
04	03	06	04F03	XAC2566	16C02	A0QH6732G01	IV.C	<i>glycosyltransferase</i>
04	04	06	04F04	XAC0540	15C02	A0UE6718F07	III.B.6	<i>ribonuclease</i>
04	05	06	04F05	XAC4334	16C03	A0QH6732G02	VIII.B	<i>hypothetical protein</i>
04	06	06	04F06	XAC2370	15C03	A0UE6718H10	VIII.A	<i>conserved hypothetical protein</i>
04	07	06	04F07	XAC2466	16C04	A0QH6732E06	V.A.4	<i>polar amino acid transporter</i>
04	08	06	04F08	XAC3025	15C04	A0QR6729C09	VIII.B	<i>hypothetical protein</i>
04	09	06	04F09	XAC3097	16C05	A0QH6732E07	VIII.A	<i>conserved hypothetical protein</i>
04	10	06	04F10	XAC3594	15C05	A0QR6729G10	II.D.17	<i>phytoene desaturase</i>
04	11	06	04F11	XAC2052	16C06	A0QH6732B03	VIII.A	<i>conserved hypothetical protein</i>
04	12	06	04F12	XAC1587	15C06	A0QR6728A05	I.B.12	<i>thiosulfate sulfurtransferase</i>
05	01	06	05F01	XAC3516	20C01	A0QR6802E03	VII.D	<i>cellulase</i>
05	02	06	05F02	XAC1673	19C01	A0UE6757F06	VIII.A	<i>conserved hypothetical protein</i>
05	03	06	05F03	XAC0937	20C02	A0UV6766B12	III.B.6	<i>ribonuclease BN</i>
05	04	06	05F04	XAC1284	19C02	A0UE6759G02	I.D.1	<i>two-component system, regulatory protein</i>
05	05	06	05F05	XAC2084	20C03	A0UV6766C01	III.A.5	<i>DNA uptake/competence protein</i>
05	06	06	05F06	XAC1428	19C03	A0UV6761A04	III.C.1	<i>methionine aminopeptidase</i>
05	07	06	05F07	XAC0495	20C04	A0UV6766C02	I.D.1	<i>two-component system, regulatory protein</i>
05	08	06	05F08	XAC2482	19C04	A0UV6761D10	I.D.2	<i>transcriptional regulator</i>
05	09	06	05F09	XAC2681	20C05	A0QR6770A01	II.B.2	<i>nicotinate-nucleotide pyrophosphorylase</i>
05	10	06	05F10	XAC3082	19C05	A0UV6762A05	VIII.A	<i>conserved hypothetical protein</i>
05	11	06	05F11	XAC2610	20C06	A0QR6770A07	VIII.B	<i>hypothetical protein</i>
05	12	06	05F12	XAC3519	19C06	A0UV6765B02	VIII.B	<i>hypothetical protein</i>
06	01	06	06F01	XAC1722	24C01	A0UV6843A09	II.E	<i>2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase</i>
06	02	06	06F02	XAC3212	23C01	A0AC6824E11	I.C.5	<i>glucose dehydrogenase</i>
06	03	06	06F03	XACa0001	24C02	A0UV6845A08	VIII.A	<i>conserved hypothetical protein</i>
06	04	06	06F04	XAC2143	23C02	A0AC6823C09	VIII.B	<i>hypothetical protein</i>
06	05	06	06F05	XAC0184	24C03	A0QR6840A01	VIII.A	<i>conserved hypothetical protein</i>
06	06	06	06F06	XAC4326	23C03	A0AC6827D10	I.A.2	<i>urea amidolyase</i>

06	07	06	06F07	XAC0416	24C04	A0UV6845E05	VII.B	<i>Hpa1 protein</i>
06	08	06	06F08	XAC3590	23C04	A0CE6829A08	I.C.3	<i>oxidoreductase</i>
06	09	06	06F09	XAC2062	24C05	A0UV9709F08	I.C.3	<i>cytochrome C552</i>
06	10	06	06F10	XAC0218	23C05	A0CE6830B11	VIII.A	<i>conserved hypothetical protein</i>
06	11	06	06F11	XAC3846	24C06	A0UV9709G03	VIII.A	<i>conserved hypothetical protein</i>
06	12	06	06F12	XAC2814	23C06	A0QH6835C10	I.D.4	<i>RNA polymerase ECF-type sigma factor</i>
07	01	06	07F01	XAC0933	28C01	A0UE9888B07	I.A.1	<i>xylanase (truncated)</i>
07	02	06	07F02	XAC1119	27C01	A0QH9865A09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	03	06	07F03	XAC3841	28C02	A0UV9900C01	VIII.A	<i>conserved hypothetical protein</i>
07	04	06	07F04	XAC1991	27C02	A0QR9859D11	I.D.3	<i>histidine kinase</i>
07	05	06	07F05	XAC1921	28C03	A0UV9903A04	VI.C	<i>ISxac3 transposase</i>
07	06	06	07F06	XAC2729	27C03	A0UT9850H09	IV.B	<i>membrane-bound lytic murein transglycosylase D precursor</i>
07	07	06	07F07	XAC1031	28C04	A0UT9891C04	VII.C	<i>colicin V production protein</i>
07	08	06	07F08	XAC3108	27C04	A0UT9846E12	VIII.A	<i>conserved hypothetical protein</i>
07	09	06	07F09	XAC2697	28C05	A0AC9896G11	I.C.1	<i>NADH-ubiquinone oxidoreductase, NQO8 subunit</i>
07	10	06	07F10	XAC2557	27C05	A0UV9871E12	VIII.A	<i>conserved hypothetical protein</i>
07	11	06	07F11	XAC2363	28C06	A0UV9901G01	VIII.B	<i>hypothetical protein</i>
07	12	06	07F12	XAC1543	27C06	A0UE9886B12	VIII.A	<i>conserved hypothetical protein</i>
08	01	06	08F01	no_clone	32C01	no_clone	NC	<i>no_gene</i>
08	02	06	08F02	no_clone	31C01	no_clone	NC	<i>no_gene</i>
08	03	06	08F03	no_clone	32C02	no_clone	NC	<i>no_gene</i>
08	04	06	08F04	no_clone	31C02	no_clone	NC	<i>no_gene</i>
08	05	06	08F05	no_clone	32C03	no_clone	NC	<i>no_gene</i>
08	06	06	08F06	no_clone	31C03	no_clone	NC	<i>no_gene</i>
08	07	06	08F07	no_clone	32C04	no_clone	NC	<i>no_gene</i>
08	08	06	08F08	no_clone	31C04	no_clone	NC	<i>no_gene</i>
08	09	06	08F09	no_clone	32C05	no_clone	NC	<i>no_gene</i>
08	10	06	08F10	no_clone	31C05	no_clone	NC	<i>no_gene</i>
08	11	06	08F11	no_clone	32C06	no_clone	NC	<i>no_gene</i>
08	12	06	08F12	no_clone	31C06	no_clone	NC	<i>no_gene</i>
09	01	06	09A11	Score_card	33A06	Score_card	SC	<i>Score_card</i>
09	02	06	09G03	Score_card	33D02	Score_card	SC	<i>Score_card</i>
09	03	06	09K19	Score_card	33F10	Score_card	SC	<i>Score_card</i>
09	04	06	09A12	Score_card	34A06	Score_card	SC	<i>Score_card</i>
09	05	06	09G04	Score_card	34D02	Score_card	SC	<i>Score_card</i>
09	06	06	09K20	Score_card	34F10	Score_card	SC	<i>Score_card</i>
09	07	06	09B12	Score_card	35A06	Score_card	SC	<i>Score_card</i>
09	08	06	09H04	Score_card	35D02	Score_card	SC	<i>Score_card</i>
09	09	06	09L20	Score_card	35F10	Score_card	SC	<i>Score_card</i>
09	10	06	09B11	Score_card	36A06	Score_card	SC	<i>Score_card</i>
09	11	06	09H03	Score_card	36D02	Score_card	SC	<i>Score_card</i>
09	12	06	09L19	Score_card	36F10	Score_card	SC	<i>Score_card</i>
01	01	07	01G01	XAC0738	01D01	A0CE0105A07	I.C.3	<i>oxidoreductase</i>
01	02	07	01G02	XAC2634	02D01	A0AR1011B09	VI.C	<i>ISxac3 transposase</i>
01	03	07	01G03	XAC1953	01D02	A0JJ0101D08	V.C	<i>flagellar protein</i>
01	04	07	01G04	XAC1687	02D02	A0JJ1049E04	VIII.A	<i>conserved hypothetical protein</i>
01	05	07	01G05	XAC0354	01D03	A0AC0115A10	I.A.2	<i>benzaldehyde dehydrogenase II</i>
01	06	07	01G06	XAC2781	02D03	A0JJ1015G02	III.B.4	<i>leucyl-tRNA synthetase</i>
01	07	07	01G07	XAC2986	01D04	A0AC0115F04	VII.D	<i>pectate lyase II</i>
01	08	07	01G08	XAC0075	02D04	A0JJ0702C08	I.D.2	<i>xylose repressor-like protein</i>
01	09	07	01G09	XAC1638	01D05	A0AC0113H02	I.A.2	<i>imidazolonepropionase</i>
01	10	07	01G10	XAC3542	02D05	A0EC1091E04	VII.H	<i>general secretion pathway protein G</i>
01	11	07	01G11	XAC3620	01D06	A0JJ0101B11	V.A.7	<i>siderophore receptor protein</i>
01	12	07	01G12	XAC0581	02D06	A0JJ1056H09	I.D.2	<i>transcriptional regulator araC family</i>

02	01	07	02G01	XAC2198	05D01	A0UV1609A01	VII.C	<i>hemolysin- type calcium binding protein</i>
02	02	07	02G02	XAC0106	06D01	A0QR5327B01	III.A.4	<i>MutT/nudix family protein</i>
02	03	07	02G03	XAC1299	05D02	A0JE1636A02	VII.C	<i>glutathione S-transferase</i>
02	04	07	02G04	XAC4229	06D02	A0QR5512C08	I.D.4	<i>starvation sensing protein</i>
02	05	07	02G05	XAC2115	05D03	A0UV1650H03	VIII.A	<i>conserved hypothetical protein</i>
02	06	07	02G06	XAC0830	06D03	A0QR5512D08	I.B.12	<i>taurine dioxygenase</i>
02	07	07	02G07	XAC2668	05D04	A0UV5307C07	IV.D	<i>pre-pilin leader sequence</i>
02	08	07	02G08	XAC4004	06D04	A0UV5515G12	III.C.3	<i>peptidase</i>
02	09	07	02G09	XAC0435	05D05	A0UV5308B10	VII.H	<i>VirK protein</i>
02	10	07	02G10	XAC3338	06D05	A0QR5508D03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	11	07	02G11	XAC0117	05D06	A0QR5206A01	VIII.A	<i>conserved hypothetical protein</i>
02	12	07	02G12	XAC2415	06D06	A0UV5503F01	VIII.A	<i>conserved hypothetical protein</i>
03	01	07	03G01	XAC1015	09D01	A0QH6313F08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	02	07	03G02	XAC1588	10D01	A0UV6351A05	VIII.A	<i>conserved hypothetical protein</i>
03	03	07	03G03	XAC1950	09D02	A0EC6321H02	V.C	<i>flagellar FljJ protein</i>
03	04	07	03G04	XAC2762	10D02	A0UV6351B08	II.D.11	<i>geranyltranstransferase (farnesyl-diphosphate synthase)</i>
03	05	07	03G05	XAC0049	09D03	A0EC6322A05	VIII.B	<i>hypothetical protein</i>
03	06	07	03G06	XAC3988	10D03	A0UV6351H04	I.D.2	<i>transmembrane regulator protein prtR</i>
03	07	07	03G07	XAC1951	09D04	A0EC6322B02	V.C	<i>flagellar protein</i>
03	08	07	03G08	XAC2495	10D04	A0QH6383G11	VIII.B	<i>hypothetical protein</i>
03	09	07	03G09	XAC1887	09D05	A0EC6322B04	III.D.1	<i>c-di-GMP phosphodiesterase A</i>
03	10	07	03G10	XAC1753	10D05	A0QR6393G02	II.D.8	<i>phosphomethylpyrimidine kinase</i>
03	11	07	03G11	XAC3001	09D06	A0EC6323A09	V.A.7	<i>MFS transporter</i>
03	12	07	03G12	XAC0198	10D06	A0UE6399D11	VIII.A	<i>conserved hypothetical protein</i>
04	01	07	04G01	XAC3151	13D01	A0QR6704A01	VIII.A	<i>conserved hypothetical protein</i>
04	02	07	04G02	XAC1378	14D01	A0UV6711G12	III.D.2	<i>delta 9 acyl-lipid fatty acid desaturase</i>
04	03	07	04G03	XAC1506	13D02	A0QR6710E02	VIII.A	<i>conserved hypothetical protein</i>
04	04	07	04G04	XAC0302	14D02	A0UV6711D12	I.D.2	<i>transcriptional regulator lysR family</i>
04	05	07	04G05	XAC2109	13D03	A0UV6708H09	VIII.A	<i>conserved hypothetical protein</i>
04	06	07	04G06	XAC3986	14D03	A0UV6710F07	IX	<i>hydrolase</i>
04	07	07	04G07	XAC2390	13D04	A0UV6708H10	III.B.5	<i>ATP-dependent RNA helicase</i>
04	08	07	04G08	XAC0665	14D04	A0UV6709H04	VIII.A	<i>conserved hypothetical protein</i>
04	09	07	04G09	XAC2752	13D05	A0UV6708F12	V.A.7	<i>transport protein</i>
04	10	07	04G10	XAC3267	14D05	A0UV6709C12	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	11	07	04G11	XAC4124	13D06	A0UV6708E07	VIII.A	<i>conserved hypothetical protein</i>
04	12	07	04G12	XAC3540	14D06	A0UV6709D02	VII.H	<i>general secretion pathway protein I</i>
05	01	07	05G01	XAC3491	17D01	A0UV6734A08	VII.C	<i>NonF-related protein</i>
05	02	07	05G02	XAC1956	18D01	A0CE6743C06	VIII.A	<i>conserved hypothetical protein</i>
05	03	07	05G03	XAC3369	17D02	A0UV6734H07	VIII.B	<i>hypothetical protein</i>
05	04	07	05G04	XAC0552	18D02	A0CE6743E08	III.C.3	<i>proteinase</i>
05	05	07	05G05	XAC1814	17D03	A0UV6739D02	VII.C	<i>outer membrane hemolysin activator protein</i>
05	06	07	05G06	XAC1866	18D03	A0CE6740E01	IV.C	<i>wall associated protein</i>
05	07	07	05G07	XAC1833	17D04	A0UV6739F05	II.A.5	<i>phosphoribosylformimino-5- aminoimidazole carboxam</i>
05	08	07	05G08	XAC3166	18D04	A0CE6742B02	V.A.4	<i>ferric enterobactin receptor</i>
05	09	07	05G09	XAC3721	17D05	A0UV6739C03	I.A.2	<i>D-amino acid oxidase</i>
05	10	07	05G10	XACb0048	18D05	A0CE6743C02	VIII.B	<i>hypothetical protein</i>
05	11	07	05G11	XAC2767	17D06	A0UV6739C08	III.C.3	<i>TldD protein</i>
05	12	07	05G12	XAC3672	18D06	A0QR6748H10	VIII.B	<i>hypothetical protein</i>
06	01	07	06G01	XAC0316	21D01	A0QR6804B09	I.D.2	<i>transcriptional regulator lysR family</i>
06	02	07	06G02	XAC2577	22D01	A0UV6821A07	VII.E	<i>GumJ protein</i>
06	03	07	06G03	XAC2738	21D02	A0UT6815D08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	04	07	06G04	XAC2626	22D02	A0UV6821E09	IV.D	<i>fimbrial biogenesis protein</i>
06	05	07	06G05	XAC2337	21D03	A0UT6817E01	I.C.1	<i>cytochrome D ubiquinol oxidase subunit II</i>
06	06	07	06G06	XAC2064	22D03	A0AC6822D07	V.A.7	<i>cation efflux system protein</i>
06	07	07	06G07	XACb0052	21D04	A0UT6817C08	VI.B	<i>partition gene repressor</i>

06	08	07	06G08	XAC2725	22D04	A0UV6818H06	II.A.4	<i>chorismate synthase</i>
06	09	07	06G09	XAC0256	21D05	A0UV6818D12	I.B.4	<i>malate synthase</i>
06	10	07	06G10	XAC2537	22D05	A0AC6822E07	III.C.3	<i>peptidase</i>
06	11	07	06G11	XAC0166	21D06	A0UV6818E01	I.D.2	<i>transcriptional regulator lacI family</i>
06	12	07	06G12	XAC1494	22D06	A0AC6827G03	VI.A	<i>phage-related protein</i>
07	01	07	07G01	XAC4170	25D01	A0QH9737E12	VIII.A	<i>conserved hypothetical protein</i>
07	02	07	07G02	XAC3779	26D01	A0QR9814H10	VIII.B	<i>hypothetical protein</i>
07	03	07	07G03	XAC0453	25D02	A0UV9750A07	VIII.A	<i>conserved hypothetical protein</i>
07	04	07	07G04	XAC0876	26D02	A0QR9815A01	VIII.A	<i>conserved hypothetical protein</i>
07	05	07	07G05	XAC0871	25D03	A0QR9717G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	06	07	07G06	XAC4076	26D03	A0UV9819G09	IV.A.1	<i>integral membrane protein</i>
07	07	07	07G07	XAC3971	25D04	A0QR9718A03	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	08	07	07G08	XAC4331	26D04	A0QR9825A08	VIII.A	<i>conserved hypothetical protein</i>
07	09	07	07G09	XAC4234	25D05	A0UV9748A07	VIII.B	<i>hypothetical protein</i>
07	10	07	07G10	XAC3907	26D05	A0UV9822F04	VII.C	<i>mitomycin resistance protein</i>
07	11	07	07G11	XAC1011	25D06	A0QR9757E02	I.C.3	<i>oxidoreductase</i>
07	12	07	07G12	XAC0046	26D06	A0UV9820D11	IV.C	<i>UDP-N-acetyl-D-mannosamine transferase</i>
08	01	07	08G01	XAC2047	29D01	genomic_DNA	IX	<i>PHA synthase subunit</i>
08	02	07	08G02	XAC0095	30D01	genomic_DNA	VII.B	<i>Hpa1 protein</i>
08	03	07	08G03	XAC2657	29D02	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	04	07	08G04	XAC0279	30D02	genomic_DNA	VII.B	<i>HpaB protein</i>
08	05	07	08G05	XAC1092	29D03	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	06	07	08G06	XAC0028	30D03	genomic_DNA	VII.B	<i>HrcU protein</i>
08	07	07	08G07	XACb0045	29D04	genomic_DNA	VII.H	<i>VirB4 protein</i>
08	08	07	08G08	XAC1568	30D04	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	09	07	08G09	XACb0030	29D05	genomic_DNA	VI.B	<i>TrwB protein</i>
08	10	07	08G10	XAC4102	30D05	genomic_DNA	II.D.17	<i>hydroxylase</i>
08	11	07	08G11	XACb0024	29D06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	07	08G12	no_clone	30D06	no_clone	NC	<i>no_gene</i>
09	01	07	09A13	Score_card	33A07	Score_card	SC	<i>Score_card</i>
09	02	07	09G05	Score_card	33D03	Score_card	SC	<i>Score_card</i>
09	03	07	09K21	Score_card	33F11	Score_card	SC	<i>Score_card</i>
09	04	07	09A14	Score_card	34A07	Score_card	SC	<i>Score_card</i>
09	05	07	09G06	Score_card	34D03	Score_card	SC	<i>Score_card</i>
09	06	07	09K22	Score_card	34F11	Score_card	SC	<i>Score_card</i>
09	07	07	09B14	Score_card	35A07	Score_card	SC	<i>Score_card</i>
09	08	07	09H06	Score_card	35D03	Score_card	SC	<i>Score_card</i>
09	09	07	09L22	Score_card	35F11	Score_card	SC	<i>Score_card</i>
09	10	07	09B13	Score_card	36A07	Score_card	SC	<i>Score_card</i>
09	11	07	09H05	Score_card	36D03	Score_card	SC	<i>Score_card</i>
09	12	07	09L21	Score_card	36F11	Score_card	SC	<i>Score_card</i>
01	01	08	01H01	XAC4278	04D01	A0JJ1553C07	VIII.A	<i>conserved hypothetical protein</i>
01	02	08	01H02	XAC2843	03D01	A0JJ1197H03	VII.C	<i>multidrug efflux transporter</i>
01	03	08	01H03	XAC0806	04D02	A0AM1540D10	I.C.7	<i>phosphoenolpyruvate carboxylase</i>
01	04	08	01H04	XAC2157	03D02	A0JJ1169A09	II.D.14	<i>uroporphyrin-III C-methyltransferase</i>
01	05	08	01H05	XAC3224	04D03	A0AM1489A12	VII.A	<i>avirulence protein</i>
01	06	08	01H06	XAC2973	03D03	A0AM1365G11	I.D.4	<i>sigma-54 modulation protein</i>
01	07	08	01H07	XAC0894	04D04	A0RN1433D11	VII.C	<i>glutathione S-transferase</i>
01	08	08	01H08	XAC0091	03D04	A0JJ1371H02	VI.C	<i>ISxac3 transposase</i>
01	09	08	01H09	XAC0616	04D05	A0RN1447G01	VIII.B	<i>hypothetical protein</i>
01	10	08	01H10	XACb0051	03D05	A0UV1356C11	VI.C	<i>ISxac2 transposase</i>
01	11	08	01H11	XAC3514	04D06	A0AM1491D12	III.C.3	<i>serine protease</i>
01	12	08	01H12	XAC2328	03D06	A0JJ1328B09	I.C.3	<i>C-type cytochrome biogenesis membrane protein</i>
02	01	08	02H01	XAC0571	08D01	A0QR5904H08	VIII.A	<i>conserved hypothetical protein</i>
02	02	08	02H02	XAC3759	07D01	A0QR5701C12	V.A.7	<i>ABC transporter ATP-binding protein</i>

02	03	08	02H03	XAC2868	08D02	A0QR6002G07	IX	<i>response regulator</i>
02	04	08	02H04	XAC1693	07D02	A0QR5701A09	IV.C	<i>glycosyl transferase</i>
02	05	08	02H05	XAC3714	08D03	A0UV6206D02	VIII.A	<i>conserved hypothetical protein</i>
02	06	08	02H06	XAC2574	07D03	A0QR5701A12	VII.E	<i>GumM protein</i>
02	07	08	02H07	XAC3150	08D04	A0QR6002E02	III.A.3	<i>holliday junction resolvase, endodeoxyribonuclease</i>
02	08	08	02H08	XAC3306	07D04	A0QR5702H08	VIII.A	<i>conserved hypothetical protein</i>
02	09	08	02H09	XAC3732	08D05	A0QR6203E09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	10	08	02H10	XAC1644	07D05	A0QR5904F07	VIII.A	<i>conserved hypothetical protein</i>
02	11	08	02H11	XAC0790	08D06	A0QR6002E10	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
02	12	08	02H12	XAC4140	07D06	A0QR5904F09	III.C.2	<i>ClpB</i>
03	01	08	03H01	XAC0570	12D01	A0QR6483H09	I.D.4	<i>anti-sigma F factor antagonist</i>
03	02	08	03H02	XAC2499	11D01	A0EC6426D07	VII.C	<i>multidrug efflux transporter</i>
03	03	08	03H03	XAC1864	12D02	A0UT6499C08	VII.H	<i>regulatory protein</i>
03	04	08	03H04	XAC3699	11D02	A0UV6434E09	V.A.7	<i>ABC-2 type transporter</i>
03	05	08	03H05	XAC1538	12D03	A0UV6511H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
03	06	08	03H06	XAC2215	11D03	A0AC6430H05	VI.A	<i>phage-related protein</i>
03	07	08	03H07	XAC0312	12D04	A0UV6509H09	I.D.2	<i>transcriptional regulator lysR family</i>
03	08	08	03H08	XAC3677	11D04	A0AC6431A04	IX	<i>conserved hypothetical protein</i>
03	09	08	03H09	XAC0584	12D05	A0UV6514F12	VIII.B	<i>hypothetical protein</i>
03	10	08	03H10	XAC4315	11D05	A0AC6440D02	VI.B	<i>plasmid stability protein</i>
03	11	08	03H11	XAC2932	12D06	A0UV6488A10	III.C.3	<i>protease</i>
03	12	08	03H12	XAC1412	11D06	A0CE6480A04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	01	08	04H01	XAC3107	16D01	A0QR6730G04	III.A.1	<i>ATP-dependent helicase</i>
04	02	08	04H02	XAC1129	15D01	A0UE6718H07	II.E	<i>3-oxoacyl-[ACP] synthase II</i>
04	03	08	04H03	XAC2969	16D02	A0UV6733A08	VIII.A	<i>conserved hypothetical protein</i>
04	04	08	04H04	XAC3852	15D02	A0UE6718H08	VIII.A	<i>conserved hypothetical protein</i>
04	05	08	04H05	XAC2806	16D03	A0UV6733A09	VII.C	<i>beta-lactamase</i>
04	06	08	04H06	XAC0346	15D03	A0QR6724G10	VII.D	<i>cellulase (degenerated)</i>
04	07	08	04H07	XAC1923	16D04	A0QH6732G03	VIII.B	<i>hypothetical protein</i>
04	08	08	04H08	XAC3803	15D04	A0QR6729G09	III.A.5	<i>DNA processing chain A</i>
04	09	08	04H09	XAC1706	16D05	A0UV6733A02	VIII.A	<i>conserved hypothetical protein</i>
04	10	08	04H10	XAC3295	15D05	A0QR6725E12	VIII.B	<i>hypothetical protein</i>
04	11	08	04H11	XAC1459	16D06	A0QH6732C10	V.A.7	<i>ABC transporter ATP-binding protein</i>
04	12	08	04H12	XAC3991	15D06	A0QR6728D01	I.C.3	<i>cytochrome B561</i>
05	01	08	05H01	XAC3271	20D01	A0UV6766B08	V.C	<i>chemotaxis transducer</i>
05	02	08	05H02	XAC0053	19D01	A0UE6759D05	IX	<i>methyltransferase</i>
05	03	08	05H03	XAC0643	20D02	A0UV6766E02	II.D.11	<i>ubiquinone/menaquinone transferase</i>
05	04	08	05H04	XAC4059	19D02	A0UE6759H12	VIII.A	<i>conserved hypothetical protein</i>
05	05	08	05H05	XAC0754	20D03	A0UV6766E12	VIII.B	<i>hypothetical protein</i>
05	06	08	05H06	XAC3236	19D03	A0UV6761B12	I.C.7	<i>succinyl-CoA synthetase, beta subunit</i>
05	07	08	05H07	XAC0507	20D04	A0UV6766F03	III.D.2	<i>2-acylglycerophosphoethanolamine acyltransferase</i>
05	08	08	05H08	XAC4046	19D04	A0UE6758A07	III.C.3	<i>dipeptidyl peptidase IV</i>
05	09	08	05H09	XAC3137	20D05	A0QR6770D01	I.D.2	<i>transcriptional regulator</i>
05	10	08	05H10	XAC3625	19D05	A0UV6762H08	II.E	<i>beta-ketoacyl-[ACP] synthase I</i>
05	11	08	05H11	XAC1820	20D06	A0QR6801A09	II.A.2	<i>bifunctional aspartokinase/homoserine dehydrogenase I</i>
05	12	08	05H12	XAC3131	19D06	A0UV6762H12	VIII.B	<i>hypothetical protein</i>
06	01	08	06H01	XAC0704	24D01	A0UV6843H03	VII.H	<i>type II secretion system protein M</i>
06	02	08	06H02	XAC4008	23D01	A0AC6824G12	VII.C	<i>entericidin A</i>
06	03	08	06H03	XACb0042	24D02	A0QR6840E02	VIII.B	<i>hypothetical protein</i>
06	04	08	06H04	XAC0630	23D02	A0AC6823G11	II.A.2	<i>aminotransferase</i>
06	05	08	06H05	XAC3848	24D03	A0QR6840E05	VII.C	<i>membrane fusion protein precursor</i>
06	06	08	06H06	XAC1264	23D03	A0AC6827F11	VIII.A	<i>conserved hypothetical protein</i>
06	07	08	06H07	XAC2367	24D04	A0UT9701C09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	08	08	06H08	XAC0303	23D04	A0CE6829E06	I.D.2	<i>transcriptional regulator</i>
06	09	08	06H09	XAC2918	24D05	A0UV9710H10	I.D.2	<i>transcriptional regulator</i>

06	10	08	06H10	XAC0021	23D05	A0CE6830F01	VIII.A	<i>conserved hypothetical protein</i>
06	11	08	06H11	XAC3668	24D06	A0UV9711C06	V.A.7	<i>ABC transporter permease</i>
06	12	08	06H12	XAC4158	23D06	A0QH6836C02	III.B.2	<i>50S ribosomal protein L33</i>
07	01	08	07H01	XAC4064	28D01	A0UE9890E10	I.D.2	<i>transcriptional regulator araC family</i>
07	02	08	07H02	XAC4079	27D01	A0UV9870E11	I.B.10	<i>a-type carbonic anhydrase</i>
07	03	08	07H03	XAC1327	28D02	A0UV9901C06	III.A.3	<i>DNA repair protein</i>
07	04	08	07H04	XAC1023	27D02	A0QR9864A12	V.A.7	<i>TonB-dependent receptor</i>
07	05	08	07H05	XAC3867	28D03	A0UV9904B06	IV.A.1	<i>membrane protein</i>
07	06	08	07H06	XAC3197	27D03	A0UE9854G12	V.A.7	<i>ABC transporter permease</i>
07	07	08	07H07	XAC1355	28D04	A0UT9894C02	VIII.A	<i>conserved hypothetical protein</i>
07	08	08	07H08	XAC0047	27D04	A0UT9849A02	IV.C	<i>galactosyltransferase</i>
07	09	08	07H09	XAC0927	28D05	A0UV9900F09	II.A.2	<i>branched-chain amino acid aminotransferase</i>
07	10	08	07H10	XAC4367	27D05	A0UV9871G04	III.D.2	<i>glycerophosphoryl diester phosphodiesterase</i>
07	11	08	07H11	XAC2140	28D06	A0UV9903G02	III.C.3	<i>D-Ala-D-Ala carboxypeptidase</i>
07	12	08	07H12	XAC3072	27D06	A0UE9889B03	I.A.1	<i>alpha-L-fucosidase</i>
08	01	08	08H01	no_clone	32D01	no_clone	NC	<i>no_gene</i>
08	02	08	08H02	no_clone	31D01	no_clone	NC	<i>no_gene</i>
08	03	08	08H03	no_clone	32D02	no_clone	NC	<i>no_gene</i>
08	04	08	08H04	no_clone	31D02	no_clone	NC	<i>no_gene</i>
08	05	08	08H05	no_clone	32D03	no_clone	NC	<i>no_gene</i>
08	06	08	08H06	no_clone	31D03	no_clone	NC	<i>no_gene</i>
08	07	08	08H07	no_clone	32D04	no_clone	NC	<i>no_gene</i>
08	08	08	08H08	no_clone	31D04	no_clone	NC	<i>no_gene</i>
08	09	08	08H09	no_clone	32D05	no_clone	NC	<i>no_gene</i>
08	10	08	08H10	no_clone	31D05	no_clone	NC	<i>no_gene</i>
08	11	08	08H11	no_clone	32D06	no_clone	NC	<i>no_gene</i>
08	12	08	08H12	no_clone	31D06	no_clone	NC	<i>no_gene</i>
09	01	08	09A15	Score_card	33A08	Score_card	SC	<i>Score_card</i>
09	02	08	09G07	Score_card	33D04	Score_card	SC	<i>Score_card</i>
09	03	08	09K23	Score_card	33F12	Score_card	SC	<i>Score_card</i>
09	04	08	09A16	Score_card	34A08	Score_card	SC	<i>Score_card</i>
09	05	08	09G08	Score_card	34D04	Score_card	SC	<i>Score_card</i>
09	06	08	09K24	Score_card	34F12	Score_card	SC	<i>Score_card</i>
09	07	08	09B16	Score_card	35A08	Score_card	SC	<i>Score_card</i>
09	08	08	09H08	Score_card	35D04	Score_card	SC	<i>Score_card</i>
09	09	08	09L24	Score_card	35F12	Score_card	SC	<i>Score_card</i>
09	10	08	09B15	Score_card	36A08	Score_card	SC	<i>Score_card</i>
09	11	08	09H07	Score_card	36D04	Score_card	SC	<i>Score_card</i>
09	12	08	09L23	Score_card	36F12	Score_card	SC	<i>Score_card</i>
01	01	09	01I01	XAC0398	01E01	A0AC0113A01	VII.B	<i>HrpD6 protein</i>
01	02	09	01I02	XAC1109	02E01	A0JE1034B09	III.A.1	<i>DNA polymerase III tau and gamma subunits</i>
01	03	09	01I03	XAC4212	01E02	A0JJ0101H09	VII.H	<i>general secretory pathway related protein</i>
01	04	09	01I04	XAC2272	02E02	A0JJ1069G04	VIII.B	<i>hypothetical protein</i>
01	05	09	01I05	XAC0963	01E03	A0AC0115C03	III.B.2	<i>50S ribosomal protein L10</i>
01	06	09	01I06	XAC0154	02E03	A0JE1034E08	I.A.1	<i>alpha-amylase</i>
01	07	09	01I07	XAC1760	01E04	A0JJ0101B08	II.A.2	<i>dihydroxydipicolinate synthase</i>
01	08	09	01I08	XAC3361	02E04	A0AM0909H04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
01	09	09	01I09	XAC0697	01E05	A0AC0115A12	VII.H	<i>type II secretion system protein F</i>
01	10	09	01I10	XAC2234	02E05	A0JJ1101B06	V.A.7	<i>MFS transporter</i>
01	11	09	01I11	XAC2802	01E06	A0JJ0101F06	IV.A.2	<i>outer membrane channel protein</i>
01	12	09	01I12	XAC2098	02E06	A0JJ1070E12	II.D.15	<i>ATP-dependent serine activating enzyme</i>
02	01	09	02I01	XAC1260	05E01	A0CE1623F05	I.C.1	<i>cytochrome O ubiquinol oxidase, subunit III</i>
02	02	09	02I02	XAC2904	06E01	A0QR5407A10	VI.A	<i>integrase/recombinase</i>
02	03	09	02I03	XAC1957	05E02	A0UV1650G05	IV.C	<i>O-antigen biosynthesis protein</i>
02	04	09	02I04	XAC3753	06E02	A0QR5520D04	VIII.A	<i>conserved hypothetical protein</i>

02	05	09	02I05	XAC2229	05E03	A0QR5205C02	I.C.3	<i>NAD(P)H dehydrogenase</i>
02	06	09	02I06	XAC3847	06E03	A0QR5518B04	III.C.3	<i>N-acyl-L-amino acid amidohydrolase</i>
02	07	09	02I07	XAC2991	05E04	A0UV5307H09	VIII.B	<i>hypothetical protein</i>
02	08	09	02I08	XAC0298	06E04	A0QR5518B07	VIII.A	<i>conserved hypothetical protein</i>
02	09	09	02I09	XAC4099	05E05	A0QR5205D08	III.D.2	<i>acyltransferase</i>
02	10	09	02I10	XAC4013	06E05	A0QR5518E12	VIII.A	<i>conserved hypothetical protein</i>
02	11	09	02I11	XAC1979	05E06	A0QR5206F07	V.C	<i>flagellar protein</i>
02	12	09	02I12	XAC0475	06E06	A0UV5516B01	VIII.A	<i>conserved hypothetical protein</i>
03	01	09	03I01	XAC1946	09E01	A0EC6321E04	V.C	<i>flagellar protein</i>
03	02	09	03I02	XAC2750	10E01	A0UV6351G03	I.C.3	<i>reductase</i>
03	03	09	03I03	XAC2386	09E02	A0EC6324C01	VII.C	<i>superoxidase dismutase</i>
03	04	09	03I04	XAC1813	10E02	A0UV6351H03	IX	<i>HmsH protein</i>
03	05	09	03I05	XAC0152	09E03	A0EC6324C02	VI.C	<i>ISxac3 transposase</i>
03	06	09	03I06	XAC3882	10E03	A0AC6358D06	VIII.A	<i>conserved hypothetical protein</i>
03	07	09	03I07	XAC0517	09E04	A0EC6324F06	VIII.A	<i>conserved hypothetical protein</i>
03	08	09	03I08	XAC1980	10E04	A0QR6389C07	V.C	<i>flagellar L-ring protein</i>
03	09	09	03I09	XAC3828	09E05	A0EC6322H04	V.A.7	<i>ABC transporter ATP-binding protein</i>
03	10	09	03I10	XAC0261	10E05	A0UE6399D10	VIII.A	<i>conserved hypothetical protein</i>
03	11	09	03I11	XAC2320	09E06	A0QH6326A11	III.C.1	<i>glutamine cyclotransferase</i>
03	12	09	03I12	XAC1990	10E06	A0AC6405G08	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
04	01	09	04I01	XAC3451	13E01	A0QR6704D11	II.A.2	<i>ketol-acid reductoisomerase</i>
04	02	09	04I02	XAC0168	14E01	A0UV6709D06	VII.D	<i>5-keto-4-deoxyuronate isomerase</i>
04	03	09	04I03	XAC3157	13E02	A0QR6704F03	V.A.1	<i>transmembrane transport protein</i>
04	04	09	04I04	XAC0591	14E02	A0UV6711G03	III.C.3	<i>dipeptidyl peptidase IV</i>
04	05	09	04I05	XAC0739	13E03	A0QR6705C04	VIII.A	<i>conserved hypothetical protein</i>
04	06	09	04I06	XAC1912	14E03	A0UV6711B07	I.D.3	<i>serine/threonine kinase</i>
04	07	09	04I07	XAC3135	13E04	A0QR6705C11	I.D.1	<i>two-component system, regulatory protein</i>
04	08	09	04I08	XAC1560	14E04	A0UV6710B08	II.A.2	<i>5-methyltetrahydrofolate-homocysteine methyl transferase</i>
04	09	09	04I09	XAC0765	13E05	A0UV6708H11	VIII.A	<i>conserved hypothetical protein</i>
04	10	09	04I10	XAC2262	14E05	A0UV6709F10	VIII.B	<i>hypothetical protein</i>
04	11	09	04I11	XAC3862	13E06	A0UV6709A02	I.A.2	<i>chloromuconate cycloisomerase</i>
04	12	09	04I12	XAC3037	14E06	A0UV6709E10	IX	<i>hydrolase</i>
05	01	09	05I01	XAC4147	17E01	A0UV6735E04	VIII.A	<i>conserved hypothetical protein</i>
05	02	09	05I02	XAC3268	18E01	A0CE6743E07	VIII.B	<i>hypothetical protein</i>
05	03	09	05I03	XAC2855	17E02	A0UV6735B06	I.D.1	<i>two-component system, regulatory protein</i>
05	04	09	05I04	XAC2754	18E02	A0CE6743G02	III.C.1	<i>peptidyl-prolyl cis-trans isomerase</i>
05	05	09	05I05	XAC2972	17E03	A0UV6739G09	I.D.4	<i>RNA polymerase sigma-54 factor</i>
05	06	09	05I06	XAC1054	18E03	A0CE6742D11	VI.A	<i>integrase</i>
05	07	09	05I07	XAC0966	17E04	A0UV6739G11	III.B.5	<i>RNA polymerase beta' subunit</i>
05	08	09	05I08	XAC2392	18E04	A0CE6743A03	VIII.A	<i>conserved hypothetical protein</i>
05	09	09	05I09	XAC0566	17E05	A0UV6736D08	II.E	<i>malonyl CoA-ACP transacylase</i>
05	10	09	05I10	XAC3201	18E05	A0CE6743E01	V.A.7	<i>TonB-dependent receptor</i>
05	11	09	05I11	XAC3279	17E06	A0UV6739D10	VIII.B	<i>hypothetical protein</i>
05	12	09	05I12	XAC2454	18E06	A0CE6746B03	I.D.2	<i>stringent starvation protein A</i>
06	01	09	06I01	XAC4157	21E01	A0QR6804D05	I.A.2	<i>4-oxalomesaconate hydratase</i>
06	02	09	06I02	XAC0138	22E01	A0UV6821C02	VIII.B	<i>hypothetical protein</i>
06	03	09	06I03	XAC0920	21E02	A0UT6817C06	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
06	04	09	06I04	XAC2082	22E02	A0UV6821G08	V.A.7	<i>ABC transporter ATP-binding protein</i>
06	05	09	06I05	XAC0376	21E03	A0UT6817G03	VIII.A	<i>conserved hypothetical protein</i>
06	06	09	06I06	XAC1405	22E03	A0UV6818H05	II.E	<i>acetyl-coenzyme A carboxylase carboxyl transferase</i>
06	07	09	06I07	XAC3123	21E04	A0UT6817G06	III.A.2	<i>DNA-binding related protein</i>
06	08	09	06I08	XAC1148	22E04	A0UV6819B12	IV.A.1	<i>bifunctional penicillin-binding protein 1C</i>
06	09	09	06I09	XAC2642	21E05	A0QH6813A04	VI.A	<i>phage-related terminase</i>
06	10	09	06I10	XAC1123	22E05	A0AC6823D11	II.E	<i>beta-ketoacyl-[ACP] synthase III</i>
06	11	09	06I11	XAC2878	21E06	A0QH6813H02	VIII.A	<i>conserved hypothetical protein</i>

06	12	09	06I12	XAC3335	22E06	A0AC6823E09	I.D.3	<i>sensor histidine kinase</i>
07	01	09	07I01	XAC1306	25E01	A0QR9739F08	VIII.A	<i>conserved hypothetical protein</i>
07	02	09	07I02	XAC2230	26E01	A0UV9819F04	VII.C	<i>glutathione S-transferase</i>
07	03	09	07I03	XAC3115	25E02	A0QR9717D03	II.D.11	<i>PqqC protein</i>
07	04	09	07I04	XAC1287	26E02	A0UV9819G02	I.B.10	<i>aldose 1-epimerase</i>
07	05	09	07I05	XAC3740	25E03	A0CE9723F10	I.A.2	<i>UDP-glucose 4-epimerase</i>
07	06	09	07I06	XAC3084	26E03	A0UV9822A05	I.A.1	<i>beta-galactosidase</i>
07	07	09	07I07	XAC3866	25E04	A0UE9721H09	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	08	09	07I08	XAC3665	26E04	A0CE9837G04	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
07	09	09	07I09	XAC1788	25E05	A0UV9751C06	I.C.4	<i>glucose-6-phosphate isomerase</i>
07	10	09	07I10	XAC1138	26E05	A0QR9825B02	I.A.2	<i>citrate synthase 2</i>
07	11	09	07I11	XAC1002	25E06	A0QR9767D03	III.B.4	<i>Glu-tRNAGln amidotransferase A subunit</i>
07	12	09	07I12	XAC0660	26E06	A0QR9825D02	IV.A.1	<i>rod shape-determining protein</i>
08	01	09	08I01	XAC3294	29E01	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	02	09	08I02	XAC0095	30E01	genomic_DNA	VII.B	<i>HpaA protein</i>
08	03	09	08I03	XAC0330	29E02	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	04	09	08I04	XAC1091	30E02	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	05	09	08I05	XAC3984	29E03	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	06	09	08I06	XAC0095	30E03	genomic_DNA	VII.B	<i>HrpB1 protein</i>
08	07	09	08I07	XACb0033	29E04	genomic_DNA	VIII.A	<i>conserved hypothetical protein</i>
08	08	09	08I08	XAC1878	30E04	genomic_DNA	VII.H	<i>RpfC protein</i>
08	09	09	08I09	XAC2622	29E05	genomic_DNA	VIII.C	<i>Xanthomonas conserved hypothetical protein</i>
08	10	09	08I10	XAC1880	30E05	genomic_DNA	VII.H	<i>RpfB protein</i>
08	11	09	08I11	XACb0025	29E06	genomic_DNA	VIII.B	<i>hypothetical protein</i>
08	12	09	08I12	no_clone	30E06	no_clone	NC	<i>no_gene</i>
09	01	09	09A17	Score_card	33A09	Score_card	SC	<i>Score_card</i>
09	02	09	09G09	Score_card	33D05	Score_card	SC	<i>Score_card</i>
09	03	09	09M01	Score_card	33G01	Score_card	SC	<i>Score_card</i>
09	04	09	09A18	Score_card	34A09	Score_card	SC	<i>Score_card</i>
09	05	09	09G10	Score_card	34D05	Score_card	SC	<i>Score_card</i>
09	06	09	09M02	Score_card	34G01	Score_card	SC	<i>Score_card</i>
09	07	09	09B18	Score_card	35A09	Score_card	SC	<i>Score_card</i>
09	08	09	09H10	Score_card	35D05	Score_card	SC	<i>Score_card</i>
09	09	09	09N02	Score_card	35G01	Score_card	SC	<i>Score_card</i>
09	10	09	09B17	Score_card	36A09	Score_card	SC	<i>Score_card</i>
09	11	09	09H09	Score_card	36D05	Score_card	SC	<i>Score_card</i>
09	12	09	09N01	Score_card	36G01	Score_card	SC	<i>Score_card</i>

ANEXO 7

**TESE DE DOUTORADO
LEANDRO MARCIO MOREIRA**

SÚMULA CURRICULAR

SÚMULA CURRICULAR

Leandro Marcio Moreira

Nascimento: 12/03/1977, Santo André - São Paulo/SP – Brasil.

FORMAÇÃO ACADÊMICA / TITULAÇÃO

1992 – 1994

Ensino médio - Técnico em Processamento de Dados

CIM “Prof. Alcina Dantas Feijão, São Caetano do Sul, São Paulo, Brasil

1995 – 1998

Graduação em Ciências Biológicas, Bacharelado.

Universidade São Judas Tadeu, USJT, São Paulo, Brasil.

1995 – 1998

Licenciatura Plena em Biologia.

Universidade São Judas Tadeu, USJT, São Paulo, Brasil.

1999 – 2000

Especialização (*Lato Sensu*) em Biologia Molecular Aplicada.

(Carga horária: 420h)

Universidade São Judas Tadeu, USJT, São Paulo, Brasil.

2000 – 2002

Mestrado em Ciências – Bioquímica.

Universidade de São Paulo, USP, São Paulo, Brasil.

Orientador: Prof^a. Dr^a. Ana Cláudia Rasera da Silva.

Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo.

2002 – 2006

Doutorado em Ciências – Bioquímica.

Universidade de São Paulo, USP, São Paulo, Brasil.

Orientador: Prof^a. Dr^a. Aline Maria da Silva.

Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo

PUBLICAÇÕES

1 - Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities.

da Silva AC, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Van Sluys MA, Almeida NF, Alves LM, do Amaral AM, Bertolini MC, Camargo LE, Camarotte G, Cannavan F, Cardozo J, Chambergo F, Ciapina LP, Cicarelli RM, Coutinho LL, Cursino-Santos JR, El-Dorry H, Faria JB, Ferreira AJ, Ferreira RC, Ferro MI, Formighieri EF, Franco MC, Greggio CC, Gruber A, Katsuyama AM, Kishi LT, Leite RP, Lemos EG, Lemos MV, Locali EC, Machado MA, Madeira AM, Martinez-Rossi NM, Martins EC, Meidanis J, Menck CF, Miyaki CY, Moon DH, **Moreira LM**, Novo MT, Okura VK, Oliveira MC, Oliveira VR, Pereira HA, Rossi A, Sena JA, Silva C, de Souza RF, Spinola LA, Takita MA, Tamura RE, Teixeira EC, Tezza RI, Trindade dos Santos M, Truffi D, Tsai SM, White FF, Setubal JC, Kitajima JP.

Nature. 2002 May 23;417(6887):459-63.

2 - Comparative genomics analyses of citrus-associated bacteria.

Moreira LM, de Souza RF, Almeida NF Jr, Setubal JC, Oliveira JC, Furlan LR, Ferro JA, da Silva AC.

Annu Rev Phytopathol. 2004; 42:163-84.

3 - DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineate genes important for bacterial virulence.

Koide T, Zaini PA, **Moreira LM**, Vencio RZ, Matsukuma AY, Durham AM, Teixeira DC, El-Dorry H, Monteiro PB, da Silva AC, Verjovski-Almeida S, da Silva AM, Gomes SL.

J Bacteriol. 2004 Aug;186(16):5442-9.

4 - Comparative analyses of *Xanthomonas* and *Xylella* complete genomes.

Moreira LM, De Souza RF, Digiampietri LA, Da Silva AC, Setubal JC.

OMICS. 2005 Spring;9(1):43-76.

5 - Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose.

Pashalidis S, **Moreira LM**, Zaini PA, Campanharo JC, Alves LM, Ciapina LP, Vencio RZ, Lemos EG, Da Silva AM, Da Silva AC.

OMICS. 2005 Spring;9(1):77-90.

6 - Bacterial Phytopathogens and Genomic Science.

Setubal JC, **Moreira LM**, da Silva ACR

Curr Opin Microbiol. 2005; 8(5): 595 - 600.