Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse III Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :

Interactions Plantes-Microorganismes

Présentée et soutenue par :
Guillaume DEJEAN

le : vendredi 2 septembre 2011

Titre :

Caractérisation et conservation d'un nouveau système CUT associé à l'utilisation du xylane chez Xanthomonas campestris pv. campestris : implications en écologie microbienne

---

Ecole doctorale :
Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :
Laboratoire des Interactions Plantes Microorganismes (UMR CNRS 2594 / INRA 441)

Directeur(s) de Thèse :
Pr. Matthieu ARLAT (Professeur de l'Université Paul Sabatier, Toulouse)
Dr. Emmanuelle LAUBER (Chargée de Recherche CNRS, Toulouse)

Rapporteurs :

Dr. Daniela BÜTTNER, Dr. Evelyne FORANO, Dr. Cindy E. MORRIS

Membre(s) du jury :

Dr. Daniela BÜTTNER (Senior Scientist, Institute of Biology, MLU, Halle-Wittenberg)
Dr. Evelyne FORANO (Directrice de Recherche INRA, Clermont-Ferrand Theix)
Dr. Cindy E. MORRIS (Directrice de Recherche INRA, Avignon)
Dr. Gabrielle POTOCKI-VERONESE (Chargée de Recherche INRA, LISBP, INSA Toulouse)
Pr. Christophe ROUX (Professeur de l'Université Paul Sabatier, Toulouse)

SOMMAIRE

LISTE DES FIGURES

INTRODUCTION GENERALE

LA GUERRE AUX PORTES DE LA PLANTE

LES TBDTS PEUVENT TOUT « FER »

RESULTATS

DISCUSSION GENERALE

LISTE DES TABLEAUX

INTRODUCTION GENERALE

RESULTATS

DISCUSSION

ABREVIATIONS

aa	acide aminé
ABC	ATP Binding Cassette
AD	Activation Domain
ADN	Acide Deoxy-ribo-Nucléique
AF	Arabinofuranosidase
AHLs	*N*-acyl Homoserine Lactones
AMP	Adénosine Monophosphate
ARN	Acide Ribo-Nucléique
Avr	Avirulence
CAZy	Carbohydrate Active Enzymes
CBD	Cellulose-binding domain
CBM	Carbohydrate-binding module
CDS	Coding DNA Sequence
CE	Carbohydrate Esterase
CERMAV	Centre de Recherche sur les Macromolécules Végétales
CLP	cAMP receptor Like Protein
CUT	Carbohydrate Utilization containing TBDT
CW	Cell Wall
Da	Dalton
DF	Diffusible Factor
DSF	Diffusible Signal Factor
ECF	Extracellular Cytoplasmic Function
EPS	Exopolysaccharide
FHA	Filamentous Hemaglutinine Antigen
Fur	Ferric Uptake Regulator
GA	Acide Galacturonique
GAX	Glucuronoarabinoxylane
GH	Glycoside Hydrolase
GlcNAc	*N*-acétylglucosamine
GMP	Guanosine Monophosphate
GOS	Global Ocean Sampling
GRP	Glycin-rich proteins
GT	Glycosyl Transferase
HG	Homogalacturonane
Hop	Hrp outer protein
HR	Hypersensitive Response
Hrc	Hrp conserved
HRGP	hydroxyprolin-rich proteins
Hrp	Hypersensitive response and pathogenicity
IS	Insertion Sequences
kb	kilobase
K_d	Constante de dissociation
kDa	Kilodalton
LPS	Lipopolysaccharides
Mb	Megabase
ME	Membrane Externe

MI	Membrane Interne
MP	Membrane Plasmique
MFS	Major Facilitator Superfamily
Mur	Manganese Uptake Regulator
NLS	Nuclear localization signal
OMP	Outer Membrane Protein
ORF	Open Reading Frame
PEG	Polyethylène Glycol
pb	paires de bases
PGA	Acide Polygalacturonique
PIP-box	Plant Inducible Promoter box
PL	Polysaccharide Lyase
PM	Plasma Membrane
PME	Pectine Méthyl Estérase
Pop	*Pseudomonas* outer protein
PRP	Prolin-rich proteins
PTI	Pathogen Triggered Immunity
PUL	Polysaccharide Utilization Locus
pv.	pathovar
QS	Quorum Sensing
RGI	Rhamnogalacturonane I
RGII	Rhamnogalacturonane II
ROS	Reactive Oxygen Species
Rpf	Regulation of pathogenicity factors
SLH	Surface Layer Homology
SLP	Surface Layer Protein
SNP	Single Nucleotide Polymorphism
sp.	species
sRNA	small Ribonucleic Acid
SST2	Système de Sécrétion de Type 2
SST3	Système de Sécrétion de Type 3
SST5	Système de Sécrétion de Type 5
SUMO	Small Ubiquitin-like Modifier
sus	Starch utilization system
sux	Sucrose utilization in *Xanthomonas*
T2S	Type 2 secretion
T3S	Type 3 secretion
TAL	Transcription-activator like
TBDT	TonB-Dependent Transporter
TPS	Two-partner secretion
Ulp	Ubiquitin-like protease
UPA	Upregulated by AvrBs3
X_2	Xylobiose
X_4	Xylotetraose
Xop	*Xanthomonas* outer protein
XTH	Xyloglucan endotransglycosylase hydrolase
Zur	Zinc uptake regulator

Bactéries :

Pst	*Pseudomonas syringae* pv. *tomato*
Rs	*Ralstonia solanacearum*
Xcc	*Xanthomonas campestris* pv. *campestris*
Xac	*Xanthomonas axonopodis* pv. *citri*
Xca	*Xanthomonas campestris* pv. *armoraciae*
Xcv	*Xanthomonas campestris* pv. *vesicatoria*
Xoo	*Xanthomonas oryzae* pv. *oryzae*
Xoc	*Xanthomonas oryzae* pv. *oryzicola*
Xalb	*Xanthomonas albilineans*

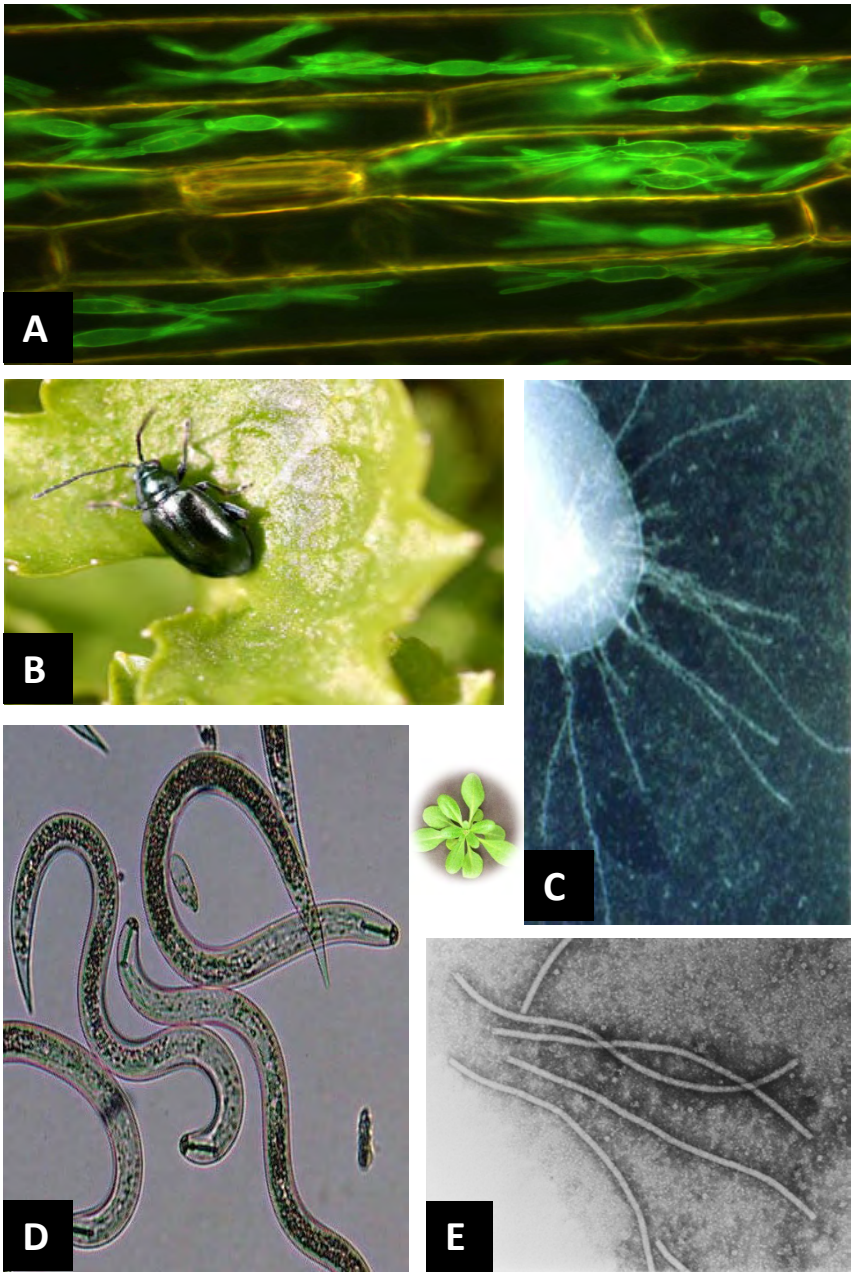


Figure 1. Les principaux organismes phytopathogènes

A. Observation par microscopie epifluorescente de la colonisation de cellules vivantes epidermiques de l'espece d'orge *Hordeum vulgare* (en jaune) par le champignon *Blumeria graminis* (en vert). Image : Pietro Spanu, Imperial College London.

B. L'insecte *Diabrotica virgifera* ou Altise noire des crucifères - photo : http://www.inra.fr

C. La bactérie *Ralstonia solanacearum* - Photo de Jacques Vasse, LIPM

D. Le nématode blanc de la pomme de terre (*Globodera pallida*), ou aussi « nématode à kystes de la pomme de terre » - Photo de Olivier Plantard, INRA

E. Le virus PVY (Potato Virus Y) – photo : www.lookfordiagnosis.com

AVANT-PROPOS

Des bactéries et des plantes

Comme l'a suggéré Edward Osborne Wilson, considéré comme l'un des pères de la notion de « biodiversité », les vertébrés, les coraux et les plantes sont relativement bien connus et constituent le fondement de la pratique actuelle de conservation (Wilson, 1992, 1994). Cependant, la science reste encore ignorante quant à l'immense majorité des insectes et autres invertébrés, tout comme à la quasi totalité des bactéries et du reste des microorganismes. Nous ne connaissons quasiment rien à propos de la vie microbienne. Pourtant, la vie sur Terre dépend de la santé et du devenir de cette population microbienne. Les microorganismes, participent pour une très large part à l'équilibre biologique existant à la surface de la Terre. Les bactéries, en particulier, jouent un rôle dans le cycle des nutriments des sols et des environnements aquatiques et sont notamment capables de fixer l'azote (Dixon and Kahn, 2004). En plus de cette activité d'autoépuration de la planète, les bactéries demeurent importantes dans l'industrie et les technologies humaines, au travers des produits de fermentation et des applications biotechnologiques ou encore dans la lutte biologique (Ahmed and Leather, 1994; Starnes *et al.*, 1993).

Malheureusement, certains microorganismes peuvent provoquer des maladies ayant un impact fort sur les animaux et les plantes. Les agents pathogènes végétaux et animaux sont très similaires et utilisent des stratégies très semblables pour infecter leurs hôtes. Les microorganismes pathogènes incluent des virus, des viroïdes, des bactéries, des champignons et des protozoaires. Il y a également des nématodes et bien sûr, les plantes sont souvent blessées par des attaques d'insectes (Figure 1). Elles sont également soumises à des stress abiotiques comme des conditions environnementales défavorables telles que la carence ou l'excès de nutriments, les variations d'humidité et de lumière, les températures extrèmes, les vents et la présence de composés toxiques dans l'air ou le sol. Les plantes souffrent aussi de la compétition avec les autres plantes. Contrairement aux animaux qui ont la capacité de se déplacer lorsque les conditions de vie ne leur sont plus favorables, les plantes, incapables de se mouvoir, doivent mettre en oeuvre des stratégies d'adaptation complexes, souvent au détriment de leur croissance. Les plantes constituent un réservoir d'énergie dont dépendent directement ou indirectement les hommes ainsi que tous les animaux.

Ainsi, une meilleure compréhension des interactions plantes-microorganismes relève d'une importance économique majeure. De manière générale, la phytopathologie s'intéresse à l'étude des organismes et des facteurs environnementaux qui causent des maladies chez les plantes. Une des finalités de la phytopathologie tend à la mise en place de stratégies de lutte durables plus efficaces et plus respectueuses de l'environnement.

De façon surprenante, malgré une sensibilisation croissante à l'importance des microorganismes, nous connaissons encore très peu de choses quant à « qui fait quoi et comment » dans le monde microbien. Ainsi, de manière plus globale, notre travail a pour objectif de trouver des éléments de réponse à des problématiques qui définiront et décrypteront la biodiversité microbienne. Nous nous sommes efforcés de comprendre le fonctionnement de modèles microbiens dans des conditions standards de « laboratoire ». Cependant, nous devons en apprendre bien davantage sur les cycles de vie et les relations écologiques, tant pour les espèces connues que pour les espèces inconnues. L'avenir s'écrit dans la compréhension du fonctionnement et du devenir des microbes dans leurs habitats naturels tels que la rhizosphère et la phyllosphère. Ainsi, cela passe par l'étude de l'expression des gènes du microbe dans un certain environnement, l'analyse de la fonction de ces gènes, l'étude des interactions entre les produits de ces gènes et la détermination de conditions (biotiques ou abiotiques) qui gouvernent ces processus. C'est dans ce contexte que nous avons choisi d'étudier les déterminants de l'adaptation à son milieu et de la mise en place du pouvoir pathogène de la bactérie phytopathogène *Xanthomonas campestris* pv. *campestris* (*Xcc*), modèle d'étude de notre groupe.

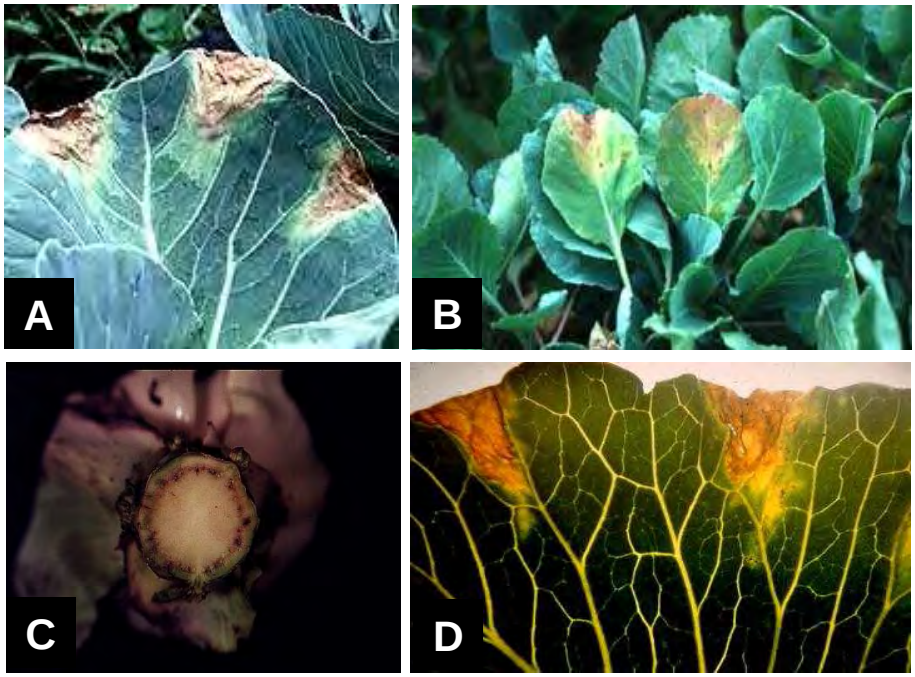


Figure 2. Symptômes de pourriture noire sur chou
Zones chlorotiques et/ou nécrotiques en forme de “V” (photos A, B et D) et noircissement des vaisseaux du xylème (photo C) produits par l’infection du chou par la bactérie *Xanthomonas campestris* pv. *campestris* (*Xcc*).
A. avrdc.org
B. http://ohioline.osu.edu/hyg-fact/3000/3125.html
C. http://ipm.uiuc.edu/vegetables/diseases/black_rot/index.html
D. http://urbanext.illinois.edu

INTRODUCTION GENERALE

I. Notre modèle d'étude, la bactérie phytopathogène *Xanthomonas campestris* pv. *campestris*

I.1. Les plantes hôtes de *Xanthomonas campestris* pv. *campestris*

Xanthomonas campestris pv. *campestris* (*Xcc*) est l'agent responsable de la pourriture noire ou nervation noire des brassicacées, maladie la plus préjudiciable de par le monde concernant cette famille de plante (Williams, 1980) (Figure 2). Cette bactérie est capable d'infecter la plupart des brassicacées d'intérêt agronomique (le chou, le chou de Bruxelles, le chou-fleur, le brocoli, le navet, le radis, la moutarde, le canola et le rutabaga sauf certains cultivars de radis et de chou frisé qui sont moins facilement infectés) ainsi que la plante modèle *Arabidopsis thaliana*. Cette dernière constitue un modèle de choix en phytopathologie pour étudier les bases moléculaires de l'interaction entre la plante et la bactérie. En effet, les génomes de ces deux organismes ont été séquencés (Arabidopsis, 2000; da Silva *et al.*, 2002; Qian *et al.*, 2005; Vorholter *et al.*, 2008) et ce pathosystème est bien caractérisé (Bent *et al.*, 1992; Lummerzheim *et al.*, 1993; Meyer *et al.*, 2005; Parker *et al.*, 1993; Simpson and Johnson, 1990; Suji, 1988).

La famille des Brassicacées, anciennement nommées Crucifères, constitue une famille importante au sein des dicotylédones. Elle comprend 3200 espèces réparties en 350 genres (Angiosperm Phylogeny Website, www.mobot.org/MOBOT/Research/APweb/welcome.html). Ce sont principalement des plantes herbacées qui se répartissent sur tous les continents mais qui sont particulièrement présentes dans l'hémisphère nord. Une des caractéristiques intéressante des Brassicacées est la production de métabolites secondaires, les glucosinolates, principalement présents dans cette famille de plantes. Ces composés organiques joueraient un rôle dans le dialogue entre l'hôte et le pathogène ou agresseur. En effet, les glucosinolates sont impliqués dans la défense des plantes contre des bactéries et des champignons phytopathogènes et jouent un rôle dans la résistance des plantes contre des insectes ravageurs (Clay *et al.*, 2009; Tierens *et al.*, 2001).

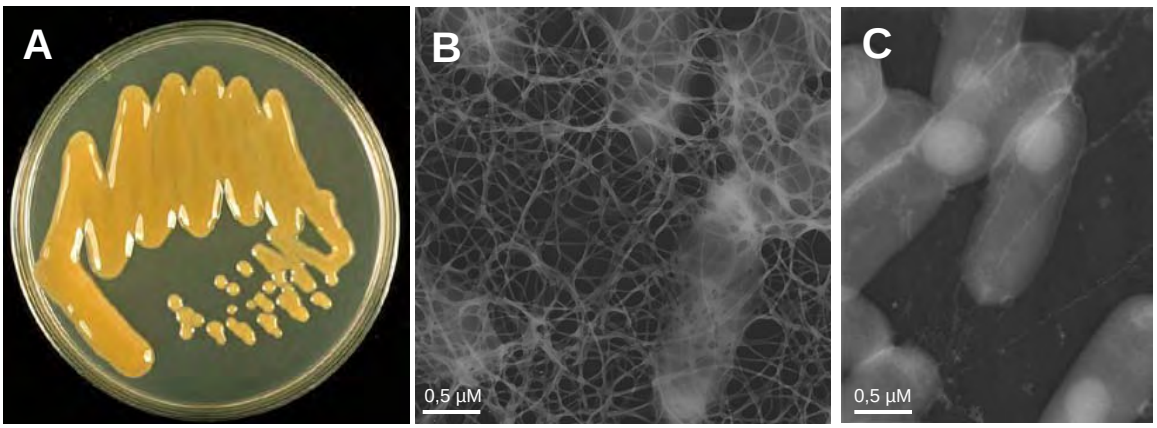


Figure 3. *Xanthomonas campestris* pv. *campestris* (*Xcc*).
A. Colonies sur milieu solide riche.
B. Bactéries enchassées dans leur matrice exopolysaccharidique (xanthane).
C. Souche mutante de *Xcc*, déficiente dans la synthèse d'exopolysaccharides.
B et **C**. Photos de microscopie électronique de Jacques Vasse, LIPM.
(D'après Blanvillain, 2007).

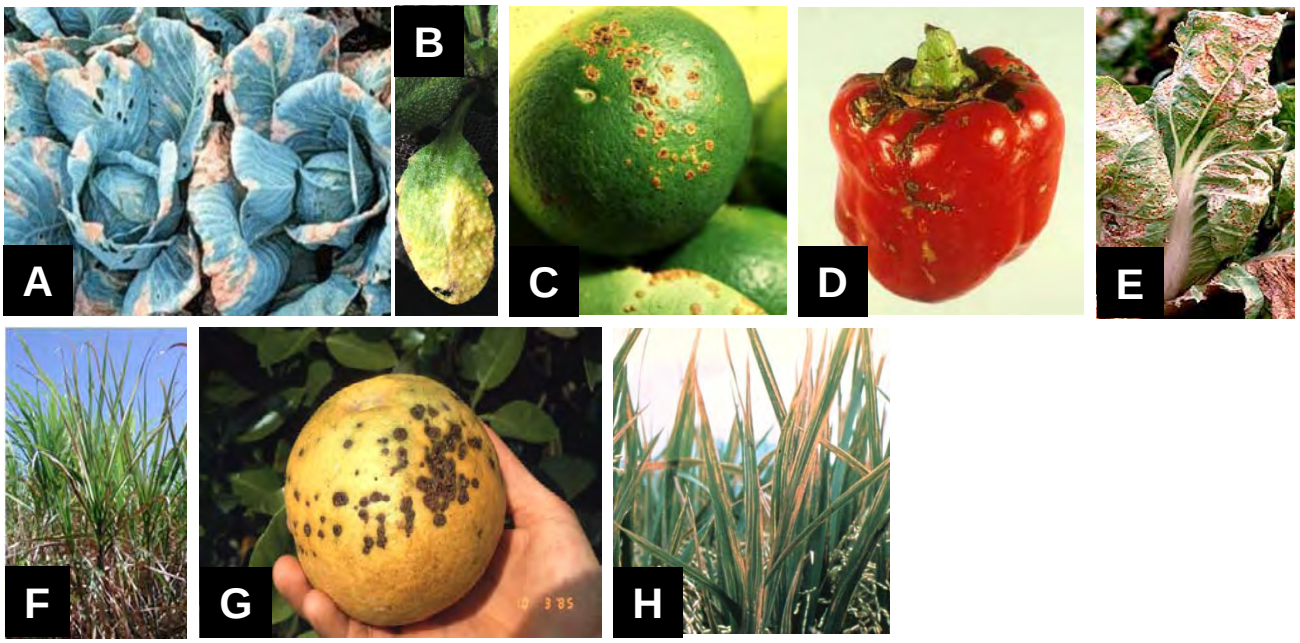


Figure 4. Le genre *Xanthomonas* : un bon modèle en génomique. Symptômes causés par différentes espèces de Xanthomonas sur des plantes d'intérêt agronomiques et sur des plantes modèles
A et B. Symptômes de pourriture noire sur chou (A) et sur *Arabidopsis thaliana* (B) par *X. campestris* pv. *campestris* (Photos, A : http://growingideas.johnnyseeds.com; B : Guillaume Déjean, LIPM). **C.** Symptômes de chancre des agrumes sur citron par *X. axonopodis* pv. *citri* (Photo : lookfordiagnosis.com). **D.** Symptômes de tâche bactérienne sur poivron par *X. campestris* pv. *vesicatoria* (Photo : ipmimages.org). **E.** Symptômes de tâche foliaire sur chou par *X. campestris* pv. *armoraciae* (Photo : bitkisagligi.net). **F.** Symptômes d'échaudure des feuilles sur canne à sucre par *X. albilineans* (Photo : umr-bgpi.cirad.fr). **G.** Symptômes de chancre des agrumes sur pamplemousse par *X. axonopodis* pv. *aurantifolii* (Photo : apsnet.org). **H.** Symptômes de nielle des feuilles sur riz par *X. oryzae* pv. *oryzae* (Photo : forestryimages.org).

I.2. Le genre *Xanthomonas*

Les *Xanthomonas* sont des bactéries à Gram négatif appartenant à la classe des γ-protéobactéries, à la famille des *Xanthomonadaceae* et à l'ordre des *Xanthomonadales*. Etymologiquement, *Xanthomonas* vient du grec « *xanthos* » qui signifie « jaune » et de « *monad* » qui signifie « unité ». La couleur jaune est due à la synthèse de caroténoïdes liés à la membrane externe, les xanthomonadines, pigments qui auraient des propriétés protectrices et antioxydantes (Rajagopal *et al.*, 1997) (Figure 3A). Les bactéries du genre *Xanthomonas* présentent une grande diversité de pouvoir pathogène et une spécificité d'hôte très étroite. Ces derniers points ont gêné pendant longtemps la mise en place d'une classification définitive. Les bactéries du genre *Xanthomonas* infectent 124 espèces de plantes monocotylédones et 268 espèces de plantes dicotylédones (Leyns *et al.*, 1984) (Figure 4). En 1981, Starr M.P. développa le concept de « new host-species » qui est basé sur le fait que chaque membre de *Xanthomonas* qui présente un spectre d'hôte différent ou qui produit des symptômes de maladie différents, est classé comme une espèce à part entière. Dans les années 90, cela a conduit à la division du genre en plus de 170 espèces (Starr, 1981). Les espèces furent subdivisées en plus de 140 pathovars (pv.) sur la base de leur spectre d'hôte (Swings, 1993) ; http://www.ncbi.nlm.nih.gov/Taxonomy). Les progrès en génotypage, biologie moléculaire et taxonomie ont permis une reclassification du genre, menant à la création de 20 espèces (Vauterin *et al.*, 1995). L'espèce *Xanthomonas campestris* présente 6 pathovars : *aberrans*, *armoraciae*, *barbareae*, *campestris*, *incanae* et *raphani* (Vauterin *et al.*, 1995) plus récemment reclassés en 3 pathovars : *campestris*, *raphani* et *incanae*, suite à des analyses de pouvoir pathogène (Fargier and Manceau, 2007).

Xanthomonas campestris pv. *campestris* (*Xcc*) est très étudiée pour sa capacité à produire le xanthane (Figure 3B), un exopolysaccharide qui donne un aspect mucoïde aux colonies (Figure 3A) et qui est très utilisé comme agent viscosant et stabilisant dans les industries cosmétiques, agroalimentaires et dans le bâtiment (Becker *et al.*, 1998). Le xanthane n'est pas requis pour le pouvoir pathogène mais est indispensable à la survie épiphyte de la bactérie (Dunger *et al.*, 2007).

Xcc est une bactérie en bâtonnet (0,7 à 1,8 µm de long et 0,4 à 0,7 µm de large), aérobie stricte, pourvue d'un flagelle polaire unique (Guerrero, 2001; Onsando, 1992).

La lutte contre *Xcc*, comme pour la majorité des phytobactérioses, est excessivement difficile. En effet, l'efficacité de la lutte chimique reste partielle et en France comme dans de nombreux

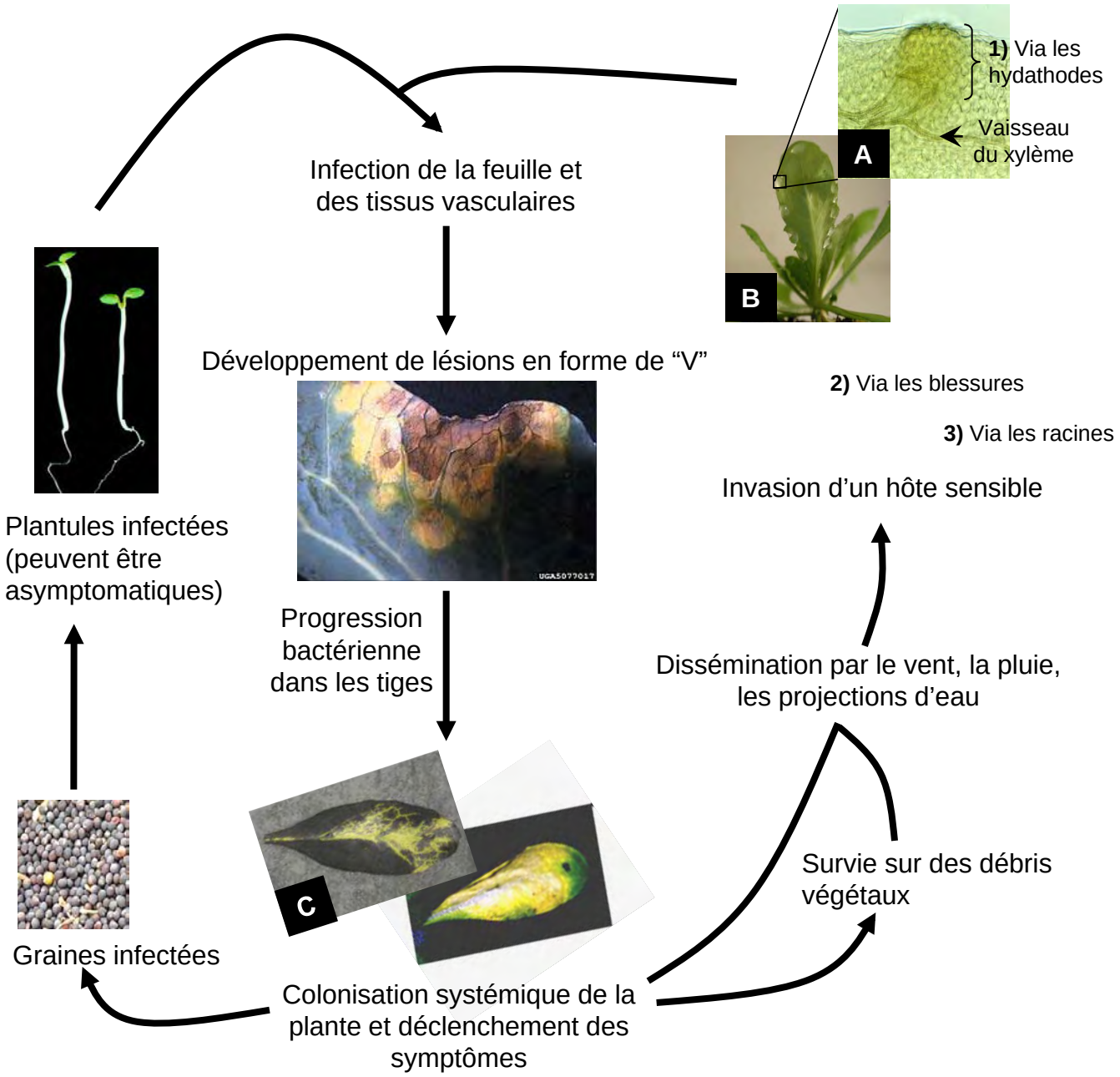


Figure 5. Représentation schématique du cycle de vie de *Xanthomonas campestris* pv. *campestris* (*Xcc*) lors d’une interaction avec la plante hôte.

Après une phase de vie épiphyte, *Xcc* va, dans certaines conditions, pénétrer à l’intérieur des tissus végétaux par les hydathodes (1) mais aussi par les blessures (2) et les racines (3), puis envahir les vaisseaux du xylème et induire les symptômes de la maladie. Enfin, la bactérie est également capable de survivre dans le sol et sur les débris végétaux après l’infection. La dissémination des bactéries s’effectue par le vent, la pluie, les insectes, mais aussi et surtout par le transport et l’utilisation de graines contaminées.

A. Vue en microscopie d’un hydathode d’*Arabidopsis thaliana* (Photo de Jacques Vasse, LIPM)

B. Gouttes de guttation présentes à la marge d’une feuille d’*Arabidopsis thaliana* (Photo d’Alice Boulanger, LIPM).

C. *Xcc* exprimant le gène de la luciférase (Lux) constitutivement (D’après Meyer *et al*., 2005). Cette photo illustre la colonisation des vaisseaux du xylème par les bactéries.

pays, les antibiotiques sont interdits en productions végétales. La lutte se réduit donc essentiellement à des méthodes prophylactiques (Porteneuve, 2002). La mise au point de ces méthodes nécessite une connaissance approfondie des étapes du cycle épidémiologique des maladies. Les semences infectées représentent la source majeure de dissémination de la maladie. Pour limiter cela, l'usage de semences indemnes de *Xcc* et l'utilisation de la rotation avec des cultures « non brassicacées » ou résistantes à *Xcc* sont préconisés.

I.3. Le cycle de vie de *Xanthomonas campestris* pv. *campestris*

Xcc est une bactérie épiphyte, c'est-à-dire que c'est une bactérie pouvant vivre et se multiplier à la surface des plantes (pour revue (Beattie and Lindow, 1999)). Les étapes d'installation de la bactérie sur la plante dépendent de nombreux éléments environnementaux dont les facteurs nutritionnels. Il est intéressant de noter que des plantes non hôtes peuvent également être colonisées par des populations épiphytes de *Xanthomonads*. Ce réservoir de population pourra alors servir d'inoculum pour développer la maladie sur des plantes sensibles. Lorsque les conditions environnementales sont favorables, soit une température comprise entre 25°C et 30°C et un taux d'humidité élevé, elle va développer un cycle infectieux sur sa plante hôte (Szczesny *et al.*, 2010) (Figure 5).

La première étape d'infection de *Xcc* correspond à l'entrée de la bactérie dans les tissus de la feuille. Plusieurs possibilités s'offrent à *Xcc* pour pénétrer dans les tissus végétaux. La bactérie peut pénétrer par des blessures causées par des vents violents, la grêle, les insectes etc. L'infection par des blessures racinaires est également possible. Outre les blessures, cette bactérie utilise également les hydathodes comme voie d'entrée pour la colonisation endophyte des brassicacées (Hugouvieux *et al.*, 1998) (Figure 5A). Les hydathodes, stomates aquifères localisés en marge des feuilles à l'aboutissement des vaisseaux du xylème, permettent la transpiration des plantes (Cook, 1952). Dans des conditions d'humidité élevée, des gouttes de guttation vont se créer en marge des feuilles, au niveau des hydathodes (Figure 5B). Les bactéries présentes à la surface des feuilles vont alors être incorporées dans ces gouttes et pénètreront dans les tissus vasculaires lorsque les gouttes seront réabsorbées. Enfin, *Xcc* peut aussi pénétrer dans la feuille par les stomates (Buell, 2002). Chez Arabidopsis, la souche 8004 de *Xcc* peut entrer dans les tissus vasculaires via les hydathodes et les stomates. Cependant, la voie d'entrée privilégiée dépend à la fois de l'écotype d'Arabidopsis et des conditions environnementales (Hugouvieux *et al.*, 1998).

Tableau 1 : Liste des bactéries phytopathogènes dont le génome est entièrement séquencé et disponible, et maladies provoquées par ces organismes.

D'après http://cpgr.plantbiology.msu.edu/cgi-bin/warehouse/cpgr_warehouse.cgi?group=Bacteria&status=Finished

Groupe	Espèce	Maladie	Taille du génome (MégaBases)
Gram - α–proteobactéries	*Agrobacterium tumefaciens* str. C58	galle du collet (crown gall)	5.67
	Agrobacterium vitis S4	galle du collet du raisin (Crown gall of grape)	6.3
Gram - β-proteobactéries	*Acidovorax avenae subsp. citrulli AAC00-1*	maladie des tâches des cucurbitacés (Bacterial Fruit Blotch disease)	5.4
	Burkholderia cenocepacia AU 1054	pourriture aigre (sour skin) de l'oignon - pathogène sur patients atteints de fibrose kystique	7.25
	Burkholderia cenocepacia HI2424	pourriture aigre (sour skin) de l'oignon - pathogène sur patients atteints de fibrose kystique	8.09
	Burkholderia cenocepacia J2315	pourriture aigre (sour skin) de l'oignon - pathogène sur patients atteints de fibrose kystique	8.07
	Burkholderia cenocepacia MC0-3	pourriture aigre (sour skin) de l'oignon - pathogène sur patients atteints de fibrose kystique	7.9
	Burkholderia glumae BGR1	pourriture des plantules et rouille des panicules du riz (seedling rot and panicle blight of rice)	7.27
	Ralstonia solanacearum GMI1000	flétrissement bactérien (bacterial wilt)	5.81
Gram - γ-proteobactéries	*Erwinia carotovora* subsp. *atroseptica* SCRI1043	pourriture molle (soft rot) et jambe noire (blackleg) de la pomme de terre	5.06
	Pseudomonas aeruginosa PA7	pourriture molle (soft rot) - pathogènes sur patients atteints de fibrose kystique	6.6
	Pseudomonas aeruginosa PAO1	pourriture molle (soft rot) - pathogènes sur patients atteints de fibrose kystique	6.26
	Pseudomonas aeruginosa UCBPP-PA14	pourriture molle (soft rot) - pathogènes sur patients atteints de fibrose kystique	6.53
	Pseudomonas syringae pv. *phaseolicola* 1448A	graisse à halo du haricot (halo blight of bean)	6.11
	Pseudomonas syringae pv. *syringae* B728a	graisse du haricot (brown spot)	6.09
	Pseudomonas syringae pv. *tomato* str. DC3000	moucheture bactérienne (bacterial speck) de la tomate	6.54
	Xanthomonas albilineans str. GPE PC73	échaudure des feuilles de canne à sucre (leaf scald)	3.85
	Xanthomonas axonopodis pv. *citri* str. 306	chancre des agrumes (citrus canker)	5.27
	Xanthomonas campestris pv. *armoraciae* str. 756C	tâche foliaire (leaf spot) des crucifères	4.94
	Xanthomonas campestris pv. *campestris* str. B100	pourriture noire (black rot) des crucifères	5.1
	Xanthomonas campestris pv. *campestris* str. 8004	pourriture noire (black rot) des crucifères	5.15
	Xanthomonas campestris pv. *campestris* str. ATCC 33913	pourriture noire (black rot) des crucifères	5.08
	Xanthomonas campestris pv. *vesicatoria* str. 85-10	tache bactérienne (bacterial spot) de la tomate et du poivron	5.42
	Xanthomonas oryzae pv. *oryzae* KACC10331	rouille des feuilles (bacterial blight) du riz	4.94
	Xanthomonas oryzae pv. *oryzae* MAFF 311018	rouille des feuilles (bacterial blight) du riz	4.94
	Xanthomonas oryzae pv. *oryzae* PXO99A	rouille des feuilles (bacterial blight) du riz	5.2
	Xanthomonas oryzae pv. *oryzicola* str. BLS256	strie des feuilles (bacterial streak) du riz	4.80
	Xylella fastidiosa 9a5c	chlorose variégée des agrumes (citrus variegated chlorosis)	2.73
	Xylella fastidiosa M12	chlorose variégée des agrumes (citrus variegated chlorosis)	2.48
	Xylella fastidiosa M23	chlorose variégée des agrumes (citrus variegated chlorosis)	2.5
	Xylella fastidiosa Temecula1	maladie de Pierce (Pierce's disease) du raisin	2.52
Gram +	*Clavibacter michiganensis* subsp. *michiganensis* NCPPB 382	chancre de la tomate (Bacterial canker of tomato)	3.4
	Clavibacter michiganensis subsp. *sepedonicus* ATCC33113	pourriture annulaire de la pomme de terre (potato ring rot)	3.45
	Leifsonia xyli subsp. xyli str. CTCB07	rabougrissement des repousses (ratoon stunting) de la canne à sucre	2.58
	Streptomyces scabies 87-22	gale commune (scab) de la pomme de terre	10.15
Phytoplasmes	*Aster yellows witches'-broom phytoplasma AYWB*	jaunisse de l'aster (aster yellows) avec symptôme en balai de sorcière (witches' broom)	0.72
	Onion yellows phytoplasma OY-M	jaunisse de l'oignon	0.86

Après pénétration dans la feuille, *Xcc* peut envahir les vaisseaux du xylème, permettant alors une colonisation systémique de la plante et le déclenchement des symptômes de la maladie (Figure 5C). Une température et une humidité élevées favorisent le développement de la pourriture noire. Les symptômes caractéristiques de cette maladie sont la formation de lésions jaunes en forme de « V » devenant brunes et nécrosées au fil de la progression des bactéries (Figure 2A, B, D). Les feuilles gravement infectées se dessèchent entièrement et peuvent alors tomber. Entre deux cycles d'infection, *Xcc* est capable de survivre pendant plusieurs années sur des débris végétaux, dans le sol et à la surface des graines issues des plantes infectées (Schaad *et al.*, 1980; Schultz and Gabrielson, 1986). La capacité de *Xcc* à exploiter les débris végétaux est possible grâce à la sécrétion d'un grand nombre d'enzymes extracellulaires (Tang *et al.*, 1991) (Cf. p33, paragraphe IV.2).

I.4. Le génome de *Xanthomonas campestris* pv. *campestris*

Les génomes de 3 souches de *Xcc* ont été entièrement séquencés et sont disponibles (Tableaux 1 et 2). Il s'agit de la souche ATCC33913 séquencée au Brésil (da Silva *et al.*, 2002), de la souche 8004 séquencée en Chine (Qian *et al.*, 2005) et de la souche B100 séquencée en Allemagne à l'université de Bielefeld (Vorholter *et al.*, 2008). Actuellement, les séquences génomiques de 11 souches de *Xanthomonas* (*Xanthomonas axonopodis* pv. *citri* (*Xac*) 306, *Xcc* 8004, ATCC33913 et B100, *Xanthomonas campestris* pv. *armoraciae* (*Xca*) 756C, *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) 85-10, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) KACC10331, MAFF311018 et PXO99A, *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) BLS256 et *Xanthomonas albilineans* (*Xalb*) GPE PC73) représentant 5 espèces, sont disponibles (Tableau 1) (da Silva *et al.*, 2002; Lee *et al.*, 2005; Ochiai *et al.*, 2005; Qian *et al.*, 2005; Salzberg *et al.*, 2008; Thieme *et al.*, 2005; Vorholter *et al.*, 2008)(Pieretti *et al.*, 2009). A ce jour, parmi les génomes séquencés, la présence de plasmides est restreinte à *Xac* souche 306, *Xcv* souche 85-10 et *Xalb* souche GPE PC73. Les trois génomes de *Xcc* ne présentent qu'un seul chromosome, ont un fort pourcentage en GC (65%) et sont riches en éléments transposables. Si le degré de conservation de séquence entre les génomes des souches ATCC33913 et 8004 est très fort, des réarrangements génomiques significatifs (translocations, inversions, insertions et délétions) différencient les deux souches. De plus, il semble que ces deux souches présentent de légères différences dans leur pouvoir pathogène sur différentes plantes : la souche ATCC33913 serait légèrement moins agressive sur certains

Tableau 2 : Caractéristiques générales et comparaison des 3 génomes de *Xanthomonas campestris* pv. *campestris* (*Xcc*) séquencés.

CDS, Coding DNA Sequences ; SNP, Single Length Polymorphism ; IS, Séquences d'insertion. (D'après Qian *et al.* , 2005 ; Vorholter *et al.* , 2008).

	***Xcc* 8004**	***Xcc* ATCC33913** [a]	***Xcc* B100**
Caractéristiques générales			
Taille du génome (pb)	5 148 708	5 076 187	5 079 002
Pourcentage en GC	64.94%	65.00%	65.00%
Nombre total de séquences codantes (CDS)	4273 (87)[b]	4181	4471
Nombre de CDS avec fonction assignée	2671 (1)[b]	2708	2878
Nombre de CDS codant de putatives protéines conservées	1523 (27)[b]	1276	1323
Nombre de CDS codant des protéines hypothétiques	79 (59)[b]	198	270
Différences de séquences dans les CDS			
Nombre de gènes parfaitement identiques	3467[c]	3408	
Nombre de gènes identiques, de même taille, avec SNP	498	500[c]	
Nombre de gènes identiques avec insertions ou délétions	200	211[c]	
Nombre de gènes spécifiques de la souche	108	62	
Taille moyenne des CDS (pb)	1023	1030	
Nombre de séquences d'insertion (IS)	115	109	59

[a] Données de da Silva *et al*. (2002) et GenBank (accession n°AE008922).
[b] Les nombres entre parenthèses indiquent le nombre de CDS possédant une protéine identique chez *Xcc* ATCC33913
[c] Traduit l'existence de duplications de CDSs.

cultivars de choux et de radis (Qian *et al.*, 2005). Excepté pour une petite région inversée, le chromosome de la souche B100 est colinéaire avec celui de la souche 8004. En revanche, ces 2 souches diffèrent de la souche ATCC33913 par une inversion d'un énorme fragment chromosomique (Vorholter *et al.*, 2008). La souche 8004 présente 92 séquences codantes putatives (CDS) supplémentaires par rapport à la souche ATCC33913 (Tableau 2). De façon intéressante, la souche B100 présente 496 CDS supplémentaires qui ne sont pas annotées dans les deux autres souches de *Xcc* (Tableau 2). Parmi ces 496 CDS, 88 ont été identifiées dans des génomes de *Xanthomonas* et 408 CDS ne sont retrouvées dans aucune des séquences de génomes de *Xanthomonas* disponibles (Vorholter *et al.*, 2008). Cependant, un TBLASTN a permis de montrer que plus de la moitié de ces « nouvelles CDS » sont également présentes dans les deux autres souches de *Xcc* (Vorholter *et al.*, 2008).

A ce jour, outre les *Xanthomonas*, les génomes de 33 bactéries phytopathogènes sont disponibles et 29 sont en cours de séquençage (http://cpgr.plantbiology.msu.edu/cgi-bin/warehouse/cpgr_warehouse.cgi?group=Bacteria&status=Finished) (Tableau 1). Au total, les génomes bactériens complètement séquencés sont au nombre de 1428 (au 20.04.2011) (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi), alors qu'ils n'étaient seulement que 209 en 2005. Cette augmentation exponentielle du nombre de données de séquences génomiques et l'analyse de ces données permettent une meilleure compréhension du comportement et des interactions régnant entre les divers organismes vivant sur Terre.

II. Les déterminants du pouvoir pathogène chez les *Xanthomonas*

Au cours de leur cycle de vie, les bactéries phytopathogènes ne se comportent pas uniquement en tant que pathogènes au sens propre. Elles possèdent notamment une phase de vie épiphyte à la surface des plantes ou/et saprophyte dans l'eau ou le sol. Ainsi, les bactéries phytopathogènes ont développé des déterminants moléculaires et des stratégies fines leur permettant de survivre et de se multiplier dans les environnements rencontrés.

Toutefois, le pouvoir pathogène étant une caractéristique importante des *Xanthomonas*, nous allons nous attacher à décrire dans cette partie les principaux déterminants génétiques et moléculaires du pouvoir pathogène bactérien, en prenant pour exemple autant que possible ceux identifiés chez les *Xanthomonadaceae* qui sont représentés sur la Figure 6.

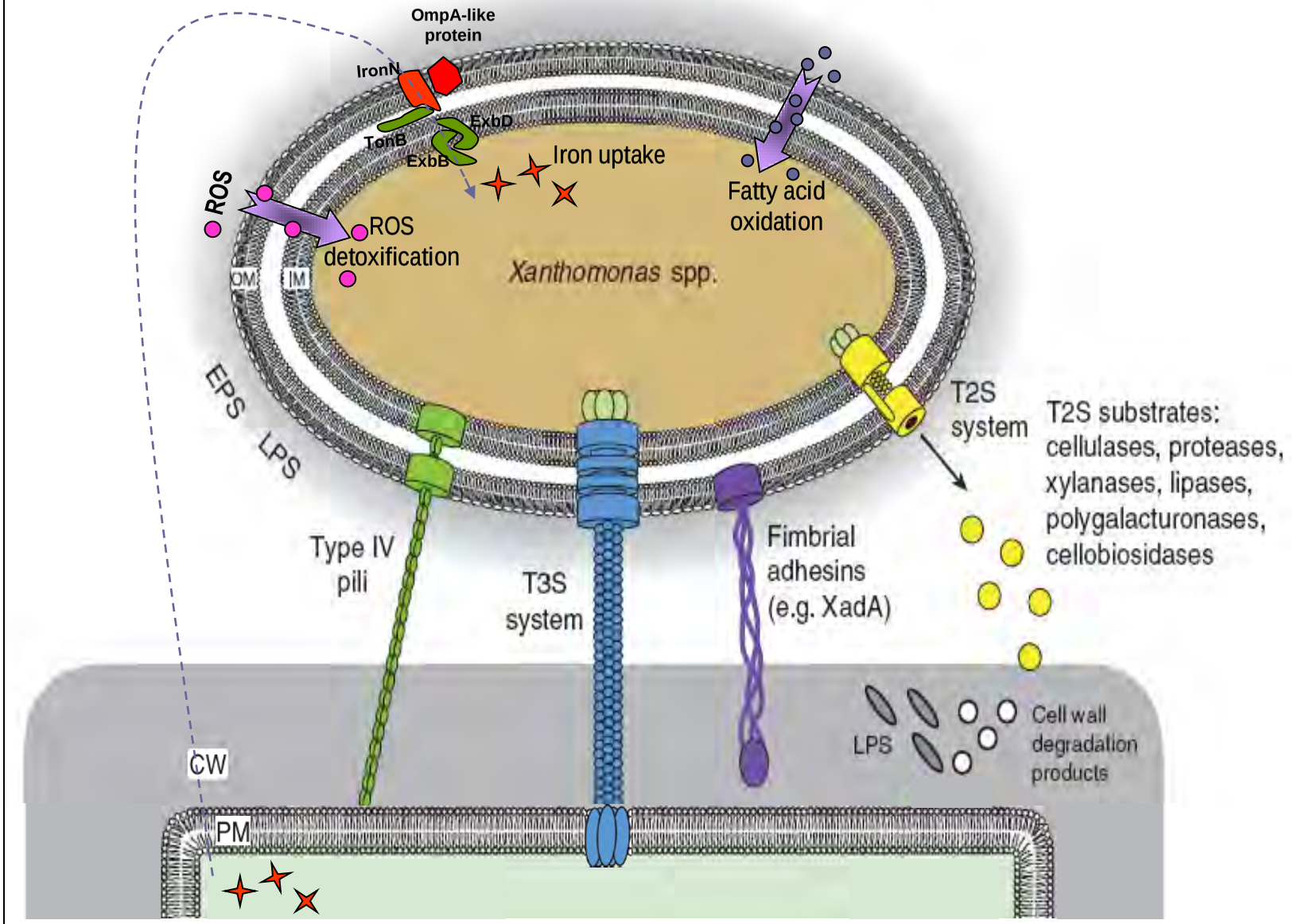


Figure 6. Illustration des facteurs de virulence connus chez les *Xanthomonas* spp. lors des interactions avec leurs plantes hôtes.

EPS, Exopolysaccharides ; LPS, Lipopolysaccharides ; ROS, Reactive Oxygen Species ; T2S, Type 2 secretion ; T3S, Type 3 secretion ; CW, Cell Wall ; PM, Plasma Membrane (Adapté de Büttner and Bonas, 2010).

II.1. Adhésion et Mobilité

Il s'agit ici des évènements initiaux de l'interaction entre l'hôte et le pathogène. Après avoir été transmis à son hôte, la bactérie doit être capable de s'attacher à la plante afin d'établir un centre d'infection et ainsi de se disséminer. Cette étape d'adhésion est assurée par des protéines bactériennes, les adhésines (Figure 6) qui sont ancrées dans la membrane externe. Ces protéines sont classées en adhésines fimbriales et non-fimbriales.

Les adhésines fimbriales ont des structures protéiques filamenteuses de type pili, tels que le pilus de type I ou aussi le pilus de type IV qui est structurellement homologue au pilus périplasmique du système de sécrétion de type 2 (SST2) (Gerlach and Hensel, 2007).

Les adhésines non-fimbriales comprennent en particulier les adhésines autotransportées par le système de sécrétion de type V (SST5) (exemple de YadA chez *Yersinia* spp.) et des substrats du TPS (Two-partner secretion) (exemples de la FHA, hémagglutinine filamenteuse chez *Bordetella pertussis* et YapH chez *Yersinia* spp.) (Gerlach and Hensel, 2007). Ces adhésines vont se lier à des récepteurs spécifiques de la cellule hôte tels que des glycolipides ou des protéoglycanes (Rostand and Esko, 1997).

On connaît peu de choses à propos de la fonction de virulence des adhésines chez les bactéries phytopathogènes. L'analyse comparative des séquences génomiques révèle que les bactéries phytopathogènes possèdent un nombre d'adhésines tel qu'il permettrait aux bactéries de se fixer à de mutliples récepteurs des cellules hôtes et contribuerait à différentes étapes du processus infectieux (Das *et al.*, 2009). Parmi les adhésines décrites chez *Xanthomonas* spp., on trouve XadA et XadB (homologues de YadA chez *Yersinia* spp.), des homologues de YapH, des protéines « hemagglutinin-like » et des protéines potentiellement impliquées dans la synthèse du pilus de type IV (da Silva *et al.*, 2002; Lee *et al.*, 2005; Ochiai *et al.*, 2005; Qian *et al.*, 2005; Salzberg *et al.*, 2008; Thieme *et al.*, 2005; Vorholter *et al.*, 2008). Chez *Xanthomonas campestris* pv. *hyacinthi*, le pilus de type IV est impliqué dans la fixation de la bactérie aux stomates des plantes hôtes (van Doorn *et al.*, 1994). A ce jour, il a été montré que les adhésines de *Xoo*, *Xac* et de *Xanthomonas fuscans* ssp. *fuscans* sont impliquées dans la virulence bactérienne et l'adhésion aux feuilles et/ou aux graines (Darsonval *et al.*, 2009; Das *et al.*, 2009; Gottig *et al.*, 2009; Ray *et al.*, 2002). L'analyse par mutagenèse des gènes codant pour les adhésines chez *X. fuscans* ssp. *fuscans* révèle que les adhésines contribuent individuellement et de manière complémentaire aux différentes étapes du processus infectieux (Darsonval *et al.*, 2009). De manière intéressante, il est à noter que chez certaines bactéries

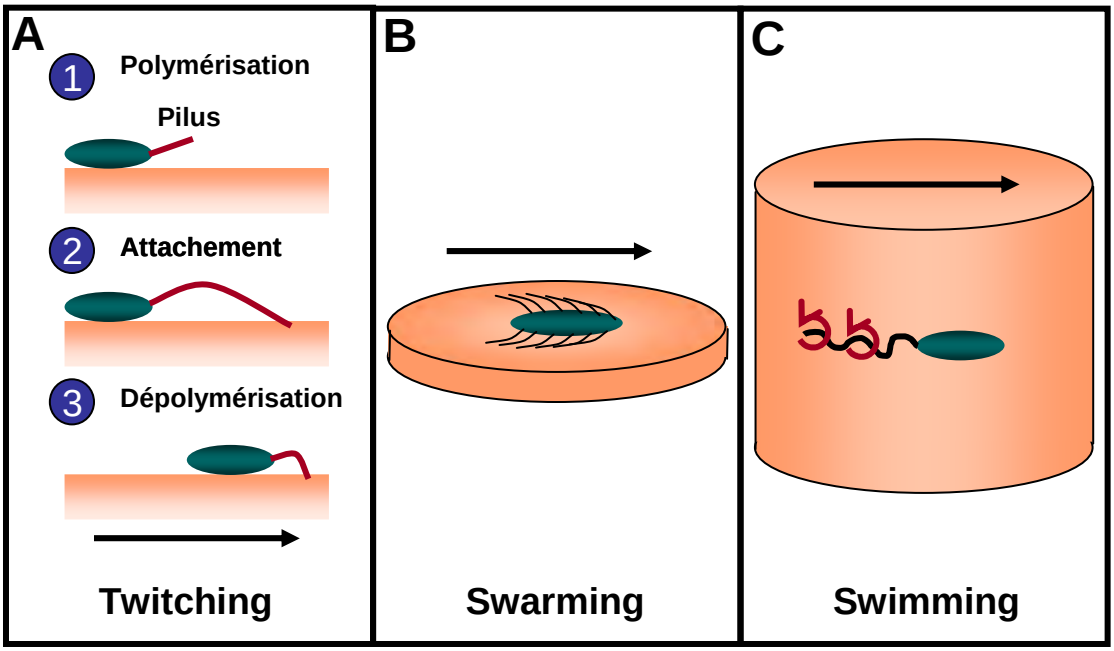


Figure 7. Représentation schématique de trois types de mobilité bactérienne. (D'après Boulanger, 2009).
A. Mobilité par « twitching » via le pilus de type IV. Ce processus peut être divisé en trois étapes : (1) la polymérisation du pilus, (2) l'attachement du pilus à la surface de migration, (3) la dépolymérisation du pilus entrainant la tractation de la bactérie.
B. Mobilité par « swarming » via la rotation de plusieurs flagelles latéraux.
C. Mobilité par « swimming » grâce à la rotation d'un flagelle polaire.

cellulolytiques telle que *Ruminococcus albus*, des pili de type IV semblent impliqués dans l'adhésion à la cellulose (Rakotoarivonina *et al.*, 2002).

Sur la surface foliaire et dans la plante, *Xcc* a besoin de se déplacer pour gagner les sites d'infection et coloniser le xylème. Il existe plusieurs types de mobilité (Figure 7). La mobilité par «twiching» ou mobilité par convulsion, représente un mode de mobilité bactérienne qui consiste en un glissement saccadé des bactéries le long de surfaces (Figure 7A). Elle est réalisée grâce aux pili de type IV (Craig and Li, 2008; Kaiser, 2007). Les mobilités par «swarming» et «swimming» sont réalisées généralement par le flagelle, long appendice protéique flexible et motorisé (Figure 7B et 7C) (pour revue (Chevance and Hughes, 2008)). Le filament flagellaire est composé d'une succession de flagelline, protéine pouvant être reconnue par la plante et induire les systèmes de défense (Felix *et al.*, 1999). La production de flagelle semble induite dans des conditions de carence nutritionnelle. Après pénétration dans la plante, le flagelle semble ne plus être un facteur de virulence. Il a été observé que *Xcc* est non-motile et non-flagellé dans la sève de xylème de feuilles de choux infectées (Kamoun and Kado, 1990).

II.2. Les phytotoxines et hormones

Les agents pathogènes sont également capables de produire des toxines nocives pour leurs hôtes. Les pathogènes bactériens et fongiques produisent un nombre important de métabolites secondaires toxiques pour les cellules végétales. Cependant, ces métabolites ne sont pas tous importants pour la pathogénicité de l'agent. Les phytotoxines peuvent être spécifiques d'un hôte et posséder le même spectre d'hôte spécifique que le pathogène qui les produit ; cependant elles peuvent également avoir un spectre d'hôte plus élargi que leur agent pathogène producteur. Les phytotoxines bactériennes les plus connues et étudiées sont celles produites par *Pseudomonas syringae* telles que la coronatine empêchant la fermeture des stomates (Melotto *et al.*, 2008) et qui est corégulée avec les gènes *hrp* (hypersensitive response and pathogenicity, Cf. p16, paragraphe II.6) ou la syringomycine et la syringopeptine qui forment des pores dans la membrane plasmique des cellules végétales entraînant une fuite d'électrolytes (Bender *et al.*, 1999). L'éventail de phytotoxines produites par les *Xanthomonas* spp. est plus restreint. Cependant *Xalb* produit l'albicidine qui bloque la réplication de l'ADN chez une large gamme de bactéries à Gram positif et négatif et dans les plastes des cellules de la canne à sucre (Birch and Patil, 1985; Birch and Patil, 1987; Huang *et*

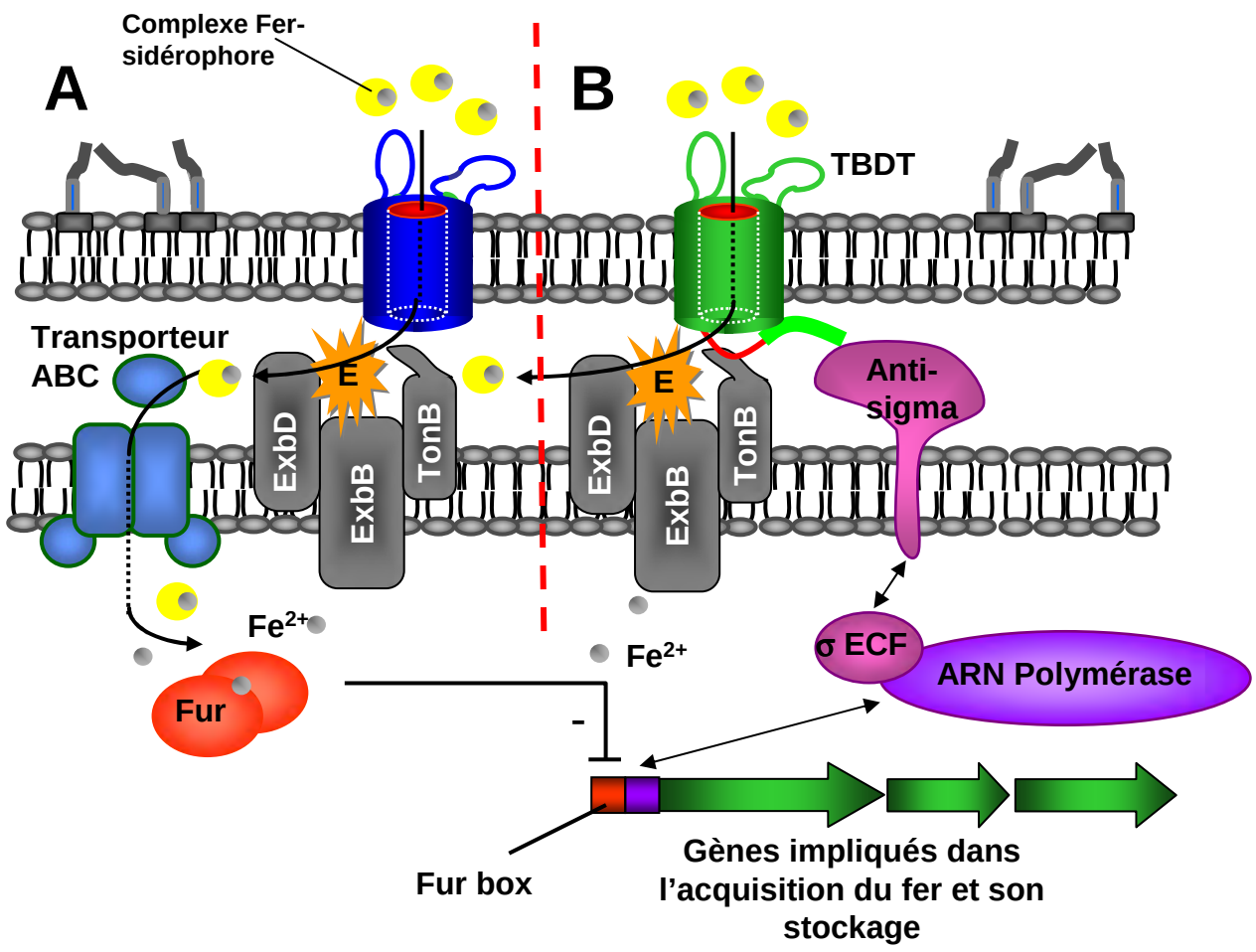


Figure 8. Modèle du système d'acquisition du fer des bactéries Gram ⁻ et de sa régulation.

Le complexe Fe^{3+}-sidérophore est reconnu par le transporteur de la membrane externe, ici un TBDT (TonB-Dependent Transporter). Suivant le cas, le TBDT aura une fonction de transport et/ou une fonction de signalisation. Le transport et la transmission du signal dépendent de l'énergie provenant du complexe TonB–ExbB–ExbD.

A. Le complexe Fe^{3+}-sidérophore est transporté vers le périplasme par le TBDT puis vers le cytoplasme via le transporteur de type ABC. Quand le niveau intracellulaire en fer est élevé le répresseur Fur complexé au fer va réprimer la transcription des gènes impliqués dans l'acquisition et le stockage du fer.

B. Le signal de fixation du complexe Fe^{3+}-sidérophore au TBDT va être transmis de l'extension N-terminale du TBDT vers le facteur anti-sigma qui va libérer un facteur sigma de type ECF permettant la transcription des gènes.

(D'après Schalk *et al.*, 2004 et Boulanger, 2009)

al., 2001). De plus, des gènes *syr* similaires aux gènes de synthèse de la syringomycine de *P. syringae* et situés entre deux transposases, ont été décrits chez *Xac* (Etchegaray *et al.*, 2004).
Les hormones végétales produites par les procaryotes pathogènes peuvent également être considérées comme des phytotoxines. En effet, elles sont produites au cours du cycle infectieux de l'agent pathogène et elles ont un effet délétère sur la plante, remplissant de ce fait les conditions décrites par Durbin en 1991 sur la définition des phytotoxines (Durbin, 1991). Un exemple chez *Ralstonia solanacearum* est l'éthylène bactérien qui est perçu par la plante et peut moduler l'expression des gènes de l'hôte de la même manière que l'éthylène végétal (Valls *et al.*, 2006).

II.3. La disponibilité en fer et l'acquisition des complexes Fer-Sidérophore par le système TBDT-TonB-ExbBD

Le fer est un élément essentiel pour le processus biologique de nombreux organismes vivants (Boelaert, 1996). Il contribue à l'adaptation de la bactérie au cours de son cycle de vie et a donc un rôle indirect dans le pouvoir pathogène. Bien que le fer soit abondant dans la nature, sa disponibilité est limitée du fait de sa faible solubilité (Braun and Killmann, 1999). L'homéostasie intracellulaire en fer doit être maintenue pour la survie cellulaire et pour une protection contre les effets toxiques du fer (Imlay and Linn, 1988). La capture du fer est restreinte à deux formes du fer, le fer ferreux Fe^{2+} (soluble et présent en très faible concentration) et le fer ferrique Fe^{3+} (espèce prédominante et insoluble). Dans l'environnement, le fer libre assimilable dans le milieu n'est disponible qu'à l'état de traces (10^{-18} M) (Raymond *et al.*, 2003). Cependant, les microorganismes requièrent une concentration en fer d'au moins 10^{-6} M pour survivre et ont donc dû développer des systèmes de capture du fer. En effet, bactéries et champignons sécrètent des chélateurs de fer, les sidérophores qui facilitent l'acquisition du Fe^{3+} (Braun and Killmann, 1999; Braun and Braun, 2002; Schryvers and Stojiljkovic, 1999). Les sidérophores sont de petites molécules organiques (Winkelmann, 2002) dont le poids moléculaire est compris entre 500 et 1500 Dalton. Les sidérophores sont nécessaires à la virulence de nombreux pathogènes (Ratledge and Dover, 2000). Par exemple, la production des sidérophores chrysobactine et achromobactine est essentielle pour la mise en place de l'infection par *Dickeya dadantii* (anciennement nommée *Erwinia chrysanthemi*) (Expert, 2005; Franza *et al.*, 2005).

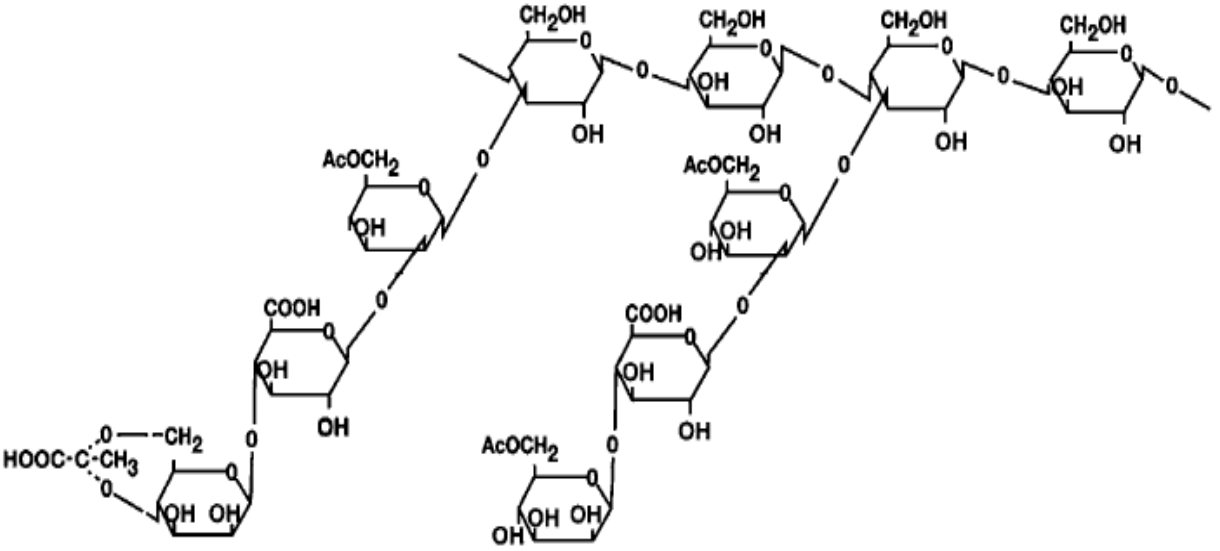


Figure 9. Le xanthane, un exopolysaccharide produit par *Xanthomonas campestris* pv. *campestris.*

Unité structurelle du xanthane produit par la souche 8004 (D'après Yun *et al*., 2006). Le xanthane est un polymère acide constitué d'une unité pentasaccharidique répétée formant un squelette de cellulose avec des chaînes latérales composées d'acide mannose(β-1,4)glucuronique(β-1,2)mannose liées en β-1,3 au squelette. Les résidus mannose peuvent être acétylés ou pyruvylés (Jansson *et al*., 1975).

Au cours d'une interaction avec un hôte eucaryote (infection pathogène ou symbiose), les bactéries peuvent également capter le fer biologiquement disponible dans certaines protéines de l'hôte telles que les protéines de stockage du fer (transferrine, lactoferrine et ferritine) et les hémoprotéines. De plus, de nombreux pathogènes animaux ont développé la capacité de reconnaître directement les hèmes de leur hôte tel que l'hémoglobine (Wagegg and Braun, 1981; Wandersman and Delepelaire, 2004). Les complexes sidérophores/Fe^{3+} sont transportés au travers des membranes bactériennes vers le cytoplasme par un système énergie-dépendant (Figure 8A). Ce système implique des récepteurs/transporteurs de la membrane externe (TonB-dependent transporter ou TBDT, Cf. p37), un transporteur de la membrane interne de type ABC et le complexe protéique TonB-ExbB-ExbD, localisé dans la membrane interne. L'énergie requise pour ce transport provient de la force protomotrice transférée au TBDT par le complexe TonB-ExbB-ExbD (Braun, 1995; Postle, 1993; Schalk *et al.*, 2004). Les TBDTs sont spécifiques d'un sidérophore et ont pour lui une très grande affinité (*Kd* de l'ordre de 0,1 μM) (Ferguson and Deisenhofer, 2002). La fixation du sidérophore au TBDT transduit un signal au travers de la membrane externe induisant une transition allostérique qui se propage jusqu'à la boîte TonB. La nature du signal transduit n'est pas clairement comprise bien que de grands changements conformationnels dans les boucles extracellulaires aient été observés chez les TBDTs FecA (Yue *et al.*, 2003) et ShuA (Cobessi *et al.*, 2010). En effet, les auteurs observent un repliement du haut du TBDT lorsque le sidérophore se lie, séquestrant ce dernier et faisant intervenir de nouveaux résidus au site de liaison.

Xcc possède dans son génome le système TonB-ExbB-ExbD essentiel à l'acquisition du fer mais avec la particularité de posséder 2 copies du gène ExbD : ExbD1 et ExbD2 (Wiggerich *et al.*, 1997). L'analyse de mutants dans ces gènes révèle que ExbD2 n'est pas essentiel pour l'acquisition du fer, la virulence et la croissance des bactéries dans une plante hôte (le chou-fleur) (Wiggerich *et al.*, 1997; Wiggerich and Puhler, 2000). En revanche, il semble que ces 4 gènes soient indispensables à la production d'une réaction hypersensible (HR) chez le poivron, une plante non-hôte. Cela dénote un découplage entre la fonction de capture du fer et la production de la HR.

II.4. Les exopolysaccharides ou EPS

Les espèces du genre *Xanthomonas* produisent un exopolysaccharide caractéristique, le xanthane qui est responsable de l'apparence mucoïde des colonies bactériennes (Figure 3A).

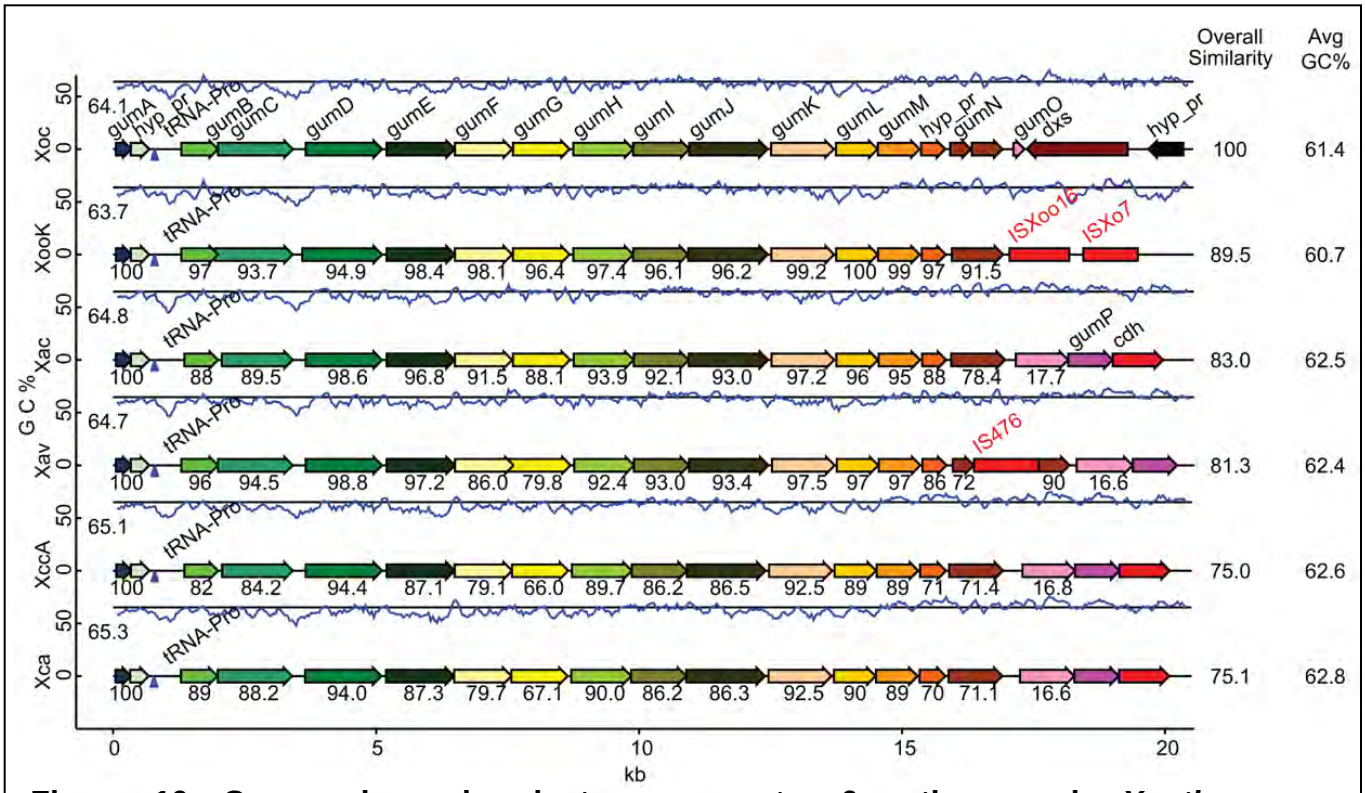


Figure 10. Comparaison du cluster gum entre 6 pathovars de *Xanthomonas* représentant 3 espèces.

Xoc, *Xanthomonas oryzae* pv. *oryzicola* str. BLS256; *XooK*, *Xanthomonas oryzae* pv. *oryzae* str. KACC100331; *Xac*, *Xanthomonas axonopodis* pv. *citri* str. 306; *Xav*, *Xanthomonas campestris* pv. *vesicatoria* str. 85-10; *XccA*, *Xanthomonas campestris* pv. *campestris* str. ATCC33913; *Xca*, *Xanthomonas campestris* pv. *armoraciae* str. 756C. Pour chaque cluster, les gènes homologues sont indiqués dans la même couleur. Le niveau d'identité est indiqué en-dessous de chaque gène. Le trait bleu en-dessous de chaque cluster représente le pourcentage en GC. La ligne noire représente la moyenne du pourcentage en GC du génome, et le chiffre est indiqué au bout de chaque cluster. Les gènes ARNt sont représentés par des triangles bleus. (D'après Lu *et al*, 2008).

Le xanthane est un polymère constitué d'un squelette de molécules de glucose liées en β−1,4 et de chaînes latérales de mannose(β-1,4)acide glucuronique(β-1,2)mannose liées en β-1,3 à la chaîne principale (Jansson *et al.*, 1975; Yun *et al.*, 2006) (Figure 9). Cet EPS est utilisé comme agent viscosant et stabilisant dans les industries cosmétiques, agroalimentaires et dans le bâtiment (Becker *et al.*, 1998). La production de xanthane est contrôlée par plusieurs loci génétiques dont le cluster de gènes *gum*, très conservé chez *Xanthomonas* spp. qui est composé de 12 gènes (*gumB* à *gumM*) (Figure 10) (Katzen *et al.*, 1998; Lu *et al.*, 2007; Vojnov *et al.*, 1998; Vorholter *et al.*, 2008). D'autres ORFs nommés *gumA* et *gumN*, *-O*, *-P*, sont localisés de part et d'autre du locus *gum* mais leur rôle dans la biosynthèse du xanthane n'a pas été démontré (Vanderslice *et al.*, 1990). Une étude récente chez *Xoo* a montré que *gumN* est co-transcrit avec le cluster *gumB-gumM* mais *gumA* est clairement dans un opéron distinct (Yoon and Cho, 2007). Le cluster est très fortemement conservé. Les différences entre les souches sont limitées à des éléments d'insertion de séquence (IS) dans ou près de *gumN* et de faibles homologies de séquence dans les gènes en dehors du cluster : notamment *gumO*, *gumP* et *chd2* (Figure 10).

Du fait de sa structure et de sa cohésion, le xanthane protège les bactéries de stress environnementaux tels que la déshydratation, les ultraviolets (UV) ou encore des composés toxiques. De plus, chez les pathogènes vasculaires, le xanthane peut causer le flétrissement des plantes hôtes en obstruant les vaisseaux du xylème et compromettant de ce fait la circulation de la sève brute (Chan and Goodwin, 1999; Denny, 1995). Les gènes *gum* de plusieurs *Xanthomonas* spp. dont *Xcc*, *Xoo*, *Xac* et *X. axonopodis* pv. *manihotis* contribuent à la survie épiphyte et/ou à la croissance bactérienne *in planta* (Chou *et al.*, 1997; Dharmapuri and Sonti, 1999; Dunger *et al.*, 2007; Katzen *et al.*, 1998; Kemp, 2004; Kim *et al.*, 2009; Rigano *et al.*, 2007). De façon intéressante, chez *Xac*, les gènes *gum* ne sont pas nécessaires au développement de la maladie et à la croissance bactérienne sur *Citrus sinensis* mais ils contribuent à la virulence bactérienne sur *Citrus limon*. Cela suggère que la contribution du xanthane à la virulence dépendrait de la plante hôte et des conditions environnementales (Dunger *et al.*, 2007; Rigano *et al.*, 2007). Des études suggèrent également que le xanthane supprime les défenses basales de la plante tel que le dépôt de callose dans la paroi végétale, vraisemblablement en chélatant les ions calcium présents dans l'apoplaste végétal (Aslam *et al.*, 2008; Yun *et al.*, 2006). De plus, il a été montré chez *Xcc* et *Xac*, que le xanthane est impliqué dans la formation de biofilms (Dow *et al.*, 2003; Rigano *et al.*, 2007; Torres *et al.*, 2007) qui fourniraient une protection contre les antibiotiques et contre les réponses de défense

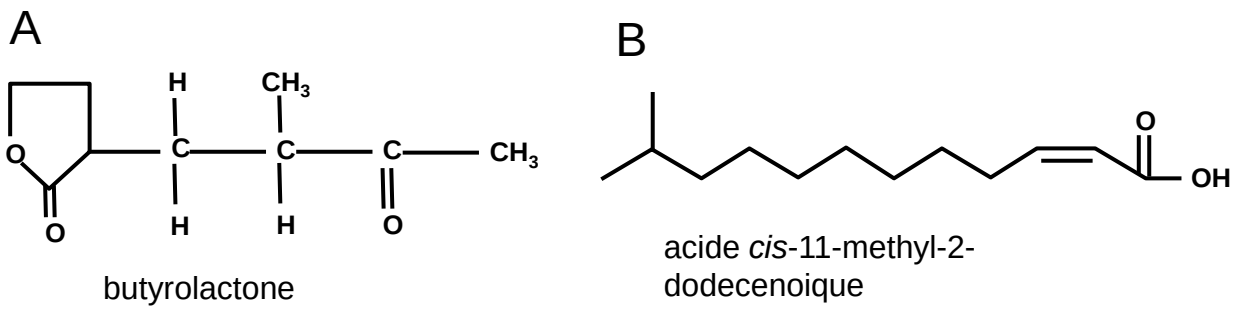


Figure 11. Structures des molécules signales des deux systèmes de quorum sensing de *Xanthomonas campestris* pv. *campestris*.

A. Facteur diffusible (DF), d'après Chun et *al*., 1997.

B. Facteur signal diffusible (DSF), d'après He and Zhang, 2008.

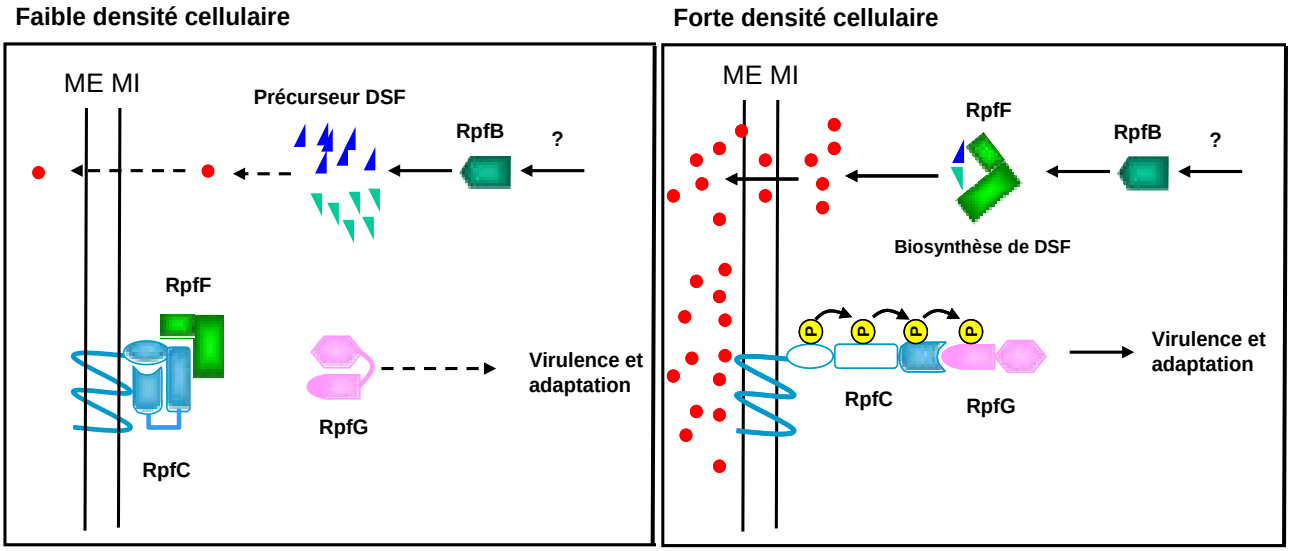


Figure 12. Modèle de la transduction du signal DSF chez *Xanthomonas campestris* pv. *campestris*.

Les protéines RpfB et RpfF sont impliquées dans la synthèse du DSF. A faible concentration cellulaire, le senseur RpfC forme un complexe avec RpfF, la protéine de synthèse du DSF, limitant ainsi la synthèse de DSF. Les flèches en pointillés représentent le taux basal de signalisation. A forte concentration cellulaire, DSF interagit avec RpfC et induit un changement conformationnel de RpfC suivi d'une autophosphorylation menant au relargage de RpfF et au phosphorelais du senseur vers le régulateur RpfG. Il y a alors augmentation de la synthèse de DSF et génération d'un signal fort d'induction du régulon DSF.

?, précurseur inconnu ; ME, membrane externe ; MI, membrane interne ; DSF, facteur signal diffusible.

(D'après He and Zhang, 2008 ; Boulanger, 2009).

de l'hôte. Le biofilm contribuerait également à la survie au cours de la vie épiphyte de la bactérie avant la colonisation de l'hôte végétal (Stoodley *et al.*, 2002). Cependant le rôle du biofilm dans la virulence des bactéries phytopathogènes n'a pas encore été clairement déterminé.

II.5. Les systèmes de Quorum Sensing (QS)

Le QS est un mécanisme de régulation dépendant de la densité de population et contrôlant l'expression de certains gènes bactériens, notamment des gènes de virulence. Le QS permet donc la synchronisation de l'expression (ou de la répression) de gènes cibles au sein d'une population bactérienne. Cela leur permet d'adopter de nouveaux comportements selon l'environnement. Le QS régule diverses fonctions telles que la division cellulaire, la production de facteurs de virulence, la mobilité, la sporulation, la nodulation, le transfert de plasmides, la production d'antibiotiques, la bioluminescence ou encore la formation de biofilms (Bassler, 2002; Tao *et al.*, 2010; Waters and Bassler, 2005; Whitehead *et al.*, 2001). Le QS repose sur la capacité des bactéries à communiquer entre elles en utilisant des signaux moléculaires appelés auto-inducteurs dont la concentration dans l'environnement est directement liée à la densité de la population bactérienne. Les molécules signal les plus communes sont des peptides modifiés chez les bactéries à Gram positif et les *N*-Acyl Homosérine Lactones (AHLs) chez les bactéries à Gram négatif mais d'autres exemples existent (Ryan and Dow, 2008).

Deux systèmes de QS ont été caractérisés chez *Xcc*. Le premier système implique un facteur diffusible (DF), codé par le gène pigB qui joue un rôle important dans la régulation de la production de xanthomonadine (Poplawsky and Chun, 1997). D'après des analyses de spectrométrie de masse, le DF aurait une structure de butyrolactone (Chun *et al.*, 1997) (Figure 11A). Le second système implique la production d'un facteur signal diffusible (DSF) (Barber *et al.*, 1997). Chez *Xcc*, le DSF a été identifié comme étant un acide *cis*-11-methyl-2-dodecanoique (Figure 11B) dont la synthèse dépend des enzymes RpfF (putative hydratase enoyl-CoA) et RpfB (ligase fatty acyl-CoA) (Figure 12). Les gènes codant ces enzymes appartiennent au cluster *rpf* (regulation of pathogenicity factor) (Tang *et al.*, 1991). Du fait de sa nature lipophile, le DSF peut diffuser au travers des membranes bactériennes (Crossman and Dow, 2004). La transduction du signal DSF s'effectue par l'intermédiaire d'un système à deux composants : un senseur à domaine kinase codé par le gène *rpfC* et le régulateur de

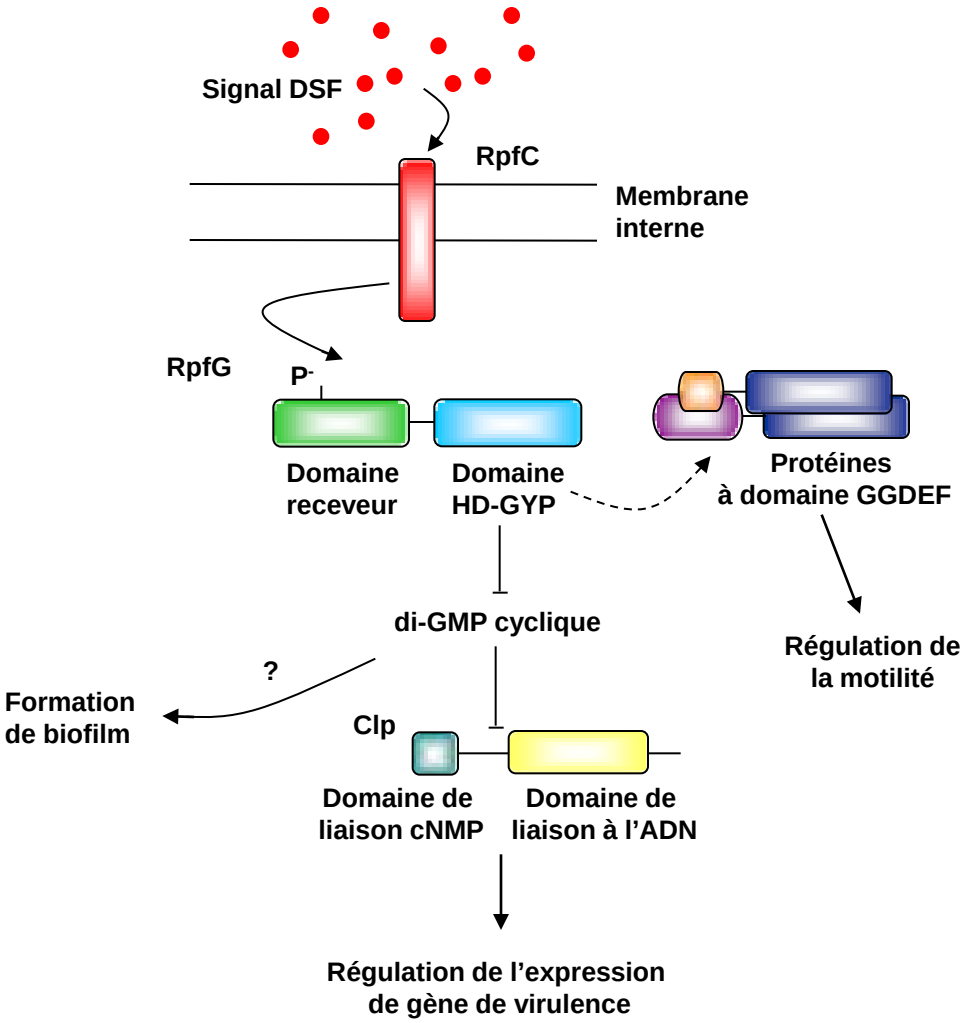


Figure 13. Représentation schématique des voies de transduction et des fonctions régulées par le signal DSF. (D'après Ryan and Dow, 2011).
Ce système de quorum sensing est impliqué dans la régulation de nombreux processus cellulaires chez *Xcc*. La perception de DSF conduit à l'autophosphorylation de RpfC et au phosphorelais du senseur vers le régulateur RpfG. Cela permet l'activation de RpfG comme une phosphodiesterase di-GMP cyclique (activité associée au domaine HD-GYP). La diminution du niveau de di-GMP cyclique permet la synthèse d'enzymes extracellulaires ainsi que d'EPS, et inhibe la formation de biofilm. La signalisation DSF induit aussi la liaison de RpfG à deux protéines possédant un domaine GGDEF. Cette interaction physique contrôle la motilité, mais n'a pas d'effet sur la synthèse d'enzymes extracellulaires, les EPS et la formation de biofilm. La régulation du niveau de di-GMP cyclique permet le contrôle de l'expression du régulon DSF via le régulateur Clp. Ce régulon comprend au moins 165 gènes incluant des gènes de virulence putatifs (He and Zhang, 2008).

réponse RpfG (Slater *et al.*, 2000) (Figure 12). Il contrôle l'expression d'au moins 165 gènes incluant des gènes de virulence putatifs (Barber *et al.*, 1997; He *et al.*, 2006; He and Zhang, 2008; Ryan *et al.*, 2007; Wang *et al.*, 2004) (Figure 13).

Des mutations dans *rpfF*, *rpfG* ou *rpfC* conduisent à une diminution de la production d'EPS, d'enzymes extracellulaires et à la formation altérée de biofilm, suggérant que le DSF est impliqué dans la régulation de l'expression de facteurs de virulence (Dow *et al.*, 2003; Jeong, 2008; Ryan *et al.*, 2007; Ryan and Dow, 2011; Slater *et al.*, 2000; Thowthampitak *et al.*, 2008; Torres *et al.*, 2007). De plus *Xcc* est capable de manipuler la fermeture des stomates au travers de sa voie de signalisation *rpf* dépendante (Gudesblat *et al.*, 2009).

Le régulateur de réponse RpfG contient un domaine phosphodiesterase HD-GYP qui est conservé chez les bactéries à Gram positif et négatif et est impliqué dans l'hydrolyse du di-GMP cyclique (Dow *et al.*, 2006; Galperin *et al.*, 2001; Ryan *et al.*, 2006) (Figure 13). De plus, RpfG interagit avec un sous-ensemble de protéines qui possèdent un domaine GGDEF (Andrade *et al.*, 2006; Ryan *et al.*, 2006). De façon intéressante, l'interaction physique entre RpfG et deux protéines à domaine GGDEF est abolie lorsque le motif est altéré. La délétion des deux domaines GGDEF conduit à une altération de la motilité de *Xcc* mais n'a pas d'effet sur la synthèse d'enzymes extracellulaires et la formation de biofilm. Ces observations élargissent notre compréhension du rôle du système Rpf/DSF dans la virulence de *Xcc* (Ryan *et al.*, 2010) (Figure 13).

De façon intéressante, chez *Stenotrophomonas maltophilia,* pathogène humain appartenant à la famille des *Xanthomonadaceae,* RpfF pourrait aussi être impliqué dans la production de lipopolysaccharides, un composant majeur de la membrane externe impliqué dans la virulence chez *Xanthomonas* spp. et offrant également à la bactérie une protection contre des environnements hostiles (Fouhy *et al.*, 2006).

II.6. Le système de sécrétion de type 3 est essentiel pour la pathogénicité bactérienne

L'objectif premier des pathogènes de plantes est d'accéder aux nutriments de leur hôte afin de se multiplier. Pauvre en eau et en nutriments, l'apoplaste végétal est considéré comme un environnement hostile pour les bactéries phytopathogènes ; de plus, il est le lieu de la sécrétion des molécules de défense des plantes. Les bactéries phytopathogènes du genre *Xanthomonas* peuvent se développer dans les espaces intercellulaires de la plante où elles

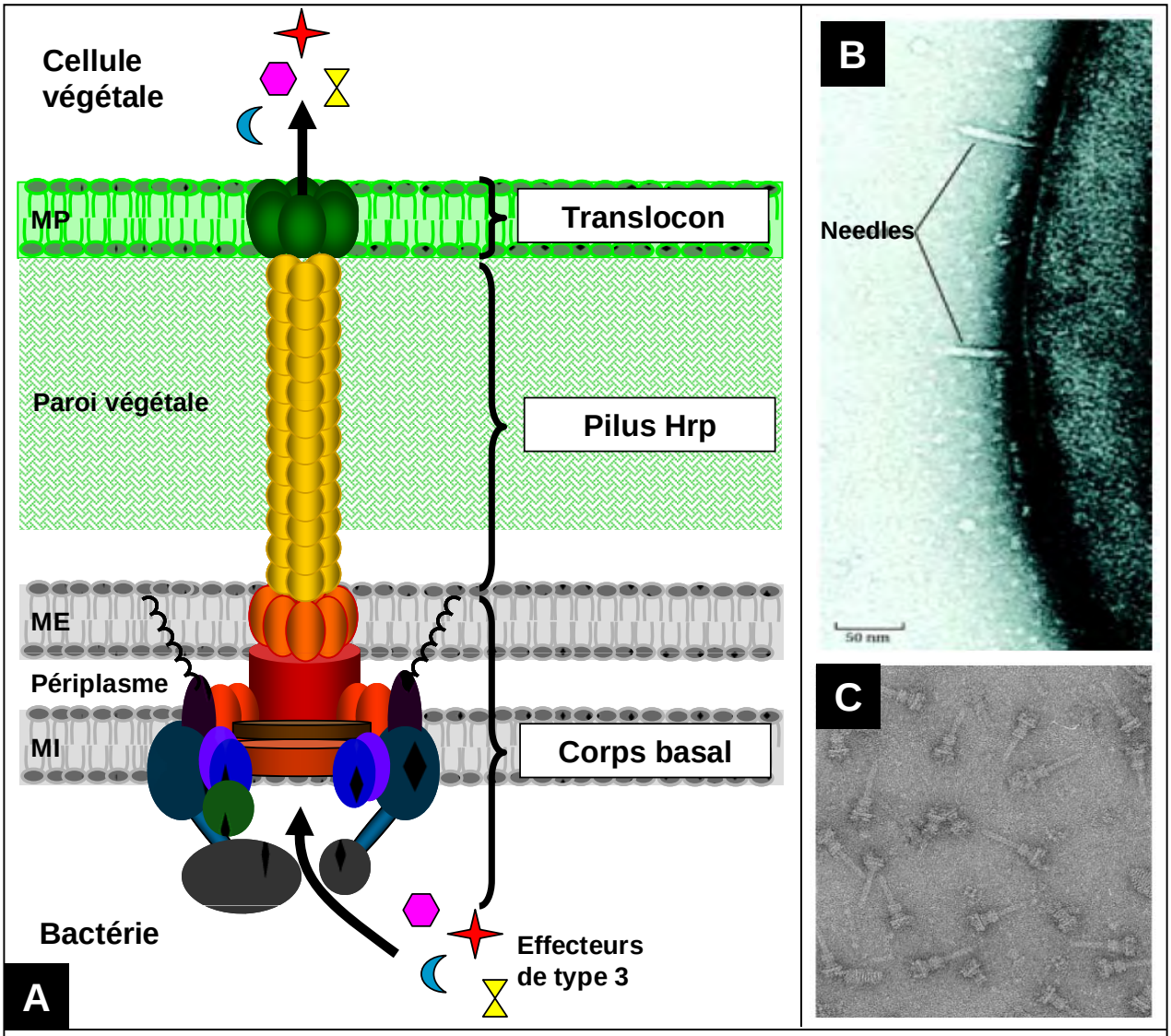


Figure 14. Le système de sécrétion de type 3 ou injectisome.
A. Le corps basal du système de sécrétion de type III est enchâssé dans la double membrane bactérienne. Il est prolongé par le « pilus Hrp » (en jaune) permettant de traverser la paroi végétale, et par le « translocon » (en vert), ancré dans la membrane plasmique de la cellule végétale. Ces structures assurent le cheminement et l'injection dans le cytoplasme végétal des substrats du système, appelés « effecteurs de type III ». MI, Membrane Interne de la bactérie ; ME, Membrane Externe de la bactérie ; MP, Membrane Plasmique de la cellule végétale.
B. Surface de *Yersinia enterocolitica* observée en microscopie électronique avec les seringues du système de sécrétion de type 3 (D'après Troisfontaine and Cornelis, 2005).
C. Seringues du système de sécrétion de type 3 isolées de *Salmonella thyphimurium* observées en microscopie électronique (D'après Galan and Wolf-Watz, 2006).

acquièrent des carbohydrates (Tang *et al.*, 2006). Les pathogènes victorieux détournent probablement les mécanismes de transport des nutriments de la plante afin d'en rediriger le flux (Patrick, 1990). Plantes et pathogènes se livrent une lutte acharnée où la plante limite l'accès du pathogène aux nutriments et initie les réponses immunes tandis que le pathogène va développer des stratégies afin d'accéder aux nutriments et supprimer l'immunité de son hôte. Les changements induits dans les cellules de l'hôte amélioreraient ainsi l'environnement des bactéries et permettraient leur prolifération (Jin *et al.*, 2003).

Peut-être une des plus remarquables adaptations de la bactérie à son environnement, le système de sécrétion de type 3 (SST3) est un des facteurs de pathogénicité clé d'un grand nombre de bactéries à Gram négatif (Ghosh, 2004). Ce système confère à la bactérie la propriété incroyable d'injecter des protéines effectrices dans le cytosol de la cellule hôte (Ghosh, 2004; Rosqvist *et al.*, 1995). Ces protéines bactériennes peuvent alors interférer avec des processus cellulaires de l'hôte.

-Structure du SST3

Le SST3 présente une structure en forme de seringue où le corps basal se prolonge par un filament protéique creux appelé « pilus » (Figure 14). Des études suggèrent que le SST3 pourrait être le fruit d'une dérive évolutive de la structure flagellaire (Macnab, 1999) bien que cette relation soit sujette à controverse (Saier, 2004). Le système de sécrétion de type 3 est nécessaire pour l'induction d'une réaction hypersensible (HR) chez les plantes résistantes permettant de confiner l'infection par un pathogène microbien et pour causer la maladie chez les plantes sensibles (Alfano and Collmer, 2004). Ce système est donc également appelé « système *hrp* » pour « hypersensitive response (HR) and pathogenicity » et les gènes codant pour ses constituants « gènes *hrp* ». Les gènes *hrp* et leur rôle dans l'interaction ont été initialement caractérisés chez le pathogène végétal *Pseudomonas syringae* pv. *phaseolicola* (Lindgren *et al.*, 1986). Depuis, les gènes *hrp* ont été identifiés chez un grand nombre de bactéries phytopathogènes, à l'exception, à ce jour, d'*Agrobacterium tumefaciens* ainsi que de *Xylella fastidiosa* et *Xalb* qui font partie toutes deux de la famille des *Xanthomonadaceae* (Bonas, 1994; Goodner *et al.*, 2001; Lindgren, 1997; Simpson *et al.*, 2000; Willis *et al.*, 1991; Wood *et al.*, 2001). Le SST3 n'est cependant pas l'apanage des bactéries pathogènes, végétales ou animales. En effet, on le retrouve également chez des bactéries symbiotes de

l'homme, des plantes et des insectes ainsi que chez des bactéries isolées de l'environnement (Dale *et al.*, 2001; Marie *et al.*, 2001; Viprey *et al.*, 1998).

Plus de 25 gènes sont nécessaires chez les pathogènes végétaux pour l'établissement d'un tel outil moléculaire. Ces gènes sont organisés en plusieurs unités transcriptionnelles et regroupés dans le cluster de gènes *hrp* (Lindgren, 1997; Van Gijsegem *et al.*, 1993). Le cluster *hrp* de *Xcc* s'étend sur 21.7 Kb et comprend 21 gènes. Le séquençage des clusters *hrp* de *R. solanacearum*, *Xcv* et de *Pseudomonas syringae* pv. *syringae* a révélé des homologies avec plusieurs composants du système de sécrétion de *Yersinia* spp. (Fenselau *et al.*, 1992; Gough *et al.*, 1992). Ces gènes *hrp* très largement conservés chez les bactéries pathogènes *Pseudomonas*, *Erwinia*, *Ralstonia*, *Xanthomonas*, *Yersinia*, *Salmonella*, et *Shigella* spp. ont été renommés *hrc* (HR and conserved) (Bogdanove *et al.*, 1996). On dénombre 9 gènes *hrc* parmi les gènes *hrp* chez *Xcc*. De façon intéressante, 8 gènes *hrc* sur 9 sont homologues aux gènes *fli/flh* codant pour le système de sécrétion des protéines du flagelle, appuyant l'hypothèse d'un mécanisme commun d'assemblage (Fenselau and Bonas, 1995; Galan and Collmer, 1999; He and Jin, 2003). Les protéines Hrc sont les protéines qui forment le corps basal du SST3 qui est une structure multiprotéique complexe se prolongeant par le pilus hrp, un canal creux extracellulaire composé majoritairement par une protéine appelée piline (Figure 14A). Cette protéine est codée chez *Pseudomonas syringae* pv. *tomato* par le gène *hrpA* (Roine *et al.*, 1997), chez *R. solanacearum* par *hrpY* (Van Gijsegem *et al.*, 2000) et chez *Xcv* par *hrpE* (Weber and Koebnik, 2005). Enfin, il est supposé que le pilus *hrp* est connecté à une structure nommée translocon, formant un pore protéique transmembranaire qui s'insérerait dans la membrane plasmique végétale et permettrait la translocation de protéines effectrices (Büttner and Bonas, 2002) (Figure 14A). Des mutations de composants du translocon conduisent à une perte totale ou à une réduction drastique de la pathogénicité (Büttner *et al.*, 2002; Meyer *et al.*, 2006; Sugio *et al.*, 2005). Chez *Xcv* et *Xoo*, HrpF serait un composant du translocon (Büttner *et al.*, 2002). Cette protéine montre des similarités avec les protéines HrpK de *P. syringae* pv. *syringae*, NolX de *Rhizobium* et PopF1/PopF2 de *R. solanacearum* (Alfano and Collmer, 2004; Meyer *et al.*, 2006).

-Régulation de l'expression des gènes du SST3

Parmi les bactéries phytopathogènes, la structure des opérons *hrp* et les systèmes de régulation de l'expression des gènes du SST3 ont permis de diviser les clusters *hrp* en deux

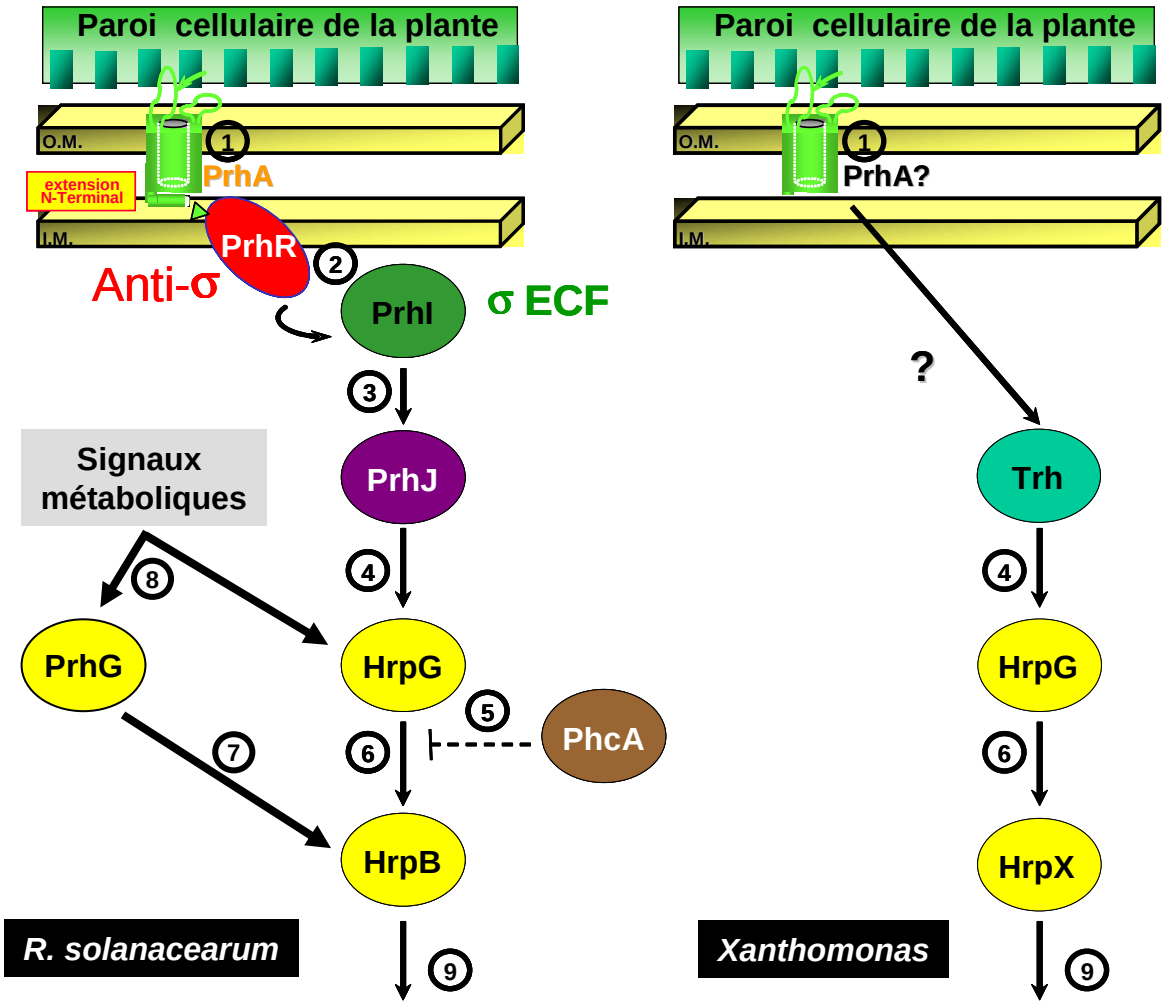


Figure 15. Modèle de régulation de l'expression des gènes du système Hrp chez les bactéries *Ralstonia solanacearum* et *Xanthomonas* (Groupe II).
1, Le TBDT PrhA est localisé dans la membrane externe et perçoit le « signal plante » pour le transférer à l'anti-sigma PrhR. 2, PrhR active PrhI. 3, PrhI active la transcription de *prhJ*. 4, PrhJ active la transcription de *hrpG* (cette transcription est activée par Trh chez *Xanthomonas oryzae)*. 5, PhcA réprime l'expression du gène *hrpG* probablement au niveau post-transcriptionnel. 6, HrpG active la transcription de *hrpB* chez *Ralstonia solanacearum* et de *hrpX* chez *Xanthomonas* spp. 7, PrhG active la transcription de *hrpB* chez *Ralstonia solanacearum.* 8, PrhG et HrpG sont régulés par des signaux métaboliques. 9, HrpX et HrpB reconnaissent les « PIP/hrpII-box » des promoteurs des gènes *hrp* et activent leur transcription. Les signaux métaboliques sont requis pour l'activation des gènes *hrp* en milieu inducteur et en présence de la plante.
(D'après Tang *et al*, 2006)

groupes : le groupe I contient les clusters *hrp* d'*Erwinia amylovora* et de *Pseudomonas syringae* et le groupe II ceux des bactéries *R. solanacearum* et *Xanthomonas* spp. (Alfano and Collmer, 1997).

Les gènes *hrp* du groupe I sont régulés par un facteur sigma de type ECF (Extracytoplasmic function), HrpL. Les facteurs sigma sont des facteurs essentiels à l'initiation de la transcription qui dirigent directement l'ARN polymérase vers des promoteurs spécifiques. L'activité transcriptionnelle de l'ARN polymérase peut être modifiée par la substitution d'un facteur sigma par un autre. Ainsi, les facteurs sigma alternatifs jouent un rôle clé dans la coordination de la transcription des gènes lors de réponses au stress ou de changements morphologiques. L'activité de HrpL est contrôlée par les régulateurs intracellulaires HrpS et HrpR ainsi que par le facteur sigma 54 (Jin *et al.*, 2003; Merighi *et al.*, 2003; Xiao and Hutcheson, 1994).

La régulation de l'expression des gènes *hrp* du groupe II est contrôlée par un activateur de la famille AraC. Chez *Xanthomonas* spp., c'est le régulateur HrpX qui contrôle de manière directe l'expression des gènes du SST3 (Figure 15) en se fixant au motif nommé « PIP-box » (Plant Inducible Promoteur-box de séquence consensus TTCGC-N_{15}-TTCGC) ou *hrpII*-box (motif dérivé de la PIP-box, de séquence consensus TTCG-N_{16}-TTCG) (Koebnik *et al.*, 2006; Wengelnik and Bonas, 1996). Chez *R. solanacearum* c'est le régulateur HrpB qui contrôle l'expression des gènes du SST3 en reconnaissant la boîte *hrpII* (Cunnac *et al.*, 2004a; Genin *et al.*, 1992). L'expression de ces régulateurs est activée par un régulateur transcriptionnel de la famille OmpR, appelé HrpG dans les deux cas (Wengelnik *et al.*, 1996; Wengelnik *et al.*, 1999) (Figure 15). Outre l'expression de *hrpB*, la protéine HrpG contrôle l'expression HrpB-indépendante d'un grand nombre de facteurs de virulence et de gènes jouant un rôle dans l'adaptation de la bactérie à son hôte végétal. Le régulon spécifique de HrpG contient notamment des gènes codant pour des enzymes de dégradation de la paroi végétale et des gènes impliqués dans la biosynthèse de phytohormones telles que l'éthylène (Valls *et al.*, 2006). Récemment, PrhG, un paralogue de HrpG (72% d'identité), a été identifié comme étant un nouveau régulateur transcriptionnel intervenant dans le contrôle des gènes du régulon du SST3 chez *R. solanacearum* (Plener *et al.*, 2010). En effet, PrhG serait spécifiquement impliqué dans la régulation du gène *hrpB* en réponse à des signaux métaboliques lorsque la bactérie croît en milieu minimum mais pas lorsque cette dernière est en contact de cellules végétales (Figure 15). Contrairement à HrpG qui est conservé parmi les *Xanthomonas* spp., PrhG semble être spécifique à *R. solanacearum*. Chez *R. solanacearum*, le gène *hrpB* est

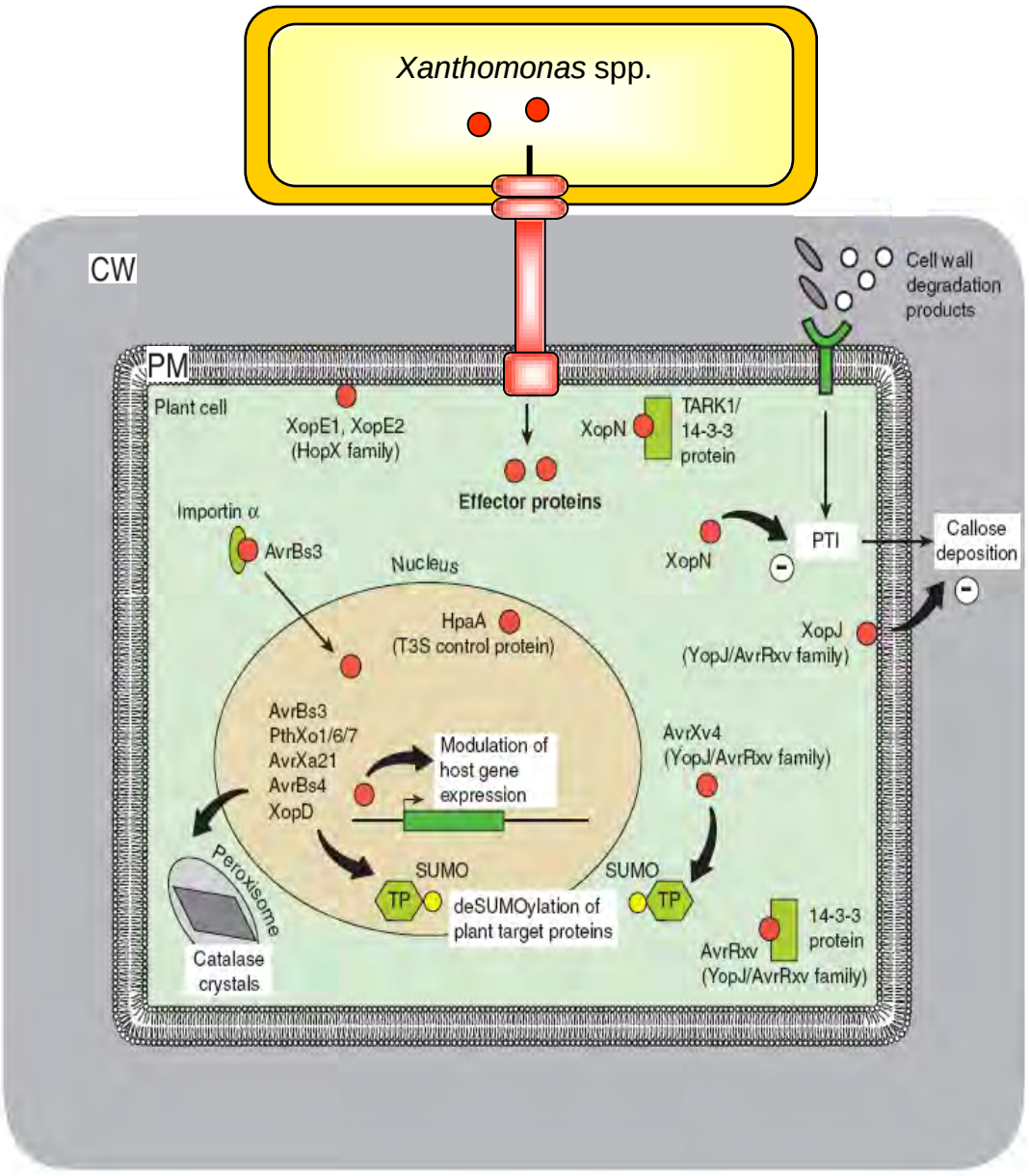


Figure 16. Mode d'action des effecteurs de type 3 caractérisés chez *Xanthomonas* spp. dans la suppression des défenses basales et/ou gène-spécifiques.

Le SST3 (système de sécrétion de type 3), qui injecte les protéines effectrices dans la cellule hôte, est essentiel pour la pathogénicité bactérienne. Les protéines effectrices de *Xanthomonas* spp. dont la localisation et/ou la fonction ainsi que les partenaires végétaux sont connus sont représentées sur cette figure. CW, cell wall ; PM, plasma membrane ; PTI, pathogen-triggered immunity (D'après Büttner and Bonas, 2010).

activé à l'issue d'une cascade de régulation très particulière mettant en jeu le transducteur TonB-dépendant (TBDT) PrhA, ancré dans la membrane externe qui perçoit un signal de la plante et le transduit, grâce à une extension N-terminale (Marenda *et al.*, 1998). Chez *Xoc*, un gène annoté *prhA* codant un TBDT conventionnel non homologue au TBDT PrhA inducteur de *R. solanacearum*, est impliqué dans l'établissement de la HR sur tabac et la virulence (Zou *et al.*, 2006). Chez *Xoo*, un régulateur transcriptionnel nommé *trh* a été identifié et est responsable de l'induction de l'expression de *hrpG* (Tsuge *et al.*, 2006). Le gène *trh*, situé loin du cluster *hrp* mais proche du gène *prhA* (gène homologue à *prhA* de *Xoc*, code pour un régulateur transcriptionnel de la famille GntR, famille qui contient des activateurs et des répresseurs. Enfin, de façon intéressante, en supplément à HrpG et HrpX, l'expression des opérons *hrpC* et *hrpE* chez *Xcc*, est contrôlée par le système de régulation à deux composants ColR/ColS. Cela suggère que plusieurs voies de transduction de signaux sont impliquées dans la régulation de l'expression des gènes *hrp* et que les opérons *hrp* peuvent être contrôlés par des voies de régulation alternatives (Zhang *et al.*, 2008).

-Les substrats du SST3

Deux grands groupes de protéines transitent par le SST3 : les protéines « helper » et les protéines effectrices. Les protéines « helper » ont pour rôle de faciliter la translocation des effecteurs en facilitant le passage au travers de la paroi végétale permettant au SST3 d'accéder à la membrane plasmique végétale. Les harpines, des protéines riches en glycine et pauvres en cystéine, inductrices de nécroses de type HR, sont une classe majeure des protéines « helper » (Alfano and Collmer, 2004).

Les protéines effectrices ou effecteurs vont interférer avec certaines fonctions de la cellule eucaryote afin d'altérer l'immunité de la plante telles que la réponse basale de l'hôte, les voies de signalisation hormonales de la plante, la résistance non hôte ou la réponse hypersensible (pour revue (Grant *et al.*, 2006; Zhou and Chai, 2008)) (Figure 16). Pour interférer avec ces fonctions certains effecteurs possèdent des domaines ayant des caractéristiques eucaryotes. C'est le cas par exemple de la protéine effectrice XopD de *Xcv*. Cet effecteur possède une région très homologue au domaine catalytique C-terminal de la famille protéique Ulp1 (Ubiquitin-like protease ; (Li and Hochstrasser, 1999)). Ulp1 est une cystéine protéase qui catalyse deux réactions critiques de la voie SUMO/Smt3. L'ubiquitination est un processus eucaryote très conservé contrôlant la dégradation des protéines. XopD possède une activité de

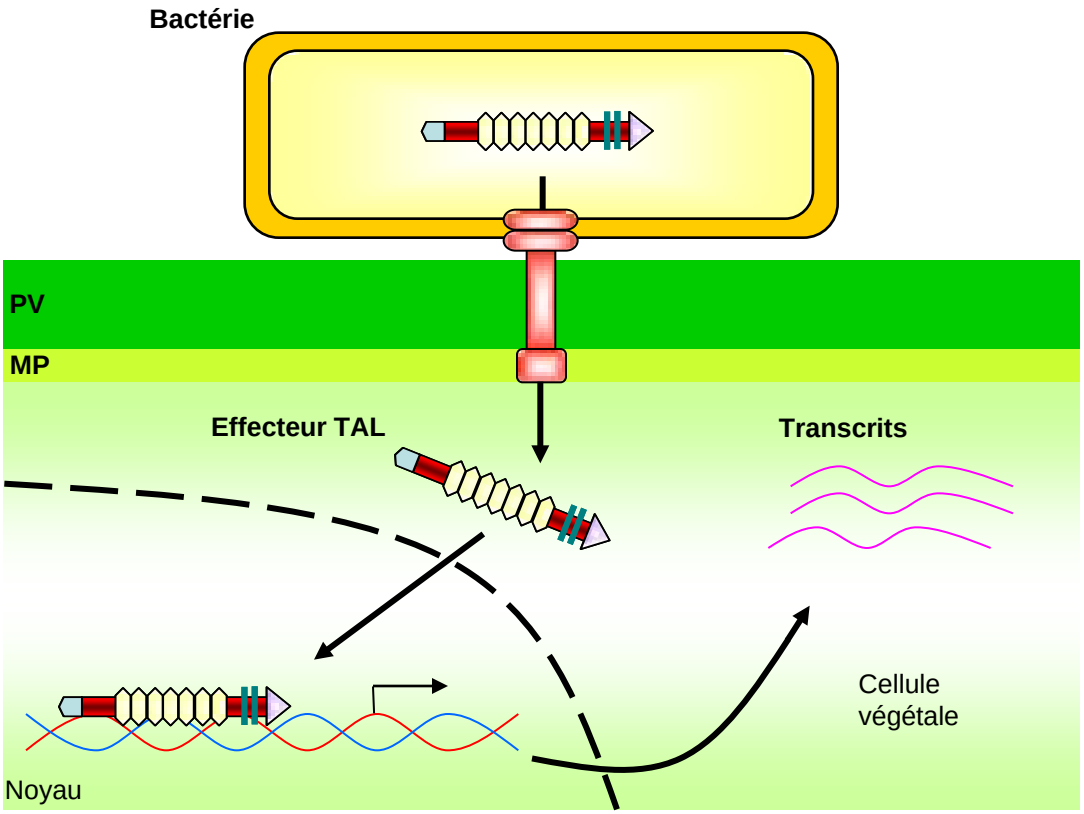


Figure 17. Modèle de fonctionnement des effecteurs TALs.
Xanthomonas injecte un coktail de protéines effectrices dans les cellules végétales par l'intermédiaire du système de sécrétion de type 3 (en rouge). Les effecteurs TAL (Transcription activator-like) vont dans le noyau de la cellule végétale où ils induisent l'expression de gènes cibles. PV, paroi végétale; MP, membrane plasmique.

déSUMOylation qui semble être impliquée dans les réponses de mort cellulaire et de défense basale (Canonne *et al.*, 2010; Hotson *et al.*, 2003; Kim *et al.*, 2008) (Figure 16). Cependant, comme pour beaucoup d'effecteurs bactériens connus, les cibles végétales directes de XopD restent inconnues.

Les effecteurs contribueraient également à entraîner la fuite d'eau et de nutriments dans l'apoplaste par les cellules végétales. Cependant, les mécanismes qu'utilisent les pathogènes afin d'altérer la physiologie de l'hôte, notamment les transports de sucres, ne sont pas très bien compris. Récemment, des travaux supportent un nouveau modèle où les effecteurs de type 3 sont impliqués dans la manipulation des ressources nutritionnelles des hôtes végétaux au profit des pathogènes en plus de l'inhibition des réponses immunes (Chen *et al.*, 2010).

L'activité de certaines protéines effectrices de type 3 aboutit à une modification de l'expression de gènes végétaux. Ces protéines sont regroupées au sein de la famille des effecteurs TALs (Transcription-activator like). Les TALs représentent la plus grande famille d'effecteurs et agissent comme des activateurs transcriptionnels au sein du noyau de la cellule végétale (Kay *et al.*, 2007; Romer *et al.*, 2007; White *et al.*, 2009) (Figure 17). Les effecteurs TALs sont principalement trouvés chez *Xanthomonas* spp. mais des homologues sont également présents chez *R. solanacearum* (Cunnac *et al.*, 2004b; Heuer *et al.*, 2007). La caractéristique la plus frappante et intéressante des TALs est leur domaine central qui est composé d'unités répétées de 30 à 42 aa de long avec le plus souvent des unités de 34 aa (Boch and Bonas, 2010) (Figure 18A), la dernière répétition étant incomplète (demi-répétition). Le nombre d'unités répétées est généralement compris entre 15.5 et 19.5 (Boch and Bonas, 2010). Les variations inter-répétitions se limitent à deux résidus situés en position 12 et 13 (Figure 18A). Les domaines répétés des TALs constituent un nouveau domaine de liaison à l'ADN ne montrant aucune similarité avec des éléments connus de liaison à l'ADN (Kay *et al.*, 2007). La liaison à l'ADN s'effectuerait par l'intermédiaire du domaine central constitué des unités répétées (Boch *et al.*, 2009; Romer *et al.*, 2009). AvrBs3, un des membres les plus étudiés de cette famille d'effecteurs TALs, induit la transcription de plus de 20 gènes chez le poivron en se fixant sur un élément conservé (*UPA* box, upregulated by AvrBs3) du promoteur des gènes cibles (Kay *et al.*, 2007). Les facteurs de l'hôte qui influencent la liaison des effecteurs TALs à l'ADN n'ont pas encore été identifiés. La découverte de cette *UPA* box a permis d'aboutir à un modèle permettant d'expliquer la spécificité de reconnaissance de l'ADN (Boch and Bonas, 2010; Bogdanove *et al.*, 2010). Ce modèle repose sur un principe qu'on peut résumer sous la forme : une répétition pour une base d'ADN. La spécificité de

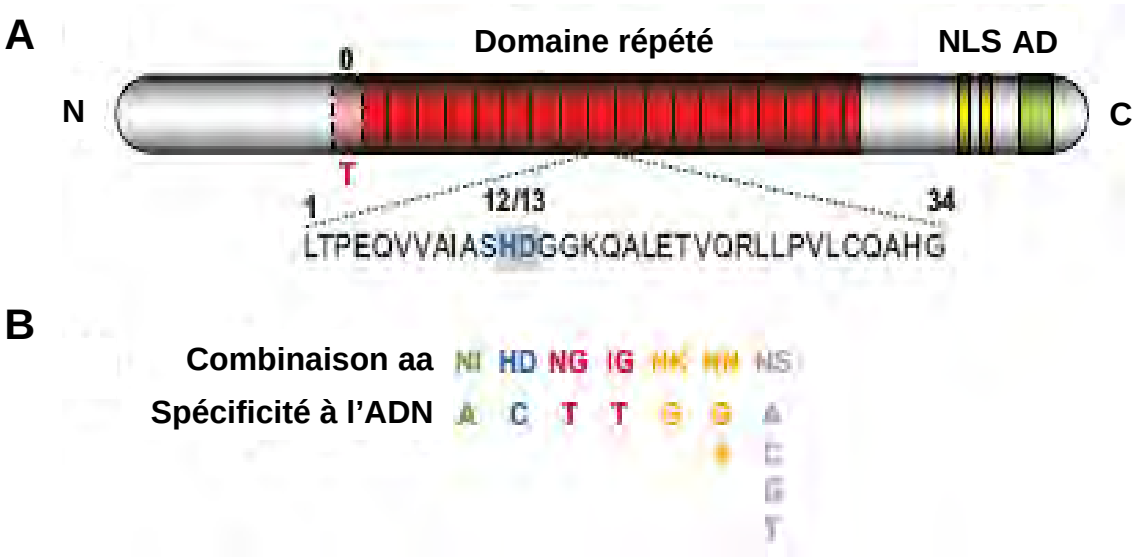


Figure 18. Code des effecteurs TALs.

A. Les TALs contiennent des domaines de localisation nucléaire (NLS, nuclear localization signal, en jaune) et un domaine d'activation (AD, activation domain, en vert) leur conférant une fonction d'activateurs transcriptionnels. Un domaine central répété (en rouge) confère la spécificité de liaison à l'ADN. Chaque répétition de 34 acides aminés (aa), dans le cas présent, se lie à une base, et la spécificité de liaison est déterminée par les aa en position 12 et 13. Une répétition est représentée sur le modèle. Les numéros représentent la position des aa à l'intérieur de la répétition. La répétition 0 est indiquée et possède une spécificité pour la base T de l'ADN.

B. Chaque combinaison des aa 12 et 13 a une spécificité pour une ou plusieurs bases d'ADN.

(D'après Scholze and Boch, 2011).

chaque répétition repose sur les deux aa variables en position 12 et 13 (Boch and Bonas, 2010; Bogdanove *et al.*, 2010). En effet, certaines combinaisons seront spécifiques d'une base d'ADN alors que d'autres en reconnaitront plusieurs (Boch and Bonas, 2010; Bogdanove *et al.*, 2010) (Figure 18B). En réponse à ce mécanisme d'action moléculaire des effecteurs TALs, les plantes ont mis en place une stratégie de reconnaissance évoluée en utilisant les gènes de résistance comme appât moléculaire. En effet, l'activation du promoteur par certains TALs conduit à l'induction de l'expression du gène de résistance et ainsi à la mort cellulaire (Gu *et al.*, 2005; Romer *et al.*, 2007).

Les effecteurs ont été nommés Hop pour «Hrp outer protein» chez *Pseudomonas* (Alfano and Collmer, 1997), Xop pour «*Xanthomonas* outer protein» chez *Xanthomonas* (Noël *et al.*, 2001) ou Pop pour «*Pseudomonas* outer protein» chez *Ralstonia* (qui faisait partie du groupe *Pseudomonas* à l'époque de cette identification) (Arlat *et al.*, 1994). Depuis 2002, l'identification de protéines effectrices a été grandement favorisée par le séquençage des génomes de nombreuses bactéries phytopathogènes dont 16 souches de *Xanthomonas* (The *Xanthomonas* ressource, www.xanthomonas.org). Les gènes candidats ressortent grâce à des analyses de génomique comparative basées sur des homologies de séquences d'effecteurs de type 3 connus ou la présence de motifs protéiques eucaryotes caractéristiques (Büttner *et al.*, 2003). De plus, des approches de bioinformatique peuvent être utilisées pour le criblage d'effecteurs de type 3. En effet, la recherche d'éléments conservés dans les promoteurs de gènes tel que la PIP-box et la composition spécifique en acides aminés dans les séquences N-terminales des protéines sont fréquemment utilisées comme critères de recherche de nouvelles protéines effectrices (Arnold *et al.*, 2009; Fouts *et al.*, 2002; Furutani *et al.*, 2009; Guttman *et al.*, 2002; Petnicki-Ocwieja *et al.*, 2002; Samudrala *et al.*, 2009; Tsuge *et al.*, 2005; Zwiesler-Vollick *et al.*, 2002).

Des études de génomique comparative ont montré que les répertoires de gènes de virulence ont une grande variabilité au sein des souches séquencées, suggérant un rôle dans la spécificité d'hôte (da Silva *et al.*, 2002; Guidot *et al.*, 2007; He *et al.*, 2007). Les répertoires d'effecteurs sont variables d'une espèce bactérienne à une autre et aussi au sein d'un même genre. Une trentaine d'effecteurs ont par exemple été identifiés chez les bactéries du genre *Xanthomonas* (Kay and Bonas, 2009) (Tableau 3) alors que la bactérie *R. solanacearum* en possède 70 à 80 (Poueymiro and Genin, 2009). Une étude pionnière conduite par Sarkar et ses

Tableau 3. Effecteurs de type 3 de *Xanthomonas campestris* pv. *campestris* str. 8004. (D'après Endrick Guy, M2R, 2009; www.xanthomonas.org/t3e.html)

Gène	Nom	HrpG/X[a]	GC%	Fonction prédite	Homologues chez[b]					
					Xcv	*Xac*	*Xoo*	*Xoc*	*Rs*	*Pst*
XC_1553	*xopAC*	oui	56	vascular protein; LRRs/Fic; adenylylation protein?	-	-	-	-	-	-
XC_2081	*avrBs1*	non	44	tyrosine phosphatase	+	-	-	-	-	+
XC_2082	*xopH*	nd	42	xopH/putative tyrosine phophatase	+	-	-	-	-	+
XC_0052	*avrBs2*	oui	68	Glycero-P-diester P-diesterase	+	+	+	+	-	-
XC_2602	*xopE2*	oui	62	avrRxv/YopJ/HopX2	+	+	-	-	+	+
XC_3802	*xopJ*	oui	60	avrRxv/YopJ/HopX2	+	-	-	-	-	-
XC_0967	*xopG*	non	51	HopPtoH	+	-	+	-	+	+
XC_1213	*xopD*	oui	59	cysteine protease C48/Nedd8 protease?	+	-	-	-	-	(+)
XC_3915-6	*xopAL2*	nd	47	HopK1/avrRps4/xopO	+	+	+	+	+	-
XC_4318	*avrXccA1*	nd	69	avrXccA/ maybe not a T3E	-	-	-	-	-	-
XC_1716	*avrXccA2*	non	62	avrXccA/ maybe not a T3E	-	-	-	-	-	-
XC_2004	*xopAH*	oui	47	homolog avrC/avrB1	-	-	-	-	-	-
XC_3018	*hpaA*	oui	68	HpaA ; NLS; Type 3 secretion control protein/maybe not a T3E	+	+	-	-	-	-
XC_4273	*XopL*	nd	57	protéine LRR	+	+	+	+	-	-
XC_0526	*xopA*	oui	69	Harpin protein? Maybe not a T3E	+	+	+	+	-	-
XC_3024	*xopF*	oui	64	?	+	(+)	+	+	-	-
XC_0241	*xopN*	oui	62	ARM/HEAT repeat	+	+	+	+	-	+
XC_2995	*xopAL1*	oui	49	HopPmaB/avrPphE	+	-	-	-	+	(+)
XC_2994	*xopP*	oui	57	?	+	+	+	+	+	-
XC_3177	*xopQ*	oui	63	Inosine-uridine nucleoside N-ribohydrolase	+	+	+	+	+	+
XC_3160	*xopAM*	non	65	transducer protein car/HopR1	-	-	-	-	+	+
XC_0541	*xopX1*	oui	63	?	+	+	+	+	+	(+)
XC_0542	*xopX2*	oui	61	?	+	+	+	+	+	-
XC_1210	*xopK*	non	63	?	+	+	+	+	-	-
XC_0268	*xopR*	oui	64	?	+	+	+	+	-	-
XC_3023	*hrpW*	oui	64	Pectate lyase/maybe not a T3E	-	+	-	-	+	+
XC_2210	*xopZ*		65	?	+	+	+	+	+	+
XC_0563	*xopAG*		57	?	-	-	-	-	-	+
?	*xopM*			?	+	?	?	?	?	?
?	*xopS*			?	+	?	?	?	?	?

[a] : Régulation par HrpG/HrpX

[b] : Les homologues chez les autres bactéries des effecteurs de *Xcc* ont été identifiés en utilisant l'algorithme BLAST.

Xcv : *Xanthomonas campestris* pv *vesicatoria* ; *Xac* : *Xanthomonas axonopodis* pv *citri* ; *Xoo* : *Xanthomonas oryzae* pv *oryzae* ; *Xoc* : *Xanthomonas oryzae* pv *oryzicola* ; *Rs* : *Ralstonia solanacearum* ; *Pst* : *Pseudomonas syringuae* pv. *tomato*.

collaborateurs (Sarkar *et al.*, 2006) a déterminé la distribution d'un large éventail de gènes de virulence chez 91 souches de *Pseudomonas syringae*, isolées de diverses plantes hôtes. Ils ont alors postulé que la distribution de ces gènes associés à la virulence pourrait indiquer leurs rôles possibles dans la spécificité d'hôte. En effet, les gènes très conservés parmi toutes les souches ne jouent probablement pas de rôle majeur dans la spécificité d'hôte. A l'opposé, les gènes distribués de façon hétérogène parmi ces mêmes souches sont de bons candidats pour expliquer la spécificité d'hôte. Parmi les déterminants de pathogénicité, les effecteurs de type 3 présentent une distribution hétérogène entre les souches (Guidot *et al.*, 2007; Sarkar *et al.*, 2006). La contribution d'un effecteur dans le pouvoir pathogène bactérien peut varier selon l'hôte étudié. La présence d'un effecteur peut donc engendrer l'élargissement de la gamme d'hôte d'un pathogène (Alavi *et al.*, 2008; Castaneda *et al.*, 2005). L'étude des répertoires des effecteurs de type 3 en association avec les pathovars d'une espèce végétale pourrait expliquer les spécificités d'hôtes et de tissus (Hajri *et al.*, 2009).

II.7. Les systèmes de sécrétion de type 2 et les enzymes de dégradation extracellulaire

Afin de faciliter sa progression dans les tissus végétaux, l'agent pathogène dégrade la paroi des cellules végétales. Cette dégradation va non seulement permettre une dissémination des agents pathogènes dans les tissus de l'hôte mais aussi permettre de libérer des nutriments essentiels à la croissance des bactéries. En effet, la paroi végétale constitue un véritable réservoir pour la plupart des microorganismes présents dans l'environnement des plantes. *Xcc* possède un répertoire important d'enzymes de dégradation de la paroi végétale impliquées dans la dégradation des différents composés de la paroi (pectines, celluloses, hémicelluloses et glycoprotéines). De plus, de nombreuses enzymes sont d'importants facteurs de virulence (Cf. p33, paragraphe IV.2). Ce « CAZyme » (enzymes répertoriées dans la base de donnée CAZy, http://www.cazy.org/) important facilite sa progression dans les tissus et lui confère un avantage non négligeable dans la compétition entre les microorganismes pour la nutrition. Généralement, les mutants individuels des enzymes de dégradation de la paroi végétale ne sont pas ou partiellement affectés au niveau de la pathogénie alors que les mutants multiples sont affectés de manière plus forte dans la virulence suggérant des redondances fonctionnelles parmi les enzymes de dégradation de la paroi végétale (Jha *et al.*, 2007; Rajeshwari *et al.*, 2005).

Les enzymes de dégradation de la paroi végétale sont sécrétées par le système de sécrétion de type 2 (SST2) dans la plupart des cas. Ce système de sécrétion permet le transport de protéines du périplasme bactérien vers le milieu extracellulaire. Le SST2 permet également la sécrétion de toxines. L'appareil du système de type 2 comprend entre 12 et 15 composants dont la plupart sont associés à la membrane interne bactérienne (Sandkvist, 2001a, b). Le SST2 est absent chez le pathogène végétal *A. tumefaciens* (Goodner *et al.*, 2001). Le SST2 joue un rôle important dans le pouvoir pathogène des bactéries *R. solanacearum*, *Erwinia* spp. et *Xanthomonas* spp. (Cianciotto, 2005; Jha *et al.*, 2005). Les analyses des séquences des génomes montrent les *Xanthomonas* spp. possèdent un (*Xoo*, *Xoc*) ou deux (*Xcv*, *Xcc*, *Xac*) SST2 qui sont codés par les clusters de gènes *xcs* et *xps*. Les mutations qui affectent la virulence bactérienne n'affectent que des gènes présents au sein du cluster *xps* (Dow *et al.*, 1987; Lu *et al.*, 2008; Qian *et al.*, 2005; Rajeshwari *et al.*, 2005; Ray *et al.*, 2000; Sun, 2005). De plus, l'expression de plusieurs gènes codant pour des enzymes sécrétées par le SST2 est régulée par les régulateurs du SST3, HrpG et HrpX, suggérant une balance fonctionnelle entre les deux systèmes de sécrétion (Furutani *et al.*, 2004; Wang *et al.*, 2008; Yamazaki *et al.*, 2008).

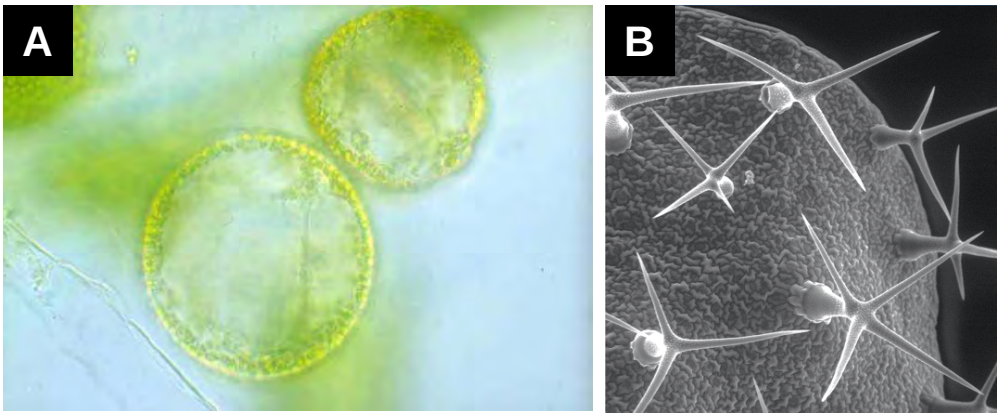


Figure 19. L'importance des parois végétales. La paroi détermine la forme des cellules végétales.
A. Protoplastes de feuille de poireau (source : snv.jussieu.fr).
B. Trichomes d'*Arabidopsis thaliana* (source : www.pnas.org, cover image, de la Paz Sanchez and Gutierez, 2009).

LA GUERRE AUX PORTES DE LA PLANTE

I. L'importance des parois végétales : plus que la taille et la forme de la cellule végétale

Le rôle le plus évident de la paroi végétale est de déterminer la taille et la forme des cellules. Sans sa paroi, la cellule végétale subirait la pression de turgescence conduisant à une forme sphérique comme dans le cas des protoplastes, cellules végétales sans paroi (Taiz, 1984) (Figure 19). Bien qu'il ait été montré que l'expansion de la paroi peut influencer la morphogenèse (Fleming *et al.*, 1997), il reste encore à déterminer si cette expansion est à l'origine ou à l'aboutissement du processus (Fleming, 2006). Toutefois, la paroi est plus que cet exosquelette entourant la cellule et doit être considérée comme un organite à part entière participant activement à la physiologie de la cellule (Robinson, 1991). Par exemple, lors d'une attaque d'un pathogène, la plante induit des dépôts de callose au niveau des sites de pénétration du pathogène, accumule des composés phénoliques et plusieurs toxines dans la paroi et synthétise des polymères « lignin-like » pour renforcer celle-ci (Huckelhoven, 2007). Ainsi, tels les remparts d'une forteresse, la paroi végétale est une structure défensive importante que les pathogènes doivent franchir avant de se confronter aux défenses intracellulaires. De plus, la paroi végétale constitue la source de carbohydrate la plus importante sur la biosphère. Les pathogènes sécrètent un grand nombre d'enzymes de dégradation de la paroi végétale libérant ainsi de petites sous unités qui serviront de source de nutriments mais qui joueront également un rôle dans l'activation de réponses de défenses chez la plante (Garcia-Brugger *et al.*, 2006). La paroi végétale joue donc un rôle d'interface active entre le cytoplasme et l'environnement de la plante. Vue du pathogène, la paroi apparaît comme une réserve de nutriments mais aussi comme la première ligne de défense à franchir en vue de la colonisation de la plante. La bataille de la paroi prend alors tout son sens.

II. Les différents types de parois végétales

L'étude de la biosynthèse et de la structure de la paroi cellulaire a permis de caractériser trois couches qui se forment successivement au sein de celle-ci : la lamelle moyenne, la paroi primaire et la paroi secondaire. Toutes les cellules comportent une paroi primaire. En

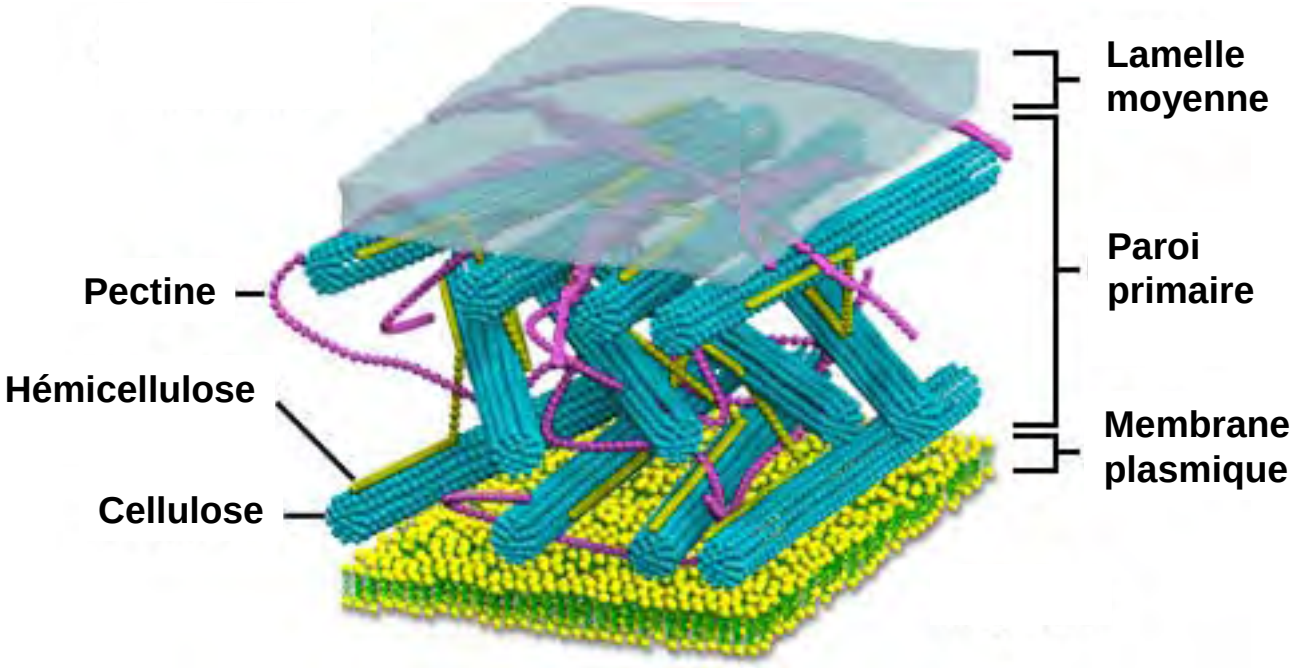


Figure 20. Modèle simplifié de la paroi primaire végétale. Elle consiste en un réseau complexe de fibres de cellulose, de pectine et d'hémicellulose. La zone de contact entre deux cellules est appelée lamelle moyenne.
http://micro.magnet.fsu.edu/cells/plants/cellwall.html

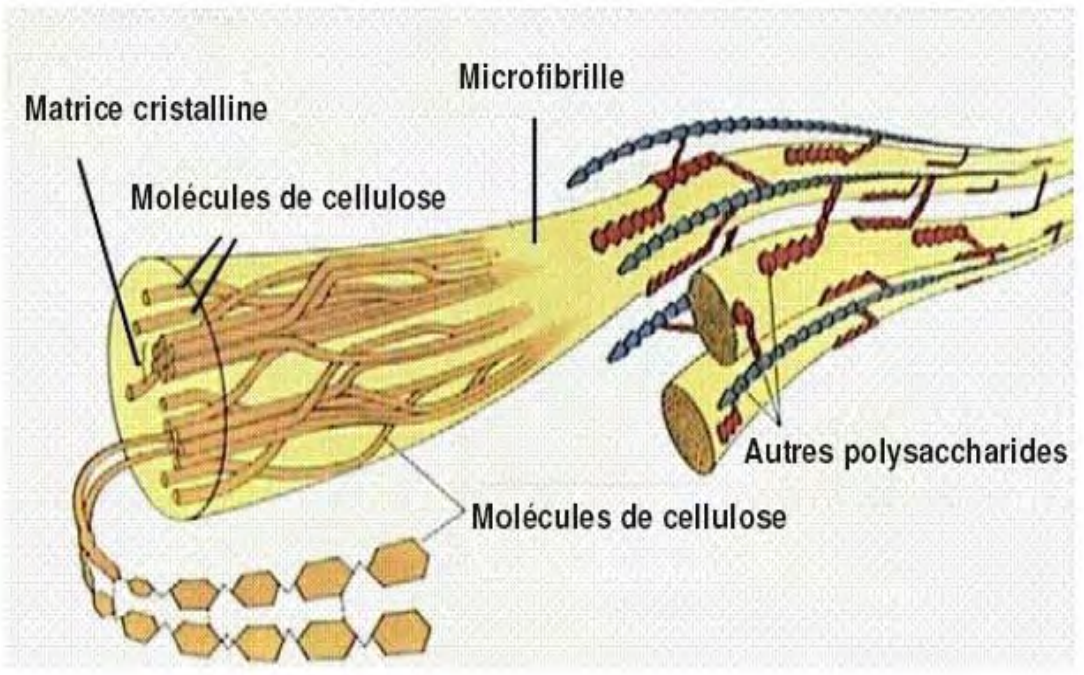


Figure 21. Microfibrilles de cellulose. Une microfibrille est composée de macromolécules de celluloses associées s'organisant en agrégats cristallins. Ces microfibrilles s'organisent autour des cellules végétales et sont réunies entre elles par une matrice constituée d'hémicelluloses et de pectines.
http://www.emc.maricopa.edu/faculty/farabee/biobk/BioBookCHEM2.html

revanche, on ne retrouvera la paroi secondaire que dans certains types cellulaires : les cellules conductrices de sève (vaisseaux du xylème) et différents tissus de soutien (sclérenchyme) ou de protection (liège). Elle est déposée par le processus d'expansion.

La paroi primaire est de type I chez les non graminées et de type II chez les graminées. Les parois de type II diffèrent des parois de type I par leur contenu élevé en glucuronoarabinoxylane (GAX), leur faible contenu en pectines et par la présence de composés aromatiques (hydroxycinnamates, ferulate et p-coumarate) (Carpita, 1996).

Tout comme la paroi primaire, la paroi secondaire est composée de polysaccharides variés et de protéines. Cependant, elle est enrichie en composés phénoliques : la lignine qui rigidifie la paroi ainsi que la cutine et la subérine qui l'imperméabilisent. Les parois secondaires forment le bois et de ce fait, leur étude présente un intérêt économique en vue d'usages industriels. La composition des parois végétales peut significativement varier d'un type cellulaire à un autre, d'une espèce à une autre ou encore entre des accessions au sein d'une espèce (Hazen *et al.*, 2003).

III. Composition et organisation de la paroi primaire

La paroi primaire est composée d'approximativement 10% de protéines et 90% de polysaccharides qui peuvent être divisés en trois groupes : la cellulose, les hémicelluloses et les pectines (McNeil *et al.*, 1984) (Figure 20). Il faut noter que la paroi contient également beaucoup d'ions (notamment du calcium) et des enzymes.

III.1. La cellulose

La cellulose est le constituant principal de la paroi des cellules végétales et la matière organique la plus abondante sur Terre. Il s'agit d'un polymère comportant 500 à 14000 résidus de β-(1,4)-D- glucopyranose (Somerville, 2006) (Figure 21). Les macromolécules de cellulose associées s'organisent en agrégats cristallins formant les microfibrilles (Bacic, 1988). Ces microfibrilles s'organisent autour des cellules végétales et sont réunies entre elles par une matrice constituée d'hémicelluloses et de pectines (Figures 20 et 21) rendant la structure finale encore plus résistante (Carpita and Gibeaut, 1993; Cosgrove, 2005).

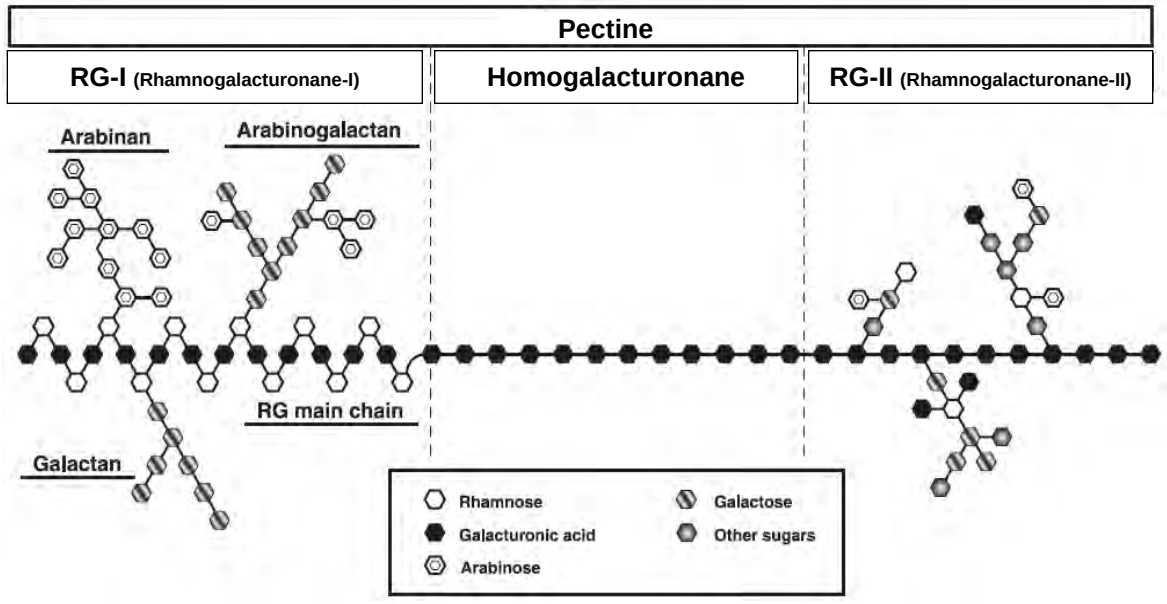


Figure 22. Structure de la pectine, de ses chaînes linéaires et de ses ramifications (D'après Ochiai *et al*., 2007 ; Blanvillain, 2007).

III.2. Les pectines

Les pectines sont des polymères de polysaccharides acides formant une chaîne linéaire ramifiée de chaînes latérales non monotones et non linéaires. Elles ont la caractéristique d'être extraites de la paroi par de l'eau chaude, des acides dilués ou des chélateurs de calcium. Il existe trois types de pectines : l'homogalacturonane (HG), le rhamnogalacturonane I (RG-I) et le rhamnogalacturonane II (RG-II) (Ridley *et al.*, 2001) (Figure 22).

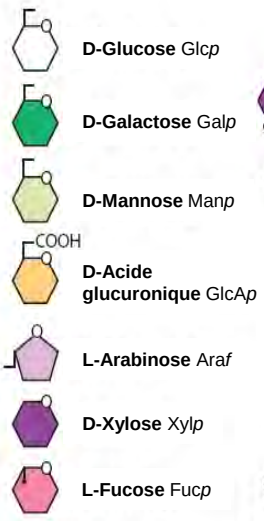
Le HG est une chaîne linéaire d'acide galacturoniques (GA) liés en α-1,4 sein desquels les groupes carboxyl peuvent être méthyl-estérifiés en C-6 et être *O*-acétylés en *O*-2 et *O*-3. Deux chaînes d'homogalacturonane peuvent interagir par des ponts Ca^{2+}.

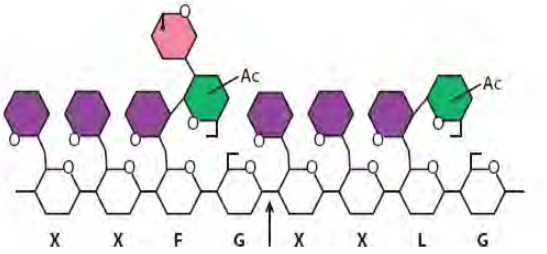
La chaîne principale du RG-I est formée d'une alternance de résidus GA liés en α-1,4 et rhamnoses liés en α-1,2 (Darvill *et al.*, 1985). Entre 20 et 80% des résidus rhamnosyl peuvent être substitués en *O*-4 avec des sucres neutres. Ces substitutions peuvent être des résidus uniques de D-galactopyranose en β-(1,4) mais aussi polymériques tels que l'arabinogalactane ou l'arabinane (50 résidus ou plus) (Lerouge *et al.*, 1993; O'Neill *et al.*, 1990). L'abondance et l'identité des chaînes latérales présentes dans le RG-I varient considérablement entre espèces (Vincken *et al.*, 2003).

Le RG-II est composé d'une colonne vertébrale d'HG substituée par des chaînes latérales diverses et variées, formées de résidus rhamnose, arabinose, galactose et*c* (Mazeau and Perez, 1998; Stevenson *et al.*, 1986). La structure du RG-II est très conservée parmi les plantes, même si certaines différences mineures ont parfois été identifiées (O'Neill *et al.*, 2004).

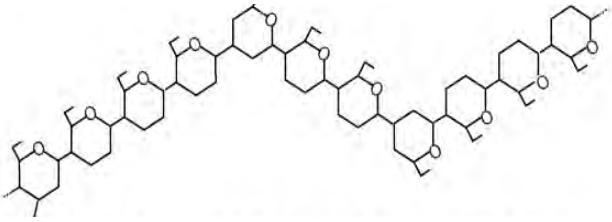
III.3. Les hémicelluloses

Les hémicelluloses sont des polysaccharides très variés formés d'une chaîne principale avec des liaisons β-(1,4), ramifiée par de courtes chaînes latérales qui empêchent la molécule de former des microfibrilles. On trouve parmi les hémicelluloses : les xyloglucanes, les xylanes, les mannanes et glucomannanes et les β-(1,3 ;1,4)-glucanes (Figure 23). Ces types d'hémicelluloses sont présents dans les parois cellulaires de toutes les plantes terrestres, excepté pour les β-(1,3 ; 1,4)-glucanes qu'on ne retrouve que chez les Poales et quelques autres groupes (Cosgrove, 2005). Le rôle biologique le plus important des hémicelluloses est leur contribution à la rigidité et à la force de la paroi cellulaire en interagissant avec la cellulose et, dans certaines parois, avec la lignine. Ce réseau est activement remodelé par les

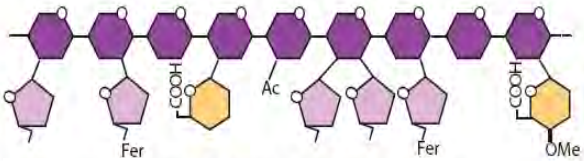




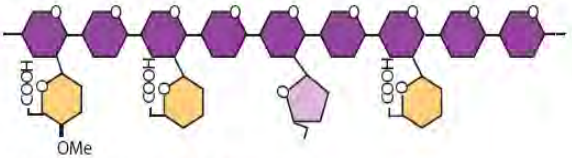
Xyloglucane constitué d'une colonne vertébrale [β-D-Glc*p*-(1,4)]$_n$ substitué de diverses chaînes latérales. La flèche indique le site de clivage classique des β-glucanases.



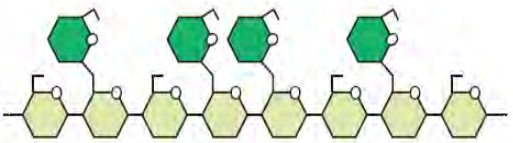
β-glucane constitué de [β-D-Glc*p*-(1,4)]$_n$-β-D-Glc*p*-(1,3)-[β-D-Glc*p*-(1,4)]$_m$, où n et m = 3 ou 4. Ne se retrouve que chez les Poales.



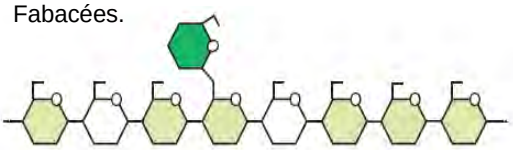
Glucuronoarabinoxylane (GAX), retrouvé classiquement chez les commelinidées (clade de monocotylédone).



Glucuronoxylane, structure typique des dicotylédones.



Galactomannane, polysaccharide qu'on retrouve dans les graines de Fabacées.



Galactoglucomannane, composant majoritaire de la paroi secondaire des gymnospermes.

Figure 23. Illustration schématique des différents types d'hémicelluloses trouvés dans les parois végétales. Les lettres sous la molécule de xyloglucane représentent les symboles utilisés pour désigner les chaînes latérales. Les hémicellulose rencontrées varient selon l'espèce végétale et le type de tissu. "Fer" représente l'estérification avec de l'acide férulique, caractéristique des xylanes chez le clade de monocotylédone, les commelinidées (D'après Scheller and Ulvskov, 2010).

enzymes végétales (xyloglucan endotransglycosylase par exemple) pendant l'expansion cellulaire (Vissenberg *et al.*, 2000).

Les mannanes et glucomannanes constituent l'hémicellulose majoritaire chez les Charophytes (algues vertes) (Popper and Fry, 2003). Le mannane peut avoir un squelette de D-mannose en β-(1,4) ou un squelette de D-mannose et de D-glucose, également unis en β-(1,4). Ils sont alors appelés glucomannanes. Ces deux types de mannanes peuvent être substitués en α-(1,6) par des résidus de D-galactose et on parle alors de galactomannanes ou galactoglucomannanes (Figure 23). Chez les gymnospermes, les galactoglucomannanes sont les composants majoritaires de la paroi secondaire (Ebringerova *et al.*, 2005). Le rôle des mannanes comme composé de réserve dans la graine a été très étudié (Buckeridge *et al.*, 2000; Hoch, 2007).

Cependant, les classes les plus étudiées correspondent au xyloglucane et au glucuronoarabinoxylane (GAX) plus couramment appelé xylane (Figure 23). Le xyloglucane est un polymère de glucoses ramifié par de courtes chaînes de xylose, galactose et fucose. Les xylanes sont composés d'un squelette de résidus de D-xylose unis en β-(1,4), qui peut être substitué par des résidus de D-glucuronate en α-(1,2) (glucuronoxylane), par des résidus de L-arabinose en α-(1,2) ou α-(1,3) (arabinoxylanes) ou par les deux (glucuronoarabinoxylane, GAX; Figure 23). Des études récentes suggèrent que les xylanes présents chez Arabidopsis, du moins dans les hampes florales, sont des xylanes ou des glucuronoxylanes, le L-arabinose n'ayant pas été détecté (Pena *et al.*, 2007). Contrairement aux xyloglucanes, les xylanes n'ont pas de structures répétées et il y a beaucoup de variations dans la structure qui ne sont pas bien connues.

Il est intéressant de noter des différences entre la paroi des monocotylédones et celle des dicotylédones au niveau de leur teneur en xylanes et xyloglucanes. Classiquement, la teneur en xyloglucanes est très faible dans la paroi primaire des monocotylédones (1 à 5%) comparée aux dicotylédones (20%). D'un autre côté, le GAX est prédominant dans les parois primaires et secondaires des monocotylédones (près de 40%). Enfin, les xylanes sont présents en très faible proportion dans la paroi primaire des dicotylédones et le glucuronoxylane est majoritaire (20 à 30%) dans les parois secondaires. Ces données structurales peuvent fournir des clés importantes sur l'écologie microbienne au sein des pathogènes de plante en étudiant leurs enzymes de dégradation de la paroi végétale.

Il a été montré que le xylane est essentiel à la formation de la paroi secondaire. En effet, des perturbations de la biosynthèse des xylanes entraînent des altérations importantes dans l'organisation des parois secondaires qui peuvent aller jusqu'à leur absence (Bauer *et al.*,

2006; Pena *et al.*, 2007; Zhou *et al.*, 2006). Les xyloglucanes peuvent jouer un rôle comme source de molécule signal. En effet, un produit de dégradation du xyloglucane, XXFG (Cf. Figure 23, xyloglucane), contrebalance l'expansion cellulaire stimulée par l'auxine (York *et al.*, 1984). A contrario, d'autres oligomères tels que XXLG et XLLG (Cf. Figure 23, xyloglucane), peuvent accélérer l'expansion cellulaire (McDougall and Fry, 1990). Cependant, des mutants affectés dans la biosynthèse des xyloglucanes peuvent présenter des phénotypes marqués (Tedman-Jones *et al.*, 2008), tandis que des mutants dépourvus de xyloglucanes ne présentent aucun phénotype notable (Cavalier *et al.*, 2008).

III.4. Autres composants

La paroi végétale est constituée également par un autre composant essentiel, les protéines pariétales. Elles peuvent être classées suivant leurs caractéristiques physico-chimiques ou suivant leur fonction. On distingue 3 classes principales de glycoprotéines pariétales : les protéines riches en glycine (GRP, Glycin-rich proteins), les protéines riches en proline (PRP, Prolin-rich proteins) et les protéines riches en hydroxyproline (HRGP, Hydroxyprolin-rich Glycoproteins).

La paroi cellulaire est dynamique et contient plusieurs enzymes et autres agents (Cosgrove, 1999) qui confèrent ces propriétés dynamiques telles que l'extension et la plasticité. Parmi ces protéines, on trouve les expansines, les enzymes de dégradation (exemple des cellulases, xylanases, pectin methylesterases, …), les transférases (exemple de la xyloglucan endotransglycosylase hydrolases (XTHs)), les lyases et les oxido-reductases.

IV. L'utilisation de la paroi végétale par les bactéries

La synthèse de carbone organique est un processus biologique majeur et la première source d'énergie pour la vie. Grâce à la photosynthèse, les plantes convertissent l'énergie solaire en carbone organique qui peut ensuite être utilisé par des organismes hétérotrophes. Ainsi, la dégradation de la paroi végétale par les microorganismes tient un rôle très important dans le cycle du carbone sur la planète. Cependant, seul un nombre restreint de microorganismes ont acquis la capacité à dégrader cet ensemble de carbohydrates. La complexité physique et chimique des parois végétales rend difficile l'action des enzymes et donc le recyclage du carbone photosynthétique est un processus biologique relativement inefficace (Brett and

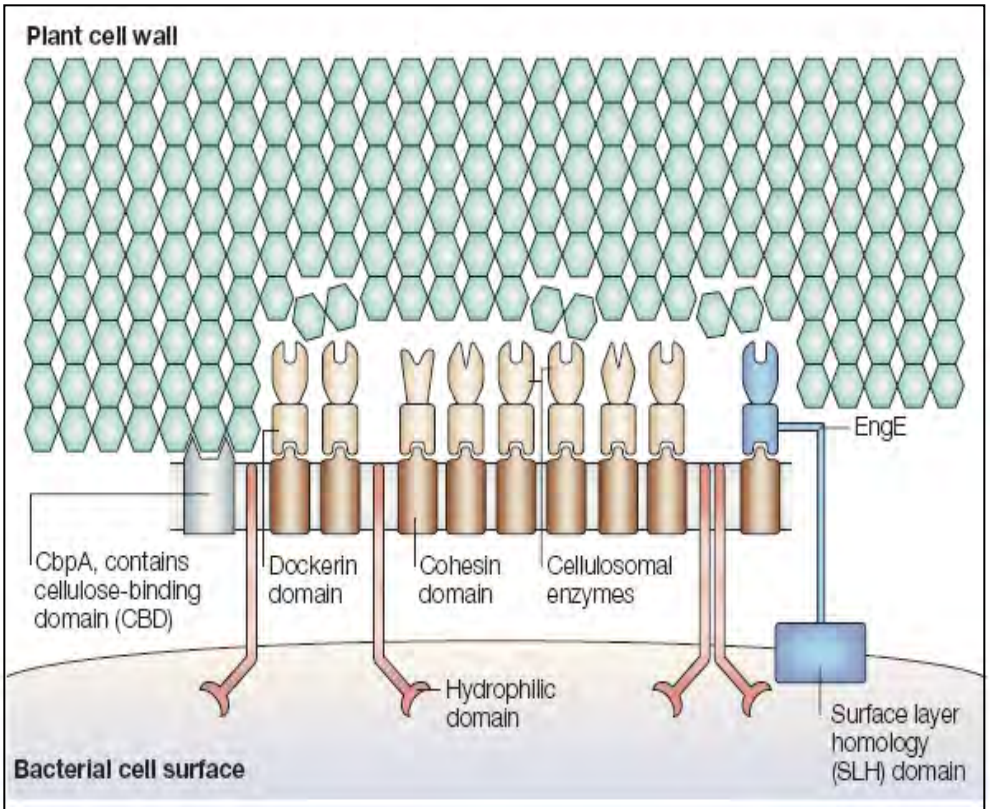


Figure 24. Modèle schématique d'un cellulosome de *Clostridium cellulovorans*. La protéine scaffoldine est représentée ici avec son domaine de liaison à la cellulose (CBD), 9 cohésines, 4 domaines hydrophiles et 9 enzymes cellulosomiques liées à la scaffoldine par l'intermédiaire des dockerines. EngE est vraisemblablement impliqué dans l'ancrage du cellulosome à la surface cellulaire bactérienne (Doi *et al.*, 2003).
(D'après Doi and Kosugi, 2004).

Waldren, 1996). Un répertoire étendu d'enzymes extracellulaires qui agissent en synergie pour dégrader ces structures composites représente une caractéristique commune des organismes capables de dégrader ces ensembles organiques (Warren, 1996).

A notre connaissance, deux types de systèmes enzymatiques permettant la dégradation des parois végétales ont été observés chez les microorganismes. Les champignons et bactéries Gram- telle que *Xcc* sécrètent plusieurs endoglucanases, exoglucanases et d'autres enzymes qui agissent de manière coordonnée et individuelle pour attaquer la paroi végétale (Cf.p33, paragrpahe IV.2) (pour revue (Lynd *et al.*, 2002; Warren, 1996)). Chez les microorganismes anaérobies, un sytème différent a émergé. Il implique la formation de grands complexes enzymatiques extracellulaires qui regroupent des protéines et des enzymes liées à ces dernières, les cellulosomes (Bayer *et al.*, 1985).

IV.1. Cellulosomes et Xylanosomes

Plusieurs auteurs avancent que les cellulosomes sont plus efficaces que les systèmes d'enzymes libres dans la dégradation des parois végétales (Demain *et al.*, 2005). Il est possible que les pressions de sélection qu'impose un environnement anaérobie aient conduit à la formation des cellulosomes bien que la nature de ces forces évolutives reste encore inconnue.

-Composition

Les cellulosomes sont constitués d'une protéine fibrillaire (la « scaffolding » protéine) avec des masses (sous-unités enzymatiques) disposées périodiquement le long des fibrilles (Madkour and Mayer, 2003). La protéine non-enzymatique « scaffolding » ou scaffoldine contient des sites de liaison (les cohésines) pour les sous-unités cellulosomiques (Bayer *et al.*, 1994). Ces sous-unités possèdent différentes fonctions et contiennent invariablement un site de liaison à la cohésine, la dockerine (Bayer *et al.*, 1994) (Figure 24). L'interaction cohésine/dockerine est un facteur important dans l'assemblage du cellulosome (Tokatlidis *et al.*, 1991). L'interaction entre ces deux unités est spécifique de 4 acides aminés de la dockerine (Mechaly *et al.*, 2000). La majorité des scaffoldines contiennent entre 6 et 9 cohésines qui peuvent lier jusqu'à 26 enzymes cellulosomiques différentes. Ainsi, les grandes possibilités de combinaisons d'enzymes liées à la scaffoldine offrent un nombre important de

cellulosomes différents au sein d'un seul microorganisme (Doi *et al.*, 2003). Cette diversité de cellulosomes est augmentée par le fait que certaines espèces bactériennes, comme *Acetivibrio cellulolyticus* possèdent plusieurs scaffoldines (Xu *et al.*, 2003b; Xu *et al.*, 2004a; Xu *et al.*, 2004b). Cette hétérogénéité dans la population des cellulosomes va permettre à la bactérie une meilleure adaptation aux différents types de parois végétales rencontrées dans l'environnement et ainsi une activité de dégradation accrue.

-Liaison à la surface cellulaire bactérienne

La plupart des scaffoldines contiennent un segment, le module SLH (surface layer homology) ou domaine SLH qui est habituellement présent chez la majorité des SLPs (surface layer protein) (Rincon *et al.*, 2003; Rincon *et al.*, 2004). Les domaines SLH permettraient l'ancrage des protéines à la paroi bactérienne (Figure 24) et pourraient se lier au peptidoglycane**.**

-Liaison à la paroi végétale

De manière opposée, un autre élément des scaffoldines permet la liaison à la paroi végétale. En effet, chaque scaffoldine contient également un domaine de liaison à la cellulose (CBD, cellulose-binding domain) ou module de liaison à un carbohydrate (CBM, carbohydrate-binding module) (Boraston *et al.*, 2004) (Figure 24). Les CBMs ont été classés en 59 familles d'après leurs séquences (Cantarel *et al.*, 2009). Des études structurelles et biochimiques divisent ces modules en 3 classes selon le type de saccharide qu'ils vont reconnaître (Boraston *et al.*, 2004). Les parois végétales sont hétérogènes et sont composées d'une grande variété de polysaccharides qui interagissent entre eux. Ainsi en plus de la liaison à la cellulose, les cellulosomes ciblent d'autres molécules pour permettre aux composants catalytiques du complexe enzymatique d'être mis à proximité de leurs substrats spécifiques. En effet plusieurs enzymes cellulosomiques possèdent des CBMs spécifiques de la cellulose, du xyloglucane, du xylane ou encore de la pectine.

-Les enzymes cellulosomiques

On compte parmi les enzymes cellulosomiques des cellulases, des hémicellulases, des pectinases, des chitinases et plusieurs autres enzymes auxiliaires capables de dégrader la paroi

végétale. Chez *C. thermocellum*, 26 enzymes ont pu être identifiées (Bayer *et al.*, 1998). Plusieurs de ces enzymes agissent de concert pour faciliter la dégradation des polymères principaux tels que les xylanes ou les mannanes par exemple.

-La régulation de l'expression des gènes cellulosomiques

Des études microscopiques pionnières ont montré la présence de protubérances qui contiennent des polycellulosomes (Lamed *et al.*, 1983) qui ne sont observables que lorsque les cellules poussent en présence de cellulose (Blair and Anderson, 1999). En présence de cellulose, 4 heures sont nécessaires à la formation de ces protubérances chez *Clostridium cellulovorans*. L'ajout de sucres solubles tels que le glucose, le cellobiose ou le methylglucose ne permet pas la formation des polycellulosomes et est responsable d'une dissociation rapide des protubérances. De plus, il est intéressant de noter que la présence d'un carbohydrate spécifique comme le glucose, le cellobiose, le xylane, le mannane ou la pectine dans le milieu de croissance influence directement la composition et l'activité enzymatique des sous-unités des cellulosomes (Ali *et al.*, 1995; Han *et al.*, 2003a). Chez *C. cellulovorans*, au niveau transcriptionnel, l'ajout de ces mêmes substrats va induire l'expression de la majorité des gènes cellulosomiques. En revanche, l'ajout de divers monosaccharides conduit à des niveaux faibles ou modérés de l'expression de ces mêmes gènes (Han *et al.*, 2003b).

-Les xylanosomes

Les xylanes, composés d'un squelette de résidus de D-xylose unis en β-(1,4) et de plusieurs types de substitutions (Cf. Figure 23, glucuronoarabinoxylane), représentent le groupe de polysaccharides non cellulosique majeur au sein des parois végétales (Dodd and Cann, 2009). La production de xylanases est une stratégie très largement répandue chez un grand nombre de microorganismes pour permettre une dégradation efficace du xylane (Kulkarni *et al.*, 1999; Sunna and Antranikian, 1997). Précédemment, nous avons pu voir que plusieurs bactéries anaérobies cellulolytiques sont capables de produire des cellulosomes (Schwarz, 2001). De la même manière, chez quelques bactéries plusieurs types de xylanases sont organisés en structures multifonctionnelles, les xylanosomes. Ces complexes jouent un rôle très important dans la dégradation de l'hémicellulose (Sunna and Antranikian, 1997). Le premier xylanosome a été identifié chez *Butyrivibrio fibrisolvens* H17c. Il est composé de 11 protéines

distinctes possédant une activité xylanase et 3 protéines ayant une activité endoglucanase, le complexe ayant une masse de plus de 669 kDa (Lin and Thomson, 1991).

-Cellulosomes et bioénergie

La société fait actuellement face à des problèmes énergétiques sans précédents et de grande envergure. Il est aujourd'hui nécessaire de développer des énergies renouvelables et alternatives afin de palier aux futures carences en énergie fossile. Cela requiert le développement de combinaisons de processus basés sur des substrats renouvelables. Actuellement, chaque année, 10^{11} tonnes de biomasse végétale, comprenant principalement les parois végétales, sont dégradées par les microorganismes et l'énergie produite par ce processus correspondrait à plusieurs centaines de milliards de barils de pétrole (Boudet *et al.*, 2003; Ragauskas *et al.*, 2006). Cependant, l'hydrolyse des polysaccharides reste l'étape limitante de la conversion de la lignocellulose en carburant. Cela nécessite le développement de systèmes enzymatiques très efficaces. Le cellulosome peut être décrit comme une des nanomachines naturelles les plus élaborées. Leur découverte a mis en lumière la nécessité d'un bloc enzymatique afin de surmonter l'hydolyse complexe des parois végétales. De plus, à l'heure de la génétique, les cellulosomes de bactéries telles que C. *thermocellum* peuvent être modifiés afin d'augmenter leur capacité à synthétiser de l'éthanol à partir de lignocellulose. Une stratégie alternative intéressante consiste à synthétiser génétiquement des cellulosomes qui seraient introduits chez des microorganismes qui sont capables de transformer par fermentation des sucres simples en molécules d'intérêt industriel tel que le butanol mais qui n'ont pas d'appareils endogènes de dégradation de la paroi végétale (Mingardon *et al.*, 2005; Perret *et al.*, 2004a; Perret *et al.*, 2004b).

IV.2. Les enzymes de dégradation de la paroi végétale chez les bactéries phytopathogènes et Xcc

L'autre système permettant la biodégradation de la paroi végétale par les microorganismes est la production et la sécrétion d'un grand nombre d'enzymes de dégradation (pour revue (Warren, 1996). *Xcc* possède un des plus grands « CAZyme » ou « glycobiome » bactérien. En effet, selon la base de données CAZy (Carbohydrate-Active enZYmes, www.cazy.org), chez *Xcc*, on trouve 173 protéines avec 93 domaines « Glycosyl Hydrolase » (GH), 48 domaines « Glycosyl Transferase » (GT), 7 domaines « Polysaccharide Lyase » (PL), 16 domaines « Carbohydrate Esterase » (CE) et 9 domaines « Carbohydrate-Binding Module » (CBM). Cet important répertoire reflète une grande capacité d'utilisation des ressources végétales, lui conférant ainsi un avantage sélectif.

-Dégradation des pectines

Les enzymes de dégradation de la pectine affaiblissent la paroi végétale en exposant les autres polymères à une dégradation par des hémicellulases et des cellulases. Il y a quatre familles principales d'enzymes impliquées dans la dégradation de la pectine : les polygalacturonases, les pectates lyases, les pectines lyases et les pectines méthylestérases (PMEs). Les PMEs catalysent la démethylesterification de l'homogalacturonane facilitant ainsi les actions des polygalacturonases et des pectates lyases. Les PMEs favorisent ainsi la macération des tissus chez *D. dadantii* (Studholme *et al.*, 2010). Les polygalacturonases hydrolysent la liaison entre les acides galacturoniques des homogalacturonanes non méthylés par un mécanisme d'hydrolyse à pH acide à neutre tandis que les lyases procèdent par un mécanisme de β-élimination calcium-dépendant à pH légèrement basique. Les pectates lyases sont spécifiques des substrats non méthylés alors que les pectines lyases dégradent les formes méthylées.
Les enzymes de dégradation de la pectine sont d'importants facteurs de virulence chez *E. chrysanthemi* (Kotoujansky, 1987). Le génome de *Xcc* possède deux gènes codant des polygalacturonases, *pghAxc* et *pghBxc* dont l'expression est régulée par les régulateurs majeurs du SST3, HrpG et HrpX ainsi que par le répresseur catabolique CLP, un homologue au répresseur de l'AMP cyclique d'*E. coli*. Ces polygalacturonases sont sécrétées via le système de sécrétion de type II (Wang *et al.*, 2008). Ces polygalacturonases sont impliquées

dans la formation de gels pectiques altérant la conduction de sève et permettant à la bactérie de collecter des nutriments d'origine végétale (Prade *et al.*, 1999).

-Dégradation de la cellulose

Parmi les cellulases, on trouve les endo-β-1,4-glucanases, les β-glucosidases et les cellobiohydrolases (aussi appelées exo-β-1,4-glucanases) qui agissent en synergie dans l'hydrolyse complète de la cellulose en glucose. Il est généralement admis que les cellulases ne sont pas des acteurs majeurs de la virulence du fait que la dégradation complète de la cellulose prend place tardivement au cours de l'infection, sinon à la fin (Cooper, 1984). Cependant, la délétion des gènes codant deux endoglucanases qui ont été décrites dans la souche 8004 de *Xcc*, *engXCA* et *engXCB*, entraîne une diminution de la virulence et de l'activité cellulolytique d'un facteur 5 (Gough *et al.*, 1990; Schroter *et al.*, 2001).

-Dégradation des xylanes

Les hémicelluloses sont composées d'une grande diversité de sucres (Cf. p27, paragraphe III.3) qui comptent parmi les plus abondants sur Terre et l'étude de la dégradation de ces polysaccharides par les pathogènes prend donc une grande place dans la compréhension des interactions entre les plantes et les microorganismes. De plus, avec la cellulose, ces composés ont été placés sur le devant de la scène pour devenir les substrats majeurs dans la bioconversion en hydrocarbones et en éthanol ou d'autres alcools de haut poids moléculaire (Somerville, 2007). En effet, les récentes problématiques de changement climatique et de la diminution des énergies fossiles ont amené la société à se tourner vers des combustibles renouvelables tels que les éthanols cellulosiques et hémicellulosiques. Le xylane est le composé hémicellulosique majoritaire des parois végétales et le second polysaccharide végétal le plus abondant dans la nature après la cellulose (Saha, 2003). Ainsi, sa dégradation en sucres simples, principalement le xylose et l'arabinose, est une étape majeure dans l'utilisation efficace de la biomasse végétale par le pathogène et pour la production de biocarburant. Coughlan avance l'hypothèse que le contenu énergétique du xylane et de la cellulose, basé sur des estimations de la biomasse végétale globale totale, serait équivalent à environ 640 milliards de tonnes de pétrole (Coughlan, 1985). Le xylane contient une grande variété de liaisons chimiques. Ainsi, sa dégradation requiert un nombre conséquent d'activités

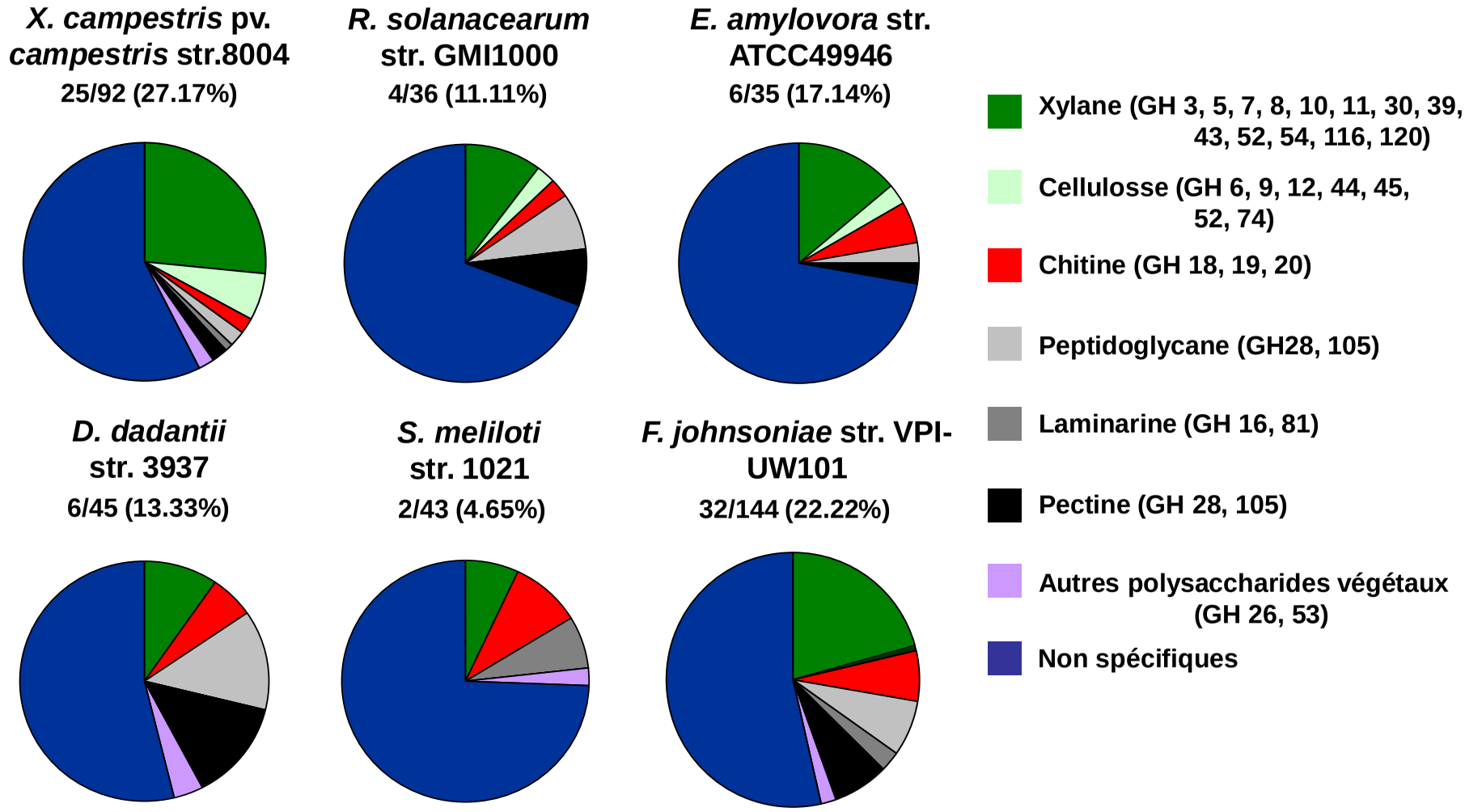


Figure 25. Répartition des domaines GH selon la spécificité de substrat. Les domaines Glycoside Hydrolase (GH) des bactéries *X. campestris* pv. *campestris* (*Xcc*), *Ralstonia solanacearum*, *Erwinia amylovora*, *Dickeya dadantii*, *Sinorhizobium meliloti* et *Bacteroides thetaiotaomicron* sont répartis selon la spécificité de substrat et sont représentés en pourcentage par rapport au nombre total de GH par génome. La proportion de GH possédant une spécificité pour le xylane est indiqué est indiqué en-dessous de chacun des noms d'espèce. On peut voir que *Xcc* contient la plus grande proportion de protéines ayant des domaines GH spécifiques du xylane en comparaison avec les génomes des autres bactéries.

enzymatiques : endo-1,4-β-xylanases, β-D-xylosidases, α-L-arabinofuranosidases (AFs), α-glucuronidases, acetyl xylan esterases et ferulic/coumaric acid esterases. La diversité structurelle des xylanes nécessite une diversité équivalente du répertoire enzymatique nécessaire à la dégradation de ces polysaccharides. Ainsi plusieurs machineries enzymatiques nécessaires à l'hydolyse du xylane ont évolué chez les bactéries et les champignons xylanolytiques. La colonne principale du xylane va être dégradée par des xylanases. On trouve ces enzymes dans les familles GH 5, 7, 8, 10, 11, 30 et 43 d'après leurs séquences en acides aminés (Cantarel *et al.*, 2009). Les xylanases des familles GH 10 et 11 sont les plus étudiées et les mieux caractérisées (Biely *et al.*, 1981; Collins *et al.*, 2005; Vrsanska *et al.*, 1982). Les produits de dégradation issus de l'activité des xylanases de la famille GH 10 sont généralement plus courts que ceux issus de l'activité des xylanases GH 11 du fait de leurs différents sites de clivage (Biely *et al.*, 1997; Kolenova *et al.*, 2006). Les xylosidases, quant à elles, permettent de produire des monomères de xylose issus des extrémités non-réductrices des xylooligosaccharides. On retrouve ces enzymes dans les familles GH 3, 30, 39, 43, 52, 54, 116 et 120. Les chaînes latérales sont retirées par l'action des arabinofuranosidases (familles GH 3, 43, 51, 54 et 62), glucuronidases (familles GH 4, 67 et 115), acetylxylanesterase (familles CE 1, 2, 3, 4, 5, 7 et 12), ferulic acid esterases (CE 1) et des *p*-coumaric acid esterases (CE 1) (Dodd and Cann, 2009). Le criblage de la base de données CAZy montre qu'une grande proportion des domaines GH de *Xcc* (27.17%) est consacrée à l'hydrolyse du xylane et de ses produits de dégradation, en comparaison à d'autres bactéries phytopathogènes telles que *R. solanacearum* (11.11%), *E. amylovora* (17.14%), *D. dadantii* (13.33%) ou la bactérie symbiotique végétale *S. meliloti* (4.65%). Cette proportion est également plus importante que celle présentée par *F. johnsoniae* (22.22%), appartenant au phylum des Bacteroidetes, qui est capable de dégrader une grande diversité de polysaccharides végétaux (McBride *et al.*, 2009) et possède un plus grand nombre de domaines GH (144) que *Xcc* (Figure 25).

Des xylanases ont été montrées comme intervenant dans la pathogénicité de 2 membres du genre *Xanthomonas*. Ces enzymes sont codées par les gènes *xynB* et *xynC* chez *Xoo* et *Xcv* respectivement (Rajeshwari *et al.*, 2005; Szczesny *et al.*, 2010).

Les bactéries se retrouvent souvent en compétition avec des microorganismes pour accéder aux nutriments dans des milieux qui peuvent aussi être pauvres en sources de carbone, azote et phosphate. De plus, les produits de dégradation d'origine végétale peuvent être reconnus par la plante comme éliciteurs de réponse de défense. Ainsi, les bactéries ont développé des

systèmes de transport évolués afin de pouvoir utiliser ces produits de dégradation avec la plus grande efficacité possible.

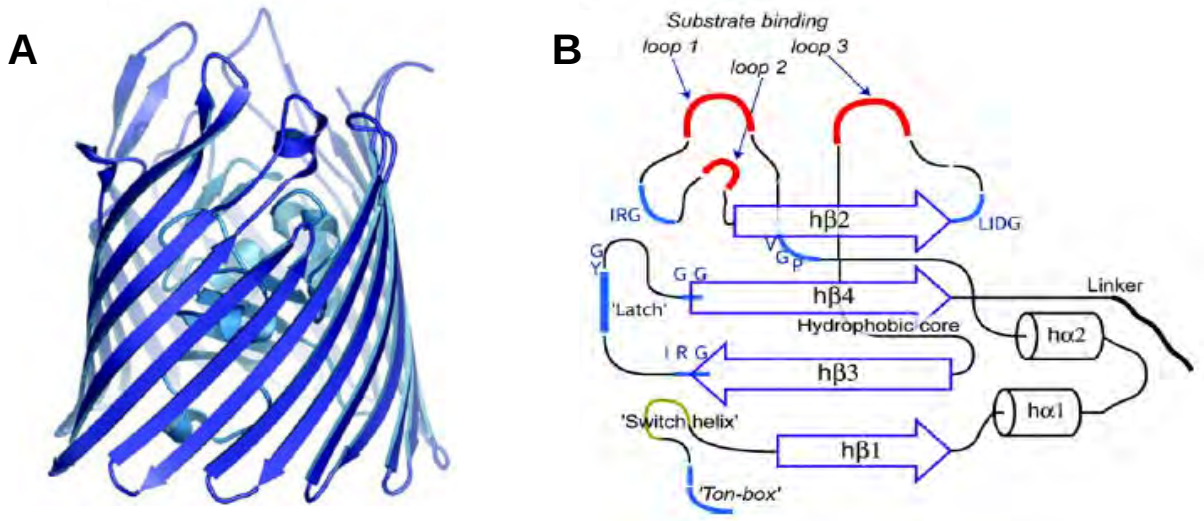


Figure 26. Structure générale des transporteurs TonB-dépendants.

A. Structure en ruban du récepteur BtuB. Les transporteurs TonB-dépendants possèdent un domaine C-terminal de 22 brins bêta antiparallèles associés en tonneau, et un domaine N-terminal (bouchon) inséré dans le tonneau. Les brins bêta sont connectés du côté périplasmique par des coudes et du côté extracellulaire par de longues et flexibles boucles.

B. Schéma en deux dimensions du domaine N-terminal des récepteurs TonB-dépendants. Les structures secondaires conservées sont indiquées (hélices alpha : cylindre, brins bêta : rectangles, boucles : traits simples). Les motifs conservés de fonction connue ou inconnue sont représentés en bleu. Les boucles apicales constituant le site de fixation du sidérophore sont indiquées en rouge. La switch hélice (FhuA, FecA) est indiquée en jaune.

(D'après Andrews *et al*., 2003 ; Huché, 2006).

LES TBDTs PEUVENT TOUT « FER »

Différentes protéines sont responsables du transport des molécules à travers la membrane externe. Les porines sont les protéines majoritaires de la membrane externe bactérienne et permettent l'entrée de molécules hydrophiles de petite taille par diffusion simple ou facilitée. *P. aeruginosa* possède une porine nommée OprB qui permet le passage lent et non spécifique de solutés de masse moléculaire inférieure à 300 Da (Trias, 1988). Chez *E. coli*, OmpF et OmpC ont une limite de diffusion d'environ 500 Da, bien qu'OmpF permette le passage de solutés légèrement plus gros qu'OmpC (Nikaido, 2003).

Enfin, pour les molécules de plus grande taille comme les complexes Fer-sidérophores, le transport s'effectue via les TBDTs (Cf. p12, paragraphe II.3). Ces protéines présentent des caractéristiques communes quant à leur structure, leur fonction et leur mécanisme moléculaire.

I. Structure

En 1998, la première structure cristallographique d'un TBDT est publiée (Ferguson *et al.*, 1998). En 2005, Chimento et ses collaborateurs ont conduit une analyse comparative portant sur les 4 structures de TBDT publiées à cette époque, FepA, BtuB, FecA et FhuA ((Buchanan *et al.*, 1999; Chimento *et al.*, 2003, 2005; Ferguson *et al.*, 1998; Ferguson *et al.*, 2002). Depuis, les structures de 8 TBDTs supplémentaires ont été déterminées. A cela s'ajoute les structures des TBDTs liés à leurs ligands, celles des TBDTs en complexe avec le domaine périplasmique de la protéine TonB et celle d'un TBDT cristallisé à partir d'une phase lipidique cubique, totalisant 45 structures cristallographiques disponibles. Ces protéines possèdent une homologie de structure importante malgré une identité de séquence assez faible. Les TBDTs sont composés de plusieurs domaines caractéristiques.

La partie C-terminale de la protéine est composée d'un grand domaine d'environ 600 résidus, le tonneau bêta ou « β-barrel ». Ce tonneau bêta se compose de 22 brins bêta antiparallèles qui sont connectés par des boucles de 25 résidus en moyenne du côté extracellulaire et par des coudes de 6 résidus en moyenne du côté périplasmique (Figure 26A). La hauteur approximative du tonneau est de 70 Å dans le cas de FepA et FhuA. Le pore formé par le tonneau a une largeur de 35 à 40 Å (domaine pfam00593) (Ferguson and Deisenhofer, 2002).

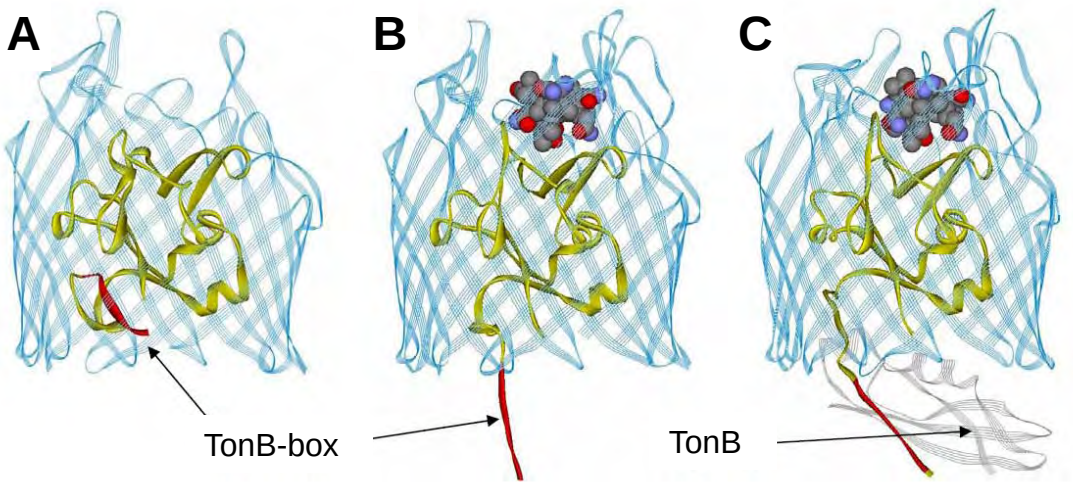


Figure 27. Structure du TBDT BtuB dans 3 configurations différentes.
A. Forme « apo » de BtuB ; la TonB-box est représentée en rouge et le domaine N-terminal de la protéine qui est replié est représenté en jaune.
B. Structure du TBDT en liaison avec son substrat, la vitamine B_{12}, avec un changement de configuration du domaine TonB-box.
C. Structure de BtuB lié au domaine C-terminal de la protéine TonB (représenté en gris clair).
(D'après Kim *et al*., 2007).

Le domaine N-terminal (Figure 26B) forme un bouchon ou « plug » qui est replié à l'intérieur du tonneau bêta. Il empêche le passage direct des nutriments dans le périplasme. La conservation des résidus est plus importante pour le bouchon que pour le tonneau (domaine pfam007715). A l'extrémité N-terminale de la séquence de ce domaine, on trouve la boîte TonB ou « TonB-box », courte séquence de 5 ou 6 acides aminés (Lundrigan and Kadner, 1986). Cette séquence constitue une signature de la famille des TBDTs. Elle permet l'interaction entre le TBDT et la protéine TonB, ancrée dans la membrane interne et associée avec les protéines ExbB et ExbD, permettant alors d'amener l'énergie nécessaire au transport via le TBDT (Cf. p12, paragraphe II.3). Cette interaction est facilitée par la transition allostérique induite par la fixation du substrat à la poche extracellulaire du TBDT (Figure 27).
Certains TBDTs présentent une extension N-terminale d'une centaine de résidus. C'est notamment le cas des TBDTs FecA d'*E. coli* et FpvA de *P. aeruginosa*. Ce domaine est impliqué dans la régulation de la transcription du gène de structure du TBDT et des gènes co-transcrits (Figure 8B). Chez *R. solanacearum*, le TBDT PrhA qui possède ce type d'extension, est responsable de l'induction de l'expression des gènes *hrp* en réponse à un contact avec les cellules végétales (Aldon *et al.*, 2000; Marenda *et al.*, 1998).
Enfin, certains TBDTs peuvent présenter une seconde extension N-terminale, intercalée entre l'extension N-terminale conventionnelle et le bouchon. Ce domaine est homologue au domaine N-terminal de la protéine Oar de *Myxococcus xanthus* impliquée dans un processus de morphogénèse au cours de la formation de myxospores. On trouve des TBDTs présentant cette extension chez *Xanthomonas* spp.

II. Régulation de l'expression des TBDTs

Bien que le fer soit indispensable aux bactéries, un excès en fer peut en revanche être toxique car cela mène à la production d'espèces réduites de l'oxygène, de radicaux superoxydes et hydroxyles (Braun, 1997). Ainsi, un contrôle de la concentration en fer chez les bactéries est essentiel. L'expression des gènes d'acquisition du fer sont régulés par la protéine régulatrice Fur (Ferric-uptake regulator). En condition d'abondance en fer, Fe^{2+} agit comme un co-répresseur. Ainsi, le complexe Fur- Fe^{2+} bloque l'expression des gènes cibles en se fixant au niveau d'un motif nucléique très conservé appelé Fur-box (Figure 8). En revanche, dans des conditions de carence en fer, la protéine Fur reste libre ; l'opérateur est donc accessible, ce qui permet la transcription des gènes impliqués dans l'acquisition et le stockage du fer (Bagg and

Neilands, 1987). La protéine Fur régule aussi l'expression des gènes impliqués dans plusieurs autres processus cellulaires tels que le stress oxydatif, la production de toxine et la virulence (Cha *et al.*, 2008; Kitphati *et al.*, 2007). Dans plusieurs bactéries, des homologues de Fur ont évolué vers des fonctions de senseurs d'autres métaux tels que le manganèse (Mur, (Diaz-Mireles *et al.*, 2004)), le zinc (Zur, (Gaballa *et al.*, 2002)) et le nickel (Nur, (Ahn *et al.*, 2006). Chez *D. dadantii*, la protéine Fur régule également l'expression des gènes codant pour des pectates lyases impliquées dans la virulence qui permettent la dégradation des parois végétales (Franza *et al.*, 2005).

Dans certains cas, il existe un second niveau de régulation : le signal de fixation du complexe fer/sidérophore au TBDT est transmis via l'extension N-terminale de ce TBDT à un facteur anti-sigma localisé dans la membrane plasmique libérant le facteur sigma nécessaire à la transcription des gènes d'acquisition du fer (Schalk *et al.*, 2004) (Figure 8B).

Il existe d'autres contrôles transcriptionnels qui affectent l'expression des gènes codant pour les TBDTs. En effet, l'expression de *fecA*, *fepA*, *cirA* et *fiuA* est augmentée dans un mutant du gène codant le régulateur transcriptionnel global Crp (Zhang *et al.*, 2005). La synthèse de ces 4 TBDTs est donc non seulement régulée par le fer mais aussi, très certainement par le niveau de carbone de la cellule.

La synthèse de BtuB, le transporteur de la vitamine B_{12}, est quant à elle régulée par un système de contrôle particulier, le riboswitch (Lundrigan *et al.*, 1991; Nahvi *et al.*, 2002). Les riboswitchs sont des éléments de l'ARN qui peuvent changer de conformation suite à une liaison spécifique avec une petite molécule (Mandal and Breaker, 2004). Ils sont classiquement positionnés à l'extrémité 5' des ARN messagers et la liaison du ligand induit des changements conformationnels qui affectent, positivement ou négativement, la transcription ou la traduction du ou des gènes situés en aval.

Enfin, chez *E. coli*, la synthèse des 3 TBDTs CirA, FecA et FepA est réprimée par les sRNAs (smalls Ribonucleic Acid) OmrA et OmrB (Guillier and Gottesman, 2006; Guillier *et al.*, 2006; Guillier and Gottesman, 2008). OmrA et OmrB sont des sRNAs de type « Hfq-binding » qui sont conservés chez plusieurs entérobactéries. Dans la majorité des cas, les sRNAs bactériens sont de courtes molécules d'ARN de 250 nucléotides ou moins et qui agissent en tant que régulateurs post-transcriptionnels. Les sRNAs constituent une classe majeure de régulateurs bactériens (Waters and Storz, 2009).

III. Des nouveaux substrats pour les TBDTs

Jusqu'à récemment, les TBDTs étaient connus pour être impliqués dans le transport actif des complexes ferriques et de la vitamine B_{12}. A cela s'ajoute d'autres fonctions pour le système TonB. En effet plusieurs phages, des colicines et antibiotiques d'origine naturelles (sideromycines) utilisent la machinerie des TBDTs afin d'accéder au périplasme bactérien (Braun, 1999; Cascales *et al.*, 2007; Killmann *et al.*, 1995). Le gène *tonA* codant pour un TBDT a été nommé ainsi car une mutation dans ce gène confère à la bactérie la résistance au phage *T one* (T1). Le TBDT a été renommé FhuA (Braun and Braun, 2002). Les TBDTs se révèlent des cibles de choix dans la lutte contre les bactéries pathogènes. Malgré les nombreuses études sur la mécanistique du transport au travers des TBDTs et sur leurs structures, seuls ou en complexes, de nombreuses études récentes montrent que notre connaissance sur les substrats des TBDTs est limitée. Au-delà des complexes ferriques et de la vitamine B_{12}, la découverte de nouveaux substrats transportés par les TBDTs a été confirmée par des données expérimentales.

Le premier exemple est le TBDT MalA, protéine de la membrane externe de la bactérie oligotrophe *Caulobacter crescentus* qui transporte les maltodextrines (Neugebauer *et al.*, 2005). Des expériences de transport montrent que MalA transporte le maltose et les maltodextrines de degré de polymérisation supérieur ou égal à 5, par un mécanisme actif, ExbB-D dépendant et TonB indépendant (Neugebauer *et al.*, 2005). Plus récemment, le transport du *N*-acétylglucosamine (GlcNAc) libre via le TBDT NagA a été montré sans ambiguïté chez cette même bactérie (Eisenbeis *et al.*, 2008).

Des analyses de génomique comparative des voies de régulation du GlcNAc ont été menées chez la bactérie marine *Shewanella oneidensis* (Yang *et al.*, 2006). Cette étude *in silico* a conduit à l'identification d'un TBDT, Chip-II qui interviendrait dans les voies d'utilisation des chito-oligosaccharides. Ce résultat découle de l'identification de 3 régulateurs, NagC, NagR et NagQ ainsi que de la prédiction de motifs de régulation dans certains promoteurs et de la caractérisation des régulons putatifs. De plus, il apparaît que d'autres TBDTs spécifiques des chito-oligosaccharides seraient aussi présents dans les groupes des Alteromonadales et des Xanthomonadales (Boulanger *et al.*, 2010).

En effet, une étude conduite chez *Xcc* a permis l'identification d'un régulon de gènes comprenant 4 TBDTs qui sont sous le contrôle du GlcNAc. Les résultats suggèrent que ces TBDTs ne seraient pas impliqués directement dans le transport de ce monosaccharide aminé

mais suggèrent qu'ils pourraient intervenir dans le transport de molécules plus complexes contenant du GlcNAc (Boulanger *et al.*, 2010).

Une étude fonctionnelle d'un TBDT contrôlant la virulence de *Xcc* révèle que ce dernier transporte le saccharose avec une affinité exceptionnelle. En effet, elle est environ 300000 fois meilleure que ce qu'on peut observer pour des porines qui permettent la diffusion du saccharose (Blanvillain *et al.*, 2007). Ces mêmes travaux montrent que plusieurs autres gènes codant pour des TBDTs sont induits par de l'acide polygalacturonique (PGA), du xylane et du maltose. Chez *Xcc*, les TBDTs sont sur-représentés. De façon intéressante, sur les 72 membres que compte cette famille protéique chez *Xcc*, seuls 9 seraient impliqués dans l'acquisition du fer et sont régulés par la protéine Fur (Blanvillain *et al.*, 2007). Ces résultats soulèvent la question du rôle des autres TBDTs encore nombreux et non caractérisés. Une hypothèse serait que plusieurs d'entre eux joueraient des rôles importants dans l'interaction avec la plante.

Enfin, chez la bactérie pathogène humaine *Helicobacter pylori*, l'expression du gène codant pour le TBDT FrpB4 est régulée par les ions nickel (Davis *et al.*, 2006). Les mutants de délétion de ce TBDT et du complexe protéique TonB-ExbB-ExbD ne transportent plus le nickel à pH faible (Schauer *et al.*, 2007).

Afin de caractériser le transport au travers des TBDTs, des molécules de saccharose et de maltodextrines radiomarquées au carbone 14 ont été utilisées. Ces expériences ont permis de montrer que ces nouveaux substrats partagent les mêmes profils de cinétique de transport que les complexes ferriques ou encore la vitamine B_{12}. Cependant, l'affinité de ces nouveaux substrats pour leurs TBDTs respectifs est de 10 à 100 fois plus faible que la liaison des substrats aux TBDTs « fer » et à BtuB (Blanvillain *et al.*, 2007; Neugebauer *et al.*, 2005). De plus une étude récente menée sur les TBDTs BtuB, FecA et FhuA montre que les transitions structurales induites par la liaison de leur substrat ne suivent pas toutes le même schéma (Kim *et al.*, 2007). Ces résultats indiquent que les mécanismes de transport peuvent différer selon le type de substrat. Il est donc intéressant de comprendre les causes des différents mécanismes de transport ; cela pourrait tenir de la structure chimique du substrat, sa disponibilité, du niveau d'énergie requis pour son transport ou encore autre chose.

De nombreuses autres études menées chez des espèces bactériennes très différentes suggèrent simplement l'implication des TBDTs dans des voies de signalisation ou des processus métaboliques de carbohydrates.

Les Sphingomonads sont connus pour dégrader une grande variété de polluants environnementaux et notamment les Polyethylène Glycol (PEGs), xénobiotiques très répandus dans l'industrie en tant que surfactants, dispersants, cosmétiques et lubrifiants. Chez *Sphingopyxis terrae* et *Sphingopyxis macrogoltabida*, les gènes de dégradation (*pegB, C, D, A, E* et *R*) sont rassemblés au sein d'un cluster. Cet opéron est réprimé par le régulateur PegR, de la famille AraC (Tani *et al.*, 2007). Le gène *pegB* code pour un TBDT qui interviendrait dans le transport du PEG mais cela est encore non démontré.

La souche *Sphingomonas* sp. A1 assimile directement l'alginate dans son cytoplasme à travers une sorte de puits, appelé « pit ». La formation de ce pit est réversible : il se forme en présence d'alginate et disparaît en son absence. Les gènes qui permettent la formation du pit sont rassemblés au sein d'un même cluster. Des études de transcriptomique et de protéomique ont permis d'identifier 8 protéines surexprimées en présence d'alginate : 4 TBDTs, 2 flagellines, une lipoprotéine et une « granule-binding protein ». Des mutants dans chacun des 4 gènes codant pour des TBDTs sont affectés en croissance sur alginate, suggérant un rôle dans l'utilisation de cette molécule. Les auteurs proposent que les TBDTs jouent un rôle structural au sein du pit (Hashimoto *et al.*, 2004; Hashimoto *et al.*, 2005).

La bactérie symbiotique du colon humain *Bacteroides thetaiotaomicron* peut utiliser un grand nombre de polysaccharides comme source de carbone et possède un « CAZyme » très important. Elle possède notamment toutes les protéines nécessaires à la dégradation de l'amidon et de ses dérivés. Le cluster sus (Starch Utilization System) comporte 4 protéines de la membrane externe (SusC, SusD, SusE et SusF) ainsi que 3 enzymes de dégradation de l'amidon (SusA, SusB et SusG). Shipman et ses collaborateurs décrivent SusC comme une porine mais ses caractéristiques font d'elle un TBDT. Un mutant dans le gène *susC* se développe normalement sur glucose et difficilement sur maltose ou maltotriose. Il ne croît en revanche plus du tout sur des maltodextrines de fort degré de polymérisation, suggérant un rôle de ce TBDT dans le transport de maltodextrines. De plus, il a été montré que les 3 protéines SusC, SusD et SusE interagissent dans la membrane externe, permettant la fixation de l'amidon (Shipman *et al.*, 2000). Des études préliminaires ont montré que la liaison de l'amidon est essentielle pour son utilisation (Anderson and Salyers, 1989a, b). Enfin et de façon intéressante, le cluster *sus* est partiellement conservé chez le pathogène opportuniste *Bacteroides fragilis*. Le TBDT fait partie d'un locus de 4 gènes, *osuABCD* qui est impliqué dans l'utilisation de l'amidon et la réponse au stress oxydatif (Spence *et al.*, 2006).

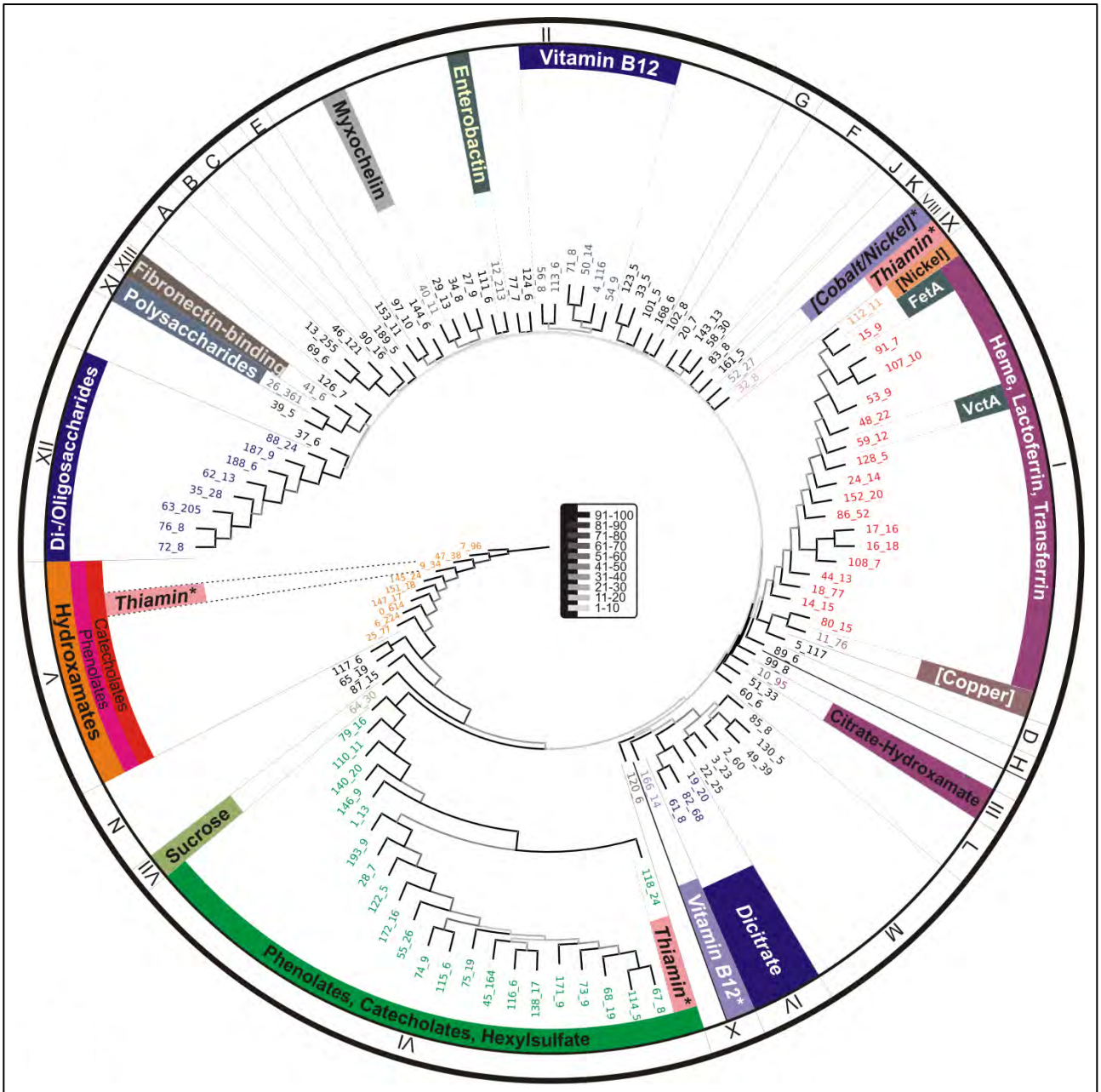


Figure 28. Regroupement des séquences de TBDTs putatifs.
Les séquences ont été analysées par CLANS (Cluster Analysis of Sequences). La figure représente l'arbre consensus construit de proche en proche montrant les distances entre les clusters. Les chiffres attribués à chaque TBDT sont représentés sous le format « x_y », où x est le numéro du cluster et y le numéro de la séquence du TBDT appartenant à ce cluster. Les substrats transportés sont regroupés par région indiquées par les chiffres romains « I à XIII » et les lettres « A à N ». Les crochets indiquent que l'ion métallique est connu, mais que le chélateur n'a pas encore été identifié. Enfin, une astérisque indique les substrats prédits (extrait de Mirus *et al*., 2009).

-classification et Distribution des TBDTs dans les génomes bactériens

Mirus et ses collaborateurs ont établi une classification de 98 TBDTs suivant la nature de leurs substrats connus ou putatifs. Afin de classer les TBDTs dont les substrats sont inconnus, les auteurs ont réalisé une analyse par cluster (méthode basée sur une co-localisation avec des gènes d'une voie métabolique spécifique ou sur une co-régulation par des facteurs de transcription ou des riboswitchs) sur 4600 séquences de TBDTs putatifs extraites de 686 génomes séquencés représentant 347 espèces bactériennes. Ainsi, 195 clusters regroupant au moins 2 séquences ont pu être définis (Mirus *et al.*, 2009) (Figure 28). Les TBDTs se regroupent selon leurs substrats et non selon leurs taxonomies, à l'exception des régions IX, XI, XIII et C (les deux dernières régions contiennent exclusivement des séquences de Bacteroidetes et γ-protéobactéries respectivement) (Figure 28). Par conséquent, il semble que la molécule transportée régisse les variations de séquence parmi les TBDTs.

Enfin, Blanvillain et ses collaborateurs ont réalisé une étude systématique des gènes codant pour des TBDTs dans les génomes séquencés de 226 bactéries Gram-. Il ressort de cette analyse que 27% des bactéries dont le génome a été analysé ne possèdent pas de TBDTs, 43% ont entre 1 et 13 TBDTs et 15.5% en possèdent plus de 30, constituant ainsi une classe particulière de bactéries où les TBDTs semblent être sur-représentés (Blanvillain *et al.*, 2007). La sur-représentation de cette famille protéique particulière confirme l'énorme diversité de substrats inconnus qu'il reste à découvrir.

Le fait que des substrats autres que des ions métalliques tels que des carbohydrates sont transportés par les TBDTs lève le voile sur la diversité chimique des molécules qui peuvent être utilisées par un mécanisme dépendant des TBDTs (pour revue (Schauer *et al.*, 2008)).

IV. Les TBDTs au cœur de nouveaux systèmes d'utilisation des glycanes complexes

IV.1. Les systèmes CUTs (Carbohydrate Utilization containing TBDT locus)

Le génome de *Xcc* possède 72 TBDTs qui peuvent être potentiellement impliqués dans le transport de divers carbohydrates en particulier. Cette bactérie possède également un des plus grands « CAZyme » avec notamment 93 « Glycosyl Hydrolase », suggérant une stratégie

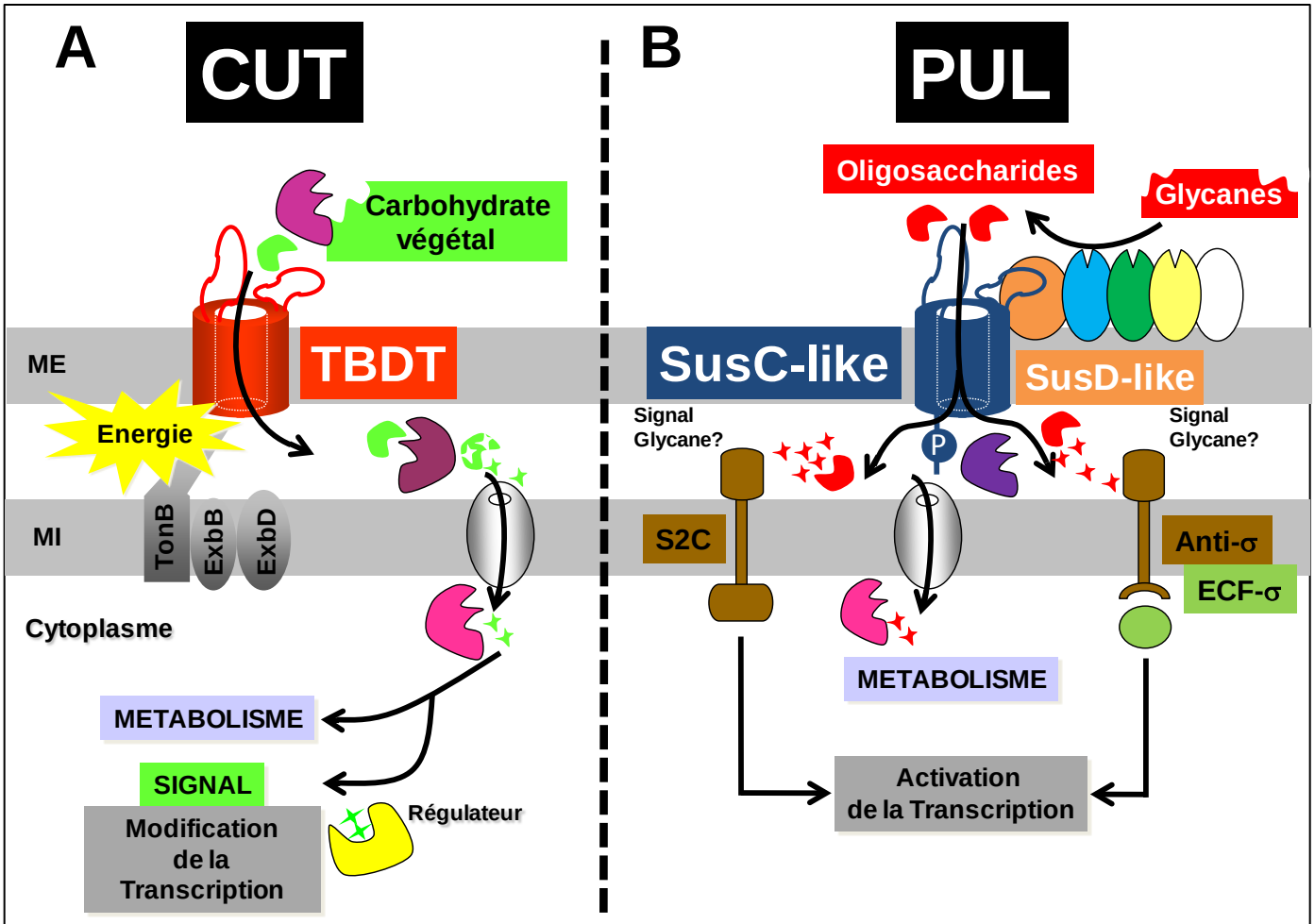


Figure 29. Modèles fonctionnels des systèmes (A) CUT (Carbohydrate Utilization containing TBDT locus) et (B) PUL (Polysaccharide Utilization Locus).

A. Dans un CUT locus, des composés issus de la dégradation d'un carbohydrate végétal sous l'action d'une enzyme extracellulaire est transporté de manière active, grâce à l'énergie amenée par le complexe protéique TonB-ExbBD, via un TBDT (en rouge). Ce ou ces produits de dégradation sont ensuite transportés du périplasme vers le cytoplasme grâce à un transporteur de la membrane interne (en gris clair). Une fois dans le cytoplasme, ce ou ces composés peuvent alors soit rentrer dans le métabolisme bactérien, soit agir comme molécule signal en interagissant avec le régulateur du système.

B. Dans un PUL, les glycanes sont liés et dégradés par diverses protéines ancrées dans la membrane externe de la bactérie telles que des polysaccharides lyases (en vert et jaune), des glycosides hydrolases (en bleu ciel), la lipoprotéine « SusD-like » (en orange) et des lipoprotéines hypothétiques (en blanc). Les oligosaccharides produits sont alors transportés via le TBDT « SusC-like » (en bleu foncé) de manière active active . Les produits issus de l'action des enzymes extracellulaires sont transportés vers le cytoplasme via des transporteurs de la membrane interne (en gris clair). Dans le périplasme, ce signal glycane est perçu par 2 genres de régulateurs : un S2C (système à deux composants) ou une paire anti-σ/ECF-σ; permettant ainsi la régulation du PUL. Pour simplifier le schéma, les protéines de la membrane interne TonB, ExbB et ExbD n'ont pas été représentées.

moléculaire spécifique pour l'utilisation de nutriments. En effet, l'étude approfondie du génome de *Xcc* a montré que parmi les gènes codant pour des TBDTs, certains d'entre eux sont présents au sein de clusters potentiellement impliqués dans le transport et la dégradation de molécules végétales. Une étude détaillée de l'acquisition du saccharose chez *Xcc* a montré la fonctionnalité d'un tel locus, définissant ainsi un nouveau concept, celui de « CUT locus » (Blanvillain *et al.*, 2007). Un tel système est classiquement composé d'au minimum un régulateur transcriptionnel, un transporteur de la membrane interne, une enzyme de dégradation spécifique d'un carbohydrate et un TBDT (Figure 29A). De façon intéressante, le CUT locus impliqué dans l'utilisation du saccharose, désigné *sux* (sucrose utilization in *Xanthomonas*), contrôle le pouvoir pathogène de *Xcc* sur *A. thaliana*, suggérant que l'acquisition de saccharose à très haute affinité constitue une étape critique dans le cycle infectieux de cette bactérie. Dans cette même étude, le criblage du génome de 226 bactéries Gram- indique que certains CUT loci de *Xcc* sont partiellement conservés chez différentes bactéries de l'environnement qui ont la particularité de présenter une surreprésentation des TBDTs. Ainsi, les CUT loci semblent non seulement participer à l'adaptation des bactéries phytopathogènes à leur hôte végétal mais joueraient également un rôle important dans le recyclage biologique des composés dérivés de plante dans un grand nombre d'environnements. L'étude de ces nouvelles voies chez les bactéries phytopathogènes, représente assurément un atout majeur pour mieux comprendre les mécanismes adaptatifs et évolutifs des bactéries.

IV.2. Les systèmes PULs (Polysaccharide Utilization Loci)

La communauté microbienne intestinale humaine est dominée par deux embranchements bactériens, les Bacteroidetes et les Firmicutes qui ont développé de grandes capacités à métaboliser des glycanes complexes qui inondent constamment leur habitat (Salyers *et al.*, 1977b; Salyers *et al.*, 1977c). Les travaux pionniers de Salyers et ses collaborateurs sur la dégradation de l'amidon chez la bactérie *B. thetaiotaomicron* ont permis la découverte d'un système multi-protéique associé à l'enveloppe cellulaire, le locus sus (IV.3). Il permet à la bactérie de lier et dégrader ce carbohydrate (Salyers *et al.*, 1977a). Ce système souligne le fait que les activités combinées de plusieurs protéines peuvent être plus sophistiquées et élaborées que le fonctionnement individuel de ces mêmes protéines. Des études de génomique comparative et fonctionnelle chez plusieurs espèces de Bacteroidetes révèlent que ces

bactéries possèdent des répertoires importants de protéines extracellulaires nommés « sus-like systems » qui sont capables de lier et de dégrader plusieurs types de glycanes (Martens *et al.*, 2008; Xu *et al.*, 2003a ; Xu *et al.*, 2007). En effet, le génome de *B. thetaiotaomicron* contient 208 homologues des gènes *susC* et *susD*, formant 101 paires de gènes « *susC*-like » et « *susD*-like ». Ces paires de gènes sont fréquemment des composants de plus larges clusters de gènes désignés PULs (Xu *et al.*, 2007) (Figure 29B). Au-delà de *susC* et *susD*, les gènes composant les PULs ne possèdent peu voire aucune homologie avec ceux du locus *sus* originel. *B. thetaiotaomicron* possède 88 PULs qui contiennent 866 gènes, représentant 18% de son génome (Martens *et al.*, 2008). La majorité de ces PULs « sus-like » peuvent être régulés soit par un système à deux composants, soit par un couple « ECF-σ factor/anti-σ factor » qui activent la transcription du PUL en réponse aux glycanes. La diversité des substrats utilisés par les systèmes « sus-like » s'étend au-delà de l'amidon. Ainsi, *Bacteroides ovatus* est capable de croître sur toutes les hémicelluloses connues grâce à la présence de PULs supplémentaires qui ciblent notamment le xylane et le galactomannane (Salyers *et al.*, 1977b; Valentine *et al.*, 1992; Valentine and Salyers, 1992a, b). Plus récemment, une étude menée chez *Prevotella bryantii*, un bacteroidete présent dans le rumen des bovins, met en évidence un cluster de gène qui comprend des TBDTs associé à la dégradation du xylane (Dodd *et al.*, 2010). De plus, l'étude de la distribution de ce système PUL indique que la majorité des gènes de ce cluster « xylane » est extrèmement bien conservée chez des Bacteroidetes environnementaux, associés aux humains et au rumen des bovins (Dodd et *al.*, 2010 ; Dodd *et al.*, 2011). Cela suggère que ce mécanisme d'utilisation du xylane est très bien conservé au sein de ce phylum. Le Bacteroidete environnemental *Flavobacterium johnsoniae* a la particularité de dégrader le glycane insoluble, la chitine. Le génome de cette bactérie possède un PUL qui contient 3 chitinases putatives (McBride *et al.*, 2003). La compréhension de ces systèmes fournit des informations clés quant à la capacité des Bacteroidetes intestinaux, terrestres mais aussi aquatiques, de survivre dans des écosystèmes compétitifs.

Cela suggère que des bactéries phylogénétiquement très éloignées et provenant d'habitats divers utilisent une même stratégie avec les TBDTs comme point d'ancrage pour dégrader la matière organique.L'étude de ces protéines, et plus largement des systèmes CUTs et PULs, permettrait une meilleure compréhension de l'écologie et de l'adaptation de la bactérie fournissant une meilleure vue de son évolution.

RESULTATS

La paroi végétale est une structure complexe composée d'un réseau de polysaccharides comprenant la cellulose, les hémicelluloses et les pectines. La majorité des phytopathogènes sécrètent des enzymes de dégradation afin de faire rompre cette barrière et d'utiliser les produits de dégradation comme source nutritive. La dégradation microbienne de la paroi végétale est un processus biologique important. Cependant, depuis la dernière décennie, elle représente un intérêt scientifique grandissant pour de nombreuses applications biotechnologiques dans la production de papier et de biocarburant par exemple. Le xylane est un composant structural majeur des parois végétales et le second polysaccharide végétal le plus abondant dans la nature après la cellulose (Dodd and Cann, 2009). L'utilisation du xylane nécessite l'action coordonnée de plusieurs enzymes afin de dépolymériser ce carbohydrate. De manière intéressante, les bactéries du genre *Xanthomonas* possèdent un arsenal enzymatique xylanolytique important.

Au cours de ce travail de thèse, nous avons caractérisé le système de dégradation et d'utilisation du xylane chez *Xanthomonas campestris* pv. *campestris*. Ce système comprend des enzymes nécessaires à la dégradation du xylane mais aussi à l'utilisation du xylose et de l'acide glucuronique et des transporteurs. Les gènes de ce système sont induits par le xylane et par les produits issus de sa dégradation, les xylo-oligosaccharides. La régulation de ce système est également dépendante d'un répresseur de type LacI, XylR, et des régulateurs du SST3, HrpG et HrpX. Ce système est requis pour la pathogénicité et nous avons montré qu'il est nécessaire pour la croissance optimale des bactéries à la surface des feuilles des plantes hôtes et non hôtes. L'une de ses particularités est la présence de 2 transporteurs spécifiques de la membrane externe (TBDT, TonB-dependent transporter) qui seraient impliqués dans le transport actif de produits d'hydrolyse du xylane. De plus, des travaux menés dans le laboratoire ont montré que cette aptitude à exploiter des composés végétaux semble être liée à la surreprésentation des TBDTs et à la présence de nombreux systèmes CUT (Carbohydrate Utilization systems containing TBDTs) (Blanvillain *et al.*, 2007). Ces systèmes CUT sont requis pour l'utilisation avec une très haute affinité des carbohydrates végétaux.

Enfin, des analyses de génomique comparative ont permis de définir un ensemble de gènes essentiel pour l'utilisation du xylane, conservé chez un grand nombre de bactéries phylogénétiquement très distinctes et nichant dans le sol, les plantes, les systèmes aquatiques ou les systèmes digestifs des animaux. Nos travaux montrent que cet ensemble de gènes est

systématiquement associé avec des TBDTs, soulignant l'importance de ces protéines dans l'utilisation du xylane. De plus, malgré la conservation de ce système xylanolytique, les xylanases produites par les espèces de Xanthomonas semblent être spécifiques aux bactéries phytopathogènes, suggérant un rôle spécifique dans la virulence.

Les carbohydrates représentent une source naturelle majeure de carbone et d'énergie pour la plupart des microbes hétérotrophes. Les variations dans les machineries cataboliques des sucres chez des microorganismes éloignés phylogénétiquement ou même très liés permettent de refléter la diversité des carbohydrates présents dans plusieurs écosystèmes, tout comme la diversité des modes de vie ou écologies microbiennes. Reconstruire ces machineries est crucial pour la compréhension de l'écophysiologie, de l'évolution, de l'adaptation et même des interactions entre les microorganismes et leurs hôtes.

Ces travaux sont regroupés dans un article récemment soumis pour publication, et présenté ci-après.

ARTICLE

TITLE

Characterization of xylan utilization system from phytopathogenic *Xanthomonas* spp. reveals common strategies with rumen and human gut symbionts.

Guillaume Déjean*[1], Servane Blanvillain*[1ψa], Alice Boulanger[1ψb], Thomas Dugé de Bernonville[1], Sébastien Carrere[1], Stevie Jamet[1], Claudine Zischek[1], Martine Lautier[1,2], Anne-Laure Girard[3], Armelle Darrasse[3], Marie-Agnès Jacques[3], Perrine David[3], Tristan Boureau[4], Stéphane Poussier[5], Emmanuelle Lauber[1] and Matthieu Arlat[1,2 Φ]

[1] INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France.

CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.

[2] Université de Toulouse, Université Paul Sabatier, France.

[3] INRA, UMR O77PaVé, 42 rue Georges Morel, BP 60057, 49071 BEAUCOUZE Cedex, France.

[4] Université d'Angers, UMR PaVé, 42 rue Georges Morel, BP 60057, 49071 BEAUCOUZE Cedex, France.

[5] Département de Sciences Biologiques, Agrocampus Ouest - centre d'Angers, Institut National d'Horticulture et de Paysage, UMR 077 PaVé, 42 rue Georges Morel, 49071 Beaucouzé cedex, France

[Φ] To whom correspondence should be addressed. E-mail: arlat@toulouse.inra.fr

* These authors contributed equally to this work.

[Ψa] Current address: Department of Plant-Microbe Interactions, Max-Planck Institute for Plant Breeding Research, Cologne, Germany

[Ψb] Current adress : Laboratory of Molecular and Cellular Biology, National Institue of Diabetes and Digestive and Kidney Diseases (NIDDK), Bldg. 8 Room 2A21, Bethesda, MD 20892, USA.

ABSTRACT

Xylan is a major structural component of plant cell wall and the second most abundant polysaccharide in nature. Several members of genus *Xanthomonas*, which cause disease in economically important crops, are able to degrade xylan as part of their pathogenesis. Here, by combining genomic and functional analyses, we define the core xylan utilization system of *Xanthomonas* species. This system comprises enzymes for xylan deconstruction, xylose and glucuronic catabolism and transporters. The extracellular xylanolytic activity of *Xanthomonas campestris* pv. *campestris* (*Xcc*) is mediated by Xyn10A xylanase but also by Agu67 putative glucuronidase and XypB, an inner membrane transporter of xylo-oligosaccharides. Genes of *Xcc* system are specifically induced by xylo-oligosaccharides and xylan. Their regulation involves the LacI repressor XylR as well as HrpG/X master regulators of *Xcc* type III secretion system. A particularity of this system is the presence of two TonB-dependent outer membrane transporters (TBDTs) which belong to loci required for efficient growth on plant leaves. Interestingly, part of *Xcc* xylanolytic system is particularly well conserved to the recently characterized xylan utilization system shared by rumen and human colonic Bacteroidetes which also comprises TBDTs. In fact, comparative genomic studies revealed that the association between TBDTs and xylan utilization is a feature conserved among plant cell wall degrading bacteria belonging to Proteobacteria, Bacteroidetes, Verrucomicrobia and Acidobacteria phyla and having very diverse niches such as animals gut or rumen, wood-boring animals, plants, soil or aquatic environments. This analysis defined a set of highly conserved genes among these bacteria that represent a core cluster for xylan utilization. Contrary to this subset of conserved genes, xylan-associated TBDTs are specific to each phylum suggesting a convergent evolution to associate TBDTs with xylan utilization. This association and its wide distribution among Gram negative bacteria certainly underscore the importance of TBDTs in xylan utilization.

Interestingly, xylanases required for *Xanthomonas* spp. virulence are not included in the conserved subset of proteins shared with non-pathogenic bacteria. They are rather specifically conserved among plant pathogenic bacteria suggestive of a specific role in pathogenesis.

AUTHOR SUMMARY

The Plant cell wall is composed of interconnected complex polysaccharides. Many plant pathogens secrete enzymes to breach this defense barrier and to use it as source of nutrients. In this study, we characterized the machinery used by phytopathogenic *Xanthomonas* species to degrade and utilize xylan, a major component of plant cell wall. This system is required for full pathogencity and optimal growth on leaves. It is characterized by the presence of specific transporters which might be involved in the active uptake of complex hydrolysis products of xylan, thus minimizing the release of monosaccharides that could be used by competitors. Morevover, xylan accounts for approximately one third of all renewable organic carbon on earth and beside plant pathogens, many microorganisms capable of degrading and exploiting this compound are found either free in nature or as part of the rumen or gut of humans and animals, where they are critical for gastrointestinal health and function. Bioconversion of xylan by these microorganisms has been intensively studied in the past decade because of its potential applications in agro-industrial processes, such as pulp and paper industry and biofuel production. Here, we show that the xylanolytic machinery and the strategy used by *Xanthomonas* spp. are conserved in a wide range of bacteria leaving in soil, aquatic systems or digestive tracts. Therefore, our work on *Xanthomonas* spp. may have an impact in industrial applications. Finally, despite the wide conservation of xylanolytic systems, xylanases produced by *Xanthomonas* spp., seem to be specific to phytopathogenic bacteria, thus suggesting a specific role in pathogenesis.

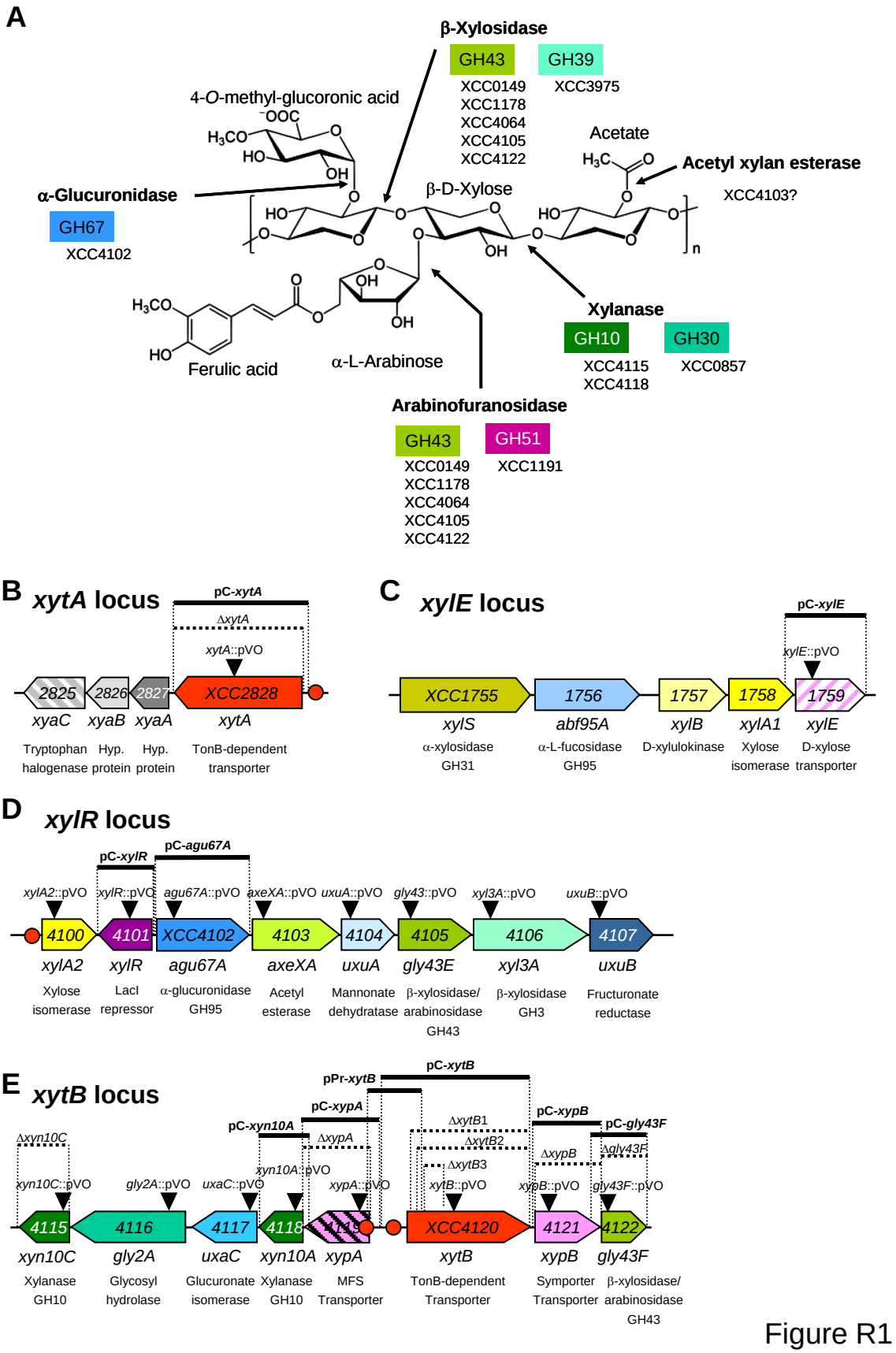

A
β-Xylosidase
GH43
XCC0149
XCC1178
XCC4064
XCC4105
XCC4122
GH39
XCC3975
4-O-methyl-glucoronic acid
Acetate
Acetyl xylan esterase
XCC4103?
α-Glucuronidase
GH67
XCC4102
β-D-Xylose
Xylanase
GH10
XCC4115
XCC4118
GH30
XCC0857
Ferulic acid
α-L-Arabinose
Arabinofuranosidase
GH43
XCC0149
XCC1178
XCC4064
XCC4105
XCC4122
GH51
XCC1191
B xytA locus
pC-xytA
ΔxytA
xytA::pVO
2825
2826
2827
XCC2828
xyaC
xyaB
xyaA
xytA
Tryptophan halogenase
Hyp. protein
Hyp. protein
TonB-dependent transporter
C xylE locus
pC-xylE
xylE::pVO
XCC1755
1756
1757
1758
1759
xylS
abf95A
xylB
xylA1
xylE
α-xylosidase GH31
α-L-fucosidase GH95
D-xylulokinase
Xylose isomerase
D-xylose transporter
D xylR locus
pC-xylR
pC-agu67A
xylA2::pVO
xylR::pVO
agu67A::pVO
axeXA::pVO
uxuA::pVO
gly43::pVO
xyl3A::pVO
uxuB::pVO
4100
4101
XCC4102
4103
4104
4105
4106
4107
xylA2
xylR
agu67A
axeXA
uxuA
gly43E
xyl3A
uxuB
Xylose isomerase
LacI repressor
α-glucuronidase GH95
Acetyl esterase
Mannonate dehydratase
β-xylosidase/ arabinosidase GH43
β-xylosidase GH3
Fructuronate reductase
E xytB locus
pPr-xytB
pC-xytB
pC-xypA
pC-xyn10A
pC-xypB
pC-gly43F
Δxyn10C
ΔxypA
ΔxytB1
ΔxytB2
ΔxytB3
ΔxypB
Δgly43F
xyn10C::pVO
gly2A::pVO
uxaC::pVO
xyn10A::pVO
xypA::pVO
xytB::pVO
xypB::pVO
gly43F::pVO
4115
4116
4117
4118
4119
XCC4120
4121
4122
xyn10C
gly2A
uxaC
xyn10A
xypA
xytB
xypB
gly43F
Xylanase GH10
Glycosyl hydrolase
Glucuronate isomerase
Xylanase GH10
MFS Transporter
TonB-dependent Transporter
Symporter Transporter
β-xylosidase/ arabinosidase GH43

Figure R1

INTRODUCTION

Plant cells are surrounded by an exoskeleton, the plant cell wall, which is composed of a highly integrated network of polysaccharides including cellulose, hemicelluloses and pectin [1,2]. This complex exoskeleton is not only important for structuring plant body but also plays a key role in growth, cell differentiation and defence. It constitutes an important obstacle that plant pathogens first encounter during their infectious process, but it also represents a substantial source of nutriment. Plant pathogens secrete numerous cell wall degrading enzymes to breach the plant cell wall barrier and to release nutrients.

Xylans are the predominant hemicelluloses in the cell wall of terrestrial plants. They constitute a family of complex, highly branched heteropolysaccharides, which vary in structure between different plant species and cell types. They comprise a conserved backbone composed of 1,4-linked β-D-xylose residues, which is branched to varying degrees by short carbohydrate chains (Figure R1). Backbones may be substituted with glucuronic acid (glucuronoxylan) or 4-*O*-methyl-glucuronic acid (methyl-glucuronoxylan), arabinose (arabinoxylan) or a combination of both types of substitutions (glucuronoarabinoxylan). The abundance and linkage types of these substitutions vary between xylans from different sources. Glucuronoxylans are major components of the secondary walls of dicots whereas glucuronoarabinoxylans are predominant hemicelluloses in the primary and secondary cell walls of monocots. In grass, some L-arabinofuranose residues may be esterified with phenolic substitutions of *p*-coumaric acid or ferulic acid (For review, see [3,4]). Altogether, xylans are the second most abundant plant polysaccharide in nature after cellulose, accounting for approximately one third of all renewable organic carbon on earth and beside plant pathogens, many microorganisms are able to degrade this hemicellulolytic substrate [5-7]. These xylanolytic microbes can be found in diverse ecological niches, which typically comprise environments where plant material accumulate and deteriorate, including plant debris, soil, aquatic environments and the digestive tract of animal, such as the human gut and the rumen of ruminants [8]. Bioconversion of xylans has been intensively studied in the past decade because of its potential applications in agro-industrial processes, such as pulp and paper industry and biofuel production. These studies have shown that depolymerisation of xylan requires the coordinated action of numerous enzymes, including endoxylanases, *β*-xylosidases, *α*-L-arabinofuranosidases, *α*-glucuronidases, acetylxylan esterases and ferulic acid esterases which belong to complex xylanolytic systems [5-7,9,10]. Although the xylanolytic systems of bacteria isolated from soil or from digestive tracts of animals have

been studied in detail, there is only limited information regarding the xylanolytic systems of plant pathogenic bacteria.

Xylanases have been shown to control the virulence of two members of the genus *Xanthomonas, Xanthomonas oryzae* pv. *oryzae* (*Xoo*), the causal agent of bacterial blight of rice, and *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) (also designated *Xanthomonas euvesicatoria*), the etiologic agent of bacterial spot disease in pepper and tomato [11,12]. The *Xanthomonas* genus comprises an important group of plant pathogenic bacteria which together affect about 124 monocotyledonous and 268 dicotyledonous plants, including agronomic important crops, such as rice (*Oriza sativa*), tomato (*Solanum lycopersicum*), pepper (*Capsicum annuum*), cabbage (*Brassica oleracea*) and Citrus species [13,14]. These pathogens have a major economical impact in regions with a warm and humid climate [13]. They produce a wide range of plant cell wall degrading enzymes (CWDEs), such as endoglucanases, polygalacturonases, pectinases and xylanases which are required for full pathogenicity on host plants [11-13,15-20]. Generally, mutation in individual CWDE genes only partially affects pathogenicity whereas mutations in multiple CWDE genes give more pronounced reduction in virulence, suggesting functional redundancies among CWDE [11,21]. Most CWDEs are secreted by the Xps type II secretion system (T2SS) which is conserved in *Xanthomonas* spp. [12,13]. *xps* mutants, of *Xcv*, *Xoo* and *Xanthomonas campestris* pv. *campestris* (*Xcc*), the causal agent of black rot disease of brassicas, are unable to provoke any disease symptoms on host plants, thus confirming that T2SS-substrates play a major role in pathogenicity [12,13]. Moreover, it appears that there is a functional interplay between Xps T2SS-substrates and the Hrp type III secretion system (T3SS) [21], which is another key virulence factor of most *Xanthomonas* species. Hrp T3SSs are injection devices which allow the delivery of effector proteins into plant cells where they play a role in defence suppression or nutritional gain [13,14,22,23]. Indeed, it was recently shown that a functional T2SS is required for optimal translocation of T3SS effectors of *Xcv* and it was speculated that CWDEs may facilitate the assembly of extracellular components of the T3SS [12]. This suggests a cooperative functionality between the two secretion machineries. This hypothesis is reinforced by the fact that several genes coding for T2SS substrates and Xps components are positively or negatively regulated by HrpG and/or HrpX, the key regulators of T3SS genes in *Xanthomonas* spp. [12,19,24]. Beside this sophisticated role, the xylan degradation process also generates various-hydrolysis products, such as xylo-oligosaccharides or decorated xylo-oligosaccharides which represent a potential source of nutriments that might be used by

Table S1. Occurrence of *xyl*-box motif upstream of *Xanthomonas* spp. and *Pseudoxanthomonas suwonensis* genes

Gene ID (Name)	Putative function	*xyl*-box motif sequence	Distance from start codon (bases)
***Xanthomonas campestris* pv *campestris* ATCC33913**			
XCC2828 (*XytA*)	TonB-dependent transporter	**TG**T**TAGCGCTA**T**CA**	282
XCC4100 (*xylA*2)	Xylose isomerase	**TG**G**TAGCGCTAACA**	122
XCC4120[a] (*xytB*)	TonB-dependent trnasporter	**TG**G**TAGCGCTAACA**	138
XCC4119 (*xypA*)	Major Facilitator Superfamily transporter	**TG**G**TAGCGCTA**T**CA**	-1
***Xanthomonas citri* pv *citri* 306**			
XAC2998	TonB-dependent transporter	**TG**T**TAGCGCTA**TAC	419
XAC4225	Xylose isomerase	**TG**G**TAGCGCTAACA**	153
XAC4256	TonB-dependent transporter	**TG**T**TAGCGCTAACA**	194
XAC4255	Major Facilitator Superfamily transporter	**TG**T**TAGCGCTAACA**	-1
XAC4194	Conserved hypothetical protein	**TG**T**TAGCGCTAACA**	260
***Xanthomonas campestris* pv *vesicatoria* 85-10**			
XCV3147[a]	TonB-dependent transporter	**TG**T**TAGCGCTA**T**CA**	411
XCV4330[a]	Xylose isomerase	**TG**G**TAGCGCTAACA**	159 in XCV4329
XCV4362	TonB-dependent transporter	**TG**T**TAGCGCTA**A**CA**	135
XCV4361	Major Facilitator Superfamily transporter	**TG**A**TAGCGCTA**C**CA**	179
***Xanthomonas fuscans* subsp *aurantifoli* ICPB11122**			
XAUC_08840	TonB-dependent transporter	**TG**T**TAGCGCTA**T**CA**	408
XAUC_34760	Xylose isomerase	**TG**G**TAGCGCTAACA**	153
XAUC_34490	TonB-dependent transporter	**TG**G**TAGCGCTAACA**	195
XAUC_34500	Major Facilitator Superfamily transporter	**TG**G**TAGCGCTA**T**CA**	-1
***Xanthomonas oryzae* pv *oryzae* MAFF311018**			
XOO_0307	Conserved hypothetical protein	**TG**A**TAGCGCTAACA**	199
XOO_1157	TonB-dependent transporter	**TG**T**TAGCGCTA**T**CA**	424
XOO_4160	Xylose isomerase	**TG**G**TAGCGCTAACA**	150
XOO_4173	Major Facilitator Superfamily transporter	**TG**G**TAGCGCTA**T**CA**	-1
XOO_4174	TonB-dependent transporter	**TG**G**TAGCGCTA**G**CA**	135
***Xanthomonas albilineans* GPE PC73**			
XALc_0037	Major Facilitator Superfamily transporter	**TG**A**TAGCGCTA**T**CA**	175
XALc_0038[a]	TonB-dependent transporter	**TG**A**TAGCGCTA**C**CA**	157
XALc_0064	Xylose isomerase	**TG**A**TAGCGCTAACA**	243
XALc_3147	TonB-dependent transporter	**TG**A**TAGCGCTA**T**CA**	258
***Pseudoxanthomonas suwonensis* 11-1**			
Psesu_0605	TonB-dependent transporter	**TG**G**TA**CC**GCTAA**CT	309
Psesu_2931	D-Xylulokinase	**TG**A**TA**CC**GCTAA**CT	204
Psesu_2909	Major Facilitator Superfamily transporter	**TG**A**TA**C**GG**T**C**C**CA**	309
Psesu_2909	Major Facilitator Superfamily transporter	**TG**G**TAGCGCT**C**TCA**	283
Psesu_2908	TonB-dependent transporter	**TG**G**GA**C**GG**T**ATCA**	132
Psesu_2908	TonB-dependent transporter	**TG**AG**A**GCGCTAC**CA**	160

[a]Start codon prediction from Blanvillain *et al*., (2007); All other start codons are from GenBank: *Xanthomonas campestris* pv. *campestris* ATCC33913: accession n° AE008922; *Xanthomonas axonopodis* pv *citri* 306: accession n° AE008923; *Xanthomonas campestris* pv *vesicatoria* 85-10: accession n° AM039952; *Xanthomonas fuscans* subsp *aurantifoli* ICPB11122: accession n° ACPX00000000; *Xanthomonas oryzae* pv *oryzae* MAFF311018: accession n° AP008229; *Xanthomonas albilineans* GPE PC73: accession n° FP565176; *Pseudoxanthomonas suwonensis* 11-1: accession n° CP002446.

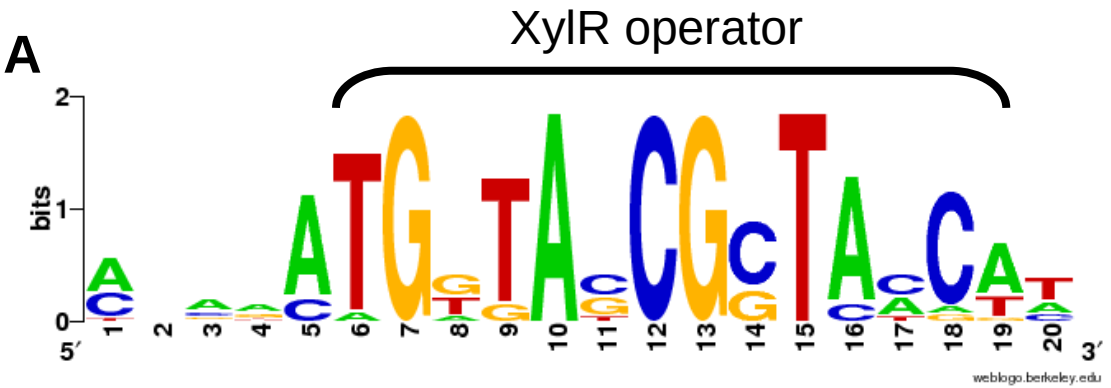
pathogens during the infection. Although the xylan deconstruction products have been characterized from different bacteria, little is known about the transport mechanisms by which these molecules enter the bacterial cell.

In *Xcc*, we recently showed that several TonB dependent transporters (TBDT), are involved in plant carbohydrate scavenging [25]. In contrast to passive transport through porins, these outer membrane transporters allow active transport which requires energy. This energy is provided by TonB-ExbB-ExbD protein complex [26,27]. TBDTs also allow uptake of larger molecules than porins and they have been shown to transport iron-siderophore complexes vitamin B12 and various carbohydrates [25,28-31]. In *Xcc*, some TBDT genes belong to CUT systems (Carbohydrate Utilization containing TBDT systems) which beside these genes comprise inner membrane transporters, degrading enzymes and/or transcriptional regulators [25,31]. The *Xcc sux* CUT locus has been studied in detail; this cluster enables *Xcc* to transport and metabolize sucrose (the main sugar present in higher plants) at very low concentrations, with a very high efficiency. The importance of such a system is highlighted by the fact that this locus is required for full virulence of the bacteria on host plants such as Cabbage or the model plant *Arabidopsis thaliana* [25].

Interestingly, we identified 2 TBDT genes (*XCC2828* and *XCC4120*) in *Xcc* genome whose expression is induced in the presence of xylose and xylan, suggesting that they might belong to a CUT system involved in the utilization of xylan [25]. The involvement of TBDTs in the transport of xylan hydrolysis products has also been suggested by two transcriptomic studies showing that TBDT genes are co-expressed with genes involved in xylose or xylan utilization in the aquatic oligotroph, *Caulobacter crescentus* CB15 and in the rumen symbiont, *Prevotella bryantii* $B_1$4 strain, respectively [10,32,33].

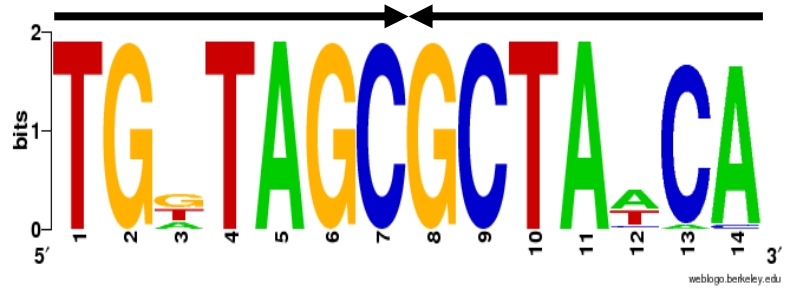
In this study, we show the existence in *Xcc* of a complex xylanolytic CUT system which comprises enzymes for the deconstruction of xylan, the utilization of glucuronic acid and the metabolism of xylose as well as TBDTs, inner membrane transporters and a regulator of the LacI family. Several genes of this system are specifically and highly induced by xylo-oligosaccharides and repressed by the LacI regulator. Moreover, one of the three endo-xylanases belonging to this system is regulated by HrpG and HrpX. Comparative genomic studies show that the presence of TBDTs is a conserved feature among xylanolytic systems of numerous Gram negative bacteria that are phylogenetically very diverse and that live in different niches such as the human or animal digestive tracts, gills of wood-boring bivalves, aquatic environment, plant debris and soil. This wide conservation underlies the importance

A



Xylose induction motif of *Caulobacter crescentus* CB15 genes

B



xyl-box motif of *Xanthomonas* spp. genes

C

xyl-box-like motif of *Pseudoxanthomonas suwonensis* genes

Figure S1

of these outer membrane transporters for the scavenging of complex plant carbohydrates by bacteria. The role of such systems in *Xanthomonas* spp. pathogenesis is discussed.

RESULTS

TBDTs genes are located in *loci* putatively involved in Xylan/xylose metabolism in *Xcc*

The global study of *Xcc* ATCC33913 (LMG568) TBDTs showed that the expression of *XCC2828* and *XCC4120* TBDT genes is specifically induced by xylan and xylose [25]. Interestingly, these two TBDT polypeptides display significant homologies with two TBDTs from *C. crescentus* strain CB15, CC_0999 (41% amino acid identity) and CC_2832 (44% amino acid identity), respectively, which gene expression is induced by xylose [32]. This induction is mediated by CC_3065, a regulator named XylR, which belongs to the LacI family [34]. XylR was shown to recognize a specific 14 bp motif which is the putative XylR operator [34](Figure S1). This motif is found upstream both CC_0999 and CC_2832 TBDT genes in *C. crescentus* CB15. It was also observed that this motif associated with xylose regulation in *C. crescentus* CB15 has several close matches in the genome of *Xcc*, two of these matching sequences being located upstream *XCC2828* and *XCC4120* TBDT genes [32]. The analysis of DNA sequences located upstream both *Xcc* TBDT genes allowed us to confirm the existence of a 14 bp palindromic motif similar to *C. crescentus* CB15 XylR operator sequence (Table S1; Figure S1).

This motif, named *xyl*-box, was used to screen *Xcc* (strain ATCC33913) genome sequence using Patscan and Predetector software [35,36]. Both software predicted two additional perfect *xyl*-box, one located upstream *XCC4119* gene which codes for a putative inner membrane transporter of the Major Facilitator Superfamily (MFS) and the other positioned upstream of *XCC4100* gene encoding a putative xylose isomerase (Table S1). Interestingly, several genes located in the region surrounding *XCC4100*, *XC4119* and *XCC4120* genes have predicted functions associated with the utilization of xylan or xylose. These observations led us to define three putative loci, associated with xylan utilization. The first locus contains *XCC2828* TBDT gene, which was named *xytA* (xyt standing for xylan TBDT). This locus was therefore called the *xytA* locus. The second locus which encompasses *xylR* regulator was named *xylR* locus and the third one which harbours *XCC4120* TBDT gene (*xytB*) was called *xytB* locus (Figure R1).

Table R1. Identification of the relevant xylanolytic-associated genes in *X. campestris* pv. *campestris* ATCC33913 genome

Locus /gene	ORF	Name	Signal peptide[a]	CAZy family	Pfam/COG/ /TIGR[b]	Annotation
***xytA* locus**						
	XCC2825	XyaC	Yes		PF04820	Putative tryptophan halogenase
	XCC2826	XyaB	No		No conserved domain	Hypothetical protein
	XCC2827	XyaA	No		PF07277	SapC-related protein
	XCC2828	XytA	Yes		PF00593-PF07715	TonB-dependent transporter
***xylR* locus**						
	XCC4100	XylA2	No		PF01261	Xylose isomerase
	XCC4101	XylR	No		PF03566-PF0532	Transcriptional regulator, LacI family
	XCC4102	Agu67A	Yes	GH67	PF03648-PF7477-PF07488	Alpha-D-glucuronidase
	XCC4103	AxeXA	Yes		PF03629	Putative acetyl esterase
	XCC4104	UxuA	no		PF02746-PF01188	Putative D-mannonate dehydratase
	XCC4105	Gly43E	Yes	GH43	PF04616	Beta-xylosidase/alpha-L-arabinofuranosidase
	XCC4106	Xyl3A	Yes	GH3	PF00933-PF01915-PF07691	Putative beta-xylosidase
	XCC4107	UxuB	No		PF01232-PF08125	Fructuronate reductase
***xytB* locus**						
	XCC4115	Xyn10C	Yes	GH10	PF00331	Putative endo-1,4-beta-xylanase
	XCC4116	Gly2A	Yes[c]	GH2	PF02836-PF02837-PF00703	Glycoside hydrolase
	XCC4117	UxaC	No		PF02614	Glucuronate isomerase
	XCC4118	Xyn10A	Yes	GH10	PF00331	Endo-1,4-beta-xylanase
	XCC4119	XypA	No		PF07690/TIGR00893	Major Facilitator Superfamily Transporter
	XCC4120	XytB	Yes		PF00593-PF07715	TonB-dependent transporter
	XCC4121	XypB	No		PF07690/COG2211/TIGR00792	Sugar:cation symporter transporter
	XCC4122	Gly43F	No	GH43	PF04616	Beta-xylosidase/alpha-L-arabinofuranosidase
***xylE* locus**						
	XCC1755	XylS	Yes	GH31	PF01055-PF07691	Alpha-xylosidase
	XCC1756	Abf95A	Yes	GH95	No conserved domain	Putative Alphal-L-fucosidase
	XCC1757	XylB	No		PF00370-PF02782	D-xylulokinase
	XCC1758	XylA1	No		PF01261	Xylose isomerase
	XCC1759	XylE	No		PF00083	D-xylose transporter
Other genes						
	XCC0149	Gly43A	Yes	GH43	PF04616/COG3507	Beta-xylosidase/alpha-L-arabinofuranosidase
	XCC0857	Xyn30A	Yes	GH30	PF02055/COG5520	Endo-1,4-beta-xylanase
	XCC1178	Gly43B	Yes[c]	GH43	PF04616/COG3507	Beta-xylosidase/alpha-L-arabinofuranosidase
	XCC1191	Abf51A	Yes[c]	GH51	PF06964/COG3534	Alpha-L-arabinofuranosidase
	XCC3975	Xyl39A	Yes	GH39	PF01229/COG3664	Putative beta-xylosidase
	XCC4064	Gly43C	Yes	GH43	PF04616/COG3507	Beta-xylosidase/alpha-L-arabinofuranosidase

[a] Signal peptide prediction using SignalP (http://www.cbs.dtu.dk/services/SignalP; [120])
[b] As determined by using the Conserved Domain Database [129] and the Pfam database [130]
[c] Start codon prediction revised in this work. All other start codons are from GenBank [58] or from Blanvillain *et al.* [25]

Characterization of the xylanolytic machinery in *xytA*, *xylR*, *xytB* loci and *Xcc* genome

The major enzymes that attack xylan are endo-1,4-*β*-D-xylanases (EC 3.2.1.8) that cleave the main backbone of xylan and generates xylo-oligosaccharides and oligosaccharides with various branched substitutions (Figure R1). These enzymes and other glycosyl hydrolases are classified in the Carbohydrate-active enzyme (CAZy) database on the basis of amino acid sequence similarities (http:://www.cazy.org/; [37]). Endo-1,4-*β*-D-xylanases are found in glycosyl hydrolase (GH) families 5, 7, 8, 10, 11, 30 and 43.

Other enzymes involved in xylan degradation are *β*-D-xylosidases (EC 3.2.1.37; GH families 3, 30, 39, 43; 52, 54, 116 and 120) that cleave xylose monomers from the non-reducing end of xylo-oligosacharides. Elimination of the side groups is catalysed by α-L-arabinofuranosidases (EC 3.2.1.55; GH families 3, 43, 51, 54 and 62), *α*-D-glucuronidases (EC 3.2.1.139; GH4, GH67 and GH115 families), acetylxylanesterase (EC 3.1.1.72) grouped in carboxyesterase (CE) families 1, 2, 3, 4, 5, 6, 7 and 12, ferulic/*p*-coumaric acid esterases (EC 3.1.1.73; CE1) (Figure R1) [7,8].

The screening of CAZy database revealed 34 *Xcc* proteins belonging to the different GH families listed above. However, several of these families include members that possess distinct enzymatic activities. Thus, all genes identified by this screening are not necessarily related to xylan degradation. Among the 34 detected genes, comparative and genome context analyses suggested that only 12 might really be involved in xylan deconstruction (Table R1), whereas the others might be implicated in xyloglucan, cellulose or pectin degradation (data not shown). Interestingly, 6 of these 12 putative "xylan-associated" genes are located in *xylR* and *xytB* loci, whereas the other 6 are scattered in the genome (Table R1; Figure R1).

***Xcc* possesses 3 genes coding for putative endo-xylanases**

Among the 12 genes identified above, two genes, *XCC4115* and *XCC4118* which are located in *xytB* locus downstream of *XCC4119* code for putative endo-1,4-*β*-xylanases of family GH10. These two genes were named *xyn10C (XCC4115*) and *xyn10A (XCC4118*) to indicate their activity and CAZy family, as previously described for *Cellvibrio japonicus* [38] and according to the nomenclature recently proposed by Potnis *et al*. [39]. These two proteins display significant similarities to functional GH10 xylanases (Table S2). A third gene, *XCC0857* coding for another putative endo-xylanase which belongs to family GH30, was

Table S2. Significative homology of the putative ORFs from *X. campestris* pv. *campestris* xylan CUT system.

ORF	Name	Protein size (aa) /Signal peptide[a]	Representative homologous protein (species/accession n°)[b]	Identity (%)/amino acid overlap	Protein size (aa) /Signal peptide[a]	Homologous gene product (reference)
***xytA* locus**						
XCC2825	XyaC	498/Yes	PyrH (*Streptomyces rugosporus*/AAU95674)	34/491	519/No	Tryptophan halogenase [132]
XCC2826	XyaB	343/No	Pass1 (*Rattus Norvegicus* /Q5BKC6)	39/126	479/No	Associated with Hsp27 [133]
XCC2827	XyaA	313/No	PHZ_c2924 (*Phenylobacterium zucineum* HLK1 /YP_002131762.1)	48/231	238/No	SapC-related protein
XCC2828	XytA	1047/Yes	Patl_3278 (*Pseudoalteromonas atlantica* T6c/YP_662838.1)	42/999	1006/Yes	TonB-dependent transporter
***xylR* locus**						
XCC4100	XylA2	446/No	XylA (*Piromyces* sp. E2 /CAB76571)	61/435	437/No	Xylose isomerase [134]
XCC4101	XylR	366/No	XylR (*C. crescentus CB15* /NP421859)	50/351	351/No	LacI family repressor [34]
XCC4102	Agu67A	739/Yes	GlcA67A (*Cellvibrio japonicus* /AAL5772)	55/719	732/Yes	Alpha-D-glucuronidase [105]
XCC4103	AxeXA	654/Yes	SiaE (*Mus musculus* /CAA67214)	32/202 32/262	541/Yes	Sialic-acid-specific 9-*O*-acetylesterase [49]
XCC4104	UxuA	419/No	ManD (*Novosphingobium aromaticivorans*/2QJJ_A)	71/401	402/No	D-mannonate dehydratase [47]
XCC4105	Gly43E	565/Yes	XylB (*Butyvibrio fibrisolvens*/P45982)	38/509	517/No	Bi functional Beta-xylosidase-alpha-L-arabinofuranosidase [135]
XCC4106	Xyl3A	896/Yes	Xyl3C (*P. Bryantii* B_14 /ADD92016)	42/821	857/Yes	Beta-D-xylosidase [42]
XCC4107	UxuB	487/No	UxuB (*Escherichia coli* K12 /BAA02591)	41/457	486/No	Mannonate oxydoreductase/Fructuronate reductase [136]
***xytB* locus**						
XCC4115	Xyn10C	385/Yes	Xyn10A (*Bacteroides xylanisolvens* XB1A/CBH32823	45/359	378/Yes	Endo-1,4-beta-xylanase [99]
XCC4116	Gly2A	900/Yes	OTER_3378 (*Opitutus terrae* PB90-1/ACB76655)	68/877	919/Yes	Putative glycoside hydrolase
XCC4117	UxaC	471/No	UxaC (*Geobacillus stearothermophilus* T6 /ABI49945)	25/453	473/No	Glucuronate isomerase [43]
XCC4118	Xyn10A	330/yes	XynB, CJA3280 (*C. japonicus* Ueda107/P23030)	39/282	599/Yes	Endo-1,4-beta-xylanase [137]
XCC4119	XypA	501/No	ExuT (*Ralstonia solanaceaum* /AAL24034)	34/389	439/No	Galacturonate transporter [138]
XCC4120	XytB	980/Yes	Caul_1838 (*Caulobacter* sp. K31/ABZ70967)	51/950	979/Yes	TonB-dependent transporter
XCC4121	XypB	495/No	GusB (*E. coli* /YP_001458395.1)	29/444	457/no	Glucuronide permease [139]
XCC4122	Gly43F	344/No	Xsa (*Bacteroides ovatus* V975 /P49943)	57/313	325/no	Bifunctional Beta-xylosidase/alpha-L-arabinofuranosidase [40]
			XynB PBR0394 (*P. bryantii* B_14 /CAA89208)	54/309	319/no	Exoxylanase [41]
***xylE* locus**						
XCC1755	XylS	967/No	XylS (*Sulfolobus solfataricus* /Q9P999)	39/588	731	Alpha-xylosidase [140]
XCC1756	Abf95A	790/Yes	Llfuc (*Lilium longiflorum* /BAF85832)	38/730	854/No	Alpha-1,2-L-fucosidase [141]
XCC1757	XylB	497/No	XylB (*Piromyces* sp. E2/CAB76752)	45/494	494/No	D-xylulokinase [134]
XCC1758	XylA1	446/No	XylA (*Piromyces* sp. E2 /CAB76571)	61/435	437/No	Xylose isomerase [134]
XCC1759	XylE	481/No	GlcP (*Synechocystis* PCC6803 /P15729.2)	52/455	468/No	Glucose/fructose transporter [142]
Other genes						
XCC0149	Gly43A	526/Yes	XynB (*Paenibacillus* sp. JDR-2 /ABV90487)	31/427	521/No	Putative Beta-xylosidase/alpha-L-arabinofuranosidase [144]
XCC0857	Xyn30A	405/Yes	XynC (*Erwinia chrysanthemi* /AAB53151)	57/397	413/Yes	Endo-1,4-beta-xylanase [143]
XCC1178	Gly43B	549/Yes	XynB (*Paenibacillus* sp. JDR-2 /ABV90487)	33/515	521/No	Putative Beta-xylosidase/alpha-L-arabinofuranosidase in aldouronate utilization cluster [144]
XCC1191	Abf51A	508/Yes	Abf51A CJA_2769 (*Cellvibrio japonicus*/AAK84947)	53/508	517/Yes	Alpha-L-arabinofuranosidase [44]
XCC3975	Xyl39A	521/Yes	XynB1 (*Geobacillus stearothermophilus* T /ABI49941)	36/483	504/No	Beta-xylosidase [43]
XCC4064	Gly43C	544/Yes	XynB (*Paenibacillus* sp. JDR-2 /ABV90487)	29/515	521/No	Putative Beta-xylosidase/alpha-L-arabinofuranosidase [144]

[a] Signal peptide prediction using SignalP (http://www.cbs.dtu.dk/services/SignalP; [120])

[b] The reported homologous proteins are those showing the highest score among proteins with an experimentally defined function. In the absence of relevant biochemical data, the most similar protein from bacteria outside the Xanthomonadaceae family was reported.

detected outside *xytA, xylR* and *xytB* loci. The polypeptide sequence of this xylanase aligns perfectly with XynC xylanase (XCV0965) (93% amino acid identity) which was recently characterized in *Xcv* [12]. This third xylanase was named Xyn30A to follow the nomenclatures proposed by Potnis *et al*. [39]. The three xylanases possess signal peptides suggesting that they are exported in the periplasm or secreted in the extracellular medium.

Xcc harbors a complex repertoire of β-D-xylosidases and/or α-L-arabinosidases

The genome of *Xcc* contains 5 genes belonging to GH43 family which might code for putative *β*-D-xylosidases and/or *α*-L-arabinofuranosidases involved in xylan deconstruction (Table R1; Table S2). Two of these genes, *XCC4105* and *XCC4122*, named *gly43E* and *gly43F*, respectively, belong to *xylR* and *xytB* loci (Figure R1). The three other genes, *XCC0149* (*gly43A*), *XCC1178* (*gly43B*), and *XCC4064* (*gly43D*) are scattered in the genome. Gly43F is the only protein of this family which has no signal peptide in *Xcc*. Interestingly, this putative enzyme displays significant homology to Xsa (57% amino acid identity) from *Bacteroides ovatus* V975 [40] and to XynD (54% amino acid identity) from *Prevotella bryantii* B_14 [41] (Table S2). These two proteins do not possess signal peptides either and XynD was shown to be a cytoplasmic exoxylanase which release xylose from substrates including xylobiose, xylopentaose, and birch wood xylan [41].

Two other putative *β*-xylosidases belonging to different GH families are also present in *Xcc* genome. The first one encoded by *XCC4106* gene maps in the *xylR* locus. This putative enzyme belongs to family GH3 and was named Xyl3A. It is notably similar to Xyl3C β-xylosidase which was recently functionally characterized in *P. bryantii* B_14 (Table S2). This enzyme displays an unusual specificity and is able to cleave xylose from xylan hydrolysis products such as aldouronic acids (methyl-D-glucuronoxylo-oligosaccharides) substituted with 4-*O*-methyl-glucuronic acid [42]. The second putative β-xylosidase, XCC3975 (Xyl39A) is grouped in family GH39. Its polypeptide sequence displays significant similarity over its all length to XynB1 (Table S2), a GH39 *β*-xylosidase which belongs to a large gene xylan utilization cluster from the Gram positive bacterium *Geobacillus stearothermophilus* [43].

Finally, *Xcc* genome exhibits a putative *α*-L-arabinosidase of family GH51. This putative enzyme, named Abf51A, aligns well to Abf51A *α*-L-arabinosidase from *Cellvibrio japonicus* (Table S2), which activity was studied in detail [44]. It also displays significant similarity to

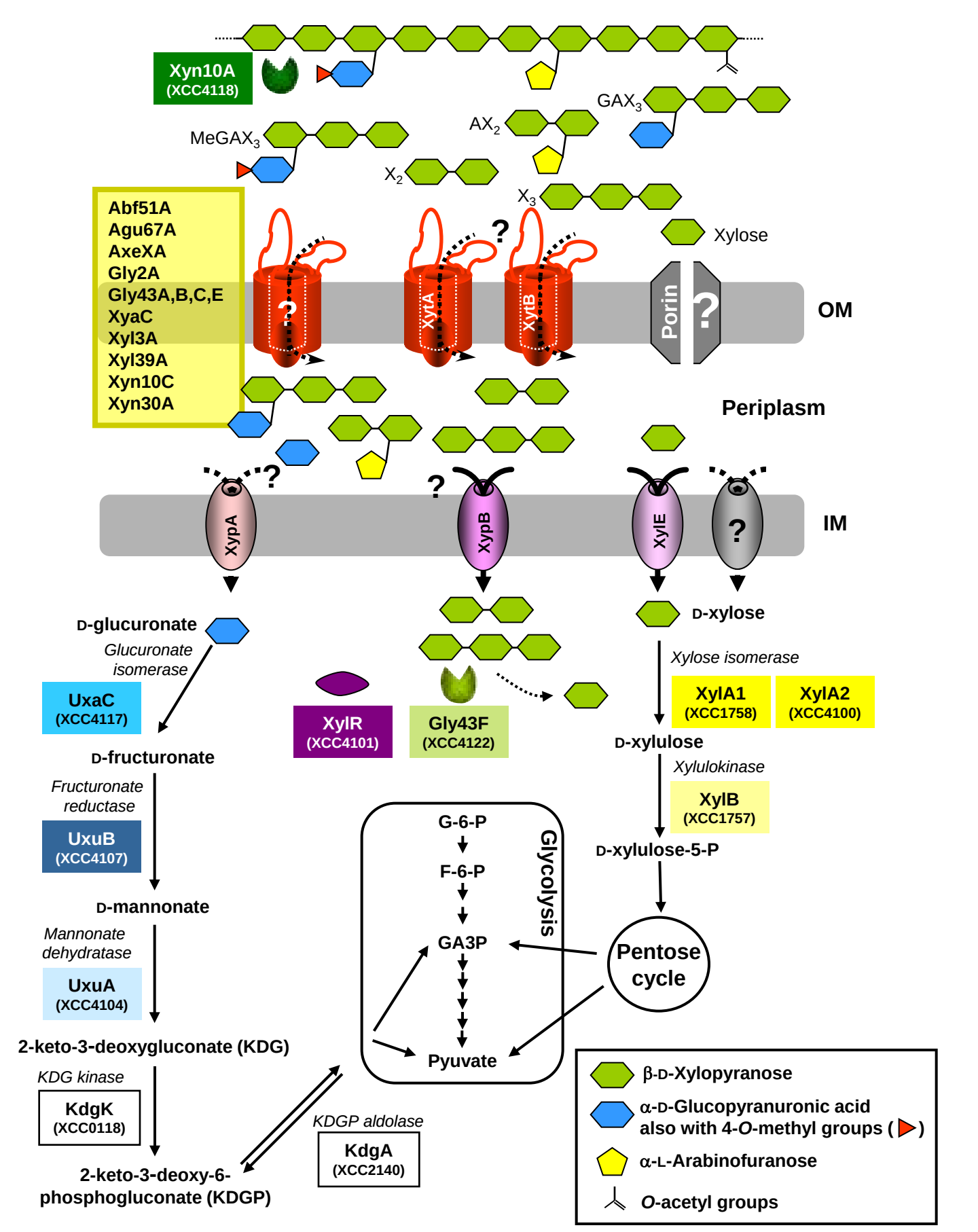

Xyn10A
(XCC4118)
GAX3
AX2
MeGAX3
X2
X3
Xylose
Abf51A
Agu67A
AxeXA
Gly2A
Gly43A,B,C,E
XyaC
Xyl3A
Xyl39A
Xyn10C
Xyn30A
?
XytA
XytB
Porin
OM
Periplasm
XypA
XypB
XylE
IM
D-glucuronate
Glucuronate isomerase
UxaC
(XCC4117)
D-fructuronate
Fructuronate reductase
UxuB
(XCC4107)
D-mannonate
Mannonate dehydratase
UxuA
(XCC4104)
2-keto-3-deoxygluconate (KDG)
KDG kinase
KdgK
(XCC0118)
2-keto-3-deoxy-6-phosphogluconate (KDGP)
KDGP aldolase
KdgA
(XCC2140)
XylR
(XCC4101)
Gly43F
(XCC4122)
D-xylose
Xylose isomerase
XylA1
(XCC1758)
XylA2
(XCC4100)
D-xylulose
Xylulokinase
XylB
(XCC1757)
D-xylulose-5-P
G-6-P
F-6-P
GA3P
Pyuvate
Glycolysis
Pentose cycle
β-D-Xylopyranose
α-D-Glucopyranuronic acid also with 4-O-methyl groups ( )
α-L-Arabinofuranose
O-acetyl groups

Figure R2

AbfB (54% amino acid identity) which is encoded in the xylan utilization cluster of *G. stearothermophilus* [43].

xytB and *xylR* loci harbour a putative α-glucuronidase and glucuronic acid utilization pathway

xylR locus encompasses *XCC4102* (*agu67A*) gene which codes for a protein listed in family GH67 (Table R1). This protein displays very high similarity to GlyA67A α-glucuronidase characterized from *C. japonicus* [45] (Table S2). Like other members of this family, GlyA67A hydrolyzes the α1,2-glycosidic bond between 4-*O*-methyl-D-glucuronic acid and the xylose moiety located at the non-reducing end of xylo-oligosaccharides (Figure R1). Therefore this suggests that α-D-glucuronic acid might be released during the degradation of xylan by *Xcc*. Interestingly, three genes, *XCC4117* (*uxaC*), *XCC4107* (*uxuB*) and *XCC4104* (*uxuA*) located in *xylR* and *xytB* loci code for putative enzymes likely involved in the utilization of this molecule (Table R1, Table S2, Figure R2). XCC4104 (UxuA) putative mannonate dehydratase does not display any homology with "canonical" mannonate dehydratases such as UxuA from *Escherichia coli* [46]. However, it has 71% identity (401-amino-acid overlap) to ManD, a mannonate dehydratase that was recently characterized from *Novosphingobium aromaticivorans* [47]. Moreover, putative enzymes involved in the utilization of 2-keto-3-deoxygluconate which might be generated by the activity of UxaC, UxuB and UxuA are present in *Xcc* genome (Figure R2).

Characterization of *Xcc* xylose utilization pathway

In addition to genes involved in glucuronate metabolism, *xylR* locus contains a gene which might be involved in xylose metabolism. This gene, *XCC4100* (*xylA2*), has a *xyl*-box motif in its promoter region and codes for a protein displaying specific features of xylose isomerase (Table R1, Figure R1 and 2). These data suggest a coupling between xylan and xylose metabolism in *Xcc*. Interestingly, *xylA2* gene seems to be duplicated in *Xcc* genome, since a gene, named *xylA1* (*XCC1758*) codes for a protein which displays 98% amino acid identity to XylA2 (*xylA1* and *xylA2* DNA coding sequences are 97% identical). In *Xcc* genome, *xylA1* gene co-localizes with other genes putatively involved in xylose utilization (Figure R1 and 2)[48]. *xylA1*, is located downstream of *xylB* which codes for a putative D-xylulokinase gene and upstream of *XCC1759* (*xylE*) gene which codes for a MFS inner membrane transporter. XCC1759 aligns well with XylE (CC_0814) (61% amino acid identity) from *C. crescentus*, a

putative xylose transporter which gene expression is regulated by XylR and induced by xylose [32,34]. The locus encompassing *xylA1*, *xylB* and *xylE* genes was named *xylE* locus and might be required for xylose entry and utilization (Figures 1 and 2). This locus is preceded by two genes, *XCC1755* and *XCC1756* which codes for a putative *α*-xylosidase and a putative *α*-L-fucosidase, respectively, which might be involved in xyloglucan deconstruction (Table S2, data not shown). We did not find any perfect or even degenerated *xyl*-box in the DNA region encompassing *XCC1750* to *XCC1759* (*xylE*) genes.

Other miscellaneous enzymatic activities encoded by *xytA*, *xytB* and *xylR* loci

The *xylR locus* harbours *XCC4103* gene which codes for a protein which displays a signal peptide and some similarity with two regions of SiaE, a sialic-acid-specific 9-*O*-acetylesterase characterized in *Mus musculus* [49]. We therefore speculate that this gene codes for a putative acetyl esterase, which is not listed in the CAZy database, and which might be involved in the removal of acetyl groups from the backbone of xylan (Table S2, Figure R1; Figure R2). This gene was named *axeXA.*

The *xytB* locus includes a putative glycoside hydrolase, XCC4116 (Gly2A), listed in family GH2 (Table R1). This family contains enzymes with different *β*-glycosidase activities including *β*-galactosidase, *β*-mannosidase, *β*-glucuronidase, mannosylglycoprotein endo-*β*-mannosidase or exo-*β*-glucosaminidase enzymes. To our knowledge, implication of members of this family in xylan utilization has not been reported yet. It is therefore difficult to place this protein in the *Xcc* xylan degradation pathway of *Xcc*. The only information concerning this protein comes from a systematic study of *Xcc* putative *β*-galactosidases which included members of family GH2. This work showed that XCC4116 ortholog in *Xcc* strain Xc17 has no *β*-galactosidase activity [50].

Finally, the genes surrounding *xytA* TBDT gene (*XCC2828*) did not reveal any features suggesting that they might code for proteins involved in xylan or xylose metabolism. However, the 3 genes, *xyaA*, *xyaB* and *xyaC*, located downstream of *xytA* and having the same transcriptional orientation display significant homologies with 3 genes, *CC_1000*, *CC_1001*, *CC_1002*, respectively, from *C. crescentus* CB15 (Table S2). In this bacterium, these three genes are located downstream of *CC_0999* TBDT gene which is preceded by a XylR motif and is induced by xylose (Table S2). *CC_1000* gene is also induced by xylose [32]. The role

Table R2. Xylanase activity of *Xanthomonas campestris* pv *campestris* strains

Strain	Xylanase activity[a] (SD)[b]
Xcc-568 (Wild type)	3.36 (0.43)
Putative xylanase mutants	
xyn30A::pVO	3.91 (0.85)
Δ*xyn10C*	4.23 (0.65)
Δ*xyn10A*	0 (0.00)
Δ*Xyn10A*/pC-*xyn10A*	7.39 (1.19)
WT/pC-*xyn10A*	8.22 (1.33)
Other xylan degradation associated mutants	
agu67A::pVO	1.05 (0.07)
agu67A::pVO/pCZ1016[c]	1.05 (0.07)
agu67A::pVO/pC-*agu67A*	3.65 (0.85)
Δgly43F	13.70 (1.30)
Δgly43F/pC-*gly43F*	1.60 (0.79)
Inner membrane transporter mutants	
Δ*xypA*	3.74 (1.06)
Δ*xypB*	0.61 (0.54)
Δ*xypB*/pC-*xypB*	4.13 (0.35)
xypB ::pVO	0.18 (0.25)
xypB ::pVO/ pC-*xypB*	9.05 (1.48)
xylE::pVO	3.68 (1.78)
TonB-dependent transporter mutants	
Δ*xytA*	4.93 (1.81)
Δ*xytB*	3.67 (0.61)
Δ*xytB*Δ*xytB*	4.56 (0.60)
Regulatory mutants	
xylR::pVO	11.57 (3.01)
xylR::pVO/pC-*xylR*	2.48 (1.23)
Xcc-568/pC-*xylR*	2.08 (0.72)
hrpX::pVO	10.13 (0.64)
Δ*hrpG*	10.70 (1.41)
hrpX::pVO Δ*xyn10A*	0 (0.00)
hrpX::pVO Δ*xyn10C*	11.83 (1.85)
hrpX::pVO Δ*xyn30A*	9.79 (1.51)

[a] Xylanase activity is given by calculating the $(H^2-C^2)/C^2$ ratio, where H is the diameter of the halo and C the diameter of the bacterial colony, measured 4 days after spotting.
[b] Standard deviation (SD) were calculated from three independent experiments.
[c] pCZ1016 is the empty expression vector that was used to perform complementation experiments.

of these three news genes in xylan metabolism could not be deduced from the analysis of their protein sequence.

The *xytB* locus contains two inner membrane transporters

xytB TBDT gene (*XCC4120*) is flanked by two inner membrane transporter genes, *XCC4119 and XCC4121,* renamed *xypA* and *xypB*, respectively (for xylan utilization plasma membrane transporters) (Figures 1 and 2). *xypA* codes for a transporter of the Major Family Superfamily (MFS) and displays homology to hexuronate transporters (Table R1; Table S2). XypB belongs to the sugar-cation symporter family (Table R1; Table S2).

Altogether these observations suggested the existence of a putative CUT system which comprises 4 major *loci* which contain enzymes for the deconstruction of xylan, the utilization of glucuronic acid and xylose as well as outer and inner membrane transporters. This Xylan CUT system also harbours a regulator of the LacI family. A model integrating all these information is shown in Figure R2.

Xyn10B, Agu67A, Gly43F and XypB control the production of extracellular xylanase activity

To study the biological relevance of the xylan CUT system, we tested the xylanolytic capability of *Xcc* ATCC33913 wild type strain. This strain produces an extracellular xylanase activity on minimal medium (MME) agar plates containing RBB xylan (Table R3).

Most genes belonging to *xytA*, *xylR* and *xytB* loci were mutated by insertion of the suicide plasmid pVO155 [51](Figure R1). To avoid polar effects of pVO155 insertion, we also constructed deletion (Δ) mutants of a subset of these genes. All these mutants were tested for the production of extracellular xylanase activity. Most of them have xylanase activities similar to that of the wild type strain (data not shown). However some mutants were significantly affected. No activity was detected for Δ*xyn10A* mutant, which carries a deletion of *xyn10A* xylanase gene (Table R2). Complementation experiments conducted by introducing pC-*xyn10A* plasmid, expressing *xyn10A*, into Δ*xyn10A* mutant confirmed that the extracellular activity detected in these conditions is coded by *xyn10A* gene (Table R2). Accordingly, we did not observe any significant reduction in extracellular xylanase activity in Δ*xyn10C* or *xyn30A*::pVO mutants affected in the two other putative xylanase genes of *Xcc* (Table R2). The level of extracellular xylanase activity was also reduced in two other mutants. It was

Table R3. Regulation of genes in the xylan degradation system by XylR.and xylan

Locus	Gene ID[a]	Name	orientation[a]	Expression ratios (SD[d])	
				Δ*xylR* mutant in MME/ Wild type in MME	MME with Xylan (0.125%) /MME
xytA locus	*XCC2825*	*xyaA*	R	5.99[e] (1.21)	nd[c]
	XCC2826	*xyaB*	R	3.98[e] (1.79)	nd
	XCC2827	*xyaC*	R	110.33[e] (31.91)	nd
	XCC2828[b]	*xytA*	R	685.25[e] (142.12)	20.85[f] (2.02)
xylR locus	*XCC4100*[b]	*xylA2*	F	7.16[e] (2.60)	1.36[f] (0.07)
	XCC4101	$xylR_{XCC}$	R	1.57[e] (0.60)	0.81[f] (0.02)
	XCC4102	*agu67A*	F	1.31[e] (0.24)	0.95[f] (0.09)
	XCC4103	*axeXA*	F	nd[c]	1.71[f] (0.04)
	XCC4104	*uxuA*	F	0.78[e] (0.19)	1.56[f] (0.21)
	XCC4105	*gly43E*	F	0.62[e] (0.06)	4.80[f] (0.04)
	XCC4106	*xyl3A*	F	1.00[e] (0.57)	3.76[f] (0.002)
	XCC4107	*uxuB*	F	nd	4.08[f] (0.02)
xytB locus	*XCC4115*	*xyn10A*	R	10.28[e] (0.75)	11.68[f] (1.31)
	XCC4116	*Gly2A*	R	61.8[e] (12.12)	12.90[f] (0.26)
	XCC4117	*uxaC*	R	129.34[e] (21.81)	11.58[f] (0.85)
	XCC4118	*xyn10B*	R	54.98[e] (12.15)	5.78[f] (1.23)
	XCC4119[b]	*xypA*	R	20.09[e] (2,27)	23.47[f] (0.4)
	XCC4120[b]	*xytB*	F	358.6[e] (74.96)	39.56[f] (5.31)
	XCC4121	*xypB*	F	85.25[e] (5.95)	1.42[f] (0.06)
	XCC4122	*gly43F*	F	47.94[e] (3.79)	48.51[f] (2.02)

[a] Gene ID and transcriptional orientation are from *Xanthomonas campestris* pv *campestris* strain ATCC33913 [58]. F, forward ; R, reverse.
[b] contains a *xyl*-box motif upstream
[c] nd : not determined
[d] SD :standard deviation calculated from values of at least three independent experiments.
[e] Activities were obtained by qRT-PCR; calculation of relative expression includes normalization against the 16S rRNA endogenous control.
[f] Activities were obtained by measuring GUS activity of pVO insertion mutant of the corresponding gene

severely decreased in Δ*xypB* inner membrane transporter mutant but not in other transporter mutants (Table R2). The level of extracellular xylanase activity was also significantly lower in *agu67*A::pVO mutant. This suggests that the putative α-glucuronidase encoded by this gene is mandatory to get full extracellular xylanase activity. For both mutants, the reduced phenotype was complemented by the introduction of pC-XypB or pC-Agu67A plasmids which constitutively express *xypB* or *agu67A* genes, respectively (Table R2).

The level of xylanase activity was significantly increased in *xylR* repressor mutant, thus suggesting that this gene represses the expression of genes required for the production of extracellular xylanase activity (Table R2). More surprisingly, the activity was also significantly higher in Δ*gly43F* deletion mutant as compared to wild-type strain (Table R2). This mutant was the only mutant of family GH43 which showed a modification in xylanase activity (data not shown). The introduction of pC-*gly43F* complementation plasmid into Δ*gly43F* mutant significantly reduced the level of xylanase activity, thus confirming the role of this enzyme in the production of xylanase activity.

Production of extracellular xylanase activity is repressed by HrpG and HrpX

As it was recently shown that the production of extracellular xylanase activity is repressed by HrpG and HrpX in *Xcv* [12], we studied the production of xylanase activity by *hrpG* and *hrpX* mutants in *Xcc*. The extracellular xylanase activity was significantly increased in both mutants when compared with the wild-type strain (Table R2). To determine which of the 3 putative xylanases of *Xcc* was responsible for this extra-activity, we constructed the three following double mutants, *hrpX*::pVO-Δ*xyn10A*, *hrpX*::pVO-Δ*xyn10C* and *hrpX*::pVO-Δ*xyn30A*, affected in *hrpX* and each of the xylanase genes. The level of extracellular xylanase activity of both *hrpX*::pVO-Δ*xyn10C* and *hrpX*::pVO-Δ*xyn30A* mutants was similar to that observed for *hrpX*::pVO mutant and is significantly higher to that measured for the wild-type strain (Table R2). No extracellular xylanase activity was detected for *hrpX*::pVO-Δ*xyn10A*, thus confirming that the increase of extracellular xylanase activity observed in *hrp* regulatory mutants depends on Xyn10A.

HrpX and HprG do not repress *xyn10A* but induce *xyn10C*

The results obtained with *hrpG* and *hrpX* mutants prompted us to study the effect of HrpG and HrpX regulators on the expression of genes belonging to *xytA*, *xytB* and *xylR* loci. For that

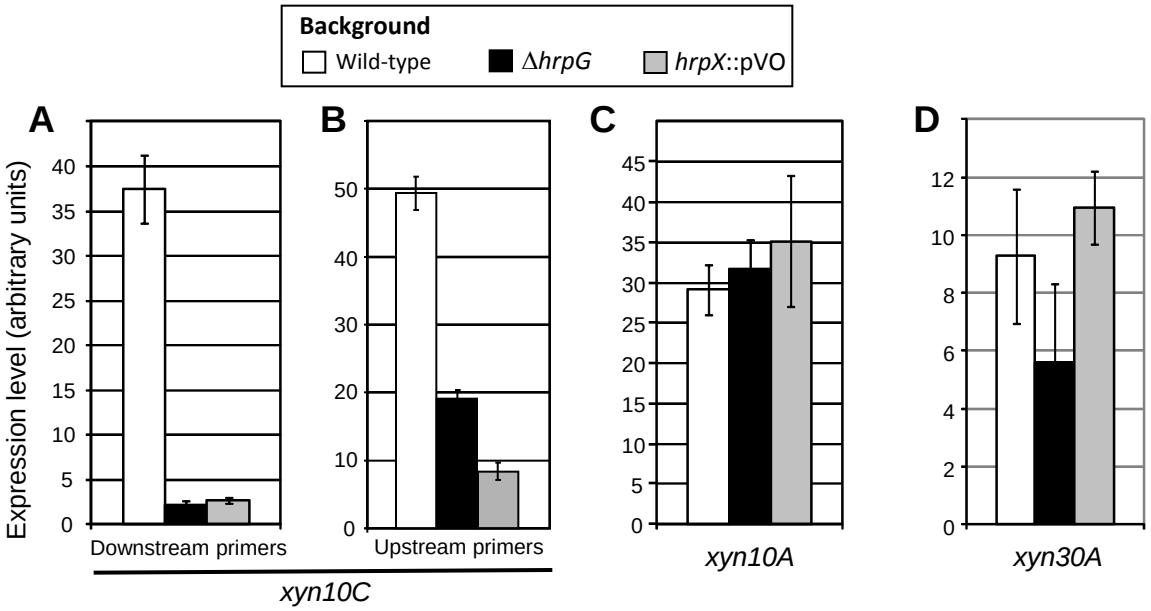

Background
Wild-type
ΔhrpG
hrpX::pVO
A
B
C
D
Expression level (arbitrary units)
Downstream primers
Upstream primers
xyn10C
xyn10A
xyn30A

Figure R3

purpose, we performed qRT-PCR analysis with the wild type strain, Δ*hrpG* and *hrpX*::pVO mutants, grown in minimal medium (MME) or in MME supplemented with xylan either in liquid or in solid media. In these conditions, we did not detect any effect of *hrpG* or *hrpX* mutation on the expression of *xyn10A* (Figure R3). In fact, none of the gene tested seemed to be regulated by *hrpG* or *hrpX* in our conditions (data not shown), with the exception of *xyn10C*, which expression is positively controlled by HrpX and HrpG (Figure R3). No PIP box (TTCGC-N_{15}-TTCGC) or *hrpII* box (TTCG-N_{16}-TTCG), two related boxes found in the promoter region of HrpX-induced genes [52-54], were detected in the 2-kb DNA region located upstream of *xyn10C* coding region. However, a degenerated PIP box TTCGC-N_{16}-TGCGC (between nucleotides 4895775 and 4895751 in *Xcc* AT33913 strain genome) was found in the middle of *xyn10C* coding sequence. To check whether the regulation of *xyn10C* by *hrpX* was not an artefact due to the presence of this degenerated PIP box inside the gene, we designed primers located upstream and downstream of this putative motif into *xyn10C* to conduct our qRT-PCR analysis. Similar results were obtained with both pair of primers (Figure R3).

The LacI regulator XCC4101 represses the expression of genes/operons preceded by a *xyl*-box

We analysed the expression of genes putatively involved in xylan degradation in the wild-type strain or in a Δ*xylR* mutant carrying a deletion of the regulatory gene. For this, we performed qRT-PCR analysis of wild-type strain and Δ*xylR* mutant grown in MME using primers specific for each of the genes belonging to the putative xylan CUT system, including genes that do not belong to *xytA*, *xytB* or *xylR* loci and which are scattered in *Xcc* genome. The expression of all genes which are located immediately downstream of putative *xyl*-boxes (*i.e.* *xytA* (*XCC2828*), *xytB* (*XCC4120*), *xypA* (*XCC4119*) and *xylA2* (*XCC4100*) was repressed by XylR (Table R3). The expression of the four genes located downstream of *xypA* (*XCC4119*) [*xyn10A* (XCC4118) to *xyn10C* (XCC4115) genes] is also repressed by XylR (Table R3), suggesting that they form an operon with *xypA*. This was confirmed by performing reverse transcription reactions in *xylR* mutant using different primers followed by PCR analyses with primers located across putative intergenic regions (Figure S2). Similarly, it appears that *xytA* (XCC2828), *xyaC* (XCC2827), *xyaB* (XCC2826) and *xyaA* (XCC2825) on one hand and *xytB*

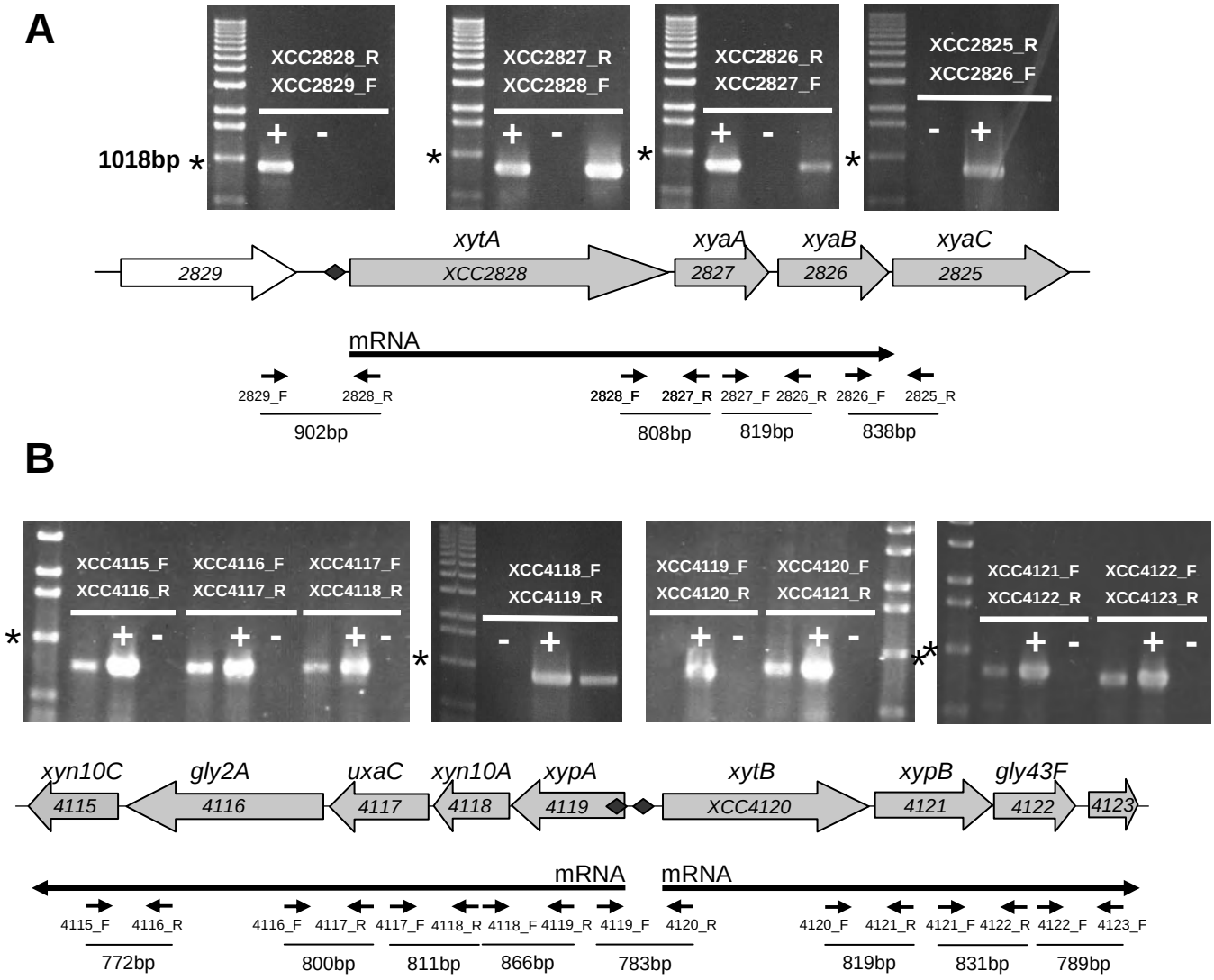

A
XCC2828_R
XCC2829_F
XCC2827_R
XCC2828_F
XCC2826_R
XCC2827_F
XCC2825_R
XCC2826_F
1018bp
*
+ -
- +
xytA
xyaA
xyaB
xyaC
2829
XCC2828
2827
2826
2825
mRNA
2829_F
2828_R
2828_F
2827_R
2827_F
2826_R
2826_F
2825_R
902bp
808bp
819bp
838bp
B
XCC4115_F
XCC4116_R
XCC4116_F
XCC4117_R
XCC4117_F
XCC4118_R
XCC4118_F
XCC4119_R
XCC4119_F
XCC4120_R
XCC4120_F
XCC4121_R
XCC4121_F
XCC4122_R
XCC4122_F
XCC4123_R
xyn10C
gly2A
uxaC
xyn10A
xypA
xytB
xypB
gly43F
4115
4116
4117
4118
4119
XCC4120
4121
4122
4123
mRNA
mRNA
4115_F
4116_R
4116_F
4117_R
4117_F
4118_R
4118_F
4119_R
4119_F
4120_R
4120_F
4121_R
4121_F
4122_R
4122_F
4123_F
772bp
800bp
811bp
866bp
783bp
819bp
831bp
789bp

Figure S2

(*XCC4120*), *xypB* (*XCC4121*) and *gly43F* (*XCC4122*) on the other hand form two operons negatively regulated by $XylR_{XCC}$ (Table R3, Figure S2).

xylA2 is the only gene of *xylR* locus regulated by XylR. The expression of genes located downstream of this gene is not affected by XylR (data not shown). Moreover, it is worth noting that the expression of *agu67A* (*XCC4102*), *axeXA* (*XCC4103*), *uxuA* (*XCC4104*), *gly43E* (*XCC4105*) and *xyl3A* (*XCC4106*) which are in the same orientation is not controlled by XylR (Table R3). No perfect *xyl*-Box was detected upstream any of these genes. However, a degenerated *xyl*-box was identified in the coding sequence of *XCC4102* (data not shown). Finally, genes belonging to the *xylE* locus, as well as those which might be involved in xylan utilization and which are scattered in *Xcc* genome (*gly43A*, *xyn30A*, *gly43B*, *abf51A*, *xyl39A* and *gly43C*), are not controlled by XylR (data not shown).

These data show that XylR regulon comprises *xylA2* gene and three putative operons, *xytA-xyaA-xyaB-xyaC*, *xypA-xyn10A-uxaC-gly2A-xyn10C* and *xytB-xypB-gly43F*, which all contain a *xyl*-box motif in their promoter region.

Xylan induces the expression of genes belonging to XylR regulon and other genes associated with xylan degradation

The expression of most genes belonging to *xytA*, *xytB*, and *xylR* loci was measured in presence of xylan (0,125%). These experiments were performed by using pVO155 insertion mutants, which carry transcriptional fusions between the targeted gene and the *uidA* reporter gene (Figure R1). Firstly, *xylR* expression is not significantly affected by the presence of xylan, whereas the expression of most genes of the XylR regulon is highly induced by this compound (Table R3). However, there are two exceptions since *xypB* and *xylA2* seem to be only very weakly induced by xylan. These results were obtained by monitoring GUS activity produced by pVO155 insertions in these genes. Therefore, these genes (or the gene located downsteam, for *xypB*) might be required for their own induction by xylan. This observation could be in agreement with the fact that the extracellular xylanase activity was significantly reduced in Δ*xypB* mutant (Table R2). However, this was not the case in *xylA2* mutant. It is thus possible that the regulation of these genes by xylan is more complex. This complexity is illustrated by the expression pattern of *xytB-xypB-gly43F* putative operon. As presented above, these three genes seem to form an operon when their expression was monitored into

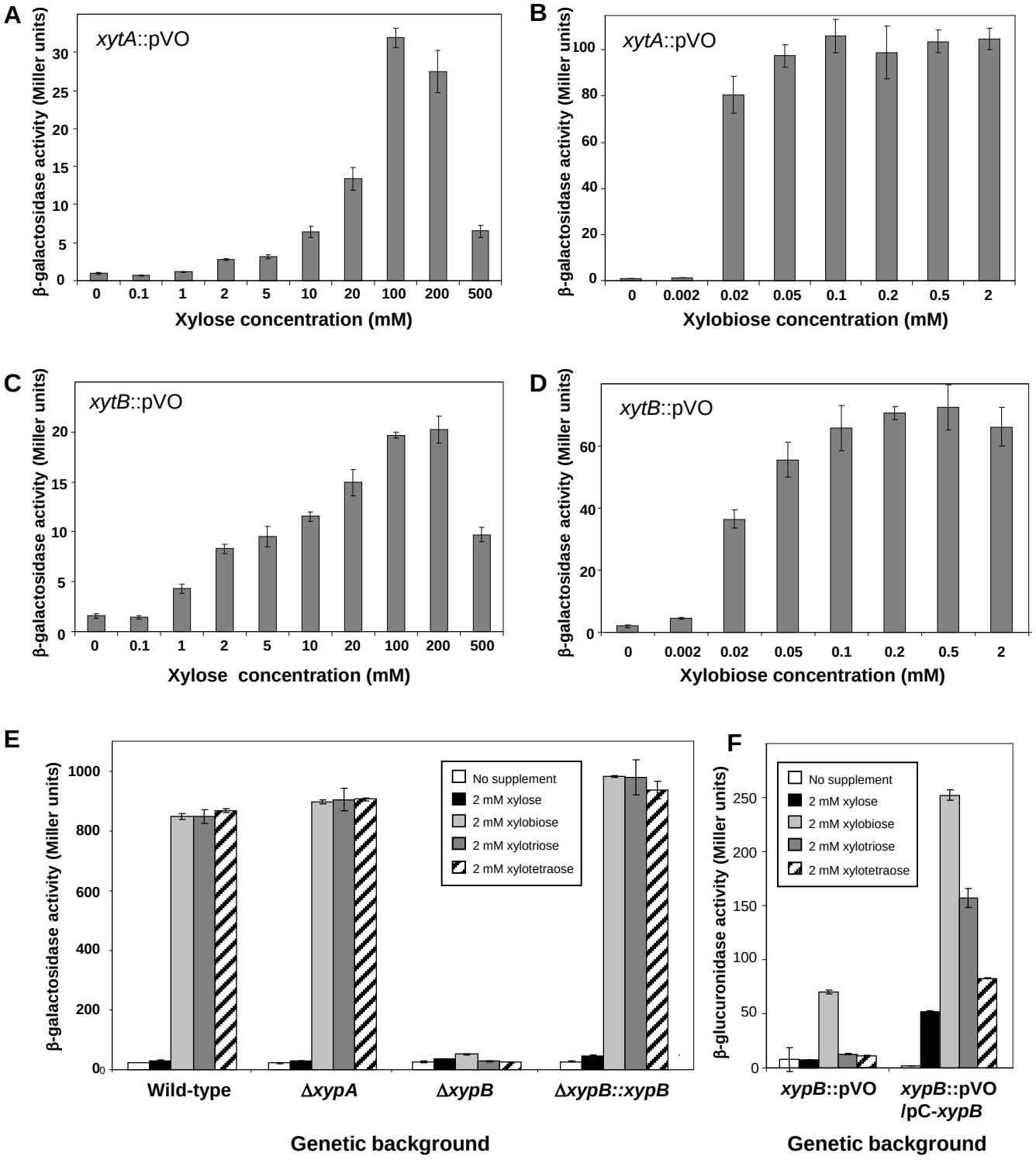

A
xytA::pVO
β-galactosidase activity (Miller units)
0
5
10
15
20
25
30
0
0.1
1
2
5
10
20
100
200
500
Xylose concentration (mM)
B
xytA::pVO
β-galactosidase activity (Miller units)
0
20
40
60
80
100
0
0.002
0.02
0.05
0.1
0.2
0.5
2
Xylobiose concentration (mM)
C
xytB::pVO
β-galactosidase activity (Miller units)
0
5
10
15
20
0
0.1
1
2
5
10
20
100
200
500
Xylose concentration (mM)
D
xytB::pVO
β-galactosidase activity (Miller units)
0
20
40
60
0
0.002
0.02
0.05
0.1
0.2
0.5
2
Xylobiose concentration (mM)
E
β-galactosidase activity (Miller units)
0
200
400
600
800
1000
No supplement
2 mM xylose
2 mM xylobiose
2 mM xylotriose
2 mM xylotetraose
Wild-type
ΔxypA
ΔxypB
ΔxypB::xypB
Genetic background
F
β-glucuronidase activity (Miller units)
0
50
100
150
200
250
No supplement
2 mM xylose
2 mM xylobiose
2 mM xylotriose
2 mM xylotetraose
xypB::pVO
xypB::pVO /pC-xypB
Genetic background

Figure R4

ΔxylR mutant (Figure S2). However, the present data suggest that this is not the case in presence of xylan. If we assume that a functional *xypB* gene is required for its own induction by xylan and by extension of its own operon, we should not observe induction of *xytB* into *xytB*::pVO mutant due to the polar effect of the mutation on *xypB*. This is not the case, since *xytB* expression is highly induced by xylan in *xytB*::pVO mutant. This result suggests either that the pVO155 insertion into *xytB* is not polar or that the operon organization previously observed in Δ*xylR* mutant is bypassed in a wild-type background in presence of xylan. This would suggest that *xypB* can be expressed independently of *xytB* in presence of xylan and that there is another regulatory circuitry which regulates the xylan system in presence of this compound. This observation is strengthened by results obtained with genes of the *xylR* locus. Three of these genes, *gly43E*, *xyl3A* and *uxuB* which are not repressed by XylR are significantly induced by xylan. This implies the existence of another level of regulation which seems to be independent of XylR. Finally, we noticed that two other genes of *xylR* locus, *uxuA* and *axeXA* are very weakly induced by xylan, whereas a third one *agu67A* is not. Once again, the expression of these genes was monitored by measuring GUS activity into pVO155 insertion mutants and we cannot rule out the possibility that functional copies of these genes (or at least one of them, if they form an operon) are required for full xylan induction. Alternatively, it is possible to consider that these genes are not regulated by xylan or that they are not associated with xylan degradation. However, the results showing that *agu67A*::pVO mutants is affected in extracellular xylanase activity seems to exclude this latter hypothesis and rather confirms an involvement of this gene in the induction of the system by xylan.

Expression of xylan associated genes in presence of xylose or xylo-oligosaccharides

The expression of genes belonging to *xytA*, *xytB*, *xylR* and *xylE* was tested in presence of xylose or xylo-oligosaccharides. Xylose was chosen because it is the main component of xylan, whereas xylo-oligosaccharides represent common degradation products of this complex molecule. We performed preliminary experiments by monitoring GUS activity produced by *xytA*::pVO and *xytB*::pVO mutants grown for 6 hours in MME in presence of various concentrations of xylose, xylobiose (X_2), xylotriose (X_3) or xylotetraose (X_4). Results obtained with xylose and xylobiose are shown in Figure R4. Results similar to that obtained with xylobiose (X_2) were observed with xylotriose (X_3) and xylotetraose (data not shown). The expression of *xytA* and *xytB* is induced at much higher levels in presence of xylo-oligosaccharides than xylose, but more importantly the threshold of induction was very

Table R4. Relative expression ratios for genes in the xylan utilization system in presence of xylose or xylo-oligosaccharides

				Expression ratios (SD[f])								
				20 mM	2 mM				50 µM			
Locus	**Gene ID[a]**	**Name**	**Orient.[a]**	**MME X_1[b]/MME**	**MME X_1[b]/MME**	**MME X_2[b]/MME**	**MME X_3[b]/MME**	**MME X_4[b]/MME**	**MME X_1[b]/MME**	**MME X_2[b]/MME**	**MME X_3[b]/MME**	**MME X_4[b]/MME**
xylE	*XCC1755*	*xylS*	F	6.16[c] (1.91)	2.74[c] (1.31)	1.53[c] (0.90)	0.97[c] (0.19)	0.51[c] (0.12)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC1757	*xylB*	F	4.55[c] (0.87)	6.00[c] (0.64)	4.91[c] (2.23)	4.03[c] (1.33)	2.67[c] (0.23)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC1758	*xylA1*	F	20.52[c] (0.64)	3.23[c] (1.70)	3.24[c] (1.94)	2.42[c] (0.99)	1.15[c] (0.49)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC1759	*xylE*	F	6.04[c] (0.54)	4.54[c] (0.25)	2.82[c] (0.80)	2.46[c] (0.43)	1.73[c] (0.15)	nd[d]	nd[d]	nd[d]	nd[d]
xytA	*XCC2825*	*xyaC*	R	nd[d]	0.92[c] (0.04)	3.61[c] (0.76)	4.78[c] (0.72)	2.53[c] (0.63)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC2826	*xyaB*	R	nd[d]	nd[d]	nd[d]	nd[d]	nd[d]	nd[d]	nd[d]	nd[d]	nd[d]
	XCC2827	*xyaA*	R	nd[d]	2.25 (1.31)	49.28 (22.50)	22.34 (3.96)	16.88 (7.19)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC2828	*xytA*	R	15.66 (2.76)	6.68 (0.27)	59.51 (4.39)	67.42 (1.10)	63.14 (4.60)	0.85 (0.0003)	38.58 (1.29)	50.26 (1.71)	21.97 (0.69)
xylR	*XCC4099*[e]		R	nd[d]	0.78 (0.02)	0.94 (0.03)	0.97 (0.06)	0.99 (0.01)	0.84 (0.08)	1.01 (0.03)	0.95 (0.06)	0.91 (0.10)
	XCC4100 [c]	*xylA2*	F	3.03 (0.39)	3.21 (0.23)	20.25 (2.70)	19.37 (2.99)	18.08 (2.61)	0.84 (0.02)	6.83 (0.50)	6.40 (0.53)	7.25 (1.44)
	XCC4101	*xylR*$_{XCC}$	R	0.37 (0.01)	0.55 (0.03)	0.73 (0.01)	0.63 (0.02)	0.64 (0.02)	0.85 (0.01)	0.93 (0.01)	0.88 (0.04)	0.90 (0.03)
	XCC4102	*agu67A*	F	6.87 (0.95)	1.61 (0.14)	27.15 (6.29)	24.79 (6.36)	23.97 (5.74)	0.77 (0.07)	24.52 (0.37)	20.38 (5.73)	15.41 (8.01)
	XCC4103	*axeXA*	F	4.31 (0.15)	1.87 (0.23)	19.69 (3.41)	19.07 (1.77)	17.79 (1.56)	0.68 (0.17)	17.20 (1.44)	16.94 (1.45)	11.03 (1.21)
	XCC4104	*uxuA*	F	3.28 (0.32)	1.33 (0.05)	11.01 (0.29) [a]	9.04 (0.12)	9.97 (0.59)	0.74 (0.13)	8.55 (0.23)	7.76 (0.28)	7.51 (2.76)
	XCC4105	*gly43E*	F	3.43 (0.38)	1.12 (0.04)	21.89 (6.53)	23.27 (3.57)	21.28 (6.04)	0.9 (0.1)	8.9 (1.5)	9.9 (0.6)	16.6 (1.5)
	XCC4106	*xyl3A*	F	3.29 (0.06)	1.20 (0.14)	13.48 (1.86)	13.03 (1.19)	13.03 (2.40)	0.94 (0.02)	10.40 (0.05)	10.73 (0.04)	10.32 (1.42)
	XCC4107	*uxuB*		3.34 (0.17)	1.40 (0.06)	15.49 (1.51)	14.84 (0.39) [a]	14.67 (0.25)	0.88 (0.03)	6.64 (0.39)	4.86 (0.17)	3.79 (0.27)
xytB	*XCC4115*	*xyn10C*	R	1.95 (0.14)	0.92 (0.12)	18.23 (1.73)	15.07 (1.03) [a]	15.57 (0.55)	0.79 (0.1)	8.19 (0.17)	7.53 (0.08)	5.43 (0.05)
	XCC4116	*gly2A*	R	2.00 (0.19)	1.21 (0.04)	29.73 (1.11)	23.42 (2.74)	26.07 (2.60)	0.81 (0.09)	8.66 (1.78)	9.87 (1.48)	6.25 (0.39)
	XCC4117	*uxaC*	R	2.42 (0.03)	1.47 (0.21)	31.39 (1.89)	27.16 (2.85)	28.05 (2.31)	0.84 (0.04)	12.62 (0.28)	11.09 (0.69)	9.02(1.59)
	XCC4118	*xyn10A*	R	1.17 (0.18)	0.97 (0.15)	4.00 (0.28)	4.59 (0.35)	5.01 (0.70)	1.16 (0.18)	10.37 (0.45)	21.98 (0.38)	13.97 (1.11)
	XCC4119 [c]	*xypA*	R	3.96 (0.62)	1.02 (0.06)	14.24 (1.25)	15.95 (1.24)	16.52 (1.48)	0.87 (0.04)	39.77 (1.01)	53.46 (0.08)	35.65 (1.5)
	XCC4120 [c]	*xytB*	F	29.20 (1.49)	1.65 (0.12)	129.75 (2.81)	152.15 (2.03)	137.63 (3.53)	0.89 (0.02)	266.09 (2.10)	268.89 (4.25)	202.87 (3.83)
	XCC4121	*xypB*	F	1.08 (0.25)	0.44 (0.01)	4.48 (0.13)	0.80 (0.02)	0.67 (0.03)	0.95 (0.07)	0.92 (0.05)	0.95 (0.02)	0.87 (0.02)
				nd[d]	1.48[c] (0.23)	7.13[c] (3.26)	5.41[c] (0.99)	3.43[c] (1.00)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC4122	*gly43F*	F	30.02 (9.18)	1.06 (0.01)	10.64 (2.42)	11.26 (1.65)	5.19 (0.91)	1.32 (0.10)	93.60 (5.25)	79.55 (0.12)	2.90 (2.1)
				2.31[c] (0.66)	0.87[c] (0.18)	25.77[c] (9.58)	20.13[c] (6.32)	15.95[c] (5.20)	nd[d]	nd[d]	nd[d]	nd[d]
	XCC4123[e]	*blc*	F	nd[d]	0.66 (0.02)	1.75 (0.10)	1.84 (0.11)	1.69 (0.0003)	0.75 (0.10)	0.98 (0.01)	1.31 (0.07)	0.90 (0.13)

[a] Gene ID and transcriptional orientation are from *Xanthomonas campestris* pv *campestris* strain ATCC33913 [58]. F, forward ; R, reverse.

[b] Minimal medium (MME) was supplemented with xylose (X_1), xylobiose (X_2), xylotriose (X_3) or xylotetraose (X_4).

[c] Expression were determined by qRT-PCRin the wild-type strain; calculation of relative expression includes normalization against the 16S rRNA endogenous control; all other ratios are from expression monitored by measuring β-glucuronidase activity of mutants carrying pVO155 insertion in the tested genes.

[d] nd : not determined

[e] These genes do not belong to the xylan/xylose CUT system and constitute controls.

[f] SD: standard deviation obtained from values of three independent experiments.

different. Induction of expression of both genes was detected at a concentration of xylobiose of 20 µM whereas a xylose effect was only detected at concentrations ranging from 1 to 2 mM. The level of induction obtained with 20 µM of xylobiose was never reached with xylose, even at 20 mM concentration. The differential effect of xylo-oligosaccharide and xylose is clearly shown in experiments monitoring the activity of *xytB* promoter (Figure R4E). The putative promoter region of *xytB* was cloned upstream of the promoterless *lacZ* gene in a reporter plasmid (see Materials and Methods). This plasmid, named pPr-*xytB*, was introduced into the *Xcc* wild-type strain and the expression of the reporter gene was monitored after 6 hours of growth in MME containing xylose, xylobiose, xylotriose or xylotetraose at a final concentration of 2 mM. At this concentration the level of induction obtained with the three xylo-oligosaccharides is 27 fold higher than that observed with xylose.

To see whether this feature was specific for the two studied genes or common to genes of our putative CUT xylan/xylose system, we compared the expression of most genes in presence of xylose (20 mM, 2mM and 50 µM) or xylo-oligosaccharides (2 mM and 50 µM) (Table R4).

Genes belonging to *xylE* locus, which are putatively involved in xylose utilization are induced in presence of xylose and xylo-oligomers. However, the expression of these genes reach higher level of induction in presence of 20 mM xylose than in presence of xylose or xylooligosaccharides at 2mM concentrations (Table R4). The differences of expression levels are not very high between xylose 20 mM, xylose 2 mM and xylobiose 2 mM conditions. Morevover, the level of induction of these genes decreases when the size of the xylo-oligosaccharides that are provided increases. *xylB* is the only gene of this locus which expression is significantly induced by xylotetraose (Table R4). It is worth noting that the expression of *xylS* which is located upstream of *xylE* locus and which is putatively involved in xyloglucan degradation follows the same pattern of induction (Table R4).

The induction pattern of genes belonging to *xytA*, *xytB* and *xylR* loci was significantly different. With very few exceptions, xylose dimer, trimer and tetramer appeared to be much better inducers than the xylose monomer for genes belonging to these loci. Most of these genes are induced when xylose is provided at a 20 mM final concentration whereas they are less induced or not induced in presence of xylose 2 mM. However, the level of induction of these genes is 3 to 300 fold higher in the presence of one of the three xylo-oligosaccharides than in the presence of xylose 20 mM. The three xylo-oligosaccharides gave approximately the same level of induction at a given concentration, and in general, the level of induction for

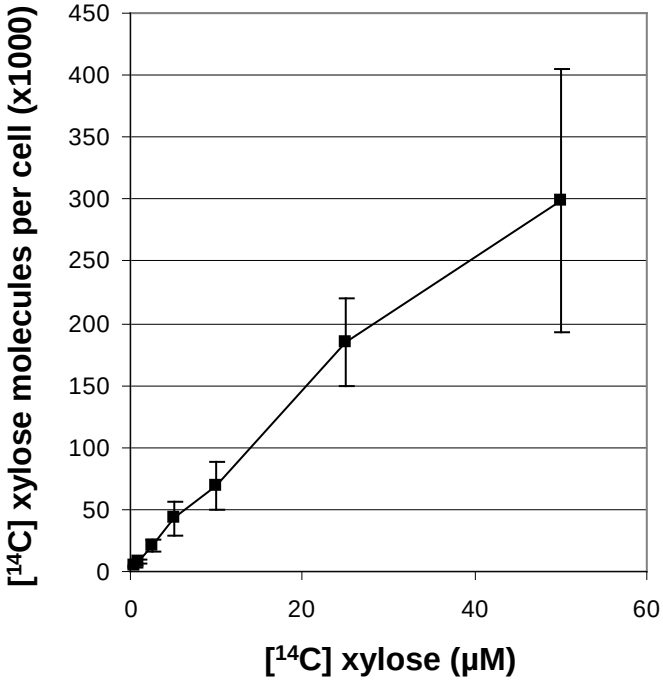

[14C] xylose molecules per cell (x1000)
450
400
350
300
250
200
150
100
50
0
0
20
40
60
[14C] xylose (µM)

Figure R5

each molecule was higher at 2 mM than at 50 µM, with the exception of *xyn10B, xytB, gly43F* and *xypA* genes which have an inverted pattern of induction. Altogether, these results confirmed the previous data obtained with *xytA* and *xytB*. It is worth noting that the two TBDT genes are the most highly induced genes that we tested in these conditions. Interestingly, we also observed that *agu67*, *axeXA* and *uxuA* genes which expression was not induced significantly by xylan in mutated background are induced by xylose and xylo-oligosaccharides and that their induction pattern is similar to that described above for most genes of *xytA*, *xytB* and *xylR* loci. Similarly, the expression of *uxuA*, *gly43E*, *xyl3A* and *uxuB* follows the same pattern of induction. These results confirm the existence of different level of regulation and the role of some of these genes in promoting the induction.

As previously observed with xylan, *xypB* gene, which codes for a putative inner membrane transporter, displayed a different behaviour. In *xypB*::pVO insertion mutant, the expression of this gene is not induced by the presence of xylose, xylotriose and xylotetraose, but a low but significative induction was observed with xylobiose. To test whether this result was due to the mutation of this gene we studied its expression by introducing (pC-*xypB*) plasmid, expressing *xypB* constituvely, into *xypB*::pVO mutant. The introduction of this plasmid modified the regulation pattern of *xytB* which in these conditions followed the general induction pattern of genes of the xylan/xylose CUT system: its expression is induced by xylobiose, xylotriose and xylotretraose as well as by xylose but the level of induction with this pentose monomer is lower than that observed with xylo-oligossacharides (Figure R4F). This confirms that the expression of *xypB* depends on a functional *xypB* gene. To further validate this result, we studied the expression of *xypB* by qRT-PCR in the wild-type strain. In this genetic background, *xypB* expression is highly induced by xylo-oligosaccharides and weakly by xylose 2mM, thus confirming the fundamental role played by this putative transporter in presence of xylo-oligosaccharides (Table R4). Moreover, we integrated pCZ1034-*xypB* plasmid which constitutively expresses *xypB* into Δ*xypB* mutant chromosome (see Material and Methods). The expression of *xytB* promoter carried by pPr-*xytB* plasmid is not induced by xylo-oligosaccharides in Δ*xypB* mutant. The integration of pCZ1034-*xypB* into Δ*xypB*/pPr-*xytB* strain restablished a normal induction pattern (Figure R4E). This result suggests that a functional XypB transporter is required for the correct induction of the xylan CUT system by xylo-oligosaccharide. These data also raise the question of *xytB-xypB-gly43F* operon organisation revealed in Δ*xylR* mutants. Indeed, as observed with xylan, the promoter of *xytB* is induced by xylo-oligosaccharides in *xytB*::pVO mutant whereas it is not induced in *ΔxypB*

Table R5. Rates of ^{14}C-labeled xylose transport of mutants compared to the rate in *Xanthomonas campestris pv. campestris* wild type strain[a]

Strain	Protein family	Mean % transport (SD)[b]
Wild type		100 (6.4)
xylR::pVO	LacI family regulator	110.3 (19.6)
Δ*xytA* Δ*xytB*	TBDT	92.5 (16.3)
Δ*xypA* Δ*xypB xylE*::pVO		20 (2.7)
Δ*xypA* Δ*xypB*		109.1 (3.8)
Δ*xypA*	MFS transporter	106.3 (9.4)
Δ*xypB*	Sugar-cation symporter	97 (4.2)
xylE::pVO	MFS transporter	19.8 (3.3)
xylE::pVO/pC-*xylE*		101.4 (7.6)

[a] Transport rates were measured 60 min after addition of ^{14}C-labeled xylose.
[b] Standard deviations were calculated from three independent experiments.

mutant. This suggests that the pVO155 insertion in *xytB* is not polar on *xypB*. Thus, the transcription of this latter gene does not exclusively depend on *xytB* promoter in presence of xylo-oligossacharides, suggesting the existence of another promoter for *xypB*.

Altogether, these results suggest that the degradation of xylan and the utilization of xylose constitute two pathways which are differentially regulated but which seem to be interwoven. The presence of TBDT and inner membrane transporters in this complex pathway indicate that they constitute a typical CUT system involved in xylan degradation.

Xylose passively diffuses through the envelope

To better understand the function of transporters in the xylan/xylose CUT system we studied the transport of xylose or xylo-oligosaccharides either directly or indirectly.

The initial concentration-dependent [^{14}C]xylose transport, reflecting the dissociation constant (K_d) for xylose uptake was determined using the previously described rapid dilution method [25,28](Figure R5). The deduced K_d (122 µM) is very similar to that previously observed for *N*-acetyl-glucosamine (GlcNAc) uptake by *Xcc* (138.9 µM) and which is in a range similar to the range for K_d values obtained for passive diffusion through porins [31]. This K_d value is 3700 higher that that obtained for sucrose uptake which is actively taken up by SuxA TBDT in *Xcc* [25]. Moreover, as described for GlcNAc, the kinetic values showed that the uptake rate was low and monophasic (Figure R5), suggesting passive diffusion, whereas it was biphasic for sucrose, a feature which is typical of active transport by TBDTs [25]. Therefore, it seems that xylose passively diffuses through the envelope of *Xcc* and that XytA and XytB TBDTs are not required for its active transport across the outer membrane. This was confirmed by comparing the rate of [^{14}C]xylose in *Xcc* wild type strain and the Δ*xytA*Δ*xytB* double mutant carrying a deletion of both *xytA* and *xytB* TBDT genes, since uptake of xylose was very similar in the double mutant and in the wild type strain (Table R5).

XylE inner membrane transporter controls xylose uptake in *Xcc*

xypA, *xypB* and *xylE* genes encoding putative inner membrane transporters were mutated individually or in combination. We generated a *xylE*::pVO155 insertion mutant, Δ*xypA* and

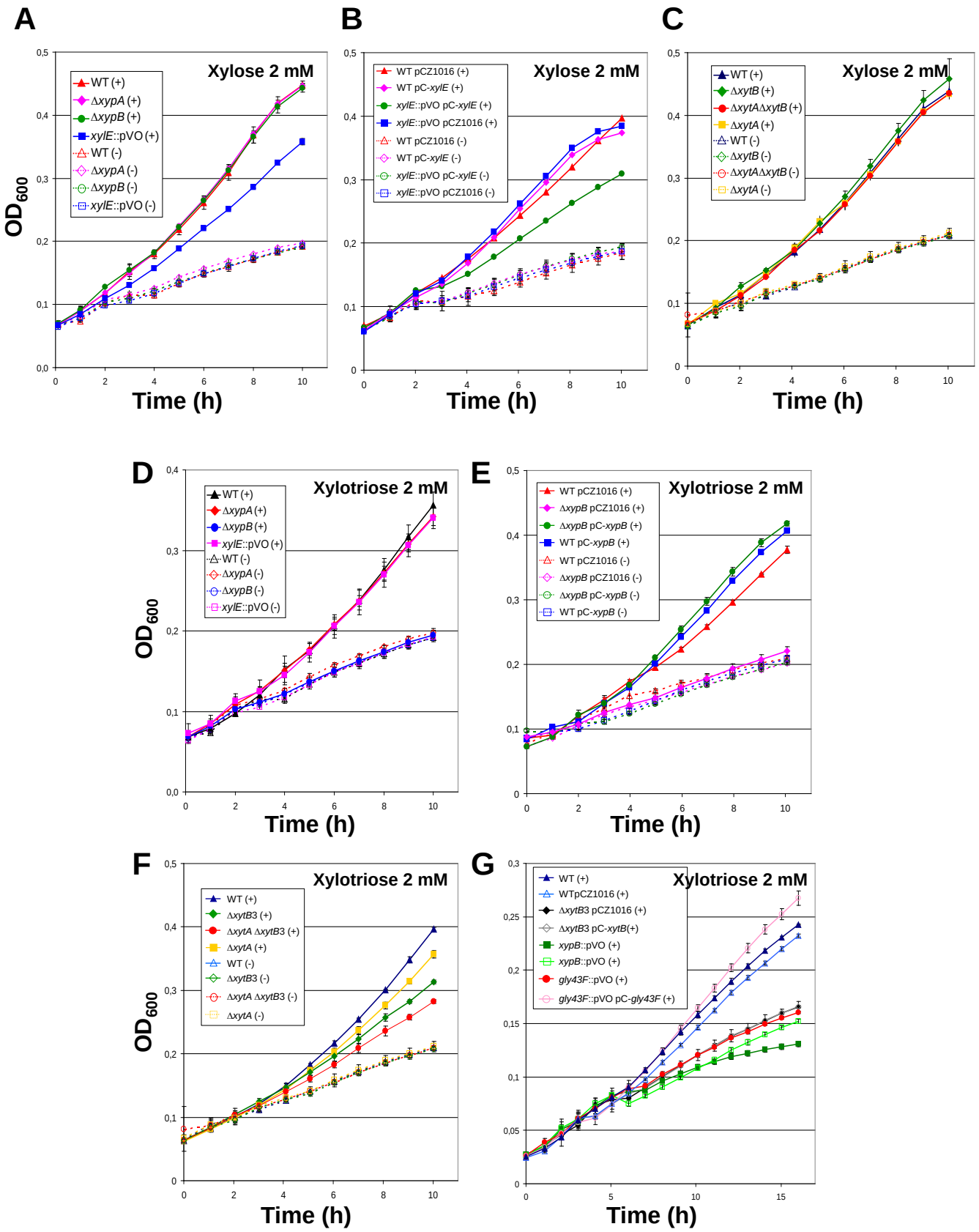

A
Xylose 2 mM
WT (+)
ΔxypA (+)
ΔxypB (+)
xylE::pVO (+)
WT (-)
ΔxypA (-)
ΔxypB (-)
xylE::pVO (-)
OD600
Time (h)
B
Xylose 2 mM
WT pCZ1016 (+)
WT pC-xylE (+)
xylE::pVO pC-xylE (+)
xylE::pVO pCZ1016 (+)
WT pCZ1016 (-)
WT pC-xylE (-)
xylE::pVO pC-xylE (-)
xylE::pVO pCZ1016 (-)
Time (h)
C
Xylose 2 mM
WT (+)
ΔxytB (+)
ΔxytAΔxytB (+)
ΔxytA (+)
WT (-)
ΔxytB (-)
ΔxytAΔxytB (-)
ΔxytA (-)
Time (h)
D
Xylotriose 2 mM
WT (+)
ΔxypA (+)
ΔxypB (+)
xylE::pVO (+)
WT (-)
ΔxypA (-)
ΔxypB (-)
xylE::pVO (-)
OD600
Time (h)
E
Xylotriose 2 mM
WT pCZ1016 (+)
ΔxypB pCZ1016 (+)
ΔxypB pC-xypB (+)
WT pC-xypB (+)
WT pCZ1016 (-)
ΔxypB pCZ1016 (-)
ΔxypB pC-xypB (-)
WT pC-xypB (-)
Time (h)
F
Xylotriose 2 mM
WT (+)
ΔxytB3 (+)
ΔxytA ΔxytB3 (+)
ΔxytA (+)
WT (-)
ΔxytB3 (-)
ΔxytA ΔxytB3 (-)
ΔxytA (-)
OD600
Time (h)
G
Xylotriose 2 mM
WT (+)
WTpCZ1016 (+)
ΔxytB3 pCZ1016 (+)
ΔxytB3 pC-xytB(+)
xypB::pVO (+)
xypB::pVO (+)
gly43F::pVO (+)
gly43F::pVO pC-gly43F (+)
Time (h)

Figure R6

Δ*xypB* single mutants carrying a deletion of *xypA* and *xypB*, respectively, a Δ*xypA*Δ*xypB* double mutant and Δ*xypA*Δ*xypBxylE*::pVO155 triple mutant. The uptake of [^{14}C]xylose was only impaired in strains carrying a mutation in *xylE* (Table R5). The uptake rate of labelled xylose obtained for the *xylE*::pVO and the Δ*xypA*Δ*xypBxylE*::pVO mutants was similar and represented only ~20% of the rate obtained for the wild-type strain. These data suggest that XylE is required for xylose transport across the inner membrane wheareas XypA and XypB are not in the conditions tested. This result was confirmed by comparing the growth rate of the wild-type strain and mutants in these genes in MME supplemented with xylose (2 mM). Growth of the *xylE*::pVO mutant was slightly but significantly impaired on MME containing xylose (2 mM) whereas that of *xypA* or *xypB* mutants was not (Figure R6A, B). In the *xylE*::pVO-complemented strain, xylose transport capacity and growth on MME-xylose (2mM) were restored (Table R5; Figure R6B). These data suggest that xylose transport through the inner membrane is mediated by XylE but also that there are other inner membrane transporters for this monomer. XypA and XypB do not seem to be these additional putative transporters suggesting the existence of other as yet uncharacterized transporters required for xylose uptake into *Xcc*. Furthermore, the growth rate of Δ*xytA* Δ*xytB* mutant in presence of 2 mM xylose was similar to that of the wild-type strain, thus confirming that these two TBDTs are not required for xylose transport (Figure R6C).

XypB inner membrane transporter and Gly43F putative cytoplasmic exoxylanase are required for optimal growth with xylo-oligosaccharides

Since our previous experiments on growth of the wild-type strain and *xylE* mutants in presence of xylose reflected well the transport status observed for [^{14}C]xylose uptake by these strains, we decided to indirectly study the transport of xylo-oligossacharides by comparing the growth rates of the wild type strain and *xypA*, *xypB* or *xylE* mutants in presence of these different compounds. This was performed by growing the strains in MME supplemented with xylobiose, xylotriose or xylotetraose at a final concentration of 2 mM, a concentration which induced the expression of most genes of the xylan CUT system. Data obtained with xylotriose are presented since they are representative of results obtained with xylobiose and xylotetraose.

We noticed that the growth rate of the wild-type strain was slower in presence of xylo-oligosaccharide than in the presence of xylose (Figure R6). *xylE*, and *xypA* mutants were not significantly affected in their growth with xylobiose, xylotriose or xylotetraose (Figure R6D).

This was not the case for Δ*xypB* (Figure R6D) and *xypB*::pVO (Figure R6G) mutants for which growth rates were significantly affected in presence of the three xylo-oligosaccharides. We observed that the growth rate of *xypB*::pVO mutant was more altered than that of Δ*xypB* (Figure R6 D, E, G). Wild-type growth rate was restored by introducing pC-*xypB* into Δ*xypB* mutant. However, the growth rate of *xypB*::pvo mutant was only slighty improved by introducing pC-*xypB* in this mutant. These data suggest that XypB is required for optimal transport of xylo-oligosaccharide across the inner membrane. They are also in good agreement with our results concerning the absence of induction by xylo-oligosaccharide of this gene in a *xypB*::pVO genetic background. The reduction of xylo-goligosaccharide transport into the cytoplasm of this mutant may limit the induction of the xylan CUT system. They also suggest that *gly43F* gene located downstream of *xypB* may also play a role in transport or induction of the system. We therefore studied the growth of *gly43F*::pVO mutant in presence of xylose or xylo-oligosaccharides. The growth rate of this mutant is not impaired in presence of xylose (data not shown) but it is significantly reduced in presence of xylo-ologosaccarides (Figure R6G). Complementation experiments carried by introducing pC-*gly43F* plasmid into this mutant showed that the growth rate of complemented strain was even higher than that of the wild-type strain.

XytA and XytB TBDTs and growth of *Xcc* in presence of xylo-oligosaccharides

To study the role of XytA and XytB TBDTs in xylo-oligosaccahride transport, we constructed *xytA* and *xytB* deletion mutants to avoid any polar effect of mutations in these genes on downstream genes. For *xytA* we constructed one mutant, named Δ*xytA* whereas 3 deletions mutants of *xytB*, Δ*xytB*1, Δ*xytB*2 and Δ*xytB*3, were constructed, by removing different parts of this gene (Figure R1). A double mutant Δ*xytA*Δ*xytB*1 was also generated.

We compared the growth rate of these mutants and of the wild-type strain in MME supplemented or not with xylobiose, xylotriose or xylotetraose at a 2 mM concentration. Growth of *xytA* and *xytB* mutants was slightly affected in presence of xylo-oligosaccharides. The growth rate of *xytA* mutant was significantly slower than that of the three *xytB* mutants which displayed similar reduced growth rates (data not shown). The double mutant Δ*xytA*Δ*xytB*1 was more affected than *xytA* and *xytB* single mutants, suggesting an additive effect of both mutations (Figures 6F). The reduction of growth rate of these mutants is lower than that observed with Δ*xypB* mutants. Our attempts to complement Δ*xytB1*, *ΔxytB2* or

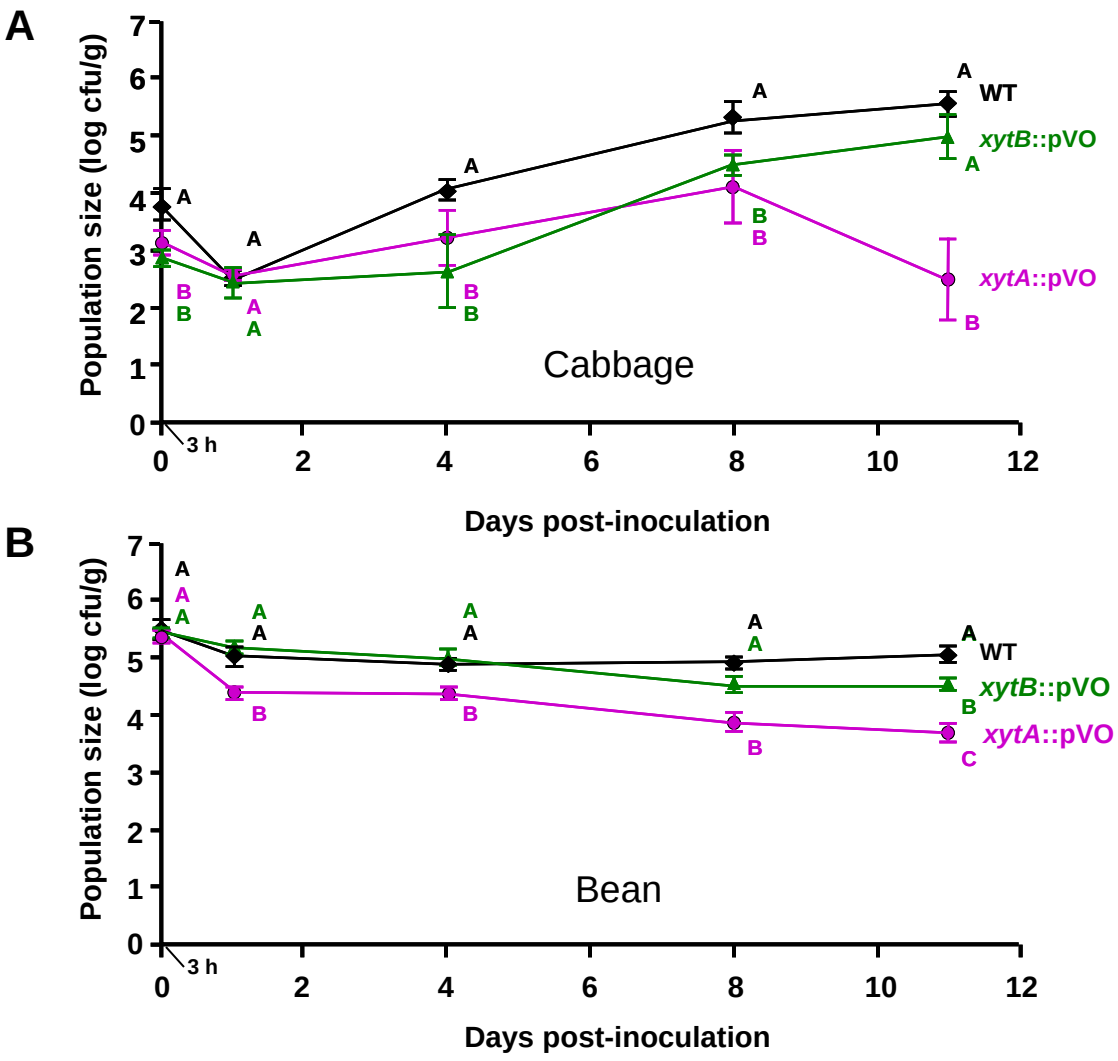

A
Population size (log cfu/g)
Cabbage
3 h
Days post-inoculation
WT
xytB::pVO
xytA::pVO
B
Population size (log cfu/g)
Bean
3 h
Days post-inoculation
WT
xytB::pVO
xytA::pVO

Figure R7

ΔxytB3 mutants by cloning *xytB* into pCZ1016 expression vector, all failed. These data suggest that the three deletions of *xytB* have probably a polar effect on *xypB* and *gly43F*. To confirm this hypothesis we compared the expression of *xypB* and *gly43F* in the wild type strain or in *xytB* deletion mutants in presence of xylo-oligosaccharides (2 mM). We observed that the expression of *xypB* and *gly43F* is slightly but significantly affected by the three deletions in *xytB* (data not shown). We also observed that the expression of *xytB* promoter is slightly affected by mutations in *xytB* (data not shown). One explanation that can be put forward to explain these results could be that the coding sequences of *xytB* harbour a cis-regulatory element which is required for the full induction of *xypB* and *gly43F* by xylo-oligosaccharides.

XytA and XytB loci are required for optimal growth on host and non-host plants

None of the mutants constructed in the frame of this work were affected in pathogenicity on cabbage or *Arabidopsis thaliana* host plants. This includes *xytA*, *xytB*, *xypB* and *xytAxytB* mutants as well as mutants affected in the three xylanase genes (data not shown). These experiments were conducted by using the wound inoculation method that allows direct delivering of bacterial cells into the xylem vessels. This inoculation method most probably bypasses several steps of the natural infectious process [55]. Therefore, we compared the survival and the multiplication of the wild-type strain and *xytA*::pVO or *xytB*::pVO mutants in the phyllosphere of cabbage (host plant) or bean (non-host plant) (Figure R7). The dynamics of bacterial population densities was followed after spray inoculation of the leaves in conditions which do not favour disease expression [56]. The multiplication of *xytB* mutant on host-plant is slightly lower than that of the wild-type strain during the first eight days following the inoculation. However, after 11 days the cell densities of both strains are not significantly different. The multiplication of *xytA* mutant is significantly more affected on this plant and the cell density measured for this mutant are clearly lower than that measured for the wild type strain and *xytA* mutant. Furthermore, the survival of both *xytA* and *xytB* mutants is significantly altered on non host plant. Once again, the defect of *xytA* mutant is more pronounced than that of *xytB* mutant.

These results suggest that the *xytA* and *xytB* loci are required for optimal growth of *Xcc* in the phyllosphere of both host plant and non-host plants.

Distribution and conservation of the xylan CUT system in *Xanthomonas* spp.

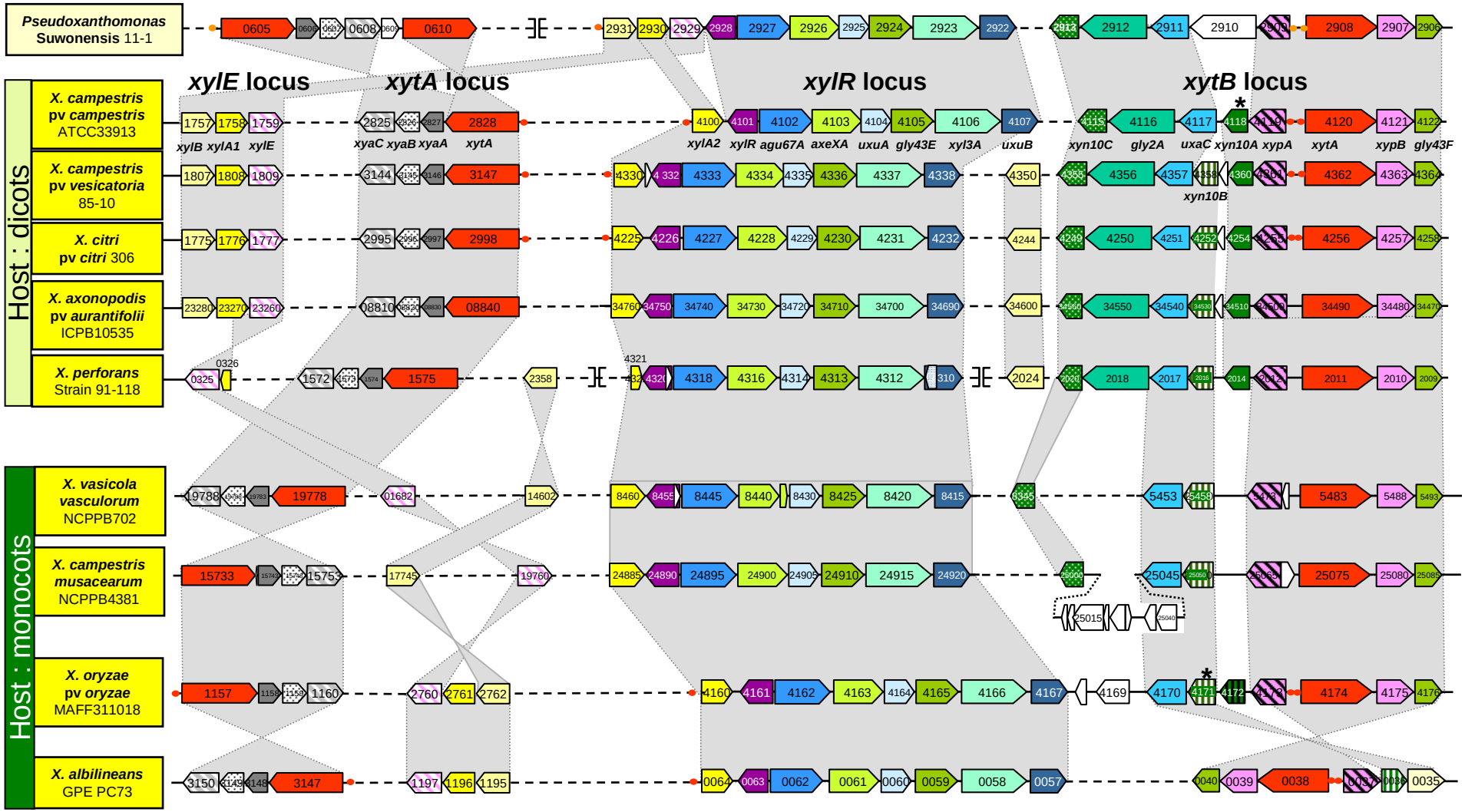

Pseudoxanthomonas Suwonensis 11-1
xylE locus
xytA locus
xylR locus
xytB locus
Host : dicots
X. campestris pv campestris ATCC33913
X. campestris pv vesicatoria 85-10
X. citri pv citri 306
X. axonopodis pv aurantifolii ICPB10535
X. perforans Strain 91-118
Host : monocots
X. vasicola vasculorum NCPPB702
X. campestris musacearum NCPPB4381
X. oryzae pv oryzae MAFF311018
X. albilineans GPE PC73
xylB xylA1 xylE
xyaC xyaB xyaA xytA
xylA2 xylR agu67A axeXA uxuA gly43E xyl3A uxuB
xyn10C gly2A uxaC xyn10A xypA xytA xypB gly43F
xyn10B

Figure S3

The genes belonging to *Xcc* Xylan CUT system were used to perform a comparative study with available *Xanthomonas* genomes. To date, the genus *Xanthomonas* is represented by eleven complete and ten draft genome sequences (Table S2) [14,39]. These genome sequences can be divided into two groups: those of strains infecting dicotyledonous hosts and those pathogenic on monocotyledonous plants. The first group comprises *Xcv* strain 85-10 [57], *Xanthomonas citri* subsp. *citri* strain 306 (*Xac*), formely, *Xanthomonas axonopodis* pv *citri* strain 306 [58], *Xanthomonas fuscans* subsp. *aurantifolii* B strain (*Xaub*) [59], *Xanthomonas fuscans* subsp. *aurantifolii* A (*Xauc*) [59], *Xanthomonas vesicatoria* strain 1111 (*Xv*), *Xanthomonas perforans* strain 91-118 (*Xp*), *Xanthomonas gardneri* strain 101 (*Xg*) [39] and 3 strains of *Xcc*, strain ATCC33913 [58], strain 8004 [60] and strain B100 [61]. The group of strains isolated from monocots includes; *Xanthomonas campestris* pv. *musacearum* NCPPB4381 (*Xcm*), *Xanthomonas vasicola* pv. *vasculorum* NCPPB702 (*Xvv*) [62], *Xanthomonas albilineans* GPE PC73 (*Xal*) [63] and three strains of *Xanthomonas oryzae* pv. *oryzae*, strain MAFF 311018 (*Xoo*MAFF), strain KACC 10331 (*Xoo*KACC) [64] and strain PXO99A (*Xoo*PXO) [65].

The *xytA* locus is conserved and identical in all these strains (Figure S3). The *xylE* is also conserved, with the exception of *Xg, Xp, Xv, Xcm* and *Xvv* which all correspond to draft genomic sequences. In these strains, the *xylE* locus is either incomplete or dislocated and the genes scattered in the genome and present in partial multicopies. It seems difficult to determine whether these differences reflect a real degeneration of this locus or if they are attributable to the draft status of these genome sequences.

The *xylR* locus is very well conserved in all strains, with the exception of *Xg* and *Xp*, in which, the *xylA2* gene is truncated. The DNA region separating *xylR* locus from *xytB* locus is variable and specific of each species (data not shown). We noticed the presence of a putative xylulokinase in the “inter-locus” region of *Xcv, Xp, Xauc*, *Xaub* and *Xac* which are phylognetically grouped together [39,66].

The *xytB* locus is the most variable locus and a recent study showed that its organization can be used to group *Xanthomonas* strains into three groups based on the presence/absence of three xylanase genes *xyn10A, xyn10B* and *xyn10C* [39]. The first group defined by Potnis et al. [39] corresponds to strains which possess *xyn10A*, *xyn10B* and *xyn10C* xylanase genes in *xytB* locus (Figure R8). Members of the second group which includes *Xcc* strains harbour *xyn10A* and *xyn10C* genes but do not possess *xyn10B* (Figure R8). The third group is

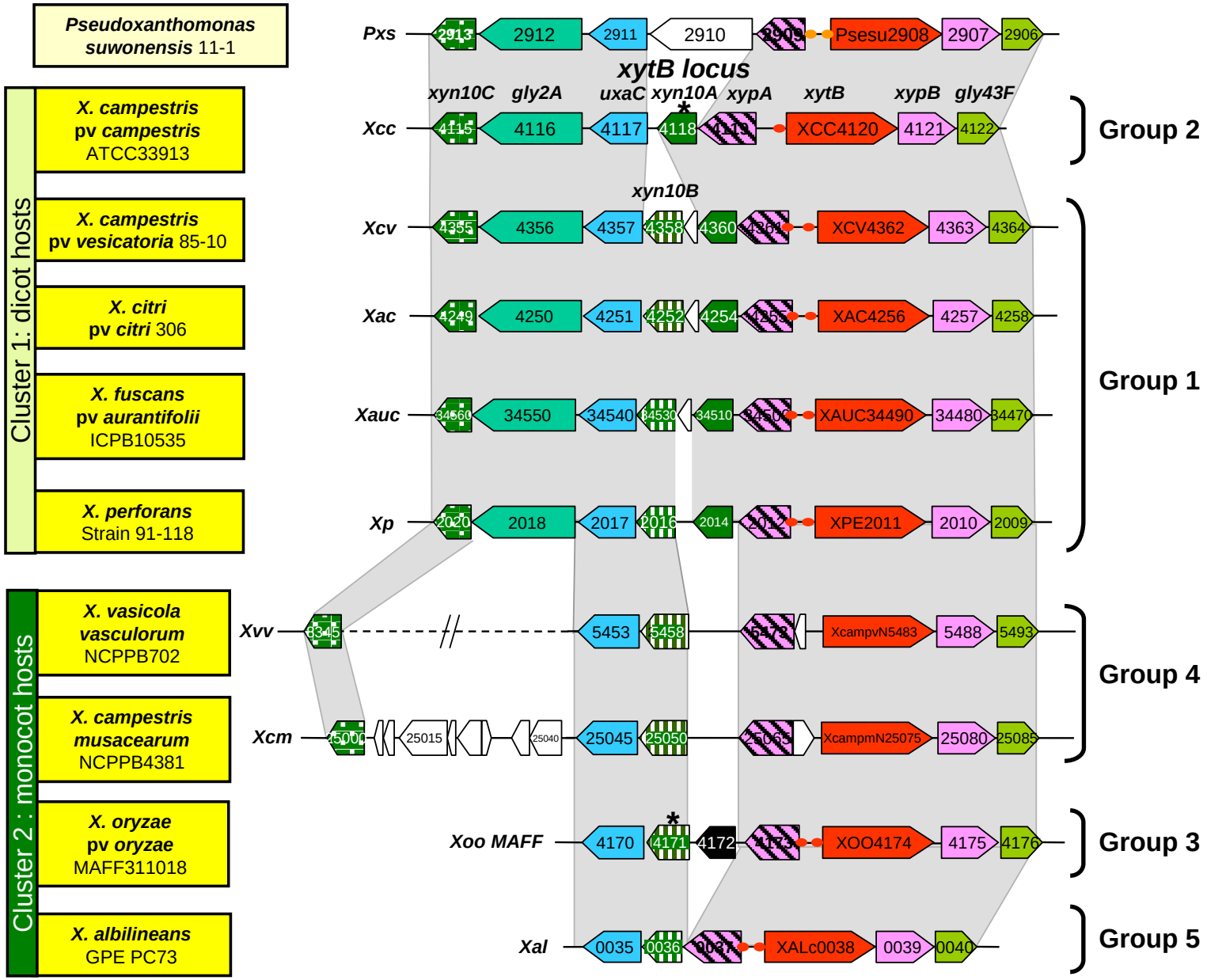

Pseudoxanthomonas suwonensis 11-1
Cluster 1: dicot hosts
X. campestris pv campestris ATCC33913
X. campestris pv vesicatoria 85-10
X. citri pv citri 306
X. fuscans pv aurantifolii ICPB10535
X. perforans Strain 91-118
Cluster 2 : monocot hosts
X. vasicola vasculorum NCPPB702
X. campestris musacearum NCPPB4381
X. oryzae pv oryzae MAFF311018
X. albilineans GPE PC73
xytB locus
xyn10C gly2A uxaC xyn10A xypA xytB xypB gly43F
xyn10B
Pxs
Xcc
Xcv
Xac
Xauc
Xp
Xvv
Xcm
Xoo MAFF
Xal
Psesu2908
XCC4120
XCV4362
XAC4256
XAUC34490
XPE2011
XcampvN5483
XcampmN25075
XOO4174
XALc0038
Group 1
Group 2
Group 3
Group 4
Group 5

Figure R8

represented by *X. oryzae* strains in which *xyn10A* and *xyn10B* are present and *xyn10C* absent. This group is also characterized by the absence of the *gly2A* gene which is contiguous to *xyn10C* in *Xanthomonas* strains of groups 1 and 2 (Figure R8). Our comparative analysis confirms this grouping, but two new groups must be added because three strains, *Xcm*, *Xvm* and *Xal* which display a different distribution pattern were not included in this previous study. The fourth group consists of *Xcm* and *Xvv* in which *xyn10B* and *xyn10C* are present and *xyn10A* and *gly2A* absent. However, it is worth noting that in these two strains the location of *xyn10C* gene is modified. In *Xvv* it is not located in the *xytB* locus but elsewhere in the genome. In *Xcm*, *xyn10C* gene is separated from the *xytB* locus by the insertion of eight genes which do not seem to be related to xylan or xylose utilization. Whether the difference of location of this gene in *Xcm* and *Xvv* affects its expression and function has to be clarified. The fifth group is represented by *Xal* in which only *xyn10B* xylanase gene is present. *gly2B* gene is also absent in this strain (Figure R8).

We noticed that all strains which belong to groups 3, 4 and 5 are pathogenic to monocots and display alteration of the left-hand end side of *xytB* locus, characterized by the loss of *gly2A* gene as well as loss or misplacing of *xyn10C* xylanase gene. Moreover, these 3 groups differ from groups 1 and 2 on the basis of *xyn10A* gene. This gene is absent in strains of groups 4 and 5 and present in strains of group 3 which exclusively comprises *X. oryzae* strains. However, we observed that the Xyn10A proteins harboured by all *X. oryzae* strains which genome have been sequenced so far do not possess a signal peptide whereas Xyn10A xylanases of groups 1 and 2 display such a signal. This observation is in agreement with data previously published on *Xoo* strain BXO1 [11]. BlastN and tBlastN analyses of *Xoo* genomes using DNA or amino acid sequences of *xyn10A* from *Xanthomonas* strains of groups 1 and 2, showed that the DNA sequences corresponding to the N-terminal signal peptide of *xyn10A* of these strains are conserved upstream *Xoo xyn10A* genes but that they are degenerated, thus inducing frameshifts in this part of the sequence and altering the start codon of the proteins. A similar observation was made for *xyn10A* gene of *Xanthomonas oryzae* pv *oryzicola* (*Xoc*) strain BLS 256, the causal agent of bacterial leaf streak of rice (data not shown). These data suggest that *xyn10A* gene is either nonfunctional or that it functions inside the cytoplasm in *Xanthomonas* strains of rice. Therefore, we can speculate that the group defined by *X. oryzae* strains (group 3) resembles groups 4 and 5 which are characterized by the absence of this gene. Altogether, these observations suggest that groups 3, 4 and 5 could be merged together to form a larger group which differs from group 1 and 2. This grouping is strengthened by the

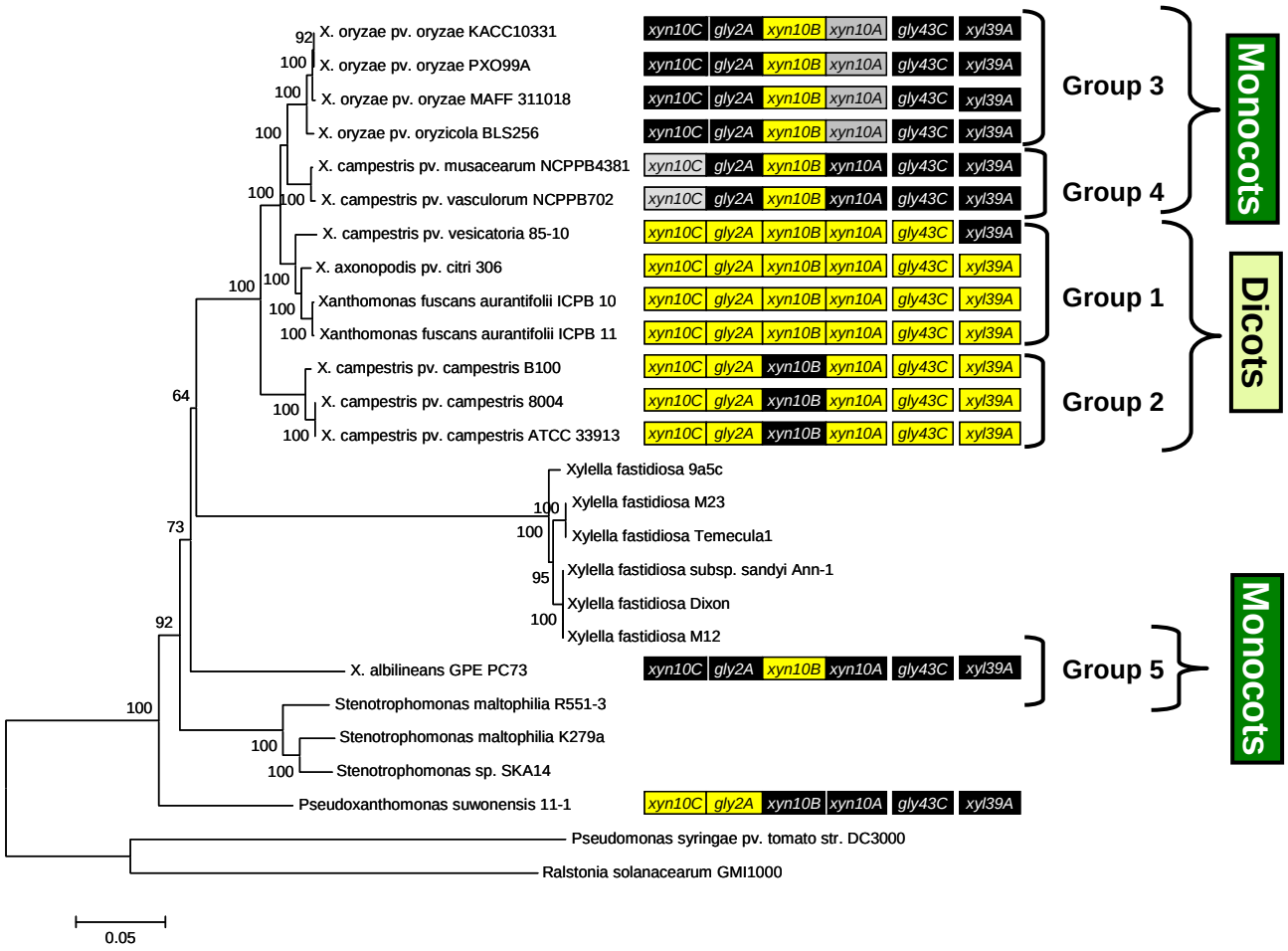

X. oryzae pv. oryzae KACC10331
X. oryzae pv. oryzae PXO99A
X. oryzae pv. oryzae MAFF 311018
X. oryzae pv. oryzicola BLS256
X. campestris pv. musacearum NCPPB4381
X. campestris pv. vasculorum NCPPB702
X. campestris pv. vesicatoria 85-10
X. axonopodis pv. citri 306
Xanthomonas fuscans aurantifolii ICPB 10
Xanthomonas fuscans aurantifolii ICPB 11
X. campestris pv. campestris B100
X. campestris pv. campestris 8004
X. campestris pv. campestris ATCC 33913
Xylella fastidiosa 9a5c
Xylella fastidiosa M23
Xylella fastidiosa Temecula1
Xylella fastidiosa subsp. sandyi Ann-1
Xylella fastidiosa Dixon
Xylella fastidiosa M12
X. albilineans GPE PC73
Stenotrophomonas maltophilia R551-3
Stenotrophomonas maltophilia K279a
Stenotrophomonas sp. SKA14
Pseudoxanthomonas suwonensis 11-1
Pseudomonas syringae pv. tomato str. DC3000
Ralstonia solanacearum GMI1000
xyn10C gly2A xyn10B xyn10A gly43C xyl39A
Group 3
Group 4
Group 1
Group 2
Group 5
Monocots
Dicots
Monocots
0.05

Figure S4

observation that *gly43C* and *xyl39A* genes are absent in all strains of groups 3, 4 and 5 and present in most strains of groups 1 and 2 (Figure S4, Table S3). Interestingly, groups 3, 4 and 5 only comprise strains which infect monocots, whereas group 1 and 2 only contains *Xanthomonas* strains pathogenic on dicots. It is worth noting that *Xal* and the other strains infecting monocots belong to two very distinct phylogenetic branches (Figure S4) [14,63]. Therefore the presence/absence pattern of genes is not strictly related to the taxonomical position of these strains and might rather be related to the class of their host plant.

Finally, *xylR* regulator is conserved in all *Xanthomonas* strains studied here and putative *xyl*-boxes were found upstream orthologs of *Xcc* genes having this box in their promoter region (Figure S3, Table S1).

The xylan CUT system but xylanases is conserved in *Pseudoxanthomonas suwonenesis* an environmental bacterium of the Xanthomonadaceae family

The complete or draft genome sequences of strains belonging to other genera of the Xanthomonadaceae family are publicly available. This comprises *Xylella*, a genus that regroups insect transmitted phytopathogenic strains which are restricted to xylem vessels of host plants and which display a reduced genome [67]. This list also includes the genome of 3 strains of *Stenotrophomonas*, a genus that occurs ubiquitously in the environment, plant and soil being its main reservoir [68]. Finally, the draft genome sequence of *Pseudoxanthomonas suwonenesis* 11-1 (*Pxs*), which was isolated from compost feedstock was recently released.

The *Xcc* Xylan CUT genes are not conserved in *Stenotrophomonas* or *Xylella* genomes. Somewhat surprisingly, it is well conserved in *Pxs* (Figure S3, Table S2), which, based on phylogenetic analyses, is less related to *Xanthomonas* genus than *Stenotrophomonas* or *Xylella* genera (Figure S4) [69]. Twenty two genes of the *Xcc* xylan/xylose CUT system are conserved in *Pxs*, encompassing xylose and glucuronate degradative enzymes, transporters and XylR regulator. The general organisation of these genes is very similar in *Xanthomonas* spp. and *Pxs* (Figure S3). However, there are some differences. *xylE* and *xylR* loci are merged together, and there is only one copy of xylose isomerase gene in *Pxs*. The *xytA* TBDT gene is duplicated in *xytA* locus of *Pxs*. The *xylR* regulatory gene is present in *Pxs* and a putative degenerated *xyl*-box motif was identified in *Pxs* genome, upstream of *Psesu_0605*, *Psesu_2909* and *Psesu_2931* genes, which are the orthologs of *xytA*, *xylB* and *xypA* from *Xcc*, respectively (Figure S1, S3 and Table S1). The location of this putative motif upstream of

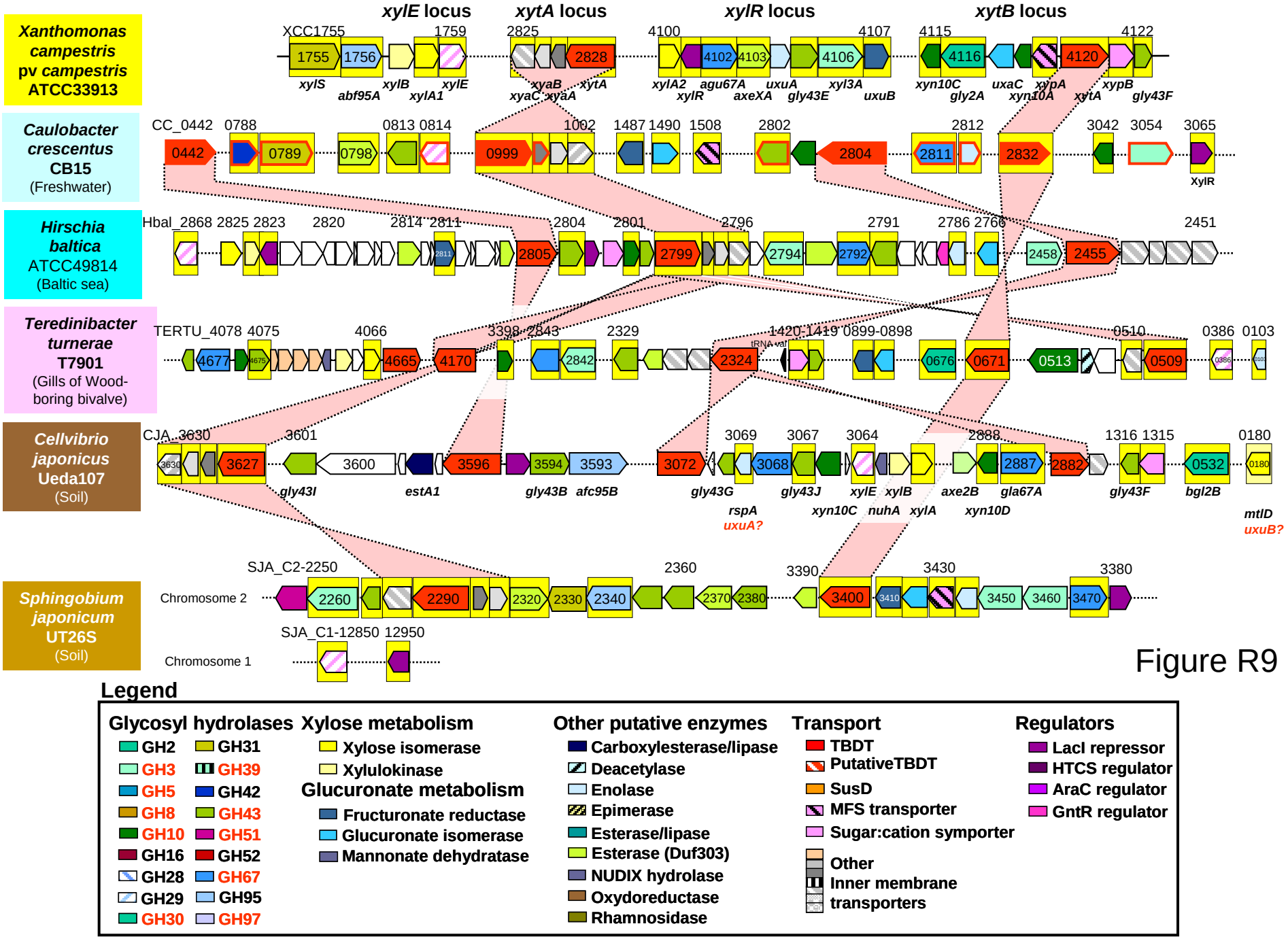

Xanthomonas campestris pv campestris ATCC33913
Caulobacter crescentus CB15 (Freshwater)
Hirschia baltica ATCC49814 (Baltic sea)
Teredinibacter turnerae T7901 (Gills of Wood-boring bivalve)
Cellvibrio japonicus Ueda107 (Soil)
Sphingobium japonicum UT26S (Soil)
xylE locus
xytA locus
xylR locus
xytB locus
Chromosome 2
Chromosome 1
Legend
Glycosyl hydrolases
GH2
GH3
GH5
GH8
GH10
GH16
GH28
GH29
GH30
GH31
GH39
GH42
GH43
GH51
GH52
GH67
GH95
GH97
Xylose metabolism
Xylose isomerase
Xylulokinase
Glucuronate metabolism
Fructuronate reductase
Glucuronate isomerase
Mannonate dehydratase
Other putative enzymes
Carboxylesterase/lipase
Deacetylase
Enolase
Epimerase
Esterase/lipase
Esterase (Duf303)
NUDIX hydrolase
Oxydoreductase
Rhamnosidase
Transport
TBDT
PutativeTBDT
SusD
MFS transporter
Sugar:cation symporter
Other
Inner membrane transporters
Regulators
LacI repressor
HTCS regulator
AraC regulator
GntR regulator

Figure R9

xytA and *xypA* orthologs corresponds to what was observed in *Xanthomonas* strains. However, the presence of this putative box upstream *xylB* ortholog in *Pxs* is different, since this motif was never detected upstream this gene in *Xanthomonas* species. This observation is probably to be correlated with the fact that, *xylE* and *xylR* loci are merged in *Pxs*, whereas they are separated in *Xanthomonas* strains.

A major and notable difference between *Xanthomonas* spp. and *Pxs* concerns xylanase encoding genes. Of the 4 xylanase genes identified in *Xanthomonas* spp., only one, *xyn10C*, is present in *Pxs* (Table S3). The location of this gene to the left-hand end side of *xytB* locus is similar in both *Pxs* and *Xanthomonas* strains of groups 1 and 2. It is worth noting that *gly2A* gene is also present beside *xyn10C* in *Pxs*. In this latter bacterium, the position of *xyn10A* and *xyn10B* missing genes which are contiguous in *xytB* loci of *Xanthomonas* spp., is occupied by a gene, *Psesu_2910*, which codes for a protein of unknown function which possesses a signal peptide (Figure S3). This protein shows the strongest similarities (around 52% amino acid identity) to proteins of phylogenetically distant bacteria belonging to different classes or phyla such as *Caulobacter* sp. K31 (Alphaproteobacteria), *Opitutus terrae* PB90-1 (Verrucomicrobia), or *Mucilaginibacter paludis* (Bacteroidetes). It is not conserved in *Xcc* strains but a protein displaying around 45% amino acid identity is present in *Xcv*, *Xcm*, *Xvv*, *Xac*, *Xauc*, *Xaub*, and *Xoo* strains. However, in these strains the homologous gene is not located in *XytB* locus. The screening of CAZy database, revealed two additional putative xylanases in *Pxs*. These two putative enzymes, Psesu_1642 and Psesu_1618, belong to GH8 and GH11 families, respectively. They are not conserved in any of the *Xanthomonas* strains which genome has been sequenced so far.

The Xanthomonas xylan CUT system is highly conserved in Alphaproteobacteria isolated from aquatic environments or soil

Of the thirty proteins belonging to *Xcc* xylan/xylose CUT system, eighteen display significant similarities to proteins of *C. crescentus* strain CB15. Most of these conserved proteins display very high levels of amino acid identity ranging from 42% to 70% (Table S3). Interestingly, a transcriptomic analysis has identified fifty one genes which expression is significantly induced by xylose in *C. crescentus* CB15. Several of the xylose-induced genes display or belong to putative operons displaying a XylR operator motif in their upstream sequences [32]. Among the eighteen genes of *C. crescentus* conserved to *Xcc* xylan CUT system, ten are induced by xylose and half of them contain a XylR operator motif upstream. Moreover, three

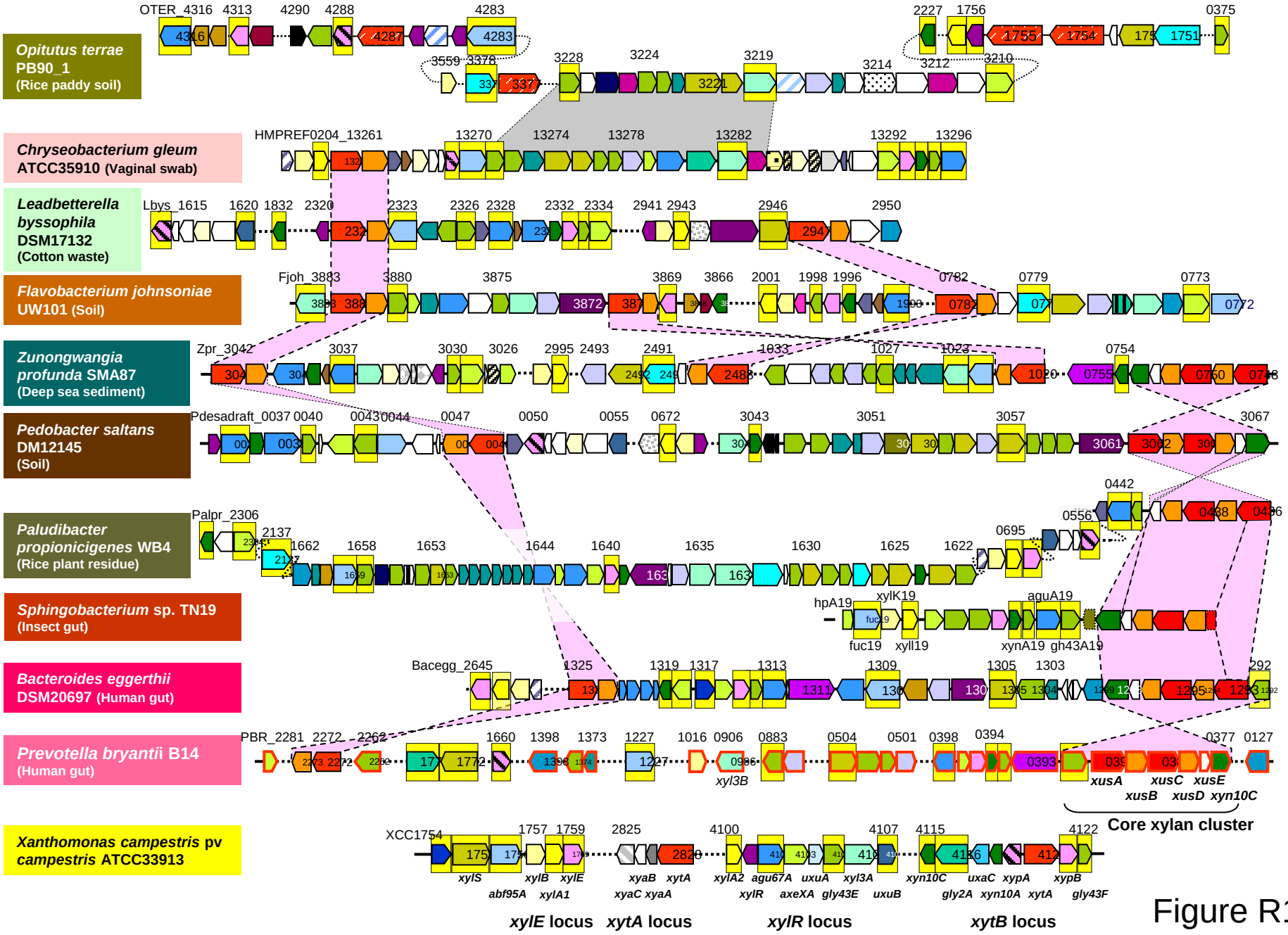

Opitutus terrae PB90_1 (Rice paddy soil)
Chryseobacterium gleum ATCC35910 (Vaginal swab)
Leadbetterella byssophila DSM17132 (Cotton waste)
Flavobacterium johnsoniae UW101 (Soil)
Zunongwangia profunda SMA87 (Deep sea sediment)
Pedobacter saltans DM12145 (Soil)
Paludibacter propionicigenes WB4 (Rice plant residue)
Sphingobacterium sp. TN19 (Insect gut)
Bacteroides eggerthii DSM20697 (Human gut)
Prevotella bryantii B14 (Human gut)
Xanthomonas campestris pv campestris ATCC33913
xylI3B
xusA xusB xusC xusD xusE xyn10C
Core xylan cluster
xylS abf95A xylB xylA1 xylE
xyaC xyaB xyaA xytA
xylA2 xylR agu67A axeXA uxuA gly43E xyl3A uxuB
xyn10C gly2A uxaC xyn10A xypA xytA xypB gly43F
xylE locus
xytA locus
xylR locus
xytB locus

Figure R10

of the eight remaining genes are in putative transcription units that have a *xylR* motif upstream and three others are located beside genes induced by xylose (Table S3, Figure R9). This conservation concerns proteins involved in the removal of substitutions or xylo-oligosaccharides degradation as well as in glucuronate metabolism. XytA and XytB TBDTs are also well conserved as well as the entire *xytA locus*. The proteins involved in xylose metabolism are not present in *C. cresenctus* CB15. This result was expected since this strain has an alternative xylose degradation pathway [70]. Another major difference concerns xylanases. As observed for *Pxs*, Xyn10C is the only *Xcc* xylanase which is conserved in *C. cresenctus* CB15. Moreover, among the conserved proteins studied here, this protein is that which shows the lowest similarity (33% identity, 47% similarity) to *Xcc* proteins. Altogether, these data clearly suggest the existence of a common pathway to utilize xylan or xylo-oligosaccharides in *Xcc* and *C. cresenctus* CB15. Such an extensive conservation is displayed by other *Caulobacter* species and is also observed for other Alphaproteobacteria. This comprises *Asticcacaulis excentricus* CB48 and *Phenylbacterium zucineum* HLK1 strains belonging to the Caulobacteraceae as well as *Sphingobium japonicum* UT26S and *Hirschia baltica* ATCC49814 strains which are members of the Sphingomonadaceae and Hyphodomonadaceae families, respectively. Most of these bacteria live in aquatic habitats, with the exception of *S. japonicum* and *Caulobacter segnis* which were isolated from soil (Table S3). None of them has been reported to be pathogenic on plants. In all these bacteria, the homology not only concerns enzymes but also comprises inner membrane transporters and TBDTs. Most genes displaying similarities with *Xcc* genes are scattered in genomes of these Alphaproteobacteria (including that of *C. crescentus* CB15) with the exception of *H. baltica* where they belong to a large cluster (Figure R10) which also contains genes putatively involved in pectin degradation (data not shown). This cluster also includes several genes of *C. crescentus* CB15 induced by xylose but which are not present in *Xcc* (Table S3).

Conservation of *Xcc* xylan CUT sytem in Gammaproteobacteria reveals bacteria living in soil or associated with plant debris as well as symbionts of wood-boring bivalves

When this comparative study was carried out, in addition to Xanthomonadaceae species, 326 complete and 490 draft genome sequences representing 251 species of the Gammaproteobacteria class were available in GenBank database. However, only three species in this class displayed a significant and extensive conservation to *Xcc* Xylan/xylose CUT

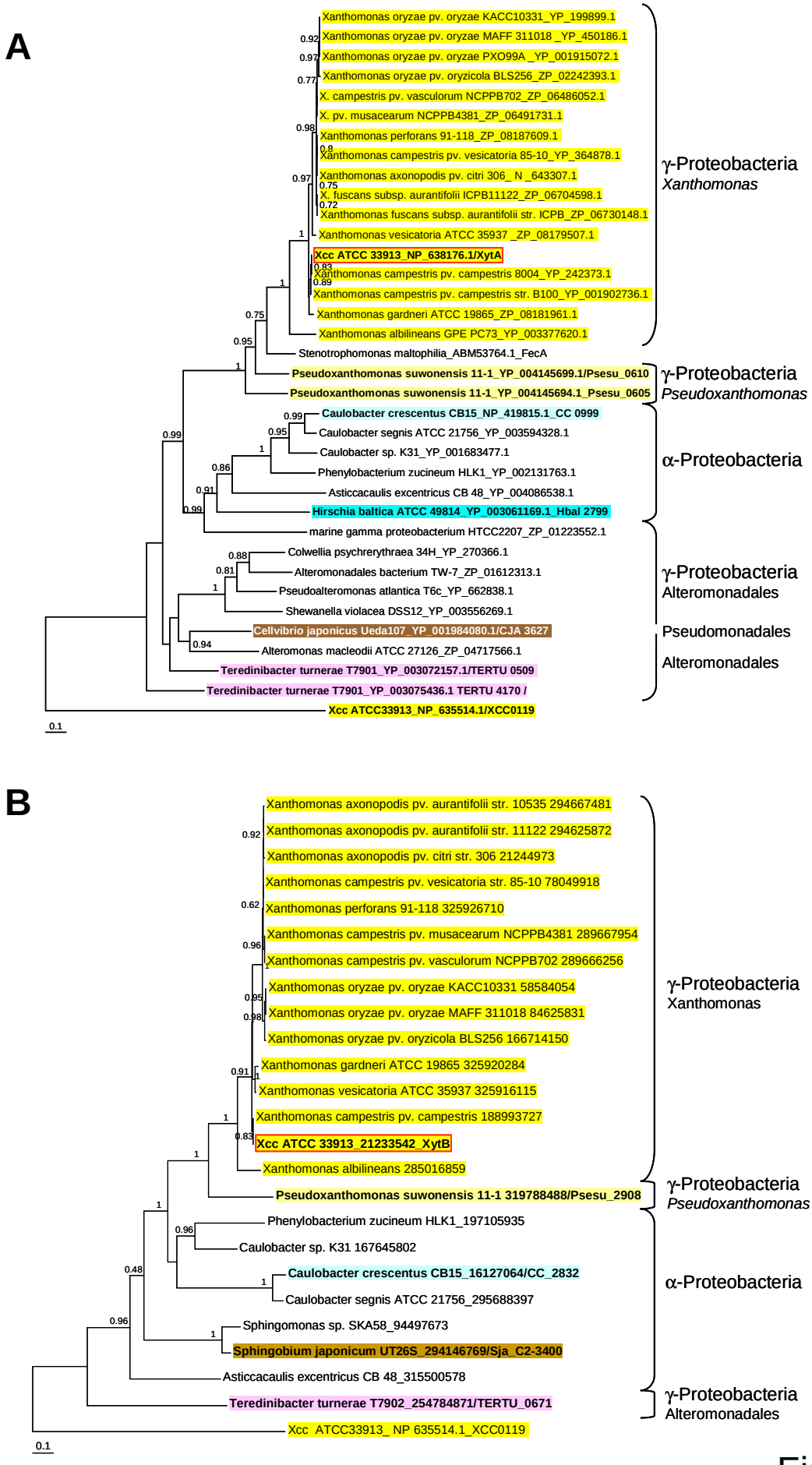

A
Xanthomonas oryzae pv. oryzae KACC10331_YP_199899.1
Xanthomonas oryzae pv. oryzae MAFF 311018 _YP_450186.1
Xanthomonas oryzae pv. oryzae PXO99A _YP_001915072.1
Xanthomonas oryzae pv. oryzicola BLS256_ZP_02242393.1
X. campestris pv. vasculorum NCPPB702_ZP_06486052.1
X. pv. musacearum NCPPB4381_ZP_06491731.1
Xanthomonas perforans 91-118_ZP_08187609.1
Xanthomonas campestris pv. vesicatoria 85-10_YP_364878.1
Xanthomonas axonopodis pv. citri 306_ N _643307.1
X. fuscans subsp. aurantifolii ICPB11122_ZP_06704598.1
Xanthomonas fuscans subsp. aurantifolii str. ICPB_ZP_06730148.1
Xanthomonas vesicatoria ATCC 35937 _ZP_08179507.1
Xcc ATCC 33913_NP_638176.1/XytA
Xanthomonas campestris pv. campestris 8004_YP_242373.1
Xanthomonas campestris pv. campestris str. B100_YP_001902736.1
Xanthomonas gardneri ATCC 19865_ZP_08181961.1
Xanthomonas albilineans GPE PC73_YP_003377620.1
γ-Proteobacteria
Xanthomonas
Stenotrophomonas maltophilia_ABM53764.1_FecA
Pseudoxanthomonas suwonensis 11-1_YP_004145699.1/Psesu_0610
Pseudoxanthomonas suwonensis 11-1_YP_004145694.1_Psesu_0605
γ-Proteobacteria
Pseudoxanthomonas
Caulobacter crescentus CB15_NP_419815.1_CC 0999
Caulobacter segnis ATCC 21756_YP_003594328.1
Caulobacter sp. K31_YP_001683477.1
Phenylobacterium zucineum HLK1_YP_002131763.1
Asticcacaulis excentricus CB 48_YP_004086538.1
Hirschia baltica ATCC 49814_YP_003061169.1_Hbal 2799
α-Proteobacteria
marine gamma proteobacterium HTCC2207_ZP_01223552.1
Colwellia psychrerythraea 34H_YP_270366.1
Alteromonadales bacterium TW-7_ZP_01612313.1
Pseudoalteromonas atlantica T6c_YP_662838.1
Shewanella violacea DSS12_YP_003556269.1
γ-Proteobacteria
Alteromonadales
Cellvibrio japonicus Ueda107_YP_001984080.1/CJA 3627
Alteromonas macleodii ATCC 27126_ZP_04717566.1
Teredinibacter turnerae T7901_YP_003072157.1/TERTU 0509
Teredinibacter turnerae T7901_YP_003075436.1 TERTU 4170 /
Pseudomonadales
Alteromonadales
Xcc ATCC33913_NP_635514.1/XCC0119
0.1
B
Xanthomonas axonopodis pv. aurantifolii str. 10535 294667481
Xanthomonas axonopodis pv. aurantifolii str. 11122 294625872
Xanthomonas axonopodis pv. citri str. 306 21244973
Xanthomonas campestris pv. vesicatoria str. 85-10 78049918
Xanthomonas perforans 91-118 325926710
Xanthomonas campestris pv. musacearum NCPPB4381 289667954
Xanthomonas campestris pv. vasculorum NCPPB702 289666256
Xanthomonas oryzae pv. oryzae KACC10331 58584054
Xanthomonas oryzae pv. oryzae MAFF 311018 84625831
Xanthomonas oryzae pv. oryzicola BLS256 166714150
Xanthomonas gardneri ATCC 19865 325920284
Xanthomonas vesicatoria ATCC 35937 325916115
Xanthomonas campestris pv. campestris 188993727
Xcc ATCC 33913_21233542_XytB
Xanthomonas albilineans 285016859
γ-Proteobacteria
Xanthomonas
Pseudoxanthomonas suwonensis 11-1 319788488/Psesu_2908
γ-Proteobacteria
Pseudoxanthomonas
Phenylobacterium zucineum HLK1_197105935
Caulobacter sp. K31 167645802
Caulobacter crescentus CB15_16127064/CC_2832
Caulobacter segnis ATCC 21756_295688397
α-Proteobacteria
Sphingomonas sp. SKA58_94497673
Sphingobium japonicum UT26S_294146769/Sja_C2-3400
Asticcacaulis excentricus CB 48_315500578
Teredinibacter turnerae T7902_254784871/TERTU_0671
γ-Proteobacteria
Alteromonadales
Xcc ATCC33913_ NP 635514.1_XCC0119
0.1

Figure S5

system (Table S3; Figure R9). The first two species, *Saccharophagus degradans* and *Teredinibacter turnerae* T7901 are closely related marine Alteromonadale bacteria. *S. degradans* which was isolated from a decaying plant in a watershed is able to degrade algae and higher plant materials, including xylan [71,72]. *T. turnerae* was isolated from the gills of wood-boring marine bivalves *Bankia gouldi* of the family *Teredinidae* (shipworms). Its genome encodes a large repertoire of plant cell wall degrading enzymes that may assist its host in using woody plant materials [73]. The third bacterium, *C. japonicus* strain Ueda107, was isolated from soil. It is intensively studied for its ability to degrade plant cell wall components and xylan in particular [38,74]. Interestingly, the cell wall degrading enzymes repertoire of these three bacteria was reported to be much conserved. Approximately 50% of the *C. japonicus* degradative enzymes are shared with *S. degradans* and *T. turnerae* [38]. Between eighteen and nineteen proteins of *Xcc* xylan CUT system are conserved in each of these bacteria, including enzymes involved in xylan and xylo-olygosaccararides degradation, xylose and glucuronate metabolism and inner membrane transporters. Genes coding for xylan-degrading enzymes homologous to proteins of the xylan CUT system of *Xcc* are scattered in the genomes of these three bacteria, as already observed for other plant cell wall degrading enzyme genes in *C. japonicus* [38]. Proteins displaying similarities to XytA and XytB are also conserved in these bacteria. However, the level of similarity of these homologs is lower than that observed with conserved TBDTs of Alphaproteobacteria (Table S3, Figure S5). XytB TBDT is nevertheless significantly conserved to Tertu_0671 from *T. turnerae* (Figure S5). Interestingly, XytA TBDT and *xytA* locus are also conserved in these three bacteria (Figure S5; Figure R9). It is also worth noting that two TBDTs induced by xylose in *C. crescentus* CB15 (CC_0442 and CC_2804) display significant similarity to TBDTs of *T. turnerae*, *S. degradans* or *C. japonicus* and that these homologs are located beside genes putatively involved in xylan degradation of these bacteria (Table S3, Figure R9).

As observed for *Pxs* and Alphaproteobacteria, the only *Xcc* xylanase which is significantly conserved in these bacteria is Xyn10C. The best homolog is found in *T. turnerae*. Xyn10C and this homologous protein, Tertu_3398, which displays 65% amino acid identity to one another, belong to the same cluster of family GH10 xylanases (Figure S6). The best homologous proteins found in *C. japonicus* and *S. degradans* are Xyn10D (Cja_2888) and Xyn10C (Sde_2633), respectively, which have been functionally characterized as endo-1,4-β-xylanases in both bacteria [72,75]. These two proteins are less conserved (~43-45% identity) to *Xcc* Xyn10C and are grouped with another putative xylanase of *T. turnerae* (Tertu_0736) in

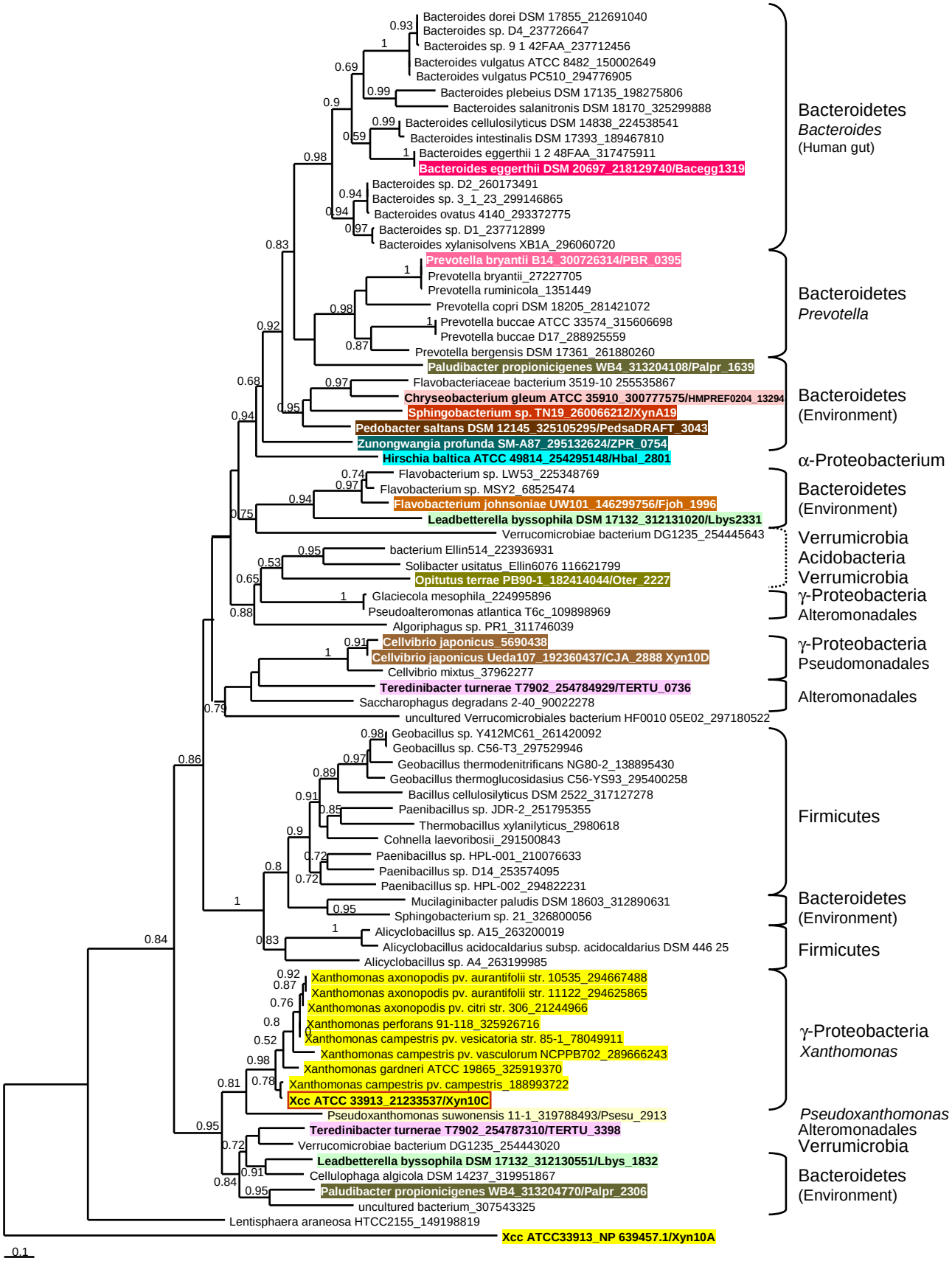

Bacteroides dorei DSM 17855_212691040
Bacteroides sp. D4_237726647
Bacteroides sp. 9 1 42FAA_237712456
Bacteroides vulgatus ATCC 8482_150002649
Bacteroides vulgatus PC510_294776905
Bacteroides plebeius DSM 17135_198275806
Bacteroides salanitronis DSM 18170_325299888
Bacteroides cellulosilyticus DSM 14838_224538541
Bacteroides intestinalis DSM 17393_189467810
Bacteroides eggerthii 1 2 48FAA_317475911
Bacteroides eggerthii DSM 20697_218129740/Bacegg1319
Bacteroides sp. D2_260173491
Bacteroides sp. 3_1_23_299146865
Bacteroides ovatus 4140_293372775
Bacteroides sp. D1_237712899
Bacteroides xylanisolvens XB1A_296060720
Prevotella bryantii B14_300726314/PBR_0395
Prevotella bryantii_27227705
Prevotella ruminicola_1351449
Prevotella copri DSM 18205_281421072
Prevotella buccae ATCC 33574_315606698
Prevotella buccae D17_288925559
Prevotella bergensis DSM 17361_261880260
Paludibacter propionicigenes WB4_313204108/Palpr_1639
Flavobacteriaceae bacterium 3519-10 255535867
Chryseobacterium gleum ATCC 35910_300777575/HMPREF0204_13294
Sphingobacterium sp. TN19_260066212/XynA19
Pedobacter saltans DSM 12145_325105295/PedsaDRAFT_3043
Zunongwangia profunda SM-A87_295132624/ZPR_0754
Hirschia baltica ATCC 49814_254295148/Hbal_2801
Flavobacterium sp. LW53_225348769
Flavobacterium sp. MSY2_68525474
Flavobacterium johnsoniae UW101_146299756/Fjoh_1996
Leadbetterella byssophila DSM 17132_312131020/Lbys2331
Verrucomicrobiae bacterium DG1235_254445643
bacterium Ellin514_223936931
Solibacter usitatus_Ellin6076 116621799
Opitutus terrae PB90-1_182414044/Oter_2227
Glaciecola mesophila_224995896
Pseudoalteromonas atlantica T6c_109898969
Algoriphagus sp. PR1_311746039
Cellvibrio japonicus_5690438
Cellvibrio japonicus Ueda107_192360437/CJA_2888 Xyn10D
Cellvibrio mixtus_37962277
Teredinibacter turnerae T7902_254784929/TERTU_0736
Saccharophagus degradans 2-40_90022278
uncultured Verrucomicrobiales bacterium HF0010 05E02_297180522
Geobacillus sp. Y412MC61_261420092
Geobacillus sp. C56-T3_297529946
Geobacillus thermodenitrificans NG80-2_138895430
Geobacillus thermoglucosidasius C56-YS93_295400258
Bacillus cellulosilyticus DSM 2522_317127278
Paenibacillus sp. JDR-2_251795355
Thermobacillus xylanilyticus_2980618
Cohnella laevoribosii_291500843
Paenibacillus sp. HPL-001_210076633
Paenibacillus sp. D14_253574095
Paenibacillus sp. HPL-002_294822231
Mucilaginibacter paludis DSM 18603_312890631
Sphingobacterium sp. 21_326800056
Alicyclobacillus sp. A15_263200019
Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446 25
Alicyclobacillus sp. A4_263199985
Xanthomonas axonopodis pv. aurantifolii str. 10535_294667488
Xanthomonas axonopodis pv. aurantifolii str. 11122_294625865
Xanthomonas axonopodis pv. citri str. 306_21244966
Xanthomonas perforans 91-118_325926716
Xanthomonas campestris pv. vesicatoria str. 85-1_78049911
Xanthomonas campestris pv. vasculorum NCPPB702_289666243
Xanthomonas gardneri ATCC 19865_325919370
Xanthomonas campestris pv. campestris_188993722
Xcc ATCC 33913_21233537/Xyn10C
Pseudoxanthomonas suwonensis 11-1_319788493/Psesu_2913
Teredinibacter turnerae T7902_254787310/TERTU_3398
Verrucomicrobiae bacterium DG1235_254443020
Leadbetterella byssophila DSM 17132_312130551/Lbys_1832
Cellulophaga algicola DSM 14237_319951867
Paludibacter propionicigenes WB4_313204770/Palpr_2306
uncultured bacterium_307543325
Lentisphaera araneosa HTCC2155_149198819
Xcc ATCC33913_NP 639457.1/Xyn10A
0.1
Bacteroidetes
Bacteroides
(Human gut)
Bacteroidetes
Prevotella
Bacteroidetes
(Environment)
α-Proteobacterium
Bacteroidetes
(Environment)
Verrumicrobia
Acidobacteria
Verrumicrobia
γ-Proteobacteria
Alteromonadales
γ-Proteobacteria
Pseudomonadales
Alteromonadales
Firmicutes
Bacteroidetes
(Environment)
Firmicutes
γ-Proteobacteria
Xanthomonas
Pseudoxanthomonas
Alteromonadales
Verrumicrobia
Bacteroidetes
(Environment)

Figure S6

a distinct branch (Figure S6). Therefore, altogether these homologies suggest that *C. japonicus*, *S. degradans* and *T. turnerae* also possess xylan CUT systems displaying similarities to that of *Xanthomonas* species.

The xylan CUT system of *Xanthomonas* displays significant homology to the xylanolytic machinery of Bacteroidetes from human gut, rumen and the environment.

Several proteins of the *Xcc* xylan/xylose CUT system display very high similarities to proteins of bacteria belonging to the Bacteroidetes phylum, the level of similarity ranging from 32 to 63% amino acid identity. These Bacteroidetes strains can be grouped into two main groups. The first group corresponds to bacteria belonging to *Bacteroides* spp. and *Prevotella* spp. which are dominant isolates found in the human gut and bovine rumen, respectively. The second group includes environmental Bacteroidetes.

Several genes coding for enzymes involved in xylan degradation have been identified and characterized in *Bacteroides* spp. and *Prevotella* spp (for review see [10]). Recently, transcriptomic studies comparing *P. bryantii* B_14 gene expression during growth in presence of soluble wheat arabinoxylan (WAX), arabinose or xylose, revealed fifty seven genes which expression is significantly higher with WAX as compared to xylose and arabinose. Interstingly, an operon of six genes, which includes two TBDT genes (*xusA* and *xusC*) and an endoxylanase gene (*xyn10C*), comprised the most highly induced genes in presence of WAX [33]. These six genes which form the *xus* cluster are widely conserved among human- and animal-associated *Bacteroides* spp. and *Prevotella* spp and it was suggested that this cluster constitutes the core set of genes required for xylan fragment uptake by gut-associated Bacteroidetes [10].

Ten proteins of *Xcc* xylan CUT system are highly conserved to *P. bryantii* B_14 proteome (Figure R10, Table S3). These conserved proteins include putative enzymes involved in xylan degradation (i.e. Abf51A, Agu67A Gly43E, Xyl3A, Xyn10C, Gly2A and Gly43F). XylE, XypA and XypB inner membrane transporters are also well conserved whereas XytA and XytB TBDT do not display homology to any proteins of *P. bryantii* B_14 even to XusA or XusB TBDTs. However, three proteins conserved to *Xcc* proteins, *PBR_0394*, *PBR_0395* and *PBR_0398*, are coded by genes that are located beside the *xus* locus in *P. bryantii* B_14 genome (Figure R10, Table S3). These three genes also correspond to genes which expression was higher in presence of WAX [33]. A fourth conserved gene, *PBR_0883*, located elsewhere in

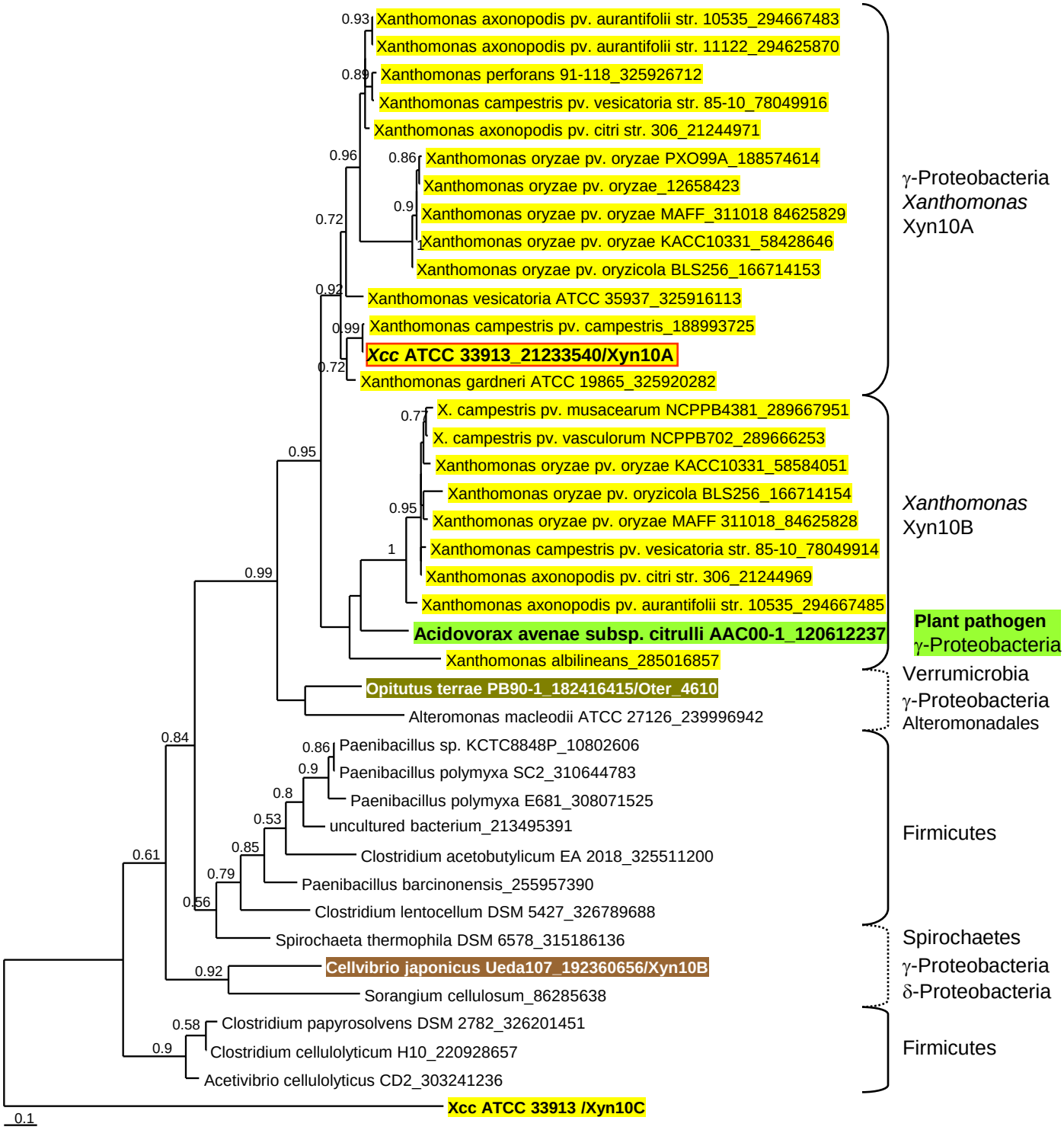

0.93
Xanthomonas axonopodis pv. aurantifolii str. 10535_294667483
Xanthomonas axonopodis pv. aurantifolii str. 11122_294625870
0.89
Xanthomonas perforans 91-118_325926712
Xanthomonas campestris pv. vesicatoria str. 85-10_78049916
Xanthomonas axonopodis pv. citri str. 306_21244971
0.96
0.86
Xanthomonas oryzae pv. oryzae PXO99A_188574614
Xanthomonas oryzae pv. oryzae_12658423
0.9
0.72
Xanthomonas oryzae pv. oryzae MAFF_311018 84625829
1
Xanthomonas oryzae pv. oryzae KACC10331_58428646
Xanthomonas oryzae pv. oryzicola BLS256_166714153
0.92
Xanthomonas vesicatoria ATCC 35937_325916113
0.99
Xanthomonas campestris pv. campestris_188993725
Xcc ATCC 33913_21233540/Xyn10A
0.72
Xanthomonas gardneri ATCC 19865_325920282
γ-Proteobacteria
Xanthomonas
Xyn10A
0.77
X. campestris pv. musacearum NCPPB4381_289667951
X. campestris pv. vasculorum NCPPB702_289666253
0.95
Xanthomonas oryzae pv. oryzae KACC10331_58584051
Xanthomonas oryzae pv. oryzicola BLS256_166714154
0.95
Xanthomonas oryzae pv. oryzae MAFF 311018_84625828
1
Xanthomonas campestris pv. vesicatoria str. 85-10_78049914
0.99
Xanthomonas axonopodis pv. citri str. 306_21244969
Xanthomonas axonopodis pv. aurantifolii str. 10535_294667485
Acidovorax avenae subsp. citrulli AAC00-1_120612237
Xanthomonas albilineans_285016857
Xanthomonas
Xyn10B
Plant pathogen
γ-Proteobacteria
Opitutus terrae PB90-1_182416415/Oter_4610
Alteromonas macleodii ATCC 27126_239996942
Verrumicrobia
γ-Proteobacteria
Alteromonadales
0.84
0.86
Paenibacillus sp. KCTC8848P_10802606
0.9
Paenibacillus polymyxa SC2_310644783
0.8
Paenibacillus polymyxa E681_308071525
0.53
uncultured bacterium_213495391
0.61
0.85
Clostridium acetobutylicum EA 2018_325511200
0.79
Paenibacillus barcinonensis_255957390
0.56
Clostridium lentocellum DSM 5427_326789688
Spirochaeta thermophila DSM 6578_315186136
Firmicutes
Cellvibrio japonicus Ueda107_192360656/Xyn10B
0.92
Sorangium cellulosum_86285638
Spirochaetes
γ-Proteobacteria
δ-Proteobacteria
0.58
Clostridium papyrosolvens DSM 2782_326201451
0.9
Clostridium cellulolyticum H10_220928657
Acetivibrio cellulolyticus CD2_303241236
Firmicutes
Xcc ATCC 33913 /Xyn10C
0.1

Figure S7

P. bryantii $B_1$4 genome also displayed a similar induction pattern. Once again, it is worth noting that as observed for Alphaproteobacteria and Gammaproteobacteria, Xyn10C is well conserved in *P. bryantii* $B_1$4 whereas Xyn10A, Xyn10B and Xyn30A are not (Table S3, Figure S6, S7).

We extended our comparative study to other *Prevotella* spp and *Bacteroides* spp. for which genome sequences were available. This analysis showed that the conservation observed with *P. bryantii* $B_1$4 is widespread among these bacteria. In several of them such as *Bacteroides eggerthii* DSM 20697, *Bacteroides ovatus* ATCC8483 or *B. plebeius* DSM 17135 for example, two or three additional *Xcc* proteins appeared to be conserved. These are XylA2 xylose isomerase, AxeXA putative acetyl esterase and UxuB putative fructuronate reductase. Moreover, we noticed that several conserved genes are clustered in some of these Bacteroides species such as *B. eggerthii* DSM 20697 and *B. ovatus* ATCC8483 (Table S3, Figure R10). In *B. eggerthii* DSM 20697, nine genes conserved to *Xcc* xylan CUT system belong to a large cluster which harbours thirty three genes putatively involved in plant cell wall degradation and which also includes the *xus* core xylan cluster (Figure R10). A similar cluster orgnaization was observed for clones obtained from a metagenomic study devoted to the characterization of xylanase activity in the human gut [76] (data not shown).

This high degree of conservation of *Xcc* Xylan CUT proteins was also observed with Bacteroidetes isolated from other environmental niches such as soil, plants, plant debris, aquatic environments, algae, choanoflagellate, human vagina and insect gut (Table S3). These additional isolates do not belong to Bacterodaceae or Prevotellaceae families and divide into 17 genera representing 7 families (Table S3). Several of these isolates, show the same or very similar conservation patterns as *B. eggerthii* DSM 20697, *B. ovatus* ATCC8483 or *P. bryantii* $B_1$4. The degree of conservation of proteins of these environmental bacteria to *Xcc* proteins was often higher than that observed for *Bacteroides* or *Prevotella* species (Table S3; Figures S7 and S8). Interestingly, as observed in *B. eggerthii* DSM 20697, several of the conserved genes are located beside or in the close vicinity of genes displaying very high similarities with the six genes of the *xus* cluster. They belong to large clusters encompassing genes involved in xylan degradation as well as other plant cell wall components. This includes *Paludibacter propionicigenes* WB4 isolated from rice plant residues, *Pedobacter saltans* DSM 12145 obtained from soil, *Mucilaginibacter paludis* DSM 18603 isolated from acidic Sphagnum peat bog, *Spirosoma linguale* DSM 74 obtained from laboratory water bath and *Zunongwangia*

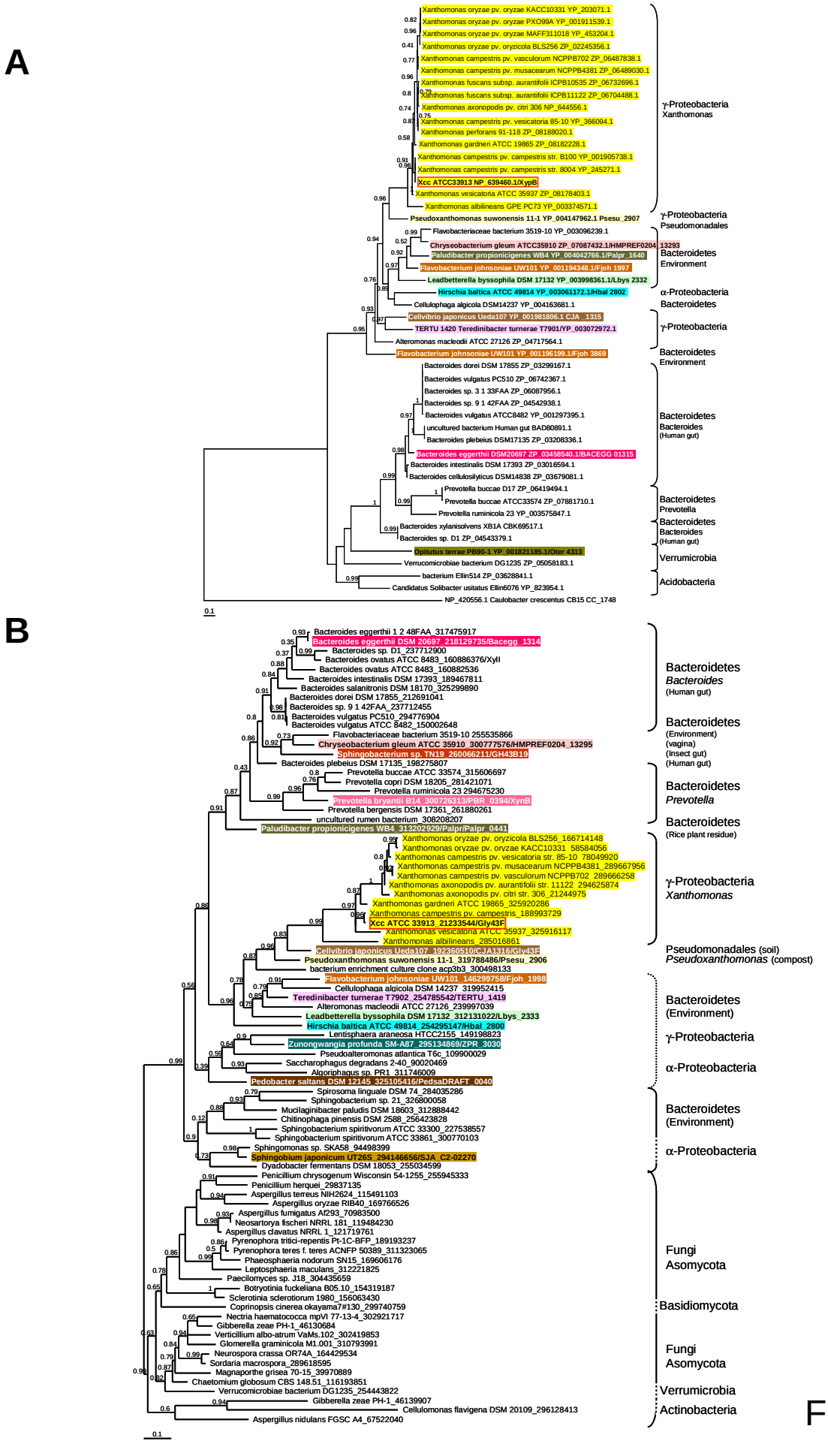

A
Xanthomonas oryzae pv. oryzae KACC10331 YP_203071.1
Xanthomonas oryzae pv. oryzae PXO99A YP_001911539.1
Xanthomonas oryzae pv. oryzae MAFF311018 YP_453204.1
Xanthomonas oryzae pv. oryzicola BLS256 ZP_02245356.1
Xanthomonas campestris pv. vasculorum NCPPB702 ZP_06487838.1
Xanthomonas campestris pv. musacearum NCPPB4381 ZP_06489030.1
Xanthomonas fuscans subsp. aurantifolii ICPB10535 ZP_06732696.1
Xanthomonas fuscans subsp. aurantifolii ICPB11122 ZP_06704488.1
Xanthomonas axonopodis pv. citri 306 NP_644556.1
Xanthomonas campestris pv. vesicatoria 85-10 YP_366094.1
Xanthomonas perforans 91-118 ZP_08188020.1
Xanthomonas gardneri ATCC 19865 ZP_08182228.1
Xanthomonas campestris pv. campestris str. B100 YP_001905738.1
Xanthomonas campestris pv. campestris str. 8004 YP_245271.1
Xcc ATCC33913 NP_639460.1/XypB
Xanthomonas vesicatoria ATCC 35937 ZP_08178403.1
Xanthomonas albilineans GPE PC73 YP_003374571.1
γ-Proteobacteria
Xanthomonas
Pseudoxanthomonas suwonensis 11-1 YP_004147962.1 Psesu_2907
γ-Proteobacteria
Pseudomonadales
Flavobacteriaceae bacterium 3519-10 YP_003096239.1
Chryseobacterium gleum ATCC35910 ZP_07087432.1/HMPREF0204_13293
Paludibacter propionicigenes WB4 YP_004042766.1/Palpr_1640
Flavobacterium johnsoniae UW101 YP_001194348.1/Fjoh 1997
Leadbetterella byssophila DSM 17132 YP_003998361.1/Lbys 2332
Bacteroidetes
Environment
Hirschia baltica ATCC 49814 YP_003061172.1/Hbal 2802
α-Proteobacteria
Cellulophaga algicola DSM14237 YP_004163681.1
Bacteroidetes
Cellvibrio japonicus Ueda107 YP_001981806.1 CJA_1315
TERTU 1420 Teredinibacter turnerae T7901/YP_003072972.1
γ-Proteobacteria
Alteromonas macleodii ATCC 27126 ZP_04717564.1
Flavobacterium johnsoniae UW101 YP_001196199.1/Fjoh 3869
Bacteroidetes
Environment
Bacteroides dorei DSM 17855 ZP_03299167.1
Bacteroides vulgatus PC510 ZP_06742367.1
Bacteroides sp. 3 1 33FAA ZP_06087956.1
Bacteroides sp. 9 1 42FAA ZP_04542938.1
Bacteroides vulgatus ATCC8482 YP_001297395.1
uncultured bacterium Human gut BAD80891.1
Bacteroides plebeius DSM17135 ZP_03208336.1
Bacteroides eggerthii DSM20697 ZP_03458540.1/BACEGG 01315
Bacteroides intestinalis DSM 17393 ZP_03016594.1
Bacteroides cellulosilyticus DSM14838 ZP_03679081.1
Bacteroidetes
Bacteroides
(Human gut)
Prevotella buccae D17 ZP_06419494.1
Prevotella buccae ATCC33574 ZP_07881710.1
Prevotella ruminicola 23 YP_003575847.1
Bacteroidetes
Prevotella
Bacteroides xylanisolvens XB1A CBK69517.1
Bacteroides sp. D1 ZP_04543379.1
Bacteroidetes
Bacteroides
(Human gut)
Opitutus terrae PB90-1 YP_001821185.1/Oter 4313
Verrumicrobia
Verrucomicrobiae bacterium DG1235 ZP_05058183.1
bacterium Ellin514 ZP_03628841.1
Candidatus Solibacter usitatus Ellin6076 YP_823954.1
Acidobacteria
NP_420556.1 Caulobacter crescentus CB15 CC_1748
0.1
B
Bacteroides eggerthii 1 2 48FAA_317475917
Bacteroides eggerthii DSM 20697_218129735/Bacegg_1314
Bacteroides sp. D1_237712900
Bacteroides ovatus ATCC 8483_160886376/XylI
Bacteroides ovatus ATCC 8483_160882536
Bacteroides intestinalis DSM 17393_189467811
Bacteroides salanitronis DSM 18170_325299890
Bacteroides dorei DSM 17855_212691041
Bacteroides sp. 9 1 42FAA_237712455
Bacteroides vulgatus PC510_294776904
Bacteroides vulgatus ATCC 8482_150002648
Bacteroidetes
Bacteroides
(Human gut)
Flavobacteriaceae bacterium 3519-10 255535866
Chryseobacterium gleum ATCC 35910_300777576/HMPREF0204_13295
Sphingobacterium sp. TN19_260065211/GH43B19
Bacteroides plebeius DSM 19275807
Bacteroidetes
(Environment)
(vagina)
(Insect gut)
(Human gut)
Prevotella buccae ATCC 33574_315606697
Prevotella copri DSM 18205_281421071
Prevotella ruminicola 23 294675230
Prevotella bryantii B14_300726313/PBR_0394/XynB
Prevotella bergensis DSM 17361_261880261
uncultured rumen bacterium_308208207
Bacteroidetes
Prevotella
Paludibacter propionicigenes WB4_316204928/Palpr/Palpr_0441
Bacteroidetes
(Rice plant residue)
Xanthomonas oryzae pv. oryzicola BLS256_166714148
Xanthomonas oryzae pv. oryzae KACC10331_58584056
Xanthomonas campestris pv. vesicatoria str. 85-10_78049920
Xanthomonas campestris pv. musacearum NCPPB4381_289667956
Xanthomonas campestris pv. vasculorum NCPPB702_289666258
Xanthomonas axonopodis pv. aurantifolii str. 11122_294625874
Xanthomonas axonopodis pv. citri str. 306_21244975
Xanthomonas gardneri ATCC 19865_325920286
Xanthomonas campestris pv. campestris_188993729
Xcc ATCC 33913_21233544/Gly43F
Xanthomonas vesicatoria ATCC 35937_325916117
Xanthomonas albilineans_285016861
γ-Proteobacteria
Xanthomonas
Cellvibrio japonicus Ueda107_192360510/CJA1316/Gly43F
Pseudoxanthomonas suwonensis 11-1_319788486/Psesu_2906
bacterium enrichment culture clone acp3b3_300498133
Pseudomonadales (soil)
Pseudoxanthomonas (compost)
Flavobacterium johnsoniae UW101_146299758/Fjoh_1998
Cellulophaga algicola DSM 14237_319952415
Teredinibacter turnerae T7902_254785542/TERTU_1419
Alteromonas macleodii ATCC 27126_239997039
Leadbetterella byssophila DSM 17132_312131022/Lbys_2333
Bacteroidetes
(Environment)
Hirschia baltica ATCC 49814_254295147/Hbal_2800
Lentisphaera araneosa HTCC2155_149198823
γ-Proteobacteria
Zunongwangia profunda SM-A87_295134869/ZPR_3030
Pseudoalteromonas atlantica T6c_109900029
Saccharophagus degradans 2-40_90020469
Algoriphagus sp. PR1_311746009
α-Proteobacteria
Pedobacter saltans DSM 12145_325105416/PedsaDRAFT_0040
Spirosoma linguale DSM 74_284035286
Sphingobacterium sp. 21_326800058
Mucilaginibacter paludis DSM 18603_312888442
Chitinophaga pinensis DSM 2588_256423828
Sphingobacterium spiritivorum ATCC 33300_227538557
Sphingobacterium spiritivorum ATCC 33861_300770103
Sphingomonas sp. SKA58_94498399
Sphingobium japonicum UT26S_294146656/SJA_C2-02270
Dyadobacter fermentans DSM 18053_255034599
Bacteroidetes
(Environment)
α-Proteobacteria
Penicillium chrysogenum Wisconsin 54-1255_255945333
Penicillium herquei_29837135
Aspergillus terreus NIH2624_115491103
Aspergillus oryzae RIB40_169766526
Aspergillus fumigatus Af293_70983500
Neosartorya fischeri NRRL 181_119484230
Aspergillus clavatus NRRL 1_121719761
Pyrenophora tritici-repentis Pt-1C-BFP_189193237
Pyrenophora teres f. teres ACNFP 50389_311323065
Phaeosphaeria nodorum SN15_169606176
Leptosphaeria maculans_312221825
Paecilomyces sp. J18_304435659
Botryotinia fuckeliana B05.10_154319187
Sclerotinia sclerotiorum 1980_156063430
Fungi
Asomycota
Coprinopsis cinerea okayama7#130_299740759
Basidiomycota
Nectria haematococca mpVI 77-13-4_302921717
Gibberella zeae PH-1_46130684
Verticillium albo-atrum VaMs.102_302419853
Glomerella graminicola M1.001_310793991
Neurospora crassa OR74A_164429534
Sordaria macrospora_289618595
Magnaporthe grisea 70-15_39970889
Chaetomium globosum CBS 148.51_116193851
Fungi
Asomycota
Verrucomicrobiae bacterium DG1235_254443822
Verrumicrobia
Gibberella zeae PH-1_46139907
Cellulomonas flavigena DSM 20109_296128413
Actinobacteria
Aspergillus nidulans FGSC A4_67522040
0.1

Figure S8

profunda SMA87 isolated from deep sea sediment. *xus* orthologs and *Xcc* conserved genes were also found on a cloned chromosomal fragment which contains a xylanase gene (XynA19) from *Sphingobacterium* sp. TN19, a symbiotic strain isolated from the gut of *Batocera horsfieldi* (*Coleoptera*) larvae [77,78] (Figure R10). Altogether, these data suggest that the *xus* xylan core cluster is widespread in Bacteroidetes phylum (Figure R10). Conservation of proteins of the xylan CUT system of *Xcc* was also found in other Bacteroidetes strains which do not harbour a *xus* locus. These bacteria comprise *Chryseobacterium gleum* which was obtained from vaginal swab, *Flavobacterium johnsoniae* UW101 isolated from soil [79] or *Leadbetterella byssophyla* DSM17132 obtained from cotton waste compounds (Figure R10). In these bacteria also, the genes conserved to *Xcc* xylan CUT system belong to cluster of genes coding for plant cell wall degrading enzymes. Interestingly these clusters also encompasses TBDT of the SusC family which are different from XusA or XusC TBDTs. Interestingly, these new TBDTs are also conserved among Bacteroidetes species, including *P. bryantii* $B_1$4 and *B. eggerthii* DSM 20697. In this latter species, these conserved TBDT genes are located in the large "xylan" cluster described above (Figure R10).

The xylan CUT system is conserved in Verrucomicrobia and Acidobacteria and reveals a new TBDT family.

Approximately 10 proteins of the *Xcc* Xylan CUT system display significant similarity to proteins of *Opitutus terrae* PB90-1, which was isolated from a rice paddy soil [80] and *Verrucomicrobiae bacterium* DG1235 which was obtained from the microflora of the marine dinoflagellate *Scrippseilla trochidea* (Table S3). These two strains belong to the recently described deeply branched phylum Verrucomicrobia which includes species isolated from soil, freshwater and human feces [80,81]. The *Xcc* conserved gene set mostly corresponds to that conserved to Bacteroidetes species. A detailed study of *O. terrae* PB90-1 homologous genes showed that three of them are located in a cluster which is partially conserved to the large cluster of *C. gleum* (Figure R10). Moreover, five of the other conserved genes are located beside four genes (*oter_1754*, *oter_175*, *oter_3377* and *oter_4287*) which code for putative outer membrane proteins (OMP), as predicted by using Pepsort software [82]. These four putative OMP proteins possess a putative plug domain (PF07715) which is specific to TBDTs. However, these four proteins do not display a typical β-barrel domain detectable by the PF00593 hidden markov model (HMM) which is also specific to TBDTs [83]. However, secondary structure analyses using PSIPRED [84,85] revealed that these proteins possess

several putative β-sheets and that their 12 C-terminal amino acids form a typical membrane anchoring β-sheet with the last residue being aromatic, as it is the rule for OMP proteins [86]. None of these four proteins display significant similarities to xylan-associated TBDTs or other TBDTs characterized in Bacteroidetes or Proteobacteria. However, they are conserved to a large set of putative OMP proteins of *O. terrae* and *V. bacterium* DG1235 (data not shown). Thus, these putative OMPs may represent a new class of TBDTs specific to the phylum Verrucomicrobia.

Finally, a similar conservation was observed with the proteome of *Candidatus Solibacter usitatus* Ellin6076, an Acidobacteria isolated from the soil of grazed pasture [87]. Four of the conserved genes in this strain are located in a large cluster which contains a putative OMP protein (Acid_2685) which displays some degree of similarity to the Oar subclass of TBDTs [83] (data not shown).

Toward the characterization of a core conserved set of proteins required for Xylan degradation and of Xanthomonas specific xylanases

The integration of data obtained by performing these comparative studies combined with phylogentic analyses allowed us to class the proteins of *Xcc* xylan cut system into three main groups based on their conservation pattern among the different phyla.

The first group consists of proteins which are highly conserved (amino acid identity ranging from 45 to 70%) in Proteobacteria, Bacteroidetes, Verrucomicrobia and Acidobacteria phyla. This group comprises XCC1191 (Abf51A), XCC4102 (Agu67A), XCC4105 (Gly43E), XCC4115 (Xyn10C), XCC4116 (Gly2A) and XCC4122 (Gly43F) involved in the degradation of xylan backbone and the release of side chains as well as enzymes involved in xylo-oligosaccharides hydrolysis to generate xylose. It also contains the inner membrane transporters XCC4119 (XypA) and XCC4121 (XypB). These eight genes might represent a core set of proteins required for xylan deconstruction.

Members of the second group are proteins that are specifically conserved in Proteobacteria, including Caulobacterales. This group comprises XytB TBDT and the four proteins of *xytA locus*.

Finally, the third group comprises Xyn10A, Xyn10B and Xyn30A xylanases which are specifically found in *Xanthomonas* species. Interestingly, these xylanases are closely

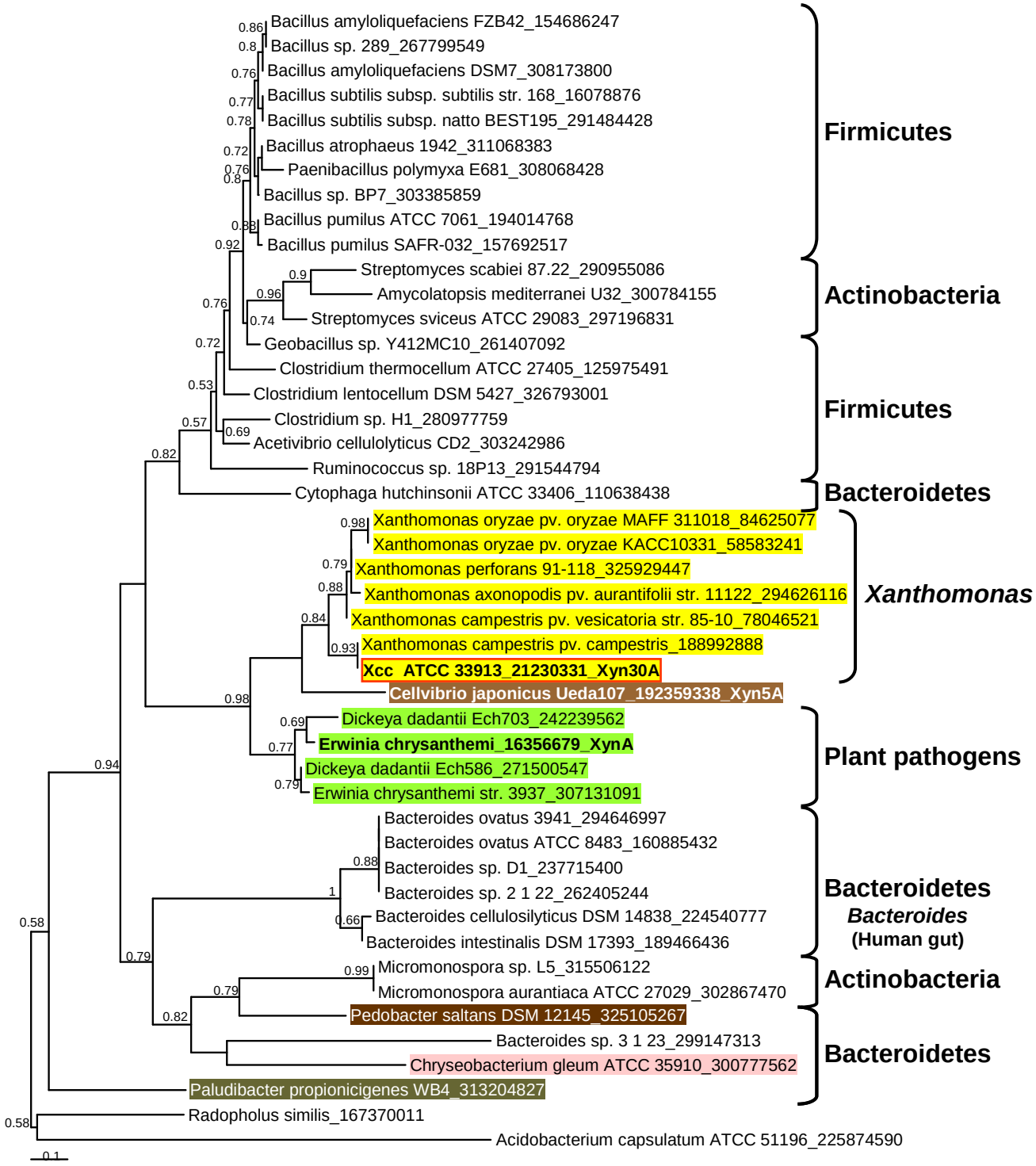

Bacillus amyloliquefaciens FZB42_154686247
Bacillus sp. 289_267799549
Bacillus amyloliquefaciens DSM7_308173800
Bacillus subtilis subsp. subtilis str. 168_16078876
Bacillus subtilis subsp. natto BEST195_291484428
Bacillus atrophaeus 1942_311068383
Paenibacillus polymyxa E681_308068428
Bacillus sp. BP7_303385859
Bacillus pumilus ATCC 7061_194014768
Bacillus pumilus SAFR-032_157692517
Streptomyces scabiei 87.22_290955086
Amycolatopsis mediterranei U32_300784155
Streptomyces sviceus ATCC 29083_297196831
Geobacillus sp. Y412MC10_261407092
Clostridium thermocellum ATCC 27405_125975491
Clostridium lentocellum DSM 5427_326793001
Clostridium sp. H1_280977759
Acetivibrio cellulolyticus CD2_303242986
Ruminococcus sp. 18P13_291544794
Cytophaga hutchinsonii ATCC 33406_110638438
Xanthomonas oryzae pv. oryzae MAFF 311018_84625077
Xanthomonas oryzae pv. oryzae KACC10331_58583241
Xanthomonas perforans 91-118_325929447
Xanthomonas axonopodis pv. aurantifolii str. 11122_294626116
Xanthomonas campestris pv. vesicatoria str. 85-10_78046521
Xanthomonas campestris pv. campestris_188992888
Xcc_ATCC 33913_21230331_Xyn30A
Cellvibrio japonicus Ueda107_192359338_Xyn5A
Dickeya dadantii Ech703_242239562
Erwinia chrysanthemi_16356679_XynA
Dickeya dadantii Ech586_271500547
Erwinia chrysanthemi str. 3937_307131091
Bacteroides ovatus 3941_294646997
Bacteroides ovatus ATCC 8483_160885432
Bacteroides sp. D1_237715400
Bacteroides sp. 2 1 22_262405244
Bacteroides cellulosilyticus DSM 14838_224540777
Bacteroides intestinalis DSM 17393_189466436
Micromonospora sp. L5_315506122
Micromonospora aurantiaca ATCC 27029_302867470
Pedobacter saltans DSM 12145_325105267
Bacteroides sp. 3 1 23_299147313
Chryseobacterium gleum ATCC 35910_300777562
Paludibacter propionicigenes WB4_313204827
Radopholus similis_167370011
Acidobacterium capsulatum ATCC 51196_225874590
0.1
Firmicutes
Actinobacteria
Firmicutes
Bacteroidetes
Xanthomonas
Plant pathogens
Bacteroidetes
Bacteroides
(Human gut)
Actinobacteria
Bacteroidetes

Figure S9

conserved to xylanases present in other plant pathogenic bacteria. Xyn10A and Xyn10B which are conserved to one another cluster with a putative xylanase from the plant pathogen *Acidovorax avenae* subsp. *citrulli* AAC00-1 (Figure S7) whereas Xyn30A is conserved to XynA xylanase from phytopathogenic *Dickeya* species (formely *Erwinia*) (Figure S9).

DISCUSSION

Xylans are a diverse group of polysaccharides which are the dominating noncellulosic polysaccharides in the cell wall of plants. They account approximately for 20-30% and 40-60% of the dicot and monocot cell walls, respectively [4,88]. They are also the most abundant polysaccharides in nature after cellulose [89]. Their natural degradation is mediated mainly by microorganisms and many xylanolytic enzymes have been characterized from numerous fungi and bacteria that are found either free in nature or in the digestive tract of animals [7,90]. Extracellular xylanolytic activities have also been identified in plant pathogenic bacteria such as *Erwinia chrysanthemi* strains (now belonging to *Dickeya* genus) isolated from corn [91], and two *Xanthomonas* species, *Xcv* and *Xoo* [11,12]. In both *Xanthomonas* species, xylanase genes were shown to be required for full pathogenicity on host plants. In this study, we show that a third *Xanthomonas* species, *Xcc*, produces an extracellular xylanase activity. Comparative genome studies and functional approaches carried out in *Xcc* showed that these xylanases belong to a complex xylanolytic CUT system which is highly conserved among *Xanthomonas* species. This system comprises enzymes for the degradation of xylan backbone and the removal of side groups, as well as enzymes involved in catabolism of glucuronate and xylose. It also harbours three inner membrane transporters and two TBDTs. Most genes encoding these proteins are located into 4 loci *xylE*, *xytA*, *xylR* or *xytB* loci. The *xylE* locus comprises genes involved in xylose transport and metabolism whereas the other three loci are devoted to xylan deconstruction and utilization as well as glucuronic acid catabolism.

The *Xcc* xylan CUT system is differentially induced by xylose, xylo-oligosaccharides and xylan and is repressed by XylR

The expression of genes belonging to *xylE* locus on one hand and to *xylR*, *xytA* and *xytB* loci on the other hand display a different induction pattern. Although the expression of genes of the four loci is induced by xylose, genes of *xylR*, *xytA* and *xytB* loci are specifically and highly activated by xylan and small oligo-saccharides (X_2-X_4). This differential expression pattern suggests the existence of two interwoven regulatory pathways controlled by xylose and xylo-

oligosaccharides. Genes of the *xytA* and *xytB* loci as well as *xylA2* which belong to *xylR* locus are under the repression of XylR. The repression mediated by this repressor is strictly correlated with the presence of a *xyl*-box motif in the promoter region of these genes or operons. Interestingly, XylR from *Xcc* is very closely related to XylR from *C. crescentus* CB15, an oligotrophic α-Proteobacterium found in aquatic environments. In this latter bacterium, XylR binds an operator sequence which is very similar to the palindromic motif of *Xanthomonas* spp. *xyl*-box. In *C. crescentus* CB15, it has been established that operator binding by XylR is greatly reduced in the presence of xylose, and a simple regulatory model has been proposed in which XylR blocks xylose-induced promoters in the absence of xylose [34]. When this monomer is present, the promoter is freed thus allowing expression of xylose-induced genes [34]. In *C. crescentus* CB15, the XylR-regulon and xylose-induced genes seems to be tightly interconnected. This regulon comprises approximately fifty one genes [32] among which eighteen are conserved to *Xcc* xylan CUT system. This regulon also includes genes involved in xylose metabolism, which form a specific and original pathway in *C. crescentus* CB15 [70] which differs from the classical pathway encoded by *xylE* locus in *Xcc*. Although the expression of *xylA1*, *xylB* and *xylE* genes of *xylE* locus is induced by xylose these genes are not repressed by XylR in *Xcc*. However, *xylA2* gene which is highly conserved to *xylA1* xylose isomerase gene is negatively regulated by XylR but also highly induced by xylo-oligosaccharides. It thus appears that the regulation exerted by xylose in *Xcc* is more complicated than that described in *C. crescentus* CB15. What molecule is recognized by XylR in *Xcc* is not known yet, but by analogy to *C. crescentus* CB15 we could speculate that xylose is the inducer of XylR repressor. However, we cannot rule out that xylo-oligosaccharides are also recognized by XylR and constitute better inducers than xylose. This is perhaps also the case in *C. crescentus* CB15. It would therefore be interesting to test and compare the effect of xylo-oligosaccharides and xylose on XylR repression in both *Xcc* and *C. crescentus* CB15. However, in *Xcc*, other data suggest the existence of at least two interconnected regulatory pathways mediated by xylose and xylo-oligosaccharides. In *xytB* locus, *xytB*, *xypB* and *gly43F* form an operon which is repressed by XylR. However, our results with xylo-oligosaccharides and xylan as well as xylanase activity tests, suggest that in presence of these compounds, *xypB* and *gly43F* form an internal operon which is transcribed independently of *xytB* transcription. Morevover, *agu67A*, *axeXA*, *uxuA*, *gly43E*, *xyl3A* and *uxuB* genes which are contiguous in *xylR* locus are not repressed by XylR but are induced by xylo-oligosaccharides. These observations suggest the occurrence of a specific regulatory

pathway mediated by xylo-oligosaccharides beside XylR and xylose regulatory pathways. The complexity of this regulation is underscored by the observation that *agu67A* gene which is induced by xylo-oligosaccharides is not induced by xylan unlike most other genes of *xytA*, *xytB* or *xylR* loci. The expression of this gene was studied in *agu67*::pVO insertion mutant, thus suggesting that its integrity is required for correct expression of the entire system. This hypothesis is strengthened by the fact that the extracellular xylanase activity was significantly reduced in this mutant. This gene codes for an enzyme putatively involved in the removal of side chains from xylan backbone. It is tempting to speculate that it thus participates in the production of inducing molecules, including xylo-oligosaccharides. This observation suggests the existence of a complex regulatory pathway involving different degradation steps to generate an array of inducing molecules. In the soil bacterium *C. japonicus*, the regulation of xylan-degrading system is controlled by such a complex regulatory system. Emami *et al*. [92] have shown that the five xylanase genes of this bacterium are differentially regulated over time. They proposed that this temporal regulation was the consequence of a spatiotemporal cascade of degradation events generating specific hydrolysis products that might influence the activity of definite regulatory proteins. This hypothesis was recently confirmed by the finding that AbfR/S two components regulatory system of *C. japonicus* is required for the induction of enzymes that remove the side chains of xylan but not of xylanase genes that hydrolyze the xylan backbone [74]. AbfS sensor domain which is probably located into the periplasm specifically binds to decorated arabinoxylo-oligosaccharides but not to undecorated xylo-oligosaccharides. Such a regulation of xylan degrading system by decorated xylo-oligosaccharides was also reported for *Geobacillus stearothermophilus*, another soil bacterium. In this Gram-positive bacterium, UxuR repressor of the GntR family represses the expression of genes coding for the degradation of glucuronoxylan, glucuronic acid catabolism and a transport system. The most probable inducer of this repressor is aldotetrauronic acid ($MeGAX_3$) a decorated xylo-oligosaccharide generated by the degradation of xylan [43]. In this bacterium too, there is another level of regulation which seems to depend on undecorated xylo-oligosaccharides. The *xynDCEFG* genes form a transcriptional unit which codes for XynDC two component system and XynEFG ATP-binding cassette (ABC) transport system which is involved in xylo-oligosaccharide transport [93]. This 6-kb transcriptional unit is repressed by XylR xylose repressor of NagC family which inducer is xylose. However, *xynEFG*, the distal genes of this transcriptional unit, form an operon which is positively induced by XynC regulator. This finding is reminiscent of what was observed for *xytB* locus

in *Xcc*, where *xytA*, *xypA* and *gly43F* seem to constitute a unique transcriptional unit repressed by XylR whereas *xypA* and *gly43F* form an internal operon in presence of xylo-oligosaccharides or xylan. This analogy between *Xcc* and *G. stearothermophilus* is in favour of the existence in *Xcc* of two regulatory pathways, one mediated by xylose and the other by xylo-oligosaccharides. In *P. bryantii* B_14, xylanase genes are also induced by xylo-oligosaccharides, but in this case, the induction is mediated by medium- to large-sized xylo-oligosaccharides rather than by small xylo-oligosaccharides [94]. This induction is controlled by XynR, a hybrid two component system regulator, which is most probably located in the inner membrane [10,95]. Therefore, as observed in these xylanolytic bacteria, it is probable that the induction of xylan utilization genes operated by xylo-oligosaccharides involves additional regulatory genes in *Xcc*. We did not find in *Xcc* any proteins displaying significant homologies to UxuR, XynDC regulators from *G. stearothermophilus*, AbfR/S from *C. japonicus* or XynR from *P. bryantii* B14. Work is now under progress to find in *Xcc* other putative regulators and to test other putative inducer molecules such as decorated xylo-oligosaccharides.

Xylanase production and regulation by HrpG and HrpX T3SS regulators

The xylan CUT system of *Xanthomonas* spp. encompasses four endo-xylanases, Xyn10A, Xyn10B, Xyn10C and Xyn30A which distribution is variable among species. Surprisingly, the extracellular xylanase activities detected in *Xcc, Xoo* and *Xcv* are controlled by different xylanases. In *Xcv* the extracellular activity is mediated by Xyn30A xylanase of family GH30, whereas it controlled by Xyn10B in *Xoo* and Xyn10A in *Xcc*, which both belong to family GH10. *xyn10B* gene is absent in *Xcc* whereas *xyn10A* has no signal peptide in *Xoo* strains. This might explain why Xyn10A- or Xyn10B-associated extracellular activities were not observed in *Xoo* and *Xcc*, respectively. *xyn30A* gene is present in *Xcc*, but we did not detect any extracellular activity associated with this gene in this strain. The Xyn30A protein of *Xcc* and *Xcv* (XCV0965, XynC) align well (81% amino acid identity, 90% similarity over entire polypeptide length) and they both possess a signal peptide. Two xylanases of family GH30 have been studied in details, XynA from *E. chrysanthemi* and XynC from *Bacillus subtilis* 168 [91,96,97]. Xyn30A proteins display a very significant homology to XynA from *E. chrysanthemi* (58% amino acid identity, 76 % similarity over entire polypeptide length) and phylogenetic analyses showed that they cluster together in the same and specific branch (Figure S9). The seven residues, including catalytic residues, which are conserved in GH30

xylanases [96] are present in Xyn10A proteins of *Xcc* and *Xcv* suggesting that they are both active. The major difference which is notable between these two species concerns *xyn30A* gene expression. In *Xcv*, *xyn30A* is repressed by HrpG and HrpX, the master regulators of *Xanthomonas* T3SS, whereas it seems that this is not the case in *Xcc.* It is worth noting that Szczesny *et al.* [12] did not reveal any activity associated to *xyn10A* gene in *Xcv*, although this gene is present and seems to be functional in this strain. Xyn10A proteins from *Xcv* and *Xcc* share 83% amino acid identity over their entire length; they both possess a signal peptide and display the putative catalytic acid base and nucleophile glutamic acid residues conserved in members of family GH10 [98]. These observations suggest that *Xanthomonas* species favour different xylanases to degrading xylan. However, it is possible that the *in vitro* tests used to detect these enzymatic activities do not correspond to the biological reality of the interaction with plant. Further work is needed to clarify the differences observed between *Xcc*, *Xoo* and *Xcv*. However, this point is not trivial, since xylanases of families GH30 and GH10 display different substrate specificities and generate different hydrolysis products. GH10 family is one of the best studied xylanase family. GH10 endoxylanases are active on decorated substrates such as arabinoxylan, glucuronoxylan, methyglucuronoxylan and glucuronoarabinoxylan as well as on xylo-oligosaccharides. They generate short xylo-oligosaccharides including xylobiose, xylotriose, aldouronates (methyl-D-glucuronoxylo-oligosaccharides) such as the aldotetrauronate, $MeGAX_3$ and probably other decorated xylo-oligosaccharides [7,74,98,99]. However, the presence of 4-*O*-methylglucuronic acid and arabinofuranose side chains usually hinder the binding and hydrolysis of xylan by GH10 enzymes [100]. For this reason, enzymes that remove decorations, such as α-1,2-glucuronidase, arabinofuranosidase and acyl esterase are required for complete utilization of these complex molecules [97]. GH30 endoxylanases unlike xylanase from GH10 have specificity for methylglucuronoxylans. They have specificity for sites of 4-*O*-methylglucuronic acid (MeGA) substitutions and generate xylooligosaccharides having a single MeGA substitution. In this case MeGA is substituted penultimate to the reducing terminus whereas it is located on the terminal nonreducing end xylose unit in aldouronates (including $MeGAX_3$) generated by xylanases of family GH10 [7]. In *E. chrysanthemi*, the most prevalent hydrolysis products obtained with XynA were either 3 to 14 xylose residues in length [91,96,101]. This length is higher than that generally observed for GH10 xylanases and it was proposed that XynA may generate substrate available for other enzymes instead of producing xylo-oligosaccharides that can be directly metabolised by the bacterium [91].

Beside Xyn10A and Xyn30A, *Xcc* possesses another xylanase, Xyn10C. No extracellular activity associated to this enzyme was detected although it harbours a signal peptide. Phylogenetic analyses showed that it is not grouped with Xyn10A and Xyn10B xylanases (Figure S6 and S7). It is more closely related to Xyn10D from *C. japonicus* and Xyn10A (formely XynC) from the related bacterium *Cellvibrio mixtus* subsp. *cellulosa*. Interestingly, in *C. mixtus* subsp. *cellulosa* Xyn10A is periplasmic [75] and it was proposed that this is also the case for Xyn10D in *C. japonicus* [92,98]. Therefore, we can speculate that Xyn10C is also cell associated in *Xcc*, thus, explaining the inability to associate an extracellular activity to this xylanase. Interestingly, although the expression of *xyn10C* follows the pattern of induction of other genes of the xylan CUT system, it is also activated by HrpG and HrpX regulators. This is the only gene of the xylan CUT system which is positively regulated at the transcriptional level by these regulators in *Xcc*. Other genes coding for plant cell wall degradative enzymes have been shown to be positively regulated by HrpG and HrpX *in Xcc*. *pghAxc* and *pghBxc* genes coding for two extracellular polygalacturonases secreted by the Xps T2SS machinery of *Xcc* (strain 8004) have been shown to be activated by HrpX and HrpG regulators [19]. This is also the case for *cysP2* gene of *Xoo*, which encodes a protease also secreted via the Xps T2SS [24]. Recently, Szczesny *et al*. [12] showed that in *Xcv*, the Xps system itself is positively regulated by HrpG and HrpX. Moreover, these authors showed that a functional Xps T2SS is required for optimal translocation of effectors by the T3SS. They suggested that enzymes secreted by this system could facilitate the assembly of extracellular components of the T3SS. If a xylanase activity is required for this process, Xyn10C could be a candidate. However, its putative localization in the periplasm is not in agreement with a role in the extracellular medium. Further work is necessary to study the exact location of this protein in *Xanthomonas* species. If this protein is really extracellular it will be interesting to understand why we could not detect any activity associated to this enzyme. One possibility being that this protein is not correctly expressed in our xylanase test conditions. Indeed, we observed that mutations in *hrpX* and *hrpG* led to a clear increase in Xyn10A-mediated extracellular xylanase activity, thus suggesting that these regulatory genes repress the expression of this gene in our tests. We did not detect any negative effect of HrpG and HrpX on *xyn10A* expression in MME minimal medium or even on RBB-xylan plates (data not shown). This data, suggest that there is another level of regulation of xylanase activity which is still unclear. However, these results illustrate the complexity of regulation of the xylan CUT system, as well as the interplay between T3SS and T2SS [21]. The degradation

of plant cell wall and the setting up of T3SS must be tightly connected and regulated to coordinate the correct assembly and functioning of the T3SS machinery, a process which might involve plant cell wall degrading enzymes. At the same time T3SS effectors might be required for the suppression of plant defence induced by plant cell wall hydrolysis products [21,102,103].

Agu67A and Gly43F modulate extracellular xylanase activity.

Beside xylanase genes, the xylan CUT system comprises genes coding for proteins putatively involved in the degradation of side chains. Among these proteins, AxeXA may represent a new family of acyl esterases. Homologs of this protein are found in cluster of genes putatively involved in the degradation of xylan and other plant cell wall components carried by several bacteria including *B. eggerthii*, *C. gleum*, *S. japonicus* or *H. baltica*. This observation suggests that AxeXA is member of a large family of putative esterase related to the deconstructrion of plant cell wall components.

In *Xcc*, *axeXA* gene is located downstream, *agu67A*, another gene putatively involved in the removal of side chain from xylan backbone. This gene codes for a putative α-D-glucuronidase of family GH67 which displays significant similarities to GlcA67A from *C. japonicus* and AguA from *G. stearothermophilus*. These enzymes cleave the α-1,2-glycosidic bond of 4-*O*-methylglucuronic acid (MeGA) and the xylose moiety located at the non-reducing end of xylan or xylo-oligosaccharide backbone [7]. Thus, these enzymes are active on aldouronates, including $MeGAX_3$, produced by the hydrolysis of xylan by GH10 xylanases but not on hydrolysis products generated by GH30 xylanases. In *C. japonicus* GlcA67A hydrolyses 4-*O*-methyl-D-glucuronoxylooligosaccharides but not 4-*O*-methyl-D-glucuronoxylan [104]. This observation and the substrate specificity of this enzyme have suggested that α-glucuronidases function downstream of xylanases to deconstruct xylan [105]. However, biochemical studies have shown that substitutions along the xylan chain reduce the ability of GH10 xylanases to degrade substituted substrates [97]. In *Xcc*, we observed that the extracellular xylanase activity is significantly reduced in *agu67A* mutants. As mentioned above, it seems that *agu67A* gene is also required for its own induction and probably by extension for the entire system. Therefore, our results confirm the existence of a synergy between Agu67A and Xyn10A enzymes: the degradation of xylan mediated by Xyn10A must generate specific substrates for Agu67A enzyme. The activity of Agu67A on hydrolysis products generated by Xyn10A, including xylo-oligosaccharides, might be required for full induction of the Xylan

CUT system and therefore the production of Xyn10A, thus explaining the observed reduction of extracellular activity in *agu67A* mutant. However we cannot rule out that Agu67A potentiates the activity of Xyn10A on xylan. It will be interesting to see whether the transcription of genes of the xylan CUT system and the production of Xyn10A is affected by *agu67A* mutation. Moreover, it will be important to determine the location of Agu67A to validate our hypothesis. This enzyme possesses a signal peptide, suggesting that it could be extracellular like Xyn10A. In *C. japonicus* GlcA67A also has a signal peptide but it is localized on the outer membrane by a mechanism which is not clarified yet. This location has strengthened the hypothesis suggesting that GlcA67A acts downstream of xylanases which are secreted in the medium [104]. It will therefore be important to determine the exact location of Agu67A in *Xcc*. The importance of glucuronic acid liberation during xylan degradation for the bacterium is underscored by the presence of enzymes involved in the metabolism of this carboxylic acid in *xylR* and *xytB* loci. In *G. stearothermophilus*, genes required for glucuronic acid utilization are also located in a xylan degradation cluster [43].

Another putative enzyme, Gly43F of the *Xcc* xylan CUT system is required for normal production of extracellular xylanase activity. This protein is closely related to XynD from *P. bryantii* $B_1$4 and Xsa from *Bacteroides ovatus*. Our phylogenetic analysis confirms that these proteins form a specific and distinct branch from GH43 family as previously reported by Gasparic *et al*. [41]. These authors showed that XynD is an intracellular exoxylanase releasing xylose progressively from the nonreducing ends of substrates ranging in size from xylan polysaccharides down to xylobiose. It is not active on $MeGAX_3$ and it was proposed to specifically degrade transported xylo-oligosaccharides inside the cells. In *Xcc*, mutation of *gly43F* led to an important increase of extracellular xylanase activity. Morevover, the growth of *gly43F* mutants was significantly and specifically reduced in presence of xylo-oligosaccharides. A simple hypothesis can be put forward to explain this dual property. In presence of xylan or xylo-oligosaccharides, this enzyme degrades transported xylo-oligosaccharides to xylose in the cytoplasm, which promotes bacterial growth. In these conditions there might be equilibrium between the regulation mediated by xylo-oligosaccharides their degradation and the derepression controlled by XylR. When *gly43F* is mutated, transported xylo-oligosaccahrides are not processed anymore and there is disequilibrium in the xylose/xylo-oligosaccahride balance and the accumulation of inducing xylo-oligosaccharides promotes a high level of induction of the system and of Xyn10A in particular. In these mutants, xylose is not produced and the growth of bacteria is impaired.

This hypothesis is in agreement with the existence of two interconnected regulatory pathways mediated by xylose and xylo-oligosaccharides in which Gly43F may play a central role by maintaining a balance between the production of inducing molecules and xylose.

Xylo-oligosaccharides and xylose are transported by different inner membrane transporters.

It is also clear that the transport of molecules generated by the hydrolysis of xylan also constitutes a key point linking regulation and xylan degradation. Although xylanolytic enzymes have been studied extensively, little is known about the transport systems involved in the uptake of hydrolysis products. In Gram positive bacteria, ABC transporters seem to be involved in xylan utilization and transport. Several, ABC transporter genes are specifically induced by xylan in the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus* [106]. In *G. stearothermophilus*, the XynEFG ABC transport system has been shown to transport xylo-oligosaccharides [93]. In Gram negative bacteria, the situation seems more complex since two membranes must be crossed. Moreover, some degradative enzymes are located into the periplasm as well as sensor domains of two components regulators embedded in the inner membrane such as XynR from *P. bryantii* $B_1$4 and AbfS from *C. japonicus*. Therefore, the substrates transported across the outer membrane are not necessarily the same as those transported through the inner membrane. Furthermore, they might be different signalling pathways related to molecules transported into the periplasm or into the cytoplasm. In this article we have identified two inner membrane transporters, XylE which is involved in xylose transport and XypB which is required for xylo-oligosaccharides uptake. Our data suggest that XylE is not the unique transporter for xylose in *Xcc* whearas XypB is the major transporter of small xylo-oligosaccharides (X_2-X_4). This transporter is required for the induction of the xylan CUT system by X_2, X_3, X_4 and xylan. It is also indispensable for growth on xylo-oligosaccharides. Interestingly, *xypB* is located just upstream of *gly43F* exoxylanase gene in *xytB* locus. Moreover, these two genes seem to form an operon specifically induced by these xylo-oligosaccharides. Therefore, this suggests that there is a tight connexion between xylo-oligosaccharides transport and degradation which are central for both signalling and physiology of *Xcc*. The importance of these two genes in the degradation of xylan is underscored by the fact that these two proteins are widely and highly conserved in numerous xylanolytic bacteria distributed among Gram negative xylanolytic bacteria (see below). Moreover, in these bacteria, the two homologous genes are very often

contiguous or close to one another and located in xylan utilization clusters. This is the case for example in *C. japonicus*, *T. turnerea*, *H. baltica*, *L. byssophila*, *C. gleum*, *Cellulophaga algicola*, *B. eggerthii*, *Bacteroides plebeius*, and several other *Bacteroides* species. Interestingly, in several of these bacteria these two genes are separated by or located beside a gene homologous to *xyn10C* xylanase gene. It is therefore tempting to speculate that the activity of this xylanase, which could be located in the periplasm generates a substrate for XypB and Gly43F.

XytA and XytB TBDTs are associated with xylan utilization.

The transport of xylo-ologosaccharides across the inner membrane raises the question about the presence of these molecules in the periplasmic space. In other words, how are these xylo-oligosaccharides or their precursors, transported across the outer membrane? The presence of *xytA* and *xytB* TBDT genes in xylan utilization loci of *Xanthomonas* species and the fact that these two genes are highly induced by xylo-oligosaccharides, suggest that these outer membrane transporters are involved in the transport of some xylan hydrolysis products. This prompted us to define the *Xanthomonas* xylanolytic system as a CUT system. Our data obtained by growing *xytA*, *xytB* or *xytAxytB* double mutant in presence of X_2, X_3 or X_4 showed that they belong to loci involved for optimal growth in presence of this oligomers. These loci are also required for epiphytic growth of *Xcc* on plant leaves. However, complementation experiments suggest that these two TBDTs are not directly required for the transport of xylo-oligosaccharides in our test conditions. It is worth noting that in *Xcc*, we obtained similar results with SuxA TBDT which transports sucrose: we did not observe any difference in growth capacities of a *suxA* mutant and the wild-type strain in minimal medium supplemented with sucrose. However, transport experiments with ^{14}C-sucrose carried out at lower sucrose concentrations clearly showed that this TBDT mediates the active uptake of sucrose with a very high affinity, which was estimated to be 3000-fold higher than that of sucrose porins [25]. Therefore, although our experiments did not show any differences with the wild-type, it is possible that XytA or XytB transports xylo-oligosaccharides. Another alternative could be that these two TBDTs transport other hydrolysis products that are the precursor of xylo-oligosaccharides which are taken up by XypB. It is worth noting that *xytB* gene is located upstream of *xypB* and *gly43F* and that these three genes seem to form an operon repressed by XylR. Similarly, *xytA* is the first gene of another operon which is also controlled by XylR and induced by xylo-ologosaccharides. This operon comprises three other

genes, *xyaA*, *xyaB*, *xyaC* which function in the xylan degradation pathway is still elusive. However, a pVO155 insertion into *xytA* affects the growth of the mutant on xylo-oligosaccharides as compared to the wild-type strain. The growth defect is not as pronounced as that observed with *xypB* or *gly43F* mutants but it is significant and similar to that obtained with *xytB* mutants. The importance of *xytA* locus is also highlighted by the fact that it is and entirely conserved in *C. crescentus* CB15 and other related Alphaproteobacteria as well as in *C. japonicus*. In *C. crescentus* CB15 the homologs of *xytA* and *xyaA* are induced by xylose and probably repressed by XylR. Moreover, the homologous locus of *H. baltica* is located within the large xylan degradation cluster harboured by this bacterium. We also noticed that *Xcc* has another locus (*XCC3963* to *XCC3966*) which has the same organization and is well conserved to *xytA* locus. Although, the TBDT gene of this second locus is not induced by xylan or xylose [25] we cannot rule out that it is involved in xylan utilization and partially redundant with *xytA* locus. Moreover, we previously reported that loci displaying a similar organization are widespread among bacteria degrading complex carbohydrates suggesting that they may play a very important function in these bacteria [25]. The importance of xytA and xytB operons is undersocre

Towards a core set of genes for xylan utilization involving TBDTs.

A large proportion of genes belonging to *Xcc* xylan CUT system, including *xytA* and *xytB* TBDT genes is conserved in *C. crescentus* and related Alphaproteobacteria. These bacteria were isolated either from aquatic environment or soil. They generally possess a wide array of plant cell wall degrading enzymes and are therefore equipped to exploit plant debris or compounds released in their environment. Surprisingly, the *Xcc* CUT system, including XytA and XytB TBDTs is only conserved in few Gammaproteobacteria outside the Xanthomonadaceae family. This comprises, *C. japonicu*s soil bacterium, *S. degradans* isolated from a decaying plant and *T. turnerae*, an intracellular endosymbiont of wood-boring bivalves. These bacteria are known for degrading various plant cell wall components [38,71,73]. Our comparative analysis show that these Alphaproteobacteria and Gammaproteobacteria share a common xylanolytic CUT system. Surprisingly, the CUT system of Xanthomonads seems to be more conserved to that of Caulobacterales and other Alphaproteobacteria than to that of Gamma-proteobacteria which are phylogentically more closely related to *Xanthomonas* genus [107]. This point is particularly illustrated by the fact that XylR regulator and xylose operator regions are specifically conserved among

Xanthomonas spp. and Caulobacterales. These observations could suggest that these systems were exchanged through horizontal transfer events. The presence of large cluster encompassing almost all of the genes of this system in *H. baltica* could be in agreement with this hypothesis. Interestingly, a recent phylogenetic study of prokaryotes based on whole proteome comparison has transferred the Xanthomonadaceae family from the base of the Gammaproteobacteria to the base of the Betaproteobacteria, thus suggesting that this family is more closely related to Alphaproteobacteria than to Gammaproteobacteria [108]. This new classification could explain the relatedness observed between *Xanthomonas* and *Caulobacter* CUT systems.

Interestingly, a specific subset of eight proteins of *Xcc* xylan CUT system, conserved among Proteobacteria, is also highly conserved in bacteria belonging to very different phyla such as Bacteroidetes, Verrucomicrobia and Acidobacteria. This subset of conserved proteins comprises enzymes for xylan deconstruction as well as XypA and XypB inner membrane transporters. It is worth noting that, Gly43F, XypB and Xyn10C, which play a key role in connecting xylan degradation and xylose production appear in this set of conserved genes. The wide distribution and the high level of conservation of these proteins suggest that they may represent a core set of proteins required for xylan utilization among bacteria belonging to these phyla.

TBDTs are associated with xylan utilization in numerous Gram negative bacteria

In *P. Bryantii* $B_1$4, most of the conserved genes described above are specifically induced by WAX and belong to the xylan utilization system carried by this xylanolytic Bacteroidete. Interestingly, four of these conserved genes are located beside the *xus* locus which was proposed to constitute the core set of genes required for xylan and xylan fragments binding and uptake in Bacteroidetes [10,33]. Interstingly, this locus comprises two TBDTs belonging to SusC TBDT family which is specific to Bacteroidetes [83,109]. The *xus* cluster also harbours two *susD*-like genes which are generally found downstream of *susC*-like TBDT genes. They code for proteins that are anchored on the external leaflet of the outer membrane and that assist SusC-like TBDT for transporting substrate molecules [109]. These proteins are also specific to Bacteroidetes. A previous comparative study has shown that the xylanolytic machinery of *P. bryantii*, including *xus* cluster, is conserved among *Bacteroides* and *Prevotella* species found in human and ruminant colons. Our comparative study shows that the *xus* locus and the xylanolytic machinery of *Prevotella* and *Bacteroides* genera is

widespread in the Bacteroidete phylum and is present in other genera of this family isolated from various environments, including soil, aquatic milieu, plant, plant debris or the gut of wood-borer insects. Thus, it appears that the co-occurrence of TBDTs genes and xylan degradation genes is not restricted to Proteobacteria but is also a shared feature of Bacteroidetes. Our comparative analysis suggests that this is also true for bacteria of Acidobacteria and Verucomicrobia phyla which form deep branches in bacterial phylogeny [107,108]. Interestingly, although a core set of enzymes and inner membrane transporters are conserved among these bacteria, the TBDTs associated with xylan utilization seem specific to each phylum. This is particularly illustrated by the fact that SusD-like proteins are not found in Proteobacteria, Verrucomicrobia or Acidobacteria. These differences suggest a convergent evolution of these systems which certainly underscores the importance of TBDT transporters in the utilization of xylan. All the bacteria identified in this comparative study share the ability to degrade and/or to utilize plant cell wall components as part of their lifestyle. They generally possess a large array of cell wall degrading enzymes as well as TBDTs over-representation (Table S3)[25,30]. A recent study has shown coevolution of some TBDT families with other families involved in metabolism of different types of saccharides and endopolygalacturonases [110]. These data suggest that TBDTs might be involved in the utilization of other plant cell wall components. In agreement with this hypothesis, in *Xcc* we identified three TBDT genes which are induced by polygalacturonate. This induction pattern and genome context analysis suggested that they might be involved in pectin utilization [25]. Moreover, in Bacteroidetes SusCD-like systems belong to polysaccharide utilization loci (PULs) involved in the metabolism of complex glycans coming from the diet or the intestinal mucosa [109]. These observations suggest the existence of numerous CUT/PUL systems involved in the exploitation of various plant molecules. The role played by TBDTs in these systems may be pivotal by allowing the binding and uptake of large hydrolysis products with a very high affinity. Such a high affinity transport may play an important role in signalling by inducing the system in presence of scarce amounts of inducing molecules which might be crucial for oligotrophic bacteria such as Caulobacterales. The coordinated expression of this active transport and of hydrolytic enzymes may optimize the utilization of substrate molecules. Moreover, the uptake of large molecules allowed by TBDTs probably prevents the release of monosaccharides in the medium that could be used by other microorganisms. These systems may thus represent a crucial competitive advantage for these bacteria in their ecosystems. In the case of phytopathogenic bacteria, such systems may also play a role in

minimizing the diffusion of plant cell wall degradation fragments which could trigger defense mechanisms [102,103].

***Xanthomonas* spp. xylanases are specifically conserved among phytopathogenic bacteria.** Xylanases of *Xoo* and *Xcv* have been shown to be required for full pathogenicity. However, in our case we could not find any implication of this system in the virulence of *Xcc*. It has already been observed that plant cell wall degradation systems are often redundant and that multiple mutants have to be constructed to see a detectable effect on virulence [11,111]. The type of pathogenicity tests and the host plant that were used to perform these experiments might also explain the absence of visible differences, since these tests are crude and bypass several steps of the natural infectious process. We therefore used another assay by assessing the growth of *Xcc* strains on host or non-host leaves. We showed that *xytA* and *xytB* loci are required for optimal phyllospheric multiplication. This observation suggests that *Xanthomonas* xylan CUT system may have multiple functions in the lifecycle of these pathogens. They may control epiphytic growth, virulence and conceivably the survival on plant debris after the death of diseased plant tissues. This latter point is probably highlighted by the fact that the xylan CUT system is conserved in *P. suwonensis* strain 11-1, a bacterium which is closely related to *Xanthomonas* genus. This strain was isolated from compost-feedstock and to our knowledge it is not phytopathogenic. The conservation of *Xanthomonas* xylan CUT system in this bacterium suggests that it is ancient and was present in the ancestor of the *Xanthomonadaceae* family. However, there is a major difference between *P. suwonenesis* and *Xanthomonas* xylan CUT systems. Indeed, Xyn10A, Xyn10B and Xyn30A xylanases are not present in *P. suwonensis* and inversely this bacterium possesses xylanases which are not found in *Xanthomonas* spp. studied so far. Interestingly, Xyn10B and Xyn30A which were shown to control virulence of *Xoo* and *Xcv* respectively, do not belong to the core set of xylan degradation genes but are specifically conserved to xylanases of other phytopathogenic bacteria, *A. avenae* subsp. *citrulli* AAC00-1 and *Dickeya* spp. (formely *Erwinia*), respectively. Although these bacteria, which belong to the γ–Proteobacteria family harbours other genes conserved to *Xanthomonas* spp. xylan CUT system, their level of similiraty to *Xcc* proteins is generally lower than that observed with α-Proteobacteria, Bacteroidetes or Verrucomicrobia homologs. Moreover, XytA and XytB seem not to be conserved in both of these bacteria. Therefore, this specific conservation of xylanases among these three phytopathogenic species suggests that they might have a specific and important

Table S4. List of plasmids and *Xanthomonas campestris* pv. *campestris* strains used or generated in this study.

Plasmids	Features	*Xcc* sequence cloned relative to the putative start codon	Reference
pVO155	pUC119 derivative, containing the promoterless *gus* (*uidA*) reporter gene encoding β-glucuronidase, used for insertion mutagenesis; Km^R Amp^R		[51]
pFAJ1700	pTR102-derived expression vector, containing a multiple cloning site and transcriptional terminators in both orientations; Tet^R Amp^R		[118]
pCZ750	pFAJ1700 containing the *Kpn* I-*Asc* I *lacZ* gene from the pCZ367 plasmid; Tet^R Amp^R		[25]
pCZ962	pCZ750 derivative, containing rrn terminator sequence and new restriction sites; Tet^R Amp^R		This study
pCZ917	pFAJ1700 derivative, containing 2094 bp of pSC150 with *LacI* gene, Ptac and T7 terminator; Tet^R Amp^R		[31]
pCZ1016	pCZ917 derivative, containing Ptac and T7 terminator, and without *LacI* gene; Tet^R Amp^R		This study
pk18*mobsacB*	mobilizable pBR322 derivative containing a genetically modified *sacB* gene; Km^R		[115]
pCZ1034	pK18mobsacB derivative with the MCS replaced by a Ptac promoter, a MCS and a T7 terminator flanked by a 700 bp fragment corresponding to the region upstream of the open reading frame *XCC0127* and a 700 bp fragment corresponding to the region downstream of the open reading frame *XCC0128* ; Km^R		This study
pPr-*xytB;* pPr-*XCC4120*	pCZ962-Pr-*XCC4120* ; Tet^R Amp^R	from -1100 to +1	This study
pPr-*xytA;* pPr-*XCC2828*	pCZ962-Pr-*XCC2828* ; Tet^R Amp^R	from -1100 to +1	This study
pC-*xylR;* pC-*XCC4101*	pCZ1016-*XCC4101* ; Tet^R Amp^R	from -28 to stop	This study
pC-*agu67A* ; pC-*XCC4102*	pCZ1016-*XCC4102* ; Tet^R Amp^R	from -30 to stop	This study
pC-*xyn10A* ; pC-*XCC4118*	pCZ1016-*XCC4118* ; Tet^R Amp^R	from -30 to stop	This study
pC-*xypA* ; pC-*XCC4119*	pCZ1016-*XCC4119* ; Tet^R Amp^R	from -37 to stop	This study
pC-*xytB* ; pC-*XCC4120*	pCZ1016-*XCC4120* ; Tet^R Amp^R	from -1110 to stop	This study
pC-*xypB* ; pC-*XCC4121*	pCZ1016-*XCC4121* ; Tet^R Amp^R	from -30 to stop	This study
pC-*gly43F* ; pC-*XCC4122*	pCZ1016-*XCC4122* ; Tet^R Amp^R	from -18 to stop	This study
pC-*xylE* ; pC-*XCC1759*	pCZ1016-*XCC1759* ; Tet^R Amp^R	from -30 to stop	This study

Strains	Genotype and/or phenotype	Location of insertion or deletion relative to the putative start codon	Reference
Xcc 568	Wild-type strain; Rifampycin resistant strain derivative of *Xanthomonas campestris* pv. *campestris* LMG568/ATCC33913		[55]
XP133	*XCC4099* ::pVO155; Rif^R Km^R	+102	This study
XP134	*xylA2* ::pVO*; XCC4100* ::pVO155; Rif^R Km^R	+64	This study
XP135	*xylR* ::pVO*; XCC4101* ::pVO155; Rif^R Km^R	+538	This study
XP136	*agu67A* ::pVO*; XCC4102* ::pVO155; Rif^R Km^R	+184	This study
XP137	*axeXA* ::pVO*; XCC4103* ::pVO155; Rif^R Km^R	+80	This study
XP138	*uxuA* ::pVO*; XCC4104* ::pVO155; Rif^R Km^R	+131	This study
XP139	*gly43E* ::pVO*; XCC4105* ::pVO155; Rif^R Km^R	+110	This study
XP140	*xyl3A* ::pVO*; XCC4106* ::pVO155; Rif^R Km^R	+168	This study
XP141	*uxuB* ::pVO*; XCC4107* ::pVO155; Rif^R Km^R	+58	This study
XP142	*xyn10C* ::pVO*; XCC4115* ::pVO155; Rif^R Km^R	+68	This study
XP143	*gly2A* ::pVO*; XCC4116* ::pVO155; Rif^R Km^R	+133	This study
XP144	*uxaC* ::pVO*; XCC4117* ::pVO155; Rif^R Km^R	+63	This study
XP145	*xyn10A* ::pVO*; XCC4118* ::pVO155; Rif^R Km^R	+78	This study
XP146	*xypA* ::pVO*; XCC4119* ::pVO155; Rif^R Km^R	+121	This study
XP068	*xytB* ::pVO*; XCC4120* ::pVO155; Rif^R Km^R	+1217	[25]
XP147	*xypB* ::pVO*; XCC4121* ::pVO155; Rif^R Km^R	+171	This study
XP148	*gly43F* ::pVO*; XCC4122* ::pVO155; Rif^R Km^R	+87	This study
XP149	*XCC4123* ::pVO155; Rif^R Km^R	+22	This study
XP038	*xytA* ::pVO*; XCC2828* ::pVO155; Rif^R Km^R	+2871	[25]
XP150	*xylE* ::pVO*; XCC1759* ::pVO155; Rif^R Km^R	+120	This study
XP151	*xyn30A* ::pVO*; XCC0857* ::pVO155; Rif^R Km^R	+103	This study
XP152	WT/pPr-*xytB; Xcc* 568/pPr-*XCC4120* ; Rif^R Tet^R Amp^R		This study
XP153	*xylR* ::pVO/pPr-*xytB; XCC4101* ::pVO155 pPr-*XCC4120* ; Rif^R Km^R Tet^R Amp^R		This study
XP154	Δ*xyn10C* ; Δ*XCC4115* ; Rif^R	from start to stop	This study
XP155	Δ*xyn10A* ; Δ*XCC4118* ; Rif^R	from start to stop	This study
XP156	Δ*xypA* ; Δ*XCC4119* ; Rif^R	from start to stop	This study
XP157	Δ*xytB1* ; Δ*XCC4120 1* ; Rif^R	from +63 to stop	This study
XP158	Δ*xytB2* ; Δ*XCC4120 2* ; Rif^R	from +226 to stop	This study
XP159	Δ*xytB3* ; Δ*XCC4120 3* ; Rif^R	from +360 to +1105	This study
XP160	Δ*xypB* ; Δ*XCC4121* ; Rif^R	from start to stop	This study
XP161	Δ*gly43F* ; Δ*XCC4122* ; Rif^R	from start to stop	This study
XP162	Δ*xytA* ; Δ*XCC2828* ; Rif^R	from start to stop	This study
XP163	Δ*xytA* Δ*xytB* ; Δ*XCC2828* Δ*XCC4120* ; Rif^R		This study
XP164	Δ*xypA* Δ*xypB* ; Δ*XCC4119* Δ*XCC4121* ; Rif^R		This study
XP165	Δ*xypA* Δ*xypB xylE* ::pVO; Δ*XCC4119* Δ*XCC4121 XCC1759* ::pVO155; Rif^R Km^R		This study
XP166	WT/pCZ1016; *Xcc* 568 pCZ1016; Rif^R Tet^R Amp^R		This study
XP167	Δ*xytB3* /pCZ1016; Δ*XCC4120-3* /pCZ1016; Rif^R Tet^R Amp^R		This study
XP168	Δ*xypB* /pCZ1016*;* Δ*XCC4121/* pCZ1016; Rif^R Tet^R Amp^R		This study
XP169	*xylE* ::pVO/pCZ1016*; XCC1759* ::pVO155/pCZ1016; Rif^R Km^R Tet^R Amp^R		This study
XP170	WT/pC-*xylR* ; *Xcc* 568/pC-*XCC4101* ; Rif^R Tet^R Amp^R		This study
XP171	WT/pC-*xyn10A* ; *Xcc* 568/pC-*XCC4118* ; Rif^R Tet^R Amp^R		This study
XP172	WT/pC-*xypB; Xcc* 568 pC-*XCC4121* ; Rif^R Tet^R Amp^R		This study
XP173	WT/pC-*xylE* ; *Xcc* 568 pC-*XCC1759* ; Rif^R Tet^R Amp^R		This study
XP174	*xylR* ::pVO/pC-*xylR; XCC4101* ::pVO155/pC-*XCC4101* ; Rif^R Km^R Tet^R Amp^R		This study
XP175	Δ*xyn10A* /pC-*xyn10A* ; Δ*XCC4118/* pC-*XCC4118* ; Rif^R Tet^R Amp^R		This study
XP176	Δ*xypB* /pC-*xypB;* Δ*XCC4121/* pC-*XCC4121* ; Rif^R Tet^R Amp^R		This study
XP177	*xylB* ::pVO/pC-*xylB; XCC4121* ::pVO155/pC-*XCC4121* ; Rif^R Km^R Tet^R Amp^R		This study
XP178	*xylE* ::pVO/pC-*xylE; XCC1759* ::pVO155/pC-*XCC1759* ; Rif^R Km^R Tet^R Amp^R		This study
XP179	Δ*gly43F* /pC-*gly43F* ; Δ*XCC4122/* pC-*XCC4122* ; Rif^R Tet^R Amp^R		This study
XP180	Δ*xytB1* /pC-*xytB* ; Δ*XCC4120-* 1/pC-*XCC4120* ; Rif^R Tet^R Amp^R		This study
XP181	Δ*xytB2* /pC-*xytB* ; Δ*XCC4120-* 2/pC-*XCC4120* ; Rif^R Tet^R Amp^R		This study
XP182	Δ*xytB3* /pC-*xytB* ; Δ*XCC4120-* 3/pC-*XCC4120* ; Rif^R Tet^R Amp^R		This study
XP183	Δ*xypB* ::*xypB;* Δ*XCC4121* ::*XCC4121* ; Rif^R Km^R		This study
XP184	Δ*xypB* ::*xypB* /pPr-*xytB;* Δ*XCC4121* ::*XCC4121* /pr-*XCC4120* ; Rif^R Km^R		This study
XP185	Δ*hrpG;* Δ*XCC1166* ; Rif^R		This study
XP083	*hrpX* ::pVO; *XCC1167* ::pVO155; Rif^R Km^R	+1013	[25]
XP186	*hrpX* ::pVO Δ*xyn10A; XCC1167* ::pVO155 Δ*XCC4115* ; Rif^R Km^R		This study
XP187	*hrpX* ::pVO Δ*xyn10C; XCC1167* ::pVO155 Δ*XCC4118* ; Rif^R Km^R		This study
XP188	*hrpX* ::pVO Δ*xyn30A; XCC1167* ::pVO155 Δ*XCC0857* Rif^R Km^R		This study

function during pathogenesis. Further work is needed to unravel the molecular bases of this specificity. Finally, this specific conservation and the connexion between *Xcc* CUT system and T3SS regulators suggest that *Xanthomonas* spp. have diverted a conserved xylan degradation strategy shared by numerous bacteria to use it as part of their pathogenicity arsenal.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions

The *Xcc* strains and plasmids used in this study are listed in Table S4. *Xcc* cells were grown at 28°C in MOKA rich medium [25] or in minimal medium (MME) [112]. *Escherichia coli* cells were grown on Luria-Bertani medium at 37°C. For solid media, agar was added at a final concentration of 1.5% (wt/vol).

Antibiotics were used at the following concentrations: for *Xcc*, 50 µg/mL rifampin, 50 µg/mL kanamycin, and 5 µg/mL tetracycline; for *E. coli*, 50 µg/mL ampicillin, 50 µg/mL kanamycin, and 10 µg/mL tetracycline.

Construction of *Xanthomonas campestris* pv. *campestris* mutants

Insertion mutants were constructed using the suicide plasmid pVO155 [51] with a 300- to 500-bp PCR amplicon internal to each open reading frame (ORF). Deletion mutants were constructed by using the *cre-lox* system adapted by Angot *et al*. [113] from the system of Marx and Lidstrom [114] or by using the *sacB* system [115]. Deleted regions are indicated in Table S4 and represented on Figure 1. Oligonucleotide primers used for PCR amplification will be provided upon request.

Plasmid constructions

DNA manipulations were performed as described previously [116].

For complementation studies, PCR amplicons (oligonucleotide primers used for PCR amplification will be provided upon request) were cloned into pCZ1016, a derivative of pFAJ1700 containing the Ptac promoter, multiple cloning sites and the T7 terminator from pSC150 [117,118]. In order to perform chromosomal complementations, PCR amplicons were also cloned into pCZ1034, a derivative of pk18*mobsacB* [115] with the MCS replaced by a Ptac promoter, a MCS and a T7 terminator flanked by a 700 bp fragment corresponding to the

region upstream of the open reading frame *XCC0127* and a 700 bp fragment corresponding to the region downstream of the open reading frame *XCC0128*.

The XCC4120 and XCC2828 promoter regions (see Table S4) were PCR amplified with appropriately designed primers. These promoter regions were cloned as *Hind*III-*Xba*I fragments, into the pCZ962 plasmid, a pFAJ1700 [118] derivative containing the *Kpn*I-*Asc*I *lacZ* gene from the pCZ367 plasmid [117].

Expression studies, RNA isolation and operon mapping

β-galactosidase and β-glucuronidase assays: bacterial cultures in the appropriate medium were harvested at different time points and β-galactosidase and β-glucuronidase (GUS) assays were performed as previously described [25].

To investigate the transcriptional organization, reverse transcription-PCR (RT-PCR) experiments were performed. Bacterial cultures from *xylR* mutant of *Xcc* strain ATCC33913 grown in minimal medium (MME) were harvested after 6 h of incubation (OD_{600}=0.6). RNAs were extracted using the RNeasy Mini Kit (Qiagen). A total of 5 µg of RNA was reverse transcribed with Transcriptor Reverse Transcriptase enzyme (Roche Applied Science) using random hexamers (Biolabs) for 10 min at 25°C and then for 40 min at 55°C. The resulting cDNAs were used as a template for PCR amplification with *Taq* polymerase using specific primer pairs for each gene (as indicated in Figure S2). RT-PCR products were analyzed on a 1.2% agarose gel along with a 1-kp DNA ladder (Invitrogen).

The methods used for quantitative reverse transcription-PCR (qRT-PCR) experiments were adapted from the methods of Blanvillain *et al.* [25]. For qRT-PCR experiments done on bacteria grown on solid medium containing 4-*O*-Methyl-D-glucurono-D-xylan dyed with Remazol brilliant blue R (RBB-Xylan, Sigma), colonies obtained after 48h growth were resuspended in 1 mL of water. A 1µg sample of RNA was treated with RNase-free DNase I (Sigma) for 20 min at room temperature. After DNase inactivation (10 min at 70°C), RNAs were reverse transcribed as indicated above. Oligonucleotide primers used for quantitative PCR amplification will be provided upon request. 16S rRNA was used as a control for real-time PCR [25,119].

Growth curves

Growth curves were generated using a FLUOStar Omega apparatus (BMG Labtech, Offenburg, Germany) with four replicates. Growth was measured using 96-well flat-bottom

microtiter plates with 200µl preparations inoculated at an optical density at 600 nm (OD_{600}) of 0.1 from 4 independent washed overnight precultures. The microplates were shaken continuously at 700 rpm using the linear-shaking mode.

[^{14}C] xylose transport experiments

[^{14}C] xylose transport assays were conducted as previously described [119]. [^{14}C] xylose (Amersham Biosciences, specific activity of 3,15 GBq/mmol) was added to a final concentration of 0.5 µM. For competition experiments, unlabeled sugars were added to [^{14}C] xylose at final concentrations of 0.5, 5, 50 and 500 µM, and cells were incubated for 1 h before collection. The initial concentration-dependent xylose transport was determined using the rapid dilution method as previously described [25,28].

Plate assays for detection of xylanase activity

The plate assay for xylanase activity was performed using MME-agar plates containing 0,1% RBB-xylan (Sigma). Overnight cultures of *Xcc* strains grown in MOKA medium were centrifuged. Pellets were resuspended in MME medium and the OD_{600} was adjusted to 4. Five microliters of bacterial suspension were spotted on plates that were incubated at 28°C. The detection of xylanase activity was examined periodically by checking the halo against the blue background.

Pathogenicity tests

Pathogenicity tests were conducted on *Arabidopsis thaliana* Sf-2 ecotype as previously described [55]. Each strain was tested on sets of 4 plants with 4 leaves per plant. Disease development was scored at days 4, 7 and 9 post-inoculation using a disease index ranging from 0 (no symptom), to 4 (leaf death).

Dynamics of bacterial population densities in the phyllosphere of cabbage and bean.

Experiments were conducted on cabbage (*Brassica oleracea* cv. Bartolo) and dry bean (*P. vulgaris* cv. Flavert) as previously described [56]. Plant inoculations were carried out under quarantine at UMR PaVe´, Centre INRA, Beaucouze´, France.

Beans at the first trifoliate stage and 5 week-old cabbages were spray inoculated until runoff with bacterial suspensions at 1×10^6 CFU/mL and with sterile distilled water as a control. The environmental conditions used for these experiments and the absence of wounding did not favour disease expression. Spray inoculation of plants is, however, satisfactory for studying bacterial colonization and dispersal. For every strain, the first trifoliate leaf of five bean plants and the third leaf of cabbages were collected 3 h and at 1, 4, and 11 days after inoculation. Each leaf was weighed and ground individually (Stomacher 80; Seward, London, United Kingdom) for 2 min at maximum power in 5 ml of distilled water. Every sample and appropriate dilutions were spiral plated (Spiral Biotech, Bethesda, MD) on selective medium to enumerate the inoculated strain. To avoid cross-contamination, plants receiving a similar treatment were grouped in the growth chamber and were separated by polypropylene walls from other treatments. In each experiment, treatments were randomly distributed, and experiments were repeated at least twice.

***In silico* analysis**

The presence of signal peptides and protein localization were determined using SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP; [120]). β-sheets were located using the PSIPRED Protein Structure Server (http://bioinf.cs.ucl.ac.uk/psipred/) [84,85]. Patscan and Predetector software [35,36] were used to identify *xyl*-boxes.

Characterization of best hits in bacterial proteomes

Target proteome have been retrieved in GenBank format from ftp://ftp.ncbi.nih.gov:21/genomes/Bacteria/ for completely sequences genomes and from ftp://ftp.ncbi.nih.gov:21/genomes/Bacteria_DRAFT/ for draft ones. We reformatted them in FASTA proteic files (files available on demand). For each candidate protein against each target proteome, we looked for best hit using NCBI blastp tool (version blastp 2.2.22, parameters: - G 11, - E 1, - F F, -b 1,-v 1). In a post processing step, we kept only the best hit using the following thresholds: i) minimum percentage of candidate protein length aligned: 60; ii) minimum percentage of best hit target protein length aligned: 60; iii) minimum percentage of identity: 30.

For each hit, locus tag, identity and positives values are reported. This pipeline has been developed with home-made PERL scripts and libraries.

Phylogenetic analyses of proteins of the xylan/xylose-CUT system

To find the closest homologs to proteins of the *Xcc* or Xanthomonas spp. xylan CUT system, best protein matches were identified by a BlastP search in the non-redundant protein database (2011-04-06) at GenBank. The protein sequences were further analysed using "Analyse your sequence neighbour" option provided by Phylogeny.fr web server (http://www.phylogeny.fr/). Phylogenetic analyses of the final dataset were performed using the Phylogeny.fr web server in default mode [121]. This server connects various bioinformatics programs to reconstruct a robust phylogenetic tree from a set of sequences. Sequences are aligned with MUSCLE (v3.7) configured for highest accuracy. The alignment are curated by Gblocks (v0.9b) [122] using the following parameters: minimum length of a block after gap cleaning: 10; no gap position allowed in the final alignment; all segments with contiguous nonconserved position bigger than 8 were rejected; minimum position for a flank position: 85%. The phylogenetic trees were reconstructed using the maximum likelihood method implemented in the PhyML program (v3.0 aLTR) [123]. The WAG was used as substitution model with 4 substitution rate categories, number of invariant sites and gamma distribution parameters were estimated to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data. Reliability of internal branches was assessed using the aLRT statistical test (SH-like) [124] which is as robust and much faster than the usual bootstrap procedure [121].

Evolutionnary relashionships based on Multi-Locus Sequence Typing

The following sequences 16SrRNA (XCVr2), *efp* (XCV2577, elongation factor P), *gyrB* (XCV0004, gyrase subunit B), *glnA* (XCV0188, glutamine synthetase A) and *rpoD* (XCV3912, RNA polymerase sigma factor) were used as queries for database searches with BlastN against the microbial whole genomes databases (http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html). Nucleotide sequences were aligned and edited for manual adjustments using Bioedit (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html). Concatenated sequences were obtained using Geneious v.4.7.6 [125]. Phylogenetic analyses were done based on the Neighbor-Joining algorithm using the MEGA 4.1 software with 1,000 bootstrap iterations [126]. The optimal tree with the sum of branch length = 1.31715582 is shown in Figure S4. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches [127]. The tree is drawn to scale, with

branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method [128] and are in the units of the number of base substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 6069 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 [126].

Acknowledgments

References

1. Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, et al. (2004) Toward a systems approach to understanding plant cell walls. Science 306: 2206-2211.
2. Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol 6: 850-861.
3. Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol 6: 724-732.
4. Scheller HV, Ulvskov P (2010) Hemicelluloses. Annu Rev Plant Biol 61: 263-289.
5. Kulkarni N, Shendye A, Rao M (1999) Molecular and biotechnological aspects of xylanases. FEMS Microbiol Rev 23: 411-456.
6. Saha BC (2003) Hemicellulose bioconversion. J Ind Microbiol Biotechnol 30: 279-291.
7. Dodd D, Cann IK (2009) Enzymatic deconstruction of xylan for biofuel production. Glob Change Biol Bioenergy 1: 2-17.
8. Collins T, Gerday C, Feller G (2005) Xylanases, xylanase families and extremophilic xylanases. FEMS Microbiol Rev 29: 3-23.
9. Beg QK, Kapoor M, Mahajan L, Hoondal GS (2001) Microbial xylanases and their industrial applications: a review. Appl Microbiol Biotechnol 56: 326-338.
10. Dodd D, Mackie RI, Cann IK (2011) Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. Mol Microbiol 79: 292-304.
11. Rajeshwari R, Jha G, Sonti RV (2005) Role of an in planta-expressed xylanase of Xanthomonas oryzae pv. oryzae in promoting virulence on rice. Mol Plant Microbe Interact 18: 830-837.
12. Szczesny R, Jordan M, Schramm C, Schulz S, Cogez V, et al. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv *vesicatoria*. New Phytol 187: 983-1002.
13. Büttner D, Bonas U (2009) Regulation and secretion of *Xanthomonas* virulence factors. FEMS Microbiol Rev 34: 107-133.
14. Ryan RP, Vorholter FJ, Potnis N, Jones JB, Van Sluys MA, et al. (2011) Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. Nat Rev Microbiol 9: 344-355.
15. Gough CL, Dow JM, Barber CE, Daniels MJ (1988) Cloning of two endoglucanase genes of *Xanthomonas campestris* pv. *campestris*: analysis of the role of endoglucanase in pathogenesis. Mol Plant Microbe Interact 1: 275-281.

16. Hu J, Qian W, He C (2007) The *Xanthomonas oryzae* pv. *oryzae eglXoB* endoglucanase gene is required for virulence to rice. FEMS Microbiol Lett 269: 273-279.
17. Jha G, Rajeshwari R, Sonti RV (2005) Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. Mol Plant Microbe Interact 18: 891-898.
18. Wang JC, Kim JH, Park YJ, Lee B-M, Kang HW (2008) Genome wide identification of pathogenicity genes in *Xanthomonas oryzae* pv. *oryzae* by transposon mutagenesis. Plant Pathology 57: 1136-1145.
19. Wang L, Rong W, He C (2008) Two *Xanthomonas* extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. Mol Plant Microbe Interact 21: 555-563.
20. Aparna G, Chatterjee A, Sonti RV, Sankaranarayanan R (2009) A cell wall-degrading esterase of *Xanthomonas oryzae* requires a unique substrate recognition module for pathogenesis on rice. Plant Cell 21: 1860-1873.
21. Jha G, Rajeshwari R, Sonti RV (2007) Functional interplay between two Xanthomonas oryzae pv,. oryzae secretion systems in modulating virulence on rice. Mol Plant Microbe Interact 20: 31-40.
22. Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, et al. (2010) Sugar transporters for intercellular exchange and nutrition of pathogens. Nature 468: 527-532.
23. Tampakaki AP, Skandalis N, Gazi AD, Bastaki MN, Sarris PF, et al. (2010) Playing the "Harp": evolution of our understanding of *hrp/hrc* genes. Annu Rev Phytopathol 48: 347-370.
24. Furutani A, Tsuge S, Ohnishi K, Hikichi Y, Oku T, et al. (2004) Evidence for HrpXo-dependent expression of type II secretory proteins in *Xanthomonas oryzae* pv. *oryzae*. J Bacteriol 186: 1374-1380.
25. Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, et al. (2007) Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS One 2: e224.
26. Cornelis P (2010) Iron uptake and metabolism in pseudomonads. Appl Microbiol Biotechnol 86: 1637-1645.
27. Krewulak KD, Vogel HJ (2011) TonB or not TonB: is that the question? Biochem Cell Biol 89: 87-97.
28. Neugebauer H, Herrmann C, Kammer W, Schwarz G, Nordheim A, et al. (2005) ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. J Bacteriol 187: 8300-8311.
29. Eisenbeis S, Lohmiller S, Valdebenito M, Leicht S, Braun V (2008) NagA-dependent uptake of *N*-acetyl-glucosamine and *N*-acetyl-chitin oligosaccharides across the outer membrane of *Caulobacter crescentus*. J Bacteriol 190: 5230-5238.
30. Schauer K, Rodionov DA, de Reuse H (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem Sci 33: 330-338.
31. Boulanger A, Dejean G, Lautier M, Glories M, Zischek C, et al. (2010) Identification and regulation of the *N*-acetylglucosamine utilization pathway of the plant pathogenic bacterium *Xanthomonas campestris* pv. *campestris*. J Bacteriol 192: 1487-1497.
32. Hottes AK, Meewan M, Yang D, Arana N, Romero P, et al. (2004) Transcriptional profiling of *Caulobacter crescentus* during growth on complex and minimal media. J Bacteriol 186: 1448-1461.
33. Dodd D, Moon YH, Swaminathan K, Mackie RI, Cann IK (2010) Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic bacteroidetes. J Biol Chem 285: 30261-30273.

34. Stephens C, Christen B, Watanabe K, Fuchs T, Jenal U (2007) Regulation of D-xylose metabolism in *Caulobacter crescentus* by a LacI-type repressor. J Bacteriol 189: 8828-8834.
35. Dsouza M, Larsen N, Overbeek R (1997) Searching for patterns in genomic data. Trends Genet 13: 497-498.
36. Hiard S, Maree R, Colson S, Hoskisson PA, Titgemeyer F, et al. (2007) PREDetector: a new tool to identify regulatory elements in bacterial genomes. Biochem Biophys Res Commun 357: 861-864.
37. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37: D233-238.
38. DeBoy RT, Mongodin EF, Fouts DE, Tailford LE, Khouri H, et al. (2008) Insights into plant cell wall degradation from the genome sequence of the soil bacterium *Cellvibrio japonicus*. J Bacteriol 190: 5455-5463.
39. Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, et al. (2011) Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. BMC Genomics 12: 146.
40. Whitehead TR (1995) Nucleotide sequences of xylan-inducible xylanase and xylosidase/arabinosidase genes from *Bacteroides ovatus* V975. Biochim Biophys Acta 1244: 239-241.
41. Gasparic A, Martin J, Daniel AS, Flint HJ (1995) A xylan hydrolase gene cluster in *Prevotella ruminicola* B(1)4: sequence relationships, synergistic interactions, and oxygen sensitivity of a novel enzyme with exoxylanase and beta-(1,4)-xylosidase activities. Appl Environ Microbiol 61: 2958-2964.
42. Dodd D, Kiyonari S, Mackie RI, Cann IK (2010) Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium *Prevotella bryantii* B14. J Bacteriol 192: 2335-2345.
43. Shulami S, Gat O, Sonenshein AL, Shoham Y (1999) The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. J Bacteriol 181: 3695-3704.
44. Beylot MH, McKie VA, Voragen AG, Doeswijk-Voragen CH, Gilbert HJ (2001) The *Pseudomonas cellulosa* glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity. Biochem J 358: 607-614.
45. Nagy T, Nurizzo D, Davies GJ, Biely P, Lakey JH, et al. (2003) The alpha-glucuronidase, GlcA67A, of *Cellvibrio japonicus* utilizes the carboxylate and methyl groups of aldobiouronic acid as important substrate recognition determinants. J Biol Chem 278: 20286-20292.
46. Ritzenthaler P, Mata-Gilsinger M, Stoeber F (1980) Construction and expression of hybrid plasmids containing *Escherichia coli* K-12 uxu genes. J Bacteriol 143: 1116-1126.
47. Rakus JF, Fedorov AA, Fedorov EV, Glasner ME, Vick JE, et al. (2007) Evolution of enzymatic activities in the enolase superfamily: D-Mannonate dehydratase from *Novosphingobium aromaticivorans*. Biochemistry 46: 12896-12908.
48. Lawlis VB, Dennis MS, Chen EY, Smith DH, Henner DJ (1984) Cloning and sequencing of the xylose isomerase and xylulose kinase genes of *Escherichia coli*. Appl Environ Microbiol 47: 15-21.
49. Stoddart A, Zhang Y, Paige CJ (1996) Molecular cloning of the cDNA encoding a murine sialic acid-specific 9-*O*-acetylesterase and RNA expression in cells of hematopoietic and non-hematopoietic origin. Nucleic Acids Res 24: 4003-4008.
50. Yang TC, Hu RM, Hsiao YM, Weng SF, Tseng YH (2003) Molecular genetic analyses of potential beta-galactosidase genes in *Xanthomonas campestris*. J Mol Microbiol Biotechnol 6: 145-154.

51. Oke V, Long SR (1999) Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. Mol Microbiol 32: 837-849.
52. Tsuge S, Terashima S, Furutani A, Ochiai H, Oku T, et al. (2005) Effects on promoter activity of base substitutions in the cis-acting regulatory element of HrpXo regulons in *Xanthomonas oryzae* pv. *oryzae*. J Bacteriol 187: 2308-2314.
53. Furutani A, Nakayama T, Ochiai H, Kaku H, Kubo Y, et al. (2006) Identification of novel HrpXo regulons preceded by two cis-acting elements, a plant-inducible promoter box and a -10 box-like sequence, from the genome database of *Xanthomonas oryzae* pv. *oryzae*. FEMS Microbiol Lett 259: 133-141.
54. Koebnik R, Kruger A, Thieme F, Urban A, Bonas U (2006) Specific binding of the *Xanthomonas campestris* pv. *vesicatoria* AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. J Bacteriol 188: 7652-7660.
55. Meyer D, Lauber E, Roby D, Arlat M, Kroj T (2005) Optimization of pathogenicity assays to study the *Arabidopsis thaliana-Xanthomonas campestris* pv. campestris pathosystem. Molecular Plant Pathology 6: 327-333.
56. Darsonval A, Darrasse A, Meyer D, Demarty M, Durand K, et al. (2008) The Type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. Appl Environ Microbiol 74: 2669-2678.
57. Thieme F, Koebnik R, Bekel T, Berger C, Boch J, et al. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. J Bacteriol 187: 7254-7266.
58. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, et al. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature 417: 459-463.
59. Moreira LM, Almeida NF, Jr., Potnis N, Digiampietri LA, Adi SS, et al. (2010) Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of *Xanthomonas fuscans* subsp. *aurantifolii*. BMC Genomics 11: 238.
60. Qian W, Jia Y, Ren SX, He YQ, Feng JX, et al. (2005) Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. Genome Res 15: 757-767.
61. Vorholter FJ, Schneiker S, Goesmann A, Krause L, Bekel T, et al. (2008) The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J Biotechnol 134: 33-45.
62. Studholme DJ, Kemen E, MacLean D, Schornack S, Aritua V, et al. (2010) Genome-wide sequencing data reveals virulence factors implicated in banana *Xanthomonas* wilt. FEMS Microbiol Lett 310: 182-192.
63. Pieretti I, Royer M, Barbe V, Carrere S, Koebnik R, et al. (2009) The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. BMC Genomics 10: 616.
64. Lee BM, Park YJ, Park DS, Kang HW, Kim JG, et al. (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res 33: 577-586.
65. Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, et al. (2008) Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics 9: 204.
66. Almeida NF, Yan S, Cai R, Clarke CR, Morris CE, et al. (2010) PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. Phytopathology 100: 208-215.

67. Puhler A, Arlat M, Becker A, Gottfert M, Morrissey JP, et al. (2004) What can bacterial genome research teach us about bacteria-plant interactions? Curr Opin Plant Biol 7: 137-147.
68. Ryan RP, Monchy S, Cardinale M, Taghavi S, Crossman L, et al. (2009) The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. Nat Rev Microbiol 7: 514-525.
69. Weon HY, Kim BY, Kim JS, Lee SY, Cho YH, et al. (2006) *Pseudoxanthomonas suwonensis* sp. nov., isolated from cotton waste composts. Int J Syst Evol Microbiol 56: 659-662.
70. Stephens C, Christen B, Fuchs T, Sundaram V, Watanabe K, et al. (2007) Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. J Bacteriol 189: 2181-2185.
71. Weiner RM, Taylor LE, 2nd, Henrissat B, Hauser L, Land M, et al. (2008) Complete genome sequence of the complex carbohydrate-degrading marine bacterium, *Saccharophagus degradans* strain 2-40 T. PLoS Genet 4: e1000087.
72. Ko JK, Jung MW, Kim KH, Choi IG (2009) Optimal production of a novel endo-acting beta-1,4-xylanase cloned from *Saccharophagus degradans* 2-40 into *Escherichia coli* BL21(DE3). N Biotechnol 26: 157-164.
73. Yang JC, Madupu R, Durkin AS, Ekborg NA, Pedamallu CS, et al. (2009) The complete genome of *Teredinibacter turnerae* T7901: an intracellular endosymbiont of marine wood-boring bivalves (shipworms). PLoS One 4: e6085.
74. Emami K, Topakas E, Nagy T, Henshaw J, Jackson KA, et al. (2009) Regulation of the xylan-degrading apparatus of *Cellvibrio japonicus* by a novel two-component system. J Biol Chem 284: 1086-1096.
75. Fontes CM, Gilbert HJ, Hazlewood GP, Clarke JH, Prates JA, et al. (2000) A novel *Cellvibrio mixtus* family 10 xylanase that is both intracellular and expressed under non-inducing conditions. Microbiology 146 (Pt 8): 1959-1967.
76. Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, et al. (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. Genome Res 20: 1605-1612.
77. Zhou J, Huang H, Meng K, Shi P, Wang Y, et al. (2009) Molecular and biochemical characterization of a novel xylanase from the symbiotic *Sphingobacterium* sp. TN19. Appl Microbiol Biotechnol 85: 323-333.
78. Zhou J, Meng K, Yang P, Shi P, Wang Y, et al. (2010) Characterization of a chromosomal segment showing xylanolytic activity from the symbiotic *Sphingobacterium*sp. TN19. World J Microbiol Biotechnol 26: 761-765.
79. McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, et al. (2009) Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. Appl Environ Microbiol 75: 6864-6875.
80. van Passel MW, Kant R, Palva A, Copeland A, Lucas S, et al. (2011) Genome Sequence of the Verrucomicrobium *Opitutus terrae* PB90-1, an Abundant Inhabitant of Rice Paddy Soil Ecosystems. J Bacteriol 193: 2367-2368.
81. Wagner M, Horn M (2006) The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. Curr Opin Biotechnol 17: 241-249.
82. Nakai K, Horton P (2007) Computational prediction of subcellular localization. Methods Mol Biol 390: 429-466.
83. Koebnik R (2005) TonB-dependent trans-envelope signalling: the exception or the rule? Trends Microbiol 13: 343-347.

84. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292: 195-202.
85. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, et al. (2005) Protein structure prediction servers at University College London. Nucleic Acids Res 33: W36-38.
86. Koebnik R (1993) Structural organization of TonB-dependent receptors. Trends Microbiol 1: 201.
87. Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, et al. (2009) Three genomes from the phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. Appl Environ Microbiol 75: 2046-2056.
88. Ebringerova A, Hromadkova Z, Heinze T (2005) Hemicellulose. Adv Polym Sci 186: 1-67.
89. Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J 3: 1-30.
90. Shallom D, Shoham Y (2003) Microbial hemicellulases. Curr Opin Microbiol 6: 219-228.
91. Hurlbert JC, Preston JF, 3rd (2001) Functional characterization of a novel xylanase from a corn strain of *Erwinia chrysanthemi*. J Bacteriol 183: 2093-2100.
92. Emami K, Nagy T, Fontes CM, Ferreira LM, Gilbert HJ (2002) Evidence for temporal regulation of the two *Pseudomonas cellulosa* xylanases belonging to glycoside hydrolase family 11. J Bacteriol 184: 4124-4133.
93. Shulami S, Zaide G, Zolotnitsky G, Langut Y, Feld G, et al. (2007) A two-component system regulates the expression of an ABC transporter for xylo-oligosaccharides in Geobacillus stearothermophilus. Appl Environ Microbiol 73: 874-884.
94. Miyazaki K, Hirase T, Kojima Y, Flint HJ (2005) Medium- to large-sized xylo-oligosaccharides are responsible for xylanase induction in *Prevotella bryantii* B14. Microbiology 151: 4121-4125.
95. Miyazaki K, Miyamoto H, Mercer DK, Hirase T, Martin JC, et al. (2003) Involvement of the multidomain regulatory protein XynR in positive control of xylanase gene expression in the ruminal anaerobe *Prevotella bryantii* B(1)4. J Bacteriol 185: 2219-2226.
96. Larson SB, Day J, Barba de la Rosa AP, Keen NT, McPherson A (2003) First crystallographic structure of a xylanase from glycoside hydrolase family 5: implications for catalysis. Biochemistry 42: 8411-8422.
97. St John FJ, Hurlbert JC, Rice JD, Preston JF, Pozharski E (2011) Ligand bound structures of a glycosyl hydrolase family 30 glucuronoxylan xylanohydrolase. J Mol Biol 407: 92-109.
98. Pell G, Taylor EJ, Gloster TM, Turkenburg JP, Fontes CM, et al. (2004) The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. J Biol Chem 279: 9597-9605.
99. Mirande C, Mosoni P, Bera-Maillet C, Bernalier-Donadille A, Forano E (2010) Characterization of Xyn10A, a highly active xylanase from the human gut bacterium *Bacteroides xylanisolvens* XB1A. Appl Microbiol Biotechnol 87: 2097-2105.
100. Biely P, Vrsanska M, Tenkanen M, Kluepfel D (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. J Biotechnol 57: 151-166.
101. Vrsanska M, Kolenova K, Puchart V, Biely P (2007) Mode of action of glycoside hydrolase family 5 glucuronoxylan xylanohydrolase from *Erwinia chrysanthemi*. Febs J 274: 1666-1677.
102. Hematy K, Cherk C, Somerville S (2009) Host-pathogen warfare at the plant cell wall. Curr Opin Plant Biol 12: 406-413.

103. Lagaert S, Belien T, Volckaert G (2009) Plant cell walls: Protecting the barrier from degradation by microbial enzymes. Semin Cell Dev Biol 20: 1064-1073.
104. Nagy T, Emami K, Fontes CM, Ferreira LM, Humphry DR, et al. (2002) The membrane-bound alpha-glucuronidase from Pseudomonas cellulosa hydrolyzes 4-O-methyl-D-glucuronoxylooligosaccharides but not 4-O-methyl-D-glucuronoxylan. J Bacteriol 184: 4925-4929.
105. Nurizzo D, Nagy T, Gilbert HJ, Davies GJ (2002) The structural basis for catalysis and specificity of the *Pseudomonas cellulosa* alpha-glucuronidase, GlcA67A. Structure 10: 547-556.
106. Vanfossen AL, Verhaart MR, Kengen SM, Kelly RM (2009) Carbohydrate utilization patterns for the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus* reveal broad growth substrate preferences. Appl Environ Microbiol 75: 7718-7724.
107. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, et al. (2009) A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature 462: 1056-1060.
108. Jun SR, Sims GE, Wu GA, Kim SH (2010) Whole-proteome phylogeny of prokaryotes by feature frequency profiles: An alignment-free method with optimal feature resolution. Proc Natl Acad Sci U S A 107: 133-138.
109. Martens EC, Koropatkin NM, Smith TJ, Gordon JI (2009) Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. J Biol Chem 284: 24673-24677.
110. Cordero OX, Snel B, Hogeweg P (2008) Coevolution of gene families in prokaryotes. Genome Res 18: 462-468.
111. Liu H, Zhang S, Schell MA, Denny TP (2005) Pyramiding unmarked deletions in *Ralstonia solanacearum* shows that secreted proteins in addition to plant cell-wall-degrading enzymes contribute to virulence. Mol Plant Microbe Interact 18: 1296-1305.
112. Arlat M, Gough CL, Barber CE, Boucher C, Daniels MJ (1991) *Xanthomonas campestris* contains a cluster of *hrp* genes related to the larger hrp cluster of *Pseudomonas solanacearum*. Mol Plant Microbe Interact 4: 593-601.
113. Angot A, Peeters N, Lechner E, Vailleau F, Baud C, et al. (2006) *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. Proc Natl Acad Sci U S A 103: 14620-14625.
114. Marx CJ, Lidstrom ME (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. Biotechniques 33: 1062-1067.
115. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, et al. (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145: 69-73.
116. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual; Laboratory CSH, editor. NY.
117. Cunnac S (2004) Identification à l'échelle génomique des effecteurs dépendant du système de sécrétion de type III de la bactérie phytopathogène *Ralstonia solanacearum*. Toulouse: Université Toulouse III Paul Sabatier. 184 p.
118. Dombrecht B, Vanderleyden J, Michiels J (2001) Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol Plant Microbe Interact 14: 426-430.
119. Morales CQ, Posada J, Macneale E, Franklin D, Rivas I, et al. (2005) Functional analysis of the early chlorosis factor gene. Mol Plant Microbe Interact 18: 477-486.

120. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2: 953-971.
121. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465-469.
122. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17: 540-552.
123. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696-704.
124. Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 55: 539-552.
125. Drummond AJ, Ashton B, Buxton S, Cheung M, Heled J, et al. (2009) Genious v4.7.6. available from http://wwwgeneiouscom/.
126. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.
127. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.
128. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101: 11030-11035.
129. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39: D225-229.
130. Finn RD, Mistry J, Tate J, Coggill P, Heger A, et al. (2010) The Pfam protein families database. Nucleic Acids Res 38: D211-222.
131. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14: 1188-1190.
132. Zehner S, Kotzsch A, Bister B, Sussmuth RD, Mendez C, et al. (2005) A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005. Chem Biol 12: 445-452.
133. Liu C, Gilmont RR, Benndorf R, Welsh MJ (2000) Identification and characterization of a novel protein from Sertoli cells, PASS1, that associates with mammalian small stress protein hsp27. J Biol Chem 275: 18724-18731.
134. Harhangi HR, Akhmanova AS, Emmens R, van der Drift C, de Laat WT, et al. (2003) Xylose metabolism in the anaerobic fungus *Piromyces* sp. strain E2 follows the bacterial pathway. Arch Microbiol 180: 134-141.
135. Utt EA, Eddy CK, Keshav KF, Ingram LO (1991) Sequencing and expression of the *Butyrivibrio fibrisolvens* xylB gene encoding a novel bifunctional protein with beta-D-xylosidase and alpha-L-arabinofuranosidase activities. Appl Environ Microbiol 57: 1227-1234.
136. Blanco C, Ritzenthaler P, Kolb A (1986) The regulatory region of the uxuAB operon in *Escherichia coli* K12. Mol Gen Genet 202: 112-119.
137. Kellett LE, Poole DM, Ferreira LM, Durrant AJ, Hazlewood GP, et al. (1990) Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes. Biochem J 272: 369-376.
138. Gonzalez ET, Allen C (2003) Characterization of a *Ralstonia solanacearum* operon required for polygalacturonate degradation and uptake of galacturonic acid. Mol Plant Microbe Interact 16: 536-544.
139. Liang WJ, Wilson KJ, Xie H, Knol J, Suzuki S, et al. (2005) The *gusBC* genes of *Escherichia coli* encode a glucuronide transport system. J Bacteriol 187: 2377-2385.

140. Moracci M, Cobucci Ponzano B, Trincone A, Fusco S, De Rosa M, et al. (2000) Identification and molecular characterization of the first alpha -xylosidase from an archaeon. J Biol Chem 275: 22082-22089.
141. Ishimizu T, Hashimoto C, Takeda R, Fujii K, Hase S (2007) A Novel alpha1,2-L-fucosidase acting on xyloglucan oligosaccharides is associated with endo-beta-mannosidase. J Biochem 142: 721-729.
142. Zhang CC, Durand MC, Jeanjean R, Joset F (1989) Molecular and genetical analysis of the fructose-glucose transport system in the cyanobacterium *Synechocystis* PCC6803. Mol Microbiol 3: 1221-1229.
143. Keen NT, Boyd C, Henrissat B (1996) Cloning and characterization of a xylanase gene from corn strains of *Erwinia chrysanthemi*. Mol Plant Microbe Interact 9: 651-657.
144. Chow V, Nong G, Preston JF (2007) Structure, function, and regulation of the aldouronate utilization gene cluster from *Paenibacillus* sp. strain JDR-2. J Bacteriol 189: 8863-8870.

Figure Legends

Figure R1 General structure of xylans, putative xylan-degrading enzymes of *Xanthomonas campestris* pv. *campestris* (*Xcc*) and their gentic organization. (A) The major enzymes and their sites of action are depicted with arrows. For each enzymatic activity, the corresponding families listed in the CAZy database are shown and *Xcc* proteins belonging to each family are listed beneath. (B, C, D, E) Genetic organization of *Xcc xytA* (B), *xylR* (C), *xytB* (D) and *xylE* (E) loci. Genes are represented by arrows, their names and putative functions are indicated beneath. Vertical black arrowheads indicate the position of pVO155 insertions. Horizontal dotted bars show deleted sequences, whereas thick bars indicate DNA segments used for plasmid constructions. Perfect *xyl*-boxes are represented by red circles.

Figure R2 Model of xylan degradation pathway in *Xanthomonas campestris* pv. *campestris*. Xyn10A is a key extracellular enzyme in the degradation of xylan. This endo-1,4-β-xylanase of family GH10 releases short to medium sized xylo-oligosaccharides which can be substituted with various side chains such as L-arabinose, D-glucoronic acid or its 4-*O*-methyl ether, thus generating decorated xylo-oligosaccharides such as methylglucuronoxylotriose ($MeGAX_3$), glucuronoxylotriose (GAX_3) or arabinoxylobiose (AX_2) for example. These compounds are either directly taken up into the periplasm or further degraded in the extracellular medium to generate transportable molecules. The transport of hydrolysis products is mediated by XytA, XytB or as yet not identified TBDTs or unknown porins. The transported degradation products are further degraded in the periplasm to generate short xylo-oligosaccharides (X_2, X_3 or X_4). The exact location of the different degradation steps is not known, yet. Enzymes displaying a signal peptide are active either in the periplasm or in the extracellular medium or even bound to membranes. They are shown in the yellow box that crosses the outer membrane (OM). The xylo-oligosaccharides are then transported into the cytoplasm by XypB inner membrane transporter. Xylose monomers present in the periplasm are taken up through XylE, whereas glucuronic acid might be transported by XypA putative hexuronate transporter. Inside the cell, xylo-oligomers are hydrolyzed to xylose by Gly43F putative exoxylanase. Xylose is converted into xylulose-5-phosphate, which can enter the pentose cycle. D-glucoronic acid is converted to glyceraldehyde 3-P and pyruvate by a five-step pathway catalyzed by three enzymes of the xylan/xylose Cut system, UxaC, UxuB, UxuA and two other enzymes KdgK, and KdgA. Glyceraldehyde 3-P and pyruvate can enter the Embden-Meyerhof-Parnas pathway.

Figure R3 Regulation of *Xanthomonas campestris* pv. *campestris* putative xylanase genes *xyn10C* (*XCC4115*), *xyn10A* (*XCC4118*) and *xyn30A* (*XCC0857*) by Hrp regulators. The relative expression was analyzed by real-time quantitative RT-PCR (qRT-PCR), perfomed on RNA extracted from the wild-type strain, Δ*hrpG* deletion mutant and *hrpX*::pVO insertion mutants grown in minimal liquid medium. For *xyn10C*, experiments were conducted with primers located either upstream (Upstream) or downstream (Downstream) the putative degenerated Pip-box found in the coding sequence. Calculation of relative expression includes normalization against the endogenous control gene 16S RNA. Bars represent the standard deviations calculated from at least two different biological repetitions.

Figure R4 Expression of the *Xanthomonas campestris* pv. *campestris xytA* and *xytB* TBDT genes and xypB inner membrane transporter gene, in presence of xylose or xylo-oligosaccharides. (A-D) The *xytA*::pVO (A, B) and *xytB*::pVO (C, D) insertion mutants were grown 6 hours in minimal medium (MME) in the presence of various concentrations of xylose (A, C) or xylobiose (B, D), before measuring *β*-glucuronidase activity. (E) The pPr-*xytB* plasmid carrying the promoterless *lacZ* reporter gene under the *xytB* promoter region was used to monitor *xytB* expression in presence of xylose or xylo-oligosaccharides in different genetic backgrounds. *β*-galactosidase activity was measured after 6 hours induction in MME supplemented with xylose or xylo-oligosaccharides at a final concentration of 2 mM. The expression of the transcriptional fusion was monitored in the wild-type strain and in the deletion mutants Δ*xypA* and Δ*xypB*. (F) The expression of *xypB* was monitored in *xypB*::pVO insertion mutant by measuring the *β*-glucuronidase activity after 6 hours of growth in MME supplemented with xylose or xylo-ologosaccharides at a final concentration of 2 mM. Bars represent the standard deviations calculated from at least two different biological repetitions.

Figure R5 Concentration-dependent transport of ^{14}C-labeled xylose into *Xanthomonas campestris* pv. *campestris*. Cells were grown in minimal medium without xylose, and transport was measured for 15 s at the [^{14}C] xylose concentrations indicated.

Figure R6 Growth of *X. campestris* pv. *campestris* wild-type and mutant strains in the presence of Xylose (X_1) or Xylotriose (X_3). (A) Growth of the wild-type strain and mutant strains in the inner membrane transporter *xypA*, *xypB* and *xylE* in the presence of X_1. (B)

Growth of XylE mutant strain and the corresponding complemented strain in the presence of X_1. (C) Growth of the wild-type strain and mutant strains in the TBDT genes *xytA* and *xytB* in the presence of X_1. (D) Growth of the wild-type strain and mutant strains in the inner membrane transporter *xypA*, *xypB* and *xylE* in the presence of X_3. (E) Growth of *xypB* mutant strain and the corresponding complemented strain in the presence of X_3. (F) Growth of the wild-type strain, *xytA* mutant and *xytB* mutant in the presence of X_3. (G) Growth of *xytB*, *xypB* and *gly43F* mutant strains and the corresponding complemented strain in the presence of X_3. After overnight growth in rich medium, cells were harvested, washed, and resuspended in minimal medium. X_1 and X_3 were added at a final concentration of 2mM. The error bars indicate the standard deviations obtained from at least 3 independent experiments.

Figure R7 Colonization of cabbage and bean by the wild-type strain *X. campestris* pv. *campestris* and strains mutated in *xytA* or *xytB*. Bacterial population densities were determined on leaves sampled at 3 h and 1, 4, and 11 days after spray inoculation (1×10^6 CFU ml^{-1}). Means and SEMs were calculated for five leaves per plant species and per sampling date. Mean population densities followed by different letters are significantly ($P < 0.05$) different on the basis of the Mann-Whitney test. These experiments were conducted two times independently and similar results were obtained.

Figure R8 Comparison of *xytB* locus of certain species within the family Xanthomonadaceae allows grouping of *Xanthomonas* species. The distribution and location of *xyn10A*, *xyn10B*, *xyn10C* xylanase genes and *gly2A* gene show the existence of five groups within the species *Xanthomonas*. These five groups can be organized into two larger clusters. Cluster 1 obtained by merging groups 1 and 2 correspond to strains infecting dicots. Groups 3, 4 and 5 which form the second cluster only comprise strains pathogenic on monocots. The organization of *xytB* locus of *Pseudoxanthomonas suwonensis* 11-1 is also shown. Stars indicate genes coding for active extracellular xylanases in various *Xanthomonas* species. ORF numbers are indicated within each of the genes as derived from the GenBank™ database. Red circles show perfect *xyl*-boxes. Degenerated *xyl*-boxes found in *xytB* locus of *P. suwonensis* are indicated by orange circles.

Figure R9 Conservation of xylan degradative enzymes among Proteobacteria reveals an association to conserved TBDTs. Genomic context of genes coding for proteins displaying significant similarities to proteins of the *Xanthomonas campestris* pv. *campestris* (*Xcc*) xylan/xylose CUT system. The genes are color-coded based on their predicted roles as indicated in the legend. *Xcc* genes coding for proteins conserved in other Proteobacteria are framed by yellow boxes. For other bacteria, genes coding for proteins showing homology to proteins of *Xcc* xylan CUT system are presented into yellow boxes. For *Caulobacter crescentus* CB15, genes induced by xylose [32] are indicated by red-edged arrows. The pink zones show conserved TBDT genes or loci. ORF numbers are from genome projects hosted in the GenBank™ database. The source of each bacterium is indicated in the color-coded boxes giving organism names.

Figure R10 Conservation of xylan degradative enzymes in Bacteroidetes and Verrucomicrobia reveals an association to TBDTs and *xus* cluster. Genomic context of genes coding for proteins displaying significant similarities to proteins of the *Xanthomonas campestris* pv. *campestris* (*Xcc)* xylan/xylose CUT system. The genes are color-coded based on their predicted roles as indicated in the legend of Figure R9. *Xcc* genes coding for proteins conserved in other bacteria are framed by yellow boxes. For other bacteria, genes coding for proteins showing homology to proteins of *Xcc* xylan CUT system are presented into yellow boxes. For *Prevotella bryantii* $B_1$4, genes induced by wheat arabinoxylan [33] are indicated by red-edged arrows. The pink zones show conserved TBDT genes or *xus* cluster conservation. The grey zone shows conserved genes between *Chrysobacterium gleum* and *Opitutus terrae*. ORF numbers are from genome projects hosted in the GenBank™ database. The source of each bacterium is indicated in the boxes giving organism names.

SUPPORTING INFORMATION

Figure S1 Xylose induction motif of *Caulobacter crescentus* genes and *xyl*-box motif of *Xanthomonas* spp. and *Pseudoxanthomonas suwonensis* genes. (A) Sequence logo generated by WebLogo (http://weblogo.berkeley.edu/, [131]) of the xylose inducing motif identified upstream of xylose induced genes in *C. crescentus* CB15. The horizontal bracket shows the extent of the putative XylR operator sequence. (B) Sequence logo of the *xyl*-box motif generated after alignment of sequences identified in *Xanthomonas* spp genomes. The horizontal arrows show the palindromic sequence. (C) Sequence logo of the *xyl*-box-like motif obtained after multiple alignment of conserved sequences detected in *P. suwonensis* genome.

Figure S2 Operon mapping of *xytA* and *xytB* loci. Solid arrows represent the different ORFs of *xytA* (A) and *xytB* (B) loci. The long arrow represent the *xytA* and *xytB* mRNA. The small arrows denote the positions and directions of primers used in RT-PCR. Lengths of expected PCR products are also shown for every primer pair. Lanes marked with a (+) indicate positive controls done by a PCR on genomic DNA. Lanes marked with a (-) indicate negative controls done on the sample without reverse transcription. (A) RT-PCR analysis indicates that *xytA*, *xyaA* and *xyaB* genes are cotranscribed as a single mRNA. The 2828_F/2827_R and 2827_F/2826_R primer pairs yield the expected 808 bp and 819 bp fragments, respectively. (B) RT-PCR analysis indicates on a one hand that *xypA*, *xyn10A*, *UxaC*, *Gly2A* and *xyn10C* genes are cotranscribed as a single mRNA. On the other hand, they indicates that *xytB*, *xypB* and *gly43F* are cotranscribed as a single mRNA. Indeed, the 4118_F/4119_R, 4117_F/4118_R, 4116_F/4117_R, 4115_F/4116_R, 4120_F/4121_R, 4121_F/4122_R and 4122_F/4123_R primer pairs yield the expected 866 bp, 811 bp, 800bp, 772 bp, 819 bp, 831 bp and 789 bp fragments, respectively.

Figure S3 Comparison of *xylE*, *xytA*, *xylR* and *xytB* loci of certain species within the family Xanthomonadaceae. ORF numbers are indicated within each of the genes as derived from the GenBank™ database. Red circles show perfect *xyl*-boxes whereas orange cicles indicate degenerated *xyl*-boxes found in *Pseudoxanthomonas suwonensis* 11-1 strain. Grey shaded zones show conserved regions.

Figure S4 Evolutionnary relashionships of 14 *Xanthomonas* species based on Multi-Locus Sequence Typing and definition of groups based on the distribution of discriminating xylan degradation associated genes. Concatenated nucleotide sequences of five genes (*efp, gyrB, gluA, rpoD* and 16rRNA) from 14 *Xanthomonas* strains, other sequenced Xanthomonadaceae and the phytopathogenic bacteria *Ralstonia solanacearum* strain GM1000 and *Pseudomonas syringae* pv. *tomato* strain DC3000 were analysed. Phylogenetic analyses were done based on the Neighbor-Joining algorithm with 1,000 bootstrap iterations. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 6069 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4.
Boxes to the right of the tree represent genes *xyn10C*, *gly2A*, *xyn10B*, *xyn10A*, *gly43C*, *xyl39A*. A yellow box indicates that the gene is present in the strain, whereas a black box signifies that it is absent. Grey boxes indicate that the considered gene is either mutated or located outside the *xytB* locus. Green boxes show host plant classes.

Figure S5 Phylogenetic tree of *Xanthomonas campestris* pv. *campestris* (*Xcc*) XytA (A) and XytB (B) TonB-dependent transporter proteins. The GenInfo identifier (gi) or NCBI sequence reference are given after the taxon names. Representative bacteria of each taxon, analysed in this study are indicated by colorful boxes. This phylogenetic analysis was performed by using the Phylogeny.fr server [121]. Significance of each node is assessed by approximate likelihood-ratio tests (aLRT) [124].

Figure S6 Phylogenetic tree of *Xanthomonas campestris* pv. *campestris* (*Xcc*) Xyn10C endo-1,4-β-xylanase protein. The GenInfo identifier (gi) or NCBI sequence reference are given after the taxon names. Representative bacteria of each taxon, analysed in this study are indicated by colorful boxes. This phylogenetic analysis was performed by using the Phylogeny.fr server [121]. Significance of each node is assessed by approximate likelihood-ratio tests (aLRT) [124].

Figure S7 Phylogenetic tree of *Xanthomonas* spp. Xyn10A and Xyn10B endo-1,4-β-xylanase proteins. The GenInfo identifier (gi) or NCBI sequence reference are given after the taxon names. Representative bacteria of each taxon, analysed in this study are indicated by colorful boxes. This phylogenetic analysis was performed by using the Phylogeny.fr server [121]. Significance of each node is assessed by approximate likelihood-ratio tests (aLRT) [124].

Figure S8 Phylogenetic tree of *Xanthomonas campestris* pv. *campestris* (*Xcc*) XypB xylo-oligosaccharide transporter protein (A) and Gly43F putative exoxylanase protein (B). The GenInfo identifier (gi) or NCBI sequence reference are given after the taxon names. Representative bacteria of each taxon, analysed in this study are indicated by colorful boxes. This phylogenetic analysis was performed by using the Phylogeny.fr server [121]. Significance of each node is assessed by approximate likelihood-ratio tests (aLRT) [124].

Figure S9 Phylogenetic tree of *Xanthomonas campestris* pv. *campestris* (*Xcc*) Xyn30A endo-1,4-β-xylanase protein. The GenInfo identifier (gi) or NCBI sequence reference are given after the taxon names. Representative bacteria of each taxon, analysed in this study are indicated by colorful boxes. This phylogenetic analysis was performed by using the Phylogeny.fr server [121]. Significance of each node is assessed by approximate likelihood-ratio tests (aLRT) [124].

Table S1 Occurrence of *xyl*-box motif upstream of *Xanthomonas* spp. and *Pseudoxanthomonas suwonensis* genes.

Table S2 Significative homology of the putative ORFs from *X. campestris* pv. *campestris* Xylanolytic CUT system.

Table S3 Conservation of *Xanthomonas campestris* pv *campestris* ATCC33913, *Caulobacter crescentus* CB15 and *Prevotella bryantii* $B_1$4 genes induced by xylan/xylo-oligosaccharides or xylose.

Table S4 List of plasmids and *Xanthomonas campestris* pv. *campestris* strains used or generated in this study.

Cluster ID	Cluster Annotation	Nonredundant Sequences	Total Sequences	NCBI-nr	PG	TGI-EST	ENS	GOS
3510	Immunoglobulin	37,227	51,944	49,206	0	1,649	1,089	0
2568	ABC transporter	34,130	69,010	8,886	6,248	150	13	53,713
49	Short chain dehydrogenase	33,406	56,266	7,607	3,055	2,852	747	42,005
4294	NAD dependent epimerase/dehydratase	29,445	35,555	2,745	1,265	1,500	111	29,934
1239	AMP-binding enzyme	22,111	37,598	3,838	1,614	2,246	613	29,287
2630	Envelope glycoprotein	21,161	41,205	41,189	2	10	0	4
157	Glycosyl transferases group 1	20,366	27,012	2,766	1,446	557	42	22,201
183	Integral membrane protein	17,627	33,079	2,154	1,298	1,198	95	28,334
530	Aldehyde dehydrogenase	15,851	30,929	3,116	1,349	1,589	388	24,487
1308	Aminotransferase class-V and DegT/DnrJ/EryC1/StrS aminotransferase	15,757	22,484	1,849	1,086	413	71	19,065
244	Kinase family, including pknb, epk, c6	15,112	21,641	6,384	83	10,809	2,761	1,604
336	Histidine kinase–, DNA gyrase B–, and HSP90-like ATPase	14,724	23,355	3,809	2,469	54	4	17,019
357	Tetratricopeptide repeat	14,323	17,058	1,598	609	1,320	315	13,216
4325	Alpha/Beta hydrolase fold	13,806	20,886	2,828	1,334	1,625	196	14,903
113	Aminotransferase class I and II	13,006	22,186	2,931	1,534	1,239	120	16,362
333	Zinc-binding dehydrogenase	12,737	22,298	4,055	1,370	2,383	269	14,221
1315	tRNA synthetases class I (I, L, M, and V)	12,545	19,992	1,152	600	472	131	17,637
26	Acyl-CoA dehydrogenase	12,150	22,340	2,081	1,152	541	179	18,387
159	ABC transporter and ABC transporter transmembrane	11,984	17,650	2,697	1,442	797	170	12,544
3357	Cytochrome P450	11,929	17,302	5,355	249	6,994	1,399	3,305
4556	Response regulator	11,928	21,903	5,387	3,320	348	5	12,843
1720	TonB-dependent receptor	11,890	17,080	1,789	1,090	34	2	14,165
514	NADH dehydrogenase (various subunits)	11,224	25,068	11,624	635	253	10	12,546
4235	Glycosyl transferase family 2	10,954	13,593	1,236	724	74	14	11,545
186	7 transmembrane receptor	10,654	22,252	13,943	0	1,475	6,829	5

Tableau D1. Liste des 25 clusters les plus représentés d'après l'expédition « Sorcerer II Global Ocean Sampling » (GOS). (Extrait de Yooseph *et al.*, 2007)

DISCUSSION GENERALE

L'aptitude des bactéries à coloniser une large variété d'habitats reflète une grande capacité d'adaptation et la présence de programmes génétiques spécifiques, dédiés à la reconnaissance et à l'exploitation des nutriments présents dans les différents environnements rencontrés (Galperin, 2004). Il existe une relation entre la niche écologique occupée par une bactérie et la proportion des gènes impliqués dans la perception et la transduction des signaux (Cases and de Lorenzo, 2005). Ainsi, les récepteurs et régulateurs bactériens seraient des marqueurs environnementaux puissants qui permettraient d'apporter un nouveau relief à l'étude de l'évolution et de l'écologie microbienne. Ce postulat est renforcé par les résultats des travaux qui résultent de l'expédition « Sorcerer II Global Ocean Sampling » (GOS). Ce projet de métagénomique sans précédent place les TBDTs avec 11890 séquences non redondantes, parmi les 25 clusters les plus représentés sur les 297254 clusters identifiés (Yooseph *et al.*, 2007) (Tableau D1). Cette étude souligne l'importance des TBDTs dans l'environnement.

En 2007, des travaux menés dans l'équipe ont permis de poser les bases de nouvelles fonctions pour les TBDTs. L'étude systématique des génomes de 226 bactéries à Gram négatif a permis de déterminer une classe particulière de bactéries, dans laquelle on retrouve les *Xanthomonas* spp., où les TBDTs semblent être sur-représentés (Blanvillain *et al.*, 2007). Ces bactéries appartiennent à différents groupes phylogénétiques, mais possèdent la capacité commune de pouvoir dégrader des carbohydrates de végétaux et d'animaux. Les TBDTs seraient donc des marqueurs spécifiques des bactéries pouvant exploiter les plantes ou les débris végétaux et autres carbohydrates complexes.

En effet, sur les 72 TBDTs que possède *Xcc*, seuls 9 seraient impliqués dans l'acquisition du fer et sont régulés par la protéine Fur (Blanvillain *et al.*, 2007). Le TBDT SuxA impliqué dans le transport actif du saccharose a été caractérisé dans ces travaux. Depuis, d'autres études ont montré l'implication des TBDTs dans le transport de substrats très diversifiés (pour revue (Schauer *et al.*, 2008)). De plus, les travaux effectués dans l'équipe au cours de ma thèse ont permis de placer 2 TBDTs au cœur d'un nouveau système CUT impliqué dans le transport et l'utilisation du xylane, constituant majeur de la paroi végétale. De manière intéressante, les bactéries qui ont une sur-représentation en TBDTs au sein de leur génome, possèdent également des CAZyme très importants. Plusieurs de ces enzymes se retrouvent au sein des systèmes CUT. Par ailleurs, l'import actif du saccharose par le TBDT SuxA participe dans le

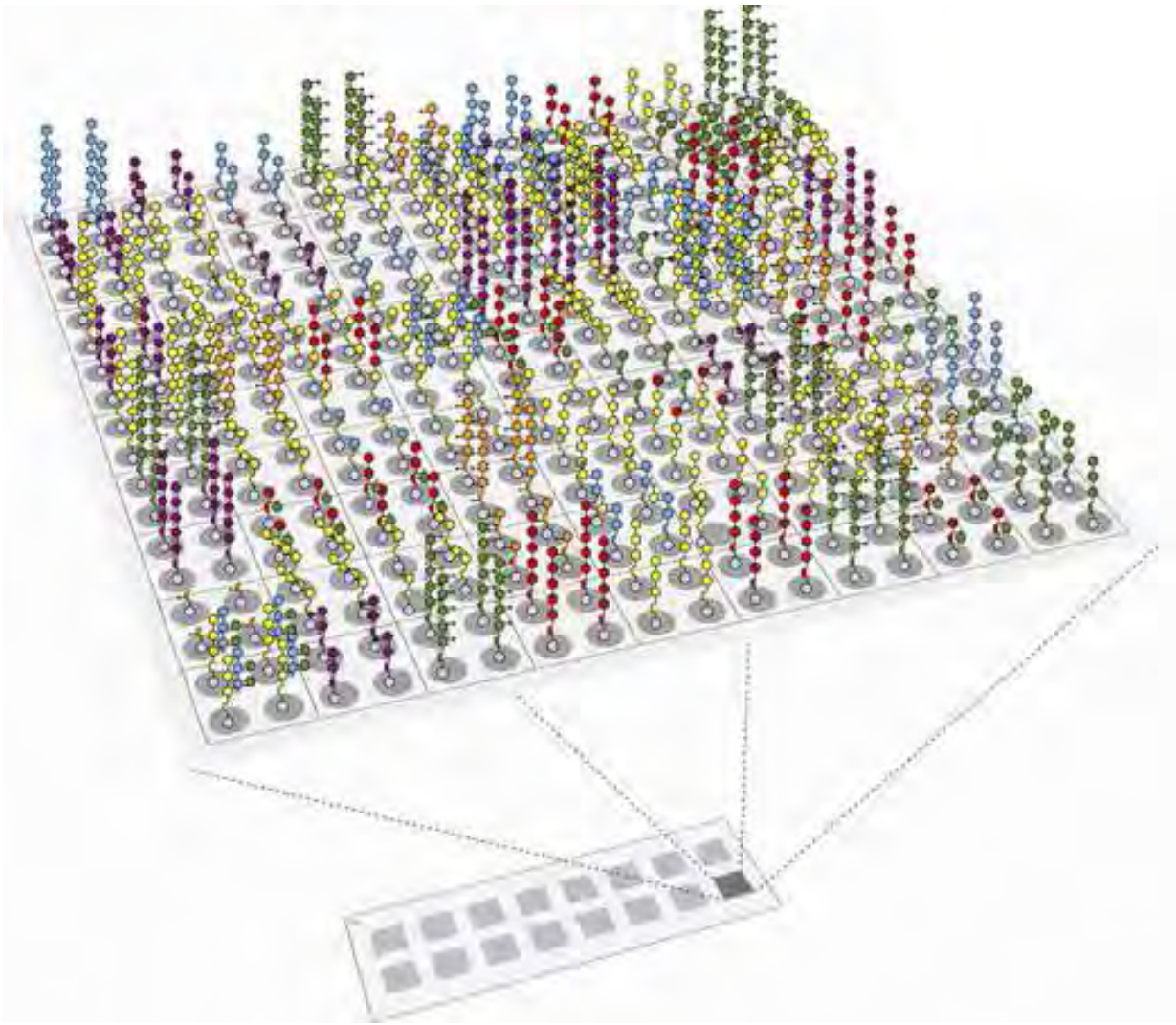


Figure D1. Puce à oligosaccharides
Puce constituée à partir d'une centaine d'oligosaccharides issus de composés de la paroi végétale qui ont été isolés pour la plupart ou produits par synthèse chimique.
Source : http://www.plbio.life.ku.dk/English/Sections/plglyc/Research/Willats/Oligosaccharide_microarrays.aspx

développement de la maladie, permet d'établir un lien entre l'adaptation pathogénique et métabolique.

Ces travaux soulèvent plusieurs questions. Est-ce que la présence de TBDTs spécifiques dans les génomes microbiens apporte une meilleure compréhension de la virulence et de l'adaptation des bactéries à leurs différents environnements ? Pour un organisme donné, pouvons-nous avoir des détails sur les environnements dans lesquels ce dernier évolue ou a évolué en appréhendant la diversité des TBDTs qu'il possède et la nature des substrats transportés ainsi qu'en étudiant leur régulation?

Pour répondre à ces questions, il est important de poursuivre la recherche des substrats potentiels des TBDTs. Un premier criblage de l'expression des TBDTs de *Xcc* a été réalisé en présence de sucres simples tels que le glucose, le xylose, l'arabinose, etc (Blanvillain *et al.*, 2007). Il serait donc utile de se constituer une banque de molécules végétales purifiées ou synthétiques afin d'aller plus avant dans ce crible transcriptionnel. Une collaboration avec l'équipe « Chimie et biotechnologie des oligosaccharides » du CERMAV (Centre de Recherche sur les Macromolécules Végétales, Grenoble, France) permettrait de profiter de leur compétence en synthèse d'oligosaccharides et de glycoconjugués afin de constituer cette banque. L'utilisation de puces à oligosaccharides s'inscrit également parfaitement dans cette approche. Le groupe « Plant cell wall evolution and diversity » de l'université de Copenhague a constitué une banque d'une centaine d'oligosaccharides issus de composés de la paroi végétale afin de construire ces puces permettant de cribler des anticorps, des CBMs, d'autres ligands de glycanes, etc (Moller *et al.*, 2008; Willats *et al.*, 2002) (Figure D1). La croissance des mutants dans les différents TBDTs en présence de cette banque fournira également des informations précieuses quant à l'implication de ces protéines dans l'utilisation des molécules végétales. De plus, des combinaisons de mutants dans les gènes codant pour les TBDTs peuvent être utilisées dans cette approche afin de décrypter la complexité ou la spécificité de fonction de ces transporteurs.

I. Vers la caractérisation d'une nouvelle classe de TBDTs

Le xylane est un composant majeur de la paroi végétale et le second polysaccharide végétal le plus abondant dans la biosphère. Avec la cellulose, le xylane représente plus de 50% de la biomasse végétale. Ces 2 polymères constituent la plus importante source de carbone organique sur la planète (Jeffries, 1990). Un grand nombre de microorganismes est capable de

dégrader ce composé hémicellulolytique. Ces organismes xylanolytiques sont présents au sein de nombreuses niches écologiques (terrestres, aquatiques, systèmes digestifs des animaux) (Collins *et al.*, 2005). Ainsi, l'étude de la dégradation du xylane représente un axe de recherche intéressant quant à l'analyse de la diversité microbienne. Des études de génomique comparative et fonctionnelle ont permis de définir un système d'utilisation du xylane chez les espèces de *Xanthomonas*. Ce système comprend des enzymes de dégradation de cet hydrate de carbone, ainsi que des enzymes impliquées dans le catabolisme du glucuronate et du xylose. On dénombre également 3 transporteurs de la membrane interne et 2 TBDTs dans ce système CUT. Au cours de ma thèse, nous avons montré qu'une grande proportion des gènes du système CUT xylane est bien conservée dans des embranchements très différents des protéobactéries : les Bacteroidetes, les Verrumicrobia et les Acidobacteria. L'étude de ces génomes bactériens a révélé que des TBDTs sont systématiquement associés à plusieurs gènes impliqués dans l'utilisation du xylane. L'analyse génomique *d'Opitutus terrae* PB90-1, bactérie isolée de rizière (van Passel *et al.*, 2011) met en lumière ce résultat. En effet, avec 12 gènes bien conservés, le système CUT xylane est bien représenté. De manière intéressante, 4 gènes localisés à côté de ces gènes conservés coderaient pour des TBDTs putatifs. Ce travail de thèse propose que ces TBDTs constitueraient une nouvelle classe de cette famille protéique. En effet, Ces 4 protéines possèdent un domaine « plug » putatif (domaine pfam07715), spécifique des TBDTs, mais ne disposeraient pas du domaine « β-barrel » (domaine pfam00593) typique des TBDTs (Koebnik, 2005). Cependant, l'analyse de la structure secondaire de ces protéines montre qu'elles possèdent plusieurs feuillets beta ainsi que 12 aa C-terminaux (dont le dernier est aromatique) qui permettent un ancrage membranaire (Koebnik, 1993). Aucune similarité significative n'est observée entre ces 4 protéines et les TBDTs associés au xylane ou à d'autres TBDTs caractérisés chez les protéobactéries. Pour un même substrat, des TBDTs très différents peuvent avoir évolué séparément. Les TBDTs étant des protéines de la membrane externe des bactéries, ils font partie des premiers acteurs potentiels impliqués dans la reconnaissance de la bactérie par son hôte et ainsi de l'activation de son système immunitaire. Chez la bactérie phytopathogène *R. solanacearum*, le TBDT PrhA est induit et est responsable de l'induction de l'expression des gènes *hrp* en réponse à un contact avec les cellules végétales (Aldon *et al.*, 2000; Marenda *et al.*, 1998). De la même manière, on peut penser que des voies de signalisation sont également activées chez l'hôte végétal. Ainsi, des variations dans la structure des TBDTs pourraient

résulter du besoin nécessaire d'échapper aux mécanismes de reconnaissance de l'hôte chez les bactéries pathogènes.

Une autre question essentielle qui émane de l'analyse de ces séquences génomiques est de comprendre si la dégradation du xylane dans les divers écosystèmes contribue ou non à la subsistance et à la dynamique des populations de la communauté microbienne dans ces environnements. Ce travail de thèse démontre la présence de complexes construits autour des TBDTs qui permettent l'utilisation du xylane, chez un grand nombre de bactéries phylogénétiquement très différentes, suggérant de ce fait l'importance de l'utilisation du xylane et le rôle central joué par la perception et le transport des sous produits du xylane par ces transporteurs.

II. Le décryptage des réseaux de régulation contrôlant l'adaptation de Xcc à la plante

Pour le système CUT xylane, nous avons pu montrer que l'expression de l'ensemble des gènes le composant est contrôlée par le xylose, les xylo-oligosaccharides de faible taille (X_2-X_4) et le xylane. Nous avons également observé des profils différentiels d'expression selon les sucres suggérant l'existence de voies de régulation entrelacées contrôlées par le xylose et les xylo-oligosaccharides (X_2-X_4). De plus, une protéine régulatrice, XylR, a pu être identifiée, notamment grâce à l'occurrence de motifs conservés dans les promoteurs de gènes du système CUT. XylR réprime l'expression des gènes appartenant aux loci *xytA* et *xytB*, ainsi que le gène *xylA2* présent dans le locus *xylR*. En revanche, il n'affecte pas l'expression des autres gènes du locus *xylR* ni ceux du locus *xylE*, augmentant ainsi la complexité du schéma régulationnel de ce système. Chez la bactérie du sol *Cellvibrio japonicus*, tout comme chez la bactérie *Prevotella bryantii* $B_1$4, les gènes impliqués dans l'utilisation du xylane sont soumis à d'autres systèmes de régulation impliquant des systèmes à deux composantes par exemple (Emami *et al.*, 2002; Emami *et al.*, 2009; Miyazaki *et al.*, 2003; Miyazaki *et al.*, 2005). Ainsi, d'après les résultats obtenus au cours de cette thèse et tout comme il a été observé chez d'autres bactéries xylanolytiques, il est probable que d'autres systèmes interviennent dans la régulation de l'expression des gènes du système CUT xylane chez *Xcc*.

Nous avons montré que l'expression du gène *xyn10C* codant pour une xylanase putative est activée par HrpG et HrpX, les régulateurs majeurs du SST3. D'autres gènes codant pour des

enzymes de dégradation de la paroi végétale sont sous le contrôle de ces 2 régulateurs. Chez la souche 8004 de *Xcc*, les 2 polygalacturonases extracellulaires codées par *pghAxc* et *pghBxc*, sont sécrétées par le SST2 et leur expression est induite par HrpG et HrpX (Wang *et al.*, 2008). Récemment, des travaux menés chez *Xcv* indiquent que le gène *xynC* codant pour une xylanase est aussi régulé par HrpG et HrpX. De plus, les auteurs montrent qu'un SST2 fonctionnel est nécessaire pour une translocation optimale des protéines effectrices de *Xcv*, suggérant une balance fonctionnelle entre les deux systèmes de sécrétion. Il est supposé que les enzymes de dégradation de la paroi végétale pourraient faciliter la mise en place des composants extracellulaires du SST3 (Szczesny *et al.*, 2010). Il sera donc intéressant d'intégrer l'étude du SST2 dans notre système afin d'aller plus en avant dans l'étude des enzymes de dégradation.

Au cours de ce travail, nous avons aussi montré que les régulateurs HrpG et HrpX répriment l'activité xylanase chez *Xcc*. Cependant, il semble que l'activité xylanase observée n'est pas dépendante de Xyn10C, mais de Xyn10A qui n'est pas sous le contrôle de HrpG et HrpX. Ces données suggèrent l'existence d'un autre niveau de régulation qui n'est pas encore totalement compris. Il reste encore à décrypter l'articulation de la régulation de ces régulateurs globaux avec celle effectuée par des régulateurs plus spécifiques.

De plus, pour plusieurs gènes du système CUT xylane, le mode de régulation n'a pas pu être complètement identifié au cours de ce travail. C'est notamment le cas des gènes du locus *xylE* et de ceux en aval du gène *xylR* dans le locus *xylR*. Bien qu'ils soient induits par les xylo-oligosaccharides, il semble que leur expression ne dépende pas de XylR. Une hypothèse intéressante serait l'intervention de petits ARNs non codant, les sRNAs (small noncoding RNAs). On retrouve les sRNAs dans tous les règnes de la vie. Chez les bactéries, ils représentent un groupe majeur de molécules régulatrices qui leur permettent de contrôler différents processus cellulaires tels que l'homéostasie du fer (Wilderman *et al.*, 2004), le métabolisme de sucres (Vanderpool and Gottesman, 2004), la réponse aux stress (Wassarman, 2002), le quorum sensing (Bejerano-Sagie and Xavier, 2007) et la virulence (Ma *et al.*, 2001; Papenfort *et al.*, 2010; Romby *et al.*, 2006). Des analyses bioinformatiques estiment qu'il y aurait plusieurs centaines de sRNAs par génome bactérien (Vogel and Sharma, 2005; Zhang *et al.*, 2004). Avec un nombre de 200 à 300 sRNAs estimés dans les génomes des entérobactéries qui ont une taille moyenne de 4 à 5 Mb, ces petits ARNs représenteraient environ 5% des ORFs totaux (Zhang *et al.*, 2004). Actuellement, plus de 200 sRNAs ont été caractérisés. Plus de la moitié l'ont été chez *E. coli*, et l'autre moitié provient majoritairement

des bactéries *Caulobacter crescentus*, *Pseudomonas aeruginosa*, *Salmonella thyphimurium*, *Sinorhizobium meliloti* et *Staphylococcus aureus*. Chez *E. coli*, environ un tiers des sRNAs dont on connaît les fonctions cellulaires contrôlent la régulation de l'expression de protéines de la membrane externe (Vogel and Papenfort, 2006). Du fait de ses propriétés physiques, la membrane externe fonctionne comme une barrière sélective et joue un rôle essentiel dans la survie de la bactérie dans divers environnements. Etant donné son importance, il n'est pas étonnant de voir que les protéines de la membrane externe, qui permettent l'adaptation bactérienne à divers stimuli environnementaux, sont très régulées au niveau transcriptionnel.

De manière intéressante, le sRNA régulateur CrfA permet l'adaptation de *C. crescentus* à une carence en carbone (Landt *et al.*, 2010). Un tiers des gènes activés par CrfA sont des transporteurs putatifs des membranes externes et internes, et la majorité d'entre eux sont des TBDTs. Les TBDTs de *C. crescentus* comme ceux de *Xcc* permettent le transport au travers de la membrane externe d'un grand nombre de substrats (Eisenbeis *et al.*, 2008; Neugebauer *et al.*, 2005; Schauer *et al.*, 2008). Ainsi, l'activation de ces protéines par CrfA en conditions de carence en carbone permettrait l'import d'une grande variété de sources potentielles de carbone. Peu de travaux portent sur la caractérisation de sRNAs chez des bactéries phytopathogènes. Pourtant, une prédiction bioinformatique estime qu'il y aurait en moyenne 36 loci potentiels par Mb qui contiendraient un sRNA chez *Xanthomonas* spp. (Livny *et al.*, 2008).

Une étude récente rapporte la caractérisation de 8 sRNAs chez *Xoo* PXO99. La comparaison des profils protéomiques entre des mutants chez 3 de ces petits ARNs et la souche sauvage révèle que 8 gènes codant pour des TBDTs sont régulés par ces derniers (Liang *et al.*, 2011). Ils constituent le plus grand groupe fonctionnel de gènes associés à la virulence ciblés par ces sRNAs. De plus, avec 41 TBDTs présents dans le génome de *Xoo* PXO99, ce groupe de gène représente 1/5ème de la totalité de cette famille de transporteurs.

En 2010, des travaux ont permis la caractérisation des 4 premiers sRNAs chez la souche 8004 de *Xcc* (Jiang *et al.*, 2010). Ces sRNAs sont parfaitement conservés chez la souche ATCC33913 de *Xcc*. Des travaux préliminaires réalisés au cours de ma thèse en utilisant le logiciel « sRNA Target » (Cao *et al.*, 2009) ont permis de déterminer les cibles potentielles de ces 4 sRNAs chez la souche ATCC33913 de *Xcc* (résultats non montrés). Plusieurs TBDTs et des régulateurs putatifs se retrouvent parmi les gènes potentiellement régulés par ces sRNAs. De façon toute aussi intéressante, on retrouve parmi ces gènes plusieurs gènes (*axeXA*, *gly43E* et *xyl3A*) appartenant au locus *xylR* et un gène (*xylS*) appartenant au locus *xylE*, identifiés au

cours de mon travail de thèse. L'expression de ces 3 gènes est induite par les xylo-oligosaccharides de faible taille, mais n'est pas soumise à la répression exercée par la protéine XylR. Bien qu'il faille prendre ces résultats avec énormément de prudence, ils permettent d'ouvrir des hypothèses quant aux autres niveaux de régulation du système CUT xylane, et plus largement aux gènes intervenant dans la perception et la régulation de signaux environnementaux chez *Xcc*.

Les sRNAs sont un groupe de molécules régulatrices qui permettent à la bactérie de s'adapter à différents stress environnementaux. La grande place occupée par les TBDTs parmi les gènes ciblés par ces petits ARNs conforte l'hypothèse du rôle joué par ces protéines dans l'écologie bactérienne.

III. Vers une meilleure compréhension du rôle des CUT loci dans le pouvoir pathogène

La méthode d'inoculation classique, qui consiste en une inoculation directe par piercing dans le système vasculaire de la plante (Meyer *et al.*, 2005), n'a pas permis de montrer une implication du système CUT xylane dans la virulence de *Xcc*. Cependant, au cours de ce travail de thèse, l'étude de la dynamique des populations en inoculant par pulvérisation les bactéries à la surface des feuilles, a permis de montrer que les locus *xytA* et *xytB* sont nécessaires pour une croissance optimale de *Xcc* dans la phyllosphère de plantes hôtes et non hôtes. En contournant les étapes de vie épiphyte et de l'infection naturelle via les hydathodes, notre principale méthode d'étude du pouvoir pathogène ne prend pas en compte la totalité du cycle de vie de *Xcc*. Les résultats de ce travail de thèse montrent que ces étapes sont pourtant cruciales. Il est donc impératif d'optimiser nos tests de pouvoir pathogène afin de refléter au mieux la fitness de la bactérie. Dans ce sens, nous développons actuellement dans le laboratoire un nouveau test de compétition qui consiste en un dosage moléculaire d'un polymorphisme entre 2 souches de *Xcc* permettant de déterminer l'abondance relative de ces 2 souches dans des tissus végétaux. Enfin, plus largement, ces données posent la question du niveau d'intervention des TBDTs et des systèmes CUT dans le cycle de vie de la bactérie. Ces systèmes interviennent-ils dans la spécificité d'hôtes, de tissus ou de symptômes ?

IV. Identification de marqueurs de spécificité d'hôtes ?

Actuellement, 11 génomes de *Xanthomonas* sont entièrement séquencés et 10 autres sont en cours (Ryan *et al.*, 2011). Ces génomes peuvent être divisés en 2 groupes : ceux qui infectent les dicotylédones et ceux qui infectent les monocotylédones. Au sein du système CUT xylane, le locus *xytB* est le plus variable et est donc potentiellement un bon outil en génomique comparative. Au cours de ce travail de thèse, nous avons montré que le gène *gly2A* est présent chez les souches infectant les dicotylédones et absent chez les souches infectant les monocotylédones. Un profil quasi similaire est observé pour le gène *xyn10C* codant pour une xylanase putative. En effet, on retrouve ce gène seulement chez *X. vasicola* pv. *vasculorum* et chez *X. campestris* pv. *musacearum*, avec la particularité de ne plus être lié génétiquement au locus *xytB*. Ces 2 derniers gènes, *gly2A* et *xyn10C*, sont conservés chez *Pseudoxanthomonas suwonensis* 11-1. Phylogénétiquement, cette bactérie est à la base de la famille des *Xanthomonadaceae*, suggérant une perte des gènes *xyn10C* et *gly2A* chez les souches infectant les monocotylédones. De façon aussi spectaculaire, le gène *xyn10A* codant pour la xylanase responsable de l'activité du même nom chez *Xcc*, est présent chez toutes les souches pathogènes de dicotylédones et absent chez les autres, à l'exception des souches de *X. oryzae*. Cependant, les protéines Xyn10A de ces souches ne possèdent pas de séquence signal, alors que celle des souches qui attaquent les dicotylédones en contiennent une. Enfin, on retrouve également ce profil de présence/absence selon l'hôte pour 2 autres gènes potentiellement impliqués dans la dégradation du xylane, *gly43C* qui code pour une β-xylosidase/arabinofuranosidase putative et *xyl39A* qui code pour une β-xylosidase. Ces gènes constitueraient donc des marqueurs de spécificité d'hôte et renforcent ainsi l'hypothèse de l'implication des systèmes CUT dans la « fitness » de la bactérie. De manière intéressante, ces gènes codent tous pour des enzymes potentiellement impliquées dans la dégradation du xylane. Il a été montré qu'il existe des différences de teneur et même de structure du xylane entre les plantes dicotylédones et les monocotylédones (Scheller and Ulvskov, 2010), suggérant une spécialisation enzymatique selon l'hôte envahi. Sonnenburg et al. ont réalisé des analyses de génomique comparative entre 6 espèces séquencées de Bacteroides, des bactéries intestinales humaines, en se concentrant sur le PUL responsable de l'utilisation du fructane (Sonnenburg *et al.*, 2010). Il ressort de cette étude que ces 6 espèces partagent la capacité d'utilisation du fructane, mais tous ne vont pas croitre sur le même type de fructane. Cette spécificité correspond aux différences de composition génétiques des PULs fructane de

ces espèces. Enfin, en utilisant des souris gnotobiotiques, les auteurs démontrent que le régime en fructane a un effet direct sur la composition de la communauté intestinale, selon la capacité d'utilisation du fructane du microbiote.

V. La métagénomique : un outil puissant pour une description efficace de l'adaptation et de l'évolution de *Xcc*

L'adaptation des microbes pathogènes à leurs hôtes repose sur l'équilibre entre 3 déterminants microbiens qui sont la virulence, la régulation et le métabolisme. Dans notre cas, à savoir l'étude de l'adaptation et du pouvoir pathogène des *Xanthomonas* et plus largement des bactéries à Gram négatif, l'étude du SST3 au sens large permettrait d'appréhender la virulence, et l'analyse des TBDTs avec les systèmes CUT donnerait une bonne vision de la régulation et du métabolisme bactérien. Ainsi, la poursuite de l'étude de la conservation des systèmes CUT chez les bactéries et du rôle précis des TBDTs nous permettra de déterminer s'ils reflètent réellement les capacités adaptatives des bactéries face à divers environnements. Cependant, ces approches mettent de côté les dynamiques des populations microbiennes au sein d'un environnement, ou encore les transferts horizontaux. Ainsi, intégrer une approche de métagénomique à ce travail permettrait d'intégrer ces problématiques et de décrypter l'évolution des pathogènes. Il a été estimé que seul 1% des procaryotes de la plupart des environnements peuvent être cultivés (Amann *et al.*, 1990), et que l'analyse écologique d'un organisme est primordiale. La métagénomique ou encore génomique environnementale ou génomique des commmunautés est donc un outil puissant qui permettrait de dévoiler les fonctions biologiques qui peuvent exister dans un milieu donné.

La métagénomique permettrait de caractériser des fonctions spécifiques associées avec un environnement particulier. Un débouché potentiel serait l'identification mais également l'exploitation de nouvelles enzymes. En effet, lors de l'expédition « Sorcerer II Global Ocean Sampling », 1200000 nouveaux gènes ont été identifiés (Kannan *et al.*, 2007; Yooseph *et al.*, 2007).Un intérêt supplémentaire de cette technique est de faire de la génomique comparative dans le but d'identifier des gènes spécifiques à un environnement donné, nous donnant ainsi des armes supplémentaires pour la description de l'évolution et de l'adaptation du microbe (Tringe *et al.*, 2005). Enfin, cela peut aussi aboutir à la caractérisation de nouvelles espèces microbiennes.

REFERENCES BIBLIOGRAPHIQUES

Ahmed, S.I., and Leather, S.R. (1994) Suitability and potential of entomopathogenic microorganisms for forest pest management - some points for consideration. *International journal of pest management* **40**: 287-292.

Ahn, B.E., Cha, J., Lee, E.J., Han, A.R., Thompson, C.J., and Roe, J.H. (2006) Nur, a nickel-responsive regulator of the Fur family, regulates superoxide dismutases and nickel transport in *Streptomyces coelicolor*. *Mol Microbiol* **59**: 1848-1858.

Alavi, S.M., Sanjari, S., Durand, F., Brin, C., Manceau, C., and Poussier, S. (2008) Assessment of the genetic diversity of *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans* as a basis to identify putative pathogenicity genes and a type III secretion system of the SPI-1 family by multiple suppression subtractive hybridizations. *Appl Environ Microbiol* **74**: 3295-3301.

Aldon, D., Brito, B., Boucher, C., and Genin, S. (2000) A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. *Embo J* **19**: 2304-2314.

Alfano, J.R., and Collmer, A. (1997) The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *J Bacteriol* **179**: 5655-5662.

Alfano, J.R., and Collmer, A. (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu Rev Phytopathol* **42**: 385-414.

Ali, B.R., Romaniec, M.P., Hazlewood, G.P., and Freedman, R.B. (1995) Characterization of the subunits in an apparently homogeneous subpopulation of *Clostridium thermocellum* cellulosomes. *Enzyme Microb Technol* **17**: 705-711.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., and Stahl, D.A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919-1925.

Anderson, K.L., and Salyers, A.A. (1989a) Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J Bacteriol* **171**: 3192-3198.

Anderson, K.L., and Salyers, A.A. (1989b) Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol* **171**: 3199-3204.

Andrade, M.O., Alegria, M.C., Guzzo, C.R., Docena, C., Rosa, M.C., Ramos, C.H., and Farah, C.S. (2006) The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis* pv *citri*. *Mol Microbiol* **62**: 537-551.

Andrews, S.C., Robinson, A.K., and Rodriguez-Quinones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**: 215-237.

Arabidopsis (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796-815.

Arlat, M., Van Gijsegem, F., Huet, J.C., Pernollet, J.C., and Boucher, C.A. (1994) PopA1, a protein which induces a hypersensitivity-like response on specific Petunia genotypes, is secreted via the Hrp pathway of *Pseudomonas solanacearum*. *Embo J* **13**: 543-553.

Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., Niinikoski, A., Mewes, H.W., Horn, M., and Rattei, T. (2009) Sequence-based prediction of type III secreted proteins. *PLoS Pathog* **5**: e1000376.

Aslam, S.N., Newman, M.A., Erbs, G., Morrissey, K.L., Chinchilla, D., Boller, T., Jensen, T.T., De Castro, C., Ierano, T., Molinaro, A., Jackson, R.W., Knight, M.R., and Cooper, R.M. (2008) Bacterial polysaccharides suppress induced innate immunity by calcium chelation. *Curr Biol* **18**: 1078-1083.

Bacic, A. (1988) Structure and function of plant cell walls. *Academic Press, San diego*.

Bagg, A., and Neilands, J.B. (1987) Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* **26**: 5471-5477.

Barber, C.E., Tang, J.L., Feng, J.X., Pan, M.Q., Wilson, T.J., Slater, H., Dow, J.M., Williams, P., and Daniels, M.J. (1997) A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol Microbiol* **24**: 555-566.

Bassler, B.L. (2002) Small talk. Cell-to-cell communication in bacteria. *Cell* **109**: 421-424.

Bauer, S., Vasu, P., Persson, S., Mort, A.J., and Somerville, C.R. (2006) Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc Natl Acad Sci U S A* **103**: 11417-11422.

Bayer, E.A., Setter, E., and Lamed, R. (1985) Organization and distribution of the cellulosome in *Clostridium thermocellum*. *J Bacteriol* **163**: 552-559.

Bayer, E.A., Morag, E., and Lamed, R. (1994) The cellulosome: a treasure-trove for biotechnology. *Trends Biotechnol* **12**: 379-386.

Bayer, E.A., Chanzy, H., Lamed, R., and Shoham, Y. (1998) Cellulose, cellulases and cellulosomes. *Curr Opin Struct Biol* **8**: 548-557.

Beattie, G.A., and Lindow, S.E. (1999) Bacterial colonization of leaves: a spectrum of strategies. *Phytopathology* **89**: 353-359.

Becker, A., Katzen, F., Puhler, A., and Ielpi, L. (1998) Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Appl Microbiol Biotechnol* **50**: 145-152.

Bejerano-Sagie, M., and Xavier, K.B. (2007) The role of small RNAs in quorum sensing. *Curr Opin Microbiol* **10**: 189-198.

Bender, C.L., Alarcon-Chaidez, F., and Gross, D.C. (1999) *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* **63**: 266-292.

Bent, A.F., Innes, R.W., Ecker, J.R., and Staskawicz, B.J. (1992) Disease development in ethylene-insensitive *Arabidopsis thaliana* infected with virulent and avirulent *Pseudomonas* and *Xanthomonas* pathogens. *Mol Plant Microbe Interact* **5**: 372-378.

Biely, P., Kratky, Z., and Vrsanska, M. (1981) Substrate-binding site of endo-1,4-beta-xylanase of the yeast *Cryptococcus albidus*. *Eur J Biochem* **119**: 559-564.

Biely, P., Vrsanska, M., Tenkanen, M., and Kluepfel, D. (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. *J Biotechnol* **57**: 151-166.

Birch, R.G., and Patil, S.S. (1985) Preliminary characterization of an antibiotic produced by *Xanthomonas albilineans* which inhibits DNA synthesis in *Escherichia coli*. *J Gen Microbiol* **131**: 1069-1075.

Birch, R.G., and Patil, S.S. (1987) Evidence that an albicidin-like phytotoxin induces chlorosis in sugarcane leaf scald disease by blocking plastid DNA replication. *Physiol Mol Plant Pathol* **30**.

Blair, B.G., and Anderson, K.L. (1999) Regulation of cellulose-inducible structures of *Clostridium cellulovorans*. *Can J Microbiol* **45**: 242-249.

Blanvillain, S. (2007) Etude des récepteurs TonB-dépendants et d'un nouvel effecteur de type III de la bactérie phytopathogène *Xanthomonas campestris* pv. *campestris. Thèse*.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., Vasse, J., Lauber, E., and Arlat, M. (2007) Plant carbohydrate scavenging through Tonb-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**: e224.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509-1512.

Boch, J., and Bonas, U. (2010) *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol* **48**: 419-436.

Boelaert, J.R. (1996) Iron and infection. *Acta clinica belgica* **51**: 213-221.

Bogdanove, A.J., Beer, S.V., Bonas, U., Boucher, C.A., Collmer, A., Coplin, D.L., Cornelis, G.R., Huang, H.C., Hutcheson, S.W., Panopoulos, N.J., and Van Gijsegem, F. (1996) Unified nomenclature for broadly conserved *hrp* genes of phytopathogenic bacteria. *Mol Microbiol* **20**: 681-683.

Bogdanove, A.J., Schornack, S., and Lahaye, T. (2010) TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* **13**: 394-401.

Bonas, U. (1994) *hrp* genes of phytopathogenic bacteria. *Curr Top Microbiol Immunol* **192**: 79-98.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *The Biochemical Journal* **382**: 769-781.

Bordes, P., Lavatine, L., Phok, K., Barriot, R., Boulanger, A., Castanie-Cornet, M.P., Dejean, G., Lauber, E., Becker, A., Arlat, M., and Gutierrez, C. (2011) Insights into the extracytoplasmic stress response of Xanthomonas campestris pv. campestris: role and regulation of {sigma}E-dependent activity. *J Bacteriol* **193**: 246-264.

Boudet, A.M., Kajita, S., Grima-Pettenati, J., and Goffner, D. (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci* **8**: 576-581.

Boulanger, A. (2009) Analyse d'un nouveau système CUT impliqué dans l'acquisition et l'utilisation du N-acétylglucosamine par Xanthomonas campestris pathovar campestris. *Thèse*.

Boulanger, A., Dejean, G., Lautier, M., Glories, M., Zischek, C., Arlat, M., and Lauber, E. (2010) Identification and regulation of the *N*-acetylglucosamine utilization pathway of the plant pathogenic bacterium *Xanthomonas campestris* pv. *campestris*. *J Bacteriol* **192**: 1487-1497.

Braun, V. (1995) Energy-coupled transport and signal transduction through the gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiol Rev* **16**: 295-307.

Braun, V. (1997) Avoidance of iron toxicity through regulation of bacterial iron transport. *Biol Chem* **378**: 779-786.

Braun, V. (1999) Active transport of siderophore-mimicking antibacterials across the outer membrane. *Drug Resist Updat* **2**: 363-369.

Braun, V., and Killmann, H. (1999) Bacterial solutions to the iron-supply problem. *Trends Biochem Sci* **24**: 104-109.

Braun, V., and Braun, M. (2002) Iron transport and signaling in *Escherichia coli*. *FEBS Lett* **529**: 78-85.

Brett, C.T., and Waldren, K. (1996) *Physiology and Biochemistry of Plant Cell Walls*: Chapman and Hall Country.

Buchanan, S.K., Smith, B.S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J. (1999) Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nat Struct Biol* **6**: 56-63.

Buckeridge, M., Santos, H., and Tiné, M. (2000) Mobilisation of storage cell wall polysaccharides in seeds. *Plant Physiol Biochem* **38**: 141-156.

Buell, C. (2002) Interactions between *Xanthomonas* species and *Arabidopsis thaliana*. In *The Arabidopsis Book*. Somerville, C., Meyerowitz EM., eds (ed). Rockville: American Society of Plant Biologists.

Büttner, D., and Bonas, U. (2002) Getting across--bacterial type III effector proteins on their way to the plant cell. *Embo J* **21**: 5313-5322.

Büttner, D., Nennstiel, D., Klusener, B., and Bonas, U. (2002) Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol* **184**: 2389-2398.

Büttner, D., Noël, L., Thieme, F., and Bonas, U. (2003) Genomic approaches in *Xanthomonas campestris* pv. *vesicatoria* allow fishing for virulence genes. *J Biotechnol* **106**: 203-214.

Büttner, D., and Bonas, U. (2010) Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol Rev* **34**: 107-133.

Canonne, J., Marino, D., Noël, L.D., Arechaga, I., Pichereaux, C., Rossignol, M., Roby, D., and Rivas, S. (2010) Detection and functional characterization of a 215 amino acid N-terminal extension in the *Xanthomonas* type III effector XopD. *PLoS One* **5**: e15773.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**: D233-238.

Cao, Y., Zhao, Y., Cha, L., Ying, X., Wang, L., Shao, N., and Li, W. (2009) sRNATarget: a web server for prediction of bacterial sRNA targets. *Bioinformation* **3**: 364-366.

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

Carpita, N.C. (1996) Structure and Biogenesis of the Cell Walls of Grasses. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 445-476.

Cascales, E., Buchanan, S.K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S., and Cavard, D. (2007) Colicin biology. *Microbiol Mol Biol Rev* **71**: 158-229.

Cases, I., and de Lorenzo, V. (2005) Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. *Int Microbiol* **8**: 213-222.

Castaneda, A., Reddy, J.D., El-Yacoubi, B., and Gabriel, D.W. (2005) Mutagenesis of all eight *avr* genes in *Xanthomonas campestris* pv. *campestris* had no detected effect on pathogenicity, but one *avr* gene affected race specificity. *Mol Plant Microbe Interact* **18**: 1306-1317.

Cavalier, D.M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., Zabotina, O.A., Hahn, M.G., Burgert, I., Pauly, M., Raikhel, N.V., and Keegstra, K. (2008) Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* **20**: 1519-1537.

Cha, J.Y., Lee, J.S., Oh, J.I., Choi, J.W., and Baik, H.S. (2008) Functional analysis of the role of Fur in the virulence of *Pseudomonas syringae* pv. *tabaci* 11528: Fur controls expression of genes involved in quorum-sensing. *Biochem Biophys Res Commun* **366**: 281-287.

Chan, J.W., and Goodwin, P.H. (1999) The molecular genetics of virulence of *Xanthomonas campestris*. *Biotechnol Adv* **17**: 489-508.

Chen, L.Q., Hou, B.H., Lalonde, S., Takanaga, H., Hartung, M.L., Qu, X.Q., Guo, W.J., Kim, J.G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F.F.,

Somerville, S.C., Mudgett, M.B., and Frommer, W.B. (2010) Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**: 527-532.

Chevance, F.F., and Hughes, K.T. (2008) Coordinating assembly of a bacterial macromolecular machine. *Nat Rev Microbiol* **6**: 455-465.

Chimento, D.P., Kadner, R.J., and Wiener, M.C. (2003) The *Escherichia coli* outer membrane cobalamin transporter BtuB: structural analysis of calcium and substrate binding, and identification of orthologous transporters by sequence/structure conservation. *J Mol Biol* **332**: 999-1014.

Chimento, D.P., Kadner, R.J., and Wiener, M.C. (2005) Comparative structural analysis of TonB-dependent outer membrane transporters: implications for the transport cycle. *Proteins* **59**: 240-251.

Chou, F.L., Chou, H.C., Lin, Y.S., Yang, B.Y., Lin, N.T., Weng, S.F., and Tseng, Y.H. (1997) The *Xanthomonas campestris gumD* gene required for synthesis of xanthan gum is involved in normal pigmentation and virulence in causing black rot. *Biochem Biophys Res Commun* **233**: 265-269.

Chun, W., Cui, J., and Poplawsky, A. (1997) Purification, characterization and biological role of a pheromone produced by *Xanthomonas campestris* pv. *campestris*. *Physiological and Molecular Plant Pathology* **51**: 1-14.

Cianciotto, N.P. (2005) Type II secretion: a protein secretion system for all seasons. *Trends Microbiol* **13**: 581-588.

Clay, N.K., Adio, A.M., Denoux, C., Jander, G., and Ausubel, F.M. (2009) Glucosinolate metabolites required for an *Arabidopsis* innate immune response. *Science* **323**: 95-101.

Cobessi, D., Meksem, A., and Brillet, K. (2010) Structure of the heme/hemoglobin outer membrane receptor ShuA from *Shigella dysenteriae*: heme binding by an induced fit mechanism. *Proteins* **78**: 286-294.

Collins, T., Gerday, C., and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol Rev* **29**: 3-23.

Cook, A., Walker, J. and Larson, R. (1952) Studies on the disease cycle of black rot of crucifers. *Phytopathology*: 162-167.

Cooper, R.M. (1984) The role of cell wall-degrading enzymes in infection and damage. In *Plant Diseases: Infection, Damage and Loss*. R.K.S. Wood and G.J. Jellis, e. (ed): Oxford: Blackwell Scientific Publications, pp. 13-27.

Cosgrove, D.J. (1999) Enzymes and other agents that enhance cell wall extensibility. *Annu Rev Plant Physiol Plant Mol Biol* **50**: 391-417.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* **6**: 850-861.

Coughlan, M.P. (1985) Cellulose hydrolysis: the potential, the problems and relevant research at Galway. *Biochem Soc Trans* **13**: 405-406.

Craig, L., and Li, J. (2008) Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* **18**: 267-277.

Crossman, L., and Dow, J.M. (2004) Biofilm formation and dispersal in *Xanthomonas campestris*. *Microbes Infect* **6**: 623-629.

Cunnac, S., Boucher, C., and Genin, S. (2004a) Characterization of the cis-acting regulatory element controlling HrpB-mediated activation of the type III secretion system and effector genes in *Ralstonia solanacearum*. *J Bacteriol* **186**: 2309-2318.

Cunnac, S., Occhialini, A., Barberis, P., Boucher, C., and Genin, S. (2004b) Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum*: identification of novel effector proteins translocated to plant host cells through the type III secretion system. *Mol Microbiol* **53**: 115-128.

da Silva, A.C., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida, N.F., Alves, L.M., do Amaral, A.M.,

Bertolini, M.C., Camargo, L.E., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L.P., Cicarelli, R.M., Coutinho, L.L., Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J., Ferreira, R.C., Ferro, M.I., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite, R.P., Lemos, E.G., Lemos, M.V., Locali, E.C., Machado, M.A., Madeira, A.M., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira, H.A., Rossi, A., Sena, J.A., Silva, C., de Souza, R.F., Spinola, L.A., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezza, R.I., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C., and Kitajima, J.P. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**: 459-463.

Dale, C., Young, S.A., Haydon, D.T., and Welburn, S.C. (2001) The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. *Proc Natl Acad Sci U S A* **98**: 1883-1888.

Darsonval, A., Darrasse, A., Durand, K., Bureau, C., Cesbron, S., and Jacques, M.A. (2009) Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Mol Plant Microbe Interact* **22**: 747-757.

Darvill, A.G., Albersheim, P., McNeil, M., Lau, J.M., York, W.S., Stevenson, T.T., Thomas, J., Doares, S., Gollin, D.J., Chelf, P., and et al. (1985) Structure and function of plant cell wall polysaccharides. *J Cell Sci Suppl* **2**: 203-217.

Das, A., Rangaraj, N., and Sonti, R.V. (2009) Multiple adhesin-like functions of *Xanthomonas oryzae* pv. *oryzae* are involved in promoting leaf attachment, entry, and virulence on rice. *Mol Plant Microbe Interact* **22**: 73-85.

Davis, G.S., Flannery, E.L., and Mobley, H.L. (2006) *Helicobacter pylori* HP1512 is a nickel-responsive NikR-regulated outer membrane protein. *Infect Immun* **74**: 6811-6820.

de la Paz Sanchez, M., and Gutierrez, C. (2009) *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. *Proc Natl Acad Sci U S A* **106**: 2065-2070.

Demain, A.L., Newcomb, M., and Wu, J.H. (2005) Cellulase, *clostridia*, and ethanol. *Microbiol Mol Biol Rev* **69**: 124-154.

Denny, T.P. (1995) Involvement of bacterial polysaccharides in plant pathogenesis. *Annu Rev Phytopathol* **33**: 173-197.

Dharmapuri, S., and Sonti, R.V. (1999) A transposon insertion in the *gumG* homologue of *Xanthomonas oryzae* pv. *oryzae* causes loss of extracellular polysaccharide production and virulence. *FEMS Microbiol Lett* **179**: 53-59.

Diaz-Mireles, E., Wexler, M., Sawers, G., Bellini, D., Todd, J.D., and Johnston, A.W. (2004) The Fur-like protein Mur of *Rhizobium leguminosarum* is a Mn(2+)-responsive transcriptional regulator. *Microbiology* **150**: 1447-1456.

Dixon, R., and Kahn, D. (2004) Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol* **2**: 621-631.

Dodd, D., and Cann, I.K. (2009) Enzymatic deconstruction of xylan for biofuel production. *Glob Change Biol Bioenergy* **1**: 2-17.

Dodd D., Moon Y.H., Swaminathan K., Mackie R.I., and Cann I.K. (2010) Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic bacteroidetes. *J Biol Chem* 285: 30261-30273.

Dodd D., Mackie R.I., and Cann I.K. (2011) Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. *Mol Microbiol* 79: 292-304.

Doi, R.H., Kosugi, A., Murashima, K., Tamaru, Y., and Han, S.O. (2003) Cellulosomes from mesophilic bacteria. *J Bacteriol* **185**: 5907-5914.

Doi, R.H., and Kosugi, A. (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol* **2**: 541-551.

Dow, J.M., Milligan, D.E., Jaison, L., Barber, C.E., and Daniels, M.J. (1987) A gene cluster in *Xanthomonas campestris* required for pathogenicity controls the excretion of polygalacturonate lyase and other enzymes. *Physiol Mol Plant P* **31**: 261-271.

Dow, J.M., Crossman, L., Findlay, K., He, Y.Q., Feng, J.X., and Tang, J.L. (2003) Biofilm dispersal in *Xanthomonas campestris* is controlled by cell-cell signaling and is required for full virulence to plants. *Proc Natl Acad Sci U S A* **100**: 10995-11000.

Dow, J.M., Fouhy, Y., Lucey, J.F., and Ryan, R.P. (2006) The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol Plant Microbe Interact* **19**: 1378-1384.

Dunger, G., Relling, V.M., Tondo, M.L., Barreras, M., Ielpi, L., Orellano, E.G., and Ottado, J. (2007) Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival. *Arch Microbiol* **188**: 127-135.

Durbin, R.D. (1991) Bacterial phytotoxins: Mechanism of action. *Experientia* **47**: 776-783.

Ebringerova, A., Hromadkova, Z., and Heinze, T. (2005) Hemicellulose. *Adv. Polym. Sci.* **186**: 1-67.

Eisenbeis, S., Lohmiller, S., Valdebenito, M., Leicht, S., and Braun, V. (2008) NagA-dependent uptake of *N*-acetyl-glucosamine and *N*-acetyl-chitin oligosaccharides across the outer membrane of *Caulobacter crescentus*. *J Bacteriol* **190**: 5230-5238.

Emami, K., Nagy, T., Fontes, C.M., Ferreira, L.M., and Gilbert, H.J. (2002) Evidence for temporal regulation of the two *Pseudomonas cellulosa* xylanases belonging to glycoside hydrolase family 11. *J Bacteriol* **184**: 4124-4133.

Emami, K., Topakas, E., Nagy, T., Henshaw, J., Jackson, K.A., Nelson, K.E., Mongodin, E.F., Murray, J.W., Lewis, R.J., and Gilbert, H.J. (2009) Regulation of the xylan-degrading apparatus of *Cellvibrio japonicus* by a novel two-component system. *J Biol Chem* **284**: 1086-1096.

Etchegaray, A., Silva-Stenico, M.E., Moon, D.H., and Tsai, S.M. (2004) I*n silico* analysis of nonribosomal peptide synthetases of *Xanthomonas axonopodis* pv. *citri*: identification of putative siderophore and lipopeptide biosynthetic genes. *Microbiol Res* **159**: 425-437.

Expert, D. (2005) Genetic regulation of iron in *Erwinia chrysanthemi* as pertains to bacterial virulence. *Barton LL, Abadia J (eds) Iron nutrition in plants and rhizospheric microorganisms.*: 215-227.

Fargier, E., and Manceau, C. (2007) Pathogenicity assays restrict the species *Xanthomonas campestris* into three pathovars and reveal nine races within *X. campestris* pv. *campestris*. *Plant Pathology* **56**: 805–818.

Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J* **18**: 265-276.

Fenselau, S., Balbo, I., and Bonas, U. (1992) Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins involved in secretion in bacterial pathogens of animals. *Mol Plant Microbe Interact* **5**: 390-396.

Fenselau, S., and Bonas, U. (1995) Sequence and expression analysis of the *hrpB* pathogenicity operon of *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. *Mol Plant Microbe Interact* **8**: 845-854.

Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K., and Welte, W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**: 2215-2220.

Ferguson, A.D., Chakraborty, R., Smith, B.S., Esser, L., van der Helm, D., and Deisenhofer, J. (2002) Structural basis of gating by the outer membrane transporter FecA. *Science* **295**: 1715-1719.

Ferguson, A.D., and Deisenhofer, J. (2002) TonB-dependent receptors-structural perspectives. *Biochim Biophys Acta* **1565**: 318-332.

Fleming, A.J., McQueen-Mason, S., Mandel, T., and Kuhlemeier, C. (1997) Induction of leaf primordia by the cell wall protein expansin. *Science* **276**: 1415-1418.

Fleming, A.J. (2006) The co-ordination of cell division, differentiation and morphogenesis in the shoot apical meristem: a perspective. *J Exp Bot* **57**: 25-32.

Fouhy, Y., Lucey, J.F., Ryan, R.P., and Dow, J.M. (2006) Cell-cell signaling, cyclic di-GMP turnover and regulation of virulence in *Xanthomonas campestris*. *Res Microbiol* **157**: 899-904.

Fouts, D.E., Abramovitch, R.B., Alfano, J.R., Baldo, A.M., Buell, C.R., Cartinhour, S., Chatterjee, A.K., D'Ascenzo, M., Gwinn, M.L., Lazarowitz, S.G., Lin, N.C., Martin, G.B., Rehm, A.H., Schneider, D.J., van Dijk, K., Tang, X., and Collmer, A. (2002) Genomewide identification of *Pseudomonas syringae* pv. *tomato* DC3000 promoters controlled by the HrpL alternative sigma factor. *Proc Natl Acad Sci U S A* **99**: 2275-2280.

Franza, T., Mahe, B., and Expert, D. (2005) *Erwinia chrysanthemi* requires a second iron transport route dependent of the siderophore achromobactin for extracellular growth and plant infection. *Mol Microbiol* **55**: 261-275.

Furutani, A., Tsuge, S., Ohnishi, K., Hikichi, Y., Oku, T., Tsuno, K., Inoue, Y., Ochiai, H., Kaku, H., and Kubo, Y. (2004) Evidence for HrpXo-dependent expression of type II secretory proteins in *Xanthomonas oryzae* pv. *oryzae*. *J Bacteriol* **186**: 1374-1380.

Furutani, A., Takaoka, M., Sanada, H., Noguchi, Y., Oku, T., Tsuno, K., Ochiai, H., and Tsuge, S. (2009) Identification of novel type III secretion effectors in *Xanthomonas oryzae* pv. *oryzae*. *Mol Plant Microbe Interact* **22**: 96-106.

Gaballa, A., Wang, T., Ye, R.W., and Helmann, J.D. (2002) Functional analysis of the *Bacillus subtilis* Zur regulon. *J Bacteriol* **184**: 6508-6514.

Galan, J.E., and Collmer, A. (1999) Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284**: 1322-1328.

Galan, J.E., and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**: 567-573.

Galperin, M.Y., Nikolskaya, A.N., and Koonin, E.V. (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* **203**: 11-21.

Galperin, M.Y. (2004) Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* **6**: 552-567.

Garcia-Brugger, A., Lamotte, O., Vandelle, E., Bourque, S., Lecourieux, D., Poinssot, B., Wendehenne, D., and Pugin, A. (2006) Early signaling events induced by elicitors of plant defenses. *Mol Plant Microbe Interact* **19**: 711-724.

Genin, S., Gough, C.L., Zischek, C., and Boucher, C.A. (1992) Evidence that the *hrpB* gene encodes a positive regulator of pathogenicity genes from *Pseudomonas solanacearum*. *Mol Microbiol* **6**: 3065-3076.

Gerlach, R.G., and Hensel, M. (2007) Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int J Med Microbiol* **297**: 401-415.

Ghosh, P. (2004) Process of protein transport by the type III secretion system. *Microbiol Mol Biol Rev* **68**: 771-795.

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B.S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C.,

Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C., and Slater, S. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**: 2323-2328.

Gottig, N., Garavaglia, B.S., Garofalo, C.G., Orellano, E.G., and Ottado, J. (2009) A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. *citri*, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS One* **4**: e4358.

Gough, C.L., Dow, J.M., Keen, J., Henrissat, B., and Daniels, M.J. (1990) Nucleotide sequence of the *engXCA* gene encoding the major endoglucanase of *Xanthomonas campestris* pv. *campestris*. *Gene* **89**: 53-59.

Gough, C.L., Genin, S., Zischek, C., and Boucher, C.A. (1992) *hrp* genes of *Pseudomonas solanacearum* are homologous to pathogenicity determinants of animal pathogenic bacteria and are conserved among plant pathogenic bacteria. *Mol Plant Microbe Interact* **5**: 384-389.

Grant, S.R., Fisher, E.J., Chang, J.H., Mole, B.M., and Dangl, J.L. (2006) Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annu Rev Microbiol* **60**: 425-449.

Gu, K., Yang, B., Tian, D., Wu, L., Wang, D., Sreekala, C., Yang, F., Chu, Z., Wang, G.L., White, F.F., and Yin, Z. (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122-1125.

Gudesblat, G.E., Torres, P.S., and Vojnov, A.A. (2009) *Xanthomonas campestris* overcomes *Arabidopsis* stomatal innate immunity through a DSF cell-to-cell signal-regulated virulence factor. *Plant Physiol* **149**: 1017-1027.

Guerrero, R. (2001) Bergey's manuals and the classification of prokaryotes. *Int Microbiol* **4**: 103-109.

Guidot, A., Prior, P., Schoenfeld, J., Carrere, S., Genin, S., and Boucher, C. (2007) Genomic structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. *J Bacteriol* **189**: 377-387.

Guillier, M., and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol Microbiol* **59**: 231-247.

Guillier, M., Gottesman, S., and Storz, G. (2006) Modulating the outer membrane with small RNAs. *Genes Dev* **20**: 2338-2348.

Guillier, M., and Gottesman, S. (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* **36**: 6781-6794.

Guttman, D.S., Vinatzer, B.A., Sarkar, S.F., Ranall, M.V., Kettler, G., and Greenberg, J.T. (2002) A functional screen for the type III (Hrp) secretome of the plant pathogen *Pseudomonas syringae*. *Science* **295**: 1722-1726.

Guy, E. (2009) Contribution de l'effectome et de l'effecteur de type III AvrAC de *Xanthomonas campestris* pathovar *campestris* au pouvoir pathogène chez *Arabidopsis thaliana*. *Rapport Master 2 Recherche Microbiologie*.

Hajri, A., Brin, C., Hunault, G., Lardeux, F., Lemaire, C., Manceau, C., Boureau, T., and Poussier, S. (2009) A "repertoire for repertoire" hypothesis: repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. *PLoS One* **4**: e6632.

Han, S.O., Yukawa, H., Inui, M., and Doi, R.H. (2003a) Regulation of expression of cellulosomal cellulase and hemicellulase genes in *Clostridium cellulovorans*. *J Bacteriol* **185**: 6067-6075.

Han, S.O., Yukawa, H., Inui, M., and Doi, R.H. (2003b) Transcription of *Clostridium cellulovorans* cellulosomal cellulase and hemicellulase genes. *J Bacteriol* **185**: 2520-2527.

Hashimoto, W., Yamasaki, M., Itoh, T., Momma, K., Mikami, B., and Murata, K. (2004) Super-channel in bacteria: structural and functional aspects of a novel biosystem for the import and depolymerization of macromolecules. *J Biosci Bioeng* **98**: 399-413.

Hashimoto, W., Miyake, O., Ochiai, A., and Murata, K. (2005) Molecular identification of *Sphingomonas* sp. A1 alginate lyase (A1-IV') as a member of novel polysaccharide lyase family 15 and implications in alginate lyase evolution. *J Biosci Bioeng* **99**: 48-54.

Hazen, S.P., Hawley, R.M., Davis, G.L., Henrissat, B., and Walton, J.D. (2003) Quantitative trait loci and comparative genomics of cereal cell wall composition. *Plant Physiol* **132**: 263-271.

He, S.Y., and Jin, Q. (2003) The Hrp pilus: learning from flagella. *Curr Opin Microbiol* **6**: 15-19.

He, Y.Q., Zhang, L., Jiang, B.L., Zhang, Z.C., Xu, R.Q., Tang, D.J., Qin, J., Jiang, W., Zhang, X., Liao, J., Cao, J.R., Zhang, S.S., Wei, M.L., Liang, X.X., Lu, G.T., Feng, J.X., Chen, B., Cheng, J., and Tang, J.L. (2007) Comparative and functional genomics reveals genetic diversity and determinants of host specificity among reference strains and a large collection of Chinese isolates of the phytopathogen *Xanthomonas campestris* pv. *campestris*. *Genome Biol* **8**: R218.

He, Y.W., Xu, M., Lin, K., Ng, Y.J., Wen, C.M., Wang, L.H., Liu, Z.D., Zhang, H.B., Dong, Y.H., Dow, J.M., and Zhang, L.H. (2006) Genome scale analysis of diffusible signal factor regulon in *Xanthomonas campestris* pv. *campestris*: identification of novel cell-cell communication-dependent genes and functions. *Mol Microbiol* **59**: 610-622.

He, Y.W., and Zhang, L.H. (2008) Quorum sensing and virulence regulation in *Xanthomonas campestris*. *FEMS Microbiol Rev* **32**: 842-857.

Heuer, H., Yin, Y.N., Xue, Q.Y., Smalla, K., and Guo, J.H. (2007) Repeat domain diversity of avrBs3-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Appl Environ Microbiol* **73**: 4379-4384.

Hoch, G. (2007) Cell wall hemicelluloses as mobile carbon stores in non-reproductive plant tissues. *Functional Ecology* **21**: 821-834.

Hotson, A., Chosed, R., Shu, H., Orth, K., and Mudgett, M.B. (2003) *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins in planta. *Mol Microbiol* **50**: 377-389.

Huang, G., Zhang, L., and Birch, R.G. (2001) A multifunctional polyketide-peptide synthetase essential for albicidin biosynthesis in *Xanthomonas albilineans*. *Microbiology* **147**: 631-642.

Huché, F. (2006) Etude structurale et fonctionnelle du récepteur de membrane externe HasR de *Serratia marcescens*, impliqué dans l'acquisition de l'hème via l'hémophore HasA. *Thèse*.

Huckelhoven, R. (2007) Cell wall-associated mechanisms of disease resistance and susceptibility. *Annu Rev Phytopathol* **45**: 101-127.

Hugouvieux, V., Barber, C.E., and Daniels, M.J. (1998) Entry of *Xanthomonas campestris* pv. *campestris* into hydathodes of *Arabidopsis thaliana* leaves: a system for studying early infection events in bacterial pathogenesis. *Mol Plant Microbe Interact* **11**: 537-543.

Imlay, J.A., and Linn, S. (1988) DNA damage and oxygen radical toxicity. *Science* **240**: 1302-1309.

Jansson, P.-E., Keene, L., and Lindberg, B. (1975) Structure of the extracellular polysaccharide from *Xanthomonas campestris*. *Carbohydrates Res* **45**: 275-282.

Jeffries, T. (1990) Biodegradation of lignin-carbohydrate complexes. *Biodegradation* **1**: 163-176.

Jeong, K.S., Lee S.E., Han J.W., Yang S.U., Lee B.M., Noh T.H. and Cha J-S (2008) Virulence reduction and differing regulation of virulence genes in *rpf* mutants of *Xanthomonas oryzae* pv. *oryzae*. *Plant Pathol J* **24**: 143-151.

Jha, G., Rajeshwari, R., and Sonti, R.V. (2005) Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. *Mol Plant Microbe Interact* **18**: 891-898.

Jha, G., Rajeshwari, R., and Sonti, R.V. (2007) Functional interplay between two *Xanthomonas oryzae* pv. *oryzae* secretion systems in modulating virulence on rice. *Mol Plant Microbe Interact* **20**: 31-40.

Jiang, R.P., Tang, D.J., Chen, X.L., He, Y.Q., Feng, J.X., Jiang, B.L., Lu, G.T., Lin, M., and Tang, J.L. (2010) Identification of four novel small non-coding RNAs from *Xanthomonas campestris* pathovar *campestris*. *BMC Genomics* **11**: 316.

Jin, Q., Thilmony, R., Zwiesler-Vollick, J., and He, S.Y. (2003) Type III protein secretion in *Pseudomonas syringae*. *Microbes Infect* **5**: 301-310.

Kaiser, D. (2007) Bacterial swarming: a re-examination of cell-movement patterns. *Curr Biol* **17**: R561-570.

Kamoun, S., and Kado, C.I. (1990) Phenotypic Switching Affecting Chemotaxis, Xanthan Production, and Virulence in *Xanthomonas campestris*. *Appl Environ Microbiol* **56**: 3855-3860.

Kannan, N., Taylor, S.S., Zhai, Y., Venter, J.C., and Manning, G. (2007) Structural and functional diversity of the microbial kinome. *PLoS Biol* **5**: e17.

Katzen, F., Ferreiro, D.U., Oddo, C.G., Ielmini, M.V., Becker, A., Puhler, A., and Ielpi, L. (1998) *Xanthomonas campestris* pv. *campestris gum* mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol* **180**: 1607-1617.

Kay, S., Hahn, S., Marois, E., Hause, G., and Bonas, U. (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**: 648-651.

Kay, S., and Bonas, U. (2009) How *Xanthomonas* type III effectors manipulate the host plant. *Curr Opin Microbiol* **12**: 37-43.

Kemp, B.P., Horne, J., Bryant, A. and Cooper, R.M. (2004) *Xanthomonas axonopodis* pv. *manihotis gumD* gene is essential for EPS production and pathogenicity and enhances epiphytic survival on cassava (*Maniho esculente*). *Physiol Mol Plant P*: 209-218.

Killmann, H., Videnov, G., Jung, G., Schwarz, H., and Braun, V. (1995) Identification of receptor binding sites by competitive peptide mapping: phages T1, T5, and phi 80 and colicin M bind to the gating loop of FhuA. *J Bacteriol* **177**: 694-698.

Kim, J.G., Taylor, K.W., Hotson, A., Keegan, M., Schmelz, E.A., and Mudgett, M.B. (2008) XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in *Xanthomonas*-infected tomato leaves. *Plant Cell* **20**: 1915-1929.

Kim, M., Fanucci, G.E., and Cafiso, D.S. (2007) Substrate-dependent transmembrane signaling in TonB-dependent transporters is not conserved. *Proc Natl Acad Sci U S A* **104**: 11975-11980.

Kim, S.Y., Kim, J.G., Lee, B.M., and Cho, J.Y. (2009) Mutational analysis of the *gum* gene cluster required for xanthan biosynthesis in *Xanthomonas oryzae* pv *oryzae*. *Biotechnol Lett* **31**: 265-270.

Kitphati, W., Ngok-Ngam, P., Suwanmaneerat, S., Sukchawalit, R., and Mongkolsuk, S. (2007) *Agrobacterium tumefaciens* fur has important physiological roles in iron and

manganese homeostasis, the oxidative stress response, and full virulence. *Appl Environ Microbiol* **73**: 4760-4768.

Koebnik, R. (1993) Structural organization of TonB-dependent receptors. *Trends Microbiol* **1**: 201.

Koebnik, R. (2005) TonB-dependent trans-envelope signalling: the exception or the rule? *Trends Microbiol* **13**: 343-347.

Koebnik, R., Kruger, A., Thieme, F., Urban, A., and Bonas, U. (2006) Specific binding of the *Xanthomonas campestris* pv. *vesicatoria* AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J Bacteriol* **188**: 7652-7660.

Kolenova, K., Vrsanska, M., and Biely, P. (2006) Mode of action of endo-beta-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides. *J Biotechnol* **121**: 338-345.

Kotoujansky, A. (1987) Molecular genetics og pathogenesis by soft-rot *erwinias*. *Ann. Rev. Phytopath.* **25**.

Kulkarni, N., Shendye, A., and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. *FEMS Microbiol Rev* **23**: 411-456.

Lamed, R., Setter, E., and Bayer, E.A. (1983) Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J Bacteriol* **156**: 828-836.

Landt, S.G., Lesley, J.A., Britos, L., and Shapiro, L. (2010) CrfA, a small noncoding RNA regulator of adaptation to carbon starvation in *Caulobacter crescentus*. *J Bacteriol* **192**: 4763-4775.

Lee, B.M., Park, Y.J., Park, D.S., Kang, H.W., Kim, J.G., Song, E.S., Park, I.C., Yoon, U.H., Hahn, J.H., Koo, B.S., Lee, G.B., Kim, H., Park, H.S., Yoon, K.O., Kim, J.H., Jung, C.H., Koh, N.H., Seo, J.S., and Go, S.J. (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res* **33**: 577-586.

Lerouge, P., O'Neill, M.A., Darvill, A.G., and Albersheim, P. (1993) The purification of commercially available endo-alpha-L-arabinanases and alpha-L-arabinosidase for use in the structural analysis of pectic polysaccharides. *Carbohydr Res* **243**: 373-378.

Leyns, F., De Cleene, M., J, S., and J, D.L. (1984) The host range of the genus *Xanthomonas*. *Bot Rev* **50**: 308-355.

Li, S.J., and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. *Nature* **398**: 246-251.

Liang, H., Zhao, Y.T., Zhang, J.Q., Wang, X.J., Fang, R.X., and Jia, Y.T. (2011) Identification and functional characterization of small non-coding RNAs in *Xanthomonas oryzae* pathovar *oryzae*. *BMC Genomics* **12**: 87.

Lin, L.L., and Thomson, J.A. (1991) Cloning, sequencing and expression of a gene encoding a 73 kDa xylanase enzyme from the rumen anaerobe *Butyrivibrio fibrisolvens* H17c. *Mol Gen Genet* **228**: 55-61.

Lindgren, P.B., Peet, R.C., and Panopoulos, N.J. (1986) Gene cluster of *Pseudomonas syringae* pv. "*phaseolicola*" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. *J Bacteriol* **168**: 512-522.

Lindgren, P.B. (1997) The role of *hrp* genes during plant-bacterial interactions. *Annu Rev Phytopathol* **35**: 129-152.

Livny, J., Teonadi, H., Livny, M., and Waldor, M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One* **3**: e3197.

Lu, G.T., Ma, Z.F., Hu, J.R., Tang, D.J., He, Y.Q., Feng, J.X., and Tang, J.L. (2007) A novel locus involved in extracellular polysaccharide production and virulence of *Xanthomonas campestris* pathovar *campestris*. *Microbiology* **153**: 737-746.

Lu, H., Patil, P., Van Sluys, M.A., White, F.F., Ryan, R.P., Dow, J.M., Rabinowicz, P., Salzberg, S.L., Leach, J.E., Sonti, R., Brendel, V., and Bogdanove, A.J. (2008)

Acquisition and evolution of plant pathogenesis-associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. *PLoS One* **3**: e3828.

Lummerzheim, M., de Oliviveira, D., Castresana, C., Miguens, F., Louzada, E., Roby, D., Van Montagu, M., and Timmerman, B. (1993) Identification of compatible and incompatible interactions between *Arabidopsis thaliana* and *Xanthomonas campestris* pv. *campestris* and characterization of the hypersensitive response. *Mol Plant Microbe Interact* **6**: 532-544.

Lundrigan, M.D., and Kadner, R.J. (1986) Nucleotide sequence of the gene for the ferrienterochelin receptor FepA in *Escherichia coli*. Homology among outer membrane receptors that interact with TonB. *J Biol Chem* **261**: 10797-10801.

Lundrigan, M.D., Koster, W., and Kadner, R.J. (1991) Transcribed sequences of the *Escherichia coli btuB* gene control its expression and regulation by vitamin B12. *Proc Natl Acad Sci U S A* **88**: 1479-1483.

Lynd, L.R., Weimer, P.J., van Zyl, W.H., and Pretorius, I.S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* **66**: 506-577, table of contents.

Ma, W., Cui, Y., Liu, Y., Dumenyo, C.K., Mukherjee, A., and Chatterjee, A.K. (2001) Molecular characterization of global regulatory RNA species that control pathogenicity factors in *Erwinia amylovora* and *Erwinia herbicola* pv. *gypsophilae*. *J Bacteriol* **183**: 1870-1880.

Macnab, R.M. (1999) The bacterial flagellum: reversible rotary propellor and type III export apparatus. *J Bacteriol* **181**: 7149-7153.

Madkour, M., and Mayer, F. (2003) Structural organization of the intact bacterial cellulosome as revealed by electron microscopy. *Cell Biol Int* **27**: 831-836.

Mandal, M., and Breaker, R.R. (2004) Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat Struct Mol Biol* **11**: 29-35.

Marenda, M., Brito, B., Callard, D., Genin, S., Barberis, P., Boucher, C., and Arlat, M. (1998) PrhA controls a novel regulatory pathway required for the specific induction of *Ralstonia solanacearum hrp* genes in the presence of plant cells. *Mol Microbiol* **27**: 437-453.

Marie, C., Broughton, W.J., and Deakin, W.J. (2001) *Rhizobium* type III secretion systems: legume charmers or alarmers? *Curr Opin Plant Biol* **4**: 336-342.

Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **4**: 447-457.

Mazeau, K., and Perez, S. (1998) The preferred conformations of the four oligomeric fragments of Rhamnogalacturonan II. *Carbohydr Res* **311**: 203-217.

McBride, M.J., Braun, T.F., and Brust, J.L. (2003) *Flavobacterium johnsoniae* GldH is a lipoprotein that is required for gliding motility and chitin utilization. *J Bacteriol* **185**: 6648-6657.

McBride, M.J., Gary, X., Martens, E.C., Lapidus, A., Henrissat, B., Rhodes R.G., Goltsman, E., Wang, W., Xu, J., Hunnicutt, D.W., Staroscik, A.M., Hoover, T.R., Cheng, Y.Q., Stein, J.L. (2009) Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol* **75**: 6864-6875

McDougall, G.J., and Fry, S.C. (1990) Xyloglucan oligosaccharides promote growth and activate cellulase: evidence for a role of cellulase in cell expansion. *Plant Physiol* **93**: 1042-1048.

McNeil, M., Darvill, A.G., Fry, S.C., and Albersheim, P. (1984) Structure and function of the primary cell walls of plants. *Annu Rev Biochem* **53**: 625-663.

Mechaly, A., Yaron, S., Lamed, R., Fierobe, H.P., Belaich, A., Belaich, J.P., Shoham, Y., and Bayer, E.A. (2000) Cohesin-dockerin recognition in cellulosome assembly: experiment versus hypothesis. *Proteins* **39**: 170-177.

Melotto, M., Underwood, W., and He, S.Y. (2008) Role of stomata in plant innate immunity and foliar bacterial diseases. *Annu Rev Phytopathol* **46**: 101-122.

Merighi, M., Majerczak, D.R., Stover, E.H., and Coplin, D.L. (2003) The HrpX/HrpY two-component system activates *hrpS* expression, the first step in the regulatory cascade controlling the Hrp regulon in *Pantoea stewartii* subsp. *stewartii*. *Mol Plant Microbe Interact* **16**: 238-248.

Meyer, D., Lauber, E., Roby, D., Arlat, M., and Kroj, T. (2005) Optimization of pathogenicity assays tu study *Arabidopsdis thaliana-Xanthomonas campestris* pv. *campestris* pathosystem. *Molecular Plant Pathology* **6**: 327-333.

Meyer, D., Cunnac, S., Gueneron, M., Declercq, C., Van Gijsegem, F., Lauber, E., Boucher, C., and Arlat, M. (2006) PopF1 and PopF2, two proteins secreted by the type III protein secretion system of *Ralstonia solanacearum*, are translocators belonging to the HrpF/NopX family. *J Bacteriol* **188**: 4903-4917.

Mingardon, F., Perret, S., Belaich, A., Tardif, C., Belaich, J.P., and Fierobe, H.P. (2005) Heterologous production, assembly, and secretion of a minicellulosome by *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol* **71**: 1215-1222.

Mirus, O., Strauss, S., Nicolaisen, K., von Haeseler, A., and Schleiff, E. (2009) TonB-dependent transporters and their occurrence in cyanobacteria. *BMC Biol* **7**: 68.

Miyazaki, K., Miyamoto, H., Mercer, D.K., Hirase, T., Martin, J.C., Kojima, Y., and Flint, H.J. (2003) Involvement of the multidomain regulatory protein XynR in positive control of xylanase gene expression in the ruminal anaerobe *Prevotella bryantii* B(1)4. *J Bacteriol* **185**: 2219-2226.

Miyazaki, K., Hirase, T., Kojima, Y., and Flint, H.J. (2005) Medium- to large-sized xylo-oligosaccharides are responsible for xylanase induction in *Prevotella bryantii* B14. *Microbiology* **151**: 4121-4125.

Moller, I., Marcus, S.E., Haeger, A., Verhertbruggen, Y., Verhoef, R., Schols, H., Ulvskov, P., Mikkelsen, J.D., Knox, J.P., and Willats, W. (2008) High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj J* **25**: 37-48.

Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. (2002) Genetic control by a metabolite binding mRNA. *Chem Biol* **9**: 1043.

Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., and Braun, V. (2005) ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. *J Bacteriol* **187**: 8300-8311.

Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**: 593-656.

Noël, L., Thieme, F., Nennstiel, D., and Bonas, U. (2001) cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. *Mol Microbiol* **41**: 1271-1281.

Ochiai, A., Itoh, T., Kawamata, A., Hashimoto, W., and Murata, K. (2007) Plant cell wall degradation by saprophytic *Bacillus subtilis* strains: gene clusters responsible for rhamnogalacturonan depolymerization. *Appl Environ Microbiol* **73**: 3803-3813.

Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A., and Kaku, H. (2005) Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *Jpn Agr Res Q*: 275-287.

O'Neill, M.A., Darvill, A.G., Albersheim, P., and Chou, K.J. (1990) Structural analysis of an acidic polysaccharide secreted by *Xanthobacter* sp. (ATCC 53272). *Carbohydr Res* **206**: 289-296.

O'Neill, M.A., Ishii, T., Albersheim, P., and Darvill, A.G. (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu Rev Plant Biol* **55**: 109-139.

Onsando, J. (1992) Black rot of crucifers. *Diseases of Vegetables and Oil Seed Crops.* **2**: 243-252.

Papenfort, K., Bouvier, M., Mika, F., Sharma, C.M., and Vogel, J. (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci U S A* **107**: 20435-20440.

Parker, J.E., Barber, C.E., Fan, M.J., and Daniels, M.J. (1993) Interaction of *Xanthomonas campestris* with *Arabidopsis thaliana*: characterization of a gene from *X. c.* pv. *raphani* that confers avirulence to most *A. thaliana* accessions. *Mol Plant Microbe Interact* **6**: 216-224.

Patrick, J. (1990) Sieve element unloading: cellular pathway, mechanism and control. *Physiologia Plantarum* **78**: 298-308.

Pena, M.J., Zhong, R., Zhou, G.K., Richardson, E.A., O'Neill, M.A., Darvill, A.G., York, W.S., and Ye, Z.H. (2007) Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* **19**: 549-563.

Perret, S., Belaich, A., Fierobe, H.P., Belaich, J.P., and Tardif, C. (2004a) Towards designer cellulosomes in *Clostridia*: mannanase enrichment of the cellulosomes produced by *Clostridium cellulolyticum*. *J Bacteriol* **186**: 6544-6552.

Perret, S., Casalot, L., Fierobe, H.P., Tardif, C., Sabathe, F., Belaich, J.P., and Belaich, A. (2004b) Production of heterologous and chimeric scaffoldins by *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* **186**: 253-257.

Petnicki-Ocwieja, T., Schneider, D.J., Tam, V.C., Chancey, S.T., Shan, L., Jamir, Y., Schechter, L.M., Janes, M.D., Buell, C.R., Tang, X., Collmer, A., and Alfano, J.R. (2002) Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. *tomato* DC3000. *Proc Natl Acad Sci U S A* **99**: 7652-7657.

Pieretti, I., Royer, M., Barbe, V., Carrere, S., Koebnik, R., Cociancich, S., Couloux, A., Darrasse, A., Gouzy, J., Jacques, M.A., Lauber, E., Manceau, C., Mangenot, S., Poussier, S., Segurens, B., Szurek, B., Verdier, V., Arlat, M., and Rott, P. (2009) The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited *Xanthomonadaceae*. *BMC Genomics* **10**: 616.

Plener, L., Manfredi, P., Valls, M., and Genin, S. (2010) PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. *J Bacteriol* **192**: 1011-1019.

Poplawsky, A.R., and Chun, W. (1997) *pigB* determines a diffusible factor needed for extracellular polysaccharide slime and xanthomonadin production in *Xanthomonas campestris* pv. *campestris*. *J Bacteriol* **179**: 439-444.

Popper, Z.A., and Fry, S.C. (2003) Primary cell wall composition of bryophytes and charophytes. *Ann Bot* **91**: 1-12.

Porteneuve, C. (2002) Lutte contre les maladies folaires. In *station SECL*.

Postle, K. (1993) TonB protein and energy transduction between membranes. *J Bioenerg Biomembr* **25**: 591-601.

Poueymiro, M., and Genin, S. (2009) Secreted proteins from *Ralstonia solanacearum*: a hundred tricks to kill a plant. *Curr Opin Microbiol* **12**: 44-52.

Prade, R.A., Zhan, D., Ayoubi, P., and Mort, A.J. (1999) Pectins, pectinases and plant-microbe interactions. *Biotechnol Genet Eng Rev* **16**: 361-391.

Qian, W., Jia, Y., Ren, S.X., He, Y.Q., Feng, J.X., Lu, L.F., Sun, Q., Ying, G., Tang, D.J., Tang, H., Wu, W., Hao, P., Wang, L., Jiang, B.L., Zeng, S., Gu, W.Y., Lu, G., Rong, L., Tian, Y., Yao, Z., Fu, G., Chen, B., Fang, R., Qiang, B., Chen, Z., Zhao, G.P., Tang, J.L., and He, C. (2005) Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. *Genome Res* **15**: 757-767.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick, W.J., Jr., Hallett, J.P., Leak, D.J., Liotta, C.L., Mielenz, J.R., Murphy, R., Templer, R., and Tschaplinski, T. (2006) The path forward for biofuels and biomaterials. *Science* **311**: 484-489.

Rajagopal, L., Sundari, C.S., Balasubramanian, D., and Sonti, R.V. (1997) The bacterial pigment xanthomonadin offers protection against photodamage. *FEBS Lett* **415**: 125-128.

Rajeshwari, R., Jha, G., and Sonti, R.V. (2005) Role of an in planta-expressed xylanase of *Xanthomonas oryzae* pv. *oryzae* in promoting virulence on rice. *Mol Plant Microbe Interact* **18**: 830-837.

Rakotoarivonina, H., Jubelin, G., Hebraud, M., Gaillard-Martini, B., Forano, E., and Mosoni, P. (2002) Adhesion to cellulose of the gram-positive bacterium *Ruminococcus albus* involves type IV pili. *Microbiology* **148**: 1871-1880.

Ratledge, C., and Dover, L.G. (2000) Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol* **54**: 881-941.

Ray, S.K., Rajeshwari, R., and Sonti, R.V. (2000) Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in general secretory pathway are virulence deficient and unable to secrete xylanase. *Mol Plant Microbe Interact* **13**: 394-401.

Ray, S.K., Rajeshwari, R., Sharma, Y., and Sonti, R.V. (2002) A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. *oryzae* exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. *Mol Microbiol* **46**: 637-647.

Raymond, K.N., Dertz, E.A., and Kim, S.S. (2003) Enterobactin: an archetype for microbial iron transport. *Proc Natl Acad Sci U S A* **100**: 3584-3588.

Ridley, B.L., O'Neill, M.A., and Mohnen, D. (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **57**: 929-967.

Rigano, L.A., Siciliano, F., Enrique, R., Sendin, L., Filippone, P., Torres, P.S., Questa, J., Dow, J.M., Castagnaro, A.P., Vojnov, A.A., and Marano, M.R. (2007) Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. *citri*. *Mol Plant Microbe Interact* **20**: 1222-1230.

Rincon, M.T., Ding, S.Y., McCrae, S.I., Martin, J.C., Aurilia, V., Lamed, R., Shoham, Y., Bayer, E.A., and Flint, H.J. (2003) Novel organization and divergent dockerin specificities in the cellulosome system of *Ruminococcus flavefaciens*. *J Bacteriol* **185**: 703-713.

Rincon, M.T., Martin, J.C., Aurilia, V., McCrae, S.I., Rucklidge, G.J., Reid, M.D., Bayer, E.A., Lamed, R., and Flint, H.J. (2004) ScaC, an adaptor protein carrying a novel cohesin that expands the dockerin-binding repertoire of the *Ruminococcus flavefaciens* 17 cellulosome. *J Bacteriol* **186**: 2576-2585.

Robinson, D.G. (1991) What Is a Plant Cell? The Last Word. *Plant Cell* **3**: 1145-1146.

Roine, E., Saarinen, J., Kalkkinen, N., and Romantschuk, M. (1997) Purified HrpA of *Pseudomonas syringae* pv. *tomato* DC3000 reassembles into pili. *FEBS Lett* **417**: 168-172.

Romby, P., Vandenesch, F., and Wagner, E.G. (2006) The role of RNAs in the regulation of virulence-gene expression. *Curr Opin Microbiol* **9**: 229-236.

Romer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U., and Lahaye, T. (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**: 645-648.

Romer, P., Recht, S., and Lahaye, T. (2009) A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proc Natl Acad Sci U S A* **106**: 20526-20531.

Rosqvist, R., Hakansson, S., Forsberg, A., and Wolf-Watz, H. (1995) Functional conservation of the secretion and translocation machinery for virulence proteins of yersiniae, *salmonellae* and *shigellae*. *Embo J* **14**: 4187-4195.

Rostand, K.S., and Esko, J.D. (1997) Microbial adherence to and invasion through proteoglycans. *Infect Immun* **65**: 1-8.

Ryan, R.P., Fouhy, Y., Lucey, J.F., Crossman, L.C., Spiro, S., He, Y.W., Zhang, L.H., Heeb, S., Camara, M., Williams, P., and Dow, J.M. (2006) Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc Natl Acad Sci U S A* **103**: 6712-6717.

Ryan, R.P., Fouhy, Y., Lucey, J.F., Jiang, B.L., He, Y.Q., Feng, J.X., Tang, J.L., and Dow, J.M. (2007) Cyclic di-GMP signalling in the virulence and environmental adaptation of *Xanthomonas campestris*. *Mol Microbiol* **63**: 429-442.

Ryan, R.P., and Dow, J.M. (2008) Diffusible signals and interspecies communication in bacteria. *Microbiology* **154**: 1845-1858.

Ryan, R.P., McCarthy, Y., Andrade, M., Farah, C.S., Armitage, J.P., and Dow, J.M. (2010) Cell-cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*. *Proc Natl Acad Sci U S A* **107**: 5989-5994.

Ryan, R.P., and Dow, J.M. (2011) Communication with a growing family: diffusible signal factor (DSF) signaling in bacteria. *Trends Microbiol* **19**: 145-152.

Ryan, R.P., Vorholter, F.J., Potnis, N., Jones, J.B., Van Sluys, M.A., Bogdanove, A.J., and Dow, J.M. (2011) Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. *Nat Rev Microbiol* **9**: 344-355.

Saha, B.C. (2003) Hemicellulose bioconversion. *J Ind Microbiol Biotechnol* **30**: 279-291.

Saier, M.H., Jr. (2004) Evolution of bacterial type III protein secretion systems. *Trends Microbiol* **12**: 113-115.

Salyers, A.A., Palmer, J.K., and Wilkins, T.D. (1977a) Laminarinase (beta-glucanase) activity in *Bacteroides* from the human colon. *Appl Environ Microbiol* **33**: 1118-1124.

Salyers, A.A., Vercellotti, J.R., West, S.E., and Wilkins, T.D. (1977b) Fermentation of mucin and plant polysaccharides by strains of *Bacteroides* from the human colon. *Appl Environ Microbiol* **33**: 319-322.

Salyers, A.A., West, S.E., Vercellotti, J.R., and Wilkins, T.D. (1977c) Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* **34**: 529-533.

Salzberg, S.L., Sommer, D.D., Schatz, M.C., Phillippy, A.M., Rabinowicz, P.D., Tsuge, S., Furutani, A., Ochiai, H., Delcher, A.L., Kelley, D., Madupu, R., Puiu, D., Radune, D., Shumway, M., Trapnell, C., Aparna, G., Jha, G., Pandey, A., Patil, P.B., Ishihara, H., Meyer, D.F., Szurek, B., Verdier, V., Koebnik, R., Dow, J.M., Ryan, R.P., Hirata, H., Tsuyumu, S., Won Lee, S., Seo, Y.S., Sriariyanum, M., Ronald, P.C., Sonti, R.V., Van

Sluys, M.A., Leach, J.E., White, F.F., and Bogdanove, A.J. (2008) Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. *BMC Genomics* **9**: 204.

Samudrala, R., Heffron, F., and McDermott, J.E. (2009) Accurate prediction of secreted substrates and identification of a conserved putative secretion signal for type III secretion systems. *PLoS Pathog* **5**: e1000375.

Sandkvist, M. (2001a) Type II secretion and pathogenesis. *Infect Immun* **69**: 3523-3535.

Sandkvist, M. (2001b) Biology of type II secretion. *Mol Microbiol* **40**: 271-283.

Sarkar, S.F., Gordon, J.S., Martin, G.B., and Guttman, D.S. (2006) Comparative genomics of host-specific virulence in *Pseudomonas syringae*. *Genetics* **174**: 1041-1056.

Schaad, N.W., Gabrielson, R.L., and Mulanax, M.W. (1980) Hot Acidified Cupric Acetate Soaks for Eradication of *Xanthomonas campestris* from Crucifer Seeds. *Appl Environ Microbiol* **39**: 803-807.

Schalk, I.J., Yue, W.W., and Buchanan, S.K. (2004) Recognition of iron-free siderophores by TonB-dependent iron transporters. *Mol Microbiol* **54**: 14-22.

Schauer, K., Gouget, B., Carriere, M., Labigne, A., and de Reuse, H. (2007) Novel nickel transport mechanism across the bacterial outer membrane energized by the TonB/ExbB/ExbD machinery. *Mol Microbiol* **63**: 1054-1068.

Schauer, K., Rodionov, D.A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends Biochem Sci* **33**: 330-338.

Scheller, H.V., and Ulvskov, P. (2010) Hemicelluloses. *Annu Rev Plant Biol* **61**: 263-289.

Scholze, H., and Boch, J. (2011) TAL effectors are remote controls for gene activation. *Curr Opin Microbiol* **14**: 47-53.

Schroter, K., Flaschel, E., Puhler, A., and Becker, A. (2001) *Xanthomonas campestris* pv. *campestris* secretes the endoglucanases ENGXCA and ENGXCB: construction of an endoglucanase-deficient mutant for industrial xanthan production. *Appl Microbiol Biotechnol* **55**: 727-733.

Schryvers, A.B., and Stojiljkovic, I. (1999) Iron acquisition systems in the pathogenic *Neisseria*. *Mol Microbiol* **32**: 1117-1123.

Schultz, T., and Gabrielson, R. (1986) *Xanthomonas campestris* pv. *campestris* in Western
Washington crucifer seed fields: occurence and survival. *Phytopathology*: 1306-1309.

Schwarz, W.H. (2001) The cellulosome and cellulose degradation by anaerobic bacteria. *Appl Microbiol Biotechnol* **56**: 634-649.

Shipman, J.A., Berleman, J.E., and Salyers, A.A. (2000) Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J Bacteriol* **182**: 5365-5372.

Simpson, A.J., and Reinach, F.C., and Arruda, P., and Abreu, F.A., and Acencio, M., and Alvarenga, R., and Alves, L.M., and Araya, J.E., and Baia, G.S., and Baptista, C.S., and Barros, M.H., and Bonaccorsi, E.D., and Bordin, S., and Bove, J.M., and Briones, M.R., and Bueno, M.R., and Camargo, A.A., and Camargo, L.E., and Carraro, D.M., and Carrer, H., and Colauto, N.B., and Colombo, C., and Costa, F.F., and Costa, M.C., and Costa-Neto, C.M., and Coutinho, L.L., and Cristofani, M., and Dias-Neto, E., and Docena, C., and El-Dorry, H., and Facincani, A.P., and Ferreira, A.J., and Ferreira, V.C., and Ferro, J.A., and Fraga, J.S., and Franca, S.C., and Franco, M.C., and Frohme, M., and Furlan, L.R., and Garnier, M., and Goldman, G.H., and Goldman, M.H., and Gomes, S.L., and Gruber, A., and Ho, P.L., and Hoheisel, J.D., and Junqueira, M.L., and Kemper, E.L., and Kitajima, J.P., and Krieger, J.E., and Kuramae, E.E., and Laigret, F., and Lambais, M.R., and Leite, L.C., and Lemos, E.G., and Lemos, M.V., and Lopes, S.A., and Lopes, C.R., and Machado, J.A., and Machado, M.A., and Madeira, A.M., and Madeira, H.M., and Marino, C.L., and

Marques, M.V., and Martins, E.A., and Martins, E.M., and Matsukuma, A.Y., and Menck, C.F., and Miracca, E.C., and Miyaki, C.Y., and Monteriro-Vitorello, C.B., and Moon, D.H., and Nagai, M.A., and Nascimento, A.L., and Netto, L.E., and Nhani, A., Jr., and Nobrega, F.G., and Nunes, L.R., and Oliveira, M.A., and de Oliveira, M.C., and de Oliveira, R.C., and Palmieri, D.A., and Paris, A., and Peixoto, B.R., and Pereira, G.A., and Pereira, H.A., Jr., and Pesquero, J.B., and Quaggio, R.B., and Roberto, P.G., and Rodrigues, V., and de, M.R.A.J., and de Rosa, V.E., Jr., and de Sa, R.G., and Santelli, R.V., and Sawasaki, H.E., and da Silva, A.C., and da Silva, A.M., and da Silva, F.R., and da Silva, W.A., Jr., and da Silveira, J.F., and Silvestri, M.L., and Siqueira, W.J., and de Souza, A.A., and de Souza, A.P., and Terenzi, M.F., and Truffi, D., and Tsai, S.M., and Tsuhako, M.H., and Vallada, H., and Van Sluys, M.A., and Verjovski-Almeida, S., and Vettore, A.L., and Zago, M.A., and Zatz, M., and Meidanis, J., and Setubal, J.C. (2000) The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *Nature* **406**: 151-159.

Simpson, R., and Johnson, L.J. (1990) *Arabidopsis thaliana* as a host for *Xanthomonas campestris* pv. *campestris*. *Mol Plant Microbe Interact* **3**: 233-237.

Slater, H., Alvarez-Morales, A., Barber, C.E., Daniels, M.J., and Dow, J.M. (2000) A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol Microbiol* **38**: 986-1003.

Somerville, C. (2006) Cellulose synthesis in higher plants. *Annu Rev Cell Dev Biol* **22**: 53-78.

Somerville, C. (2007) Biofuels. *Curr Biol* **17**: R115-119.

Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010) Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141**: 1241-1252.

Spence, C., Wells, W.G., and Smith, C.J. (2006) Characterization of the primary starch utilization operon in the obligate anaerobe *Bacteroides fragilis*: Regulation by carbon source and oxygen. *J Bacteriol* **188**: 4663-4672.

Starnes, R.L., Liu, C.L., and Marrone, P.G. (1993) History, use and future of microbial insecticides. *American entomologist* **39**: 83-91(89).

Starr, M.P. (1981) The genus *Xanthomonas*. *In The Prokaryotes*: 742-763.

Stevenson, T.T., McNeil, M., Darvill, A.G., and Albersheim, P. (1986) Structure of Plant Cell Walls : XVIII. An Analysis of the Extracellular Polysaccharides of Suspension-Cultured Sycamore Cells. *Plant Physiol* **80**: 1012-1019.

Stoodley, P., Sauer, K., Davies, D.G., and Costerton, J.W. (2002) Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**: 187-209.

Studholme, D.J., Kemen, E., MacLean, D., Schornack, S., Aritua, V., Thwaites, R., Grant, M., Smith, J., and Jones, J.D. (2010) Genome-wide sequencing data reveals virulence factors implicated in banana *Xanthomonas* wilt. *FEMS Microbiol Lett* **310**: 182-192.

Sugio, A., Yang, B., and White, F.F. (2005) Characterization of the hrpF pathogenicity peninsula of *Xanthomonas oryzae* pv. *oryzae*. *Mol Plant Microbe Interact* **18**: 546-554.

Suji, J., and Somerville, S. (1988) *Xanthomonas campestris* pv. *campestris* induces chlorosis in *Arabidopsis thaliana*. *Arabidopsis Information Service* **26**: 1-8.

Sun, Q.H., Hu, J., Huang, G.X., Ge, C., Fang, R.X., He, C.Z. (2005) Type- II secretion pathway structural gene *xpsE*, xylanase- and cellulase secretion and virulence in *Xanthomonas oryzae* pv. *oryzae*. *Plant Pathol J* **54**: 15-21.

Sunna, A., and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. *Crit Rev Biotechnol* **17**: 39-67.

Swings, J.G., Vauteriun L., Kersters K. (1993) The bacterium *Xanthomonas*. In *Xanthomonas*, pp. 121-156.

Szczesny, R., Jordan, M., Schramm, C., Schulz, S., Cogez, V., Bonas, U., and Büttner, D. (2010) Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv *vesicatoria*. *New Phytol* **187**: 983-1002.

Taiz, L. (1984) Plant cell expansion: regulation of cell wall mechanical properties. *Annu Rev Plant Physiol* **35**: 585-657.

Tang, J.L., Liu, Y.N., Barber, C.E., Dow, J.M., Wootton, J.C., and Daniels, M.J. (1991) Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. *Mol Gen Genet* **226**: 409-417.

Tang, X., Xiao, Y., and Zhou, J.M. (2006) Regulation of the type III secretion system in phytopathogenic bacteria. *Mol Plant Microbe Interact* **19**: 1159-1166.

Tani, A., Charoenpanich, J., Mori, T., Takeichi, M., Kimbara, K., and Kawai, F. (2007) Structure and conservation of a polyethylene glycol-degradative operon in *sphingomonads*. *Microbiology* **153**: 338-346.

Tao, F., Swarup, S., and Zhang, L.H. (2010) Quorum sensing modulation of a putative glycosyltransferase gene cluster essential for *Xanthomonas campestris* biofilm formation. *Environ Microbiol*.

Tedman-Jones, J.D., Lei, R., Jay, F., Fabro, G., Li, X., Reiter, W.D., Brearley, C., and Jones, J.D. (2008) Characterization of *Arabidopsis mur3* mutations that result in constitutive activation of defence in petioles, but not leaves. *Plant J* **56**: 691-703.

Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A., Kay, S., Kirchner, O., Lanz, C., Linke, B., McHardy, A.C., Meyer, F., Mittenhuber, G., Nies, D.H., Niesbach-Klosgen, U., Patschkowski, T., Ruckert, C., Rupp, O., Schneiker, S., Schuster, S.C., Vorholter, F.J., Weber, E., Puhler, A., Bonas, U., Bartels, D., and Kaiser, O. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. *J Bacteriol* **187**: 7254-7266.

Thowthampitak, J., Shaffer, B.T., Prathuangwong, S., and Loper, J.E. (2008) Role of *rpfF* in virulence and exoenzyme production of *Xanthomonas axonopodis* pv. *glycines*, the causal agent of bacterial pustule of soybean. *Phytopathology* **98**: 1252-1260.

Tierens, K.F., Thomma, B.P., Brouwer, M., Schmidt, J., Kistner, K., Porzel, A., Mauch-Mani, B., Cammue, B.P., and Broekaert, W.F. (2001) Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of *Arabidopsis* to microbial pathogens. *Plant Physiol* **125**: 1688-1699.

Tokatlidis, K., Salamitou, S., Beguin, P., Dhurjati, P., and Aubert, J.P. (1991) Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. *FEBS Lett* **291**: 185-188.

Torres, P.S., Malamud, F., Rigano, L.A., Russo, D.M., Marano, M.R., Castagnaro, A.P., Zorreguieta, A., Bouarab, K., Dow, J.M., and Vojnov, A.A. (2007) Controlled synthesis of the DSF cell-cell signal is required for biofilm formation and virulence in *Xanthomonas campestris*. *Environ Microbiol* **9**: 2101-2109.

Trias, J. (1988) Specificity of the glucose channel formed by protein D1 of *Pseudomonas aeruginosa*. *Biochim Biophys Acta* **938**: 493-496.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., Bork, P., Hugenholtz, P., and Rubin, E.M. (2005) Comparative metagenomics of microbial communities. *Science* **308**: 554-557.

Troisfontaines, P., and Cornelis, G.R. (2005) Type III secretion: more systems than you think. *Physiology (Bethesda)* **20**: 326-339.

Tsuge, S., Terashima, S., Furutani, A., Ochiai, H., Oku, T., Tsuno, K., Kaku, H., and Kubo, Y. (2005) Effects on promoter activity of base substitutions in the cis-acting regulatory element of HrpXo regulons in *Xanthomonas oryzae* pv. *oryzae*. *J Bacteriol* **187**: 2308-2314.

Tsuge, S., Nakayama, T., Terashima, S., Ochiai, H., Furutani, A., Oku, T., Tsuno, K., Kubo, Y., and Kaku, H. (2006) Gene involved in transcriptional activation of the hrp regulatory gene *hrpG* in *Xanthomonas oryzae* pv. *oryzae*. *J Bacteriol* **188**: 4158-4162.

Valentine, P.J., Arnold, P., and Salyers, A.A. (1992) Cloning and partial characterization of two chromosomal loci from *Bacteroides ovatus* that contain genes essential for growth on guar gum. *Appl Environ Microbiol* **58**: 1541-1548.

Valentine, P.J., and Salyers, A.A. (1992a) Use of inducible disaccharidases to assess the importance of different carbohydrate sources for *Bacteroides ovatus* growing in the intestinal tracts of germfree mice. *Appl Environ Microbiol* **58**: 2698-2700.

Valentine, P.J., and Salyers, A.A. (1992b) Analysis of proteins associated with growth of *Bacteroides ovatus* on the branched galactomannan guar gum. *Appl Environ Microbiol* **58**: 1534-1540.

Valls, M., Genin, S., and Boucher, C. (2006) Integrated regulation of the type III secretion system and other virulence determinants in *Ralstonia solanacearum*. *PLoS Pathog* **2**: e82.

van Doorn, J., Boonekamp, P.M., and Oudega, B. (1994) Partial characterization of fimbriae of *Xanthomonas campestris* pv. *hyacinthi*. *Mol Plant Microbe Interact* **7**: 334-344.

Van Gijsegem, F., Genin, S., and Boucher, C. (1993) Conservation of secretion pathways for pathogenicity determinants of plant and animal bacteria. *Trends Microbiol* **1**: 175-180.

Van Gijsegem, F., Vasse, J., Camus, J.C., Marenda, M., and Boucher, C. (2000) *Ralstonia solanacearum* produces hrp-dependent pili that are required for PopA secretion but not for attachment of bacteria to plant cells. *Mol Microbiol* **36**: 249-260.

van Passel, M.W., Kant, R., Palva, A., Copeland, A., Lucas, S., Lapidus, A., Glavina Del Rio, T., Pitluck, S., Goltsman, E., Clum, A., Sun, H., Schmutz, J., Larimer, F.W., Land, M.L., Hauser, L., Kyrpides, N., Mikhailova, N., Richardson, P.P., Janssen, P.H., de Vos, W.M., and Smidt, H. (2011) Genome Sequence of the *Verrucomicrobium Opitutus terrae* PB90-1, an Abundant Inhabitant of Rice Paddy Soil Ecosystems. *J Bacteriol* **193**: 2367-2368.

Vanderpool, C.K., and Gottesman, S. (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54**: 1076-1089.

Vanderslice, R.W., Doherty, D.H., Capage, M.A., Betlach, M.R., Hassler, R.A., Henderson, N.M., Ryan-Graniero, J., and Tecklenburg, M. (1990) Genetic engineering of polysaccharide structure in *Xanthomonas campestris*. *In Biomedical and Biotechnological Advances in Industrial Polysaccharides*: 145-156.

Vauterin, L., Hoste, B., Kersters, K., and Swings, J. (1995) Reclassification of *Xanthomonas*. *Int J Syst Bacteriol* **45**: 472-489.

Vincken, J.P., Schols, H.A., Oomen, R.J., McCann, M.C., Ulvskov, P., Voragen, A.G., and Visser, R.G. (2003) If homogalacturonan were a side chain of rhamnogalacturonan I. Implications for cell wall architecture. *Plant Physiol* **132**: 1781-1789.

Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J., and Perret, X. (1998) Symbiotic implications of type III protein secretion machinery in *Rhizobium*. *Mol Microbiol* **28**: 1381-1389.

Vissenberg, K., Martinez-Vilchez, I.M., Verbelen, J.P., Miller, J.G., and Fry, S.C. (2000) In vivo colocalization of xyloglucan endotransglycosylase activity and its donor substrate in the elongation zone of *Arabidopsis* roots. *Plant Cell* **12**: 1229-1237.

Vogel, J., and Sharma, C.M. (2005) How to find small non-coding RNAs in bacteria. *Biol Chem* **386**: 1219-1238.

Vogel, J., and Papenfort, K. (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr Opin Microbiol* **9**: 605-611.

Vojnov, A.A., Zorreguieta, A., Dow, J.M., Daniels, M.J., and Dankert, M.A. (1998) Evidence for a role for the *gumB* and *gumC* gene products in the formation of xanthan from its pentasaccharide repeating unit by *Xanthomonas campestris*. *Microbiology* **144 (Pt 6)**: 1487-1493.

Vorholter, F.J., Schneiker, S., Goesmann, A., Krause, L., Bekel, T., Kaiser, O., Linke, B., Patschkowski, T., Ruckert, C., Schmid, J., Sidhu, V.K., Sieber, V., Tauch, A., Watt, S.A., Weisshaar, B., Becker, A., Niehaus, K., and Puhler, A. (2008) The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. *J Biotechnol* **134**: 33-45.

Vrsanska, M., Gorbacheva, I.V., Kratky, Z., and Biely, P. (1982) Reaction pathways of substrate degradation by an acidic endo-1,4-beta-xylanase of *Aspergillus niger*. *Biochim Biophys Acta* **704**: 114-122.

Wagegg, W., and Braun, V. (1981) Ferric citrate transport in *Escherichia coli* requires outer membrane receptor protein FecA. *J Bacteriol* **145**: 156-163.

Wandersman, C., and Delepelaire, P. (2004) Bacterial iron sources: from siderophores to hemophores. *Annu Rev Microbiol* **58**: 611-647.

Wang, L., Rong, W., and He, C. (2008) Two *Xanthomonas* extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. *Mol Plant Microbe Interact* **21**: 555-563.

Wang, L.H., He, Y., Gao, Y., Wu, J.E., Dong, Y.H., He, C., Wang, S.X., Weng, L.X., Xu, J.L., Tay, L., Fang, R.X., and Zhang, L.H. (2004) A bacterial cell-cell communication signal with cross-kingdom structural analogues. *Mol Microbiol* **51**: 903-912.

Warren, R.A. (1996) Microbial hydrolysis of polysaccharides. *Annu Rev Microbiol* **50**: 183-212.

Wassarman, K.M. (2002) Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. *Cell* **109**: 141-144.

Waters, C.M., and Bassler, B.L. (2005) Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**: 319-346.

Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615-628.

Weber, E., and Koebnik, R. (2005) Domain structure of HrpE, the Hrp pilus subunit of *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol* **187**: 6175-6186.

Wengelnik, K., and Bonas, U. (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol* **178**: 3462-3469.

Wengelnik, K., Van den Ackerveken, G., and Bonas, U. (1996) HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. *Mol Plant Microbe Interact* **9**: 704-712.

Wengelnik, K., Rossier, O., and Bonas, U. (1999) Mutations in the regulatory gene *hrpG* of *Xanthomonas campestris* pv. *vesicatoria* result in constitutive expression of all *hrp* genes. *J Bacteriol* **181**: 6828-6831.

White, F.F., Potnis, N., Jones, J.B., and Koebnik, R. (2009) The type III effectors of *Xanthomonas*. *Mol Plant Pathol* **10**: 749-766.

Whitehead, N.A., Barnard, A.M., Slater, H., Simpson, N.J., and Salmond, G.P. (2001) Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**: 365-404.

Wiggerich, H.G., Klauke, B., Koplin, R., Priefer, U.B., and Puhler, A. (1997) Unusual structure of the tonB-exb DNA region of *Xanthomonas campestris* pv. *campestris*: *tonB*, *exbB*, and *exbD1* are essential for ferric iron uptake, but *exbD2* is not. *J Bacteriol* **179**: 7103-7110.

Wiggerich, H.G., and Puhler, A. (2000) The *exbD2* gene as well as the iron-uptake genes *tonB*, *exbB* and *exbD1* of *Xanthomonas campestris* pv. *campestris* are essential for the induction of a hypersensitive response on pepper (*Capsicum annuum*). *Microbiology* **146 (Pt 5)**: 1053-1060.

Wilderman, P.J., Sowa, N.A., FitzGerald, D.J., FitzGerald, P.C., Gottesman, S., Ochsner, U.A., and Vasil, M.L. (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc Natl Acad Sci U S A* **101**: 9792-9797.

Willats, W.G., Rasmussen, S.E., Kristensen, T., Mikkelsen, J.D., and Knox, J.P. (2002) Sugar-coated microarrays: a novel slide surface for the high-throughput analysis of glycans. *Proteomics* **2**: 1666-1671.

Williams, P. (1980) Black rot: a continuing threat to world crucifers. *Plant Dis* **64**: 736-742.

Willis, D.K., Rich, J.J., and Hrabak, E.M. (1991) *hrp* genes of phytopathogenic bacteria. *Molecular plant Microbe Interaction* **4**: 132-138.

Wilson, E.O. (1992) *The Diversity of Life*: Harvard University Press.

Wilson, E.O. (1994) *Naturalist.* Washington: Island Press.

Winkelmann, G. (2002) Microbial siderophore-mediated transport. *Biochem Soc Trans* **30**: 691-696.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Jr., Woo, L., Chen, Y., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee, D., Sr., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Olson, M.V., and Nester, E.W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317-2323.

Xiao, Y., and Hutcheson, S.W. (1994) A single promoter sequence recognized by a newly identified alternate sigma factor directs expression of pathogenicity and host range determinants in *Pseudomonas syringae*. *J Bacteriol* **176**: 3089-3091.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., and Gordon, J.I. (2003a) A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* **299**: 2074-2076.

Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho, P.M., Minx, P., Latreille, P., Cordum, H., Van Brunt, A., Kim, K., Fulton, R.S., Fulton, L.A., Clifton, S.W., Wilson, R.K., Knight, R.D., and Gordon, J.I. (2007) Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol* **5**: e156.

Xu, Q., Gao, W., Ding, S.Y., Kenig, R., Shoham, Y., Bayer, E.A., and Lamed, R. (2003b) The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. *J Bacteriol* **185**: 4548-4557.

Xu, Q., Barak, Y., Kenig, R., Shoham, Y., Bayer, E.A., and Lamed, R. (2004a) A novel *Acetivibrio cellulolyticus* anchoring scaffoldin that bears divergent cohesins. *J Bacteriol* **186**: 5782-5789.

Xu, Q., Bayer, E.A., Goldman, M., Kenig, R., Shoham, Y., and Lamed, R. (2004b) Architecture of the *Bacteroides cellulosolvens* cellulosome: description of a cell surface-anchoring scaffoldin and a family 48 cellulase. *J Bacteriol* **186**: 968-977.

Yamazaki, A., Hirata, H., and Tsuyumu, S. (2008) HrpG regulates type II secretory proteins in *Xanthomonas axonopodis* pv. *citri*. *j gen plant pathol* **74**: 138-150.

Yang, C., Rodionov, D.A., Li, X., Laikova, O.N., Gelfand, M.S., Zagnitko, O.P., Romine, M.F., Obraztsova, A.Y., Nealson, K.H., and Osterman, A.L. (2006) Comparative genomics and experimental characterization of N-acetylglucosamine utilization pathway of *Shewanella oneidensis*. *J Biol Chem* **281**: 29872-29885.

Yoon, K.H., and Cho, J.Y. (2007) Transcriptional analysis of the *gum* gene cluster from *Xanthomonas oryzae* pathovar *oryzae*. *Biotechnol Lett* **29**: 95-103.

Yooseph, S., Sutton, G., Rusch, D.B., Halpern, A.L., Williamson, S.J., Remington, K., Eisen, J.A., Heidelberg, K.B., Manning, G., Li, W., Jaroszewski, L., Cieplak, P., Miller, C.S., Li, H., Mashiyama, S.T., Joachimiak, M.P., van Belle, C., Chandonia, J.M., Soergel, D.A., Zhai, Y., Natarajan, K., Lee, S., Raphael, B.J., Bafna, V., Friedman, R., Brenner, S.E., Godzik, A., Eisenberg, D., Dixon, J.E., Taylor, S.S., Strausberg, R.L., Frazier, M., and Venter, J.C. (2007) The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biol* **5**: e16.

York, W.S., Darvill, A.G., and Albersheim, P. (1984) Inhibition of 2,4-dichlorophenoxyacetic Acid-stimulated elongation of pea stem segments by a xyloglucan oligosaccharide. *Plant Physiol* **75**: 295-297.

Yue, W.W., Grizot, S., and Buchanan, S.K. (2003) Structural evidence for iron-free citrate and ferric citrate binding to the TonB-dependent outer membrane transporter FecA. *J Mol Biol* **332**: 353-368.

Yun, M.H., Torres, P.S., El Oirdi, M., Rigano, L.A., Gonzalez-Lamothe, R., Marano, M.R., Castagnaro, A.P., Dankert, M.A., Bouarab, K., and Vojnov, A.A. (2006) Xanthan induces plant susceptibility by suppressing callose deposition. *Plant Physiol* **141**: 178-187.

Zhang, S.S., He, Y.Q., Xu, L.M., Chen, B.W., Jiang, B.L., Liao, J., Cao, J.R., Liu, D., Huang, Y.Q., Liang, X.X., Tang, D.J., Lu, G.T., and Tang, J.L. (2008) A putative colR(XC1049)-colS(XC1050) two-component signal transduction system in *Xanthomonas campestris* positively regulates *hrpC* and *hrpE* operons and is involved in virulence, the hypersensitive response and tolerance to various stresses. *Res Microbiol* **159**: 569-578.

Zhang, Y., Zhang, Z., Ling, L., Shi, B., and Chen, R. (2004) Conservation analysis of small RNA genes in *Escherichia coli*. *Bioinformatics* **20**: 599-603.

Zhang, Z., Gosset, G., Barabote, R., Gonzalez, C.S., Cuevas, W.A., and Saier, M.H., Jr. (2005) Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*. *J Bacteriol* **187**: 980-990.

Zhou, G.K., Zhong, R., Richardson, E.A., Morrison, W.H., 3rd, Nairn, C.J., Wood-Jones, A., and Ye, Z.H. (2006) The poplar glycosyltransferase GT47C is functionally conserved with *Arabidopsis Fragile* fiber8. *Plant Cell Physiol* **47**: 1229-1240.

Zhou, J.M., and Chai, J. (2008) Plant pathogenic bacterial type III effectors subdue host responses. *Curr Opin Microbiol* **11**: 179-185.

Zou, L.F., Wang, X.P., Xiang, Y., Zhang, B., Li, Y.R., Xiao, Y.L., Wang, J.S., Walmsley, A.R., and Chen, G.Y. (2006) Elucidation of the *hrp* clusters of *Xanthomonas oryzae*

pv. *oryzicola* that control the hypersensitive response in nonhost tobacco and pathogenicity in susceptible host rice. *Appl Environ Microbiol* **72**: 6212-6224.

Zwiesler-Vollick, J., Plovanich-Jones, A.E., Nomura, K., Bandyopadhyay, S., Joardar, V., Kunkel, B.N., and He, S.Y. (2002) Identification of novel hrp-regulated genes through functional genomic analysis of the *Pseudomonas syringae* pv. *tomato* DC3000 genome. *Mol Microbiol* **45**: 1207-1218.

ANNEXES

Mon travail de thèse a essentiellement consisté à la caractérisation d'un nouveau système d'utilisation du xylane chez *Xcc* permettant la mise en évidence d'un groupe de gènes (incluant des TBDTs) essentiels pour l'utilisation de ce carbohydrate chez des bactéries symbiotiques humaines, aquatiques et terrestres. Cependant, j'ai également participé à deux autres études au cours de ma thèse. L'ensemble de ces travaux sont regroupés dans deux articles publiés.

La première étude a permis la caractérisation d'une nouvelle voie d'utilisation du *N*-acetylglucosamine chez *Xcc* faisant intervenir un transporteur de la famille MFS pour le passage à travers la membrane interne et 2 glucokinases pour la phosphorylation du GlcNAc.
De plus, les enzymes NagA et NagB-II sont impliquées dans le catabolisme du GlcNAc. Enfin, les gènes impliqués dans le transport et le métabolisme du GlcNAc sont sous le contrôle de deux régulateurs, NagQ et NagR. De manière intéressante, la séquence consensus reconnue par NagR est présente en amont de quatre gènes codant des TBDTs (*XCC0531*, *XCC2887*, *XCC2944* et *XCC3408*) suggérant un rôle de ces transporteurs dans l'utilisation du GlcNAc. Cependant, bien que ces 4 TBDTs soient sous le contrôle du GlcNAc via NagR, nos données indiquent qu'ils ne seraient pas directement impliqués dans le transport du GlcNAc, mais plutôt dans le transport de molécules plus complexes contenant du GlcNAc.
Ce travail suggère donc l'existence d'un CUT système impliqué dans l'utilisation de molécules complexes contenant du GlcNAc.
Au cours de ce travail, j'ai participé à la construction de mutants ainsi qu'aux études d'expression. L'ensemble des données sont rassemblées dans un article publié en 2010 dans la revue Journal of Bacteriology (Boulanger *et al.*, 2010).

JOURNAL OF BACTERIOLOGY, Mar. 2010, p. 1487–1497 Vol. 192, No. 6
0021-9193/10/$12.00 doi:10.1128/JB.01418-09

Identification and Regulation of the *N*-Acetylglucosamine Utilization Pathway of the Plant Pathogenic Bacterium *Xanthomonas campestris* pv. campestris▽†

Alice Boulanger,[1] Guillaume Déjean,[1] Martine Lautier,[1,2] Marie Glories,[1] Claudine Zischek,[1] Matthieu Arlat,[1,2] and Emmanuelle Lauber[1]*

Laboratoire des Interactions Plantes Micro-organismes (LIPM), UMR CNRS-INRA 2594/441, F-31326 Castanet-Tolosan, France,[1] *and Université de Toulouse, UPS, 118 Route de Narbonne, F-31062 Toulouse, France*[2]

Received 29 October 2009/Accepted 4 January 2010

Xanthomonas campestris **pv. campestris, the causal agent of black rot disease of brassicas, is known for its ability to catabolize a wide range of plant compounds. This ability is correlated with the presence of specific *c*arbohydrate *u*tilization loci containing *T*onB-dependent transporters (CUT loci) devoted to scavenging specific carbohydrates. In this study, we demonstrate that there is an *X. campestris* pv. campestris CUT system involved in the import and catabolism of *N*-acetylglucosamine (GlcNAc). Expression of genes belonging to this GlcNAc CUT system is under the control of GlcNAc via the LacI family NagR and GntR family NagQ regulators. Analysis of the NagR and NagQ regulons confirmed that GlcNAc utilization involves NagA and NagB-II enzymes responsible for the conversion of GlcNAc-6-phosphate to fructose-6-phosphate. Mutants with mutations in the corresponding genes are sensitive to GlcNAc, as previously reported for *Escherichia coli*. This GlcNAc sensitivity and analysis of the NagQ and NagR regulons were used to dissect the *X. campestris* pv. campestris GlcNAc utilization pathway. This analysis revealed specific features, including the fact that uptake of GlcNAc through the inner membrane occurs via a major facilitator superfamily transporter and the fact that this amino sugar is phosphorylated by two proteins belonging to the glucokinase family, NagK-IIA and NagK-IIB. However, NagK-IIA seems to play a more important role in GlcNAc utilization than NagK-IIB under our experimental conditions. The *X. campestris* pv. campestris GlcNAc NagR regulon includes four genes encoding TonB-dependent active transporters (TBDTs). However, the results of transport experiments suggest that GlcNAc passively diffuses through the bacterial envelope, an observation that calls into question whether GlcNAc is a natural substrate for these TBDTs and consequently is the source of GlcNAc for this nonchitinolytic plant-associated bacterium.**

Xanthomonas campestris pv. campestris, the causal agent of black rot disease of brassicas, produces extracellular plant cell wall-degrading enzymes which contribute to its pathogenicity by facilitating its spread through plant tissues and give the bacterium access to a ready source of nutrients via the *c*arbohydrate *u*tilization loci containing *T*onB-dependent transporters (CUT loci) (7, 16, 35). The CUT loci are characterized by the presence of genes encoding regulators, degradative enzymes, inner membrane transporters, and outer membrane TonB-dependent transporters (TBDTs), which have been identified as active carbohydrate transporters (7, 33, 44). However, recently, an example of passive diffusion through a TBDT in *Caulobacter crescentus* was described (17). *X. campestris* pv. campestris has 72 TBDTs and belongs to a class of bacteria in which TBDTs are overrepresented (7). Our previous study suggested that there are several CUT loci or systems in this bacterium (7).

N-Acetylglucosamine (GlcNAc) is an amino sugar that is used for the synthesis of cell surface structures in bacteria and plays an important role in supplying carbon and energy by entering the glycolytic pathway after it is converted into fructose-6-phosphate (fructose-6P) (1, 9). In a recent comparative study of bacterial GlcNAc utilization pathways and regulatory networks, Yang and coworkers identified conserved and distinct features of the GlcNAc utilization pathway in proteobacteria (48). The expression of *X. campestris* pv. campestris GlcNAc-specific genes was proposed to be controlled by NagR and NagQ regulators belonging to the LacI and GntR families, respectively. In *X. campestris* pv. campestris strain ATCC 33913, one predicted binding motif specific for NagQ (designated the NagQ box) consists of two imperfect repeats of the TGGTATT sequence separated by 4 bp and is located upstream of the *nagQ* gene (*XCC3414*) (Fig. 1A) (48). This gene is part of the *nag* cluster and is followed by genes encoding the major facilitator superfamily (MFS) inner membrane transporter NagP (*XCC3413*), the regulator NagR (*XCC3412*), the GlcN-6P deaminase NagB-II (*XCC3411*), and the GlcNAc-6P deacetylase NagA (*XCC3410*) (Fig. 1A). NagR boxes contain the palindromic sequence **AA**TGACA**R**CG**Y**TGTCA**TT** (bold type indicates less highly conserved nucleotides) and are upstream of genes encoding two glucokinase-like NagK-II proteins (*XCC2886* [*nagK-IIA*] and *XCC2943* [*nagK-IIB*]), as well as 5 genes encoding TBDTs (*XCC0531*, *XCC2887*, *XCC3045*, *XCC3408*, and *XCC2944* located downstream of *XCC2943*) (Fig. 1A). All of the *X. campestris* pv. campestris genes located

* Corresponding author. Mailing address: Laboratoire des Interactions Plantes Micro-organismes (LIPM), UMR CNRS-INRA 2594/441, F-31320 Castanet-Tolosan, France. Phone: (33) 561 28 50 47. Fax: (33) 561 28 50 61. E-mail: elauber@toulouse.inra.fr.

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

▽ Published ahead of print on 14 January 2010.

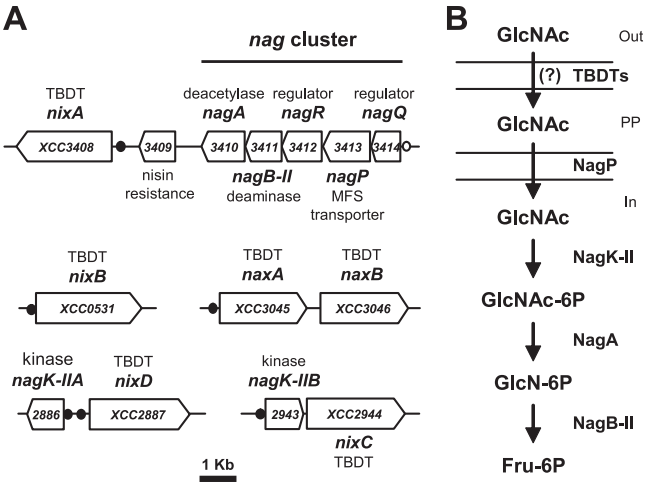


FIG. 1. *X. campestris* pv. campestris *N*-acetylglucosamine (GlcNAc) utilization pathway. (A) Organization of genes in the proposed GlcNAc utilization pathway. NagR boxes are indicated by filled circles, and the NagQ box is indicated by an open circle. (B) GlcNAc is proposed to be transported through the outer membrane by TBDTs and then transported across the inner membrane by the MFS transporter NagP. GlcNAc would then be phosphorylated by *nagK-II*-encoded enzymes. Subsequent metabolism via the *nagA*-encoded (GlcNAc-6P deacetylase) and *nagB-II*-encoded (GlcN-6P deaminase) enzymes results in fructose 6-phosphate (Fru-6P) (48). MFS, major facilitator superfamily; PP, periplasm; TBDT, TonB-dependent transporter.

downstream of NagR or NagQ boxes were proposed to belong to a GlcNAc utilization pathway involved in uptake of GlcNAc through the bacterial envelope and subsequent phosphorylation, deacetylation, and deamination, which finally leads to the common metabolic intermediate fructose-6-phosphate (Fig. 1B) (48). It was recently demonstrated that in *C. crescentus* the TBDT CC0446 gene, which is clustered with other *nag* genes, is responsible for the uptake of GlcNAc (17). The presence of TBDTs in the GlcNAc regulon, which has been observed in *Alteromonadales* and *Xanthomonadales* (48), suggests that genes belonging to the GlcNAc utilization pathway define a new CUT system.

Here we describe characterization of the *X. campestris* pv. campestris GlcNAc utilization pathway and regulatory network, which involves at least the repressors NagR and NagQ. TBDTs are associated with this pathway, confirming the presence of a GlcNAc CUT system in *X. campestris* pv. campestris. In this bacterium, GlcNAc entry and catabolism imply that novel families containing a GlcNAc inner membrane transporter and GlcNAc kinases are involved.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The *X. campestris* pv. *campestris* strains and plasmids used in this study are listed in Table 1. *X. campestris* pv. campestris cells were grown at 30°C in MOKA (7) or KADO (4) rich medium or in minimal medium (MME) (3). Sodium-free minimal medium contained 10.5 g/liter K_2HPO_4 and 4.5 g/liter KH_2PO_4. *Escherichia coli* cells were grown on Luria-Bertani medium at 37°C. For solid media, agar was added at a final concentration of 1.5% (wt/vol).

Antibiotics were used at the following concentrations: for *X. campestris* pv. campestris, 50 μg/ml rifampin, 50 μg/ml kanamycin, and 5 μg/ml tetracycline; for *E. coli*, 50 μg/ml ampicillin, 50 μg/ml kanamycin, and 10 μg/ml tetracycline.

Mutagenesis of *X. campestris* pv. campestris. *X. campestris* pv. campestris insertion mutants were constructed using the suicide plasmid pVO155 (34) with a 300- to 500-bp PCR amplicon internal to each open reading frame (ORF) (Table 1). Deletion mutants were constructed by using the *cre-lox* system adapted by Angot et al. (2) from the system of Marx and colleagues (30) or by using the *sacB* system (43). Deleted regions are indicated in Table 1. Oligonucleotide primers used for PCR amplification will be provided upon request.

Plasmids were introduced into *E. coli* by electroporation and into *Xanthomonas* strains by triparental conjugation, as described by Turner et al. (45).

Plasmid constructs. DNA manipulations were performed as described previously (42). For complementation studies, PCR amplicons (oligonucleotide primers used for PCR amplification will be provided upon request) were cloned into pCZ917, a derivative of pFAJ1700 (15) containing a 2,094-bp fragment of pSC150 (13) with the *lacI* gene, Ptac promoter, and T7 terminator.

Expression studies. Bacterial cultures grown in the appropriate medium were harvested after 6 h of incubation for β-glucuronidase assays (25).

The methods used for quantitative reverse transcription-PCR (qRT-PCR) experiments were adapted from the methods of Blanvillain et al. (7). A 2-μg sample of RNA was treated with RNase-free DNase I (Amersham) for 30 min at 37°C. After DNase inactivation (10 min at 75°C), RNAs were reverse transcribed with Superscript II (Invitrogen) using random hexamers (Biolabs) for 10 min at room temperature and then for 1 h at 42°C. Oligonucleotide primers used for quantitative PCR amplification will be provided upon request. 16S rRNA was used as a control for real-time PCR (7, 32).

Growth curves. Growth curves were generated using a FLUOStar Omega apparatus (BMG Labtech, Offenburg, Germany) with four replicates. Growth was measured using 96-well flat-bottom microtiter plates with 200-μl preparations inoculated at an optical density at 600 nm (OD_{600}) of 0.1 from 4 independent washed overnight precultures. The microplates were shaken continuously at 700 rpm using the double-orbital-shaking mode.

[^{14}C]GlcNAc transport experiments. Transport experiments with radiolabeled GlcNAc (specific activity, 2.04 GBq/mmol; PerkinElmer) were performed as previously described (7). For competition experiments, unlabeled sugars were added to [^{14}C]GlcNAc at final concentrations of 50 and 500 μM, and cells were incubated for 1 h before collection. The initial concentration-dependent GlcNAc transport was determined using the rapid dilution method as previously described (7, 33).

GlcNAc phosphorylation assays. GlcNAc kinase activity assays were performed using an enzyme-linked assay based on the NAD^+/NADH ratio (19). Fifty milliliters of an overnight culture in minimal medium supplemented with 10 mM GlcNAc was centrifuged and resuspended in 2 ml of resuspension buffer (0.05 mM Tris HCl [pH 8], 13.3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol). Cells were disrupted with a French press and centrifuged, and 100 μl of supernatant was added to 900 μl of reaction buffer (0.1 M Tris HCl [pH 7.5], 10 mM $MgCl_2$, 1 mM phosphoenolpyruvate, 4 mM ATP, 0.2 mM NADH, 10 mM GlcNAc, 4 U lactate dehydrogenase [Sigma], 4 U pyruvate kinase [Sigma]) prewarmed for 5 min at 37°C. The OD_{340} was determined every 10 s for 5 min at 37°C. A decrease in the OD_{340} corresponded to production of NAD^+ from NADH and was enzymatically coupled to GlcNAc phosphorylation to form GlcNAc-6P. Protein concentrations of cell lysates were determined using the Bradford assay (Bio-Rad).

RESULTS

GlcNAc and chitobiose, but not chitin, are carbon and nitrogen sources for *X. campestris* pv. campestris. The presence in *X. campestris* pv. campestris of genes proposed to belong to a GlcNAc utilization pathway suggests that GlcNAc can be metabolized by *X. campestris* pv. campestris. Therefore, the growth rates of *X. campestris* pv. campestris cultures in MME supplemented with GlcNAc and with other carbon sources were compared. After sucrose and glucose, GlcNAc and the GlcNAc dimer chitobiose were among the best carbon sources for *X. campestris* pv. campestris (Fig. 2A and B). In the presence of the GlcNAc homopolymer chitin, slight growth was reproducibly observed (Fig. 2B), probably due to the presence of small amounts of free GlcNAc or chitobiose molecules. This result suggests that *X. campestris* pv. campestris is not able to efficiently degrade chitin, a suggestion corroborated by the absence of any obvious chitinase-encoding gene in the genome of *X. campestris* pv. campestris strain ATCC 33913 (14, 48).

TABLE 1. Plasmids and *X. campestris* pv. campestris strains used or generated in this study

Strain or plasmid	Characteristics[a]	Location[b]	Designation	Reference
Xanthomonas strains				
Wild type	Wild-type strain; rifampin-resistant derivative of *X. campestris* pv. campestris LMG568 (= ATCC 33913)			31
nixB::pVO	*XCC0531*::pVO155; Rif^r Km^r	736	XP010	7
nixC::pVO	*XCC2944*::pVO155; Rif^r Km^r	334	XP041	7
nixD::pVO	*XCC2887*::pVO155; Rif^r Km^r	1558	XP040	7
naxA::pVO	*XCC3045*::pVO155; Rif^r Km^r	1572	XP044	7
naxB::pVO	*XCC3046*::pVO155; Rif^r Km^r	702	XP045	7
nixA::pVO	*XCC3408*::pVO155; Rif^r Km^r	1897	XP059	7
nagQ::pVO	*XCC3414*::pVO155; Rif^r Km^r	659	XP108	This study
nagR::pVO	*XCC3412*::pVO155; Rif^r Km^r	519	XP109	This study
nagA::pVO	*XCC3410*::pVO155; Rif^r Km^r	236	XP110	This study
nagK-IIA::pVO	*XCC2886*::pVO155; Rif^r Km^r	303	XP111	This study
Δ*nagQ*	Δ*XCC3414*; Rif^r	From 203 to stop	XP112	This study
Δ*nagR*	Δ*XCC3412*; Rif^r	From start to stop	XP113	This study
Δ*nagP*	Δ*XCC3413*; Rif^r	From start to stop	XP114	This study
Δ*nagA*	Δ*XCC3410*; Rif^r	From start to stop	XP115	This study
Δ*nagB-II*	Δ*XCC3411*; Rif^r	From start to stop	XP116	This study
Δ*nagK-IIA*	Δ*XCC2886*; Rif^r	From start to stop	XP117	This study
Δ*nagK-IIB*	Δ*XCC2943*; Rif^r	From start to stop	XP118	This study
Δ*nagK-IIAB*	Δ*XCC2886* Δ*XCC2943*; Rif^r	From start to stop	XP119	This study
Δ*nagA* pC-*nagA*	Δ*XCC3410* pC-*XCC3410*; Rif^r Tet^r		XP120	This study
Δ*nagB-II* pC-*nagB-II*	Δ*XCC3411* pC-*XCC3411*; Rif^r Tet^r		XP121	This study
Δ*nagB-II* pC-*nagB-II-nagA*	Δ*XCC3411* pC-*XCC3411-XCC3410*; Rif^r Tet^r		XP122	This study
Δ*nagK-IIA* pC-*nagK-IIA*	Δ*XCC2886* pC-*XCC2886*; Rif^r Tet^r		XP123	This study
Δ*nagK-IIB* pC-*nagK-IIB*	Δ*XCC2943* pC-*XCC2943*; Rif^r Tet^r		XP124	This study
Δ*nagK-IIAB* pC-*nagK-IIA*	Δ*XCC2886* Δ*XCC2943* pC-*XCC2886*; Rif^r Tet^r		XP125	This study
Δ*nagK-IIAB* pC-*nagK-IIB*	Δ*XCC2886* Δ*XCC2943* pC-*XCC2943*; Rif^r Tet^r		XP126	This study
Δ*nagR nagK-IIA*::pVO	Δ*XCC3412 XCC2886*::pVO155; Rif^r Km^r		XP127	This study
Δ*nagQ nagK-IIA*::pVO	Δ*XCC3414 XCC2886*::pVO155; Rif^r Km^r		XP128	This study
Δ*nagP nagA*::pVO	Δ*XCC3413 XCC3410*::pVO155; Rif^r Km^r		XP129	This study
Δ*nagK-IIAB nagA*::pVO	Δ*XCC2886* Δ*XCC2943 XCC3410*::pVO155; Rif^r Km^r		XP130	This study
Δ*nagP* Δ*nagA*	Δ*XCC3413* Δ*XCC3410*; Rif^r		XP131	This study
Δ*nagK-IIAB* Δ*nagA*	Δ*XCC2886* Δ*XCC2943* Δ*XCC3410*; Rif^r		XP132	This study
Plasmids				
pVO155	pUC119 derivative containing the promoterless *gus* (*uidA*) reporter gene encoding β-glucuronidase, used for insertion mutagenesis; Km^r Amp^r			34
pFAJ1700	pTR102-derived expression vector containing a multiple-cloning site and transcriptional terminators in both orientations; Tet^r Amp^r			15
pSC150	pET-26b(+) derivative vector with a Ptac promoter sequence; Km^r			13
pCZ917	pFAJ1700 derivative containing 2,094 bp of pSC150 with *lacI*, Ptac promoter, and T7 terminator; Tet^r Amp^r			This study
pC-*nagA*	pCZ917-*XCC3410*; Tet^r Km^r	From −20 to stop		This study
pC-*nagB-II*	pCZ917-*XCC3411*; Tet^r Km^r	From −19 to stop		This study
pC-*nagB-II-nagA*	pCZ917-*XCC3411-XCC3410*; Tet^r Km^r	From −19 of *XCC3411* to stop of *XCC3410*		This study
p*C-nagK-IIA*	pCZ917-*XCC2886*; Tet^r Km^r	From −21 to stop		This study
pC-*nagK-IIB*	pCZ917-*XCC2943*; Tet^r Km^r	From −19 to stop		This study

[a] Rif: rifampin; Km: kanamycin; Tet: tetracycline.
[b] Position of insertion, deletion, or *X. campestris* pv. campestris sequence cloned relative to the putative start codon.

GlcNAc is also a nitrogen source for *X. campestris* pv. campestris, since this bacterium grows in nitrogen-depleted MME (MME without Casamino Acids and NH_4SO_4 [see Materials and Methods]) in the presence of GlcNAc, whereas no growth was observed in the presence of glucose (data not shown).

GlcNAc pathway genes are induced by GlcNAc. The expression of genes located downstream of putative NagR or NagQ boxes was measured to assess the relationship of these genes to utilization of GlcNAc. The TBDT gene *XCC3046* located downstream of the TBDT gene *XCC3045* might belong to the same operon (Fig. 1B) and was therefore included in this

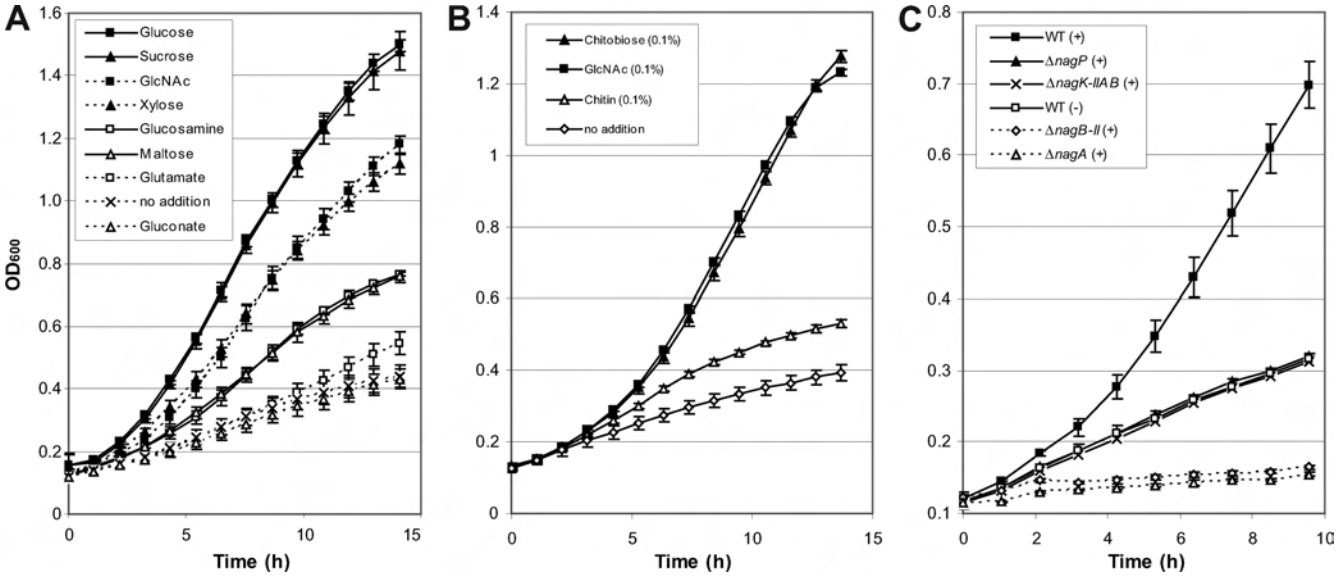


FIG. 2. Growth of *X. campestris* pv. campestris wild-type and mutant strains in the presence of *N*-acetylglucosamine (GlcNAc) or other carbon sources. (A) Growth of the wild-type strain in the presence of carbon sources and in the presence of carbon and nitrogen sources. (B) Growth of the wild-type strain in the presence of GlcNAc, chitobiose, or chitin. (C) Growth of the wild type (WT) and mutant strains in the presence (+) or in the absence (−) of GlcNAc. After overnight growth in complete medium, cells were harvested, washed, and resuspended in minimal medium. Carbohydrates were added at a final concentration of 10 mM (A and C) or 0.1% (B). The error bars indicate the standard deviations obtained from 4 independent experiments.

analysis. None of the primer pairs designed for qRT-PCR experiments yielded reliable results for the *XCC2886* and *XCC3046* genes. Therefore, transcriptional fusions with the promoterless *uidA* gene were constructed by pVO155 insertion mutagenesis (34). The resulting mutants were used to measure the expression of these genes with a β-glucuronidase assay. All of the ORFs in the proposed GlcNAc CUT system, including the *nagQ* and *nagR* putative regulatory genes, were clearly induced in the presence of GlcNAc (Table 2). However, *XCC3045* and *XCC3046* were repressed. Based on these results, the TBDT genes *XCC3408*, *XCC0531*, *XCC2944*, and *XCC2887* were designated *nixA*, *nixB*, *nixC*, and *nixD*, respectively (*N*-acetylglucosamine-*i*nduced genes in *Xanthomonas*), while the TBDT genes *XCC3045* and *XCC3046* were designated *naxA* and *naxB*, respectively (*N*-acetylglucosamine-*a*ssociated genes in *Xanthomonas*).

To determine whether expression of the other TBDT genes of *X. campestris* pv. campestris is affected by GlcNAc, β-glucuronidase assays were performed using pVO155 insertion mutants with mutations in each of the 72 TBDT genes (7). The

TABLE 2. Relative expression ratios for genes in the *N*-acetylglucosamine utilization pathway

Gene	Designation	Function	Expression ratios (SD)		
			Wild type with GlcNAc/ wild type in MME	Δ*nagQ* mutant in MME/ wild type in MME	Δ*nagR* mutant in MME/ wild type in MME
XCC3414[a]	*nagQ*	GntR repressor	7.11 (0.70)[d]	ND[e]	0.47 (0.04)
XCC3413[a]	*nagP*	MFS transporter	6.43 (0.62)[d]	2.28 (0.90)[d]	0.83 (0.10)
XCC3412[a]	*nagR*	LacI repressor	6.99 (0.73)[d]	7.91 (3.34)[d]	ND
XCC3411[a]	*nagB-II*	Deaminase	6.72 (0.46)[d]	3.17 (0.16)[d]	0.88 (0.06)
XCC3410[a]	*nagA*	Deacetylase	3.16 (0.16)[d]	8.33 (2.76)[d]	0.41 (0.28)
XCC2886[b]	*nagK-IIA*	GlcNAc kinase	2.70 (0.12)[d]	0.88 (0.02)	4.88 (1.06)[d]
XCC2943[a]	*nagK-IIB*	GlcNAc kinase	3.00 (0.81)[d]	1.70 (0.73)[c]	27.11 (7.00)[c,d]
XCC3408[a]	*nixA*	TBDT	6.32 (0.84)[d]	0.31 (0.08)	7.77 (1.17)[d]
XCC0531[a]	*nixB*	TBDT	3.97 (0.49)[d]	0.63 (0.24)	1.08 (0.47)
XCC2944[a]	*nixC*	TBDT	9.21 (1.17)[d]	0.42 (0.15)	40.14 (14.35)[d]
XCC2887[a]	*nixD*	TBDT	106.70 (12.56)[d]	0.86 (0.55)	47.82 (15.16)[d]
XCC3045[a]	*naxA*	TBDT	0.42 (0.12)[d]	0.61 (0.34)	0.26 (0.08)
XCC3046[b]	*naxB*	TBDT	0.25 (0.06)[d]	ND	ND

[a] Data from real-time quantitative reverse transcriptase PCR performed in at least three independent experiments. Calculation of the relative expression included normalization with the 16S rRNA data.

[b] Data from β-glucuronidase assays performed in at least three independent experiments using pVO155 insertion mutations leading to transcriptional fusions with the promotorless *uidA* gene. Insertions were made in the wild-type strain and, for *XCC2886*, in the Δ*nagR* and Δ*nagQ* strains.

[c] qRT-PCR expression values were obtained from *nagR* and *nagQ* pVO155 insertion mutants instead of deletion mutants.

[d] The levels of expression in the conditions compared were significantly different ($P < 0.05$) as determined using a Student *t* test.

[e] ND, not determined.

nixA, *nixB*, *nixC*, and *nixD* genes were the only TBDT genes induced by GlcNAc (data not shown). It is worth noting that the expression of the *nix* TBDT genes was not as strongly induced in the pVO155 insertion mutant (the induction levels ranged from 1.8-fold for *nixB* to 36-fold for *nixD* [data not shown]) as was expected based on the results of qRT-PCR for a wild-type background (for which the induction levels ranged from 3.97-fold for *nixB* to 106.7-fold for *nixD* [Table 2]).

The expression of GlcNAc-induced genes was then measured after growth in MME supplemented with a range of GlcNAc concentrations. Representative results obtained with the *nixD*::pVO mutant are reported here because this mutant displayed one of the highest levels of induction in the presence of GlcNAc and because its growth was not impaired in MME supplemented with GlcNAc (see below). The reporter gene was induced with 5 μM to 20 mM GlcNAc. The maximal induction (around 30-fold) was observed with 50 μM GlcNAc (data not shown). Induction was also observed with high concentrations of glucosamine (GlcN), but the maximal induction was only 3-fold (data not shown).

NagQ and NagR are GlcNAc pathway-specific regulators. The involvement of two presumptive regulators, NagQ and NagR, was evaluated by comparing *nix* gene expression in the wild-type strain and *nix* gene expression in *nagQ* and *nagR* mutants in MME without added GlcNAc. Mutants with insertions and deletions of these two regulatory genes were constructed, but deletion mutants were chosen to avoid possible polar effects, since both regulatory genes may be expressed as part of an operon (Fig. 1A).

The levels of expression of the *nagP*, *nagR*, *nagB-II*, and *nagA* genes were clearly higher in the Δ*nagQ* deletion mutant than in the wild-type strain (Table 2). These genes are downstream of the *nagQ* gene, which is itself downstream of the unique putative NagQ box detected in the *X. campestris* pv. campestris genome (Fig. 1A). This result suggests that NagQ regulates its own expression and that the genes from *nagQ* to *nagA* form an operon. The expression of the other GlcNAc pathway genes was not significantly affected by deletion of *nagQ*.

The expression of NagQ-regulated genes was not affected by deletion of *nagR*. The expression of *nixA*, *nixC*, *nixD*, *nagK-IIA*, and *nagK-IIB* was derepressed in the Δ*nagR* deletion mutant compared to the expression in the wild-type strain (Table 2). This is in agreement with the presence of putative NagR boxes in the promoter regions of these genes or operons, as determined by Yang and coworkers (48). Surprisingly, the GlcNAc-induced TBDT *nixB* gene located downstream of a NagR box seemed not to be regulated by NagR under our conditions (Table 2), suggesting that four of the five putative NagR boxes are functional. This result prompted us to generate a position weight matrix with the PREDetector program (23) using the four functional NagR boxes for screening the *X. campestris* pv. campestris genome. Of the 61 predicted targets, 16 are located in intergenic regions (see Table S1 in the supplemental material). Sequences upstream of *nixA*, *nixB*, *nixD*, *nagK-IIA*, and *nagK-IIB* each had strong predicted NagR-binding sites. However, the score obtained for the *nixB* promoter site was close to the scores for weak sites (see Table S1 in the supplemental material). This low score might explain the poor NagR regulation of *nixB*, a gene which is nevertheless induced by GlcNAc.

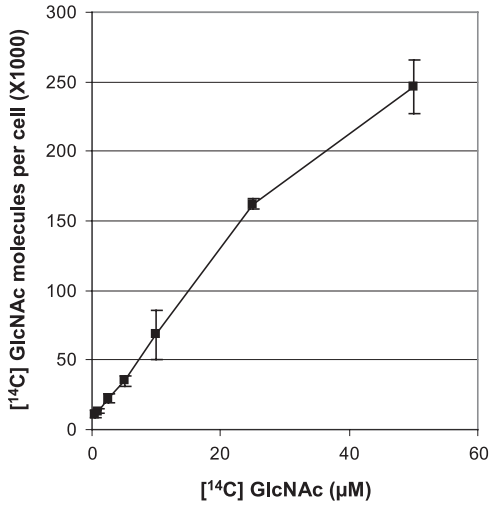


FIG. 3. Concentration-dependent transport of ^{14}C-labeled *N*-acetylglucosamine (GlcNAc) into *X. campestris* pv. campestris. Cells were grown in minimal medium without GlcNAc, and transport was measured for 15 s at the [^{14}C]GlcNAc concentrations indicated.

A sequence logo was generated by WebLogo (http://weblogo.berkley.edu/; 11) from the alignment of the four putative functional NagR boxes, which resulted in discovery of a new NagR box (GTTGACARCGYTGTCANC). This NagR box differed at positions 1, 2, and 18 from the previously proposed NagR box (AATGACARCGYTGTCATT) (48).

Together, these results show that NagR and NagQ are functional repressors of genes belonging to the GlcNAc CUT system. Proteins encoded by NagR- and NagQ-regulated genes can be classified into two main categories: transport and metabolism of GlcNAc.

Transport of GlcNAc in *X. campestris* pv. campestris. (i) Free GlcNAc passively diffuses through the envelope. GlcNAc uptake rates in the *X. campestris* pv. campestris wild-type strain were compared after overnight preculture in the presence of GlcNAc (induced) and after overnight preculture in the presence of xylose (uninduced), a substrate that results in a growth rate similar to that obtained with GlcNAc (Fig. 2A) but does not affect the expression of GlcNAc-induced TBDT genes (data not shown). Before transport experiments were performed with [^{14}C]GlcNAc, cells were washed to remove nonradiolabeled GlcNAc from the medium. The GlcNAc uptake rates under the two conditions were not significantly different (data not shown), suggesting that GlcNAc import is limited by a GlcNAc-independent transport step.

The initial concentration-dependent [^{14}C]GlcNAc transport, reflecting the dissociation constant (K_d) for GlcNAc uptake, was determined using the previously described rapid dilution method (7, 33). The kinetic values revealed that the uptake rate was low and monophasic (Fig. 3), suggesting either that the outer and inner membrane transporters have similar affinities for GlcNAc or that transport through the outer membrane is limiting and masks transport through the inner membrane. The deduced K_d (138.9 μM) is more than 100-fold higher than the K_d estimated for passive uptake of GlcNAc through the

TABLE 3. Inhibition of uptake of 0.5 μM [^{14}C]GlcNAc by various carbohydrates in *X. campestris* pv. campestris wild-type strain after 1 h of incubation

Carbohydrate	% of control uptake (SD) at carbohydrate concn of[a]:	
	50 μM	500 μM
N-Acetylglucosamine	65 (1)	18 (2)
N-Acetylglucosamine-6P	101 (4)	105 (4)
Glucosamine	66 (3)	41 (0)
Chitobiose	58 (1)	34 (4)
Glucose	98 (3)	84 (5)
Galactose	105 (5)	103 (5)
Mannose	95 (1)	94 (5)
Fructose	100 (3)	101 (4)
Xylose	99 (5)	102 (11)
Sucrose	100 (5)	101 (7)

[a] Standard deviations were calculated from the results of three independent experiments.

TABLE 4. Rates of ^{14}C-labeled *N*-acetylglucosamine transport compared to the rate in *X. campestris* pv. campestris wild-type strain[a]

Strain	Transporter family	Mean % transport (SD)[b]
Wild type		100 (9.9)
nixA::pVO	TBDT	96.6 (8.8)
nixB::pVO	TBDT	115.2 (6.4)
nixC::pVO	TBDT	111.2 (12.5)
nixD::pVO	TBDT	106.8 (11.8)
naxA::pVO	TBDT	114 (11.3)
naxB::pVO	TBDT	115 (15.7)
Δ*nagP*	MFS	1.2 (0.4)
Δ*nagP* pC-*nagP*		98 (4)

[a] Transport rates were measured 60 min after addition of ^{14}C-labeled *N*-acetylglucosamine.
[b] Standard deviations were calculated from three independent experiments.

CC0446 TBDT in *C. crescentus* (17) and is in a range similar to the range for K_d values obtained for passive diffusion through porins (18). Therefore, free GlcNAc uptake through the *X. campestris* pv. campestris envelope seems to occur *via* passive diffusion rather than by active uptake, although the *X. campestris* pv. campestris GlcNAc regulon contains at least four TBDT genes encoding active outer membrane transporters.

A 100-fold excess of unlabeled glucose, galactose, sucrose, mannose, xylose, fructose, or GlcNAc-6P had no effect on radiolabeled GlcNAc uptake (Table 3). With unlabeled GlcNAc, the concentration for inhibition of the transport rate to one-half of the control rate was estimated to be 193.7 μM, which is in accordance with the K_d deduced from the results of the initial concentration-dependent GlcNAc transport assays. Glucosamine and chitobiose both inhibit radioactive GlcNAc uptake as much as unlabeled GlcNAc (Table 3). Inhibition of GlcNAc uptake by chitobiose could be due either to the chitobiose molecule itself or to degradation of this molecule to GlcNAc. These competition experiments suggest that GlcNAc, glucosamine, and probably chitobiose are transported across the envelope *via* the same transporters.

(ii) None of the GlcNAc-induced TBDTs seems to play a major role in utilization of free GlcNAc. The rates of [^{14}C]GlcNAc uptake in the *X. campestris* pv. campestris wild-type strain and in GlcNAc regulon TBDT insertion mutants were compared, and none of the TBDT mutants exhibited a significant effect in GlcNAc uptake (Table 4). Furthermore, the growth rates of strains with mutations in the *nix* and *nax* TBDT genes in the presence of 10 mM GlcNAc were similar to the growth rate of the wild-type strain (data not shown). The absence of a phenotype for mutants with single mutations in TBDT genes could be due to the redundant functions of the transporters.

(iii) NagP is the major GlcNAc inner membrane transporter in *X. campestris* pv. campestris. The *nagP* gene was deleted to test the putative role of NagP, which belongs to the major facilitator superfamily (MFS), in the transport of GlcNAc. Growth of the Δ*nagP* strain was impaired on MME containing GlcNAc as the sole carbon source (Fig. 2C), suggesting that this transporter could be involved in the uptake of GlcNAc through the inner membrane. The rate of uptake of radiolabeled GlcNAc obtained for the Δ*nagP* strain was only 1.2% of the rate obtained for the wild-type strain (Table 4). In the *nagP*-complemented strain, GlcNAc transport capacity (Table 4) and growth on GlcNAc-containing MME (Fig. 4A) were restored, confirming that NagP is the major transporter of GlcNAc across the inner membrane.

Catabolism of GlcNAc in *X. campestris* pv. campestris. (i) NagK-IIA and NagK-IIB phosphorylate GlcNAc. In the cytoplasm, the first step in the *X. campestris* pv. campestris GlcNAc utilization pathway is phosphorylation of GlcNAc (Fig. 1B). Two genes, *nagK-IIA* and *nagK-IIB*, coding for proteins belonging to the glucokinase family, belong to the GlcNAc regulon (Table 2), and their products have been proposed to act as putative GlcNAc kinases in *Xanthomonas* (48). To test the function of these proteins in the phosphorylation of GlcNAc, Δ*nagK-IIA* and Δ*nagK-IIB* single mutants, as well as a Δ*nagK-IIAB* double mutant, were constructed. The GlcNAc kinase activity of the Δ*nagK-IIAB* double mutant was about 41% of the wild-type activity (Fig. 5A). The GlcNAc kinase activity of the double mutant could have been due to the presence of residual ADP or pyruvate in the crude extracts used in the experiments. Wild-type GlcNAc kinase activity was restored when either *nagK-IIA* or *nagK-IIB* was supplied in *trans* on an expression plasmid, suggesting that both proteins phosphorylate GlcNAc. However, the activities obtained for each single mutant did not differ significantly from the wild-type activity (Fig. 5A), suggesting that these two proteins are functionally redundant.

The growth of the Δ*nagK-IIAB* double mutant was clearly impaired in GlcNAc-containing minimal medium (Fig. 2C and Fig. 5B). The growth of the Δ*nagK-IIA* single mutant was also affected, but to a lesser extent, whereas the Δ*nagK-IIB* mutant grew like the wild-type strain (Fig. 5B). Growth of the Δ*nagK-IIAB* double mutant in GlcNAc minimal medium was partially restored when *nagK-IIA* or *nagK-IIB* was overexpressed in *trans* on an expression plasmid; however, better complementation was observed with *nagK-IIA* (Fig. 4B). These results suggest that although both GlcNAc kinases are enzymatically functional, NagK-IIA plays a major role in GlcNAc utilization, whereas NagK-IIB, which is apparently not essential, can also function in this capacity.

(ii) GlcNAc is toxic for *nagA* and *nagB-II* mutants. NagA deacetylase (XCC3410) catalyzes the conversion of GlcNAc-6P

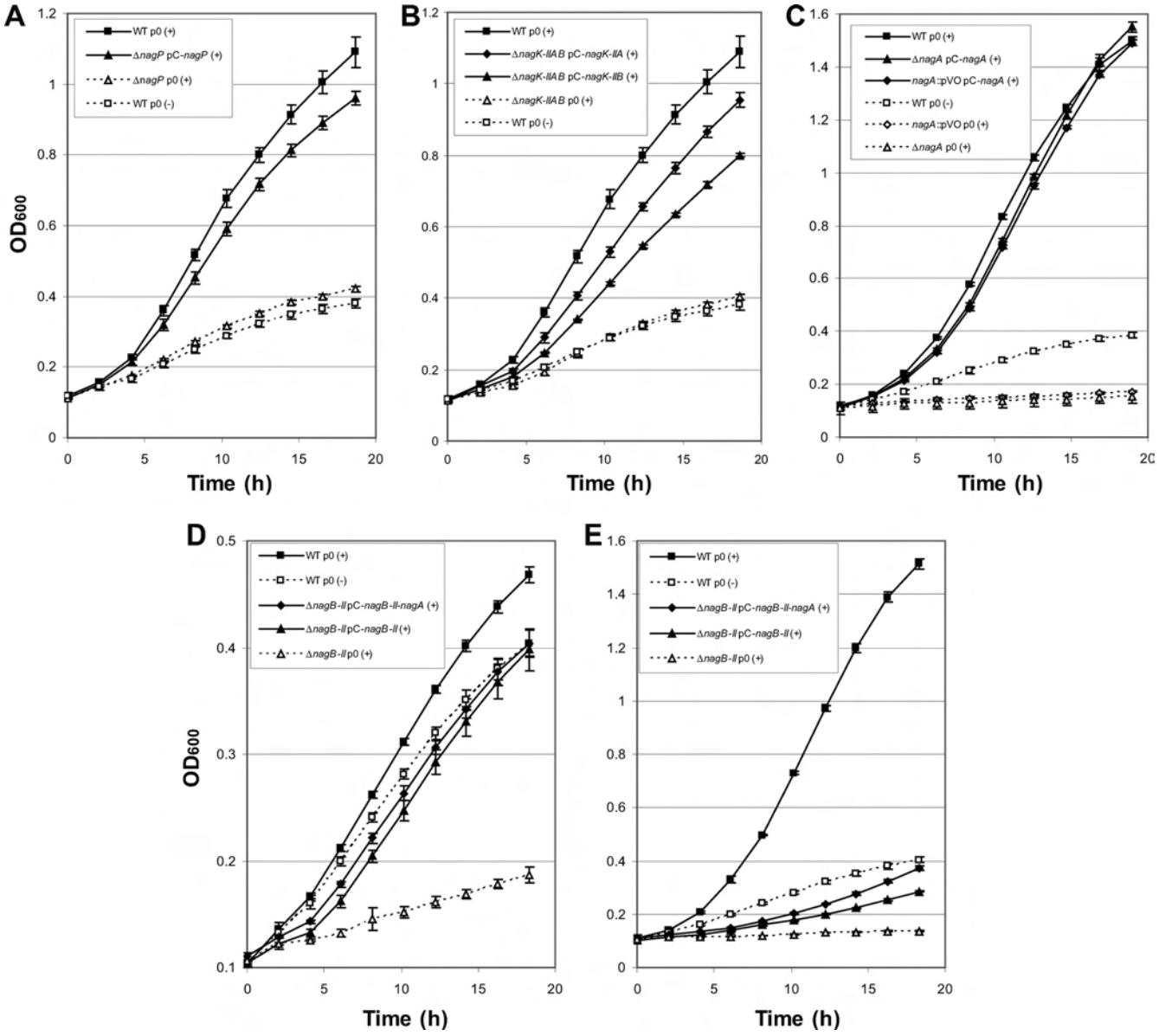


FIG. 4. Complementation of mutants with mutations in genes encoding proteins in the *X. campestris* pv. campestris *N*-acetylglucosamine (GlcNAc) utilization pathway. After overnight growth in complete medium, cells were harvested, washed, and resuspended in minimal medium containing 200 μM isopropyl-β-thiogalactoside (IPTG), 5 μg/ml of tetracycline, and 10 mM (A, B, C, and E) or 100 μM (D) GlcNAc. The error bars indicate the standard deviations obtained from 4 independent experiments. WT, wild-type strain.

to GlcN-6P, and the NagB-II deaminase (XCC3411) deaminates and isomerizes GlcN-6P to fructose-6P (Fig. 1B). Because the NagA and NagB-II proteins of *X. campestris* pv. campestris are very similar to those of *Shewanella oneidensis* strain MR-1 (48), their enzymatic activities were not studied *in vitro*. However, their biological importance for the *X. campestris* pv. campestris GlcNAc utilization pathway was studied genetically. Addition of GlcNAc to the medium resulted in rapid inhibition of growth of the Δ*nagA* mutant (Fig. 2C and 6A). Complete inhibition was observed with a GlcNAc concentration of 1 mM (data not shown). This "amino sugar sensitivity" phenomenon, previously observed in *E. coli* (6, 46), has been proposed to be due to accumulation of GlcNAc-6P, leading to pentose starvation (6). In *X. campestris* pv. campestris, GlcNAc sensitivity was still observed when gluconate or glucose was added to the medium (Fig. 6A). Intriguingly, a *nagA*::pVO insertion mutant seemed to be less sensitive to GlcNAc than the Δ*nagA* deletion mutant, since significant differences in inhibition of growth between these two strains were observed for GlcNAc concentrations ranging from 10 μM to 250 μM (Fig. 6B). This difference is likely not due to any polar effect on the expression of a downstream gene because GlcNAc sensitivity was abolished when *nagA* was supplied in *trans* in both mutants (Fig. 4C).

Growth of the Δ*nagB-II* mutant was also inhibited in the presence of GlcNAc (Fig. 2C), but the sensitivity of this mutant

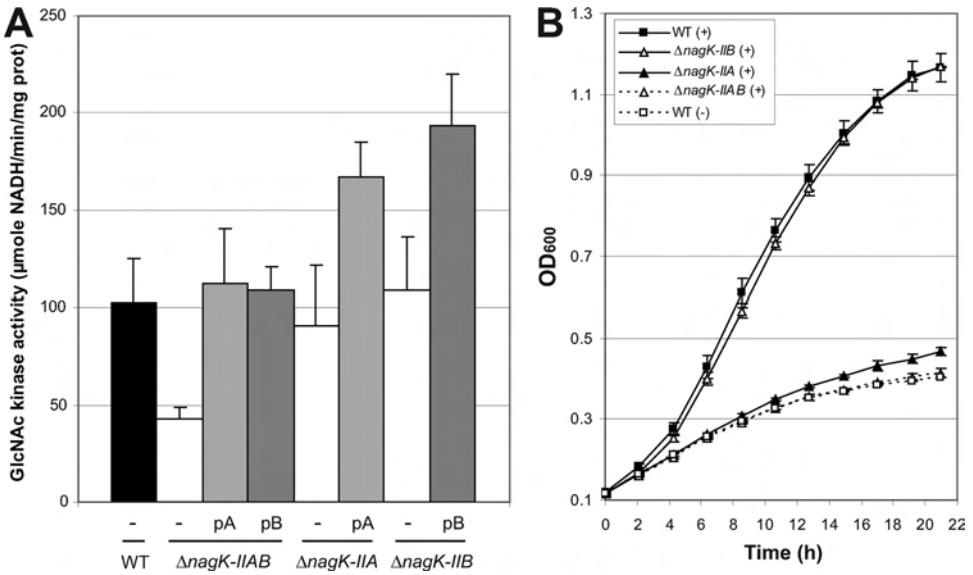


FIG. 5. Phosphorylation of *N*-acetylglucosamine by *X. campestris* pv. campestris NagK-II enzymes. (A) *In vitro* GlcNAc kinase assay based on the NAD^+/NADH ratio of the wild-type strain (WT), strains with single deletions (Δ*nagK-IIA* and Δ*nagK-IIB*), or a strain with a double deletion (Δ*nagK-IIAB*) containing plasmid pC-*nagK-IIA* (pA), plasmid pC-*nagK-IIB* (pB), or no plasmid (−). The activity observed in ATP-depleted medium was subtracted to normalize the assay results. Strains were cultured in minimal medium supplemented with 10 mM *N*-acetylglucosamine (GlcNAc) to induce expression of the *nagK-II* genes. The error bars indicate the standard deviations obtained from 4 independent experiments. (B) Growth curves for mutant strains cultivated in minimal medium supplemented (+) or not supplemented (−) with GlcNAc. After overnight growth in complete medium, cells were harvested, washed, and resuspended in minimal medium containing 10 mM *N*-acetylglucosamine. The error bars indicate the standard deviations obtained from 4 independent experiments.

was less pronounced than that of the Δ*nagA* strain since no inhibition was observed at GlcNAc concentrations below 10 μM with the Δ*nagB-II* mutant (data not shown). A difference in sensitivity of *nagA* and *nagB-II* mutants has also been observed for *E. coli* and could be due to the gradual assimilation of GlcN-6P in the *nagB-II* mutant, whereas GlcNAc-6P that accumulates in the *nagA* mutant cannot be metabolized further (6). When *nagB-II* or *nagB-II* and *nagA* were supplied in *trans*, GlcNAc sensitivity was abolished in the presence of 100 μM GlcNAc (Fig. 4D), but the growth rate of each complemented strain was lower than that of the wild-type strain. Surprisingly, very faint complementation was observed in the presence of 10 mM GlcNAc, even with isopropyl-β-D-thiogalactopyranoside (IPTG) induction of complementing genes (Fig. 4E).

(iii) NagP, NagK-II, and NagA enzymes are in the same metabolic pathway. The GlcNAc sensitivity of *nagA* mutants with a second mutation in either *nagP* or the *nagK-IIA* and *nagK-IIB* genes was tested to confirm that the NagP transporter and NagK-IIA and NagK-IIB GlcNAc kinases are located upstream of NagA deacetylase in the GlcNAc utilization pathway. The absence of GlcNAc entry or phosphorylation should prevent the formation of GlcNAc-6P and therefore reduce or eliminate its accumulation in *nagA* mutants. A Δ*nagP* Δ*nagA* double mutant and a Δ*nagK-IIAB* Δ*nagA* triple mutant were constructed, and their sensitivities to GlcNAc were tested. Growth inhibition of *nagA* mutants by GlcNAc was abolished when the inner membrane transporter was absent (Δ*nagP* Δ*nagA* mutant) or when the two GlcNAc kinases were mutated (Δ*nagK-IIAB* Δ*nagA* mutant) (Fig. 6C). Identical results were obtained for *nagA*::pVO mutants (data not shown). This epistatic effect confirms that the NagP, NagK-II, and NagA enzymes are all part of the same metabolic pathway.

DISCUSSION

Efficient bacterial exploitation (sensing, uptake, and catabolism) of nutrients requires close regulation of specific genetic programs. In *X. campestris* pv. campestris, carbohydrate utilization loci containing TonB-dependent transporters (CUT loci) have been proposed to play a major role in carbohydrate scavenging (7). In this paper, we characterize the *X. campestris* pv. campestris GlcNAc utilization pathway and propose that this pathway comprises a CUT system since the TBDTs were coregulated with the other genes in this pathway.

The MFS transporters comprise a novel family of GlcNAc transporters in bacteria. The Major facilitator superfamily (MFS) inner membrane transporters comprise the largest family of secondary active transporters and have a diverse range of substrates, including ions, sugars, drugs, nucleosides, and peptides, but individual members of this family show stringent specificity. NagP is the major GlcNAc transporter in *X. campestris* pv. campestris and is required for bacterial growth on GlcNAc as the sole carbon source. To our knowledge, this is the first report of bacterial GlcNAc uptake through an MFS transporter. Previously, two important families of inner membrane transporters for the uptake of GlcNAc in bacteria were identified: the phosphotransferase transporter systems (PTS), analogous to NagE in *E. coli* (28), and the ATP-binding cassette transporters (ABC), such as NgcEFG in *Streptomyces olivaceoviridis* (47). Our results indicate that the MFS is a third

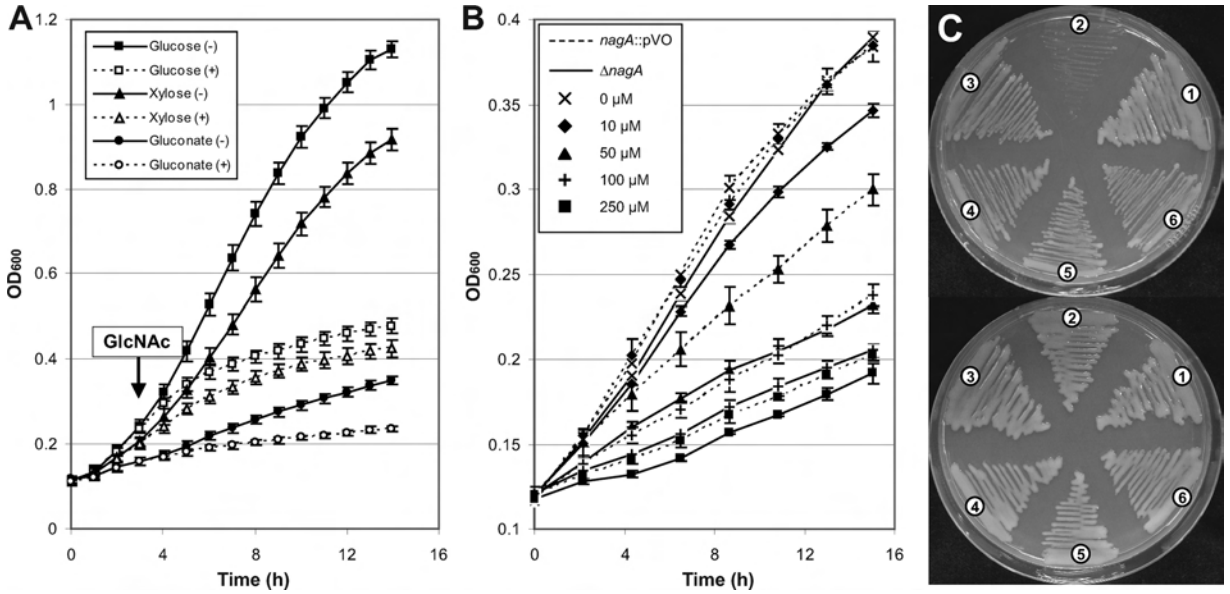


FIG. 6. *X. campestris* pv. campestris *nagA* mutants are sensitive to *N*-acetylglucosamine. (A) Effect of addition of *N*-acetylglucosamine (GlcNAc) on growth of the Δ*nagA* mutant cultivated in minimal medium containing various carbon sources at a concentration of 20 mM. After 3 h of growth, media were supplemented (+) or not supplemented (−) with *N*-acetylglucosamine at a concentration of 10 mM. The error bars indicate the standard deviations obtained from 4 independent experiments. (B) Effects of different concentrations of *N*-acetylglucosamine (GlcNAc) on growth of *nagA*::pVO (dashed lines) and Δ*nagA* (solid lines) mutants. Cells were grown in minimal medium containing the concentrations of GlcNAc indicated. The error bars indicate the standard deviations obtained from 4 independent experiments. (C) Effects of secondary mutations on the GlcNAc sensitivity of the Δ*nagA* mutant. Strains were streaked on complete medium plates containing 10 mM *N*-acetylglucosamine (upper plate) or 10 mM xylose (lower plate). Region 1, wild-type strain; region 2, Δ*nagA* mutant; region 3, Δ*nagP* mutant; region 4, Δ*nagP* Δ*nagA* mutant; region 5, Δ*nagK-IIAB* mutant; region 6, Δ*nagK-IIAB* Δ*nagA* mutant. The photograph was taken after 96 h of growth. Identical results were obtained on minimal medium plates.

inner membrane transporter family involved in the uptake of GlcNAc. The gene encoding NagP is present in all *Xanthomonadales* strains sequenced so far, suggesting that it has a conserved role in GlcNAc uptake in these bacteria.

The MFS NagP transporter is a nonphosphorylating transport system. Therefore, in contrast to *E. coli*, in which GlcNAc uptake through the PTS is coupled with phosphorylation (28), the first step in the *X. campestris* pv. campestris GlcNAc utilization pathway in the cytoplasm is phosphorylation of GlcNAc.

GlcNAc: a new substrate for glucokinase enzymes. NagK-IIA and NagK-IIB, which are very similar (54% identity), can both phosphorylate GlcNAc. These glucokinase family proteins (PF02685 in the Pfam database [5]) possess a ROK domain (PF0480) and might be members of group B of the hexokinase superfamily (27). Although the physiological functions of most members of this group are unknown, several enzymes belonging to this family have been shown to phosphorylate glucose and other hexoses (8, 21). Our work provides the first evidence of a role in phosphorylation of GlcNAc for this protein family. Therefore, these glucokinase family proteins represent a new subfamily of GlcNAc kinases, as proposed by Yang and coworkers (48). *X. campestris* and *Xanthomonas axonopodis* strains possess both *nagK-II* genes. However, in the other sequenced *Xanthomonadaceae* (*Stenotrophomonas*, *Xylella*, and *Xanthomonas oryzae*) an ortholog of *nagK-IIB* is not present or not functional (data not shown), raising a question about the role of NagK-IIB in strains that also possess NagK-IIA. Although NagK-IIB is not required for *X. campestris* pv. campestris growth on GlcNAc-containing media, this protein can substitute for NagK-IIA in GlcNAc phosphorylation. The glucokinase family protein GlcK from *Bacillus sphaericus* has been reported to have substrate ambiguity, phosphorylating not only glucose but also fructose and mannose (21). It is thus possible that GlcNAc is an alternative substrate for NagK-IIB. It was recently shown that in *X. campestris* pv. campestris strain 8004 NagK-IIB was not involved in the phosphorylation of glucose, arabinose, xylose, sorbitol, mannose, mannitol, sorbose, fructose, galactose, rhamnose, sucrose, or maltose (29). Since this enzyme is under the control of NagR, its substrate is probably related to GlcNAc metabolism. Therefore, it could be interesting to determine the role of NagK-IIB in the phosphorylation of other hexosamines, such as glucosamine or even chitobiose.

The *X. campestris* pv. campestris *nag* cluster encodes two functional repressors, NagR and NagQ. At least three nonorthologous types of transcriptional regulators were proposed previously to control the expression of GlcNAc utilization genes in proteobacteria: the ROK family protein NagC, the LacI family protein NagR, and the GntR-type protein NagQ (48). Interestingly, bacteria in the genus *Xanthomonas* are the only bacteria known to have two genes coding for these GlcNAc-specific regulators (i.e., *nagR* and *nagQ*) (48), both of which are located in the *nag* cluster. The work performed here showed that both of the regulators are functional. Repression by NagQ seems to be restricted to the *nag* operon. Therefore, NagQ can be considered a local transcription factor. On the other hand, NagR likely acts at multiple sites, since several NagR boxes were identified and are scattered throughout the

X. campestris pv. campestris genome. The NagR regulon could be broader, and it would be interesting to identify additional targets of this repressor in global transcriptomic analyses.

In *X. campestris* pv. campestris, NagQ represses *nagR*. The importance of this regulatory loop in GlcNAc catabolism is a matter of conjecture. Furthermore, the GlcNAc regulatory network could be much more complicated, and additional regulatory elements could participate in this network. Indeed, *nixA*, *naxA*, and *nagK-IIA* belong to the Clp regulon, a conserved global regulator shown to play a central role in the regulation of virulence factors in *X. campestris* pv. campestris (22).

The *nagQ* gene is specific for *Xanthomonas* strains. Indeed, in the sequenced plant-pathogenic *Xylella* strains, which have reduced genomes (37), *nagQ* is partially deleted, suggesting that this gene has become vestigial and can therefore be lost during evolution. This gene is also not present in *Stenotrophomonas maltophilia*, a non-plant-pathogenic species belonging to the *Xanthomonadaceae* family that includes free-living as well as endophytic isolates and opportunistic human pathogens (12, 41). Interestingly, in both bacteria, the NagQ box located upstream of the *nag* cluster is replaced by a candidate NagR box (48).

GlcNAc passively diffuses through the outer membrane. Although four *X. campestris* pv. campestris active outer membrane TBDT genes are induced by GlcNAc, three of which are under the control of NagR, GlcNAc passively diffuses through the outer membrane. Furthermore, none of the GlcNAc-induced TBDTs was individually required for growth on GlcNAc and for uptake of this molecule. This implies either that several of these TBDTs allow diffusion of GlcNAc or, alternatively, that GlcNAc can diffuse through the outer membrane via other transporters, such as porins.

TBDTs, which are well known for their role in iron and vitamin B_{12} uptake (36), were shown previously to be involved in the active uptake of carbohydrates, such as maltodextrins in *C. crescentus* (33) or sucrose in *X. campestris* pv. campestris (7). In *C. crescentus*, TBDT CC0446 is essential for growth on the chito-oligosaccharides $(GlcNAc)_3$ and $(GlcNAc)_5$ in a TonB-dependent manner, but transport of GlcNAc apparently occurs by passive diffusion through this transporter, a process which is TonB independent (17). Therefore, we propose that *X. campestris* pv. campestris TBDTs belonging to the GlcNAc regulon are involved in the active transport of complex molecules containing GlcNAc, as observed for *C. crescentus* (17). Consequently, this implies that the source of GlcNAc could be more complex than monomeric GlcNAc.

What is the source of GlcNAc for *X. campestris* pv. campestris in the environment? There are two principal sources of GlcNAc in nature: chitin and bacterial cell walls. In *E. coli*, about 50% of cell wall peptidoglycan is broken down each generation (20, 24). Recent bioinformatic analyses identified the gene coding for the D-Ala-D-Ala aminopeptidase that is responsible for catabolism of the cell wall precursor D-Ala-D-Ala as a part of the *Streptomyces coelicolor* DasR regulon (39). This pleiotropic regulator belonging to the GntR family is essential for development and is involved in the regulation of genes encoding both the GlcNAc PTS and the GlcN-6P deaminase NagB (38). Therefore, there is a direct relationship between peptidoglycan recycling and GlcNAc utilization. Although this link was not studied in *X. campestris* pv. campestris, a role for an *X. campestris* pv. campestris GlcNAc-induced TBDT(s) in the active uptake of peptidoglycan degradation products can be readily envisaged.

Alternatively, *X. campestris* pv. campestris could have developed a system to exploit chito-oligosaccharides derived from fungal and insect chitin degraded by plant chitinases (26) or by chitinolytic bacteria. Occupation of niches in the plant phyllosphere by epiphytic, saprophytic, and pathogenic fungi and bacteria and the interactions of these organisms are important for the *Xanthomonas* life cycle (40). In this environment, the growth of *X. campestris* pv. campestris on chitin or its polymeric subunits relies on other organisms for chitin degradation, as has recently been proposed for the nonchitinolytic aquatic bacterium *C. crescentus* (17). To exploit chito-oligosaccharides, *X. campestris* pv. campestris must produce enzymes involved in degradation of these compounds. This bacterium has one of the largest glycobiomes, as determined using the CAZy database (http://www.cazy.org/) (10). Interestingly, downstream of the *nixD* TBDT gene (*XCC2887*), we identified several putative enzyme-encoding genes which could be involved in the degradation of oligosaccharides that contain GlcNAc. Among these, *XCC2889* codes for a protein belonging to the GH-18 family, which includes endo-beta-*N*-acetylglucosaminidase (EC 3.2.1.96), and *XCC2890* codes for a protein belonging to the glycoside hydrolase GH-20 family, which includes the β-hexosaminidases (EC 3.2.1.52). Work is now under way to further characterize these genes, to determine whether they are part of the GlcNAc CUT system, and to determine which molecules are targeted by these enzymes.

ACKNOWLEDGMENTS

We are grateful to Scott Soby for critical reading of the manuscript. We also acknowledge Laurent Noël and Servane Baufumé for valuable discussions and critical reading of the manuscript.

A. Boulanger and G. Déjean were supported by a grant from the French Ministère de la Recherche et de l'Enseignement Supérieur. We gratefully acknowledge financial support from the Département Santé des Plantes et Environnement de l'Institut National de la Recherche Agronomique (grant 2007_0441_02) and from the French Agence Nationale de la Recherche (grant ANR-08-BLAN-0193-01).

REFERENCES

1. **Alvarez-Anorve, L. I., M. L. Calcagno, and J. Plumbridge.** 2005. Why does *Escherichia coli* grow more slowly on glucosamine than on *N*-acetylglucosamine? Effects of enzyme levels and allosteric activation of GlcN6P deaminase (NagB) on growth rates. J. Bacteriol. **187:**2974–2982.
2. **Angot, A., N. Peeters, E. Lechner, F. Vailleau, C. Baud, L. Gentzbittel, E. Sartorel, P. Genschik, C. Boucher, and S. Genin.** 2006. *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. Proc. Natl. Acad. Sci. U. S. A. **103:**14620–14625.
3. **Arlat, M., C. L. Gough, C. E. Barber, C. Boucher, and M. J. Daniels.** 1991. *Xanthomonas campestris* contains a cluster of *hrp* genes related to the larger *hrp* cluster of *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. **4:**593–601.
4. **Arlat, M., F. Van Gijsegem, J. C. Huet, J. C. Pernollet, and C. A. Boucher.** 1994. PopA1, a protein which induces a hypersensitivity-like response on specific *Petunia* genotypes, is secreted via the Hrp pathway of *Pseudomonas solanacearum*. EMBO J. **13:**543–553.
5. **Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. Sonnhammer, D. J. Studholme, C. Yeats, and S. R. Eddy.** 2004. The Pfam protein families database. Nucleic Acids Res. **32:**D138–D141.
6. **Bernheim, N. J., and W. J. Dobrogosz.** 1970. Amino sugar sensitivity in *Escherichia coli* mutants unable to grow on *N*-acetylglucosamine. J. Bacteriol. **101:**384–391.
7. **Blanvillain, S., D. Meyer, A. Boulanger, M. Lautier, C. Guynet, N. Denance, J. Vasse, E. Lauber, and M. Arlat.** 2007. Plant carbohydrate scavenging

through tonb-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS One **2:**e224.
8. **Brigham, C. J., and M. H. Malamy.** 2005. Characterization of the RokA and HexA broad-substrate-specificity hexokinases from Bacteroides fragilis and their role in hexose and N-acetylglucosamine utilization. J. Bacteriol. **187:** 890–901.
9. **Calcagno, M., P. J. Campos, G. Mulliert, and J. Suastegui.** 1984. Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from *Escherichia coli*. Biochim. Biophys. Acta **787:**165–173.
10. **Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. **37:**D233–D238.
11. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res. **14:**1188–1190.
12. **Crossman, L. C., V. C. Gould, J. M. Dow, G. S. Vernikos, A. Okazaki, M. Sebaihia, D. Saunders, C. Arrowsmith, T. Carver, N. Peters, E. Adlem, A. Kerhornou, A. Lord, L. Murphy, K. Seeger, R. Squares, S. Rutter, M. A. Quail, M. A. Rajandream, D. Harris, C. Churcher, S. D. Bentley, J. Parkhill, N. R. Thomson, and M. B. Avison.** 2008. The complete genome, comparative and functional analysis of Stenotrophomonas maltophilia reveals an organism heavily shielded by drug resistance determinants. Genome Biol. **9:**R74.
13. **Cunnac, S.** 2004. Identification à l'échelle génomique des effecteurs dépendant du système de sécrétion de type III de la bactérie phytopathogène *Ralstonia solanacearum*. Ph.D. thesis. Université Toulouse III Paul Sabatier, Toulouse, France.
14. **da Silva, A. C. R., J. A. Ferro, F. C. Reinach, C. S. Farah, L. R. Furlan, R. B. Quaggio, C. B. Monteiro-Vitorello, M. A. V. Sluys, N. F. Almeida, L. M. C. Alves, A. M. do Amaral, M. C. Bertolini, L. E. A. Camargo, G. Camarotte, F. Cannavan, J. Cardozo, F. Chambergo, L. P. Ciapina, R. M. B. Cicarelli, L. L. Coutinho, J. R. Cursino-Santos, H. El-Dorry, J. B. Faria, A. J. S. Ferreira, R. C. C. Ferreira, M. I. T. Ferro, E. F. Formighieri, M. C. Franco, C. C. Greggio, A. Gruber, A. M. Katsuyama, L. T. Kishi, R. P. Leite, E. G. M. Lemos, M. V. F. Lemos, E. C. Locali, M. A. Machado, A. M. B. N. Madeira, N. M. Martinez-Rossi, E. C. Martins, J. Meidanis, C. F. M. Menck, C. Y. Miyaki, D. H. Moon, L. M. Moreira, M. T. M. Novo, V. K. Okura, M. C. Oliveira, V. R. Oliveira, H. A. Pereira, A. Rossi, J. A. D. Sena, C. Silva, R. F. de Souza, L. A. F. Spinola, M. A. Takita, R. E. Tamura, E. C. Teixeira, R. I. D. Tezza, M. Trindade dos Santos, D. Truffi, S. M. Tsai, F. F. White, J. C. Setubal, and J. P. Kitajima.** 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature **417:**459–463.
15. **Dombrecht, B., J. Vanderleyden, and J. Michiels.** 2001. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol. Plant-Microbe Interact. **14:**426–430.
16. **Dow, J. M., and M. J. Daniels.** 1994. Pathogenicity determinants and global regulation of pathogenicity of Xanthomonas campestris pv. campestris. Curr. Top. Microbiol. Immunol. **192:**29–41.
17. **Eisenbeis, S., S. Lohmiller, M. Valdebenito, S. Leicht, and V. Braun.** 2008. NagA-dependent uptake of *N*-acetyl-glucosamine and *N*-acetyl-chitin oligosaccharides across the outer membrane of *Caulobacter crescentus*. J. Bacteriol. **190:**5230–5238.
18. **Freundlieb, S., U. Ehmann, and W. Boos.** 1988. Facilitated diffusion of p-nitrophenyl-alpha-D-maltohexaoside through the outer membrane of *Escherichia coli*. Characterization of LamB as a specific and saturable channel for maltooligosaccharides. J. Biol. Chem. **263:**314–320.
19. **Gonzali, S., L. Pistelli, L. De Bellis, and A. Alpi.** 2001. Characterization of two *Arabidopsis thaliana* fructokinases. Plant Sci. **160:**1107–1114.
20. **Goodell, E. W.** 1985. Recycling of murein by *Escherichia coli*. J. Bacteriol. **163:**305–310.
21. **Han, B., H. Liu, X. Hu, Y. Cai, D. Zheng, and Z. Yuan.** 2007. Molecular characterization of a glucokinase with broad hexose specificity from Bacillus sphaericus strain C3-41. Appl. Environ. Microbiol. **73:**3581–3586.
22. **He, Y. W., A. Y. Ng, M. Xu, K. Lin, L. H. Wang, Y. H. Dong, and L. H. Zhang.** 2007. Xanthomonas campestris cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network. Mol. Microbiol. **64:**281–292.
23. **Hiard, S., R. Maree, S. Colson, P. A. Hoskisson, F. Titgemeyer, G. P. van Wezel, B. Joris, L. Wehenkel, and S. Rigali.** 2007. PREDetector: a new tool to identify regulatory elements in bacterial genomes. Biochem. Biophys. Res. Commun. **357:**861–864.
24. **Jacobs, C., L. J. Huang, E. Bartowsky, S. Normark, and J. T. Park.** 1994. Bacterial cell wall recycling provides cytosolic muropeptides as effectors for beta-lactamase induction. EMBO J. **13:**4684–4694.
25. **Jefferson, R. A., T. A. Kavanagh, and M. W. Bevan.** 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:**3901–3907.
26. **Kasprzewska, A.** 2003. Plant chitinases—regulation and function. Cell. Mol. Biol. Lett. **8:**809–824.
27. **Kawai, S., T. Mukai, S. Mori, B. Mikami, and K. Murata.** 2005. Hypothesis: structures, evolution, and ancestor of glucose kinases in the hexokinase family. J. Biosci. Bioeng. **99:**320–330.
28. **Lengeler, J. W., K. Jahreis, and U. F. Wehmeier.** 1994. Enzymes II of the phosphoenol pyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. Biochim. Biophys. Acta **1188:** 1–28.
29. **Lu, G. T., Z. J. Yang, F. Y. Peng, Y. N. Tan, Y. Q. Tang, J. X. Feng, D. J. Tang, Y. Q. He, and J. L. Tang.** 2007. The role of glucose kinase in carbohydrate utilization and extracellular polysaccharide production in *Xanthomonas campestris* pathovar campestris. Microbiology **153:**4284–4294.
30. **Marx, C. J., and M. E. Lidstrom.** 2002. Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. Biotechniques **33:** 1062–1067.
31. **Meyer, D., E. Lauber, D. Roby, M. Arlat, and T. Kroj.** 2005. Optimization of pathogenicity assays to study the *Arabidopsis thaliana-Xanthomonas campestris* pv. campestris pathosystem. Mol. Plant Pathol. **6:**327–333.
32. **Morales, C. Q., J. Posada, E. Macneale, D. Franklin, I. Rivas, M. Bravo, J. Minsavage, R. E. Stall, and M. C. Whalen.** 2005. Functional analysis of the early chlorosis factor gene. Mol. Plant-Microbe Interact. **18:**477–486.
33. **Neugebauer, H., C. Herrmann, W. Kammer, G. Schwarz, A. Nordheim, and V. Braun.** 2005. ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. J. Bacteriol. **187:**8300–8311.
34. **Oke, V., and S. R. Long.** 1999. Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. Mol. Microbiol. **32:**837–849.
35. **Onsando, J.** 1992. Black rot of crucifers. Dis. Veg. Oil Seed Crops **2:**243–252.
36. **Postle, K., and R. J. Kadner.** 2003. Touch and go: tying TonB to transport. Mol. Microbiol. **49:**869–882.
37. **Puhler, A., M. Arlat, A. Becker, M. Gottfert, J. P. Morrissey, and F. O'Gara.** 2004. What can bacterial genome research teach us about bacteria-plant interactions? Curr. Opin. Plant Biol. **7**:137–147.
38. **Rigali, S., H. Nothaft, E. E. Noens, M. Schlicht, S. Colson, M. Muller, B. Joris, H. K. Koerten, D. A. Hopwood, F. Titgemeyer, and G. P. van Wezel.** 2006. The sugar phosphotransferase system of Streptomyces coelicolor is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. Mol. Microbiol. **61:**1237–1251.
39. **Rigali, S., F. Titgemeyer, S. Barends, S. Mulder, A. W. Thomae, D. A. Hopwood, and G. P. van Wezel.** 2008. Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces. EMBO Rep. **9:**670–675.
40. **Rudolph, K.** 1993. Infection of the plant by *Xanthomonas*, p. 193–264. *In* J. G. Swings and E. L. Civerolo (ed.), *Xanthomonas*. Chapman and Hall, London, United Kingdom.
41. **Ryan, R. P., S. Monchy, M. Cardinale, S. Taghavi, L. Crossman, M. B. Avison, G. Berg, D. van der Lelie, and J. M. Dow.** 2009. The versatility and adaptation of bacteria from the genus Stenotrophomonas. Nat. Rev. Microbiol. **7:**514–525.
42. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
43. **Schafer, A., A. Tauch, W. Jager, J. Kalinowski, G. Thierbach, and A. Puhler.** 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene **145:**69–73.
44. **Schauer, K., D. A. Rodionov, and H. de Reuse.** 2008. New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem. Sci. **33:**330–338.
45. **Turner, P., C. E. Barber, and M. J. Daniels.** 1985. Evidence for clustered pathogenicity genes in *Xanthomonas campestris* pv. campestris. Mol. Gen. Genet. **199:**338–343.
46. **White, R. J.** 1968. Control of amino sugar metabolism in *Escherichia coli* and isolation of mutants unable to degrade amino sugars. Biochem. J. **106:**847–858.
47. **Xiao, X., F. Wang, A. Saito, J. Majka, A. Schlosser, and H. Schrempf.** 2002. The novel *Streptomyces olivaceoviridis* ABC transporter Ngc mediates uptake of *N*-acetylglucosamine and *N,N′*-diacetylchitobiose. Mol. Genet. Genomics **267:**429–439.
48. **Yang, C., D. A. Rodionov, X. Li, O. N. Laikova, M. S. Gelfand, O. P. Zagnitko, M. F. Romine, A. Y. Obraztsova, K. H. Nealson, and A. L. Osterman.** 2006. Comparative genomics and experimental characterization of *N*-acetylglucosamine utilization pathway of *Shewanella oneidensis*. J. Biol. Chem. **281:**29872–29885.

TABLE S1. Computational prediction of the NagR regulon of *Xanthomonas campestris* pv. campestris.

Numbers highlighted in yellow indicates genes validated experimentally as part of the NagR regulon of *Xanthomonas campestris* pv. campestris. Numbers highlighted in green indicates cis-elements displaying a threshold for weak sites. The prediction has been performed with PREDetector using the position weight matrix obtained with *nixD* , *nixA* , *nagK-IIA* and *nagK-IIB* NagR-boxes.

#	GENE ID	FUNCTION	CIS-ELEMENT	POSITION	SCORE	CO-TRANSC GENE	CO-TRANSC FUNCTION
1	XCC2887, iroN / nixD	TonB-dependent receptor	gttgacaacgctgtcaac	-355	19.32	-	-
	XCC2886c, glk / nagK-IIA	glucose kinase		-369			
2	XCC2886c, glk / nagK-IIA	glucose kinase	gtagacaacgttgtcagc	-82	19.26	-	-
	XCC2887, iroN / nixD	TonB-dependent receptor		-642			
3	XCC3408c, fyuA / nixA	TonB-dependent receptor	gctgacaacgttatcagc	-341	18.72	-	-
4	XCC2943, glk / nagK-IIB	glucose kinase	attgacagcgttgtcatc	-83	18.27	XCC2944	TonB-dependent receptor
5	XCC0531 / nixB	conserved hypothetical protein	gttgacaccgatgtcatc	-92	14.39	-	-
	XCC0530c	conserved hypothetical protein	gttgacaccgatgtcatc	-108	14.39	-	-
8	XCC2136, zwf	glucose-6-phosphate 1-dehydrogenase	gctgataacgatgtctcc	-37	13.41	XCC2137/XCC2138	glucose kinase/6-phosphogluconolactonase
	XCC2135c, ugpC	sugar ABC transporter ATP-binding protein		-403		-	-
9	XCC1217, hrpF	HrpF protein	gtagacagcattttcagc	-858	12.85	-	-
12	XCC3284c	conserved hypothetical protein	gttgaaaacgatttcagc	-496	12.44	-	-
15	XCC3450c, prc	tail-specific protease	gttgctggcgttgtcatc	-90	11.93	-	-
20	XCC1880c, tsr	chemotaxis protein	gtggacaaggtggtcatc	-152	11.26	-	-
21	XCC3045, bfeA / naxA	ferric enterobactin receptor	gatgacaacgatggcact	-145	11.17	-	-
29	XCC1265, czcD	cobalt-zinc-cadmium resistance protein	gaaggcagcgtggtcagc	-196	10.81	-	-
50	XCC4237, oar	Oar protein	gatgacatctctgtaaat	-166	10.19	-	-
51	XCC3322c, ggt	gamma-glutamyltranspeptidase	gttgccaacgataccagt	-51	10.16	-	-
	XCC3323, ilvC	ketol-acid reductoisomerase		-607		XCC3324	acetolactate synthase isozyme II large subunit
56	XCC3470c, mreB	rod shape-determining protein	gcagaaaacgaagtcaac	-76	10.04	-	-
	XCC3471	sugar kinase		-219	10.04		
60	XCC1413, gpo	glutathione peroxidase-like protein	gatgaccacgctctccac	-1	10.03	-	-

Afin de s'adapter aux changements environnementaux rapides, les bactéries ont développé des régulateurs tels que les facteurs sigma, pour orienter l'expression des gènes en réponse à diverses variations extracellulaires.

La deuxième étude à laquelle j'ai participé au cours de ma thèse a décrypté le rôle et le mécanisme d'activation du facteur sigma de type ECF, σ^E qui est codé par le gène *rpoE,* chez *Xcc*. Le gène *rpoE* est présent au sein d'une unité transcriptionnelle où on trouve le gène *rseA* qui code pour le facteur anti-σ et le gène *mucD* qui code pour une protéase. L'opéron est sous le contrôle de σ^E. De plus, σ^E est nécessaire pour la survie de *Xcc* en phase stationnaire, pour la résistance au cadmium et l'adaptation aux températures élevées. Des analyses microarrays menées au cours de ces travaux ont permis d'élucider le régulon de σ^E. Ce régulon comprend des protéines membranaires, des protéines du métabolisme, des protéines effectrices et de structure du SST3, et le facteur σ^H impliqué dans la réponse à la chaleur. Ces analyses ont permis de déterminer un motif consensus de liaison de σ^E, $GGAACTN_{15\text{-}17}GTCNNA$. Enfin, cette étude montre que l'activité de σ^E est régulée par un mécanisme RIP (Regulated Intramembrane Proteolysis) impliquant les protéases RseP et DegS.

Au cours de ce travail, j'ai réalisé les expériences de microarrays et l'ensemble des données sont rassemblées dans un article publié en 2011 dans la revue Journal of Bacteriology (Bordes *et al.*, 2011).

JOURNAL OF BACTERIOLOGY, Jan. 2011, p. 246–264 Vol. 193, No. 1
0021-9193/11/$12.00 doi:10.1128/JB.00884-10

Insights into the Extracytoplasmic Stress Response of *Xanthomonas campestris* pv. *campestris*: Role and Regulation of σ^E-Dependent Activity▽‡

Patricia Bordes,[1,2]* Laure Lavatine,[1,2] Kounthéa Phok,[1,2] Roland Barriot,[1,2] Alice Boulanger,[3,4]† Marie-Pierre Castanié-Cornet,[1,2] Guillaume Déjean,[3,4] Emmanuelle Lauber,[4] Anke Becker,[5] Matthieu Arlat,[3,4] and Claude Gutierrez[1,2]

Université de Toulouse, UPS, Laboratoire de Microbiologie et Génétique Moléculaires, F-31000 Toulouse, France[1]; Centre National de la Recherche Scientifique, LMGM, F-31000 Toulouse, France[2]; Université de Toulouse, UPS, F-31000 Toulouse, France[3]; Laboratoire des Interactions Plantes Micro-Organismes, UMR CNRS-INRA 2594/441, F-31326 Castanet-Tolosan, France[4]; and Institute of Biology III, Faculty of Biology, University of Freiburg, D-79104 Freiburg, Germany[5]

Received 28 July 2010/Accepted 15 October 2010

***Xanthomonas campestris* pv. *campestris* is an epiphytic bacterium that can become a vascular pathogen responsible for black rot disease of crucifers. To adapt gene expression in response to ever-changing habitats, phytopathogenic bacteria have evolved signal transduction regulatory pathways, such as extracytoplasmic function (ECF) σ factors. The alternative sigma factor σ^E, encoded by *rpoE*, is crucial for envelope stress response and plays a role in the pathogenicity of many bacterial species. Here, we combine different approaches to investigate the role and mechanism of σ^E-dependent activation in *X. campestris* pv. *campestris*. We show that the *rpoE* gene is organized as a single transcription unit with the anti-σ gene *rseA* and the protease gene *mucD* and that *rpoE* transcription is autoregulated. *rseA* and *mucD* transcription is also controlled by a highly conserved σ^E-dependent promoter within the σ^E gene sequence. The σ^E-mediated stress response is required for stationary-phase survival, resistance to cadmium, and adaptation to membrane-perturbing stresses (elevated temperature and ethanol). Using microarray technology, we started to define the σ^E regulon of *X. campestris* pv. *campestris*. These genes encode proteins belonging to different classes, including periplasmic or membrane proteins, biosynthetic enzymes, classical heat shock proteins, and the heat stress σ factor σ^H. The consensus sequence for the predicted σ^E-regulated promoter elements is $GGAACTN_{15\text{-}17}GTCNNA$. Determination of the *rpoH* transcription start site revealed that *rpoH* was directly regulated by σ^E under both normal and heat stress conditions. Finally, σ^E activity is regulated by the putative regulated intramembrane proteolysis (RIP) proteases RseP and DegS, as previously described in many other bacteria. However, our data suggest that RseP and DegS are not only dedicated to RseA cleavage and that the proteolytic cascade of RseA could involve other proteases.**

Bacteria often encounter diverse and rapidly changing environments. To overcome harmful situations, they must be capable of sensing external changes and transmitting this information across biological membranes into the cell, which results in the appropriate redirection of gene expression to prevent or repair cellular damages caused by stress. Extracytoplasmic function (ECF) σ factors provide one common means of bacterial signal transduction to regulate gene expression in response to various extracellular changes (65). ECF σ factors represent the largest and most diverse subfamily of σ^{70} proteins. They generally recognize a −35 box with a clear bias toward a GAAC in their target promoters, while the −10 region tends to be highly variable between ECF subfamily members (65). One of the best-studied ECF σ factors is the key regulator of the extracytoplasmic stress response factor σ^E from *Escherichia coli*, encoded by the *rpoE* gene (56). ECF proteins were recently divided into 43 major phylogenetically distinct groups named ECF01 to ECF43 (65). RpoE-like ECF σ factors are part of one predominant subgroup found in most bacterial phyla and comprise ECF01 to -04 proteins. RpoE-like ECF σ factors are autoregulated and are required for a wide range of functions. For instance, the *E. coli* σ^E factor is essential for growth and promotes the expression of factors that help to preserve and/or restore cell envelope integrity (2). *Salmonella enterica* serovar Typhimurium σ^E is required for protection against reactive oxygen species and antimicrobial peptides and for stationary-phase survival (20, 67). *Bacillus subtilis* σ^W seems to constitute an antibiosis regulon acting against cell envelope stress (34). *S.* Typhimurium σ^E, *Pseudomonas aeruginosa* AlgU, and *Vibrio cholerae* σ^E are required for virulence (5). ECFs can thus be considered models to understand how bacteria sense and respond to their environment both during their interaction with their host and in their free-living state.

* Corresponding author. Mailing address: LMGM, CNRS, Université Paul Sabatier, Bât. IBCG, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France. Phone: 33 561335941. Fax: 33 561335886. E-mail: Patricia.Bordes@ibcg.biotoul.fr.

† Present address: NIDDK, National Institutes of Health, Building 8A, Room 2A21, 8 Center Dr., Bethesda, MD 20892.

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

▽ Published ahead of print on 22 October 2010.

RpoE-like ECF σ factors are tightly regulated in order to coordinate their activation with the appropriate environmental cues. In most cases, the σ^E factor is cotranscribed with a cognate transmembrane anti-σ factor possessing an extracytoplasmic domain and an intracellular σ-binding domain. In the absence of stimulus, the membrane-bound anti-σ binds tightly to the σ factor, thereby keeping it inactive (33). Upon receiving a proper signal, the anti-σ factor is inactivated by regulated intramembrane proteolysis (RIP), resulting in the release and subsequent activation of the σ^E factor. This mechanism has been well studied for the anti-σ factors RseA, MucA, and RsiW, regulating the activity of *E. coli* σ^E, *P. aeruginosa* AlgU, or *B. subtilis* σ^W, respectively (1, 32, 75). In *E. coli*, the accumulation of C-terminal domains of unfolded porins is the activating signal of the RpoE response by triggering the activation of the inner-membrane-anchored protease, DegS (site 1 protease), and the subsequent cleavage of RseA within its periplasmic domain by DegS. The resulting truncated anti-σ factor is then a suitable substrate for a second inner-membrane protease, RseP/YaeL (site 2 protease), which cleaves RseA near the cytoplasmic face of the inner membrane, releasing an $RseA_{cyto}$-σ^E complex into the cytoplasm, where the remaining RseA fragment is degraded by cytoplasmic proteases, resulting in the active σ^E (1). Another important mediator of the extracytoplasmic stress response is the periplasmic protease DegP, also known as HtrA and DO in *E. coli* or MucD in *P. aeruginosa* (22, 55). DegP binds to and degrades misfolded proteins and acts as a chaperone to direct the proper folding of some envelope proteins (66). As such, this family of proteases regulates the σ^E stress response system by removing misfolded proteins in the periplasm that could activate the degradation pathway of the anti-σ^E factor, even in the absence of stress (27).

The Gram-negative phytopathogenic bacterium *Xanthomonas campestris* pv. *campestris* is an epiphytic bacterium that can become a vascular pathogen, causing black rot disease of crucifers (52). The bacterium produces a large amount of extracellular polysaccharide (EPS) that plays an important role during bacterial infection, and *X. campestris* pv. *campestris* has been used as a model organism for investigating the mechanism of bacterial pathogenesis. *X. campestris* pv. *campestris* flourishes in and adapts to a wide range of habitats: during epiphytic life, *X. campestris* pv. *campestris* is exposed to harsh stresses, such as oligotrophic conditions, desiccation, or large changes in temperature. Upon entry into plant tissues, *X. campestris* pv. *campestris* cells must face defense reactions of the host, including oxidative conditions. Finally, the natural life cycle of *X. campestris* pv. *campestris* includes long periods of survival on seeds or plant scraps or in the soil, where again it must survive a variety of stressful conditions before it can infect a new host plant. Its ability to manage variable and often lethal external conditions can be partly attributed to its large repertoire of alternative σ factors. Of the 4,179 open reading frames (ORFs) comprising the large 5.1-Mb *X. campestris* pv. *campestris* strain ATCC 33913 genome, 15 ORFs encode characterized or putative σ factors, 10 of which belong to the ECF subfamily (23). Little is known about which σ factors are required for the survival of *X. campestris* pv. *campestris* under stress and the contribution of these factors to virulence. The classification of ECF σ factors strongly suggested that the *XCC1267* gene encodes the σ^E factor of *X. campestris* pv. *campestris* (65). Moreover, previous work by Cheng et al. (17) described the biochemical characterization of the σ^E factor of *X. campestris* pv. *campestris* strain 11 and suggested that it could have a role in the heat shock response. Therefore, we aimed at deciphering the roles and regulation mechanisms of the extracytoplasmic stress response regulator σ^E in *X. campestris* pv. *campestris*.

In the present work we characterized the *rpoE* operon genes, *rpoE*, *rseA*, and *mucD*. Using primer extension and *lacZ* transcriptional reporter fusions, we show that *rpoE* transcription is autoregulated and that RseA and MucD are negative regulators of σ^E activity. We identified 45 putative members of the σ^E regulon by a transcriptome analysis, including the heat stress σ factor σ^H and a number of periplasmic or membrane proteins. We provided evidence that σ^E is an important regulator of stress responses in *X. campestris* pv. *campestris*, since it has a role in heat adaptation, resistance to cadmium, and stationary-phase survival. Furthermore, our results strongly suggest that σ^E is regulated by a RIP mechanism involving RseP (XCC1366) and DegS (XCC3898) putative proteases, as in many other bacteria. However, our data suggest that the RIP proteases RseP and DegS are not only dedicated to RseA cleavage and that the proteolytic cascade of RseA could involve other proteases.

MATERIALS AND METHODS

Strains and growth conditions. The *X. campestris* pv. *campestris* strains, plasmids, and oligonucleotides used or generated in this study are listed in Table 1. *X. campestris* pv. *campestris* cells were grown at 30°C in MOKA (yeast extract, 4 g/liter; Casamino Acids, 8 g/liter; K_2HPO_4, 2 g/liter; $MgSO_4 \cdot 7H_2O$, 0.3 g/liter) (9) or in KADO (MOKA plus 1% sucrose) (39) medium. *E. coli* cells were grown at 37°C in LB medium (44). Antibiotics were used at the following concentrations for *X. campestris* pv. *campestris*: rifampin, 50 μg/ml; kanamycin, 50 μg/ml; tetracycline, 5 μg/ml. Antibiotics were used at the following concentrations for *E. coli*: ampicillin, 50 μg/ml; kanamycin, 50 μg/ml; tetracycline, 10 μg/ml. The following supplements were added when required: sucrose, 1%, and isopropyl-β-D-thiogalactopyranoside (IPTG), 0.5 mM.

Recombinant DNA procedures. Genomic DNA from *X. campestris* pv. *campestris* was extracted using the DNeasy Blood and Tissue kit according to the instructions of the manufacturer (Qiagen). Plasmid DNA and PCR products were purified with the Qiagen mini-plasmid purification kits and PCR purification kits, respectively. *E. coli* strain DH5α was used for cloning. Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase, and Phusion High-Fidelity DNA polymerase were used as specified by the manufacturer (New England Biolabs).

Construction of in-frame deletion mutant strains in *X. campestris* pv. *campestris*. The two-step recombination system (59), based on the inability of *X. campestris* pv. *campestris* carrying the *sacB* gene to grow in media with high sucrose concentrations, was used for the chromosomal inactivation of the *rpoE*, *rseA*, *mucD*, *rseP*, *degS*, and *XCC1664* genes of *X. campestris* pv. *campestris*. For each planned inactivation experiment, a mobilizable *X. campestris* pv. *campestris* integration vector was constructed, which contained two 1,000-bp fragments on each side of the gene to be deleted, comprising the first and the last 18 nucleotides of the selected gene, thus providing two homology regions for recombination.

Genomic DNA of *X. campestris* pv. *campestris* strain 568 and primer pairs rpoE F-Sma/rpoE R-Xba and rseA F-Xba/rpoE R-Hind were used to obtain two PCR products of ~1 kb from the region upstream of *rpoE* and downstream of *rpoE*, respectively. The products were sequentially cloned into the appropriate sites of the pK18*mobsacB* vector, starting with the upstream region, to finally obtain the mobilizable plasmid pK18-$rpoE_{U+D}$. The same cloning procedure was used for the other genes, using the primer pairs indicated in Table 1. The resulting plasmids were verified by sequencing from both ends with standard primers. The plasmids were transformed into *E. coli* DH5α and mobilized into *X. campestris* pv. *campestris* by conjugation, as described previously (68). Successful first re-

TABLE 1. *X. campestris* pv. *campestris* strains, plasmids, and oligonucleotides used or generated in this study

Strain, plasmid, or oligonucleotide	Sequence (5′ to 3′)	Features or purpose	Source or reference
Plasmids			
pVO155		pUC119 derivative containing the promoterless *gus* (*uidA*) reporter gene encoding β-glucuronidase; used for insertion mutagenesis; Km^r Amp^r	51
pFAJ1700		pTR102-derived expression vector containing a multiple-cloning site and transcriptional terminators in both orientations; Tet^r Amp^r	25
pCZ750		pFAJ1700 containing the KpnI-AscI *lacZ* gene from the pCZ367 plasmid; Tet^r Amp^r	9
p917		pFAJ1700 derivative containing 2,094 bp of pSC150 with *lacI*, *tacp* promoter, and T7 terminator; Tet^r Amp^r	11
p917-lacI		p917 derivative containing 1,526 bp of pMF533 with *lacI*, *tacp* promoter, ribosome binding site, and T7 terminator; used for protein overexpression; Tet^r Amp^r	This study
pCZ-rpoEp		pCZ750 derivative; *rpoEp-lacZ*; Tet^r Amp^r	This study
pCZ-rseAp		pCZ750 derivative; *rseAp-lacZ*; Tet^r Amp^r	This study
pCZ-rpoHp		pCZ750 derivative; *rpoHp-lacZ*; Tet^r Amp^r	This study
pCZ-prcp		pCZ750 derivative; *prcp-lacZ*; Tet^r Amp^r	This study
pCZ-ompWp		pCZ750 derivative; *ompWp-lacZ*; Tet^r Amp^r	This study
pCZ-hrpFp		pCZ750 derivative; *hrpFp-lacZ*; Tet^r Amp^r	This study
pCZ-xcc0401p		pCZ750 derivative; *xcc0401p-lacZ*; Tet^r Amp^r	This study
p917-rpoE		p917-lacI derivative; *tacp-rpoE*; Tet^r Amp^r	This study
pMF533		pMALc2E (New England Biolabs) containing *malE-gadE* fusion	16
pK18*mobsacB*		Mobilizable cloning vector containing a modified *sacB* gene from *B. subtilis*; used for gene disruption; Kan^r	61
pK18-$rpoE_U$		pK18 containing ~1 kb upstream of the *rpoE* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$rpoE_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of the *rpoE* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$rseA_U$		pK18 containing ~1 kb upstream of the *rseA* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$rseA_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of the *rseA* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$mucD_U$		pK18 containing ~1 kb upstream of the *mucD* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$mucD_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of the *mucD* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$rpoE_U$$rseA_D$		pK18 containing ~1 kb upstream of the *rpoE* gene (up to 15 nucleotides downstream of the ATG start codon) and ~1 kb downstream of the *rseA* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$rpoE_U$$mucD_D$		pK18 containing ~1 kb upstream of the *rpoE* gene (up to 15 nucleotides downstream of the ATG start codon) and ~1 kb downstream of the *rseA* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$degS_U$		pK18 containing ~1 kb upstream of the *degS* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$degS_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of of the *degS* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$rseP_U$		pK18 containing ~1 kb upstream of the *rseP* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$rseP_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of the *rseP* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study
pK18-$xcc1664_U$		pK18 containing ~1 kb upstream of the *XCC1664* gene of *X. campestris* pv. *campestris* (up to 15 nucleotides downstream of the ATG start codon)	This study
pK18-$xcc1664_{U+D}$		pK18 containing ~1 kb upstream and ~1 kb downstream of the *XCC1664* gene of *X. campestris* pv. *campestris* (from 15 nucleotides upstream of the stop codon)	This study

Continued on following page

TABLE 1—*Continued*

Strain, plasmid, or oligonucleotide	Sequence (5′ to 3′)	Features or purpose	Source or reference
Strains			
E. coli			
DH5α		λ− ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(r_K^- m_K^-) *supE44 thi-1 gyrA relA1*	Laboratory collection
X. campestris pv. *campestris*			
568		Wild-type strain; rifampin-resistant derivative of *X. campestris* pv. *campestris* LMG568/ATCC33913	9
XcPB1		Δ*rpoE*; Rif^r	This study
XcPB2		Δ*rseA*; Rif^r	This study
XcPB3		Δ*mucD*; Rif^r	This study
XcPB4		Δ*rpoE-rseA*; Rif^r	This study
XcPB5		Δ*rpoE-rseA-mucD*; Rif^r	This study
XcPB6		Δ*xcc1664*; Rif^r	This study
XcPB7		Δ*degS*; Rif^r	This study
XcPB8		Δ*rseP*; Rif^r	This study
rpoE::pVO		*XCC1267*::pVO155; Rif^r Km^r	This study
Oligonucleotides[a]			
rpoE F-Sma	TTTCCCGGGCATACGGCGCGGGATAGTGTTC	*rpoE* gene deletion	
rpoE R-Xba	TTTTCTAGAGACTTCGGCCATGTCGGG		
rpoE F-Xba	TTTTCTAGACACCGTGTATGACCAATAACC		
rpoE R-Hind	TTTAAGCTTGGCCGAACATCTGGGTGCGG		
rseA F-Sma	TTTCCCGGGTGACATCGCCCAGTTCGAG	*rseA* gene deletion	
rseA R-Xba	TTTTCTAGAGTTATTGGTCATACACGGTG		
rseA F-Xba	TTTTCTAGACCGCAGGACTGATGTTCTCGCC		
rseA R-Hind	TTTAAGCTTGAAAGCCAGGTCGATCGGGATC		
mucD F-Eco (B)	TTTGAATTCGAAAGCGCTACCCGTGAGCGAC	*mucD* gene deletion	
mucD R-Xba	TTTTCTAGAGCGGGGATTCATCAGGTTG		
mucD F-Xba	TTTTCTAGAGCGGGCGGCTGAGACGCAGGG		
mucD R-Hind	TTTAAGCTTGGGTGTCGCCGACCGGCGCGCC		
rseP F-Sma	TTTCCCGGGACCCAGGCGCATGCCGGTGATC	*rseP* gene deletion	
rseP R-Xba	TTTTCTAGAGAAATCACCCATGGATGCAAC		
rseP F-Xba	TTTTCTAGAGTTCCACGATGAAGCTGCTCC		
rseP R-Hind	TTTAAGCTTGAAGATGTCGCCGGCCTTGGGG		
degS F-Eco	TTTGAATTCAACGCTGTTCTTGGCCACCAC	*degS* gene deletion	
degS R-Xba	TTTTCTAGACAGCGGTCGCATGCAACGGATTC		
degS F-Xba	TTTTCTAGACTCATGCGTTGATCCGGCGTG		
degS R-Hind	TTTAAGCTTCATGGCGCCGAATTTCATGGG		
xcc1664 F-Eco	TTTGAATTCGCCGCCCAGATCGGGCGTG	*XCC1664* gene deletion	
xcc1664 R-Bam	TTTGGATCCCCCACCGGGCACTGCATGATTTC		
xcc1664 F-Bam	TTTGGATCCGCGCCATGACCACGCCCTG		
xcc1664 R-Hind	TTTAAGCTTGGAAAGCCATCCAGGCGC		
rpoE RT-F (A)	GCGGCGTTCGATGTGTTGG		
rpoE-EXT (I)	CCAGCTCCTGAGGTGTATC	Mapping of *rpoEp*	
mucD RT-F (D)	CCATCACCCGCAAGGACGCC		
mucD-EXT (G)	GCCAACGGCAGGGTCATGG		
mucD(2)-EXT (K)	GAACATCTGGGTGCGGATG		
mucD(3)-EXT (H)	GGAAATGATGAAGCCCGAAC		
rseA-EXT (E)	CTTTCGGTCTCCAGCAGAGG		
rseA(2)-EXT (J)	GGACATGTCAGGGTTATTG	Mapping of *rseAp*	
rseA(3)-EXT (F)	GTTCGCGGGACACGAACAAG		
rseA RT-F (C)	GACGAAGAGTTGGCCGGCTG		
lacI-FseI	TTTGGCCGGCCATCGAATGGTGC		
PtaqNcoBam	TTGGATCCATGGGCTATGGTCCTTGTTG		
P1268-Hind	TTTAAGCTTCGGTGGCAGACAGG	Upstream region of *rseA* (fusion to *lacZ*)	
P1268-Xba	TTTCTAGATGGCCGACCGGAGTTC		
P1267-Hind	TTTAAGCTTGGGGCAGGGCAGCTCGG	Upstream region of *rpoE* (fusion to *lacZ*)	
P1267-Xba	TTTTCTAGAAGTCGGGCAATGAGACC		

Continued on following page

TABLE 1—*Continued*

Strain, plasmid, or oligonucleotide	Sequence (5′ to 3′)	Features or purpose	Source or reference
ORF1267-Nco	TTTCCATGGCCGAAGTCGATACACC	*rpoE* gene overexpression	
ORF1267-Hind	TTTAAGCTTCATACACGGTGTCGCTGAC		
xcc1535-EXT	GTTCACCACGTCCGGTGAGCTC	Mapping of *XCC1535p*	
xcc0964-EXT	CTCGGCCAGCATAATCTTGCC	Mapping of *XCC0964p*	
hrpF-EXT	CCTCGCAGTGACAGAGCAG	Mapping of *hrpFp*	
hpa1-EXT	CTGCAGGTTGATGAAGTTGG	Mapping of *hpaIp*	
ompW-EXT	GGAAATGGAACGCATGAGGG	Mapping of *ompWp*	
pqqB-EXT	CCGATCCCAAAACGATGATG	Mapping of *pqqAp*	
xcc0401-EXT	GTCGGTCTGGATCTGGATCTG	Mapping of *XCC0401p*	
prc-EXT	CCAGCAATGCCATCGGAGTG	Mapping of *prcp*	
xcc4186-EXT	GAAGCTGCGTACTGCGTTGCAG	Mapping of *XCC4186p*	
rpoH-EXT	GATTGTTTGCCACAAGGGCAGTC	Mapping of *rpoHp*	
xcc3227-EXT	GGAAGGAGCGGCGGTGCGTTC	Mapping of *XCC3227p*	
xcc1246-EXT	GGCAAGTCTGATCCTCTTGG	Mapping of *XCC1246p*	
P0539-Hind	TTTAAGCTTGGCCGGGCTGGTCGAGGG	Upstream region of *ompW* (fusion to *lacZ*)	
P0539-Xba	TTTTCTAGATTGTTAGGCCCTTGGAGTGG		
PhrpF-Hind	TTTAAGCTTCGGGGTCCAATCCAAGCC	Upstream region of *hrpF* (fusion to *lacZ*)	
PhrpF-Xba	TTTCTAGAGCAGTGGCGGAGCTG		
Pprc-Hind	TTTAAGCTTGTCACCCTGCGCGACCTG	Upstream region of *prc* (fusion to *lacZ*)	
Pprc-Xba	TTTTCTAGATGAAAGAAGGGCGGTGATC		
P0401-Hind	TTTAAGCTTCGTCCTGCGCCACCGCAG	Upstream region of *XCC0401* (fusion to *lacZ*)	
P0401-Xba	TTTTCTAGATTCGATTCTGCGACAGCC		
P3771-Hind	TTTAAGCTTCAATGCGGTGCTGACGGTGG	Upstream region of *rpoH* (fusion to *lacZ*)	
P3771-Xba	TTTTCTAGAACTATTGGACTGCTGGGTAC		
P1230-Hind	TTTAAGCTTCGGCATCGGCGTCCTCTTC	Upstream region of *XCC1230* (fusion to *lacZ*)	
P1230-Xba	TTTTCTAGATTGCTGCACCCCCATTCTG		

[a]Letters in parentheses after oligonucleotide designations indicate letter codes used for primers in Fig. 1 and 2 and Fig. S1 in the supplemental material.

combinants (chromosomal integration mutants) were selected by plating cells on MOKA containing kanamycin. For selection of the second recombination event, the integration mutants were plated on MOKA containing 5% (wt/vol) sucrose. Clones that had lost kanamycin resistance were screened by colony PCR using primers outside the cloned regions. Two of the confirmed transconjugants were randomly chosen for further study.

Construction of *rpoE*-overexpressing *X. campestris* pv. *campestris* strains for complementation. To construct the overexpression plasmid for the *rpoE* gene, an FseI-BamHI fragment encompassing the *lacI* gene and the *tacp* promoter was amplified by PCR using plasmid pMF533 (16) as a template, together with primers lacI-FseI and *tacp* NcoBam. The fragment generated was subsequently cloned into the pCZ917 plasmid (11) to yield p917-lacI. The resulting plasmid contains the *lacI* gene and the *tacp* promoter, followed by a ribosome binding site and the restriction sites for NcoI and BamHI. For complementation of deletion mutants of *X. campestris* pv. *campestris*, a 621-bp DNA fragment containing the entire *rpoE* gene was amplified by PCR using total DNA of *X. campestris* pv. *campestris* strain 568 as a template and primers ORF1267-Nco and ORF1267-Hind and cloned into the p917-lacI plasmid using the NcoI and HindIII enzymes. The resulting plasmid, p917-rpoE, was verified by sequencing and subsequently transferred to *X. campestris* pv. *campestris* by triparental conjugation. For the complementation experiments, different concentrations of IPTG were tested, and in order to avoid deleterious effects of σ^E overexpression, we chose not to add IPTG because the expression of σ^E from the leaky *tacp* promoter was sufficient.

Construction of promoter-reporting plasmids and β-galactosidase assays. The promoter regions (~500 bp upstream of the ATG) of genes *rpoE*, *rseA*, *prc*, *ompW*, *hrpF*, *XCC0401*, *rpoH*, and *XCC1230* were PCR amplified with primer sets shown in Table 1, using genomic DNA from *X. campestris* pv. *campestris* strain 568 as a template. These promoter regions were cloned as HindIII-XbaI fragments into the pCZ750 plasmid (9) upstream of a promoterless *lacZ* gene. The resulting plasmids were confirmed by sequencing and introduced into *X. campestris* pv. *campestris* strains by triparental conjugation. Overnight cultures were diluted to an optical density at 600 nm (OD_{600}) of 0.05 and grown at 30°C for 9 h (~3 generations). β-Galactosidase activities were assayed as described previously (44). The data shown are the averages of at least two independent cultures, each measured in triplicate. Negligible β-galactosidase activity was derived from the control promoterless plasmid pCZ750 (less than 9 Miller units) (data not shown).

RNA purification. Overnight cultures were diluted to an OD_{600} of 0.05 and grown at 30°C for 9 h (~3 generations) to reach an OD_{600} of 0.4. If the cells required a temperature upshift, it was done for 60 min at 35°C. For all experiments, 10-ml culture samples were harvested by immediately adding 1.25 ml of ice-cold 5% water-saturated phenol in ethanol and centrifugation at 5,000 rpm. The cell pellets were flash frozen in liquid N_2 and stored at −80°C. Total RNA was extracted from the cell pellets by use of TRIzol reagent following the manufacturer's specifications (Invitrogen). A further treatment with 0.03 U RQ1 DNase I (Promega) per μg of RNA for 30 min at 37°C, followed by phenol extraction and ethanol precipitation, was carried out. RNA was evaluated for quantity and quality using the NanoDrop1000 spectrophotometer (NanoDrop Technologies) and agarose gel electrophoresis. The absence of DNA contamination was confirmed by PCR.

RT-PCR. One microgram of total RNA was reverse transcribed into cDNA by using PrimeScript reverse transcriptase (RT) (Takara Bio) with random hexamers according to the manufacturer's instructions, and 0.5 μl of each retrotranscription reaction mixture was subjected to PCR using GoTaq DNA polymerase (Promega). Positive controls were performed with genomic DNA, and negative controls were performed with RNA that had not been subjected to retrotranscription (data not shown). The cycling conditions used were 95°C for 2 min and 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min 30 s, followed by incubation at 72°C for 5 min.

Microarray data collection and statistical analyses. Briefly, total RNA was extracted from four sets of identically treated batch cultures of Δ*rpoE* (no σ^E) and Δ*rseA* (overactivation of σ^E) mutant *X. campestris* pv. *campestris* strains grown up to mid-exponential phase in MOKA medium at 30°C. Total RNA (10 μg) was reverse transcribed in the presence of aminoallyl dUTP using Superscript II (Invitrogen) and random hexamers for priming, according to the manufacturer's instructions. The resulting amine-modified cDNA was then chemically labeled at the aminoallyl group using Alexa Fluor 555 and 647 reactive dyes (Invitrogen). Hybridization took place underneath a coverslip in 60 μl of warmed DIG Easy Hybridization buffer (Roche) at 42°C for 16 h in a sealed humidified chamber.

Xcc5kOLI microarrays (62) based on the genome sequence of *X. campestris* pv. *campestris* strain B100 (70) were used for hybridizations. The array contains 4,441 50-mer to 70-mer oligonucleotides representing the predicted protein-encoding genes. In addition, it contains 15 stringency controls of the genes *gapA*, *rpsA*, *rpsB*, *rpsL*, and *rpmI* (70%, 80%, and 90% identity to the native sequence), 12 alien DNA oligonucleotides, and 5 spiking control oligonucleotides. Each probe was spotted in three replicates.

After being washed, the hybridized microarray slides were scanned using an Axon GenePix 4100A scanner (Molecular Devices). The acquired microarray images were analyzed with GenePix Pro (version 3.0.6.90). Preprocessing of raw data and statistical analyses were performed using the Bioconductor and

LIMMA packages in the R programming environment (64). Spots marked as "bad" (flags ≤ −49) by GenePix were excluded from further analysis. Intensity normalization was performed using the robustspline method (within arrays) and the aquantile method (between arrays). Within array replicates, correlation was estimated and incorporated in the linear model was inferred by LIMMA, which was further used to build an empirical Bayes moderated *t* test statistic to assess differential expression. *P* values were adjusted for multiple testing using the method of Benjamini and Hochberg (8) to control the false-discovery rate at a level of 0.05, and the statistical significance threshold to decide differential expression was set to 0.05.

Primer extension analyses. For these experiments, we used one of the four RNA sample sets that were used for the microarray experiments. Primer extension experiments were performed at 50°C using PrimeScript Reverse Transcriptase (Takara Bio) and primers hybridizing in the 5′ region of the coding sequence of the respective genes (Table 1). Ten micrograms of total RNA and 1 pmol of ^{32}P end-labeled primers were used per reaction. The extension products were loaded on a 6% denaturing polyacrylamide gel adjacent to a sequencing ladder obtained with the ^{32}P-labeled universal cycle primer and pUC18 plasmid provided in the Thermo Sequenase cycle-sequencing kit (USB).

Microarray data accession numbers. Fully annotated microarray data from this study have been deposited in ArrayExpress under accession no. E-MEXP-2935. The array design is available in ArrayExpress under accession no. A-MEXP-1909.

RESULTS

***rpoE* genomic organization.** The *rpoE* (*XCC1267*) genomic organization found in the *X. campestris* pv. *campestris* genome is similar to that of other gammaproteobacteria, such as *E. coli*, *P. aeruginosa*, and *Xanthomonas fastidiosa* (Fig. 1A). The gene immediately downstream of *rpoE* is predicted to encode an alanine-rich protein of 286 amino acids that contains an N-terminal anti-σ^E protein RseA domain (pfam03872; RseA_N). Taken together with the fact that the genes encoding the σ^E factors are contiguous to a coding region specifying an anti-σ factor, this strongly suggested that XCC1268 was the putative anti-σ^E factor, and *XCC1268* has been renamed *rseA*. The third gene of the *rpoE* cluster is the *mucD* orthologue *XCC1269*. MucD is a periplasmic serine protease of *P. aeruginosa* that belongs to the HtrA protein family and alleviates periplasmic stress by degrading misfolded outer membrane proteins (OMPs) selectively (75). Hence, MucD is an indirect negative regulator of σ^E activity. Interestingly, *P. aeruginosa* and *X. fastidiosa* carry a *mucD* gene in the *rpoE* operon, but *E. coli* does not (Fig. 1A). In other bacteria, there may be one or two genes downstream of *rseA* (namely, *rseB-mucB* and *rseC-mucC*), which encode accessory σ^E regulatory proteins (Fig. 1A). RseB is a negative regulator of σ^E that binds to the periplasmic domain of RseA, probably in order to block the access of DegS to the cleavage site of RseA (40). The function of RseC is generally unknown, but it has been reported to be both a positive and a negative regulator (10, 46). Strikingly, there are no orthologues of *rseB* and *rseC* in the *X. campestris* pv. *campestris* genome (Fig. 1A), suggesting that σ^E activity might be regulated only by RseA binding in this bacterium.

To investigate whether *rpoE-rseA-mucD* constituted a single transcription unit, RT-PCR experiments were carried out using appropriate primers located within the *rpoE-rseA-mucD* region. As shown in Fig. 1B, we could identify specific transcripts encompassing *rpoE* and *rseA* (Fig. 1B, primer pairs AF and BF), *rseA* and *mucD* (Fig. 1B, primer pair CG), and *rpoE* and *mucD* (Fig. 1B, primer pairs AG and BG). These data strongly suggest that *rpoE-rseA-mucD* form a single transcription unit.

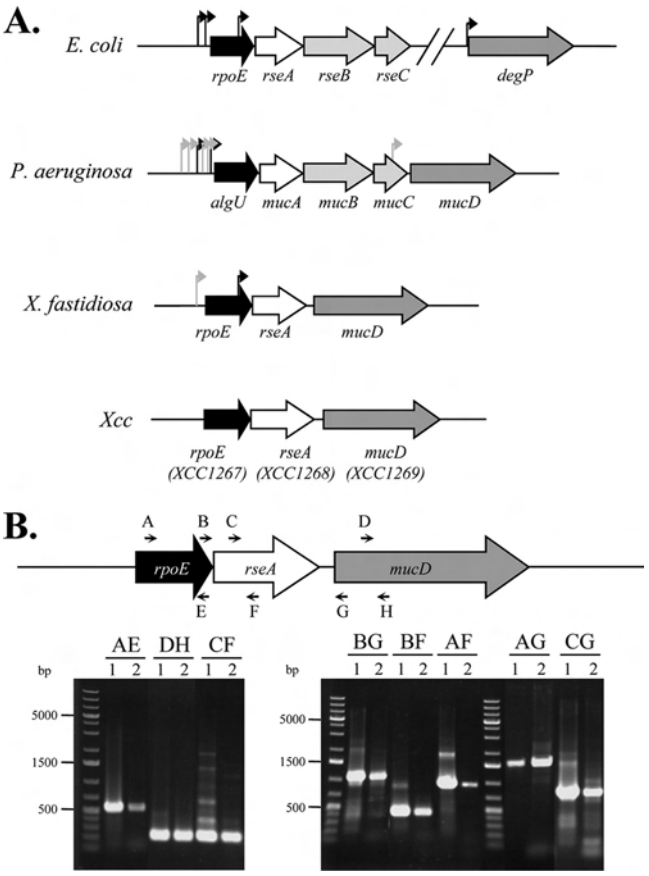


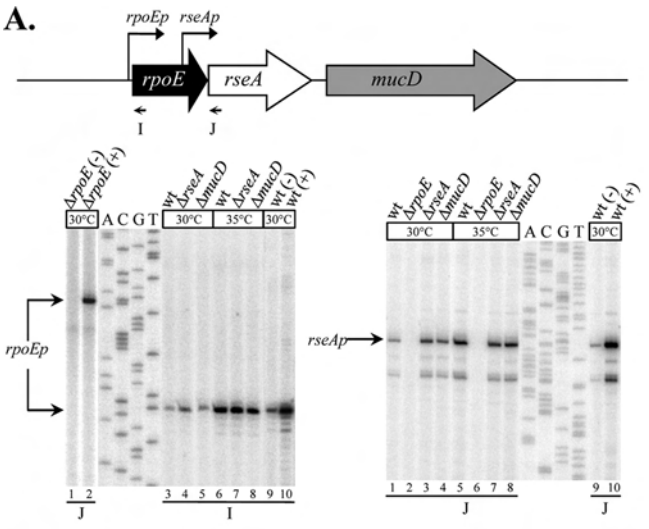
FIG. 1. Transcriptional organization of the *X. campestris* pv. *campestris rpoE* region. (A) Schematic (to scale) showing the organization of the *rpoE* region in *E. coli*, *P. aeruginosa*, *X. fastidiosa*, and *X. campestris* pv. *campestris*. The arrows indicate experimentally demonstrated promoter start sites (black, σ^E-dependent promoter; gray, σ^E-independent promoter). (B) Organization of the *rpoE* region showing the primer pairs used for the amplifications (Table 1 shows primer sequences) and agarose gel of the RT-PCR amplification products. For each primer pair (named according to the letter code of each primer), two lanes are shown (lane 1, positive control using genomic DNA templates; lane 2, RT-PCR using RNA extracted from cells in exponential phase). The molecular size marker is the O'GeneRuler 1-kb Plus DNA Ladder (Fermentas). In each case, the main extension product migrated at the expected size: AE, 524 bp; DF, 228 bp; CF, 235 bp; BG, 1,058 bp; BF, 393 bp; AF, 912 bp; CG, 900 bp; and AG, 1,577 bp.

Transcriptional regulation of the *rpoE* operon. A typical σ^E-dependent promoter has been identified upstream of the *rpoE* gene in *X. campestris* pv. *campestris* strain 11 (Fig. 2B) (17). Since this promoter sequence is conserved in *X. campestris* pv. *campestris*, we postulated that *rpoE* was autoregulated in *X. campestris* pv. *campestris*. A single transcription start site upstream of the *rpoE* gene (*rpoEp*) was identified by primer extension (Fig. 2A, left, lane 3), and it matched P1 of *X. campestris* pv. *campestris* strain 11 (Fig. 2B). To test the transcriptional regulation of *rpoEp*, we conducted primer extension experiments using RNAs purified from strains with *rpoE*, *rseA*, or *mucD* deleted (chromosomal unmarked in-frame deletions; see Materials and Methods for details). The growth of the mutants was comparable to that of the parental strain in rich

medium (MOKA) at 30°C, evidencing a doubling time of ~3 h (data not shown), suggesting that *rpoE* plays a nonessential role under ordinary growth conditions. As expected, *rpoEp* expression was strongly reduced when *rpoE* was inactivated (Fig. 2A, left, lane 1), and conversely *rpoEp* expression was increased when *rseA* or *mucD* was deleted (Fig. 2A, left, compare lanes 4 and 5 with lane 3). Note that the effect of *mucD* deletion is minor but reproducible. Further, the wild-type (WT) or Δ*rpoE* mutant strain was complemented with plasmid p917-rpoE, carrying the *rpoE* gene under the control of the IPTG-inducible promoter *tacp*. Overexpression of σ^E from plasmid p917-rpoE significantly increased the amount of *rpoEp* expression in the wild-type strain (Fig. 2A, left, lanes 9 and 10) and restored *rpoEp* activity, while no activity was seen when the control plasmid p917 was introduced into the Δ*rpoE* mutant strain (Fig. 2A, left, lanes 1 and 2).

These results were confirmed by the analysis of the β-galactosidase activity expressed from the reporter plasmid pCZ-rpoEp carrying a fusion between a DNA fragment containing *rpoEp* and the *lacZ* gene (see Materials and Methods). This plasmid was introduced into the *X. campestris* pv. *campestris* wild-type strain and the Δ*rpoE*, Δ*rseA*, Δ*mucD*, Δ*rpoE-rseA*, and Δ*rpoE-mucD* mutants. We observed that deletion of *rpoE* caused an ~4-fold reduction in *rpoEp*-driven β-galactosidase activity (Fig. 2C), showing that the expression of σ^E is autoregulated in *X. campestris* pv. *campestris*, as in many other bacteria and as previously reported in *X. campestris* pv. *campestris* strain 11 (17). However, significant β-galactosidase activity remained when *rpoE* was deleted (~256 Miller units) (Fig. 2C), suggesting that the *rpoE* operon could also be controlled by a σ^E-independent promoter. This is consistent with the regulation of *rpoE* transcription by a combination of σ^E-dependent and σ^E-independent promoters in other bacteria (Fig. 1). Nevertheless, we were unable to map such a promoter using primer extension, possibly because its expression was too low under our experimental conditions. The deletion of *rseA* or *mucD* caused ~5-fold and ~2-fold activation, respectively, of *rpoEp* expression (Fig. 2C), indicating that RseA and MucD are negative regulators of σ^E-dependent activity. The modest effect of *mucD* deletion on the transcriptional activation of *rpoEp* compared with the impact of *rseA* deletion (Fig. 2A, left, and C) was in agreement with the predicted role of MucD as an indirect negative regulator of σ^E that acts by removing misfolded proteins that activate proteases for degradation of anti-σ^E.

To check the effect of unfolding stress on *rpoEp* transcription, we tested the effect of temperature stress, since σ^E has been shown to be involved in cell survival after a heat shock stress in *X. campestris* pv. *campestris* strain 11 and the closely related bacterium *X. fastidiosa* (17, 24). More generally, σ^E is involved in the transcription of a set of heat shock response genes and the heat shock sigma factor σ^H in several bacterial species (3, 28, 69). We conducted primer extension experiments using RNAs extracted from *X. campestris* pv. *campestris* strains that had been shifted from 30°C to 35°C. As shown in Fig. 2A (left), there was a strong increase in *rpoEp* transcription upon heat treatment (compare lane 6 with lane 3). Similar transcription levels were obtained in WT, Δ*rseA*, and Δ*mucD* strains (compare lanes 7 and 8 with lane 6). These results showed that *rpoE* expression was induced by temperature



C.

Strain	β-galactosidase activity (Miller Units)	
	rpoEp	*rseAp*
WT	981±176	3082±992
Δ*rpoE*	256±75	36±6
Δ*rpoE-rseA*	244±95	33±9
Δ*rpoE-rseA-mucD*	225±17	29±4
Δ*rseA*	5193±303	16075±5952
Δ*mucD*	2236±584	5914±1317

FIG. 2. Expression of the *rpoE* operon genes. (A) Determination of the transcription start site of the *rpoE* operon genes by primer extension. The schematic (to scale) shows the primers used for transcriptional start site mapping experiments, and the black arrows indicate the positions of the two identified σ^E-dependent promoters. Total RNAs from the WT, Δ*rpoE*, Δ*rseA*, Δ*mucD*, and WT or Δ*rpoE* strains containing the control plasmid p917 (−) or the σ^E-overexpressing plasmid p917-rpoE (+) were used as templates in primer extension experiments when they were suited. Total RNAs were obtained from cells incubated at 30°C or after a 60-min shift at 35°C. The arrows indicate the bands corresponding to the observed start site. (B) Sequence alignment depicting the relevant features of *rpoEp* and *rseAp* compared to σ^E-binding sites of P1 of *X. campestris* pv. *campestris* strain 11 (Xc11) and P2 of *X. fastidiosa* (Xf) (17, 24). The transcription start sites identified by primer extension are indicated in boldface, and the putative −10 and −35 regions are boxed. The consensus ECF02 group of ECF σ factor sites is indicated below (65). (C) Determination of *rpoEp* and *rseAp* activities in different strains. Plasmids containing a transcriptional fusion of the upstream region of *rpoEp* or *rseAp* to the *lacZ* gene were transferred into *X. campestris* pv. *campestris* strains. Overnight cultures of these strains grown in MOKA medium were diluted in the same medium and grown for 9 h before determination of β-galactosidase activity. The results represent the mean values of at least two independent experiments, each performed in triplicate, with the standard errors.

stress. Moreover, these strongly suggest that RseA and MucD are negative regulators of σ^E in the absence of activating signal, which could be the presence of nonfolded proteins in the periplasm.

To go further into the regulation of the *rpoE* operon, we tested for the presence of other promoters using primer extension experiments with primers located at the 3′ end of *rpoE* and the 5′ end of *rseA* or *mucD*. As for *X. fastidiosa* and *E. coli*, a typical σ^E-dependent promoter was found within the *rpoE* coding sequence and was named *rseAp* (Fig. 2A, right, and B). Note that we used a primer further upstream to show that the upper band corresponded to the transcription start site of *rseAp* (data not shown). The *rseAp* regulation pattern followed that of *rpoEp*, being activated by σ^E (Fig. 2A, right, compare lane 10 to lane 9), repressed by RseA and MucD (Fig. 2A, right, compare lanes 3 and 4 to lane 1), and induced by heat stress (Fig. 2A, right, compare lane 5 to lane 1). As a negative control, we checked that there was no transcript corresponding to *rseAp* when RNAs purified from the Δ*rpoE* strain were used (Fig. 2A, right, lanes 2 and 6). Moreover, we could not detect any independent transcription start site for *mucD* if we used a primer located in the *mucD* gene, while we could still see a band corresponding to the *rseAp* start site (see Fig. S1 in the supplemental material). Taken together with the RT-PCR results (Fig. 1B), this strongly suggested that the *mucD* gene was transcribed from both *rpoEp* and *rseAp*. Moreover, as shown in Fig. 2C, the β-galactosidase activity driven from a plasmidic *rseAp-lacZ* fusion was higher than that of *rpoEp* (~3 fold) and was strictly dependent on σ^E, since almost no activity was detected in the absence of *rpoE*. Our results suggest that σ^E positively autoregulates itself and upregulates RseA and MucD. Since RseA and MucD are negative regulators of σ^E activity, this could set up a negative feedback loop to ensure rapid downregulation of the σ^E-dependent response for a return to the baseline level after a protein-folding stress condition or during normal growth.

σ^E is involved in stationary-phase survival and in response to diamide and cadmium. Several reports have indicated that *rpoE* mutants of Gram-negative bacteria are more susceptible to environmental stresses (60). To evaluate the role of σ^E in the physiology and environmental stress response of *X. campestris* pv. *campestris*, bacterial growth and the resistance levels of isogenic WT, Δ*rpoE*, Δ*rpoE-rseA*, Δ*rpoE-rseA-mucD*, Δ*rseA*, and Δ*mucD* strains against different stresses were determined. All of the strains exhibited mucoid phenotypes when grown on KADO plates, suggesting that the production of exopolysaccharides was not affected by the absence or the overactivation of σ^E (data not shown), as previously described in *X. campestris* pv. *campestris* strain 11 (17). Moreover, the morphology of the mutant cells (determined by phase-contrast microscopy) grown on MOKA medium was unaffected in stationary phase, but Δ*rseA* mutant cells were significantly smaller than WT cells in exponential phase (data not shown). This change in cell shape could indicate a link between σ^E and peptidoglycan assembly, as suggested for *E. coli* (31).

Bacteria were cultivated aerobically in MOKA medium and exposed to different treatments during mid-exponential growth for several periods of time, and viable-cell counts were determined. Resistance to acidic pH, 0.1% SDS, alkaline stress caused by 40 mM NaOH, oxidative stress caused by 5 mM H_2O_2, and hyperosmotic stress caused by 1.5 M NaCl was assayed. No statistically significant differences between the WT and Δ*rpoE* mutant strains were observed (data not shown). In addition, the different strains were tested for survival upon oxidative stress (1 M diamide and 200 mM paraquat) and against cell wall-active antibiotics (50 mg/ml vancomycin, 10 mg/ml polymyxin B sulfate, and 100 mM chlorpromazine) using a disk diffusion assay. Briefly, *X. campestris* pv. *campestris* cultures in mid-exponential phase were plated on MOKA agar, and sterile paper disks saturated with 10 μl of chemicals were layered on top prior to incubation at 30°C. Strains with *rpoE* deleted showed no significant differences (data not shown) apart from increased sensitivity to diamide: the diameter of growth inhibition for Δ*rpoE*, Δ*rpoE-rseA*, and Δ*rpoE-rseA-mucD* mutants was 2.85 cm (±0.20 cm) compared with 1.9 cm (±0.02 cm) for the WT strain. This suggests that σ^E is involved in the oxidative-stress response via a thiol oxidation pathway in *X. campestris* pv. *campestris*.

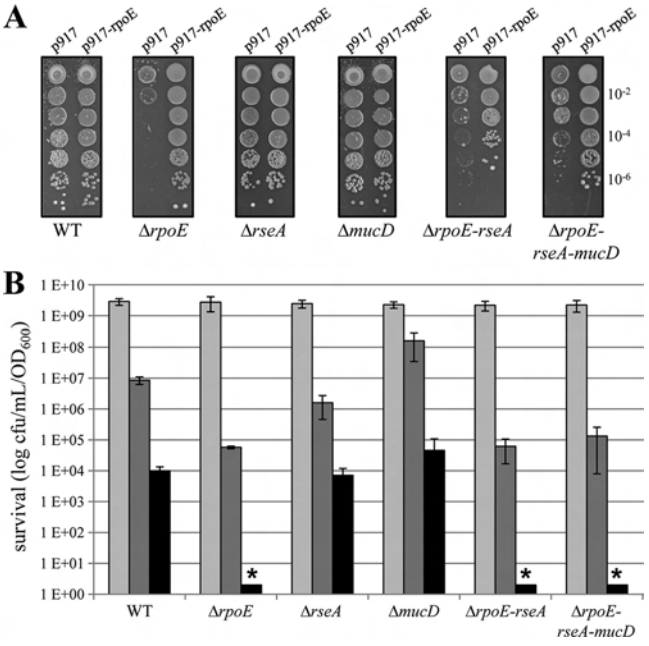


FIG. 3. Stress sensitivity of Δ*rpoE* mutants. (A) Influence of cadmium on growth of *X. campestris* pv. *campestris* WT and mutant strains containing control plasmid p917 or σ^E-overexpressing plasmid p917-rpoE. Serial 10-fold dilutions of late-exponential-phase bacteria were spotted on MOKA plates containing 40 μM cadmium and incubated at 30°C for 72 h. Each experiment was repeated three times. (B) Effect of *rpoE* deletion on stationary-phase survival of *X. campestris* pv. *campestris*. Overnight cultures in MOKA medium were diluted to an OD_{600} of 0.05 (time zero), and survival was monitored by viable-cell counting 24 h (light-gray bars), 48 h (dark-gray bars), and 96 h (black bars) postinoculation. The data are the mean values from four experiments, with the error bars representing standard deviations. The asterisks indicate no detectable cells.

Further, σ^E factors have been reported to be required for metal resistance in *E. coli* (26). We submitted *X. campestris* pv. *campestris* cells to cadmium stress by spotting serial dilutions of cultures of WT and mutant strains in the late exponential phase of growth on MOKA agar plates containing cadmium (40 μM) and incubating them at 30°C for 72 h. The results of a representative spot dilution experiment are shown in Fig. 3A.

We checked that control plasmid p917 had no impact on growth in the different *X. campestris* pv. *campestris* strains (data not shown). The deletion of *rpoE* was detrimental for resistance to cadmium, since there was a clear growth defect compared to the WT strain. Normal growth was restored when σ^E was overexpressed from plasmid p917-rpoE. These results indicate that σ^E is required for full adaptation to cadmium stress in *X. campestris* pv. *campestris*, as in *E. coli*.

It has been shown that σ^E activity increases upon entry into stationary phase in *E. coli* (19) and that *S.* Typhimurium σ^E is required for stationary-phase survival (67). To monitor the stationary-phase survival of *X. campestris* pv. *campestris* WT and mutant strains, the cells were grown at 30°C on a continuously shaking platform, and when the cells reached stationary phase, viable-cell counts were determined periodically by plating cells onto MOKA plates. As shown in Fig. 3B, the survival of Δ*rpoE* mutants was severely impaired, since the mutants displayed a 100-fold reduction in viability with respect to the WT after 48 h in stationary phase and there was no survival after 72 h, whereas the Δ*rseA* mutant did not show any appreciable phenotype under the conditions tested. These results strongly suggest that σ^E is required for stationary-phase survival of *X. campestris* pv. *campestris*.

Heat sensitivity of the *X. campestris* pv. *campestris* ΔrpoE mutant. Since we had shown that exposure to heat (35°C) activated σ^E, we tested the effect of temperature stress in isogenic WT, Δ*rpoE*, Δ*rpoE-rseA*, Δ*rpoE-rseA-mucD*, Δ*rseA*, and Δ*mucD* strains. Heat shocks from 42°C to 50°C were applied to exponentially growing cultures for several time points up to 30 min. Viable-cell counts were determined for all strains, and all the mutants gave similar responses to heat-killing treatment compared to the WT (data not shown), indicating that σ^E is not essential for survival after a heat shock in *X. campestris* pv. *campestris*. To check if σ^E could be involved in adaptation to heat stress, the growth of *rpoE* mutants was compared to that of the WT strain by spotting serial dilutions of bacterial cultures in the late exponential phase of growth on MOKA agar plates and incubating them at 30°C and 35°C. The results of a representative experiment are shown in Fig. 4A. Growth at 35°C was strongly impaired for the Δ*rpoE* mutants (Δ*rpoE*, Δ*rpoE-rseA*, and Δ*rpoE-rseA-mucD*), while there was no or little effect on growth for mutants that overactivate σ^E (Δ*rseA*) or for mutants of other ECF σ factors (Δ*XCC1664*). Conversely, there was no difference between the WT and mutant strains in plating efficiency when the strains were incubated for 72 h at 30°C. However, it must be noted that the colonies formed at 35°C with the Δ*mucD* mutant are reproducibly smaller than those formed at 30°C. This could be due to its putative protective role during protein folding stress by degrading misfolded proteins in the periplasm.

Complementation of the Δ*rpoE* mutants with plasmid p917-rpoE partially restored growth at 35°C (Fig. 4A). When we introduced the control plasmid p917 into the WT strain, we observed a small decrease in growth at 35°C compared to 30°C that could be the result of the combination of heat and antibiotic stresses (since we added tetracycline in the plates to maintain the plasmid). This inhibitory effect was even stronger when plasmid p917-rpoE was introduced into the WT strain, probably due to the toxicity of σ^E overexpression, as has been described for *E. coli* (58). This could explain why complementation of Δ*rpoE-rseA* and Δ*rpoE-rseA-mucD* mutants is less efficient than that of the Δ*rpoE* mutant, since the lack of appropriate posttranslational regulation of σ^E could be detrimental to bacterial fitness. We also tested the effect of σ^E on adaptation to cold stress at 14°C, and there was no difference between the WT and mutant strains (data not shown). Our results show that deletion of *rpoE* makes *X. campestris* pv. *campestris* cells extremely vulnerable to temperature adaptation at 35°C.

Given that ethanol is an amphiphilic compound that mimics the effects of high-temperature stress and that σ^E is required for growth in the presence of ethanol in several bacteria, such as *X. fastidiosa* or *V. cholerae* (24, 42), we tested the growth of WT and mutant *X. campestris* pv. *campestris* strains in the presence of 1.5% ethanol. As shown in Fig. 4B, the deletion of the *rpoE* gene strongly impaired growth, while deletion of *rseA* or *mucD* had no significant effect. Overall, these results indicate that σ^E contributes to cell envelope stress adaptation of *X. campestris* pv. *campestris*.

Identification of genes regulated by the σ^E factor. To gain insight into the functions regulated by σ^E in *X. campestris* pv. *campestris*, a global approach was chosen by comparing transcriptomes of the Δ*rpoE* strain (no σ^E activity) to that of the Δ*rseA* strain (overactivation of σ^E) using Xcc5kOLI microarrays (62). Given that there are nine ECFs in *X. campestris* pv. *campestris* and that they recognize close promoter sequences (65), we chose these setups in order to mimic a physiological activation of σ^E and to minimize nonspecific promoter recognition by σ^E. In the microarray analyses, any gene with a *P* value of ≤0.05 showing an increase of expression of 1.3-fold or more and a reduction in expression of 1.5-fold or more in the Δ*rseA* strain in comparison to the Δ*rpoE* strain was defined as being regulated by σ^E, either directly or indirectly (Tables 2 and 3). We have chosen relatively low thresholds to identify the significantly σ^E-regulated genes, since we have done four independent biological replicates, including a dye swap experiment. A total of 45 genes comprising 37 putative transcription units (TUs) were induced (Table 2), and 20 genes were repressed (Table 3). As expected, the *rpoE* and *mucD* transcripts were included in the group of upregulated genes. Nevertheless, the signal for the *rseA* transcript was in the range of the background under the conditions employed (data not shown), suggesting that there could be a problem with the detection of the oligonucleotide probe.

Of the upregulated genes, σ^E transcribes an array of biosynthetic enzymes that are involved in fatty acid metabolism (*fadE2* and *XCC3937*), redox metabolic functions (*XCC1588*), and electron transport systems (*bioI* and *XCC3906*). This raised electron-transport system activity could possibly compensate for proton leakage across the membrane when its integrity is compromised or could be a response to the formation of reactive oxygen species generated by a perturbation of the electron transport chain. This is further supported by the induction of *pqqA*, since the redox cofactor PQQ can act as an antioxidant metabolite to detoxify reactive oxygen species (45).

A large number of σ^E-upregulated genes bear signal sequences or transmembrane domains, which is consistent with the σ^E response having a role in monitoring and preserving the membrane during stress. In addition, three genes encode cell envelope proteins. Among them, *XCC0539* encodes a pre-

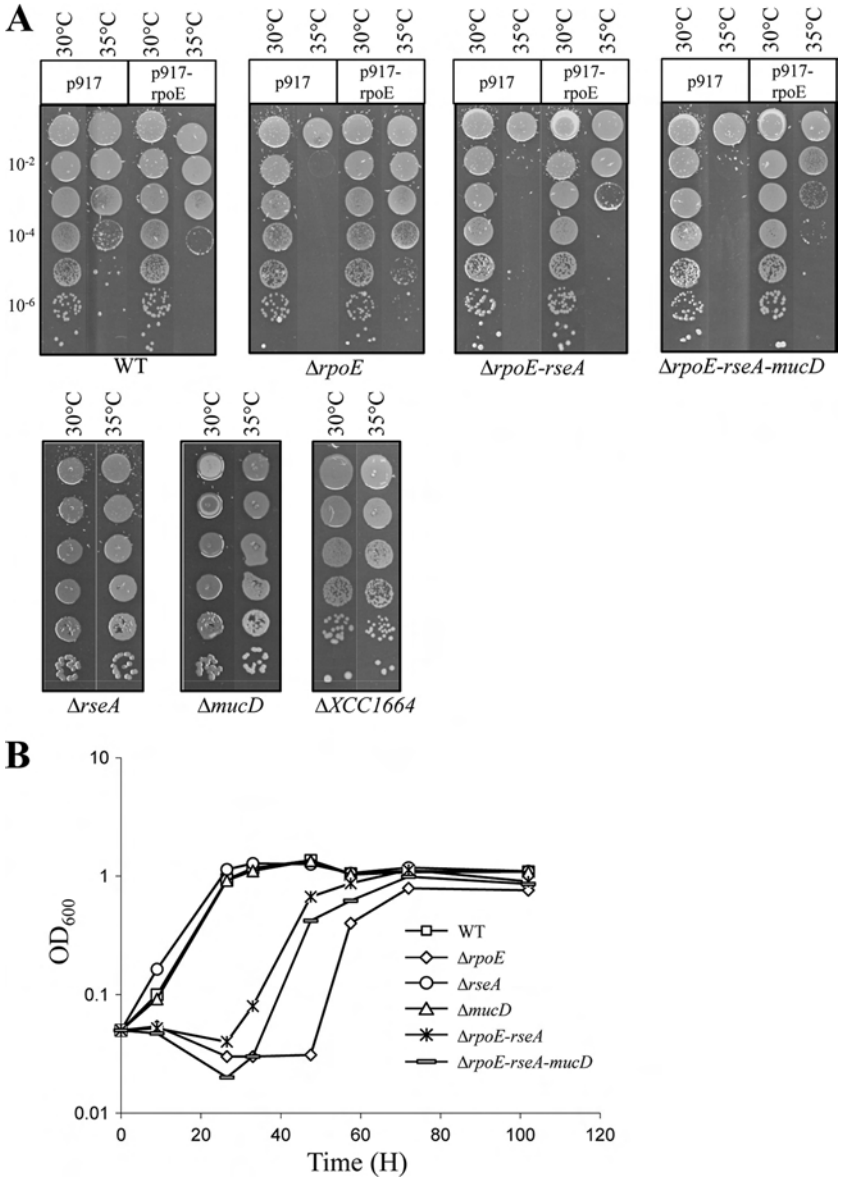


FIG. 4. The Δ*rpoE* mutant of *X. campestris* pv. *campestris* is sensitive to heat and ethanol. (A) Heat sensitivities of *X. campestris* pv. *campestris* WT and mutant strains containing no plasmid, control plasmid p917, or σ^E-overexpressing plasmid p917-rpoE. Serial 10-fold dilutions of late-exponential-phase bacteria were spotted on plates and incubated at 35°C for 72 h. Each experiment was repeated three times. (B) Growth curve of *X. campestris* pv. *campestris* WT and mutant strains in the presence of 1.5% ethanol. Cells were precultured in MOKA medium overnight and then diluted to an OD_{600} of 0.05 in MOKA medium containing 1.5% ethanol. One representative experiment out of three independent replicates is shown.

dicted member of the OmpW/AlkL family that is found in all Gram-negative bacteria and is involved in the protection of bacteria against various forms of environmental stress (36). *pglA* encodes a predicted polygalacturonase (PG) to selectively degrade the pectic polymers of the plant cell walls, and PGs are virulence factors in closely related bacterial species (73). Interestingly there is also an overrepresentation of type III secretion (T3S)-related genes within the σ^E regulon of *X. campestris* pv. *campestris*. In plant-pathogenic bacteria, the T3S system (T3SS) is one of the key pathogenicity factors and is encoded by the chromosomal *hrp* gene cluster. Here, we identified 12 genes whose products are involved in all aspects of T3S machinery and effector proteins that are under the control (at least partial) of σ^E. In xanthomonads, the *hrp* cluster is organized into at least six transcriptional units and is under the positive control of HrpG and HrpX (30). In our experiments, only two *hrp* operons (*hrpE* and *hrpF*) were upregulated, suggesting that σ^E-dependent regulation of T3S genes did not occur via HrpG or HrpX. To check this, we tested the effect of *rpoE* inactivation on plasmid-driven *hrpXp*- or *hrpGp-lacZ* transcriptional fusion reporters, and there was no significant change in their activity (data not shown).

As expected from the known σ^E-dependent stress responses, a heat shock response was induced in *X. campestris* pv. *campes-*

TABLE 2. Genes with increased expression in the Δ*rseA* strain compared with the Δ*rpoE* strain

ID	Gene	Ratio[a]	Description of gene product	Putative TU[b]	Characteristics
Regulatory function					
XCC1267	*rpoE2*	4.28	RNA polymerase σ factor σ^E	XCC1267, -68, -69	Transcription initiation
XCC3771	*rpoH*	1.52	Heat shock RNA polymerase factor σ^H		Response to heat; transcription initiation
XCC3348		1.48	Putative sensor-response regulator hybrid	XCC3347, -48	Two-component signal transduction system (phosphorelay)
Metabolism					
XCC2432	*fadE2*	3.42	Putative acyl-coenzyme A dehydrogenase	XCC2430, -31, -32	Oxidation reduction
XCC3937	*baf*	3.14	Putative type III pantothenate kinase	XCC3938, -37, -36	Positive regulation of transcription
XCC3047	*bioI*	1.45	Putative cytochrome P450 hydroxylase		Heme b metabolic process
(NC_003902: 3492579-3492659, plus strand[c])	*pqqA*	1.41	Putative coenzyme PQQ biosynthesis protein A	XCC2937, -38, -39, -40, -41	Pyrroloquinoline quinone biosynthetic process
XCC1588		1.32	Putative sulfite oxidase subunit YedY		Electron carrier activity
XCC3906		1.31	Putative cytochrome B561	XCC3905, -06	Respiratory electron transport chain
Protein synthesis and fate					
XCC1269	*mucD*	2.91	Putative periplasmic protease	XCC1267, -68, -69	Proteolysis
XCC1047	*hspA*	2.02	Low-molecular-wt heat shock protein		Response to stress
XCC3493-like degradation complex	*hslU*	1.83	Chaperone subunit of a proteasome	XCC3493, -94	Response to stress
XCC3227	*moxR*	1.67	MoxR-like AAA^+ ATPase chaperone	XCC3227, -26, -25, -, 24, -23, -22	ATPase activity
XCC2393	*htpG*	1.53	Molecular chaperone Hsp90 family		Protein folding
XCC1535	–[d]	1.41	FKPB-type peptidyl-prolyl *cis-trans* isomerase		Protein folding
XCC1475	*dnaJ*	1.36	Molecular chaperone Hsp40 family	XCC1474, -75	Response to heat; protein folding
XCC3450	*prc*	1.34	Putative carboxyl-terminal processing protease		Proteolysis
XCC1474	*dnaK*	1.31	Molecular chaperone Hsp70 family	XCC1474, -75	Response to heat; protein folding
T3S, translocation, and regulation machinery					
XCC1240	*hpa1*	2.15	Harpin-like elicitor protein (T3SS-dependent secreted protein)		
XCC1217	*hrpF*	2.04	Type III translocon protein		Interaction with host via protein secreted by T3SS
XCC1246		1.99	Type III effector protein (XopAL class)		Pathogenesis
XCC1241	*hpa2*[d]	1.78	Lytic transglycosylase-like		
XCC1222	*hrpD6*	1.64	Type III secretion system component	XCC1222, -21, -20, -19	Protein secretion by the T3SS
XCC1220	*hpaB*	1.54	Global T3S chaperone	XCC1222, -21, -20, -19	T3SS
XCC2565		1.50	Leucin-rich-repeat-containing protein/type III effector protein (XopAC class)		Pathogenesis
XCC4186		1.49	Leucin-rich-repeat-containing protein/type III effector (XopL class)		Pathogenesis
XCC1221	*hrpE*	1.47	Type III secretion system pilus protein	XCC1222, -21, -20, -19	T3SS complex
XCC2896	*psvA*	1.42	Peptidase C48 family/type III effector protein (XopD class)		Proteolysis
XCC1224	*hpaA*	1.34	Type III secreted virulence factor	XCC1222, -21, -20, -19	Interaction with host via protein secreted by T3SS
XCC1219	*hrpW*	1.33	Harpin pectate lyase	XCC1222, -21, -20, -19	Interaction with host via protein secreted by T3SS
Cell envelope					
XCC0539	*ompW3*	1.39	Putative outer membrane protein		Cell outer membrane
XCC2266	*pglA*[d]	1.31	Putative polygalacturonase		Carbohydrate metabolic process
XCC3925	*ecnA*[d,e]	1.31	Putative entericine A		Response to toxin
Unknown function					
XCC1308	–[d]	2.22	Hypothetical protein	XCC1306, -07, -08	
XCC0401		2.06	Hypothetical protein/ribosomal protein S30Ae/σ^{54} modulation protein		Primary metabolic process
(NC_003902: 3290964-3290833, minus strand[c])		2.05	Small putative membrane protein		
XCC3224		2.03	Hypothetical membrane protein	XCC3227, -26, -25, -24, -23, -22	
XCC1244		1.95	Hypothetical protein		
XCC0944	–[d]	1.83	Conserved hypothetical protein		
XCC3226		1.54	Hypothetical protein	XCC3227, -26, -25, -24, -23, -22	
XCC3798	–[d]	1.47	Calcium-binding EF hand motif		Calcium ion binding
XCC3887		1.35	Hypothetical protein		
XCC2566		1.35	Putative carboxymethylenebutenolidase		Hydrolase activity
XCC0863	–[d,e]	1.33	Putative membrane protein		
XCC1736	–[d]	1.33	Putative secreted protein	XCC1737, -36, -35, -34	Catalytic activity

[a] Ratio, averaged expression ratio of σ^E induced (Δ*rseA*)/no σ^E (Δ*rpoE*).
[b] TUs are listed in chromosomal order.
[c] Nonannotated gene in the genome of *X. campestris* pv. *campestris* ATCC 33913. Genomic location contains *X. campestris* pv. *campestris* ATCC 33913 chromosome accession number followed by the start coordinate, end coordinate, and strand.
[d] A predicted cleavable signal sequence.
[e] Predicted transmembrane helices.

TABLE 3. Genes with decreased expression in Δ*rseA* strain, as compared with Δ*rpoE*

ID	Gene	Ratio[a]	Description of gene product	Characteristic
Transporter activity				
XCC2867	*btuB*[b]	−1.97	TonB-dependent transporter	Transport; membrane
XCC3316	–[b]	−1.73	TonB-dependent transporter	Transport; membrane
XCC2497	–[b]	−1.61	Pseudo-TonB-dependent transporter/Oar like	Transport; membrane
XCC3271	–[b]	−1.54	Pseudo-TonB-Dependent Transporter/Oar like	Transport; membrane
XCC1892	*cirA*[b]	−1.39	TonB-dependent transporter	Transport; membrane
Regulatory function				
XCC3677		−9.64	Putative two-component system sensor kinase	Two-component signal transduction system (phosphorelay)
XCC1935	*rpoN2*	−1.54	RNA polymerase σ^{54} factor	Transcription initiation
XCC1276		−1.50	Putative sensor/response regulator hybrid	Two-component signal transduction system (phosphorelay)
Metabolism				
XCC2410		−23.07	Ketoglutarate semialdehyde dehydrogenase	Metabolic process
XCC2295		−1.68	Putative polyhydroxyalkanoate synthesis repressor	
XCC3324	*ilvB*	−1.57	Acetolactate synthase	Branched-chain family amino acid biosynthetic process
XCC0550	*atpF*	−1.55	F0F1 ATP synthase subunit B	ATP synthesis-coupled proton transport
Nucleic acid metabolic process				
XCC2904	*hsdS*	−7.86	Type I restriction enzyme (specificity chain) homologue	DNA modification
XCC0377	*mnmG*	−1.38	Pyridine nucleotide-disulfide oxidoreductase, class II	tRNA processing
Catalytic activity				
XCC0700	–[b]	−1.39	Putative peptidase S15	Proteolysis
Motility and attachment				
XCC3232	*pilM*	−1.35	Type IV pilus assembly protein PilM	Pilus assembly
Cell envelope				
XCC3017	*ompP6*[b]	−.53	Outer membrane protein P6 precursor (OmpA family)	Cell outer membrane
Oxidative stress function				
XCC1109	*katE*	−5.64	Catalase	Response to oxidative stress
Unknown function				
XCC1080		−5.74	Hypothetical protein	
XCC2823		−1.37	Hypothetical protein	

[a] Ratio, averaged expression ratio of σ^E induced (Δ*rseA*)/no σ^E (Δ*rpoE*).
[b] A predicted cleavable signal sequence.

tris, since several upregulated gene products are highly conserved heat shock proteins (HSPs), such as *dnaK-dnaJ*, *hspA*, *htpG*, and *hslU*, and the heat stress σ^H factor encoded by *rpoH*. Not only important during heat stress, many HSPs assist protein folding and homeostasis and are general stress proteins (49). Two other heat shock protease genes, *hslV* (forming a putative TU, *hslU-hslV*) and *lon*, showed a small but statistically significant 1.23-fold induction (data not shown). In many bacteria, most of these HSP genes are transcribed by the σ^H factor and *rpoH* expression is regulated by σ^E (50), and our data strongly suggest that heat stress regulation is the same in *X. campestris* pv. *campestris*. This is further supported by the presence of σ^H promoter elements (CTTGAAN$_{13\text{-}14}$CCCCA TNT) (41) within the 300 nucleotides upstream of the start codons of *dnaK*, *hspA*, *htpG*, *lon*, and *hslV* (data not shown). Nine σ^E-regulated genes encode proteins involved in protein quality control in the cytoplasm (e.g., DnaK and HslU) and in the periplasm (e.g., Prc and XCC1535), possibly to cope with the defects in membrane protein insertion, folding, and assembly resulting from σ^E-activating stress signals. Interestingly, the predicted periplasmic protease Prc has been reported to be implicated in the proteolysis of the anti-σ^E factors MucA in *P. aeruginosa* and RsiW in *B. subtilis* (32, 57). It is also worth noting that we expected to find *XCC1535* in our analysis, since

it shares the same promoter elements with its orthologue in *X. fastidiosa* (*xf0644*), which has been shown to be a direct target of σ^E (24).

Among the functions encoded by 20 σ^E-downregulated genes, there was an overrepresentation of proton motive force (PMF) consumers that could compete with putative proton leakage in response to localized disturbances in membrane integrity: (i) TonB-dependent transporters (TBDTs), which are in the outer membrane and are mainly involved in iron, vitamin B_{12}, or plant-derived carbohydrate uptake (9); (ii) ATP synthase; and (iii) flagella, since the transcript for the alternative σ factor RpoN2, which is responsible for the transcription of the flagellar T3S system (76), was downregulated upon σ^E overactivation. Thus, σ^E could function to maintain the membrane potential component of the PMF in *X. campestris* pv. *campestris* as in *S.* Typhimurium (7). In addition, the *pilM* gene was repressed by σ^E, and type IV pili could also represent an energy burden for the cell under stressful conditions.

One of the most downregulated genes was *hsdS*, possibly to provide restriction alleviation in response to stress stimuli, in agreement with work showing that the signal of chromosomal DNA damage might be transmitted to the cell surface via activation of the σ^E regulon (4). Among genes negatively regulated by σ^E, there was only one OMP. This minor inhibitory effect on OMP biosynthesis contrasts with other bacteria in which the σ^E response limits the expression of a number of OMPs via small RNAs (54) to prevent the accumulation of misfolded intermediates. This discrepancy may be due to the setup of our experiments, since we did not overexpress σ^E (instead, σ^E was constantly activated), but it could also suggest that the σ^E signaling system is regulated differently in *X. campestris* pv. *campestris*.

Validation of microarray experiments and determination of a σ^E-binding consensus motif. In order to validate the results of the microarray data, a subset of five genes from the σ^E regulon was randomly chosen for verification of σ^E dependence *in vivo*, using *lacZ* transcriptional reporter fusion assays. The upstream regions (~500 bp) of the chosen genes were fused to a promoterless *lacZ* reporter gene in a broad-host-range vector as described in Materials and Methods. We also included in this study a non-σ^E-dependent promoter region as a negative control. To test the expression of these promoters with different levels of σ^E activity, the plasmids were transferred to the isogenic WT, Δ*rpoE*, and Δ*rseA* strains. As expected, the gene fusions showed dramatically increased expression in the Δ*rseA* mutant strain containing constitutively high σ^E activity (Fig. 5). However, there was only a small decrease of β-galactosidase activity upon loss of σ^E, probably because the cells were grown under nonactivating conditions. As expected, the expression levels of the control gene *XCC1230* were similar in all strains.

To confirm our transcriptome analysis, we also checked some phenotypes associated with predicted σ^E-regulated functions. In *X. campestris* pv. *campestris*, the alternative σ^{54} factor RpoN2 is responsible for the transcription of the flagellar genes (76). Since RpoN2 was negatively regulated by σ^E, it was possible that *rseA* mutants possessed reduced motility. To test this, we measured the motility of the WT, Δ*rpoN2*, and Δ*rseA* strains (data not shown). The Δ*rpoN2* mutant cells had impaired motility, and Δ*rseA* mutant motility was delayed, since its swimming zone was reproducibly smaller than that of the WT after 2 days at 30°C. These results are consistent with the microarray analyses. In addition, we tested the Δ*rpoE* mutants for pathogenicity on a host plant, since σ^E controlled virulence-associated genes. The *X. campestris* pv. *campestris* WT and mutant strains were inoculated into *Arabidopsis thaliana Sf-2* ecotype leaves by using the leaf-clipping method. The deletion of *rpoE* did not reproducibly alter symptom development (data not shown), suggesting that σ^E is not required for the virulence of *X. campestris* pv. *campestris* under these conditions.

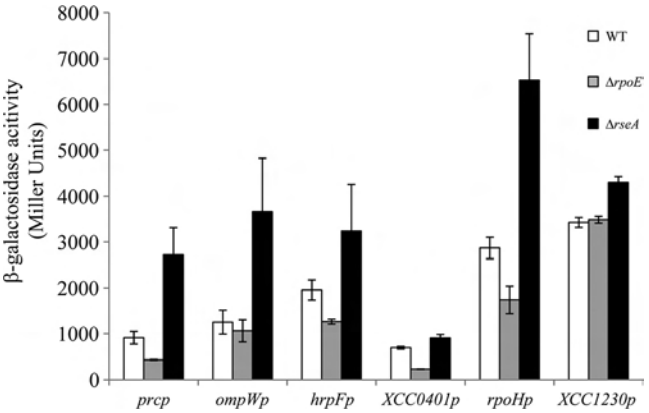


FIG. 5. Expression of σ^E-regulated genes. Reporter plasmids carrying the 5′ ends of selected genes and their upstream promoter regions fused to *lacZ* were transferred to three *X. campestris* pv. *campestris* strains with different σ^E activities: WT, Δ*rpoE*, and Δ*rseA*. Overnight cultures of these strains grown in MOKA medium were diluted in the same medium and grown for 9 h before determination of β-galactosidase activity. The results represent the mean values of at least two independent experiments, each performed in triplicate, and the error bars indicate the standard errors.

To further verify the transcriptome data and to identify σ^E promoter elements, we selected 11 σ^E-upregulated genes from nearly all the functional categories identified, and their promoters were mapped by primer extension (Fig. 6; see Fig. S1 in the supplemental material). We added to our study the *XCC0964* gene (encoding a putative membrane metalloprotease), since it shares promoter elements with its *X. fastidiosa* orthologue (*xf2594*), which has been shown to be a direct target of σ^E (24). We could not detect any reverse transcription product for *XCC1246* unless σ^E was overexpressed from plasmid p917-rpoE, suggesting that the amount of *XCC1246* transcripts was below the detection threshold. For *rpoH*, two primer extension products (*rpoHp1* and *rpoHp2*) were detected. All primer extension experiments confirmed the trend of changes in transcript levels following σ^E activation observed in microarray experiments (Fig. 6, compare lanes Δ*rseA* and Δ*rpoE*), with the exception of the distal promoter of *rpoH* (*rpoHp2*) (see below). Moreover, all promoters except *rpoHp2* were induced when σ^E was overexpressed from the p917-rpoE plasmid [Fig. 6, compare lanes wt(+) and wt(−)]. The proximal promoter *rpoHp1* was controlled by σ^E, since its activity increased when *rseA* or *mucD* was deleted and decreased when *rpoE* was deleted. In addition, the loss of *rpoHp1* activity was complemented by the overexpression of σ^E from plasmid p917-

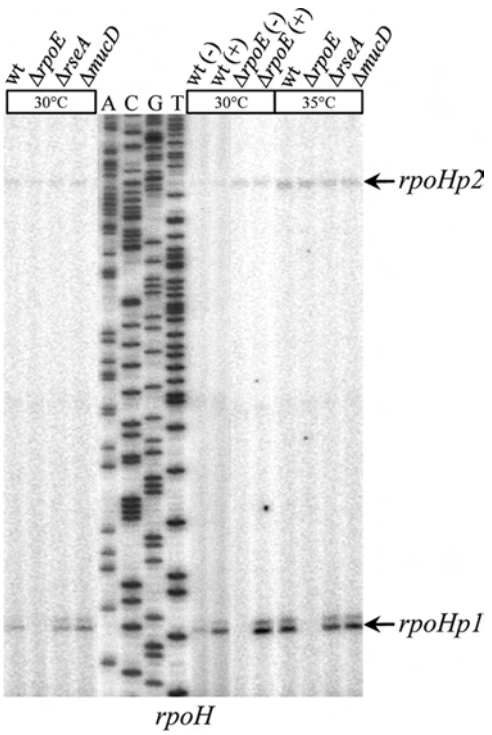


FIG. 6. Determination of the transcription start sites of *rpoH*. Total RNAs from WT, Δ*rpoE*, Δ*rseA*, Δ*mucD*, and WT or Δ*rpoE* strains containing control plasmid p917 (−) or σ^E-overexpressing plasmid p917-rpoE (+) were used as templates in primer extension experiments. Total RNAs were obtained from cells incubated at 30°C or after a 60-min shift at 35°C. Primer rpoH-EXT was 5′ end labeled with ^{32}P and extended with reverse transcriptase to map its corresponding gene promoter sequences. The arrows indicate the bands corresponding to the main start sites observed.

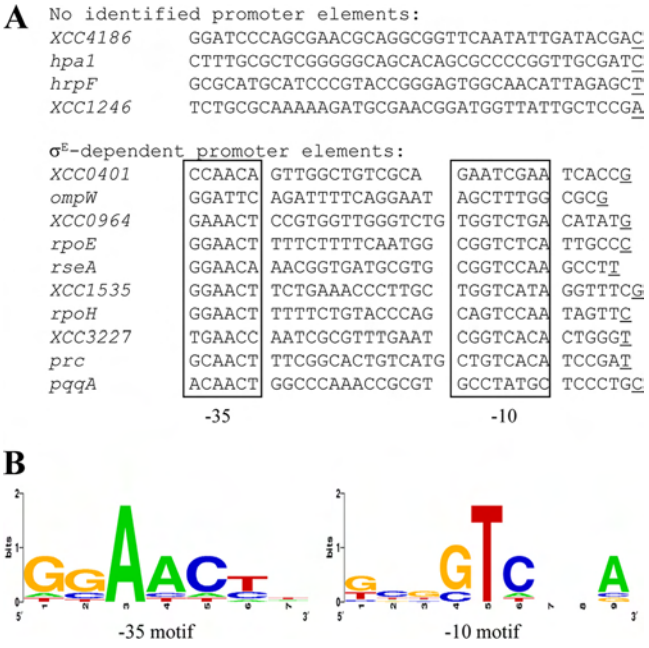


FIG. 7. Identification of the *X. campestris* pv. *campestris* σ^E promoter recognition sequence. (A) Alignment of σ^E-dependent promoters identified by primer extension assays. The transcription start site is underlined, and the −35 and −10 motifs are boxed. (B) The aligned promoter sequences were analyzed using the WebLogo program (http://weblogo.berkeley.edu/). The height of a stack indicates sequence conservation (2 = 100% conservation), and the height of each individual nucleotide within the stack indicates its relative frequency at that position.

rpoE [Fig. 6, compare lanes Δ*rpoE*(+) and Δ*rpoE*(−)]. The distal *rpoHp2* was σ^E independent, since its activity did not vary in mutants of the *rpoE* operon. Overall, these results indicated good verification of the microarray data.

A consensus for the σ^E-binding motif was obtained by alignment of the 40 nucleotides upstream of the experimentally determined start sites and a search for conserved promoter elements identified from other bacterial σ^E regulons and for *rpoEp* and *rseAp* (65) (Fig. 2B). As shown in Fig. 7A, 8 out of the 12 selected genes had these conserved promoter elements, indicating that they could be directly regulated by σ^E. The proposed consensus for *X. campestris* pv. *campestris* σ^E target promoter motifs ($GGAACTN_{15\text{-}17}GTCNNA$) is very similar to the σ^E-binding sequence of homologous ECF σ factors from the ECF02 group (65) (Fig. 2B and 7B). It harbors the hallmark feature of an ECF-type promoter, the AAC motif in the −35 region, and the −10 region contains highly conserved TC and A residues. Interestingly, only half of the 14 mapped promoters depended solely on σ^E for their transcription: *rpoEp*, *rseAp*, *XCC1535p*, *rpoHp2*, *XCC3227p*, *XCC0964p*, and *prcp* (Fig. 2A and 6; see Fig. S1 in the supplemental material). This suggests that partially σ^E-dependent genes could be controlled by other σ factors that overlap functionally in *X. campestris* pv. *campestris*. We were unable to identify σ^E conserved promoter elements for the four T3S-related genes *XCC1246*, *XCC4186*, *hpaI*, and *hrpF*, suggesting that the effect of σ^E on their expression is indirect.

To gain further insight into σ^E activity under heat stress conditions and to further validate our microarray data, we extracted total RNAs from bacterial cultures after a shift to 35°C. We included only the genes whose expression was heat inducible (Fig. 6 and data not shown; see Fig. S1 in the supplemental material). The relative levels of *rpoH*, *ompW*, *XCC3227*, *hspA*, *XCC1535*, and *prc* transcripts dramatically increased upon heat treatment, indicating that a subset of the σ^E regulon was heat responsive. Both promoters of *rpoH* were induced by a temperature upshift, but only *rpoHp1* was upregulated in a σ^E-dependent manner. We failed to identify conserved promoter elements upstream of the *rpoHp2* start site, so we could not predict which σ factor could be in charge of its transcription. The heat inducibility of *XCC3227p*, *XCC1535p*, and *prcp* expression was dependent on σ^E (see Fig. S1 in the supplemental material), as for *rpoEp* and *rseAp* (Fig. 2A). Hence, this suggests that their induction is likely to be accomplished by activation of σ^E upon heat exposure, probably mediated by anti-σ RseA cleavage. In contrast, *ompWp* heat induction was σ^E independent, and this will need further investigation. Interestingly, we could not directly test the impact of σ^H on gene expression because, despite numerous attempts, we could not inactivate the *rpoH* gene. It is thus likely (although not formally proven) that *rpoH* is essential in *X. campestris* pv. *campestris*, as it is thought to be in other bacteria (29, 63).

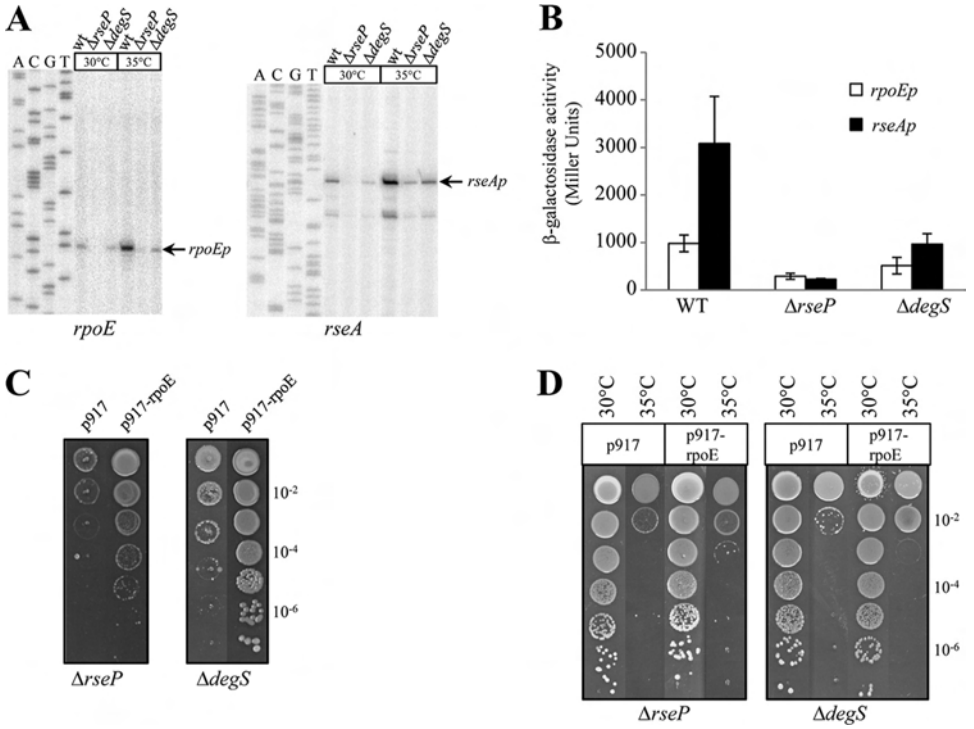


FIG. 8. The deletion of *rseP* or *degS* impairs σ^E activation. (A) Expression of σ^E-dependent promoters using primer extension assays. Total RNA was prepared from the WT, Δ*rseP*, and Δ*degS* mutant strains grown at 30°C or exposed to a temperature upshift at 35°C for 60 min. Primers rpoE-EXT and rseA(2)-EXT were used. (B) Determination of *rpoEp* and *rseAp* activities in WT, Δ*degS*, and Δ*rseP* strains. Plasmids containing a transcriptional fusion of the upstream region of *rpoEp* or *rseAp* to the *lacZ* gene were transferred into *X. campestris* pv. *campestris* strains. Overnight cultures of these strains grown in MOKA medium were diluted in the same medium and grown for 9 h before determination of β-galactosidase activity. The results represent the mean values of at least two independent experiments, each performed in triplicate, and the error bars indicate the standard errors. (C) Influence of 40 μM cadmium on growth of *X. campestris* pv. *campestris* WT, Δ*degS*, and Δ*rseP* strains containing control plasmid p917 or σ^E-overexpressing plasmid p917-rpoE. Serial 10-fold dilutions of late-exponential-phase bacteria were spotted on MOKA plates containing cadmium and incubated at 30°C for 72 h. (D) Heat sensitivities of *X. campestris* pv. *campestris* WT, Δ*degS*, and Δ*rseP* strains containing control plasmid p917 or σ^E-overexpressing plasmid p917-rpoE. Serial 10-fold dilutions of late-exponential-phase bacteria were spotted on plates and incubated at 35°C for 72 h. For panels C and D, each experiment was repeated three times.

Regulation of σ^E-mediated response by RseP (XCC1366) and DegS (XCC3898). Upon cell envelope stress in *E. coli*, RseA is sequentially cleaved by the RIP proteases DegS and YaeL (1). Specifically, the recognition of the C termini of unfolded OMPs allows DegS to cleave the periplasmic C terminus (site 1) of RseA. This converts RseA into a substrate for RseP, which cleaves the transmembrane segment (site 2) of RseA. To gain insights into the σ^E activation pathway, we asked whether the σ^E-activating signal generated in *X. campestris* pv. *campestris* also requires DegS and RseP. To find the *X. campestris* pv. *campestris* homologues of DegS and RseP, BLASTP (http://blast.ncbi.nlm.nih.gov/Blast.cgi) searches were performed using DegS and RseP homologues from *E. coli* and *P. aeruginosa*. The *X. campestris* pv. *campestris* genome contains a homologue of RseP (XCC1366, with its predicted amino acid sequence sharing 56% and 62% similarity with *E. coli* RseP and *P. aeruginosa* MucP, respectively; XCC1366 will be referred to as RseP protease), and XCC3898, annotated as protease DO, had the highest level of homology to DegS proteins (XCC3898, with its predicted amino acid sequence sharing 61% similarity with *E. coli* DegS and *P. aeruginosa* AlgW; XCC3898 will be referred to as DegS protease). In *X. campestris* pv. *campestris*, there was also significant similarity (62%) between DegS and MucD, since they belong to the widely conserved family of HtrA proteins (18).

To probe the role of the RseP and DegS proteases in σ^E-dependent transcription in *X. campestris* pv. *campestris*, we constructed chromosomal in-frame deletions of *rseP* or *degS*. As shown in Fig. 8A, upon deletion of *degS* or *rseP*, there was a strong decrease in the expression of the σ^E-dependent promoters *rpoEp* and *rseAp* at both 30°C and 35°C compared to the WT levels. Moreover, σ^E-dependent transcriptional activity from *rpoEp-lacZ* or *rseAp-lacZ* in the Δ*rseP* and Δ*degS* mutants was significantly reduced during normal growth (Fig. 8B). These results indicated that *degS* and *rseP* are required for the activation of σ^E, thus suggesting a role for DegS and RseP proteases in cleaving RseA in *X. campestris* pv. *campestris*. Interestingly, the two proteases do not have the same impact on σ^E-dependent activity, since the Δ*rseP* mutant was the most affected. If the RIP protease cascade is conserved in *X. campestris* pv. *campestris*, this raises the possibility that, at least under certain conditions, site 1 proteases other than DegS could initiate the cleavage of RseA or that the second-site cleavage of RseA by RseP could occur independently of a site 1 protease. However, σ^E-dependent activity was still induced in Δ*rseP* and Δ*degS* mutants following a temperature upshift to

35°C (Fig. 8A), indicating that RseP and DegS had little or no impact on the heat stress response of σ^E activation.

The protease mutants were then tested for σ^E-dependent stress responses, such as resistance to cadmium (Fig. 8C) and adaptation to increased temperature (Fig. 8D). As expected, deletion of *rseP* or *degS* resulted in a strong growth defect in the presence of cadmium, similar to the results obtained with a Δ*rpoE* mutant (Fig. 3A). The overexpression of σ^E from plasmid p917-rpoE restored resistance to cadmium, indicating that overexpression of σ^E bypasses the need for RseP and DegS for metal resistance. These data point to a role of proteolytically active DegS and RseP in the activation of σ^E in *X. campestris* pv. *campestris* during cadmium stress. At 35°C, growth of the Δ*degS* and the Δ*rseP* mutants was strongly impaired, similar to the results obtained with a Δ*rpoE* mutant (Fig. 4A). However, the overexpression of σ^E from plasmid p917-rpoE did not fully restore the growth defect phenotype of the protease mutant strains, probably because DegS and RseP are essential for growth at 35°C. These results strongly suggest that DegS and RseP degrade cytoplasmic-membrane-localized substrates (other than RseA) that could be involved in the heat stress response. Furthermore, it must be noted that, as for *X. fastidiosa*, there are no σ^E-dependent promoter elements upstream of the *degS* gene, but instead, its promoter presents a putative σ^H consensus sequence (data not shown). Unfortunately, we have been unable to experimentally identify *degS* or *rseP* promoters, so that the regulation of these genes under stress conditions is still under investigation. All these data show that DegS and RseP are required to activate σ^E, and this is probably achieved by two-step proteolysis to liberate σ^E from RseA. However, a residual induction of σ^E activity could still occur without these proteases, depending on growth conditions.

DISCUSSION

Here, we have combined different approaches to investigate the role and mechanism of σ^E-dependent activation in *X. campestris* pv. *campestris*. We demonstrated that σ^E is an important regulator of the stress response of *X. campestris* pv. *campestris*, since it is involved in stationary-phase survival and resistance to membrane-perturbing stresses. However, we were unable to find any increased sensitivity of the *rpoE* mutant to a variety of chemical stresses (data not shown). This limited response was not surprising, since there are 9 other ECF σ factors in *X. campestris* pv. *campestris*. Each one could respond to a specific stress stimulus and could substitute functionally for the others, since ECFs tend to share overlapping promoter specificity, as has been well described in *B. subtilis* for σ^M, σ^W, and σ^X (43). Knowledge of the remaining ECF σ factors and their inducing signals is sparse, and it is crucial to address these issues in future studies.

As in *X. campestris* pv. *campestris* strain 11 and *X. fastidiosa*, the *rpoE* gene of *X. campestris* pv. *campestris* is organized as a single transcription unit with the anti-σ gene *rseA* and the protease gene *mucD*. However, the transcriptional control of *rpoE* expression is different in these bacterial species, since it is controlled by the housekeeping σ factor in *X. fastidiosa* (24), while transcription of *rpoE* is autoregulated in *X. campestris* pv. *campestris*, like most of its known orthologues (our results and reference 17). This indicates that regulatory pathways may have evolved differently in these closely related bacterial species in order to be able to process different environmental parameters, since *X. fastidiosa* appears to have an *in planta*-restricted lifestyle while *X. campestris* pv. *campestris* is able to survive in the environment between infections. Our results strongly suggest that RseA is the anti-σ^E factor and that MucD is a negative regulator of σ^E-dependent activity. The transcription of *rseA* and *mucD* is controlled by a highly conserved σ^E-dependent promoter within the σ^E gene (24, 59). This negative feedback loop would ensure the tight control of σ^E-dependent activity in *X. campestris* pv. *campestris* in order to avoid the energy cost of an inappropriate σ^E response, as we observed that the overexpression of σ^E was deleterious to *X. campestris* pv. *campestris* fitness.

We used microarrays to define the *X. campestris* pv. *campestris* σ^E regulon by comparing the levels of the transcripts of all annotated ORFs in the *rpoE* deletion mutant with those in the *rseA* deletion mutant in order to be closer to physiological levels of induction of the σ^E response. σ^E upregulates at least 45 genes, including *rpoE* regulon members themselves, and 20 genes are downregulated upon σ^E overactivation. We identified the *X. campestris* pv. *campestris* σ^E promoter consensus motif ($GGAACTN_{15\text{-}17}GTCNNA$), which has the conserved characteristics of promoter elements of ECF σ factors from the ECF02 group (65). Most σ^E-dependent genes fall broadly into the same categories previously described for the core and extended σ^E regulon members (59) with structural components of the cell envelope, OMP assembly, the periplasmic chaperone, fatty acid biosynthesis, and two-component systems. The higher electron transport system activity could possibly compensate for proton leakage across the membrane when its integrity is compromised in order to maintain the PMF. In agreement with this notion, a functional σ^E regulon was essential for the maintenance of PMF in *S.* Typhimurium (7), and we observed that genes encoding functions that are PMF-energized processes (ATP synthesis, motility, and active transport) were downregulated when σ^E was overactive.

In contrast to core σ^E regulons of enterobacteria, we did not find genes involved in cell envelope biogenesis (59). Variation in growth conditions, treatment conditions, and microarray platforms could be the reasons for this discrepancy. A subset of σ^E-dependent genes in *X. campestris* pv. *campestris* belongs to the highly conserved heat shock regulon involved in aiding protein folding, protein disaggregation, and proteolysis and comprising the heat shock sigma factor (*rpoH*). We also showed that *rpoH* transcription was directly regulated by σ^E under both normal and heat stress conditions, so that σ^H could in turn direct transcription of the heat shock regulon. The ability of σ^E to mediate the expression of *rpoH* has been reported in several other bacteria, particularly in gammaproteobacteria (3, 69). It is noteworthy that the *rpoH* gene of *X. campestris* pv. *campestris* strain 11 has σ^E promoter elements identical to those in *X. campestris* pv. *campestris*, but previous work did not correctly predict the σ^E-dependent promoter of *rpoH* in *X. campestris* pv. *campestris* strain 11 (37). We have also identified a distal σ^E-independent promoter for *rpoH*, and its expression was slightly induced by temperature. These data underline the complex regulation of *rpoH*, as in many other bacteria, e.g., in *E. coli* there are five promoters upstream of

the *rpoH* gene, recognized by σ^{70}, σ^{S}, σ^{E}, and σ^{54} (38). This ensures that the protective heat shock response can be triggered by several environmental cues or during the cell cycle. In *X. campestris* pv. *campestris* it seems particularly relevant, since we could not inactivate the *rpoH* gene under normal growth conditions (data not shown). Moreover, we observed that *rpoEp*, *rseAp*, *XCC1535p*, *XCC3227p*, and *prcp* were directly activated by σ^{E} following heat treatment at 35°C. XCC1535 and Prc are predicted to be periplasmic proteins and to have roles, respectively, in protein folding and proteolysis. This suggests that they could prevent aggregation of misfolded periplasmic protein derivatives under increased export stress, which could occur when cellular envelope integrity is compromised. Taken together, our results support the essential role of *rpoE* in heat stress protection of *X. campestris* pv. *campestris* and suggest that the heat shock response in *X. campestris* pv. *campestris* is similar to that of *E. coli* and other gammaproteobacteria.

We also found that σ^{E} is involved in the transcriptional activity of a subset of genes encoding proteins involved in the T3S translocation and regulation machinery. Only two of the six *hrp* operons of *X. campestris* pv. *campestris*, *hrpE* and *hrpF*, were induced. This is reminiscent of the regulation of the *hrpC* and *hrpE* operons by the two-component regulatory system ColR/ColS in *X. campestris* pv. *campestris* 8004 (77), suggesting that individual *hrp* operons might be targeted by alternative regulatory networks integrating diverse environmental signals. Taken together with the contribution of σ^{E} to the functionality of T3S systems of the enteropathogens *Yersinia pseudotuberculosis* and *S.* Typhimurium (15, 53), this regulation suggests a link between cell envelope perturbation and T3S induction in *X. campestris* pv. *campestris*. This could occur during the pathogenesis process in response to host factors (oxidative antimicrobial systems and membrane-targeted host defense peptides) that compromise bacterial membrane integrity. We tested the hypothesis that σ^{E} could be involved in *X. campestris* pv. *campestris* pathogenesis, but we were unable to see significant differences between the WT and the Δ*rpoE* mutant when we inoculated them into *A. thaliana* (data not shown). These results are in agreement with previous studies based on *X. campestris* pv. *campestris* strain 11 (17). This lack of a phenotype does not exclude the possibility that σ^{E} has a role under natural infection conditions of *X. campestris* pv. *campestris* that may be discovered using a different plant model or route of infection. Moreover, T3S is involved not only in pathogenesis, but also in pathogen dissemination. Mutation of genes encoding structural components and regulatory genes of the T3SS of *Xanthomonas fuscans* subsp. *fuscans* altered the ability of the bacterium to transmit to bean seeds through both vascular and floral pathways (21). The importance of *rpoE* in resisting external stresses suggests that σ^{E} may be involved in survival in the environment, which constitutes a prerequisite to plant infection and disease development (35). Bacterial epiphytic fitness and seed transmission experiments will be crucial to decipher the role of σ^{E} in the dispersion of plant-pathogenic bacteria, colonization, and survival on their hosts.

In *E. coli*, the σ^{E} response is triggered when periplasmic protein folding and assembly are compromised. The accumulation of unassembled OMPs is the activation signal for the initiation of a sequential proteolytic cascade of the membrane-spanning anti-σ RseA by DegS and RseP (72). DegS and RseP homologues have been identified in *X. campestris* pv. *campestris* based on sequence homologies. The deletion of their corresponding genes led to a decrease in σ^{E}-dependent activity at both 30°C and 35°C. Our results strongly suggest that the predicted proteases DegS and RseP regulate σ^{E} by the mechanism of RIP as described in *E. coli*. However, there were differences in the proteolytic process: when proteases were inactivated, residual σ^{E}-dependent transcription under normal growth conditions remained, and there was still induction of the σ^{E} response by a temperature upshift. This was rather unexpected, because it would be the first example of σ^{E} activation by a temperature upshift that did not require DegS and RseP proteases. Consequently, it must be assumed that RseA/σ^{E} cleavage can be performed by other proteases, depending on stress stimuli, in *X. campestris* pv. *campestris*. This is supported by the recent findings that acid stress activation of σ^{E}-dependent gene expression in *S.* Typhimurium is independent of the unfolded OMP signal or the DegS protease (48). These observations unravel a new activation pathway for the σ^{E} response. In addition, in *E. coli*, it has been shown that the HtrA family members DegQ and DegS can complement the loss of the periplasmic protease Prc (6). Interestingly, *prc* was one of the σ^{E} target genes identified in our microarray analysis, and we also showed that its expression was heat inducible. Tail-specific proteases like Prc have been reported to be implicated in RIP: CtpB of *B. subtilis* triggers activation of σ^{K} indirectly by cleaving the SpoIVFA protein (14), and Prc of *P. aeruginosa* is implicated in the AlgU/MucA σ/anti-σ system (57). Hence, it is conceivable that Prc is involved in RseA/σ^{E} cleavage in *X. campestris* pv. *campestris* under some conditions, and it will be the aim of future work to address these hypotheses. These unusual cleavage features of the RseA/σ^{E} system in *X. campestris* pv. *campestris* could also be due to the absence of the *rseB* gene in the *rpoE* operon. RseB is a negative regulator of the σ^{E}-dependent envelope stress response that binds to the periplasmic domain of RseA (1). RseB prevents DegS and RseP from cleaving RseA, suggesting that interaction of RseB with RseA must be altered before the signal transduction cascade is activated. There is no homologue of RseB in the *X. campestris* pv. *campestris* genome, and it is unknown if another protein could fulfill its protective function for inappropriate cleavage of RseA in the absence of activating stress.

Other observations were that *degS* and *rseP* mutants were strongly affected for growth at 35°C and in the presence of cadmium but that σ^{E} overexpression in these mutant strains restored growth only in the presence of heavy metals. These data implied that both DegS and RseP are required for growth at 35°C, possibly because they could have other substrates. It is the case for *E. coli* RseP, which could use the signal-transducing protein FecR as a substrate (13), and the *B. subtilis* RseP orthologue, RasP, which can attack the cell division protein FtsL (12). In addition, DegS and MucD belong to the HtrA protein family. In *E. coli*, the HtrA family member DegQ can rescue the conditional lethality of *degP* deletion (6, 71). It is tempting to speculate that since there is no homologue of DegQ in *X. campestris* pv. *campestris*, DegS could also substitute for the *X. campestris* pv. *campestris* DegP orthologue, MucD. This could explain why the *mucD* mutant is not affected by temperature stress in *X. campestris* pv. *campestris*, whereas

in most bacterial species *degP* mutants show a thermosensitive phenotype *in vitro* (47, 74).

Taken together, the data show that while the general function of σ^E is regulation of the extracellular stress response in *X. campestris* pv. *campestris*, its specific functions and regulation pathways are complex and multifaceted. A lot of questions remain to be answered, and continued work on the role and RIP mechanism of the σ^E response will be of major importance in ultimately understanding the nature and complexity of the envelope stress response, as these regulatory proteins can be found in a great variety of prokaryotes.

ACKNOWLEDGMENTS

We are grateful to Claude Bruand for critical reading of the manuscript. We thank Denis Bidot, Manuela Meyer, and Eva Schulte-Berndt for technical assistance.

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and the Université de Toulouse (UPS) and by grants 0313805A (Bundesministerium für Bildung und Forschung, Germany) and SPP 1316 (German Research Foundation) to A.B.

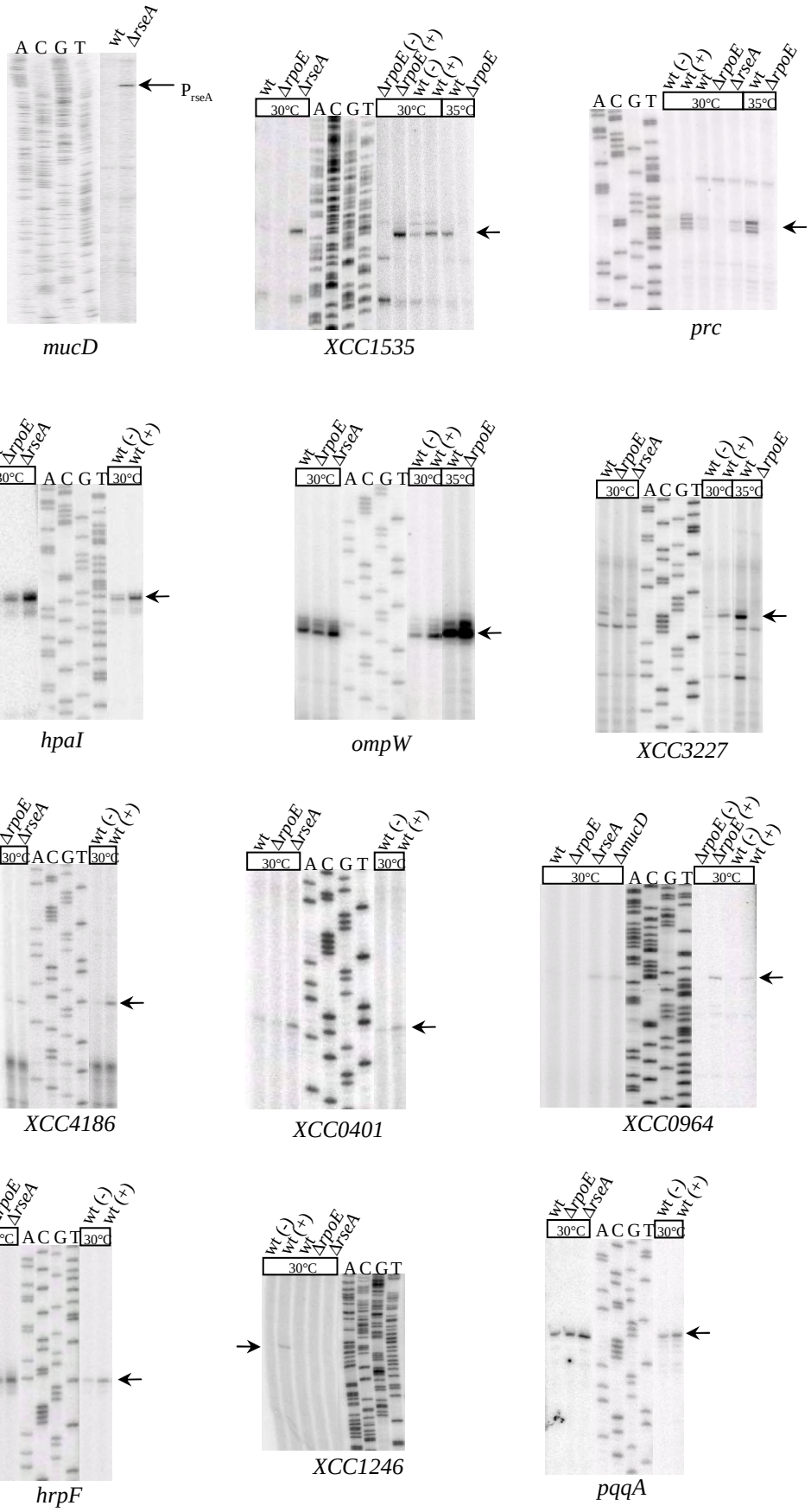
REFERENCES

1. **Ades, S. E.** 2008. Regulation by destruction: design of the σ^E envelope stress response. Curr. Opin. Microbiol. **11:**535–540.
2. **Alba, B. M., and C. A. Gross.** 2004. Regulation of the *Escherichia coli* σ-dependent envelope stress response. Mol. Microbiol. **52:**613–619.
3. **Aramaki, H., T. Hirata, C. Hara, M. Fujita, and Y. Sagara.** 2001. Transcription analysis of *rpoH* in *Pseudomonas putida*. FEMS Microbiol. Lett. **205:** 165–169.
4. **Asakura, Y., and I. Kobayashi.** 2009. From damaged genome to cell surface: transcriptome changes during bacterial cell death triggered by loss of a restriction-modification gene complex. Nucleic Acids Res. **37:**3021–3031.
5. **Bashyam, M. D., and S. E. Hasnain.** 2004. The extracytoplasmic function sigma factors: role in bacterial pathogenesis. Infect. Genet. Evol. **4:**301–308.
6. **Bass, S., Q. Gu, and A. Christen.** 1996. Multicopy suppressors of *prc* mutant *Escherichia coli* include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. J. Bacteriol. **178:**1154–1161.
7. **Becker, L. A., I. S. Bang, M. L. Crouch, and F. C. Fang.** 2005. Compensatory role of PspA, a member of the phage shock protein operon, in *rpoE* mutant *Salmonella enterica* serovar Typhimurium. Mol. Microbiol. **56:**1004–1016.
8. **Benjamini, Y., and Y. Hochberg.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B **57:**289–300.
9. **Blanvillain, S., D. Meyer, A. Boulanger, M. Lautier, C. Guynet, N. Denancé, J. Vasse, E. Lauber, and M. Arlat.** 2007. Plant carbohydrate scavenging through *tonB*-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS One **21:**e224.
10. **Boucher, J. C., M. J. Schurr, H. Yu, D. W. Rowen, and V. Deretic.** 1997. Pseudomonas aeruginosa in cystic fibrosis: role of *mucC* in the regulation of alginate production and stress sensitivity. Microbiology **143:**3473–3480.
11. **Boulanger, A., G. Déjean, M. Lautier, M. Glories, C. Zischek, M. Arlat, and E. Lauber.** 2010. Identification and regulation of the N-acetylglucosamine utilization pathway of the plant pathogenic bacterium *Xanthomonas campestris* pv. *campestris.* J. Bacteriol. **192:**1487–1497.
12. **Bramkamp, M., L. Weston, R. A. Daniel, and J. Errington.** 2006. Regulated intramembrane proteolysis of FtsL protein and the control of cell division in *Bacillus subtilis*. Mol. Microbiol. **62:**580–591.
13. **Braun, V., S. Mahren, and A. Sauter.** 2006. Gene regulation by transmembrane signaling. Biometals **19:**103–113.
14. **Campo, N., and D. Z. Rudner.** 2006. A branched pathway governing the activation of a developmental transcription factor by regulated intramembrane proteolysis. Mol. Cell **23:**25–35.
15. **Carlsson, K. E., J. Liu, P. J. Edqvist, and M. S. Francis.** 2007. Extracytoplasmic-stress-responsive pathways modulate type III secretion in *Yersinia pseudotuberculosis*. Infect. Immun. **75:**3913–3924.
16. **Castanié-Cornet, M.-P., K. Cam, B. Bastiat, A. Cros, P. Bordes, and C. Gutierrez.** 2010. Acid stress response in *Escherichia coli*: mechanism of regulation of *gadA* transcription by RcsB and GadE. Nucleic Acids Res. **38:**3546–3554.
17. **Cheng, C. Y., S. Y. Shieh, C. C. Hsu, and M. T. Yang.** 2008. Characterization and transcriptional analysis of an ECF sigma factor from *Xanthomonas campestris* pv. *campestris.* FEMS Microbiol. Lett. **289:**250–257.
18. **Clausen, T., C. Southan, and M. Ehrmann.** 2002. The HtrA family of proteases: implications for protein composition and cell fate. Mol. Cell **10:**443–455.
19. **Costanzo, A., and S. E. Ades.** 2006. Growth phase-dependent regulation of the extracytoplasmic stress factor, σ^E, by guanosine 3′,5′-bispyrophosphate (ppGpp). J. Bacteriol. **188:**4627–4634.
20. **Crouch, M. L., L. A. Becker, I. S. Bang, H. Tanabe, A. J. Ouellette, and F. C. Fang.** 2005. The alternative sigma factor σ^E is required for resistance of *Salmonella enterica* serovar Typhimurium to anti-microbial peptides. Mol. Microbiol. **56:**789–799.
21. **Darsonval, A., A. Darrasse, D. Meyer, M. Demarty, K. Durand, C. Bureau, C. Manceau, and M. A. Jacques.** 2008. The type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. Appl. Environ. Microbiol. **74:**2669–2678.
22. **Dartigalongue, C., H. Loferer, and S. Raina.** 2001. EcfE, a new essential inner membrane protease: its role in the regulation of heat shock response in *Escherichia coli*. EMBO J. **20:**5908–5918.
23. **da Silva, A. C. R., J. A. Ferro, F. C. Reinach, C. S. Farah, L. R. Furlan, et al.** 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature **417:**459–463.
24. **da Silva Neto, J. F., T. Koide, S. L. Gomes, and M. V. Marques.** 2007. The single extracytoplasmic-function σ factor of *Xylella fastidiosa* is involved in the heat shock response and presents an unusual regulatory mechanism. J. Bacteriol. **189:**551–560.
25. **Dombrecht, B., J. Vanderleyden, and J. Michiels.** 2001. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol. Plant Microbe Interact. **14:**426–430.
26. **Egler, M., C. Grosse, G. Grass, and D. H. Nies.** 2005. Role of the extracytoplasmic function protein family sigma factor RpoE in metal resistance of *Escherichia coli*. J. Bacteriol. **187:**2297–2307.
27. **Ehrmann, M., and T. Clausen.** 2004. Proteolysis as a regulatory mechanism. Annu. Rev. Genet. **38:**709–724.
28. **Erickson, J. W., and C. A. Gross.** 1989. Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternative sigma factor involved in high-temperature gene expression. Genes Dev. **3:**1462–1471.
29. **Grall, N., J. Livny, M. Waldor, M. Barel, A. Charbit, and K. L. Meibom.** 2009. Pivotal role of the *Francisella tularensis* heat-shock sigma factor RpoH. Microbiology **155:**2560–2572.
30. **Gürlebeck, D., F. Thieme, and U. Bonas.** 2006. Type III effector proteins from the plant pathogen *Xanthomonas* and their role in the interaction with the host plant. J. Plant Physiol. **163:**233–255.
31. **Hayden, J. D., and S. E. Ades.** 2008. The extracytoplasmic stress factor, σ^E, is required to maintain cell envelope integrity in *Escherichia coli*. PLoS One **3:**e1573.
32. **Heinrich, J., K. Hein, and T. Wiegert.** 2009. Two proteolytic modules are involved in regulated intramembrane proteolysis of *Bacillus subtilis* RsiW. Mol. Microbiol. **74:**1412–1426.
33. **Helmann, J. D.** 1999. Anti-sigma factors. Curr. Opin. Microbiol. **2:**135–141.
34. **Helmann, J. D.** 2006. Deciphering a complex genetic regulatory network: the *Bacillus subtilis* sigmaW protein and intrinsic resistance to antimicrobial compounds. Sci. Prog. **89:**243–266.
35. **Hirano, S. S., and C. D. Upper.** 1983. Ecology and epidemiology of foliar bacterial plant pathogens. Annu. Rev. Phytopathol. **21:**243.
36. **Ho, E. M., H. W. Chang, S. I. Kim, H. Y. Kahng, and K. H. Oh.** 2004. Analysis of TNT (2,4,6-trinitrotoluene)-inducible cellular responses and stress shock proteome in *Stenotrophomonas* sp. OK-5. Curr. Microbiol. **49:** 346–352.
37. **Huang, L. H., Y. H. Tseng, and M. T. Yang.** 1998. Isolation and characterization of the *Xanthomonas campestris rpoH* gene coding for a 32-kDa heat shock sigma factor. Biochem. Biophys. Res. Commun. **244:**854–860.
38. **Janaszak, A., B. Nadratowska-Wesołowska, G. Konopa, and A. Taylor.** 2009. The P1 promoter of the *Escherichia coli rpoH* gene is utilized by σ^{70}-RNAP or σ^S-RNAP depending on growth phase. FEMS Microbiol. Lett. **291:**65–72.
39. **Kado, C. I., and M. G. Heskett.** 1970. Selective media for isolation of *Agrobacterium*, *Corynebacterium*, *Erwinia*, *Pseudomonas*, and *Xanthomonas*. Phytopathology **60:**969–976.
40. **Kim, D. Y., E. Kwon, J. Choi, H. Y. Hwang, and K. K. Kim.** 2010. Structural basis for the negative regulation of bacterial stress response by RseB. Protein Sci. **19:**1258–1263.
41. **Koo, B. M., V. A. Rhodius, E. A. Campbell, and C. A. Gross.** 2009. Dissection of recognition determinants of *Escherichia coli* σ^{32} suggests a composite −10 region with an ′extended −10′ motif and a core −10 element. Mol. Microbiol. **72:**815–829.
42. **Kovacikova, G., and K. Skorupski.** 2002. The alternative sigma factor σ^E plays an important role in intestinal survival and virulence in *Vibrio cholerae*. Infect. Immun. **70:**5355–5362.
43. **Luo, Y., and J. D. Helmann.** 2009. Extracytoplasmic function sigma factors with overlapping promoter specificity regulate sublancin production in *Bacillus subtilis*. J. Bacteriol. **191:**4951–4958.
44. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
45. **Misra, H. S., N. P. Khairnar, A. Barik, K. Indira Priyadarsini, H. Mohan, and S. K. Apte.** 2004. Pyrroloquinoline-quinone: a reactive oxygen species scavenger in bacteria. FEBS Lett. **578:**26–30.

46. **Missiakas, D., M. P. Mayer, M. Lemaire, C. Georgopoulos, and S. Raina.** 1997. Modulation of the *Escherichia coli* σ^E (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. Mol. Microbiol. **24:**355–371.
47. **Mo, E., S. E. Peters, C. Willers, D. J. Maskell, and I. G. Charles.** 2006. Single, double and triple mutants of *Salmonella enterica* serovar Typhimurium *degP* (*htrA*), degQ (*hhoA*) and *degS* (*hhoB*) have diverse phenotypes on exposure to elevated temperature and their growth *in vivo* is attenuated to different extents. Microb. Pathog. **41:**174–182.
48. **Muller, C., I. S. Bang, J. Velayudhan, J. Karlinsey, K. Papenfort, J. Vogel, and F. C. Fang.** 2009. Acid stress activation of the σ^E stress response in *Salmonella enterica* serovar Typhimurium. Mol. Microbiol. **71:**1228–1238.
49. **Nollen, E. A., and R. I. Morimoto.** 2002. Chaperoning signaling pathways: molecular chaperones as stress-sensing 'heat shock' proteins. J. Cell Sci. **115:**2809–2816.
50. **Nonaka, G., M. Blankschien, C. Herman, C. A. Gross, and V. A. Rhodius.** 2006. Regulon and promoter analysis of the *E. coli* heat-shock factor, σ^{32}, reveals a multifaceted cellular response to heat stress. Genes Dev. **20:**1776–1789.
51. **Oke, V., and S. R. Long.** 1999. Bacterial genes induced within the nodule during the Rhizobium-legume symbiosis. Mol. Microbiol. **32:**837–849.
52. **Onsando, J. M.** 1992. Black rot of crucifers, p. 243–252. *In* H. S. Chaube, J. Kumar, A. N. Mukhopadhyay, and U. S. Singh (ed.), Plant diseases of international importance. II. Diseases of vegetable and oil seed crops. Prentice Hall, Englewood Cliffs, NJ.
53. **Osborne, S. E., and B. K. Coombes.** 2009. RpoE fine tunes expression of a subset of SsrB-regulated virulence factors in *Salmonella enterica* serovar Typhimurium. BMC Microbiol. **9:**45.
54. **Papenfort, K., V. Pfeiffer, F. Mika, S. Lucchini, J. C. D. Hinton, and J. Vogel.** 2006. σ^E dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. Mol. Microbiol. **62:**1674–1688.
55. **Qiu, D., V. M. Eisinger, D. W. Rowen, and H. D. Yu.** 2007. Regulated proteolysis controls mucoid conversion in *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. U. S. A. **104:**8107–8112.
56. **Raivio, T. L., and T. J. Silhavy.** 2001. Periplasmic stress and ECF sigma factors. Annu. Rev. Microbiol. **55:**591–624.
57. **Reiling, S. A., J. A. Jansen, B. J. Henley, S. Singh, C. Chattin, M. Chandler, and D. W. Rowen.** 2005. Prc protease promotes mucoidy in *mucA* mutants of *Pseudomonas aeruginosa*. Microbiology **151:**2251–2261.
58. **Rezuchova, B., and J. Kormanec.** 2001. A two-plasmid system for identification of promoters recognized by RNA polymerase containing extracytoplasmic stress response σ^E in *Escherichia coli*. J. Microbiol. Methods **45:**103–111.
59. **Rhodius, V. A., W. C. Suh, G. Nonaka, J. West, and C. A. Gross.** 2006. Conserved and variable functions of the σ^E stress response in related genomes. PLoS Biol. **4:**e2.
60. **Rowley, G., M. Spector, J. Kormanec, and M. Roberts.** 2006. Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. Nat. Rev. Microbiol. **4:**383–394.
61. **Schäfer, A., A. Tauch, W. Jäger, J. Kalinowski, G. Thierbach, and A. Pühler.** 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene **145:**69–73.
62. **Serrania, J., F.-J. Vorhölter, K. Niehaus, A. Pühler, and A. Becker.** 2008. Identification of *Xanthomonas campestris* pv. *campestris* galactose utilization genes from transcriptome data. J. Biotechnol. **135:**309–317.
63. **Slamti, L., J. Livny, and M. K. Waldor.** 2007. Global gene expression and phenotypic analysis of a *Vibrio cholerae rpoH* deletion mutant. J. Bacteriol. **189:**351–362.
64. **Smyth, G. K.** 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. **3:**3.
65. **Staroń, A., H. J. Sofia, S. Dietrich, L. E. Ulrich, H. Liesegang, and T. Mascher.** 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. Mol. Microbiol. **74:**557–581.
66. **Strauch, K. L., and J. Beckwith.** 1988. An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. Proc. Natl. Acad. Sci. U. S. A. **85:**1576–1580.
67. **Testerman, T. L., A. Vazquez-Torres, Y. Xu, J. Jones-Carson, S. J. Libby, and F. C. Fang.** 2002. The alternative sigma factor σ^E controls antioxidant defences required for *Salmonella* virulence and stationary-phase survival. Mol. Microbiol. **43:**771–782.
68. **Turner, P., C. E. Barber, and M. J. Daniels.** 1985. Evidence for clustered pathogenicity genes in *Xanthomonas campestris* pv. *campestris.* Mol. Gen. Genet. **199:**338–343.
69. **Vanaporn, M., P. Vattanaviboon, V. Thongboonkerd, and S. Korbsrisate.** 2008. The *rpoE* operon regulates heat stress response in *Burkholderia pseudomallei*. FEMS Microbiol. Lett. **284:**191–196.
70. **Vorhölter, F. J., S. Schneiker, A. Goesmann, L. Krause, T. Bekel, O. Kaiser, B. Linke, T. Patschkowski, C. Rückert, J. Schmid, V. K. Sidhu, V. Sieber, A. Tauch, S. A. Watt, B. Weisshaar, A. Becker, K. Niehaus, and A. Pühler.** 2008. The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J. Biotechnol. **134:**33–45.
71. **Waller, P. R., and R. T. Sauer.** 1996. Characterization of *degQ* and *degS*, Escherichia coli genes encoding homologs of the DegP protease. J. Bacteriol. **178:**1146–1153.
72. **Walsh, N. P., B. M. Alba, B. Bose, C. A. Gross, and R. T. Sauer.** 2003. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. Cell **113:**61–71.
73. **Wang, L., W. Rong, and C. He.** 2008. Two *Xanthomonas* extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. Mol. Plant Microbe Interact. **21:**555–563.
74. **Wood, L. F., and D. E. Ohman.** 2006. Independent regulation of MucD, an HtrA-like protease in *Pseudomonas aeruginosa*, and the role of its proteolytic motif in alginate gene regulation. J. Bacteriol. **188:**3134–3137.
75. **Wood, L. F., and D. E. Ohman.** 2009. Use of cell wall stress to characterize σ^{22} (AlgT/U) activation by regulated proteolysis and its regulon in *Pseudomonas aeruginosa*. Mol. Microbiol. **72:**183–201.
76. **Yang, T. C., Y. W. Leu, H. C. Chang-Chien, and R. M. Hu.** 2009. Flagellar biogenesis of *Xanthomonas campestris* requires the alternative sigma factors RpoN2 and FliA and is temporally regulated by FlhA, FlhB, and FlgM. J. Bacteriol. **191:**2266–2275.
77. **Zhang, S. S., Y. Q. He, L. M. Xu, B. W. Chen, B. L. Jiang, J. Liao, J. R. Cao, D. Liu, Y. Q. Huang, X. X. Liang, D. J. Tang, G. T. Lu, and J. L. Tang.** 2008. A putative *colR*(XC1049)-*colS*(XC1050) two-component signal transduction system in *Xanthomonas campestris* positively regulates *hrpC* and *hrpE* operons and is involved in virulence, the hypersensitive response and tolerance to various stresses. Res. Microbiol. **159:**569–578.

Supplemental Figure 1: Determination of transcription start sites of a subset of σ^E target promoters by primer extension assays. Total RNA from WT, Δ*rpoE*, Δr*seA*, Δ*mucD* and WT or Δ*rpoE* strains containing control plasmid p917 (-) or σ^E-overexpressing plasmid p917-rpoE (+) were used as templates in primer extension experiments when suited. Total RNAs were obtained from cells incubated at 30°C or after a 60 min. shift at 35°C. Primers K, xcc1535-EXT, xcc0964-EXT, hrpF-EXT, hpa1-EXT, xcc1246, ompW-EXT, pqqA-EXT, xcc0401-EXT, prc-EXT, xcc4186-EXT and xcc3227-EXT were 5' end labeled with ^{32}P and extended with reverse transcriptase to map their corresponding gene promoter sequences. The arrowheads indicate the bands corresponding to the main start sites observed.

Fig. S1


A C G T
wt
ΔrseA
P rseA
mucD
wt
ΔrpoE
ΔrseA
30°C
A C G T
ΔrpoE (-)
ΔrpoE (+)
wt (-)
wt (+)
35°C
XCC1535
prc
hpaI
ompW
XCC3227
XCC4186
XCC0401
ΔmucD
XCC0964
hrpF
XCC1246
pqqA

AUTEUR : Guillaume Déjean

TITRE : Caractérisation et conservation d'un nouveau système CUT associé à l'utilisation du xylane chez *Xanthomonas campestris* pv. *campestris* : implications en écologie microbienne

DIRECTEUR DE THESE : Matthieu Arlat

CO-DIRECTEUR DE THESE : Emmanuelle Lauber

Résumé

La dégradation microbienne de la paroi végétale n'est pas seulement un processus biologique important, elle présente également un intérêt scientifique grandissant pour de nombreuses applications biotechnologiques. Au cours de ces travaux, nous avons identifié le système de dégradation et d'utilisation du xylane, un composant structural majeur des parois végétales, chez les espèces de *Xanthomonas* phytopathogènes. Ce système est requis pour la pathogénicité et nous avons montré qu'il est nécessaire pour la croissance optimale des bactéries à la surface des feuilles des plantes hôtes et non hôtes. L'une de ses particularités est la présence de 2 transporteurs spécifiques de la membrane externe (TBDT, TonB-dependent transporter) qui seraient impliqués dans le transport actif de produits d'hydrolyse du xylane. Enfin, des analyses de génomique comparative ont permis de définir un ensemble de gènes essentiel pour l'utilisation du xylane, conservé chez un grand nombre de bactéries phylogénétiquement très distinctes et nichant dans le sol, les plantes, les systèmes aquatiques ou les systèmes digestifs des animaux. Nos travaux montrent que cet ensemble de gènes est systématiquement associé avec des TBDTs, soulignant l'importance de ces protéines dans l'utilisation du xylane, le second polysaccharide végétal le plus abondant dans la nature.

MOTS-CLES : *Xanthomonas campestris* pv. *campestris*, xylane, Transporteurs TonB-Dépendants, système CUT, pouvoir pathogène, dégradation de la paroi végétale

---

AUTHOR: Guillaume Déjean

TITLE: Characterization and conservation of a new xylan CUT system from *Xanthomonas campestris* pv. *campestris* : insights in microbial ecology

PhD SUPERVISOR: Matthieu Arlat

PhD CO-SUPERVISOR: Emmanuelle Lauber

Abstract

Microbial degradation of plant cell walls is not only an important biological process but it also has a growing scientific interest for many biotechnological applications. In this work, we identified the degradation and utilization system of xylan, a major structural component of plant cell walls, in phytopathogenic *Xanthomonas* species. This system is required for pathogenicity, and we have shown the need of this system for optimal bacterial growth on the leaf surface of host and non host plants. One of the features of this system is the presence of two specific outer membrane transporters (TBDTs, TonB-dependent transporters) that would be involved in active uptake of xylan hydrolysis products. Finally, genomic comparative analysis have identified a set of genes essential for xylan utilization, conserved in many phylogenetically distinct bacteria belonging to diverse habitats such as soil, plants, aquatic systems or digestive tracts systems. Our work shows that this set of genes is systematically associated with TBDTs, confirming the importance of these proteins in xylan utilization, the second most abundant plant polysaccharide in nature.

KEYWORDS: *Xanthomonas campestris* pv. *campestris*, xylan, TonB-Dependent Transporters, CUT system, pathogenicity, plant cell wall degradation

---