

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL

**ESTUDO DE ISOENZIMAS DO METABOLISMO DE
CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA
ENOLASE DE *Xylella fastidiosa***

Agda Paula Facincani

Orientadora: Profa. Dra. Maria Inês Tiraboschi Ferro

Co-orientador: Prof. Dr. João Martins Pizauro Junior

Dissertação apresentada à Faculdade de
Ciências Agrárias e Veterinárias do
Campus de Jaboticabal – UNESP, para
obtenção do título de Mestre em
Agronomia – área de Concentração em
Genética e Melhoramento de Plantas

Jaboticabal - São Paulo

Julho/2002

DADOS CURRICULARES DO AUTOR

AGDA PAULA FACINCANI - nascida em 2 de maio de 1976, em São Paulo – S.P, é Bióloga (Bacharel e Licenciada) formada pelas Faculdades Integradas Riopretense – FIRP, São José do Rio Preto, SP, em dezembro de 1997. Atualmente é aluna do Programa em Genética e Melhoramento de Plantas da FCAV/UNESP, em nível de mestrado, iniciado em 1998.

“Um texto anônimo da Tradição diz que cada pessoa em sua existência pode ter duas atitudes:

Construir ou Plantar.

Os construtores podem demorar anos em suas tarefas, mas um dia terminam aquilo que estavam fazendo. Então param, e ficam limitados por suas próprias paredes. A vida perde o sentido quando a construção acaba.

Mas existem os que plantam. Estes às vezes sofrem com tempestades, estações e raramente descansam. Mas ao contrário de um edifício o jardim jamais para de crescer. E ao mesmo tempo em que exige a atenção do jardineiro, também permite que para ele a vida seja uma grande aventura.

Os jardineiros se reconhecerão entre si porque sabem que na história de cada planta está o crescimento de toda a Terra”

(Paulo Coelho)

"Ao senhor meu Deus

Mesmo na tempestade,

Mesmo que se agite o mar.

Te louvo, Te louvo em verdade.

Mesmo longe dos meus,

Mesmo na solidão.

Te louvo, Te louvo em verdade.

Pois somente tenho a ti, tu és a minha Herança

Te louvo, Te louvo em verdade.

Mesmo que me faltem palavras.

Mesmo que eu não saiba louvar.

Te louvo, Te louvo em verdade."

(Hino da Igreja Católica)

Uma pessoa é a soma de seus conhecimentos e experiências adquiridas durante toda uma vida. Tais conhecimentos e experiências são depositados sobre um cerne moldado desde sua geração por um par de pessoas que se eximiu de seus sonhos para poder ver forjados os sonhos de um novo outro. A meus pais, *Antonio e Maria José*, que nunca mediram esforços para que pudesse ter o que consideraram como o melhor do conhecimento, inúmeras experiências de vida e ser o que sou hoje, não posso somente agradecer, mas dedicar-lhes esse novo sonho realizado e retribuir-lhes sempre o amor que têm por mim. Em especial, a minha mãe por ser a professora dos meus ensinamentos da escola da vida e principalmente pela força de viver, demonstrado mesmo nos momentos difíceis em que pareciam sem soluções,

Dedico

Aos meus irmãos *Claudiner e Márcio*, a minha cunhada *Luciane*, minha nova irmã, e ao mais novo integrante da nossa família o meu sobrinho *Thiago*, pela força e ajuda sempre presente e pela alegria de convivência,

Ofereço

AGRADECIMENTOS

A Profa. Dra. Maria Inês Tiraboschi Ferro pela orientação durante estes anos, pelas oportunidades concedidas, disposição constante e pela amizade.

Ao Prof. Dr. Jesus Aparecido Ferro, pelas perguntas instigantes que me pegavam geralmente de surpresa e que fizeram com que eu aprendesse ainda mais e pelo exemplo de profissionalismo.

Ao Prof. Dr. João Martins Pizauro Junior, pela co-orientação, conselhos e ensinamentos sempre presentes.

A Profa. Dra. Eliana Gertrudes de Macedo Lemos, pelos ensinamentos e disponibilidade de seu laboratório, para realização de parte deste experimento, e de todas as pessoas do LBMP que de certa forma contribuíram para a realização deste.

Ao Prof. Dr. Haroldo Alves Pereira Junior, devo muito do interesse pela Biologia Molecular, conselhos e principalmente pela grande amizade.

Aos membros da Comissão Examinadora pela disponibilidade, pelas valiosas sugestões e pelo esmero na correção deste trabalho.

A CAPES, pela bolsa de estudo concedida.

Aos meus amigos do LBM, que por medo de minha memória falhar, não vou citar nomes, pois no meu coração estão os nomes de todos aqueles que convivi e fiz muitas amizades nessa passagem de minha vida. A minha eterna gratidão por tudo “*Pessoal*”.

Aos meus primos Elenice, Ademir e Cláudio e ao amigo Luís Fernando pelo companheirismo de morada, pelas brincadeiras e sorrisos arrancados nos momentos de tristeza, deixo aqui o que o ser humano tem de mais precioso e eterno, a amizade, o meu muito obrigado, “valeu galera”.

Aos Amigos de Pós Graduação pela eventual ajuda e amizade.

Aos meus professores durante a Pós Graduação, pela oportunidade do aprendizado.

Aos funcionários da FCAV, em especial ao pessoal do Departamento de Tecnologia, cuja disponibilidade permite a realização do nosso trabalho.

A todos que de alguma forma colaboraram na realização deste e que por falha de minha memória não foram citados.

A Deus, primeiramente, por sempre ter iluminado meu caminho e permitir que chegasse até aqui.

SUMÁRIO

LISTA DE FIGURAS.....	iv
LISTA DE TABELAS.....	v
RESUMO.....	vi
ABSTRACT.....	vii
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	1
A citricultura e a Clorose Variegada dos Citros (CVC).....	1
<i>Xylella fastidiosa</i>	3
Localização e características nutricionais do habitat.....	5
Vias glicolíticas: Embden-Meyerhof (Glicólise) e Entner-Doudoroff.....	6
2-Fosfo- D- glicerato hidrolase (Enolase).....	8
REFERÊNCIAS.....	11
 CAPÍTULO 2 - SEQÜENCIAMENTO E ANOTAÇÃO DO COSMÍDEO 02F10 DA BACTÉRIA <i>Xylella fastidiosa</i>	15
Introdução.....	16
Materiais e Métodos.....	17
Construção de biblioteca randômica a partir de DNA do cosmídeo 02F10 da bactéria <i>X. fastidiosa</i>	17
Reações de seqüenciamento.....	19
Montagem do cosmídeo, determinação e anotação as ORFs.....	20
Resultados e Discussão.....	20
Anotação do cosmídeo 02F10.....	23
Conclusões.....	25
REFERÊNCIAS.....	26
 CAPÍTULO 3 - ESTUDO DE ISOENZIMAS DO METABOLISMO DE CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA ENOLASE DE <i>Xylella fastidiosa</i>	28
Introdução.....	29
Material e Métodos.....	30
Clonagem, expressão e purificação da enolase.....	30
Crescimento de <i>X. fastidiosa</i> e linhagens de <i>E. coli</i> DH5 α e JM101,e extração de suas enolases.....	33
Extração da enolase de músculo e fígado de frango.....	33
Estudo da ativação da enolase.....	34
Determinação da atividade e estabilidade da enolase.....	34
Estudo de isoenzimas extraídas de <i>X. fastidiosa</i>	34

Resultados e Discussão.....	35
Clonagem, expressão e atividade das enolases recombinates e nativas.....	35
Análise da atividade de isoenzimas da via Glicolítica.....	38
Conclusões.....	44
REFERÊNCIAS.....	44
APÊNDICE.....	48

LISTA DE FIGURAS

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

Figura 1. Vias Glicolíticas: Glicolise e Entner-Doudoroff. EC 2.1.12 -Glucoquinase, E.C 5.3.1.9 – Fosfoglicose isomerase, EC 2.7.1.11 – Fosfofrutoquinase, E.C 4.1.2.1.3 – Frutose-1-6-bifosfato aldolase, E.C 1.2.1.1.2 – Gliceraldeído-3-fosfato desidrogenase, EC 2.7.2.3 – 3-Fosfoglicerato quinase, EC 5.4.2.1 – Fosfoglicerato mutase, E.C 4.2.1.11 – 2-Fosfo-D-glicerato hidrolase, E.C 2.7.1.40 – Piruvato quinase, E.C 1.1.1.49 – Glicose-6-fosfato desidrogenase, E.C 3.1.1.31 – 6-Fosfogluconolactonase, E.C 4.2.1.12 – 6-Fosfogluconato desidratase, E.C 4.1.2.14 – 2-Desidro-3-desoxifosfogluconato aldolase, E.C 2.7.1.12 – Gluconoquinase.....

8

Figura 2. Mecanismo catalítico da enolase. A enolase cataliza a desidratação reverssível de 2-fosfo-D-glicerato (PGA) para fosfoenolpiruvato (PEP), usando como cofator o íon magnésio (Mg^{+2}).....

9

CAPÍTULO 2 - SEQÜENCIAMENTO E ANOTAÇÃO DO COSMÍDEO 02F10 DA BACTÉRIA *Xylella fastidiosa*

Figura 1. Gel de agarose 0,8% em 1xTBE corado com brometo de etídeo (0,5µg/mL). Digestão de DNA (1µg) do cosmídeo 02F10 com *EcoRI* (1), *HindIII* (2) e *PstI* (3). O cosmídeo não digerido (ND) foi aplicado ao lado das amostras. Os marcadores de tamanho molecular são o λ *EcoRI/HindIII* (λ) e o DNA Ladder (L). A corrida foi de 90 V por 3 horas.....

21

Figura 2. Eletroforese em gel de agarose (0,8%) dos fragmentos de DNA nebulizados, reparados, fosforilados, recuperados e purificados a partir da agarose LMP. Utilizou-se tampão 1x TBE, a eletroforese procedeu-se por 2 horas a 90 V e o gel foi corado com brometo de etídeo. λ - DNA do fago lambda digerido com *EcoRI/HindIII*. 1- 1µL do DNA purificado. 2- 2µL do DNA purificado.....

22

CAPÍTULO 3 - ESTUDO DE ISOENZIMAS DO METABOLISMO DE CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA ENOLASE DE *Xylella fastidiosa*

Figura 1. Perfil eletroforético das etapas de purificação da enolase. A eletroforese foi realizada em gel de poliacrilamida a 12% de acordo com o procedimento descrito por Laemmli (1970) e o gel foi corado com comassie blue. Foram aplicados 15µL da amostra não induzida (NI), 5µL da amostra induzida (I), 15µL do sobrenadante 2 (S2), 10µL das frações protéicas de interesse (1) após a

passagem pela coluna de DEAE, 15 μ L das frações protéicas de interesse (2) após a passagem pela coluna de CM. Pb-padrão de peso molecular baixo em kDa. A flexa da esquerda indica a indução da proteína recombinante.....	37
Figura 2. Vias glicolíticas: Glicólise, Glioneogênese, Entner-Doudoroff (A.J.S.Ferreira).....	42

LISTA DE TABELAS

CAPÍTULO 2 - SEQUENCIAMENTO E ANOTAÇÃO DO COSMÍDEO 02F10 DA BACTÉRIA *Xylella fastidiosa*

Tabela 1: ORFs anotadas no cosmídeo X0JJ-02F10. São mostrados os códigos das ORFs e as informações referentes ao resultado de maior similaridade da busca realizada através do BLASTP: o nome da proteína, o organismo a que pertence, seu código no GenBank (gi) ou no SwissProt (sp) e a probabilidade esperada para a similaridade ser ao acaso (E-value).....

24

CAPÍTULO 3 - ESTUDO DE ISOENZIMAS DO METABOLISMO DE CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA ENOLASE DE *Xylella fastidiosa*

Tabela 1. Atividade das enolases extraídas de músculo e fígado de frango, de bactérias DH5 α , JM 101 e de *X. fastidiosa* usando dois tampões diferentes (músculo e fígado de frango) e dois substratos, o ácido 2-fosfoglicérico (2-PGA) e o fosfoenolpiruvato (PEP) e 10 μ L do sobrenadante dessas soluções. ND – não detectado.....

38

Tabela 2. Atividade enzimática das enzimas Fosfoglicose Isomerase (*PGI*), Glicose 6-Fosfato Desidrogenase (*G6PDH*), Aldolase ou Gliceraldeído-3-Fosfato-Liase (*ALD*), Gliceraldeído-3-Fosfato Desidrogenase (*GAPDH*) e Piruvato Quinase (*PK*), usando as bactérias *X. fastidiosa* crescidas nos meios 1, 3, 4, e 5, e a bactéria *E. coli* DH5 α como padrão, crescida em meio mínimo (meio 6). Sinal (+) indica presença, sinal (–) indica ausência.....

39

ESTUDO DE ISOENZIMAS DO METABOLISMO DE CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA ENOLASE DE *Xylella fastidiosa*

RESUMO - Muitas das questões relacionadas a como *Xylella fastidiosa* vive num ambiente cuja concentração de nutrientes é baixa, e como e porquê ela forma agregados nos vasos do xilema, devem estar relacionados com seu metabolismo. O seqüenciamento deste organismo forneceu informações para serem aplicadas em estudos genéticos funcionais, visando à elucidação da funcionalidade de vias metabólicas que possam estar intimamente ligadas à etiologia do amarelinho. O presente estudo foi conduzido com o objetivo de verificar a funcionalidade da via Glicolítica de *X. fastidiosa* através de estudo das isoenzimas fosfoglicose isomerase, aldolase ou gliceraldeído-3-fosfato liase, gliceraldeído-3-fosfato desidrogenase e piruvato quinase presentes na via Glicolítica e a glicose 6-fosfato desidrogenase presente na Entner–Doudoroff, bem como a clonagem, expressão e a determinação da atividade da enolase. O gene da enolase é uma das 35 ORFs anotadas no cosmídeo 02F10, denominada XF1291. Estes estudos sugerem que a bactéria *X. fastidiosa* não utiliza a via Glicolítica no metabolismo de carboidratos, inferindo assim no elevado tempo de duplicação deste fitopatógeno. As atividades enzimáticas da enolase recombinante, da enolase nativa e das isoenzimas aldolase e gliceraldeído-3-fosfato desidrogenase não foram detectadas, sugerindo que a *X. fastidiosa* utiliza a via Entner-Doudoroff para a formação de piruvato. São apresentadas evidências de que a regulação gênica e a presença de isoformas, com propriedades regulatórias diferentes, podem dificultar a compreensão do metabolismo de carboidratos em *X. fastidiosa*.

Palavras-Chave: enolase, Entner-Doudoroff, Glicólise, isoenzimas, seqüenciamento, *Xylella fastidiosa*

STUDY OF THE ISOZYMES INVOLVED IN CARBOHYDRATE METABOLISM AND CLONING AND EXPRESSION OF *Xylella fastidiosa* ENOLASE

ABSTRACT - Many of the questions related to how *Xylella fastidiosa* lives in an environmental with low concentration of nutrients, and the reason why she forms aggregates in xylem vessels should be related with its metabolism. However, the sequencing this organism provided information to be applied in functional genetic studies. These studies focus on elucidating the functionality of metabolic pathways that might be intimately involved in “amarelinho” etiology. The objective of this study was to verify the functionality of the Glycolytic pathways by studying the isozymes phosphoglucose isomerase, aldolase or glyceraldehyde-3-phosphato-lyase, glyceraldehyde-3-phosphate dehydrogenase and pyruvate kinase in the Glycolytic pathway and glucose-6-phosphate dehydrogenase in Entner–Doudoroff pathway, as well as the cloning and expression of enolase and determination of its activity. The enolase gene is one of 35 ORFs annotated in 02F10 cosmid and denominated XF 1291. These studies suggested that *X. fastidiosa* bacterium does not use the Glycolitic pathway in the metabolism of carbohydrates, which might explain the long time of duplication presented by this phytopathogen. The enzymatic activities of recombinant enolase, native enolase and isozymes aldolase and glyceraldehyde-3-phosphate dehydrogenase have not been detected, suggesting that *X. fastidiosa* uses the Entner-Doudoroff pathway to produce pyruvate. Evidences are presented that the regulation of genes and the presence of isoforms with different properties of regulation can make it difficult to understand the metabolism of carbohydrates in *X. fastidiosa*.

Keywords: enolase, Entner-Doudoroff, Glycolysis, isozymes, sequencing, *Xylella fastidiosa*

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

A citricultura e a Clorose Variegada dos Citros (CVC)

Desde a década de 80, a citricultura vem se destacando como a cultura mais importante na agropecuária do Estado de São Paulo correspondendo a 87% da produção brasileira e 30% da produção mundial. A produção de suco de laranja concentrado consome 70% da produção, sendo que o Estado de São Paulo e o Estado da Flórida nos Estados Unidos da América, são responsáveis por 90% da produção mundial de suco concentrado (MACHADO, 1997).

Em 1987 uma nova doença do citros, conhecida popularmente como “Amarelinho” e denominada posteriormente de Clorose Variegada dos Citros (CVC), foi identificada nos Estados de São Paulo e Minas Gerais. Inicialmente encontrada em pomares de Colina a doença foi posteriormente detectada nas regiões Norte e Nordeste do Estado de São Paulo e no Estado de Minas Gerais (ROSSETTI et al., 1998).

A doença se alastrou rapidamente pela região produtora e, atualmente, é citada como um dos fatores responsáveis pela retração na produtividade de citros do país. Em 2000, pesquisas demonstraram que do total estimado de laranjeiras, 35% apresentavam sintomas de CVC (FUNDECITRUS, 2000). O volume de recursos movimentados pelo agronegócio citrícola supera R\$ 5 bilhões por ano, gerando cerca de 400 mil empregos diretos, somente no estado de São Paulo, o maior produtor do Brasil (BESPALHOK FILHO et al., 2002).

A CVC também ocorre no Paraguai, constatada em 1997, e na Argentina onde foi encontrada pela primeira vez na região de Misiones em 1984 (ROSSETTI & DE NEGRI, 1990a).

Os sintomas da doença se caracterizam por clorose foliar, apresentando manchas amarelas na face superior das folhas, correspondendo à manchas marrons na face inferior ou em ambas; redução no tamanho da laranjeira, com a ocorrência de copa plana, quando esta é infectada ainda jovem, ou redução da velocidade de crescimento

devido aos efeitos sobre os ramos em desenvolvimento; e a ocorrência de frutos pequenos com sabor muito ácido, de amadurecimento precoce, coloração amarela intensa com manchas marrons e constituição rígida. Tais características tornam os frutos impróprios tanto para o mercado de frutas frescas como para a produção de suco concentrado (MACHADO, 1997; WELLS et al., 1987).

Os sintomas da doença estão relacionados com o transporte irregular de nutrientes nos vasos do xilema, que são bloqueados por células bacterianas, tiloses e gomas, pela presença de fitotoxinas (HOPKINS, 1989) e por desbalanço hormonal (FRENCH & STASSI, 1978). Os sintomas podem ser facilmente confundidos com estresse hídrico ou deficiências nutricionais (MALAVOLTA et al., 1990).

Em 1989 foi identificada, nos vasos do xilema de tecidos de plantas afetadas pela CVC, uma bactéria similar ao gênero *Xylella* (ROSSETTI et al., 1990). Três anos de pesquisas foram necessários após esta constatação para que o agente etiológico da CVC fosse identificado, com o fechamento dos postulados de Koch por dois grupos distintos (CHANG et al., 1993 & LEE et al., 1993). A partir da identificação da bactéria restrita ao xilema, *Xylella fastidiosa*, (WELLS et al., 1987) como causadora da CVC, iniciou-se uma nova fase na pesquisa da doença, uma vez que estudos começaram a ser dirigidos ao fitopatógeno e suas associações com o hospedeiro e vetores.

A *X. fastidiosa* é transmitida por material propagativo contaminado e por insetos vetores, particularmente cigarrinhas das famílias *Cicadellidae* (sub-família *Cicadellinae*) e *Cercopidae*, além das cigarras (*Cicadidae*). (CARLOS et al., 1998; HOPKINS, 1989). A transmissão da doença através de borbulhas contaminadas também foi comprovada por HARTUNG et al, 1994 e por LI et al, 1997. Portanto, a rápida disseminação da doença parece ser feita através de material propagativo contaminado. Uma vez introduzido na área, vetores disseminam a doença para novas plantas.

Xylella fastidiosa

Xylella fastidiosa (WELLS et al., 1987) é uma bactéria Gram-negativa limitada aos vasos do xilema de uma grande diversidade de plantas hospedeiras. É considerada espécie única do gênero *Xylella*, grupo Xanthomonas, família Lysobacteriaceae, ordem Lysobacterales, subdivisão Gama das Eubactérias.

X. fastidiosa é uma espécie de bactéria limitada ao xilema, cujo nome é derivado do termo XLB, ou em inglês, “xylem-limited bacteria”. Este termo engloba bactérias que vivem no xilema, como as bactérias parasitas endofíticas, *Xylella* e outras que habitam exclusivamente os elementos dos vasos ou traqueídeos, constituintes do xilema da planta hospedeira. Tais organismos necessitam de meios complexos para seu crescimento e ocorrem somente no xilema de plantas infectadas, são transmitidos por insetos que se alimentam do xilema e causam doenças de difícil controle em diversas culturas economicamente importantes como a citricultura (WELLS et al., 1987; PURCELL, 1999).

As bactérias limitadas ao xilema têm formato cilíndrico com paredes celulares enrugadas, não são flageladas nem formam esporos. Medem de 0,3 a 0,5 μm de diâmetro e de 1 a 5 μm de comprimento. Crescem bem a temperatura de 20 a 25 °C e pH 6,7 a 7,0. Suas colônias são brancas ou levemente pigmentadas sendo, na maioria, circulares com margens lisas ou ásperas e podem ser vistas 1-2 semanas após o isolamento primário (RAJU & WELLS., 1986). Os tamanhos dos genomas estariam nas faixas de 1,4 e entre 2,42 Mpb (MACHADO, 1997).

Estudos iniciais, em 1985, indicavam que o conteúdo GC do genoma variava de 50,1 a 54,0 %, enquanto que para o gênero *Rickettsiae* sp estão ao redor de 30%. Tal dado forneceu base para a distinção entre as bactérias Gram-negativas limitadas ao xilema (XLB) e o gênero *Rickettsiae* sp. Estudos bioquímicos, sorológicos, de enzimas extracelulares, de ácidos graxos, de DNA e RNA de 25 isolados de XLB de videira, pessegueiro, amendoeira, ameixeira, olmo, plátano, carvalho e amoreira concluíram que todo o grupo era homogêneo e que havia diferenças de crescimento *in vitro* entre os isolados (CARLOS et al., 1998).

A bactéria *X. fastidiosa* é responsável por doenças de grande importância econômica em diversas culturas como: Mal de Pierce em videiras, a escaldadura foliar em ameixeiras, uma doença de má formação em pessegueiros, clorose variegada do citros e requeima do cafeeiro. Foram relatados também em alfafa, pereira, nogueira pecã e outros hospedeiros que englobam espécies de pelo menos 28 famílias de plantas mono e dicotiledôneas (HOPKINS, 1989; PURCELL & HOPKINS, 1996; CARLOS et al., 1998). Em estudos de relação genética entre cinco cepas de XLB, as cepas não puderam ser distinguidas no nível de espécie devido à elevada homologia de DNA (75 a 100%) determinada por hibridizações (HOPKINS, 1989; PURCELL & HOPKINS, 1996).

Os resultados dos estudos de hibridizações de DNA e análise das sequências de RNA ribossomal 16S determinaram similaridades às de outras espécies, sugerindo que as XLB pertenciam ao subgrupo Gama das Eubactérias, que inclui os fitopatógenos do gênero *Xanthomonas*, excluindo deste modo qualquer relacionamento com *Rickettsiae*. Entretanto as XLB diferiam genotipica e fenotipicamente das espécies do gênero *Xanthomonas* o suficiente, para pertencerem a um gênero separado. Foi proposto então o nome de *X. fastidiosa* para estabelecer um novo gênero com uma única espécie incluindo todas as cepas de bactérias Gram-negativas fastidiosas limitadas ao xilema (XLB ou FXLB) (HOPKINS, 1989; PURCELL & HOPKINS, 1996).

Há, entretanto suficientes diferenças para uma separação no nível de subespécie em *X. fastidiosa*. Entre os isolados desta bactéria tem sido verificado variações fisiológicas, bioquímicas e moleculares, possivelmente devido à associação com seus hospedeiros. Entre os isolados de parreiras de uvas (Doença de Pierce) e de pessegueiro, existem diferenças quanto a requerimentos nutricionais (FRY et al, 1990). Uma análise utilizando a técnica de RAPD (Randomly Amplified Polymorphic DNA) entre 11 cepas de *X. fastidiosa* de ameixeira, amendoeira, amoreira, citros, olmo e videira sugeriu a existência de cinco grupos de bactérias de acordo com a origem: citros, ameixeira-olmo, videira –*Ambrósia* sp., amendoeira e amoreira.

A associação das linhagens desta bactéria com seus hospedeiros ainda não está completamente estabelecida quanto à atividade dos genes envolvidos na patogenicidade e seu nicho endofítico (CARLOS et al., 1998).

Proteínas de duas cepas de *X. fastidiosa* presentes em citros foram analisadas através de eletroforese para comparação com cepas presentes em videira, *Ambrósia* sp, plátano e ameixeira, sendo detectadas similaridades. Estudos sorológicos entre cepas de *X. fastidiosa* utilizando a técnica DIBA (Dot Immuno Binding Assay) mostraram detecções positivas em amostras da cepa causadora da CVC e nos do grupo de pessegueiros, e com menor intensidade nas do grupo da doença de Pierce. Além disso, compatibilidade sorológica entre cepas de CVC e da doença de Pierce foi detectada. Tais dados sugerem que a cepa causadora da CVC possa ser uma intermediária entre esses dois grupos (CARLOS et al., 1998).

Localização e características nutricionais do habitat

Xylella fastidiosa forma normalmente agregados nos vasos do xilema dos pecíolos, nervuras foliares e do tronco. Esses agregados aparentemente se mantêm coesos através de fibras de polissacarídeos extracelulares produzidas pela bactéria. Tais fibras são descritas por microscopia eletrônica, como microfibrilas e fímbrias elétrons densas, sendo mais abundantes nas extremidades do corpo cilíndrico das bactérias e assemelhando-se à massa entrelaçada de fibras de polissacarídeos que se estende da superfície de várias outras bactérias denominada glicocálix.

As principais funções do glicocálix que se destacam são: a capacidade de adesão entre as bactérias e destas a um tecido do hospedeiro em um ambiente favorável, a possível conservação e concentração de enzimas digestivas e o possível papel de resina de troca iônica, uma vez que tem densidade eletrônica elevada, o que tornaria o glicocálix um reservatório nutricional ligando íons ao agregado bacteriano e tornando-os disponíveis num ambiente nutricional diluído. Há evidências de que as fibras das cepas de *X. fastidiosa* causadora da doença de Pierce funcionem como um glicocálix (HOPKINS, 1989; PURCELL & HOPKINS, 1996).

Os elementos traqueais do xilema das plantas hospedeiras de *X. fastidiosa* oferecem um ambiente físico e nutricional único para seu crescimento. A seiva bruta, suspensão de nutrientes proveniente das raízes que corre pelo xilema em direção às folhas, é um meio nutricionalmente diluído, embora contenha uma grande variedade de nutrientes essenciais. A quantidade e composição de tais nutrientes varia com o ciclo circadiano e sazonal, a localização, a idade, a saúde e a espécie da planta (HOPKINS, 1989; PURCELL & HOPKINS, 1996).

Estudos feitos sobre a seiva bruta de *Vitis rotundifolia*, espécie de videira, que também pode ser afetada por *X. fastidiosa*, indicam que ela consiste de uma solução pouco concentrada de moléculas orgânicas pequenas (aminoácidos, outros ácidos orgânicos e açúcares, sendo que estes apresentaram o menor perfil total, menos de 0,2 mmol/L), ânions e cátions de micro e macroelementos, e concentrações muito baixas de reguladores de crescimento como citocininas e giberelinas. Todavia, todos os micro e macro elementos essenciais analisados (NH_4^+ , NO_3^- , P, K, Ca, Mg, Mn, Fe, Cu, Zn, Na e Cl) foram detectados, todos os aminoácidos exceto triptofano também foram encontrados. Os compostos orgânicos em maiores concentrações foram: glutamina, ácido málico, ácido tartárico e ácido cítrico (ANDERSEN & BRODBECK, 1989). Muitas das questões relacionadas a como *X. fastidiosa* vive num ambiente cuja concentração de nutrientes é baixa, e como e porquê ela forma agregados nos vasos do xilema devem estar intimamente relacionados com seu metabolismo e pode também ser umas das explicações ao longo tempo de geração desse fitopatógeno.

Vias glicolíticas: Embden-Meyerhof (Glicólise) e Entner-Doudoroff

A diversidade dos procariotos está intimamente ligada a sua grande variação na produção de energia e metabolismo. Os procariotos, como um grupo, possuem basicamente três vias metabólicas para a degradação de monossacarídeos sendo elas: Embden-Meyerhof-Parnas (glicólise), a Entner-Doudoroff e a Via das Pentoses.

A glicólise é uma via central quase universal do catabolismo da glicose, também chamada de via Embden-Meyerhof. Esta é uma via metabólica antiga usada

provavelmente a 3,5 bilhões de anos atrás pelas primeiras bactérias conhecidas. Esta foi a prévia para a presença de uma grande quantidade de oxigênio na atmosfera. Entretanto, a via funcionava sob condições anaeróbicas. A glicólise é um processo que pode ocorrer na presença ou ausência de oxigênio. A glicose é quebrada em duas moléculas de ácido pirúvico e envolve duas fases: a fase de “investimento” de energia e a fase de “pagamento”. Durante este processo a energia livre é liberada e é conservada na forma de ATP. Enzimas específicas são envolvidas na via catabólica com seu substrato específico produzindo produtos intermediários na conversão de glicose a piruvato (Figura 1).

Nos organismos em que a via glicolítica não está presente, e mesmo em grande parte dos organismos que a utilizam como principal via de degradação de monossacarídeos, existe uma via metabólica, ou variações da mesma, denominada via de Entner-Doudoroff. Essa via tem como substrato inicial, glicose-6-fosfato, seguindo para glucono-1,5-lactona-6-fosfato, 6-fosfogluconato e 2-desidro-3-desoxifosfogluconato. Normalmente é utilizada em bactérias para o metabolismo de glucoronato, com um balanço energético líquido de um mol de ATP por mol de glucoronato (ou glicose, dependendo do substrato inicial) (Figura 1) (ROMANO & CONWAY, 1996).

A Via Entner-Doudoroff foi descoberta em *Pseudomonas saccharophila* em 1952 (ENTNER & DOUDOROFF, 1952) e mais tarde (1967) em *Escherichia coli* (ESIENBERG & DOBROGOSZ, 1967), e está presente em todos os três domínios filogenéticos, incluindo o *Archaea* (CONWAY, 1992). ROMANO & CONWAY (1996) em ensaios sobre a evolução da via glicolítica, sugeriram que a Entner-Doudoroff predata da via Embden-Meyerhof.

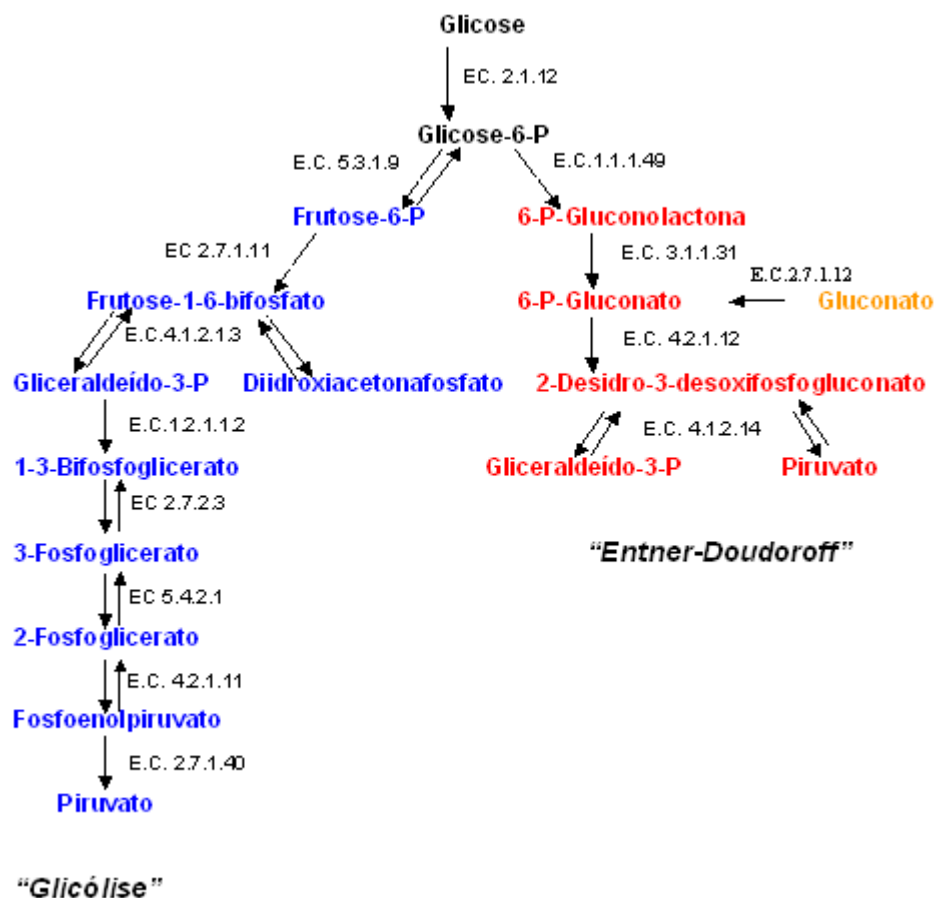


Figura 1. Vias Glicolíticas: Glicólise e Entner-Doudoroff. EC 2.1.12 -Glucocinase, E.C 5.3.1.9 – Fosfoglicose isomerase, EC 2.7.1.11 – Fosfofrutoquinase, E.C 4.1.2.1.3 – Frutose-1-6-bifosfato aldolase, E.C 1.2.1.1.2 – Gliceraldeído-3-fosfato desidrogenase, EC 2.7.2.3 – 3-Fosfoglicerato quinase, EC 5.4.2.1 – Fosfoglicerato mutase, E.C 4.2.1.11 – 2-Fosfo-D-glicerato hidrolase, E.C 2.7.1.40 – Piruvato quinase, E.C 1.1.1.49 – Glicose-6-fosfato desidrogenase, E.C 3.1.1.31 – 6-Fosfogluconolactonase, E.C 4.2.1.12 – 6-Fosfogluconato desidratase, E.C 4.1.2.14 – 2-Desidro-3-desoxifosfogluconato aldolase, E.C 2.7.1.12 - Gluconocinase

2-Fosfo- D- glicerato hidrolase (Enolase)

A 2 fosfo- D- glicerato hidrolase (enolase, EC 4.2.1.11) é uma enzima vital da via glicolítica, inteiramente citoplasmática, produzida por uma variedade de organismos diferentes, estando envolvida na fase de “pagamento” da glicólise, é uma “metaloenzima - enzima ativada por íons metais” (BREWER, 1981) que cataliza a

deidratação de 2-fosfo-D-glicerato (PGA) para fosfoenolpiruvato (PEP) na via glicolítica (Figura 2) e a reação reversa, a hidratação do PEP para PGA, na gliconeogênese (WOLD, 1971).

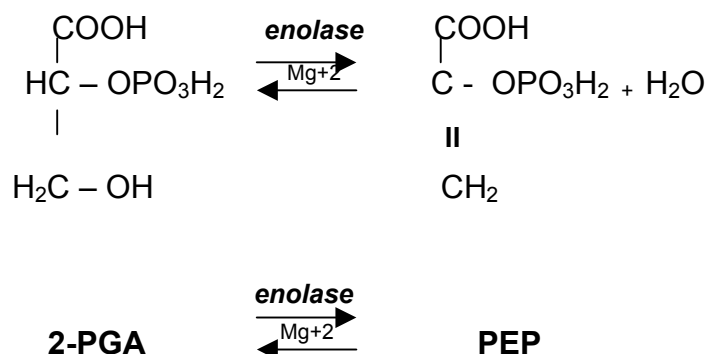


Figura 2. Mecanismo catalítico da enolase. A enolase cataliza a deidratação reverssível de 2-fosfo-D-glicerato (PGA) para fosfoenolpiruvato (PEP), usando como cofator o íon magnésio (Mg^{+2})

A reação envolve somente uma pequena quantidade de energia livre, cerca de 1 Kcal/mol.

VAN DER STRAENTEN (1991) verificou que a enolase de uma variedade de fontes, inclindo bactérias, leveduras, drosófilas, anfíbios, pássaros, plantas e humanos são altamente conservadas.

A enolase tem um requerimento absoluto de certos íons metais divalentes para atividade. O cofator natural é Mg^{+2} (íon magnésio) que confere maior atividade (WOLD & BALEON, 1957; BREWER, 1985). A enzima pode ser deionizada e o Mg^{+2} pode ser substituído por muitos outros íons metais divalentes como Zn^{+2} , Mn^{+2} , Co^{+2} , Ni^{+2} e Cu^{+2} , que conferem atividade menor que o íon Mg^{+2} e Ca^{+2} e o íon Sm^{+3} que não apresentam atividade mensurável.

A enolase é uma enzima dimérica, encontrada em mamíferos como um homodímero ou heterodímero consistindo de três subunidades: α , β e γ codificadas por três loci distintos. O peso molecular da maioria das enolases estudadas estão dentro de uma faixa que vai de 80 – 100 kDa, entretanto a massa de uma subunidade simples varia geralmente de 40 a 50 kDa. Enzimas de células procarióticas parecem ser uma exceção, por exemplo a enolase de bactérias termofílicas é um octâmero, a enolase

isolada de *Streptococcus rattus* tem um peso molecular de 49 kDa (o peso da subunidade neste caso é aproximado a 22 kDa) (HÿTHER et al., 1990). Em *Escherichia coli* a enolase possui aproximadamente 90 kDa e é compsta de duas subunidades de peso molecular idênticos (WOLD, 1971).

Os pHs ótimos da enolase de vários organismos são similares. Para as proteínas de *Candida albicans* ele é 6,8 (KUSTRZEBA-WOJCICKA & GOLCZAK, 2000) e é semelhante ao músculo de carpa (6,8) (PIETKIEWICZ & KUSTRZEBA-WOJCICKA, 1983) e músculo de coelho (pH 7,0) (WOLD, 1971). *Pyrococcus furiosus* tem, no entanto um pH ótimo mais alto (8,1) (PEAK et al., 1994). O perfil para atividade da enolase parece característico; a atividade diminui excessivamente para valores de pH abaixo do pH ótimo, enquanto para valores de pH acima do ótimo a atividade diminui, mas não tão abruptamente.

Em *Sacharomyces cerevisiae* a enolase e outras enzimas glicolíticas são as proteínas mais abundantes na célula (HOLLAND & HOLLAND, 1978; MAITRA & LOBO, 1971). A enolase de *S. cerevisiae* é a mais estudada de todas as enolases com respeito as características bioquímicas, biofísicas e regulação genética. A enolase também é a mais abundante enzima glicolítica em *Zymomonas mobilis*. A relativa abundância da enzima enolase pode ser devido a altos níveis de mRNA_{eno} (RNA mensageiro da enzima enolase), tradução elevada de mRNA_{eno} ou resistência da enzima a proteólise (BURNETT et al., 1992).

Estudos em levedura mostraram a presença de três sítios de ligação íon metal por subunidade na enolase. O sítio de ligação I, tradicionalmente chamado “conformacional”, induz a mudança conformacional no sítio ativo e possibilita a ligação do substrato ou substrato análogo (FALLER & JOHNSON, 1974). A ligação do íon metal neste sítio é o mais forte dos três sítios. Seguindo a ligação do substrato ou substrato análogo, o segundo íon metal chamado “catalítico” pode se ligar no sítio II e então a reação catalítica ocorre (FALLER et al., 1977). Se a concentração de íons metais é alta, o terceiro íon metal, chamado “inibitório”(sítio III), pode se ligar. Este íon metal ligado inibe a atividade enzimática (FALLER et al., 1977). O sítio inibitório pode ser eliminado a valores de pH baixos. A ligação de íons metálicos no sítio catalítico é mais fraca

quando comparado com a ligação no sítio conformacional. A nomenclatura do sítio íon metal é derivada de efeitos macroscópicos incluindo associação na subunidade (BREWER & WEBER, 1968) de ligação dos íons metais nestes sítios, embora não reflète a função do íon metal em resolução atômica.

Referências

- ANDERSEN, P.C., BRODBECK, B.V., Diurnal and temporal changes in the chemical profile of xylem exudate from *Vitis rotundifolia*. **Physiologia Plantarum**, v. 75, p. 63-70, 1989.
- BESPALHOK-FILHO, J. C. et al. Laranja TRANSGÊNICA: Transformação de laranja visando resistência ao cancro cítrico usando genes de peptídeos antibacterianos. **Biotecnologia**. Ed. IV, v. 23, p 62, 2002.
- BREWER, J.M. & WEBER, G. The reversible dissociation of yeast enolase. **Proc Natl Acad Sci USA**, v. 59, n. 1, p. 216-223, 1968.
- BREWER, J. M. Yeast enolase: mechanism of activation by metal ions. **Crit. Rev. Biochem**, v. 11, p. 209-254, 1981.
- BREWER, J. M. Specificity and mechanism of action of metal-ions in yeast enolase. **FEBS Lett**, v. 182, p. 8-14, 1985.
- BURNETT, M. E. et al. Molecular characterization of the *Zymomonas mobilis* Enolase (eno) Gene. **J. Bacteriol**, v. 174, p. 6548-6553, 1992.
- CARLOS, E.F., et al. The bacterium *Xylella fastidiosa*. In: DONADIO, L.C. AND MOREIRA, C.S. (eds). **Citrus Variegated Chlorosis**, Bebedouro, 1998.
- CHANG, C.J., GARNIER, M., ZREIK, L., ROSSETTI, V., BOVE, J.M. CHANG, C. J. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. **Curr. Microbiol**, v.27, p.137-42, 1993.
- CONWAY, T. The Entner-Doudoroff pathway: history, physiology, and molecular biology. **FEMS Microbiol. Rev**, v. 103, p. 1-27, 1992.

- EISENBERG, R. C., DOBROGOSZ, W. J. Gluconate metabolism in *Escherichia coli*. *J. Bacteriol*, v. 93, p. 941-949, 1967.
- ENTNER, N., DOUDOROFF, M. Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. *J. Biol.Chem*, v. 196, p. 852-862, 1952.
- FALLER, L. D. et al. Magnesium ion requirements for yeast enolase activity. **Biochemistry**, v. 16, p. 3864-3869, 1977.
- FALLER, L. D., JOHNSON, A.M. Calorimetric studies of the role of magnesium ions in yeast enolase catalysis. **Proc. Natl. Acad. Sci. USA**, v. 71, p. 1083-1087, 1974.
- FRENCH, W. I., STASSI, D. L. Response of phony-infected peach trees with gibberillic acid. **HortScience**, v.13, p.158-159, 1978.
- FRY, S. M., MILHOLLAND, R. D. Multiplication and translocation of *Xylella fastidiosa* in petioles and stems of grapevine resistant, tolerant, and susceptible to Pierce's disease. **Phytopathology**, v. 80, p. 61-65, 1990.
- FUNDECITRUS, (<http://www.fundecitrus.com.br>), 2000.
- HARTUNG, J. S. et al. Citrus variegated chlorosis bacterium: axenic culture, pathogenicity and relationships with other strains of *Xylella fastidiosa*. **Phytopathol**, v. 64, p. 591-597, 1994.
- HOLLAND, M. J., J. P. HOLLAND. Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde -3- phosphate dehydrogenase, and phosphoglycerate kinase. **Biochemistry**, v. 17, p. 4900-4907, 1978.
- HOPKINS, D.L., *Xylella fastidiosa*: xylem limited bacteria pathogen of plants. **Annual Reviews of Phytopathology**, v. 27 p. 271-290, 1989.
- HÜTHER, F. J. et al. Isolation, characterization and inhibition kinetics of enolase from *Streptococcus rathus*. FA-1. **Infect. Immun**, v. 58, p. 1043-1047, 1990.
- KUSTRZEBA-WÓJCICKA, I., GOLCZAK, M. Enolase from *Candida albicans* – purification and characterization. **Comp. Biochem. Physiol. B**. v. 126, p. 109-120, 2000.
- LEE, R. F. et al. Citrus variegated chlorosis: confirmation of a *Xylella fastidiosa*: as the causal agent. **Summa Phytopathologica**, v.19, n.2, p.123-125,1993.

- LI, W. B. et al. Métodos de avaliação de resistência e seleção de variedades à clorose variegada dos citros . **Laranja**, v. 17, n. 1, p. 55-56, 1997.
- MACHADO, M. Citrus Variegated Chlorosis (CVC), a new destructive citrus disease in Brazil, and the xylem limited bacteria, *Xylella fastidiosa*. <<http://www.dcc.unicamp.br/genoma/xylella.html/>>. (1997).
- MAITRA, P. K., Z. LOBO. A kinetic study of glycolitic enzymes synthesis in yeast. **J. Biol. Chem**, v. 246, p. 475-488, 1971.
- MALAVOLTA, E. et al. Nova anomalia dos citros – Estudos preliminares. **Laranja**, v.11, n.1, p.15-38, 1990.
- PEAK, M. J. et al. The hyperthermophilic glycolitic enzyme enolase in the archaeon, *Pyrococcus furiosus*: Comparison with mesophilic enolase. **Arch. Biochem. Biophys**, v. 313, p. 280-286, 1994.
- PIET KIEWICZ, J., KUSTRZEBA-WÓJCICKA, I. Purification and properties of enolase from carp. Comparason with enolase from mammals muscles and yeast. **Comp. Biochem. Physiol, C**. v. 75B, p. 693-698, 1983.
- PURCELL, A. <<http://nature.berkeley.edu/xylella/portugese/brazil.html>>. (1999).
- PURCELL, A.H., HOPKINS, D.L. Fastidious xylem-limited bacterial plant pathogens. **Annu. Rev. Phytopathol**, v.34, p.131-151, 1996.
- RAJU, B. C., WELLS, J. M. Disease caused by fastidious xylem-limited. **Plant Dis**, v. 70, n. 3, p.182-86, 1986.
- ROMANO, A.H., CONWAY, T. Evolution of carbohydrate metabolic pathways, 14th Forum in Microbiology. **Res. Microbiol**, v. 147, p. 448-455, 1996.
- ROSSETTI, V., DE NEGRI, D. Clorose Variegada dos Citros: revisão. **Laranja**, v.11, n1, p.1-14, 1990a.
- ROSSETTI, V., GARNIER, M., BOVÉ, J.M., BERETA, M.J.G., TEIXEIRA, A.R.R., QUAGIO, J.A., DE NEGRI, J.D. Présence de bactéries dans le xylème d'orangers atteints de chlorose variégée, une nouvelle maladie des agrumes au Brésil. **C. R. Acad. Sci.**, Paris, v. 310, p. 345-349, 1990.
- ROSSETTI, V., GONZALES, M.A., DONADIO, L.C. History. In: DONADIO, L.C. AND MOREIRA, C.S. (eds). **Citrus Variegated Chlorosis**, Bebedouro, 1998.

VAN DER STRAETEN, D. Plant enolase: gene structure, expression and evolution. **Plant Cell**, v. 3, p. 719-735, 1991.

WELLS, J., RAJU, B., JUNG, H., WEISBURG, W., MANDELCO-PAUL, L., BRENNER, D. *Xylella fastidiosa* gen nov, sp nov gramnegative, xylem limited fastidious plant bacteria related to *Xanthomonas* ssp., **Int. J. of Syst. Bacteriol**, v. 37, p. 136-143, 1987.

WOLD, F. **In the Enzymes** (Boyer, P.D. Ed.) v. 5, 3rd ed. New York. Academic Press. 1971. p.499-538.

WOLD, F., BALLOW, C. E. Studies on the enolase. **J. Biol. Chem**, v. 227, p. 313-328, 1957.

CAPÍTULO 2 - SEQÜENCIAMENTO E ANOTAÇÃO DO COSMÍDEO 02F10 DA BACTÉRIA *Xylella fastidiosa*

Seqüenciamento e anotação do cosmídeo 02F10 da bactéria *Xylella fastidiosa*

A. P. Facincani¹, H. A. Pereira Jr¹, L. R. Furlan¹, J. A. Ferro¹.

¹*Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias/
UNESP – Jaboticabal – SP - Brasil.*

RESUMO: Em 1987 uma nova doença de citros, conhecida popularmente como “Amarelinho” e denominada posteriormente de Clorose Variegada dos Citros (CVC), foi identificada nos estados de São Paulo e Minas Gerais. Atualmente 35% de laranjeiras apresentam sintomas da CVC. Em 1989 foi identificada nos vasos do xilema de tecidos de plantas afetadas pela CVC a bactéria *Xylella fastidiosa*. *X. fastidiosa* é uma bactéria Gram-negativa limitada aos vasos do xilema de uma grande diversidade de plantas hospedeiras causando perdas em diversas culturas. Em 1997 a FAPESP, conjuntamente ao FUNDECITRUS, desenvolveu o projeto de seqüenciamento do genoma do fitopatógeno *X. fastidiosa*. Em nosso laboratório foram seqüenciados onze cosmídeos, sendo o cosmídeo 02F10 o objeto deste estudo. Dentro desse contexto o objetivo deste trabalho foi o seqüenciamento completo do cosmídeo 02F10 e a sua anotação. A montagem do inserto compreendeu 856 seqüências numa extensão de 35.283 pb e a taxa de erro final do inserto do cosmídeo foi determinada em 0,02/10.000 pb. Também ficou constatado que o 02F10 está localizado na posição 519065, de acordo com a base número 1 (origem de replicação), e se estendendo até 554348 na direção c, e mediado pelos cosmídeos 6C10 e 7G7. A análise da seqüência do inserto do cosmídeo 02F10, determinou 55 ORFs das quais 35 foram analisadas para anotação

Palavras-Chave: anotação, clorose variegada dos citros (CVC), cosmídeo, seqüenciamento, *Xylella fastidiosa*

Introdução

A determinação da sequência nucleotídica de genomas bacterianos tem se destacado pela potencialidade de controlar as doenças causadas por estas bactérias.

Desde a publicação da sequência completa do genoma da bactéria *Haemophilus influenzae* em 1987, várias outras bactérias estão sendo alvo de seqüenciamento do seu genoma. Hoje mais de 100 genomas bacterianos estão para ser completamente seqüenciados ou em andamento.

A maioria das bactérias alvo de seqüenciamento são patogênicas do homem ou de animais domésticos. As bactérias fitopatogênicas são responsáveis por perdas de mais de 20% das culturas de interesse do homem nas regiões agrícolas (DOW & DANIELS, 2000).

O conhecimento da sequência de nucleotídeos de um genoma tem interesse econômico e social devido à possibilidade do controle de pragas ou na utilização deste conhecimento em benefício do homem, como no desenvolvimento de vacinas e diagnósticos mais confiáveis e precoces. O interesse na determinação da sequência completa do genoma de um organismo baseia-se na possibilidade de identificação de genes que concedam as características únicas do mesmo.

Em 1997 foi feito um consórcio entre a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), e o Fundo de Defesa da Citricultura de São Paulo (FUNDECITRUS), para o desenvolvimento do projeto de seqüenciamento do genoma do fitopatógeno *Xylella fastidiosa*, causador da doença Clorose Variegada dos Citros (CVC) que aflige a citricultura paulista. Dentre os diversos motivos relacionados para a realização desse projeto de seqüenciamento estavam a inexistência de um modelo de interação entre o patógeno e os citros e questões como: qual a diversidade genética entre as diferentes cepas que compõem a espécie *X. fastidiosa*, qual o tamanho real do genoma, se haveria a ocorrência de plasmídeos na cepa que ataca os citros e a possível associação destes com a patogenicidade, e se haveria genes de virulência ou

avirulência que estariam relacionados aos sintomas da CVC ou aos vetores (MACHADO, 1997).

A publicação da primeira seqüência do genoma completo de uma bactéria fitopatogênica, *X. fastidiosa*, pelo consórcio de pesquisadores no Brasil (SIMPSON et al., 2000) refletiu um trabalho pioneiro na área genômica. Esse fitopatógeno foi escolhido por considerações econômicas sendo, o de maior importância para a citricultura tanto brasileira como mundial.

X. fastidiosa é uma bactéria Gram-negativa limitada aos vasos do xilema de uma grande diversidade de plantas hospedeiras que causa perdas em diversas culturas. Em 1989 foi identificada como o agente causal da Clorose Variegada do Citros (CVC), que atinge atualmente 35% das laranjeiras. Atualmente a CVC é citada como um dos fatores responsáveis pela retração na produtividade de citros no país.

Em nosso laboratório foram seqüenciados onze cosmídeos, sendo o cosmídeo 02F10 o objeto deste estudo. Dentro desse contexto o objetivo deste trabalho foi o seqüenciamento completo do cosmídeo 02F10 e a sua anotação, pois os dados do seqüenciamento da bactéria *X. fastidiosa* geraram informações para serem aplicadas em estudos genéticos funcionais. Esse trabalho foi publicado dentro do contexto do artigo de SIMPSON et al., 2000, podendo ser constatado no Apêndice.

Materiais e Métodos

Construção de biblioteca randômica a partir de DNA do cosmídeo 02F10 da bactéria Xylella fastidiosa.

Duas estratégias foram utilizadas para o seqüenciamento do genoma da bactéria *Xylella fastidiosa*: 1) biblioteca randômica de DNA genômico total em plasmídeo (vetor pBS+) e, 2) biblioteca de cosmídeo, totalizando 113 cosmídeos selecionados para o sequenciamento, sendo o cosmídeo 02F10 escolhido como objeto deste estudo.

Para a construção da biblioteca randômica cosmidial com aproximadamente 40 Kb de inserto de DNA genômico da bactéria *X. fastidiosa* no vetor *Lawrist 4*

(<http://fugu.hgmp.mrc.ac.uk/fugu/Lawrist4.html>), foi realizada a extração em larga escala utilizando-se o Kit Wizard^R Plus Maxipreps DNA Purification System (Promega), segundo as especificações do fabricante e a purificação através da ultracentrifugação em gradiente de cloreto de cério (SAMBROOK et al., 1989), sendo que 1 µg do DNA purificado foi digerido com as enzimas *Hind* III, *Eco*R I e *Pst* I (BioLabs) segundo as especificações do fabricante e aplicados em gel de agarose 0,8%. A fragmentação randômica do DNA foi feita por nebulização segundo BIRREN et al (1997) utilizando as variantes de 20 segundos e pressão de 3 Kgf/cm² para a obtenção de fragmentos de 0,9 a 2,0 Kb. O reparo das extremidades dos fragmentos foi feito em combinação da T4 DNA polimerase (Pharmacia) e Klenow da DNA polimerase I de *E. coli* (Pharmacia) e a fosforilação das extremidades 5' dos fragmentos randômicos pela T4 polinucleotídeo quinase (T4 PNK- Biolabs), seguindo as especificações do fabricante. Os fragmentos com as pontas reparadas e fosforiladas foram aplicados em gel de agarose 0,7% Low Melting Point (LMP) (BioRad) para recuperação e purificação de acordo com BIRREN et al (1997), e então, clonados em vetor pUC 18 utilizando o Kit Ready-To-GoTM pUC18 smal/BAP+Ligase (Pharmacia) seguindo as especificações do fabricante.

Após a ligação, bactérias da linhagem JM109 de *E. coli* competentes (Promega) foram transfectadas, através da metodologia de choque térmico, segundo as especificações do fabricante. Foram selecionados clones contendo plasmídeos com inserto através da expressão de *lacZ*, após plaqueamento em meio sólido 2xTY (SAMBROOK et al., 1989) contendo 100 µg/mL de carbenicilina e 5'-bromo-4-cloro-3-indolil-galactopiranosídeo (X-gal). As colônias obtidas foram transferidas para placas de Petri (150 x 15mm) contendo meio 2xTY acrescido de 100 µg/mL de carbenicilina, e nomeadas (A, B, C, etc) e cada clone numerado (1 a 96). O DNA plasmidial foi obtido através de minipreparações segundo GRIFFIN & GRIFFIN (1993).

Reações de seqüenciamento

O Kit utilizado para o seqüenciamento das amostras de DNA plasmidial foi ABI Prism® Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Applied Biosystems, Perkin-Elmer) e primer universal que flanqueiam o sítio de clonagem do plasmídeo pUC18 (#1201 ou #1211). O termociclador utilizado para o desenvolvimento das reações de seqüenciamento cíclico foi o MJ-PTC 1197. Para a reação de extensão foram utilizados 400 ng de DNA plasmidial; 1,0 µL BigDye™ premix; 3,0 µL de 2,5x Buffer (200 mmol/L Tris-HCl pH 9,0; 5,0 mmol/L MgCl₂) e 3,2 pmol de primer único. A extensão foi realizada por um programa com um período de desnaturação inicial de 2 minutos a 94°C seguido de 35 ciclos de 94°C, 90 segundos; 54°C, 30 segundos; 72°C, 90 segundos. A purificação dos produtos das reações de seqüenciamento foi realizada através de precipitação com isopropanol 75% e lavadas duas vezes com etanol 70%. As amostra purificadas e concentradas foram ressuspensas em 4,0 µL de tampão de aplicação (1 parte de Blue Dextran em solução de EDTA 25 mmol/L para 5 partes de formamida desionizada) e aplicadas (1,5 ou 1,8 µL) em gel desnaturante de seqüenciamento (36,0 ou 48cm de comprimento respectivamente), sendo a eletroforese desenvolvida em seqüenciador automático ABI 377 DNA Sequencer (Applied Biosystems, Perkin-Elmer). O géis de seqüenciamento foram preparados utilizando-se a mistura concentrada de acrilamida/bisacrilamida Long Ranger™ 50% (FMC). Os géis tinham composição final de acrilamida 5% (% m/v), tampão TBE 1% (% v/v) e uréia 0,36 g/mL (5,0 mol/L). As amostras foram analisadas no seqüenciador utilizando dois parâmetros de corrida, ou seja, para corridas com placa de 36 cm foi usado no Run Module-SeqRun 36E-1200 e o tempo de corrida ajustado pelo próprio aparelho para 7 horas, já com placas de 48cm o Run Module-SeqRun foi 48E-1200 e 10 horas de corrida.

Montagem do cosmídeo, determinação e anotação das ORFs

As seqüências geradas foram analisadas com o software Sequencher 3.1 (Mac) e então transferidas para uma Workstation Sun–Ultra 5, na qual foram feitas as análises e montagens com o pacote Phred-Phrap-Consed (GORDON et al., 1998) sob sistema UNIX. Os critérios para o seqüenciamento e montagem do cosmídeo foram: a seqüência consenso da montagem não deveria ter regiões com qualidade menor ou igual a 20; a taxa de erro esperada para a montagem deveria ser menor ou igual a 1 em cada 10.000 bases (0,01%) e cada base (ou espaço) do consenso da montagem deveria ser confirmado por pelo menos uma seqüência em cada fita. Estando a montagem de acordo com tais critérios, o cosmídeo em questão era considerado terminado (“finished”). A determinação das ORFs (“open reading frames”) foram realizada inicialmente pelo programa Sequin (BENSON et al, 1999) e posteriormente pelo programa Glimmer (SALZBERG et al, 1998), fazendo buscas em seqüências de DNA por regiões com alta probabilidade estatística de codificarem proteínas que se iniciem pelos códons: ATG, GTG e TTG. A busca de similaridades entre seqüências protéicas foi realizada através do serviço BLASTP (v. 2.0) (ALTSCHUL & GISH, 1996), no banco de dados disponibilizado pelo National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST/>). Cada ORF anotada foi inicialmente classificada de acordo com a classificação existente para os genes de *Escherichia coli* (RILEY, 1993).

Resultados e Discussão

O DNA obtido da larga escala e purificação do cosmídeo 02F10 apresentou uma relação de $A_{260}/A_{280}=1,8898$ e uma concentração de $0,334\mu\text{g}/\mu\text{L}$. A Digestão com as enzimas *Hind* III, *Eco*R I e *Pst* I (BioLabs) do DNA cosmidial purificado (Figura 1) foi feito para verificar a presença do cosmídeo correto, presença ou ausência de contaminates, possíveis recombinações e ainda ter uma estimativa do tamanho do inserto do DNA clonado, através de medição das distâncias migradas e a plotação

desses dados em papel semi logarítmico tendo como padrão a migração do λ E-H e DNA Ladder (dados não apresentados).

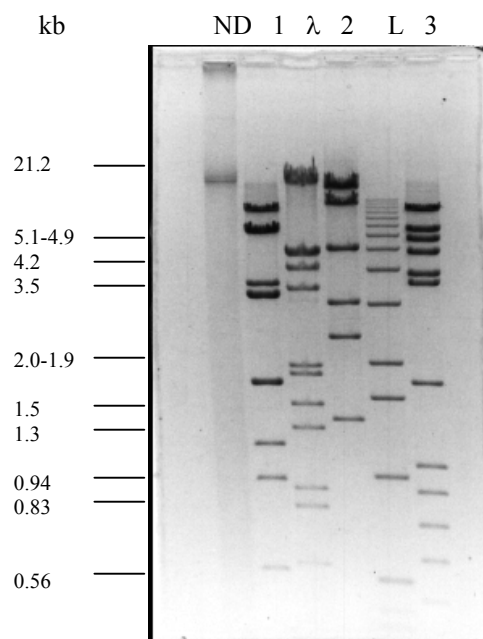


Figura 1. Gel de agarose 0,8% em 1xTBE corado com brometo de etídeo (0,5 μ g/mL). Digestão de DNA (1 μ g) do cosmídeo 02F10 com *EcoRI* (1), *HindIII* (2) e *PstI* (3). O cosmídeo não digerido (ND) foi aplicado ao lado das amostras. Os marcadores de tamanho molecular são o λ *EcoRI/HindIII* (λ) e o DNA Ladder (L). A corrida foi de 90 V por 3 horas.

A fragmentação randômica do DNA por nebulização produz quebras randômicas (ao acaso) em qualquer ponto da seqüência, tendo-se no final uma mistura de fragmentos com tamanhos variados. Como a clonagem de cada fragmento foi feita no sítio *SmaI* do vetor pUC18, havia a necessidade de que as duas extremidades dos fragmentos randômicos fossem abruptas. Isso só foi conseguido utilizando a combinação da T4 DNA polimerase com a Klenow DNA polimerase. Após a recuperação dos fragmentos desejados, a quantidade de DNA presente nesta solução foi determinada a partir da absorbância a 260 nm. Esta solução de DNA apresentou uma relação $A_{260}/A_{280}=1,602$ e uma concentração de 166 ng/ μ L. A constatação da recuperação dos fragmentos desejados pode ser observada na figura 2.

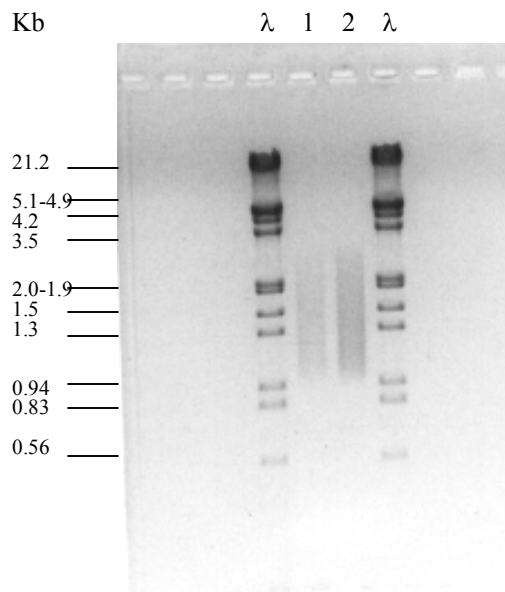


Figura 2. Eletroforese em gel de agarose (0,8%) dos fragmentos de DNA nebulizados, reparados, fosforilados, recuperados e purificados a partir da agarose LMP. Utilizou-se tampão 1x TBE, a eletroforese procedeu-se por 2 horas a 90 V e o gel foi corado com brometo de etídeo. λ - DNA do fago lambda digerido com *EcoRI/HindIII*. 1- 1 μ L do DNA purificado. 2- 2 μ L do DNA purificado.

Para a clonagem destes fragmentos utilizou-se um vetor com os fosfatos da extremidade 5' removidos pela fosfatase o que evita ou minimiza muito a recircularização do vetor sem a presença de inserto, aumentando a eficiência da biblioteca.

O sucesso da reação de ligação é de fundamental importância para a obtenção de uma boa biblioteca. Entende-se como boa biblioteca a que, neste caso, randômica, apresenta clones contendo insertos com tamanhos que cobrem toda a faixa selecionada. (no caso, 0,9 a 2,0 kb). Dois fatores principais estão envolvidos na reação de ligação: porcentagem de vetor digerido e desfosforilado e proporção entre vetor e inserto. Garantindo que todo o DNA esteja digerido e desfosforilado, impede-se que o mesmo se religue e volte a ser circular (sem a presença de inserto). Já a proporção inserto:vetor, apesar de vários estudos a respeito, permanece mais ou menos empírico.

A reação de seqüenciamento cíclico utilizado no projeto é baseada no método de extensão de "primer" terminado por didesoxinucleotídeos trifosfato (dNTPs), método de

Sanger (VOET & VOET, 1995) consistindo de etapas sucessivas de desnaturação do molde (plasmídeo onde está inserido o fragmento a ser seqüenciado), pareamento do primer a ser estendido e extensão do primer pela ação de uma DNA polimerase termoestável, utilizando-se um termociclador, resultando numa amplificação linear dos produtos de extensão.

A montagem final do inserto compreendeu 856 seqüências numa extensão de 35.283 pb, contendo 760 seqüências providas de plasmídeos da biblioteca do cosmídeo 02F10, 9 seqüências de bibliotecas do cosmídeo 07G07, 4 seqüências de bibliotecas do cosmídeo 03E01, 3 seqüências de bibliotecas do cosmídeo 08A02, 25 seqüências de bibliotecas do cosmídeo 09E09 e 55 seqüências provenientes de bibliotecas de “shotgun” genômico. A taxa de erro final do inserto do cosmídeo foi determinada como 0,02/10.000 pb. Após o término do seqüenciamento total do Genoma da *X. fastidiosa* ficou determinado que o cosmídeo 02F10 está localizado na posição 519065, de acordo com a base número 1 determinada pela coordenação (origem de replicação), e se estendendo até 554348 na direção c (totalizando 35.283pb), e mediado pelos cosmídeos 6C10 e 7G7.

A baixa taxa de erro final do inserto do cosmídeo foi possível graças ao uso de 17 “primers” que foram sintetizados pela Bio Synthesis, com o propósito de sanar as regiões de “single strands” de acordo com as normas da coordenação do projeto.

Anotação do cosmídeo 02F10

A análise da seqüência do inserto do cosmídeo 02F10, realizada pelo programa Sequin, determinou inicialmente 55 ORFs, posteriormente com a utilização do programa Glimmer ficou constatado a presença de 35 ORFs, valor este devido às sobreposições de ORFs ou ORFs hipotéticas, ou seja, não classificada de acordo com a coordenação como codificadoras de proteínas. Dentre essas 35 ORFs anotadas, 11 são hipotéticas, 8 hipotéticas conservadas e 16 foram atribuídas funções com base na similaridade pelo programa BLASTP, cujas descrições podem ser observadas na tabela 1.

Tabela 1: ORFs anotadas no cosmídeo X0JJ-02F10. São mostrados os códigos das ORFs e as informações referentes ao resultado de maior similaridade da busca realizada através do BLASTP: o nome da proteína, o organismo a que pertence, seu código no GenBank (gi) ou no SwissProt (sp) e a probabilidade esperada para a similaridade ser ao acaso (E-value).

ORF	Proteína	Organismo	Código # (Acession #)	E-value (BLASTP)
XF-1282	Carboxipeptidase	<i>Deinococcus radiodurans</i>	gi 6458686	5e-53
XF-1286	Topoisomerase IV subunitB	<i>Salmonella typhimurium</i>	sp P31598	0.0
XF-1288	CTP synthetase	<i>Escherichia coli</i>	sp P08398	0.0
XF-1289	2-dehydro-3-deoxyphosphooctonate aldolase	<i>Escherichia coli</i>	sp P17579	1e-63
XF-1291	Enolase (2-Phosphoglycerate Dehydratase) (2-Phospho-D-Glycerate Hydro-liase).	<i>Bacillus subtilis</i>	sp P37869	e-147
XF-1297	Gluconolactonase Precursor	<i>Streptomices coelicolor</i>	gi 6137026	4e-46
XF-1298	Electron Transfer Flavoprotein Ubiquinone Oxidoreductase	<i>Zimomonas mobilis</i>	gi 4511983	e-138
XF-1299	DNA repair system specific for alkylated DNA	<i>Brucella melitensis</i>	gi 6759911	5e-04
XF-1302	ABC Transporter ATP-Binding Protein	<i>Helicobacter pylori</i>	gi 2314642	1e-32
XF-1309	Rod Shape-Determining Protein MREB	<i>Escherichia coli</i>	sp P13519	e-147
XF-1310	Rod Shape-Determining Protein mreC	<i>Escherichia coli</i>	sp P16926	6e-43
XF-1311	Rod Shape-Determining Protein MRED	<i>Escherichia coli</i>	sp P16927	3e-11
XF-1312	Penicillin-Binding Protein 2	<i>Haemophilus influenzae</i>	sp P44469	e-113
XF-1313	Rod Shape-Determining Protein rodA	<i>Escherichia coli</i>	sp P15035	1e-81
XF-1314	S-Adenosylmethionine: tRNA Ribosyltransferase-Isomerase	<i>Escherichia coli</i>	sp P21516	1e-91
XF-1316	ATP: GTP3' - pyrophosphotranferase	<i>Vibrio sp</i>	sp P55133	e-143

Nesse segmento seqüenciado e anotado de 35.283 pares de bases do genoma de *X. fastidiosa* foi encontrado um “operon” denominado *mre*, cujos genes são responsáveis pela determinação da forma celular e sensibilidade das células crescidas na presença de uma amidopenicilina, a mecilina. O “operon” é formado por quatro genes, sendo eles o *mrdA* que codifica a peptidoglicana sintetase, proteína de ligação à

penicilina, *mrdB* que codifica para a proteína *rodA* e os genes *mreC* e *mreD* que estão diretamente envolvidos na determinação do formato e sensibilidade a mecilina das células (WACHI, et al., 1987)

Com o grande número de projetos de seqüenciamento em andamento, há sempre uma busca pela determinação de uma estratégia que seja a mais eficaz e eficiente, fazendo com que a etapa de determinação da seqüência do genoma de um organismo se torne cada vez menos importante em vista das descobertas realizadas pela anotação e pelos projetos funcionais decorrentes.

A utilização de cosmídeos para o seqüenciamento deste fitopatógeno teve como principal ponto favorável o fato de estabelecer rapidamente uma base para um banco de dados de anotação de ORFs. O fato de o inserto ter um tamanho aproximado a 40 quilobases também foi importante para o seqüenciamento e a montagem correta de regiões contendo repetições de grande número de nucleotídeos.

Um problema relacionado à utilização de cosmídeo, como de qualquer outro vetor que aceite insertos de grandes proporções, é a clonagem de insertos quiméricos formados por fragmentos de diferentes regiões do genoma que se ligam, formando um fragmento de DNA de tamanho compatível a ser inserido no vetor.

O seqüenciamento do cosmídeo 02F10, não apresentou grandes problemas, pois a constituição de sua seqüência não portava o problema anteriormente descrito, e a presença de alguns "GAPs" foram sanadas pela construção de 17 "primers" nas regiões de "single strands"

Conclusões

Os resultados obtidos no presente trabalho nos permitiram concluir que:

- a metodologia utilizada para a construção da biblioteca randômica se demonstrou eficaz e eficiente, tornando-se cada vez menos importantes em vista das descobertas realizadas pela anotação,
- na região seqüenciada e anotada foi encontrado o "operon" *mre*, responsável pela determinação do formato e sensibilidade a mecilina das células.

Referências

- ALTSCHUL, S. F. and GISH, W. Local alignment statistics, **Methods of Enzymology**, v. 266, p. 460-480, 1996.
- BENSON, D. A. et al. Genbank. **Nucleic Acids Research**, v. 27, n.1, p. 12-17, 1999.
- BIRREN, B. et al. **Genome Analysis. A Laboratory Manual**. Cold Spring Harbor Laboratory Press. 1997.
- DOW, J.M., DANIELS, M.J. Xylella genomics and bacterial pathogenicity to plants. **Yeast**, v. 17, p. 263-271, 2000.
- GORDON, D., ABAJIAN, C., GREEN, P., "Consed: a graphical tool for sequence finishing", **Genome Research**, v. 8, n. 3, p. 195-202, 1998.
- GRIFFIN, H.G. and GRIFFIN, A.M. **Methods in molecular biology. DNA sequencing protocols**. Totowa, Humana Press Inc., NJ, 1993. p. 23.
- MACHADO, M. Citrus Variegated Chlorosis (CVC), a new destructive citrus disease in Brazil, and the xylem limited bacteria, *Xylella fastidiosa*. <<http://www.dcc.unicamp.br/genoma/xylella.html>>. (1997).
- RILEY, M. Functions of the gene products of *Escherichia coli*. **Microbiological Reviews**, v. 57, n. 4, p. 862-952, 1993.
- SALZBERG, S. et al. Microbial gene identification using interpolated Markov models, **Nucleic Acids Research**, v. 26, n. 2, p. 544-548, 1998.
- SAMBROOK, J., FRITSCH, E.E. and MANIATIS, T. **Molecular Cloning: A Laboratory Guide**, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 1989.
- SIMPSON, A.J. et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. **Nature**, v. 406, p.151-157, 2000.
- WACHI, M., et al. Mutant isolation and molecular cloning of *mre* genes, which determine cell shape, sensitivity to mecillinam, and amount of penicillin-binding proteins in *Escherichia coli*. **J. Bacteriol**, v. 169, p. 4935-4940, 1987.
- VOET, D., VOET, J.G. **Biochemistry**, 2nd edition, John Wiley & Sons, New York. 1995.

CAPÍTULO 3 - ESTUDO DE ISOENZIMAS DO METABOLISMO DE CARBOIDRATOS E CLONAGEM E EXPRESSÃO DA ENOLASE DE *Xylella fastidiosa*

Estudo de isoenzimas do metabolismo de carboidratos e clonagem e expressão da enolase de *Xylella fastidiosa*

A. P. Facincani¹, J. A. Ferro¹, J. M. Pizauro. Jr¹, H. A. Pereira Jr¹, E. G. M. Lemos¹, A. L. Prado¹, M. I. T. Ferro¹.

¹Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias/UNESP- Jaboticabal - SP - Brasil.

RESUMO: O presente estudo foi conduzido com o objetivo de verificar a funcionalidade da via Glicolítica da bactéria *Xylella fastidiosa* através de estudo das isoenzimas fosfoglicose isomerase, aldolase ou gliceraldeído-3-fosfato liase, gliceraldeído-3-fosfato desidrogenase e piruvato quinase presentes na via Glicolítica e a glicose 6-fosfato desidrogenase presente na Entner-Doudoroff, bem como a clonagem, expressão e a determinação da atividade da enolase. Estes estudos sugerem que a bactéria *X. fastidiosa* não utiliza a via Glicolítica no metabolismo de carboidratos, inferindo assim no elevado tempo de duplicação deste fitopatógeno. A enolase recombinante foi expressa em bactéria *E. coli* BL21(DE3) pLYS-s na forma de corpos de inclusão e sua solubilização foi realizada com uréia, Triton X-100 ou TCA. A uréia foi o extrator mais eficiente e o Triton X-100 e TCA os menos eficientes. A enolase extraída de *X. fastidiosa*, músculo e fígado de frango é inativada irreversivelmente pela uréia. A enolase foi purificada parcialmente e apresentou baixo rendimento. As atividades enzimáticas da enolase recombinante, da enolase nativa e das isoenzimas aldolase e gliceraldeído-3-fosfato desidrogenase não foram detectadas, sugerindo que a *X. fastidiosa* utiliza a via Entner-Doudoroff para a formação de piruvato. São apresentadas evidências de que a regulação gênica e a presença de isoformas, com propriedades regulatórias diferentes, podem dificultar a compreensão do metabolismo de carboidratos em *X. fastidiosa*.

Palavras-Chave: enolase, Entner-Doudoroff, Glicólise, isoenzimas, *Xylella fastidiosa*.

Introdução

Xylella fastidiosa é uma bactéria Gram negativa, de formato cilíndrico, com paredes celulares enrugadas, aflageladas e não formam esporos. Está presente no xilema de plantas infectadas e foi detectada não somente em laranjas doces, mas também em outras plantas de importância econômica, como videira, ameixeira, pessegueiros, amendoadeiras, café, alfafa, pereira, noqueira e outros (PURCELL & HOPKINS, 1996).

O desenvolvimento da bactéria leva a obstrução do sistema vascular da planta, causando estresse hídrico e uma desordem nutricional na planta. A *X. fastidiosa* é transmitida por material propagativo contaminado, por diferentes insetos vetores, principalmente cigarrinhas (Cicadellidae e Cercopidae) e por borbulhas contaminadas (HARTUNG et al., 1994).

A compreensão do mecanismo de infecção e das interações planta-patógeno pode contribuir para o controle deste fitopatógeno. Neste contexto, a bactéria *X. fastidiosa* estirpe 9a5c foi o primeiro fitopatógeno de plantas a ser completamente seqüenciado, por um consórcio de pesquisadores no Brasil (SIMPSON et al., 2000).

Uma das estratégias de seqüenciamento deste fitopatógeno foi o seqüenciamento de cosmídeos, sendo um deles o 02F10, cuja seqüência é o objeto de interesse deste estudo. As informações sobre o seqüenciamento e anotação desse cosmídeo, bem como do genoma desse fitopatógeno, podem ser observada no “site” <http://onsona.lbi.dcc.unicamp.br/xf/>.

Os dados do seqüenciamento da bactéria *X. fastidiosa* geraram informações para serem aplicadas em estudos genéticos funcionais. Estudos da análise de metabolismo de carboidratos no genoma de *X. fastidiosa*, FERREIRA (2000) demonstraram que todos os genes das enzimas que participam da glicólise, do ciclo do ácido tricarboxílico e da cadeia transportadora de elétrons estão presentes no genoma deste fitopatógeno. Já a via das pentoses e da gliconeogênese estão aparentemente incompletas, podendo ser uma possível explicação do elevado tempo de duplicação da

bactéria, pois os carboidratos disponíveis também serão destinados a síntese da parede celular.

Dentro deste contexto, os objetivos deste trabalho são os de estudar se estas vias são funcionais ou se existem vias alternativas. O estudo do metabolismo de carboidratos foi feito através do estudo das isoenzimas fosfoglicose isomerase, aldolase ou gliceraldeído-3-fosfato-liase, gliceraldeído-3-fosfato desidrogenase e piruvato quinase, presentes na via Glicolítica e a glicose 6-fosfato desidrogenase presente na Entner–Doudoroff, bem como a clonagem, expressão e a determinação da atividade da 2 fosfo-D-glicerato hidrolase (enolase - EC 4.2.1.11).

A enolase é uma enzima inteiramente citoplasmática, e catalisa a hidratação de 2-fosfo-D-glicerato (PGA) para fosfoenolpiruvato (PEP) na glicólise e a reação reversa, a hidratação do PEP para PGA, na gliconeogênese (WOLD, 1971), identificada por similaridade no cosmídeo 02F10 e denominada como XF1291. A enolase é uma metaloenzima, cujo peso molecular é cerca de 40 a 50 kDa e apresenta um requerimento absoluto de magnésio para sua atividade.

Material e Métodos

Clonagem, expressão e purificação da enolase

A seqüência de nucleotídeos da enolase de *X. fastidiosa* foi clonada por produto de PCR. Para tanto foram sintetizados dois “primers” diretos com inserção de sítio de restrição *Nde*I (Enolase F- 5’ CGT TTT CTT ACA TAT GAT G 3’ e Enolase 32F- 5’ GTT GAC TGG ACA TAT GAC CGC TAT TGC C 3’) e um “primer” reverso com inserção do sítio de *Hind* III (Enolase R- 5’ GCC AGT TAC AAG CTT ATC AGG ACT T 3’), para facilitar a clonagem no vetor de expressão. Essas duas combinações de “primers” possibilitaram a amplificação do gene a partir das duas possíveis metioninas iniciadoras, a primeira na posição 1 e a segunda na posição 32. Um quarto “primer” (Enolase 396- 5’ AAT AAG GGG CGT TTG GGG GCT 3’) foi sintetizado na posição 396 para auxílio no seqüenciamento. A amplificação foi feita com 100 ng de DNA cosmidial,

5 pmoles de cada “primer”, 1U de Taq DNA Polimerase (Gibco), 200 μ M de dNTPs, 50 mM de $MgCl_2$ e tampão 10x Taq , utilizando o termociclador Perkin Elmer 2400 e um programa com uma desnaturação inicial de 2 minutos a 94°C seguido de 35 ciclos de 94°C, 90 segundos; 54°C, 30 segundos; 72°C, 90 segundos. Os fragmentos amplificados foram purificados de acordo com BIRREN et al (1997), e então, clonados em vetor T utilizando o Kit pGEMT System II (Promega), segundo as especificações do fabricante.

Estas construções foram utilizadas para transformar as bactérias da linhagem DH10B de *E. coli* por eletroporação (Bio Rad Gene Pulser II) usando a capacitância de 50 μ F, “low range” a 200 Ω , “high range” a $\infty \Omega$ e ajuste para 1,5 k volts.

A partir das colônias transformadas foram feitos minipreparações (GRIFFIN & GRIFFIN, 1993) e 0,5 μ g do DNA obtido, digerido com as endonucleases de restrição *Hind* III e *Nde* I (Biolabs) segundo as especificações do fabricante para confirmação do inserto. Os clones com tamanhos de inserto esperado foram seqüenciados com auxílio do Kit Big Dye Terminator (ABI-Perkin Elmer) utilizando 5 pmoles dos “primers” T7 e SP6, específicos para o vetor T, 400 ng de DNA, 2 μ L de “Big Dye Terminator”, 6 μ L de tampão 2,5x (200 mM de Tris-HCl pH 9,0; 5,0 mM de $MgCl_2$) e um programa com um tempo de desnaturação inicial de 2 minutos a 96°C e 39 ciclos de 96°C, 30 segundo; 50°C, 15 segundos; 60°C, 4 minutos. O produto da amplificação foi analisado em um “ABI Prism 377 DNA Automated Sequencer (Perkin-Elmer)”. Após confirmação por seqüenciamento do clone correto de cada construção, foi feita uma preparação de DNA em larga escala utilizando-se o Kit Wizard^R Plus Maxipreps DNA Purification System (Promega), segundo as especificações do fabricante sendo que, 5 μ g de cada amostra de DNA digerido com *Hind* III e *Nde* I e aplicado em gel 0,8% LMP para posterior recuperação e purificação (BIRREN et al., 1997). Esses fragmentos foram subclonados no vetor de expressão pET 3a, também digerido com as endonucleases *Hind* III e *Nde* I, usando uma proporção de 4 volumes de vetor para 1 de inserto, 2 μ L de tampão para T4 DNA ligase 10 x concentrado, 0,5 μ L de T4 DNA ligase Bio Labs (400 U/ μ L), num volume final de 20 μ L e mantidas a 15°C por 16 horas. Em seguida, 4 μ L dessa solução

de DNA foi utilizada para eletroporar 40 μ L de células de *E. coli* BL21(DE3) pLYS-S. As colônias obtidas foram crescidas em 3 mL de meio 2xTY (SAMBROOK et al., 1989) acrescido de 200 μ g/mL de carbenicilina e cloranfenicol, até $OD_{600nm} = 0,8 - 1,0$ e induzidas com Isopropil-1-Tiogalactopiranosídeo (IPTG) numa concentração de 0,4 mM por 3 horas. As células foram centrifugadas (16.000 x g por 5 minutos a 4°C) e rompidas com 50 μ L de tampão de lise 2x concentrado (125 mM Tris-HCl pH 6,8; 4% w/v SDS; 20% v/v glicerol; 5% v/v β -mercaptoetanol; 0,01% w/v azul de bromofenol) sendo aquecidas a 95°C por 5 minutos, centrifugadas (16.000 x g por 10 minutos) e aplicados 15 μ L em gel de poliacrilamida SDS –PAGE 12%. A expressão da enolase em larga escala foi feita segundo Monteiro et al (1994), e a solubilização dos corpos de inclusão foi feita segundo GLOVER & HAMES (1995) utilizando uréia. Sendo a uréia, um agente desnaturante, também foram utilizados Triton X-100 a 3% e Tricloroacetato de Sódio 2,5 M (TCA) como desnaturantes menos caotrópicos.

Ao precipitado das bactérias induzidas e lisadas foram adicionados 10 volumes de tampão pH 8,0 contendo 50 mM TrisHCl; 2,0 mM $MgSO_4$; 0,1 mM EDTA; 8,0 M uréia e 1,4 mM β -mercaptoetanol, centrifugado por 45 minutos a 6500 x g a temperatura ambiente. O sobrenadante obtido foi aplicado em coluna de DEAE-Sepharose Fast Flow (Pharmacia) (18 x 2 cm) equilibrada com 10mM Tris HCl, pH 9,0 contendo 5mM $MgSO_4$, 1mM EDTA, 1 mM 2-mercaptoetanol e 8M uréia e fluxo de 8 mL/h. A enolase foi eluída em gradiente de NaCl (0-250 mM). As frações com enolase foram reunidas e dialisadas em tampão 15 mM NaH_2PO_4/Na_2HPO_4 , pH 6,2 contendo 3 mM $MgSO_4$, 1mM 2- β mercaptoetanol e quantidades decrescentes, 6, 4, 2, 1 e 0 M de uréia até sua ausência total. Finalmente, o extrato obtido foi aplicado em coluna de CM-Sepharose (Pharmacia) (18 x 2 cm) e fluxo de 8 mL/h a 4°C. A absorbância das frações eluídas foram determinadas a 280 nm e o perfil protéico das amostras de interesse foi analisado por SDS-PAGE 12% (LAEMMLI, 1970).

Crescimento de Xylella fastidiosa e linhagens de E. coli DH5 α e JM101, e extração de suas enolases.

As bactérias *X. fastidiosa* foram crescidas em 100 mL de três meios contendo K_2HPO_4 (1,5 g/L), KH_2PO_4 (1,25 g/L), $MgSO_4 \cdot 7H_2O$ (0,5 g/L), $CaCl_2$ (0,00625 g/L), KNO_3 (0,2 g/L), Pirofosfato Fe^{+3} (0,25 g/L), Solução de vitaminas (10 mL/L) e Glicose (10 g/L) ou Frutose (10 g/L) ou Manose (10 g/L), pH 6,8, só diferenciando a fonte energética entre glicose (meio 1), frutose (meio 2) e manose (meio 3) a 30°C, 150 rpm por 7 dias. As células coletadas por centrifugação (6.500 x g por 15 minutos a 4°C) foram ressuspensas em 1mL tampão 50 mM Tris HCl pH 8,0; 1 mM EDTA pH 8,0; 25 % w/v sacarose e submetidos a ação de um aparelho de ultrassom Branson Sonifier 250 (pulsos de 10 segundos, 60 W de potência por 6 minutos), e então, centrifugadas a 16.000 x g por 5 minutos a 4°C.

As bactérias das linhagens de *E. coli* DH5 α e JM101, foram crescidas em 500mL de meio mínimo (SAMBROOK et al, 1989) a 37°C, 250 rpm por 16 horas e as células coletadas por centrifugação a 6.500 x g por 15 minutos a 4°C, foram ressuspensas em 20 mL de tampão descrito anteriormente e lisadas em “French Pressure Cell Press” (16.000 psi), por 4 vezes, e então centrifugadas a 6.500 x g por 15 minutos a 4°C. O precipitado foi descartado e o sobrenadante obtido foi utilizado no estudo.

Extração da enolase de músculo e fígado de frango

Duas amostras de tecidos de músculo com 0,5 g cada, foram homogeneizadas com auxílio de um homogeneizador de tecido (Ultra Turrax - Marconi) em 5 mL de dois tampões, sendo o tampão 1 contendo 10 mM Tris HCl, pH 9,0 contendo 5 mM $MgSO_4$, 1 mM EDTA, 1 mM β -mercaptoetanol e tampão 2, 15 mM NaH_2PO_4/Na_2HPO_4 , pH 6,2 contendo 3 mM $MgSO_4$, 1 mM β -mercaptoetanol. O mesmo método foi feito com tecidos de fígado, sendo centrifugados 1 ml das soluções por 10 minutos, 16.000 x g a 4°C.

Estudo da ativação da enolase

O estudo da ativação da enolase “in vitro” foi feita a partir das proteínas expressas em corpos de inclusão segundo VUILLARD, L., & FREEMAN, A., 2001. (<http://www.nwfrc.noaa.gov/protocols/inclusion.html>).

Determinação da atividade e estabilidade da enolase

A atividade de enolase foi determinada segundo KURSTRZEBA et al (2000), utilizando 10 µL do sobrenadante dos extratos de *X. fastidiosa*, de *E. coli* e de músculo e fígado de frango e usando como substratos o fosfoenolpiruvato (PEP) e o ácido fosfoglicérico (2-PGA). A quantificação da proteína total no sobrenadante de todas as amostras foi feita pelo método de HARTREE (1972). A atividade das amostras de músculo e fígado de frango foi monitorada durante cinco dias. O estudo da estabilidade da enolase na presença de uréia foi realizada utilizando-se 8 M de uréia aos sobrenadantes, e retirando-os por diálise com o tampão 2. A atividade da enolase foi determinada antes e após a adição e remoção da uréia.

Estudo de isoenzimas das vias de degradação de carboidratos extraídas de *Xylella fastidiosa*

As bactérias *X. fastidiosa* foram crescidas em 100 mL de quatro tipos de meios líquidos a 30°C, 150 rpm por 10 dias e um quinto tipo sólido (30°C por 10 dias). Todos os meios eram compostos de: K₂HPO₄ (2,1 g/L), KH₂PO₄ (0,8 g/L), MgSO₄ 7H₂O (0,4 g/L), Pirofosfato Fé⁺⁺ (0,125 g/L), Glicose (10 g/L), Solução de vitaminas (10 mL/L), Vermelho de Fenol (0,004 g/L) sendo o meio 1 adicionado de Glutamina (0,04 g/L), Metionina (0,004 g/L) e Asparagina (0,01 g/L), o meio 2 sem a adição desses três aminoácidos, o meio 3 adicionado de KNO₃ (0,2 g/L), o meio 4 adicionado de NH₄NO₃ (0,2 g/L) e o meio 5 adicionado de KNO₃ (0,2 g/L) e Ágar (15 g/L), pH 6,8. A *E. coli* DH5α crescida em 100 mL meio mínimo foi usada como padrão.

As bactérias foram lavadas em solução salina (0,85%) e centrifugadas (16.000 x g por 5 minutos a 4°C) duas vezes, sendo ressuspensas em 0,5 mL de tampão TRIS-HCl (0,1 M pH 7,2) contendo 5 mM MgCl₂, 15% de glicerol e 0,1% de 2-β mercaptoetanol, e as células coletadas, ressuspensas novamente nesse mesmo tampão contendo 1 mg/mL de lisozima, e então, incubadas a temperatura ambiente por 10 minutos. A suspensão foi novamente centrifugada (16.000 x g por 5 minutos a 4°C) e o sedimentado ressuspendido em tampão de sonicação (0,1 M TRIS-HCl pH 8,8 com 15% de glicerol), e submetidos a ação de um aparelho de ultrasom Branson Sonifier 250 (pulsos de 10 segundos, 60 W de potência por 10 minutos), centrifugados a 16.000 x g por 5 minutos a 4°C sendo, 50 µL do sobrenadante das amostras de *X. fastidiosa* e 5 µL das amostras de *E. coli* DH5α, aplicados em gel de poliacrilamida 10%.

As atividades das isoenzimas Fosfoglicose Isomerase (*PG* – *EC* 5.3.1.9), Glicose 6-Fosfato Desidrogenase (*G6PDH* – *EC* 1.1.1.49), Aldolase ou Gliceraldeído-3-Fosfato-Liase (*ALD* - *EC* 4.1.2.1.3), Gliceraldeído-3-Fosfato Desidrogenase (*GAPDH* - *EC* 1.2.1.1.2) e Piruvato Quinase (*PK* - *EC* 2.7.1.40) foram analisadas de acordo com os procedimentos descritos em ALFENAS (1998).

Resultados e Discussão

Clonagem, expressão e atividade das enolases recombinantes e nativas

A obtenção de proteínas solúveis na expressão em procariotos (*E. coli*) é geralmente difícil, podendo levar à produção de corpos de inclusão, que são agregados protéicos insolúveis contendo proteínas inativas, com enovelamento irregular e de fácil purificação. Contudo, a solubilização das proteínas expressas, pode ser obtida usando condições desnaturantes e o maior problema, neste caso, é obter uma proteína na sua configuração nativa e com atividade biológica (www.nwfsv.noaa.gov/protocols/inclusion.html). A expressão da enolase da bactéria *X. fastidiosa* não foi uma exceção, e a solubilização destes corpos de inclusão foi feita com uréia 8 M, Triton X-100 ou TCA. Os resultados obtidos mostraram que Triton X-100

e TCA não foram eficientes para romper a membrana dos corpos de inclusão. Assim, a extração da enolase dos corpos de inclusão foi realizada utilizando uréia, que foi o agente solubilizador mais eficiente.

Comparações feitas, por similaridade, pela ferramenta BLASTP da enolase de *X. fastidiosa* com uma gama de outros organismo tanto procariotos como eucariotos confirmou ser a metionina inicial, a localizada na posição do 32º aminoácido, ficando assim com uma constituição na sua sequência de 1290 pares de bases. Esta metionina é o códon iniciador de todos os organismos comparados.

O fato da enolase ser expressa na forma de corpos de inclusão das duas construções, fez com que prosseguissem os experimentos somente com a enolase descrita no Projeto, contendo 1386 pb. Entretanto, após o rompimento dos corpos de inclusão, pelo agente desnaturante uréia, a enzima foi purificada parcialmente, apresentando baixo rendimento.

O perfil eletroforético obtido durante a expressão e purificação da enolase é mostrado na Figura 1. Como pode ser observado, a cromatografia de troca iônica não retirou todas as proteínas contaminantes e a enolase obtida por este procedimento, não apresentou alto grau de pureza. Além disso, a enzima enolase expressa e purificada não apresentou atividade enzimática, fato este que nos levou a realização de testes paralelos, a fim de investigar outros motivos pelos quais a atividade não era obtida.

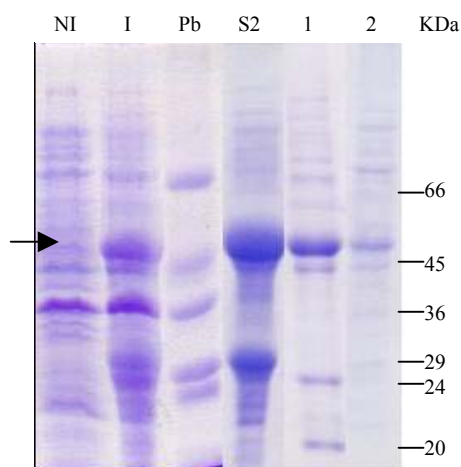


Figura 1. Perfil eletroforético das etapas de purificação da enolase. A eletroforese foi realizada em gel de poliacrilamida a 12% de acordo com o procedimento descrito por Laemmli (1970) e o gel foi corado com comassie blue. Foram aplicados 15 μ L da amostra não induzida (NI), 5 μ L da amostra induzida (I), 15 μ L do sobrenadante 2 (S2), 10 μ L das frações protéicas de interesse (1) após a passagem pela coluna de DEAE, 15 μ L das frações protéicas de interesse (2) após a passagem pela coluna de CM. Pb-padrão de peso molecular baixo em kDa. A flexa da esquerda indica a indução da proteína recombinante.

O crescimento da bactéria *X. fastidiosa* nos meios de cultivo acrescidos de manose e frutose não foi detectado. Segundo SIMPSON et al (2000), a bactéria possui todos os genes que fazem parte da via Glicolítica e uma série de transportadores de carboidratos, sugerindo assim a utilização de glicose, frutose, manose, galactose, glicerol e outros sacarídeos. O não crescimento de *X. fastidiosa* nos meios contendo como fonte de carbono, os carboidratos, manose e frutose são um indício experimental de que o gene da frutose-1,6-bifosfatase não está presente em *X. fastidiosa*, pois sem a presença deste não é possível a metabolização de manose e frutose que são transformadas em frutose-1,6-bifosfato, para entrarem na via glicolítica. A bactéria crescida em meio com glicose teve a atividade da enolase nativa não detectada, sugerindo que a *X. fastidiosa* não utiliza a via glicolítica no metabolismo dos carboidratos. Esta suposição é baseada no fato de que a enolase extraída de tecidos de músculo e fígado de frango apresentaram atividades em relação aos substratos fosfoenolpiruvato (PEP) e ácido fosfoglicérico (2PGA), que permaneceu estável durante cinco dias. As bactérias *E. coli* DH5 α e JM101 também apresentaram atividade da

enolase usando os dois substratos, indicando assim a bifuncionalidade da enolase (Tabela 1).

Tabela 1. Atividade das enolases extraídas de músculo e fígado de frango, de bactérias DH5 α , JM 101 e de *X. fastidiosa* usando dois tampões diferentes (músculo e fígado de frango) e dois substratos, o ácido 2-fosfoglicérico (2-PGA) e o fosfoenolpiruvato (PEP) e 10 μ L do sobrenadante dessas soluções. ND – não detectado.

	Atividade (10^{-4}) (min. μ g de proteína)			
	Tampão 1 (2-PGA)	Tampão 1 (PEP)	Tampão 2 (2-PGA)	Tampão 2 (PEP)
Fígado	1,91	1,39	3,72	2,51
Músculo	1,73	3,68	2,85	5,47
DH5 α	4,53	3,514	-	-
JM101	2,50	2,50	-	-
<i>X. fastidiosa</i>	ND	ND	ND	ND

O estudo do efeito da adição e retirada, por diálise, da uréia nestas amostras inativou a enolase de músculo e fígado de frango, demonstrando assim que a uréia é um desnaturante irreversível da enolase e que ela não pode ser utilizada na extração da enolase dos corpos de inclusão, embora tenha sido o extrator mais eficiente dentre os testados.

O estudo do efeito da guanidina na ativação das proteínas dos corpos de inclusão revelou que a guanidina também é um eficiente agente desnaturante da enolase.

Após a não constatação da atividade da enzima enolase recombinante e da enolase nativa de *X. fastidiosa*, a atividade de algumas isoenzimas da via Glicolítica e da via Entner-Doudoroff foram escolhidas para verificar os possíveis caminhos do metabolismo de carboidratos na bactéria *X. fastidiosa*.

Análise da atividade de isoenzimas da via Glicolítica

O crescimento da bactéria *X. fastidiosa*, para análise de isoenzimas, foi bem sucedido em todos os meios com exceção do meio dois, que não houve crescimento. As atividades das isoenzimas extraídas das bactérias crescidas nesses quatro tipos de meios são mostradas na Tabela 2.

Tabela 2. Atividade enzimática das enzimas Fosfoglicose Isomerase (*PGI*), Glicose 6-Fosfato Desidrogenase (*G6PDH*), Aldolase ou Gliceraldeído-3-Fosfato-Liase (*ALD*), Gliceraldeído-3-Fosfato Desidrogenase (*GAPDH*) e Piruvato Quinase (*PK*), usando as bactérias *X. fastidiosa* crescidas nos meios 1, 3, 4, e 5, e a bactéria *E. coli* DH5 α como padrão, crescida em meio mínimo (meio 6). Sinal (+) indica presença, sinal (–) indica ausência.

Meios de cultura	Atividade enzimática				
	G6PDH	PGI	GAPDH	ALD	PK
1	+	+	–	–	+
3	+	+	–	–	+
4	+	+	–	–	+
5	+	+	–	–	+
6	+	+	+	+	+

Todos os meios eram compostos de: K₂HPO₄ (2,1g/L), KH₂PO₄ (0,8g/L), MgSO₄ 7H₂O (0,4g/L), Pirofosfato Fé⁺⁺ (0,125g/L), Glicose (10g/L), Solução de vitaminas (10mL/L), Vermelho de Fenol (0,1%) (4mL/L), sendo o meio 1 adicionado de Glutamina (0,04g/L), Metionina (0,004g/L) e Asparagina (0,01g/L), o meio 2 sem a adição desses três aminoácidos, o meio 3 adicionado de KNO₃ (0,2g/L), o meio 4 adicionado de NH₄NO₃ (0,2g/L) e o meio 5 adicionado de KNO₃ (0,2g/L) e Ágar (15g/L), pH 6,8.

Os resultados obtidos foram surpreendentes, pois não foram constatadas as atividades das isoenzimas da via Glicolítica, aldolase e gliceraldeído-3-fosfato desidrogenase. Entretanto, a observação da atividade da Glicose-6-Fosfato Desidrogenase sugere que a *X. fastidiosa* utiliza a via Entner-Doudoroff para a formação de piruvato. Um outro indício, que confirma tal fato, é a não identificação de ORFs (Open Reads Frames) cujos produtos são similares a fosfogluconato desidrogenase, no genoma da *X. fastidiosa*. Dessa forma seria impossível utilizar a via das Pentoses para a obtenção de ribose-5-fosfato. Outra possibilidade que não pode ser descartada seria algum tipo de regulação dos genes relacionados ao metabolismo de carboidratos, devido às condições de crescimento da bactéria. Além disso, as enzimas glicolíticas são difíceis de serem detectadas, devido a uma gama de fatores que influenciam na sua detecção como, por exemplo, a metodologia utilizada, estágio específico de desenvolvimento do organismo, componentes do meio nutritivo, estresse, condições de crescimento e outros que podem resultar em uma situação complexa elucidando as relações quantitativas e respectivas funções metabólicas, e ainda propriedades regulatórias de diferentes isoformas (RIVOAL et al., 2002).

A utilização de carboidratos, principalmente glicose, como fonte de carbono e energia é praticamente universal entre as bactérias, arqueobactérias e eucariotos. A forma como tais carboidratos são metabolizados pode, entretanto, diferir consideravelmente (ROMANO & CONWAY, 1996).

A rota metabólica presente na quase totalidade das bactérias modernas e eucariotos para o metabolismo de monossacarídeos, provindos diretamente do meio ambiente ou da degradação dos diversos tipos de carboidratos, é a glicólise, ou via de Embden-Meyerhoff-Parnas. Os substratos iniciais de tal rota são geralmente glicose-6-fosfato ou frutose-6-fosfato. O produto final da rota é o piruvato que constitui um dos principais intermediários do metabolismo central devido a sua utilização em diversas vias, tanto catabólicas como anabólicas.

Em termos de metabolismo energético, foram identificadas todas as enzimas da glicólise em *X. fastidiosa*. Entre as vias aparentemente ausentes estão a gliconeogênese, pois não foram identificadas a piruvato carboxilase, fosfoenolpiruvato carboxiquinase nem frutose 1,6 bisfosfatase, que são requeridas para retornar ao passo irreversível na glicólise (SIMPSON et al., 2000).

A principal importância da gliconeogênese seria a obtenção de glicose, ou derivados fosforilados nas posições 1 ou 6, para serem utilizados em processos biossintéticos, como a síntese da parede celular, a partir de compostos não-glicídicos como: intermediários do ciclo dos ácidos tricarboxílicos, glicerol, lactato e aminoácidos gliconeogênicos. Assim, na falta de glicose, a degradação de tais compostos pode supri-la ao metabolismo (VOET & VOET, 1995).

É possível que entre um grande número de genes não identificados de *X. fastidiosa* haja genes não homólogos que compensem os passos em vias críticas (SIMPSON et al., 2000). A tendência inicial é considerar que seja um caso de substituição de genes não homólogos. Entretanto, se confirmada a real ausência da enzima frutose 1-6-bifosfatase, em *X. fastidiosa*, essa também poderia ser uma possível explicação do elevado tempo de duplicação da bactéria. Visto que, sem uma gliconeogênese funcional, *X. fastidiosa* não poderia dispor de outros substratos para a síntese da parede celular que não os próprios carboidratos que absorvesse, haveria uma constante competição para a utilização dos carboidratos, diminuindo a taxa de síntese da parede celular e aumentando o tempo de duplicação (FERREIRA, 2000).

Considerando uma linha comparativa, visto que já foram seqüenciadas e anotadas duas bactérias próximas a *X. fastidiosa*: *E. coli* e *Haemophilus influenzae*,

podemos ter que, o fato de terem sido identificadas as ORFs das enzimas da maioria das vias metabólicas pode somente estar relacionado a uma maior similaridade entre as proteínas de *X. fastidiosa* e as já presentes no banco de dados. De outra forma, poderíamos inferir que, pelo fato do xilema ser um meio pobre em comparação aos habitats de patógenos animais, *X. fastidiosa* necessita do metabolismo central completo, pois precisa sintetizar senão todos, pelo menos a grande maioria dos intermediários.

Nos organismos em que a glicólise não está presente, e mesmo em grande parte dos organismos que a utilizam como principal via de degradação de monossacarídeos, existe uma via metabólica, ou variações da mesma, denominada via de Entner-Doudoroff. Essa via utiliza glicose-6-fosfato como substrato inicial, seguindo para glucono-1,5-lactona-6-fosfato, 6-fosfogluconato e 2-desidro-3-desoxifosfogluconato (Figura 2). Normalmente é utilizada em bactérias para o metabolismo de glucoronato, com um balanço energético líquido de um mol de ATP por mol de glucoronato ou de glicose, dependendo do substrato inicial (ROMANO & CONWAY, 1996).

Foram descobertas, entretanto, na arqueobactéria hipertermofílica *Pyrococcus furiosus* tanto uma hexoquinase como uma fosfofrutoquinase dependentes de ADP (formadoras de AMP), o que indica a utilização da glicólise como a principal via de metabolismo de glicose, como comprovado por estudos de atividade enzimática e marcação com carbono 13 (KENGEN et al., 1994). Tal descoberta reforça a suposição de que a simples comparação de seqüências de proteínas homólogas é um mecanismo preliminar para a identificação de vias metabólicas, mas que não deve ser tomado como totalmente certo, visto a possibilidade da existência de enzimas análogas que desempenhem funções conhecidas de forma completamente diferente da já descrita.

Foram identificadas no genoma de *X. fastidiosa*, ORFs codificando para as enzimas mais comuns na glicólise e também para as enzimas da via de Entner-Doudoroff, o que está de acordo com sua similaridade com *E. coli*. Assim, havendo carboidratos disponíveis, a obtenção de piruvato está de certa forma garantida por tais vias.

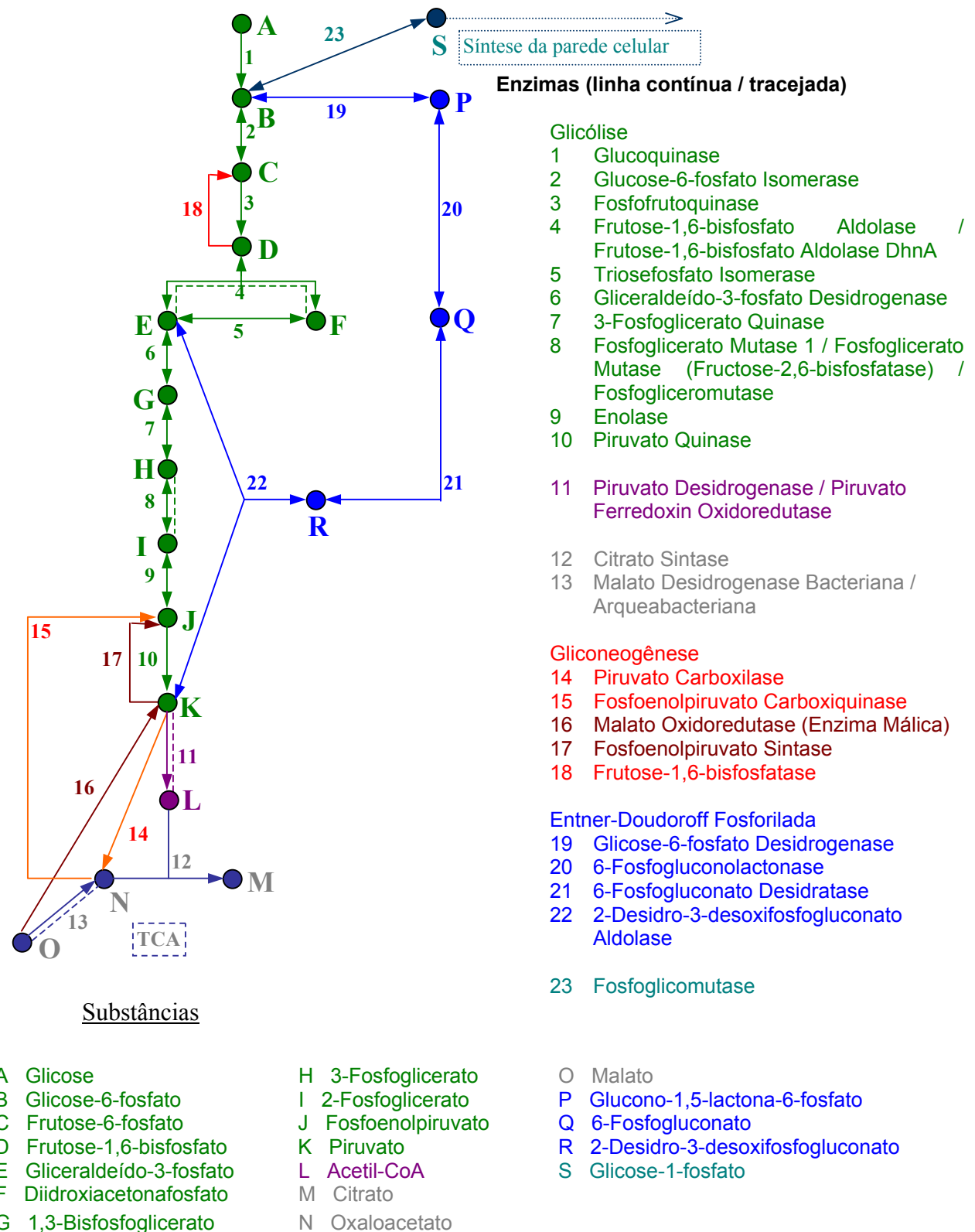


Figura 2. Vias glicolíticas: Glicólise, Glioneogênese, Entner-Doudoroff (A.J.S.Ferreira).

O papel fisiológico da via Entner-Doudoroff em *E. coli* ainda não está claro. A descoberta da pirroloquinolina quinona (PQQ) dependente da glicose desidrogenase, que catalisa a oxidação da glicose a ácido glucônico no periplasma, sugere uma rota alternativa para catabolismo da glicose em *E. coli* (HOMMES et al., 1984). Em *X. fastidiosa* foram encontrados duas orfs (XF-1740 e XF-2259) que codificam esse gene.

A baixa concentração do fosfato promove o uso da via Entner-Doudoroff indiretamente pelo acesso a PQQ no periplasma, ao contrário da diretamente, por ativação dos genes *edd*-codifica a enzima 6-fosfogluconato desidratase, e *eda*, codifica a enzima 2-ceto-3-deoxi-6-fosfogluconato (KDPG) aldolase. (FLIEGE et al., 1992).

Um levantamento dos genomas publicados e genomas parcialmente completados indicam que a via Entner-Doudoroff está presente em vários organismos. Isto não é surpresa, dado que levantamentos bioquímicos indicam que as enzimas desta via são amplamente distribuídas entre as bactérias e estão presentes em todos os três domínios filogenéticos, incluindo o mais profundamente estabelecido, o *Archae* (PEEKHAUS & CONWAY, 1998). Recentemente, com o seqüenciamento de três organismos filogeneticamente próximos a *X. fastidiosa*, a *Xanthomonas campestris* pv. *campestris* e *Xanthomonas axonopodis* pv. *citri* (DA SILVA et al., 2002) e a estirpe Termecula de *X. fastidiosa* responsável pela Doença de Pierce em videira (comunicação pessoal), foi proposto que todos possuem as duas enzimas descritas anteriormente, responsáveis pelo funcionamento desta via. Fato este verificado pela grande semelhança de “habitats” desses organismos fitopatogênicos e seus constituintes.

Em análises comparativas das enzimas glicolíticas de 17 genomas seqüenciados completamente (*Archae*, bactéria e levedura), revelaram uma surpreendente plasticidade da via glicolítica, mostrando que as enzimas da Via Entner-Doudoroff não são tão bem conservadas, quanto à da segunda fase da glicólise, nos organismos estudados. Fato este que pode ser explicado pela coexistência das duas vias em várias espécies (DANDEKAR et al., 1999).

Em estudos de genes repressores de glicose em *Bacillus subtilis*, que são genes induzidos após o consumo de fontes de carbono glicolítico, constatou que a expressão

destes genes permaneceu inibida pela adição de glicose. Este fato é devido a presença em *B. subtilis*, da CcpA, que é uma proteína de controle catabólico. Em análises de “microarray” os autores detectaram 66 genes repressores de glicose, dos quais 22 estão, pelo menos, parcialmente sob controle independentes de CcpA (YOSHIDA et al., 2001).

Conclusões

Os resultados obtidos no presente trabalho nos permitiram concluir que:

- a enolase foi expressa em corpos de inclusão, no sistema de expressão utilizado nesse estudo;
- a uréia é um agente desnaturante irreversível da enolase, além de ter sido o extrator mais eficiente dos corpos de inclusão;
- a não detecção da atividade da enolase nativa, aldolase e gliceraldeído-3-fosfato sugere que a *X. fastidiosa* não utiliza a via glicolítica no metabolismo de carboidratos,
- a *X. fastidiosa* provavelmente utiliza a via Entner-Doudoroff na metabolização dos carboidratos.

Referências

- ALFENAS, A.C. (Ed). **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. Viçosa: Sociedade de Investigações Florestais. 1998.
- BIRREN, B., GREEN, E.D. KLAPHOLZ, MYERS, R.M., ROSKAMS, J. **Genome Analysis. A Laboratory Manual**. Cold Spring Harbor Laboratory Press. 1997.
- CHANG, C.J., GARNIER, M., ZREIK, L., ROSSETTI, V., BOVE, J.M. CHANG, C. J. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. **Curr. Microbiol**, v.27, p.137-42, 1993.

- DA SIVA, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B. et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. **Nature**, v. 417, p. 459-463, 2002.
- DANDEKAR, T., SCHUSTER, S., SNEL, B., HUYNEN, M., BORK, P. Pathway alignment: application to the comparative analysis of glycolytic enzymes. **J. Biochem**, v. 343, p. 115-124, 1999.
- FERREIRA, A.J.S. **Seqüenciamento de parte do genoma de *Xylella fastidiosa* e análise do metabolismo de carboidratos no genoma anotado**. São Paulo, 2000. 120 f. Dissertação (mestrado em Bioquímica). Instituto de Química, Universidade de São Paulo. São Paulo.
- FLIEGE, R., TONG, S., SHIBATA, A., NICKERSON, K.W., CONWAY, T. The Entner-Doudoroff pathway in *Escherichia coli* is induced for oxidative glucose metabolism via pyrroloquinoline quinone- dependent glucose dehydrogenase. **Appl. and Environ. Microbiol**, v. 58 n. 12, p. 3826-3829, 1992.
- GLOVER, D.M., HAMES, B.D. **DNA Cloning 2: A practical approach – Expression systems**. 2nd. Ed. New York: Oxford University Press, 1995. p. 67-69.
- GRIFFIN, H.G., GRIFFIN, A.M. **Methods in molecular biology. DNA sequencing protocols**. Totowa, Humana Press Inc., NJ, 1993. p. 23.
- HARTREE, E.F. Determination of protein: a modification of the Lowry method that gives a linear photometric response. **Anal. Biochem**, v. 48, p. 422-427, 1972.
- HARTUNG, J.S., BERETA, M.J.G., BRLANSKY, R.H., SPISSO, J., LEE, R.F. Citrus variegated chlorosis bacterium: axenic culture, pathogenicity and relationships with other strains of *Xylella fastidiosa*. **Phytopathol**, v. 64, p. 591-597, 1994.
- HOMMES, R.W.J., POSTMA, R.W., NEIJSEL, O.M., TEMPEST, D.W., DOKTER, P., DUINE, J.A. Evidence for quinoprotein glucose dehydrogenase apoenzyme in several strains of *Escherichia coli*. **FEMS Microbiol. Lett**, v. 24, p. 1904-1910, 1984.
- KENGEN, S.W.M., DE BOK, F.A.M., VAN LOO, N.D., DIJKEMA, C., STAMS, A.J.M., DE VOS, W.M. Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. **J. Biol. Chem**, v. 269, n. 26, p. 17537-17541, 1994.

- KUSTRZEBA-WÓJCICKA, I., GOLCZAK, M. Enolase from *Candida albicans* – purification and characterization. **Comp. Biochem. Physiol**, v. B 126, p. 109-120, 2000.
- LAEMMLI, V.K. Cleavage of structural proteins during the assembly of the head bacteriophage T4. **Nature**, v. 227, p. 680-685, 1970.
- MACHADO, M. Citrus Variegated Chlorosis (CVC), a new destructive citrus disease in Brazil, and the xylem limited bacteria, *Xylella fastidiosa*. <<http://www.dcc.unicamp.br/genoma/xylella.html>>. 1997.
- MONTEIRO, P.B., LATARO, R.C., FERRO, J.A., REINACH, F.C. Functional α -tropomyosin produced in *E.coli*: a dipeptide extension can substitute the amino terminal acetyl group. **J. Biol. Chem**, v.269, p. 10461-10466, 1994.
- PEEKHAUS, N., CONWAY, T. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. **J. bacterial**, v. 180, n. 14, p. 3495-3502, 1998.
- PURCELL, A.H., HOPKINS, D.L. Fastidious xylem-limited bacterial plant pathogens. **Annu. Rev. Phytopathol**, v. 34, p. 131-151, 1996.
- RIVOAL, J., SMITH, C.R., MORAES, T.F., TURPIN, D.H., PLAXTON, W.C. A method for activity stain after native polyacrylamide gel electrophoresis using coupled enzyme assay and fluorescence detection: application to the analysis of several glycolytic enzymes. **Anal. Biochem**, v. 300, p. 94-99, 2002.
- ROMANO, A.H., CONWAY, T. Evolution of carbohydrate metabolic pathways, 14th Forum in Microbiology, **Res. Microbiol**, v. 147, p. 448-455, 1996.
- ROSSETTI, V., GARNIER, M., BOVÉ, J.M., BERETA, M.J.G., TEIXEIRA, A.R.R., QUAGIO, J.A., DE NEGRI, J.D. Présence de bactéries dans le xylème d'orangers atteints de chlorose variégée, une nouvelle maladie des agrumes au Brésil. **C. R. Acad. Sci.**, Paris, v. 310, p. 345-349, 1990.
- ROSSETTI, V., GONZALES, M.A., DONADIO, L.C. History. In: DONADIO, L.C. AND MOREIRA, C.S. (eds). **Citrus Variegated Chlorosis**, Bebedouro, 1998.
- SAMBROOK, J., FRITSCH, E.E., MANIATIS, T. **Molecular Cloning: A Laboratory Guide**, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 1989.

SIMPSON, A.J., REINACH, F.C., ARRUDA, P., ABREU, F.A., ACENCIO, M., ALVARENGA, R. et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. **Nature**, v. 406, p. 151-157, 2000.

VOET, D., VOET, J.G. **Biochemistry**, 2nd edition, John Wiley & Sons, New York. 1995.

WELLS, J., RAJU, B., JUNG, H., WEISBURG, W., MANDELCO-PAUL, L., BRENNER, D. *Xylella fastidiosa* gen nov, sp nov gramnegative, xylem limited fastidious plant bacteria related to *Xanthomonas* ssp., **Int. J. of Syst. Bacteriol**, v. 37, p. 136-143, 1987.

WOLD, F. **In the Enzymes** (Boyer, P.D. Ed.) Vol. 5, 3rd ed. New York. Academic Press. p. 499-538. 1971.

YOSHIDA, K., KOBAYASHI, K., MIWA, Y., KANG, C-M., MATSUNAGA, M. et al. Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. **Nucleic Acids Res**, v. 29, n. 3, p. 683-692, 2001.

APÊNDICE

The genome sequence of the plant pathogen *Xylella fastidiosa*

The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis*, São Paulo, Brazil

* A full list of authors appears at the end of this paper

Xylella fastidiosa is a fastidious, xylem-limited bacterium that causes a range of economically important plant diseases. Here we report the complete genome sequence of *X. fastidiosa* clone 9a5c, which causes citrus variegated chlorosis—a serious disease of orange trees. The genome comprises a 52.7% GC-rich 2,679,305-base-pair (bp) circular chromosome and two plasmids of 51,158 bp and 1,285 bp. We can assign putative functions to 47% of the 2,904 predicted coding regions. Efficient metabolic functions are predicted, with sugars as the principal energy and carbon source, supporting existence in the nutrient-poor xylem sap. The mechanisms associated with pathogenicity and virulence involve toxins, antibiotics and ion sequestration systems, as well as bacterium–bacterium and bacterium–host interactions mediated by a range of proteins. Orthologues of some of these proteins have only been identified in animal and human pathogens; their presence in *X. fastidiosa* indicates that the molecular basis for bacterial pathogenicity is both conserved and independent of host. At least 83 genes are bacteriophage-derived and include virulence-associated genes from other bacteria, providing direct evidence of phage-mediated horizontal gene transfer.

Citrus variegated chlorosis (CVC), which was first recorded in Brazil in 1987, affects all commercial sweet orange varieties¹. Symptoms include conspicuous variegations on older leaves, with chlorotic areas on the upper side and corresponding light brown lesions, with gum-like material on the lower side. Affected fruits are small, hardened and of no commercial value. A strain of *Xylella fastidiosa* was first identified as the causal bacterium in 1993 (ref. 2) and found to be transmitted by sharpshooter leafhoppers in 1996 (ref. 3). CVC control is at present limited to removing infected shoots by pruning, the application of insecticides and the use of healthy plants for new orchards. In addition to CVC, other strains of *X. fastidiosa* cause a range of economically important plant diseases including Pierce's disease of grapevine, alfalfa dwarf, phony peach disease, periwinkle wilt and leaf scorch of plum, and are also associated with diseases in mulberry, pear, almond, elm, sycamore, oak, maple, pecan and coffee⁴. The triply cloned *X. fastidiosa* 9a5c, sequenced here, was derived from the pathogenic culture 8.1b obtained in 1992 in Bordeaux (France) from CVC-affected Valencia sweet orange twigs collected in Macaubal (São Paulo, Brazil) on May 21, 1992 (ref. 2). Strain 9a5c produces typical CVC symptoms on inoculation into experimental citrus plants⁵, and into *Nicotiana tabacum* (S. A. Lopes, personal communication) and *Catharantus roseus* (P. Brant-Monteiro, personal communication)—two novel experimental hosts.

General features of the genome

The basic features of the genome are listed in Table 1 and a detailed map is shown in Fig. 1. The conserved origin of replication of the large chromosome has been identified in a region between the putative 50S ribosomal protein L34 and *gyrB* genes containing *dnaA*, *dnaN* and *recF*⁶. The *Escherichia coli* DnaA box consensus sequence TTATCCACA is found on both DNA strands close to *dnaA*. In addition, there are typical 13-nucleotide (ACCACCAC-CACCA) and 9-nucleotide (two TTTCATTGG and two TTTTAT-TATT) sequences in other intergenic sequences of this region. This region is coincident with the calculated GC-skew signal inversion⁷. We have designated base 1 of the *X. fastidiosa* genome as the first T of the only TTTTAT sequence found between the ribosomal protein L34 gene and *dnaA*.

The overall percentage of open reading frames (ORFs) for which a putative biological function could be assigned (47%) was slightly below that for other sequenced genomes such as *Thermotoga maritima*⁸ (54%), *Deinococcus radiodurans*⁹ (52.5%) and *Neisseria*

*meningitidis*¹⁰ (53.7%). This may reflect the lack of previous complete genome sequences from phytopathogenic bacteria. Plasmid pXF1.3 contains only two ORFs, one of which encodes a replication-associated protein. Plasmid pXF51 contains 64 ORFs, of which 5 encode proteins involved in replication or plasmid stability and 20 encode proteins potentially involved in conjugative transfer. One ORF encodes a protein similar to the virulence-associated protein D (VapD), found in many other bacterial pathogens¹¹. Four regions of pXF51 present significant DNA similarity to parts of transposons found in plasmids from other bacteria, suggesting interspecific horizontal exchange of genetic material.

The principal paralogous families are summarized in Table 2. The complete list of ORFs with assigned function is shown in Table 3. Seventy-five proteins present in the 21 completely sequenced genomes in the COG database¹² (as of 15th March 2000) were also found in *X. fastidiosa*. Each of these sequences was used to

Table 1 General features of the *Xylella fastidiosa* 9a5c genome

Main chromosome	
Length (bp)	2,679,305
G+C ratio	52.7%
Open reading frames (ORFs)	2,782
Coding region (% of chromosome size)	88.0%
Average ORF length (bp)	799
ORFs with functional assignment	1,283
ORFs with matches to conserved hypothetical proteins	310
ORFs without significant data base match	1,083
Ribosomal RNA operons	2
	(16SrRNA-Ala-TGC-tRNA-Ile-GAT-tRNA-23SrRNA-5SrRNA)
tRNAs	49 (46 different sequences corresponding to all 20 amino acids)
tmRNA	1
Plasmid pXF51	
Length (bp)	51,158
G+C ratio	49.6%
Open reading frames (ORFs)	64
Protein coding region (% of plasmid size)	86.9%
ORFs with functional assignment	30
ORFs with matches to conserved hypothetical proteins	8
ORFs without significant data base match	24
Plasmid pXF1.3	
Length (bp)	1,285
G+C ratio	55.6%
Open reading frames (ORFs)	2
ORFs with functional assignment	1

Table 2 Largest families of paralogous genes

Family (total number of families = 312)	Number of genes (total number of genes = 853)
ATP-binding subunits of ABC transporters	23
Reductases/dehydrogenases	12
Two-component system, regulatory proteins	12
Hypothetical proteins	10
Transcriptional regulators	9
Fimbrial proteins	9
Two-component system, sensor proteins	9

generate a phylogenetic tree of the 22 organisms. In 69% of such trees, *X. fastidiosa* was grouped with *Haemophilus influenzae* and *E. coli*, consistent with a phylogenetic analysis undertaken with the 16S rRNA gene¹³.

One ORF, a cytosine methyltransferase (XF1774), is interrupted by a Group II intron. The intron was identified on the basis of the presence of a reverse transcriptase-like gene (as in other Group II introns), conserved splice sites, conserved sequence in structure V and conserved elements of secondary structure¹⁴. Group II introns are rare in prokaryotes, but have been found in different evolutive lineages including *E. coli*, cyanobacteria and proteobacteria¹⁵.

Transcription, translation and repair

The basic transcriptional and translational machinery of *X. fastidiosa* is similar to that of *E. coli*¹⁶. Recombinational repair, nucleotide and base-excision repair, and transcription-coupled repair are present with some noteworthy features. For example, no photolyase was found, indicating exclusively dark repair. Although the main genes of the SOS pathway, *recA* and *lexA*, are present, ORFs corresponding to the three DNA polymerases induced by SOS in *E. coli* (DNA polymerases II, IV and V)¹⁷ are missing, indicating that the mutational pathway itself may be distinct.

Energy metabolism

Even though *X. fastidiosa* is, as its name suggests, a fastidious organism, energy production is apparently efficient. In addition to all the genes for the glycolytic pathway, all genes for the tricarboxylic acid cycle and oxidative and electron transport chains are present. ATP synthesis is driven by the resulting chemiosmotic proton gradient and occurs by an F-type ATP synthase. Fructose, mannose and glycerol can be utilized in addition to glucose in the glycolytic pathway. There is a complete pathway for hydrolysis of cellulose to glucose, consisting of 1,4- β -cellobiosidase, endo-1,4- β -glucanase and β -glucosidase, suggesting that cellulose breakdown may supplement the often low concentrations of monosaccharides in the xylem¹⁸. Two lipases are encoded in the genome, but there is no β -oxidation pathway for the hydrolysis of fatty acids, presumably precluding their utilization as an alternative carbon and energy source. Likewise, although enzymes required for the breakdown of threonine, serine, glycine, alanine, aspartate and glutamate are present, pathways for the catabolism of the other naturally occurring amino acids are incomplete or absent.

The gluconeogenesis pathway appears to be incomplete. Phosphoenolpyruvate carboxykinase and the gluconeogenic enzyme fructose-1,6-bisphosphatase, which are required to bypass the irreversible step in glycolysis, are not present. The absence of the first is compensated by the presence of phosphoenolpyruvate synthase and malate oxidoreductase, which together can generate phosphoenolpyruvate from malate. There appears, however, to be no known compensating pathway for the absence of fructose-1,6-bisphosphatase. It is possible that among the large number of unidentified *X. fastidiosa* genes there are non-homologous genes that compensate for steps in such critical pathways. Barring this possibility, however, the absence of a functional gluconeogenesis pathway implies a strict dependence on carbohydrates both as a

source of energy and anabolic precursors. The glyoxylate cycle is absent and the pentose phosphate pathway is incomplete. In the latter pathway, genes for neither 6-phosphogluconic dehydrogenase nor transaldolase were identified.

Small molecule metabolism

X. fastidiosa exhibits extensive biosynthetic capabilities, presumably an absolute requirement for a xylem-dwelling bacterium. Most of the genes found in *E. coli* necessary for the synthesis of all amino acids from chorismate, pyruvate, 3-phosphoglycerate, glutamate and oxaloacetic acid¹⁶ were identified. However, some genes in *X. fastidiosa* are bi-functional, such as phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphatase (XF2213), aspartokinase/homoserine dehydrogenase I (XF2225), imidazole-glycerolphosphate dehydratase/histidinol-phosphate phosphatase (XF2217) and a new diaminopimelate decarboxylase/aspartate kinase (XF1116) that would catalyse the first and the last steps of lysine biosynthesis. In addition, the gene for acetylglutamate kinase (XF1001) has an acetyltransferase domain at its carboxy-terminal end that would compensate for the missing acetyltransferase in the arginine biosynthesis pathway. Other missing genes include phosphoserine phosphatase, cystathionine β -lyase, homoserine *O*-succinyltransferase and 2,4,5-methyltetrahydrofolate-homocysteine methyltransferase. The first two enzymes are also absent in the *Bacillus subtilis* genome, the third is absent in *Haemophilus influenzae* and the fourth is missing in both genomes¹². We thus presume that alternative, unidentified enzymes complete the biosynthetic pathways in these organisms and in *X. fastidiosa*.

The pathways for the synthesis of purines, pyrimidines and nucleotides are all complete. *X. fastidiosa* is also apparently capable of both synthesizing and elongating fatty acids from acetate. Again, however, some *E. coli* enzymes were not found, such as holo acyl-carrier-protein synthase (also absent in *Synechocystis* sp., *H. influenzae* and *Mycoplasma genitalium*) and enoyl-ACP reductase (NADPH) (*FabI*) (also absent from *M. genitalium*, *Borrelia burgdorferi* and *Treponema pallidum*)¹².

X. fastidiosa appears to be capable of synthesizing an extensive variety of enzyme cofactors and prosthetic groups, including biotin, folic acid, pantothenate and coenzyme A, ubiquinone, glutathione, thioredoxin, glutaredoxin, riboflavin, FMN, FAD, pyrimidine nucleotides, porphyrin, thiamin, pyridoxal 5'-phosphate and lipolate. In a number of the synthetic pathways, one or more of the enzymes present in *E. coli* are absent, but this is also true for at least one other sequenced Gram-negative bacterial genome in each case¹². We therefore again infer that the missing enzymes are either not essential or replaced by unknown proteins with novel structures.

Transport-related proteins

A total of 140 genes encoding transport-related proteins were identified, representing 4.8% of all ORFs. For comparison, *E. coli*, *B. subtilis* and *M. genitalium* have around 10% of genes encoding transport proteins, whereas *Helicobacter pylori*, *Synechocystis* sp. and *Methanococcus jannaschii* have 3.5–5.4% (ref. 19). Transport systems are central components of the host–pathogen relationship (Fig. 2). There are a number of ion transporters and transporters for the uptake of carbohydrates, amino acids, peptides, nitrate/nitrite, sulphate, phosphate and vitamin B12. Many different transport

Figure 1 Linear representation of the main chromosome and plasmids pXF51 and pXF1.3 of the *Xylella fastidiosa* genome. Genes are coloured according to their biological role. Arrows indicate the direction of transcription. Genes with frameshift and point mutations are indicated with an X. Ribosomal RNA genes, the tmRNA, the principal repeats, prophages and the group II intron are indicated by coloured lines. Transfer RNAs are identified by a single letter identifying the amino acid. Pie chart represents the distribution of the number of genes according to biological role. The numbers below protein-producing genes correspond to gene IDs.

families are represented and include both small and large mechanosensitive conductance ion channels, a monovalent cation:proton antiporter (CAP-2) and a glycerol facilitator belonging to the major intrinsic protein (MIP) family. In addition, 23 ABC transport systems comprising 41 genes can be identified. *X. fastidiosa* appears to possess a phosphotransferase system (PTS) that typically mediates small carbohydrate uptake. There are both the enzyme I and HPr components of this system, as well as a gene supposedly involved in its regulation (*pstK* or *hprK*); however, there is no PTS permease—an essential component of the phosphotransferase complex. The functionality of the system therefore remains in question.

There are five outer membrane receptors, including siderophores, ferrichrome-iron and haemin receptors, which are all associated with iron transport. The energizing complexes, TonB–ExbB–ExbD and the paralogous TolA–TolR–TolQ, essential for the functioning of the outer membrane receptors, are also present. In all, 67 genes encode proteins involved in iron metabolism. We propose that in *X. fastidiosa* the uptake of iron and possibly of other transition metal ions such as manganese causes a reduction in essential micronutrients in the plant xylem, contributing to the typical symptoms of leaf variegation.

The *X. fastidiosa* genome encodes a battery of proteins that mediate drug inactivation and detoxification, alteration of potential drug targets, prevention of drug entry and active extrusion of drugs and toxins. These include ABC transporters and transport processes driven by a proton gradient. Of the latter, eight belong to the hydrophobe/amphiphile efflux-1 (HAE1) family, which act as multidrug resistance factors.

Adhesion

X. fastidiosa is characteristically observed embedded in an extracellular translucent matrix *in planta*²⁰. Clumps of bacteria form within the xylem vessels leading to their blockage and symptoms of the disease such as water-stress leaf curling. We deduce, from our analysis of the complete genome sequence, that the matrix is composed of extracellular polysaccharides (EPSs) synthesized by enzymes closely related to those of *Xanthomonas campestris* pv *campestris* (Xcc) that produce what is commercially known as xanthan gum. In comparison with Xcc, however, we did not find *gumI* (encoding glycosyltransferase V, which incorporates the terminal mannose), *gumL* (encoding ketolase which adds pyruvate to the polymer) or *gumG* (encoding acetyltransferase which adds acetate), suggesting that *Xylella* gum may be less viscous than its *Xanthomonas* counterpart.

Positive regulation of the synthesis of extracellular enzymes and EPS in *Xanthomonas* is effected by proteins coded by the *rfp* (regulation of pathogenicity factors) gene cluster²¹. Mutations in any of these genes in *Xanthomonas* results in failure to synthesize the EPS. In consequence, the strain becomes non-pathogenic²¹. *X. fastidiosa* contains genes that encode RpfA, RpfB, RpfC and RpfF, suggesting that both bacteria may regulate the synthesis of pathogenic EPS factors through similar mechanisms.

Fimbria-like structures are readily apparent upon electron microscopical observation of *X. fastidiosa* within both its plant and insect hosts²². Because of the high velocity of xylem sap passing through narrow portions of the insect foregut, fimbria-mediated attachment may be essential for insect colonization. Indeed, in the insect mouthparts the bacteria are attached in ordered arrays, indicating specific and polarized adhesion²³. In addition, fimbriae are thought to be involved in both plant–bacterium and bacterium–bacterium interactions during colonization of the xylem itself. We identified 26 genes encoding proteins responsible for the biogenesis and function of Type 4 fimbria filaments. This type of fimbria is found at the poles of a wide range of bacterial pathogens where they act to mediate adhesion and translocation along epithelial surfaces²⁴. The genes include *pilS* and *pilR* homologues, which encode a two-component

system controlling transcription of fimbrial subunits, presumably in response to host cues, and *pilG*, *H*, *I*, *J* and *chpA*, which encode a chemotactic system transducing environmental signals to the pilus machinery.

In addition to the EPS and fimbriae, which are likely to have central roles in the clumping of bacteria and in adhesion to the xylem walls, we also identified outer membrane protein homologues for afimbrial adhesins. Although fimbrial adhesins are well characterized as crucial virulence factors in both plant and human pathogens²⁵, afimbrial adhesins, which are directly associated with the bacterial cell surface, have been hitherto associated only with human and animal pathogens, where they promote adherence to epithelial tissue. Of the three putative adhesins of this kind identified in *X. fastidiosa*, two exhibit significant similarity to each other (XF1981, XF1529) and to the *hsf* and *hia* gene products of *H. influenzae*²⁶. The third (XF1516) is similar to the *uspA1* gene product of *Moraxella catarrhalis*²⁷. All these proteins share the common C-terminal domain of the autotransporter family²⁸. Direct experimentation will be required to establish whether these adhesins promote binding to plant cell structures or components of the insect vector foregut, or both. Nevertheless, their presence in the *X. fastidiosa* genome adds to the increasing evidence for the generality of mechanisms of bacterial pathogenicity, irrespective of the host organism²⁹.

We also identified three different haemagglutinin-like genes. Again, similar genes have not previously been identified in plant pathogens. These genes (XF2775, XF2196, XF0889) are the largest in the genome and exhibit highest similarity to a *Neisseria meningitidis* putative secreted protein¹⁰.

Intervessel migration

Movement between individual xylem vessels is crucial for effective colonization by *X. fastidiosa*. For this to occur, degradation of the pit membrane of the xylem vessel is required. Of the known pectolytic enzymes capable of this function, a polygalacturonase precursor and a cellulase were identified, although the former contains an authentic frameshift. These genes exhibited highest similarity to orthologues in *Ralstonia solanacearum*—which causes wilt disease in tomatoes—where the polygalacturonase genes are required for wild-type virulence.

Toxicity

We identified five haemolysin-like genes: haemolysin III (XF0175), which belongs to an uncharacterized protein family, and four others (XF0668, XF1011, XF2407, XF2759) which belong to the RTX toxin family that contains tandemly repeated glycine-rich nonapeptide motifs at the C-terminal domain. One of these ORFs is closely related to bacteriocin, an RTX toxin also found in the plant bacterium *Rhizobium leguminosarum*³⁰. RTX or RTX-like proteins are important virulence factors widely distributed among Gram-negative pathogenic bacteria³¹.

There are two Colicin-V-like precursor proteins. Colicin V is an antibacterial polypeptide toxin produced by *E. coli*, which acts against closely related sensitive bacteria³². The precursors consist of 102-amino-acid peptides (XF0262, XF0263) that have the typical conserved leader 15-amino-acid motif, and have some similarity with Colicin V from *E. coli* at the remaining C-terminal portion. The necessary apparatus for Colicin biosynthesis and secretion is also present. Interestingly, in *E. coli* most of the genes necessary for biogenesis and export of Colicin V are in a gene cluster present in a plasmid, whereas in *X. fastidiosa* these genes are dispersed in the chromosome.

We found four genes that may function in polyketide biogenesis: polyketide synthase (PKS), pteridine-dependent deoxygenase, daunorubicin C-13 ketoreductase and a NonF-related protein. These genes belong to the synthesis pathways of frenolicin, rapamycin, daunorubicin and nonactin, respectively. These pathways

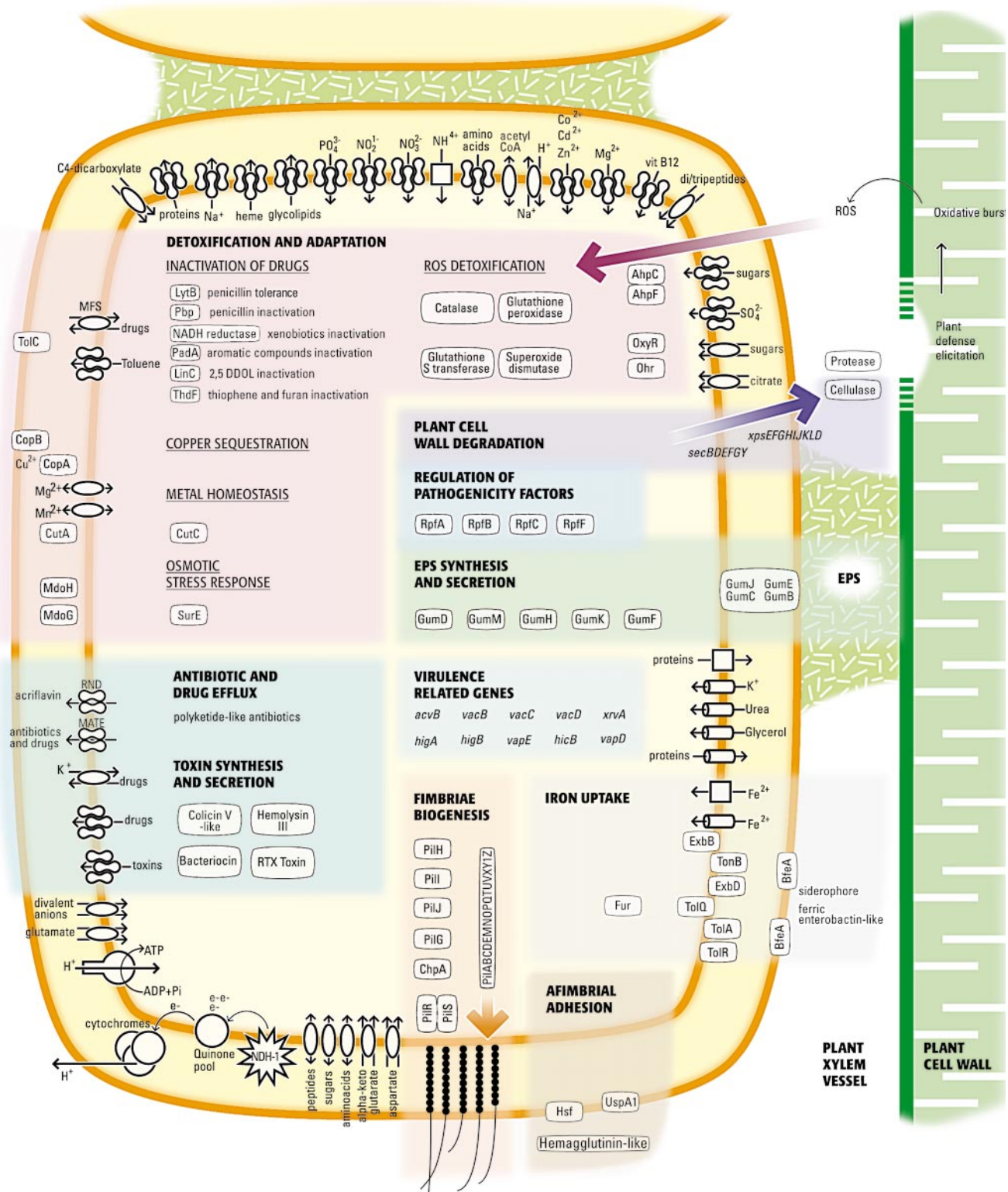


Figure 2 A comprehensive view of the biochemical processes involved in *Xylella fastidiosa* pathogenicity and survival in the host xylem. The principal functional categories are shown in bold, and the bacterial genes and gene products related to that function are arranged within the coloured section containing the bold heading. Transporters are indicated as follows: cylinders, channels; ovals, secondary carriers, including the MFS family; paired dumbbells, secondary carriers for drug extrusion; triple dumbbells, ABC transporters; bulb-like icon, F-type ATP synthase; squares, other transporters. Icons with two arrows

represent symporters and antiporters (H^+ or Na^+ porters, unless noted otherwise). 2,5DDOL, 2,5-dichloro-2,5-cyclohexadiene-1,4-diol; EPS, exopolysaccharides; MATE, multi-antimicrobial extrusion family of transporters multidrug efflux gene (XF2686); MFS, major facilitator superfamily of transporters; Pbp, β -lactamase-like penicillin-binding protein (XF1621); RND, resistance-nodulation-cell division superfamily of transporters; ROS, reactive oxygen species.

include many more enzymes, which we did not find; however, some of the genes listed lie close to ORFs without significant database matches, suggesting that at least one (as yet undiscovered) polyketide pathway may be functional.

Prophages

Bacteriophages can mediate the evolution and transfer of virulence factors and occasional acquisition of new traits by the bacterial host. Because as much as 7% of the *X. fastidiosa* genome sequenced corresponds to double-stranded (ds) DNA phage sequences, mostly from the Lambda group, we suspect that this route may have been of particular importance for this bacterium. It is noteworthy that a very high percentage of phage-related sequences has also been detected in a second vascular-restricted plant pathogen, *Spiroplasma citri*³³. We identified four regions, with a high density of ORFs homologous to phage sequences, that we considered to be prophages, in addition to isolated phage sequences dispersed throughout the genome. Two of these prophages (each ~42 kbp, designated Xfp1 and Xfp2) are similar to each other, lie in opposite orientations in distinct regions and appear to belong to the dsDNA, tailed-phage group. Both appear to contain most of the genes responsible for particle assembly, although we know of no reports of phage particle release from *X. fastidiosa* cultures. In prophage Xfp1, we found two ORFs between tail genes V and W that are similar to ORF118 and *vapA* from the virulence-associated region of the animal pathogen *Dichelobacter nodosus*, which by homology encode a killer and a suppressor protein³⁴. Interestingly, in prophage Xfp2, we found two other ORFs also between tail genes V and W that are similar to hypothetical ORFs of *Ralstonia eutropha* transposon Tn4371 (ref. 35). The other two identified prophages, Xfp3 and Xfp4, are also similar in sequence to each other and to the *H. influenzae* cryptic prophage ϕ flu (ref. 36). They both contain a 14,317-bp exact repeat. Few particle-assembly genes were found in these regions, suggesting that these prophages are defective. An ORF similar to *hicB* from *H. influenzae*, a component of the major pilus gene cluster in some isolates, was found in Xfp4 (ref. 37).

The presence of virulence-associated genes from other organisms within the prophage sequences is strong evidence for a direct role for bacteriophage-mediated horizontal gene transfer in the definition of the bacterial phenotype.

Absence of avirulence genes

Phytopathogenic bacteria generally have a limited host range, often confined to members of a single species or genus. This specificity is defined by the products of the so-called avirulence (*avr*) genes present in the pathogen, which are injected directly into host cells, on infection, through a type III secretory system^{38–40}. BLAST⁴¹ searches with all known *avr* and type III secretory system sequences failed to identify genes encoding proteins with significant similarities in the genome of *X. fastidiosa*. Although the variability of *avr* genes amongst bacteria might account for this apparent lack, the high level of similarity of some components of the type III secretory system argues against this. We suspect that these genes are, in fact, not required because of the insect-mediated transmission and vascular restriction of the bacterium that obviates the necessity of host cell infection. Furthermore, if the differing host ranges of *X. fastidiosa* are molecularly defined, this may be by a quite different mechanism not involving *avr* proteins.

Conclusions

Before the elucidation of its complete genome sequence, very little was known of the molecular mechanisms of *X. fastidiosa* pathogenicity. Indeed, this bacterium was probably the least characterized of all organisms that have been fully sequenced. Our complete genetic analysis has determined not only the basic metabolic and replicative characteristics of the bacterium, but also a number of potential

pathogenicity mechanisms. Some of these have not previously been postulated to occur in phytopathogens, providing new insights into the generality of these processes. Indeed, the availability of this first complete plant pathogen genome sequence will now allow the initiation of the detailed comparison of animal and plant pathogens at the whole-genome level. In addition, the information contained in the sequence should provide the basis for an accelerated and rational experimental dissection of the interactions between *X. fastidiosa* and its hosts that might lead to fresh insights into potential approaches to the control of CVC. □

Methods

The sequencing and analysis in this project were carried out by a network of 34 biology laboratories and one bioinformatics centre. The network is called the Organization for Nucleotide Sequencing and Analysis (ONSA)⁴², and is entirely located in the state of São Paulo, Brazil.

Sequencing and assembly

The sequence was generated using a combination of ordered cosmid and shotgun strategies⁴³. A cosmid library was constructed, providing roughly 15-fold genome coverage, containing 1,056 clones with average insert size of 40 kilobases (kb). High-density colony filters of the library were made, and a physical map of the genome was constructed using a strategy of hybridization without replacement⁴⁴. A total of 113 cosmid clones was selected for sequencing on the basis of the hybridization map and end-sequence analysis. The cosmid sequences were assembled into 15 contigs covering 90% of the genome. Additionally, shotgun libraries with different insert sizes (0.8–2.0 kb and 2.0–4.5 kb) were constructed from nebulized or restricted genomic DNA cloned into plasmids, and sequenced to achieve a 3.74-fold coverage of high-quality sequence (29,140 reads). Most of the sequencing was performed with BigDye terminators on ABI Prism 377 DNA sequencers.

Cosmid and shotgun sequences were assembled into six contigs. We identified sequence gaps by linking information from forward and reverse reads, and closed either by primer walking or insert subcloning. The remaining physical gaps were closed by combinatorial PCR and by lambda clones selected from a λ Dash library by end-sequencing. The collinearity between the genome and the obtained sequence was confirmed by digestion of genomic DNA with *AscI*, *NotI*, *SfiI*, *SmaI* and *SrfI*, followed by comparison of the digestion pattern with the electronic digestion of the generated sequence. In addition, sequences from both ends of most cosmid clones and 236 λ clones were used to confirm the orientation and integrity of the contigs. The sequence was assembled using phred+phrap+consed⁴⁵. All consensus bases have quality with Phred value of at least 20. There are no unexplained high quality discrepancies, each consensus base is confirmed by at least one read from each strand, and the overall error estimate is less than 1 in every 10,000 bases.

ORF prediction and annotation

ORFs were determined using glimmer 2.0 (ref. 46) and the glimmer post-processor RBSfinder (S. L. Salzberg, personal communication). A few ORFs were found by hand guided by BLAST⁴¹ results. Annotation was carried out in a cooperative way, mostly by comparison with sequences in public databases, using BLAST⁴¹ and tRNAscan-SE (ref. 47) and was based on the functional categories for *E. coli*⁴⁸. Only one tmRNA was located (K. Williams, personal communication). To help annotate transport proteins, we built a custom BLAST⁴¹ database using sequences from <http://www-biology.ucsd.edu/~msaier/transport/toc.html> and compared our ORFs with these sequences. Phylogenetic trees for conserved COG⁴² were built using ClustalX⁴⁹ for multiple alignment and Phylip⁵⁰. Paralogous gene families (Table 2) were determined using BLASTX with the E-value cut-off equal to e^{-5} and such that at least 60% of the query sequence and at least 30% of the subject sequence were aligned.

Received 24 March; accepted 24 May 2000.

- Rossetti, V. et al. Présence de bactéries dans le xylème d'orangers atteints de chlorose variée, une nouvelle maladie des agrumes au Brésil. *C. R. Acad. Sci. Paris série III*, **310**, 345–349 (1990).
- Chang, C. J. et al. 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 (1993).
- Roberto, S. R., Coutinho, A., De Lima, J. E. O., Miranda, V. S. & Carlos, E. F. Transmissão de *Xylella fastidiosa* pelas cigarrinhas *Dilobopterus costalimai*, *Acrogonia terminalis* e *Oncometopia facialis* em citros. *Fitopatol. Bras.* **21**, 517–518 (1996).
- Purcell, A. H. & Hopkins, D. L. Fastidious xylem-limited bacterial plant pathogens. *Annu. Rev. Phytopathol.* **34**, 131–151 (1996).
- Li, W. B. et al. A triply cloned strain of *Xylella fastidiosa* multiplies and induces symptoms of citrus variegated chlorosis in sweet orange. *Curr. Microbiol.* **39**, 106–108 (1999).
- Ye, F., Renaudin, J., Bové, J. M. & Laigret, F. Cloning and sequencing of the replication origin (*oriC*) of the *Spiroplasma citri* chromosome and construction of autonomously replicating artificial plasmids. *Curr. Microbiol.* **29**, 23–29 (1994).
- Francino, M. P. & Ochman, H. Strand asymmetries in DNA evolution. *Trends Genet.* **13**, 240–245 (1997).
- Nelson, K. E. et al. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323–329 (1999).

9. White, O. *et al.* Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* **286**, 1571–1577 (1999).
10. Tettelin, H. *et al.* Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science* **287**, 1809–1815 (2000).
11. Katz, M. E., Strugnell, R. A. & Rood, J. I. Molecular characterization of a genomic region associated with virulence in *Dichelobacter nodosus*. *Infect. Immun.* **60**, 4586–4592 (1992).
12. Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: a tool for genome scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000).
13. Preston, G. M., Haubold, B. & Rainey, P. B. Bacterial genomics and adaptation to life on plants: implications for the evolution of pathogenicity and symbiosis. *Curr. Opin. Microbiol.* **1**, 589–597 (1998).
14. Knoop, V., Kloska, S. & Brennicke, A. On the identification of group II introns in nucleotide sequence data. *J. Mol. Biol.* **242**, 389–396 (1994).
15. Ferat, J. L. & Michel, F. Group II self-splicing introns in bacteria. *Nature* **364**, 358–361 (1993).
16. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
17. Bridges, B. A. DNA repair: Polymerases for passing lesions. *Curr. Biol.* **9**, R475–R477 (1999).
18. Brodbeck, B. V., Andersen, P. C. & Mizell, R. F. Effects of total dietary nitrogen and nitrogen form on the development of xylophagous leafhoppers. *Arch. Insect Biochem. Physiol.* **42**, 37–50 (1999).
19. Paulsen, I. T., Sliwinski, M. K. & Saier, M. H. J. Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *J. Mol. Biol.* **277**, 573–592 (1998).
20. Chagas, C. M., Rossetti, V. & Beretta, M. J. G. Electron-microscopy studies of a xylem-limited bacterium in sweet orange affected with citrus variegated chlorosis disease in Brazil. *J. Phytopathol.* **134**, 306–312 (1992).
21. Tang, J. L. *et al.* 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 (1991).
22. Raju, C. B. & Wells, J. M. Diseases caused by fastidious xylem-limited bacteria. *Plant Disease* **70**, 182–186 (1986).
23. Bransky, R. H., Timmer, I. W., French, W. J. & McCoy, R. E. Colonization of the sharpshooter vectors, *Oncometopia igracans* and *Homalodisca coagulata*, by xylem-limited bacteria. *Phytopathology* **73**, 530–535 (1983).
24. Fernandez, L. A. & Berenguer, J. Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiol. Rev.* **24**, 21–44 (2000).
25. Soto, G. E. & Hultgren, S. J. Bacterial adhesins: common themes and variations in architecture and assembly. *J. Bacteriol.* **181**, 1059–1071 (1999).
26. Geme, J. W. Molecular determinants of the interaction between *Haemophilus influenzae* and human cells. *Am. J. Respir. Crit. Care Med.* **154**, S192–S196 (1996).
27. Cope, L. D. *et al.* Characterization of the *Moraxella catarrhalis* *uspA1* and *uspA2* genes and their encoded products. *J. Bacteriol.* **181**, 4026–4034 (1999).
28. Henderson, I. R., Navarro-Garcia, F. & Nataro, J. P. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol.* **6**, 370–378 (1998).
29. Rahme, L. G. *et al.* Common virulence factors for bacterial pathogenicity in plants and animals. *Science* **268**, 1899–1902 (1995).
30. Oresnik, I. J., Twelker, S. & Hynes, M. F. Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins. *Appl. Environ. Microbiol.* **65**, 2833–2840 (1999).
31. Lally, E. T., Hill, R. B., Kieba, I. R. & Korostoff, J. The interaction between RTX toxins and target cells. *Trends Microbiol.* **7**, 356–361 (1999).
32. Havarstein, L. S., Holo, H. & Nes, I. F. The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. *Microbiology* **140**, 2383–2389 (1994).
33. Ye, F. *et al.* A physical and genetic map of the *Spiroplasma citri* genome. *Nucleic Acids Res.* **20**, 1559–1565 (1992).
34. Billington, S. J., Johnston, J. L. & Rood, J. I. Virulence regions and virulence factors of the ovine footrot pathogen, *Dichelobacter nodosus*. *FEMS Microbiol. Lett.* **145**, 147–156 (1996).
35. Merlin, C., Springael, D. & Toussaint, A. Tn4371: A modular structure encoding a phage-like integrase, a Pseudomonas-like catabolic pathway, and RP4/Ti-like transfer functions. *Plasmid* **41**, 40–54 (1999).
36. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl Acad. Sci. USA* **96**, 2192–2197 (1999).
37. Mhlanga-Mutagadura, T., Morlin, G., Smith, A. L., Eisenstark, A. & Golomb, M. Evolution of the major pilus gene cluster of *Haemophilus influenzae*. *J. Bacteriol.* **180**, 4693–4703 (1998).
38. Alfano, J. R. & Collmer, A. The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *J. Bacteriol.* **179**, 5655–5662 (1997).
39. Galan, J. E. & Collmer, A. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284**, 1322–1328 (1999).
40. Young, G. M., Schmiel, D. H. & Miller, V. L. A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc. Natl Acad. Sci. USA* **96**, 6456–6461 (1999).
41. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
42. Simpson, A. J. & Perez, J. F. ONSA, the São Paulo Virtual Genomics Institute. *Nature Biotechnol.* **16**, 795–796 (1998).
43. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
44. Hoheisel, J. D. *et al.* High resolution cosmid and P1 maps spanning the 14 Mb genome of the fission yeast *S. pombe*. *Cell* **73**, 109–120 (1993).
45. Gordon, D., Abajian, C. & Green, P. Consed: a graphical tool for sequence finishing. *Genome Res.* **8**, 195–202 (1998).
46. Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636–4641 (1999).
47. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequences. *Nucleic Acids Res.* **25**, 955–964 (1997).
48. Riley, M. Functions of the gene products of *Escherichia coli*. *Microbiol. Rev.* **57**, 862–952 (1993).
49. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882 (1997).
50. Felsenstein, J. PHYLIP—Phylogeny Inference Package (Version 3. 2). *Cladistics* **5**, 164–166 (1989).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>), or on the author's World-Wide Web site (<http://www.lbi.ic.unicamp.br/xf>) or as paper copy from the London editorial office of Nature.

Acknowledgements

The consortium is indebted to J. F. Perez, Scientific Director of Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), for his strategic vision in creating and nurturing this project as well as C. A. de Pian and Juçara Parra for their administrative coordination. We thank our Steering Committee: S. Oliver, A. Goffeau, J. Sgouros, A. C. M. Paiva and J. L. Azevedo for their critical accompaniment of the work. We also thank R. Fulton and P. Minx for their timely contribution and advice. Project funding was from FAPESP, the RHAE programme of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundecitrus. For the full list of individuals who contributed to the completion of this project see (<http://www.lbi.ic.unicamp.br/xf>).

Correspondence and requests for materials should be addressed to J.C.S. (e-mail: setubal@ic.unicamp.br). The sequence has been deposited in GenBank with accession numbers AE003849 (chromosome), AE003850 (pXF1.3) and AE003851 (pXF51).

Authors: A. J. G. Simpson¹, F. C. Reinach², P. Arruda³, F. A. Abreu⁴, M. Acencio⁵, R. Alvarenga², L. M. C. Alves⁶, J. E. Araya⁷, G. S. Baia², C. S. Baptista⁸, M. H. Barros⁸, E. D. Bonaccorsi², S. Bordin⁹, J. M. Bové¹⁰, M. R. S. Briones⁷, M. R. P. Bueno¹¹, A. A. Camargo¹, L. E. A. Camargo¹², D. M. Carraro¹², H. Carrer¹², N. B. Colauto¹³, C. Colombo¹⁴, F. F. Costa⁹, M. C. R. Costa¹⁵, C. M. Costa-Neto¹⁶, L. L. Coutinho¹², M. Cristofani¹⁷, E. Dias-Neto¹, C. Docena², H. El-Dorry², A. P. Facincani⁶, A. J. S. Ferreira², V. C. A. Ferreira¹⁸, J. A. Ferro⁶, J. S. Fraga⁴, S. C. França¹⁹, M. C. Franco²⁰, M. Frohme²¹, L. R. Furlan²², M. Garnier¹⁰, G. H. Goldman²³, M. H. S. Goldman²⁴, S. L. Gomes², A. Gruber⁴, P. L. Ho²⁵, J. D. Hoheisel²¹, M. L. Junqueira²⁶, E. L. Kemper³, J. P. Kitajima²⁷, J. E. Krieger²⁶, E. E. Kuramae²⁸, F. Laigret¹⁰, M. R. Lambais¹², L. G. C. Leite²⁵, E. G. M. Lemos⁶, M. V. F. Lemos²⁹, S. A. Lopes¹⁹, C. R. Lopes¹³, J. A. Machado^{30†}, M. A. Machado¹⁷, A. M. B. N. Madeira⁴, H. M. F. Madeira^{12†}, C. L. Marino¹³, M. V. Marques⁸, E. A. L. Martins²⁵, E. M. F. Martins¹⁸, A. Y. Matsukuma², C. F. M. Menck⁸, E. C. Miracca⁵, C. Y. Miyaki¹¹, C. B. Monteiro-Vitorello¹², D. H. Moon²⁰, M. A. Nagai⁵, A. L. T. O. Nascimento²⁵, L. E. S. Netto¹¹, A. Nhani Jr⁶, F. G. Nobrega^{8†}, L. R. Nunes³¹, M. A. Oliveira³², M. C. de Oliveira³³, R. C. de Oliveira³¹, D. A. Palmieri¹³, A. Paris¹³, B. R. Peixoto², G. A. G. Pereira³², H. A. Pereira Jr⁶, J. B. Pesquero¹⁶, R. B. Quaggio², P. G. Roberto¹⁹, V. Rodrigues³⁴, A. J. de M. Rosa³⁴, V. E. de Rosa Jr²⁸, R. G. de Sá³⁴, R. V. Santelli², H. E. Sawasaki¹⁴, A. C. R. da Silva², A. M. da Silva², F. R. da Silva^{3,27}, W. A. Silva Jr¹⁵, J. F. da Silveira⁷, M. L. Z. Silvestri², W. J. Siqueira¹⁴, A. A. de Souza¹⁷, A. P. de Souza³, M. F. Terenzi²³, D. Truffi¹², S. M. Tsai²⁰, M. H. Tshako¹⁸, H. Vallada³⁵, M. A. Van Sluys³³, S. Verjovski-Almeida², A. L. Vettore³, M. A. Zago¹⁵, M. Zatz¹¹, J. Meidanis²⁷ & J. C. Setubal²⁷.

Addresses: 1, Instituto Ludwig de Pesquisa sobre o Câncer, Rua Prof. Antonio Prudente, 109 – 4º andar, 01509-010, São Paulo - SP, Brazil; 2, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes, 748, 05508-900, São Paulo - SP, Brazil; 3, Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, Caixa Postal 6010, 13083-970, Campinas - SP, Brazil; 4, Laboratório de Biologia Molecular, Departamento de Patologia, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, Av. Prof. Dr. Orlando Marques de Paiva, 87, 05508-000, São Paulo - SP, Brazil; 5, Disciplina de Oncologia, Departamento de Radiologia, Faculdade de Medicina, Universidade de São Paulo, Av. Dr. Arnaldo, 455, 01296-903, São Paulo - SP, Brazil; 6, Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, Universidade Estadual Paulista, Via

de Acesso Prof. Paulo D. Castellane s/n, Km 5, 14884-900, Jaboticabal – SP, Brazil; 7, Departamento de Microbiologia, Imunologia e Parasitologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, Rua Botucatu, 862, 04023-062, São Paulo – SP, Brazil; 8, Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, Av. Prof. Lineu Prestes, 1374, 05508-900, São Paulo – SP, Brazil; 9, Hemocentro, Faculdade de Ciências Médicas, Universidade Estadual de Campinas, 13083-97, Campinas – SP, Brazil; 10, Institut National de la Recherche Agronomique et Université Victor Ségalen Bordeaux 2, 71 Avenue Edouard Bourleaux, Laboratoire de Biologie Cellulaire et Moléculaire, 33883 Vilenave d'Ornon Cedex, France; 11, Departamento de Biologia, Instituto de Biociências, Universidade de São Paulo, Rua do Matão, 277, 05508-900, São Paulo – SP, Brazil; 12, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias, 11, 13418-900, Piracicaba – SP, Brazil; 13, Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista, Distrito de Rubião Junior, 18618-000, Botucatu – SP, Brazil; 14, Centro de Genética, Biologia Molecular e Fitoquímica, Instituto Agrônomo de Campinas, Av. Barão de Itapura, 1481, Caixa Postal 28, 13001-970, Campinas – SP, Brazil; 15, Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14049-900, Ribeirão Preto – SP, Brazil; 16, Departamento de Biofísica, Escola Paulista de Medicina, Universidade Federal de São Paulo, Rua Botucatu, 862, 04023-062, São Paulo – SP, Brazil; 17, Centro de Citricultura Sylvio Moreira, Instituto Agrônomo de Campinas, Caixa Postal 04, 13490-970, Cordeirópolis – SP, Brazil; 18, Laboratório de Bioquímica Fitopatológica e Laboratório de Imunologia, Instituto Biológico, Av. Cons. Rodrigues Alves, 1252, 04014-002, São Paulo – SP, Brazil; 19, Departamento de Biotecnologia de Plantas Medicinais, Universidade de Ribeirão Preto, Av. Costabile Romano, 2201, 14096-380, Ribeirão Preto – SP, Brazil; 20, Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Av. Centenário, 303, Caixa Postal 96, 13400-970, Piracicaba – SP, Brazil; 21, Funktionelle Genomanalyse, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120, Heidelberg, Germany; 22, Departamento de Melhoramento e Nutrição Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista, Fazenda Lageado, Caixa Postal 560, 18600-000, Botucatu – SP, Brazil; 23, Departamento de Ciências Farmacêuticas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Av. do Café s/n, 14040-903, Ribeirão Preto – SP, Brazil; 24, Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14040-901, Ribeirão Preto – SP, Brazil; 25, Centro de Biotecnologia, Instituto Butantan, Av. Vital Brasil, 1500, 05503-900, São Paulo – SP, Brazil; 26, Laboratório de Genética e Cardiologia Molecular/LIM 13, Instituto do Coração (InCor), Faculdade de Medicina, Universidade de São Paulo, Av. Dr. Enéas de Carvalho Aguiar, 44, 05403-000, São Paulo – SP, Brazil; 27, Instituto de Computação, Universidade Estadual de Campinas, Caixa Postal 6176, 13083-970, Campinas – SP, Brazil; 28, Departamento de Produção Vegetal, Faculdade de Ciências Agrônomicas, Universidade Estadual Paulista, Fazenda Lageado, Caixa Postal 237, 18603-970, Botucatu – SP, Brazil; 29, Departamento de Biologia Aplicada à Agropecuária, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, Universidade Estadual Paulista, Via de Acesso Prof. Paulo Donato Castellane, 14884-900, Jaboticabal – SP, Brazil; 30, Hospital do Câncer – A. C. Camargo, R. Antonio Prudente, 211, 01509-010, São Paulo – SP, Brazil; 31, Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, Av. Dr. Cândido Xavier de Almeida Souza, 200, 08780-911, Mogi das Cruzes – SP, Brazil; 32, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, Caixa Postal 6010, 13083-970, Campinas – SP, Brazil; 33, Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Rua do Matão, 277, 05508-900, São Paulo – SP, Brazil; 34, Departamento de Parasitologia, Microbiologia e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14049-900, Ribeirão Preto – SP, Brazil; 35, Departamento de Psiquiatria, Instituto de Psiquiatria, Faculdade de Medicina, Universidade de São Paulo, Rua Dr. Ovídeo Pires de Campos s/n, Sala 4051 (3º andar), 05403-010, São Paulo – SP, Brazil.

† Present addresses: Novartis Seeds LTDA, Av. Prof. Vicente Rao, 90, 04706-900, São Paulo – SP, Brazil (J. A. Machado); Centro de Ciências Agrárias e Ambientais, Pontifícia Universidade Católica do Paraná, BR-376, Km 14, Caixa Postal 129, 83010-500, São José dos Pinhais – PR, Brazil (H. M. F. Madeira); Instituto de Pesquisa e Desenvolvimento, Universidade do Vale do Paraíba, Av. Shishimi Hifumi, 2911, 12244-000, São José dos Campos – SP, Brazil (F. G. Nobrega).



[illegible]