

**JONAS WEISSMANN GAIARSA**

**HISTÓRIA EVOLUTIVA DE CARBO-HIDROLASES LIGNO-CELULÓSICAS DA  
FAMÍLIA XANTHOMONADACEAE**

Tese apresentada ao Programa de Pós-Graduação Interunidades em Biotecnologia USP/Instituto Butantan/IPT, para obtenção do Título de Doutor em Biotecnologia.

Área de concentração: Biotecnologia

Orientadora: Profa. Dra. Marie-Anne Van Sluys

Versão original

São Paulo  
2013

## RESUMO

Gaiarsa JW. História evolutiva de carbo-hidrolases ligno-celulósicas da família *Xanthomonadaceae*. [tese (Doutorado em Biotecnologia)]. São Paulo: Instituto de Ciências Biomédicas, Universidade de São Paulo; 2013.

O presente trabalho visa compreender o processo de degradação da parede celular vegetal de hospedeiros de fitopatógenos da família *Xanthomonadaceae*. Criamos e aperfeiçoamos uma técnica de enumeração dos genes relacionados ao metabolismo de polissacarídeos, com enfoque na distinção entre aqueles que agem sobre os componentes da parede celular vegetal e sobre outros polissacarídeos. A história evolutiva desse conjunto de enzimas foi delineada através de inferências sobre as relações de homologia entre os genes enumerados, sua presença ou ausência nos diversos genomas abordados e comparação das taxas de mutação entre grupos de homólogos. Além disso, procuramos também, com essa etapa de bioinformática e a etapa seguinte, incrementar a anotação desses genes, muitos descritos como hipotéticos ou com vaga definição de sua função. Na segunda parte do desenvolvimento do projeto foram feitos experimentos de expressão heteróloga e verificação da atividade enzimática para validação da anotação de alguns dos genes identificados. Foram exploradas diferentes técnicas *in silico* para procurar diferenças de potencial de maceração da parede celular vegetal. Quatro técnicas de enumeração de genes codificantes para enzimas de degradação da parede celular vegetal (PCWDE) foram construídas em sucessão, para melhor distinguir a atuação sobre polissacarídeos da parede celular vegetal e outros polissacarídeos. A primeira técnica procurou termos relativos à PCWDE sobre a anotação dos genomas; a segunda comparou esses genes com mais genomas; a terceira baseou-se na escolha de modelos de Markov (HMM) de conservação de aminoácidos do projeto Pfam para a seleção de genes e a quarta estratégia no uso do banco dbCAN de HMMs das sequências do projeto CAZy (*Carbohydrate Active enZymes*) em conjunto com um levantamento das atividades enzimática (ECs) relevantes, de forma a correlacionar polissacarídeos, ECs e domínios proteicos. Foi possível, então, a descoberta de novas putativas PCWDE, além da recuperação das já descritas. Os resultados aqui apresentados revelam que o repertório genômico PCWDE varia entre os genomas de *Xanthomonadaceae*. Procuramos aperfeiçoar a técnica de inferência automática de homologia de sequências de aminoácidos baseando-se em alinhamentos múltiplos globais. Essa inferência foi usada para traçar-se um perfil de ação do aparato de PCWDE de cada genoma sobre os diferentes polissacarídeos. Árvores filogenéticas de homólogos foram comparadas para suas as taxas de mutação no intuito de verificar movimentos adaptativos semelhantes entre cada grupo de homólogos. Além disso, um levantamento dos genes de PCWDE pertencentes a regiões de transferência horizontal gênica (HGT) foi feito caracterizando suas contribuições aos genomas. As análises indicam que alguns genomas, em especial algumas espécies do gênero *Xanthomonas*, são detentores da maior parte da diversidade de PCWDE, tendo outros genomas apenas um subconjunto. No intuito de validar o trabalho de bioinformática fizemos um processo de seleção de genes de PCWDE candidatos para uma caracterização molecular. Os genes envolvidos no início da cascata de degradação, com sinal para secreção e que ainda não tivessem nenhum relato específico na literatura foram selecionados.

Dois dos genes submetidos a esse processo chegaram ao final com retificação de suas anotações pela atividade enzimática verificada.

Palavras-chave: *Xanthomonas*. Parede celular vegetal. Fitopatógenos. Polissacarídeos. Evolução molecular.

## ABSTRACT

Gaiarsa JW. Evolutionary history of lignocellulosic carbo-hydrolases of the *Xanthomonadaceae* family. [Ph. D. thesis (Biotechnology)]. São Paulo: Instituto de Ciências Biomédicas, Universidade de São Paulo; 2013.

This study aims to understand the process of degradation of host plant cell walls by plant pathogens of the *Xanthomonadaceae* family. We created and perfected a technique for enumeration of genes related to the metabolism of polysaccharides, focusing on the distinction between those who act on components of plant cell wall and on other polysaccharides. The evolutionary history of this group of enzymes has been outlined through inferences about the relations of homology between the genes listed, their presence or absence in different genomes and comparison of mutation rates between groups of homologues. Moreover, we also intended with this bioinformatics step and the next step, to enhance the annotation of these genes, many described as hypothetical or vague in the determination of its function. In the second part of the project development heterologous expression and enzymatic activity assays were made to validate the annotation of some of the genes identified. We explored different *in silico* techniques to assess differences in plant cell wall maceration potential. Four techniques for enumeration of genes coding for plant cell wall degrading enzymes (PCWDE) were built in succession to better distinguish between activities on plant cell wall polysaccharides and other polysaccharides. The first technique searched for keywords relating to PCWDE in genome annotations, the second comparing these genes with a bigger set of genomes, the third was based on the selection of amino acid conservation Markov models (HMM) from the Pfam project used for gene selection and the fourth in the use of dbCAN HMMs constructed from sequences of the CAZy project (Carbohydrate Active enZymes) together with a survey of relevant enzymatic activities (ECs), in order to correlate polysaccharides, ECs and protein domains. It was then possible to discover new putative PCWDE, besides the recovery of already described ones. The results presented here show that the genomic PCWDE repertoire varies among *Xanthomonadaceae*. We seek also to improve the technique of automatic inference of homologous amino acid sequences through the use of multiple global alignments. This inference was used to outline the potential of action of the apparatus of PCWDE of each genome on the different polysaccharides. Phylogenetic trees of homologous groups were compared through their mutation rates in order to verify adaptive movements similar between each group. In addition, a survey of PCWDE genes belonging to horizontal gene transfer (HGT) regions was made assessing their contributions to the genomes. The analysis indicates that some genomes, especially certain species of the genus *Xanthomonas*, harbor most of the PCWDE diversity, with other genomes having only a subset. In order to validate the bioinformatics work done, a process of selection of candidate PCWDE genes was made for molecular characterization. The genes involved in the early steps of the degradation cascade, with secretion signal and not having yet had any specific reports in the literature were selected. Two of the genes undergoing this process have had their annotation corrected based on the enzymatic activity observed.

Keywords: *Xanthomonas*. Plant cell wall. Phytopathogens. Polysaccharides. Molecular evolution.

## 1 INTRODUÇÃO

### 1.1 *Xanthomonadaceae*

A ordem *Xanthomonadales*, à qual pertence sua maior família *Xanthomonadaceae*, é um grupo que contém dois gêneros de fitopatógenos com diversos genomas sequenciados desde 2000 (Simpson et al., 2000). Dentre os gêneros, *Xanthomonas* se destaca por possuir uma ampla gama de hospedeiros, sendo o gênero composto por 20 espécies classificadas em 80 patovares (Rademaker et al., 2005) (Tabela 1). Perdas mundiais de safra causada por patógenos somam tipicamente 10% da produção (Strange, Scott, 2005). Três das espécies de destaque dentre as que causam os maiores prejuízos são *Xanthomonas campestris*, *Xanthomonas axonopodis* e *Xanthomonas oryzae*. No entanto, já há quatro décadas vem se empregando a espécie *Xanthomonas campestris* de forma industrial na produção de goma xantana, um importante biopolímero empregado nas indústrias de alimentos, cosméticos e petróleo. O gênero *Stenotrophomonas*, embora tenha dois genomas sequenciados de cepas causadoras de infecções oportunistas em pacientes imunodeprimidos, também pode ser encontrada associada a plantas apresentando atividade antifúngica (*Stenotrophomonas rhizophila*) (Wolf et al., 2002) e na forma de espécie endofítica fixadora de nitrogênio (*Stenotrophomonas pavanii*) (Ramos et al., 2011).

### 1.2 Mecanismos de instalação do patógeno

Existem duas fases sequenciais distintas no início da patogênese no que tange a instalação da infecção. A primeira é à entrada do agente patogênico em um órgão da planta, vencendo quaisquer barreiras físicas que se interponham, e a segunda, constituída em uma de duas possibilidades, a detecção da invasão pelos sistemas de resistência da planta e a resposta de hipersensibilidade ou a de fato colonização, fruto da não detecção do patógeno, nesse caso considerando-se o hospedeiro susceptível a infecção. Existe um período de incubação antes da manifestação dos sintomas patogênicos, mais longo em algumas variedades do que em outras, tempo onde acontece a colonização. Esse tempo pode ser necessário, dependendo da variedade, para aumento do número de células no biofilme, ativação do metabolismo

derivado da sinalização de *quorum sensing* e atuação dos mecanismos de evasão da detecção como os atuadores secretados pelo sistema de secreção do tipo 3, que modificam o metabolismo da célula vegetal impedindo uma resposta de hipersensibilidade.

**Tabela 1** - Espécies e patovares do gênero *Xanthomonas* (Rademaker et al., 2005).

Espécie	Patovar	Espécie	Patovar	Espécie	Patovar
<i>fragariae</i>		<i>vasicola</i>	<i>vasculorum</i>	<i>axonopodis</i>	<i>Sesbaniae</i>
<i>populi</i>		<i>axonopodis</i>	<i>begoniae</i>	<i>axonopodis</i>	<i>Vignaeradiatae</i>
<i>arboricola</i>		<i>axonopodis</i>	<i>alfalfae</i>	<i>axonopodis</i>	<i>Vignicola</i>
<i>cassavae</i>		<i>axonopodis</i>	<i>cassavae</i>	Sublinhados táxons com	
<i>vodiaei</i>		<i>axonopodis</i>	<i>cassiae</i>	genomas sequenciados.	
<i>bromi</i>		<i>axonopodis</i>	<i>citrumelo</i>		
<i>cucurbitae</i>		<i>axonopodis</i>	<i>coracanae</i>		
<i>psi</i>		<i>axonopodis</i>	<i>cyamopsidis</i>		
<i>melonis</i>		<i>axonopodis</i>	<i>desmodii</i>		
<i>vesicatoria</i>		<i>axonopodis</i>	<i>desmodiiiganetici</i>		
<i>hyacinthi</i>		<i>axonopodis</i>	<i>desmodiirotundifolii</i>		
<i>theicola</i>		<i>axonopodis</i>	<i>erythrinae</i>		
<i>sacchari</i>		<i>axonopodis</i>	<i>lespedezae</i>		
<i>albilineans</i>		<i>axonopodis</i>	<i>patelii</i>		
<i>arboricola</i>	<i>corylina</i>	<i>axonopodis</i>	<i>phyllanthi</i>		
<i>arboricola</i>	<i>juglandis</i>	<i>axonopodis</i>	<i>poinsettiicola</i>		
<i>arboricola</i>	<i>poinsettiicola</i>	<i>axonopodis</i>	<i>ricini</i>		
<i>arboricola</i>	<i>populi</i>	<i>axonopodis</i>	<i>tamarindi</i>		
<i>arboricola</i>	<i>pruni</i>	<i>axonopodis</i>	<i>vesicatoria</i>		
<i>campestris</i>	<i>aberrans</i>	<i>axonopodis</i>	<i>axonopodis</i>		
<i>campestris</i>	<i>armoraciae</i>	<i>axonopodis</i>	<i>vasculorum</i>		
<i>campestris</i>	<i>barbareae</i>	<i>axonopodis</i>	<i>manihotis</i>		
<i>campestris</i>	<i>campestris</i>	<i>axonopodis</i>	<i>phaseoli</i>		
<i>campestris</i>	<i>incanae</i>	<i>axonopodis</i>	<i>nova citri E</i>		
<i>campestris</i>	<i>raphani</i>	<i>axonopodis</i>	<i>dieffenbachiae</i>		
<i>campestris</i>	<i>musacearum</i>	<i>axonopodis</i>	<i>bauhiniae</i>		
<i>hortorum</i>	<i>hederae</i>	<i>axonopodis</i>	<i>cajani</i>		
<i>hortorum</i>	<i>pelargonii</i>	<i>axonopodis</i>	<i>citri</i>		
<i>hortorum</i>	<i>vitians</i>	<i>axonopodis</i>	<i>clitoriae</i>		
<i>oryzae</i>	<i>oryzae</i>	<i>axonopodis</i>	<i>desmodiilaxiflori</i>		
<i>oryzae</i>	<i>oryzicola</i>	<i>axonopodis</i>	<i>glycines</i>		
<i>translucens</i>	<i>arrhenatheri</i>	<i>axonopodis</i>	<i>malvacearum</i>		
<i>translucens</i>	<i>cerealis</i>	<i>axonopodis</i>	<i>vitians</i>		
<i>translucens</i>	<i>graminis</i>	<i>axonopodis</i>	<i>nova vitians</i>		
<i>translucens</i>	<i>bromi</i>	<i>axonopodis</i>	<i>nova phaseoli</i>		
<i>translucens</i>	<i>hordei</i>	<i>axonopodis</i>	<i>cajani</i>		
<i>translucens</i>	<i>phlei</i>	<i>axonopodis</i>	<i>aurantifolii</i>		
<i>translucens</i>	<i>phleipratensis</i>	<i>axonopodis</i>	<i>aurantifolii</i>		
<i>translucens</i>	<i>poae</i>	<i>axonopodis</i>	<i>aurantifolii</i>		
<i>translucens</i>	<i>secalis</i>	<i>axonopodis</i>	<i>dieffenbachiae</i>		
<i>translucens</i>	<i>translucens</i>	<i>axonopodis</i>	<i>phaseoli</i>	<i>var.</i>	
<i>translucens</i>	<i>undulosa</i>	<i>axonopodis</i>	<i>fuscans</i>		
<i>vasicola</i>	<i>holcicola</i>	<i>axonopodis</i>	<i>rhyngosiae</i>		

Na fase de entrada, a atuação das PCWDE (Tabela 2) pode estar presente no rompimento da barreira física imposta pela parede celular, permitindo um contato mais direto entre o patógeno e as células do hospedeiro ou o acesso ao sistema vascular da planta hospedeira. Já na fase de colonização da planta essas enzimas podem atuar na hidrólise da parede proporcionando carboidratos nutrientes para o patógeno. Os dois gêneros fitopatogênicos tratados nesse trabalho, *Xanthomonas* e *Xylella*, apesar de compartilharem boa parte de seus genes, possuem diferentes estratégias de dispersão, invasão e colonização de seus hospedeiros. *Xanthomonas* com genomas sequenciados possuem três estratégias principais de dispersão e invasão. Trazidos através de agentes físicos ambientais, como chuva e vento, e entrando por estômatos foliares, se estabelecendo nas câmaras estomatares ou no mesófilo e depois atingindo o sistema vascular; ou entrando por hidatódios, entrando diretamente no sistema vascular e lá se estabelecendo e depois se dispersando pelos órgãos da planta. Também pode invadir pela injúria mecânica, em geral fruto da manipulação no cultivo, estabelecendo-se possivelmente diretamente junto a células e também sendo disperso pelo sistema vascular. *Xylella* possui um mecanismo que envolve a atuação de um inseto vetor sugador, que ao se alimentar da planta injeta a bactéria, que carrega no seu trato digestivo, diretamente no sistema vascular da planta.

Dentre as espécies do gênero *Xanthomonas* existe também uma variação de estratégias, sendo a maior distinção encontrada entre patovares que colonizam sistemicamente o hospedeiro e patovares que se restringem ao sistema vascular.

Tabela 2 - Fenótipos de maceração.

Nome	Hospedeiros	Tecidos	Maceração
<i>Xanthomonas campestris patovar vesicatoria</i>	Tomate e <i>Capsicum</i>	Mesófilo e parênquima (EPPO XANTVE)	Manchas superficiais granulosas ou crostas, com margens encharcadas de água, oval ou irregular, 2-10 mm de diâmetro.
<i>Xanthomonas axonopodis patovar citri</i>	Citros	Mesófilo e parênquima (EPPO XANTCI)	Lesões nas folhas, caules e frutos com margens levantadas, marrom, banhadas de água, geralmente com um halo amarelo ou anel em torno da lesão.
<i>Xanthomonas campestris patovar armoraciae</i>	<i>Brassica oleracea</i>	Mesófilo e parênquima (Kamoun; Kado, 1990).	Numerosas pequenas manchas pretas, circulares, encharcadas de água, eventualmente, com halos amarelos na parte inferior das folhas, dando uma cor amarelada a folhagem e aparência irregular. Manchas foram também visíveis na superfície superior das folhas e em pecíolos
<i>Xanthomonas campestris patovar campestris</i>	Crucíferas e Brassicáceas	Vascular (Kamoun; Kado, 1990).	Interno (enegrecimento vascular)
<i>Xanthomonas campestris patovar musacearum</i>	Ensete, banana e milho.	Mesófilo e vascular	Folhas infectadas tendem a quebrar ao longo do limbo foliar ( <a href="http://www.cabi.org/isc/?compid=5&amp;dsid=56917&amp;loadmodule=datasheet&amp;page=481&amp;site=144">http://www.cabi.org/isc/?compid=5&amp;dsid=56917&amp;loadmodule=datasheet&amp;page=481&amp;site=144</a> )( <a href="http://www.promusa.org/tiki-index.php?page=Xanthomonas+wilt">http://www.promusa.org/tiki-index.php?page=Xanthomonas+wilt</a> )
<i>Xanthomonas oryzae patovar oryzae</i>	Arroz, gramíneas e sedges.	Mesófilo, parênquima e vascular (EPPO XANTOR).	Folhas de plantas jovens, como verde-claro a verde-cinza, estrias encharcadas de água perto da ponta da folha e margens. Essas lesões se aglutinam e se tornar branco-amarelada, com bordas onduladas.
<i>Xanthomonas oryzae patovar oryzicola</i>	Todas as espécies selvagens do gênero <i>oryza</i>	Coloniza o parênquima e se limita a regiões vasculares da planta (EPPO XANTOR)	A doença aparece inicialmente como leves estrias intervenais encharcadas de água. As estrias são de início de cor verde escuro, mas depois se tornam translúcidas. As estrias ampliam e aglutinam-se, eventualmente tornando-se castanho claro.
<i>Xylella fastidiosa</i>	Citros, parreira, espiiradeira, amêndoa.	Vascular (EPPO XYLEFA)	Interno

Hospedeiros, tecidos afetados e características da maceração de algumas espécies e patovares com genomas sequenciados. EPPO se refere aos códigos de patógenos da Organização Europeia e Mediterrânea para a proteção de plantas (<http://www.eppo.int/QUARANTINE/quarantine.htm>) e suas listas de descrições.



### 1.3 Enzimas de degradação da parede celular vegetal (PCWDE)

Genes codificadores para enzimas com atividade sobre carboidratos ou mais especificamente, enzimas de degradação de polissacarídeos da parede celular de plantas são conhecidos por ser uma característica particular de genomas de microrganismos associados a plantas incluindo a família *Xanthomonadaceae*. Em parasitas, PCWDE têm um papel fundamental no processo de infecção do hospedeiro, de modo que já foram implicados como fatores de virulência. Essas enzimas também desempenham um papel muito importante na indústria nascente de etanol lignocelulósico também conhecido como o bioetanol de segunda geração. Dentro da família *Xanthomonadaceae*, apenas algumas das cepas de *Xanthomonas* foram analisadas quanto à expressão, atividade enzimática de alguns dos genes PCWDE e sobre o papel que os sistemas de secreção têm na atuação das enzimas (Hsiao et al., 2007; Hu et al., 2007; Rajeshwari et al., 2005; Ray et al., 2000; Schröter et al., 2001; Sun et al., 2005; Xiao et al., 2008).

Uma das restrições da atividade de uma enzima sobre múltiplos polissacarídeos é o tipo de ligação, sendo que enzimas que agem sobre ligações betas não conseguem agir sobre ligações alfa e vice-versa. Portanto PCWDE em geral não agem sobre glicose alfa ligada, preponderante em amilose, amilopectina, glicogênio e dextrana, mas não em PCWPS.

### 1.4 Polissacarídeos da parede celular vegetal (PCWPS)

A parede celular vegetal (Figura 1) é composta por três ambientes: a membrana celular, a parede e o apoplasto. São enumerados três grandes grupos de polissacarídeos que compõe a parede celular vegetal, celulose, hemicelulose e pectina, sendo celulose um polissacarídeo de composição invariável de glicopirranose  $\beta$ -(1 $\rightarrow$ 4) ligada e as outras duas classes compostas por diversas variações de estrutura (Tabela 3). Dentre os polissacarídeos hemicelulósicos a literatura enumera 4 tipos com ainda outras variações de composição: mananos (manano linear, glucomanano, galactoglucomanano, galactomanano), xilanos (xilano, glucuronoxilano, glucuronoarabinoxilano), xiloglucano (xiloglucano, fucogalactoxiloglucano) e glucanos  $\beta$ -(1 $\rightarrow$ 3,1 $\rightarrow$ 4) (liquenano). Já os polissacarídeos

pécticos podem ser classificados em até 6 tipos: homogalacturonanos, ramnogalacturonano-I, galacturonanos substituídos (ramnogalacturonano-II, xilogalacturonano, apiogalacturonano), arabinano, arabinogalactano-I, arabinogalactano-II.

Na distribuição desses polissacarídeos entre as espécies vegetais celulose e mananos ocorrem ambos em dicotiledôneas e monocotiledôneas. Xilanos ocorrem principalmente na parede de monocotiledôneas, mas em pequenas quantidades em dicotiledôneas. Xiloglucanos, ao contrário, ocorrem principalmente em dicotiledôneas, mas em pequenas quantidades na parede primária de monocotiledôneas. Já glucanos  $\beta(1\rightarrow3,1\rightarrow4)$  ocorrem única e exclusivamente na parede primária de monocotiledôneas. Pectinas em geral ocorrem principalmente em dicotiledôneas, mas em pequena quantidade na parede celular primária de monocotiledôneas.

**Tabela 3** – Monômeros e ligações de hemiceluloses e pectinas.

			Hemicelulose	Pectina
<b>Ligação principal</b>	da	cadeia	$\beta(1\rightarrow4)$	$\alpha(1-4)$
<b>Açúcares principal</b>	da	cadeia	D-Glcp, D-Xylp, D-Galp, D-Manp	D-GalpA, L-Rhap, D-GlcpA.
<b>Ligações secundárias</b>	nas	cadeias	$\alpha(1-2)$ $\beta(1\rightarrow2)$ $\beta(1\rightarrow6)$	$\alpha(1-2)$ $\alpha(1-3)$ $\alpha(1-5)$ $\alpha(3-5)$ $\alpha(2-3)$ $\beta(1\rightarrow4)$ $\beta(1\rightarrow2)$ $\beta(1\rightarrow3)$ $\beta(1\rightarrow5)$ $\beta(2\rightarrow3)$ [ $\beta(1\rightarrow3)$ -D-Apif]
<b>Outros açúcares nas cadeias secundárias</b>			L-Fucp, L-Araf.	L-Araf, L-Arap, L-Fucp, D-Dha, D-Kdo, L-AcefA, D-Xylp.
<b>Modificações de açúcares de cadeia lateral</b>			O-Acetil	4-O-Metoxi

. As composições de hemicelulose e pectina podem variar de acordo com tecido, estágio de desenvolvimento, contexto ambiental e também com a espécie.



### 1.5 PCWDE de *Xanthomonadaceae*.

Dois dos artigos se voltaram para o sistema de secreção do tipo 2 a influência que ele pode ter no processo de infecção. Ray et al. (2000) e Sun et al. (2005) demonstraram a atividade de uma xilanase e uma celulase em isolados de *Xanthomonas oryzae* pv. *oryzae* (Xoo) de Beijing e das Filipinas (cepa sequenciada PXO99), respectivamente. Principalmente, que mutantes para genes do sistema de secreção de proteínas de *Xanthomonas* (xps; sistema tipo dois) apresentavam virulência reduzida e ausência de bandas de 42 kDa e 50 kDa (xilanase e celulase, respectivamente) num gel de proteína da porção extracelular. Não houve, no entanto, alteração na resposta de hipersensibilidade em espécies não suscetíveis e ausência da banda de xilanase para o selvagem da cepa 8004 de Xcc.

Com foco no melhoramento da produção de goma xantana o grupo de Schröter et al. (2001) produziu um mutante da cepa FC2 (NRLLB-1459 derivada) de *Xanthomonas campestris* pv. *campestris* (Xcc) para os genes engXC A e B. O duplo mutante possuía redução de 80% da atividade endoglucanásica sobre celulose.

No trabalho de Rajeshwari et al. (2005) foi exposto o efeito de diminuição de virulência da cepa BXO43 na mutagênese da xilanase xynB e da lipase/esterase lipA, conjuntamente e individualmente. Foi também verificada a atividade de ambas sobre xilano e tween80/tributirina, respectivamente.

Hsiao et al. (2007) exploraram o papel regulatório do fator sinalizador difusível (DSF) sintetizado pelo gene rpfF e da proteína fator de transcrição análoga à proteína receptora de cAMP (clp). Mostraram a existência de sítio promotor clp à montante de uma série de genes codificantes para pectatolases e poligalacturonases de diversas *Xanthomonas*. Além disso, fizeram uma caracterização enzimática da poligalacturonase pehA da cepa Xc17 de Xcc e averiguaram a indução dessa atividade na presença em meio de cultura de ácido poligalacturônico e a inibição na presença de catabólitos da reação (glicose), na limitação de oxigênio e na falta de nitrogênio.

Já Hu et al. (2007) fizeram ensaios sobre Xoo e expressão heteróloga em *E. coli* da celulase eglXoB da cepa PXO99. O mutante para esse gene tinha virulência reduzida, a expressão do gene foi comprovada *in planta*, porém a verificação de atividade em placa de cultura teve retorno negativo, assim como a atividade só foi verificada nas porções intracelulares no heterólogo em *E. coli*.

No trabalho de Xiao et al. (2008) fez-se um trabalho biotecnológico de mineração dos genomas de *Xcc* e *Streptomyces coelicolor* para pectinases. Verificou-se a atividade em expressão heteróloga em *E. coli* de três pectatolases e duas poligalacturonases de *Xcc*. Uma possível pectatolase não apresentou atividade, fato atribuído a uma especificidade da enzima que não estava apta a degradar a pectina de citros utilizada. Dentre os genes contemplados nesse trabalho temos os mesmos de Hsiao et al. (2007).

Por fim, Aparna et al. (2009) divulgaram uma análise cristalográfica, enzimática e de virulência do gene *lipA* de *Xoo* BXO43 mostrando que este difere do dobramento canônico de outras lipases. Sua atividade se restringe a triacilglicerídeos de cadeia curta, não agindo sobre cadeias mais longas que C6 e possui características que indicam a uma ação sobre alquil ésteres de ligação cruzada entre polissacarídeos da parede vegetal. A mutação de alguns resíduos particulares dessa proteína inibem respostas de imunidade inata na planta de arroz (deposição de calose e morte celular programada). Esse gene não foi incluído na análise por se tratar de um dobramento proteico pouco usual para PCWDE e novo na literatura. Porém, foram encontrados outros dois grupos homólogos: um com um gene de cada um dos genomas de *Stenotrophomonas* e o outro em todas as *Xanthomonadaceae*. Esses genes possuem o domínio Pfam LIP, pertencente ao clã AB\_hydrolase, mas ainda não é contemplado no CAZy.

## 2 CONCLUSÕES

Nesse projeto procuramos estabelecer um processo de validação de anotação de genes PCWDE enumerando-os e prevendo seus substratos *in silico* e testando suas atividades enzimáticas *in vitro*. Procuramos também delinear uma provável história evolutiva das enzimas de *Xanthomonadaceae* pertencentes a esse aparato de catabolismo.

Cerca de 20% dos genes enumerados estavam anotados como proteínas hipotéticas. No restante, alguns genes tinham uma anotação ainda superficial ou defasada e havia grande descompasso no vocabulário utilizado para descrição das anotações dos genes. Algumas dificuldades no início da enumeração foram encontradas em casos em que encontramos elementos de transposição fusionados aos genes de PCWDE, especialmente em *X. oryzae*.

Com a otimização do processo de enumeração eliminamos genes com indícios de ação sobre não-PCWPS, especialmente genes relacionados aos processos da parede celular bacteriana. Apesar disso, dada à complexidade dos PCWPS, procuramos abranger ao máximo as PCWDE correspondentes, talvez até abarcando algumas das cazys não-PCWDE, sendo que há partes compartilhadas da composição de PCWPS com outros polissacarídeos.

Desenvolvemos um método que liga os termos encontrados na literatura a modelos estatísticos que refletem a conservação da estrutura terciária de proteínas PCWDE. Para se adquirir a sensibilidade alcançada, os bancos de dados do projeto CAZy foram de grande valia, especialmente na forma de HMMs montados no projeto paralelo dbCAN. CAZy e dbCAN são em conjunto, em seu estado atual, um grande exemplo do potencial que os HMM tem a oferecer.

Apesar da riqueza de números EC recuperados do banco da IUBMB encontramos uma limitação no número de sequências moleculares de referência no banco do KEGG. Isso provavelmente não reflete o estado atual do KEGG, pois se descobriu posteriormente que o projeto foi fechado para acesso pago por falta de verbas de pesquisa, sendo disponibilizada publicamente uma versão que se torna mais defasada a cada dia.

Uma característica das PCWDE que se tornou mais clara ao longo do desenvolvimento é a inespecificidade de ação perante diferentes substratos, que também pode ser averiguada nos testes enzimológicos. Uma consulta ao site

BRENDA a respeito de PCWDE já caracterizadas, mostra que de fato podem existir diversos substratos para uma mesma enzima, porém talvez com diferentes níveis de afinidade. Por outro lado, há restrições de ação das enzimas envolvendo o micro ambiente próximo contendo ramificações e outras modificações existentes nos PCWPS.

A diferença no número de genes PCWDE levantados para *Xanthomonas* e *Xylella* está provavelmente correlacionada com a diferença de estilo de vida observada para esses dois patógenos. Enquanto uma pode ser encontrada na forma de vida livre no ambiente, a outra fica restrita ao xilema de seu hospedeiro e ao trato digestivo do inseto vetor. Já *Stenotrophomonas* ainda tem laços fortes com seus ancestrais que provavelmente tinham um contato bastante próximo com plantas. O espectro de ação de *Xylella* parece ser voltado para celulose, mananos, xiloglucanos e  $\beta$  1-3,1-4 glucanos enquanto o ancestral de *Stenotrophomonas* aparenta ter tido maior contato com xilanos, galacturonanos, homogalacturonanos e ramnogalacturonano-I. *Xanthomonas* não aparenta ter uma tendência no seu potencial de hidrólise. Talvez por isso, e pelo grande número de genes, seja tão mais capaz de macerar o tecido de seus hospedeiros.

O grupo mais abrangente de genes que compõem o aparato de PCWDE deve ter ocorrido em concerto próximo da derivação de *Xanthomonas axonopodis citri*, contando com alguns eventos de HGT que ocorreram antes da divergência entre *Xanthomonas* e *Stenotrophomonas* e após divergência de *Xylella*. Alguns eventos isolados de HGT possivelmente também ocorreram nas *Xanthomonas* mais basais (*X. sachari*, *X. vesicatoria* e *X. glycines*) que não são compartilhados com quase nenhum outro genoma. *X. campestris* teve duplicações recentes que devem contribuir para o fenótipo de maceração mais agressiva. Algumas diferenças mais sutis de polimorfismos foram implicadas na adaptação à arquitetura e composição dos diferentes PCWPS de cada hospedeiro.

Estabelecemos também critérios de seleção de genes candidatos para testes enzimáticos com o intuito de contribuir para a anotação e para o levantamento com fins tecnológicos de uso dessas enzimas. Foram selecionados 24 genes candidatos dos quais 3 permaneceram no início do processo, 11 foram clonados a partir do material genômico, 3 puderam ser subclonados no vetor de expressão. 5 produziram proteína e 2 demonstraram atividade enzimática. Pudemos ao final fazer a refino da anotação de dois genes ortólogos de *Xanthomonas oryzae* pv. *oryzae* e

*Xanthomonas oryzae* pv. *oryzicola* (xop04177 e xoz00638). Originalmente eram descritos como celulase, uma anotação pouco específica, mas após os ensaios verificamos que se tratava de uma endo-glucoronoarabinoxilanase. Os processos de enumeração e predição de substrato foram validados, permitindo também um refino. Apesar de esses genes terem sido recuperados por serem glicosil hidrolases do tipo 5, com um amplo espectro de atuação sobre os PCWPS, o resultado dos ensaios indicou a existência de um subtipo de GH5 com atuação específica sobre arabinoxilano.

Com o advento das tecnologias de sequenciamento massivo de DNA ganhando momento à mesma época do início do projeto, foi natural a incorporação de vários genomas ao longo do projeto. Ao iniciarmos as primeiras análises contávamos com cerca de 10 genomas disponíveis publicamente em diferentes níveis de fechamento e anotação. Já a última análise incorporou 36 genomas com vistas a dezenas ainda por virem a público em um período de poucos meses. Tornando-se o completo sequenciamento de genomas um evento cada vez mais trivial muda-se a restrição do processo experimental da leitura do DNA para sua interpretação e para as derivações dessa interpretação. Isto é, para a bioinformática e para a biologia molecular, aonde esse trabalho vem, portanto, a contribuir.



## REFERÊNCIAS\*

Aparna G, et al. A cell wall-degrading esterase of *Xanthomonas oryzae* requires a unique substrate recognition module for pathogenesis on Rice. *The Plant Cell*. 2009;21:1860-73.

Carpita NC, Gibeaut DM. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*. 1993;3(1):1-30.

Da Silva ACR, et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*. 2002;417:459-63.

Darvill JE, et al. Structure of Plant Cell Walls - XI. Glucuronoarabinoxylan, a second hemicellulose in the primary cell walls of suspension-cultured sycamore cells. *Plant Physiol*. 1980;66:1135-9.

De Vries RP. Synergy between enzymes from *Aspergillus* involved in the degradation of plant cell wall polysaccharides. *Carbohydrate Research*. 2000;327:401–10.

Dongen SV. Graph clustering via a discrete uncoupling process. *SIAM Journal on Matrix Analysis and Applications*. 2008;30(1):121-41.

Encarnação TBC. Modulação da degradação enzimática de galactomanano por sua própria estrutura fina. [dissertação (Mestrado em Ciências Biológicas)]. São Paulo: Instituto de Biociências, Universidade de São Paulo; 2012.

Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle; 2005. [cited from 2008 Jun 11]. Available from: [evolution.genetics.washington.edu/phylip.html](http://evolution.genetics.washington.edu/phylip.html).

Gardy JL. PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics*. 2005;21(5):617-23.

Gilbert HJ, Stalbrand H, Brumer H. How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Current Opinion in Plant Biology*. 2008;11:338–48.

Hsiao YM, et al. Regulation of the *pehA* gene encoding the major polygalacturonase of *Xanthomonas campestris* by Clp abnd RpfF. *Microbiology*. 2008;154:705-13.

Hsieh YSY, Harris PJ. Xyloglucans of Monocotyledons Have Diverse Structures. *Molecular Plant*. 2009;2(5):943–65.

\*De acordo com: International Committee of Medical Journal Editors. [Internet]. Uniform requirements for manuscripts submitted to Biomedical Journal: sample references. [updated 2011 Jul 15]. Available from: <http://www.icmje.org>

Hu J, Qian W, He C. The *Xanthomonas oryzae* pv. *oryzae* eglXoB endoglucanase gene is required for virulence to Rice. FEMS Microbiol Lett. 2007;269:273-9.

Imai K, et al. SOSUI-GramN: high performance prediction for subcellular localization of proteins in Gram-negative bacteria. Bioinformatics. 2008;2(9):417-21.

Kamoun S, Kado CI. Phenotypic switching affecting chemotaxis, xanthan production, and virulence in *xanthomonas campestris*. Appl Environ Microbiol. 1990;56(12):3855-60

Kirchberg J, Büttner D, Thiemer B, Sawers RG. Aconitase B is required for optimal growth of *Xanthomonas campestris* pv. *vesicatoria* in pepper plants. PLoS One. 2012;7(4):e34941.

Lagaert S, Beliën T, Volckaert G. Plant cell walls: Protecting the barrier from degradation by microbial enzymes. Seminars in Cell & Developmental Biology. 2009;20:1064–73.

Lee BM, et al. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Research. 2005;33(2):577–86.

Lee ST, Lee JJ. Insoluble dye substrate for screening and assay of xylan-degrading enzymes. Journal of Microbiological Methods. 1997;29:1-5.

Li L, Stoeckert Jr. CJ, Roos DS. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res. 2003;13:2178-89.

Lima WC, et al. Non-gamma-proteobacteria gene islands contribute to the *Xanthomonas* genome. OMICS. 2005;9(2):160-72.

Lin RJ, Capage M, Hill CW. A repetitive DNA sequence, rhs, responsible for duplications within the *Escherichia coli* K-12 chromosome. Journal of Molecular Biology. 1984;177(1):1–18.

Löwer M, Schneider G. Prediction of type III secretion signals in genomes of gram-negative bacteria. PLoS One. 2009;4(6):e5917.

Lu Z, et al. Predicting subcellular localization of proteins using machine learned classifiers. Bioinformatics. 2004;20(4):547–56.

Lu H, et al. Acquisition and evolution of plant pathogenesis-associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. PLoS One. 2008;3(11):e3828

Moreira LRS, Filho EXF. An overview of mannan structure and mannan-degrading enzyme systems. Appl Microbial Biotechnology. 2008;79:165–78.

Moreira LM, et al. Comparative analyses of *Xanthomonas* and *Xylella* complete genomes. *OMICS*. 2005;9(1):43-76.

Moreira LM, et al. Comparative genomics analyses of citrus-associated bacteria. *Annu Rev Phytopathol*. 2004;42:163–84.

Ochiai H, et al. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *JARQ*. 2005;39(4):275–87.

Pradhan BB, Ranjan M, Chatterjee S. XadM, a novel adhesin of *Xanthomonas oryzae* pv. *oryzae*, exhibits similarity to Rhs family proteins and is required for optimum attachment, biofilm formation and virulence. *Mol Plant Microbe Interact*. 2012;25(9):1157-70.

Rademaker JLW, et al. A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology*. 2005;95:1098-111.

Rajeshwari R, Jha G, Sonti RV. Role of an in planta-expressed xylanase of *Xanthomonas oryzae* pv. *oryzae* in promoting virulence on rice. *Molecular Plant-Microbe Interactions*. 2005;18(8):830-7.

Ramos PL, et al. Screening for endophytic nitrogen-fixing bacteria in Brazilian sugar cane varieties used in organic farming and description of *Stenotrophomonas pavanii* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. 2011;61:926–31.

Ray SK, Rajeshwari RR, Sonti RV. Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in General Secretory Pathway are virulence deficient and unable to secrete xylanase. *Mol Plant Microbe Interact*. 2000;13(4):394-401.

Ridley BL, O'Neill MA, Mohnen D. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry*. 2001;57:929–67.

Sambrook J, MacCallum P, Russell D. *Molecular cloning: a laboratory manual*. 3rd ed. New York: CSH Press; 2001.

Scheller HV, Ulvskov P. Hemicelluloses. *Annu Rev Plant Biol*. 2010;61:263–89.

Schröter K, et al. *Xanthomonas campestris* pv. *campestris* secretes the endoglucanases ENGXCA and ENGXCB: construction of an endoglucanase-deficient mutant for industrial xanthan production. *Appl Microbiol Biotechnol*. 2001;55:727-33.

Simpson AJ, et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature*. 2000;406:151-60.

Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22(21):2688-90.

Strange RN, Scott PR. Plant Disease: A Threat to Global Food Security. *Annu Rev Phytopathol.* 2005;43:83–116.

Sun QH, et al. Type-II secretion pathway structural gene *xpsE*, xylanase and cellulose secretion and virulence in *Xanthomonas oryzae* pv. *oryzae*. *Plant Pathology.* 2005;54:15-21.

Suzek BE, et al. A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics.* 2001;17:1123–30.

Ten LN, et al. Development of a plate technique for screening of polysaccharide-degrading microorganisms by using a mixture of insoluble chromogenic substrates. *J. Microbiol. Methods.* 2004;56(3):375-82.

Thieme F, et al. Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. *Journal of bacteriology.* 2005;187(21):7254–66.

Toth IK, Pritchard L, Birch PRJ. Comparative genomics reveals what makes an enterobacterial plant pathogen. *Annual Review of Phytopathology.* 2006;44:305-36.

Untergasser A, et al. Primer3 - new capabilities and interfaces. *Nucleic. Acids. Res.* 2012;40(15):e115.

Van Sluys MA, et al. Comparative analyses of the complete genome sequences of Pierce's Disease and Citrus Variegated Chlorosis strains of *Xylella fastidiosa*. *Journal of Bacteriology.* 2003;185(3):1018-26.

Van Sluys MA, et al. Comparative genomic analysis of plant-associated bacteria. *Annu. Rev. Phytopathol.* 2002;40:169–89.

Vincken JP, et al. If homogalacturonan were a side chain of rhamnogalacturonan i. implications for cell wall architecture. *Plant Physiology.* 2003;132:1781–9.

Vorhölter FJ, et al. Comparison of two *Xanthomonas campestris* pathovar *campestris* genomes revealed differences in their gene composition. *Journal of Biotechnology.* 2003;106:193–202.

Wolf A, et al. *Stenotrophomonas rhizophila* sp. nov., a novel plant-associated bacterium with antifungal properties. *International Journal of Systematic and Evolutionary Microbiology.* 2002;52:1937–44.

Xiao Z, et al. Mining *Xanthomonas* and *Streptomyces* genomes for new pectinase-encoding sequences and their heterologous expression in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 2008;78:973-981.

Yamada T, et al. Prediction and identification of sequences coding for orphan enzymes using genomic and metagenomic neighbours. *Molecular Systems Biology.* 2012;8:581.

Ye J, et al. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;18:13-134.

Yin Y, et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*. 2012;40:1–7.

Yu C, et al. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci*. 2004;13:1402–06

Zhang D, et al. Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct*. 2012;7(1):18.