



Déterminisme de la spécificité d'hôte et rôle des effecteurs TAL dans l'interaction Xanthomonas - haricot

Mylène Ruh

► To cite this version:

Mylène Ruh. Déterminisme de la spécificité d'hôte et rôle des effecteurs TAL dans l'interaction Xanthomonas - haricot. Biologie végétale. Agrocampus Ouest, 2017. Français. NNT : 2017NSARC131 . tel-02082457

HAL Id: tel-02082457

<https://theses.hal.science/tel-02082457>

Submitted on 28 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AGRO CAMPUS

OUEST

Mylène RUH • 19 décembre 2017

Thèse AGROCAMPUS OUEST
sous le label de l'Université européenne de Bretagne
pour obtenir le grade de
DOCTEUR D'AGROCAMPUS OUEST
Spécialité Sciences agronomiques

ÉCOLE DOCTORALE • Écologie, Géosciences, Agronomie,
ALimentation (EGAAL)

LABORATOIRE D'ACCUEIL • UMR INRA / Agrocampus Ouest /
Université d'Angers Institut de Recherche en Horticulture et
Semences (IRHS) - Équipe EmerSys

**Déterminisme de la
spécificité d'hôte et
rôle des effecteurs TAL
dans l'interaction
Xanthomonas – haricot**

Mathilde BRIARD

Professeure Agrocampus Ouest, UMR IRHS / *présidente du jury*

Valérie GEFFROY

Directrice de recherche Institute of Plant Sciences Paris Saclay /
rapporteur

Laurent NOËL

Directeur de recherche CNRS, UMR LIPM / *rapporteur*

Alia DELLAGI

Maître de conférence AgroParisTech, UMR IJPB / *examineur*

Thomas GUILLEMETTE

Professeur Université d'Angers, UMR IRHS / *examineur*

Boris SZUREK

Chargé de recherche IRD, UMR IPME / *examineur*

Marie-Agnès JACQUES

Directrice de recherche INRA, UMR IRHS / *directrice de thèse*

Nicolas CHEN

Maître de conférence Agrocampus Ouest, UMR IRHS /
co-encadrant

A ma Maman.

A Benjamin.

« Les cicatrices nous rappellent d'où l'on vient.

Elles n'ont pas à dicter où nous allons. »

Harrison Scott

Remerciements

Ecrire des remerciements n'est pas la tâche la plus facile lors de l'écriture d'un manuscrit de thèse, mais c'est sans aucun doute la plus plaisante car c'est une belle occasion d'exprimer sa reconnaissance envers toutes les personnes qui ont contribué, à leur manière, à ce projet de thèse.

Tout d'abord, j'adresse mes remerciements à ceux qui ont financé cette thèse : Angers Loire Métropole pour le projet X-HOT et l'ANR pour le projet CropTAL.

Je remercie les membres de mon jury, Valérie Geffroy et Laurent Noël, les rapporteurs, ainsi que Mathilde Briard, Alia Dellagi, Thomas Guillemette et Boris Szurek, les examinateurs, pour avoir accepté d'évaluer mes travaux de thèse. Je remercie également les membres de mon comité de suivi de thèse, Marie-Anne Barny et Boris Szurek, pour avoir posé un regard éclairé sur mes travaux.

Je tiens tout particulièrement à remercier ma directrice de thèse, Marie-Agnès Jacques, et mon encadrant, Nicolas Chen. Tout d'abord, merci d'avoir pensé à ce projet de recherche qui m'a tout de suite séduite et de m'avoir fait confiance. Grâce à vous, j'ai vécu cette expérience avec beaucoup de plaisir, vous m'avez à la fois guidé et laissé entreprendre mes propres recherches, j'ai pu m'épanouir dans mon travail, apprendre beaucoup de nouvelles choses et avoir le sentiment d'être exactement là où je devais me trouver. Merci pour vos précieuses corrections et relectures riches d'enseignements pour les publications et pour le manuscrit. Merci aussi de m'avoir laissé présenter mes travaux à de nombreuses reprises et notamment à Valencia. Marie-Agnès, merci pour ta disponibilité, malgré un emploi du temps très chargé, pour tes conseils lors de nos réunions et dès que je passais ma tête dans ton bureau et pour nos discussions enrichissantes, où tu m'as souvent convaincu mais où tu m'as aussi laissé te convaincre. Merci pour ton enthousiasme dans les bons moments et pour ton soutien dans les moins bons. Merci pour ton rire communicatif, inimitable et inoubliable. Nicolas, merci pour ton aide et ta générosité durant ces trois ans. Merci pour nos petites réunions, pas toujours planifiées, qui duraient toujours plus longtemps que prévu alors que tu avais beaucoup d'autres choses à faire. Merci pour nos échanges où tu as toujours fait preuve d'enthousiasme et où ton esprit foisonnait toujours de mille et une idées. Merci pour tes traits d'humour et tes jeux de mots inattendus mais qui survenaient toujours à point nommé.

Un immense merci à Sophie B. Merci d'avoir toujours été présente quand j'avais « juste une petite question ». Merci pour ton investissement dans ce projet alors que ce n'était pas le seul pour toi. Je présume que... ;-) non je suis sûre que je n'oublierai pas nos franches rigolades où les larmes te montaient aux yeux.

Merci à Martial, pro de la bio-info sans qui l'équipe ne tournerait plus tout à fait rond, à Anne, experte des inoculations sur plante, des comptages bactériens et des tableaux croisés dynamiques et à Armelle, spécialiste des manip à grande envergure. Merci à vous trois, vous avez toujours répondu présent dès que j'ai eu besoin de vous et vous m'avez beaucoup appris. Merci aussi à Daniel, pour avoir toujours pris soin de mes petits haricots, à Céline, pour ton aide sous R lors des analyses d'imagerie, et à Jean-François, pour tous tes conseils qui m'ont été très utiles et pour nos conversations quand on travaillait sous la hotte.

Merci à tout le reste de l'équipe EmerSys : Chrystelle, Karine, Géraldine, et Sophie C. pour votre aide dans vos différents domaines d'expertises et pour toutes les commandes que j'ai pu passer grâce à vous, Audrey pour tout ce que tu fais pour tout le monde malgré toutes tes galères avec l'autoclave et pour penser à tout le monde au moment du déjeuner, Cécile, Perrine et Marion pour toutes les souches que j'ai commandé à la CFBP et votre expertise en taxonomie et pour les règles liées à nos petits micro-organismes, Matthieu et Tristan, pour vos conseils avisés avec toujours une petite pointe d'humour, Quentin et Anne-Lise, pour toutes ces petits échanges chaleureux entre deux manip, je vous souhaite à tous les deux une belle route vers de nouvelles aventures et Brice et Coralie, vous êtes les derniers arrivés et nous n'avons pas encore eu beaucoup de temps pour apprendre à se connaître alors je vous souhaite de bons moments dans cette super équipe.

Merci aussi à tous les doctorants, post-docs et stagiaires que j'ai côtoyé au cours de ces trois années. Samir, mon coloc de bureau, je suis ravie d'avoir partagé cette expérience avec toi, j'espère que nos chemins se croiseront par la suite pour qu'on se rappelle ensemble de tout ça. Valérian, je me rappelle du jour de ton entretien, j'ai tout de suite su qu'on s'entendrait bien et j'ai eu raison. Merci pour toutes nos conversations sérieuses ou non, particulièrement ces dernières semaines, je suis sûre que tu feras une super troisième année et qu'un bel avenir t'attend. Justine, ma stagiaire préférée, merci de m'avoir aidé à deux reprises et d'avoir rendu cette expérience d'encadrement agréable. Je suis ravie de te passer le flambeau des TALE et te souhaite une belle thèse. Nicolas D., merci pour ton expertise sur les TALE et pour ton généreux don de publications à ce sujet dès ton arrivée, et Gloria, merci de toujours porter un grand intérêt à ce que fait toute l'équipe, tu impressionnes tout le monde avec ton niveau en français, je vous souhaite à tous les deux un avenir plein de belles surprises aussi bien sur le plan professionnel que sur le plan personnel avec vos petites familles. Merci à Déborah, Amandine et Arnaud, vous avez été un exemple à suivre pour moi. Caroline, Enora et Felipe, vous venez tout juste d'arriver dans l'équipe et je vous souhaite bonne chance pour vos thèses et post-doc respectifs. Merci à Lucien, Romain, Ellen, Elodie et Julia, doctorants de l'IRHS que j'ai eu plaisir de rencontrer et avec qui j'ai partagé de bons moments de partage que ce soit à la cafétéria, dans le labo, dans le bus ou à la salle de sport. Merci aussi à Nassima et Vincent, doctorants Agreenium (la

crème du gratin), j'ai adoré vous croiser lors des séminaires de notre promo mais aussi en dehors et j'espère très sincèrement que nos chemins se croiseront encore.

Je remercie tous ceux du bâtiment C avec qui j'ai pu interagir. Merci à Christophe, Ellen, Ludovic, Thibaut, Marie-Noëlle Be., Mélanie et Jason, accompagnés de Chrystelle, Valérian, Anne, Armelle, Tristan, Déborah, Amandine et Enora, qui ont fait de ces pauses déjeuners à la cafétéria de vrais moments de détente où la bonne humeur et l'humour sont au rendez-vous. Merci à Alexandre, Mathieu, Romain, Marie-Noëlle Br. et Raphaël qui me sont venus en aide quand j'en ai eu besoin.

Je souhaite aussi remercier les membres du projet CropTAL, j'ai adoré faire partie de ce projet. Laurent, Lionel, Boris et Nicolas merci d'avoir pensé et conçu le projet CropTAL tout simplement. Merci aussi pour tous les échanges, les entraides et les partages de connaissances. Sandrine et Tu merci pour cette super formation sur les ArtTAL, Alvaro merci pour ton aide et tes supers outils bio-informatiques toujours très bien nommés et Sébastien merci pour le traitement de mes données RNAseq. Un merci tout particulier à Sophie B., Sandrine et Ivanna pour votre persévérance pour le clonage des TALE.

Merci aussi aux membres de l'IRD et du CIRAD de Montpellier, Sandrine, Tu, Alvaro, Céline, Mathilde, Kossi, Carolina, Daniel, Servane, Ralf, Hinda et Florence, de m'avoir accepté comme l'un des vôtres pendant deux semaines. C'était un vrai plaisir de vous rencontrer et de vous recroiser par la suite, j'espère que ça arrivera encore.

Je remercie également les membres de l'institut de génétique végétale de l'université Leibniz à Hanovre de m'avoir si chaleureusement accueilli pendant ces trois mois, malgré le froid. Merci à Jens, Sebastian et Jana pour votre aide au quotidien, merci aussi à Stephanie, Maik, Max, Claudia, Luisa, Thomas, Gerrich, Merlin, Stanislaus, Ana et Maren pour tous ces échanges dans les labo et à la mensa. Merci pour cette journée où nous avons célébré Noël avant l'heure. Je vous ai quitté avec beaucoup d'émotion et j'espère très bientôt vous revoir.

Merci à Patricia Vandaële, Patricia Raffoux, Magali Tabuteau, Heidi Eppe, Marie-Pierre Argoud, Pierre Langlais et Mathilde Ricou pour votre aide pour tous les aspects administratifs de l'INRA. Un merci tout particulier à Marie-Pierre Louage, qui ne s'arrête jamais de travailler, et à Sylvie Guyader, vous êtes vraiment aux petits soins pour tous les doctorants Agrocampus Ouest. Merci aussi à Margaux Caborderie et Cécile Dubas pour votre aide avec le dossier ERASMUS.

A tous ceux que j'ai pu croiser durant ces trois années et que je n'ai pas cités, désolée, mais sachez que je n'en pense pas moins : merci.

Sur un plan plus personnel, je tiens à remercier mes fidèles amis et leurs moitiés : Natacha, Amélie, Henri, Cécile, Cindy, Nicolas, Vincent et Megumi. Malgré les années et la distance nous restons toujours en contact je vous remercie pour ça. Durant ces trois années, il y a eu beaucoup de grands événements et de beaux souvenirs pour notre petit groupe. Je suis ravie pour vous tous et d'avoir (presque) tout partagé avec vous. J'espère qu'il y en aura encore beaucoup d'autres.

Il n'y a pas de mot assez fort pour exprimer toute ma gratitude et mon respect pour mon héroïne : ma maman. Maman, je ne te dirai jamais assez combien je t'aime et que tu comptes pour moi, alors je t'écris ces quelques mots qui nous survivront. Merci pour tout ce que tu m'as apporté, pour ton soutien inconditionnel, ton amour infini et d'être ce que tu es, la maman parfaite pour moi. Merci de m'avoir encouragé à toujours aller plus loin, sans toi je me serai sans doute arrêtée avant, et de m'avoir appris tout ce qu'il faut savoir pour être quelqu'un de bien. J'espère un jour faire preuve d'autant de force, de dignité et de courage que toi. Tu as toujours été là pour moi, prête à déplacer des montagnes, alors sache que jamais je ne t'abandonnerai, même s'il y a de la distance entre nous, je serai toujours là pour toi, prête à décrocher des étoiles. Je t'aime.

Et je terminerai ces quelques mots pour remercier ma moitié, Benjamin. Merci d'être entré dans ma vie il y a maintenant 10 ans. Tu es tout pour moi, mon âme sœur, mon meilleur ami, celui qui cherche toujours à me séduire, à me faire rire et à me faire plaisir. Je t'ai trouvé alors je ne te lâche plus. Merci d'avoir accepté la distance et pour tous ces week-ends que nous avons partagé. Merci de faire de moi ta seule et unique préoccupation dès que nous sommes ensemble. Merci de croire en moi. J'aime voir qui je suis dans tes yeux et j'espère te donner autant que ce que tu m'apportes. Une nouvelle page se tourne et je suis impatiente de vivre à tes côtés, au quotidien. Nous avons grandi et évolué ensemble et j'espère que nous allons vieillir ensemble, je t'aime.

TABLE DES MATIÈRES

ABRÉVIATIONS	v
---------------------	----------

INTRODUCTION	1
---------------------	----------

I. Introduction générale	3
---------------------------------	----------

II. Synthèse bibliographique	5
-------------------------------------	----------

1. L'interaction hôte-bactérie phytopathogène	5
---	---

1.1. L'attraction	7
-------------------	---

1.2. L'installation	11
---------------------	----

1.3. La multiplication	15
------------------------	----

1.4. La dispersion	25
--------------------	----

2. L'adaptation et la spécificité d'hôte chez les <i>Xanthomonas</i>	31
--	----

2.1. Les gènes candidats impliqués dans la spécificité l'hôte	33
---	----

2.2. La notion de répertoire	39
------------------------------	----

2.3. Les événements évolutifs mis en jeu	41
--	----

3. Les effecteurs TAL	51
-----------------------	----

3.1. Structure protéique des effecteurs TAL	51
---	----

3.2. Le code responsable de l'interaction effecteur TAL – ADN cible	55
---	----

3.3. La mission première des effecteurs TAL: jouer le rôle d'effecteur de virulence	61
---	----

3.4. Leur implication dans la réaction de défense de la plante	63
--	----

3.5. Les outils biotechnologiques et de lutte	67
---	----

4. Le pathosystème <i>Phaseolus vulgaris</i> – <i>Xanthomonas</i> spp.	73
--	----

III. Projet de thèse	105
-----------------------------	------------

CHAPITRE 1 : La spécificité d'hôte chez les agents responsables de la graisse commune du haricot	109
---	------------

I. Introduction	111
------------------------	------------

II. Comparative genomics analyses identified a subset of genes associated with host specificity in <i>Xanthomonas</i> strains pathogenic on common bean.	115
III. Perspectives et discussion	155

CHAPITRE 2 : Diversité et évolution des gènes *tal* des agents responsables de la graisse commune du haricot et d'autres pathovars de *Xanthomonas*

I. Introduction	159
II. Implication des effecteurs TAL dans l'adaptation des <i>Xanthomonas</i> au haricot commun	161
III. Contenu génomique d'une diversité de souches de <i>X. citri</i> pv. <i>vignicola</i>	209
IV. Données et résultats complémentaires	215
V. Perspectives et discussion	221

CHAPITRE 3 : TAL18H : un effecteur TAL atypique partagé entre des lignées phylogénétiquement éloignées de *Xcf-Xpp*

I. Introduction	227
II. Deciphering the role of TAL18H effector in the interaction between the common bacterial blight agents and common bean	231
III. Résultats complémentaires et perspectives	273
1. Expériences complémentaires à réaliser pour caractériser le rôle de TAL18H dans l'interaction <i>Xanthomonas</i> / haricot	273
2. Différences dans l'interaction <i>Xanthomonas</i> / haricot entre les deux variétés de haricot BAT93 et JaloEEP558	275
3. Analyse d'une potentielle cible intergénique de TAL18H	277
4. Production de mutants permettant l'étude de l'ensemble des effecteurs TAL des souches responsables de CBB	281

DISCUSSION GÉNÉRALE_____ 285

I. Aspects méthodologiques et limites de l'étude de la spécificité d'hôte___ 287

1. Importance du jeu de données et des méthodes d'analyse_____ 287
2. Validation fonctionnelle de gènes candidats _____ 291

II. L'adaptation des *Xanthomonas* au haricot commun : un processus évolutif en plusieurs étapes_____ 293

1. Importance du transfert horizontal _____ 293
2. Scénarios évolutifs pour l'acquisition de la spécificité d'hôte des *Xanthomonas* au haricot commun_____ 293

III. Spécificités adaptatives des effecteurs TAL _____ 299

1. L'adaptation est inhérente à la fonction des effecteurs TAL _____ 299
2. Spécificités adaptatives de TAL18H _____ 301

IV. Développement de la lutte contre la graisse commune du haricot_____ 305

1. De nouveaux outils d'épidémiosurveillance ? _____ 305
2. Quelle durabilité pour les résistances basées sur les effecteurs TAL ? _____ 305

CONCLUSION GÉNÉRALE_____ 311

RÉFÉRENCES BIBLIOGRAPHIQUES_____ 315

ABRÉVIATIONS

AD : acidic Activation Domain

ADN : Acide DésoxyriboNucléique (**DNA** : DeoxyriboNucleic Acid)

AFLP : Amplified Fragment Length Polymorphism

AIM : Agrobacterium Infiltration Medium

Amp^R : Résistance à l'Ampicilline

ANI : Average Nucleotide Identity

ARN : Acide RiboNucléique (**RNA** : RiboNucleic Acid)

ARR : Arabidopsis Response Regulator

ArtTAL ou **artTAL** : *tal* ou effecteur TAL artificiel

A. thaliana : *Arabidopsis thaliana*

A. tumefaciens : *Agrobacterium tumefaciens*

Avr ou **avr** : avirulence

CBB : Common Bacterial Blight of bean (graisse commune du haricot)

CDS : CoDing Sequence

CFBP : Collection Française de Bactéries associées aux Plantes

CIAT : Centro Internacional de Agricultura Tropical

CRSP : Collaborative Research Support Program

C-ter : C-terminal

DO : Densité Optique (**OD** : Optique Density)

dpi : day post-inoculation

EBE : Effector Binding Element

E. coli : *Escherichia coli*

EFSA : European Food Safety Authority

EPPO : European and Mediterranean Plant Protection Organization (**OEPP** : Organisation Européenne et Méditerranéenne pour la Protection des Plantes)

EPS : ExoPolySaccharide

ETI : Effector-Triggered Immunity

ETS : Effector-Triggered Susceptibility

ET3 : Effecteur de Type 3 (**T3E** : Type 3 Effector)

FAO : Food and Agriculture Organization of the United Nations

Genta^R : Résistance à la Gentamycine

gFM : gram of Fresh Material (**gMF** : gramme de Matière Fraîche)

GL : Genetic Lineage

GUS : β -glucuronidase

HAMP : Histidine kinases, Adenylyl cyclases Methyl-accepting chemotaxis proteins and Phosphatases

HR : Hypersensitive Response (ou Reaction)

hra : *hrp* associated

hrc : *hrp* conserved

hrp : hypersensitive response and pathogenicity

INRA : Institut National de la Recherche Agronomique

IRHS : Institut de Recherche en Horticulture et Semences

IS : Insertion Sequence

ITS : InTergenic Spacer

Kana^R : Résistance à la Kanamycine

Ks : value of pairwise nucleotide synonymous substitution rates at silent sites

LOB : Lateral Organ Boundaries

logFC : log de Fold Change

LPS : LipoPolySaccharide

MAMP : Microbe-Associated Molecular Pattern

MAPK : Mitogen-Associated Protein Kinase

MCP : Methyl-accepting Chemotaxis Protein

MIC : Cassette d'Insertion Mobile

ML : Maximum Likelihood

MUG : 4-Methylumbelliferyl- β -D-Glucuronide Trihydrate

N. benthamiana : *Nicotiana benthamiana*

NDH : NAD(P)H Déshydrogénase

NF : Non-Fuscos

N.I. : Non Inoculated

NJ : Neighbor Joining

NLS : Nuclear Localisation Signal

N-ter : N-terminal

PAMP : Pathogen-Associated Molecular Pattern

PCR : Polymerase Chain Reaction

PIP : Plant Inducible Promoter

PL : Pectate Lyase

PRR : Pattern Recognition Receptor

PTI : PAMP-Triggered Immunity

pv. : pathovar

P. vulgaris : *Phaseolus vulgaris*

qRT-PCR : quantitative Reverse Transcription PCR

R : Résistance

R. solanacearum : *Ralstonia solanacearum*

RVD : Repeat Variable Diresidue

S : Susceptibilité ou Sensibilité

SMRT : Single Molecule, Real-Time

SNP : Single-Nucleotide Polymorphism ou polymorphisme nucléotidique

sST3 : signal de Sécrétion de Type 3

SST3/2/4 : Système de Sécrétion de Type 3/2/4 (**T3SS** : Type 3 Secretion System)

subsp. : subspecies

TAL : Transcription Activator-Like

tal : gène codant un effecteur TAL

TALEN : TALE Nuclease

TBDT : Transporteur TonB-Dépendant

THG : Transfert Horizontal de Gène (**HGT** : Horizontal Gene Transfer)

Tn-seq : Transposon sequencing

UFC : Unité Formant Colonie (**CFU** : Colony Forming Unit)

upa : up-regulated by AvrBs3

Vir ou ***vir*** : virulence

WT : Wild-Type

X. : *Xanthomonas*

Xcf : *Xanthomonas citri* pv. *fuscans*

Xpp : *Xanthomonas phaseoli* pv. *phaseoli*

INTRODUCTION

I. Introduction générale

Le secteur des productions végétales est en pleine mutation et doit relever des défis majeurs tels que l'explosion démographique ou le changement climatique. En même temps, les maladies des plantes d'intérêt agronomique sont de véritables fléaux et peuvent engendrer annuellement d'importantes pertes économiques partout dans le monde, pertes pouvant être responsables de considérables crises socio-économiques. En particulier, il est difficile de lutter contre les maladies causées par des bactéries en raison d'un manque de traitements curatifs efficaces et non toxiques. Pour ces raisons, le développement de nouvelles méthodes de lutte contre les agents phytopathogènes est nécessaire. Notamment, le développement de résistances durables pour les plantes cultivées (lutte génétique) est l'un des challenges des prochaines décennies (Dangl *et al.*, 2013).

Les bactéries du genre *Xanthomonas* sont responsables de maladies sur plusieurs centaines d'espèces végétales. Cependant, chaque souche individuelle n'attaque qu'un nombre très limité d'espèces végétales (Vauterin *et al.*, 1995; Jacques *et al.*, 2016). Cela indique qu'il existe des barrières efficaces et durables chez les plantes non-hôtes, et que seules certaines souches possèdent l'arsenal adéquat pour passer ces barrières (Senthil-Kumar & Mysore, 2013). Identifier les déterminants de la spécificité d'hôte permettrait donc de proposer des méthodes de lutte génétique spécifiques et durables ayant peu d'effets sur les autres membres des biotopes.

Le haricot commun est cultivé sur tous les continents en tant que culture vivrière ou commerciale et est le légume sec le plus consommé à travers le monde. Sa production peut subir jusqu'à 40 % de pertes à cause d'une maladie bactérienne nommée la graisse commune du haricot. Cette bactériose est causée par deux espèces distinctes de bactéries : *Xanthomonas phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans*. Ainsi, les bactéries responsables de la graisse commune du haricot, bien qu'ayant la même spécificité d'hôte, forment un pathovar de souches polyphylétiques, ce qui suggère une convergence pathologique de ces souches probablement via l'acquisition de déterminants communs de la spécificité d'hôte.

Le projet de recherche développé au cours de ma thèse et rapporté dans ce manuscrit traite de la spécificité d'hôte des agents pathogènes infectant le haricot commun en mettant la lumière sur une classe d'effecteurs de type III caractéristiques du genre *Xanthomonas* : les effecteurs TAL (Transcription Activator-Like). Ces effecteurs sont des déterminants majeurs du pouvoir pathogène des souches qui en possèdent et de potentiels déterminants de la spécificité d'hôte. La synthèse bibliographique portera sur (i) l'interaction hôte-bactérie phytopathogène, (ii) l'adaptation et la spécificité d'hôte chez les *Xanthomonas*, (iii) les effecteurs TAL et (iv) le pathosystème haricot-*Xanthomonas*.

Type d'interaction	NON-HÔTE	HÔTE	
		INCOMPATIBLE	COMPATIBLE
Plante	Immunisée	Résistante	Sensible
Pathogène	Avirulent	Avirulent	Virulent
Défense de la plante	Barrières de défenses constitutives Défenses basales : PTI	Défenses basales : PTI Protéines de résistance activant une réponse de défense spécifique : ETI	Défenses basales : PTI
Mécanisme d'interaction associé	Barrières physique et biochimique MAMP détectés	Effecteurs avirulents reconnus : relation gène pour gène	Interactions spécifiques entre effecteurs / toxines et leurs cibles hôtes
Conséquence	Pas de maladie	Pas de maladie ou fortement réduite	Maladie : multiplication et dissémination de l'agent pathogène

Figure 1 : Les différents types d'interaction hôte-pathogène. Cette illustration est adaptée de Hammond-Kosack and Kanyuka 2007. PTI : PAMP-Triggered Immunity ; MAMP : Microbe-Associated Molecular Pattern ; ETI : Effector-Triggered Immunity.

II. Synthèse bibliographique

1. L'interaction hôte-bactérie phytopathogène

Il existe trois types d'interactions hôte-pathogène définis selon l'issue de l'infection : *(i)* l'interaction non-hôte, *(ii)* l'interaction hôte incompatible et *(iii)* l'interaction hôte compatible (Hammond-Kosack & Kanyuka, 2007). *(i)* L'interaction non-hôte correspond à une interaction où, dans n'importe quelle situation, un agent pathogène donné est incapable d'infecter une espèce végétale donnée. On parle aussi de résistance non-hôte, celle-ci permettant de définir la gamme d'hôte des agents pathogènes. *(ii)* L'interaction hôte incompatible correspond à une interaction où une plante est résistante car elle est capable de reconnaître l'agent pathogène et de lui résister spécifiquement. On parle également de résistance spécifique et l'agent pathogène est avirulent. *(iii)* L'interaction hôte compatible correspond à une interaction où une plante est sensible soit parce qu'elle ne parvient pas à reconnaître l'agent pathogène, soit parce que, bien que la plante reconnaisse l'agent pathogène, ce dernier parvient à contourner les défenses de la plante. Dans tous les cas, ce type d'interaction permet le développement de la maladie et l'agent phytopathogène est virulent (**Figure 1**).

Ainsi, dans le cas d'une interaction hôte compatible, l'aptitude d'une bactérie phytopathogène à coloniser une plante serait le résultat d'une adaptation de l'agent pathogène à son hôte. Cette adaptation s'exprimerait tout au long du cycle infectieux composé de plusieurs étapes successives (**Figure 2**). La première et la seconde étapes de ce cycle correspondent à **(i) l'attraction** vers l'hôte puis à **(ii) l'installation** au sein de l'hôte. Ces deux étapes initiant le cycle infectieux constituent la phase asymptomatique. La troisième étape du cycle correspond à **(iii) la multiplication** bactérienne et à l'expression du pouvoir pathogène (Alfano & Collmer, 1996). Cette étape constitue donc la phase symptomatique du cycle infectieux. Ce cycle se termine par la quatrième étape qui correspond à **(iv) la dispersion** des agents pathogènes. Cette dispersion permet la transmission de cet agent pathogène à une nouvelle plante-hôte et/ou à la semence afin qu'un nouveau cycle infectieux démarre. Au cours de ce cycle, les bactéries phytopathogènes sont confrontées à de multiples stress. En effet, dans un environnement tel que la phyllosphère, correspondant aux feuilles et à l'environnement qui leur est soumis, les bactéries, en plus de subir des stress abiotiques tels que des changements de température et d'humidité relative, vont aussi devoir faire face aux défenses mises en place par la plante.

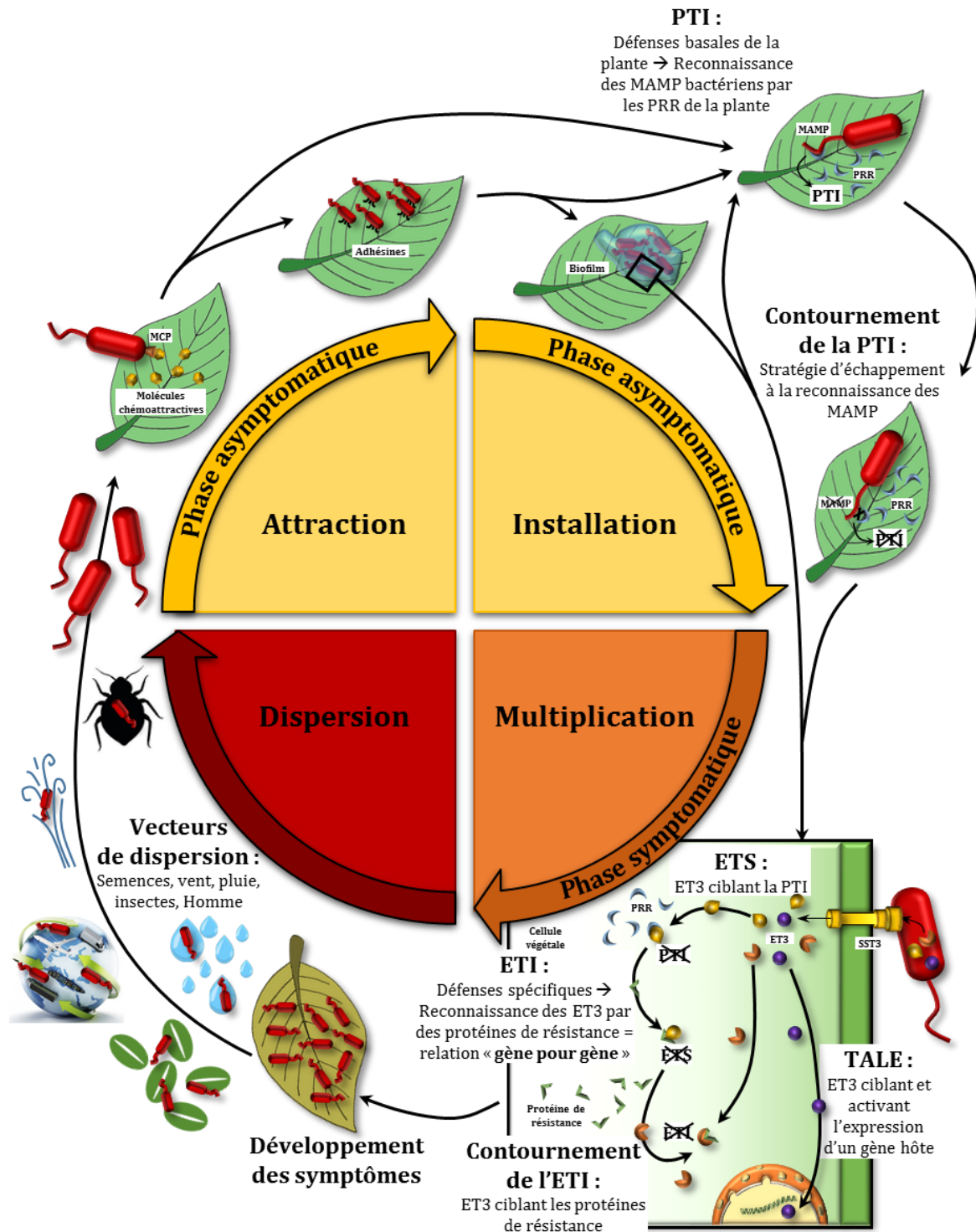


Figure 2: Cycle infectieux des bactéries phytopathogènes. Durant l'étape d'attraction, le chimiotactisme permet à la bactérie de détecter des sites favorables grâce aux MCP (Methyl-accepting chemotaxis protein), lui permettant d'entrer dans l'étape d'installation où le biofilm se forme grâce aux adhésines. La plante peut reconnaître des MAMP (Microbe-Associated Molecular Patterns) grâce à ces PRR (Pattern Recognition Receptors) ce qui déclenche la PTI (PAMP-Triggered Immunity). Le cycle s'arrête alors, à moins que la bactérie puisse contourner la PTI. L'étape de multiplication va dépendre du SST3 (Système de Sécrétion de Type 3) qui injecte des ET3 (Effecteurs de Type 3) dans la cellule hôte. Certains ET3, nommés effecteur TAL (Transcription Activator-Like), activent l'expression de gènes de l'hôte, d'autres ciblent la PTI, on parle alors d'ETS (Effector-Triggered Susceptibility). La plante peut reconnaître ces ET3, ce qui déclenche l'ETI (Effector-Triggered Immunity), que la bactérie peut contourner grâce à d'autres ET3. L'étape de multiplication permet le développement de symptômes de la maladie et la bactérie entre alors dans l'étape de dispersion, notamment grâce au vent, à l'eau, aux insectes ou à l'homme.

1.1. L'attraction

L'étape d'attraction correspond à la première phase du cycle infectieux, où la bactérie va chercher à coloniser les sites favorables à sa survie dès son arrivée sur la plante hôte. En effet, la phyllosphère est, certes un environnement stressant, mais elle comporte aussi une diversité de microhabitats pour les microorganismes, comme par exemple la base des trichomes (Lindow & Brandl, 2003). La surface des feuilles possède également des voies d'entrée pour la bactérie, comme les stomates, les hydathodes ou les blessures, permettant à celle-ci de pénétrer dans la plante (Pérez-Bueno *et al.*, 2015). Par conséquent, lors de cette étape d'attraction, la bactérie a besoin de pouvoir détecter les sites favorables à son installation et ceci est possible grâce au chimiotactisme permettant notamment l'activation de la mobilité de la bactérie (Indiana, 2014).

Le chimiotactisme est défini comme le mouvement dirigé de cellules mobiles dans des gradients de divers paramètres physico-chimiques (Alexandre, 2010). Il permet d'optimiser l'accession aux nutriments (molécules chémoattractives) et l'évitement des molécules toxiques (molécules chémorépulsives) en contrôlant le sens de rotation du moteur flagellaire (Berg, 2003; Wadhams & Armitage, 2004). L'ensemble de ces signaux est détecté au niveau de la membrane plasmique de la bactérie par des chémorécepteurs appartenant à la famille des Methyl-accepting Chemotaxis Proteins ou MCP. Ces chémorécepteurs réagissent alors en déclenchant dans le cytoplasme une cascade de transduction du signal chimiotactique jusqu'au moteur flagellaire afin de contrôler les changements de son sens de rotation. Il existe deux cascades de transduction du signal chimiotactique (**Figure 3**). La première, basée sur la déphosphorylation des protéines kinases CheA et CheY, est spécifique de l'attraction de la bactérie vers une source chémoattractrice et la seconde, basée sur la phosphorylation de ces deux mêmes protéines, est spécifique de sa répulsion d'une source chémorépulsive. En effet, la détection des signaux de l'environnement se fait via la fixation des molécules chémoattractives ou chémorépulsives sur les MCP qui peuvent alors réguler l'activité d'autophosphorylation de CheA positivement ou négativement. Cette activité d'autophosphorylation est inhibée en présence de molécules chémoattractives et activée en absence de ces molécules. Ainsi, dans le contexte d'une répulsion, CheA transfère son groupement phosphate à CheY qui diffuse ensuite librement dans le cytoplasme pour aller se fixer au moteur flagellaire afin de modifier son sens de rotation d'antihoraire à horaire. Ce changement de sens de rotation du moteur flagellaire entraîne la désolidarisation des flagelles les uns des autres et la bactérie culbute, c'est-à-dire qu'elle change de direction (Vladimirov & Sourjik, 2009). La bactérie peut ensuite reprendre une course rectiligne vers une nouvelle direction en changeant à nouveau le sens de rotation du moteur flagellaire, cette fois-ci d'horaire à antihoraire (Belas *et al.*, 2008). Ce sens de rotation permet quant à lui d'organiser les flagelles en un seul faisceau qui pousse la bactérie dans une longue course rectiligne (**Figure 4**).

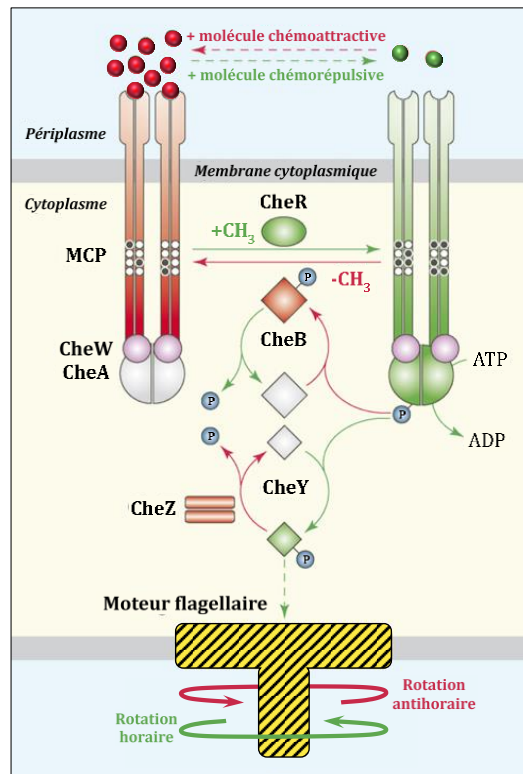


Figure 3 : Cascades de transduction du signal chimiotactique. Ce schéma est adapté de Hazelbauer et al. 2008. Les cascades induites par des molécules chémoattractives et ou par une des molécules chémorépulsives sont respectivement représentées en rouge et en vert. Les composants en gris représentent des formes inactives. Les flèches pleines représentent des réactions enzymatiques et celles en pointillées indiquent des interactions de liaison. Le groupe phosphorylé dérivé de CheA est représenté sous la forme d'une sphère bleue.

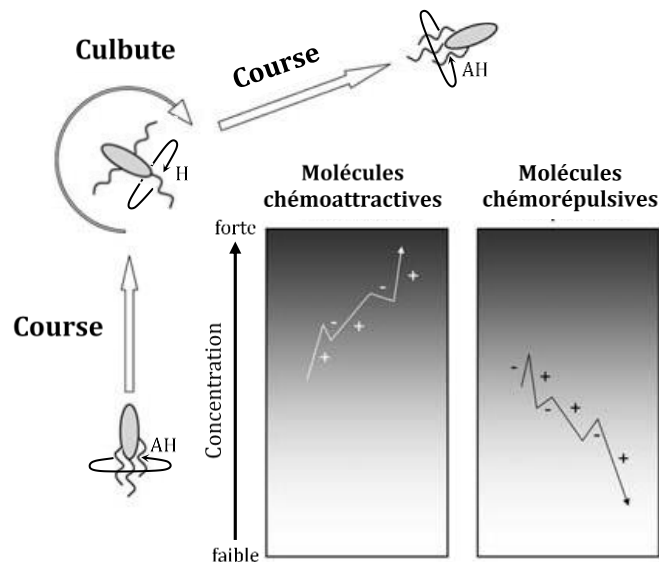


Figure 4 : Chimiotactisme chez *E. coli*. Cette illustration est adaptée de Vladimirov and Sourjik 2009. La bactérie se déplace aléatoirement en effectuant de longues courses rectilignes d'environ 1 seconde entrecoupées de petites culbutes d'environ 0,1 seconde. La direction d'une nouvelle course est choisie au hasard, mais sa durée est plus longue si la bactérie rencontre des concentrations croissantes de molécules attractives ou des concentrations décroissantes de molécules répulsives. Les courses rectilignes sont indiquées par '+' et les courses aléatoires par '-'. Les courses rectilignes sont la conséquence d'une rotation antihoraire (AH) du moteur flagellaire et les culbutes se produisent lorsque le moteur flagellaire est en rotation horaire (H).

Les principaux acteurs moléculaires du chimiotactisme sont donc les MCP qui se composent de différents domaines essentiels pour la perception de l'environnement par la bactérie (**Figure 3**). En effet, un chémorécepteur est constitué d'un ou plusieurs domaines de fixation extracellulaire jouant le rôle de senseur, d'un domaine de signalisation cytoplasmique permettant la transduction du signal chimiotactique et d'un domaine HAMP (pour Histidine kinases, Adenylyl cyclases Methyl-accepting chemotaxis proteins and Phosphatases) contrôlant la phosphorylation du domaine cytoplasmique (Wadhams & Armitage, 2004; Hulko *et al.*, 2006; Alexandre, 2010; Stewart, 2014). Selon le domaine senseur, les MCP se répartissent en quatre types, chacun spécialisé dans la détection d'un type de molécule. Il y a le type Tsr, capable de reconnaître la sérine, le type Tar, pour la détection de l'aspartate et du maltose, le type Trg, pour des sucres comme le ribose, le galactose et le glucose et le type Tap pour les dipeptides et les pyrimidines (Manson *et al.*, 1998; Hazelbauer *et al.*, 2008; Hazelbauer, 2012). Par exemple, chez *X. campestris*, un senseur du chimiotactisme nommé Hsb1 est suspecté de percevoir la sinigrine et la L-phénylalanine, émise par les blessures des feuilles de chou, ce qui permettrait à la bactérie de s'internaliser dans les tissus de la plante hôte (Indiana, 2014). Le rôle primordial de ces MCP dès le début de l'interaction hôte-pathogène a pour conséquence que les gènes codant ces protéines sont soumis à une forte pression de sélection. Cependant, cette pression n'est pas la même selon les domaines qui les composent. En effet, de précédentes analyses ont montré que le domaine cytoplasmique est très conservé alors que le domaine senseur est très variable (Wuichet & Zhulin, 2003; Alexander & Zhulin, 2007). Cette différence de conservation semble être directement liée à la fonction de ces domaines. En effet, comme le domaine senseur est en interaction directe avec une grande diversité de molécules de l'environnement, il subit une pression de sélection positive (diversifiante) qui aboutit à une plus grande diversité au niveau protéique, alors que le domaine cytoplasmique, qui interagit uniquement avec les protéines du chimiotactisme de la bactérie, se doit d'être conservé et subit donc une pression de sélection négative (purifiante) qui aboutit à une conservation des résidus au cours du temps (Wuichet *et al.*, 2007; Krell *et al.*, 2010). D'autre part, les gènes codant les MCP sont majoritairement localisés dans un îlot génomique dédié au chimiotactisme et à la mobilité flagellaire plus ou moins conservé et, chez les bactéries du genre *Xanthomonas*, le nombre de gène codant les MCP est élevé car une grande partie de ces gènes sont dupliqués et répétés en tandem, ce qui n'est pas courant chez les bactéries. Cela suggère donc que, pour le genre *Xanthomonas*, les MCP sont importants pour leur mode de vie (Thieme *et al.*, 2005; Mhedbi-Hajri *et al.*, 2011; Darrasse *et al.*, 2013b).

Lors de cette étape d'attraction, la bactérie doit aussi faire face aux défenses basales de la plante. Une partie de ces défenses est passive et correspond aux barrières protectrices de la plante. Ces barrières sont à la fois physiques et biochimiques, notamment grâce à la paroi cellulaire et/ou la cuticule d'une part et à divers métabolites secondaires d'autre part.

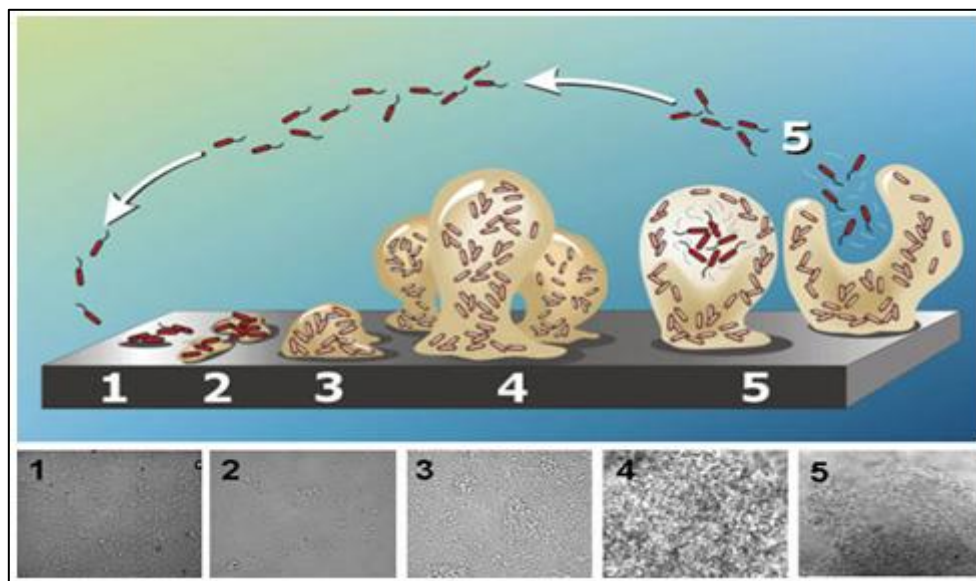


Figure 5 : Formation d'un biofilm en cinq étapes. Cette illustration est adaptée de Stoodley et al. 2002. Etape 1 : Fixation initiale des bactéries à la surface. Etape 2 : Production d'exopolysaccharides favorisant l'attachement irréversible des bactéries. Etape 3 : Développement précoce de l'architecture du biofilm. Etape 4 : Maturation de l'architecture du biofilm. Etape 5 : Dispersion de bactéries isolées à partir du biofilm mature. En dessous, photographies de *Pseudomonas aeruginosa* au microscope électronique à transmission à chaque étape de la formation du biofilm. Les bactéries sont cultivées dans des conditions d'écoulement continu sur un substrat de verre.

Cette première ligne de défense est donc dynamique et constitue un obstacle contre la plupart des agents pathogènes. Ces barrières ne sont pas infranchissables, et parfois des agents phytopathogènes parviennent à les passer. Cependant, dès leur entrée dans les tissus végétaux, ils devront faire face à un deuxième niveau de défenses basales correspondant au système immunitaire inné de la plante qui lui permet d'enclencher immédiatement des mécanismes de défense contre les agents pathogènes (Grennan, 2006; Zipfel, 2008; Doughari, 2015). Ce système immunitaire est détaillé plus loin (1.3 La multiplication). Afin de se préparer à l'assaut et déborder efficacement le système immunitaire des plantes, de nombreux agents phytopathogènes, dont les *Xanthomonas*, ont développé une stratégie qui consiste à se multiplier de façon épiphyte avant d'entrer dans la plante : c'est la phase d'installation.

1.2. L'installation

L'étape d'installation se déroule une fois que la bactérie a trouvé un site favorable afin de s'y fixer et de former un biofilm, se définissant comme une communauté de microorganismes sécrétant une matrice adhésive et protectrice constituée de polysaccharides leur permettant d'adhérer à une surface et entre eux (Costerton *et al.*, 1995). L'adhésion aux tissus est considérée comme une étape nécessaire et initiatrice de la formation d'un biofilm, et est essentielle pour une colonisation efficace de l'hôte. En effet, il a été démontré que le biofilm est indispensable à la bonne croissance des bactéries et au développement des symptômes de la maladie (Costerton *et al.*, 1987; Danhorn & Fuqua, 2007). C'est pourquoi ce mécanisme est conservé chez les bactéries à Gram-négatif pathogènes de plantes, d'invertébrés ou de mammifères (Kline *et al.*, 2009; Amano, 2010; Cao *et al.*, 2011). De plus, un biofilm peut être composé d'une seule espèce ou de plusieurs espèces microbiennes et constitue un " hot spot " d'échange de matériel génétique (Danhorn & Fuqua, 2007). Le biofilm protège également les bactéries des stress environnementaux (Morris & Monier, 2003; Jacques *et al.*, 2005) et, éventuellement, des mécanismes de défense de la plante. En effet, une étude réalisée sur le pathosystème basilic - *Pseudomonas aeruginosa* a montré que le biofilm confère à l'agent pathogène une résistance contre les composés antimicrobiens sécrétés par la plante (Walker *et al.*, 2004).

La formation de biofilm dépend en premier lieu de l'adhésion des microorganismes à la surface de la plante (**Figure 5**). Pour ce faire, les bactéries possèdent deux types de protéines spécialisées : les adhésines non fibrillaires et les adhésines fibrillaires. Ces deux types de protéines interviennent à différents stades de la formation du biofilm. En effet, les adhésines non fibrillaires sont essentielles à la première étape de la formation d'un biofilm, c'est-à-dire la fixation initiale des bactéries à une surface (Stoodley *et al.*, 2002), alors que les adhésines fibrillaires, en plus de jouer également un rôle dans l'adhésion des bactéries à la surface, sont impliquées dans l'adhésion

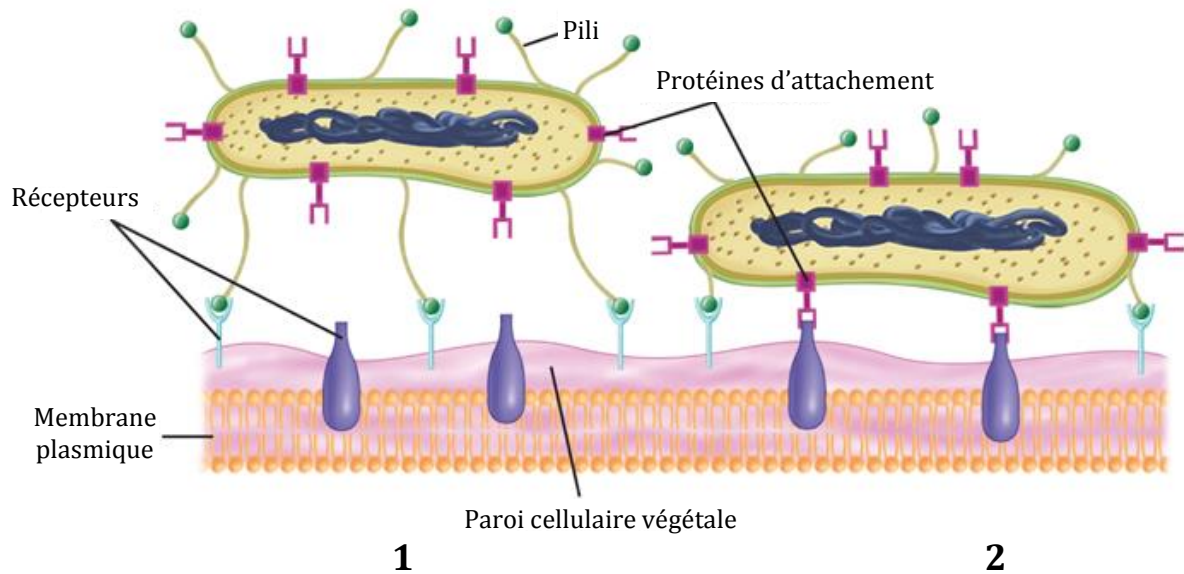


Figure 6 : Fixation initiale de bactéries phytopathogènes à la surface d'une cellule hôte. Cette figure est adaptée de Pottinger et al. 2004. Adhésion de la bactérie via l'utilisation d'adhésines fibrillaires (1) et non fibrillaires (2).

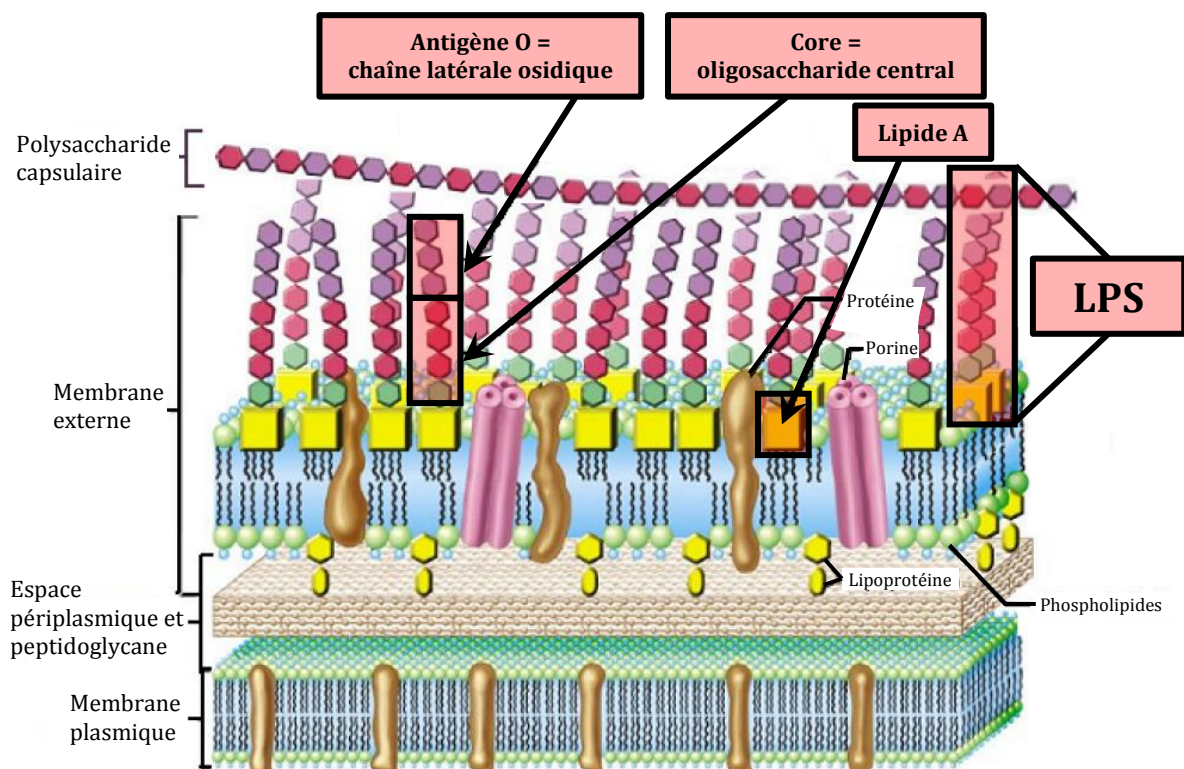


Figure 7 : Lipopolysaccharides sur la surface d'une bactérie à Gram-négatif. Cette figure est adaptée de Poinot et al. 2012. Un lipopolysaccharide se compose de trois structures : l'antigène O, le core et le lipide A. Les exopolysaccharides ne sont pas représentés ici mais ils sont au niveau des polysaccharides capsulaires. LPS : lipopolysaccharide.

des bactéries entre elles dans le but de former des colonies (**Figure 6**) (Danhorn & Fuqua, 2007; Darsonval *et al.*, 2009). Les adhésines non fibrillaires appartiennent à une famille très large de protéines de surface (Gerlach & Hensel, 2007) comprenant entre autres des systèmes de sécrétion à deux partenaires tel que YapH, une hémagglutinine filamenteuse. Cet exemple est intéressant car, chez *X. citri* pv. *fuscans*, il a été démontré que la mutation du gène codant cette protéine supprime la capacité d'adhésion de la bactérie et par conséquent sa capacité à former un biofilm (Darsonval *et al.*, 2009). Cette étude de la protéine YapH reflète donc ici l'importance des adhésines non fibrillaires dans le processus d'adhésion de la bactérie à une surface. Une fois que la bactérie est fixée, le processus de formation du biofilm à proprement parler peut commencer. Pour ce faire, les bactéries utilisent de façon successive les adhésines non-fibrillaires puis les adhésines fibrillaires. Parmi les adhésines fibrillaires, il y a les pili de type IV, qui jouent un rôle important dans l'agressivité des bactéries, notamment via leur capacité d'attachement à une surface. En effet, de précédentes études ont montré que le pilus de type IV de *X. campestris* pv. *hyacinthi* se fixait au niveau des stomates de feuilles de jacinthe (Doorn *et al.*, 1994) et que le mutant *fimA* de *X. axonopodis* pv. *vesicatoria* était affecté dans sa capacité à adhérer et à coloniser efficacement aux trichomes des feuilles de tomates (Ojanen-Reuhs *et al.*, 1997). D'autre part, les mutants *pilB* et *pilC* de *X. campestris* pv. *campestris* sont affectés dans leur virulence lors de l'inoculation sur feuille de choux (Qian *et al.*, 2005).

Par ailleurs, la paroi bactérienne est majoritairement composée de polysaccharides dont il existe deux types : les exopolysaccharides (EPS) et les lipopolysaccharides (LPS) (**Figure 7**). Ces polysaccharides interviennent également dans l'adhésion des cellules bactériennes et dans la formation de biofilm (Danhorn & Fuqua, 2007). En effet, il a été démontré que la mutation du gène *gpsX* entraîne une diminution significative de la production d'EPS et de LPS chez *X. citri* pv. *citri*, ce qui affecte la formation de biofilm, mais aussi la croissance bactérienne et les symptômes de la maladie sur feuilles d'agrumes (Li & Wang, 2012). Les polysaccharides jouent un rôle d'autant plus important chez certaines bactéries vasculaires des plantes qui en produisent de grandes quantités. Cette importante production peut conduire à la formation de biofilms obstruant les vaisseaux de la plante, ce qui empêche l'alimentation des tissus végétaux à proximité et entraîne leur mort (Danhorn & Fuqua, 2007). D'autre part, les polysaccharides confèrent le rôle protecteur au biofilm pour les bactéries. En effet, le confinement créé par les polysaccharides autour des bactéries agrégées en biofilm assure la résistance aux stress environnementaux. Par exemple, la délétion du locus *xagABC*, codant un système glycosyltransférase associé à la biosynthèse de polysaccharides, affecte la production d'EPS et réduit la résistance au stress oxydatif (Tao *et al.*, 2010). Ceci a également été observé pour le mutant du gène *gpsX* de *X. citri* pv. *citri* (Li & Wang, 2012). Ainsi, les bactéries agrégées en biofilm sont protégées et jouent un rôle très important dans la colonisation de l'hôte et dans le développement de la maladie (Costerton *et al.*, 1987).

En plus d'être impliqués dans la fixation initiale des bactéries à une surface, dans l'adhésion des bactéries entre elles et dans leur protection (Stoodley *et al.*, 2002; Danhorn & Fuqua, 2007; Darsonval *et al.*, 2009; Tao *et al.*, 2010), les adhésines et les polysaccharides sont également des facteurs de virulence des bactéries du genre *Xanthomonas* car ils sont indispensables pour une infection réussie et pour la multiplication bactérienne (Alavi *et al.*, 2008; Büttner & Bonas, 2010). En effet, chez la souche CFBP4834-R de *X. citri* pv. *fuscans*, la mutation de gènes codant des adhésines a altéré leurs capacités d'adhérence et à se transmettre à la semence ainsi que dans leur pouvoir pathogène. Cela signifie donc que les adhésines sont impliquées dans la colonisation et dans la transmission à la semence (Darsonval *et al.*, 2009). Egalement, chez *X. oryzae* pv. *oryzicola*, la mutation du gène *wxocB*, impliqué dans la synthèse de LPS, entraîne une diminution de la quantité d'EPS et une perte de virulence de la souche qui serait directement liée à la perte des chaînes latérales osidiques ou antigène O du LPS (Wang *et al.*, 2013). Ainsi, en tant que facteurs de virulence, les adhésines et polysaccharides sont également impliqués dans la phase suivante de l'infection : la multiplication.

1.3. La multiplication

L'étape de multiplication correspond à la phase symptomatique du cycle infectieux des agents pathogènes. La bactérie est arrivée sur son hôte, elle a trouvé un site favorable à sa survie, elle y a adhéré et s'y est multipliée de façon épiphyte. Pour coloniser la plante de façon efficace, elle va pénétrer dans les tissus. Les principales voies d'entrée des bactéries du genre *Xanthomonas* dans la plante sont les ouvertures naturelles (stomates, hydathodes, lenticelles, nectaires) ou les blessures (Jacques *et al.*, 2016). Dès l'entrée dans les tissus végétaux, il s'ensuit une course de vitesse entre la plante qui enclenche immédiatement ses mécanismes de défense active, et l'agent pathogène qui doit contourner ces défenses pour exprimer son pouvoir pathogène (Grennan, 2006; Doughari, 2015). Pour y parvenir, la bactérie a besoin d'exprimer un arsenal de facteurs de virulence, dont les plus importants, en plus des adhésines et polysaccharides décrits précédemment, sont le système de sécrétion de type 3 (SST3) et ses effecteurs, les transporteurs TonB-dépendants, les toxines, les enzymes pectinolytiques, ou encore le système de sécrétion de type 2 (SST2).

La première ligne de défense active de la plante, nommée PTI pour PAMP-Triggered Immunity, utilise des récepteurs extracellulaires, appelés PRR pour Pattern Recognition Receptors, capables de reconnaître des éléments présents chez tous les agents pathogènes : les MAMP (anciennement nommés PAMP) pour Microbe- (Pathogen-) Associated Molecular Patterns (Zipfel, 2009). Chaque PRR est spécifique d'un MAMP particulier et les MAMP sont des motifs situés dans des domaines très conservés de protéines elles-mêmes très conservées chez les agents

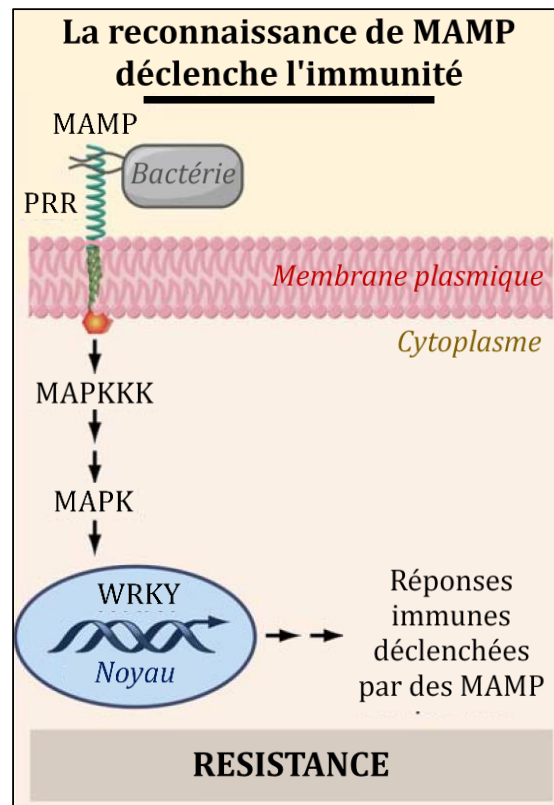


Figure 8 : Description de la PTI. Ce schéma est adapté de Chisholm et al. 2006. Reconnaissance d'un MAMP (Microbe-Associated Molecular Pattern), tel que la flagelline des bactéries phytopathogènes, par un récepteur extracellulaire de la plante de type PRR (Pattern Recognition Receptor). Cette reconnaissance active les protéines MAPK (Mitogen-Associated Protein Kinases), qui à leur tour activent les facteurs de transcription WRKY.

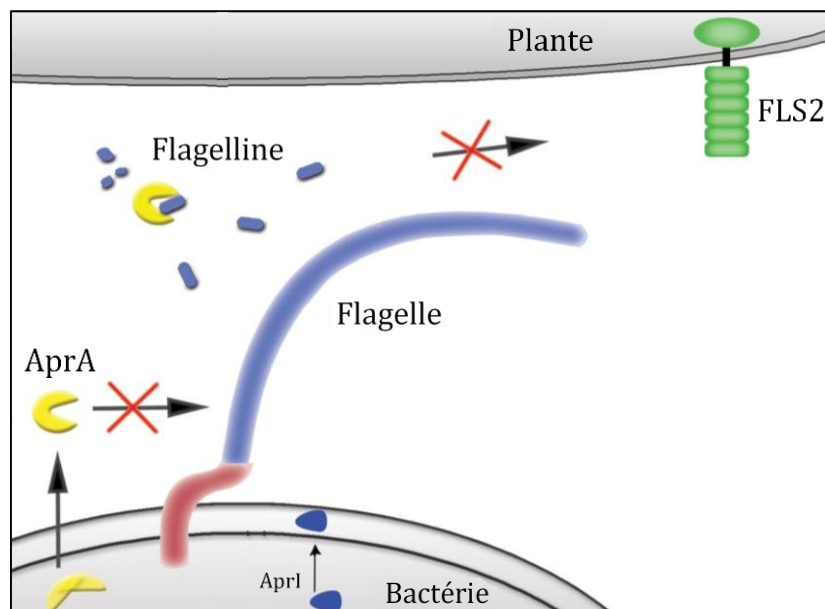


Figure 9 : Représentation d'une stratégie de contournement de la PTI mise en place par l'agent pathogène *Pseudomonas aeruginosa*. Ce schéma est adapté de Bardoel et al. 2011. La bactérie possède la protéine AprA sous sa forme inactive dans le cytoplasme. Lorsqu'AprA est sécrétée dans le milieu extracellulaire, elle devient active et clive la flagelline sans affecter la fonction du flagelle. Ainsi, FLS2 (PRR) n'est plus capable de reconnaître flg22 (MAMP) de la flagelline et la PTI n'est pas activée.

pathogènes. En effet, la perte ou modification de ces domaines peut entraîner une perte de virulence ou de fitness (Boller & He, 2009). Lorsqu'un MAMP est reconnu par un PRR, des protéines MAPK (Mitogen-Associated Protein Kinases) sont activées et sont transloquées dans le noyau où elles activent à leur tour des facteurs de transcription de type WRKY ciblant notamment des gènes de défense (**Figure 8**) (Eulgem *et al.*, 2000). L'exemple le plus connu de MAMP est le peptide flg22 constitué de 22 acides aminés situés dans la partie N-terminale de la flagelline des bactéries à Gram-négatif, qu'il s'agisse d'individus pathogènes ou commensaux (Felix *et al.*, 1999). Ce motif est reconnu par le récepteur FLS2, présent notamment chez *Arabidopsis* et chez la tomate, qui induit alors la réaction de défense précédemment décrite (Chinchilla *et al.*, 2006).

Afin de contourner ce type de défense, les bactéries ont développé des stratégies d'échappement à la reconnaissance des MAMP. En effet, elles peuvent perdre ou modifier les motifs reconnus par les PRR de la plante. Par exemple, des souches de *Pseudomonas aeruginosa* sécrètent une protéine AprA qui clive la flagelline. Ainsi, la mobilité de la bactérie n'est pas altérée de façon durable, et les défenses d'*Arabidopsis* ne sont pas activées, FLS2 ne pouvant plus reconnaître flg22 (**Figure 9**) (Bardoel *et al.*, 2011). De façon plus drastique, certains individus d'une population d'agents pathogènes peuvent perdre leur flagelle et il est tentant d'émettre l'hypothèse qu'il s'agit d'une stratégie de contournement de la PTI. En effet, il a été démontré que certaines souches de *X. citri* pv. *fuscans* ont perdu leur capacité de mobilité flagellaire sans altération de leur fitness *in planta*. Par exemple, la souche CFBP4834-R, bien que très agressive sur haricot, a perdu le cluster de gènes codant les protéines flagellaires, et n'est donc pas motile. Bien que cela reste à démontrer, il pourrait s'agir d'une stratégie mise en place par la population d'agents pathogènes pour que leur présence sur la plante ne déclenche pas des réactions de défenses proportionnées à la taille réelle de la population (Darrasse *et al.*, 2013b; Indiana, 2014).

Le principal facteur de virulence des bactéries à Gram-négatif est le SST3, qui est mis en place par la bactérie dans son hôte grâce à l'expression des gènes *hrp* (hypersensitive response and pathogenicity). Ces gènes ont été découverts chez *Pseudomonas savastanoi* pv. *phaseolicola* (Lindgren *et al.*, 1986) et sont indispensables au déclenchement de la maladie (Lindgren, 1997). Par exemple, la mutation de quelques gènes *hrp* (*hrpG*, *hrcC* et *hrpN*) chez *Erwinia chrysanthemi* impacte fortement la virulence de la bactérie sur saintpaulia (Yang *et al.*, 2002). L'expression du SST3 n'est pas constitutive et dépend de l'environnement de la bactérie. En effet, il a été démontré que *Pseudomonas syringae* pv. *tomato* est capable de percevoir de nombreux signaux provenant de la plante afin d'initier la mise en place de son SST3 (Anderson *et al.*, 2014). L'expression des gènes *hrp* est régulée par une protéine nommée HrpX qui reconnaît directement des motifs conservés chez ces gènes, comme notamment des boîtes PIP pour Plant Inducible Promoter (Wengelnik & Bonas, 1996). L'expression de *hrpX* est quant à elle régulée par HrpG, un régulateur

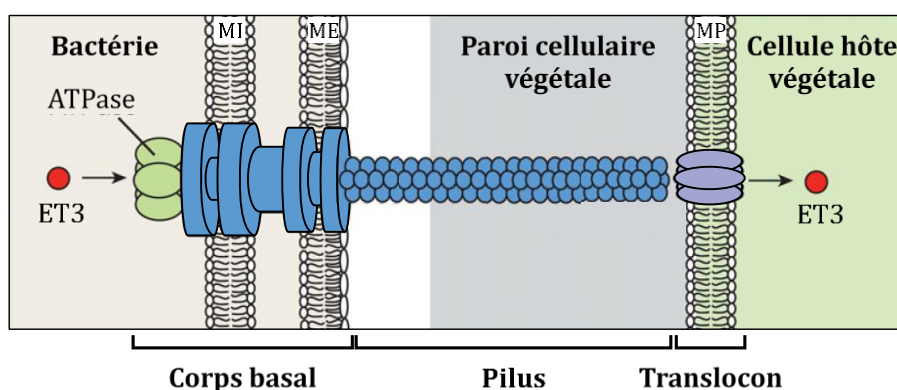


Figure 10 : Représentation schématique du système de sécrétion de type III des bactéries phytopathogènes. Ce schéma est adapté de Büttner and He 2009. Le corps basal du SST3 (Système de Sécrétion de Type 3) est inclus dans les deux membranes bactériennes et est associé à une ATPase cytoplasmique. Ce SST3 est relié à un pilus extracellulaire qui s'étend à travers la paroi cellulaire de la plante et qui joue le rôle de seringue. La seringue est reliée au translocon, un canal intégré dans la membrane plasmique de la cellule hôte. Le SST3 permet donc le transport et l'injection d'ET3 (Effecteur de Type 3) bactérien directement dans le cytoplasme de la cellule hôte. MI : Membrane Interne ; ME : Membrane Externe ; MP : Membrane Plasmique.

Tableau 1 : Fonction et distribution d'une liste non exhaustive d'ET3 dans le genre *Xanthomonas*. Ce tableau est adapté de Ryan et al. 2011 et de Hutin, 2015. *Xal* : *Xanthomonas albilineans* ; *Xau* : *X. citri* pv. *aurantifolii* ; *Xca* : *Xanthomonas campestris* pv. *armoraciae* ; *Xcc* : *Xanthomonas campestris* pv. *campestris* ; *Xcf* : *X. citri* pv. *fuscans* ; *Xcg* : *Xanthomonas citri* pv. *glycines* ; *Xci* : *Xanthomonas citri* pv. *citri* ; *Xcm* : *X. citri* pv. *malvacearum* ; *Xcr* : *X. campestris* pv. *raphani* ; *Xcu* : *X. axonopodis* pv. *citrumelo* ; *Xcv* : *Xanthomonas campestris* pv. *vasculorum* ; *Xeu* : *Xanthomonas euvesicatoria* ; *Xg* : *Xanthomonas gardneri* ; *Xmu* : *Xanthomonas campestris* pv. *musacearum* ; *Xoc* : *Xanthomonas oryzae* pv. *oryzicola* ; *Xoo* : *Xanthomonas oryzae* pv. *oryzae* ; *Xp* : *X. perforans* ; *Xpm* : *X. phaseoli* pv. *manihotis* ; *Xpp* : *X. phaseoli* pv. *phaseoli* ; *Xtu* : *X. translucens* pv. *undulosa*.

Effecteurs	Domaines fonctionnel ou motif	Distribution
Effecteurs « core »		
XopR	Inconnu	Toutes sauf <i>Xal</i>
AvrBs2	Homologie de séquence avec une agrocinopine synthase et une glycérol phosphodiesterase	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopK	Inconnu	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopL	Répétions riches en leucine (LRR)	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopN	Protéine à motif répété ARM/Heat	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopP	Inconnu	Toutes sauf <i>Xal</i>
XopQ (HopQ1)	Inosine-uridine nucléoside N-ribohydrolase	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopX	Protéine riche en méthionine	Toutes sauf <i>Xca</i> et <i>Xal</i>
XopZ	Inconnu	Toutes sauf <i>Xca</i> et <i>Xal</i>
Effecteurs "variables"		
AvrBsT (XopJ2)	Acetyltransferase Ser/Thr, protéase à cystéine de la famille C55	Quelques de souches de <i>Xeu</i>
XopC	Domaine phosphorybosyl transferase et haloacid dehalogenase-like hydrolase	<i>Xeu</i> et quelques souches de <i>Xoo</i>
XopF1	Inconnu	<i>Xeu</i> , <i>Xca</i> , <i>Xoo</i> , <i>Xoc</i> , <i>Xcv</i> et <i>Xmu</i>
XopF2	Inconnu	<i>Xeu</i> , <i>Xmu</i> et <i>Xcv</i>
XopJ	Acetyltransferase Ser/Thr, protéase à cystéine de la famille C55	<i>Xeu</i>
XopO	Inconnu	<i>Xeu</i> et <i>Xoc</i>
XopAE	Répétions riches en leucine (LRR)	<i>Xcg</i> , <i>Xci</i> , <i>Xoo</i> , <i>Xoc</i> , <i>Xmu</i> et <i>Xcv</i>
AvrXccC (XopAH)	Inconnu	<i>Xcc</i>
XopE2	Transglutaminase putative	<i>Xeu</i> , <i>Xci</i> et <i>Xcc</i>
XopAM	Inconnu	<i>Xcc</i> , <i>Xmu</i> et <i>Xcv</i>
XopD	Cystéine protéase de type SUMO avec un domaine NLS, un site de liaison à l'ADN et un motif EAR possiblement impliqué dans l'inhibition de la transcription	<i>Xeu</i> et certaines souches de <i>Xcc</i>
Effecteur TAL	Facteur de transcription nucléaire de type AvrBs3	<i>Xau</i> , <i>Xci</i> , <i>Xcg</i> , <i>Xca</i> , <i>Xcc</i> , <i>Xcf</i> , <i>Xcu</i> , <i>Xcm</i> , <i>Xcr</i> , <i>Xeu</i> , <i>Xg</i> , <i>Xoc</i> , <i>Xoo</i> , <i>Xp</i> , <i>Xpm</i> , <i>Xpp</i> , <i>Xtu</i> .
XopB	Inconnu	<i>Xeu</i>
XopH	Tyrosine phosphatase putative	<i>Xeu</i> , <i>Xcc</i> , <i>Xmu</i> et <i>Xcv</i>
XopI	Protéine F-box	<i>Xci</i> , <i>Xeu</i> , <i>Xoo</i> , <i>Xoc</i> , <i>Xmu</i> et <i>Xcv</i>
XopAA	Impliqué dans la chlorose	<i>Xeu</i> , <i>Xoo</i> et <i>Xoc</i>
XopAC	Protéine LRR	<i>Xcc</i>
XopU	Inconnu	<i>Xoo</i> et <i>Xoc</i>
XopAD	Répétitions SKWP	<i>Xci</i> , <i>Xeu</i> , <i>Xca</i> , <i>Xoo</i> et <i>Xoc</i>

transcriptionnel (Wengelnik *et al.*, 1996). L'ensemble des gènes *hrp* forme un îlot de pathogénie, l'îlot *hrp*, constitué d'une vingtaine de gènes (Büttner & Bonas, 2002). Par exemple, chez la souche de *X. oryzae* pv. *oryzae* KACC10859, l'îlot *hrp* se compose de neuf gènes *hrp*, neuf gènes *hrc* (*hrp* conserved) et huit gènes *hpa* (*hrp* associated) ce qui fait un total de 26 gènes (Cho *et al.*, 2008). Les gènes *hrc* sont ainsi nommés car ils sont conservés entre des bactéries pathogènes de plantes et d'animaux (Bogdanove *et al.*, 1996). En effet, le SST3 a été rapporté chez des bactéries phytopathogènes, comme *Xanthomonas* sp., *Ralstonia* sp., *Erwinia* sp., *Burkholderia* sp. et *Pantoea* sp., chez des bactéries pathogènes de l'homme et des animaux, comme *Salmonella* sp., *Yersinia* sp., *Shigella* sp., *Escherichia coli* et *Citrobacter* sp., mais aussi chez des bactéries non-pathogènes du genre *Pseudomonas* ou encore des bactéries symbiotiques du genre *Rhizobium* (He *et al.*, 2004; Cornelis, 2006; Grant *et al.*, 2006). Cette conservation des gènes *hrp* entre des bactéries pathogènes diversifiées conforte l'idée que le SST3 est un acteur majeur du pouvoir pathogène des bactéries. Le SST3 se compose de trois structures : un corps basal, inclus dans la membrane interne et externe de la bactérie, un pilus situé dans la partie extracellulaire, et enfin des protéines formant le translocon constituant la partie du SST3 qui s'intègre dans la membrane de la cellule hôte (**Figure 10**) (Izoré *et al.*, 2011). L'ensemble de ces structures constitue un complexe protéique formant une sorte de seringue, capable de percer la membrane d'une cellule hôte eucaryote afin d'y injecter des effecteurs de type III (ET3) (Galán & Collmer, 1999; Izoré *et al.*, 2011; Portaliou *et al.*, 2016).

Les ET3 sont des protéines sécrétées directement dans le cytoplasme de la cellule hôte grâce au SST3. Ces effecteurs possèdent un signal de sécrétion de type 3 en position N-terminale pour être injectés dans la cellule hôte via le SST3, et un ou plusieurs domaines fonctionnel au centre et/ou en position C-terminale de la protéine (Sory *et al.*, 1995). On distingue deux types d'ET3 : les effecteurs « core », qui sont conservés au sein du genre *Xanthomonas* et dont la fonction est le plus souvent dirigée contre la PTI, et les effecteurs « variables », qui sont présents chez seulement quelques souches et majoritairement impliqués dans le contournement des défenses de la plante, mais aussi dans la spécificité d'hôte (**Tableau 1**) (Ryan *et al.*, 2011). Les ET3 sont très diversifiés et peuvent être classés selon leur mode d'action (Mudgett, 2005). En effet, il a été dénombré au moins 65 gènes différents codant des ET3 au sein du genre *Xanthomonas* (<http://xanthomonas.org/t3e.html>). La plupart des ET3 ont des activités de type protéase ou phosphatase. Certains sont capables de manipuler les signaux de défense hormonaux de la plante (Kazan & Lyons, 2014), d'autres peuvent inhiber l'induction de gènes de défense (Kim *et al.*, 2008), certains parviennent à inverser les processus physiologiques mis en place par la plante pour contrer l'entrée des agents pathogènes, comme la fermeture des stomates (Melotto *et al.*, 2006). Il existe même des ET3 capables de jouer le rôle de facteur de transcription eucaryote, les effecteurs TAL (Transcription Activator-Like) qui seront détaillés au chapitre 3 (Mak *et al.*, 2013).

Tableau 2 : Modèle de relation « gène pour gène ». Ce tableau est adapté de Flor 1971. La résistance de la plante ne peut être mise en place que si elle possède le gène *R* (gène de résistance) et que la bactérie possède le gène *Avr* (gène d'avirulence) correspondant, il s'agit d'une interaction incompatible représentée en vert. Dans tous les autres cas, si la plante possède le gène *r* et/ou que la bactérie possède le gène *avr*, la plante est sensible à l'infection provoquée par la bactérie, il s'agit d'une réaction compatible représentée en rouge. *R* : allèle dominant du gène de résistance ; *r* : allèle récessif du gène de résistance ; *Avr* : allèle dominant du gène d'avirulence ; *avr* : allèle récessif du gène d'avirulence.

		Génotype de la plante	
		<i>R</i>	<i>r</i>
Génotype de l'agent pathogène	<i>Avr</i>	Résistance	Sensibilité
	<i>avr</i>	Sensibilité	Sensibilité

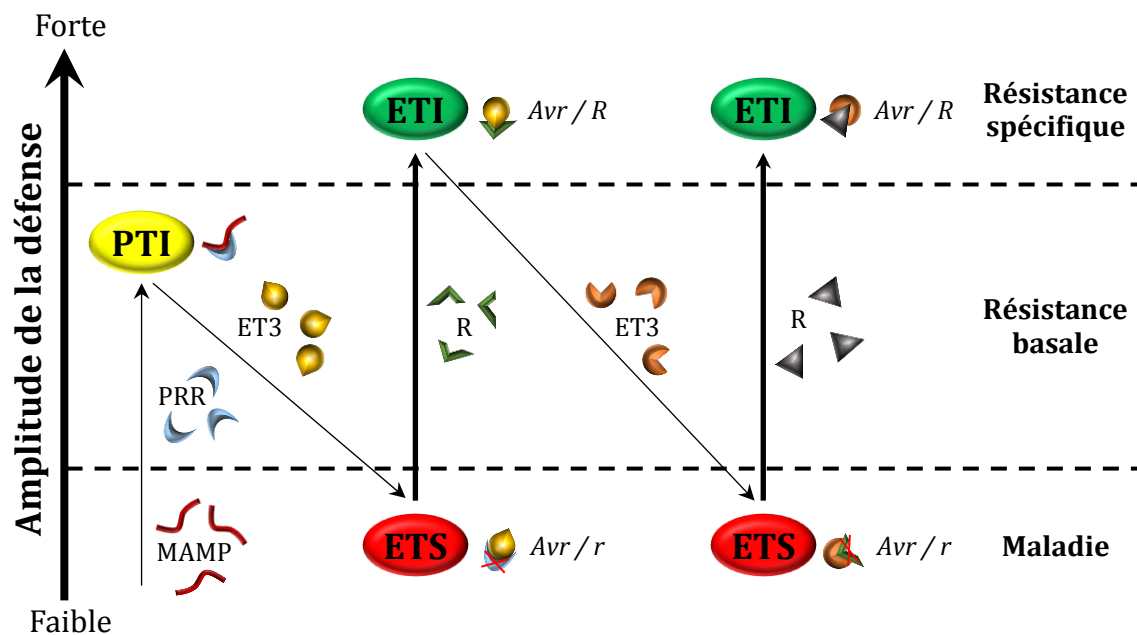


Figure 11 : Modèle de co-évolution en zigzag de l'interaction hôte-pathogène. Ce schéma est adapté de Jones and Dangl 2006. La plante est capable de reconnaître des MAMP (Microbe-Associated Molecular Patterns) bactériens grâce à des PRR (Pattern Recognition Receptors) ce qui déclenche la PTI (PAMP-Triggered Immunity), correspondant à la résistance basale de la plante avec une faible amplitude. Pour contourner ce niveau de défense, la bactérie peut sécréter des ET3 (Effecteur de Type 3) ciblant la PTI, on entre alors dans une phase de sensibilité induite par ces ET3, appelés ETS (Effector-Triggered Susceptibility). Ensuite, la plante peut produire des protéines *R* (résistance) qui reconnaissent les ET3 ciblant la PTI, devenant donc des protéines d'avirulence (*Avr*), ce qui déclenche l'ETI (Effector-Triggered Immunity), correspondant à une résistance spécifique dont l'amplitude est plus importante que celle de la PTI. En réponse, le pathogène peut évoluer en acquérant de nouveaux ET3 non reconnus et lui permettant de contourner l'ETI, jusqu'à ce que de nouvelles protéines *R* apparaissent. *R* : allèle dominant du gène de résistance ; *r* : allèle récessif du gène de résistance ; *Avr* : allèle dominant du gène d'avirulence.

Quand des ET3 ont pour rôle d'inhiber la PTI, on parle d'ETS pour Effector-Triggered Susceptibility. Par exemple, *X. oryzae* pv. *oryzae* possède deux ET3, XopAA et XopY, qui permettent à la bactérie de survivre sur le riz, car ces deux ET3 ciblent des protéines impliquées dans la PTI et empêchent la mise en place de cette défense basale (Yamaguchi *et al.*, 2013). Egalement, chez *Pseudomonas syringae* pv. *tomato* l'ET3 AvrPtoB agit en tant que suppresseur de la mort cellulaire programmée (Alfano & Collmer, 2004). La superfamille AvrE-T3E, qui est répandue chez les phytobactéries, joue quant à elle un rôle crucial dans la pathogenèse bactérienne, notamment, en supprimant la PTI (Degraeve *et al.*, 2015). L'ETS peut également inhiber les dépôts de callose (Hauck *et al.*, 2003), l'activation des cascades MAPK (He *et al.*, 2006), ainsi que la synthèse de protéines ou de micro-ARN impliqués dans la PTI (Navarro *et al.*, 2008). D'autres ET3 sont quant à eux responsables du déclenchement d'un second mécanisme de défense de la plante nommé ETI pour Effector-triggered immunity (Jones & Dangl, 2006). En effet, certaines plantes sont capables de reconnaître directement ou indirectement les ET3 qui ciblent la PTI. La reconnaissance directe des ET3 correspond à une relation "gène pour gène", où l'agent pathogène possède une protéine d'avirulence qui va être ciblée par une protéine de résistance de la plante permettant ainsi à la résistance de s'exprimer (**Tableau 2**) (Flor, 1971). Par exemple, chez la tomate, le gène de résistance *Bs4*, codant une protéine de type NB-LRR, est capable de reconnaître l'effecteur TAL AvrBs4 de *X. euvesicatoria* pv. *euvesicatoria*, ce qui déclenche une réponse hypersensible (HR), conduisant à une mort cellulaire programmée au point, ce qui permet de limiter la propagation de l'agent pathogène (Schornack *et al.*, 2004). Il existe aussi d'autres ET3 capables d'inhiber l'ETI. Par conséquent, au cours de cette étape de multiplication, l'interaction hôte-pathogène va suivre une dynamique évolutive décrite comme suivant un modèle en zigzag, qui permet d'expliquer, du moins en partie, la coévolution entre l'agent pathogène et la plante hôte (Jones & Dangl, 2006). Ce modèle d'évolution propose que la sélection naturelle ait abouti à une alternance de phases de sensibilité et de résistance de la plante, avec pour commencer la reconnaissance des MAMP par la plante et le déclenchement de la PTI, puis le contournement de la PTI par la bactérie via une modification des MAMP ou à l'aide d'ET3, suivi de la reconnaissance par la plante des ET3 ciblant la PTI, permettant ainsi le déclenchement de l'ETI, elle-même potentiellement contournée par la bactérie grâce à d'autres ET3 (**Figure 11**). La réalité, est sans doute bien plus complexe que ce modèle. Notamment, la distinction entre les MAMP et les ET3 n'est pas toujours évidente. En effet, les ET3 sont très répandus et peuvent être reconnus par la plante dès la PTI, ils peuvent donc parfois correspondre à des MAMP. Les MAMP, quant à eux, peuvent jouer à la fois un rôle dans la fitness et dans la virulence de la bactérie (Boller & He, 2009; Thomma *et al.*, 2011). Ainsi, les ET3 sont des facteurs de virulence majeurs des bactéries phytopathogènes et des acteurs majeurs de l'interaction hôte-pathogène, leur diversité fonctionnelle leur permettant de jouer de multiples rôles (**Figure 12**).

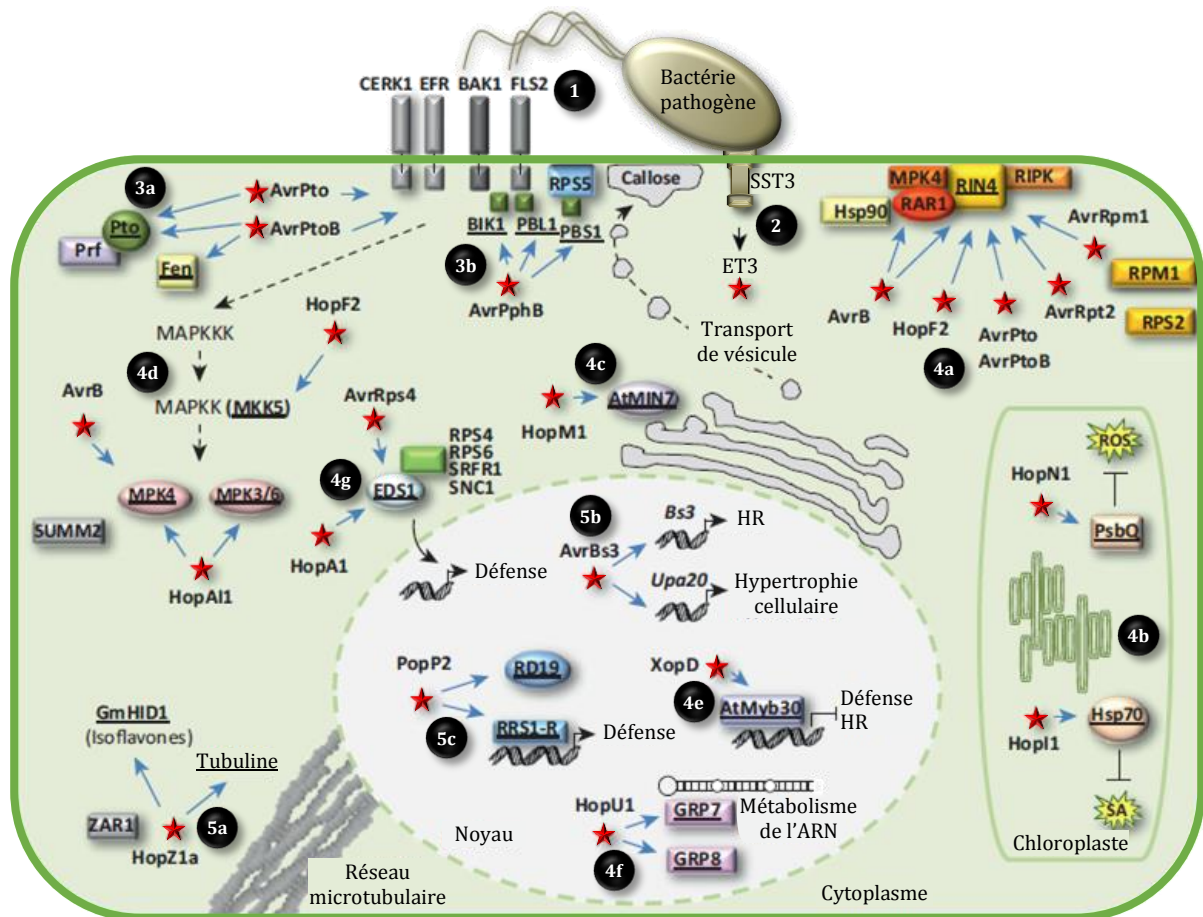


Figure 12: Cibles des effecteurs de type III bactériens. Cette figure est adaptée de Deslandes and Rivas 2012. **(1)** L'activation de complexes de récepteurs associés à la membrane plasmique, comme FLS2/BAK1 et EFR/BAK1, par des MAMP bactériens conservés, déclenche la PTI. **(2)** Les bactéries phytopathogènes injectent des effecteurs de type III (ET3) dans les cellules végétales en utilisant leur système de sécrétion de type III (SST3) et après la translocation des ET3 dans les cellules végétales, ils peuvent être adressés à différents compartiments sous-cellulaires où ils manipulent une diversité de fonctions cellulaires hôtes. **(3)** Cependant, la reconnaissance des ET3 par des protéines de résistance déclenche des ETI. Par exemple, **(3a)** AvrPto et AvrPtoB ciblent des récepteurs associés à la membrane plasmique mais Pto et Fen agissent comme des mimétiques moléculaires de ces récepteurs pour activer l'ETI. **(3b)** Le clivage de PBS1, BIK1 et PBL1 par AvrPphB est reconnu par la protéine RPS5, ce qui déclenche l'ETI. **(4)** Néanmoins, beaucoup d'ET3 sont impliqués dans le contournement des ETI. Par exemple, **(4a)** AvrPto, AvrPtoB, AvrB, AvrRpm1, AvrRpt2 et HopF2 ciblent le régulateur négatif de la défense RIN4 associé à la membrane plasmique. **(4b)** HopI1 et HopN1 sont adressés au chloroplaste où ils ciblent respectivement Hsp70, pour supprimer l'accumulation d'acide salicylique (SA) et PsbQ associée au Photosystème II, pour diminuer la production de ROS (espèces réactives de l'oxygène). **(4c)** HopM1 s'accumule dans le réseau « trans-Golgi / endosome précoce » où il cible AtMIN7, un élément clé du trafic vésiculaire, supprimant ainsi les défenses associées à la paroi cellulaire. **(4d)** HopF2, HopA1 et AvrB inhibe l'activation des cascades MAPK. **(4e)** XopD cible AtMYB30, un régulateur positif des réponses de défense, afin de supprimer les HR et les réponses de défense. **(4f)** HopU1 cible GRP7/8 ce qui modifie probablement le métabolisme de l'ARN liée à l'immunité. **(4g)** HopA1 et AvrRps4 visent EDS1 afin de perturber son association avec diverses protéines de résistance comme RPS4. **(5)** En dehors de ces fonctions liées à la défense de la plante, les ET3 ciblent beaucoup d'autres fonctions cellulaires hôtes, mais sont toujours potentiellement reconnus par des protéines de résistance. Par exemple, **(5a)** HopZ1a, reconnu par la protéine de résistance ZAR1, cible GmHID1, qui est impliqué dans la biosynthèse d'isoflavones et de la tubuline, ce qui affecte le réseau de microtubules cellulaires. **(5b)** Dans le noyau, AvrBs3 (un effecteur TAL) active directement la transcription du gène hôte *Upa20* ce qui induit une hypertrophie des cellules végétales et contribue au développement de la maladie. En revanche, dans les plantes résistantes, AvrBs3 active l'expression du gène de résistance *Bs3* ce qui conduit au développement d'une HR. **(5c)** PopP2 induit la relocalisation nucléaire de RD19, mais la perception de son activité par la protéine de résistance RRS1-R active l'immunité.

D'autres facteurs importants du pouvoir pathogène sont les transporteurs TonB-dépendants ou TBDT, qui permettent aux bactéries d'interagir avec leur environnement (Koebnik, 2005). Les TBDT correspondent à des protéines intégrées dans la membrane externe de la bactérie qui sont capables de lier et d'effectuer un transport actif et à haute affinité de sidérophores, de la vitamine B12 mais aussi de glucides (Blanvillain *et al.*, 2007). Pour obtenir la force proton-motrice nécessaire au transport des substrats de la membrane interne vers la membrane externe de la bactérie, le TBDT a besoin d'interagir avec un complexe protéique composé de trois protéines associées à la membrane interne : TonB, ExbB et ExbD (**Figure 13**) (Noinaj *et al.*, 2010; Celia *et al.*, 2016). Ce type de transporteur est essentiel à la virulence de la bactérie. Par exemple, chez la souche KT2440 de *Pseudomonas putida*, la mutation du TBDT empêche l'absorption du fer présent dans les graines de maïs par la bactérie, ce qui a pour effet de diminuer la capacité des bactéries à initier l'étape d'installation sur ses graines et par conséquent leur capacité à les coloniser (Molina *et al.*, 2005). Egalement, le TBDT peut être impliqué dans la formation de HR. En effet, chez *X. campestris* pv. *campestris* la structure du cluster de gènes *tonB* est un peu particulière, car elle possède deux paralogues du gène *exbD*, nommés *exbD1* et *exbD2*, et il a été démontré qu'*exbD2* n'est pas nécessaire pour l'absorption de fer, mais essentiel à l'induction d'une HR sur piment, une plante non-hôte. En fait, *exbD2* est essentiel à l'induction de l'activité de la pectate lyase bactérienne, et lorsqu'il est déficient, la capacité de *X. campestris* pv. *campestris* à dégrader la paroi cellulaire végétale est fortement réduite, ce qui empêche le déclenchement d'une HR sur une plante non-hôte (Vorhölter *et al.*, 2012). Lorsqu'une enzyme de dégradation de la paroi cellulaire et un TBDT s'associent, un système spécifique nommé CUT se forme. Ce système CUT est impliqué dans l'utilisation des glucides végétaux par les bactéries, mais aussi dans la virulence bactérienne. En effet, il a été démontré que, chez *X. campestris* pv. *campestris*, ce système CUT est impliqué dans l'utilisation du saccharose et est indispensable aux souches pour qu'elles aient une pathogénie complète sur *Arabidopsis* (Blanvillain *et al.*, 2007).

Le SST2 permet aux bactéries phytopathogènes de sécréter un vaste répertoire d'enzymes de dégradation des parois végétales, comportant des cellulases, des pectinases ou des xylanases, mais il peut aussi sécréter des toxines et des enzymes hydrolytiques extracellulaires telles que les protéases, les lipases et les phospholipases. Ainsi, ce système de sécrétion et les enzymes qu'il sécrète contribue à l'interaction hôte-pathogène (Sandkvist, 2001; Büttner & Bonas, 2010). Le mécanisme de translocation à travers le SST2 se déroule en deux étapes. Lors de la première étape, les protéines sont reconnues et transportées à travers la membrane interne de la bactérie via un système de transport : le Sec-dépendant et/ou le Tat-dépendant (Cianciotto, 2005). Les deux systèmes de transport diffèrent par leur source d'énergie. En effet, le Sec-dépendant contient une

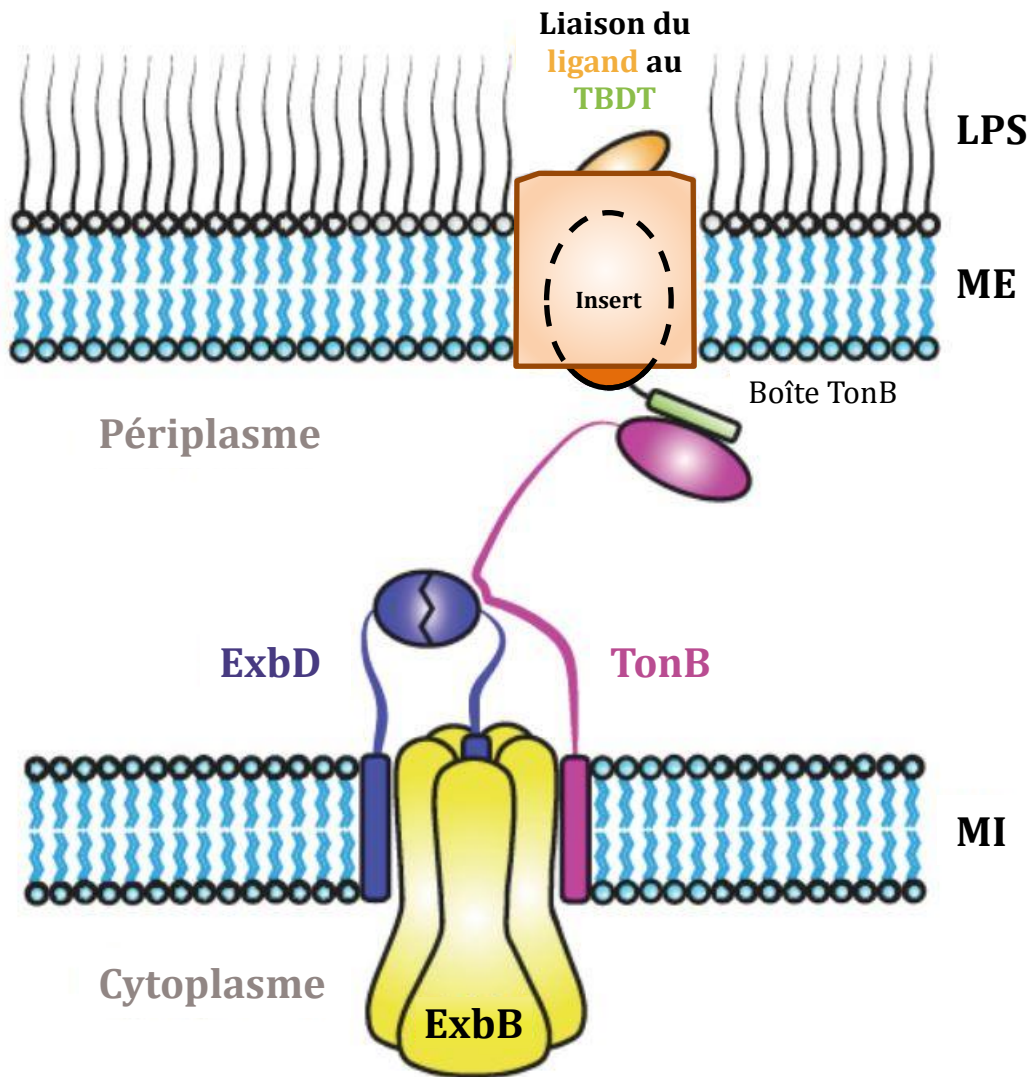


Figure 13 : Modèle structurel du complexe Ton. Ce schéma est adapté de Celia et al. 2016. Le système Ton est composé d'un pentamère d'ExbB, d'un dimère d'ExbD et d'au moins un TonB. TBDT : transporteur dépendant de TonB ; LPS : lipopolysaccharide ; ME : membrane externe ; MI : membrane interne.

protéine cytoplasmique nommée SecA capable d'hydrolyser l'ATP, alors que le Tat-dépendant utilise la différence de pH transmembranaire (Santini *et al.*, 1998; Cao & Saier, 2003). Ainsi, le système Tat-dépendant permet la translocation de protéines particulières repliées et activées dans le cytoplasme. Pour emprunter le système Tat-dépendant, les protéines présentent une séquence signal en N-terminale contenant deux arginines consécutives, ce qui explique le nom du système, Tat étant pour « Twin Arginine Translocation ». Lors de la seconde étape, les protéines sont transloquées du périplasma vers l'extérieur via le SST2 (Cianciotto, 2005). Le SST2 est composé de 12 à 16 protéines, dont notamment une ATPase, une peptidase et différentes pseudopilines, constituant un pseudopilus directement impliqué dans la translocation. En effet, la sécrétion des protéines via le SST2 serait dépendante de ce pseudopilus périplasmique qui serait assemblé et désassemblé continuellement afin de pousser les protéines vers l'extérieur à travers le canal de sécrétion (Cianciotto, 2005; Johnson *et al.*, 2006). Une analyse de génomique comparative regroupant des pathogènes vasculaires et non vasculaires de Brassicacées, d'agrumes, du poivron, du riz et de la tomate, a révélé qu'il existait deux groupes de gènes codant le SST2 : le groupe Xps, présent chez toutes les souches étudiées, et le groupe Xcs, absent des souches infectant le riz (Lu *et al.*, 2008). Ce système est donc conservé à travers le genre *Xanthomonas*, car il est important pour la virulence de ces bactéries, tout comme les protéines qu'il sécrète. En effet, il a été démontré que plusieurs gènes *xps* sont essentiels à la virulence de *X. campestris* pv. *campestris* (Lu *et al.*, 2008) et qu'une mutation dans *xpsD* de *X. citri* pv. *citri* affecte sa capacité à coloniser les tissus, à hydrolyser la cellulose et donc à infecter l'oranger (Baptista *et al.*, 2010). Egalement, tout comme le SST3 et ses effecteurs, le SST2 et ses effecteurs peuvent induire des mécanismes de défense de plante. Par exemple, chez *X. oryzae* pv. *oryzae*, il a été démontré que ClsA (une cellulase), CbsA (une cellobiosidase) et LipA (une lipase), toutes identifiées comme étant des effecteurs du SST2, induisent des mécanismes de défense du riz. De manière intéressante, il a aussi été démontré que ces réactions de défense sont contournées par des ET3 de la bactérie, ce qui permet l'apparition de symptômes sur riz (Jha *et al.*, 2007).

1.4. La dispersion

La dernière étape du cycle infectieux des bactéries phytopathogènes correspond à la dispersion de celles-ci. Cette dispersion peut être médiée par le vent, l'eau, le sol ou encore les semences et les plants contaminés, mais aussi par les insectes et l'homme. Cette étape de dispersion permet la transmission de l'agent pathogène à une nouvelle plante-hôte et/ou à la semence. Ainsi, cette nouvelle source d'inoculum permet à un nouveau cycle infectieux de démarrer sur un nouvel hôte.

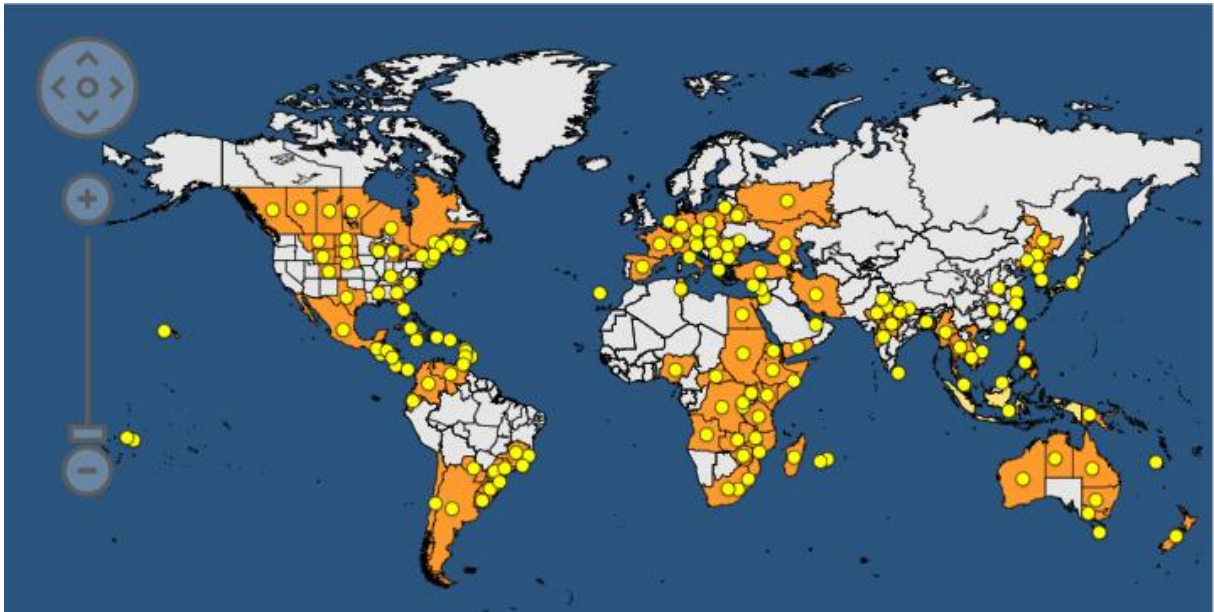


Figure 14 : Distribution mondiale des agents responsables de la graise commune du haricot. La présence de l'agent pathogène est représentée par un cercle jaune. La dernière actualisation de ces données de distribution a été faite le 13 Septembre 2017 par l'EPPPO ([European and Mediterranean Plant Protection Organization](https://www.eppo.int/)).

L'eau (de pluie ou d'irrigation), le vent, les tempêtes tropicales ou le cas échéant les insectes sont les vecteurs majeurs de dispersion des agents pathogènes, notamment à la surface des feuilles d'une même plante et entre plantes voisines, voire entre parcelles (Hayward, 1993; Hirano & Upper, 2000; Purcell, 2013). Par exemple, il a été démontré que la pluie est un facteur majeur de dispersion de *Pseudomonas syringae* pv. *tomato* qui, de plus, favorise la croissance des populations bactériennes sur tomate (Pietrarelli *et al.*, 2006). Certains agents phytopathogènes des genres *Liberibacter*, *Phytoplasma*, *Spiroplasma* et *Xylella* sont transmis entre plantes hôtes par des insectes vecteurs, essentiellement des hémiptères (Perilla-Henao & Casteel, 2016).

La transmission à une échelle plus importante, régionale voire intercontinentale, ou la survie entre deux cultures de la plante hôte est principalement le fait des semences contaminées. Ce mode de dispersion des agents pathogènes est presque général chez les *Xanthomonas* (Hayward, 1993). De plus, la dispersion des agents pathogènes peut se faire par la combinaison de plusieurs de ces facteurs. Ainsi, la dispersion de *X. euvesicatoria* pv. *allii* est favorisée par la succession d'un épisode venteux et pluvieux suivi d'une augmentation de la température (Roumagnac *et al.*, 2004). Chez *X. campestris* pv. *campestris*, les semences contaminées sont, certes, la source majeure d'inoculum et le principal facteur de dispersion, mais la dispersion de cet agent pathogène peut être également favorisée par d'autres facteurs tels que les insectes (Shelton & Hunter, 1985), le vent (Kuan *et al.*, 1986) et la pluie (Williams, 1980). Via ses différentes activités (taille, commerce, ...), l'homme est également le principal facteur de dissémination à grande échelle des bactéries phytopathogènes, ce qui favorise souvent l'émergence de phytobactérioses. En effet, avec les échanges de matériel végétal qui se sont largement intensifiés durant ces dernières décennies, les agents responsables de la graisse commune du haricot (*X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans*), originaires d'Amérique, ont pu se disséminer sur tous les continents par le biais de semences de haricot contaminées (**Figure 14**) (Hayward, 1993) et l'apparition de *X. arboricola* pv. *corylina* en France serait due à l'introduction de plants contaminés en provenance des Etats-Unis (Fischer-Le Saux *et al.*, 2015).

La majorité des *Xanthomonas* sont capables de se transmettre d'une génération de plantes à une autre par l'intermédiaire des graines. La transmission à la graine peut se faire selon trois voies différentes. La première voie emprunte les tissus conducteurs de la plante mère. Par exemple, *X. citri* pv. *fuscans* peut se transmettre aux semences de haricot en colonisant les organes reproducteurs à travers le système vasculaire (Darsonval *et al.*, 2008). Ce type de transmission à la graine par voie vasculaire est dépendant d'un SST3 fonctionnel et partiellement des adhésines (Darsonval *et al.*, 2008, 2009). La deuxième voie est indirecte et se fait par l'intermédiaire du stigmate de la plante. L'utilisation de cette voie a été démontrée pour le soja (Kauffman & Leben, 1974), le haricot (Darsonval *et al.*, 2008, 2009) et la pastèque (Walcott *et al.*, 2003). La troisième

voie correspond à un contact direct entre les téguments d'une graine saine et une surface symptomatique comme des graines ou des fruits présentant des symptômes ou encore des débris végétaux ou des pailles contaminées lors des étapes de récolte et de stockage (Weller & Saettler, 1980). Par ailleurs, il a été démontré que les bactéries ont une importante capacité de survie. Par exemple, *X. campestris* pv. *campestris* a été détecté dans des lots commerciaux de semences datant de plus de trois ans (Schaad & Dianese, 1981; Schultz & Gabrielson, 1986) et *X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans* ont été isolés de semences de haricot âgées de plus de 15 ans (Schuster & Coyne, 1974). Même après de longues périodes de latence, ces agents pathogènes transmis par les semences sont capables de se multiplier sur la graine dès lors qu'elle germera. De fortes densités de population bactérienne ne sont pas indispensables à la transmission des bactéries aux semences. En effet, des expérimentations au champ ont montré que des semences de haricot contaminées avec des inocula faiblement concentrés (de 10^2 à 10^4 UFC par semence) peuvent être à l'origine d'une colonisation de la phyllosphère et permettre la transmission de *X. citri* pv. *fuscans* aux graines de haricot (Darrasse *et al.*, 2007). De plus, les bactéries phytopathogènes sont capables de se maintenir dans des débris végétaux pendant plus ou moins longtemps selon les organismes et les conditions environnementales. Par exemple, *X. phaseoli* pv. *phaseoli* peut se maintenir dans les débris végétaux en climat tropical et subtropical (Arnaud-Santana *et al.*, 1991), mais cette survie ne dépassera pas plus de trois mois dans un climat correspondant à celui du Michigan (Saettler *et al.*, 1986). Egaleme nt, *X. campestris* pv. *campestris* peut survivre dans des résidus de culture pendant plus de 500 jours et durant cette période, l'inoculum est toujours susceptible de provoquer des épidémies (Schultz & Gabrielson, 1986). Ainsi, la survie est un élément majeur de la dispersion locale et de la dissémination lointaine des agents phytopathogènes.

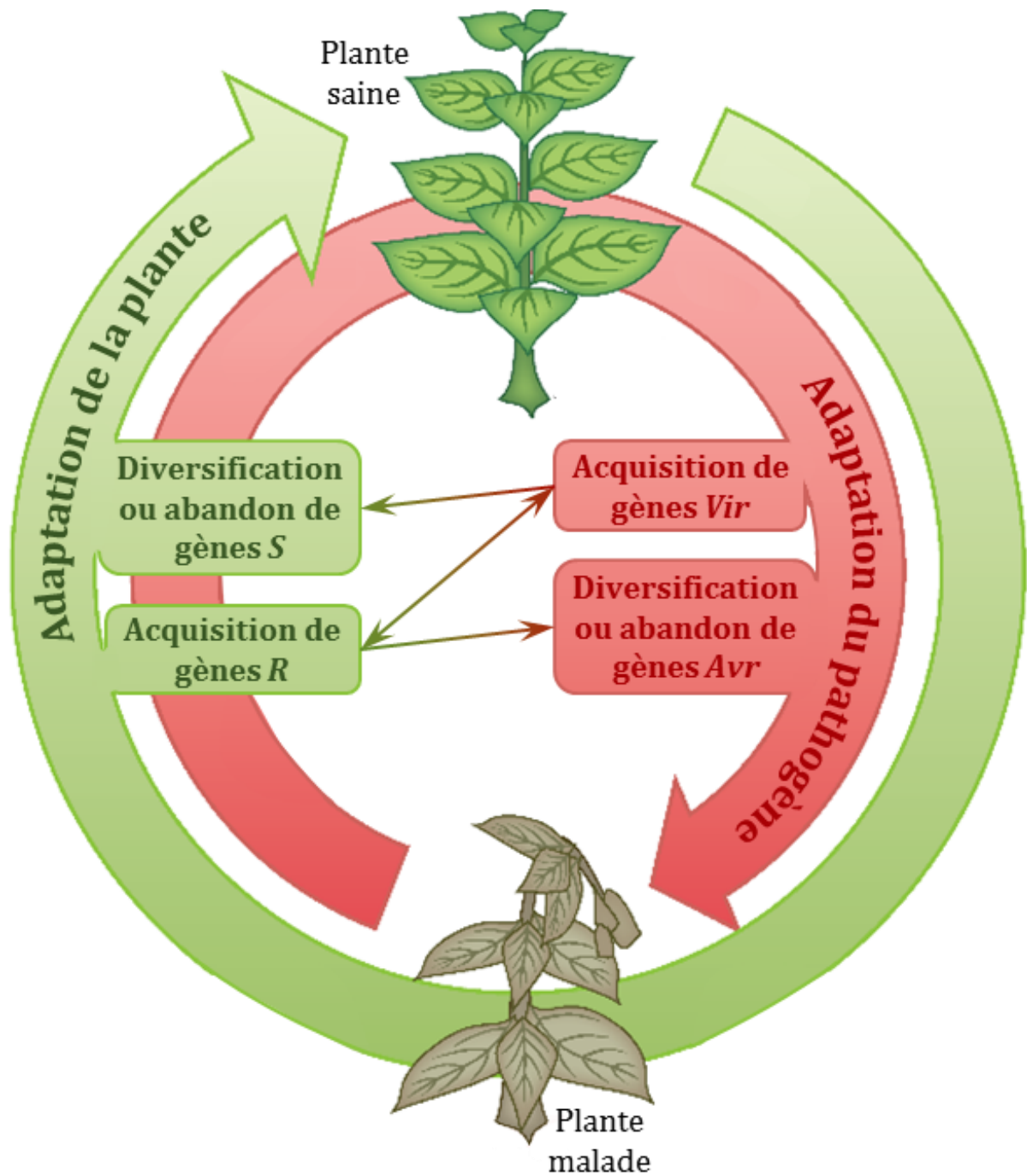
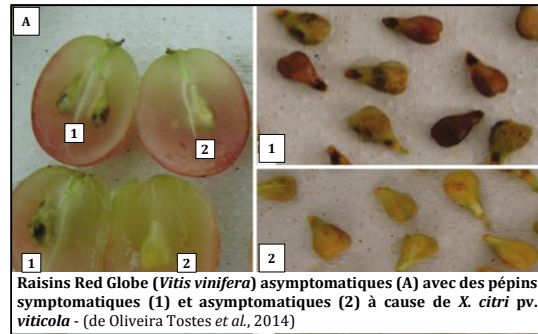
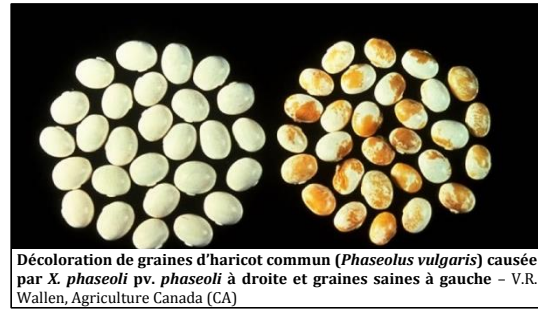


Figure 15 : Mécanismes adaptatifs sous l'influence de l'interaction plante-pathogène. Cette figure est adaptée de Vinatzer et al. 2014. Les agents pathogènes acquièrent des gènes de virulence (*Vir*) et diversifient ou abandonnent leurs gènes d'avirulence (*Avr*) afin de s'adapter à leurs plantes hôtes qui acquièrent des gènes de résistance (*R*) et diversifient ou abandonnent leurs gènes de sensibilité (*S*) pour répondre aux mécanismes adaptatifs des agents pathogènes.

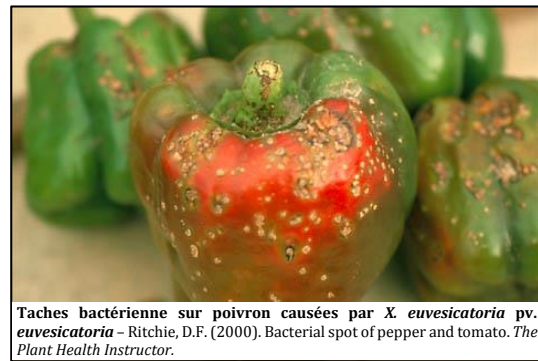
2. L'adaptation et la spécificité d'hôte chez les *Xanthomonas*

Dans le monde du vivant, l'adaptation correspond à une ou plusieurs modifications sélectionnées au cours de l'évolution car ayant pour résultat d'augmenter les chances de survie et de reproduction d'un individu, d'une lignée ou d'une espèce dans le milieu où ils vivent. Par exemple, pour les bactéries phytopathogènes, leurs milieux de vie et de reproduction correspondent pour une grande partie de leur cycle, aux plantes sur lesquelles elles s'établissent. Ainsi, les bactéries phytopathogènes doivent s'adapter à l'environnement qui les entoure et qui est régi par de multiples paramètres comme la température, l'hygrométrie, mais aussi et surtout la plante et ses différentes barrières de défense, comme décrit précédemment (**Figure 15**). L'adaptation implique en général initialement des événements de mutations géniques produisant de nouveaux gènes ou versions alléliques qui confèrent à la bactérie qui les porte un avantage sur les autres individus et qui seront alors transmis aux générations futures par transfert vertical ou à leurs congénères par transfert horizontal (Koonin *et al.*, 2001). L'adaptation peut conduire les bactéries phytopathogènes à se spécialiser pour coloniser spécifiquement un nombre plus ou moins limité d'espèces végétales, on parle alors de spécificité d'hôte. Cette spécificité d'hôte peut être plus ou moins étroite selon la taille du spectre d'hôtes de l'organisme.

Le genre *Xanthomonas* est un excellent modèle pour étudier les mécanismes régissant la spécificité d'hôte, car il est responsable de maladies sur plusieurs centaines d'espèces végétales différentes, mais individuellement, chaque souche n'attaque qu'un nombre très limité d'entre elles, avec parfois une spécificité à un type de tissus au sein de la plante (Vauterin *et al.*, 1995). Les bactéries du genre *Xanthomonas* sont réparties en 27 espèces capables d'infecter un large éventail d'espèces végétales, comprenant au moins 124 monocotylédones et 268 dicotylédones. Parmi les espèces végétales touchées, il y a notamment de nombreuses plantes cultivées d'intérêt agronomique, comme le riz, le poivron, le soja ou le coton, mais aussi de nombreuses plantes ornementales, telle que la jacinthe, le géranium ou le saintpaulia (Hayward, 1993; Ryan *et al.*, 2011; Jacques *et al.*, 2016). En plus d'être responsables de maladies sur une large gamme d'hôte, les bactéries du genre *Xanthomonas* provoquent des symptômes très différents, tels que des nécroses, des chancres ou encore des flétrissements, sur divers tissus et organes végétaux, comme les feuilles, les tiges, les fruits ou les semences (**Figure 16**). Par exemple, *X. hyacinthi* est l'agent pathogène responsable de la jaunisse bactérienne de la jacinthe, *X. oryzae* pv. *oryzae* s'est adapté aux tissus vasculaires du riz et cause la brûlure bactérienne du riz et *X. oryzae* pv. *oryzicola* infecte le parenchyme des feuilles de riz et provoque des symptômes sous forme de stries (Hayward, 1993). Cette spécificité suggère qu'il existe des barrières efficaces et durables chez les plantes non-hôtes, et que chaque souche possède l'arsenal adéquat, parfois spécifique d'un tissu, pour passer les barrières de sa plante hôte (Senthil-Kumar & Mysore, 2013). Ainsi, la majorité de ces



Pustules sur gousses de haricot causées par la graisse commune du haricot (*X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans*) – INRA Angers (France)



Chancre bactérien des agrumes sur pamplemousse causé par *X. citri* pv. *citri* – Shabbir A. Rizvi & Timothy Schubert (Etats-Unis)



Figure 16 : Exemples de maladies provoquées par des bactéries du genre *Xanthomonas*.
(1/2)

bactéries possède une spécificité d'hôte très étroite et cet aspect adaptatif a été traduit au niveau taxonomique par la notion de pathovar (pv.), définie comme un groupe infra-sous-spécifique de souches capables d'infecter et de produire les mêmes symptômes sur une même gamme d'hôte (Dye et al. 1980). Certains pathovars présentent néanmoins une large gamme d'hôtes, tels que *X. campestris* pv. *campestris*, capable d'infecter un grand nombre d'espèces végétales cultivées, ornementales et sauvages de la famille des Brassicacées, et *X. citri* pv. *citri*, responsable du chancre bactérien des agrumes qui infecte les parties aériennes (jeunes feuilles, rameaux, tiges, tronc, épines et fruits) de quelques espèces d'agrumes comme le pamplemoussier, le citronnier, l'oranger ou encore le mandarinier (OEPP/CABI, 1996; Vicente & Holub, 2013).

Aujourd'hui, les bases moléculaires et génétiques de la spécificité d'hôte ne sont pas encore intégralement définies mais des approches de type « génétique d'association » sur de larges collections de souches suggèrent que la spécificité d'hôte se dessine progressivement à toutes les étapes du cycle infectieux. Ainsi, **(i)** certains **gènes candidats** impliqués dans la spécificité d'hôte de bactéries du genre *Xanthomonas* ont été déterminés, **(ii) la notion de répertoire** de gènes candidats pour la spécificité d'hôte a été définie et **(iii)** de nombreux **événements évolutifs** ont été caractérisés comme étant impliqué dans ce processus adaptatif.

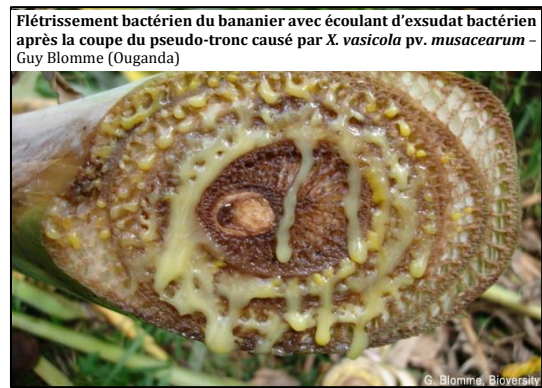
2.1. Les gènes candidats impliqués dans la spécificité l'hôte

Des approches de type « génétique d'association » ont permis de montrer que les MCP, les adhésines, les TBDT et les ET3 sont de potentiels déterminants de la spécificité d'hôte (Hajri *et al.*, 2009, 2012a,b; Mhedbi-Hajri *et al.*, 2011). De plus, des analyses des séquences de ces gènes ont conforté ce résultat et, notamment, il a été démontré que la composition en ET3 dans un génome variait selon les pathovars au sein de l'espèce *X. axonopodis* (sensu Vauterin *et al.*, 1995, 2000; Hajri *et al.*, 2009). Par ailleurs, d'autres analyses ont montré que certains de ces gènes étaient conservés chez toutes les souches capables d'infecter le haricot et absents chez celles qui n'en sont pas capables, alors que ces souches sont phylogénétiquement éloignées (Alavi *et al.*, 2008).

Les MCP permettent à la bactérie de reconnaître des molécules provenant de son environnement afin d'emprunter la direction la plus appropriée pour sa survie. De manière intéressante, la distribution des gènes codant des MCP est d'avantage liée à la gamme d'hôte des souches qu'à leur phylogénie. Par exemple, *X. axonopodis* pv. *vasculorum* et *X. vasicola* pv. *vasculorum* sont phylogénétiquement distantes (Rademaker *et al.*, 2005), mais elles partagent un hôte commun, la canne à sucre, et possèdent aussi des gènes codant des MCP en commun (Mhedbi-Hajri *et al.*, 2011). Cela confirme donc que, pour le genre *Xanthomonas*, les MCP sont importants pour leur mode de vie et sont sélectionnés en fonction de celui-ci afin que la bactérie s'adapte à



Chancres verticaux suintant sur noyer (*Juglans* sp.) avec production d'exsudat bactérien noir colorant le tronc causé par *X. arboricola* pv. *juglandis* – (Hajri *et al.*, 2010)



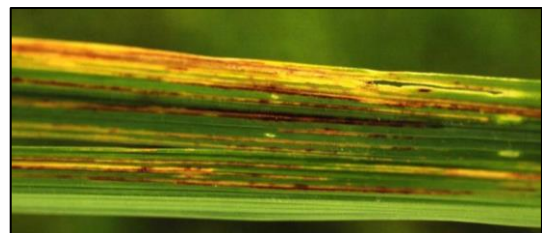
Fétriissement bactérien du bananier avec écoulement d'exsudat bactérien après la coupe du pseudo-tronc causé par *X. vasicola* pv. *musacearum* – Guy Blomme (Ouganda)



Chancres bactériens des agrumes sur tige de Natsumikan (*Citrus natsudaia*) causé par *X. citri* pv. *citri* – M. Goto (Japon)



Taches nécrotiques et halos chlorotiques sur feuilles de haricot causés par *X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans* – (Akhavan *et al.*, 2013)



Bactériose à stries foliaires du riz (*Oryza sativa*) causée par *X. oryzae* pv. *oryzicola* – Donald Groth (Etats-Unis)



Pustules bactériennes sur feuille de soja (*Glycine max* (L.) Merr.) causées par *X. citri* pv. *glycines* – Daren Mueller (Etats-Unis)



Bactériose vasculaire du riz avec gouttes d'exsudat bactérien causée par *X. oryzae* pv. *oryzae* – www.knowledgebank.irri.org

Taches angulaires du fraisier causées par *X. fragariae* – P. Nitzsche (Etats-Unis)



Nervation noire des crucifères sur feuille de choux causée par *X. campestris* pv. *campestris* – Gerald Holmes (Etats-Unis)

Figure 17 : Exemples de maladies provoquées par des bactéries du genre *Xanthomonas*.
(2/2)

son environnement voire se spécialise à son hôte (Thieme *et al.*, 2005). De plus, une analyse des signatures de sélection de gènes codant des MCP chez différents pathovars a confirmé que les différences présentes entre les compositions en MCP dans les génomes correspondent à de la divergence adaptative spécifiques à la plante hôte (Mhedbi-Hajri *et al.*, 2011).

L'adhésion des bactéries à leur plante hôte est un phénomène précoce qui pourrait être également impliqué dans l'adaptation à l'hôte. En effet, cette étape du cycle infectieux leur permet de s'agréger en biofilm, afin qu'elles puissent résister à diverses contraintes, mettre en place des réponses adaptées aux changements environnementaux et entrer en phase de multiplication, autant d'évènements nécessitant une adaptation à l'hôte efficace (Danhorn & Fuqua, 2007; Amano, 2010). Par conséquent, les adhésines et les polysaccharides semblent être de bons candidats pour la spécificité d'hôte des souches de *Xanthomonas*. En effet, la comparaison de trois espèces phylogénétiquement éloignées de *Xanthomonas* infectant le poivron et / ou la tomate a révélé des gènes codant des LPS communs chez ces agents pathogènes, ce qui indique que la distribution de ces gènes est d'avantage liée au spectre d'hôte des souches qu'à leurs phylogénies (Potnis *et al.*, 2011). D'autre part, un mutant du gène *opsX* impliqué dans la biosynthèse d'EPS et de LPS produit moins d'EPS et des LPS dont la structure est altérée, ce qui a pour conséquences de diminuer le taux de croissance de *X. campestris* pv. *citrumelo* et une perte de virulence sur agrume mais, de façon surprenante, pas sur haricot (Kingsley *et al.*, 1993). Ainsi, cet exemple suggère que les polysaccharides sont directement impliqués dans la spécificité d'hôte des *Xanthomonas*, et leurs altérations peuvent modifier et / ou réduire la gamme d'hôte d'une souche. D'autre part, on peut imaginer que la modification de la structure d'un LPS permettrait à la bactérie de contourner les défenses basales de la plante, étant donné qu'ils sont le principal composant de la membrane externe des bactéries à Gram-négatif et peuvent être potentiellement reconnus en tant que MAMP par les PRR de la plante, responsables du déclenchement de la PTI.

Le transport TonB-dépendant au sens large, dont notamment le système CUT, semble être aussi un élément clé de l'adaptation des bactéries du genre *Xanthomonas* à leurs plantes hôtes (Jacques *et al.*, 2016). Cette hypothèse est notamment soutenue par le fait que les gènes impliqués dans ce type de transport sont surreprésentés chez les bactéries du genre *Xanthomonas*, alors qu'il est consommateur d'énergie, ce qui devrait être défavorable pour la survie des bactéries (Mhedbi-Hajri *et al.*, 2011). Par exemple, une analyse du génome de *X. campestris* pv. *campestris* a prédit 72 gènes codant des récepteurs TonB-dépendants, ce qui semble confirmer que les avantages apportés par les TBDT sont bien supérieurs au coût de leurs utilisations (Blanvillain *et al.*, 2007). De la même façon que pour les MCP, les répertoires de TBDT diffèrent au sein des génomes de pathovars et de lignées génétiques appartenant aux *Xanthomonas* et ayant une gamme d'hôte différente (Mhedbi-Hajri *et al.*, 2011). Par exemple, un récepteur TonB-dépendant a été identifié

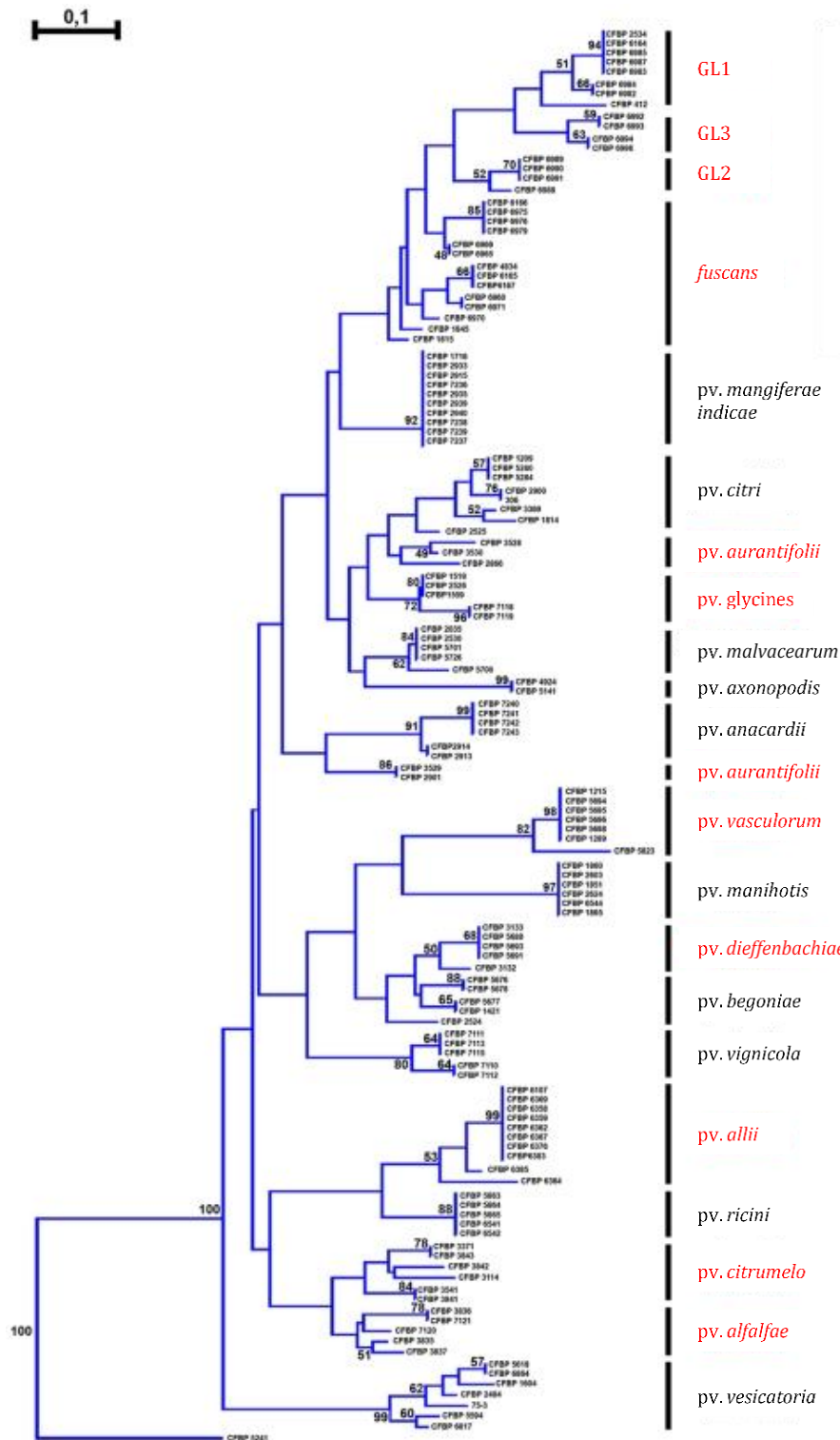


Figure 18 : Corrélation entre les compositions en ET3 et les spectres d'hôtes chez l'espèce *X. axonopodis*. Cette figure est adaptée de Hajri et al. 2009. Le dendrogramme a été obtenu à partir de la matrice présence/absence de 35 ET3 chez 132 souches de *X. axonopodis* et a été construit par la méthode de Neighbor Joining, utilisant les distances de Jaccard, avec la souche CFBP5241 de *X. campestris* pv. *campestris* comme outgroup. Seules les valeurs significatives (>50) de bootstraps (1000 itérations) sont indiquées. Les pathovars polyphylétiques sont en rouge et les monophylétiques, en noir. GL1, GL2, GL3 et *fuscans* correspondent aux quatre lignées génétiques responsables de la gousse commune du haricot et appartiennent aux souches *X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans*, anciennement nommées *X. axonopodis* pv. *phaseoli* et *X. fuscans* subsp. *fuscans*.

parmi les traits génomiques spécifiques des souches de *X. translucens* pv. *graminis* et aucun homologue de ce récepteur n'a été identifié chez les autres pathovars de *X. translucens*, ce qui suggère que ce récepteur est impliqué dans la spécificité d'hôte (Hersemann *et al.*, 2017).

Les ET3 sont des acteurs majeurs de l'interaction hôte-pathogène, leur présence étant parfois indispensable pour la mise en place de la maladie et leur diversité fonctionnelle leur permettant de jouer un rôle dans l'adaptation et la spécificité d'hôte des bactéries phytopathogènes. En effet, les ET3 « core » sont d'avantage engagés dans le cadre d'une spécificité d'hôte large, car ils permettent aux bactéries de survivre sur une large gamme de plantes hôtes sensibles et non-sensibles, tandis que les ET3 « variables » interviennent majoritairement dans le cadre d'une spécificité d'hôte plus étroite, car leur action est dirigée spécifiquement contre l'hôte en faveur des bactéries, ce qui permet aux agents pathogènes d'exprimer leurs pouvoirs pathogènes et de se disséminer pour s'installer sur un nouvel hôte (Ryan *et al.*, 2011). Egalement, il a été démontré que certains ET3 sont directement impliqués dans la gamme d'hôte d'un agent pathogène. Par exemple, il a été démontré que la perte d'AvrBst par *X. perforans* lui permet d'étendre son spectre d'hôte sur le poivron et que si elle perd en plus un autre ET3, nommé XopQ, elle est également capable d'infecter *Nicotiana benthamiana*. Ces diminutions du spectre d'hôte induites par AvrBst ou XopQ sont la conséquence du fait qu'ils sont ciblés par l'ETI chez le poivron ou *N. benthamiana*, respectivement (Schwartz *et al.*, 2015; Boulanger & Noël, 2016). Egalement, il a été démontré que, chez une souche de *X. citri* pv. *aurantifolii* de type C, la présence de l'ET3 avrGf2 entraîne une HR sur pamplemousse et que sa mutation rétablit une réaction compatible. Cet effet direct d'avrGf2 sur la gamme d'hôte a été d'autant plus confirmé par la présence d'avrGf1, présentant 45 % d'homologie de séquence avec avrGf2, chez *X. citri* pv. *citri* de type A^w, qui engendre aussi une HR sur pamplemousse à cause de cet ET3 (Gochez *et al.*, 2015). Ainsi, les espèces du genre *Xanthomonas* partageant la même gamme d'hôte partagent également les mêmes ET3. Ceci a en effet été confirmé chez des souches de *X. axonopodis* où les compositions en ET3 dans les génomes sont corrélés à leurs spectres d'hôte (**Figure 18**) (sensu Vauterin *et al.*, 1995, 2000; Hajri *et al.*, 2009). Par exemple, différentes analyses de génomique comparative ont montré que 65 souches de *X. axonopodis* pv. *manihotis* possèdent neuf ET3 très conservés en commun (Bart *et al.*, 2012). Egalement, 13 ET3 sont conservés chez les trois espèces de *Xanthomonas* infectant le poivron et / ou la tomate (Schwartz *et al.*, 2015). Ce partage d'ET3 entre souches phylogénétiquement éloignées mais ayant le même spectre d'hôte a notamment pu être exploité afin de développer un outils d'identification moléculaire des souches infectant le haricot commun par PCR multiplex en ciblant deux gènes codant des ET3 : *avrBst* et *xopL* (Boureau *et al.*, 2013).

2.2. La notion de répertoire

Au regard de tous ces exemples, il apparaît peu probable que l'adaptation et la spécificité d'hôte des bactéries phytopathogènes se résume à un unique gène. En effet, comme explicité ci-dessus, les souches du genre *Xanthomonas* possèdent une grande quantité de gènes de virulence différents parfois partagés en grand nombre entre souches ayant la même gamme d'hôte. Ainsi, l'implication de ces différents facteurs de virulence au cours des étapes du cycle infectieux suggère que la spécificité d'hôte repose sur un répertoire de gènes communs à toutes les souches d'un même pathovar. En effet, il a été démontré qu'une petite proportion de la variation génétique totale de 60 souches de *Pseudomonas* pouvait être expliquée par leur gamme d'hôte (Sarkar & Guttman, 2004), et il a été proposé une hypothèse selon laquelle un répertoire de gènes de résistance de la plante correspond à un répertoire de gènes de virulence chez le genre bactérien *Xanthomonas*, et que c'est cette combinaison qui définit la spécificité d'hôte. Ainsi, la spécificité d'hôte peut être expliquée par l'association de gènes de virulence permettant une convergence pathologique de souches phylogénétiquement éloignées (Hajri *et al.*, 2009).

Aujourd'hui, cette notion de répertoire de gènes impliqué dans la spécificité d'hôte a été confirmée à plusieurs reprises, notamment chez des bactéries du genre *Xanthomonas*. En effet, il a été démontré qu'il existe différents répertoires d'ET3 proches mais qui varient en taille et en contenu et qui permettent de différencier les sept pathovar de *X. arboricola* possédant des spécificités d'hôte différentes (Hajri *et al.*, 2012b). De plus, chez cette même espèce, deux types de répertoires d'adhésines permettant de différencier les souches pathogènes et des souches non pathogènes ont été identifiés et sont certainement à l'origine de cette différence de pathogénie (Cesbron *et al.*, 2015). Egalement, une étude a montré que 55 souches représentant la diversité mondiale de *X. citri* pv. *citri* possèdent 28 ET3, dont sont 26 partagés par toutes les souches (Escalon *et al.*, 2013). De manière intéressante, les répertoires de gènes peuvent aussi présenter une légère variabilité qui permet une caractérisation plus fine de la spécificité d'hôte. Par exemple, les répertoires d'ET3 des souches de *X. oryzae* pv. *oryzae* et *X. oryzae* pv. *oryzicola* sont très proches mais différenciables et sont probablement impliqués dans la spécialisation tissulaire de ces agents pathogènes du riz (Hajri *et al.*, 2012a).

Néanmoins, pour un même pathovar le répertoire d'un type de gène n'est parfois pas suffisant pour définir sa spécificité d'hôte. En effet, il existe des exemples où des répertoires de gènes ne sont pas parfaitement corrélés à la gamme d'hôte. Par exemple, une analyse de type présence / absence de gènes a montré qu'il existe un nombre extrêmement faible de gènes spécifiques des souches de *X. citri* pv. *citri* (Gordon *et al.*, 2015). Egalement, chez l'espèce *X. campestris* seuls trois ET3 putatifs communs ont été déterminés (Roux *et al.*, 2015). Ainsi, le répertoire de gènes caractérisant la spécificité d'hôte serait composé de plusieurs types de gène.

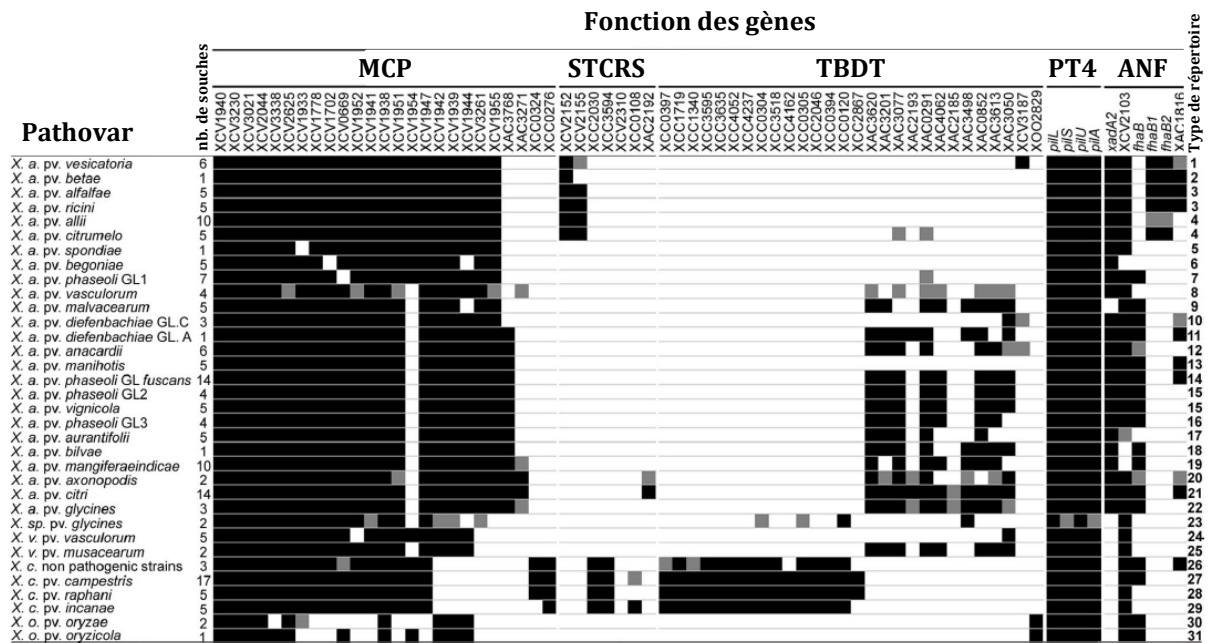


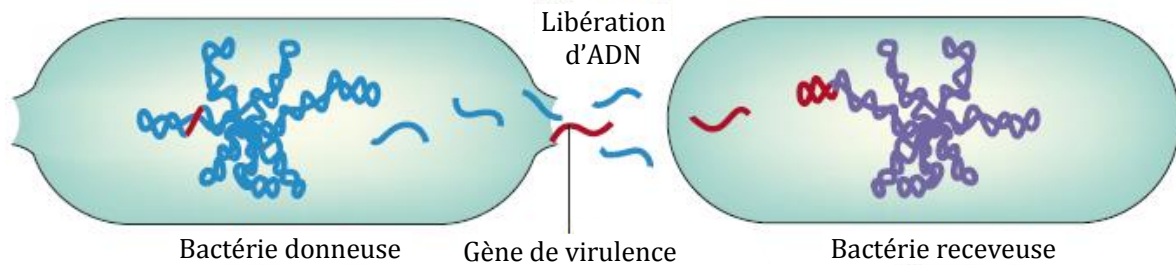
Figure 19 : Distribution de gènes potentiellement impliqués dans la spécificité d'hôtes de bactéries du genre *Xanthomonas*. Cette figure est adaptée de Mhedbi-Hajri et al. 2011. Les carrés noir et blanc représentent, respectivement, la présence et l'absence du gène correspondant dans au moins 80% des souches d'un pathovar ou d'une lignée génétique. Les carrés gris indiquent que la présence du gène est variable parmi les souches de la lignée génétique. Cette analyse a permis d'identifier 31 types de répertoires différents parmi une diversité de bactéries du genre *Xanthomonas*. MCP : Methyl-accepting Chemotaxis Proteins ; STCRS : Sensors of Two-Component Regulatory System ; TBDT : TonB-Dependent Transporters ; PT4 : Pili de Type 4 ; ANF : Adhésines Non Fibrillaires.

En effet, il a été démontré que les traits génomiques conservés chez *X. translucens* pv. *graminis* et absent des autres *X. translucens* comprennent une multitude de gènes, dont notamment des gènes codant un SST3, un pilus de type 4, un récepteur TonB-dépendant, des transporteurs et des ET3 (Hersemann *et al.*, 2017). De même, plus de 100 gènes ont été identifiés comme étant de potentiels déterminants de la spécificité d'hôte et de la convergence pathologique de deux lignées phylogénétiquement distantes responsables de la graisse commune du haricot (Aritua *et al.*, 2015). Egalement, de façon plus large, chez 173 souches appartenant à différents pathovars de *Xanthomonas* et ayant différentes gammes d'hôtes, il a été démontré que c'est la combinaison des cinq répertoires de gènes codant des MCP, des senseurs des systèmes de régulation à deux composants, des TBDT et des adhésines fibrillaires et non fibrillaires qui corrèle avec le regroupement des souches en pathovars (**Figure 19**) (Mhedbi-Hajri *et al.*, 2011). Ainsi, la spécificité d'hôte est le résultat de l'interaction de plusieurs répertoires de gènes de virulence de la bactérie qui correspondent certainement à des répertoires de gènes de défenses chez la plante hôte (Hajri *et al.*, 2009).

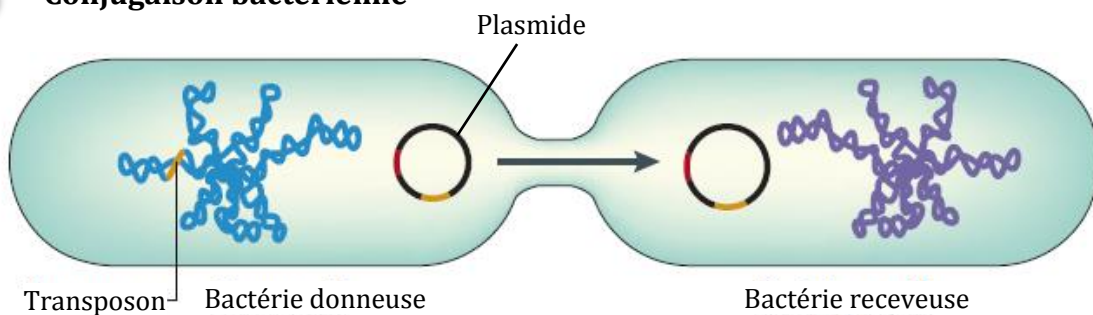
2.3. Les évènements évolutifs mis en jeu

Comme détaillé ci-dessus, la spécificité d'hôte est un processus complexe engageant plusieurs répertoires de gènes impliqués dans différents processus essentiels à la bactérie. Des analyses comparatives de génomes complets sur de large collection de souches ont récemment permis de mettre à jour de potentiels nouveaux déterminants de la spécificité d'hôte dont les fonctions peuvent, notamment, n'avoir jamais été explicitées comme étant impliquées dans ce processus adaptatif. La spécificité d'hôte serait donc caractérisée, non pas par un gène de virulence ou un répertoire de gènes, mais par une large collection de gènes contenant des signatures spécifiques pour chaque pathovar. Ces signatures, comme par exemple le polymorphisme nucléotidique (SNP) ou la trace de sélection positive, sont également des traces potentielles de l'adaptation à l'hôte dans le génome des bactéries (Jacques *et al.*, 2016). Les mécanismes évolutifs les plus relatés comme étant moteur de l'évolution sont la mutation, le transfert horizontal de gènes (THG), et la recombinaison homologue. Ces mécanismes moléculaires engendrent diverses modifications dans le génome. Par exemple, la mutation ponctuelle permet l'apparition de nouveaux allèles ou l'inactivation d'un gène, le THG peut permettre l'acquisition de nouveaux gènes et la recombinaison homologue peut être responsable de réarrangements génomiques importants mais aussi, comme la mutation, de l'apparition de nouveaux allèles ou de l'inactivation d'un gène. Ainsi, ces mécanismes évolutifs sont les moteurs de l'adaptation à l'hôte et de la convergence pathologique de souches phylogénétiquement éloignées, mais aussi de l'émergence de maladies (Bartoli *et al.*, 2016).

1 Transformation bactérienne



2 Conjugaison bactérienne



3 Transduction bactérienne

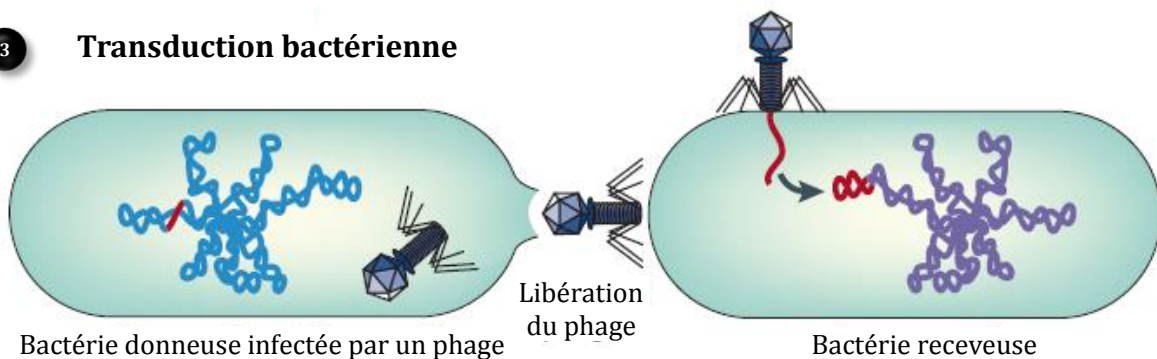


Figure 20 : Mécanismes régissant le transfert horizontal de gènes entre bactéries. Cette figure est adaptée de Furuya and Lowy 2006. **(1)** La transformation bactérienne correspond à de l'ADN nu libéré par une bactérie donneuse lors d'une lyse qui est récupérée par une bactérie receveuse. Le gène de virulence ainsi récupéré peut être ensuite intégré au chromosome ou au plasmide de la bactérie receveuse. **(2)** La conjugaison bactérienne se produit via un contact direct entre deux bactéries. Les plasmides provoquent la formation de ponts de conjugaison entre les bactéries par lesquels de l'ADN, notamment plasmidique, est échangé, ce qui peut entraîner l'acquisition de gènes de virulence par la bactérie receveuse. Les transposons sont des séquences d'ADN qui portent leurs propres enzymes de recombinaison permettant la transposition d'un endroit à l'autre du génome et entre différents génomes. Les transposons peuvent être associés à des gènes de virulence qui peuvent être transposés dans le génome receveur lors d'une conjugaison bactérienne. **(3)** La transduction bactérienne permet le transfert de gènes de virulence entre une bactérie donneuse et une bactérie receveuse grâce à des bactériophages. De la même façon, les gènes récupérés peuvent être intégrés au chromosome de la bactérie receveuse, c'est ce qu'on appelle la lysogénie.

La mutation

La mutation correspond à une modification de l'ADN qui peut n'avoir aucune conséquence ou entraîner l'apparition d'un nouveau phénotype potentiellement transmissible aux générations futures. Il existe trois types de mutations : l'insertion, la délétion ou la substitution nucléotidique. Ainsi, ce mécanisme est une source de diversité génétique, elle-même moteur de l'évolution et par voie de conséquence moteur de l'adaptation et de la spécificité d'hôte. En effet, lors d'études de génomique comparative la mutation est souvent analysée car elle représente un indicateur de l'évolution (Bart *et al.*, 2012; Mhedbi-Hajri *et al.*, 2013). Par exemple, il a été démontré que dans le genre *Xanthomonas*, la mutation était la principale force évolutive des gènes *hrp*, ce qui a amené du polymorphisme et généré de nouveaux allèles (Merda *et al.*, 2017). Egalement, une analyse de génomique comparative des trois pathotypes de *X. citri* pv. *citri*, présentant différentes gammes d'hôtes, a permis d'identifier 426 Single Nucleotide Polymorphisms (SNP), en comparant les pathotypes deux à deux (Gordon *et al.*, 2015). D'autre part, la mutation peut être aussi responsable de l'inactivation d'un gène ou de la diminution de la virulence. Notamment, il existe de nombreux exemples de mutants, déficients dans leur virulence, produits lors d'expérimentations où l'on mute un gène afin d'en caractériser le rôle. Par exemple, chez *X. oryzae* pv. *oryzae*, la mutation des gènes *gltB/D*, qui codent respectivement les sous-unités α et β de la glutamate synthase, engendre une avirulence, tout comme la mutation de *CbsA*, une exoglucanase sécrétée par ces bactéries (Pandey *et al.*, 2014; Tayi *et al.*, 2017).

Le transfert horizontal de gènes

Le THG (ou recombinaison hétérologue) correspond à un échange de matériel génétique entre un organisme donneur et un organisme receveur, alors que ces deux organismes ne sont pas parents. Ainsi, le THG facilite la diversité et est un moteur essentiel et puissant de l'évolution adaptative des bactéries et de l'émergence de nouveaux agents pathogènes (Wiedenbeck & Cohan, 2011). Par exemple, il a été montré que deux gènes codant des ET3 potentiellement impliqués dans la spécificité d'hôte ont été transférés horizontalement entre *X. citri* pv. *bilvae* et certaines souches de *X. citri* pv. *citri* (Escalon *et al.*, 2013). L'émergence des pathovars *pruni*, *juglandis* et *corylina* de l'espèce *X. arboricola* semble également être liée à l'acquisition de huit gènes codant des ET3 par THG (Merda *et al.*, 2016). Il existe trois mécanismes régissant le THG : la transformation, la transduction et la conjugaison (**Figure 20**). La transformation correspond à un mécanisme où, lors d'une lyse, de l'ADN nu est libéré par une bactérie donneuse et est récupérée par une bactérie receveuse. Cet ADN peut être ensuite intégré au génome de la bactérie receveuse, notamment s'il lui confère un avantage adaptatif, comme par exemple s'il correspond à un gène de virulence. La transduction est un mécanisme similaire à la transformation mais le transfert de

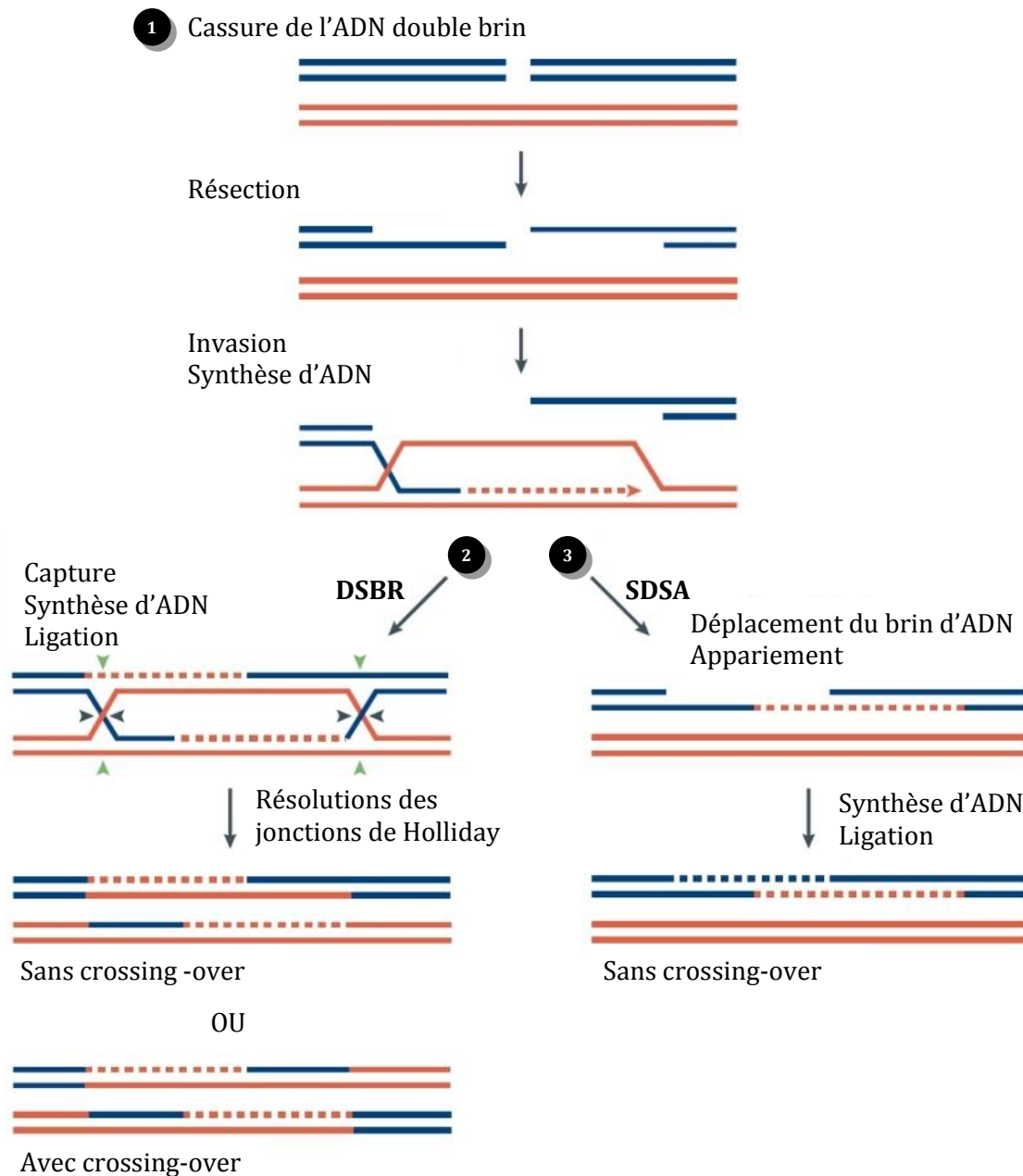


Figure 21 : Mécanisme de la recombinaison homologue. Cette figure est adaptée de Sung and Klein 2006. **(1)** La cassure de l'ADN double brin est à l'origine de la recombinaison homologue. Cette cassure engendre la résection, c'est-à-dire la dégradation du brin d'ADN de part et d'autre de la cassure, ce qui expose un simple brin d'ADN. L'ADN simple brin généré par la résection va alors envahir une autre séquence d'ADN homologue et une fois hybridés la synthèse de l'ADN est déclenchée. Cette synthèse d'ADN est suivie par un processus de réparation de l'ADN. Il existe deux types de réparation de l'ADN : le DSBR (double-strand break repair) et le SDSA (synthesis-dependent strand annealing). **(2) DSBR** : Après l'invasion suivie de la synthèse d'ADN, le deuxième simple brin d'ADN généré peut être capturé pour former un intermédiaire avec deux jonctions de Holliday. Cet intermédiaire engendre un processus de synthèse et de ligature de l'ADN. La structure obtenue est alors résolue au niveau des jonctions de Holliday avec ou sans crossing-over. La résolution sans crossing-over se fait au niveau des flèches noires alors qu'avec crossing-over elle se fait au niveau des flèches vertes à une jonction et au niveau des flèches noires à l'autre jonction. **(3) SDSA** : Après l'invasion suivie de la synthèse d'ADN, le simple brin d'ADN peut glisser et s'apparier à l'ADN complémentaire sur l'autre extrémité de la cassure. Après cet appariement, un processus de synthèse et de ligature de l'ADN est déclenché pour combler l'espace vide. Le crossing-over n'est donc pas utilisé pour le SDSA.

gènes entre les bactéries se fait grâce à des bactériophages. La conjugaison correspond à un mécanisme où de l'ADN est échangé entre deux bactéries via un pont de conjugaison, c'est-à-dire un contact direct entre bactéries. Ce pont correspond à un système de sécrétion de type 4 (SST4) qui peut être impliqué dans l'échange de diverses macromolécules, et dont les gènes ont souvent une localisation plasmidique. L'ADN échangé est souvent plasmidique et peut concerner des plasmides entiers. Les plasmides sont connus pour être porteurs de gènes conférant un avantage adaptatif comme des déterminants de la virulence et leur acquisition par THG permet une évolution rapide de l'organisme receveur. Par exemple, une bactérie épiphyte commensale répandue, *Pantoea agglomerans*, est devenue un pathogène tumorigène avec une spécificité d'hôte définie après l'acquisition d'un plasmide contenant un îlot de pathogénie par THG (Barash & Manulis-Sasson, 2009). Il est intéressant de noter que la plupart des plasmides provenant de bactéries phytopathogènes sont autotransmissibles et possèdent souvent des gènes codant des SST4 (Lavigne *et al.*, 2006), mais aussi des éléments génétiques mobiles, comme les transposons ou les séquences d'insertion (IS), qui contribuent aux mouvements de séquences au sein et entre génomes en favorisant la recombinaison homologue. Ainsi, après l'acquisition d'un plasmide apportant un avantage adaptatif, la bactérie peut l'intégrer dans son chromosome afin de stabiliser ces gènes nouvellement acquis (Sundin, 2007). Bien que le THG adaptatif soit très souvent avantageux pour la bactérie, il peut engendrer une réduction de la fitness bactérienne, qui peut être contrebalancée par la sélection naturelle sur les traits bénéfiques apportés par les gènes transférés (Wiedenbeck & Cohan, 2011). Par exemple, il a été démontré que le pathotype A de *X. citri* pv. *citri* a une gamme d'hôte plus large que les pathotypes A* et A^w en raison d'une acquisition de gènes bénéfiques et d'une perte de gènes préjudiciables combinées à une sélection positive de près de 14 % des gènes core, notamment impliqués dans le transport et le métabolisme des sucres ; la réplication, la recombinaison et la réparation de l'ADN ; et la virulence bactérienne (Zhang *et al.*, 2015a).

La recombinaison homologue

La recombinaison homologue (ou transfert horizontal de remplacement) correspond à un échange de matériel génétique entre deux séquences d'ADN homologues provenant d'un même génome ou de deux génomes différents et permet l'apparition de gènes présentant une nouvelle organisation, donc de nouveaux allèles (**Figure 21**). La recombinaison homologue peut se produire après l'acquisition de séquences d'ADN exogènes par THG. En effet, une fois que l'ADN est acquis par la bactérie receveuse, il peut recombiner avec l'ADN déjà présent. Ainsi, une nouvelle séquence allélique peut apparaître et potentiellement contribuer à la spécificité d'hôte. En effet, la spéciation bactérienne est généralement associée à des événements de recombinaison

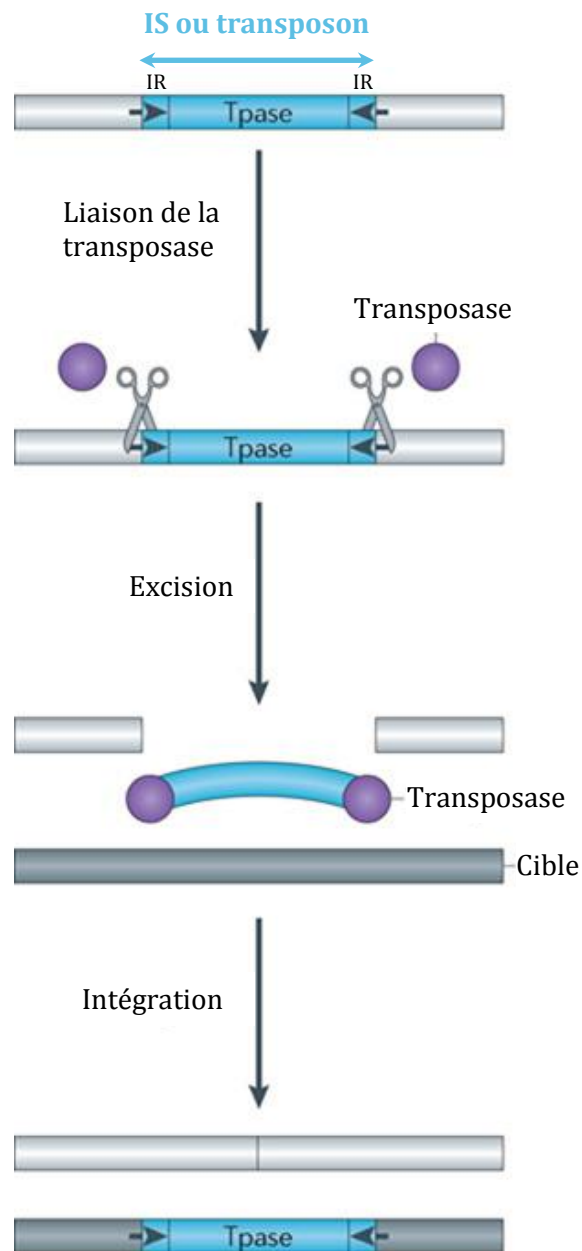


Figure 22 : Mécanisme de type « coupé-collé » mobilisant un transposon bactérien. Cette figure est adaptée de Levin and Moran 2011. Les transposons bactériens sont aussi appelés transposons à ADN et le mécanisme de type « coupé-collé » dont ils sont responsables est représenté par des ciseaux. Généralement, les transposons bactériens ou IS (Insertion Sequence) sont composés d'un gène codant une transposase (Tpase) flanqué par des répétitions inversées ou IR représentées par des flèches noires. Les transposases produites, représentées par des cercles violet, se lient à proximité ou au niveau des IR, pour extraire le transposon de son emplacement génomique existant, représenté en gris clair, et le coller dans un nouvel emplacement génomique, représenté en gris foncé.

homologue durables, qui peuvent agir comme une force de convergence évolutive entre lignées phylogénétiquement distantes, et qui peuvent ainsi modifier la gamme d'hôte d'un agent pathogène en lui apportant de nouveaux allèles avantageux qui seront alors conservés par sélection naturelle. Par exemple, une comparaison des espèces *X. campestris* et *X. citri* a révélé que 10 % du core génome impliquant des événements de recombinaison homologue a contribué à la diversification de ces espèces, que 179 gènes montrent des traces de sélection positive, et que 98 d'entre eux sont impliqués dans la recombinaison homologue, ce qui suggère que des fragments d'ADN étrangers sont à l'origine de la diversification adaptative de ces deux espèces (Huang *et al.*, 2015). Egalement, il a été démontré qu'au cours de l'histoire évolutive de *X. axonopodis* les événements de recombinaison étaient aussi fréquents que les mutations ponctuelles, mais que leur impact sur la diversité était trois fois plus important (Mhedbi-Hajri *et al.*, 2013).

Les gènes nécessaires pour qu'une bactérie devienne pathogène ou soit spécifique d'un hôte donné ne sont pas tous acquis en une fois et via un unique type de mécanisme évolutif. En effet, il apparaît plus probable que l'émergence d'un nouvel agent pathogène se fasse en plusieurs étapes. Par exemple, les souches de *Xanthomonas* pathogènes sur cannabis ne possèdent ni SST3, ni ET3, mais possèdent des gènes régulateurs du SST3. Ainsi, il a été suggéré que la pathogénie aurait été acquise via un processus évolutif progressif, qui implique (i) l'acquisition de gènes régulateurs clés et d'enzymes de dégradation de la paroi cellulaire, suivis de (ii) l'acquisition d'un SST3, qui est finalement accompagnée (iii) d'une succession d'acquisitions d'ET3 (Jacobs *et al.*, 2015). De plus, après acquisition de gènes conférant un avantage adaptatif, ces derniers peuvent évoluer par la suite permettant ainsi à la bactérie de s'adapter voire de se spécialiser à son hôte. Par exemple, pour répondre au système de défense de son hôte, la bactérie pathogène *Pseudomonas syringae* a acquis deux ET3 par THG, HopZ2 et HopZ3, et a diversifié, grâce à diverses mutations, un autre ET3, HopZ1, en trois formes fonctionnelles et deux formes dégénérées. Ainsi, la combinaison de ces différents événements évolutifs a permis à la bactérie de contourner le mécanisme de défense de la plante tout en conservant ces ET3 (Ma *et al.*, 2006). Egalement, une analyse de génomique comparative des trois pathotypes de *X. citri* pv. *citri*, présentant différentes gammes d'hôtes, a montré qu'un ensemble de THG, de SNP et d'événements de recombinaison permettait de différencier ces trois pathotypes (Gordon *et al.*, 2015).

D'autre part, comme évoqué ci-dessus, les éléments génétiques mobiles contribuent aux mouvements de séquences dans les génomes, ce qui peut conduire à la conservation de séquences d'intérêts, transmises par THG, dans le chromosome de la bactérie receveuse. Les transposons ou IS (Insertion Sequence) comportent un gène codant une transposase flanqué par des répétitions inversées et, une fois produite, la transposase se lie au niveau de ces répétitions pour extraire le transposon de son emplacement génomique et le coller dans un nouvel emplacement (**Figure 22**).

Par ailleurs, il a été observé que plusieurs gènes codant des déterminants de la pathogénie sont souvent situés dans des régions correspondant à des cassettes d'insertion mobiles ou MIC, ce qui faciliterait la transposition de ces gènes et donc leur dispersion au sein du génome par le même mécanisme que celui utilisé pour la transposition répllicative d'un transposon de type Tn3 (Ferreira *et al.*, 2015). Ainsi, ces MIC pourraient être impliqués dans l'adaptation à l'hôte mais aussi dans l'émergence de nouveaux agents pathogènes. D'autre part, l'insertion d'un transposon dans le génome d'une souche pathogène peut aussi impacter sa gamme d'hôte. Par exemple, chez le pathotype B de *X. citri* pv. *aurantifolli*, la présence d'un transposon dans un gène codant un ET3 lui permet de produire des symptômes sur pamplemousse, cet ET3 étant impliqué dans le déclenchement d'une HR sur cette plante (Gochez *et al.*, 2015). Ainsi, ces éléments mobiles participent aussi à l'évolution adaptative des bactéries phytopathogènes, leur présence dans un génome reflétant l'acquisition de séquences d'ADN étrangères et permettant la diversification du génome receveur (Sundin, 2007). Par exemple, il a été démontré que les souches de *X. arboricola* pathogènes possèdent une gamme plus large d'éléments génétiques mobiles que les souches non pathogènes. Ainsi, la pathogénie de ces souches pourrait avoir été acquise, du moins en partie, grâce à ces éléments mobiles qui leur ont probablement apporté des gènes de virulence (Cesbron *et al.*, 2015). Cet exemple conforte donc l'idée que les éléments génétiques mobiles sont impliqués dans la spécificité d'hôte.

Tableau 3 : Nombre de gènes codant des effecteurs TAL par souche dans le genre *Xanthomonas*. Ce tableau est adapté de Scholze and Boch 2011 et de Hutin, 2015. En dehors des pathovars (pvs.) de *X. campestris* et des espèces *X. gardneri* et *X. translucens*, la nomenclature adoptée est celle détaillée par Constantin et al. 2016. Les gènes codant des effecteurs TAL sont nommés *tal*.

Espèce ou pathovar	Hôte	Nombre de <i>tal</i>
<i>X. campestris</i> pv. <i>musacearum</i>	Bananier	0
<i>X. axonopodis</i> pv. <i>vasculorum</i>	Canne à sucre	0
<i>X. gardneri</i>	Poivron et tomate	0-1
<i>X. euvesicatoria</i> pv. <i>euvesicatoria</i>	Poivron et tomate	0-2
<i>X. campestris</i> pv. <i>armoraciae</i>	Brassicacées	0-3
<i>X. campestris</i> pv. <i>campestris</i>	Brassicacées	0-5
<i>X. translucens</i> pvs.	Céréales	0-8
<i>X. phaseoli</i> pv. <i>phaseoli</i>	Haricot	1-2
<i>X. citri</i> pv. <i>fuscans</i>	Haricot	1-2
<i>X. citri</i> pv. <i>citri</i>	Agrumes	2-4
<i>X. citri</i> pv. <i>aurantifolii</i>	Agrumes	2-4
<i>X. phaseoli</i> pv. <i>manihotis</i>	Manioc	2-6
<i>X. citri</i> pv. <i>malvacearum</i>	Coton	6-10
<i>X. oryzae</i> pv. <i>oryzae</i>	Riz	9-16
<i>X. oryzae</i> pv. <i>oryzicola</i>	Riz	12-28

3. Les effecteurs TAL

Comme décrit précédemment, la virulence des bactéries à Gram-négatif, repose en partie sur leur SST3 et leur répertoire d'ET3, l'ensemble représentant l'un des moteurs de l'interaction hôte-pathogène. Chez les bactéries du genre *Xanthomonas*, on trouve des ET3, les effecteurs TAL (Transcription Activator-Like), qui ont la particularité d'activer l'expression de gènes cibles de la plante hôte en se liant directement au promoteur de ces gènes (Mak *et al.*, 2013). Ce type d'ET3 a été découvert chez *X. euvesicatoria* pv. *euvesicatoria* (Bonas *et al.*, 1989) et est très présent au sein du genre *Xanthomonas*. Un nombre très variable de gènes codant des effecteurs TAL (*tal*) a été identifié chez différentes souches de *Xanthomonas* (**Tableau 3**). Par exemple, 28 *tal* ont été dénombrés chez la souche BLS256 de *X. oryzae* pv. *oryzicola* (Booher *et al.*, 2015), seulement huit *tal* sont présents chez la souche XT4699 de *X. translucens* pv. *undulosa* (Peng *et al.*, 2016) et aucun *tal* n'a été identifié chez différentes souches de *X. campestris* pv. *musacearum* (Wasukira *et al.*, 2012). Les effecteurs TAL permettent à la bactérie de détourner la machinerie transcriptionnelle de la plante hôte, en ciblant et en activant l'expression de gènes de l'hôte (Boch & Bonas, 2010). Pour ce faire, les effecteurs TAL possèdent **(i) une structure protéique particulière** totalement adaptée à leur fonction mimant les facteurs de transcription eucaryotes. Leur spécificité d'action met en jeu **(ii) un code responsable de l'interaction spécifique entre l'effecteur TAL et l'ADN de la plante**. Dans une relation compatible, les effecteurs TAL induisent l'expression de gènes de sensibilité (*S*) qui favorisent l'infection bactérienne (Cox *et al.*, 2017) et jouent donc **(iii) le rôle d'effecteur de virulence**. Cependant, dans une relation incompatible, ils peuvent être impliqués dans **(iv) la réaction de défense de la plante**, par exemple, en activant l'expression de gènes de résistance (*R*) alors appelés gènes exécuteurs qui favorisent la résistance de la plante (Römer *et al.*, 2007). La capacité des effecteurs TAL à se lier à une séquence d'ADN spécifique a été très étudiée et a permis le développement **(v) de divers outils biotechnologiques et de lutte**.

3.1. Structure protéique des effecteurs TAL

Les effecteurs TAL ont principalement été décrits chez les bactéries du genre *Xanthomonas*, chez qui la structure de ces effecteurs est extrêmement conservée avec au minimum 78 % d'identité de séquence. Des homologues des effecteurs TAL ont aussi été identifiés chez des bactéries des genres *Ralstonia* (RipTAL) et *Burkholderia* (Bats) (Heuer *et al.*, 2007; de Lange *et al.*, 2014a) et plus récemment dans une analyse du métagénome marin (MOrTL) (de Lange *et al.*, 2015). Les effecteurs TAL de *Xanthomonas* ont environ 60 % d'identité avec les RipTAL de *Ralstonia*, 45 % avec les Bat de *Burkholderia* et 35 % avec les MOrTL marins récemment identifiés (de Lange *et al.*, 2015). Cette structure conservée se décompose en quatre domaines fonctionnels permettant à ces effecteurs d'accomplir les quatre étapes constituant leur mode

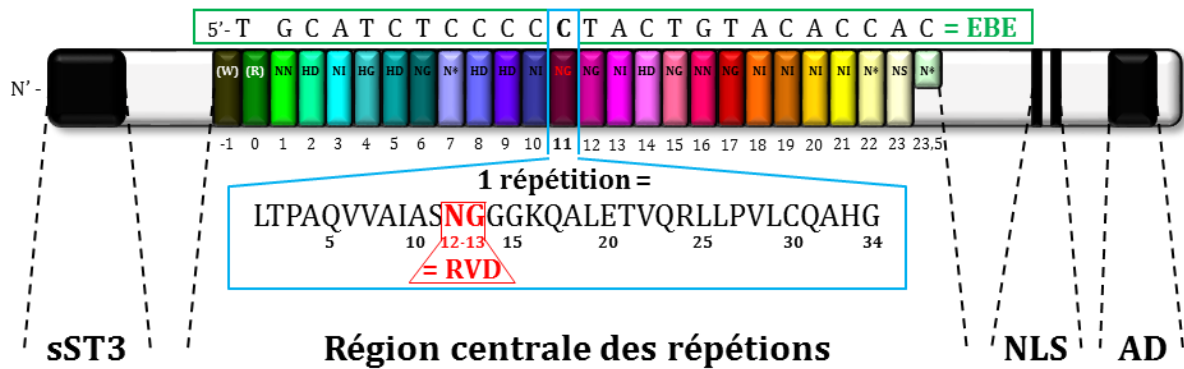


Figure 23 : Organisation des domaines fonctionnels d'un effecteur TAL. Cette figure est adaptée de Mak et al. 2012. Les effecteurs TAL possèdent quatre domaines fonctionnels : un signal de sécrétion de type 3 (sST3) ; une région centrale de répétitions de 34 acides aminés ; un domaine NLS contenant deux ou trois signaux de localisation nucléaire ; et un domaine AD (acidic Activation Domain). La région centrale se compose d'un nombre variable de répétitions quasi-identiques entre elles, qui sont représentées par les différentes cases de couleur. Encadrée en bleu, la séquence typique d'une répétition dont le RVD (Repeat Variable Di-residues) est indiqué en rouge aux positions 12 et 13. Un RVD a la capacité de se lier spécifiquement à un nucléotide, ici le RVD NG lie une cytosine, et par conséquent la séquence de RVD d'un effecteur TAL (indiquée dans les cases de différentes couleurs) définit la séquence d'ADN de l'hôte ciblée ou EBE (Effector Binding Element) encadrée en vert. Les répétitions -1 et 0 sont différentes des autres répétitions et se lient à une thymine. La dernière répétition d'un effecteur TAL est considérée comme une demi-répétition car elle ne contient généralement que 20 acides aminés. La séquence protéique correspond à celle de PthXo1, un effecteur TAL de *X. oryzae* pv. *oryzae* contenant 23,5 répétitions et ciblant le gène *Os8N3* du riz codant pour un transporteur de sucre.

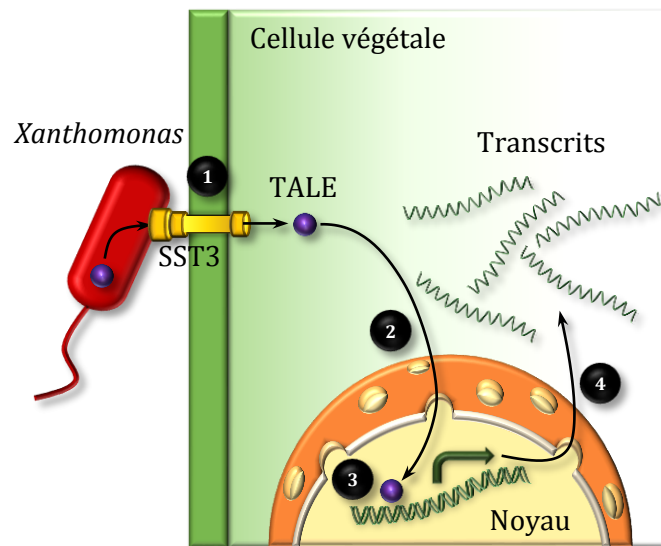


Figure 24 : Mode d'action des effecteurs TAL. Cette figure est adaptée de Boch and Bonas 2010. Une bactérie du genre *Xanthomonas* est représentée en rouge. Elle possède un SST3, en jaune, qui lui permet de sécréter des ET3, dont notamment des effecteurs TAL (TALE), dans la cellule végétale hôte, en vert. (1) L'effecteur TAL, illustré par une sphère violette, est injecté dans la cellule hôte grâce à son signal de sécrétion de type 3, (2) puis il migre jusqu'au noyau, en orange, à l'aide de ses signaux de localisation nucléaire, (3) afin de se lier spécifiquement à l'EBE, généralement située dans la séquence promotrice d'un gène cible dans le génome de l'hôte, (4) pour activer l'expression de ce gène cible. Le gène ciblé peut être de sensibilité ou de résistance, ce qui favorise la maladie ou déclenche une réaction de défense, respectivement.

d'action : (i) la sécrétion de l'effecteur dans la cellule hôte, (ii) la migration jusqu'au noyau de la plante, (iii) la fixation spécifique au promoteur d'un gène cible et (iv) l'activation de la transcription du gène ciblé (**Figures 23 et 24**).

(i) Ainsi, tout commence par l'injection des effecteurs TAL dans la plante hôte via le SST3 de la bactérie. Pour y parvenir, l'effecteur TAL possède un signal de sécrétion de type 3 (sST3) en position N-terminale. Ce type de signal est fréquemment retrouvé dans de nombreux autres ET3, également en position N-terminale (Sory *et al.*, 1995; Mudgett *et al.*, 2000; Lloyd *et al.*, 2001). De plus, il est indispensable à la translocation de l'ET3 dans la cellule hôte. En effet, sa délétion empêche la sécrétion de l'effecteur TAL AvrBs3 par le SST3 de *X. euvesicatoria* *in vitro* (Szurek *et al.*, 2002).

(ii) Une fois dans la cellule hôte, l'effecteur TAL migre et pénètre dans le noyau. Pour ce faire, il possède deux ou trois signaux de localisation nucléaire ou NLS (Nuclear Localisation Signal) en position C-terminale. Par exemple, trois NLS ont été identifiés dans les séquences de AvrB6 et PthA, des effecteurs TAL de *X. euvesicatoria* et *X. citri* respectivement (Yang & Gabriel, 1995a). Ces NLS permettent l'adressage des effecteurs TAL vers le noyau en se liant à l'importine α , qui est capable de transporter les protéines porteuses d'au moins un NLS depuis le cytoplasme vers l'intérieur du noyau à travers les pores nucléaires. Ces NLS sont bien entendu indispensables à la fonction des effecteurs TAL (Szurek *et al.*, 2001).

(iii) Après la pénétration de l'effecteur TAL dans le noyau, il se fixe à l'ADN de l'hôte au niveau du promoteur d'un gène spécifiquement ciblé. Cette spécificité d'accroche à l'ADN de l'hôte est conférée par la région centrale de l'effecteur TAL, composée d'un nombre variable de répétitions quasi-identiques entre elles dont la longueur est en général de 33 ou 34 acides aminés (**Figure 25**). L'identité des deux acides-aminés situés en positions 12 et 13 de chaque répétition varie plus particulièrement que celle des autres résidus, c'est pourquoi ils ont été nommés RVD pour Repeat Variable Di-residues. Ce sont les RVD qui permettent à l'effecteur TAL de se lier spécifiquement à sa séquence cible. En effet, chaque RVD a la capacité de se lier à un nucléotide donné (Doyle *et al.*, 2013a). Ainsi, l'ensemble des RVD d'un effecteur TAL forme une séquence dont la combinaison est spécifique de la séquence d'ADN ciblée. La région centrale des effecteurs TAL forme donc un domaine de liaison à l'ADN tout à fait unique en son genre dans le vivant, qui possède une structure tridimensionnelle parfaitement adaptée à cette liaison. En effet, chaque répétition forme une structure en épingle à cheveux projetant le RVD sur l'ADN hôte, et la succession de répétitions aboutit à une structure tridimensionnelle globale formant une super hélice enroulée autour de l'ADN de la plante qui se fixe grâce aux liaisons RVD-nucléotide (**Figure 26**) (Mak *et al.*, 2012; Doyle *et al.*, 2013b).

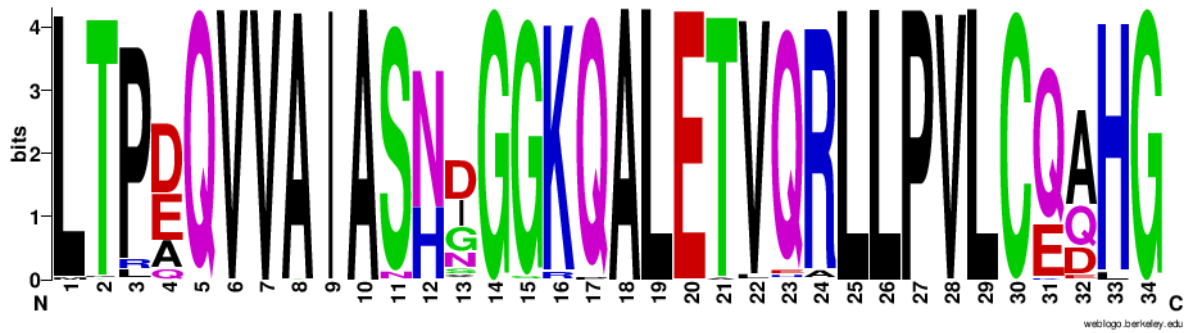


Figure 25 : Composition des répétitions d'effecteurs TAL appartenant à des souches du genre *Xanthomonas*. Cet alignement de séquences a été réalisé avec WebLogo (Schneider & Stephens, 1990; Crooks *et al.*, 2004). Alignement représentant la séquence consensus et le polymorphisme en acides aminés de 1824 répétitions issues de 113 effecteurs TAL provenant de diverses souches du genre *Xanthomonas*.

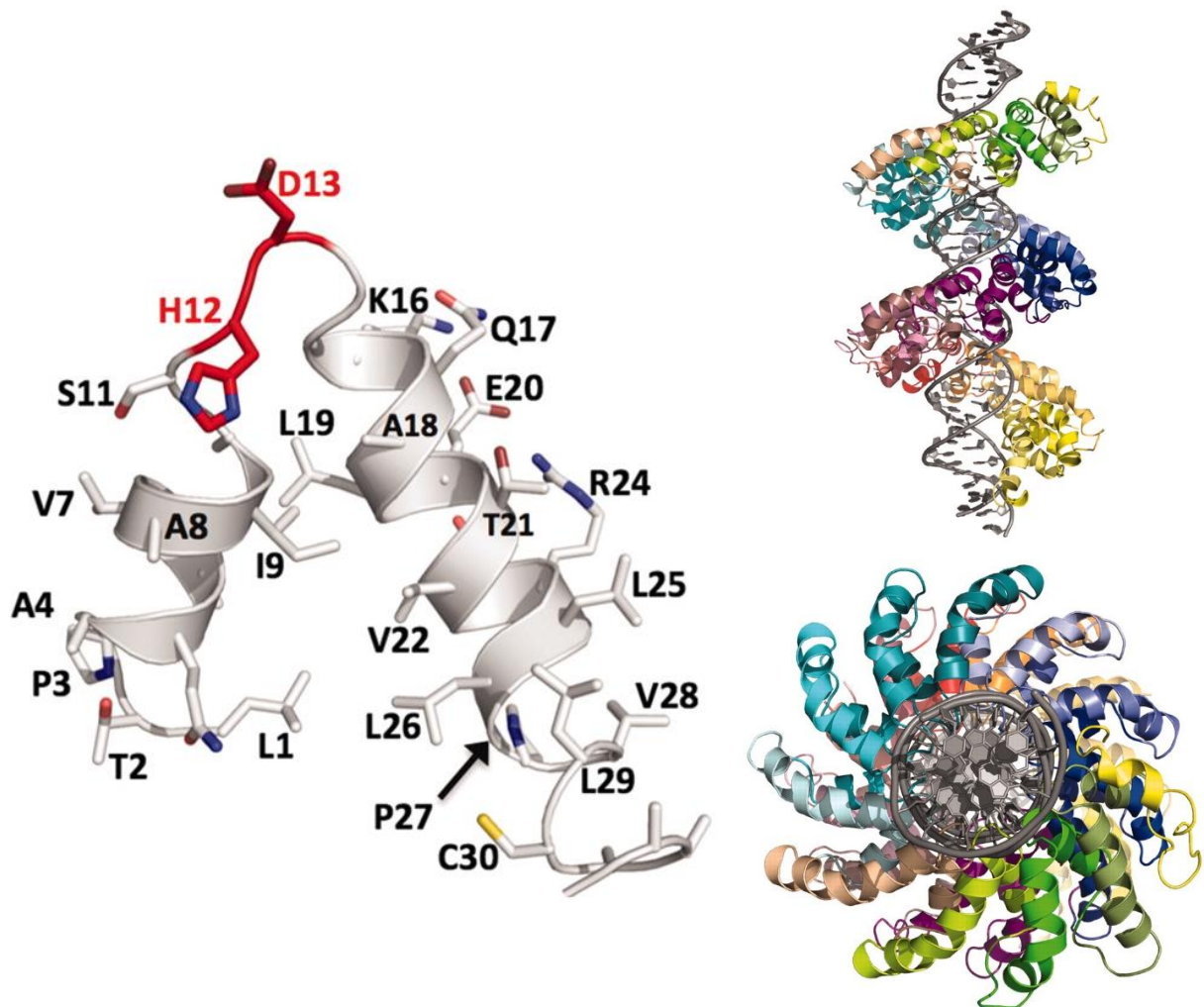


Figure 26 : Structure tridimensionnelle d'un effecteur TAL lié à sa cible et d'une répétition isolée. Cette figure est adaptée de Mak et al. 2012. A gauche, structure en épingle à cheveux d'une répétition provenant de l'effecteur TAL PthXo1 de *X. oryzae* pv. *oryzae*, avec le RVD (HD) en rouge. A droite, structures tridimensionnelles de PthXo1 lié à sa séquence d'ADN cible (en gris) sous deux angles. Chaque répétition possède une couleur différente pour les différencier.

(iv) La séquence cible d'un effecteurs TAL se situe en général au niveau du promoteur d'un gène. Ainsi, l'effecteur TAL peut jouer son rôle de facteur de transcription eucaryote. Pour ce faire, il possède en position C-terminale, après les NLS, un domaine AD (acidic Activation Domain), capable de recruter l'ARN polymérase de la plante pour activer l'expression du gène hôte ciblé (Marois *et al.*, 2002). Bien évidemment, ce domaine fonctionnel est essentiel à la fonction de l'effecteur TAL et permet donc à la bactérie de détourner la machinerie transcriptionnelle de l'hôte en sa faveur (Szurek *et al.*, 2001).

3.2. Le code responsable de l'interaction effecteur TAL – ADN cible

Les RVD sont responsables de la spécificité d'interaction de l'effecteur TAL avec l'ADN de l'hôte et chaque RVD possède sa propre spécificité de liaison aux bases azotées de l'ADN. Ainsi, le nombre de RVD définit la longueur de la séquence d'ADN ciblée nommée EBE pour Effector Binding Element, et la combinaison de RVD détermine la spécificité d'accroche à cette EBE spécifiquement ciblée (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). Par conséquent, connaître la séquence de RVD d'un effecteur TAL permet de prédire ses cibles potentielles dans le génome de l'hôte grâce à l'utilisation d'outils bio-informatiques dédiés (Doyle *et al.*, 2012; Grau *et al.*, 2013; Pérez-Quintero *et al.*, 2013).

Aujourd'hui, plusieurs outils de prédictions des cibles d'effecteur TAL existent, mais avant d'élaborer ces programmes bio-informatiques, il a fallu décrypter le code responsable de l'interaction effecteur TAL – ADN régit par les RVD et ce ne fut pas une mince affaire. En effet, il s'est écoulé 20 ans entre la découverte du premier effecteur TAL en 1989 et le décryptage de ce code en 2009. Ce code a été découvert simultanément par deux équipes à travers le monde. La première équipe a décrypté le code via l'utilisation d'un plasmide contenant un gène rapporteur *GUS* dans le cadre d'une transformation transitoire de *Nicotiana benthamiana* par *Agrobacterium tumefaciens* (Boch *et al.*, 2009). La seconde équipe a quant à elle déterminé la spécificité de liaison des RVD grâce à une matrice générée en fonction de la fréquence d'interaction entre un RVD et un nucléotide (Moscou & Bogdanove, 2009). Les deux équipes ont présenté des résultats de spécificité de liaison des RVD très similaires et ont démontré que chaque RVD possède une spécificité d'interaction à l'ADN qui lui est propre. Par exemple, HD reconnaît la cytosine, NG la thymine, NI l'adénine tandis que NN peut lier à la fois la guanine et l'adénine (Boch & Bonas, 2010) (**Figure 27**). Ainsi, chaque RVD possède une affinité de liaison pour une ou plusieurs bases azotées. Ces résultats ont été confortés par la validation fonctionnelle de cibles de plusieurs effecteurs TAL, ainsi que par la description de la structure cristalline d'un effecteur TAL lié à son brin d'ADN. En effet, cette structure cristalline, permettant de visualiser l'accroche des RVD aux bases de l'ADN, a permis de découvrir que c'est le résidu en position 13 qui se lie à l'ADN et qui

RVD		Freq.	Spé.
aa12	aa13		
N	I	391	A T GA ACGT
	G	334	
	N	312	
	S	163	
	—	138	
	D	12	
	K	6	
	C	3	
	V	2	
	A	1	
	H	1	
H	D	527	C
	G	106	
	A	6	
	—	4	
	H	4	
	I	4	
	N	2	
S	—	2	
	N	1	
	S	1	
I	G	2	T
Y	G	1	

Figure 27 : Spécificité de liaison de RVD retrouvés dans des effecteurs TAL naturels. Ce tableau est adapté de Boch and Bonas 2010. Les RVD (Repeat Variable Di-residues) et leur fréquence (Freq.) ont été déterminés à partir de l'ensemble des répétitions de 113 effecteurs TAL provenant de souches du genre *Xanthomonas*. Un tiret jaune indique l'absence de l'acide aminé en position 13 (aa13). Les spécificités de liaison (Spé.) indiquées sont celles qui ont été observées dans la nature. aa12 : acide aminé en position 12.

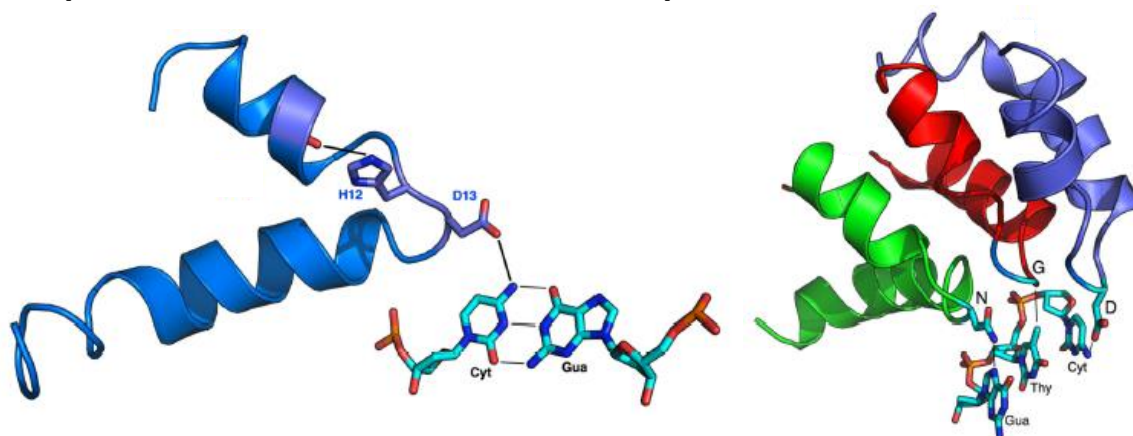


Figure 28 : Structures tridimensionnelles de répétitions d'un effecteur TAL dont les RVD sont liés à aux bases de l'ADN. Cette figure est adaptée de Doyle et al. 2013. A gauche, la structure montre l'accroche du RVD HD à une cytosine, elle-même reliée à une guanine. L'histidine en position 12 (H12) forme une liaison hydrogène avec le résidu 8 de la première hélice alpha ce qui permet de stabiliser la structure de la répétition alors que l'aspartate en position 13 (D13) est lié à la cytosine avec une liaison hydrogène. A droite, la structure représente des répétitions successives d'un effecteur TAL en interaction avec l'ADN. Les RVD HD, NG et NN (dans l'ordre en bleu, rouge et vert) sont respectivement liés à une cytosine, une thymine et une guanine consécutives dans la séquence d'ADN ciblée, la deuxième interaction étant une liaison de Van der Waals, tandis que les deux autres sont des liaisons hydrogène.

régit la spécificité de liaison alors que le résidu en position 12 assure sa stabilisation (**Figure 28**) (Mak *et al.*, 2012). Egalement, des différences de forces d'interaction entre RVD et nucléotides ont été démontrées. En effet, les RVD HD et NN ont une forte affinité de liaison pour leur cible, car l'interaction implique des liaisons hydrogène, alors que les RVD NI et NG ont une faible affinité de liaison pour leur cible, car l'interaction implique des force de Van der Walls (Deng *et al.*, 2012; Mak *et al.*, 2012). Par ailleurs, il a été démontré que le nombre de répétitions minimal nécessaire à l'accroche de l'effecteur TAL à l'ADN et donc à sa capacité d'induire l'expression du gène ciblé est de six (Boch *et al.*, 2009).

Ainsi, le décryptage du code responsable de l'interaction effecteur TAL – ADN cible a permis de générer des programmes bio-informatiques de prédiction des potentielles cibles des effecteurs TAL dans le génome des plantes hôtes. A ce jour, il existe quatre outils de prédictions des cibles d'effecteurs TAL : TargetFinder, TALgetter, TALVEZ et Storyteller (Doyle *et al.*, 2012; Grau *et al.*, 2013; Pérez-Quintero *et al.*, 2013). Chacun de ces outils vise à prédire des cibles d'effecteurs TAL, mais pour ce faire ils n'utilisent pas les mêmes matrices informatiques et ne produisent donc pas les mêmes résultats. C'est pourquoi, la prédiction *in silico* des cibles potentielles d'un effecteur TAL ne suffit pas pour déterminer cette cible et il est nécessaire de compléter ces prédictions avec des analyses transcriptomiques et des validations fonctionnelles de ces cibles via la création de mutant. De plus, l'efficacité de prédiction des EBE est limitée par plusieurs points :

Le premier point est que, même si la spécificité de liaison des RVD les plus fréquemment retrouvés dans la nature est bien décrite, il existe théoriquement 400 combinaisons différentes de deux acides aminés. Ces 400 RVD ont été testées pour déterminer leur affinité de liaison aux nucléotides et la majorité n'avait pas ou peu d'affinité (Yang *et al.*, 2014). Face à ce fait, Target Finder permet de prendre en compte différentes matrices de notation dont celle décrite par Yang et al., (2014), TALgetter donne la possibilité de ne considérer que la position 13 du RVD pour déterminer sa spécificité de liaison et TALVEZ propose pour des RVD dont la spécificité est inconnue, d'attribuer une spécificité de liaison équivalente aux quatre bases azotées afin de parer à toutes les éventualités.

Le deuxième point concerne la présence, en amont des répétitions classiques de l'effecteur TAL, de répétitions conservées en positions 0 et -1. Ces répétitions, qui sont différentes des répétitions classiques, présentent une préférence de liaison à la thymine (Schreiber & Bonas, 2014). La séquence ciblée par les effecteurs TAL commence donc quasi-systématiquement par une thymine. C'est pourquoi lors de la prédiction des cibles des effecteurs TAL un des postulats est qu'en position 0, immédiatement en amont de l'EBE, il y a une thymine nommée T0. Cependant, il a été démontré que TalC de *X. oryzae* pv. *oryzae* cible une EBE qui n'est pas précédée

d'une thymine mais d'une cytosine, nommée alors C0 (Yu *et al.*, 2011). Cette possibilité a été prise en compte par Target Finder qui permet la prédiction de cibles précédées de T0 et/ou C0.

Le troisième point est en rapport avec le nombre de répétition et la position des RVD dans la séquence de RVD. En effet, nous savons qu'il existe un nombre minimal de six répétitions nécessaires à la liaison de l'effecteur TAL à l'ADN, mais plus récemment, il a été démontré que le nombre de RVD n'impacte pas l'affinité de liaison de l'effecteur TAL de façon linéaire. En effet, l'augmentation progressive du nombre de RVD engendre une augmentation puis une diminution de la spécificité de liaison avec un pic entre 15 et 19 répétitions (Rinaldi *et al.*, 2017). De plus, la position du RVD dans la séquence de RVD est importante. En effet, les RVD situés à la fin de la séquence ont moins d'impact sur la spécificité de liaison finale de l'effecteur TAL. Par exemple, la mutation des trois dernières bases du site de liaison d'AvrBs3 n'a aucun effet sur l'activité de l'effecteur TAL (Kay *et al.*, 2009). Ce fait a notamment été pris en compte par TALVEZ, où l'on peut choisir de réduire le poids dans la matrice des RVD situées à la fin de la séquence de l'effecteur TAL et il a été démontré qu'utiliser cette option pour des RVD situées à partir de la position 15 permettait de nettement améliorer les prédictions des cibles des effecteurs TAL (Pérez-Quintero *et al.*, 2013).

Le quatrième point concerne une découverte récente : la capacité des effecteurs TAL à activer la transcription de gènes cibles, en se fixant sur le brin sens ou sur le brin anti-sens, et ceci quel que soit son orientation par rapport au gène. En effet, des effecteurs TAL ont été capables d'activer l'expression du gène cible *OsSWEET14* lorsque les EBE étaient situées sur chaque brin du promoteur, qu'elles soient dans le même sens ou en sens inverse du gène (Wang *et al.*, 2017). Cette capacité des effecteurs TAL nouvellement découverte a été prise en compte par les trois outils de prédictions qui permettent d'effectuer des prédictions en sens et en anti-sens.

Pour finir, il a été signalé l'existence de répétitions aberrantes, comme par exemple des répétitions plus longues que la normale. C'est le cas chez AvrXa7 et PthXo3, des effecteurs TAL de *X. oryzae* pv. *oryzae*, qui comportent tous les deux une répétition longue de 39 acides aminés. A ce jour, nous ne connaissons pas encore exactement l'impact de ce type de répétitions sur la liaison des RVD et par conséquent cet élément ne peut pas encore être pris en compte pour la prédiction des cibles des effecteurs TAL. A priori, ces répétitions plus longues que la normale, cassent le mode de liaison strict des effecteurs TAL, basé sur une série de RVD se liant à une séquence nucléotidiques, mais n'empêchent pas pour autant la liaison de l'effecteur TAL à l'ADN. Une des hypothèses émises permettant d'expliquer cela est que ce type de répétition pourrait conférer une certaine flexibilité de liaison aux effecteurs TAL qui en possèdent (Richter *et al.*, 2014).

Tableau 4 : Gènes de virulence ciblés par des effecteurs TAL de souches de *Xanthomonas*.

Ce tableau est adapté de Hutin, 2015 et de Pérez-Quintero, 2017. TALE : effecteur TAL ; *Xau* : *X. citri* pv. *aurantifolii* ; *Xci* : *Xanthomonas citri* pv. *citri* ; *Xeu* : *Xanthomonas euvesicatoria* ; *Xg* : *Xanthomonas gardneri* ; *Xoc* : *Xanthomonas oryzae* pv. *oryzicola* ; *Xoo* : *Xanthomonas oryzae* pv. *oryzae* ; *Xpm* : *X. phaseoli* pv. *manihotis*. Un astérisque indique l'ensemble des effecteurs TAL de type PthA provenant d'une diversité de *X. citri* pv. *citri*.

Fonction ciblée	Cible	TALE	Espèce	Hôtes	Références
Facteur de transcription	<i>UPA20 (bHLH)</i>	AvrBs3	<i>Xeu</i>	Poivron et Tomate	Kay et al. 2007
	<i>bHLH</i>	AvrHah1	<i>Xg</i>	Tomate	Schwartz et al. 2017
	<i>CsLOB1</i>	PthA*	<i>Xci</i>	Agrumes	Al-Saadi et al. 2007; Hu et al. 2014
		PthB	<i>Xau</i>	Agrumes	
		PthC	<i>Xau</i>	Agrumes	
	<i>CsLOB2</i>	PthA1*	<i>Xci</i>	Agrumes	Abe and Benedetti 2016
	<i>TFX1</i>	PthXo6	<i>Xoo</i>	Riz	Sugio et al. 2007
	<i>TFIIAγI</i>	PthXo7	<i>Xoo</i>	Riz	
Transporteur de sucre	<i>CsSWEET1</i>	PthA*	<i>Xci</i>	Agrumes	Hu et al. 2014
	<i>OsSWEET11</i>	PthXo1	<i>Xoo</i>	Riz	Yang et al. 2006
	<i>OsSWEET13</i>	PthXo2	<i>Xoo</i>	Riz	Zhou et al. 2015
	<i>OsSWEET14</i>	PthXo3	<i>Xoo</i>	Riz	Yang and White 2004
		AvrXa7	<i>Xoo</i>	Riz	Antony et al. 2010
		Tal5	<i>Xoo</i>	Riz	Streubel et al. 2013
		TalC	<i>Xoo</i>	Riz	Yu et al. 2011
	<i>UPA16</i>	AvrBs3	<i>Xeu</i>	Poivron et Tomate	Kay et al. 2009
	<i>MeSWEET10a</i>	TAL20	<i>Xpm</i>	Manioc	Cohn et al. 2014
Transporteur de sulfate	<i>OsSULTR3;6</i>	Tal2g	<i>Xoc</i>	Riz	Cernadas et al. 2014
Stabilités des miRNA	<i>OsHEN1</i>	PthXo8	<i>Xoo</i>	Riz	Moscou and Bogdanove 2009; Pérez-Quintero et al. 2013
		Tal1C	<i>Xoc</i>	Riz	
Dioxygénase	<i>CsDIOX</i>	PthA*	<i>Xci</i>	Agrumes	Abe and Benedetti 2016

3.3. La mission première des effecteurs TAL: jouer le rôle d'effecteur de virulence

Les effecteurs TAL jouent un rôle primordial dans la virulence des souches de *Xanthomonas*. Pour ce faire, ils ciblent des gènes *S* pour en activer l'expression. Ces gènes *S* correspondent à des gènes de plantes dont l'expression favorise et est parfois nécessaire au développement d'une maladie provoquée par un agent pathogène. Ces gènes peuvent être impliqués dans n'importe quelle étape du cycle infectieux de l'agent pathogène : son attraction, son installation, sa multiplication et sa dispersion (Lapin & Van den Ackerveken, 2013). Le décryptage du code responsable de l'interaction effecteur TAL – ADN, et le développement d'outils de prédiction correspondant, a largement facilité l'identification de gènes ciblés par les effecteurs TAL, dont l'induction facilite la multiplication bactérienne, le développement des symptômes ou encore la dissémination de la bactérie (**Tableau 4**).

Les gènes *S* les plus décrits sont induits par les effecteurs TAL de souches de *X. oryzae* pv. *oryzae* pathogènes sur riz. Ces souches peuvent contenir de 9 à 16 effecteurs TAL dans leur génome, mais seul un ou deux de ces effecteurs TAL jouent un rôle majeur dans la pathogénie, leur mutation rendant les souches incapables de produire des symptômes sur riz (Yang & White, 2004). Pour ces effecteurs TAL les gènes *S* ciblés sont donc particulièrement importants pour le développement de la maladie. Notamment, les gènes *S* les plus caractérisés sont les gènes *SWEET* codant pour des transporteurs de sucre permettant l'export du saccharose dans l'apoplasme (Chen, 2014). En effet, trois gènes *OsSWEET* du riz sont ciblés par des effecteurs TAL : *OsSWEET11*, ciblé par PthXo1 de la souche PX099A ; *OsSWEET13*, ciblé par PthXo2 présent dans les souches JX01A et MAFF311018 ; et *OsSWEET14*, ciblé par AvrXa7, PthXo3, TalC et Tal5 provenant respectivement des souches PX086, PX061, BAI3 et MAI1 (Yang & White, 2004; Chu *et al.*, 2006a; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013; Zhou *et al.*, 2015). De façon intéressante, des gènes homologues à ces gènes *OsSWEET* et appartenant à d'autres espèces végétales sont également ciblés par des effecteurs TAL provenant d'autres *Xanthomonas*. Par exemple, AvrBs3 de *X. euvesicatoria* induit l'expression de UPA16 chez le poivron ; TAL20 de *X. phaseoli* pv. *manihotis* cible *MeSWEET10a* chez le manioc ; et la série d'effecteurs TAL PthA de *X. citri* pv. *citri* active la transcription de *CsSWEET1* chez les agrumes (Kay *et al.*, 2009; Cohn *et al.*, 2014; Hu *et al.*, 2014). Ainsi, au regard de la diversité de souches et d'effecteurs TAL ciblant les gènes *SWEET*, il apparaît clairement que l'export de sucre issu de l'induction de ces gènes est capital pour le développement de la maladie provoquée par ces bactéries, notamment pour l'étape de multiplication, le sucre étant une source nutritive pour les bactéries.

Un second type de cibles d'effecteurs TAL grandement décrit correspond aux facteurs de transcription. Parmi les facteurs de transcription ciblés, il y a le type LOB qui est induit par des effecteurs TAL de *X. citri* pv. *citri* et *X. citri* pv. *aurantifolii*. En effet, *CsLOB1* présent chez différentes

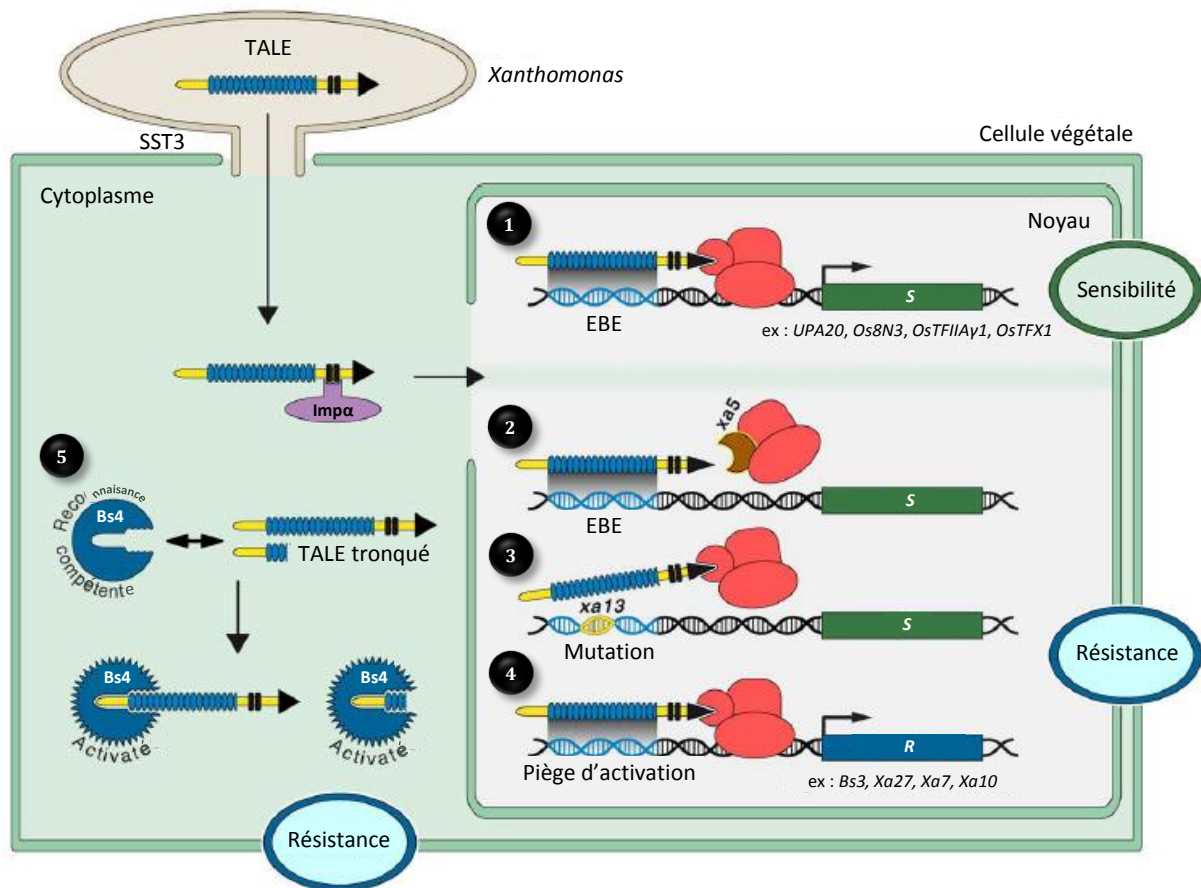


Figure 29 : Modes d'action des effecteurs TAL dans des contextes de réactions compatibles et incompatibles. Ce schéma est adapté de Bogdanove et al. 2010. Les effecteurs TAL (TALE) sont injectés dans la cellule hôte via le *SST3* puis ils sont adressés au noyau grâce à leurs signaux de localisation nucléaire. **(1)** Dans un contexte de relation compatible, où la plante est sensible, l'effeteur TAL se fixe spécifiquement à l'EBE (Effector Binding Element) d'un promoteur de gène *S* (sensibilité), afin d'activer son expression pour favoriser l'infection. Dans un contexte de relation incompatible, où la plante est résistante, **(2)** l'effeteur TAL peut être incapable de recruter l'ARN polymérase car celle-ci est mutée ou **(3)** être incapable de se fixer à son EBE car celle-ci peut elle aussi être mutée. **(4)** Il est également envisageable que les effecteurs TAL activent des gènes *R* (résistance) à la place des gènes *S*. **(5)** Les effecteurs TAL peuvent aussi être reconnus dans le cytoplasme de la cellule hôte par une protéine de résistance selon les voies classiques de l'ETI.

variétés d'agrumes est ciblé par 5 effecteurs TAL différents : PthA4, PthAw et PthA* de *X. citri* pv. *citri*, et PthB et PthC de *X. citri* pv. *aurantifolii* (Al-Saadi *et al.*, 2007; Hu *et al.*, 2014). Le rôle de *CsLOB1* dans le développement de la maladie n'est pas encore caractérisé, mais il est avéré que ce facteur de transcription est nécessaire à la formation des pustules caractéristiques du chancre bactérien des agrumes. Par ailleurs, il a été démontré que *CsLOB2* peut être induit par PthA2 et PthA4 de *X. citri* pv. *citri* et PthC1 de *X. citri* pv. *aurantifolii* (Pereira *et al.*, 2014), ce qui confirme l'importance de ce type de facteur de transcription dans le développement de la maladie. Les facteurs de transcription de type bHLH ont également été identifiés comme cibles d'effecteurs TAL. L'exemple le plus connu est celui d'UPA20, facteur de transcription du poivron, dont l'expression est induite par AvrBs3 de *X. euvesicatoria* (Kay *et al.*, 2007). Ce facteur de transcription cible quant à lui le gène *upa7*, dont l'expression entraîne une hypertrophie des cellules de mésophylle, ce qui contribue à la dispersion des bactéries puisque les tissus infectés se rompent ainsi plus facilement. Un autre facteur de transcription de type bHLH de la tomate a été identifié comme cible de l'effecteur TAL AvrHah1 de *X. gardneri* (Schornack *et al.*, 2008). Ce facteur de transcription induit quant à lui l'expression d'un gène codant une pectate lyase qui favorise l'absorption de l'eau par les cellules de tomate, ce qui permettrait à la bactérie d'augmenter les dommages aux tissus et faciliterait sa dissémination (Schwartz *et al.*, 2017). Pour finir, PthXo6 et PthXo7 de la souche PXO99 de *X. oryzae* pv. *oryzae* ciblent respectivement les gènes codant les facteurs de transcription OsTFXI et OsTFIIAγ1a (Sugio *et al.*, 2007). Cependant, le rôle de ces facteurs de transcription dans le développement de la maladie n'est pas encore connu.

D'autres catégories de gènes *S* ciblés par les effecteurs TAL existent. Par exemple, la mutation systématique de tous les gènes codant des effecteurs TAL chez la souche BLS256 de *X. oryzae* pv. *oryzicola* a montré que Tal2g induit l'expression de *OsSULTR3;6* du riz codant pour un transporteur de sulfate. Le rôle de ce transporteur dans le développement de la maladie n'est pas connu, mais il a été confirmé que l'induction de son expression contribue à l'expansion des lésions et à l'exsudation bactérienne (Cernadas *et al.*, 2014).

3.4. Leur implication dans la réaction de défense de la plante

Les effecteurs TAL sont certes des facteurs de virulence majeurs des *Xanthomonas*, mais il existe aussi chez les plantes plusieurs mécanismes de résistance ciblant les effecteurs TAL (**Figure 29**).

(i) La plante peut produire une protéine qui va reconnaître spécifiquement l'effecteur TAL pour le rendre non-fonctionnel. Une seule protéine de résistance ciblant un effecteur TAL est connue à ce jour. Cette protéine est codée par le gène *Bs4* chez la tomate qui lui permet d'être

résistante à *X. euvesicatoria*. Ce gène *R* code une protéine de type NB-LRR capable de reconnaître les domaines NLS ou AD des effecteurs TAL (Schornack *et al.*, 2004). Ainsi, la reconnaissance d'AvrBS4 par Bs4 serait d'avantage une reconnaissance structurale qu'une reconnaissance fonctionnelle (Schornack *et al.*, 2004, 2005).

(ii) Des mutations de l'ARN polymérase de la plante peuvent empêcher l'effecteur TAL de la recruter, empêchant la surexpression du gène ciblé. Par exemple, chez le riz, le gène *xa5* code un variant de substitution de la sous-unité γ du facteur de transcription TFIIA, qui est un composant de l'ARN polymérase II (Iyer & McCouch, 2004). Cette mutation permet aux variétés de riz portant l'allèle *xa5* d'être résistante aux souches *X. oryzae* pv. *oryzae* grâce à l'incapacité des effecteurs TAL à recruter l'ARN polymérase II, ce qui affecte la surexpression de leurs gènes cibles (Gu *et al.*, 2009).

(iii) Des mutations peuvent également avoir lieu dans l'EBE, ce qui peut empêcher l'effecteur TAL de s'y fixer et empêcher par conséquent la surexpression du gène ciblé. Des EBE mutées situées en amont de gènes *S* ont été identifiés chez le riz. En effet, deux gènes *SWEET* : *OsSWEET11* et *OsSWEET13* possèdent des allèles de résistance récessifs nommés *xa13* et *xa25* respectivement (Chu *et al.*, 2006b; Liu *et al.*, 2011; Zhou *et al.*, 2015). Les variétés de riz contenant l'allèle *xa13* sont résistantes à la souche PX099 de *X. oryzae* pv. *oryzae*. Cette résistance se caractérise par une absence de symptômes, qui est due à l'impossibilité pour PthXo1 d'induire l'expression de *OsSWEET11* (Yuan *et al.*, 2009, 2011). De façon intéressante, de nombreuses variétés de riz portent naturellement une version d'un allèle de résistance de *OsSWEET11* et tous ces allèles diffèrent de l'EBE originelle par une insertion, une délétion ou une substitution au niveau du second nucléotide de l'EBE, ce qui indique qu'un unique évènement de substitution est suffisant pour empêcher la liaison de l'effecteur TAL et rendre ces variétés résistantes (Römer *et al.*, 2010). De la même façon, les variétés de riz portant l'allèle *xa25* sont résistantes à la souche PX0339 de *X. oryzae* pv. *oryzae*, PthXo2 ne pouvant plus activer l'expression de *OsSWEET13* essentielle à la bactérie pour provoquer la maladie (Liu *et al.*, 2011; Zhou *et al.*, 2015).

(iv) Pour finir, une EBE peut se situer en amont d'un gène *R* appelé alors exécuter, dont l'expression par l'effecteur TAL déclenche les défenses de la plante. La résistance due à l'expression d'un gène exécuter par un effecteur TAL se nomme TAL^{AD}-ETI pour TAL Activation Domain – ETI (Bogdanove *et al.*, 2010; Zhang *et al.*, 2015b). De nombreux exemples de ce type ont été identifiés. Le plus connu d'entre eux est le gène *Bs3* du poivron ciblé par AvrBs3 de *X. euvesicatoria* et dont l'expression engendre une HR. Ce gène code une protéine qui est homologue des protéines de la famille FMO (Flavin-dépendant Mono-Oxygénase), connue pour jouer un rôle dans la biosynthèse de l'auxine et dans la réponse immunitaire de la plante (Römer *et al.*, 2007).

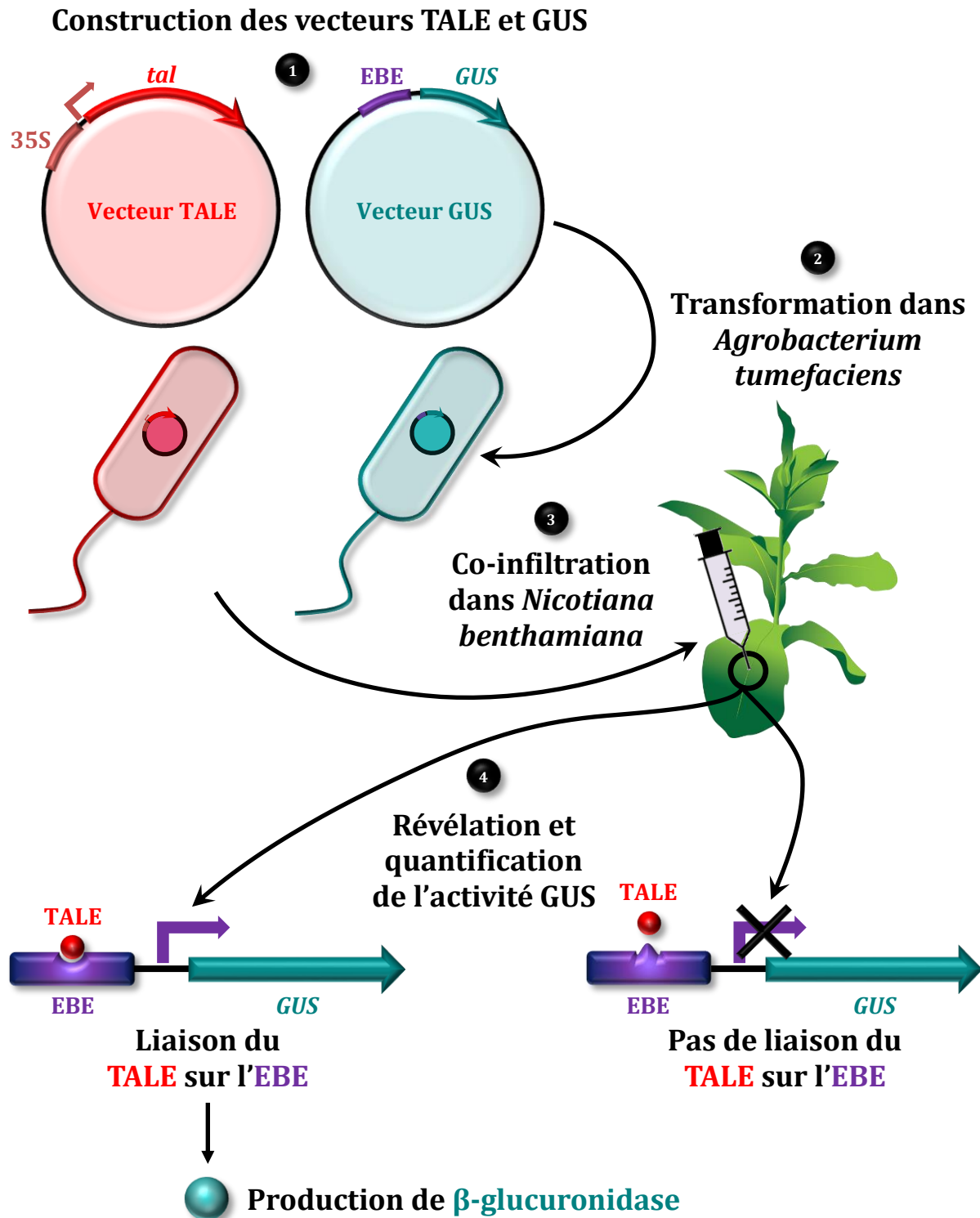


Figure 30 : Méthode pour étudier l'interaction effecteur TAL – ADN basée sur l'utilisation du gène rapporteur GUS. (1) Construction de deux vecteurs, l'un contenant un gène *tal* exprimant constitutivement l'effecteur TAL (TALE) à tester et l'autre possédant un gène rapporteur *GUS* précédé de l'EBE désirée. (2) Transformation de souches d'*Agrobacterium tumefaciens* avec les deux vecteurs produits. (3) Co-inoculation sur tabac d'un mélange des deux souches d'*Agrobacterium* transformées. (4) Prélèvement des zones infiltrées pour révéler et quantifier l'activité *GUS* engendrée par la liaison de l'effecteur TAL sur l'EBE. Si l'effecteur TAL lie l'EBE, la transcription du gène *GUS* est activée et de la β -glucuronidase est produite, alors que si l'effecteur TAL ne parvient à s'accrocher sur l'EBE, aucune β -glucuronidase n'est produite. L'activité *GUS* se quantifie en incubant l'échantillon prélevé avec le substrat de la β -glucuronidase, tel que le X-gluc. Si l'enzyme est présente, le substrat sera alors clivé, ce qui produira un précipité bleu.

Un autre exemple est celui de *Xa27*, un gène du riz codant une protéine sécrétée dans l'apoplaste et qui est toxique pour la cellule (Wu *et al.*, 2008). Ainsi, lorsque l'effecteur TAL AvrXa27 sécrété par PX099 de *X. oryzae* pv. *oryzae* induit l'expression de *Xa27*, cela se traduit par une HR qui bloque totalement le développement de la maladie (Gu *et al.*, 2005). Il y a également le gène *Bs4C* du poivron exprimé en présence d'AvrBs4 de *X. euvesicatoria* ce qui provoque également une HR (Strauß *et al.*, 2012). Par ailleurs, le gène *Xa23* a été identifié chez une espèce de riz sauvage résistante aux souches de *X. oryzae* pv. *oryzae* portant AvrXa23. Le gène *Xa23* code une protéine dont la fonction est inconnue (Wang *et al.*, 2015). Cependant, il a été établi que son induction par AvrXa23 entraîne une forte HR chez le riz empêchant le développement de la maladie. Pour finir, le gène *Xa10* du riz est spécifiquement induit par AvrXa10 de la souche PX086 de *X. oryzae* pv. *oryzae* (Hopkins *et al.*, 1992). Ce gène *R* code une protéine comportant quatre hélices transmembranaires et dont on ne connaît pas la fonction. Le gène *Xa10* confère une résistance aux variétés de riz le portant contre les souches possédant AvrXa10 en provoquant une HR (Gu *et al.*, 2008).

3.5. Les outils biotechnologiques et de lutte

Le décryptage du code responsable de l'interaction effecteur TAL – ADN a en partie nécessité l'élaboration d'une méthode utilisant un plasmide contenant un gène rapporteur *GUS* dans le cadre d'une transformation transitoire de *Nicotiana benthamiana* par *Agrobacterium tumefaciens* (**Figure 30**) (Boch *et al.*, 2009). Cette méthode demande la construction de deux vecteurs, le premier contenant le gène rapporteur *GUS* précédé d'une EBE choisie et le deuxième permettant l'expression de l'effecteur TAL à tester. L'inoculation d'un mélange de deux souches d'*Agrobacterium* contenant respectivement les deux vecteurs sur tabac permet de quantifier l'activité GUS engendrée par l'accroche de l'effecteur TAL sur son EBE. Cette méthode permet donc de déterminer la spécificité de liaison d'un RVD particulier, mais permet aussi de valider une EBE prédite comme cible d'un effecteur TAL. La construction du vecteur contenant le gène *tal* peut se faire de deux manières. La première est de cloner un *tal* naturel, cette possibilité permet de valider les EBE prédites de l'effecteur TAL correspondant. La deuxième possibilité consiste à construire un *tal* artificiel (*ArtTAL*) par l'intermédiaire de kits basés sur la stratégie Golden Gate, qui permettent l'expression d'un effecteur ArtTAL possédant n'importe quelle combinaison de RVD (Geißler *et al.*, 2011). Par conséquent, les ArtTAL permettent de cibler n'importe quel séquence ou promoteur d'un génome donné pour activer l'expression du gène associé (Bogdanove & Voytas, 2011). C'est pourquoi aujourd'hui, ils sont couramment utilisés, en particulier pour identifier et confirmer les gènes ciblés par des effecteurs TAL naturels (Li *et al.*, 2013b; Bogdanove, 2014). De plus, des ArtTAL modifiés ont été générés pour être fonctionnels dans divers génomes eucaryotes

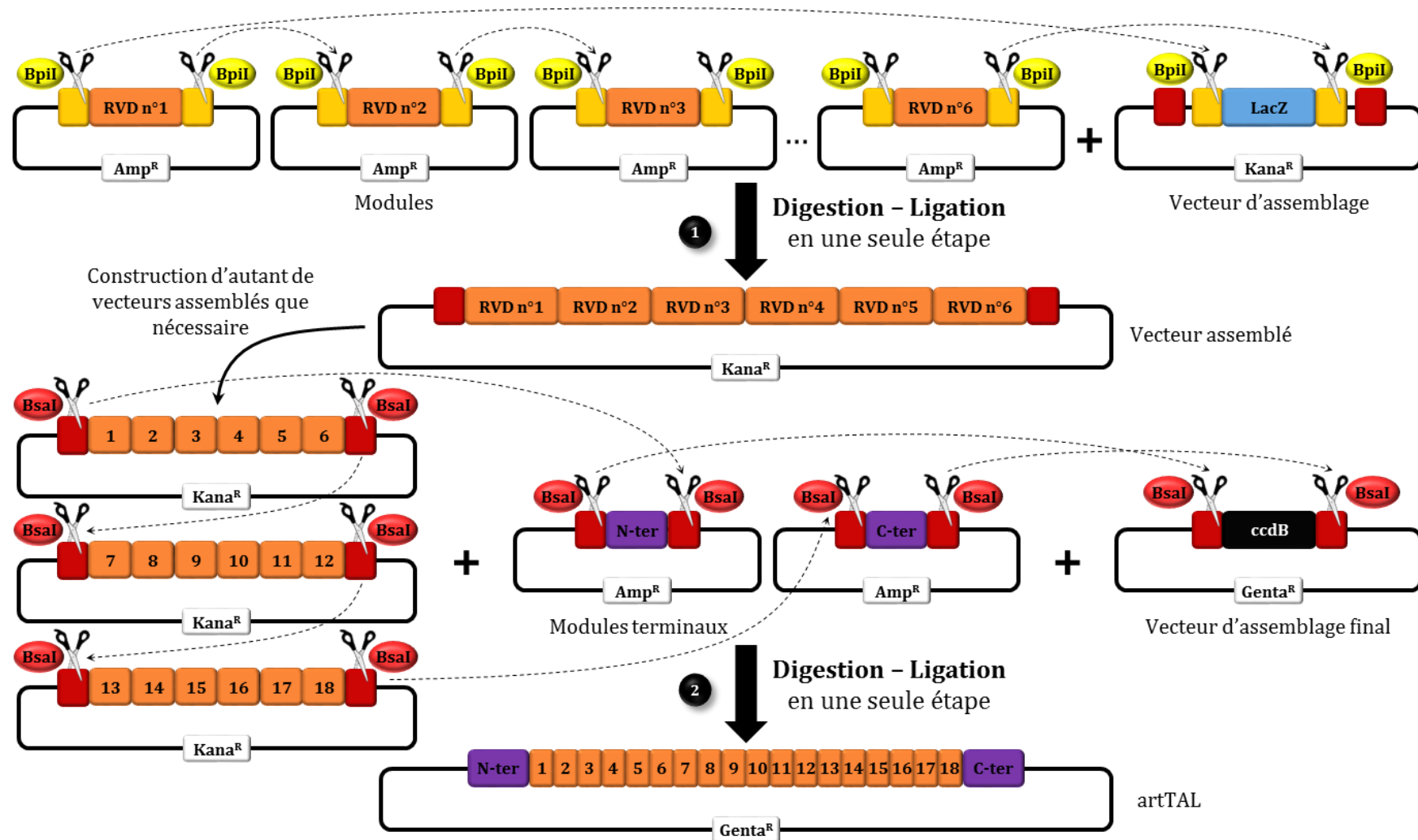


Figure 31 : Construction d'un effecteur TAL artificiel. Ce schéma est adapté de Cermak *et al.*, 2011. Un kit de construction d'effecteurs TAL artificiels (ArtTAL) se compose de vecteurs d'assemblage et de modules, contenant les répétitions, avec différents RVD, et les régions N-terminale et C-terminale. **(1)** Assemblage d'une série de six répétitions (RVD n°1 à 6) dans un vecteur contenant un gène *LacZ* par digestion / ligation utilisant BspI, suivi d'une transformation bactérienne et d'une sélection des bons clones avec un crible blanc / bleu sur un milieu sélectif contenant de la Kanamycine. **(2)** Assemblage de vecteurs assemblés avec des modules terminaux dans un vecteur contenant un gène *ccdB* en utilisant BsaI, suivi d'une transformation bactérienne et d'une sélection des bons clones, contenant des ArtTAL, avec un crible positif sur un milieu sélectif contenant de la Gentamycine.

et peuvent même activer la transcription de gènes humains (Geißler *et al.*, 2011; Zhang *et al.*, 2011). Ainsi, l'*ArtTAL* est le premier outil biotechnologique développé grâce à ces effecteurs (Boch *et al.*, 2009; Cermak *et al.*, 2011). Pour construire des *ArtTAL*, divers kits ont été développés (Cermak *et al.*, 2011; Weber *et al.*, 2011). Bien qu'un peu différents, ils sont tous basés sur le même principe : l'utilisation d'une collection de vecteurs contenant respectivement les séquences encodant les régions N-terminale, C-terminale et les répétitions d'un effecteur TAL. Ainsi, avec la diversité de répétitions appropriées on peut construire un gène codant un *ArtTAL* avec la séquence de RVD désirée via plusieurs étapes de digestion, ligation, transformation et vérification (**Figure 31**).

L'application biotechnologique la plus connue en lien avec les effecteurs TAL est celle des TALEN pour TAL Effector Nuclease (Voytas & Joung, 2009; Bogdanove & Voytas, 2011). La création des TALEN découle directement de celle des *ArtTAL*. En effet, un TALEN correspond au produit issu de la fusion du domaine d'activation d'un *ArtTAL* et d'une nucléase (**Figure 32**). Ainsi, les TALEN ciblent une séquence non pas pour activer l'expression d'un gène en aval mais pour générer une coupure au niveau de l'EBE ce qui peut induire une modification génique lors de la réparation de cette coupure (Christian *et al.*, 2010). Par conséquent, les TALEN permettent de modifier génétiquement et précisément des sites spécifiques dans un génome donné. Cette manipulation génétique a largement été exploitée sur divers génomes de plantes, d'invertébrés et de vertébrés (Carroll, 2014). Par exemple, des variétés de riz résistantes aux souches de *X. oryzae* pv. *oryzae* ont été générées en utilisant la technologie des TALEN (Li *et al.*, 2012a; Blanvillain-Baufumé *et al.*, 2017). En effet, des TALEN ont été construits dans le but de muter des EBE ciblées par divers effecteurs TAL en amont du gène *OsSWEET14*, ce qui empêche la fixation de ces effecteurs et donc la surexpression du transporteur de sucre nécessaire au développement de la maladie (Li *et al.*, 2012a; Blanvillain-Baufumé *et al.*, 2017). Un autre mécanisme de défense peut être mis en place via l'utilisation de TALEN en plaçant une ou plusieurs EBE, synthétisées *in vitro*, en amont d'un gène exécuteur pour piéger un ou plusieurs effecteurs TAL. En effet, le clonage de 6 EBE différentes en amont d'un gène *R* chez le riz, le rend résistant aux souches de *X. oryzae* pv. *oryzae* et *X. oryzae* pv. *oryzicola* portant au moins l'un des 6 effecteurs TAL piégés. Les 6 EBE en question sont ciblées par les effecteurs TAL PthXo1, PthXo6 et Tal9a de *X. oryzae* pv. *oryzae* et Tal2g, Tal4a et Tal4c de *X. oryzae* pv. *oryzicola*. Ainsi, en plus de montrer que cette méthode permet de produire des variétés résistantes, les auteurs ont montrés qu'il était également possible d'ajouter des EBE ciblées par des effecteurs TAL différents et de les cloner en amont d'un même gène *R* dans le but d'étendre la résistance de la variété végétale à un ensemble de souches pathogènes (Hummel *et al.*, 2012).

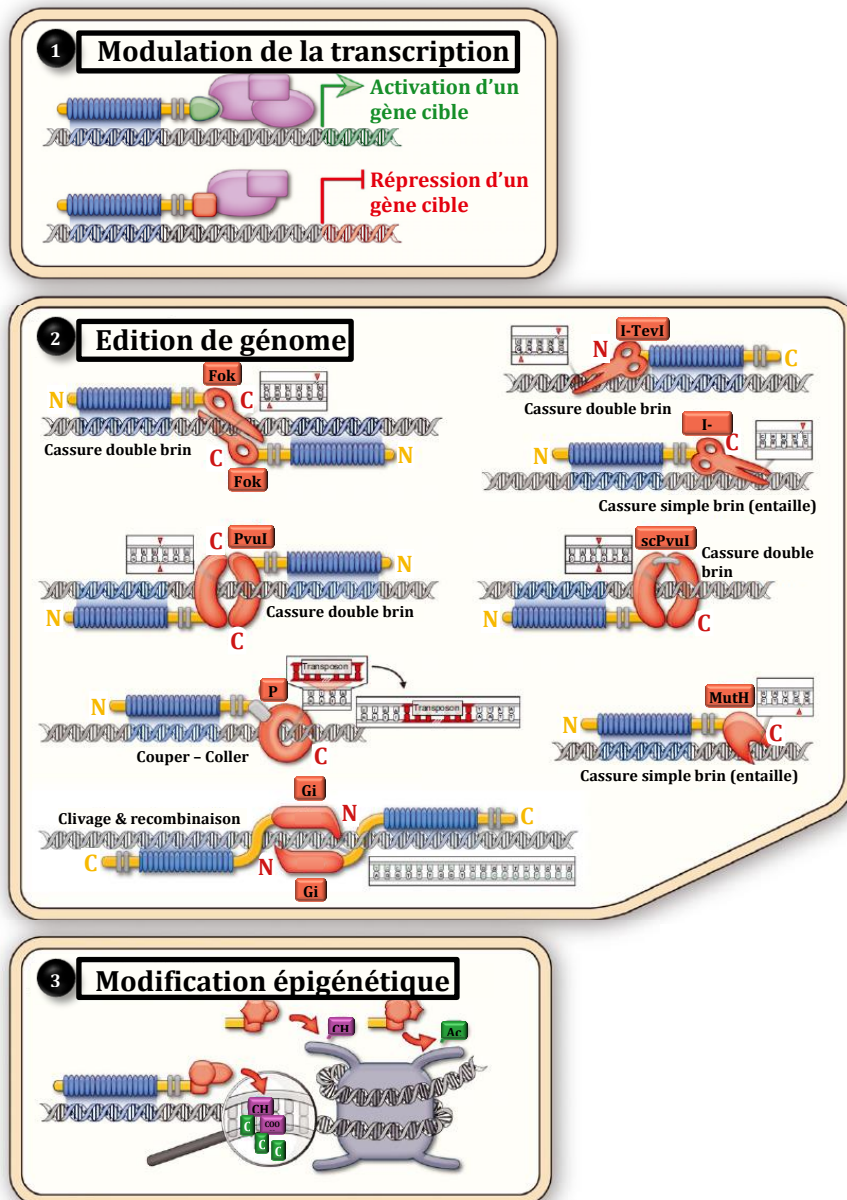


Figure 32 : Les protéines de fusion avec un effecteur TAL. Cette figure est adaptée de de Lange et al. 2014. Les différentes protéines de fusion avec un effecteur TAL (TALE) ciblent des séquences d'ADN particulières grâce à leurs répétitions représentées par des ovales bleus. **(1)** Les fusions avec des domaines d'activation (vert) ou de répression (rouge) sont respectivement utilisées pour la transcription ou la régulation négative de la transcription. **(2)** Les protéines de fusion TALE-FokI et TALE-PvuII clivent l'ADN lors de la dimérisation du domaine FokI et de l'endonuclease PvuII. FokI n'a pas de préférence pour le site cible, alors que PvuII ne fixe que les sites CAGCTG. Deux enzymes PvuII peuvent être fusionnées et ainsi une seule protéine de fusion TALE-scPvuII suffit pour rompre l'ADN double brin. Deux domaines I-TevI fusionnés avec les régions N- et C-terminales de l'effecteur TAL mènent respectivement à des cassures à double ou simple brin. I-TevI a une forte affinité avec les séquences CNNNG. MutH est une nickase, coupant sur un site GATC adjacent à la séquence cible de l'effecteur TAL. Les spécificités de séquence d'I-TevI, PvuII et MutH limitent quelque peu le choix des cibles mais permettent d'augmenter la spécificité de liaison. Gin favorise le clivage et la recombinaison, évitant ainsi l'activation des systèmes de réparation de l'ADN, mais cette recombinase cible une séquence de 20 pb spécifique, ce qui limite significativement les choix de cibles. La protéine de fusion TALE-PB (pour transposase piggyback) permet des insertions de transposons dans des sites TTAA à proximité de la cible. **(3)** Les fusions avec le domaine de liaison de l'ADN d'un effecteur TAL peuvent provoquer l'élimination du groupe méthyle des cytosines ou altérer les modifications de la queue des histones-queue dans les cibles. Certains de ces 'epiTALE' fonctionnent indirectement en recrutant des enzymes modifiant les histones.

La capacité des ArtTAL à se lier à des séquences spécifiques d'ADN a également été exploitée dans un tout autre intérêt. En effet, des ArtTAL ont été fusionnés avec des domaines répresseurs pour que ces derniers répriment l'expression de gènes ciblés (**Figure 32**) (de Lange *et al.*, 2014a). Cette méthodologie a notamment été utilisée avec succès chez *Arabidopsis* (Mahfouz *et al.*, 2012; Lin *et al.*, 2016) ce qui infique que cette répression spécifique de la transcription pourrait être utilisée pour des études de génomique fonctionnelle. D'autre part, des ArtTAL ont également été couplés à d'autres domaines fonctionnels pour des applications très variées comme la recombinaison dirigée (Mercer *et al.*, 2012), la visualisation de la chromatine (Miyanari *et al.*, 2013; Miyanari, 2014) ou encore des modifications épigénétiques ciblées (Mendenhall *et al.*, 2013; Cho *et al.*, 2015). Ainsi, les outils biotechnologiques dérivés des ArtTAL laissent imaginer de multiples perspectives en biotechnologie.

4. Le pathosystème *Phaseolus vulgaris* – *Xanthomonas* spp.

Le haricot (*Phaseolus vulgaris*) est cultivé sur tous les continents en tant que culture vivrière ou commerciale, et représente la principale source de protéine pour cinq cent millions de personnes à travers le monde, ce qui en fait une ressource majeure pour la nutrition humaine. La maladie commune du haricot est la bactériose la plus dévastatrice de la culture de haricot, en raison de sa distribution mondiale et des pertes de rendement significatives qu'elle engendre. Les souches capables d'attaquer le haricot sont groupées en quatre lignées génétiques appartenant à deux espèces de *Xanthomonas* : *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* (Alavi *et al.*, 2008; Constantin *et al.*, 2016). Ainsi, l'ensemble de ces lignées forme un clade polyphylétique et toutes sont capables d'induire les mêmes symptômes sur haricot, suggérant une convergence pathologique entre ces souches. Au cours des dernières décennies, de nombreuses publications sur ce pathosystème ont contribué à l'amélioration de la croissance et à la résistance aux diverses maladies du haricot et à la compréhension des mécanismes régissant la spécificité d'hôte. L'ensemble des informations collectées sur le pathosystème *P. vulgaris* – *Xanthomonas*, leurs évolutions respectives, la maladie commune du haricot, son épidémiologie et les moyens de contrôle et de gestion utilisés sont présentés sous la forme d'une ébauche d'article dont je suis le premier auteur.

Future prospects on the management of common bacterial blight of bean.

Mylène Ruh, Marie-Agnès Jacques and Nicolas W.G. Chen*

IRHS, INRA, AGROCAMPUS OUEST, Université d'Angers, SFR4207 QUASAV, 42, rue Georges Morel,
49071 Beaucouzé, France

* Corresponding author

E-mail addresses of all authors:

Mylène Ruh: mylene.ruh@agrocampus-ouest.fr

Marie-Agnès Jacques: marie-agnes.jacques@inra.fr

Nicolas W.G. Chen: nicolas.chen@agrocampus-ouest.fr

Abstract

Common bean (*Phaseolus vulgaris*) is a major crop and the most important source of protein for about five hundred million people worldwide. Common bacterial blight of bean (CBB) due to *Xanthomonas phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* is one of the most important diseases affecting common bean production. CBB agents are classified as quarantine pests in Europe, while are widespread in tropical and subtropical areas. In America, the use of resistant common bean accession has been successful for years, but in most other countries CBB management is restricted to cultural practices and/or testing of seed lots. Here, we describe the pathosystem and its epidemiology and the means of control and management used against the CBB as well as future prospects based on recent *Phaseolus* and *Xanthomonas* genomic data.

Keywords: Common bacterial blight of bean; *Xanthomonas citri* pv. *fuscans*; *Xanthomonas phaseoli* pv. *phaseoli*; bean crop; pathosystem; worldwide distribution

THE HOST PLANT

The legume family

The legume family (Fabaceae) includes dicotyledonous plants (herbaceous plants, shrubs, trees or lianas) whose fruit is a pod. This economically important family is highly diversified morphologically and comprises 751 genera and around 19,500 species of which 90 % are cultivated and exploited, particularly for animal and human consumption, as ornamental plant, and for wood (Cronk *et al.*, 2006; Christenhusz & Byng, 2016). As a result, the legume family corresponds to the world second largest food crop family for human consumption and, according

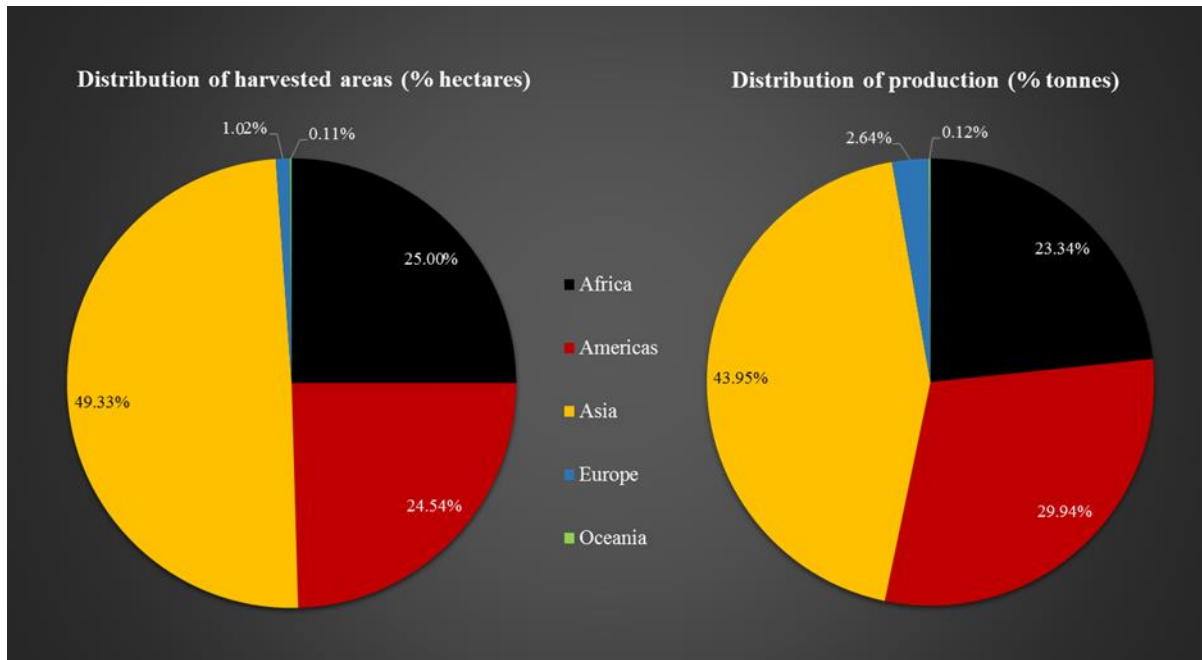


Figure 1: Worldwide distribution of common bean cultivation. These graphs were created using data available at <http://www.fao.org/faostat/en/#data/QC>. At the left, worldwide distribution of harvested areas in 2014. At the right, worldwide production distribution in 2014.

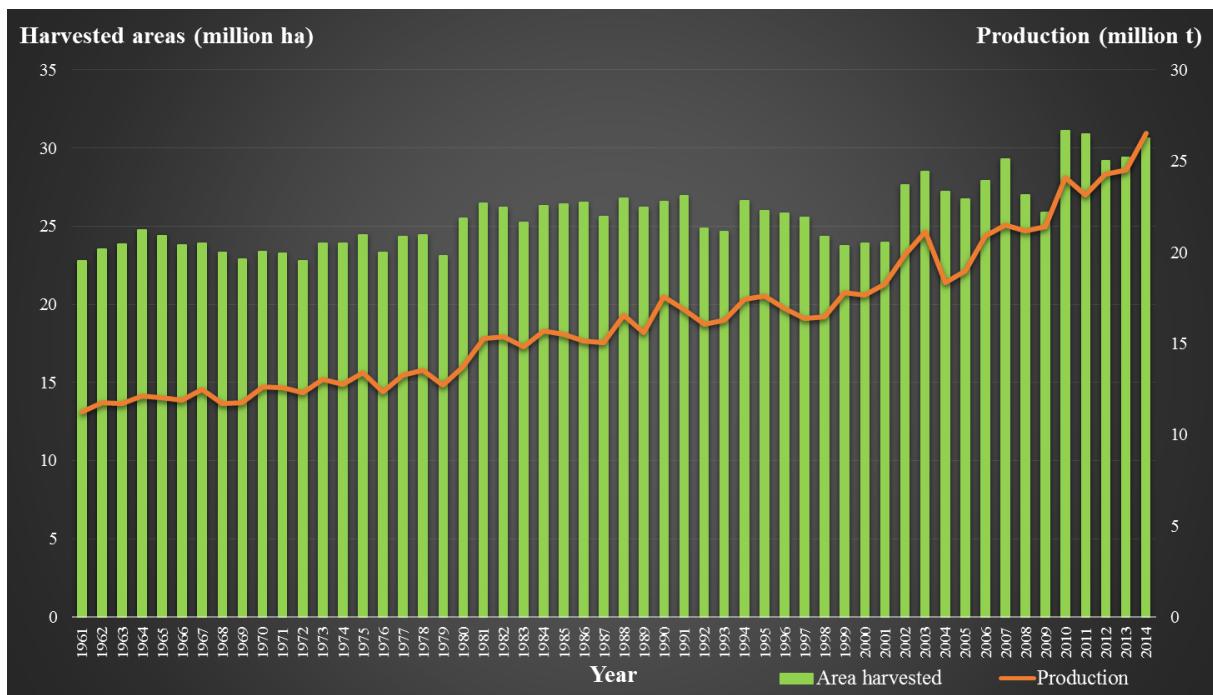


Figure 2: Evolution of worldwide cultivation of common bean since 1961. This graph was created using data available at <http://www.fao.org/faostat/en/#data/QC>. The harvested areas (green bars) are expressed in million hectares and the production (orange line) is expressed in million tonnes.

to the FAO, about 300 million tons of legumes have been produced worldwide in 2014 (Foyer *et al.*, 2016). In addition, the legume plants have the unique capacity of fixing nitrogen from the environment, through symbiosis with diverse rhizobacteria from the *Rhizobium* and *Sinorhizobium* genera living in the soil and/or transmitted by seeds (Zahran, 1999; Mora *et al.*, 2014; Ormeño-Orrillo *et al.*, 2015). Nitrogen fixation is a very beneficial trait for legume plants in natural habitats, and has been a major subject for plant research. In particular, the coevolution and physiological processes leading to this symbiosis have been extensively studied (Oldroyd *et al.*, 2011; Guan *et al.*, 2013), and breeders have long dreamed of engineering nonlegumes nitrogen-fixing crops such as cereals for reducing dependence on nitrogenous fertilizers in agriculture (Charpentier & Oldroyd, 2010; Mus *et al.*, 2016). Also, cultivated legume plants are commonly exploited in crop rotation for soil nitrogen enrichment. Large efforts have been made during the last decade to sequence the whole genome of legume plants including *Lotus japonicus* (Sato *et al.*, 2008), *Glycine max* (Schmutz *et al.*, 2010), *Medicago truncatula* (Young *et al.*, 2011; Tang *et al.*, 2014), *Cicer arietinum* (Varshney *et al.*, 2013), *P. vulgaris* (Schmutz *et al.*, 2014), *Trifolium pratense* (Ištvánek *et al.*, 2014), *Vigna angularis* (Yang *et al.*, 2015), *Arachis duranensis* (Bertioli *et al.*, 2016), or *Lupinus angustifolius* (Hane *et al.*, 2017). These genomic resources have brought important information about the natural evolution and domestication processes within the legume species, and are a stepping stone for further genomic and genetic analyses within this family.

Common bean

The genus *Phaseolus* comprises about 70 species spread throughout the world. Many *Phaseolus* species have been domesticated and are now cultivated as ornamentals, for human or animal consumption such as *P. vulgaris*, *P. coccineus*, *P. dumosus*, *P. acutifolius* or *P. lunatus*. Common bean (*P. vulgaris* L.) is grown worldwide (**Figure 1**), and is the most important legume crop for direct human consumption (Bitocchi *et al.*, 2017). It is the most important source of protein for nearly five hundred million people in the world, and is of major economic importance to many developing countries, particularly in Asia, Africa and Latin America (Ndakidemi *et al.*, 2006; Katungi *et al.*, 2011; Romero-Arenas *et al.*, 2013; De Lúque & Creamer, 2014; Nascimento *et al.*, 2016). In 2014, the FAO estimated that the worldwide production of dry beans corresponded to 27 million tonnes for 30 million harvested hectares (**Figure 2**). This legume combines many advantages: it is inexpensive to produce, nutritionally rich and needs reduced use of fertilizers due to its ability to fix nitrogen, which is in favour of sustainable agriculture (Jones, 1999; Ndakidemi *et al.*, 2006; Castro Guerrero *et al.*, 2016). This economic and social importance is also supported by nutritional and medical facts. Indeed, common bean is rich in potassium, magnesium, folic acid, starch, dietary fibers, minerals and vitamins and is also the best non-meat

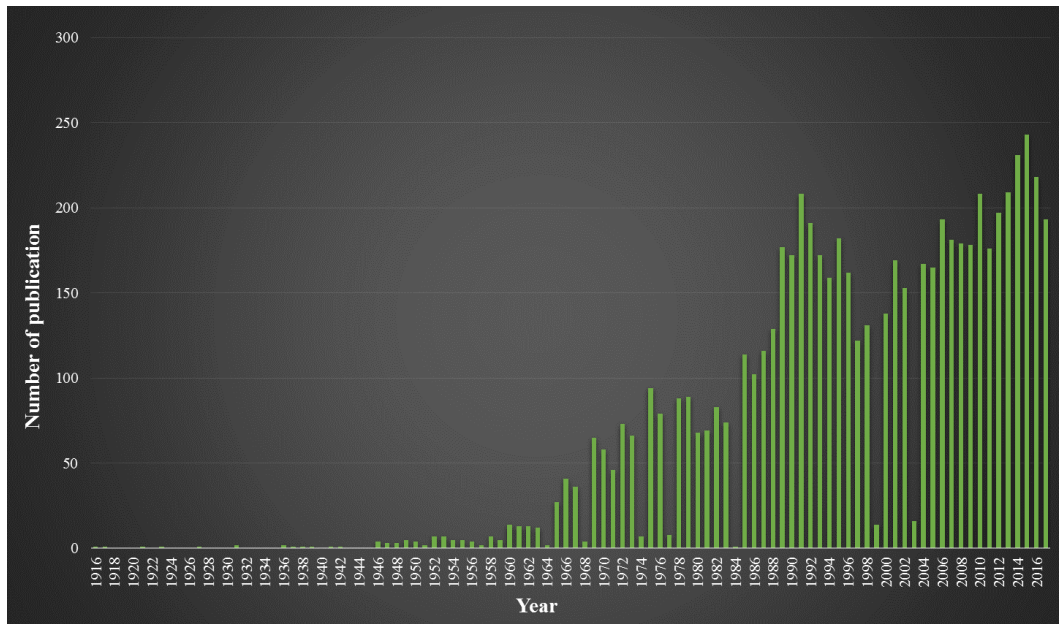


Figure 3: Evolution of common bean publications over the last century. This graph was created using data available at <https://www.ncbi.nlm.nih.gov/pubmed>. The number of publications was retrieved by searching the keyword “Phaseolus vulgaris” on NCBI.

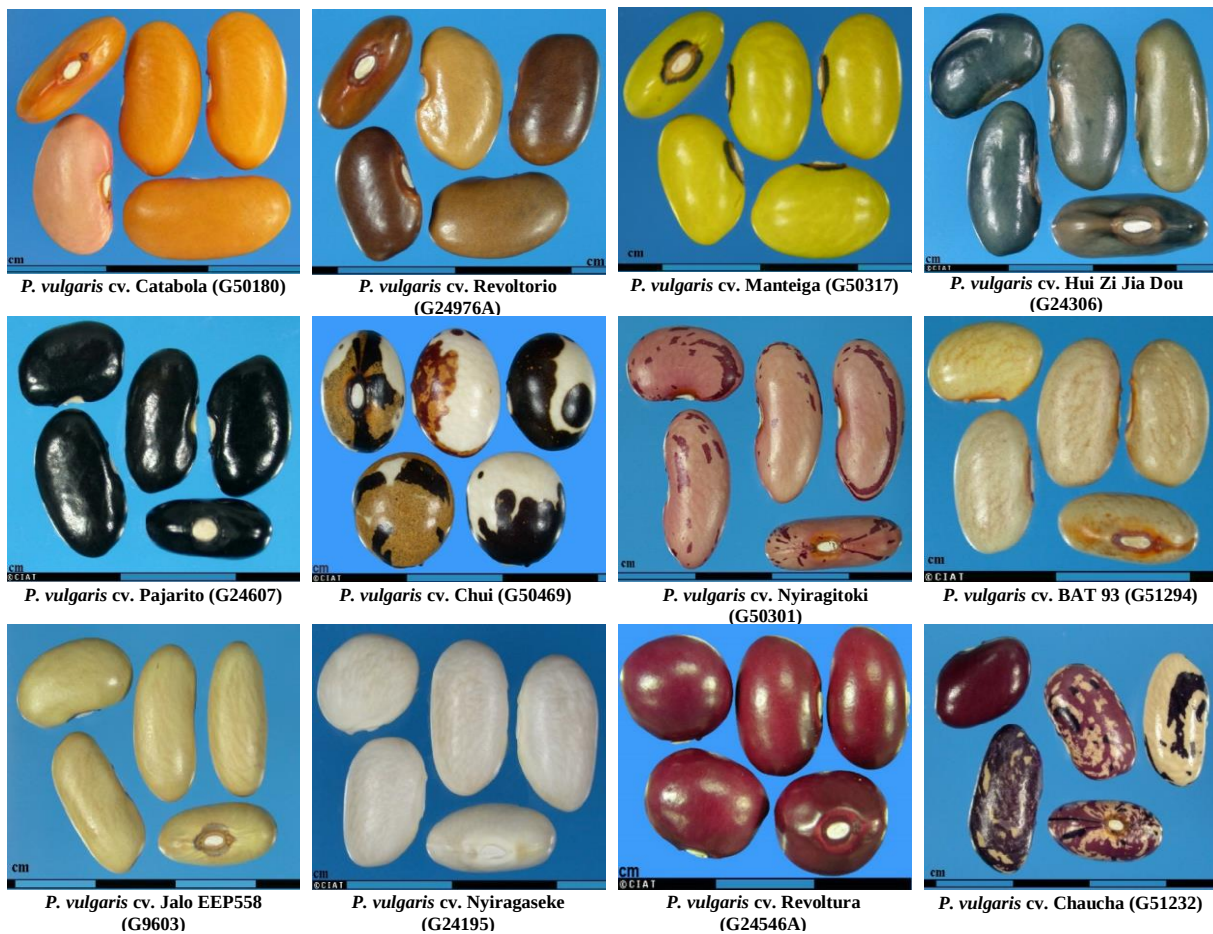


Figure 4: Diversity of common bean seeds. All these photos are from CIAT (<http://ciat.cgiar.org/what-we-do/crop-conservation-and-use/bean-diversity/>). CIAT accession numbers are indicated in parentheses.

source of iron and a major source of proteins (Lareo & González V, 1988; Messina, 2014). Also, common bean is rich in antioxidants and contain a wide range of flavonoids (Hayat *et al.*, 2014). Consequently, bean consumption would prevent the risks of chronically degenerative diseases, cardiovascular diseases, obesity, diabetes and cancer (Helmstädter, 2010; Messina, 2014; He *et al.*, 2015). However, grain legumes including common bean form only a minor part of most current human diets compared to cereals, which is in part due to the current lack of coordinated focus on grain legumes (Foyer *et al.*, 2016). The global economic importance of common bean has resulted in numerous research projects on *P. vulgaris*, which is one of the most studied legumes. Research work concerning common bean has been steadily increasing over the last 50 years, with an annual number of publications recently reaching around 200 publications per year (**Figure 3**).

The origin of *P. vulgaris* is estimated 4 to 6 million years ago in Mexico, then a subpopulation diverged in South America about 165,000 years ago (Schmutz *et al.*, 2014; Cichy *et al.*, 2015). This geographical separation, led to the formation of two genetically independent gene pools for wild common beans: the Mesoamerican gene pool, covering a surface from northern Mexico to Colombia and the Andean gene pool, located along the Pacific coast of South America from Peru to Argentina. Then, two independent episodes of domestication occurred in both the Mexican and South American gene pools around 8,000 years ago, followed by local adaptations resulting in different landraces (Chacón S *et al.*, 2005; Chacón S. *et al.*, 2007; Schmutz *et al.*, 2014). Around 500 years ago and over a period of about a century, common beans from these two gene pools were introduced into the Iberian peninsula before being disseminated in Europe, Africa and Asia (Santalla *et al.*, 2002, 2010; Angioi *et al.*, 2010; Castro Guerrero *et al.*, 2016). This rich evolutionary history has led to different genetic traits in wild and domesticated common bean genotypes, and make *P. vulgaris* a choice model for studying the genetic and phenotypic evolution of cultivated species after domestication (Bitocchi *et al.*, 2017).

Today close to 38,000 wild and landrace accessions are conserved in the global bean collection of CIAT (Centro Internacional de Agricultura Tropical) in Colombia. Interestingly, the important diversification of common beans to adapt to a wide variety of environments and the identification of the genes involved in the various phenotypic modifications and in the plant adaptations to the environment enable the development of new varieties according to the desired aesthetics, nutritional needs, climate change, food security and disease control (Gaut, 2014; Cichy *et al.*, 2015). Today, there are significant international bean improvement programs managed by CIAT in Colombia and the CRSP (Collaborative Research Support Program) in the United States allowing the development and conservation of a wide diversity of bean varieties (**Figure 4**). For example, through these breeding programs, 20 improved varieties were selected in Tanzania to control pests and diseases and increase tolerance to some edaphic stress factors (Hillocks *et al.*, 2006). Also, folate amount produced by common beans can be increased to improve the nutritional



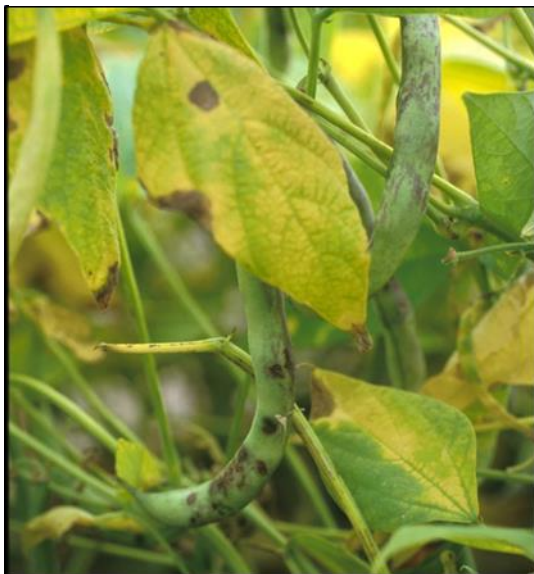
Necrotic lesion and chlorotic halo on leaves - INRA Angers (France)



Necrotic lesion and chlorotic halo on leaves (Akhavan *et al.*, 2013)



Water-soaked on pods (Akhavan *et al.*, 2013)



Pustules on pods - INRA Angers (France)



© AAC - Saint-Jean-sur-Richelieu

Longitudinally extending reddish streaks on stem - René Crête (France)



Seed discoloration at right and healthy seeds at left - V.R. Wallen, Agriculture Canada (CA)

Figure 5: Diversity of symptoms of common bacterial blight of bean caused by *X. citri* pv. *fuscans* or *X. phaseoli* pv. *phaseoli* on different common bean tissues.

status and vitamin intake (Ramírez Rivera *et al.*, 2016). Similarly, decreasing the tannin content of beans increases their nutritional quality (Islam *et al.*, 2003). However, selecting beans containing less tannins can impact their resistance to diseases and pests. Thus, further selection must be made to improve common bean nutritional quality without affecting resistance. Many other traits have been selected to improve early maturity, adaptation to high latitudes, vertical growth, higher pods quality, better yield and resistance to various diseases (Singh, 2001; Singh & Schwartz, 2010). Most of these improvements were done using varietal selection, but *in vitro* cultivation or transformation are also possible (Svetleva *et al.* 2003; Rech *et al.* 2008). For example, a transgenic variety of *P. vulgaris* resistant to bean golden mosaic virus has been obtained and is marketed in Brazil (Bonfim *et al.*, 2007), and a nutritional evaluation of this variety has been carried out to ensure its good quality (Aragão *et al.*, 2013; Carvalho *et al.*, 2015). Therefore, transgenesis and genome editing are promising ways to further generate novel common bean varieties, although technical improvements are still needed as common bean is particularly recalcitrant to *in vitro* transformation and regeneration (Hnatuszko-Konka *et al.*, 2014).

Common bean production is impaired by many pests and diseases including insects, fungi, bacteria and viruses, which significantly reduces yields (Singh & Muñoz, 1999; Miklas *et al.*, 2006). Among the most significant diseases impacting common bean yields are anthracnose caused by *Colletotrichum lindemuthianum*, angular leaf spot caused by *Phaeoisariopsis griseola*, bean common mosaic virus, bean golden mosaic virus, and common bacterial blight of bean (CBB) (Graham & Ranalli, 1997).

COMMON BACTERIAL BLIGHT OF BEAN

Symptoms

Common bacterial blight of bean (CBB) was described for the first time in 1893. Foliar symptoms occur as water-soaked spots evolving into dry and brown necrotic lesions encircled by a narrow chlorotic halo. Necrotic lesions later coalesce until leaves appear burned, which can cause total defoliation of the plant. Infected stems and pods also present water-soaked spots that turn from yellow to dark red in the later stages. Inside the pods, seed symptoms occur as butter yellow spots that turn into brown spots distributed throughout the seed coat, or restricted to the hilum area. Severely affected seeds can eventually shrivel, which strongly affects their germination rate and vigour. If the infected seeds are sown, they may not germinate or produce dying plants or plants with lesions on cotyledons, stems and trifoliates (**Figure 5**) (EPPO/CABI, 1986; Hayward, 1993; EFSA PLH Panel (EFSA Panel on Plant Health), 2014).

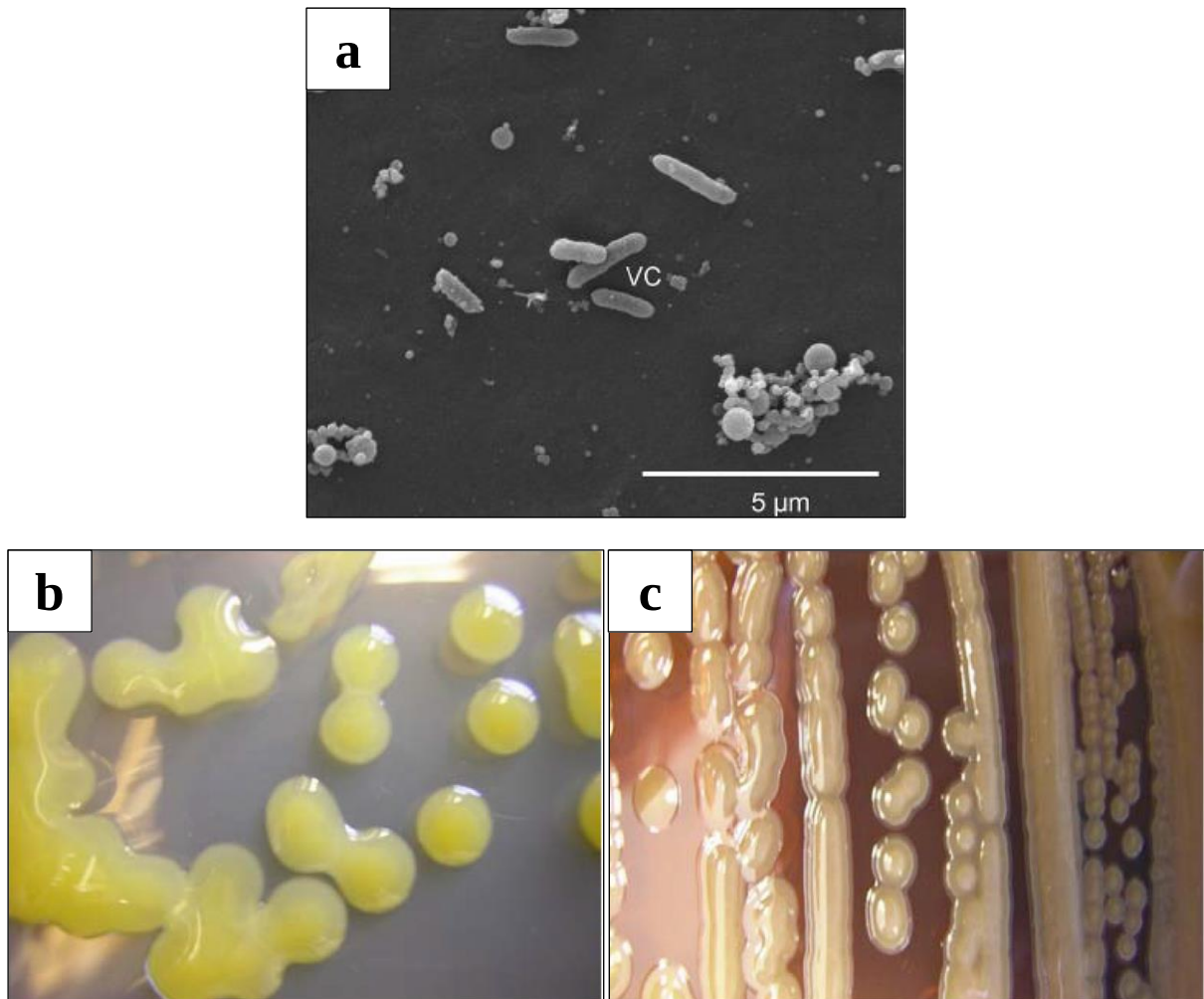


Figure 6: Morphology of the bacteria responsible for common bacterial blight of bean. (a) Scanning electron micrographs of *X. phaseoli* pv. *phaseoli* (Spago *et al.*, 2014). (b) Typical yellow colonies of *X. phaseoli* pv. *phaseoli* on culture medium (Duncan *et al.*, 2011). (c) Brown pigment produced and solubilised in culture medium by *X. citri* pv. *fuscans* (Duncan *et al.*, 2011).

The bacterial strains responsible for this disease were formally named *Bacillus phaseoli* by Smith in 1897, then were renamed *Xanthomonas phaseoli* 42 years later, as they presented all common characteristics of this newly proposed genus (Dowson, 1939). These strains were described as conventional aerobic *Xanthomonas* strains with sizes between 0.4-0.9 µm wide and 0.6-2.6 µm long (**Figure 6a**) and which form convex, yellow and shining colonies on agar medium (**Figure 6b**) (Corey & Starr, 1957). Strains producing a brown pigment, which is soluble in medium containing tyrosine, were identified later as the fuscous strains (**Figure 6c**) (EPPO/CABI, 1986; Goodwin & Sopher, 1994). This brown pigment is the result of the disruption of the tyrosine catabolism, possibly due to the pseudogenization of the *HmgA* gene encoding a homogentisate oxygenase in the fuscous variants (Darrasse *et al.*, 2013b). Four decades later, fuscous and non-fuscous strains were grouped under the name *X. campestris* pv. *phaseoli* (Dye *et al.*, 1980). It was not until 1995 that all strains responsible for CBB were renamed *X. axonopodis* pv. *phaseoli* (Vauterin *et al.*, 1995). In 2005, the work conducted by Rademaker and colleagues (2005), using PCR amplifications of the conserved repetitive sequences (BOX), enterobacterial repetitive intergenic consensus (ERIC) and repetitive extragenic palindromic sequences (REP) forming the rep-PCR method, revealed that strains responsible for CBB were genetically diverse and belonged to two separate rep-PCR groups. Non-fuscous strains belonged to *X. axonopodis* Rep-PCR group 9.4, while the fuscous variant belonged to *X. axonopodis* Rep-PCR group 9.6. The same year, Schaad and colleagues (2005) based on DNA-DNA hybridizations, Amplified Fragment Length Polymorphism (AFLP), and sequencing of the 16S-23S intergenic spacer (ITS) regions proposed to reallocate the fuscous strains into a new species / subspecies entity named *X. fuscans* subsp. *fuscans*, while the non-fuscous lineage conserved the previous name *X. axonopodis* pv. *phaseoli*. In 2008, the genetic diversity of a large collection of strains pathogenic on common bean was studied by AFLP and led to the description of four genetic lineages for CBB agents (Alavi *et al.*, 2008). Within Rep-PCR group 9.6, the fuscous strains grouped within the genetic lineage *fuscans* that was closely related to two genetic lineages of non-fuscous (NF) strains named NF2 and NF3, while the non-fuscous strains from Rep-PCR group 9.4 grouped in a single genetic lineage named NF1. More recent works led to the taxonomical revision of the *Xanthomonas axonopodis* species complex into four distinct species (Ah-You *et al.*, 2009; Rodriguez-R *et al.*, 2012). In this proposal, the *fuscans*, NF2 and NF3 lineages belong to the newly described *X. citri* species in a pv. *fuscans* (*Xcf*) while the lineage NF1 belongs to *X. phaseoli* pv. *phaseoli* (*Xpp*) (Constantin *et al.*, 2016).

The main host plant for *Xcf* and *Xpp* strains is *P. vulgaris*, but other legumes are naturally infected by these two bacteria, including *P. lunatus*, *Vigna aconitifolia*, *V. radiata*, *Lablab purpureus* and *Mucuna deeringiana*. *P. coccineus*, *P. acutifolius* and *Lupinus polyphyllus* can be infected by these bacteria after artificial inoculation (EFSA PLH Panel (EFSA Panel on Plant Health), 2014).

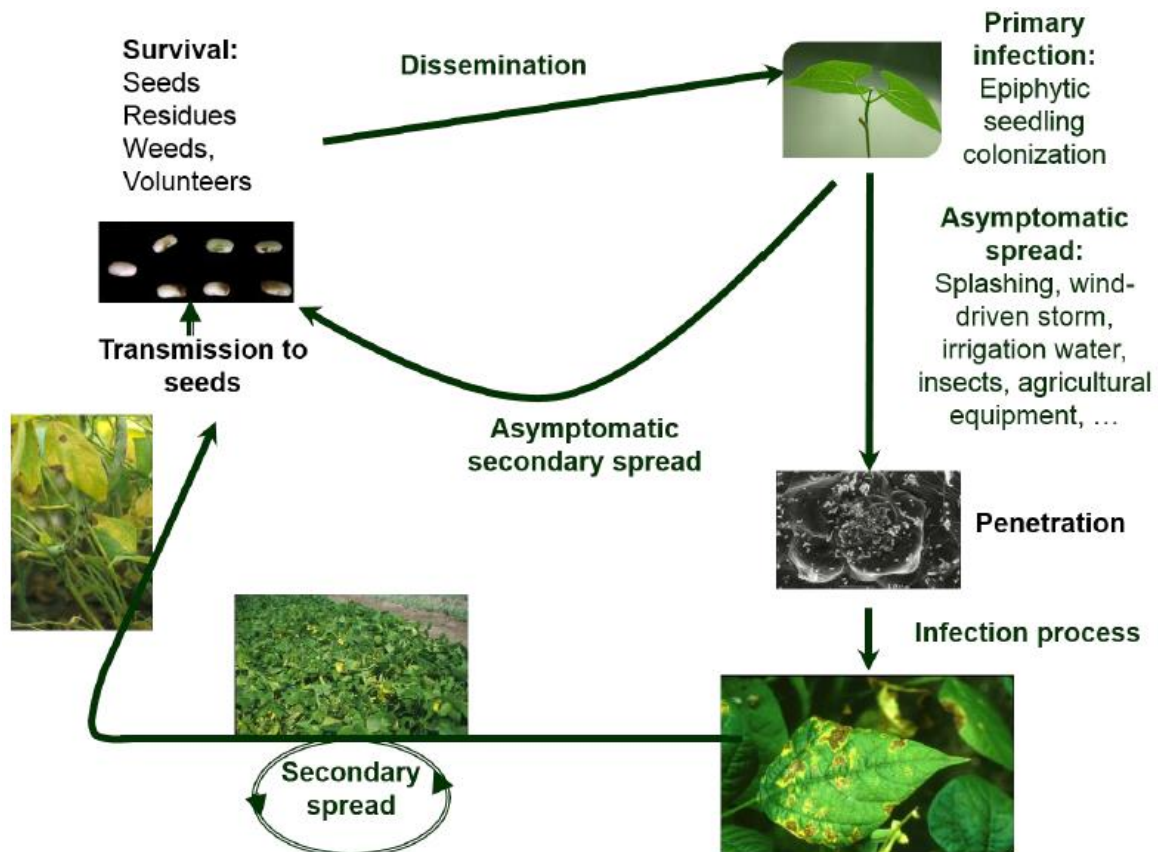


Figure 7: Disease cycle of common bacterial blight of bean (EFSA PLH Panel (EFSA Panel on Plant Health), 2014).

Xcf and *Xpp* cause identical symptoms, but *Xcf* has often been reported to be more aggressive than *Xpp* (Schuster *et al.*, 1973; Opio *et al.*, 1996; EFSA PLH Panel (EFSA Panel on Plant Health), 2014). Bean varieties can be more or less resistant to the disease and it has been suggested that a gene-for-gene relationship influencing the aggressiveness of the strain exists between the pathogen and its host (Opio *et al.*, 1996; Bai *et al.*, 1997; Singh & Muñoz, 1999; Hillocks *et al.*, 2006; Alladassi *et al.*, 2017). Geographical origin also seems to have an effect on the aggressiveness of the strain. Indeed, it was demonstrated that when *Xcf* and *Xpp* come from the same geographical origin, *Xcf* induce greater symptoms than *Xpp*, but that African strains were the most pathogenic and that the variation in pathogenicity was more important for *Xpp* than for *Xcf* and that this variation was all the more important when the strains originated from the Caribbean and South America (Mutlu *et al.*, 2008).

Modes of transmission

The disease cycle of CBB is divided into three main phases that can explain the symptom developments described above. The cycle starts with an epiphytic phase, followed by a pathogenic phase and a survival phase (**Figure 7**). Primary infection usually start with an epiphytic phase where CBB agents are able to grow on the plant surface without invading the internal tissues of the host for long periods of time (Belete & Bastas, 2017). After growing on the bean leaf surface bacteria penetrate into the tissues of the host plant through openings such as stomata, hydathodes or woundings. Inside the host tissues the bacteria multiply exponentially and may express their pathogenicity when conditions are favourable (Darsonval *et al.*, 2008; Akhavan *et al.*, 2013). This phase therefore corresponds to the expression of the symptoms of the disease, which, thanks to the deterioration of the plant tissues and the production of bacterial exudates, allows a rapid second spread of the disease. The secondary inoculum can spread by direct contact between infected leaves, or by the irrigation waters, the rain and the wind, as well as leaf-feeding insects, animals, people and seeds (Kaiser & Vakili, 1978; EPPO/CABI, 1986; Akhavan *et al.*, 2013; Jacques *et al.*, 2016; Belete & Bastas, 2017). The spread of disease as well as its incidence and severity are favoured by warm temperature, high humidity and low rain conditions (Belete & Bastas, 2017). Finally, the survival phase corresponds to the overwintering of the bacteria on previously infected plant debris or agricultural equipment. During this phase, bacteria produce exopolysaccharides and form biofilms acting as protective barriers against unfavourable conditions and favouring the survival of the bacteria in the soil, living host or non-host plants, plant debris and/ or in the host seeds (Jacques *et al.*, 2005; Darsonval *et al.*, 2009). In tropical and subtropical areas, these latent inocula represent a particularly effective strategy for the dissemination of *Xcf* and *Xpp*, because they can survive from one culture to another and perpetually contaminate the same field if not properly treated. For example, in Ethiopia, CBB incidence has been assessed at the flowering stage,

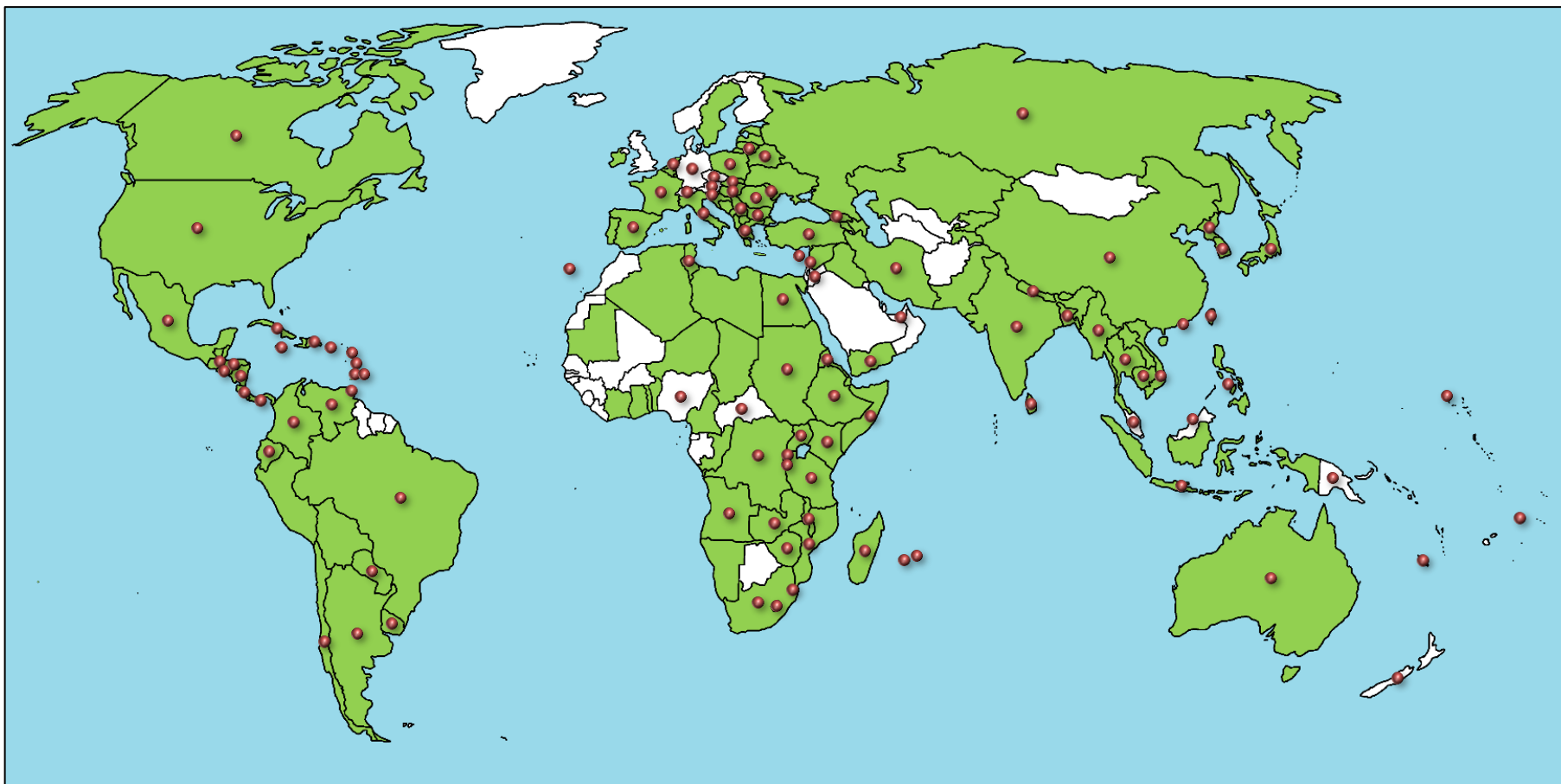


Figure 8: Worldwide distribution of common bacterial blight of bean. In green are represented the countries where dry beans are harvested, according to the last data of the FAO (<http://www.fao.org/faostat/en/#data/QC>). Orange points indicate countries where common bacterial blight of bean has been reported according to the last data of the EPPO (<https://gd.eppo.int/taxon/XANTPH/distribution>).

and on the contaminated debris, it was higher by 59 % compared to the bean from treated seed plots. Also, the bacterial survival is higher in the debris at or near the soil surface than in the residue turned during plowing (Fininsa & Tefera, 2001; Belete & Bastas, 2017). This is probably why phytopathogenic bacteria have developed many other survival mechanisms such as those associated with seeds, leaf-feeding insects or epiphytic life on non-host plants (Schuster & Coyne, 1974; Darrasse *et al.*, 2010).

The main source of CBB inoculum corresponds to infected seeds. Indeed, *Xcf* and *Xpp* are able to enter seeds via the vascular system or through the stigma (Darsonval *et al.*, 2008, 2009). Then, both the seed and the pathogen can usually survive for several years, and when the infected seed germinates, the bacteria are transmitted from the seed to the seedling (Baker & Smith, 1966; Darrasse *et al.*, 2007; Akhavan *et al.*, 2013; Shade *et al.*, 2017). This mechanism of dissemination is a major constraint for bean production, because contaminated seed lots often present no visible symptoms (Abd-Alla *et al.*, 2010). Also, low contamination levels (1×10^3 cfu/ml) of sown seeds are sufficient to contaminate a field (Opio *et al.*, 1993; Darrasse *et al.*, 2007). In addition, the transport of infected seed by human exchanges is a major way for *Xcf* and *Xpp* to travel long distances and spread around the world. Thus, one can imagine that CBB is a worldwide disease in part because of its ability of being transmitted to the seed, but also because of the expansion of international trade, which relies on global exports and imports of potentially contaminated plant material.

Significance and worldwide distribution

According to the latest available EPPO data, strains responsible for CBB are distributed worldwide, on all continents, especially in dry bean producing countries (**Figure 8**) (EPPO/CABI, 1986). For example, the top ten bean producers in 2014 (according to FAO): Myanmar, India, Brazil, United States, Mexico, Tanzania, China, Uganda, Kenya and Ethiopia are all affected by CBB. Under warm to high temperatures and high humidity, CBB can cause yield losses of up to 40 % and Africa is the most affected with 220 thousand tons loss by year (Wortmann *et al.*, 1998; Opio *et al.*, 2002; Mutlu *et al.*, 2008; Belete & Bastas, 2017). Recently, local losses of up to 75 % have been reported in Kenya (Belete & Bastas, 2017). The global distribution of CBB would not have been possible without an industrial common bean seed production and commercialization on a worldwide scale. The international trade in seeds play thus an important role in pathogen dissemination (EFSA PLH Panel (EFSA Panel on Plant Health), 2014). Moreover, this intensification of international trade favours contacts between strains, which may allow horizontal transfer of genes associated with virulence (Aritua *et al.*, 2015; Ruh *et al.*, 2017a).

DISEASE MANAGEMENT

Methods used

In Europe, including countries from the European and Mediterranean Plant Protection Organization (EPPO), as well as Turkey and Bahrain, CBB is under the A2 list of quarantine pests, meaning that it is locally present (EPPO (2017) EPPO Global Database (available online at <https://gd.eppo.int>); EFSA PLH Panel (EFSA Panel on Plant Health), 2014). It is absent or no longer present in Azerbaijan and Israel, respectively, and under strict quarantine status in both countries. In quarantine areas, seed lots are routinely tested using a method from the International Seed Testing Association involving isolation of bacterial strains, pathogenicity tests, and specific PCR assays (Audy *et al.*, 1994; Boureau *et al.*, 2013; Grimault *et al.*, 2014). Seed lots are then rejected and/or eliminated if contaminated (at a rate of one seed out of 5000), which is an efficient way to control the disease in these countries.

In many producing countries, chemical control is restricted to the use of copper-based solutions, such as copper sulphate or copper hydroxide, but also of potassium N-methyldithiocarbamate, for precocious foliar application in order to limit pathogen multiplication on plant surfaces (Singh & Muñoz, 1999; Fininsa, 2003). However, this method is costly and in most cases inefficient, because the pathogen is vascular and copper treatments are not systemic. Also the massive use of copper-based chemicals can result in the selection of copper-resistant strains. Antibiotics could be efficient, but their use is authorized only in some countries, such as the USA, because of the risk of developing resistant set (EFSA PLH Panel (EFSA Panel on Plant Health), 2014; Belete & Bastas, 2017). It has also been reported that the use of essential oils and antimicrobial compounds has a significant impact on the pathogens (Kelemu *et al.*, 2004; Lo Cantore *et al.*, 2009; Sati & Joshi, 2011; Mageshwaran *et al.*, 2012; Gakuubi *et al.*, 2016; Todorović *et al.*, 2016).

Given the lack of efficient treatments, cultural practices are essential to control CBB. The most important practice correspond to the reduction of the initial CBB inoculum, especially since it is able to survive in association with bean debris in the soil for several month (Arnaud-Santana *et al.*, 1991). It is therefore recommended to avoid overwintering of *Xcf* and *Xpp* by eliminating weeds, infected beans and other potential hosts (EFSA PLH Panel (EFSA Panel on Plant Health), 2014). The disposal of contaminated debris can reduce significantly the survival of CBB responsible for the primary inoculum. Likewise, it is necessary to clean and sterilize all harvesting equipment regularly (Belete & Bastas, 2017). Moreover, it has been shown that the type of irrigation plays an important role in the spread of the disease and it is preferable to favour irrigation by furrow compared to spray irrigation (Akhavan *et al.*, 2013). Intercropping is a proven practice to reduce the incidence of CBB agents (Fininsa, 1996, 2003). For example, the growing of

common beans with maize allows to reduce CBB incidence because the maize appears to provide a physical barrier to the movement of *Xcf* and *Xpp* between beans plants (Fininsa, 1996). However, the use of varietal mixture including different plant species can reduce the quality of the harvest, therefore crop rotation is privileged. Indeed, CBB epidemics can effectively be reduced through employing long crop rotations (at least three years) because it can limit the risk of contamination of common bean crops by pathogens surviving on alternative inoculum sources. However, the choice of the other crop is very important, for example, onion should be avoided, since it can provide a source of inoculum by asymptomatic epiphytic colonization (EFSA PLH Panel (EFSA Panel on Plant Health), 2014; Belete & Bastas, 2017). Another way to limit the development of the disease is to grow common bean under climatic conditions not suitable for *Xpp* and *Xcf*, such as arid tropical areas, which are poorly affected by CBB (EFSA PLH Panel (EFSA Panel on Plant Health), 2014).

In America, the use of resistant varieties is the most effective strategy and environmentally sound approach to controlling CBB, and is pivotal to all other CBB control measures. The identification of at least 24 quantitative trait loci (QTL) for resistance to CBB spread in different lines across the 11 common bean chromosomes offers a solid source for resistance genes (Singh & Muñoz, 1999; Santos *et al.*, 2003; Miklas *et al.*, 2006; Vandemark *et al.*, 2008; Liu *et al.*, 2010; Duncan *et al.*, 2011; Shi *et al.*, 2011a,b; Alladassi *et al.*, 2017). Since the early 2000s, the use of a major QTL from common bean great northern landrace cultivar Montana No. 5 has been effectively used for generating CBB resistant varieties used in north and south Americas (Miklas *et al.*, 2003). However, resistance to CBB is often moderate in common bean, compared to levels of resistance in *P. coccineus*, *P. lunatus*, and *P. acutifolius*. It is thus important to make use of the genetic diversity present in the gene pools of wild and landrace populations of *Phaseolus* species as a potential source for resistance genes. For example, a germplasm line with a high level of CBB resistance was developed from an interspecific cross between *P. vulgaris* and *P. acutifolius* and has been used in white and coloured bean breeding programs in both the US and Canada (Belete & Bastas, 2017). The development of genetic maps and molecular markers has facilitated the selection of resistant varieties by molecular-assisted breeding (Bai *et al.* 1997; Freyre *et al.*, 1998; Yu *et al.* 2000; Shi *et al.* 2011a, 2012; Meziadi *et al.* 2016). In addition, the sequencing of Andean and Mesoamerican common bean genomes brings even more materials for unveiling novel resistance genes against CBB (Perry *et al.* 2013; Schmutz *et al.* 2014, Vlasova, 2016). Today, all this information are exploited by breeders, and by the scientific community, and in particular by international bean improvement programs managed by the CIAT and CRSP in order to develop and conserve a wide diversity of bean varieties including CBB resistant varieties.

Future prospects

Finally, more detailed knowledge of the strains responsible for CBB is essential to understand the host-pathogenic interaction governing this pathosystem and to control CBB. In particular, the sequencing of various strains of *Xpp* and *Xcf* allowed to obtain further knowledge about the virulence genes of the bacteria that are important for the development of CBB. At present, 45 whole genomes of strains responsible for CBB are available in public databases (Darrasse *et al.*, 2013b; Indiana *et al.*, 2014; Aritua *et al.*, 2015; Ruh *et al.*, 2017a). On average, the total genome size is about 5.2 Mb, and comprises about 4,500 genes. Using these data, comparative genomics analyses have pointed out one hundred genes conserved with 100 % nucleotidic identity between phylogenetically distant strains from the *Xpp* NF1 and *Xcf* fuscans lineages, suggesting massive horizontal gene transfer (HGT) events between these lineages (Aritua *et al.*, 2015). In particular, genes encoding Transcription Activator-Like (TAL) type III effectors have been recently transferred by HGT between all lineages of CBB agents (Ruh *et al.*, 2017a). These results suggest that a large set of horizontally transferred genes could be the genetic basis determining the pathological convergence of *Xcf* and *Xpp* strains for infecting common bean. Previous work based on the sequencing of PCR products in a large diversity of *Xanthomonas* strains have shown that repertoires of particular gene families were associated with host specificity (Hajri *et al.*, 2009; Mhedbi-Hajri *et al.*, 2011). This was the case of chemotaxis receptors, capable of detecting attractive or repulsive molecules to orient the bacteria; TonB-dependent transporters, capable of liking and carrying out an active and high affinity transport of siderophores, vitamin B12 and carbohydrates; two-component regulatory systems, involved in sensing and response to changes in many different environmental conditions; adhesins, involved in the adhesion of strains to a surface and in the biofilm formation; or type III effectors, possessing various functions such as bypassing the plant defences or strain aggressiveness (Hajri *et al.*, 2009; Mhedbi-Hajri *et al.*, 2011).

Information on conserved genes important for CBB can be exploited to enhance the management of the disease in at least two ways. First, reliable detection methods, using PCR, specific markers of specific primers, MLSA or AFLP, have already been developed to efficiently detect *Xcf* and *Xpp*, in contaminated materials such as seed lots (Toth *et al.*, 1998; Parkinson *et al.*, 2007, 2009; Young *et al.*, 2008; Mayer *et al.*, 2010; Albuquerque *et al.*, 2011; Boureau *et al.*, 2013; EFSA PLH Panel (EFSA Panel on Plant Health), 2014). Another step for improving detection would be to search for genes that are indispensable for infecting common bean, and develop specific PCR primers on these genes for detecting CBB agents. The main advantage of this would be to forecast the apparition of novel recombinant strains potentially dangerous for common bean, by detecting horizontal transfers of these genes in non-*Xcf*, non-*Xpp* *Xanthomonas* populations.

A second way would be to use the genes encoding TAL effectors. Indeed, TAL effectors are known to play a major role in the virulence of bacteria of the genus *Xanthomonas*. These are type III effectors injected into the host cells via the type III secretion system. There, they have the ability to migrate to the nucleus in order to bind specifically to the promoter of a host susceptibility gene to activate its expression (Boch & Bonas, 2010). Different strategies using TAL effectors have been successfully performed for engineering resistance against *Xanthomonas* strains pathogenic on rice, either by mutating the DNA sequence targeted by the TAL effector, or by placing this target sequence in the promoter of a resistance gene (Hummel *et al.*, 2012, 2016; Li *et al.*, 2012b; Schornack *et al.*, 2013; Boch *et al.*, 2014; Wang *et al.*, 2015; Zeng *et al.*, 2015; Hutin *et al.*, 2016; Blanvillain-Baufumé *et al.*, 2017). These TAL-based strategies would require the development of genetically modified common bean varieties, and the use of genome editing strategies such as the CRISPR/Cas9 system (Ran *et al.*, 2013; Hsu *et al.*, 2014; Nishimasu *et al.*, 2014; Belhaj *et al.*, 2015; Bortesi & Fischer, 2015).

CONCLUSION

Common bacterial blight of bean, caused by *Xcf* and *Xpp* strains, causes considerable yield losses on common bean crops all over the world. In view of the economic, social and nutritional importance of common bean cultivation and the fact that CBB agents are quarantine pests, a great deal of research was undertaken to understand the mechanisms governing host-pathogen interaction of this pathosystem and in order to effectively control CBB. All these researches, have brought crucial information on the diversity and evolutionary history of both the plant and the bacteria. In particular, the analysis of the diversity of *Xanthomonas* strains responsible for this disease shed light on pathological convergence events involving HGT. Control of CBB is often limited to good cultural practices, and to the systematic detection of pathogens in seeds lots. However, CBB is far from being eradicated, and due to the ability of strains to be transmitted by seeds and to survive for several months in debris, this disease still produces significant yield losses in many of the largest producers of beans. Moreover, with the intensification of agriculture and the development of international trade the risks that this disease emerges in novel areas are important. It is therefore important to exploit the data acquired, such as the genomes of the *Xcf* and *Xpp* strains and of the common bean to develop new methods for genetic improvement of common bean against CBB. Finally, the ability of the bacteria to acquire new virulence factors by HGT as well as global climate change must both be taken into consideration in order to adapt the means of control, especially since this disease is favoured by a warm and humid climate that is predicted to spread throughout the world.

LITERATURE CITED

- Abd-Alla MH, Bashandy SR, Schnell S. 2010.** Occurrence of *xanthomonas axonopodis* pv. *phaseoli*, the causal agent of common bacterial blight disease, on seeds of common bean (*Phaseolus vulgaris* L.) in upper Egypt. *Folia Microbiologica* **55**: 47–52.
- Ah-You N, Gagnevin L, Grimont PAD, Brisse S, Nesme X, Chiroleu F, Bui Thi Ngoc L, Jouen E, Lefeuvre P, Vernière C, et al. 2009.** Polyphasic characterization of xanthomonads pathogenic to members of the Anacardiaceae and their relatedness to species of *Xanthomonas*. *International Journal of Systematic and Evolutionary Microbiology* **59**: 306–318.
- Akhavan A, Bahar M, Askarian H, Lak MR, Nazemi A, Zamani Z. 2013.** Bean common bacterial blight: pathogen epiphytic life and effect of irrigation practices. *SpringerPlus* **2**: 41.
- Alavi SM, Sanjari S, Durand F, Brin C, Manceau C, 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 h. *Applied and Environmental Microbiology* **74**: 3295–3301.
- Albuquerque P, Caridade CMR, Marcal ARS, Cruz J, Cruz L, Santos CL, Mendes M V., Tavares F. 2011.** Identification of *Xanthomonas fragariae*, *Xanthomonas axonopodis* pv. *phaseoli*, and *Xanthomonas fuscans* subsp. *fuscans* with novel markers and using a dot blot platform coupled with automatic data Analysis. *Applied and Environmental Microbiology* **77**: 5619–5628.
- Alladassi BME, Nkalubo ST, MukanKusi C, Mwale ES, Gibson P, Edema R, Urrea CA, Kelly JD, Rubaihayo PR. 2017.** Inheritance of resistance to common bacterial blight in four selected common bean (*Phaseolus vulgaris* L.) genotypes. *Journal of Plant Breeding and Crop Science* **9**: 71–78.
- Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Zeuli PLS, Papa R. 2010.** Beans in Europe: Origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theoretical and Applied Genetics* **121**: 829–843.
- Aragão FJL, Nogueira EOPL, Tinoco MLP, Faria JC. 2013.** Molecular characterization of the first commercial transgenic common bean immune to the Bean golden mosaic virus. *Journal of Biotechnology* **166**: 42–50.
- Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. 2015.** Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Frontiers in Microbiology* **6**: 1–18.
- Arnaud-Santana E, Pena-Matos E, Coyne DP, Vidaver AK. 1991.** Longevity of *Xanthomonas campestris* pv. *phaseoli* in naturally infested dry bean (*Phaseolus vulgaris*) debris. *Plant Disease* **75**: 952–953.
- Audy P, Laroche A, Saindon G, Huang HC, Gilbertson RL. 1994.** Detection of the Bean Common Blight Bacteria, *Xanthomonas campestris* pv. *phaseoli* and *X. c. phaseoli* var. *fuscans*, Using the Polymerase Chain Reaction. *Molecular Plant Pathology* **84**: 1185–92.
- Bai Y, Michaels TE, Pauls KP. 1997.** Identification of RAPD markers linked to common bacterial blight resistance genes in *Phaseolus vulgaris* L. *Genome* **40**: 544–551.
- Baker KF, Smith SH. 1966.** Dynamics of Seed Transmission of Plant Pathogens. *Annual Review of Phytopathology* **4**: 311–332.
- Belete T, Bastas K. 2017.** Common Bacterial Blight (*Xanthomonas axonopodis* pv. *phaseoli*) of Beans with Special Focus on Ethiopian Condition. *Journal of Plant Pathology & Microbiology* **8**.
- Belhaj K, Chaparro-garcia A, Kamoun S, Patron NJ, Nekrasov V. 2015.** Editing plant genomes with CRISPR / Cas9. *Current Opinion in Biotechnology* **32**: 76–84.
- Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, Liu X, Gao D, Clevenger J, Dash S, et al. 2016.** The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature genetics* **48**: 438–446.

- 437 **Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Santo D, Nanni L, Attene G, Papa R. 2017.** Beans (*Phaseolus*
438 *ssp.*) as a Model for Understanding Crop Evolution. *Frontiers in Plant Science* **8**: 1–21.
- 439 **Blanvillain-Baufumé S, Reschke M, Solé M, Auguy F, Doucoure H, Szurek B, Meynard D, Portefaix M, Cunnac S,**
440 **Guiderdoni E, et al. 2017.** Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals
441 differential activities for SWEET14-inducing TAL effectors. *Plant biotechnology journal* **15**: 306–317.
- 442 **Boch J, Bonas U. 2010.** *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of*
443 *Phytopathology* **48**: 419–36.
- 444 **Boch J, Bonas U, Lahaye T. 2014.** TAL effectors - pathogen strategies and plant resistance engineering. *New Phytologist*
445 **204**: 823–832.
- 446 **Bonfim K, Faria JC, Nogueira EOPL, Mendes E a, Aragão FJL, Mendes É a, Aragão FJL. 2007.** RNAi-Mediated
447 Resistance to Bean golden mosaic virus in Genetically Engineered Common Bean (*Phaseolus vulgaris*). *Molecular plant-*
448 *microbe interactions : MPMI* **20**: 717–726.
- 449 **Bortesi L, Fischer R. 2015.** The CRISPR / Cas9 system for plant genome editing and beyond. *Biotechnology Advances*
450 **33**: 41–52.
- 451 **Boureau T, Kerkoud M, Chhel F, Hunault G, Darrasse A, Brin C, Durand K, Hajri A, Poussier S, Manceau C, et al.**
452 **2013.** A multiplex-PCR assay for identification of the quarantine plant pathogen *Xanthomonas axonopodis* pv. *phaseoli*.
453 *Journal of Microbiological Methods* **92**: 42–50.
- 454 **Lo Cantore P, Shanmugaiah V, Iacobellis NS. 2009.** Antibacterial activity of essential oil components and their
455 potential use in seed disinfection. *Journal of agricultural and food chemistry* **57**: 9454–61.
- 456 **Carvalho JLV, de Oliveira Santos J, Conte C, Pacheco S, Nogueira EOPL, Souza TLPO, Faria JC, Aragão FJL. 2015.**
457 Comparative analysis of nutritional compositions of transgenic RNAi-mediated virus-resistant bean (event EMB-
458 PV051-1) with its non-transgenic counterpart. *Transgenic Research* **24**: 813–819.
- 459 **Castro Guerrero NA, Isidra-Arellano MC, Mendoza-Cozatl DG, Valdes-Lopez O. 2016.** Common bean: a legume
460 model on the rise for unraveling responses and adaptations to iron, zinc and phosphate deficiencies. *Frontiers in Plant*
461 *Science* **7**: 1–7.
- 462 **Chacón S. MI, Pickersgill B, Debouck DG, Arias JS. 2007.** Phylogeographic analysis of the chloroplast DNA variation
463 in wild common bean (*Phaseolus vulgaris* L.) in the Americas. *Plant Systematics and Evolution* **266**: 175–195.
- 464 **Chacón S MI, Pickersgill B, Debouck DG. 2005.** Domestication patterns in common bean (*Phaseolus vulgaris* L.) and
465 the origin of the Mesoamerican and Andean cultivated races. *Theoretical and Applied Genetics* **110**: 432–444.
- 466 **Charpentier M, Oldroyd G. 2010.** How close are we to nitrogen-fixing cereals? *Current Opinion in Plant Biology* **13**:
467 556–564.
- 468 **Christenhusz MJM, Byng JW. 2016.** The number of known plants species in the world and its annual increase.
469 *Phytotaxa* **261**: 201–217.
- 470 **Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, Grusak MA, Kamfwa K, Katuuramu DN, McClean P,**
471 **et al. 2015.** A *Phaseolus vulgaris* diversity panel for andean bean improvement. *Crop Science* **55**: 2149–2160.
- 472 **Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B. 2016.** Genetic
473 characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the
474 *X. axonopodis* species complex. *Plant Pathology* **65**: 792–806.
- 475 **Corey RR, Starr MP. 1957.** Colony types of *Xanthomonas phaseoli*. *Journal of Bacteriology* **74**: 137–140.
- 476 **Cronk Q, Ojeda I, Pennington RT. 2006.** Legume comparative genomics: Progress in phylogenetics and
477 phylogenomics. *Current Opinion in Plant Biology* **9**: 99–103.
- 478 **Darrasse A, Bureau C, Samson R, Morris CE, Jacques MA. 2007.** Contamination of bean seeds by *Xanthomonas*
479 *axonopodis* pv. *phaseoli* associated with low bacterial densities in the phyllosphere under field and greenhouse
480 conditions. *European Journal of Plant Pathology* **119**: 203–215.

- 481 **Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, Briand M, Brin C, Cociancich S, Durand**
 482 **K, et al. 2013.** Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is
 483 not a general feature of xanthomonads. *BMC Genomics* **14**: 761.
- 484 **Darrasse A, Darsonval A, Boureau T, Brisset MN, Durand K, Jacques MA. 2010.** Transmission of plant-pathogenic
 485 bacteria by nonhost seeds without induction of an associated defense reaction at emergence. *Applied and Environmental*
 486 *Microbiology* **76**: 6787–6796.
- 487 **Darsonval A, Darrasse A, Durand K, Bureau C, Cesbron S, Jacques M-A. 2009.** Adhesion and fitness in the bean
 488 phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Molecular plant-microbe interactions* :
 489 *MPMI* **22**: 747–757.
- 490 **Darsonval A, Darrasse A, Meyer D, Demarty M, Durand K, Bureau C, Manceau C, Jacques M. 2008.** The Type III
 491 Secretion System of *Xanthomonas fuscans* subsp. *fuscans* Is Involved in the Phyllosphere Colonization Process and in
 492 Transmission to Seeds of Susceptible Beans ☐. *Applied and environmental microbiology* **74**: 2669–2678.
- 493 **Dowson DW. 1939.** On the systematic position and generic names of the Gram negative bacterial plant pathogens.
 494 *Zentralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten, 2* **100**: 177–193.
- 495 **Duncan RW, Singh SP, Gilbertson RL. 2011.** Interaction of common bacterial blight bacteria with disease resistance
 496 quantitative trait Loci in common bean. *Phytopathology* **101**: 425–435.
- 497 **Dye DW, Bradbury JF, Goto M, Hayward AC, Lelliott RA, Schroth MN. 1980.** International standards for naming
 498 pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Review of Plant Pathology*
 499 **59**: 153–159.
- 500 **EFSA PLH Panel (EFSA Panel on Plant Health). 2014.** Scientific Opinion on the pest categorisation of *Xanthomonas*
 501 *axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *EFSA Journal* **12**: 3856.
- 502 **EPPO/CABI. 1986.** *Xanthomonas axonopodis* pv. *phaseoli*. *EPPO quarantine pest*.
- 503 **Fininsa C. 1996.** Effect of intercropping bean with maize on bean common bacterial blight and rust diseases.
 504 *International Journal of Pest Management* **42**: 51–54.
- 505 **Fininsa C. 2003.** Relationship between common bacterial blight severity and bean yield loss in pure stand and bean-
 506 maize intercropping systems. *International Journal of Pest Management* **49**: 177–185.
- 507 **Fininsa C, Tefera T. 2001.** Effect of primary inoculum sources of bean common bacterial blight on early epidemics,
 508 seed yield and quality aspects. *International Journal of Pest Management* **47**: 221–225.
- 509 **Foyer CH, Lam H-M, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA,**
 510 **Hodgson JM, et al. 2016.** Neglecting legumes has compromised human health and sustainable food production. *Nature*
 511 *Plants* **2**: 16112.
- 512 **Gakuubi MM, Wagacha JM, Dossaji SF, Wanzala W. 2016.** Chemical Composition and Antibacterial Activity of
 513 Essential Oils of *Tagetes minuta* (Asteraceae) against Selected Plant Pathogenic Bacteria. *International Journal of*
 514 *Microbiology* **2016**.
- 515 **Gaut BS. 2014.** The complex domestication history of the common bean. *Nature genetics* **46**: 663–4.
- 516 **Goodwin PH, Sopher CR. 1994.** Water stress in leaves of *Phaseolus vulgaris* infected with *Xanthomonas campestris*
 517 pv. *phaseoli*. *Journal of Phytopathology* **140**: 219–226.
- 518 **Graham PH, Ranalli P. 1997.** Common bean (*Phaseolus vulgaris* L.). *Field Crops Research* **53**: 131–146.
- 519 **Grimault V, Olivier V, Rolland M, Darrasse A, Jacques MA. 2014.** Seed health testing methods. 7-021: Detection of
 520 *Xanthomonas axonopodis* pv. *phaseoli* on *Phaseolus vulgaris*. *ISTA International rules for seed testing. Annexe to Chapter*
 521 *7: Seed health methods 7-021. International Seed Testing Association, Basserdorf, Switzerland, 1-20*.
- 522 **Guan SH, Gris C, Cruveiller S, Pouzet C, Tasse L, Leru A, Maillard A, Médigue C, Batut J, Masson-Boivin C, et al.**
 523 **2013.** Experimental evolution of nodule intracellular infection in legume symbionts. *The ISME Journal* **7**: 1367–1377.
- 524 **Hajri A, Brin C, Hunault G, Lardeux F, Lemaire C, Manceau C, Boureau T, 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.
- Hane JK, Ming Y, Kamphuis LG, Nelson MN, Garg G, Atkins CA, Bayer PE, Bravo A, Bringans S, Cannon S, et al. 2017.** A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. *Plant Biotechnology Journal* **15**: 318–330.
- Hayat I, Ahmad A, Masud T, Ahmed A, Bashir S. 2014.** Nutritional and health perspectives of beans (*Phaseolus vulgaris* L.): an overview. *Critical reviews in food science and nutrition* **54**: 580–92.
- Hayward AC. 1993.** The hosts of *Xanthomonas*. *Xanthomonas*.1–119.
- He S, Simpson BK, Sun H, Ngadi MO, Ma Y, Huang T. 2015.** *Phaseolus vulgaris* Lectins: A Systematic Review of Characteristics and Health Implications. *Critical reviews in food science and nutrition* **19**.
- Helmstädter A. 2010.** Beans and diabetes: *Phaseolus vulgaris* preparations as antihyperglycemic agents. *Journal of medicinal food* **13**: 251–4.
- Hillocks RJ, Madata CS, Chirwa R, Minja EM, Msolla S. 2006.** *Phaseolus* bean improvement in Tanzania, 1959-2005. *Euphytica* **150**: 215–231.
- Hnatuszko-Konka K, Kowalczyk T, Gerszberg A, Wiktorek-Smagur A, Kononowicz AK. 2014.** *Phaseolus vulgaris* - Recalcitrant potential. *Biotechnology Advances* **32**: 1205–1215.
- Hsu PD, Lander ES, Zhang F. 2014.** Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**: 1262–1278.
- Hummel AW, Doyle EL, Bogdanove AJ. 2012.** Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**: 883–893.
- Hummel AW, Wilkins KE, Wang L, Cernadas RA, Bogdanove AJ. 2016.** A transcription activator-like effector from *Xanthomonas oryzae* pv. *oryzicola* elicits dose-dependent resistance in rice. *Molecular plant pathology* **18**: 55–56.
- Hutin M, Césari S, Chalvon V, Michel C, Tran TT, Boch J, Koebnik R, Szurek B, Kroj T. 2016.** Ectopic activation of the rice NLR heteropair RGA4/RGA5 confers resistance to bacterial blight and bacterial leaf streak diseases. *Plant Journal* **88**: 43–55.
- Indiana A, Briand M, Arlat M, Gagnevin L, Koebnik R, Noël L, Portier P, Darrasse A, Jacques MA. 2014.** Draft genome sequence of the flagellated *Xanthomonas fuscans* subsp. *fuscans* strain CFBP 4884. *Genome announcements* **2**: e00966-14.
- Islam FMA, Renjifo J, Redden RJ, Basford KE, Beebe S. 2003.** Association Between Seed Coat Polyphenolics (Tannins) and Disease Resistance in Common Bean. *Plant Foods for Human Nutrition* **58**: 285–297.
- Ištvánek J, Jaroš M, Křenek A, Řepková J. 2014.** Genome assembly and annotation for red clover (*Trifolium pratense*; Fabaceae). *American journal of botany* **101**: 327–37.
- Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al. 2016.** Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas*: French Network on *Xanthomonads* (FNX). *Annual Review of Phytopathology* **54**: 163–187.
- Jacques M-A, Josi K, Darrasse A, Samson R. 2005.** *Xanthomonas axonopodis* pv. *phaseoli* ar. *fuscans* Is Aggregated in Stable Biofilm Population Sizes in the Phyllosphere of Field- Grown Beans. *Applied and Environmental Microbiology* **71**: 2008–2015.
- Jones AL. 1999.** PHASEOLUS BEAN: Post-harvest Operations. *AGSI/FAO*.
- Kaiser WJ, Vakili NG. 1978.** Insect transmission of pathogenic *Xanthomonads* to bean and cowpea in Puerto Rico. *Phytopathology* **68**: 1057–1063.
- Katungi E, Sperling L, Karanja D, Farrow A, Beebe S. 2011.** Relative importance of common bean attributes and variety demand in the drought areas of Kenya. *Journal of Development and Agricultural Economics* **3**: 411–422.

- 569 **Kelemu S, Cardona C, Segura G. 2004.** Antimicrobial and insecticidal protein isolated from seeds of *Clitoria ternatea*,
570 a tropical forage legume. *Plant Physiology and Biochemistry* **42**: 867–873.
- 571 **Lareo LR, González V F. 1988.** Acceptability and nutritional quality of common beans (*Phaseolus vulgaris* L.): A
572 bibliography. *Cali Colombia: Centro Internacional de Agricultura Tropical*: 1–279.
- 573 **Li C, Wei J, Lin Y, Chen H. 2012.** Gene silencing using the recessive rice bacterial blight resistance gene *xa13* as a new
574 paradigm in plant breeding. *Plant Cell Reports* **31**: 851–862.
- 575 **Liu SY, Yu K, Huffner M, Park SJ, Banik M, Pauls KP, Crosby W. 2010.** Construction of a BAC library and a physical
576 map of a major QTL for CBB resistance of common bean (*Phaseolus vulgaris* L.). *Genetica* **138**: 709–716.
- 577 **De Lúque JJR, Creamer B. 2014.** Major constraints and trends for common bean production and commercialization ;
578 establishing priorities for future research. *Agronomía Colombiana* **32**: 423–431.
- 579 **Mageshwaran V, Walia S, Annapurna K. 2012.** Isolation and partial characterization of antibacterial lipopeptide
580 produced by *Paenibacillus polymyxa* HKA-15 against phytopathogen *Xanthomonas campestris* pv. *phaseoli* M-5. *World*
581 *Journal of Microbiology and Biotechnology* **28**: 909–917.
- 582 **Mayer L, da Silva WP, Moura AB, Vendruscolo CT. 2010.** AFLP analysis of *Xanthomonas axonopodis* and *X. arboricola*
583 strains used in xanthan production studies reveal high levels of polymorphism. *Brazilian Journal of Microbiology* **41**:
584 741–748.
- 585 **Messina V. 2014.** Nutritional and health benefits of dried beans. *The American Journal of Clinical Nutrition* **100**: 437–
586 42.
- 587 **Meziadi C, Richard MMS, Derquennes A, Thareau V, Blanchet S, Gratias A, Pflieger S, Geffroy V. 2016.**
588 Development of molecular markers linked to disease resistance genes in common bean based on whole genome
589 sequence. *Plant Science* **242**: 351–357.
- 590 **Mhedbi-Hajri N, Darrasse A, Pigné S, Durand K, Fouteau S, Barbe V, Manceau C, Lemaire C, Jacques M-A. 2011.**
591 Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. *BMC evolutionary biology* **11**: 67.
- 592 **Miklas PN, Coyne DP, Grafton KF, Mutlu N, Reiser J, Lindgren DT, Singh SP. 2003.** A major QTL for common bacterial
593 resistance derives from the common bean great northern landrace cultivar Montana No. 5. *Euphytica* **131**: 137–146.
- 594 **Miklas PN, Kelly JD, Beebe SE, Blair MW. 2006.** Common bean breeding for resistance against biotic and abiotic
595 stresses: From classical to MAS breeding. *Euphytica* **147**: 105–131.
- 596 **Mora Y, Díaz R, Vargas-Lagunas C, Peralta H, Guerrero G, Aguilar A, Encarnación S, Girard L, Mora J. 2014.**
597 Nitrogen-fixing rhizobial strains isolated from common bean seeds: Phylogeny, physiology, and genome analysis.
598 *Applied and Environmental Microbiology* **80**: 5644–5654.
- 599 **Mus F, Crook MB, Garcia K, Costas AG, Geddes BA, Kouri ED, Paramasivan P, Ryu M-H, Oldroyd GED, Poole PS, et**
600 **al. 2016.** Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental*
601 *Microbiology* **82**: 3698–3710.
- 602 **Mutlu N, Vidaver AK, Coyne DP, Steadman JR, Lambrecht PA, Reiser J. 2008.** Differential Pathogenicity of
603 *Xanthomonas campestris* pv. *phaseoli* and *X. fuscans* subsp. *fuscans* Strains on Bean Genotypes with Common Blight
604 Resistance. *Plant Disease* **92**: 546–554.
- 605 **Nascimento V do, Arf O, Tarsitano MAA, Garcia NFS, Penteado M de S, Sabundjian MT. 2016.** Economic analysis of
606 rice and common bean production in succession to green manure crops and mechanical soil decompression in Brazilian
607 Savannah. *Revista Ceres* **63**: 315–322.
- 608 **Ndakidemi PA, Dakora FD, Nkonya EM, Ringo D, Mansoor H. 2006.** Yield and economic benefits of common bean
609 (*Phaseolus vulgaris*) and soybean (*Glycine max*) inoculation in northern Tanzania. *Australian Journal of Experimental*
610 *Agriculture* **46**: 571–577.
- 611 **Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. 2014.** Crystal
612 Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**: 935–949.

- 613 **Oldroyd GED, Murray JD, Poole PS, Downie JA. 2011.** The Rules of Engagement in the Legume-Rhizobial Symbiosis.
614 *Annual Review of Genetics* **45**: 119–144.
- 615 **Opio AF, Allen DJ, Teri JM. 1996.** Pathogenic variation in *Xanthomonas campestris* pv. *phaseoli*, the causal agent of
616 common bacterial blight in *Phaseolus* beans. *Plant Pathology (United Kingdom)* **45**: 1126–1133.
- 617 **Opio F, Kimani PM, Musaana SM, Buruchara R. 2002.** Research on common blight and halo blight of common beans
618 in east and central Africa. *Bean Improvement Cooperative. Annual Report (USA)*. **45**: 160–161.
- 619 **Opio A., Teri J., Allen D. 1993.** Studies on Seed Transmission of *Xanthomonas campestris* pv *phaseoli* in Common Beans
620 in Uganda. *African Crop Science Journal* **1**: 59–67.
- 621 **Ormeño-Orrillo E, Servín-Garcidueñas LE, Rogel MA, González V, Peralta H, Mora J, Martínez-Romero J,
622 Martínez-Romero E. 2015.** Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics.
623 *Systematic and Applied Microbiology* **38**: 287–291.
- 624 **Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D. 2007.** Phylogenetic analysis of *Xanthomonas* species by
625 comparison of partial gyrase B gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **57**:
626 2881–2887.
- 627 **Parkinson N, Cowie C, Heeney J, Stead D. 2009.** Phylogenetic structure of *Xanthomonas* determined by comparison
628 of *gyrB* sequences. *International Journal of Systematic and Evolutionary Microbiology* **59**: 264–274.
- 629 **Perry G, Dinatale C, Xie W, Navabi A, Reinprecht Y, Crosby W, Yu K, Shi C, Pauls KP. 2013.** A comparison of the
630 molecular organization of genomic regions associated with resistance to common bacterial blight in two *Phaseolus*
631 *vulgaris* genotypes. *Frontiers in plant science* **4**: 318.
- 632 **Rademaker JLW, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J, de Bruijn FJ. 2005.** A comprehensive
633 species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* **95**: 1098–1111.
- 634 **Ramírez Rivera NG, García-Salinas C, Aragão FJL, Díaz de la Garza RI. 2016.** Metabolic engineering of folate and its
635 precursors in Mexican common bean (*Phaseolus vulgaris* L.). *Plant Biotechnology Journal* **14**: 2021–2032.
- 636 **Ran FA, Hsu PD, Lin C, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S. 2013.** Double
637 Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**: 1380–1389.
- 638 **Rodríguez-R LM, Grajales A, Arrieta-Ortiz M, Salazar C, Restrepo S, Bernal A. 2012.** Genomes-based phylogeny of
639 the genus *Xanthomonas*. *BMC Microbiology* **12**: 1–14.
- 640 **Romero-Arenas O, Damián Huato MA, Rivera Tapia JA, Báez SA, Huerta LM, Huerta, Cabrera E. 2013.** The
641 Nutritional value of Beans (*Phaseolus vulgaris* L.) and its importance for Feeding of Rural communities in Puebla-
642 Mexico. *International Research Journal of Biological Sciences* **2**: 59–65.
- 643 **Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017.** *Xanthomonas* adaptation to common bean is associated
644 with horizontal transfers of genes encoding TAL effectors. *BMC Genomics* **18**: 670.
- 645 **Santalla M, Rodiño AP, De Ron AM. 2002.** Allozyme evidence supporting southwestern Europe as a secondary center
646 of genetic diversity for the common bean. *Theoretical and Applied Genetics* **104**: 934–944.
- 647 **Santalla M, de Ron AM, de La Fuente M. 2010.** Integration of genome and phenotypic scanning gives evidence of
648 genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe.
649 *Theoretical and Applied Genetics* **120**: 1635–1651.
- 650 **Santos AS, Bressan-Smith RE, Pereira MG, Rodrigues R, Ferreira CF. 2003.** Genetic linkage map of *Phaseolus*
651 *vulgaris* and identification of QTLs responsible for resistance to *Xanthomonas axonopodis* pv. *phaseoli*. *Fitopatologia*
652 *Brasileira* **28**: 5–10.
- 653 **Sati SC, Joshi S. 2011.** Antibacterial activities of *Ginkgo biloba* L. leaf extracts. *TheScientificWorldJournal* **11**: 2237–42.
- 654 **Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, et al.**
655 **2008.** Genome structure of the legume, *Lotus japonicus*. *DNA Research* **15**: 227–239.
- 656 **Schaad NW, Postnikova E, Lacy GH, Sechler A, Agarkova I, Stromberg PE, Stromberg VK, Vidaver AK. 2005.**

- 657 Reclassification of *Xanthomonas campestris* pv. *citri* (ex Hasse 1915) Dye 1978 forms A, B/C/D, and E as *X. smithii*
 658 subsp. *citri* (ex Hasse) sp. nov. nom. rev. comb. nov., *X. fuscans* subsp. *aurantifolii* (ex Gabriel 1989) sp. nov. nom. rev.
 659 comb. nov., and *X. Systematic and applied microbiology* **28**: 494–518.
- 660 **Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, et al. 2010.**
 661 Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178–83.
- 662 **Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C, et al.**
 663 **2014.** A reference genome for common bean and genome-wide analysis of dual domestications. *Nature genetics* **46**:
 664 707–13.
- 665 **Schornack S, Moscou MJ, Ward ER, Horvath DM. 2013.** Engineering Plant Disease Resistance Based on TAL Effectors.
 666 *Annual Review of Phytopathology* **51**: 383–406.
- 667 **Schuster ML, Coyne DP. 1974.** Survival Mechanisms of Phytopathogenic Bacteria. *Annual Review of Phytopathology*
 668 **12**: 199–221.
- 669 **Schuster ML, Coyne DP, Hoff B. 1973.** Comparative Virulence of *Xanthomonas phaseoli* Strains from Uganda,
 670 Colombia, and Nebraska. *Plant Disease Reporter* **57**: 74–75.
- 671 **Shade A, Jacques MA, Barret M. 2017.** Ecological patterns of seed microbiome diversity, transmission, and assembly.
 672 *Current Opinion in Microbiology* **37**: 15–22.
- 673 **Shi C, Chaudhary S, Yu K, Park SJ, Navabi A, McClean PE. 2011a.** Identification of candidate genes associated with
 674 CBB resistance in common bean HR45 (*Phaseolus vulgaris* L.) using cDNA-AFLP. *Molecular Biology Reports* **38**: 75–81.
- 675 **Shi C, Navabi A, Yu K. 2011b.** Association mapping of common bacterial blight resistance QTL in Ontario bean breeding
 676 populations. *BMC Plant Biology* **11**: 1–11.
- 677 **Shi C, Yu K, Xie W, Perry G, Navabi A, Pauls KP, Miklas PN, Fourie D. 2012.** Development of candidate gene markers
 678 associated to common bacterial blight resistance in common bean. *Theoretical and Applied Genetic* **125**: 1525–37.
- 679 **Singh SP. 2001.** Broadening the genetic base of common bean cultivars: A review. *Crop Science* **41**: 1659–1675.
- 680 **Singh SP, Muñoz CG. 1999.** Resistance to Common Bacterial Blight among *Phaseolus* Species and Common Bean
 681 Improvement. *Crop Science* **39**: 80–89.
- 682 **Singh SP, Schwartz HF. 2010.** Breeding common bean for resistance to diseases: A review. *Crop Science* **50**: 2199–
 683 2223.
- 684 **Svetleva D, Velcheva M, Bhowmik G. 2003.** Biotechnology as a useful tool in common bean (*Phaseolus vulgaris* L.)
 685 improvement. *Euphytica* **131**: 189–200.
- 686 **Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, et**
 687 **al. 2014.** An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC genomics* **15**:
 688 312.
- 689 **Todorović B, Potočnik I, Rekanović E, Stepanović M, Kostić M, Ristić M, Milijašević-Marčić S. 2016.** Toxicity of
 690 twenty-two plant essential oils against pathogenic bacteria of vegetables and mushrooms. *Journal of environmental*
 691 *science and health, Part B* **51**: 832–839.
- 692 **Toth IK, Hyman LJ, Taylor R, Birch PRJ. 1998.** PCR-based detection of *Xanthomonas campestris* pv. *phaseoli* var.
 693 *fuscans* in plant material and its differentiation from *X. c. pv. phaseoli*. *Journal of Applied Microbiology* **85**: 327–336.
- 694 **Vandemark GJ, Fourie D, Miklas PN. 2008.** Genotyping with real-time PCR reveals recessive epistasis between
 695 independent QTL conferring resistance to common bacterial blight in dry bean. *Theoretical and Applied Genetics* **117**:
 696 513–522.
- 697 **Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe a G, Cannon S, Baek J, Rosen BD, Tar'an B, et al. 2013.**
 698 Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nature Biotechnology*
 699 **31**: 240–246.
- 700 **Vauterin L, Hoste B, Kersters K, Swings J. 1995.** Reclassification of *Xanthomonas*. *International Journal of Systematic*

- 701 *Bacteriology* **45**: 472–489.
- 702 **Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, Qin T, Li Y, Che J, Zhang M, et al. 2015.** XA23 Is an executor r protein
703 and confers broad-spectrum disease resistance in rice. *Molecular Plant* **8**: 290–302.
- 704 **Wortmann CS, Kirkby RA, Eledu CA, Allen DJ. 1998.** Atlas of common bean (*Phaseolus vulgaris* L.) production in
705 Africa. *Centro Internacional de Agricultura Tropical (CIAT)* **297**: 133.
- 706 **Yang K, Tian Z, Chen C, Luo L, Zhao B, Wang Z, Yu L, Li Y, Sun Y, Li W, et al. 2015.** Genome sequencing of adzuki bean
707 (*Vigna angularis*) provides insight into high starch and low fat accumulation and domestication. *Proceedings of the*
708 *National Academy of Sciences of the United States of America* **112**: 13213–8.
- 709 **Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof**
710 **H, et al. 2011.** The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* **480**: 520–4.
- 711 **Young JM, Park DC, Shearman HM, Fargier E. 2008.** A multilocus sequence analysis of the genus *Xanthomonas*.
712 *Systematic and Applied Microbiology* **31**: 366–377.
- 713 **Yu K, Park SJ, Poysa V. 2000.** Marker-assisted selection of common beans for resistance to common bacterial blight:
714 Efficacy and economics. *Plant Breeding* **119**: 411–415.
- 715 **Zahran HH. 1999.** Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate.
716 *Microbiology and molecular biology reviews : MMBR* **63**: 968–89.
- 717 **Zeng X, Tian D, Gu K, Zhou Z, Yang X, Luo Y, White FF, Yin Z. 2015.** Genetic engineering of the Xa10 promoter for
718 broad-spectrum and durable resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant Biotechnology Journal* **13**: 993–1001.

III. Projet de thèse

La spécificité d'hôte se dessine dès les étapes précoces de l'interaction hôte-pathogène, mais aussi lors de l'étape plus tardive de multiplication et d'expression des symptômes dépendant du SST3 et de ses effecteurs. Les quatre lignées génétiques de *Xanthomonas* responsables de la graise commune du haricot possèdent des répertoires d'ET3 similaires mais non identiques. Ceci souligne bien que la spécificité d'hôte ne repose pas sur un gène isolé ou même sur un répertoire d'une famille de gènes donnée, mais plutôt sur une combinaison d'une collection de gènes avec diverses fonctions. Récemment, il a été démontré que plus de 100 gènes ont récemment été transférés entre les lignées NF1 et fuscans et sont de ce fait de potentiels déterminants de la spécificité d'hôte et de la convergence pathologique de ces deux lignées phylogénétiquement distantes.

Mon équipe d'accueil s'est intéressé à cette question de la spécificité d'hôte des souches infectant le haricot et, précédemment à ma thèse, des analyses de génomique comparative sans a priori ont été réalisées pour trouver des gènes ayant été transmis horizontalement entre ces souches polyphylétiques infectant le haricot. Parmi ces gènes, certains sont conservés chez toutes les souches capables d'infecter le haricot, et absents chez celles n'ayant pas le haricot comme hôte. Ces gènes sont donc également de potentiels déterminants de la spécificité d'hôte entre le haricot et *Xcf-Xpp*. Le premier objectif de ma thèse était donc d'approfondir les analyses réalisées sur ces gènes candidats pour la spécificité d'hôte. D'autre part, bien qu'absents des analyses précédemment réalisées, les gènes codant des effecteurs TAL semblent être aussi de bons candidats pour la spécificité d'hôte. En effet, d'autres analyses préliminaires de mon équipe ont révélé que les souches infectant le haricot possèdent ce type de gène et ont mis en évidence un transfert horizontal d'un variant spécifique de *tal* entre les lignées NF1 et fuscans. Le second objectif de ma thèse était donc d'effectuer une étude approfondie des effecteurs TAL de ces souches en mettant l'accent sur leur implication dans la spécificité d'hôte. Ainsi, ces deux objectifs m'ont permis de répondre à la question de recherche suivante :

QUELS SONT LES GENES IMPLIQUES DANS LA SPECIFICITE D'HOTE DES AGENTS RESPONSABLE DE LA GRAISSE COMMUNE DU HARICOT ET QUEL EST LE ROLE DES EFFECTEURS TAL DANS L'INTERACTION HARICOT / *XANTHOMONAS* ?

A présent, je vais développer les résultats obtenus de mon projet de thèse en trois chapitres. Le premier chapitre est présenté sous la forme d'un article consacré à l'ensemble des analyses de génomiques comparatives effectuées pour identifier les gènes candidats pour la spécificité d'hôte des souches infectant le haricot. Le second chapitre comporte un article de recherche traitant de la diversité et de l'histoire évolutive des gènes *tal* des souches de *Xcf-Xpp* et de leur potentielle implication dans l'adaptation au haricot, ainsi que la publication du génome

complet de trois souches de *X. citri* pv. *vignicola* où l'on évoque notamment la diversité des gènes *tal* portés par ces souches. Le troisième chapitre est présenté sous la forme d'une ébauche d'article dédié à l'analyse globale d'un effecteur TAL partagé entre les lignées NF1 et fuscans. Cet article présente notamment des résultats obtenus au cours d'une mobilité de trois mois en Allemagne dans l'équipe de Jens Boch (Leibniz Universität Hanover), et qui ont notamment permis de caractériser la spécificité de liaison d'un RVD et l'impact d'un motif atypique et spécifique d'un effecteur TAL des souches de *Xcf-Xpp*.

CHAPITRE 1 :

La spécificité d'hôte chez les agents responsables de la graisse commune du haricot

I. Introduction

Les bactéries du genre *Xanthomonas* sont responsables de maladies sur un large éventail de plantes. Cependant, individuellement, chaque souche n'infecte généralement que quelques espèces végétales, ce qui sous-entend qu'elles sont très adaptées à leurs hôtes. Ainsi, ce genre bactérien est un modèle de choix pour l'étude de la spécificité de l'hôte. Les agents responsables de la CBB sont particulièrement intéressants pour ce type d'étude car ils sont répartis en quatre lignées génétiques appartenant à deux espèces distinctes du genre *Xanthomonas*, la lignée NF1 appartient à *X. phaseoli* pv. *phaseoli*, les lignées NF2, NF3 et fuscans appartiennent quant à elles à *X. citri* pv. *fuscans*. Par conséquent, bien qu'ayant apparemment la même spécificité d'hôte, les bactéries infectant le haricot commun forment un groupe polyphylétique, suggérant une convergence pathologique, probablement grâce à l'acquisition de déterminants communs de la spécificité d'hôte.

Des études antérieures ont montré que les récepteurs chimiotactiques, les adhésines, les régulateurs transcriptionnels ou encore les effecteurs de type III jouaient un rôle clé dans la spécificité de l'hôte des bactéries du genre *Xanthomonas* (Hajri et al., 2009; Mhedbi-Hajri et al., 2011; Schwartz et al., 2015). Par ailleurs, nous savons que le transfert horizontal est une force motrice majeure de l'adaptation écologique et de la diversification des bactéries (El Yacoubi et al., 2007; Wiedenbeck & Cohan, 2011; Darmon & Leach, 2014; Jacques et al., 2016). Plus récemment, une étude a apporté les premiers éléments montrant l'impact de la recombinaison hétérologue adaptative chez des bactéries phylogénétiquement éloignées infectant le haricot commun. En effet, elle a démontré que la convergence pathologique entre les lignées de NF1 et fuscans était associée à de grands événements de transfert horizontaux impliquant une centaine de gènes (Aritua et al., 2015).

Pour rechercher les déterminants potentiels de la spécificité d'hôte des agents responsables de la CBB, nous avons mené une étude de génomique comparative avec 72 génomes représentant la diversité des pathogènes de *Xanthomonas*. Pour cette analyse, nous avons associé plusieurs approches afin de retenir les gènes communs aux quatre lignées génétiques infectant le haricot commun et d'affiner au maximum la liste de ces gènes candidats. Notamment, nous avons produit une matrice de présence / absence de groupes orthologues afin de déterminer les gènes partagés par différents clades de *Xanthomonas*, nous avons utilisé une approche phylogénétique pour rechercher les gènes pour lesquels les souches de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* formaient un groupe monophylétique, nous avons utilisé une approche basée sur les k-mer pour rechercher les gènes contenant de courtes séquences spécifiques présentes chez toutes les souches responsables de CBB et absentes chez toutes les autres souches de *Xanthomonas* et nous avons aussi rechercher les gènes possédant du polymorphisme non synonyme afin de trouver des

acides aminés spécifiques des souches infectant le haricot commun. D'autre part, nous avons aussi recherché les fonctions de tous les gènes candidats ainsi que les événements de recombinaison et de transferts horizontaux de ces gènes impliqués dans la spécificité d'hôte des agents responsables de la CBB.

La majorité de ces analyses ont été menées dans un premier temps par Laurana Serres-Giardi puis dans un second temps par Nicolas Chen et l'outil SkIf, basé sur la recherche de k-mer, a été développé par Martial Briand. Pour ma part, je me suis chargée de trouver les sites non synonymes spécifiques des souches de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* et de rechercher les fonctions des gènes candidats. L'ensemble de ces résultats a fait l'objet d'une ébauche d'article, présentée ci-après, dont Nicolas Chen est le premier auteur et dont je fais également partie des auteurs en raison des analyses que j'ai mené mais aussi parce que j'ai participé à la correction de cette ébauche. Après quelques corrections supplémentaires, il est prévu de soumettre cet article au journal BMC Genomics.

II. Comparative genomics analyses identified a subset of genes associated with host specificity in *Xanthomonas* strains pathogenic on common bean.

Nicolas W.G. Chen¹, Laurana Serres-Giardi¹, Mylène Ruh¹, Martial Briand¹, Sophie Bonneau¹, Armelle Darrasse¹, Sébastien Carrère², Valérie Barbe³, Lionel Gagnevin^{4,5}, Ralf Koebnik⁵, Marie-Agnès Jacques^{1*}

¹ IRHS, INRA, AGROCAMPUS OUEST, Université d'Angers, SFR4207 QUASAV, 42, rue Georges Morel, 49071 Beaucouzé, France. ² INRA, CNRS, UMR 441/2594 Laboratoire des Interactions Plantes-Microorganismes (LIPM), F-31326 Castanet-Tolosan, France. ³ CEA/DSV/IG/Genoscope, 2 rue Gaston Crémieux, BP5706, 91057 Evry, France. ⁴ CIRAD, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), F-97410 Saint-Pierre, La Réunion, France. ⁵ Institut de Recherche pour le Développement (IRD), UMR IRD-CIRAD-UM2 Interactions Plantes-Microorganismes-Environnement (IPME), F-34394 Montpellier, Cedex 5, France.

* Corresponding author

ABSTRACT

Background: Host specialization is a hallmark of numerous plant pathogens including bacteria, fungi, oomycetes and viruses. Yet, the molecular and evolutionary bases of host specificity are still poorly understood. In some cases, pathological convergence is observed for individuals not belonging to a single monophyletic group. This is the case for *Xanthomonas* strains responsible for common bacterial blight of bean, which are spread across four phylogenetically distant genetic lineages. Every strain belonging to each of these lineages is able to cause similar symptoms on common bean, suggesting that these strains possess a common arsenal of genes conferring the ability to infect common bean.

Results: To search for genes involved in common bean specificity, we used a combination of approaches, including a genome scan based on *k*-mer search. Analysis of 72 genomes from a diversity of *Xanthomonas* pathovars unveiled 116 genes bearing DNA sequences specific of strains responsible for common bacterial blight. Of these 116 genes, 89 were involved in consecutive events of horizontal gene transfers among the four genetic lineages and 45 contained at least one amino acid specific to the causal agents of common bacterial blight.

Conclusions: Our study reveals that host specificity of common bacterial blight agents is associated with a combination of horizontally transferred genes, and highlights the role of plasmids in these horizontal transfers.

Keywords: *Xanthomonas*; common bean; TAL effectors; host adaptation; horizontal gene transfer

BACKGROUND

In nature, most pathogens are generalist, meaning that they are able to infect multiple hosts, while other pathogens are specialists, meaning that they are highly adapted to a single or few host species (Bäumler & Fang, 2013). Understanding the molecular basis of host specificity can provide new insights into the evolution and ecology of specialist pathogens, and their potential to cross the species barrier and to infect new hosts. Bacteria from the genus *Xanthomonas* infect a large diversity of at least 392 plant species including important crops and ornamentals (Hayward, 1993). Yet, each strain individually is able to infect only one or few plant species. Strains able to cause the same symptoms on the same host range are grouped into pathovars (Dye *et al.*, 1980). Adaptation of a given pathovar to its host plant is a complex process possibly involving diverse molecular determinants (Büttner & Bonas, 2010; Jacques *et al.*, 2016). Although our understanding of the molecular basis of host specificity is still limited, chemotactic sensors, adhesins and type III effectors emerge as key determinants for shaping host specificity in *Xanthomonas* (Hajri *et al.*, 2009; Mhedbi-Hajri *et al.*, 2011; Schwartz *et al.*, 2015). Chemotactic sensors enable the bacteria to detect attractant or repellent molecules and trigger flagellar motility towards entry sites of the host plant, while adhesins allow the attachment on the host plant surface and biofilm formation, and type III effectors are delivered into the plant cells where they can have different functions including pathogen-associated molecular pattern triggered immunity (PTI). These different examples show that host specificity spans all the main different steps of host colonization.

The *Xanthomonas axonopodis* species complex (*sensu* Vauterin *et al.*, 1995, 2000) groups more than 30 pathovars infecting a wide diversity of plants including economically important crops and ornamentals such as *Citrus*, *Anthurium* and *Dieffenbachia* species as well as pepper, cotton, citrus, soybean, or common bean. Based on repetitive-sequence-based (rep)-PCR fingerprints, *X. axonopodis* has been subdivided in six subclusters named 9.1 to 9.6 (Rademaker *et al.*, 2005). More recently, this species complex has been split in four different species including *X. citri*, *X. euvesicatoria*, *X. phaseoli* and *X. axonopodis* (*sensu* Constantin *et al.*, 2016). Common bacterial blight of bean (CBB) is the most devastating bacterial disease infecting common bean

Table 1. Information on the sequenced strains used in this study.

Identifier	Nomenclature	Strain	Host of isolation	Country and date of isolation	Contigs	Total size	% GC	Assembly accession number	Reference
stemma5513	<i>Stenotrophomonas maltophilia</i>	R551-3	<i>Populus trichocarpa</i>	na	1	4,573,969	66.29	GCA_000020665.1	Taghavi <i>et al.</i> , 2009
albiliPC73	<i>X. albilineans</i>	GPE PC73	<i>Saccharum</i> spp. cv. H63-1418	Guadeloupe (France), 2003	1	3,768,695	62.97	GCA_000087965.1	Pieretti <i>et al.</i> , 2009
aberra6865	<i>X. campestris</i> pv. <i>campestris</i>	CFBP6865	<i>Brassica oleracea</i> var. <i>capitata</i>	Australie, 1975	4	5,169,518	64.93	Unpublished	Noël, comm. pers
barbar5825	<i>X. campestris</i> pv. <i>barbareae</i>	CFBP5825R	<i>Barbarea vulgaris</i>	USA, 1939	2	5,055,186	65.09	ATNQ00000000	Roux <i>et al.</i> , 2015
campes8004	<i>X. campestris</i> pv. <i>campestris</i>	8004	<i>Brassica oleracea</i> var. <i>botrytis</i>	Royaume-Uni, 1958	1	5,148,708	64.95	GCA_000012105.1	Qian <i>et al.</i> , 2005
campesAT33	<i>X. campestris</i> pv. <i>campestris</i>	ATCC 33913	<i>Brassica oleracea</i> var. <i>gemmifera</i>	Royaume-Uni, 1957	1	5,076,188	65.06	GCA_000007145.1	da Silva <i>et al.</i> , 2002
campesB100	<i>X. campestris</i> pv. <i>campestris</i>	B100	na	na	1	5,079,002	65.04	GCA_000070605.1	Vorhölter <i>et al.</i> , 2008
incana2527	<i>X. campestris</i> pv. <i>incanae</i>	CFBP2527R	<i>Matthiola incana</i>	USA, 1950	1	4,926,205	65.12	ATNQ00000000	Roux <i>et al.</i> , 2015
incana1606	<i>X. campestris</i> pv. <i>incanae</i>	CFBP1606R	<i>Matthiola incana</i>	France, 1974	1	4,967,651	65.17	ATNQ00000000	Roux <i>et al.</i> , 2015
musace4381	<i>X. campestris</i> pv. <i>musacearum</i>	NCPBP4381	<i>Musa</i> sp.	Uganda, 2005	115	4,810,038	61.96	ACHT00000000	Studholme <i>et al.</i> , 2010
raphan5828	<i>X. campestris</i> pv. <i>raphani</i>	CFBP5828R	<i>Raphanus sativus</i>	USA, na	2	4,912,709	65.35	ATNP00000000	Roux <i>et al.</i> , 2015
raphan756C	<i>X. campestris</i> pv. <i>raphani</i>	756C	<i>Brassica oleracea</i> var. <i>capitata</i>	east asia, na	1	4,941,214	65.28	GCA_000221965.1	Bogdanove <i>et al.</i> , 2011
cassav4642	<i>X. cassavae</i>	CFBP4642	<i>Manihot esculenta</i> Crantz	Malawi, 1951	7	5,278,192	65.22	GCA_000454545.1	Bolot <i>et al.</i> , 2013
axanac2913	<i>X. citri</i> pv. <i>anacardii</i>	CFBP2913	<i>Mangifera indica</i>	Brazil, na	1	5,203,496	64.57	CP024057	this study
axaura1035	<i>X. citri</i> pv. <i>aurantifolii</i>	ICPB10535	<i>Citrus aurantifolia</i>	Brazil, na	351	5,060,896	64.81	ACPY00000000	Moreira <i>et al.</i> , 2010
axcitr1083	<i>X. citri</i> pv. <i>citri</i>	FDC1083	<i>Citrus reticulata</i>	Brazil, 1980	1	5,218,314	64.72	CCVZ01000000	Gordon <i>et al.</i> , 2015
axcitr1902	<i>X. citri</i> pv. <i>citri</i>	JF90-2	<i>Citrus aurantifolia</i>	Oman, 1986	1	5,251,225	64.66	CCWA01000000	Gordon <i>et al.</i> , 2015
axcitr1238	<i>X. citri</i> pv. <i>citri</i>	J1238-24	<i>Citrus aurantifolia</i>	Thailand, 1989	1	5,282,622	64.78	CCVX01000000	Gordon <i>et al.</i> , 2015
axcitrLC80	<i>X. citri</i> pv. <i>citri</i>	LC80	<i>Citrus reticulata</i> x <i>C. sinensis</i>	Mali, 2006	1	5,229,127	64.71	CCWJ01000000	Gordon <i>et al.</i> , 2015
axcitrLE20	<i>X. citri</i> pv. <i>citri</i>	LE20-1	<i>Citrus aurantifolia</i>	Ethiopia, 2008	1	5,309,240	64.71	CCWK01000000	Gordon <i>et al.</i> , 2015
axcitr9322	<i>X. citri</i> pv. <i>citri</i>	LMG9322	<i>Citrus aurantifolia</i>	Florida, 1986	1	5,194,131	64.75	CCVY01000000	Gordon <i>et al.</i> , 2015
axcitr306	<i>X. citri</i> pv. <i>citri</i>	306	na	na	1	5,175,554	64.77	GCA_000007165.1	da Silva <i>et al.</i> , 2002
phafus1815	<i>X. citri</i> pv. <i>fuscans</i>	CFBP1815	<i>Phaseolus</i> sp	Greece, 1978	1	4,957,446	64.82	GCA_900234415	this study
phafus4834	<i>X. citri</i> pv. <i>fuscans</i>	CFBP4834-R	<i>Phaseolus vulgaris</i> cv. <i>Michelet</i>	France, 1998	1	4,981,995	64.81	FO681494	Darrasse <i>et al.</i> , 2013b
phafus6166	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6166	<i>Phaseolus vulgaris</i>	South Africa, 1963	1	4,987,587	64.8	GCA_900234455	this study
phafus6960	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6960	<i>Phaseolus vulgaris</i>	Reunion island, France, 2000	1	4,992,131	64.79	GCA_900234515	this study
phafus6970	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6970	<i>Phaseolus</i> sp	USA, 1990	1	5,006,656	64.9	GCA_900234505	this study
pha2GL6988	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6988	<i>Phaseolus vulgaris</i> cv. <i>marla</i>	Reunion island, France, 2000	1	5,132,433	64.64	Deposited	this study
pha2GL6990	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6990	<i>Phaseolus vulgaris</i> cv. <i>marla</i>	Reunion island, France, 2000	1	5,124,653	64.63	GCA_900234485	this study
pha2GL6991	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6991	<i>Phaseolus vulgaris</i> cv. <i>marla</i>	Reunion island, France, 2000	1	5,342,565	64.35	GCA_900234525	this study
pha3GL6992	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6992	<i>Phaseolus vulgaris</i> cv. <i>marla</i>	Reunion island, France, 2000	1	5,517,999	64.48	GCA_900234495	this study
pha3GL6994	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6994	<i>Phaseolus vulgaris</i>	Tanzania, 1990	1	5,250,266	64.57	GCA_900234565	this study
pha3GL6996	<i>X. citri</i> pv. <i>fuscans</i>	CFBP6996	<i>Phaseolus vulgaris</i> cv. <i>marla</i>	Reunion island, France, 2000	4	5,095,420	64.72	Deposited	this study
phafus7766	<i>X. citri</i> pv. <i>fuscans</i>	CFBP7766	<i>Phaseolus vulgaris</i>	Cameroon, 2009	1	5,185,891	64.63	GCA_900234475	this study
phafus7767	<i>X. citri</i> pv. <i>fuscans</i>	CFBP7767	<i>Phaseolus vulgaris</i>	Cameroon, 2009	1	5,107,388	64.7	GCA_900234465	this study
axglyc2526	<i>X. citri</i> pv. <i>glycines</i>	CFBP2526	<i>Glycine hispida</i>	Sudan, 1956	4	5,255,152	64.63	GCA_000495275.1	Darrasse <i>et al.</i> , 2013a
axglyc7119	<i>X. citri</i> pv. <i>glycines</i>	CFBP7119	<i>Glycine max</i>	Brazil, 1981	4	5,521,783	64.39	GCA_000488895.1	Darrasse <i>et al.</i> , 2013a
axmalv1386	<i>X. citri</i> pv. <i>malvacearum</i>	GSPB1386	na	na	127	4,991,411	64.7	GCA_000309905.1	Hainan University
axmalv2388	<i>X. citri</i> pv. <i>malvacearum</i>	GSPB2388	na	na	61	5,127,016	64.53	GCA_000309925.1	Hainan University
axmalvaX18	<i>X. citri</i> pv. <i>malvacearum</i>	X18	<i>Gossypium</i> spp.	Burkina Faso, 1980s	4	4,993,692	64.7	ATMA00000000	Cunnac <i>et al.</i> , 2013
axmalvaX20	<i>X. citri</i> pv. <i>malvacearum</i>	X20	<i>Gossypium</i> spp.	Burkina Faso, 1980s	4	5,219,648	64.49	ATMB00000000	Cunnac <i>et al.</i> , 2013
axmang5610	<i>X. citri</i> pv. <i>mangiferaeindicae</i>	LG56-10	<i>Mangifera indica</i>	na	4	5,281,952	64.56	PEBY00000000	this study
axmang8127	<i>X. citri</i> pv. <i>mangiferaeindicae</i>	LG81-27	<i>Mangifera indica</i>	na	6	5,209,594	64.68	PEBZ00000000	this study
axmangL941	<i>X. citri</i> pv. <i>mangiferaeindicae</i>	LMG941	<i>Mangifera indica</i>	India, 1957	195	5,144,323	64.84	CAHO01000001	Midha <i>et al.</i> , 2012
axalfa3836	<i>X. euvesicatoria</i> pv. <i>alfalfae</i>	CFBP3836	<i>Medicago sativa</i>	Sudan, na	6	5,081,438	64.74	GCA_000488955.1	Jacques <i>et al.</i> , 2013
axalli6369	<i>X. euvesicatoria</i> pv. <i>allii</i>	CFBP6369	<i>Allium cepa</i>	Reunion island, France, 1996	3	5,427,488	64.36	GCA_000730305.1	Gagnevin <i>et al.</i> , 2014
axmeloniF1	<i>X. euvesicatoria</i> pv. <i>citrumelonis</i>	F1	<i>Citrus</i> sp.	FL, USA, 1984	1	4,967,469	64.91	GCA_000225915.1	Jalan <i>et al.</i> , 2011
axeuve8510	<i>X. euvesicatoria</i> pv. <i>euvesicatoria</i>	85-10	na	na	1	5,178,466	64.74	GCA_000009165.1	Thieme <i>et al.</i> , 2005
perfor9118	<i>X. euvesicatoria</i> pv. <i>perforans</i>	91-118	<i>Lycopersicon esculentum</i>	FL, USA, na	291	5,296,241	65.04	AEQW00000000	Potnis <i>et al.</i> , 2011
gardne1965	<i>X. gardneri</i>	ATCC19865	<i>Lycopersicon esculentum</i>	Yugoslavia, 1953	552	5,594,687	63.68	AEQX00000000	Potnis <i>et al.</i> , 2011
carotaM081	<i>X. hortorum</i> pv. <i>carotae</i>	M081	<i>Daucus carota</i>	OR, USA, na	154	5,039,269	63.83	GCA_000505565.1	Kimbel <i>et al.</i> , 2011
oryzaeKA10	<i>X. oryzae</i> pv. <i>oryzae</i>	KACC10331	<i>Oryza sativa</i>	Korea, na	1	4,941,439	63.69	GCA_000007385.1	Lee <i>et al.</i> , 2005
oryzaeMA31	<i>X. oryzae</i> pv. <i>oryzae</i>	MAFF 311018	<i>Oryza sativa</i>	Japan, na	1	4,940,217	63.7	GCA_000010025.1	unpublished
oryzaePX99	<i>X. oryzae</i> pv. <i>oryzae</i>	PXO99A	<i>Oryza sativa</i>	Philippines, na	1	5,240,075	63.63	GCA_000019585.1	Salzberg <i>et al.</i> , 2008
oryzicBA15	<i>X. oryzae</i> pv. <i>oryzicola</i>	BA115	<i>Oryza sativa</i>	Burkina faso, 2009	1	4,315,327	64.16	GCA_002189395.1	unpublished
oryzicBA20	<i>X. oryzae</i> pv. <i>oryzicola</i>	BA120	<i>Oryza sativa</i>	Burkina faso, 2009	1	4,498,643	64.05	GCA_002189435.1	unpublished
oryzicBA21	<i>X. oryzae</i> pv. <i>oryzicola</i>	BA121	<i>Oryza sativa</i>	Burkina faso, 2009	1	4,419,870	64.07	GCA_002189465.1	unpublished
oryzicB256	<i>X. oryzae</i> pv. <i>oryzicola</i>	BLS256	<i>Oryza sativa</i>	Philippines, 1984	1	4,831,739	64.05	AAQN01000001	Bogdanove <i>et al.</i> , 2011
axdieff695	<i>X. phaseoli</i> pv. <i>diffenbachiae</i>	LMG695	<i>Anthurium andreaeanum</i>	Brazil, 1965	1	5,037,357	64.88	GCA_001564415.1	Robène <i>et al.</i> , 2016
axmaniC151	<i>X. phaseoli</i> pv. <i>manihotis</i>	CIO151	<i>Manihot</i> sp.	Colombia, na	36	5,154,532	64.7	GCA_000265845.1	Bart <i>et al.</i> , 2012
pha1GLC412	<i>X. phaseoli</i> pv. <i>phaseoli</i>	CFBP412	<i>Phaseolus vulgaris</i>	USA, na	1	5,028,351	64.94	GCA_900234435	this study
pha1GL6164	<i>X. phaseoli</i> pv. <i>phaseoli</i>	CFBP6164	<i>Phaseolus vulgaris</i>	Romania, 1996	1	5,319,061	64.64	GCA_900234535	this study
pha1GL6546	<i>X. phaseoli</i> pv. <i>phaseoli</i>	CFBP6546	<i>Phaseolus vulgaris</i>	USA, na	4	4,994,665	64.92	Deposited	this study
pha1GL6984	<i>X. phaseoli</i> pv. <i>phaseoli</i>	CFBP6984	<i>Phaseolus vulgaris</i>	Reunion island, France, 2000	1	5,106,663	64.82	GCA_900234425	this study
pha1GL7430	<i>X. phaseoli</i> pv. <i>phaseoli</i>	CFBP7430	<i>Phaseolus vulgaris</i> cv. <i>CFO16</i>	Iran, 2006	1	5,085,978	64.89	GCA_900234445	this study
axsyng9055	<i>X. phaseoli</i> pv. <i>syngonii</i>	LMG9055	<i>Syngonium podophyllum</i>	na	6	5,010,905	64.82	GCA_001640215.1	Robène <i>et al.</i> , 2016
saccha4393	<i>X. sacchari</i>	NCPBP4393	<i>Musa</i> sp.	Tanzania, 2007	470	4,955,099	69.03	AGDB00000000	Studholme <i>et al.</i> , 2011
cereal2541	<i>X. translucens</i> pv. <i>cerealis</i>	CFBP2541	<i>Bromus inermis</i>	USA, 1941	10	4,524,870	67.37	GCA_000807145.1	Pesce <i>et al.</i> , 2015
trgramX129	<i>X. translucens</i> pv. <i>graminis</i>	ART-Xig29	<i>Lolium multiflorum</i>	Switzerland, na	788	4,203,855	68.58	ANGG00000000	Wichmann <i>et al.</i> , 2013
gramin2053	<i>X. translucens</i> pv. <i>graminis</i>	CFBP2053	<i>Dactylis glomerata</i>	Switzerland, 1973	2	4,344,936	68.33	LHSI00000000	Pesce <i>et al.</i> , 2013
trtran1874	<i>X. translucens</i> pv. <i>translucens</i>	DSM18974	<i>Hordeum vulgare</i>	USA, 1958	551	4,550,921	67.67	GCA_000331775.1	unpublished
vasculN702	<i>X. vasicola</i> pv. <i>vasculorum</i>	NCPBP702	<i>Saccharum officinarum</i>	Zimbabwe, 1959	97	5,491,457	56.86	ACHS00000000	Studholme <i>et al.</i> , 2010
vesica3537	<i>X. vesicatoria</i>	ATCC35937	<i>Lycopersicon esculentum</i>	New Zealand, 1955	296	5,567,561	64.06	AEQV01000000	Potnis <i>et al.</i> , 2011
xylefaTem1	<i>Xylella fastidiosa</i> subsp. <i>fastidiosa</i>	Temecula1	<i>Vitis vinifera</i>	USA, na	1	2,519,802	51.77	GCA_000007245.1	Van Sluys <i>et al.</i> , 2003
xylefa9a5c	<i>Xylella fastidiosa</i> subsp. <i>pauca</i>	9a5c	<i>Valencia sweet orange</i>	Brazil, 1992	1	2,679,306	52.67	GCA_000006725.1	Simpson <i>et al.</i> , 2000

na: not applicable

(*Phaseolus vulgaris* L.). CBB occurs everywhere common bean is cultivated and may cause up to 75% yield loss in the most severe cases (EFSA Panel on Plant Health, 2015; Belete & Bastas, 2017). *Xanthomonas* strains responsible for CBB are spread across four different genetic lineages (Alavi *et al.*, 2008). The fuscans lineage and the non-fuscans lineages 2 (NF2) and 3 (NF3) belong to *X. citri* pv. *fuscans* while the non-fuscans lineage 1 (NF1) belongs to *X. phaseoli* pv. *phaseoli* (Vauterin *et al.*, 1995; Rademaker *et al.*, 2005; Constantin *et al.*, 2016). Pathological convergence between the NF1 and fuscans lineages is associated with large events of horizontal gene transfer (HGT) involving dozens of genes (Aritua *et al.*, 2015). Horizontal transfer of genes encoding Transcription Activator-Like (TAL) type III effectors was also observed between the four lineages of CBB agents (Ruh *et al.*, 2017a). In particular, all strains from the four genetic lineages display an allele of the *tal23A* gene, suggesting that this gene is important for *Xanthomonas* adaptation to common bean. Unveiling *Xanthomonas* genes involved in common bean specificity is key to better understanding the molecular bases of *Xanthomonas*-common bean interaction.

In order to search for genes putatively involved in common bean specificity in *Xanthomonas*, we have generated the whole genome sequences of 18 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. A combination of approaches including a comparisons of individual gene phylogenies compared to the phylogeny of organisms, an analyses of gene gains and losses based on parsimony, and a genome scan analysis to search for specific *k*-mers, was used to search for genes presenting common traits unique to strains belonging to the four lineages of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*.

RESULTS

Genome sequencing and phylogeny

In order to obtain genomic data representative of the diversity of *Xanthomonas* strains responsible for CBB, we produced whole genomes sequences from 18 strains representing the diversity of the four genetic lineages of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*. In addition, we used 54 other publically available genomes for a total of 72 whole genome sequences from *Xanthomonas* (**Table 1**). *Stenotrophomonas maltophilia* strain R551-3 and *Xylella fastidiosa* strains 9a5c and Temecula1 were used as outgroups for further analyses (Simpson *et al.*, 2000; Van Sluys *et al.*, 2003; Taghavi *et al.*, 2009). Annotation revealed from 3,029 to 5,405 coding sequences (CDS) per *Xanthomonas* genome (**Additional file 1 Fig. S1**). In *Xanthomonas* strains responsible for CBB, chromosome size ranged from 4,957,446 bp in strain CFBP1815 to 5,517,999 bp in strain CFBP6992, and the average GC% ranged from 64.3% to 64.9%. Strain CFBP4834-R contained less CDS (4,086) than any other strain responsible for CBB (4,454 to 5,405), which was

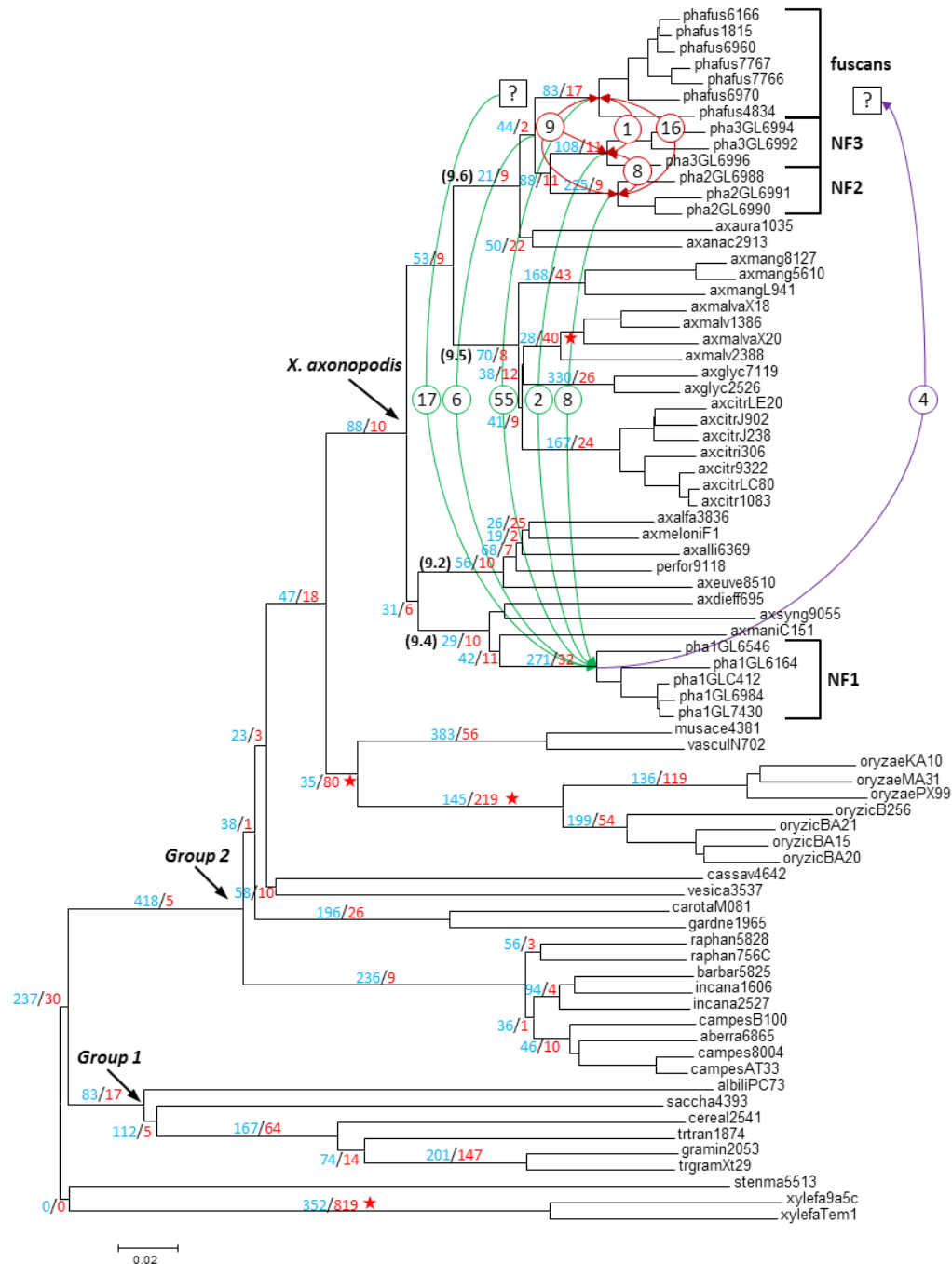


Fig. 1. Phylogeny of *Xanthomonas* strains used in this study with indication of gene gains and losses in lineages responsible for common bacterial blight of bean. The phylogenetic tree is based on whole proteome analysis using CVTree (Qi *et al.*, 2004) with default parameters. Strain aliases are described in **Table 1**. *Stenotrophomonas* and *Xylella* proteomes have been used as outgroups. *Xanthomonas* main phylogenetic groups 1 and 2 (Young *et al.*, 2008) and the *X. axonopodis* species complex (Vauterin *et al.*, 1995, 2000) are indicated by arrows. Groups 9.2, 9.4, 9.5 and 9.6 (Rademaker *et al.*, 2005) are indicated in brackets at branching. Fuscans, NF2, NF3 and NF1 refer to the four genetic lineages of strains responsible for CBB. A parsimony approach has been performed to infer gene gains (blue) and losses (red) at levels higher than the pathovar rank, and numbers are displayed at each branch. Red stars highlight cases where gene loss was greater than gene gain. Curved arrows represent cases of HGT retrieved by Ks analysis on alignments of 116 candidate genes for bean specificity, with HGT from *X. citri* pv. *fuscans* to *X. phaseoli* pv. *phaseoli* in green, HGT from *X. phaseoli* pv. *phaseoli* to *X. citri* pv. *fuscans* in purple, and HGT between *X. citri* pv. *fuscans* lineages in red. Numbers in circles correspond to the numbers of candidate genes involved for each HGT. Question marks indicate events for which origin or end of the HGT was not precise enough to assign any particular lineage.

consistent with previous analyses showing that CFBP4834-R genome presented a deletion of 33 kbp affecting flagellum biosynthesis (Darrasse *et al.*, 2013a). The phylogeny of strains was assessed based on the whole proteomes from the 75 strains studied using CVTree (**Fig. 1**). The overall topology of this tree was congruent with previous *Xanthomonas* phylogenies (Rademaker *et al.*, 2005; Constantin *et al.*, 2016; Jacques *et al.*, 2016). As described previously, strains responsible for CBB split into four distinct genetic lineages now grouped in two different species *X. citri* and *X. phaseoli* (Alavi *et al.*, 2008; Mhedbi-Hajri *et al.*, 2013).

***Xanthomonas* evolution is shaped by numerous gene gains and losses**

To study the genes shared by different clades of *Xanthomonas*, we constructed an orthology matrix using OrthoMCL (**Additional File 2 Table S1**, *non présenté ici en raison de sa taille*). Based on this orthology matrix, we performed a parsimony approach to infer gene gains and losses at each branch of the phylogenetic tree (**Fig. 1**). We did not take into account the most distal branches because of possible biases due to the difference of quality between sequenced genomes. At every branch, from one to several hundred genes were either gained or lost. A general observation was that gene gains were higher than gene losses except in four cases (red stars in **Fig. 1**). These four cases of genome reduction were observed (i) at the origin of the *Xylella* genus, (ii and iii) at two consecutive branches before and after the split between the *X. vasculorum* and *X. musacearum* species and the *X. oryzae* and *X. oryzicola* species, and (iv) at the origin of *X. citri* pv. *malvacearum*. The most important gene gain (418) was observed at the origin of *Xanthomonas* phylogenetic group 2 as defined by Young *et al.* (2008), while the most important gene loss (819) was observed at the origin of the *Xylella* genus. Few gene losses were observed before diversification of each of the four genetic lineages involved in CBB. The NF1 lineage had the most gained genes (271) followed by the NF2 (225), NF3 (108) and fuscans (83) lineages, respectively.

Pan and core genomes are usually used to assess the divergence in gene content between individuals. Individuals that are closely related to each other typically share more orthologs than unrelated individuals, therefore groups of closely related individuals tend to have a smaller pan genome and a larger core genome than groups of more divergent individuals. As such, the pan and core genomes for the 72 *Xanthomonas* strains comprised 32,602 and 1,144 CDS, respectively, while the pan and core genomes for the 75 strains used in this study comprised 34,723 and 816 CDS, respectively (**Fig. 2**). Similarly, each group 9.2, 9.4, 9.5 and 9.6 had smaller pan genomes (6,578, 8,222, 9,387 and 9,437 CDS, respectively) and larger core genomes (3,493, 2,949, 3,056 and 3,213 CDS, respectively), than the *X. axonopodis* species complex that had pan and core genomes of 19,010 and 2,297 CDS, respectively. Strikingly, when grouping strains responsible for CBB from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*, both the pan and core genomes (10,750 and 3,222 CDS, respectively) were larger than the pan and core genomes from groups 9.4 (8,222

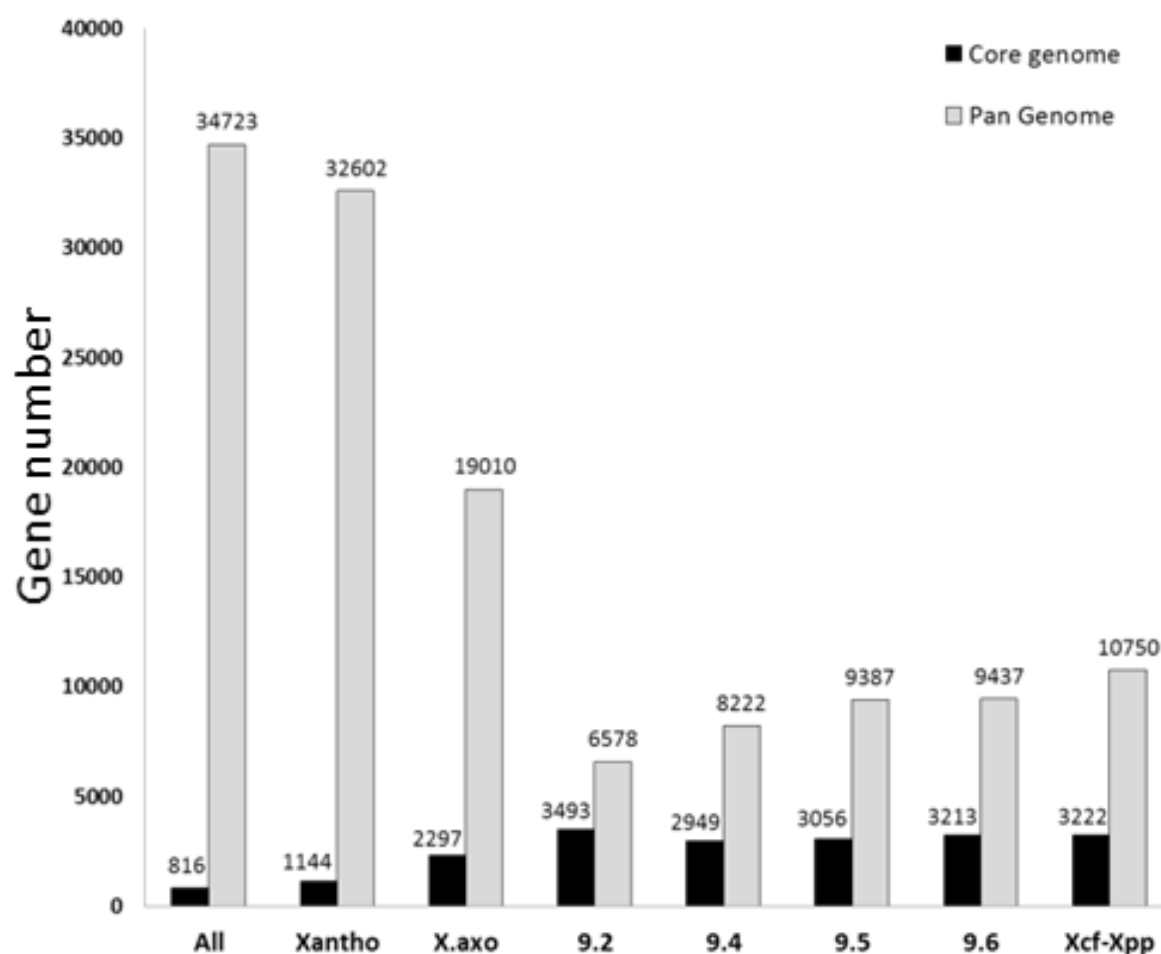


Fig. 2. The core- and pan-genome of *Xanthomonas*. Gene numbers correspond to the number of ortholog groups retrieved for each group of strains. All: all strains used in this study (n=75); Xantho: strains from the *Xanthomonas* genus (n=72); X.axo: strains from the *Xanthomonas axonopodis* species complex (n=44). 9.2, 9.4, 9.5, 9.6: strains belonging to rep-PCR groups 9.2 (n=5), 9.4 (n=8), 9.5 (n=16) and 9.6 (n=15), respectively, as defined in Rademaker *et al.*, 2000. Xcf-Xpp: strains pathogenic on common bean belonging to *Xanthomonas citri* pv. *fuscans* or *Xanthomonas phaseoli* pv. *phaseoli* (n=18).

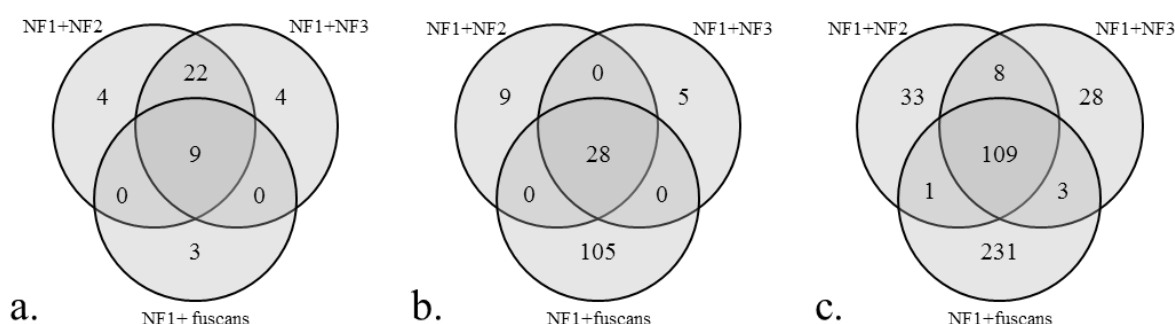


Fig. 3. Venn diagrams of genes presenting different types of similarities among the NF1 plus NF2, the NF1 plus NF3, and/or the NF1 plus fuscans lineages. a. Ortholog groups of genes shared by the NF1 and NF2 lineages, the NF1 and NF3 lineages, and/or the NF1 and fuscans lineages, and that were gained in at least one of these lineages. b. Genes monophyletic within the NF1 and NF2 lineages, the NF1 and NF3 lineages, and/or the NF1 and fuscans lineages. c. Genes containing 24-mers specifically retrieved in, or absent from the NF1 and NF2 lineages, the NF1 and NF3 lineages, and/or the NF1 and fuscans lineages.

and 2,949 CDS, respectively) or 9.6 (9,437 and 3,213 CDS, respectively) (**Fig. 2**). Thus, strains responsible for CBB, although being phylogenetically diverse, had a larger core genome together than with any other strains belonging to their respective clades, which is an evidence of extensive HGT among these strains. This result is reminiscent of previous comparative analyses showing that dozens of genes have been horizontally transferred between the fuscans and the NF1 lineages (Aritua *et al.*, 2015).

Strains pathogenic on common bean share 116 CDS presenting unique characteristics

To search for genes potentially involved in common bean specificity, we performed a combination of analyses for unveiling genes with features unique to the causal agents of CBB. First, to search for genes corresponding to new acquisitions in *X. phaseoli* pv. *phaseoli* (*i.e.* the NF1 lineage) and/or in the three lineages from *X. citri* pv. *fuscans* (*i.e.* NF2, NF3 and fuscans), we used the same parsimony approach as described above to infer CDS gains and losses within the core genome of 3,222 CDS shared by the four genetic lineages (**Fig. 3a**). Second, we used a phylogenetic approach to search for genes for which strains from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* formed a monophyletic group. For this, we constructed phylogenetic trees on 3,202 CDS present in every *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains and in at least one other strain group 9.2 or 9.4 and one other strain from group 9.5 or 9.6 as outgroups (**Fig 3b**). Finally, we used the SkIf tool (Briand *et al.*, 2016) on the 72 *Xanthomonas* annotated genomes to search for genes containing short 24-bp sequences (24-mers) specific for strains responsible for CBB, or alternatively genes from strains responsible for CBB lacking 24-mers present in all strains from the *X. axonopodis* species complex (**Fig. 3c**).

The first approach unveiled nine CDS that were shared by all four genetic lineages, and gained in at least one of these lineages (**Fig. 3a**). Moreover, 22 gained CDS were specifically shared by the NF1, NF2 and NF3 lineages, while only three were specifically shared by the NF1 and fuscans lineages. This indicates that more gained genes were shared among the NF1, NF2 and NF3 lineages than among the NF1 and fuscans lineages. The second approach unveiled 28 CDS for which the four genetic lineages form a monophyletic group, suggesting that they were horizontally transferred among these lineages (**Fig. 3b**). Nine CDS were specifically monophyletic for the NF1 and NF2 lineages, five for the NF1 and NF3 lineages, and 105 for the NF1 and fuscans lineages, suggesting that most horizontal transfers occurred among the NF1 and fuscans lineages. Finally, 109 CDS comprised 24-mers either specifically present in the four lineages, or absent from these lineages but present in all other strains from the *X. axonopodis* species complex (**Fig. 3c**). Thirty-three CDS contained 24-mers specific for the NF1 and NF2 lineages, 28 for the NF1 and NF2 lineages and 231 for the NF1 and fuscans lineages.

Table 2. Overview of the 116 genes putatively involved in bean specificity.

Identifier in Fig. 4	Accession number in strain CFBP6546	Predicted function	Gene name	Gained	Monophyletic	Specific 24-mers	Recombinant	HGT between NF1 and others	Nonsynonymous sites specific for CBB	In the list from Aritua et al., 2015
1	m00100560	TonB-dependent transporter	-	-	-	Yes	-	Yes	-	-
2	m00100580	type III effector	<i>avrBs2</i>	-	-	Yes	Yes	Yes	-	-
3	m00100590	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
4	m00101230	two component system sensor protein	-	-	-	Yes	-	-	-	-
5	m00101980	two component system protein	<i>glnG</i>	-	-	Yes	-	-	-	-
6	m00102200	two component system response regulator	-	-	Yes	Yes	-	Yes	-	Yes
7	m00104250	type III secretion system protein	<i>hrpF</i>	-	-	Yes	Yes	Yes	-	-
8	m00104520	type III effector	<i>xopA</i>	-	-	Yes	-	Yes	Yes	-
9	m00104530	lytic transglycosylase-like protein	<i>hpaH</i>	-	-	Yes	-	Yes	Yes	Yes
10	m00104540	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
11	m00104640	diguanylate cyclase	-	-	-	Yes	-	Yes	Yes	-
12	m00104690	4-alpha-glucanotransferase	<i>malQ</i>	-	-	Yes	-	Yes	-	-
13	m00105290	threonine aldolase	-	-	-	Yes	-	Yes	-	-
14	m00105330	flavin reductase	-	-	Yes	Yes	-	Yes	Yes	-
15	m00105340	anthranilate phosphoribosyl transferase	<i>trpD</i>	-	-	Yes	-	Yes	-	-
16	m00105390	S-adenosylmethionine decarboxylase	<i>speD</i>	-	-	Yes	-	Yes	-	-
17	m00105750	hypothetical membrane protein	-	-	-	Yes	-	-	-	-
18	m00107500	TonB-dependent transporter	-	-	-	Yes	Yes	Yes	-	-
19	m00138140	two-component system sensor protein	-	-	-	Yes	-	Yes	-	-
20	m00109610	TonB-dependent transporter	-	-	-	Yes	Yes	Yes	-	-
21	m00109620	phosphoanhydride phosphohydrolase	<i>appA</i>	-	Yes	Yes	Yes	Yes	Yes	-
22	m00109670	hypothetical protein	-	-	-	Yes	-	Yes	-	-
23	m00109990	bis(5' nucleosyl)-tetraphosphatase	<i>apaH</i>	-	-	Yes	-	-	-	-
24	m00110970	peptidyl-tRNA hydrolase	-	-	Yes	Yes	-	Yes	-	Yes
25	m00111160	50S ribosomal protein	<i>rplW</i>	-	Yes	-	-	Yes	-	-
26	m00112070	two-component system sensor protein	<i>phoR</i>	-	-	Yes	-	-	-	-
27	m00136550	hypothetical protein	-	-	-	Yes	-	-	-	-
28	m00136050	phosphoglycerate kinase	<i>pgk</i>	-	-	Yes	-	Yes	-	-
29	m00135980	siroheme synthase	<i>cysG</i>	-	Yes	Yes	-	Yes	-	-
30	m00135970	transcriptional regulator protein	<i>cysB</i>	-	-	Yes	-	Yes	-	-
31	m00135880	sulfite reductase (NADPH) subunit	<i>cysI</i>	-	-	Yes	-	Yes	-	-
32	m00135870	sulfite reductase (NADPH) subunit	<i>cysJ</i>	-	-	Yes	Yes	Yes	-	-
33	m00135840	membrane fusion protein	-	-	-	Yes	-	-	-	-
34	m00130690	inner membrane protein	-	-	-	Yes	-	Yes	-	-
35	m00130670	RNA-binding protein	-	-	-	Yes	-	Yes	-	Yes
36	m00130660	hypothetical protein	-	-	Yes	Yes	-	Yes	Yes	-
37	m00130650	RNA polymerase ECF-type sigma factor	<i>rpoE2</i>	-	-	Yes	-	Yes	Yes	-
38	m00121810	threonine synthase	<i>thrC</i>	-	Yes	Yes	-	Yes	Yes	-
39	m00121850	hisdyl tRNA synthetase	<i>hisS</i>	-	-	Yes	-	Yes	-	-
40	m00121900	hisdinol phosphatase	<i>hisB</i>	-	-	Yes	-	Yes	-	Yes
41	m00123210	recombination factor protein	<i>rarA</i>	-	-	Yes	-	Yes	-	-
42	m00126530	ethanolamine permease	<i>eutP</i>	-	Yes	Yes	-	Yes	-	-
43	m00126540	ethanolamine ammonia-lyase	<i>eutB</i>	-	-	Yes	-	Yes	-	-
44	m00127140	membrane protein	-	-	-	Yes	-	Yes	-	-
45	m00127150	cGMP specific phosphodiesterase	-	-	Yes	Yes	-	Yes	Yes	-
46	m00127220	hypothetical protein	-	-	-	Yes	-	Yes	-	-
47	m00127250	ubiquinol-cytochrome c reductase subunit	-	-	-	Yes	-	Yes	-	-
48	m00127260	ubiquinol-cytochrome c reductase subunit	-	-	-	Yes	-	Yes	-	-
49	m00127850	1-phosphofructokinase	<i>fruK</i>	-	-	Yes	-	Yes	-	-
50	m00127860	PTS fructose porter IIBC component	<i>fruA</i>	-	-	Yes	-	Yes	-	-
51	m00119700	putative secreted protein	-	-	Yes	-	-	-	-	-
52	m00119970	phosphomethylpyrimidine kinase	<i>thiD</i>	-	-	Yes	-	Yes	-	Yes
53	m00118890	hypothetical protein	-	-	Yes	-	-	Yes	-	-
54	m00117610	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
55	m00117600	transcriptional regulator	-	-	Yes	Yes	Yes	Yes	Yes	-
56	m00117590	acetyltransferase	-	-	-	Yes	-	Yes	-	-
57	m00116790	NAD/FAD binding protein	-	-	Yes	-	-	Yes	-	-
58	m00114680	two component system sensor protein	-	-	-	Yes	Yes	Yes	-	-
59	m00113910	membrane protein	-	-	Yes	-	-	Yes	-	Yes
60	m00113490	xanthine dehydrogenase subunit	-	-	-	Yes	-	Yes	-	-
61	m00113470	xanthine dehydrogenase subunit	-	-	Yes	Yes	Yes	Yes	Yes	-
62	m00108600	TonB-dependent transporter	-	-	-	Yes	Yes	Yes	Yes	-
63	m00108570	hypothetical protein	-	-	Yes	Yes	-	Yes	-	-
64	m00108560	general secretion pathway protein D	<i>xpsD</i>	-	Yes	Yes	Yes	Yes	Yes	-
65	m00139330	dihydrolipoamide acetyltransferase	<i>aceF</i>	-	-	Yes	Yes	Yes	-	-
66	m00139350	DNA glycosylase	-	-	Yes	Yes	-	Yes	-	-
67	m00139420	membrane protein	-	-	-	Yes	-	Yes	Yes	Yes
68	m00139860	putative secreted protein	-	-	-	Yes	-	-	-	-
69	m00139930	flavodoxin protein	-	-	-	Yes	-	Yes	-	-
70	m00140880	hypothetical protein	-	-	-	Yes	-	-	-	-
71	m00140860	peptidoglycan binding protein	<i>lysM</i>	-	Yes	-	-	Yes	-	-
72	m00141310	PhnB like protein	-	-	-	Yes	-	Yes	Yes	-
73	m00141340	hypothetical protein	-	-	-	Yes	-	Yes	-	-
74	m00141400	hypothetical protein	-	-	-	Yes	-	Yes	-	-
75	m00141430	transmembrane protein	-	-	-	Yes	-	Yes	-	Yes
76	m00141440	hypothetical protein	-	-	Yes	Yes	-	Yes	-	-
77	m00144630	enolase	-	-	-	Yes	-	Yes	-	-
78	m00142140	putative glucosyltransferase	-	-	-	Yes	Yes	Yes	-	-
79	m00114910	type VI secretion system protein	<i>icmF</i>	-	-	Yes	Yes	Yes	Yes	-
80	m00144240	type VI secretion system kinase protein	<i>tagE</i>	-	-	Yes	Yes	Yes	-	-
81	m00144270	transmembrane sensor protein	<i>fecR</i>	-	-	Yes	-	Yes	-	-
82	m00144420	type VI secretion system virulence protein	<i>impE</i>	-	-	Yes	Yes	Yes	-	-
83	m00144930	type III effector	<i>xopAD</i>	-	-	Yes	Yes	Yes	Yes	-
84	m00145330	membrane protein	-	-	-	Yes	Yes	Yes	-	-
85	m00146940	LysR family transcriptional regulator	-	-	-	Yes	-	Yes	-	-
86	m00146970	short chain dehydrogenase	-	-	-	Yes	-	Yes	-	-
87	m00146980	RNA 2'-phosphotransferase family protein	-	-	-	Yes	-	Yes	Yes	-
88	m00146990	hypothetical protein	-	-	Yes	Yes	-	Yes	-	-
89	m00147020	esterase family protein	-	-	-	Yes	-	Yes	-	-
90	m00147060	putative secreted protein	-	-	-	Yes	-	Yes	-	-
91	m00147090	putative secreted protein	-	-	-	Yes	-	Yes	Yes	-
92	m00147100	ribosomal pseudouridine synthase	-	-	Yes	Yes	-	Yes	Yes	-
93	m00147120	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
94	m00147130	putative secreted protein	-	-	-	Yes	-	Yes	Yes	-
95	m00147140	membrane protein	-	-	-	Yes	-	-	-	-
96	m00200790	transposase	<i>tnpA</i>	-	-	Yes	-	Yes	Yes	-
97	m00200820	resolvase	<i>tnpR</i>	-	-	Yes	-	Yes	Yes	-
98	m00200130	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
99	m00200150	hypothetical protein	-	Yes	-	Yes	-	na	Yes	-
100	m00200160	hypothetical protein	-	-	-	Yes	-	na	Yes	-
101	m00248020	hypothetical protein	-	Yes	-	-	-	na	Yes	-
102	m00248040	hypothetical protein	-	-	-	Yes	-	na	Yes	-
103	m00248080	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
104	m00200880	DNA topoisomerase IA	-	-	-	Yes	-	na	Yes	-
105	m00132430	DNA topoisomerase IA	-	-	-	Yes	-	na	Yes	-
106	m00248630	type IV secretion system protein	<i>trbI</i>	Yes	Yes	Yes	-	na	Yes	-
107	m00201220	type IV conjugal transfer protein	<i>traG</i>	-	-	Yes	-	na	Yes	-
108	m00201230	hypothetical protein	-	-	-	Yes	-	na	Yes	-
109	m00201250	lytic transglycolase	-	Yes	Yes	Yes	-	na	Yes	-
110	m00201260	type IV conjugal transfer protein	<i>traF</i>	Yes	-	Yes	-	na	Yes	-
111	m00201270	hypothetical protein	-	Yes	Yes	Yes	-	na	Yes	-
112	m00248870	hypothetical protein	-	Yes	Yes	Yes	-	na	Yes	-
113	m00200020	plasmid partitioning protein	-	-	Yes	Yes	-	na	Yes	-
114	m00200060	hypothetical protein	-	Yes	-	Yes	-	na	Yes	-
115	m00200080	hypothetical protein	-	Yes	-	Yes	-	na	Yes	-
116	m00247980	hypothetical protein	-	-	-	Yes	-	Yes	Yes	-
TOTAL	-	116	-	9	28	109	18	89	45	9

- : not found in corresponding analysis

na: not applicable to the absence of corresponding CDS in outgroups

In all, it was striking that the number of gained CDS was higher when comparing the NF1 lineage to the NF2 or NF3 strains (than to the fuscans strains while the vast majority of monophyletic genes was retrieved between the NF1 and fuscans strains compared to the NF2 or NF3 strains. Similarly, most CDS retrieved using the 24-mer approach were shared between the NF1 and fuscans lineage compared to the NF2 or NF3 lineages. Together, these results suggest that different types of events occurred independently between the NF1 lineage and the NF2, NF3 or fuscans lineages. Non-homologous transfers corresponding to the acquisition of novel genes were predominant when comparing the NF1 lineage to the NF2 and/or NF3 lineages. By contrast, the high numbers of monophyletic CDS and CDS sharing 24-mers found among the NF1 and fuscans lineages were suggestive of homologous HGT resulting in the replacement of ancient genes by foreign alleles by recombination between the NF1 and fuscans ancestors.

Together, the three aforementioned analyses unveiled 116 non-redundant CDS presenting features unique to CBB agents (**Table 2**). The most represented functions encoded by these 116 predicted CDS were hypothetical proteins (27 CDS), followed by membrane-related proteins (10 CDS), two-component system proteins (6 CDS), putative secreted proteins (5 CDS), reductases (5 CDS), RNA-related proteins (5 CDS), TonB-dependent proteins (4 CDS), Type III secretion system-related proteins (4 CDS), Type IV secretion system-related proteins (3 CDS), Type VI secretion system-related proteins, DNA-related proteins (3 CDS) and transcription regulators (3 CDS) (**Table 2**). We hypothesized that the CDS potentially involved in adaptation to common bean should encode proteins bearing specific amino acid sequences compared to the other *Xanthomonas* strains. Analysis of amino acid alignments for the 116 candidate CDS highlighted 45 predicted proteins with amino acids retrieved exclusively in *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* (**Table 2**). More than one third of these predicted proteins (17/45) were hypothetical proteins. Other candidate proteins included three type IV secretion system proteins TrbI, TraG and TraF, two putative secreted proteins, two type III effectors XopA and XopAD, two DNA topoisomerases, and other proteins of various functions (**Table 2**).

Strain CFBP6546 from the NF1 lineage was used as reference for further analyses. Its genome contained one chromosome and three extrachromosomal plasmids formerly described as plasmids a, b and c in strain CFBP4834-R (Darrasse *et al.*, 2013a). Most candidate genes (96/116) were located on the chromosome, while 20 were located on plasmid a (**Fig.4**). This corresponds to a density of one candidate gene per 50.9 kbp in the chromosome, and one per 3.5 kbp for plasmid a, while none were retrieved in plasmids b or c. Interestingly, all CDS found in plasmid a contained specific amino acid sequences (**Table 2**). Thus, plasmid a appeared as an important vector of genes involved in the adaptation to common bean. Another observation was that the chromosome contained regions with various 24-mers shared by the NF1 lineage and any of the fuscans, NF2 or NF3 lineages (in green, blue or black in **Fig. 4**, respectively). This suggests

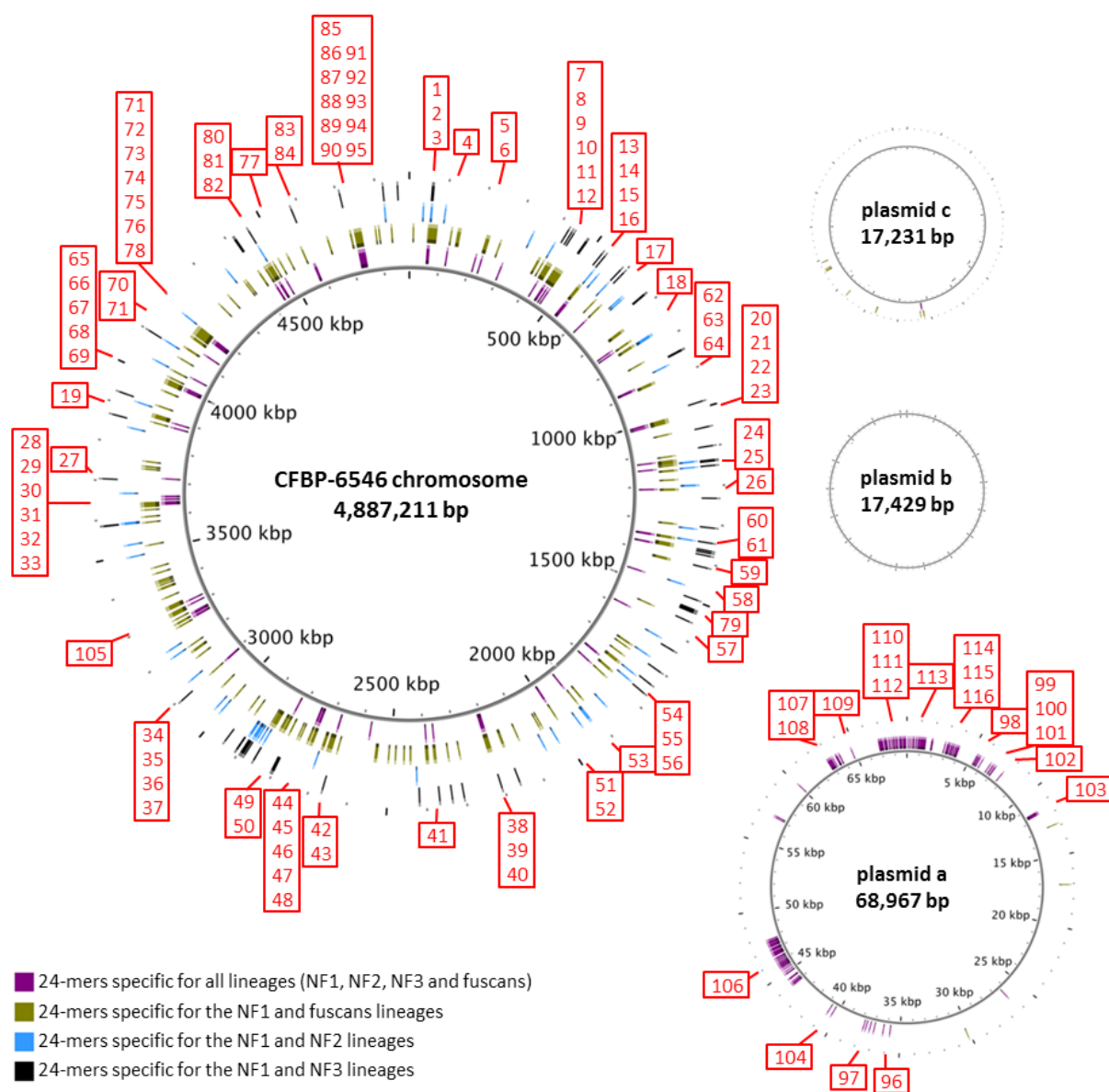


Fig. 4. Mapping of the 24-mers specific for strains pathogenic on common bean. The 24-mers specific for *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains, or specific for both the NF1 lineage plus another lineage from *X. citri* pv. *fuscans*, were aligned to the genome of strain CFBP6546 (NF1) using BLASTN (100% identity over 24 bp). The innermost rings indicate the positions on the reference chromosome or plasmids. Red numbers correspond to the identifiers of the 116 genes putatively involved in bean specificity, as listed in **Table 2**.

that the regions shared by the NF1 and the other lineages diverged since the split between the NF2, NF3 and fuscans lineages. By contrast to what was observed for the chromosome, plasmid a contained exclusively 24-mers shared simultaneously by the four genetic lineages of strains responsible for CBB (in purple in **Fig. 4**), indicating that these regions have been shared between the four lineages recently enough to keep sharing 100% identity between each other. Together, these results suggest that HGT events occurred at different times during the evolution of groups 9.4 and 9.6.

Bean specificity is associated with consecutive waves of horizontal gene transfers

To detect and estimate the time of HGT for the 116 genes shared between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*, we compared the values of pairwise nucleotide synonymous substitution rates at silent sites (Ks), with higher Ks value meaning longer time of divergence between two sequences. Among these 116 genes, 18 were recombinant genes according to RDP software analysis (**Table 2**). For these 18 recombinants, different Ks values were independently calculated on either side of the breakpoint. Moreover, 16 genes were absent from the outgroups (**Table 2**), therefore we couldn't perform this analysis on these genes. In all, we tested 100 out of 116 genes. After alignment, we compared the mean Ks values for 28 different pairwise combinations of strains (**Additional file 3 Table S2**). For example, for gene *m00105290* from strain CFBP6546, the mean Ks value was 0.0 +/- 0.0 between strains from the NF1 and fuscans lineages. This value was inferior to the Ks value for strains from the NF1 lineage versus strains from *X. citri* pv. *anacardii* and *X. citri* pv. *aurantifolii* (Ks = 3.10e-03 +/- 1.90e-03), which was in turn inferior to the Ks value for strains from the NF1 lineage versus *X. phaseoli* pv. *manihotis* strain CIO151 (Ks = 7.13e-2 +/- 0.0). Therefore, in NF1 strains the *m00105290* gene was closer to orthologs from group 9.6 than to other strains from group 9.4, meaning that it was horizontally transferred from the fuscans lineage to the NF1 lineage. Using the same approach of Ks comparisons, we confirmed HGT for 93 out of 100 genes tested (**Fig. 1, Table 2, Additional file 3 Table S2**). The majority of HGT events (88) were directed from the *X. citri* pv. *fuscans* lineages to the NF1 lineage, while only four occurred from the NF1 lineage to *X. citri* pv. *fuscans* lineages. In particular, 55 HGT events were detected from the fuscans to the NF1 lineages. In addition to having been transferred between distant lineages, 16 and 1 genes were also transferred between the fuscans and the NF2 lineages, or the fuscans and the NF3 lineages, respectively (**Fig. 1**). Moreover, 17 genes had null Ks values between the NF2 and NF3 lineages, suggesting that HGT events also occurred between both lineages (**Fig. 1, Additional file 3 Table S2**). Interestingly, among these 17 genes, nine also shared null Ks values when compared to the fuscans strains, and five of those were identical among the four genetic lineages. Together, our results show that several more or less important waves of HGT occurred between phylogenetically distant strains

responsible for CBB.

DISCUSSION

Here, we performed a comparative genomics analysis to mine for genes putatively involved in *Xanthomonas* specificity to common bean. For this, we generated the whole genome sequence from 15 strains representing the diversity of the four genetic lineages belonging to *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*. We used a combination of comparative genomics approaches that led to the discovery of 116 genes bearing features unique to CBB agents. Out of these 116 genes, 109 were retrieved using the SkIf tool based on specific 24-mer search (Briand *et al.*, 2016). Previous analyses based on identity percentage unveiled 63 genes sharing 100% identity over at least 95% of their length among strains from the NF1 and *fuscans* lineages (Aritua *et al.*, 2015). Strikingly, only nine of these genes were retrieved within our list of 116 genes (**Table 2**). These different results can be explained by the fact that we discarded most of the genes shared only by the NF1 and *fuscans* lineages for keeping genes simultaneously similar among the four genetic lineages of CBB agents (**Fig. 3**). On the other hand, we unveiled numerous genes that did not share 100% identity among the NF1 and *fuscans* lineages within their whole length, but instead shared small specific sequences of 24 nucleotides or more. Whether these similarities lie within functionally important domains of the encoded proteins remains to be studied.

Ks comparisons, showed that a majority of these genes were involved in HGT between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*. Therefore, HGT was the predominant force leading to similarities between the genomes of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*. Finding HGT events within these genes validated our approach. In particular, SkIf was an interesting tool because in addition to being more sensitive than gene gain and loss or phylogenetic approaches, it was not based on gene alignments, thus less sensitive to annotation and/or sequencing biases. Genes not involved in HGT between *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* lineages were retrieved, suggesting that other evolutionary events occurred, possibly including convergent evolution and/or HGT with strains other than CBB agents.

HGT events occurred at different moments of the evolution of *Xanthomonas* strains having common bean as host. The vast majority of these HGT were directed from *X. citri* pv. *fuscans* to *X. phaseoli* pv. *phaseoli*. This strongly suggest that the NF2, NF3 and *fuscans* lineages from *X. citri* pv. *fuscans* were originally pathogenic on common bean, and that the NF1 lineage from *X. phaseoli* pv. *phaseoli* subsequently acquired the ability to cause CBB on bean thanks to consecutive acquisitions of novel gene and/or novel alleles coming from the three *X. citri* pv. *fuscans* lineages. This can be contrasted with our knowledge on the origin of the lineages and their genetic diversity.

The causal agent of CBB was first isolated and identified by Smith in 1897 (Zaumeyer, 1930) as a yellow pigmented strain, later shown as belonging to the NF1 lineage. The first fuscous strains were later isolated by Burkholder from beans grown in Switzerland in 1924 (Burkholder, 1930). There are however no data to document a putative pre-existence of one, the other, or both types of strains prior to their first identifications. The diversity of the yellow and fuscous strains was demonstrated by various methods. Amplified or restriction fragment length polymorphism (AFLP or RFLP) analyses (Lazo *et al.*, 1987; Birch *et al.*, 1997; Chan & Goodwin, 1999; Alavi *et al.*, 2008), amplified polymorphic DNA fragments (Birch *et al.*, 1997; Mkandawire *et al.*, 2004), pulse field gel electrophoresis (Chan & Goodwin, 1999), and more recently MLSA (Mhedbi-Hajri *et al.*, 2013) revealed that both types of strains are more or less equivalent in terms of genetic diversity, which suggests that diversification of both lineages occurred around the same time frame, and thus that the ancestors of these two lineages coexisted. In consequence, *X. citri* pv. *fuscans* may be the original CBB agent that had transferred determinants useful for adaptation on common bean to the ancestor of *X. phaseoli* pv. *phaseoli*. Very diverse functions were retrieved among the proteins encoded by the 116 candidate genes. Interestingly, 45 genes contained nonsynonymous polymorphisms leading to amino acids specific for strains responsible for CBB, suggesting that they may play an important role in common bean specificity. Although 17 of these 45 genes encoded hypothetical proteins, it appears that most other genes encoded proteins involved in pathogenicity or in the interaction with the plant environment. The type IV secretion system was particularly represented with genes encoding traF, traG and traI. Type IV secretion system is involved in the translocation of macromolecules that can be proteins important for pathogenicity, or DNA for mediating HGT (Fronzes *et al.*, 2009; Wallden *et al.*, 2010). Thus, sharing similar proteins of the type IV secretion system could favour HGT among strains responsible for CBB. Indeed, strains from the NF1, NF2, NF3 and fuscans lineages have been found in La Réunion Island in 2000 (**Table 1**), indicating that sympatry exists among all these lineages, rendering further HGT events possible (Ruh *et al.*, 2017a). Three genes encoding proteins related to the type III secretion system were retrieved. XopA (Noël *et al.*, 2002) and HpaH (Hausner *et al.*, 2017) are two proteins that may be involved in the structure of the type III secretion system, while XopAD is a type III effector of unknown function consisting of multiple SKW repeats of semi-conserved 42 amino acids (Escalon *et al.*, 2013; Teper *et al.*, 2016). The type III secretion system is pivotal for the virulence of most Gram negative plant pathogenic bacteria, and repertoires of type III effectors have been described as potentially important factors for host specificity and host adaptation in *Xanthomonas* (Hajri *et al.*, 2009; Barak *et al.*, 2016; Hersemann *et al.*, 2017) and other genera such as *Pseudomonas* (Baltrus *et al.*, 2011). Moreover, our analysis pointed out one diguanilate cyclase and one cGMP specific phosphodiesterase, two proteins involved in the metabolism of cyclic di-GMP that may play a role in biofilm formation (Newell *et al.*, 2011) and pathogenicity (Su *et al.*,

2016). One TonB-dependent transporter was retrieved. TonB-dependent transporter are outer membrane receptors involved in molecule uptake such as iron siderophore complexes or nutrients and may play a role in host specificity (Blanvillain *et al.*, 2007; Jacques *et al.*, 2016). Other proteins putatively involved in pathogenicity were retrieved, such as ThrC, a threonine synthase involved in the virulence of *X. oryzae* pv. *oryzicola* in rice (Guo *et al.*, 2012), XpsD, an outer membrane protein from the type II secretion system that is putatively involved in the secretion of cell wall degradative enzymes during infection (Solé *et al.*, 2015), or IcmF, a protein of the type VI secretion system, which is involved in the interaction with other bacteria and may participate in pathogenicity (Records & Gross, 2010; Alteri & Mobley, 2016). One flavine reductase and one xanthine dehydrogenase, two proteins putatively involved in oxydoreduction pathways were retrieved, and may be involved in response to stress during the interaction with common bean. In addition to genes putatively involved in pathogenicity or virulence, our analysis unveiled genes involved in more general metabolism pathways, such as PhnB involved in the biosynthesis of tryptophan (Xie *et al.*, 2003), or two DNA topoisomerases involved in the relaxation of the supercoiled DNA molecule during transcription, replication or recombination (Hashimi *et al.*, 2008). In all, the genes retrieved here encompass almost all the steps of host plant colonization by the bacteria, from the interaction with other microorganisms, to biofilm formation, response to oxidative stress, and inhibition of plant defences. Therefore, the ability to infect a particular plant species may require fine tuning of every aspects of the bacterial arsenal.

Interestingly, 19 of these 45 candidate genes were retrieved on plasmid a, suggesting that plasmid a played a major role for pathological convergence of CBB agents. Plasmids a and c carry additional type III effector genes encoding effectors from the Transcription Activator-Like (TAL) family that were horizontally transferred between the NF1, NF2, NF3 and/or fuscans lineages (Ruh *et al.*, 2017a). Plasmids are genetic elements that favour HGT, but transfers of whole plasmids often induces a fitness cost for the bacteria (Soucy *et al.*, 2015). More generally, horizontally transferred genes tend to be lost if not providing selective advantages for recipient strains (Koskiniemi *et al.*, 2012). Interestingly, the nine candidate genes retrieved by the gain and loss approach were all located on plasmid a (**Table 2**). The maintenance of these novel genes in the four genetic lineages of CBB agents is a testament to the importance of these genes for common bean specificity. Except for two genes encoding proteins involved in the type IV secretion system, the other seven genes encoded proteins of unknown function. It would be interesting to perform functional characterization of these genes, and to further analyse their implication in common bean specificity. After validation, specific PCR primers defined on genes required for infecting common bean could then be used for improving PCR-based monitoring of CBB agents. Moreover, such primers could potentially detect novel HGT events involving these genes in other strains

unrelated to *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans*, thus allowing to forecast new threads potentially dangerous for common bean production.

CONCLUSION

Together, our results indicate that consecutive waves of HGT occurred between phylogenetically distant *Xanthomonas* strains able to cause the same symptoms on the same host plant: common bean. These HGT led to specific combination of genes only retrieved in strains responsible for CBB, which provided new insights into the evolution of these bacteria towards infecting common bean. Mining for candidate genes for host specificity could be generalized to other polyphyletic pathovars such as *pathovars euvesicatoria*, *vesicatoria*, *perforans*, and *gardneri* forming a polyphyletic group of strains pathogenic on pepper and tomato (Schwartz *et al.*, 2015). Such analyses could both give new information on the molecular bases of host specificity, and provide new tools for enhancing epidemiological surveillance of strains pathogenic on a given host, or detecting recombinant strains presenting a high potential of emergence through the acquisition of novel genes.

MATERIALS AND METHODS

Bacterial genomes and strains

All strains used in this study are listed in **Table 1**. Genomes published in this study were sequenced using the following procedure. Genomic DNA was prepared from overnight liquid cultures of bacteria previously grown on 10% TSA medium (tryptone at 1.7 g/L, soybean peptone at 0.3 g/L, glucose at 0.25 g/L, NaCl at 0.5 g/L, K₂HPO₄ at 0.5 g/L, agar at 15 g/L, pH 7.2) for 2 days at 28°C. DNA extraction was performed using the QIAamp DNA Minikit (Qiagen) following the manufacturer's instructions. Genomes of strains CFBP6988 and CFBP6996 were shotgun sequenced on HiSeq2000 Illumina platform (GATC Biotech, Germany or Genoscope) using paired-end (ca. 250-bp inserts) and mate-pair (ca. 3-kb inserts) libraries. Genome assembly was performed using a combination of SOAPdenovo (Short Oligonucleotide Analysis Package, version 2.04) (Li *et al.*, 2010), SOAPGapCloser (version 1.12) (Luo *et al.*, 2012) and Velvet (version 1.2.02) (Zerbino & Birney, 2008) assemblers.

Annotation and phylogenomic analyses

Structural and functional annotation of whole genome assemblies was performed using the automated pipeline Eugene-PP (Sallet *et al.*, 2014), using SWISS-PROT as protein database and training protein database (<http://www.uniprot.org/>). Additional functional annotation of all predicted CDS was performed with InterProScan (Quevillon *et al.*, 2005). A presence/absence matrix of orthologs groups was constructed using OrthoMCL on amino acid sequences from all predicted CDS at an inflation index of 1.5 (Li *et al.*, 2003). This matrix was then used for defining core and pan genomes. Phylogenetic trees were constructed using CVTree (Qi *et al.*, 2004) either on the pan proteome or on the core proteome composed of 625 ortholog groups corresponding to genes present as single copy in all 75 genomes used in this study. CDS gains and losses were analysed by using a home-made R script to search for the most parsimonious succession of events explaining the repertoire of ortholog groups at each node of the phylogenetic tree.

Searching for genes monophyletic for *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains

A phylogenetic approach was used to search for genes for which strains from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* form a monophyletic group. For this we selected 3,202 CDS present in all *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains and in at least one outgroup strain from Rep-PCR group 9.2 or 9.4 (Rademaker *et al.*, 2005) and one additional outgroup strain from Rep-PCR group 9.5 or 9.6 (Fig. 1). CDS were aligned using MAFFT with L-INS-I strategy (Kato *et al.*, 2005). Neighbour-joining trees were constructed using APE under the Kimura 80 model (Paradis *et al.*, 2004). CDS monophyletic for all *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains, or alternatively for the NF1 and another lineage (*i.e.* NF2, NF3, or *fuscans*), were retrieved thanks to a home-made R script.

Searching for genes containing *k*-mers specific for *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains

A *k*-mer-based approach was used to search for genes containing short specific sequences present in all strains from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* but absent in other strains. For this, we used SkIf (<https://sourcesup.renater.fr/wiki/skif/start>) with a *k*-mer size of 24 (or 24-mer), and using *X. citri* pv. *fuscans* strain CFBP4834-R genome as reference. The same approach was used to search for genes containing 24-mers absent in strains belonging to *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* but conserved in all other strains from the *X. axonopodis* species complex, using *X. citri* pv. *anacardii* strain CFBP2913 genome as reference.

Recombination and HGT analyses

The 116 genes presenting specific traits of adaptation to common bean were aligned using MAFFT with L-INS-I strategy (Katoh *et al.*, 2005). Intragenic recombination events were then searched using a suite of programs implemented in RDP version 4.16 (Martin *et al.*, 2015), RDP (Martin & Rybicki, 2000), Geneconv (Padidam *et al.*, 1999), MaxChi (Smith, 1992), Chimaera (Posada & Crandall, 2001), Bootscan (Martin *et al.*, 2005) and 3seq (Boni *et al.*, 2007). Default parameters were used for each method except for Bootscan (window = 150, step = 20, neighbor joining trees, 200 replicates, 95% cut-off, J&N model with Ti:Tv = 2, coefficient of variation = 2). The number of synonymous substitution per synonymous site (K_s) was calculated using DNAsp version 5 (Librado & Rozas, 2009). For each gene, the occurrence and dating of HGT events were estimated by comparing K_s values from 28 different pairwise combinations listed in table S3. For example, NF1 and fuscans strains belong to phylogenetically distant strains, thus if the K_s between strains from genetic lineages NF1 and fuscans was the lower than the mean K_s between other lineages, it is indicative of recent HGT between the ancestors of NF1 and fuscans. Direction of events were assessed by comparing the K_s values for outgroups belonging to Rep-PCR groups 9.4 and 9.6 (**Fig.1**). For recombinants, separate analyses were performed for each region on both sides of the recombination point.

REFERENCES

- Alavi SM, Sanjari S, Durand F, Brin C, Manceau C, 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 h. *Applied and Environmental Microbiology* **74**: 3295–3301.
- Alteri CJ, Mobley HLT. 2016. The versatile type VI secretion system. *Microbiology Spectrum* **4**.
- Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. 2015. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Frontiers in Microbiology* **6**: 1–18.
- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**.
- Barak JD, Vancheva T, Lefeuvre P, Jones JB, Timilsina S, Minsavage G V., Vallad GE, Koebnik R. 2016. Whole-Genome Sequences of *Xanthomonas euvesicatoria* Strains Clarify Taxonomy and Reveal a Stepwise Erosion of Type 3 Effectors. *Frontiers in Plant Science* **7**.
- Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, *et al.* 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E1972–E1979.

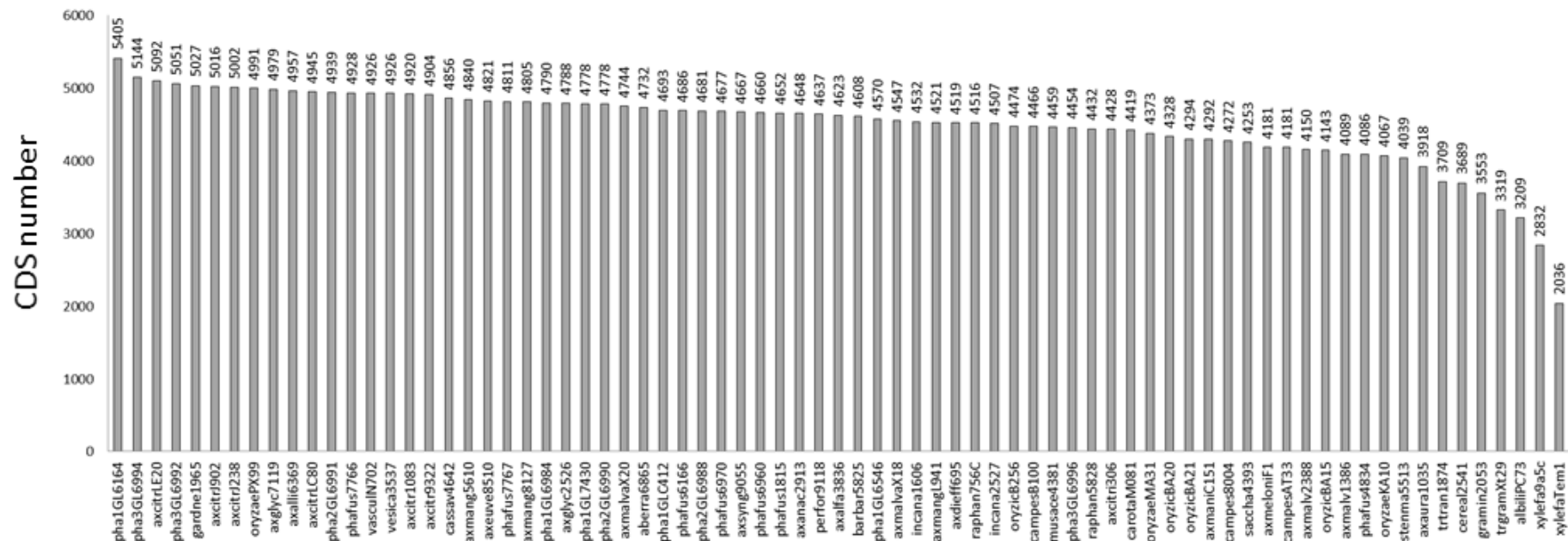
- Bäumler A, Fang FC. 2013.** Host specificity of bacterial pathogens. *Cold Spring Harbor perspectives in medicine* **3**: a010041.
- Belete T, Bastas K. 2017.** Common Bacterial Blight (*Xanthomonas axonopodis* pv. *phaseoli*) of Beans with Special Focus on Ethiopian Condition. *Journal of Plant Pathology & Microbiology* **8**.
- Birch PRJ, Hyman LJ, Taylor R, Opio AF, Bragard C, Toth IK. 1997.** RAPD PCR-based differentiation of *Xanthomonas campestris* pv. *phaseoli* and *Xanthomonas campestris* pv. *phaseoli* var. *fuscans*. *European Journal of Plant Pathology* **103**: 809–814.
- Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M. 2007.** Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**.
- Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli S V., Patil PB, Van Sluys MA, Ryan RP, Meyer DF, Han SW, et al. 2011.** Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *Journal of Bacteriology* **193**: 5450–5464.
- Bolot S, Bodnar AM, Cunnac S, Ortiz E, Szurek B, Noel LD, Arlat M, Jacques M-A, Gagnevin L, Portier P, et al. 2013.** Draft Genome Sequence of the *Xanthomonas cassavae* Type Strain CFBP 4642. *Genome announcements* **1**: e00679-13.
- Boni MF, Posada D, Feldman MW. 2007.** An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**: 1035–1047.
- Briand M, Gaborieau R, Jacques M, Barret M, Boureau T, Gaillard S, Chen NWG. 2016.** Sklf: a tool for rapid identification of genes or regulators of interest [version 1; not peer reviewed]. *F1000Research* **5**: (poster).
- Burkholder WH. 1930.** The bacterial diseases of the bean. Memoirs. *Cornell University Agricultural Experiment Station* **127**: 1–88.
- Büttner D, Bonas U. 2010.** Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Reviews* **34**: 107–133.
- Chan JWYF, Goodwin PH. 1999.** Differentiation of *Xanthomonas campestris* pv. *phaseoli* from *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* by PFGE and RFLP. *European Journal of Plant Pathology* **105**: 867–878.
- Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B. 2016.** Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathology* **65**: 792–806.
- Cunnac S, Bolot S, Forero Serna N, Ortiz E, Szurek B, Noël LD, Arlat M, Jacques M-A, Gagnevin L, Carrere S, et al. 2013.** High quality draft genome sequences of two *Xanthomonas citri* pv. *malvacearum* strains. *Genome announcements* **1**: e00674-13.
- Darrasse A, Bolot S, Serres-Giardi L, Charbit E, Boureau T, Saux MF-L, Briand M, Arlat M, Gagnevin L, Koebnik R, et al. 2013a.** High-Quality Draft Genome Sequences of *Xanthomonas axonopodis* pv. *glycines* Strains CFBP 2526 and CFBP 7119. *Genome announcements* **1**: e01036-13.
- Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, Briand M, Brin C, Cociancich S, Durand K, et al. 2013b.** Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. *BMC Genomics* **14**: 761.
- Dye DW, Bradbury JF, Goto M, Hayward AC, Lelliott RA, Schroth MN. 1980.** International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Review of Plant Pathology* **59**: 153–159.
- EFSA PLH Panel (EFSA Panel on Plant Health). 2015.** Scientific Opinion on the pest categorisation of *Spodoptera littoralis*. *EFSA Journal* **13**: 3987.
- Escalon A, Javegny S, Vernière C, Noël LD, Vital K, Poussier S, Hajri A, Boureau T, Pruvost O, Arlat M, et al. 2013.** Variations in type III effector repertoires, pathological phenotypes and host range of *Xanthomonas citri* pv. *citri*

- pathotypes. *Molecular Plant Pathology* **14**: 483–496.
- Fronzes R, Christie PJ, Waksmas G. 2009.** The structural biology of type IV secretion systems. *Nature Review Microbiology* **7**: 703–14.
- Gagnevin L, Bolot S, Gordon JL, Pruvost O, Vernière C, Robène I, Arlat M, Noël LD, Carrère S, Jacques M-A, et al. 2014.** Draft Genome Sequence of *Xanthomonas axonopodis* pv. *allii* Strain CFBP 6369. *Genome announcements* **2**: e00727-14.
- Gordon JL, Lefeuvre P, Escalon A, Barbe V, Cruveiller S, Gagnevin L, Pruvost O. 2015.** Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. *BMC Genomics* **16**: 1098.
- Guo W, Cui Y ping, Li Y rong, Che Y zhou, Yuan L, Zou L fang, Zou H song, Chen G you. 2012.** Identification of seven *Xanthomonas oryzae* pv. *oryzicola* genes potentially involved in pathogenesis in rice. *Microbiology* **158**: 505–518.
- Hajri A, Brin C, Hunault G, Lardeux F, Lemaire C, Manceau C, Boureau T, 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.
- Hashimi H, Zíková A, Panigrahi AK, Stuart KD, Lukeš J. 2008.** TbRGG1, an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex. *RNA* **14**: 970–980.
- Hausner J, Hartmann N, Jordan M, Büttner D. 2017.** The predicted lytic transglycosylase HpaH from *Xanthomonas campestris* pv. *vesicatoria* associates with the type III secretion system and promotes effector protein translocation. *Infection and Immunity* **85**: e00788-16.
- Hayward AC. 1993.** The hosts of *Xanthomonas*. *Xanthomonas*. 1–119.
- Hersemann L, Wibberg D, Blom J, Goesmann A, Widmer F, Vorholter F-J, Kolliker R. 2017.** Comparative genomics of host adaptive traits in *Xanthomonas translucens* pv. *graminis*. *BMC genomics* **18**: 35.
- Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al. 2016.** Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas*: French Network on *Xanthomonads* (FNX). *Annual Review of Phytopathology* **54**: 163–187.
- Jacques MA, Bolot S, Charbit E, Darrasse A, Briand M, Arlat M, Gagnevin L, Koebnik R, Noël LD, Portier P, et al. 2013.** High-Quality Draft Genome Sequence of *Xanthomonas alfalfae* subsp. *alfalfae* Strain CFBP 3836. *Genome Announcements* **1**: e01035-13.
- Jalan N, Aritua V, Kumar D, Yu F, Jones JB, Graham JH, Setubal JC, Wang N. 2011.** Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. *Journal of Bacteriology* **193**: 6342–6357.
- Katoh K, Kuma KI, Toh H, Miyata T. 2005.** MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research* **33**: 511–518.
- Kimbrel JA, Givan SA, Temple TN, Johnson KB, Chang JH. 2011.** Genome sequencing and comparative analysis of the carrot bacterial blight pathogen, *Xanthomonas hortorum* pv. *carotae* M081, for insights into pathogenicity and applications in molecular diagnostics. *Molecular Plant Pathology* **12**: 580–594.
- Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012.** Selection-driven gene loss in bacteria. *PLoS Genetics* **8**: 1–7.
- Lazo GR, Roffey R, Gabriel DW. 1987.** Pathovars of *Xanthomonas campestris* are distinguishable by restriction fragment-length polymorphism. *International Journal of Systematic Bacteriology* **37**: 214–221.
- Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS, et al. 2005.** The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Research* **33**: 577–586.
- Li L, Stoeckert C, Roos DS. 2003.** OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research* **13**: 2178–2189.

- Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, *et al.* 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**: 265–72.
- Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, *et al.* 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**: 18.
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**: 1–5.
- Martin DP, Posada D, Crandall K a, Williamson C. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS research and human retroviruses* **21**: 98–102.
- Martin D, Rybicki E. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics (Oxford, England)* **16**: 562–563.
- Mhedbi-Hajri N, Darrasse A, Pigné S, Durand K, Fouteau S, Barbe V, Manceau C, Lemaire C, Jacques M-A. 2011. Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. *BMC evolutionary biology* **11**: 67.
- Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, Saux MF Le, Manceau C, Poussier S, Pruvost O, *et al.* 2013. Evolutionary History of the Plant Pathogenic Bacterium *Xanthomonas axonopodis*. *PLoS ONE* **8**: e58474.
- Midha S, Ranjan M, Sharma V, Pinnaka AK, Patil PB. 2012. Genome sequence of *xanthomonas citri* pv. *mangiferaeindicae* strain LMG 941. *Journal of Bacteriology* **194**: 3031.
- Mkandawire ABC, Mabagala RB, Guzmán P, Gepts P, Gilbertson RL. 2004. Genetic Diversity and Pathogenic Variation of Common Blight Bacteria (*Xanthomonas campestris* pv. *phaseoli* and *X. campestris* pv. *phaseoli* var. *fuscans*) Suggests Pathogen Coevolution with the Common Bean. *Phytopathology* **94**: 593–603.
- Moreira LM, Almeida NF, Potnis N, Digiampietri L a, Adi SS, Bortolossi JC, da Silva AC, da Silva AM, de Moraes FE, de Oliveira JC, *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.
- Newell PD, Boyd CD, Sondermann H, O'Toole GA. 2011. A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. *PLoS Biology* **9**: e1000587.
- Noël L, Thieme F, Nennstiel D, Bonas U. 2002. Two Novel Type III-Secreted Proteins of *Xanthomonas campestris* pv. *vesicatoria* Are Encoded within the *hrp* Pathogenicity Island Two Novel Type III-Secreted Proteins of *Xanthomonas campestris* pv. *vesicatoria* Are Encoded within the *hrp* Pathogenicity Island. *Journal of Bacteriology* **184**: 1340–1348.
- Padidam M, Sawyer S, Fauquet CM. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**: 218–225.
- Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.
- Pesce C, Bolot S, Berthelot E, Bragard C, Cunnac S, Saux MF-L, Portier P, Arlat M, Gagnevin L, Jacques M-A, *et al.* 2013. Draft Genome Sequence of *Xanthomonas translucens* pv. *graminis* Pathotype Strain CFBP 2053. *Genome announcements* **3**: e01174-15.
- Pesce C, Bolot S, Cunnac S, Portier P, Saux MF-L, Jacques M-A, Gagnevin L, Arlat M, Noël LD, Carrère S, *et al.* 2015. High-Quality Draft Genome Sequence of the *Xanthomonas translucens* pv. *cerealis* Pathotype Strain CFBP 2541. *Genome announcements* **3**: e01574-14.
- Pieretti I, Royer M, Barbe V, Carrere S, Koebnik R, Cociancich S, Couloux A, Darrasse A, Gouzy J, Jacques MA, *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**: 1–15.
- Posada D, Crandall KA. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of the National Academy of Sciences* **98**: 13757–13762.

- Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, Ryan RP, Sharlach M, Behlau F, Dow JM, Momol M, *et al.* 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC Genomics* **12**: 146.
- Qi J, Luo H, Hao B. 2004. CVTree: A phylogenetic tree reconstruction tool based on whole genomes. *Nucleic Acids Research* **32**: 45–47.
- 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, *et al.* 2005. Comparative and functional genomic analyses of the pathogenicity of phytopathogen. *Genome Research* **15**: 757–767.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: Protein domains identifier. *Nucleic Acids Research* **33**: 116–120.
- Rademaker JLW, Hoste B, Louws FJ, Kersters K, Swings J, Vauterin L, Vauterin P, Bruijn FJ De. 2000. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA – DNA homology studies : Xanthomonas as a model system. *International Journal of Systematic and Evolutionary Microbiology* **50**: 665–677.
- Records AR, Gross DC. 2010. Sensor kinases RetS and LadS regulate Pseudomonas syringae Type VI secretion and virulence factors. *Journal of Bacteriology* **192**: 3584–3596.
- Robène I, Bolot S, Pruvost O, Arlat M, Noël L, Carrère S, Jacques M, Koebnik R, Gagnevin L. 2016. High-Quality Draft Genome Sequences of Two Xanthomonas Pathotype Strains Infecting Aroid Plants. *Genome announcements* **4**: e00902-16.
- Roux B, Bolot S, Guy E, Denancé N, Lautier M, Jardinaud M-F, Fischer-Le Saux M, Portier P, Jacques M-A, Gagnevin L, *et al.* 2015. Genomics and transcriptomics of Xanthomonas campestris species challenge the concept of core type III effectome. *BMC genomics* **16**: 975.
- Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017. Xanthomonas adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors. *BMC Genomics* **18**: 670.
- Sallet E, Gouzy J, Schiex T. 2014. EuGene-PP: A next-generation automated annotation pipeline for prokaryotic genomes. *Bioinformatics* **30**: 2659–2661.
- Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, Furutani A, Ochiai H, Delcher AL, Kelley D, *et al.* 2008. Genome sequence and rapid evolution of the rice pathogen Xanthomonas oryzae pv. oryzae PXO99A. *BMC Genomics* **9**: 204.
- Schwartz AR, Potnis N, Timilsina S, Wilson M, Patané J, Martins J, Minsavage G V., Dahlbeck D, Akhunova A, Almeida N, *et al.* 2015. Phylogenomics of Xanthomonas field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. *Frontiers in Microbiology* **6**: 1–17.
- da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Van Sluys MA, Almeida NF, Alves LMC, *et al.* 2002. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. *Nature* **417**: 459–463.
- Simpson A, Reinach F, Arruda P, Abreu F, Acencio M, Alvarenga R, Alves L, Araya J, Baia G, Baptista C, *et al.* 2000. The genome sequence of the plant pathogen Xylella fastidiosa. *Nature* **406**: 151–157.
- Van Sluys M, de Oliveira M, Monteiro-Vitorello C, Miyaki C, Furlan L, Camargo L, da Silva A, Moon D, Takita M, Lemos E, *et al.* 2003. Comparative Analyses of the Complete Genome Sequences of Pierce's Disease and Citrus Variegated Chlorosis Strains of Xylella fastidiosa. *Journal of bacteriology* **185**: 1018–1026.
- Smith JM. 1992. Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* **34**: 126–129.
- Solé M, Scheibner F, Hoffmeister AK, Hartmann N, Hause G, Rother A, Jordan M, Lautier M, Arlat M, Büttner D. 2015. Xanthomonas campestris pv. vesicatoria secretes proteases and xylanases via the Xps type II secretion system and outer membrane vesicles. *Journal of Bacteriology* **197**: 2879–2893.
- Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. *Nature Reviews Genetics* **16**: 472–482.

- 639 **Studholme DJ, Kemen E, MacLean D, Schornack S, Aritua V, Thwaites R, Grant M, Smith J, Jones JDG. 2010.**
 640 Genome-wide sequencing data reveals virulence factors implicated in banana *Xanthomonas* wilt. *FEMS Microbiology*
 641 *Letters* **310**: 182–192.
- 642 **Studholme DJ, Wasukira A, Paszkiewicz K, Aritua V, Thwaites R, Smith J, Grant M. 2011.** Draft genome sequences
 643 of *xanthomonas sacchari* and two banana-associated *xanthomonads* reveal insights into the *xanthomonas* group 1
 644 clade. *Genes* **2**: 1050–1065.
- 645 **Su J, Zou X, Huang L, Bai T, Liu S, Yuan M, Chou S-H, He Y-W, Wang H, He J. 2016.** DgcA, a diguanylate cyclase from
 646 *Xanthomonas oryzae* pv. *oryzae* regulates bacterial pathogenicity on rice. *Scientific reports* **6**: 25978.
- 647 **Taghavi S, Garafola C, Monchy S, Newman L, Hoffman A, Weyens N, Barac T, Vangronsveld J, Van Der Lelie D.**
 648 **2009.** Genome survey and characterization of endophytic bacteria exhibiting a beneficial effect on growth and
 649 development of poplar trees. *Applied and Environmental Microbiology* **75**: 748–757.
- 650 **Teper D, Burstein D, Salomon D, Gershovitz M, Pupko T, Sessa G. 2016.** Identification of novel *Xanthomonas*
 651 *euvasicatoria* type III effector proteins by a machine-learning approach. *Molecular Plant Pathology* **17**: 398–411.
- 652 **Thieme F, Koebnik R, Bekel T, Berger C, Boch J, Büttner D, Caldana C, Gaigalat L, Goesmann A, Kay S. 2005.**
 653 Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv.
 654 *vasicatoria* revealed by the complete genome sequence. *Journal of bacteriology* **187**: 7254–7266.
- 655 **Vauterin L, Hoste B, Kersters K, Swings J. 1995.** Reclassification of *Xanthomonas*. *International Journal of Systematic*
 656 *Bacteriology* **45**: 472–489.
- 657 **Vorhölter F-J, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Patschkowski T, Rückert C, Schmid**
 658 **J, et al. 2008.** The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of
 659 metabolic pathways involved in xanthan biosynthesis. *Journal of Biotechnology* **134**: 33–45.
- 660 **Wallden K, Rivera-Calzada A, Waksman G. 2010.** Type IV secretion systems: versatility and diversity in function.
 661 *Cellular Microbiology* **12**: 1203–1212.
- 662 **Wichmann F, Vorhölter F-J, Herseemann L, Widmer F, Blom J, Niehaus K, Reinhard S, Conradin C, Kölliker R. 2013.**
 663 The noncanonical type III secretion system of *Xanthomonas translucens* pv. *graminis* is essential for forage grass
 664 infection. *Molecular Plant Pathology* **14**: 576–588.
- 665 **Xie G, Bonner CA, Brettin T, Gottardo R, Keyhani NO, Jensen RA. 2003.** Lateral gene transfer and ancient paralogy
 666 of operons containing redundant copies of tryptophan-pathway genes in *Xylella* species and in heterocystous
 667 cyanobacteria. *Genome biology* **4**: R14.
- 668 **Young JM, Park DC, Shearman HM, Fargier E. 2008.** A multilocus sequence analysis of the genus *Xanthomonas*.
 669 *Systematic and Applied Microbiology* **31**: 366–377.
- 670 **Zaumeyer WJ. 1930.** The bacterial blight of beans caused by *Bacterium phaseoli*. *Technical Bulletin US Department of*
 671 *Agriculture*: 1–35.
- 672 **Zerbino DR, Birney E. 2008.** Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome*
 673 *Research* **18**: 821–829.



Additional file 1 Fig. S1. CDS content per strain

Additional file 3 Table S2. Ks analysis on 100 genes presenting characteristics unique to CBB agents.

1	2	m00100560		m00100580a		m00100580b		m00100590		m00101230		m00101980	
		mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
NF1	NF2,NF3,fuscans	1,17E-02	5,83E-03	3,63E-02	1,52E-02	2,33E-02	1,72E-02	2,74E-02	8,77E-03	2,85E-02	8,37E-04	4,65E-02	1,68E-03
NF1	NF2	5,10E-03	1,54E-10	1,52E-02	3,68E-10	4,36E-02	1,04E-09	2,65E-02	0,00E+00	2,89E-02	3,92E-04	4,57E-02	3,92E-04
NF1	NF3	5,80E-03	0,00E+00	2,54E-02	6,59E-10	4,01E-02	0,00E+00	4,29E-02	0,00E+00	2,71E-02	4,41E-04	4,41E-02	3,92E-04
NF1	fuscans	1,71E-02	1,40E-04	5,00E-02	0,00E+00	7,40E-03	2,85E-10	2,11E-02	0,00E+00	2,89E-02	3,92E-04	4,80E-02	3,92E-04
NF1	X.citri pv. anacardii and aurantifolii	5,10E-03	0,00E+00	2,81E-02	8,50E-04	4,40E-02	4,00E-04	3,19E-02	8,07E-10	3,24E-02	3,92E-04	4,41E-02	3,92E-04
X.phaseoli pv. manihotis	NF2,NF3,fuscans	na	na	5,63E-02	1,10E-02	5,31E-01	2,37E-03	na	na	4,66E-02	7,58E-04	4,76E-02	1,65E-03
X.phaseoli pv. manihotis	NF2	na	na	6,43E-02	0,00E+00	5,28E-01	0,00E+00	na	na	4,70E-02	0,00E+00	4,67E-02	0,00E+00
X.phaseoli pv. manihotis	NF3	na	na	7,16E-02	0,00E+00	5,35E-01	0,00E+00	na	na	4,52E-02	0,00E+00	4,52E-02	0,00E+00
X.phaseoli pv. manihotis	fuscans	na	na	4,64E-02	0,00E+00	5,30E-01	0,00E+00	na	na	4,70E-02	6,59E-10	4,91E-02	9,31E-10
X.phaseoli pv. manihotis	X.citri pv. anacardii and aurantifolii	na	na	7,44E-02	9,00E-04	5,30E-01	7,00E-04	na	na	4,52E-02	0,00E+00	4,36E-02	0,00E+00
X.phaseoli pv. syngonii and diffebachiae	NF2,NF3,fuscans	3,90E-02	2,69E-03	5,91E-02	1,11E-02	3,29E-02	1,29E-02	2,16E-01	9,04E-03	4,41E-02	3,79E-04	4,88E-02	2,01E-03
X.phaseoli pv. syngonii and diffebachiae	NF2	4,04E-02	1,80E-03	6,71E-02	9,00E-04	4,73E-02	1,45E-03	2,05E-01	3,40E-03	4,43E-02	6,59E-10	4,79E-02	1,20E-03
X.phaseoli pv. syngonii and diffebachiae	NF3	4,21E-02	1,85E-03	7,44E-02	9,00E-04	4,58E-02	2,15E-03	2,26E-01	3,55E-03	4,34E-02	0,00E+00	4,64E-02	1,15E-03
X.phaseoli pv. syngonii and diffebachiae	fuscans	3,71E-02	1,45E-03	4,91E-02	9,00E-04	2,12E-02	2,75E-03	2,16E-01	6,95E-03	4,43E-02	6,59E-10	5,03E-02	1,15E-03
X.phaseoli pv. syngonii and diffebachiae	X.citri pv. anacardii and aurantifolii	4,21E-02	1,85E-03	7,72E-02	1,31E-03	4,69E-02	2,18E-03	1,99E-01	3,40E-03	4,34E-02	0,00E+00	4,48E-02	1,20E-03
NF2	NF3	2,30E-03	0,00E+00	1,01E-02	1,16E-10	1,98E-02	2,33E-10	2,65E-02	0,00E+00	5,90E-03	0,00E+00	2,90E-02	0,00E+00
NF2	fuscans	1,60E-02	1,40E-04	5,53E-02	0,00E+00	4,58E-02	1,32E-09	2,11E-02	0,00E+00	0,00E+00	0,00E+00	6,60E-03	1,84E-10
NF3	fuscans	1,68E-02	1,05E-04	6,62E-02	0,00E+00	4,08E-02	1,54E-09	2,65E-02	0,00E+00	5,90E-03	0,00E+00	5,10E-03	1,93E-10
X.citri pv. anacardii and aurantifolii	NF2,NF3,fuscans	9,88E-03	6,66E-03	4,02E-02	2,71E-02	3,08E-02	1,71E-02	2,28E-02	9,90E-03	1,04E-02	1,05E-03	8,97E-03	2,24E-03
X.citri pv. anacardii and aurantifolii	NF2	2,30E-03	0,00E+00	1,27E-02	8,50E-04	6,45E-03	3,50E-04	5,20E-03	0,00E+00	1,10E-02	2,33E-10	7,30E-03	0,00E+00
X.citri pv. anacardii and aurantifolii	NF3	3,10E-03	0,00E+00	9,25E-03	8,50E-04	1,95E-02	3,50E-04	3,19E-02	8,07E-10	8,50E-03	0,00E+00	5,90E-03	0,00E+00
X.citri pv. anacardii and aurantifolii	fuscans	1,60E-02	1,40E-04	6,53E-02	9,50E-04	4,62E-02	3,50E-04	2,65E-02	0,00E+00	1,10E-02	1,16E-10	1,10E-02	1,16E-10
NF1	X.phaseoli pv. manihotis	na	na	5,53E-02	0,00E+00	5,33E-01	0,00E+00	na	na	2,01E-02	4,41E-04	1,67E-02	3,43E-04
NF1 and X.phaseoli pv. manihotis	X.phaseoli pv. syngonii and diffebachiae	4,08E-02	1,80E-03	5,36E-02	1,80E-02	1,02E-01	1,89E-01	2,02E-01	6,80E-03	1,92E-02	4,09E-03	2,07E-02	5,73E-03
NF1	NF1	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	4,80E-04	3,92E-04	4,20E-04	3,43E-04
NF2	NF2	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
NF3	NF3	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
fuscans	fuscans	1,14E-04	1,81E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
HGT between the NF1 and the NF2, NF3 and/or fuscans lineages and associated Ks values		NF2 to NF1		Undetected		fuscans to NF1		fuscans to NF1		Undetected		Undetected	
HGT between the NF2, NF3 and/or fuscans lineages and associated Ks values		Undetected		Undetected		Undetected		Undetected		Undetected		Undetected	

Pairwise Ks values were calculated between strains from the lineages and/or pathovars listed in column 1 versus column 2

CDS are identified by accession numbers from strain CFBP6546 in bold

The different fractions from each recombinant CDS are indicated by different letters (a, b or c) at the end of CDS accession numbers

HGT: Horizontal Gene Transfer

SD: Standard Deviation

na: not applicable due to the absence of corresponding CDS in one lineage or another

Additional file 3 Table S2. (continued)

m00102200		m00104250a		m00104250b		m00104520		m00104530		m00104540		m00104640		m00104690		m00105290	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
1,52E-03	1,65E-03	5,70E-02	7,93E-02	8,44E-02	1,36E-02	4,70E-02	3,40E-03	3,39E-03	4,75E-03	3,81E-03	4,52E-03	3,22E-03	2,52E-03	2,56E-03	2,80E-03	9,54E-04	1,11E-03
3,30E-03	7,13E-11	1,98E-01	3,73E-09	9,20E-02	0,00E+00	4,82E-02	0,00E+00	1,18E-02	0,00E+00	5,50E-03	0,00E+00	6,88E-03	3,43E-04	5,00E-03	8,23E-11	2,50E-03	4,12E-11
3,30E-03	7,13E-11	4,56E-02	9,31E-10	5,97E-02	1,14E-09	5,24E-02	0,00E+00	2,90E-03	0,00E+00	1,10E-02	0,00E+00	4,68E-03	3,43E-04	6,10E-03	4,24E-04	1,63E-03	6,13E-04
0,00E+00	0,00E+00	1,40E-03	6,17E-11	9,16E-02	2,06E-03	4,41E-02	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,02E-03	3,92E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00
5,50E-03	1,10E-03	4,60E-02	3,50E-04	6,37E-02	5,50E-04	4,83E-02	4,15E-03	8,80E-03	0,00E+00	7,10E-03	8,23E-11	5,43E-03	8,25E-04	5,85E-03	1,65E-03	3,10E-03	1,90E-03
6,07E-02	5,98E-04	7,77E-02	7,12E-02	1,16E-02	1,16E-02	1,68E-01	2,02E-03	1,13E-01	1,43E-03	1,10E-01	1,71E-03	4,36E-02	5,10E-04	5,05E-02	3,25E-04	7,23E-02	1,20E-03
6,01E-02	0,00E+00	2,06E-01	0,00E+00	1,52E-01	0,00E+00	1,65E-01	0,00E+00	1,13E-01	0,00E+00	1,09E-01	1,32E-09	4,44E-02	0,00E+00	5,06E-02	0,00E+00	7,40E-02	0,00E+00
6,01E-02	0,00E+00	6,17E-02	0,00E+00	1,25E-01	0,00E+00	1,70E-01	0,00E+00	1,10E-01	0,00E+00	1,14E-01	1,32E-09	4,29E-02	0,00E+00	5,00E-02	4,24E-04	7,31E-02	6,60E-04
6,13E-02	0,00E+00	2,98E-02	3,29E-10	1,52E-01	0,00E+00	1,70E-01	1,86E-09	1,13E-01	2,95E-09	1,10E-01	1,32E-09	4,36E-02	0,00E+00	5,06E-02	6,59E-10	7,13E-02	0,00E+00
6,13E-02	0,00E+00	6,21E-02	3,50E-04	1,29E-01	6,00E-04	1,65E-01	4,80E-03	1,17E-01	0,00E+00	1,10E-01	9,50E-04	4,37E-02	7,50E-04	4,97E-02	0,00E+00	7,40E-02	1,40E-03
6,19E-02	1,34E-03	1,66E-01	5,14E-02	1,13E-01	1,56E-02	1,52E-01	6,02E-03	9,75E-02	8,74E-03	1,05E-01	1,07E-02	3,90E-02	9,48E-04	4,98E-02	8,02E-04	9,23E-02	1,15E-03
6,13E-02	1,20E-03	7,88E-02	1,87E-02	1,08E-01	1,21E-02	1,41E-01	0,00E+00	1,00E-01	6,70E-03	1,05E-01	1,03E-02	3,98E-02	8,00E-04	5,02E-02	4,50E-04	9,33E-02	1,40E-03
6,13E-02	1,20E-03	2,01E-01	1,73E-02	1,33E-01	8,10E-03	1,55E-01	1,86E-09	9,67E-02	6,65E-03	1,06E-01	1,08E-02	3,83E-02	7,50E-04	4,87E-02	6,18E-04	9,24E-02	1,55E-03
6,25E-02	1,20E-03	1,89E-01	1,88E-02	1,07E-01	1,23E-02	1,55E-01	0,00E+00	9,67E-02	1,00E-02	1,05E-01	1,08E-02	3,90E-02	8,00E-04	5,02E-02	4,50E-04	9,19E-02	0,00E+00
6,22E-02	1,56E-03	2,02E-01	1,78E-02	1,34E-01	9,37E-03	1,50E-01	4,80E-03	1,03E-01	6,70E-03	1,05E-01	9,91E-03	3,90E-02	1,10E-03	4,93E-02	4,50E-04	9,33E-02	1,98E-03
0,00E+00	0,00E+00	1,95E-01	0,00E+00	5,07E-02	0,00E+00	1,18E-02	0,00E+00	8,80E-03	0,00E+00	1,18E-02	0,00E+00	3,70E-03	0,00E+00	4,47E-03	3,77E-04	1,63E-03	6,13E-04
3,30E-03	9,20E-11	1,96E-01	0,00E+00	4,94E-03	1,82E-03	1,18E-02	0,00E+00	1,18E-02	0,00E+00	5,50E-03	0,00E+00	5,90E-03	0,00E+00	5,00E-03	0,00E+00	2,50E-03	0,00E+00
3,30E-03	9,20E-11	4,41E-02	0,00E+00	5,04E-02	1,96E-03	1,57E-02	3,29E-10	2,90E-03	4,12E-11	1,10E-02	0,00E+00	3,70E-03	8,23E-11	6,10E-03	4,24E-04	1,63E-03	6,13E-04
4,99E-03	1,23E-03	7,05E-02	6,98E-02	3,81E-02	1,67E-02	1,67E-02	8,13E-03	9,52E-03	3,12E-03	7,82E-03	1,81E-03	4,56E-03	1,37E-03	5,91E-03	1,52E-03	2,91E-03	1,95E-03
4,40E-03	1,10E-03	1,95E-01	1,40E-03	4,91E-02	5,50E-04	1,98E-02	7,95E-03	1,48E-02	0,00E+00	6,30E-03	0,00E+00	6,60E-03	0,00E+00	5,85E-03	1,65E-03	3,10E-03	1,90E-03
4,40E-03	1,10E-03	6,80E-03	1,10E-03	7,85E-03	5,50E-04	1,58E-02	7,95E-03	5,90E-03	0,00E+00	1,10E-02	8,00E-04	2,90E-03	0,00E+00	6,12E-03	9,30E-04	2,27E-03	1,96E-03
5,50E-03	1,10E-03	4,45E-02	3,50E-04	4,65E-02	2,04E-03	1,58E-02	7,95E-03	8,80E-03	0,00E+00	7,10E-03	1,16E-10	4,40E-03	7,00E-04	5,85E-03	1,65E-03	3,10E-03	1,90E-03
6,13E-02	0,00E+00	3,13E-02	0,00E+00	7,22E-02	0,00E+00	1,27E-01	0,00E+00	1,13E-01	1,86E-09	1,10E-01	1,86E-09	4,47E-02	3,92E-04	5,06E-02	0,00E+00	7,13E-02	0,00E+00
5,51E-02	1,65E-02	1,90E-01	2,05E-02	1,24E-01	2,62E-02	1,43E-01	1,55E-02	9,32E-02	2,43E-02	9,76E-02	2,39E-02	3,54E-02	1,06E-02	4,49E-02	1,17E-02	8,45E-02	1,72E-02
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	4,20E-04	3,43E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	5,33E-04	3,77E-04	8,00E-04	5,66E-04
0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,49E-03	2,35E-03	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
fusans to NF1		fusans to NF1		Undetected		fusans to NF1		fusans to NF1		fusans to NF1		fusans to NF1		fusans to NF1		fusans to NF1	
NF2 and NF3		Undetected		fusans and NF2		Undetected		Undetected		Undetected		Undetected		Undetected		Undetected	

m00105330		m00105340		m00105390		m00105750		m00107500a		m00107500b		m00138140		m00109610a		m00109610b	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
6,13E-03	1,52E-03	5,85E-03	1,55E-03	2,45E-03	1,25E-03	1,01E-01	2,65E-03	2,45E-03	2,64E-03	1,24E-01	1,06E-03	3,44E-03	2,22E-03	4,36E-02	7,24E-04	2,08E-03	2,27E-03
5,30E-03	0,00E+00	3,64E-03	8,80E-04	3,80E-03	9,20E-11	9,63E-02	0,00E+00	5,30E-03	0,00E+00	1,25E-01	2,80E-04	4,60E-03	0,00E+00	4,36E-02	1,04E-09	4,00E-03	0,00E+00
8,90E-03	0,00E+00	5,76E-03	8,45E-04	3,80E-03	9,20E-11	1,03E-01	0,00E+00	5,30E-03	0,00E+00	1,22E-01	2,80E-04	6,80E-03	1,16E-10	4,24E-02	0,00E+00	5,00E-03	8,23E-11
5,30E-03	1,30E-10	6,84E-03	8,80E-04	1,30E-03	0,00E+00	1,03E-01	4,93E-09	0,00E+00	0,00E+00	1,24E-01	2,80E-04	1,50E-03	0,00E+00	4,42E-02	0,00E+00	0,00E+00	0,00E+00
5,05E-02	2,73E-02	8,96E-03	1,38E-03	5,10E-03	1,30E-03	9,95E-02	0,00E+00	7,25E-03	6,50E-04	1,25E-01	4,92E-04	4,20E-03	4,00E-04	4,42E-02	6,00E-04	9,05E-03	2,05E-03
8,28E-02	1,64E-03	5,79E-02	5,98E-04	9,70E-03	6,48E-04	1,01E-01	2,65E-03	1,02E-01	2,46E-03	1,24E-01	6,32E-04	4,56E-02	6,74E-04	4,24E-02	7,24E-04	4,12E-02	1,37E-03
8,37E-02	0,00E+00	5,73E-02	0,00E+00	9,00E-03	0,00E+00	9,63E-02	0,00E+00	9,80E-02	0,00E+00	1,24E-01	0,00E+00	4,52E-02	0,00E+00	4,24E-02	0,00E+00	4,21E-02	0,00E+00
7,98E-02	0,00E+00	5,73E-02	0,00E+00	9,00E-03	0,00E+00	1,03E-01	2,28E-09	1,01E-01	0,00E+00	1,23E-01	0,00E+00	4,68E-02	0,00E+00	4,12E-02	4,66E-10	4,32E-02	6,59E-10
8,37E-02	9,31E-10	5,85E-02	6,59E-10	1,03E-02	0,00E+00	1,03E-01	0,00E+00	1,04E-01	1,32E-09	1,24E-01	1,86E-09	4,52E-02	0,00E+00	4,30E-02	8,07E-10	4,00E-02	4,66E-10
5,30E-02	1,51E-02	5,97E-02	1,15E-03	1,10E-02	6,50E-04	9,95E-02	0,00E+00	1,03E-01	7,50E-04	1,25E-01	3,50E-04	4,56E-02	4,00E-04	4,30E-02	6,00E-04	4,64E-02	1,05E-03
7,99E-02	3,36E-03	6,20E-02	2,93E-03	1,23E-02	1,41E-03	1,04E-01	4,12E-03	9,78E-02	4,35E-03	1,29E-01	2,25E-03	4,50E-02	1,89E-03	4,21E-02	2,22E-03	4,39E-02	2,08E-03
8,08E-02	2,95E-03	6,14E-02	2,85E-03	1,16E-02	1,25E-03	9,95E-02	3,15E-03	9,43E-02	3,70E-03	1,30E-01	2,25E-03	4,44E-02	1,60E-03	4,21E-02	2,10E-03	4,48E-02	1,60E-03
7,69E-02	2,90E-03	6,14E-02	2,85E														

Additional file 3 Table S2. (continued)

m00109620a		m00109620b		m00109620c		m00109670		m00109990		m00110970		m00111160		m00112070		m00136550	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
3,00E-03	3,25E-03	5,68E-03	8,89E-03	5,69E-03	3,12E-03	5,83E-03	7,40E-03	5,06E-02	0,00E+00	2,08E-03	2,33E-03	5,78E-03	2,83E-03	4,35E-02	1,29E-03	3,54E-02	0,00E+00
6,10E-03	8,23E-11	0,00E+00	0,00E+00	0,00E+00	0,00E+00	5,20E-03	1,01E-10	5,06E-02	1,14E-09	5,40E-03	0,00E+00	1,03E-02	0,00E+00	4,27E-02	1,14E-09	3,54E-02	6,59E-10
6,90E-03	0,00E+00	2,11E-02	0,00E+00	7,40E-03	0,00E+00	1,89E-02	4,66E-10	5,06E-02	1,14E-09	3,60E-03	9,20E-11	6,80E-03	1,16E-10	4,17E-02	8,07E-10	3,54E-02	6,59E-10
0,00E+00	0,00E+00	1,50E-03	3,67E-03	7,40E-03	2,85E-10	5,00E-04	0,00E+00	5,06E-02	6,59E-10	0,00E+00	0,00E+00	3,40E-03	0,00E+00	4,47E-02	1,32E-09	3,54E-02	4,66E-10
6,10E-03	0,00E+00	2,30E-01	0,00E+00	1,12E-02	3,75E-03	4,95E-03	7,50E-04	5,19E-02	0,00E+00	1,10E-02	1,85E-03	5,10E-03	1,70E-03	4,17E-02	0,00E+00	3,54E-02	6,59E-10
9,03E-02	1,73E-03	2,28E-01	7,51E-03	5,95E-02	3,37E-03	1,00E-01	7,24E-04	5,33E-02	0,00E+00	5,02E-02	2,49E-03	1,96E-02	2,87E-03	4,45E-02	1,29E-03	3,71E-02	0,00E+00
9,26E-02	0,00E+00	2,30E-01	0,00E+00	5,33E-02	9,31E-10	1,00E-01	0,00E+00	5,33E-02	9,31E-10	4,66E-02	0,00E+00	2,42E-02	0,00E+00	4,37E-02	8,07E-10	3,71E-02	0,00E+00
9,16E-02	0,00E+00	2,16E-01	0,00E+00	6,13E-02	0,00E+00	9,91E-02	0,00E+00	5,33E-02	9,31E-10	4,86E-02	0,00E+00	2,07E-02	0,00E+00	4,27E-02	8,07E-10	3,71E-02	0,00E+00
8,87E-02	0,00E+00	2,32E-01	5,00E-03	6,13E-02	0,00E+00	1,01E-01	0,00E+00	5,33E-02	0,00E+00	5,24E-02	0,00E+00	1,72E-02	2,33E-10	4,57E-02	0,00E+00	3,71E-02	0,00E+00
9,01E-02	3,40E-03	0,00E+00	0,00E+00	6,53E-02	4,00E-03	9,88E-02	9,00E-04	5,46E-02	0,00E+00	4,86E-02	1,95E-03	1,90E-02	1,75E-03	4,27E-02	0,00E+00	3,71E-02	0,00E+00
8,99E-02	3,16E-03	2,35E-01	1,04E-02	4,37E-02	3,29E-03	9,76E-02	7,72E-03	4,86E-02	2,00E-03	4,92E-02	2,67E-03	1,79E-02	3,37E-03	4,46E-02	2,42E-03	3,97E-02	4,30E-03
9,16E-02	1,95E-03	2,37E-01	7,15E-03	3,77E-02	0,00E+00	9,74E-02	7,70E-03	4,86E-02	2,00E-03	4,57E-02	9,50E-04	2,25E-02	1,75E-03	4,37E-02	2,00E-03	3,97E-02	4,30E-03
9,36E-02	1,95E-03	2,23E-01	7,00E-03	4,55E-02	6,59E-10	9,68E-02	7,70E-03	4,86E-02	2,00E-03	4,76E-02	1,00E-03	1,90E-02	1,75E-03	4,27E-02	2,00E-03	3,97E-02	4,30E-03
8,77E-02	1,90E-03	2,39E-01	8,76E-03	4,55E-02	6,59E-10	9,80E-02	7,70E-03	4,86E-02	2,00E-03	5,15E-02	9,50E-04	1,55E-02	1,75E-03	4,58E-02	2,05E-03	3,97E-02	4,30E-03
8,92E-02	3,92E-03	5,25E-03	5,25E-03	4,94E-02	3,90E-03	5,59E-02	7,72E-03	4,99E-02	2,00E-03	4,76E-02	2,17E-03	1,72E-02	2,47E-03	4,27E-02	2,00E-03	3,97E-02	4,30E-03
1,13E-02	0,00E+00	2,11E-02	0,00E+00	7,40E-03	0,00E+00	1,78E-02	2,33E-10	2,50E-03	2,91E-11	1,80E-03	0,00E+00	3,40E-03	4,12E-11	9,00E-04	0,00E+00	0,00E+00	0,00E+00
6,10E-03	8,23E-11	1,50E-03	3,67E-03	7,40E-03	1,65E-10	4,70E-03	8,23E-11	0,00E+00	0,00E+00	5,40E-03	0,00E+00	6,80E-03	1,65E-10	3,80E-03	8,23E-11	0,00E+00	0,00E+00
6,90E-03	0,00E+00	2,26E-02	3,78E-03	1,49E-02	0,00E+00	1,84E-02	4,66E-10	2,50E-03	0,00E+00	3,60E-03	0,00E+00	3,40E-03	8,23E-11	2,80E-03	1,01E-10	0,00E+00	0,00E+00
7,48E-03	2,11E-03	2,28E-01	7,51E-03	1,52E-02	5,30E-03	7,57E-03	5,55E-03	6,30E-03	0,00E+00	8,83E-03	3,01E-03	4,07E-03	3,29E-03	5,52E-03	1,21E-03	0,00E+00	0,00E+00
6,90E-03	0,00E+00	2,30E-01	0,00E+00	1,12E-02	3,75E-03	4,95E-03	7,50E-04	6,30E-03	0,00E+00	5,45E-03	1,85E-03	5,10E-03	1,70E-03	3,80E-03	4,12E-11	0,00E+00	0,00E+00
1,13E-02	0,00E+00	2,16E-01	0,00E+00	1,12E-02	3,75E-03	1,76E-02	8,00E-04	6,30E-03	0,00E+00	7,25E-03	1,85E-03	5,10E-03	1,70E-03	3,80E-03	4,12E-11	0,00E+00	0,00E+00
6,10E-03	0,00E+00	2,32E-01	5,00E-03	1,87E-02	3,75E-03	4,40E-03	8,00E-04	6,30E-03	0,00E+00	1,10E-02	1,85E-03	5,10E-03	1,70E-03	3,80E-03	4,12E-11	0,00E+00	0,00E+00
8,87E-02	0,00E+00	2,30E-01	0,00E+00	5,33E-02	0,00E+00	1,02E-01	0,00E+00	1,27E-02	0,00E+00	5,24E-02	0,00E+00	1,37E-02	0,00E+00	4,70E-03	0,00E+00	5,00E-03	0,00E+00
7,96E-02	1,83E-02	1,98E-01	8,66E-02	3,77E-02	3,16E-03	8,85E-02	2,36E-02	1,14E-02	2,55E-03	4,62E-02	1,17E-02	1,09E-02	3,08E-03	1,34E-02	2,72E-03	1,86E-02	3,41E-03
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	3,00E-03	4,74E-03	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
fusans to NF1		NF2 to NF1		NF2 to NF1		fusans to NF1		Undetected		fusans to NF1		fusans to NF1		Undetected		Undetected	
Undetected		Undetected		Undetected		fusans and NF2		fusans and NF3		Undetected		Undetected		Undetected		Undetected	

m00136050		m00135980		m00135970		m00135880		m00135870a		m00135870b		m00135840		m00130690		m00130670	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
4,72E-03	1,02E-03	2,20E-03	1,07E-03	6,28E-03	2,04E-03	3,25E-03	3,52E-03	3,49E-03	3,78E-03	9,26E-02	3,22E-03	3,14E-02	9,69E-04	2,31E-03	2,49E-03	1,18E-03	2,15E-03
3,20E-03	0,00E+00	2,60E-03	5,04E-11	4,10E-03	0,00E+00	6,70E-03	1,65E-10	7,80E-03	0,00E+00	9,28E-02	0,00E+00	3,09E-02	8,07E-10	5,00E-03	8,23E-11	0,00E+00	0,00E+00
4,30E-03	8,23E-11	3,90E-03	0,00E+00	4,10E-03	0,00E+00	7,40E-03	0,00E+00	7,33E-03	3,30E-04	9,81E-02	0,00E+00	3,32E-02	6,59E-10	5,00E-03	8,23E-11	5,10E-03	1,54E-10
5,56E-03	3,85E-04	1,30E-03	0,00E+00	8,14E-03	3,85E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	9,01E-02	0,00E+00	3,09E-02	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
4,59E-02	5,50E-04	7,20E-03	2,00E-03	4,65E-03	5,50E-04	6,70E-03	7,00E-04	7,45E-03	3,50E-04	9,28E-02	0,00E+00	3,21E-02	1,15E-03	7,55E-03	2,55E-03	5,10E-03	0,00E+00
4,86E-02	1,60E-03	na	na	3,97E-02	9,81E-04	6,69E-02	9,38E-04	6,90E-02	1,49E-03	9,52E-02	3,22E-03	4,31E-02	2,42E-03	4,52E-02	2,23E-03	2,20E-02	2,23E-03
4,76E-02	0,00E+00	na	na	4,01E-02	0,00E+00	6,55E-02	0,00E+00	6,73E-02	0,00E+00	9,54E-02	0,00E+00	4,26E-02	0,00E+00	4,64E-02	0,00E+00	2,08E-02	0,00E+00
4,64E-02	0,00E+00	na	na	4,11E-02	0,00E+00	6,63E-02	0,00E+00	6,76E-02	3,77E-04	1,01E-01	0,00E+00	4,74E-02	0,00E+00	4,11E-02	0,00E+00	2,61E-02	0,00E+00
5,01E-02	3,85E-04	na	na	3,88E-02	3,85E-04	6,77E-02	0,00E+00	7,04E-02	0,00E+00	9,28E-02	0,00E+00	4,14E-02	4,66E-10	4,64E-02	0,00E+00	2,08E-02	0,00E+00
1,47E-02	5,50E-04	na	na	3,74E-02	5,50E-04	6,63E-02	0,00E+00	7,08E-02	4,00E-04	9,54E-02	0,00E+00	4,38E-02	2,40E-03	4,38E-02	2,65E-03	2,61E-02	0,00E+00
4,66E-02	1,61E-03	5,41E-02	1,40E-03	4,13E-02	1,12E-03	7,26E-02	1,60E-03	6,94E-02	1,55E-03	9,66E-02	5,16E-03	4,61E-02	2,35E-03	5,32E-02	3,51E-03	1,67E-02	2,23E-03
4,53E-02	1,15E-03	5,41E-02	1,40E-03	4,17E-02	5,50E-04	7,10E-02	4,00E-04	6,77E-02	4,00E-04	9,68E-02	4,00E-03	4,50E-02	1,20E-03	5,44E-02	2,70E-03	1,55E-02	0,00E+00
4,53E-02	0,00E+00	5,41E-02	1,40E-03	4,28E-02	5,50E-04	7,18E-02	3,50E-04	6,80E-02	5,50E-04	1,02E-01	4,05E-03	4,98E-02	1,20E-03	4,91E-02	2,65E-03	2,08E-02	0,00E+00
4,77E-02	1,22E-03	5,41E-02	1,40E-03	4,05E-02	6,26E-04	7,36E-02	1,50E-03	7,08E-02	4,00E-04	9,41E-02	4,00E-03	4,50E-02	1,20E-03	5,44E-02	2,70E-03	1,55E-02	0,00E+00
2,12E-02	3,39E-03	5,20E-02	1,21E-03	3,90E-02	7,78E-04	7,18E-02	3,50E-04	7,12E-02	5,66E-04	9,68E-02	4,00E-03	4,68E-02	3,23E-03	5,17E-02	3,78E-03	2,08E-02	0,00E+00
5,40E-03	0,00E+00	6,60E-03	0,00E+00	8,30E-03	0,00E+00	3,30E-03	0,00E+00	3,77E-03	3,77E-04	9,50E-03	0,00E+00	9,00E-03	1,16E-10	1,01E-02	1,16E-10	5,10E-03	0,00E+00
4,46E-03	3,85E-04	3,90E-03	0,00E+00	9,16E-03	3,50E-04	6,70E-03	1,16E-10	7,80E-03	0,00E+00	2,40E-03	0,00E+00	4,50E-03	0,00E+00	5,00E-03	0,00E+00	0,00E+00	0,00E+00
5,56E-03	3,85E-04	5,20E-03	1,16E-10	1,23E-02	3,50E-04	7,40E-03	1,65E-10	7,33E-03	3,30E-04	7,10E-03	1,84E-10	9,00E-03	0,00E+00	5,00E-03	0,00E+00	5,10E-03	1,93E-10</

Additional file 3 Table S2. (continued)

m00130660		m00130650		m00121810		m00121850		m00121900		m00123210		m00126530		m00126540		m00127140	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
2,22E-03	2,39E-03		1,93E-03	3,12E-03	3,38E-03	1,07E-02	1,15E-02	1,15E-03	1,33E-03	1,73E-02	1,03E-03	2,28E-02	7,79E-04	4,68E-03	1,77E-03	3,42E-03	3,69E-03
4,80E-03	0,00E+00	6,60E-04	1,32E-03	7,20E-03	1,84E-10	5,70E-03	1,16E-10	1,90E-03	4,60E-11	1,66E-02	3,29E-10	2,30E-02	6,80E-04	5,66E-03	3,20E-04	7,40E-03	0,00E+00
4,80E-03	0,00E+00	3,98E-03	1,36E-03	6,30E-03	0,00E+00	3,88E-02	1,04E-09	3,10E-03	4,24E-04	1,58E-02	0,00E+00	2,21E-02	6,80E-04	7,26E-03	3,20E-04	7,40E-03	0,00E+00
0,00E+00	0,00E+00	6,60E-04	1,32E-03	0,00E+00	0,00E+00	8,00E-04	3,41E-11	0,00E+00	0,00E+00	1,82E-02	3,29E-10	2,30E-02	6,80E-04	3,16E-03	4,03E-04	0,00E+00	0,00E+00
2,40E-03	0,00E+00	na	na	6,75E-03	4,50E-04	8,90E-03	0,00E+00	1,90E-03	0,00E+00	1,78E-02	4,00E-04	2,73E-02	1,09E-03	6,86E-03	1,24E-03	1,12E-02	3,75E-03
8,41E-02	1,32E-09	7,02E-02	1,56E-03	5,81E-02	1,60E-03	6,94E-02	1,14E-03	6,80E-02	9,73E-04	4,87E-02	7,03E-04	6,77E-02	9,44E-04	8,05E-02	7,14E-04	5,76E-02	3,99E-03
8,41E-02	0,00E+00	6,93E-02	0,00E+00	5,55E-02	0,00E+00	6,73E-02	0,00E+00	6,82E-02	0,00E+00	4,76E-02	0,00E+00	6,93E-02	0,00E+00	7,93E-02	0,00E+00	5,33E-02	9,31E-10
8,41E-02	0,00E+00	7,30E-02	0,00E+00	5,75E-02	0,00E+00	7,00E-02	0,00E+00	6,95E-02	4,71E-04	4,84E-02	0,00E+00	6,66E-02	0,00E+00	8,10E-02	0,00E+00	5,33E-02	9,31E-10
8,41E-02	1,32E-09	6,93E-02	0,00E+00	5,94E-02	6,59E-10	7,00E-02	0,00E+00	6,72E-02	0,00E+00	4,93E-02	0,00E+00	6,75E-02	0,00E+00	8,09E-02	2,80E-04	6,13E-02	0,00E+00
8,14E-02	0,00E+00	na	na	5,51E-02	4,50E-04	7,09E-02	0,00E+00	6,92E-02	0,00E+00	4,89E-02	4,50E-04	7,03E-02	9,50E-04	7,97E-02	1,30E-03	5,73E-02	4,00E-03
8,14E-02	2,70E-03	6,29E-02	3,91E-03	5,23E-02	2,46E-03	6,56E-02	2,11E-03	6,34E-02	5,63E-03	4,49E-02	7,69E-04	7,16E-02	5,62E-03	7,84E-02	2,29E-03	6,96E-02	5,69E-03
8,14E-02	2,70E-03	6,21E-02	3,60E-03	4,98E-02	1,90E-03	6,39E-02	1,75E-03	6,37E-02	5,55E-03	4,39E-02	4,00E-04	7,39E-02	5,50E-03	7,71E-02	2,20E-03	6,53E-02	4,00E-03
8,14E-02	2,70E-03	6,57E-02	3,60E-03	5,17E-02	1,90E-03	6,65E-02	8,50E-04	6,50E-02	5,57E-03	4,47E-02	4,00E-04	7,03E-02	5,45E-03	7,89E-02	2,15E-03	6,53E-02	4,00E-03
8,14E-02	2,70E-03	6,21E-02	3,60E-03	5,36E-02	1,90E-03	6,60E-02	2,20E-03	6,27E-02	5,55E-03	4,55E-02	4,00E-04	7,12E-02	5,45E-03	7,87E-02	2,18E-03	7,34E-02	4,05E-03
7,87E-02	2,70E-03	na	na	4,93E-02	1,96E-03	6,56E-02	1,75E-03	6,47E-02	5,55E-03	4,51E-02	5,66E-04	7,35E-02	5,09E-03	7,76E-02	2,51E-03	6,93E-02	5,69E-03
4,80E-03	0,00E+00	3,30E-03	0,00E+00	4,50E-03	5,82E-11	3,63E-02	0,00E+00	3,10E-03	4,24E-04	2,30E-03	0,00E+00	5,80E-03	0,00E+00	1,60E-03	0,00E+00	0,00E+00	0,00E+00
4,80E-03	0,00E+00	0,00E+00	0,00E+00	7,20E-03	0,00E+00	4,80E-03	0,00E+00	1,90E-03	4,12E-11	4,70E-03	8,23E-11	5,80E-03	8,23E-11	5,39E-03	2,80E-04	7,40E-03	1,65E-10
4,80E-03	0,00E+00	3,30E-03	9,20E-11	6,30E-03	0,00E+00	3,97E-02	4,66E-10	3,10E-03	4,24E-04	5,50E-03	0,00E+00	5,00E-03	0,00E+00	6,99E-03	2,80E-04	7,40E-03	1,65E-10
2,40E-03	0,00E+00	na	na	5,92E-03	1,18E-03	1,44E-02	1,39E-02	3,03E-03	1,31E-03	3,75E-03	7,69E-04	7,93E-03	1,26E-03	5,53E-03	1,58E-03	7,71E-03	5,26E-03
2,40E-03	0,00E+00	na	na	5,85E-03	4,50E-04	4,00E-03	8,00E-04	3,70E-03	0,00E+00	2,70E-03	4,00E-04	7,05E-03	1,25E-03	3,50E-03	4,00E-04	3,70E-03	3,70E-03
2,40E-03	0,00E+00	na	na	4,05E-03	4,50E-04	3,97E-02	8,50E-04	5,00E-03	4,24E-04	3,50E-03	4,00E-04	8,75E-03	4,50E-04	5,10E-03	4,00E-04	3,70E-03	3,70E-03
2,40E-03	0,00E+00	na	na	6,75E-03	4,50E-04	8,10E-03	2,02E-10	1,90E-03	4,60E-11	4,30E-03	4,00E-04	7,95E-03	1,25E-03	6,59E-03	1,23E-03	1,12E-02	3,75E-03
8,41E-02	0,00E+00	7,00E-02	1,48E-03	5,94E-02	9,31E-10	6,91E-02	0,00E+00	6,72E-02	0,00E+00	4,35E-02	0,00E+00	4,91E-02	7,20E-04	8,12E-02	3,60E-04	6,13E-02	0,00E+00
7,17E-02	2,19E-02	5,51E-02	1,77E-02	4,84E-02	1,20E-02	5,83E-02	1,54E-02	5,48E-02	1,83E-02	3,86E-02	9,89E-03	5,78E-02	1,16E-02	6,99E-02	2,05E-02	6,30E-02	2,35E-02
0,00E+00	0,00E+00	1,32E-03	1,62E-03	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	6,80E-04	8,33E-04	3,20E-04	3,92E-04	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	6,00E-04	4,24E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	2,29E-04	3,61E-04	0,00E+00	0,00E+00
fuscons to NF1		fuscons or NF2 to NF1		fuscons to NF1		fuscons to NF1		fuscons to NF1		fuscons+NF2+NF3 to NF1		fuscons+NF2+NF3 to NF1		fuscons to NF1		fuscons to NF1	
Undetected		fuscons and NF2		Undetected		Undetected		Undetected		Undetected		Undetected		Undetected		Undetected	

m00127150		m00127220		m00127250		m00127260		m00127850		m00127860		m00119700		m00119970		m00118890	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
4,35E-03	3,10E-03	0,00E+00	0,00E+00	4,46E-03	1,12E-03	2,94E-03	2,21E-03	1,39E-02	1,89E-03	5,24E-03	5,68E-03	2,43E-02	2,38E-03	2,28E-03	2,71E-03	1,85E-02	1,15E-03
9,60E-03	0,00E+00	0,00E+00	0,00E+00	3,00E-03	0,00E+00	5,70E-03	1,16E-10	1,71E-02	0,00E+00	1,21E-02	0,00E+00	2,35E-02	1,04E-03	3,30E-03	7,13E-11	1,68E-02	0,00E+00
4,80E-03	0,00E+00	0,00E+00	0,00E+00	5,60E-03	1,41E-03	4,90E-03	0,00E+00	1,40E-02	0,00E+00	1,06E-02	0,00E+00	2,09E-02	1,04E-03	6,60E-03	1,43E-10	2,02E-02	3,29E-10
1,90E-03	5,04E-11	0,00E+00	0,00E+00	4,60E-03	1,75E-10	9,14E-04	2,80E-04	1,24E-02	0,00E+00	0,00E+00	0,00E+00	2,61E-02	1,04E-03	0,00E+00	0,00E+00	1,85E-02	5,21E-10
8,60E-03	1,16E-10	6,00E-03	8,23E-11	5,35E-03	7,50E-04	4,90E-03	8,00E-04	1,71E-02	0,00E+00	1,59E-02	8,00E-04	2,48E-02	1,66E-03	1,60E-03	0,00E+00	1,68E-02	0,00E+00
8,27E-02	2,14E-03	8,89E-02	0,00E+00	5,83E-02	1,20E-03	4,74E-02	7,97E-04	5,56E-02	7,16E-04	5,12E-02	3,22E-03	3,70E-02	2,18E-03	4,37E-02	2,13E-03	4,62E-02	1,19E-03
8,65E-02	1,61E-09	8,89E-02	0,00E+00	5,67E-02	0,00E+00	4,62E-02	0,00E+00	5,60E-02	0,00E+00	4,56E-02	0,00E+00	3,61E-02	0,00E+00	4,39E-02	0,00E+00	4,44E-02	0,00E+00
8,22E-02	0,00E+00	8,89E-02	0,00E+00	5,95E-02	1,51E-03	4,70E-02	0,00E+00	5,43E-02	9,31E-10	5,12E-02	6,59E-10	3,35E-02	0,00E+00	4,74E-02	0,00E+00	4,79E-02	0,00E+00
8,12E-02	0,00E+00	8,89E-02	1,32E-09	5,84E-02	0,00E+00	4,80E-02	3,15E-04	5,60E-02	0,00E+00	5,36E-02	9,31E-10	3,88E-02	0,00E+00	4,21E-02	9,31E-10	4,62E-02	0,00E+00
8,22E-02	0,00E+00	9,57E-02	0,00E+00	5,92E-02	8,00E-04	4,88E-02	9,00E-04	5,76E-02	0,00E+00	4,84E-02	4,00E-04	3,75E-02	1,35E-03	4,04E-02	0,00E+00	4,44E-02	0,00E+00
8,00E-02	2,67E-03	6,90E-02	6,55E-03	4,28E-02	1,43E-03	4,26E-02	5,72E-04	5,56E-02	1,78E-03	6,21E-02	5,56E-03	3,83E-02	2,54E-03	3,95E-02	1,43E-03	4,35E-02	1,50E-03
8,38E-02	1,60E-03	6,90E-02	6,55E-03	4,13E-02	8,00E-04	4,23E-02	4,00E-04	5,60E-02	1,65E-03	6,56E-02	4,45E-03	3,75E-02	1,35E-03	3,87E-02	0,00E+00	4,18E-02	8,50E-04
7,96E-02	1,60E-03	6,90E-02	6,55E-03	4,40E-02	1,74E-03	4,32E-02	4,50E-04	5,44E-02	1,65E-03	6,21E-0E							

Additional file 3 Table S2. (continued)

m00117610		m00117600a		m00117600b		m00117590		m00116790		m00114680a		m00114680b		m00113910		m00113490	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
7.20E-03	3.19E-10	6.20E-03	1.35E-03	8.20E-03	1.43E-03	1.19E-02	1.66E-03	1.31E-02	2.28E-03	6.25E-03	1.75E-03	5.20E-02	7.11E-04	9.46E-04	1.73E-03	2.93E-03	2.61E-03
7.20E-03	1.84E-10	2.10E-03	0.00E+00	1.03E-02	0.00E+00	1.45E-02	6.37E-04	1.73E-02	0.00E+00	6.32E-03	2.40E-04	5.11E-02	0.00E+00	0.00E+00	0.00E+00	1.50E-03	0.00E+00
7.20E-03	1.84E-10	5.30E-03	0.00E+00	6.10E-03	8.23E-11	1.19E-02	6.37E-04	1.19E-02	1.65E-10	3.22E-03	2.40E-04	5.14E-02	3.77E-04	4.10E-03	0.00E+00	7.70E-03	1.65E-10
7.20E-03	0.00E+00	2.10E-03	0.00E+00	8.20E-03	0.00E+00	1.07E-02	5.88E-04	1.19E-02	0.00E+00	7.52E-03	2.40E-04	5.26E-02	1.86E-09	0.00E+00	0.00E+00	1.50E-03	0.00E+00
2.19E-02	0.00E+00	9.23E-02	1.39E-02	4.10E-03	0.00E+00	9.42E-03	6.37E-04	2.13E-02	4.03E-10	3.52E-03	3.84E-04	5.30E-02	3.50E-04	2.05E-03	2.05E-03	6.20E-03	1.16E-10
1.92E-01	7.75E-03	4.63E-02	4.63E-04	6.48E-02	1.15E-03	na	na	4.99E-02	1.47E-03	4.67E-02	8.56E-04	5.19E-02	9.20E-04	5.65E-02	1.85E-03	6.16E-02	2.71E-03
1.96E-01	0.00E+00	3.82E-02	0.00E+00	6.60E-02	0.00E+00	na	na	5.16E-02	0.00E+00	4.62E-02	0.00E+00	5.33E-02	9.31E-10	5.55E-02	0.00E+00	6.06E-02	0.00E+00
1.78E-01	0.00E+00	3.93E-02	0.00E+00	6.60E-02	0.00E+00	na	na	4.74E-02	0.00E+00	4.55E-02	0.00E+00	5.06E-02	3.30E-04	5.99E-02	0.00E+00	5.73E-02	0.00E+00
1.96E-01	4.56E-09	3.82E-02	0.00E+00	6.37E-02	1.61E-09	na	na	5.02E-02	0.00E+00	4.75E-02	6.59E-10	5.19E-02	0.00E+00	5.55E-02	0.00E+00	6.39E-02	0.00E+00
2.06E-01	0.00E+00	1.02E-01	1.28E-02	6.82E-02	0.00E+00	na	na	5.51E-02	2.10E-03	4.72E-02	3.50E-04	5.23E-02	3.50E-04	5.77E-02	2.20E-03	6.22E-02	0.00E+00
1.87E-01	8.99E-03	3.40E-02	4.47E-03	7.72E-02	2.72E-03	6.43E-02	4.83E-03	5.03E-02	5.39E-03	4.76E-02	1.84E-03	5.39E-02	1.22E-03	4.99E-02	2.88E-03	6.33E-02	7.06E-03
1.92E-01	4.65E-03	3.38E-02	4.45E-03	7.95E-02	2.25E-03	6.31E-02	5.50E-03	5.09E-02	6.30E-03	4.72E-02	1.65E-03	5.61E-02	9.31E-10	4.89E-02	2.20E-03	6.31E-02	5.80E-03
1.73E-01	4.55E-03	3.49E-02	4.45E-03	7.50E-02	2.25E-03	6.18E-02	4.15E-03	4.81E-02	4.90E-03	4.62E-02	1.30E-03	5.32E-02	4.81E-04	5.33E-02	2.20E-03	5.98E-02	5.80E-03
1.92E-01	4.65E-03	3.38E-02	4.45E-03	7.72E-02	2.25E-03	6.59E-02	4.15E-03	5.09E-02	4.90E-03	4.85E-02	1.65E-03	5.33E-02	0.00E+00	4.89E-02	2.20E-03	6.48E-02	7.50E-03
2.01E-01	4.70E-03	9.62E-02	1.34E-02	7.72E-02	2.25E-03	6.59E-02	2.75E-03	5.41E-02	6.26E-03	4.78E-02	1.33E-03	5.33E-02	1.11E-03	5.11E-02	3.11E-03	6.48E-02	5.85E-03
1.45E-02	0.00E+00	3.20E-03	0.00E+00	1.23E-02	0.00E+00	7.60E-03	0.00E+00	7.90E-03	0.00E+00	6.80E-03	8.23E-11	9.53E-03	3.30E-04	4.10E-03	0.00E+00	6.20E-03	1.16E-10
0.00E+00	0.00E+00	0.00E+00	0.00E+00	6.10E-03	8.23E-11	8.90E-03	0.00E+00	7.90E-03	1.65E-10	3.70E-03	8.23E-11	1.26E-02	0.00E+00	0.00E+00	0.00E+00	3.10E-03	9.20E-11
1.45E-02	0.00E+00	3.20E-03	0.00E+00	1.03E-02	1.16E-10	3.80E-03	8.23E-11	2.60E-03	5.82E-11	8.00E-03	0.00E+00	1.02E-02	2.83E-04	4.10E-03	0.00E+00	9.20E-03	2.33E-10
1.79E-02	6.24E-03	9.02E-02	1.38E-02	1.14E-02	3.30E-03	5.58E-03	2.59E-03	1.10E-02	1.28E-03	6.38E-03	1.34E-03	1.22E-02	3.87E-04	3.01E-03	2.70E-03	6.27E-03	1.55E-03
1.45E-02	0.00E+00	8.99E-02	1.38E-02	1.44E-02	2.03E-10	1.02E-02	0.00E+00	1.06E-02	1.35E-03	7.10E-03	3.00E-04	1.23E-02	3.00E-04	2.05E-03	2.05E-03	4.60E-03	0.00E+00
2.93E-02	0.00E+00	9.11E-02	1.39E-02	6.10E-03	0.00E+00	5.10E-03	0.00E+00	9.20E-03	0.00E+00	4.00E-03	3.00E-04	1.19E-02	4.53E-04	6.20E-03	2.10E-03	4.60E-03	0.00E+00
1.45E-02	0.00E+00	8.99E-02	1.38E-02	1.23E-02	0.00E+00	3.80E-03	9.20E-11	1.19E-02	1.65E-10	7.10E-03	3.00E-04	1.23E-02	3.00E-04	2.05E-03	2.05E-03	7.70E-03	1.65E-10
1.87E-01	0.00E+00	4.04E-02	0.00E+00	6.82E-02	0.00E+00	na	na	3.77E-02	0.00E+00	4.63E-02	2.40E-04	1.26E-02	0.00E+00	5.55E-02	0.00E+00	6.22E-02	0.00E+00
1.56E-01	5.86E-02	3.28E-02	8.13E-03	6.88E-02	5.96E-02	5.51E-03	5.96E-02	3.60E-02	9.63E-03	4.10E-02	1.34E-02	1.97E-02	3.95E-03	4.28E-02	1.38E-02	5.89E-02	1.61E-02
0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	7.80E-04	6.37E-04	0.00E+00	0.00E+00	2.40E-04	2.94E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.67E-04	3.30E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00
0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
fuscars or NF2 or NF3 to NF1		NF1 to fuscars or NF2		NF3 to NF1		fuscars+NF2+NF3 to NF1		fuscars+NF2+NF3 to NF1		NF3 to NF1		Undetected		fuscars or NF2 to NF1		fuscars or NF2 to NF1	
fuscars and NF2		fuscars and NF2		fuscars and NF2		Undetected		Undetected		Undetected		Undetected		fuscars and NF2		fuscars and NF2	

m00113470a		m00113470b		m00108600a		m00108600b		m00108570		m00108560a		m00108560b		m00139330a		m00139330b	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
2.04E-02	1.00E-02	1.90E-02	1.70E-03	7.81E-02	8.14E-03	8.07E-03	8.71E-03	0.00E+00	0.00E+00	1.32E-02	2.36E-02	4.43E-03	8.09E-03	1.57E-03	1.43E-03	4.96E-02	7.04E-03
9.20E-03	0.00E+00	1.68E-02	3.60E-04	6.77E-02	0.00E+00	1.25E-02	1.65E-10	0.00E+00	0.00E+00	2.80E-04	3.43E-04	0.00E+00	0.00E+00	9.00E-04	7.35E-04	5.67E-02	1.47E-09
1.01E-02	1.65E-10	1.77E-02	4.00E-04	7.12E-02	0.00E+00	2.13E-02	4.03E-10	0.00E+00	0.00E+00	5.62E-02	3.92E-04	1.92E-02	0.00E+00	3.80E-03	7.35E-04	5.77E-02	1.14E-09
2.97E-02	8.71E-10	2.05E-02	4.00E-04	8.56E-02	3.73E-09	5.00E-04	0.00E+00	0.00E+00	0.00E+00	2.80E-04	3.43E-04	0.00E+00	0.00E+00	9.00E-04	7.35E-04	4.31E-02	1.47E-09
2.70E-02	2.03E-02	6.94E-02	6.40E-04	7.12E-02	0.00E+00	2.50E-02	8.50E-04	9.50E-03	3.20E-03	5.42E-02	5.60E-04	2.69E-01	3.13E-02	5.30E-03	7.35E-04	6.04E-02	1.55E-03
4.10E-02	1.03E-02	6.19E-02	4.64E-03	8.54E-02	8.22E-03	4.24E-02	8.28E-04	2.10E-03	0.00E+00	4.25E-02	2.01E-02	3.03E-01	8.85E-04	na	na	na	na
5.08E-02	9.31E-10	6.47E-02	0.00E+00	7.48E-02	0.00E+00	4.33E-02	0.00E+00	2.10E-03	0.00E+00	5.35E-02	0.00E+00	3.03E-01	0.00E+00	na	na	na	na
5.35E-02	0.00E+00	6.87E-02	0.00E+00	7.84E-02	0.00E+00	4.10E-02	0.00E+00	2.10E-03	0.00E+00	5.80E-03	0.00E+00	3.05E-01	0.00E+00	na	na	na	na
3.15E-02	0.00E+00	5.78E-02	0.00E+00	9.29E-02	0.00E+00	4.27E-02	0.00E+00	2.10E-03	0.00E+00	5.35E-02	0.00E+00	3.03E-01	6.45E-09	na	na	na	na
4.20E-02	9.70E-03	2.98E-02	0.00E+00	7.84E-02	0.00E+00	3.96E-02	3.00E-04	7.40E-03	3.20E-03	2.47E-02	1.96E-02	1.21E-01	1.19E-01	na	na	na	na
4.29E-02	8.65E-03	6.21E-02	3.73E-03	8.36E-02	8.39E-03	3.94E-02	1.20E-03	5.00E-02	6.75E-03	1.25E-01	6.32E-02	2.52E-01	9.86E-02	3.63E-02	9.90E-04	4.98E-02	2.00E-03
4.95E-02	1.35E-03	6.52E-02	5.00E-04	7.30E-02	1.80E-03	4.11E-02	1.15E-03	5.00E-02	6.75E-03	1.26E-01	6.41E-02	2.52E-01	9.96E-02	3.59E-02	4.66E-10	4.73E-02	0.00E+00
4.95E-02	4.50E-04	6.67E-02	1.00E-03	7.66E-02	1.80E-03	3.88E-02	1.15E-03	5.00E-02	6.75E-03	1.21E-01	5.98E-02	2.52E-01	9.55E-02	3.75E-02	1.55E-03	4.83E-02	0.00E+00
3.72E-02	8.30E-03	5.88E-02	9.50E-04	9.11E-02	1.85E-03	3.91E-02	2.50E-04	5.00E-02	6.75E-03	1.26E-01	6.41E-02	2.52E-01	9.96E-02	3.59E-02	4.66E-10	5.15E-02	1.05E-03
4.52E-02	1.44E-02	2.74E-02	2.81E-03	7.66E-02	1.80E-03	3.85E-02	1.16E-03	4.55E-02	9.09E-03	1.13E-01	6.00E-02	3.05E-01	4.37E-02	3.75E-02	0.00E+00	4.99E-02	5.50E-04
5.90E-03	0.00E+00	1.38E-02	0.00E+00	3.20E-03	0.00E+00	8.70E-03	0.00E+00	0.00E+00	0.00E+00	5.59E-02	0.00E+00	1.92E-02	0.00E+00	2.90E-03	0.00E+00	2.90E-03	0.00E+00
3.41E-02	0.00E+00	2.41E-02	0.00E+00	1.64E-02	0.00E+00	1.20E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.28E-02	0.00E+00
3.49E-02	0.00E+00	2.88E-02	0.00E+00	1.31E-02	0.00E+00	2.08E-02	4.66E-10	0.00E+00	0.00E+00	5.59E-02	9.31E-10	1.92E-02	0.00E+00	2.90E-03	4.12E-11	1.58E-02	3.29E-10
2.90E-02	1.45E-02	6.30E-02	3.25E-03	7.79E-03	5.84E-03	1.78E-02	7.35E-03	9.50E-03	3.20E-03	4.70E-02	1.60E-02	2.70E-01	3.13E-02	5.07E-03	1.22E-03	1.23E-02	6.57E-03
2.62E-02	2.03E-02	6.52E-02	1.50E-03	3.20E-03	0.00E+00	1.23E-02	8.50E-04	9.50E-03	3.20E-03	5.39E-02	4.00E-04	2.69E-01	3.13E-02	4.40E-03	0.00E+00	5.40E-03	1.50E-03
2.70E-02	2.20E-02	6.72E-02	1.50E-03	0.00E+00	0.00E+00	7.85E-03	8.50E-04	9.50E-03	3.20E-03	2.39E-02	2.03E-02	2.71E-01	3.14E-02	7.30E-03	0.00E+00	5.40E-03	5.00E-04
3.11E-02	4.50E-04	6.02E-02	5.00E-04	1.31E-02	0.00E+00	2.44E-02											

Additional file 3 Table S2. (continued)

m00139350		m00139420		m00139860		m00139930		m00140880		m00140860		m00141310		m00141340		m00141400	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	a41501	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
2,12E-03	2,29E-03	9,02E-03	9,76E-03	4,59E-02	1,49E-03	3,50E-03	0,00E+00	7,18E-02	2,34E-03	5,55E-03	1,89E-03	2,41E-03	2,68E-03	2,03E-03	2,36E-03	1,08E-02	1,17E-02
4,60E-03	0,00E+00	2,02E-02	3,29E-10	4,72E-02	1,32E-03	3,50E-03	0,00E+00	6,79E-02	1,32E-09	7,60E-03	1,84E-10	4,80E-03	0,00E+00	3,50E-03	1,27E-03	2,34E-02	5,70E-10
4,60E-03	0,00E+00	1,89E-02	4,66E-10	4,35E-02	0,00E+00	3,50E-03	0,00E+00	7,46E-02	0,00E+00	7,60E-03	1,84E-10	5,63E-03	1,18E-03	5,30E-03	0,00E+00	2,34E-02	5,70E-10
0,00E+00	0,00E+00	0,00E+00	0,00E+00	4,63E-02	1,32E-09	3,50E-03	0,00E+00	7,23E-02	1,32E-09	3,80E-03	1,01E-10	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
4,60E-03	0,00E+00	2,57E-02	0,00E+00	4,21E-02	1,40E-03	9,80E-03	9,00E-04	7,35E-02	1,15E-03	7,60E-03	0,00E+00	4,80E-03	0,00E+00	0,00E+00	0,00E+00	2,82E-02	4,66E-10
3,29E-02	2,44E-03	2,45E-02	3,13E-03	6,42E-02	1,34E-03	5,11E-02	0,00E+00	6,74E-02	2,31E-03	4,83E-02	1,99E-03	na	na	4,95E-02	1,01E-03	4,61E-02	1,22E-02
3,55E-02	0,00E+00	2,85E-02	0,00E+00	6,43E-02	1,37E-03	5,11E-02	6,59E-10	6,35E-02	0,00E+00	5,05E-02	0,00E+00	na	na	5,10E-02	1,32E-03	3,30E-02	0,00E+00
3,55E-02	0,00E+00	2,71E-02	0,00E+00	6,62E-02	0,00E+00	5,11E-02	6,59E-10	7,01E-02	0,00E+00	5,05E-02	0,00E+00	na	na	4,91E-02	0,00E+00	3,30E-02	0,00E+00
3,06E-02	4,66E-10	2,16E-02	0,00E+00	6,33E-02	0,00E+00	5,11E-02	6,59E-10	6,79E-02	9,31E-10	4,65E-02	1,14E-09	na	na	4,91E-02	9,31E-10	5,74E-02	6,59E-10
3,55E-02	0,00E+00	2,71E-02	0,00E+00	6,48E-02	1,45E-03	5,78E-02	9,50E-04	6,90E-02	1,10E-03	5,05E-02	0,00E+00	na	na	4,91E-02	0,00E+00	3,78E-02	0,00E+00
3,89E-02	4,36E-03	2,80E-02	1,90E-03	6,70E-02	3,15E-03	5,11E-02	1,32E-09	6,80E-02	2,16E-03	5,04E-02	2,86E-03	na	na	4,67E-02	2,98E-03	6,33E-02	1,26E-02
4,16E-02	3,65E-03	3,05E-02	7,00E-04	6,71E-02	3,16E-03	5,11E-02	0,00E+00	6,46E-02	1,10E-03	5,26E-02	2,05E-03	na	na	4,82E-02	3,10E-03	5,26E-02	9,85E-03
4,16E-02	3,65E-03	2,92E-02	6,50E-04	6,91E-02	2,85E-03	5,11E-02	0,00E+00	6,90E-02	1,10E-03	5,26E-02	2,05E-03	na	na	4,63E-02	2,80E-03	5,26E-02	9,85E-03
3,67E-02	3,60E-03	2,64E-02	7,00E-04	6,62E-02	2,85E-03	5,11E-02	0,00E+00	6,90E-02	1,10E-03	4,85E-02	2,00E-03	na	na	4,63E-02	2,80E-03	7,26E-02	5,05E-03
4,16E-02	3,65E-03	2,92E-02	6,50E-04	6,76E-02	3,20E-03	5,78E-02	9,50E-04	6,79E-02	1,56E-03	5,26E-02	2,05E-03	na	na	4,63E-02	2,80E-03	5,76E-02	9,95E-03
0,00E+00	0,00E+00	6,70E-03	1,16E-10	8,87E-03	1,23E-03	0,00E+00	0,00E+00	1,64E-02	0,00E+00	0,00E+00	0,00E+00	8,00E-04	1,13E-03	5,33E-03	3,77E-03	0,00E+00	0,00E+00
4,60E-03	1,16E-10	2,02E-02	2,33E-10	8,67E-04	1,23E-03	0,00E+00	0,00E+00	8,20E-03	0,00E+00	1,14E-02	2,85E-10	4,80E-03	0,00E+00	3,50E-03	1,27E-03	2,34E-02	8,07E-10
4,60E-03	1,16E-10	1,89E-02	0,00E+00	8,00E-03	0,00E+00	0,00E+00	0,00E+00	1,64E-02	0,00E+00	1,14E-02	2,85E-10	5,63E-03	1,18E-03	5,30E-03	0,00E+00	2,34E-02	8,07E-10
6,77E-03	2,34E-03	1,97E-02	6,47E-03	6,23E-03	1,96E-03	6,20E-03	9,00E-04	1,44E-02	3,15E-03	9,65E-03	1,89E-03	7,25E-03	2,72E-03	2,03E-03	2,36E-03	1,73E-02	1,18E-02
9,30E-03	0,00E+00	1,34E-02	0,00E+00	7,53E-03	1,83E-03	6,20E-03	9,00E-04	1,75E-02	1,05E-03	7,60E-03	8,23E-11	9,70E-03	0,00E+00	3,50E-03	1,27E-03	4,60E-03	0,00E+00
9,30E-03	0,00E+00	1,21E-02	0,00E+00	3,95E-03	1,35E-03	6,20E-03	9,00E-04	9,20E-03	1,00E-03	7,60E-03	8,23E-11	1,05E-02	1,18E-03	5,30E-03	0,00E+00	4,60E-03	0,00E+00
4,60E-03	0,00E+00	2,57E-02	0,00E+00	6,65E-03	1,35E-03	6,20E-03	9,00E-04	1,54E-02	1,05E-03	1,14E-02	2,33E-10	4,80E-03	0,00E+00	0,00E+00	0,00E+00	2,82E-02	0,00E+00
3,06E-02	3,29E-10	2,16E-02	2,33E-10	3,24E-02	0,00E+00	5,11E-02	0,00E+00	1,23E-02	0,00E+00	4,25E-02	0,00E+00	na	na	4,91E-02	0,00E+00	5,74E-02	9,31E-10
3,63E-02	4,21E-03	2,44E-02	4,64E-03	3,43E-02	9,87E-03	4,48E-02	1,41E-02	1,33E-02	1,00E-03	3,87E-02	1,32E-02	na	na	4,04E-02	1,36E-02	6,68E-02	1,43E-02
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,73E-03	1,23E-03	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	5,33E-03	3,77E-03	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,60E-03	1,13E-03	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
fuscarins to NF1		fuscarins to NF1		Undetected		fuscarins or NF2 or NF3 to NF1		Undetected		fuscarins to NF1		fuscarins to NF1		fuscarins to NF1		fuscarins to NF1	
NF2 and NF3		Undetected		fuscarins and NF2		fuscarins and NF2 and NF3		fuscarins and NF2		NF2 and NF3		Undetected		Undetected		NF2 and NF3	

m00141430		m00141440		m00144630		m00142140a		m00142140b		m00114910a		m00114910b		m00144240a		m00144240b	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
2,26E-03	2,62E-03	0,00E+00	0,00E+00	5,01E-03	5,42E-03	2,13E-01	8,47E-04	0,00E+00	0,00E+00	3,42E-03	3,30E-03	5,53E-02	1,19E-03	4,24E-03	8,48E-04	7,27E-02	2,64E-03
3,90E-03	1,41E-03	0,00E+00	0,00E+00	1,03E-02	0,00E+00	2,14E-01	5,27E-09	0,00E+00	0,00E+00	6,20E-03	1,65E-10	5,64E-02	5,39E-04	5,40E-03	0,00E+00	6,98E-02	9,31E-10
5,90E-03	0,00E+00	0,00E+00	0,00E+00	1,14E-02	2,33E-10	2,14E-01	5,27E-09	0,00E+00	0,00E+00	7,67E-03	2,36E-04	5,64E-02	5,39E-04	4,60E-03	0,00E+00	6,98E-02	9,31E-10
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	2,13E-01	8,33E-09	0,00E+00	0,00E+00	4,00E-04	1,71E-11	5,43E-02	6,62E-04	3,59E-03	5,25E-04	7,51E-02	2,46E-09
1,18E-02	3,00E-03	9,90E-03	3,30E-03	7,20E-03	2,10E-03	1,71E-02	6,50E-04	0,00E+00	0,00E+00	1,18E-02	2,00E-04	5,47E-02	1,71E-03	7,70E-03	8,00E-04	7,25E-02	2,65E-03
5,84E-02	1,15E-03	3,34E-02	8,07E-10	5,43E-02	1,33E-03	2,13E-01	8,47E-04	0,00E+00	0,00E+00	na	na	na	na	5,05E-02	2,72E-03	8,86E-02	2,64E-03
6,00E-02	1,51E-03	3,34E-02	0,00E+00	5,42E-02	0,00E+00	2,14E-01	0,00E+00	0,00E+00	0,00E+00	na	na	na	na	4,65E-02	0,00E+00	8,57E-02	0,00E+00
5,79E-02	0,00E+00	3,34E-02	0,00E+00	5,20E-02	0,00E+00	2,14E-01	0,00E+00	0,00E+00	0,00E+00	na	na	na	na	4,90E-02	0,00E+00	8,57E-02	0,00E+00
5,79E-02	0,00E+00	3,34E-02	8,07E-10	5,53E-02	0,00E+00	2,13E-01	4,56E-09	0,00E+00	0,00E+00	na	na	na	na	5,29E-02	5,95E-04	9,10E-02	1,32E-09
6,43E-02	3,20E-03	3,00E-02	3,40E-03	5,53E-02	0,00E+00	1,84E-02	6,50E-04	0,00E+00	0,00E+00	na	na	na	na	5,14E-02	0,00E+00	8,84E-02	2,65E-03
4,11E-02	1,91E-03	3,69E-02	3,45E-03	5,32E-02	2,57E-03	2,14E-01	1,20E-03	2,91E-02	0,00E+00	6,39E-02	2,20E-03	5,52E-02	2,91E-03	5,13E-02	4,25E-03	8,59E-02	3,78E-03
4,27E-02	2,13E-03	3,69E-02	3,45E-03	5,31E-02	2,20E-03	2,15E-01	8,50E-04	2,91E-02	0,00E+00	6,38E-02	2,15E-03	5,63E-02	2,75E-03	4,74E-02	3,25E-03	8,30E-02	2,70E-03
4,07E-02	1,55E-03	3,69E-02	3,45E-03	5,09E-02	2,20E-03	2,15E-01	8,50E-04	2,91E-02	0,00E+00	6,34E-02	2,16E-03	5,63E-02	2,75E-03	4,98E-02	3,30E-03	8,30E-02	2,70E-03
4,07E-02	1,55E-03	3,69E-02	3,45E-03	5,42E-02	2,20E-03	2,14E											

Additional file 3 Table S2. (continued)

m00144270		m00144420a		m00144420b		m00144930a		m00144930b		m00145330a		m00145330b		m00146940		m00146970	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
7,38E-04	7,98E-04	1,11E-01	1,17E-03	1,79E-03	1,73E-03	4,49E-02	4,29E-02	1,32E-01	1,12E-01	1,37E-02	2,08E-02	2,79E-02	3,17E-02	2,68E-03	2,71E-03	2,49E-03	2,69E-03
1,60E-03	0,00E+00	1,10E-01	0,00E+00	3,24E-03	1,08E-03	5,29E-02	2,16E-02	9,69E-02	1,18E-01	5,12E-02	0,00E+00	4,57E-02	0,00E+00	6,12E-03	4,40E-04	5,40E-03	0,00E+00
1,60E-03	0,00E+00	1,11E-01	1,13E-03	3,24E-03	1,08E-03	3,68E-02	5,81E-02	1,67E-01	7,29E-02	8,10E-03	2,02E-10	7,51E-02	2,08E-09	4,94E-03	4,80E-04	5,40E-03	0,00E+00
0,00E+00	0,00E+00	1,12E-01	3,95E-09	5,40E-04	1,08E-03	4,81E-02	3,36E-02	1,31E-01	1,19E-01	7,14E-05	1,75E-04	0,00E+00	0,00E+00	2,40E-04	4,80E-04	0,00E+00	0,00E+00
3,10E-03	5,82E-11	1,15E-01	0,00E+00	9,90E-03	1,75E-03	1,11E-01	4,71E-02	2,80E-01	1,80E-02	4,90E-02	1,15E-03	6,56E-02	6,70E-03	4,94E-03	4,80E-04	4,70E-03	7,00E-04
9,79E-02	1,19E-03	1,09E-01	1,12E-03	6,98E-02	2,00E-03	1,20E-01	2,70E-02	2,98E-01	1,64E-02	1,17E-01	1,34E-02	6,03E-03	3,86E-02	1,02E-03	5,12E-02	6,32E-04	
9,96E-02	1,86E-09	1,08E-01	1,32E-09	6,69E-02	0,00E+00	9,61E-02	0,00E+00	2,91E-01	1,77E-02	9,27E-02	0,00E+00	1,14E-01	0,00E+00	3,83E-02	0,00E+00	5,15E-02	0,00E+00
9,61E-02	0,00E+00	1,08E-01	1,08E-03	7,28E-02	0,00E+00	1,51E-01	9,48E-03	2,99E-01	4,15E-03	1,26E-01	0,00E+00	1,09E-01	0,00E+00	3,70E-02	0,00E+00	5,00E-02	0,00E+00
9,79E-02	0,00E+00	1,10E-01	0,00E+00	6,98E-02	1,32E-09	1,06E-01	1,92E-02	3,00E-01	1,82E-02	1,24E-01	2,10E-04	1,23E-01	0,00E+00	3,95E-02	4,66E-10	5,15E-02	0,00E+00
9,79E-02	0,00E+00	1,12E-01	0,00E+00	7,72E-02	1,50E-03	1,66E-02	0,00E+00	2,27E-02	4,50E-04	9,27E-02	1,20E-03	1,07E-01	1,45E-03	3,77E-02	6,50E-04	5,08E-02	7,50E-04
8,47E-02	4,51E-03	1,03E-01	3,71E-03	7,00E-02	7,47E-03	9,85E-02	3,69E-02	2,90E-01	1,82E-02	1,13E-01	1,32E-02	1,03E-01	6,56E-03	4,29E-02	2,13E-03	5,19E-02	5,05E-03
8,65E-02	4,35E-03	1,02E-01	3,50E-03	6,84E-02	7,30E-03	8,70E-02	1,84E-02	2,85E-01	1,76E-02	8,94E-02	3,50E-04	1,00E-01	2,80E-03	4,26E-02	1,85E-03	5,22E-02	5,00E-03
8,30E-02	4,35E-03	1,02E-01	3,69E-03	6,84E-02	7,30E-03	1,15E-01	4,75E-02	2,94E-01	1,66E-02	1,23E-01	1,30E-03	9,44E-02	2,80E-03	4,14E-02	1,85E-03	5,08E-02	5,05E-03
8,48E-02	4,35E-03	1,04E-01	3,55E-03	7,14E-02	7,35E-03	9,10E-02	2,79E-02	2,91E-01	1,86E-02	1,19E-01	1,27E-03	1,08E-01	2,85E-03	4,38E-02	1,90E-03	5,22E-02	5,00E-03
8,48E-02	4,35E-03	1,06E-01	3,50E-03	7,28E-02	7,49E-03	8,04E-02	6,38E-02	2,31E-01	1,09E-02	1,29E-03	1,33E-03	9,30E-02	2,85E-03	4,20E-02	1,94E-03	5,15E-02	5,08E-03
3,10E-03	5,82E-11	6,67E-04	9,43E-04	5,40E-03	0,00E+00	5,09E-02	1,25E-02	1,63E-01	6,95E-02	5,30E-02	0,00E+00	4,83E-02	0,00E+00	5,90E-03	0,00E+00	5,40E-03	0,00E+00
1,60E-03	0,00E+00	6,20E-03	1,84E-10	2,70E-03	0,00E+00	8,42E-03	1,68E-02	1,26E-01	1,19E-01	5,13E-02	2,10E-04	4,75E-02	9,31E-10	5,90E-03	0,00E+00	5,40E-03	0,00E+00
1,60E-03	0,00E+00	6,87E-03	9,43E-04	2,70E-03	0,00E+00	4,24E-02	2,10E-02	1,57E-01	6,29E-02	8,19E-03	2,10E-04	7,51E-02	2,08E-09	4,70E-03	8,23E-11	5,40E-03	0,00E+00
3,12E-03	1,05E-03	7,60E-03	2,96E-03	9,37E-03	2,28E-03	1,10E-01	2,67E-02	2,93E-01	2,01E-02	3,96E-02	1,82E-02	4,93E-02	2,18E-02	4,70E-03	9,98E-04	4,39E-03	8,93E-04
4,70E-03	0,00E+00	4,10E-03	0,00E+00	1,21E-02	1,40E-03	8,68E-02	0,00E+00	2,85E-01	2,17E-02	3,60E-03	1,10E-03	4,71E-02	3,95E-03	5,90E-03	0,00E+00	3,35E-03	6,50E-04
1,60E-03	0,00E+00	4,80E-03	9,90E-04	6,70E-03	1,30E-03	1,41E-01	9,38E-03	2,94E-01	4,15E-03	5,07E-02	1,15E-03	1,37E-02	6,30E-03	3,50E-03	1,20E-03	4,70E-03	7,00E-04
3,10E-03	7,13E-11	1,03E-02	0,00E+00	9,35E-03	1,35E-03	9,63E-02	1,90E-02	2,24E-02	2,96E-01	1,16E-03	6,90E-02	6,70E-03	6,56E-02	4,70E-03	0,00E+00	4,70E-03	7,00E-04
9,79E-02	0,00E+00	2,20E-02	1,03E-03	7,04E-02	1,20E-03	1,15E-01	5,77E-02	2,87E-01	1,46E-02	1,24E-01	1,32E-09	1,23E-01	1,86E-09	3,97E-02	4,80E-04	5,15E-02	0,00E+00
7,69E-02	1,81E-02	1,84E-02	1,81E-03	6,38E-02	1,95E-02	9,98E-02	5,63E-02	2,73E-01	2,55E-02	1,03E-01	3,63E-02	9,56E-02	2,88E-02	3,84E-02	1,28E-02	4,59E-02	1,49E-02
0,00E+00	0,00E+00	1,20E-03	9,80E-04	1,08E-03	1,32E-03	5,77E-02	7,06E-02	9,71E-02	1,19E-01	0,00E+00	0,00E+00	0,00E+00	0,00E+00	4,80E-04	5,88E-04	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	na	na	1,60E-01	1,13E-01	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	1,33E-03	9,43E-04	0,00E+00	0,00E+00	1,67E-02	1,18E-02	1,38E-01	9,73E-02	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,68E-02	2,06E-02	1,37E-01	1,19E-01	1,43E-04	2,26E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
fuscons to NF1		Undetected		fuscons to NF1		Undetected		NF2 to NF1		fuscons to NF1		fuscons to NF1		fuscons to NF1		fuscons to NF1	
Undetected		Undetected		Undetected		fuscons and NF2		Undetected		Undetected		Undetected		Undetected		Undetected	

m00146980		m00146990		m00147020		m00147060		m00147090		m00147100		m00147120		m00147130		m00147140	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
7,67E-03	7,04E-03	1,15E-03	2,11E-03	1,03E-02	1,27E-02	3,55E-03	3,84E-03	6,90E-03	7,59E-03	5,51E-03	3,96E-03	8,11E-03	9,00E-03	3,07E-03	3,35E-03	1,24E-01	6,91E-02
1,41E-02	1,03E-03	0,00E+00	0,00E+00	3,21E-02	3,92E-04	7,70E-03	1,65E-10	1,28E-02	0,00E+00	1,02E-02	6,86E-04	1,62E-02	4,03E-10	5,90E-03	0,00E+00	1,21E-01	8,24E-02
1,62E-02	1,08E-03	5,00E-03	8,23E-11	1,14E-02	3,92E-04	7,70E-03	1,65E-10	1,71E-02	0,00E+00	9,52E-04	1,89E-02	3,87E-03	7,40E-03	0,00E+00	1,21E-01	8,25E-02	
1,26E-03	1,03E-03	0,00E+00	0,00E+00	4,80E-04	3,92E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,92E-03	6,37E-04	0,00E+00	0,00E+00	0,00E+00	0,00E+00	1,26E-01	5,49E-02
1,95E-02	7,64E-03	5,00E-03	1,16E-10	8,28E-03	3,92E-04	7,70E-03	0,00E+00	2,59E-02	0,00E+00	8,13E-02	7,35E-04	7,57E-02	0,00E+00	4,46E-02	0,00E+00	1,28E-01	7,71E-02
1,29E-01	1,05E-03	5,87E-02	2,32E-03	9,53E-02	2,15E-03	4,36E-02	4,04E-03	4,46E-02	1,90E-03	8,05E-02	2,44E-03	8,47E-02	9,95E-03	4,85E-02	9,31E-10	na	na
1,29E-01	0,00E+00	5,74E-02	0,00E+00	9,92E-02	0,00E+00	3,92E-02	0,00E+00	4,81E-02	0,00E+00	8,37E-02	0,00E+00	9,36E-02	0,00E+00	4,85E-02	0,00E+00	na	na
1,31E-01	0,00E+00	6,29E-02	0,00E+00	9,48E-02	0,00E+00	3,92E-02	0,00E+00	4,36E-02	0,00E+00	7,66E-02	7,54E-04	9,66E-02	4,29E-03	4,85E-02	0,00E+00	na	na
1,29E-01	0,00E+00	5,74E-02	6,59E-10	9,39E-02	0,00E+00	4,73E-02	0,00E+00	4,36E-02	0,00E+00	8,07E-02	0,00E+00	7,57E-02	0,00E+00	4,85E-02	9,31E-10	na	na
1,26E-01	2,50E-03	6,29E-02	0,00E+00	9,56E-02	0,00E+00	3,92E-02	0,00E+00	1,71E-02	0,00E+00	1,52E-02	0,00E+00	0,00E+00	0,00E+00	1,11E-02	0,00E+00	na	na
1,20E-01	2,68E-03	4,78E-02	2,28E-03	8,81E-02	1,28E-02	3,95E-02	5,69E-03	3,35E-02	4,15E-03	8,66E-02	3,95E-03	8,93E-02	1,68E-02	5,13E-02	2,75E-03	1,71E-01	6,12E-03
1,17E-01	3,70E-03	4,66E-02	0,00E+00	9,18E-02	1,01E-02	3,52E-02	4,00E-03	3,25E-02	6,60E-03	8,99E-02	3,10E-03	9,83E-02	1,37E-02	5,13E-02	2,75E-03	1,77E-01	0,00E+00
1,22E-01	1,25E-03	5,20E-02	0,00E+00	8,79E-02	1,31E-02	3,52E-02	4,00E-03	3,69E-02	2,20E-03	8,28E-02	3,1						

Additional file 3 Table S2. (continued)

m00200790		m00200820		m00200130		m00248080		m00247980	
mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD	mean Ks	SD
0,00E+00	0,00E+00	7,00E-04	1,40E-03	4,54E-02	1,61E-09	3,58E-02	8,80E-04	3,74E-02	3,01E-03
0,00E+00	0,00E+00	7,00E-04	1,40E-03	4,54E-02	0,00E+00	3,58E-02	8,80E-04	3,65E-02	8,07E-10
0,00E+00	0,00E+00	7,00E-04	1,40E-03	4,54E-02	0,00E+00	3,58E-02	8,80E-04	3,65E-02	8,07E-10
0,00E+00	0,00E+00	7,00E-04	1,40E-03	4,54E-02	0,00E+00	3,58E-02	8,80E-04	3,81E-02	3,95E-03
na	na	4,68E-01	2,60E-03	na	na	na	na	na	na
na	na	na	na	4,76E-01	1,39E-08	5,71E-02	0,00E+00	4,82E-01	1,79E-03
na	na	na	na	4,76E-01	5,27E-09	5,71E-02	0,00E+00	4,83E-01	0,00E+00
na	na	na	na	4,76E-01	5,27E-09	5,71E-02	0,00E+00	4,83E-01	0,00E+00
na	na	na	na	4,76E-01	0,00E+00	5,71E-02	0,00E+00	4,82E-01	2,34E-03
na	na	na	na	na	na	na	na	na	na
1,40E-03	3,56E-11	3,50E-03	0,00E+00	na	na	na	na	na	na
1,40E-03	0,00E+00	3,50E-03	0,00E+00	na	na	na	na	na	na
1,40E-03	0,00E+00	3,50E-03	0,00E+00	na	na	na	na	na	na
1,40E-03	0,00E+00	3,50E-03	0,00E+00	na	na	na	na	na	na
na	na	4,75E-01	0,00E+00	na	na	na	na	na	na
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	3,10E-03	7,59E-03
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	3,10E-03	7,59E-03
na	na	4,69E-01	7,45E-09	na	na	na	na	na	na
na	na	4,69E-01	0,00E+00	na	na	na	na	na	na
na	na	4,69E-01	0,00E+00	na	na	na	na	na	na
na	na	4,69E-01	7,45E-09	na	na	na	na	na	na
na	na	na	na	4,61E-01	7,45E-09	5,53E-02	8,80E-04	4,76E-01	0,00E+00
1,40E-03	0,00E+00	4,22E-03	1,44E-03	na	na	na	na	na	na
0,00E+00	0,00E+00	1,40E-03	1,71E-03	0,00E+00	0,00E+00	8,00E-04	9,80E-04	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00
0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	0,00E+00	6,20E-03	9,80E-03
fuscans or NF2 or NF3 to NF1		NF1 to fuscans or NF2 or NF3		fuscans or NF2 or NF3 to NF1		fuscans or NF2 or NF3 to NF1		NF2 or NF3 to NF1	
fuscans and NF2 and NF3		fuscans and NF2 and NF3		fuscans and NF2 and NF3		fuscans and NF2 and NF3		NF2 and NF3	

III. Perspectives et discussion

Grâce à l'utilisation de notre combinaison d'approches de génomique comparative, nous avons mis à jour 116 gènes portant des caractéristiques uniques aux agents responsables de la CBB. Parmi ces 116 gènes, 89 gènes ont été impliqués dans des événements successifs de HGT entre les quatre lignées génétiques NF1, NF2, NF3 et *fuscans* et 45 gènes contenaient au moins un acide aminé spécifique des souches infectant le haricot commun. Fait intéressant, une grande partie de ces 45 gènes candidats codent des protéines impliquées dans la pathogénie ou dans l'interaction avec l'environnement de la plante. Ainsi, ces gènes constituent une piste intéressante pour l'identification et la caractérisation de déterminants de l'adaptation au haricot. Pour confirmer notre hypothèse, j'ai entrepris la construction de mutants de délétion de pour trois gènes candidats afin de les caractériser fonctionnellement, mais ces expériences n'ont pas marché. Par la suite, il serait intéressant de continuer cette production de mutants et effectuer des expérimentations *in planta* pour caractériser le rôle de ces gènes dans l'établissement de la maladie ou dans l'aggravation des symptômes.

Nous avons montré que le HGT était probablement la force motrice prédominante conduisant à la convergence pathologique des souches de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* sur le haricot. Ceci est en accord avec des études antérieures, mais nous avons également montré que des plasmides sont impliqués dans ces transferts horizontaux et dans cette convergence pathologique des souches infectant le haricot commun. En effet, 19 des 45 gènes candidats ont été retrouvés sur le plasmide a et les neuf gènes candidats récupérés par l'approche gain et perte étaient tous situés sur ce plasmide. Généralement, les gènes acquis par HGT ont tendance à être perdus et leur conservation dans les quatre lignées génétiques de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* témoigne de l'importance de ces gènes pour l'adaptation au haricot commun. Ainsi, en plus d'être de potentiels gènes candidats de la spécificité d'hôte, ces gènes plasmidiques constituent également une source d'émergence de la CBB. En effet, des bactéries non pathogènes du haricot commun pourraient acquérir facilement ces déterminants génétiques par transfert horizontal de plasmide.

Par ailleurs, nous avons démontré que Sklf était un outil particulièrement intéressant car il semble être plus sensible que les approches recherchant les gains et pertes de gènes et les approches phylogénétiques. En effet, parmi les 116 gènes portant des caractéristiques uniques aux souches infectant le haricot commun, 109 ont été retrouvés avec cet outil en cherchant des k-mer spécifiques de 24 pb. Cet outil ouvre de nouvelles perspectives pour les études de génomiques comparatives comme par exemple lors de la comparaison de génomes entiers, lors de la recherche de gènes communs et ou spécifiques ou encore lors d'analyse de données RNAseq où l'on peut visualiser l'accumulation de transcrits particuliers.

CHAPITRE 2 :

Diversité et évolution des gènes *tal* des agents responsables de la graisse commune du haricot et d'autres pathovars de *Xanthomonas*

I. Introduction

Les bactéries du genre *Xanthomonas* sont connues pour posséder des ET3 particuliers : les effecteurs TAL. Les effecteurs TAL sont injectés dans la cellule hôte via le SST3 de la bactérie et, une fois à l'intérieur de la plante, ils sont capables de jouer le rôle de facteur de transcription eucaryote en ciblant spécifiquement un gène de la plante. Leur capacité à se lier à une séquence d'ADN hôte spécifique suggère qu'ils représentent de potentiels candidats pour la spécificité d'hôte. De précédentes analyses, ainsi que l'étude présentée dans le chapitre 1, ont montré qu'un grand nombre de gènes avec des fonctions très diverses étaient potentiellement impliqués dans ce processus adaptatif (Hajri *et al.*, 2009; Mhedbi-Hajri *et al.*, 2011; Aritua *et al.*, 2015; Jacques *et al.*, 2016). Malheureusement, l'ensemble de ces analyses n'inclut jamais les gènes *tal*. Ceci est dû à leur séquence d'ADN qui, parce qu'elle comporte beaucoup de répétitions, n'est pas correctement séquencée, ou est assemblée de façon erronée avec les méthodes de séquençage classique. Cependant, depuis quelques années une stratégie de séquençage nommée SMRT (Single Molecule Real-Time) permettant le séquençage de longs fragments d'ADN (~20kbp), s'est avérée efficace pour l'obtention des séquences de *tal* de bonnes qualité (Booher *et al.*, 2015; Wilkins *et al.*, 2015; Peng *et al.*, 2016).

Ainsi, dans le cadre du projet ANR CROpTAL, nous avons séquencé des souches infectant le haricot commun à l'aide du séquençage SMRT et exploité ces nouveaux génomes pour déterminer le rôle des effecteurs TAL dans l'adaptation des *Xanthomonas* au haricot commun. L'ensemble de ces analyses vous est présenté dans la deuxième partie de ce chapitre sous la forme d'un article publié dans la revue BMC Genomics et dont je suis le premier auteur. D'autre part, nous avons séquencé d'autres pathovars de *Xanthomonas* en séquençage SMRT. Nous avons choisi de séquencer huit souches phylogénétiquement proches des souches infectant le haricot et/ou responsables de maladie sur légumineuses afin d'analyser et de comparer leur contenu en effecteurs TAL. Les données de séquençage de trois souches de *X. citri* pv. *vignicola* ont notamment fait l'objet d'une publication dans la revue Genome Announcement, qui vous est présenté en troisième partie de ce chapitre. Des analyses préliminaires sur l'ensemble de ces génomes vous sont présentées dans la quatrième partie de ce chapitre.

II. Implication des effecteurs TAL dans l'adaptation des *Xanthomonas* au haricot commun

Pour déterminer le rôle des effecteurs TAL dans l'adaptation des souches de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* au haricot commun, nous avons séquencé 17 de ces souches, que nous avons soigneusement sélectionnées au préalable, afin de représenter au mieux la diversité des quatre lignées génétiques responsable de CBB. Les génomes ont été séquencés et assemblés par l'Icahn School of Medicine at Mount Sinai de New York. Avec ces nouveaux génomes, enrichis de diverses données génomiques disponibles dans les bases de données publiques, j'ai mené plusieurs analyses. Notamment, avec les conseils de Sophie Bonneau et Nicolas Chen, j'ai mené des analyses phylogénétiques en utilisant les génomes complets des souches ou seulement des séquences de leurs *tal*, j'ai aussi calculé des ANI et construit des CGView à l'aide des conseils de Martial Briand et j'ai mené une analyse de la séquence centrale de répétitions de ces effecteurs, avec l'aide d'Alvaro L. Pérez-Quintero et de Nicolas Chen. Pour compléter ces analyses, Sophie Bonneau a effectué un test de Kishino-Hasegawa-Templeton, Nicolas Chen a estimé des Ks et nous avons repris d'anciens résultats de pouvoir pathogène des 17 souches sur haricot qui ont été produit par mon équipe. L'ensemble de ces travaux a été conduit sous la direction de Marie-Agnès Jacques et Nicolas Chen et ont donc pu faire l'objet d'une publication dont je suis le premier auteur et qui a été accepté dans le journal BMC Genomics en Août 2017 (Ruh *et al.*, 2017a).

RESEARCH ARTICLE

Open Access



Xanthomonas adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors

Mylène Ruh, Martial Briand, Sophie Bonneau, Marie-Agnès Jacques and Nicolas W.G. Chen*

Abstract

Background: Common bacterial blight is a devastating bacterial disease of common bean (*Phaseolus vulgaris*) caused by *Xanthomonas citri* pv. *fuscans* and *Xanthomonas phaseoli* pv. *phaseoli*. These phylogenetically distant strains are able to cause similar symptoms on common bean, suggesting that they have acquired common genetic determinants of adaptation to common bean. Transcription Activator-Like (TAL) effectors are bacterial type III effectors that are able to induce the expression of host genes to promote infection or resistance. Their capacity to bind to a specific host DNA sequence suggests that they are potential candidates for host adaption.

Results: To study the diversity of *tal* genes from *Xanthomonas* strains responsible for common bacterial blight of bean, whole genome sequences of 17 strains representing the diversity of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* were obtained by single molecule real time sequencing. Analysis of these genomes revealed the existence of four *tal* genes named *tal23A*, *tal20F*, *tal18G* and *tal18H*, respectively. While *tal20F* and *tal18G* were chromosomic, *tal23A* and *tal18H* were carried on plasmids and shared between phylogenetically distant strains, therefore suggesting recent horizontal transfers of these genes between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. Strikingly, *tal23A* was present in all strains studied, suggesting that it played an important role in adaptation to common bean. In silico predictions of TAL effectors targets in the common bean genome suggested that TAL effectors shared by *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains target the promoters of genes of similar functions. This could be a trace of convergent evolution among TAL effectors from different phylogenetic groups, and comforts the hypothesis that TAL effectors have been implied in the adaptation to common bean.

Conclusions: Altogether, our results favour a model where plasmidic TAL effectors are able to contribute to host adaptation by being horizontally transferred between distant lineages.

Keywords: *Xanthomonas*, Common bean, TAL effectors, Host adaptation, Horizontal gene transfer

Background

Bacterial pathogens from the genus *Xanthomonas* harbour singular type III effectors called transcription activator-like (TAL) effectors that are able to mimic eukaryotic transcription factors [1–3]. After being injected into host cells via the type III secretion system, TAL effectors migrate to the nucleus where they act as transcription factors thanks to an activation domain localized in the C-terminal (C-ter) region [4, 5]. The central region of the protein is composed of a variable

number of quasi-identical, 34 amino acids long repeats, except for the last repeat of the central region usually corresponding to a half-repeat of 18 amino acids. The pair of residues at positions 12 and 13 of each repeat is named repeat variable diresidue (RVD) and determines the binding specificity of a repeat to a single nucleotide. Thus, the RVD sequence of a TAL effector determines its binding specificity to a target DNA sequence called Effector Binding Element (EBE) [6, 7]. A TAL code linking each RVD to their nucleotide binding affinity has been deciphered, enabling in silico prediction of EBE for a given TAL effector in the host genome [6–11]. The EBE is usually located in the promoter region of susceptibility genes, enabling TAL effectors to promote

* Correspondence: nicolas.chen@agrocampus-ouest.fr
IRHS, INRA, AGROCAMPUS OUEST, Université d'Angers, SFR4207 QUASAV, 42, rue Georges Morel, 49071 Beaucouzé, France



© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

bacterial infection [12–15]. In some cases, the EBE is located upstream of resistance genes termed *executor*, leading to the defence reaction of the host [16, 17].

Because the function of a TAL effector requires its specific binding to DNA, each EBE is a potential source of resistance that can be exploited for developing innovative disease resistance strategies [18, 19]. For instance, any modification of an EBE can lead to loss of susceptibility because the TAL effector is not able to bind to the EBE anymore. Such strategy has been successfully used to engineer resistance of rice to *X. oryzae* pv. *oryzae* by genome editing [20, 21]. Another strategy consists in trapping TAL effectors by placing EBE of conserved TAL effectors in the promoter regions of resistance genes [22–25]. This has successfully been done in rice [26, 27]. These examples are proofs of concepts showing that TAL-based resistance engineering is possible. However, the durability of such resistances and their effectiveness against more diverse pathogens in field has still to be assessed. This requires broadening our knowledge of *tal* genes diversity in *Xanthomonas* populations. However, because of the repeated nature of their central domain, *tal* genes can be misassembled after whole genome sequencing based on Illumina, and/or 454 technologies [28]. Long-read sequencing strategies such as PacBio Single Molecule, Real-Time (SMRT) sequencing have been proved efficient for getting TAL effector sequences [29, 30].

Common bacterial blight of bean (CBB) is the most devastating bacterial disease on common bean. It is due to four different genetic lineages belonging to two distinct species within the genus *Xanthomonas* [31, 32]. The fuscous lineage (fuscans) and the non-fuscous lineages 2 (NF2) and 3 (NF3) all belong to *X. citri* pv. *fuscans* (formerly *X. fuscans* subsp. *fuscans* and *X. axonopodis* pv. *phaseoli* Rep-PCR group 9.6) while the non-fuscous lineage 1 (NF1) belongs to *X. phaseoli* pv. *phaseoli* (formerly *X. axonopodis* pv. *phaseoli* Rep-PCR group 9.4) [33–35]. Pathological convergence of strains responsible for CBB is apparently linked to large events of horizontal gene transfer (HGT) involving a hundred of genes [36]. High-quality genome sequencing of *X. citri* pv. *fuscans* strain CFBP4885 (synonym 4834-R) revealed the existence of two *tal* genes in this strain: *Xfital1* and *Xfital2*, both located on plasmids [37]. PCR assays suggested that both *Xfital1* and *Xfital2* exist in the fuscans and NF1 lineages [36]. However, the diversity of TAL effectors in *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* has not yet been described.

In order to study the diversity of TAL effectors in *Xanthomonas* strains responsible for CBB, we have generated the whole genome sequence of 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains by using PacBio SMRT sequencing. The sequence of *tal* genes

was extracted and phylogenetic analyses including the *tal* gene sequences from other *Xanthomonas* strains were performed. The comparison of the phylogeny of *tal* genes with the corresponding phylogeny of strains allowed us to predict potential HGT events between the four *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages, therefore suggesting a role of *tal* genes in host adaptation.

Results

Pathogenicity, genome quality and *tal* genes content

To estimate the diversity of *tal* genes present in strains responsible for CBB, the genome sequence of 17 strains representing the diversity of the four genetic lineages of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* was obtained through PacBio SMRT sequencing (Table 1). All these strains were pathogenic on common bean, although with different degrees of aggressiveness (Table 1). The majority of strains (12/17) were highly aggressive on common bean while the other (5/17) were less aggressive. Overall, the average genome size and gene content (5.24 $\pm$ 0.08 Mbp; 4596 $\pm$ 160 genes) were quite classical for *Xanthomonas* strains [36–42]. Each genome sequence consisted of two to six contigs, with a maximum contig size corresponding to the expected chromosome size (5.1 Mbp) and additional circularized contigs corresponding to extra-chromosomal plasmids. Plasmids A, B and C were named based on similarity with corresponding plasmids previously described in strain CFBP4885, also known as 4834-R [37] (Fig. 1).

Following BLASTN searches in the 17 SMRT-sequenced genomes, a set of 26 complete *tal* genes and 14 putative *tal* pseudogenes were retrieved. Four pseudogenes corresponded to partial N-ter-encoding sequences, two corresponded to degenerated *tal* gene sequences bearing several frameshifts leading to premature stop codons, and eight corresponded to full-length *tal* genes containing a deletion of one nucleotide in either the N-ter- or the C-ter-encoding region (Additional file 1). Given the high rate of indels observed after SMRT sequencing [43], we sequenced the PCR products corresponding to these eight full-length *tal* pseudogenes with Sanger sequencing technology and all were validated as complete *tal* genes. This led to a total of 34 complete *tal* genes and six pseudogenes. Only complete *tal* genes were retained for further analyses (Additional file 2).

Strain CFBP4885 was previously sequenced using a combination of Illumina, 454 and Sanger technologies [37]. This high-quality genome sequencing revealed the existence of two *tal* genes (*Xfital1* and *Xfital2*) in this strain, therefore allowing comparison of *Xfital1* and *Xfital2* to the *tal* sequences obtained here by SMRT

Table 1 Genome features of the 17 bacterial strains responsible for CBB sequenced in this study

Pathovar (lineage)	Strain	Country (date of isolation)	Disease index ^a	Contig number	Genome size (bp)	Max. contig length (bp)	% GC	Gene number	GenBank Accession numbers
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	France (1998)	3	6	5,173,315	5,012,288	64.66	4519	CP020992 to CP020997
	CFBP6165	Canada (1957)	3	3	5,118,412	5,054,301	64.80	4412	CP020998 to CP021000
	CFBP6166	South Africa (1963)	3	5	5,200,310	5,025,712	64.67	4566	CP021001 to CP021005
	CFBP6167	USA (1954)	2	5	5,388,310	5,176,780	64.54	4755	CP021018 to CP021022
	CFBP6975	France (1994)	3	5	5,244,382	5,092,018	64.64	4634	CP021006 to CP021010
	CFBP7767R	Cameroon (2009)	3	3	5,332,674	5,227,850	64.56	4893	CP021012 to CP021014
<i>X. citri</i> pv. <i>fuscans</i> (NF2)	CFBP6988R	La Réunion (2000)	3	2	5,199,272	5,122,265	64.62	4526	CP020979 to CP020980
	CFBP6989	La Réunion (2000)	3	2	5,179,726	5,111,201	64.62	4498	CP020981 to CP020982
	CFBP6990	La Réunion (2000)	3	2	5,193,607	5,123,413	64.64	4512	CP020983 to CP020984
	CFBP6991	La Réunion (2000)	3	3	5,309,897	5,107,678	64.59	4673	CP021015 to CP021017
<i>X. citri</i> pv. <i>fuscans</i> (NF3)	CFBP6992	La Réunion (2000)	1	2	5,299,787	5,232,548	64.57	4478	CP020985 to CP020986
	CFBP6994R	Tanzania (1990)	3	2	5,191,148	5,124,593	64.70	4399	CP020987 to CP020988
	CFBP6996R	La Réunion (2000)	2	2	5,145,832	5,078,988	64.73	4352	CP020989 to CP020990
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP412	USA (NA)	2	3	5,186,811	5,071,497	64.89	4568	CP020964 to CP020966
	CFBP6164	Romania (1966)	2	4	5,399,156	5,254,225	64.63	4833	CP020967 to CP020970
	CFBP6546R	USA (NA)	3	4	5,258,370	5,081,709	64.78	4801	CP020971 to CP020974
	CFBP6982	La Réunion (2000)	3	4	5,291,156	5,100,676	64.75	4712	CP020975 to CP020978
Average				3	5,241,892	5,117,514	64.67	4596	

^aDisease symptoms were scored 11 days after bathing the first trifoliate leaf into bacterial suspensions at 1×10^7 CFU/mL according to the following scale: 0 = no symptoms, 1 = 1 to 50 spots per leaf, 2 = 51 to 200 spots per leaf, necrosis and sagging and 3 = more than 200 spots per leaf, necrosis, sagging or leaf death. The average of pathogenicity scores were calculated from the values of three plants per strain

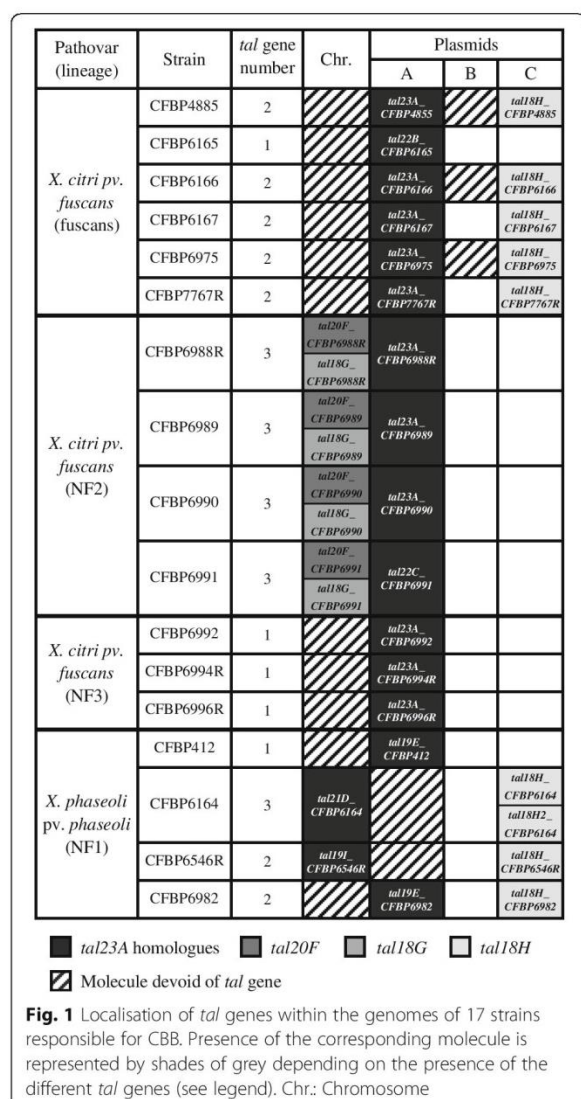
sequencing for the same strain. Unexpectedly, the distinct sequencing technologies differed in the predicted length of *tal* genes. The lengths of *Xfutal1* and *Xfutal2* in strain CFBP4885 have been previously estimated at 4935 bp and 3405 bp, respectively [37]. However, lengths of 4017 bp and 3507 bp were found for *tal* genes obtained after SMRT sequencing. These sequences differed only by their number of repeats, indicating that the observed differences in length were due to assembly errors in their repeat regions. To validate *Xfutal1* and *Xfutal2* sequence lengths, we designed PCR primers on conserved sequences corresponding to the N-ter- and C-ter-encoding regions of XfuTAL1 and XfuTAL2 proteins. PCR profiles were consistent with the lengths obtained after SMRT sequencing, and Sanger sequencing of the PCR products further confirmed that *Xfutal1* and *Xfutal2* sequences have been successfully sequenced and assembled in the SMRT-sequenced genomes (Additional file 3). These results corroborate previous analyses showing that SMRT sequencing led to accurate assembly of *tal* genes [29, 30].

The length of the 34 complete *tal* genes found in our 17 genome sequences ranged between 3501 bp and 4017 bp, which corresponded to TAL effectors possessing from 18 to 23 repeats. The overall RVD composition was similar to that of previously sequenced *tal* genes from other *Xanthomonas* pathovars [4, 29] with

more than 97% of RVD being HD, NN, NG, NI, N* or NK (Additional file 4). The only unusual RVD was HY, retrieved in nine TAL effectors and previously reported in only six other TAL effectors from *X. oryzae* pv. *oryzicola* and *Ralstonia solanacearum* [29, 44]. Nine unique RVD combinations ranging from 18 to 23 RVD were identified (Fig. 2). We named these nine unique RVD combinations using their number of repeats followed by a letter corresponding to each unique RVD combination, from A to I (Fig. 2). For example, for strain CFBP4885, corrected sequences of XfuTAL1 with 23 RVD and XfuTAL2 with 18 RVD corresponded to RVD combinations TAL23A and TAL18H, respectively. Following this nomenclature, *Xfutal1* and *Xfutal2* genes were renamed as *tal23A_CFBP4885* and *tal18H_CFBP4885*, respectively (Figs. 1, 2).

***tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages originate from a single ancestor**

To study the evolution of *tal* genes from strains responsible for CBB, we compared the phylogeny of *tal* genes to the phylogeny of *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* strains. For this, we first reconstructed the phylogeny of organisms using the whole predicted proteomes of the 17 *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* strains from our study plus 31 strains representing 10 *Xanthomonas* pathovars from which complete



TAL effectors have previously been found, and 26 additional strains responsible for CBB sequenced previously (Additional file 5). The CVTree topology (Fig. 3a) was congruent with previously published *Xanthomonas* phylogenetic trees [33, 35, 45]. As previously shown, strains responsible for CBB split into four distinct genetic lineages [31, 46]. Three lineages, fuscans, NF2 and NF3 belonged to the *X. citri* species and one lineage, NF1, belonged to the *X. phaseoli* species (Fig. 3a). These groupings were confirmed by an average nucleotide identity above 99% within each genetic lineage, and under 95% between *X. phaseoli* pv. *phaseoli* and the three other lineages (Additional file 6).

In evolutionary studies of *tal* genes, the central region is usually discarded because it is composed of a variable

number of highly-conserved repeats that are able to recombine with each other, which can obscure alignments and phylogeny [29, 30]. Therefore, we reconstructed two phylogenies of *tal* genes using DNA sequences corresponding either to the N-ter- or to the C-ter-encoding region of complete TAL effectors (Figs. 3b, 3c). Both trees had a topology different from the organism CVTree, suggesting that recombination and/or horizontal transfers of *tal* genes occurred during the diversification of xanthomonads. In particular, *tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* clustered into a monophyletic group, together with *tal* genes from *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis*. For both trees, *tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*, were divided in three clades (Fig. 3b, c).

The CDS corresponding to RVD combinations TAL23A, TAL22B, TAL22C, TAL21D, TAL19E and TAL19I appeared to be derived from a recent common ancestor (Fig. 3b, c). Interestingly, every strain responsible for CBB possessed a single CDS corresponding to one of these six RVD combinations. These CDS were located on plasmid A for a majority of strains (15/17) except for strains CFBP6164 and CFBP6546R where it was located on the chromosome (Fig. 1). Also, all these CDS were surrounded by a conserved 17 kbp region whatever their location (see detailed explanation in the fifth section below). Together, these observations indicate that these CDS corresponded to allelic versions of a single gene, thus they were all considered as *tal23A* homologues. These *tal23A* homologues diverged by indels of complete internal repeats and/or by substitutions or recombinations affecting their RVD content (Fig. 2). Two additional genes, *tal20F* and *tal18G*, were only found in the four strains belonging to the NF2 lineage and had a chromosomal location (Fig. 1). These two *tal* genes were very similar to each other (Fig. 3b and c) and diverged by an indel involving two contiguous RVD (Fig. 2). Both *tal* genes were retrieved in every NF2 strain, indicating that they corresponded to paralogues resulting from a duplication predating the diversification of the NF2 lineage. Finally, *tal18H* was retrieved in most strains belonging to the fuscans (5/6) or NF1 (3/4) lineages, and was always located on plasmid C (Fig. 1). Notably, the three strains from the NF3 lineage did not present any additional *tal* gene other than *tal23A* (Fig. 1).

Partial *tal* sequences retrieved in previously sequenced genomes from the NF1 and fuscans lineages [36] confirmed that *tal* genes found in both lineages were all related to *tal23A* or *tal18H* genes (Additional file 7). Also, strains responsible for CBB isolated on lablab bean (*Dolichos lablab*), a legume plant phylogenetically close to common bean, have previously been described [36].

Name	RVD sequence alignment	Pathovar	Genetic lineages
TAL23A	NI NG NI NG NI NN NG NN HD HD NN NN HD NI HD NI HD HD HD HD HD NG NG	<i>X. citri</i> pv. <i>fuscans</i>	fuscans + NF2 + NF3
TAL22B	NI NG NI NG NI NN NG NN HD NN NN NN HD NI HD NI HD HD HD HD HD NG	<i>X. citri</i> pv. <i>fuscans</i>	fuscans
TAL22C	NI NG NI NG NI NN NG NN HD HD NN NN HD NI HD NI HD HD HD HD HD NG NG	<i>X. citri</i> pv. <i>fuscans</i>	NF2
TAL21D	NI NG NI NG NI NN NG NN HD HD NN NN HD NI HD NI HD HD HD HD HD NG HD	<i>X. phaseoli</i> pv. <i>phaseoli</i>	NF1
TAL19E	NI NG NI NG NI NN NG NN HD HD NN NN HD NI HD NI HD HD HD HD HD NG	<i>X. phaseoli</i> pv. <i>phaseoli</i>	NF1
TAL19I	NI NG NI NG NI NN NG NN HD HD NN NN HD NI HD NI HD HD HD HD HD NG HD	<i>X. phaseoli</i> pv. <i>phaseoli</i>	NF1
TAL20F	N* HD HD NG NG NI NG NI NI NN NI HD HD HD HD HD NI HD NK NG	<i>X. citri</i> pv. <i>fuscans</i>	NF2
TAL18G	N* HD HD NG NG NI NG NI NI NN NI HD HD HD HD NI HD NK NG	<i>X. citri</i> pv. <i>fuscans</i>	NF2
TAL18H	NI NG HD NG HD NI NG NI HD HY NN N* HD NG HY NN HD NG	<i>X. citri</i> pv. <i>fuscans</i> and <i>X. phaseoli</i> pv. <i>phaseoli</i>	fuscans and NF1

Fig. 2 Alignment of the nine unique RVD sequences retrieved in TAL effectors from 17 strains responsible for CBB. TAL effectors that could be aligned are grouped together, indels are represented by black rectangles and putative substitutions are marked with grey colour

These strains clustered with *X. citri* pathovars other than *X. citri* pv. *fuscans* and *X. citri* pv. *aurantifolii* (Fig. 3a). Partial *tal* sequences were retrieved in the genomes from these lablab-associated strains. Strikingly, these *tal* genes did not cluster with any of the three aforementioned clades, further suggesting that TAL effectors from *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* were specific for strains having common bean as natural host.

Recombination occurred between *tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*

We observed discrepancies between the trees generated with the N-ter- or C-ter-encoding sequences. Indeed, in the tree generated using the C-ter-encoding sequence, *tal18G* and *tal20F* formed a clade external to the other *tal* genes from strains responsible for CBB, while in the tree generated using the N-ter-encoding sequence, *tal18H* clustered outside from the others (Fig. 3b, c). Tree topologies were significantly different ($P < 0.05$) based on the Kishino-Hasegawa-Templeton test [47]. This suggested that a recombination event occurred between the N-ter and C-ter encoding regions of these *tal* genes. To confirm this observation, we estimated recombination using seven nonparametric detection programs from the RDP4 package [48]. Two recombination events were retrieved on an alignment of the concatenated N-ter- and C-ter-encoding regions of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains, but none were retrieved on the separated N-ter- or C-ter-encoding regions, suggesting that recombination break-points were located within the central repeat region of *tal* genes (data not shown). The first event was detected in *tal18H*, predicted as recombinants between the common ancestor of *tal18G* and *tal20F*, and *pthB* from the *X. phaseoli* pv. *manihotis* strain. The second event was

detected within *pthB*, predicted as a recombinant between *pthC2* and other *tal* genes from *X. citri* pv. *aurantifolii*. Both recombination events were in accordance with the differences observed between the tree topologies (Fig. 3b, c).

tal genes have undergone multiple horizontal transfers between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages

To estimate the time of divergence between *tal* genes, we compared the values of pairwise nucleotide synonymous substitution rates at silent sites (Ks) from the N-ter or C-ter-encoding regions of *tal* genes to Ks values from seven housekeeping genes (*atpD*, *dnaK*, *efp*, *glnA1*, *gyrB*, *rpoD* and *suxA*) [46] after removal of recombinant regions. In absence of codon usage bias, and for genes under constant evolutionary rates, Ks is an estimation of neutral evolution because it does not take into account the nonsynonymous sites that can be under selection pressure [49, 50]. Here, codon usage was constant among the different studied taxa, and no significant divergence from neutrality was retrieved ($P > 0.05$) after Tajima's or Fu and Li's relative rate tests (data not shown). Therefore, Ks could be used as an approximation of the time of divergence between genes or taxa [51], with higher Ks meaning longer time of evolution between two sequences.

The mean Ks for *tal* genes from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains were 0.0930 and 0.1844 for the N-ter- and C-ter-encoding regions, respectively (Fig. 3b, c). Thus, Ks value for the C-ter-encoding region was approximately twice as high as the Ks value for the N-ter-encoding region, indicating that the C-ter-encoding region has accumulated more synonymous substitutions at silent sites than the N-ter-encoding

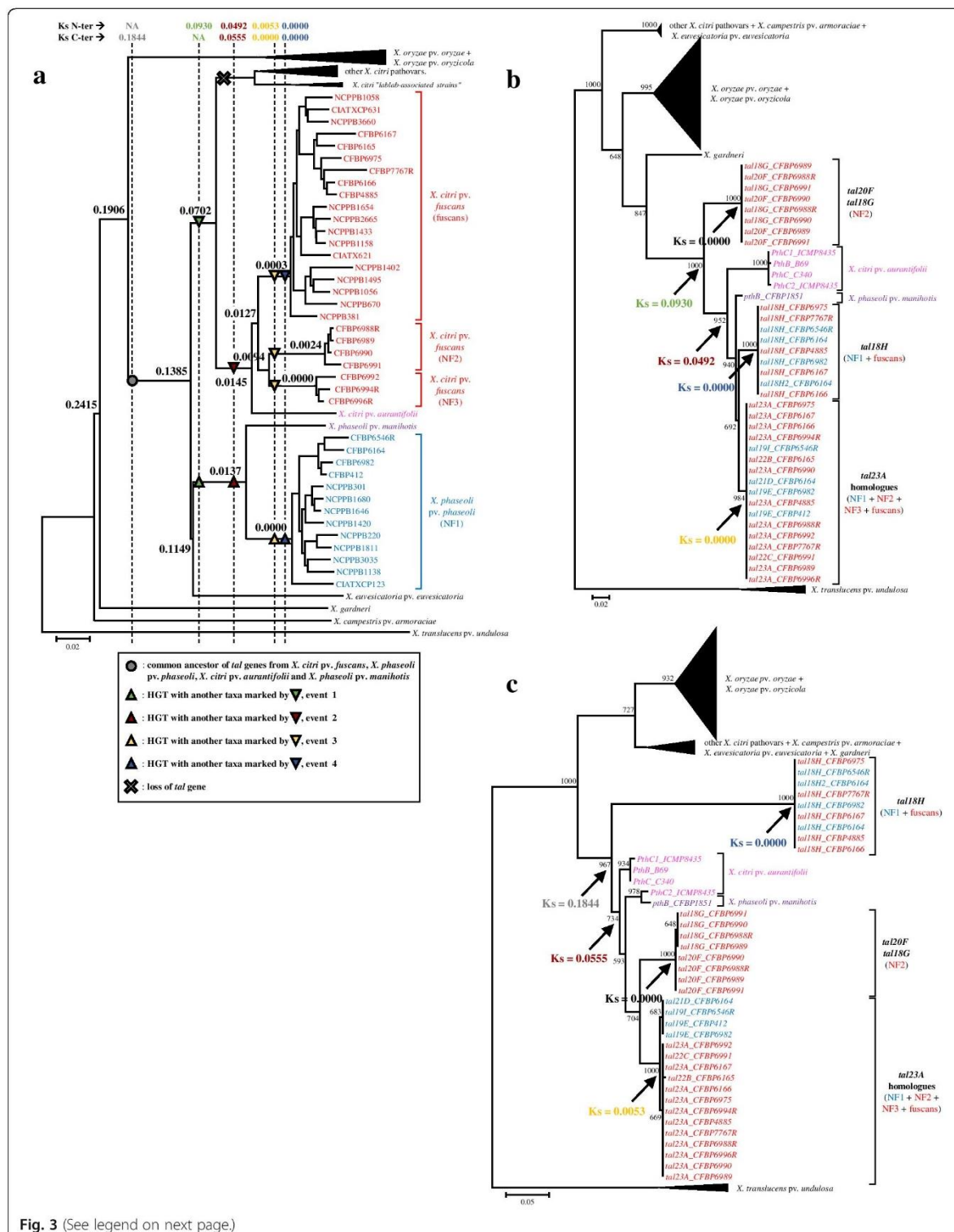


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Phylogeny of *Xanthomonas* strains and *tal* genes. Strains or *tal* genes from the *X. citri* pv. *fuscans* genetic lineages fuscans, NF2 and NF3 are indicated in red, *X. phaseoli* pv. *phaseoli* NF1 lineage in blue, *X. citri* pv. *aurantifolii* in pink and *X. phaseoli* pv. *manihotis* in purple. Ks values corresponding to the divergence of *tal* genes are indicated in the different trees using the same colours. **a** Phylogenetic tree representing the evolution of *Xanthomonas* strains constructed using CTree on the whole proteome of strains. Numbers at node correspond to mean Ks values calculated on seven housekeeping genes. Vertical dotted lines represent time of *tal* gene divergence estimated by Ks values written above. The most parsimonious course of events for the evolution of *tal* genes is highlighted according to the legend. NA: Not applicable. **b** ML tree constructed on the nucleotide sequences of N-ter-encoding region of *tal* genes. **c** ML tree constructed on the nucleotide sequences of C-ter-encoding region of *tal* genes. For **(b)** and **(c)**, bootstrap values greater than 50% are shown for 1000 replicates and horizontal scale bars represent number of nucleotide substitutions per site

region did. This suggested that the C-ter-encoding region had a more ancient origin than the N-ter-encoding region in these strains. The Ks value for the C-ter-encoding region (0.1844) was in the same range as the Ks value calculated for housekeeping genes corresponding to the split between the *X. oryzae* species and the *X. axonopodis* species complex (0.1906), suggesting that the C-ter-encoding region has been transmitted vertically from the ancestor of *X. axonopodis* to the present days (Fig. 3a). In contrast, Ks value for the N-ter-encoding region (0.0930) was lower than the Ks value for housekeeping genes corresponding to the split between *X. euvesicatoria* and *X. phaseoli* species (Ks = 0.1149) and higher than the Ks value corresponding to the emergence of the *X. citri* species (Ks = 0.0702). This suggested that the N-ter-encoding region of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* diverged after the split between *X. euvesicatoria* and *X. phaseoli* species and before the emergence of the *X. citri* species. This can be explained by a HGT between the ancestors of the *X. phaseoli* and *X. citri* species (event 1 on Fig. 3). Then, absence of orthologous *tal* genes in other *X. citri* pathovars would be due to a *tal* gene loss (Fig. 3a).

For both N-ter- and C-ter-encoding regions, *tal* genes from *X. citri* pv. *aurantifolii*, *X. citri* pv. *fuscans*, *X. phaseoli* pv. *manihotis* and *X. phaseoli* pv. *phaseoli* strains were grouped in a clade having *tal18G* and *tal20F* or *tal18H* as direct outgroups, respectively (Fig. 3b and c). In absence of HGT, the time of divergence for this clade should reflect the time of divergence between the *X. citri* and *X. phaseoli* species (Ks = 0.1385). However, the N-ter- and C-ter-encoding regions diverged at Ks values of 0.0492 and 0.0555, respectively (Fig. 3b and c). These Ks values were similar to each other, and lower than the Ks at the split between *X. citri* pathovars *aurantifolii* plus *fuscans* and the other *X. citri* pathovars (0.0702), but higher than Ks at the split between *X. citri* pv. *aurantifolii* and *X. citri* pv. *fuscans* (0.0145) and at the split between *X. phaseoli* pv. *manihotis* and *X. phaseoli* pv. *phaseoli* (0.0137; Fig. 3a). This suggested that a HGT of *tal* genes occurred between the ancestor of *X. citri* pv. *aurantifolii*

and *X. citri* pv. *fuscans*, and the ancestor of *X. phaseoli* pv. *manihotis* and *X. phaseoli* pv. *phaseoli* (event 2 on Fig. 3a).

Presence of *tal23A* homologues in the phylogenetically distant *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains suggested that these *tal* genes were horizontally transferred between the ancestor of the NF1 lineage and the ancestor of one of the three *X. citri* pv. *fuscans* lineages (event 3 on Fig. 3a). Ks values for *tal23A* homologues between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages were zero for the N-ter-encoding region and 0.0053 for the C-ter-encoding region, indicating that *tal23A* homologues accumulated very few synonymous substitutions between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages, and indicating that this HGT was very recent. These low Ks values were similar to the values found within each lineage for housekeeping genes (NF1: 0.0000; NF2: 0.0024; NF3: 0.0000; fuscans: 0.0003), further confirming that this HGT occurred in the same time frame as the divergence of strains within each lineage. Moreover, most *tal23A* homologues were located on plasmid A. This plasmid was present in all *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains but not in *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains [52, 53] (Fig. 1). Plasmid A was highly conserved with, on average, more than 97% nucleotide identity on more than 36% of the total plasmid length between the strains from the three *X. citri* pv. *fuscans* lineages and the strains from the NF1 lineage (Additional file 8, Additional file 9). This suggested that the event responsible for the horizontal transfer of *tal23A* homologues corresponded in fact to the horizontal transfer of a large portion of plasmid A. The largest conserved portion of plasmid A between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains corresponded to 76% of plasmid A from strain CFBP6546R (NF1 lineage) and 77% of plasmid A from strain CFBP6988R (NF2 lineage) (Additional file 9). Thus, the most probable event would correspond to a HGT between the NF2 and NF1 lineages. The region between strains CFBP6988R and CFBP6546R corresponded to 39 kbp comprising 38 predicted protein-encoding genes (Additional file 10a,

Additional file 11). These predicted proteins included 11 hypothetical proteins and 27 proteins with predicted functions including seven proteins involved in the type IV conjugal transfer system, three transposases, and two DNA topoisomerases (Additional file 11). In strain CFBP6546R, the *tal23A* homologue was located on the chromosome. Therefore, it was not retrieved in the 39 kbp region shared by both strains. In strain CFBP6988R, this *tal23A* homologue was located contiguous to the 39kbp region for a total of ~44 kbp (Additional file 10a). This suggests that a region of ~44 kbp from plasmid A was horizontally transferred between the NF2 and NF1 lineages, followed by the integration of *tal23A* homologues in the chromosome of some strains belonging to the NF1 lineage (this event is explained with more details in the next section).

Finally, *tal18H* was common to most strains from the NF1 and fuscans lineages but absent in the other lineages, suggesting that *tal18H* has been horizontally transferred between the ancestors of NF1 and fuscans strains (event 4 on Fig. 3a). In accordance with this, Ks values between NF1 and fuscans lineages were zero for both N-ter and C-ter-encoding regions, indicating that *tal18H* did not accumulate any nonsynonymous substitutions between the NF1 and fuscans lineages. Moreover, *tal18H* was located on plasmid C, and its presence was directly correlated with the presence of plasmid C (Fig. 1). Plasmid C was highly conserved with, on average, more than 99% nucleotide identity on more than 59% of the total plasmid length between NF1 and fuscans strains (Additional file 8, Additional file 12), suggesting that *tal18H* has been horizontally transferred along with a large portion of plasmid C. The largest conserved portion between *X. citri* pv. *fuscans* (fuscans) and *X. phaseoli* pv. *phaseoli* (NF1) strains corresponded to 68% of plasmid C from strain CFBP6982 (NF1 lineage) and 76% of plasmid C from strain CFBP6166 (fuscans lineage) (Additional file 12). The horizontally transferred region corresponded to ~30 kbp that were split in different smaller regions and/or duplicated depending on the strain, suggesting that recombination and/or local duplications occurred after the HGT event (Additional file 10b, Additional file 13). In addition to the *tal18H* gene, these regions comprised 26 predicted protein-encoding genes including eight hypothetical proteins and 18 proteins with predicted functions including the type III effector XopC1, two proteins involved in the type IV conjugal transfer system, three ribonucleases and two type II-like restriction endonucleases (Additional file 13). Strains CFBP412 and CFBP6165 were the only strains from the NF1 and fuscans lineages not possessing plasmid C, suggesting that they recently lost this plasmid, together with the *tal18H* gene (Fig. 1).

Recent *tal* gene duplications and movements are associated with the presence of insertion sequences (IS) and transposons

Presence of IS and/or Tn3-like transposons has been proved as a hallmark of *tal* gene movements and duplication [54]. Here, all *tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* were associated with IS3 and/or Tn3-like transposons. To estimate if these IS3 and/or Tn3-like transposons were responsible for *tal* gene movements or duplications, we analysed the genomic environment of relocated or duplicated *tal* genes.

As described above, our analyses indicated that *tal23A* homologues have been transmitted by a HGT between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains involving plasmid A. The majority of *tal23A* homologues (15/17) was located on plasmid A except for two *X. phaseoli* pv. *phaseoli* strains, CFBP6164 and CFBP6546R, where *tal21D_CFBP6164* and *tal19I_CFBP6546R* had a chromosomal location (Fig. 1). Strains CFBP6164 and CFBP6546R were closer to each other than any other strains (Fig. 3a). This suggested that the *tal* gene moved from the plasmid to the chromosome in the common ancestor of strains CFBP6164 and CFBP6546R. Analysis of the genomic environment indicated that this movement was associated with the relocation of a 17 kbp region surrounding this *tal* gene (Additional file 14a). Presence of intact IS3 sequences framing the receiving region of the chromosome of CFBP6164 and CFBP6546R strains, and presence of an IS3 remnant at the donor site in plasmid A in these strains suggested that recombination between IS3 sequences was involved in this movement, although we could not define the breakpoints precisely (Additional file 14a).

In each NF2 strain, *tal* genes *tal20F* and *tal18G* had nearly-identical sequences. These *tal* genes were not retrieved outside from the NF2 lineage, and the Ks value between these two copies was zero for both N-ter- and C-ter-encoding regions, indicating that these two genes resulted from a duplication in the direct ancestor of the NF2 lineage (Figs. 1 and 3). These *tal* genes were oriented back to back on the chromosome, and separated by around 720 kb. Both *tal* genes were surrounded by two Tn3-like transposons that were intact except for the region downstream *tal18G* (Additional file 14b). This suggested that this duplication involved a recombination between Tn3-like transposons, and indicated that *tal18G* was a copy of *tal20F*.

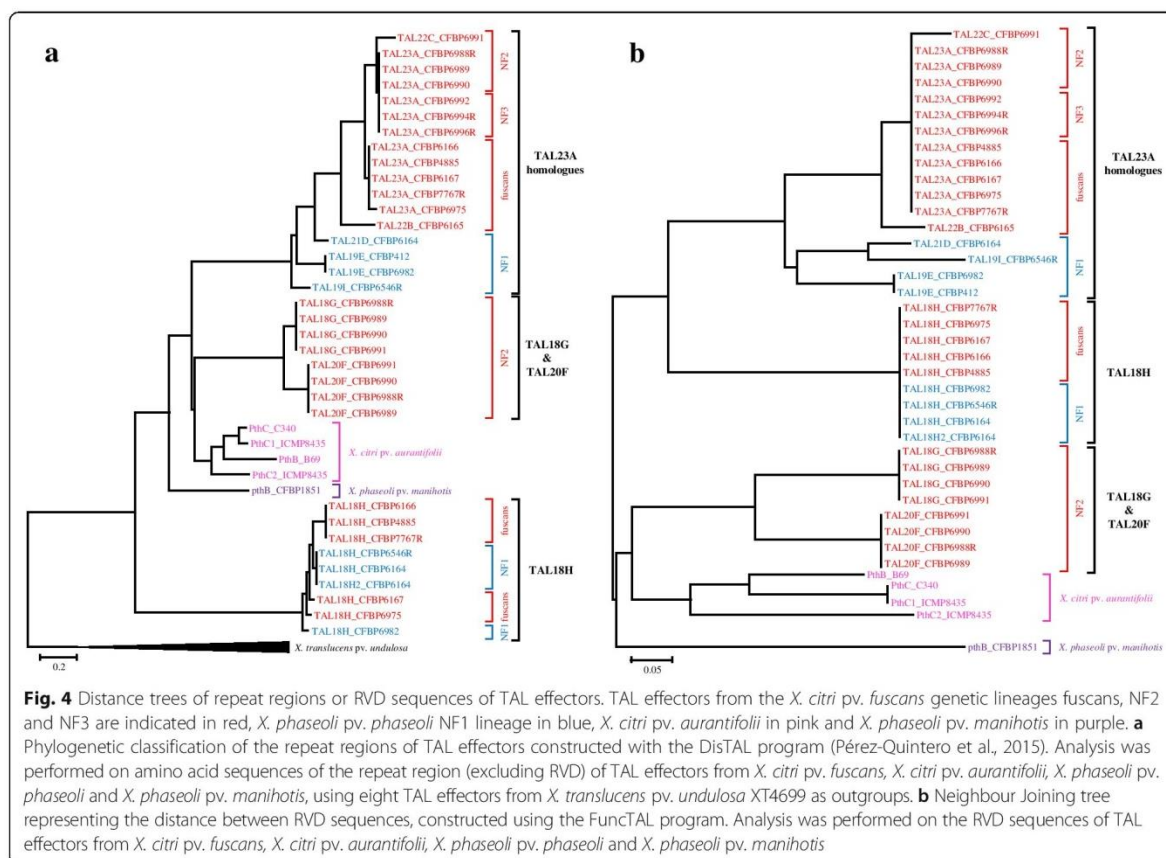
In plasmid C, *tal18H* was present as a single copy except for the CFBP6164 strain where two copies were found on plasmid C (*tal18H_CFBP6164* and *tal18H2_CFBP6164*). This duplication corresponded to the tandem duplication of a seven kbp region including the *tal* gene and flanked by two IS3 sequences, suggesting that this duplication was the consequence of an

unequal crossing-over between these two IS3 (Additional file 14c). Interestingly, the N-ter-encoding region of *tal18H2_CFBP6164* comprised the insertion of a cytosine 88 bp after the start codon. This led to a frameshift resulting in the relocation of the start codon 27 aminoacids upstream of the classical N-ter-encoding region and modifying the first 30 aminoacids of the *TAL18H2_CFBP6164* effector.

In silico predictions suggest that *tal18H* and *tal23A* homologues were involved in *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* adaptation to common bean

The repeat region of TAL effectors is important because it is involved in the binding specificity of TAL effectors to the promoters of targeted host plant genes [6, 7]. Previous studies suggested that recombination, duplication and deletion events within repeats were involved in *tal* gene evolution [54–56]. To study the evolution of the repeat regions of *tal* genes from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli*, we used the DisTAL program that is dedicated to the analysis of the evolution of TAL repeats excluding the RVD [57]. *TAL23A* homologues were distributed according to their genetic lineages

(Fig. 4a). An important diversification was observed for the NF1 lineage that comprised genes encoding RVD combinations *TAL19E*, *TAL19I* and *TAL21D* (Figs. 2, 4a). In fuscans, NF2 and NF3 lineages, repeats were less diversified and *TAL23A* RVD combination was retrieved in 11 strains out of 13. An interesting point is that genes encoding *TAL23A* RVD combination presented a lineage-dependant diversification of their repeats outside of the RVD, while keeping the same RVD combination. Indeed, repeats encoding *TAL23A* in the NF2 and NF3 lineages were divergent from those found in the fuscans lineage (Fig. 4a). Also, within genes encoding *TAL23A* RVD combination in the fuscans lineage, *tal23A_CFBP6975* presented a repeat region with a slight divergence from its orthologues in other strains from the fuscans lineage. Similarly, diversification occurred in the repeats of *tal18H*, leading to two different repeat profiles in the NF1 lineage and three in the fuscans lineage (Figs. 2, 4a). Thus, for *TAL23A* and *TAL18H* RVD combinations, RVD were retained despite repeat diversification, suggesting that these RVD combinations were functionally important for CBB causal agents.



We performed a comparative analysis of the 772 individual repeats from the 39 TAL effectors retrieved in *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains after removal of the RVD. These repeats were distributed in 48 unique repeat sequences (Additional file 15). All repeats comprised 34 amino-acids, or 18 amino-acids for the last half-repeat, and no repeat of unusual length was retrieved [58]. One predominant repeat was found 259 times (in red in Additional file 15, Additional file 16), while 14 repeats were retrieved only once. Interestingly, the predominant repeat was conserved in all TAL effectors whatever the associated RVD. Also, this repeat differed by at least one amino acid from any other repeat found in other pathovars described so far (data not shown), further highlighting that *tal* genes from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains share a common evolutionary history different from the evolutionary history of *tal* genes found in other *Xanthomonas* strains. A second repeat found 33 times was shared by all TAL effectors whatever the associated RVD, except for *X. phaseoli* pv. *manihotis* (in green in Additional file 15, Additional file 16). A third repeat found 31 times was shared among TAL23A homologues, TAL20F, TAL18G, and PthC from *X. citri* pv. *aurantifolii* strain (in purple in Additional file 15, Additional file 16). Two additional repeats were specific for TAL effectors from *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains (in blue and yellow in Additional file 15, Additional file 16). This is in agreement with the DisTAL tree general topology (Fig. 4a) and suggested that TAL effectors of *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains had a specific evolutionary history compared to *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains.

As mentioned before, RVD diversification of TAL effectors from CBB causal agents led to nine different RVD combinations (Fig. 2). To analyse the functional relationships between these TAL effectors, we used the FuncTAL program that compares DNA binding specificities using RVD sequences [57]. Despite the diversification of RVD sequences observed among TAL23A homologues, or between TAL20F and TAL18G, the FuncTAL tree showed that TAL effectors coming from a same clade clustered together (Fig. 4b). Interestingly, TAL23A homologues were closer to TAL18H than to TAL20F and TAL18G, while their repeats were more distant from TAL18H according to the DisTAL tree. This could be a trace of functional convergence between TAL23A homologues and TAL18H.

To search for putative common targets of TAL effectors within and between the three phylogenetic groups of *tal* genes, we used the TALVEZ EBE prediction tool on the promoterome (3 kb-promoter regions) of the

common bean genome sequence [11, 59]. For each single RVD combination found in TAL effectors from *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* strains, the top 10 predicted targets according to the scores calculated by TALVEZ did not correspond to any other previously described target of TAL effectors from other *Xanthomonas* pathovars [12, 18, 60–67] (Additional file 17). However, several top 10 predicted targets were shared by TAL effectors coming from a same phylogenetic group. Indeed, eight predicted targets were shared by at least two of the six TAL23A homologues, including two targets predicted for four RVD combinations (Additional file 17). This was the case for a peptidase (Phvul.006G032900) predicted for TAL23A, TAL22B, TAL19E and TAL19I, and a 40S ribosomal subunit (Phvul.002G052200) predicted for TAL23A, TAL22C, TAL21D and TAL19I. In addition, two different genes encoding nucleolar ribonucleoprotein subunits (Phvul.008G022975 and Phvul.003G201700) were predicted as targets for TAL23A, TAL22B and TAL21D (Additional file 17). Also, four out of the first 10 predicted targets were shared by TAL20F and TAL18G. This included a protein of unknown function (Phvul.005G097300), a carboxyl methyltransferase (Phvul.006G048600), a nucleic acid deaminase (Phvul.003G292800) and an oligopeptide transporter-related protein (Phvul.010G009400). Interestingly, two genes encoding UDP-glycosyl transferases (Phvul.003G097300 and Phvul.009G148800) were predicted as targets for TAL23A, TAL22C and TAL21D as well as for TAL18H, further suggesting that a functional convergence exist between TAL23A homologues and TAL18H (Additional file 17).

Discussion

To study the diversity of TAL effectors from *Xanthomonas* strains responsible for CBB, we generated whole genome sequences for 17 strains representing the diversity of the four *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* genetic lineages. PacBio SMRT sequencing assemblies allowed us to retrieve 40 sequences of complete *tal* genes and pseudogenes. Previously, only two *tal* genes, *Xfital1* and *Xfital2* from CFBP4885 strain were reported for CBB agents [37]. Here, we show that *Xfital1* and *Xfital2* repeat regions contained assembly errors despite having been sequenced using a combination of Sanger, Illumina and 454 sequencing strategies. We thus confirmed that SMRT sequencing is a good methodology to obtain the sequence of TAL effector repeats [29, 30, 68, 69], which is essential for good prediction of their potential targets in host genomes [6, 7, 11]. However, corrections had to be made for deletions observed in the N-ter- and/or C-ter-encoding regions. These deletions were observed only in

G- or C-rich stretches, which could correspond to error hotspots for SMRT sequencing.

Recombination and HGT are important driving forces of ecological adaptation and diversification in bacteria such as xanthomonads [45, 70–72]. Here, phylogenetic analyses revealed that the overall topology of *tal* genes differed from the topology of strains, indicating that the evolution of *tal* genes has been impacted by different events of recombination and/or HGT (Fig. 3). It was particularly interesting to observe that whether for the repeats, the N-ter-, or the C-ter-encoding regions, *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains clustered in a single clade (Figs. 3b, c, 4a). This strongly supports a scenario where *tal* genes from these four pathovars share a common ancestor different from *tal* genes found in other *Xanthomonas* species. Ks values for the N-ter- and C-ter-encoding regions suggested that two HGT involving *tal* genes occurred between the ancestors of *X. citri* and *X. phaseoli* species, and between the ancestors of *X. citri* pv. *aurantifolii* and *X. citri* pv. *fuscans* strains and the ancestor of *X. phaseoli* pv. *manihotis* and *X. phaseoli* pv. *phaseoli* strains (events 1 and 2 on Fig. 3a). Absence of orthologues in *X. citri* pathovars other than *X. citri* pv. *aurantifolii* and *X. citri* pv. *fuscans* certainly should be a result of a specific gene loss in these other *X. citri* pathovars. In particular, *X. citri* strains that were isolated on lablab and able to cause CBB on common bean [36, 73] did not possess any *tal* gene orthologous to the *tal* genes found in strains isolated on common bean. This further strengthens that the *tal* genes described here are specific to *Xanthomonas* strains having common bean as a natural host. Also, it showed that these lablab-associated strains were able to cause CBB symptoms on common bean without any contribution of the TAL effectors found in the genuine common bean pathogens, suggesting that these TAL effectors were not fully needed for pathogenicity and would rather contribute to aggressiveness. It would be interesting to obtain the RVD combination of TAL effectors from lablab-associated strains to analyse the contribution of these TAL effectors to the aggressiveness of *Xanthomonas* strains on common bean.

In strains responsible for CBB, *tal* gene diversification led to three distinct clades of *tal* genes. Subsequent to this diversification, we observed *tal* gene duplications and movements associated with IS and Tn3-like transposons, in accordance with previous observations in *X. citri* subsp. *citri* and other xanthomonads [54]. IS and transposons appear to have facilitated recombination, thus contributing to *tal* gene duplications and movements in strains responsible for CBB. It is tempting to speculate that the transfer of *tal23A* homologues that was observed from the plasmid A to the chromosome in

strains CFBP6164 and CFBP6546 (Additional file 14) would allow a stable maintenance of these *tal* genes in the genome, and thus would underline the primary importance of *tal23A* homologues in adaptation of *Xanthomonas* to common bean.

Combination of phylogeny and Ks analyses unveiled events of recombination and HGT that shaped the evolution of *tal* genes in strains responsible for CBB. First, these analyses showed that the different *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages bear different combinations of *tal* genes, suggesting that these lineages evolved in different ways for manipulating the common bean transcriptome. Strains from *X. citri* pv. *fuscans* NF3 lineage contained *tal23A* homologues only. Strains from *X. citri* pv. *fuscans* NF2 lineage also contained *tal23A* homologues, and were the only strains carrying *tal20F* and *tal18G*, suggesting that these genes emerged specifically in the *X. citri* pv. *fuscans* NF2 lineage. Finally, most strains from the *X. citri* pv. *fuscans* *fuscans* and *X. phaseoli* pv. *phaseoli* NF1 lineages, despite being phylogenetically distant, contained *tal23A* homologues and *tal18H* which is highly suggestive of HGT (Fig. 1). Ks analyses confirmed that *tal18H* has been horizontally transferred between these two lineages and that *tal23A* homologues have been transmitted by HGT between the NF1 lineage and the NF2 lineage. Moreover, plasmid conservation among phylogenetically distant lineages showed that HGT involving both *tal18H* and *tal23A* homologues were not restricted to these *tal* genes alone, and highlight the role of plasmids for these two HGT. These results are reminiscent of studies showing that numerous genes including type three effectors have been horizontally transferred between *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages [36, 74]. This also suggests that *tal18H* and *tal23A* homologues have been involved in *Xanthomonas* adaptation to common bean. In particular, *tal23A* homologues were conserved in all strains responsible for CBB, suggesting that they could play a crucial role in *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* pathogenicity.

The study of repeats and RVD evolution shed new light on the importance of *tal* genes for the interaction between *Xanthomonas* strains and common bean. The DisTAL tree revealed that TAL23A and TAL18H presented variations in their repeats for aminoacids other than the RVD, resulting in three variants of TAL23A and five variants of TAL18H (Fig. 4a). As RVD sequence is responsible for DNA-binding specificity, these observations suggested that selection pressure acted against RVD diversification in both TAL23A and TAL18H, highlighting the importance of TAL23A and TAL18H functional maintenance in different *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages. In particular, TAL18H RVD combination was completely conserved in

both NF1 and fuscans lineages, which is in line with TAL18H being involved in *Xanthomonas* adaptation to common bean. To our knowledge, this is the first case of RVD sequence conservation across distant *Xanthomonas* lineages described so far.

FuncTAL analysis and EBE predictions showed that the function of TAL effectors appeared relatively conserved within each *tal* clade (Fig. 4b, Additional file 17). Another interesting observation was that predicted targets of TAL23A homologues and TAL18H both comprised genes encoding an UDP-glycosyltransferase. So far, UDP-glycosyltransferases have not been described as TAL effectors targets [12, 18, 60–67]. UDP-glycosyltransferase transcription has previously been associated with plant resistance to fungi and bacteria [75–77]. Further analyses including mutant construction, phenotyping and transcriptomics are underway to functionally validate the targets of TAL effectors of strains responsible for CBB.

The origins of the strains used for this study gave additional information on *tal* gene evolution in the different *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* lineages (Table 1). Indeed, all four lineages have been found in La Réunion Island in 2000 (<http://catalogue-cfbp.inra.fr/recherche.php>), indicating that sympatry exists among all these lineages in this place, rendering possible further HGT events between strains from these four lineages. Also, in contrast to the NF2 and NF3 lineages, fuscans and NF1 strains have been collected across very distant geographical areas spanning American, European, and African continents. This widespread distribution could have been facilitated by transport of contaminated seeds [78, 79], in frame with agricultural intensification and expansion [46]. Finally, two fuscans strains collected in 1954 and 2009 and two NF1 strains collected in 1966 and 2000 all contained *tal18H* and *tal23A* homologues, indicating that these *tal* genes have been maintained in both lineages for dozens of years. This further highlights the importance of these TAL effectors in these two lineages.

Our finding that plasmidic localisation has played a role in *tal* gene dissemination across phylogenetically distant *Xanthomonas* lineages adapted to common bean shed new light in *tal* genes potential of evolution. Indeed, it appears that two different ways of evolution exist for *tal* genes in *Xanthomonas*. First, it is striking that *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* pathovars rely on numerous (up to 28) chromosomal *tal* genes [29, 68, 80–82]. Also, several cases of convergent targeting of *SWEET* family genes by *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* TAL effectors have been described [12, 58, 61, 83, 84]. Moreover, polymorphism in EBE, or localisation of EBE upstream of *executor* genes led to the resistance of several rice accession

[15, 22–25]. Together, these observations in rice suggested that consecutive events of coevolution occurred between rice and *X. oryzae* pv. *oryzae* / *X. oryzae* pv. *oryzicola* TAL effectors. An interpretation to this would be that presence of numerous *tal* genes localised in the chromosome would be an evidence of long-termed co-evolution between TAL effectors and the host plant genome. In sharp contrast with what was observed in *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*, other pathovars rely solely on few *tal* genes localised on plasmids. For example, *X. euvesicatoria* contains only two *tal* genes, *AvrBs3* and *AvrBs4*, both encoded on plasmids [85], while *X. citri* subsp. *citri* bears up to four *pthA* homologues on two different plasmids [86]. In *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans*, we observed only one to three *tal* genes per strain and the majority was plasmidic. The *tal* genes with chromosomal localisation were the result of a recent duplication in the NF2 lineage, which could be a trace of recent coevolution with common bean. On the other hand, plasmidic TAL effectors were conserved in different lineages and some of them bore the exact same RVD sequence, arguing for a recent horizontal transfer of these *tal* genes. It is tempting to speculate that RVD conservation in horizontally-transferred *tal* genes reflects the fact that the host plant did not have enough time to evolve any resistance against these TAL effectors. An interpretation to this would be that a short number of *tal* genes with plasmidic localisation would be an evidence of short-termed evolution without coevolution between the TAL effector and the host genome. As such, horizontal transfers of *tal* genes can lead to the rapid modification of TAL repertoires, and could participate to the acquisition of novel traits for host adaptation.

Conclusion

Together, our observations favour a model where plasmidic *tal* genes are able to contribute to host adaptation by being horizontally transferred between distant lineages. Such horizontal gene transfer can take place either on a host plant or on another asymptomatic plant where a broader *Xanthomonas* diversity is expected to live [45]. Because we are not completely aware of this diversity, we may still underestimate the evolutionary potential of TAL effector to circumvent TAL-based resistance engineering strategies. To anticipate such potential and elaborate durable resistances against diseases caused by xanthomonads, it is important to make use of third-generation sequencing strategies on a much broader diversity of *Xanthomonas* strains than what has been made so far, and to explore more in-depth the whole TAL repertoire of xanthomonads.

Methods

Bacterial strains, plant material and growing conditions

Xanthomonas strains used and sequenced for this study are listed in Table 1. Strains were grown 48 h at 28 °C on TSA medium (17.0 g/L pancreatic digest of casein; 3.0 g/L enzymatic digest of soya bean; 5.0 g/L NaCl; 2.5 g/L K₂HPO₄; 2.5 g/L glucose; 15 g/L agar), then 24 h at 28 °C on a 1/10 dilution of TSA medium to obtain fresh grown bacteria. When needed, rifamycin was added at a final concentration of 50 mg/L.

The seeds from common bean cultivar Flavert used in this study were kindly provided by Vilmorin (La Méritré, France) and were considered free of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* following the analysis of approximately 100,000 seed per lot with standard tests (International Seed Testing Association 2007). Bean plants were grown in plastic square pots (8 cm) containing compost (Neuhaus humin substrat S NF 44–551) in a controlled climatic room under 95% relative humidity with an alternation of 16 h of light at 23 °C and 8 h of darkness at 20 °C until the first trifoliate leaf was fully expanded. Plants were watered every 2 days and supplemented with N-P-K (18:14:18) at 0.3 g/L once a week.

Pathogenicity assays

The day before inoculation, and for the duration of pathogenicity assays, plants were incubated under 95% relative humidity with an alternation of 16 h of light at 28 °C and 8 h of darkness at 25 °C. Pathogenicity assays were performed with bacterial suspensions calibrated at 1×10^7 CFU/mL in sterile distilled H₂O. Pathogenicity assays were performed by bathing the first trifoliate leave for 30 s into bacterial suspensions. One trifoliate per plant and three plants per strain were inoculated. Symptoms were scored on the 11th day following inoculation according to the following scale: 0 = no symptoms, 1 = 1 to 50 spots per leaf, 2 = 51 to 200 spots per leaf, necrosis and sagging and 3 = more than 200 spots per leaf, necrosis, sagging or leaf death. Average scores were calculated from the values of three plants per strain. Pathogenicity assays were carried out under quarantine at UMR1345 IRHS, Centre INRA, Beaucozéz, France.

Genomic DNA extraction, genome sequencing, assembly and annotation

Bacterial cells were scraped from agar plates, suspended in sterile distilled H₂O and collected by centrifugation. Genomic DNA was extracted with The Wizard® Genomic DNA Purification Kit (Promega) according to manufacturer's recommendations. Whole genomes were sequenced using the PacBio SMRT technology [87] at Icahn School of Medicine at Mount Sinai (New York, NY). One SMRT cell was used per strain to achieve ~100× coverage. De novo assembling was performed

using HGAP assembler [88] version 3.0 (Pacific Biosciences, Menlo Park, CA). Annotation of whole genome assemblies was performed with EuGene-PP version 1.2 automated pipeline [89], using SWISS-PROT as protein and training protein databases (<http://www.uniprot.org/>). The sequences of *tal* genes were manually retrieved using BLASTN search with the N-ter- and C-ter-encoding regions of *Xfuta1* and *Xfuta2* as queries [37]. In case of indels leading to *tal* pseudogenes, sequence verifications were performed by PCR amplifications using specific primers on genomic DNA, followed by Sanger sequencing of PCR products (Additional file 18).

Phylogenetic analyses and detection of recombination

The phylogeny of organisms was performed using CVTtree version 4.2.1 [90] on the whole predicted proteome of strains listed in Table 1, plus 31 proteomes available in public database that belong to *Xanthomonas* pathovars where *tal* genes have previously been described (Additional file 5).

For *tal* genes phylogeny, we used the *tal* genes from strains listed in Table 1, plus 151 *tal* genes encoding unique RVD sequences, available in public databases and sequenced by SMRT sequencing or functionally validated (Additional file 5). The amino acid sequences of the N-ter and C-ter regions of these TAL effectors were aligned with MAFFT version 7 using the L-ins-I strategy [91], then this alignment was used as template for generating a codon-based nucleotide alignment with RevTrans 2.0 [92]. For both alignments, the best model of evolution determined using ModelTest 3.7 in Paup was the TVM + I + G model. Under this model, Maximum Likelihood (ML) phylogenetic trees were generated using PhyML 3.0 version 20,110,919 [93] and bootstraps analyses were done with 1000 iterations. Trees were visualised and manually edited with Mega 7.0.14 [94].

For detecting recombination, two additional trees constructed using nucleotide alignments of the N-ter- or C-ter-encoding regions of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli*, *X. phaseoli* pv. *manihotis* strains and a *tal* gene from *X. translucens* pv. *undulosa* as outgroup were produced using the aforementioned method. For the N-ter-encoding region tree, the model of evolution was K81uf + I and for the C-ter-encoding region tree, the model of evolution was TVM + G (Additional file 19). The Kishino-Hasegawa-Templeton test [47] implemented in the DNAML program from PHYLIP version 3.69 [95] was used to test the congruence of tree topologies. Also, nucleotide alignments from both N-ter- and C-ter-encoding regions of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains were concatenated. Detection of potential recombination events were searched for

on the three nucleotide alignments (separated N-ter- and C-ter-encoding regions, and concatenated) using a set of seven nonparametric detection programs implemented in RDP version 4 beta 69 [48]: RDP [96]; Geneconv [97]; Bootscan [98]; MaxChi [99]; Chimera [100]; SiScan [101] and 3Seq [102]. Default parameter settings were used for each method except as follows: RDP (internal reference sequence); Bootscan (window size = 150, neighbour joining trees, 200 bootstrap replicates, 95% cutoff, Jin and Nei model, variation coefficient = 2). For general settings, the maximum P value for accepting recombination was set at 0.001 after a Bonferroni correction. Only recombination events retrieved by at least three programs with a P value < 0.001 were considered.

Average nucleotide identity calculations

Pairwise comparisons of *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* genome and plasmids nucleotide identities were performed by calculating the Average Nucleotide Identity by using BLASTn with JSpecies using default parameters [103] (Additional file 6, Additional file 8, Additional file 9).

Ks calculations

Ks was estimated on nucleotide alignments of N-ter- and C-ter-encoding regions of *tal* genes from *X. citri* pv. *aurantifolii*, *X. citri* pv. *fuscans*, *X. phaseoli* pv. *manihotis* and *X. phaseoli* pv. *phaseoli* strains, as well as on a concatenated alignment of seven previously described housekeeping genes from *X. oryzae*, *X. citri*, *X. phaseoli*, and *X. euvesicatoria* species [46]. Alignments and recombination analyses were done as described in the “Phylogenetic analyses and detection of recombination” section. After removal of recombinant regions, Ks was determined using DNAsp version 5 with default parameters [104]. On the same alignment, codon usage bias was calculated using DNAsp version 5 and no major difference was observed between the different *Xanthomonas* species (data not shown). Also Tajima’s and Fu and Li’s relative rate test were performed with DNAsp version 5 and no significant divergence from neutrality was retrieved ($P > 0.05$).

Repeats and RVD analyses

Amino acid sequences of every repeats from TAL effectors of *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains were manually extracted and aligned using MAFFT version 7 [91], after RVD exclusion. Then, unique repeats were searched, and repeats found multiple times were counted using Geneious version 9.1.3 [105]. For Fig. 1, amino acid RVD sequence of TAL effectors from *X. citri* pv. *fuscans* and *X. phaseoli* pv.

phaseoli strains were extracted and an alignment was performed with MAFFT version 7 [91].

The evolution of TAL effector repeat regions and RVD sequences were analysed using the QueTAL suite [57]. First, amino acid sequences of repeat regions from TAL effectors of *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli*, *X. phaseoli* pv. *manihotis* strains were extracted and aligned manually according to the QueTAL’s instructions, using TAL effectors from *X. translucens* pv. *undulosa* as outgroup for the DisTAL analysis and no outgroup for the FuncTAL analysis. Then, these alignments were used to construct a distance tree of TAL effector repeat regions using DisTAL version 1.1 and a distance tree of RVD sequences using FuncTAL version 1.1 [57]. DisTAL program considers repeat sequences without RVD whereas FuncTAL program only considers RVD sequences.

With the nine unique RVD combinations, predictions of EBE were performed using TALVEZ version 3.2 with correction matrix starting from the 14th RVD [11] on the common bean promoterome consisting of the 3 kb region upstream of the transcriptional start of each transcript annotated on the genome of common bean genotype G19833 version 2.1 [59] available on Phytozome 12 (<https://phytozome.jgi.doe.gov/>).

Additional files

Additional file 1: Table S1. Correction of sequences from *tal* pseudogenes containing indels in the SMRT sequence. (XLSX 10 kb)

Additional file 2: Table S2. Amino acid sequences of TAL effector from the 17 strains responsible for CBB obtained in this study. (DOCX 19 kb)

Additional file 3: Figure S1. Gel photograph showing the sizes obtained for *tal23A_CFBP4885* (*Xfufal1*) and *tal18H_CFBP4885* (*Xfufal2*) after PCR amplification. Amplification was performed using XapF2 and XapR primers on plasmidic DNA from the CFBP4885 strain (Additional file 18: **Table S11**). Expected PCR product sizes for *tal23A_CFBP4885* and *tal18H_CFBP4885* were 4604 bp and 3074 bp for Sanger sequencing of the entire genome, respectively, and 3686 bp and 3176 bp for SMRT sequencing of the entire genome, respectively. (DOCX 2788 kb)

Additional file 4: Figure S2. RVD frequencies in TAL effectors from the 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains used in this study. (DOCX 20 kb)

Additional file 5: Table S3. References for previously published *tal* genes and whole genome sequences from *Xanthomonas* spp. strains used in this study. (XLSX 19 kb)

Additional file 6: Table S4. Average nucleotide identity and percentage of alignment between whole genome sequences from 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. BLAST calculation of average nucleotide identity (ANIb) and alignment percentages were estimated using Jspecies (Richter and Rossello-Mora, 2009). (XLSX 15 kb)

Additional file 7: Figure S3. Phylogenetic trees of N-ter and C-ter-encoding regions of *tal* genes with sequences published by Aritua et al. (2015). Bootstrap values greater than 50% (100 replicates) are shown and horizontal scale bars represent the number of nucleotide substitutions per site. The *tal* genes from the *X. citri* pv. *fuscans* genetic lineages *fuscans*, NF2 and NF3 are indicated in red, *X. phaseoli* pv. *phaseoli* NF1 lineage in blue, *X. citri* pv. *aurantifolii* in pink and *X. phaseoli* pv. *manihotis* in purple. The sequences published by Aritua et al. (2015) are in bold. Both ML trees were

constructed using *tal* genes from *X. translucens* pv. *undulosa* XT4699 as outgroups. **a** ML tree constructed on a nucleotide alignment of the N-ter-encoding region of *tal* genes. **b** ML tree constructed on a nucleotide alignment of the C-ter-encoding region of *tal* genes. (PPTX 137 kb)

Additional file 8: Figure S4. Circular representation of plasmids A and C. Genomic sequences were compared and converted in a graphical map using CGView (Grant et al., 2012) with strain CFBP4885 as reference. Colours differ according to identity percentage (see legend). Strains order, ANIb and alignment percentages are indicated in the center of each graphical map. Localisation of *tal* genes are highlighted by green zones. Absence of *tal* genes in plasmid A from strains CFBP6164 and CFBP6546R is indicated by a black cross. (PDF 444 kb)

Additional file 9: Table S5. Average nucleotide identity and percentage of alignment between sequences of plasmid A from 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. BLAST calculation of average nucleotide identity (ANIb) and alignment percentages were estimated using Jspecies (Richter and Rossello-Mora, 2009). (XLSX 14 kb)

Additional file 10: Figure S5. Circular representation of the regions shared by phylogenetically distant lineages in plasmids A and C. Genomic sequences were compared and converted in a graphical map using CGView (Grant et al., 2012). Colours differ according to identity percentage (see legend). Regions shared by both strains are surrounded by a black line. Numbers represent genes conserved in both strains (see Additional files 11 and 13 for more details). **a** Plasmids A from strains CFBP6988R (outer circle = reference) and CFBP6546R (inner circle). **b** Plasmids C from strains CFBP6166 (outer circle = reference) and CFBP6982 (inner circle). (PDF 168 kb)

Additional file 11: Table S6. Genes localized in Plasmid A shared by strains CFBP6988R and CFBP6546R. Please refer to Additional file 10a for graphic representation of the data. (XLSX 15 kb)

Additional file 12: Table S7. Average nucleotide identity and percentage of alignment between sequences of plasmid C from 8 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* NF1 strains. BLAST calculation of average nucleotide identity (ANIb) and alignment percentages were estimated using Jspecies (Richter and Rossello-Mora, 2009). (XLSX 11 kb)

Additional file 13: Table S8. Genes localized in Plasmid C shared by strains CFBP6982 and CFBP6166. Please refer to Additional file 10b for graphic representation of the data. (XLSX 15 kb)

Additional file 14: Figure S6. Duplications and movements of *tal* genes associated with IS3 and Tn3-like transposons. Arrows represent *tal* genes (yellow), IS3 sequences (orange) or Tn3-like transposons (dark blue). Arrowheads correspond to truncated IS3 (orange) or Tn3-like transposon (dark blue). Striped areas correspond to conserved regions. Dashed lines represent the borders of duplication events between different regions of a single molecule. **a** Movement of *tal23A* homologues from the plasmid A to the chromosome associated with IS3 sequences. **b** Duplication of *tal20F* and *tal18G* associated with Tn3-like transposons. **c** Duplication of *tal18H_CFBP6164* associated with IS3 sequences in strain CFBP6164. (PPTX 74 kb)

Additional file 15: Table S9. Repeats found within all TAL effectors from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains. Sequences were compared after removal of RVD. Xcf = *X. citri* pv. *fuscans*; Xpp = *X. phaseoli* pv. *phaseoli*; Xca = *X. citri* pv. *aurantifolii*; Xpm = *X. phaseoli* pv. *manihotis*. A coloured line indicates that the repeat is shared between TAL effectors from different groups, one colour being associated with a single repeat and corresponding to the colour of repeats in Additional file 16: **Figure S7**. (XLSX 12 kb)

Additional file 16: Figure S7. Repeats conserved and shared between different TAL effectors from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains. Repeats conserved between TAL effectors from different groups are indicated by the same colour (red, purple, green, blue or yellow). Half-coloured boxes indicate that polymorphism exist among TAL effectors homologues from different strains. Please see Additional file 15: **Table S9** for more detailed information. (PPTX 69 kb)

Additional file 17: Table S10. Top 10 predicted targets for the nine TAL effectors with unique RVD sequences retrieved in 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. Target names correspond to

locus tags from the *P. vulgaris* genome available at <https://phytozome.jgi.doe.gov/>. Similar targets predicted for different TAL effectors are highlighted by colours corresponding to the predicted function of the target (see the legend below). (XLSX 17 kb)

Additional file 18: Table S11. Primers used in this study. (DOCX 19 kb)

Additional file 19: Figure S8. Phylogenetic trees of N-ter- and C-ter-encoding regions of *tal* genes used for the Kishino-Hasegawa-Templeton test. Bootstrap values greater than 50% are shown for 1000 replicates and horizontal scale bars represent the number of nucleotide substitutions per site. **a** ML tree constructed on a nucleotide alignment of the N-ter-encoding region of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains using a *tal* gene from *X. translucens* pv. *undulosa* XT4699 as outgroup. **b** ML tree constructed on a nucleotide alignment of the C-ter-encoding region of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains using a *tal* gene from *X. translucens* pv. *undulosa* XT4699 as outgroup. (PPTX 56 kb)

Abbreviations

ANI: Average nucleotide identity; CBB: Common bacterial blight; CFBP: Collection Française de Bactéries associées aux Plantes; CFU: Colony forming unit; EBE: Effector binding element; HGT: Horizontal gene transfer; IS: Insertion sequence; ML: Maximum likelihood; NF: Non-fuscous; PCR: Polymerase chain reaction; RVD: Repeat variable diresidue; SMRT: Single molecule, real-time; TAL: Transcription activator-like

Acknowledgements

The authors thank Alvaro L. Pérez-Quintero for helping in TAL effector target predictions, Chrystelle Brin and Armelle Darrasse for sharing preliminary data on the diversity of TAL effectors from *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* and Matthieu Barret for critical review of the manuscript. The authors also thank the French Network on Xanthomonads (FNX) (<https://www.reseau-xantho.org/>) and particularly Laurent D. Noël, Boris Szurek, Lionel Gagnevin and Ralf Koebernik for recurrent scientific exchanges on *Xanthomonas* within the framework of the ANR project CROPTAL led by Laurent D. Noël. PacBio sequencing was performed at Kahn School of Medicine at Mount Sinai (NY, USA) under the supervision of Robert Sebra and preliminary quality control of DNA was performed at the ANAN platform from the SFR 4207 QUASAV (Angers, France). We thank the CIRM-CFBP (Angers, France) for strain preservation and supply and the members of the EmerSys and INEM teams (IRHS, Angers) for their participation in pathogenicity assays.

Funding

This study was supported by the French National Research Agency (ANR-14-CE19-0002-04 CROPTAL project to NWGC. MR is funded by a PhD grant (X-HOT project) from Angers-Loire Metropole, France. The funding bodies did not participate in any of the study design, data collection and analysis, or writing the manuscript.

Availability of data and materials

The datasets (genome sequences) generated through various projects and used in this study have been deposited in GenBank under accession numbers listed in Table 1.

Authors' contributions

MR, MAJ and NWGC designed the study and wrote the manuscript. MR, MB, SB and NWGC contributed to data analyses. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 3 May 2017 Accepted: 24 August 2017
Published online: 30 August 2017

References

- Bonas U, Stall RE, Staskawicz B. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet*. 1989;218:127–36.
- de d LO, Wolf C, Dietze J, Elsaesser J, Morbitzer R, Lahaye T. Programmable DNA-binding proteins from Burkholderia provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Res*. 2014;42:7436–49.
- Heuer H, Yin YN, Xue QY, Smalla K, Guo JH. Repeat domain diversity of *avrBs3*-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Appl Environ Microbiol*. 2007;73:4379–84.
- Boch J, Bonas U. *Xanthomonas AvrBs3* family-type III effectors: discovery and function. *Annu Rev Phytopathol*. 2010;48:419–36.
- Mak ANS, Bradley P, Bogdanove AJ, Stoddard BL. TAL effectors: Function, structure, engineering and applications. *Curr Opin Struct Biol*. 2013;23:93–9.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* (80-). 2009;326:1509–12.
- Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science* (80-). 2009;326:1501.
- Yang J, Zhang Y, Yuan P, Zhou Y, Cai C, Ren Q, et al. Complete decoding of TAL effectors for DNA recognition. *Cell Res*. 2014;24:628–31.
- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, et al. TAL Effector-Nucleotide Targeter (TALEN) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res*. 2012;40:117–22.
- Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J. Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Comput Biol*. 2013;9.
- Pérez-Quintero AL, Rodríguez-R LM, Dereeper A, López C, Koebnik R, Szurek B, et al. An Improved Method for TAL Effectors DNA-Binding Sites Prediction Reveals Functional Convergence in TAL Repertoires of *Xanthomonas oryzae* Strains. *PLoS One*. 2013;8:e68464.
- Yang B, Sugio A, White FF. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc Natl Acad Sci U S A*. 2006;103:10503–8.
- Kay S, Hahn S, Marois E, Hause G, Bonas U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* (80-). 2007;318:648–51.
- Chen L-Q, Hou B-H, Lalonde S, Takanaga H, Hartung ML, Qu X-Q, et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*. 2010;468:527–32.
- Römer P, Recht S, Strauss T, Elsaesser J, Schornack S, Boch J, et al. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol*. 2010;187:1048–57.
- Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, et al. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature*. 2005;435:1122–5.
- Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T. Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science* (80-). 2007;318:645–8.
- Boch J, Bonas U, Lahaye T. TAL effectors - pathogen strategies and plant resistance engineering. *New Phytol*. 2014;204:823–32.
- Schornack S, Moscou MJ, Ward ER, Horvath DM. Engineering Plant Disease Resistance Based on TAL Effectors. *Annu Rev Phytopathol*. 2013; 51:383–406.
- Blanvillain-Baufumé S, Reschke M, Solé M, Auguy F, Doucoure H, Szurek B, et al. Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for SWEET14-inducing TAL effectors. *Plant Biotechnol J*. 2017;15:306–17.
- Li C, Wei J, Lin Y, Chen H. Gene silencing using the recessive rice bacterial blight resistance gene *xa13* as a new paradigm in plant breeding. *Plant Cell Rep*. 2012;31:851–62.
- Hummel AW, Wilkins KE, Wang L, Cernadas RA, Bogdanove AJ. A transcription activator-like effector from *Xanthomonas oryzae* pv. *oryzicola* elicits dose-dependent resistance in rice. *Mol. Plant Pathol*. 2016;18:55–6.
- Hutin M, Césari S, Chalvon V, Michel C, Tran TT, Boch J, et al. Ectopic activation of the rice NLR heteropair RGA4/RGA5 confers resistance to bacterial blight and bacterial leaf streak diseases. *Plant J*. 2016;88:43–55.
- Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, et al. *XA23* is an executor protein and confers broad-spectrum disease resistance in rice. *Mol Plant*. 2015;8:290–302.
- Zhang J, Yin Z, White F. TAL effectors and the executor R genes. *Front Plant Sci*. 2015;6:641.
- Hummel AW, Doyle EL, Bogdanove AJ. Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytol*. 2012;195:883–93.
- Zeng X, Tian D, Gu K, Zhou Z, Yang X, Luo Y, et al. Genetic engineering of the *Xa10* promoter for broad-spectrum and durable resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant Biotechnol J*. 2015;13:993–1001.
- Gardiner DM, Upadhyaya NM, Stiller J, Ellis JG, Dodds PN, Kazan K, et al. Genomic analysis of *Xanthomonas translucens* pathogenic on wheat and barley reveals cross-kingdom gene transfer events and diverse protein delivery systems. *PLoS One*. 2014;9.
- Wilkins KE, Booher NJ, Wang L, Bogdanove AJ. TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. *Front Plant Sci*. 2015;6:1–15.
- Peng Z, Hu Y, Xie J, Potnis N, Akhunova A, Jones J, et al. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics*. 2016;17:21.
- Alavi SM, Sanjari S, Durand F, Brin C, Manceau C, Poussier S. 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. *Appl. Environ. Microbiol*. 2008;74:3295–301.
- EFSA Panel on Plant Health. Scientific Opinion on the pest categorisation of *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *EFSA J*. 2014;12:3856.
- Constantin EC, Cleenwerck I, Maes M, Baeyens S, Van Malderghem C, De Vos P, et al. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathol*. 2016;65:792–806.
- Vauterin L, Hoste B, Kersters K, Swings J. Reclassification of *Xanthomonas*. *Int J Syst Bacteriol*. 1995;45:472–89.
- Rademaker J, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J, et al. A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology*. 2005;95:1098–111.
- Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Front. Microbiol*. 2015;6:1–18.
- Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, et al. Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. *BMC Genomics*. 2013;14:761.
- Indiana A, Briand M, Arlat M, Gagnevin L, Koebnik R, Noël L, et al. Draft genome sequence of the flagellated *Xanthomonas fuscans* subsp. *fuscans* strain CFBP 4884. *Genome Announc*. 2014;2:e00966–14.
- Jalan N, Kumar D, Yu F, Jones JB, Graham JH, Wang N. Complete Genome Sequence of *Xanthomonas citri* subsp. *citri* Strain Aw12879, a Restricted-Host-Range Citrus Canker-Causing Bacterium. *Genome Announc*. 2013;1: e00235–13.
- Bolot S, Guy E, Carrère S, Barbe V, Arlat M, Noël LD. *Xanthomonas campestris* pv. *campestris* Strain Xca5. *Genome Announc*. 2013;1:e00032–12.
- Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, et al. Comparative Genomics of Pathogenic and Nonpathogenic Strains of *Xanthomonas arboricola* Unveils Molecular and Evolutionary Events Linked to Pathoadaptation. *Front Plant Sci*. 2015;6:1126.
- Gagnevin L, Bolot S, Gordon JL, Pruvost O, Vernière C, Robène I, et al. Draft Genome Sequence of *Xanthomonas axonopodis* pv. *allii* Strain CFBP 6369. *Genome Announc*. 2014;2:e00727–14.
- Quail M, Smith ME, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*. 2012;13:341.
- de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinski J, et al. Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol*. 2013;199:773–86.

45. Jacques M-A, Ariat M, Boulanger A, Boureau T, Carrère S, Cesbron S, et al. Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas*: French Network on *Xanthomonads* (FNX). *Annu Rev Phytopathol.* 2016;54:163–87.
46. Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, et al. Evolutionary History of the Plant Pathogenic Bacterium *Xanthomonas axonopodis*. *PLoS One.* 2013;8:e58474.
47. Kishino H, Hasegawa M. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoids. *J Mol Evol.* 1989;29:170–9.
48. Martin DP, Williamson C, Posada D. RDP2: recombination detection and analysis from sequence alignments. *Bioinform Appl Note.* 2005;21:260–2610.
49. Ochman H, Elwyn S, Moran NA. Calibrating bacterial evolution. *Proc Natl Acad Sci U S A.* 1999;96:12638–43.
50. Kuo C-H, Ochman H. Inferring clocks when lacking rocks: the variable rates of molecular evolution in bacteria. *Biol Direct.* 2009;4:35.
51. Schuenzel EL, Scally M, Stouthamer R, Nunney L. A Multigene Phylogenetic Study of Clonal Diversity and Divergence in North American Strains of the Plant Pathogen *Xylella fastidiosa* A Multigene Phylogenetic Study of Clonal Diversity and Divergence in North American Strains of the Plant Pathogen *Xylella*. *Appl Environ Microbiol.* 2005;71:3832–9.
52. Moreira LM, Almeida NF, Potnis N, Digiampietri L a, Adi SS, Bortolossi JC, et al. Novel insights into the genomic basis of citrus canker based on the aureantifolii. *BMC Genomics.* 2010;11:238.
53. Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, et al. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc Natl Acad Sci U S A.* 2012;109:E1972–9.
54. Ferreira RM, de Oliveira ACP, Moreira LM, Belasque JJ, Goubeyre E, Siguier P, et al. A TALE of Transposition: Tn3-like transposons play a major role in the spread of pathogenicity determinants of *xanthomonas citri* and other *xanthomonads*. *MBio.* 2015;6:1–12.
55. Vera Cruz CM, Bai J, Ona I, Leung H, Nelson RJ, Mew TW, et al. Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc Natl Acad Sci U S A.* 2000;97:13500–5.
56. Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli SV, Patil PB, et al. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *J. Bacteriol.* 2011;193:5450–64.
57. Pérez-Quintero AL, Lamy L, Gordon JL, Escalon A, Cunnac S, Szurek B, et al. QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Front Plant Sci.* 2015;6:545.
58. Richter A, Streubel J, Blücher C, Szurek B, Reschke M, Grau J, et al. A TAL effector repeat architecture for frameshift binding. *Nat Commun.* 2014;5:3447.
59. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet.* 2014;46:707–13.
60. Sugio A, Yang B, Zhu T, White FF. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OstF1Agamma1* and *OstFX1* during bacterial blight of rice. *Proc Natl Acad Sci U S A.* 2007;104:10720–5.
61. Streubel J, Pesce C, Hutin M, Koebnik R, Boch J, Szurek B. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* 2013;200:808–19.
62. Cernadas RA, Doyle EL, Niño-Liu DO, Wilkins KE, Bancroft T, Wang L, et al. Code-Assisted Discovery of TAL Effector Targets in Bacterial Leaf Streak of Rice Reveals Contrast with Bacterial Blight and a Novel Susceptibility Gene. *PLoS Pathog.* 2014;10:e1003972.
63. Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, et al. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc Natl Acad Sci U S A.* 2014;111:E521–9.
64. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, et al. *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. *Mol Plant-Microbe Interact.* 2014;27:1186–98.
65. Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom JS, et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* 2015;82:632–43.
66. Schwartz AR, Morbitzer R, Lahaye T, Staskawicz BJ. TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. *Proc Natl Acad Sci U S A.* 2017;114: E897–903.
67. Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, et al. TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. *Nat Commun.* 2017;8:1–14.
68. Huguet-Tapia JC, Peng Z, Yang B, Yin Z, Liu S, White FF. Complete Genome Sequence of the African Strain AXO1947 of *Xanthomonas oryzae* pv. *oryzae*. *Genome Announc.* 2016;4:10–1.
69. Booher NJ, Carpenter SCD, Sebra RP, Wang L, Salzberg SL, Leach JE, et al. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Microb Genomics.* 2015;1:1–22.
70. Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev.* 2011;35:957–76.
71. Darmon E, Leach DRF. Bacterial genome instability. *Microbiol Mol Biol Rev.* 2014;78:1–39.
72. El Yacoubi B, Brunings AM, Yuan Q, Shankar S, Gabriel DW. In planta horizontal transfer of a major pathogenicity effector gene. *Appl Environ Microbiol.* 2007;73:1612–21.
73. Sabet KA. Studies in the bacterial diseases of Sudan crops: III. On the occurrence, host range and taxonomy of the bacteria causing leaf blight diseases of certain leguminous plants. *Ann Appl Biol.* 1959;47:318–31.
74. Hajri A, Brin C, Hunault G, Lardeux F, Lemaire C, Manceau C, et al. A «repertoire for repertoire» hypothesis: Repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. *PLoS One.* 2009;4:e6632.
75. Le Roy J, Huss B, Creach A, Hawkins S, Neutelings G. Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants. *Front Plant Sci.* 2016;7:735.
76. Langlois-Meurinne M, Gachon CMM, Saindrenan P. Pathogen-responsive expression of glycosyltransferase genes UGT73B3 and UGT73B5 is necessary for resistance to *Pseudomonas syringae* pv. *tomato* in *Arabidopsis*. *Plant Physiol.* 2005;139:1890–901.
77. Simon C, Langlois-Meurinne M, Didierlaurent L, Chaouch S, Bellvert F, Massoud K, et al. The secondary metabolism glycosyltransferases UGT73B3 and UGT73B5 are components of redox status in resistance of *Arabidopsis* to *Pseudomonas syringae* pv. *tomato*. *Plant Cell Environ.* 2014;37:1114–29.
78. Darrasse A, Darsonval A, Boureau T, Brisset MN, Durand K, Jacques MA. Transmission of plant-pathogenic bacteria by nonhost seeds without induction of an associated defense reaction at emergence. *Appl Environ Microbiol.* 2010;76:6787–96.
79. Akhavan A, Bahar M, Askarian H, Lak MR, Nazemi A, Zamani Z. Bean common bacterial blight: pathogen epiphytic life and effect of irrigation practices. *Spring.* 2013;2:41.
80. Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, et al. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* 2005;33:577–86.
81. Ochiai H, Inoue Y, Takeya M, Sasaki A, Kaku H. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *Japan Agric Res Q.* 2005;39:275–87.
82. Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, et al. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. *BMC Genomics.* 2008;9:204.
83. Antony G, Zhou J, Huang S, Li T, Liu B, White F, et al. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell.* 2010;22:3864–76.
84. Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, et al. Colonization of rice leaf blades by an african strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. *Mol Plant-Microbe Interact.* 2011;24:1102–13.
85. Barak JD, Vancheva T, Lefeuvre P, Jones JB, Timilsina S, Minsavage G V, et al. Whole-Genome Sequences of *Xanthomonas euvesicatoria* Strains Clarify Taxonomy and Reveal a Stepwise Erosion of Type 3 Effectors. *Front Plant Sci.* 2016;7:1805.
86. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature.* 2002;417:459–63.
87. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Single Polymerase Molecules. *Science (80-).* 2009;323:133–8.

88. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods*. 2013;10:563–9.
89. Sallet E, Gouzy J, Schiex T. EuGene-PP: A next-generation automated annotation pipeline for prokaryotic genomes. *Bioinformatics*. 2014;30:2659–61.
90. Qi J, Luo H, Hao B. CVTree: A phylogenetic tree reconstruction tool based on whole genomes. *Nucleic Acids Res*. 2004;32:45–7.
91. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.
92. Wernersson R, Pedersen AG. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res*. 2003;31:3537–9.
93. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol*. 2010;59:307–21.
94. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:msw054.
95. Felsenstein J. PHYLIP - Phylogeny inference package - v3.2. *Cladistics*. 1989. p. 164–6.
96. Martin D, Rybicki E. RDP: detection of recombination amongst aligned sequences. *Bioinformatics*. 2000;16:562–3.
97. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 1999;265:218–25.
98. Martin DP, Posada D, Crandall KA, Williamson C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retrovir*. 2005;21:98–102.
99. Maynard-Smith J. Analyzing the mosaic structure of genes. *J Mol Evol*. 1992;34:126–9.
100. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc Natl Acad Sci*. 2001;98:13757–62.
101. Gibbs MJ, Armstrong JS, Gibbs A. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*. 2000;16:573–82.
102. Boni MF, Posada D, Feldman MW. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*. 2007;176:1035–47.
103. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A*. 2009;106:19126–31.
104. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25:1451–2.
105. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Additional file 1: Table S1. Correction of sequences from *tal* pseudogenes containing indels in the SMRT sequence.

Strain	<i>tal</i> gene	Size after SMRT sequencing (bp)	Number and localisation of sequencing errors	Sequence with error	Coordinates on <i>tal</i> gene	Verification	Sequence after correction (corrected bases in red)	Size of <i>tal</i> gene after correction (bp)
4885	<i>tal23A_CFBP4885</i>	4016	1 in C-ter	CATTTCCTCCCCCTCAGC	3805-3822	Alignment + PCR/Sanger	CATTTCCTCCCCCTCAGC	4017
	<i>tal18H_CFBP4885</i>	3506	1 in C-ter	ACACTCTCCCCAGCCTCG	3010-3027	Alignment + PCR/Sanger	ACACTCTCCCCAGCCTCG	3507
6546R	<i>tal19I_CFBP6546R</i>	3608	1 in C-ter	CATTTCCTCCCCCTAAGC	3397-3414	Alignment + PCR/Sanger	CATTTCCTCCCCCTAAGC	3609
	<i>tal18H_CFBP6546R</i>	3504	1 in N-ter and 2 in C-ter	TCGTGGGGGTCTCCGCC ACACTCTCCCCAGCCTCG AATCGGGGCGGCCTTGC	93-110 3010-3027 3336-3353	Alignment + PCR/Sanger	TCGTGGGGGTCTCCGCC ACACTCTCCCCAGCCTCG AATCGGGGCGGCCTTGC	3507
6988R	<i>tal18G_CFBP6988R</i>	3500	1 in N-ter and 2 in C-ter	TCGTGGGGGGCTCCGCC	93-110	Alignment + PCR/Sanger	TCGTGGGGGGCTCCGCC	3501
6994R	<i>tal23A_CFBP6994R</i>	4016	1 in C-ter	CATTTCCTCCCCCTCAGC	3805-3822	Alignment + PCR/Sanger	CATTTCCTCCCCCTCAGC	4017
7767R	<i>tal23A_CFBP7767R</i>	4015	2 in C-ter	ACGCTCTCCCCAGCCCCG CATTTCCTCCCCCTCAGC	3523-3540 3805-3822	Alignment + PCR/Sanger	ACGCTCTCCCCAGCCCCG CATTTCCTCCCCCTCAGC	4017
	<i>tal18H_CFBP7767R</i>	3503	2 in N-ter and 2 in C-ter	CAGCCCCTGGATGGCTT AGGAGAGGGGATGAGGT ACACTCTCCCCAGCCTCG GCTTCTCCCCAGCTGGGG	117-134 309-326 3010-3027 3297-3314	PCR/Sanger	CAGCCCCTGGATGGCTT AGGAGAGGGGATGAGGT ACACTCTCCCCAGCCTCG GCTTCTCCCCAGCTGGGG	3507

Additional file 2: Table S2. Amino acid sequences of TAL effectors from the 17 strains responsible for CBB obtained in this study.

TAL23A_CFBP4885
MDPIRPRTPSPAHELLAGFPQPDVQPPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMS AFGAPHTEAASGEGDEVQSGLRADDPQATVQVAVTAARPPRAKVPVPRRAAHTSDASPAQVVDLCTLGYSQQQQEKIKLKARSTVAQHHEA LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTGQLLKIAKRGGVTAVEA VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRL LLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV LCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQ LETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTP IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPAPQRWHRILQASGMKRAE PSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGG LDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEMAWLMELFPQ
TAL18H_CFBP4885
MDPIRPRAPSPAHELLAEFPQPDVQPPQPTADRGGSPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFRDILLRQFDPSSLDTSLFDSMS AFGAPHTEAASGEGDEVQSGLRADDPQATVQVAVTAARPPRAKVPVPRRAAHTSDASPAQVVDLCTLGYSQQQQEKIKLKARSTVAQHHEA LIGHGFTRAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTGQLLKIAKRGGVTAVEA VHAWRNALTGVPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA NHDGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIANHDGGKQALETVQRLLPVLCEQHGL LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNIGGKQALETVQRL LLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTP CEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRL EAVKKGLPHALTILKRTSRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMAQFAMSRHGLLQLFRRVGVTELEACNGTLPPASQRW HRILQASGVRTATPSRASAQTPDQASLDAFADALERGLDALSPIDQAVQAQASSRRKRSRSESSVTSSAHYAVEVPVREQHAALDSLPPSW GAKRPRTRIGGGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMELFPQ
TAL22B_CFBP6165
MDPIRPRTPSPAHELLAGFPQPDVQPPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMS AFGAPHTEAASGEGDEVQSGLRADDPQATVQVAVTAARPPRAKVPVPRRAAHTSDASPAQVVDLCTLGYSQQQQEKIKLKARSTVAQHHEA LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTGQLLKIAKRGGVTAVEA VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRL LLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV LCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQ LETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTP PAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPAPQRWHRILQASGMKRAE PSGASAQTPDQASLHAFADALERELDAPSPIDRA GQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGGLDPGTPMDADLAASSTMIWEQDAAPFVGEEDDF AFNEEMAWLMELFPQ

TAL23A_CFBP6166
MDPIRPRTPSPAHELLAGPQPDRVQPQTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMS AFGAPHTEAASGEDEVQSGLRAADDPQATVQVAVTAARPPRAKVPFRRRAHTSDASPAQVVDLCTLGYSQQQEQKIKLKARSTVAQHHEA LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTQQLLAKRGVGTAVEA VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA SNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRL LLPVLCEQHGLTPEQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQ ALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLT IKRNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGGLQLFRRVGVTELEARSGLTPPAQRWHRILQASGMKRAE PSGASQTPDQASLHAFADALERELDAPSPIDRAGQALASSRRKRSSESVTGSFAQQAVEVVRVPEQHEGLHFPPLSGVGRPRTRIGGLG PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFPQ
TAL18H_CFBP6166
MDPIRPRAPSPAHELLAEPPQDRVQPQTADRGGSPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMS AFGAPHTEAASGEDEVQSGLRAADDPQATVQVAVTAARPPRAKVPFRRRAHTSDASPAQVVDLCTLGYSQQQEQKIKLKARPIVAQHHEA LIGHGFTRAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTQQLLAKRGVGTAVEA VHAWRNALTGVPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIA NHDGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPDQVVAIANHDGGKQALETVQRLLPVLCKDHGL LTPDQVVAIASNIGGKQALETVQRLLPVLCKDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNIGGKQALETVQRL LLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQ VVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVL EAVKKGPHALTLIKRTRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMAQFAMSRHGLQLFRRVGVTELEACNGTLPPASQRW HRILQASGVRTATPSRASAQTPDQASLDAFADALERGLDALSPIDQAVQAQASSRRKRSSESVTRSSAHYAVEVPVREQHAALDSLPPSW GAKRPRTRIGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMELFPQ

TAL23A_CFBP6167
MDPIRPTPSPAHELLAGPQPDVQPOPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSLLDTSLFDMS AFGAPHTEAASGEDEVQSGLRAADDPQATVQVAVTAARPPRAKPVPRRAAHTSDASPAQVDLCTLGYSQQQEKIKLKARSTVAQHHEA LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTGQLLAKRGGVTAVEA VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA SNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQH LTPDQGVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQ LLPVLCEQHGLTPEQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQ ALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA IASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTKRSGMRHGLLQFRRVGVTELEARSGLTPAPQRRWHRIQLQASGMKRAE PSGASQTPDQASLHAFADALERELDAPSPIDRAGQALASSRKRSSSVTGSFAQQAVEVVRPEQHEGLHFPPLSWGVKRPRTRIGGGL PDGTPMDADLAASSTMIWEQDAAAPFVGEEDDFPAFNEEMAWLMELFPQ
TAL18H_CFBP6167
MDPIRPRAPSPAHELLAEPPQPDVQPOPTADRGSSPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSLLDTSLFDMS AFGAPHTEAASGEDEVQSGLRAADDPQATVQVAVTAARPPRAKPVPRRAAHTSDASPAQVDLCTLGYSQQQEKIKLKARPIVAQHHEA LIGHGFTRAHVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLDGTGQLLAKRGGVTAVEA VHAWRNALTGVPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIA NHDGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPDQVVAIANHDGGKQALETVQRLLPVLCEQHGL LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQ LLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA VVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVL CEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA EAVKGLPHALTLIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTKRSGMRHGLLQFRRVGVTELEACNGTLPASQRW HRILQASGVRTATPSRASAQTPDQASLDAFADALERGLDALSPIDQAVQAQASSRKRSSSVTRSSAHYAVEVVRPEQHAALDSLPPSW GAKRPRTRIGGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMELFPQ

[illegible][illegible]

TAL23A_CFBP6988R

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDDTSLFDSMS
 AFGAPHTAASGEGDEVQSGLRADDDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASPAQVVDLCTLGYSQQQEQEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPLQLDGTGQLLKIARGGVTAVEA
 VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAI
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTREQVVAIASNGGGKQALETVQ
 LLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQ
 ALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVV
 IASNGGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLT
 IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPAPQRWHRIQLASGMKRAE
 PSGASAQTPDQASLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGL
 PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFHQ

TAL20F_CFBP6988R

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAGSFSDLLRQFDPSSLDDTSLFDSMP
 AFGAHTAAPGEADEVQSGLRVDDPHPTVQVAVTAARPPRAKPAARRRAAHPPDASPAQVVDLCTLGYSQQQEQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIARGGVTAVEAV
 HAWRNALTGAPLNLTDPQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQ
 ALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQV
 TVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVV
 SLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLPDPGTPMDADLA
 ASSTMIWEQDAAPFAGAADDFTAFFNEEEMAWLMELFHQ

TAL18G_CFBP6988R

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAGSFSDLLRQFDPSSLDDTSLFDSMP
 AFGAHTAAPGEADEVQSGLRVDDPHPTVQVAVTAARPPRAKPAARRRAAHPPDASPAQVVDLCTLGYSQQQEQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIARGGVTAVEAV
 HAWRNALTGAPLNLTDPQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASNIGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQ
 ALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQV
 AVKGLPQAPTILIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSRPAQAFDDAMTQFEMSRHGLLQLFRRAGVTELEARSALPPASQRWH
 RILQASGMKRAEPSRASAQTPNQASLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWG
 VKRPRTRIGGGLPDPGTPMDADLAASSTMIWEQDAAPFAGAADDFTAFFNEEEMAWLMELFHQ

[illegible]

TAL23A_CFBP6990

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAQSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTAASGEGDEVQSGLRADDDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASPAQVVDLCTLGYSQQQEQEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPLQLDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAI
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTREQVVAIASNGGGKQALETVQ
 LLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQ
 ALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVV
 IASNGGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLT
 IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPAPQRWHRIQLASGMKRAE
 PSGASAQTPDQASLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGL
 PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFHQ

TAL20F_CFBP6990

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAQSFSDDLRLQFDPSSLDTSLFDSMP
 AFGAHTAAPGEADEVQSGLRVDDPHPTVQVAVTAARPPRAKPAARRRAAHPDASPAQVVDLCTLGYSQQQEQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIAKRGGVTAVEAV
 HAWRNALTGAPLNLTDPQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQ
 ALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 TVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVV
 IASNIGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 SLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLPDPGTPMDADLA
 ASSTMIWEQDAAPFAGAADDFTAFNEEEMAWLMELFHQ

TAL18G_CFBP6990

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAQSFSDDLRLQFDPSSLDTSLFDSMP
 AFGAHTAAPGEADEVQSGLRVDDPHPTVQVAVTAARPPRAKPAARRRAAHPDASPAQVVDLCTLGYSQQQEQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIAKRGGVTAVEAV
 HAWRNALTGAPLNLTDPQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASNIGGKQALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQ
 ALETVHRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVHRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 AVKKGLPQAPTILIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSRPAQAFDDAMTQFEMSRHGLLQLFRRAGVTELEARSALPPASQRWH
 RILQASGMKRAEPSRASAQTPNQASLHAFADALERELEDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWG
 VKRPRTRIGGGLPDPGTPMDADLAASSTMIWEQDAAPFAGAADDFTAFNEEEMAWLMELFHQ

TAL22C_CFBP6991

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGSLRAADDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHAPALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPLQLDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGAPLNLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIAS
 SNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQH
 GLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETLQR
 LLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPD
 QVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPAQVVA
 IASNGGGKQALESIFAQLSRPDQALAAALTNDHLVALACLGGRPALEAVKKGPHAPTLLIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSH
 PAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPAPQRWHRILOASGMKRAEPSGASQTPDQASLHAFADALERELDAPSPIDRA
 GQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTTRIGGGLPDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFP
 AFNEEEMAWLMELFPQ

TAL20F_CFBP6991

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMP
 AFGAHTAAPGEADEVQSGSLRAVDDPHPTVQVAVTAARPPRAKPAAPRRRAAHPDASPAAQVDLCTLGYSQQQQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIAKRGGVTAVEAV
 HAWRNALTGAPLNLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGRQALETVHRLLPVLCQAHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVHRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLC
 QAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALE
 TVHRLLPVLCQAHGLTPDQVVAIASNGGGKQALESIFAQLSRPDQALAAALTNDHLVALACLGGRPALEAVKKGLPQAPTLLIKRTNRRLPERT
 SHRVADHAQVARVLGFFQCHSRPAQAFDDAMTQFEMSRHGLLQLFRRAGVTELEARSGLPPASQRWHRILOASGMKRAEPSRASAQTPDQA
 SLHAFADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTTRIGGGLPDPGTPMDADLA
 ASSTMIWEQDAAPFAGAADDFTAFNEEEMAWLMELFHQ

TAL18G_CFBP6991

MDPIRPRTPSPAHELLAEFPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPTPLPAFSAGSFSDLLRQFDPSSLDTSLFDSMP
 AFGAHTAAPGEADEVQSGSLRAVDDPHPTVQVAVTAARPPRAKPAAPRRRAAHPDASPAAQVDLCTLGYSQQQQEKIKPKVRSTVAQHHEAL
 VGHGFTHAHIVALSQHPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGLPLQLDGTGQLLKIAKRGGVTAVEAV
 HAWRNALTGAPLNLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASH
 DGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGTQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLT
 PAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRL
 PVLCAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGRQALETVHRLLPVLCQAHGLTPDQVVAIASNIGGK
 QALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVHRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQV
 VAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLC
 QAHGLTPDQVVAIASNIGGKQALETVHRLLPVLCQAHGLTPDQVVAIASNGGGKQALESIFAQLSRPDQALAAALTNDHLVALACLGGRPALE
 AVKKGLPQAPTLLIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSRPAQAFDDAMTQFEMSRHGLLQLFRRAGVTELEARSGLPPASQRWH
 RILOASGMKRAEPSRASAQTPNQASLHAFADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWG
 VKRPRTTRIGGGLPDPGTPMDADLAASSTMIWEQDAAPFAGAADDFTAFNEEEMAWLMELFHQ

TAL23A_CFBP6992

MDPIRPRTPSPAHELLAGFPQPDVQPPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRADDDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASAGQVDLCTLGYSQQQKEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPAALGTAVVYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPLQLDGTGQLLKIAKRGVTAVEA
 VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRL
 LLPVLCEQHGLTPEQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA
 IASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLT
 IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRVGVTELEARSGLTPPAPQRWHRILQASGMKRAE
 PSGASAQTPDQASLHAFADALERELEDPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGL
 PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFPQ

TAL23A_CFBP6994R

MDPIRPRTPSPAHELLAGFPQPDVQPPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRADDDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASAGQVDLCTLGYSQQQKEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPAALGTAVVYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPLQLDGTGQLLKIAKRGVTAVEA
 VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRL
 LLPVLCEQHGLTPEQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA
 IASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLT
 IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRVGVTELEARSGLTPPAPQRWHRILQASGMKRAE
 PSGASAQTPDQASLHAFADALERELEDPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGL
 PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFPQ

TAL23A_CFBP6996R

MDPIRPRTPSPAHELLAGFPQPDVQPPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRADDDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASAGQVDLCTLGYSQQQKEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPAALGTAVVYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPLQLDGTGQLLKIAKRGVTAVEA
 VHAWRNALTGAPLNLTDPQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQRL
 LLPVLCEQHGLTPEQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPV
 LCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCEQHGLTPEQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPEQVVA
 IASNNGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLT
 IKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRVGVTELEARSGLTPPAPQRWHRILQASGMKRAE
 PSGASAQTPDQASLHAFADALERELEDPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGL
 PDPGTPMDADLAASSTMIWEQDAAPFVGEEDDFPAFNEEEMAWLMELFPQ

TAL19E_CFBP412

MDPIRPTPSPAHELLAGPQPDRVQPQTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPVPRRRAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQHHEA
LIGHGFTRAQIVALSQHPAALGTVAVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLQDGTQQLLKIARGGVTAVEA
VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
SNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQH
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQR
LLPVLCEQHGLTPEQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG
GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTRE
QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGKGKQALETVQRLLP
LCQAHGLTPEQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTPEQVVAIASHDGKGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQ
LESIFAQLSRPDQALAAALTNHLVALACLGGRPALAEAVKKGLPHAPTLIKRTNRRLPERTSHRVADHAQVAVRLGFFQCCHSHPAQAFDEAMT
QFGMSRHGLLQLFRVRGVTELEARSGLTPAPQQRWHIRLQASGMKRAEPSGASAGTPDQASLHAFADALERELDPSPIDRAGQALASSRK
RSRSESSVTGSFAQQAVEVRVPEQHEGLHFPLPSWGVKRPTRIGGGLPDGTPMDADLAASSTMIWEQDAAPFVGEADDFPAFNEEMAWL
MELFPQ

TAL21D_CFBP6164
MDPIRPTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDDTSLFDSMS AFGAPHTEAASGEGDEVQSSGLRAADDPQATVQVAVTAARPPRAKPVPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQHHEA LIGHGFTRAQIVALSHPAALGTVAIVKYQVMI AALPEATHEDIVGVGKQWSGARALEALLTVSGGELRGPLLQDLTGQLLKIAKRGVTVAVEA VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA SNIGGKQALETVQRLLPVLCEQHGLTPEQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNTGGKQALETVQRLLPVLCEQH LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQR LLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPD QVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNTGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPV LCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQ LETQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVA ALACGGRPALAEAVKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTLEAR SGTLPPAPQRWHRIIQASGMKRAEPSGASQTPDQASLHAFADALERDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRPEQ HEGLHFPPLSWGVKRPRTTRIGGGLDPGTPMDADLAASSTMIWEQDAAPFVGEADDFPAFNEEEMAWLMELFPQ

TAL18H_CFBP6164
MDPIRPRAPSPAHELLAEPPQDRVQPQPTADRGGSPPAGSLDGLPARRTMSRTRLPSPPAPLPAFSAGSFRDLLRQFDPSLLDTSLFDMS AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQEQEKIKLKARPIVAQHHEA LIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPLLQDLTGQLLKIARGGVTAVEA VHAWRNALTGVPNLNTPDQVVAIASNIGGKQALETVQLLLPLVLCBQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIA NHDGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQLLLPLVCKDHGLTPDQVVAIANHDGKQALETVQRLLPVLCKELG LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGAKQALETVQRLLPVLCEQLGLTPAQVVAIASNIGGKQALETVQ LLPVLCEQHGLTPDQVVAIASHDGGKQALETVRLLPVLCEQLGMLTQVVAIASHYGGKQALEAVQRLLPVLCEQLGLTPDQVVAIASNNG GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVRLLPVLCEQLGMLTQ VVAIASNNGGKQALETVRLLPVLCEQLGMLTQVVAIASHYGGKQALEAVQRLLPVLCEQLGLTPDQVVAIASNNGGKQALETVQRLLPVL CEQHGLTPDQVVAIASHDGGKQALEAVQRLLPVLCEQHGLNPDQVVAIASNNGGGRQALESIFAQLSRPDQALAAALTNHDHLVALACLGCCPAL EAVKKGFLPHALTIKRTSRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEMAQFAMSRHGLLQLFRRVGVTELEACNGTLPPASQRW HRIQLASGVGTATPSRASATPDQASLDAFADALERGLDALSPIDQAVQQAQRSSRRRSSESSVTRSSAHYAVEVPVQRQAALDLSPPSW GAKRPRTRIGGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMELFPPQ

TAL18H2_CFBP6164
MRIRQGVGEPETFSL EATIFSIKRYARWIPFVRARQVLP TNFWPNPSRIGFSRSRLPDRGGSPPAGSPLDGLPARTMSRTRL PSPPAPLPA FSAGSFRDLLRQFDPSLLDTSLFDMSAFAFGAPHT EEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDAS PAGQVDLCTLGYSQQQEQEKILKARPVIAQHH EALIGHGFTRAHIVALSQHPAAITVAVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVS GELRGPPQLDGTGQLLKIARKGGVTAVEAVHAWRNAL TGVPNLNTPDQVVAIASNIGGKQALETVQLLPLVLC EQHGLTPDQVVAIASN GGKQALETVQRLLPVLCKDHGLTPAQVVA IANHDGGKQALETVQRLLPVLCKDHGLTPAQVVA IASNGGKQALETVQLLPLVCKDHGL TPDQVVAIANHDGGKPALETVQRLLPVL CQELGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNGGAKQALETV QRLLPVLCEQ ELGLTPAQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVA IASHDGGKQALETVRLLPVL CQELGMTLAQVVA IASHYGGKQ ALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQV VAIASH DGGKQALETVRLLPVL CQELGMTLAQVVAIASNNGGKQALETVRLLPVL CQELGMTLAQVVA IASHYGGKQALEAVQRLLPVL CQDLGL TPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVA IASHDGGKQALEAVQRLLPVL CQDHGLNPDQVVAIASNGGGRQALESI FAQLSRPDQALAAALNTDHLVALACLGCPAL EAVKKGLPVHALTLIKRSRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMAQFAM SRHGLLQFLFRVGVTELEACNGTLPPASQRWHRI LQASGVRTATPSRASAQTPDQASLDADFADALERGLDALSPIDQAVQAQASSRRKR SRSSSVTRSSAHYAVEVPVREQHAALDSLPPSWGAKRPRTRIGGGLADPGTPMHGD LAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIA WLME LFPQ

TAL19I_CFBP6546R

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQHHEA
 LIGHGFTRAQIVALSQHPPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLQDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTREQVVAIASNNGGKQALETVQR
 LLPVLCQAHLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPD
 QVVAIASHDGGKQALETVQRLLPVLCEQHGLTREQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPV
 LCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQA
 LESIFAQLSRPDQALAAALTNDHLVALACLGGRPALEAVKKGPHAPTILIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMT
 QFGMSRHGLLQLFRRVGVTELEARSGLTPAPQQRWHRIQASGMKRAEPPSGASQTPDQASLHAFADALERELDAPSPIDRAGQALASSSRK
 RSRSESVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGPLDPGTPMDADLAASSTMIWEQDAAPFVGEADDFPAFNNEEMAWL
 MELFPQ

TAL18H_CFBP6546R

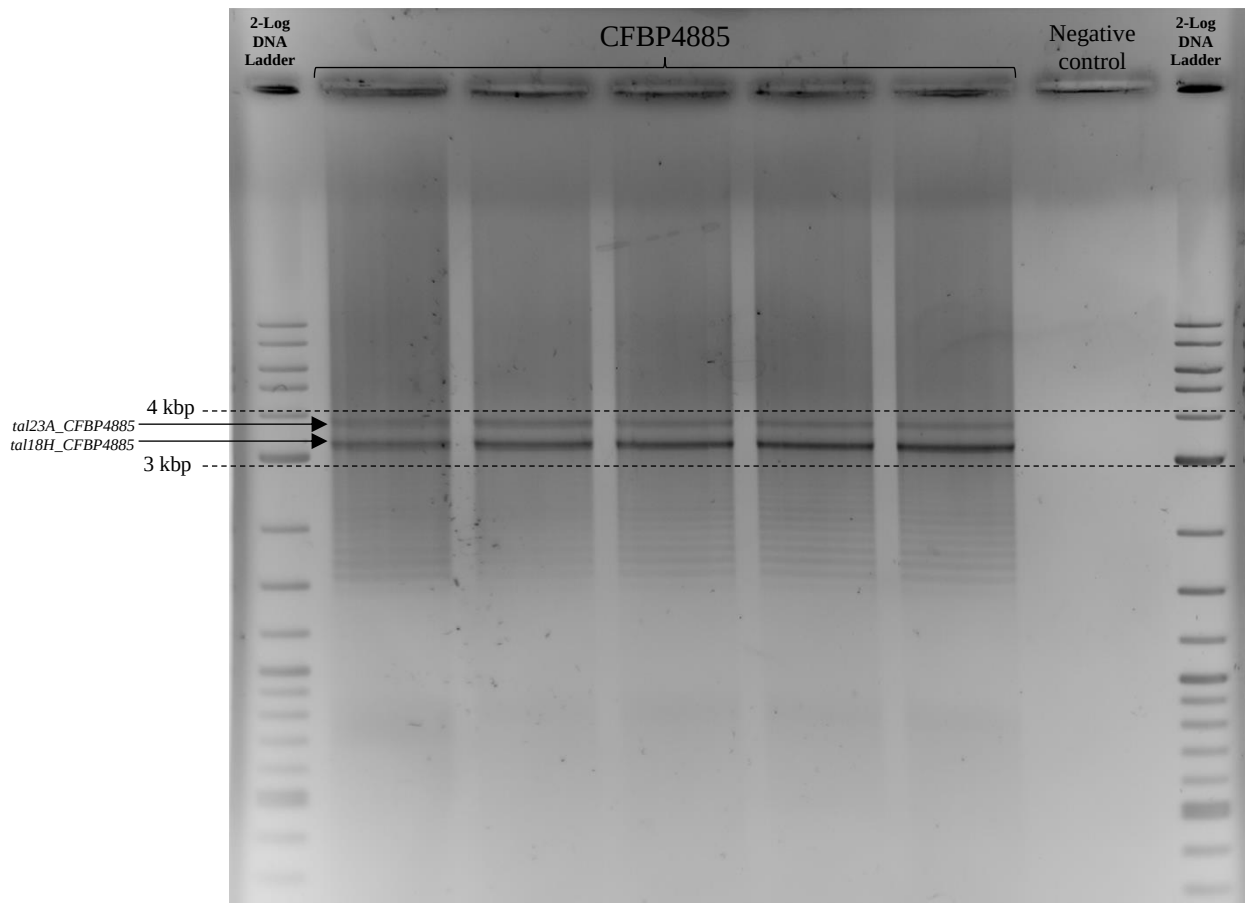
MDPIRPRAPSPAHELLAEFPQPDRVQPQPTADRGGSPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFRDILLRQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARPIVAQHHEA
 LIGHGFTRAHIVALSQHPPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLQDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGVPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIA
 NHDGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPDQVVAIANHDGGKPALETQRLLPVLCQELG
 LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNIGGKQALETVQR
 LLPVLCQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQELGMLTAQVVAIASHYGGKQALEAVQRLLPVLCQELGMLTPDQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQELGMLTAQ
 VVAIASNNGGKQALETVQRLLPVLCQELGMLTAQVVAIASHYGGKQALEAVQRLLPVLCQDGLTPDQVVAIASNNGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASHDGGKQALEAVQRLLPVLCQDHGLNPDQVVAIASNNGGKQALESIQAQLSRPDQALAAALTNDHLVALACLGCPAL
 EAVKKGPHALTILIKRTSRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMAQFAMSRHGLLQLFRRVGVTELEACNGTLPPASQRW
 HRILQASGVRTATPSRASAQTPDQASLDAFADALERGLDALSPIDQAVQAQASSRRKRSRSESVTRSSAHYAVEVPVREQHAALDSLPPSW
 GAKRPRTRIGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMEFLFPQ

TAL19E_CFBP6982

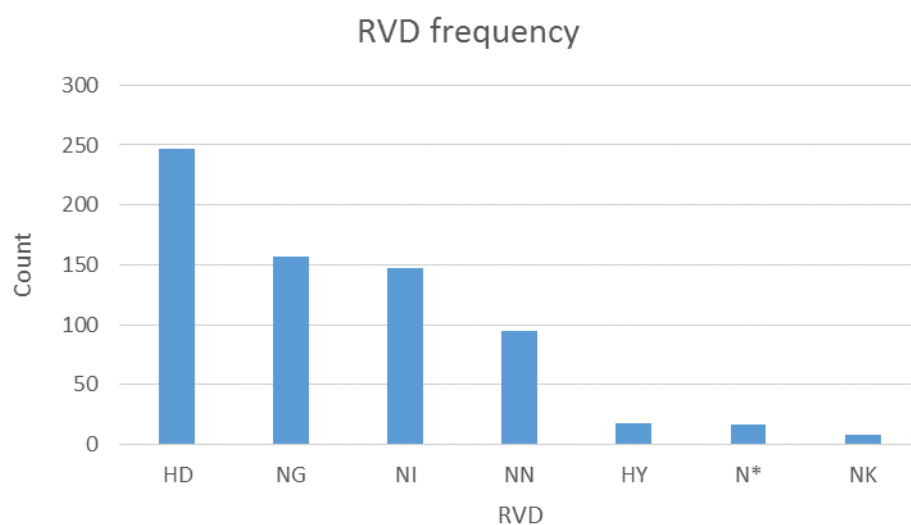
MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDDLRLQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARPIVAQHHEA
 LIGHGFTRAQIVALSQHPPAALGTAVVKYQVMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLQDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIA
 SNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGL
 LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQR
 LLPVLCQAHLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETLQRLLPVLCEQHGLTREQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETLQRLLPVLCEQHGLTRE
 QVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPV
 LCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHLTPDQVVAIASHDGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNNGGKQALE
 LESIFAQLSRPDQALAAALTNDHLVALACLGGRPALEAVKKGPHAPTILIKRTNRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMT
 QFGMSRHGLLQLFRRVGVTELEARSGLTPAPQQRWHRIQASGMKRAEPPSGASQTPDQASLHAFADALERELDAPSPIDRAGQALASSSRK
 RSRSESVTGSFAQQAVEVRVPEQHEGLHFPPLSWGVKRPRTRIGGGPLDPGTPMDADLAASSTMIWEQDAAPFVGEADDFPAFNNEEMAWL
 MELFPQ

TAL18H_CFBP6982

MDPIRPRAPSPAHELLAEFPQPDRVQPQPTADRGGSPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFRDILLRQFDPSSLDTSLFDSMS
 AFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARPIVAQHHEA
 LIGHGFTRAHIVALSQHPPAALGTAVVKYQAMIAALPEATHEDIVGVGKQWSGARALEALLTVSGELRGPPQLQDGTGQLLKIAKRGGVTAVEA
 VHAWRNALTGVPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPAQVVAIA
 NHDGGKQALETVQRLLPVLCKDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCKDHGLTPDQVVAIANHDGGKPALETQRLLPVLCQELG
 LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPAQVVAIASNIGGKQALETVQR
 LLPVLCQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQELGMLTAQVVAIASHYGGKQALEAVQRLLPVLCQELGMLTPDQVVAIASNNG
 GKQALETVQRLLPVLCEQHGLTPDQVVAIASNNGGKQALETVQRLLPVLCEQHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQELGMLTAQ
 VVAIASNNGGKQALETVQRLLPVLCQELGMLTAQVVAIASHYGGKQALEAVQRLLPVLCQDGLTPDQVVAIASNNGGKQALETVQRLLPVL
 CEQHGLTPDQVVAIASHDGGKQALEAVQRLLPVLCQDHGLNPDQVVAIASNNGGKQALESIQAQLSRPDQALAAALTNDHLVALACLGCPAL
 EAVKKGPHALTILIKRTSRRLPERTSHRVADHAQVARVLGFFQCHSHPAQAFDEAMAQFAMSRHGLLQLFRRVGVTELEACNGTLPPASQRW
 HRILQASGVRTATPSRASAQTPDQASLDAFADALERGLDALSPIDQAVQAQASSRRKRSRSESVTRSSAHYAVEVPVREQHAALDSLPPSW
 GAKRPRTRIGGLADPGTPMHGDLAASSTVAWEHDAAPFAAAEAGDFPAFNDEEIAWLMEFLFPQ



Additional file 3: Figure S1. Gel photograph showing the sizes obtained for *tal23A_CFBP4885* (*Xfuta1*) and *tal18H_CFBP4885* (*Xfuta2*) after PCR amplification. Amplification was performed using XapF2 and XapR primers on plasmidic DNA from the CFBP4885 strain (**Additional file 18: Table S11**). Expected PCR product sizes for *tal23A_CFBP4885* and *tal18H_CFBP4885* were 4604 bp and 3074 bp after Sanger genome sequencing, respectively, and 3686 bp and 3176 bp for SMRT sequencing, respectively.



Additional file 4: Figure S2. RVD frequencies in TAL effectors from the 17 *X. citri* pv. *fusca* and *X. phaseoli* pv. *phaseoli* strains used in this study.

Additional file 5: Table S3. References for previously published *tal* genes (a) and whole genome (b) sequences from *Xanthomonas* spp. strains used in this study. (1/2)

(a)		<i>tal</i> genes		
Species	Strains	<i>tal</i> gene	Accession number	Reference
<i>Xanthomonas campestris</i> pv. <i>armoraciae</i>	Xca5	<i>Hax2</i> ; <i>Hax3</i> ; <i>Hax4</i>	AY993937; AY993938; AY993939	Kay <i>et al.</i> , 2005
<i>Xanthomonas citri</i> pv. <i>aurantifolii</i>	B69	<i>PthB</i>	AAO72098	El Yacoubi <i>et al.</i> , 2007
	C340	<i>PthC</i>	EF473088	Al-Saadi <i>et al.</i> , 2007
	ICMP8435	<i>PthC1</i> ; <i>PthC2</i>	ADI48327; ADI48328	Pereira <i>et al.</i> , 2014
<i>Xanthomonas citri</i> pv. <i>citri</i>	306	<i>PthA1</i> ; <i>PthA2</i> ; <i>PthA3</i> ; <i>PthA4</i>	AAM39226; AAM39243; AAM39261; AAM39311	da Silva <i>et al.</i> , 2002
	3213	<i>PthA</i>	U28802	Yang & Gabriel, 1995
	270	<i>PthA*</i>	EF473086	Al-Saadi <i>et al.</i> , 2007
	X0053	<i>PthAW</i>	EF473085	
	049	<i>TalC</i>	KF682432	Yu <i>et al.</i> , 2011
<i>Xanthomonas citri</i> pv. <i>fuscans</i>				
<i>Xanthomonas citri</i> pv. <i>malvacearum</i>	XcmH	<i>avrb6</i>	L06634	De Feyter <i>et al.</i> , 1993
<i>Xanthomonas citri</i> "lablab-associated strains"				
<i>Xanthomonas euvesicatoria</i> pv. <i>euvesicatoria</i>	71-21	<i>avrBs3</i>	X16130	Bonas <i>et al.</i> , 1989
	82-8	<i>avrBs4</i>	X68781	Bonas <i>et al.</i> , 1993
<i>Xanthomonas gardneri</i>	XV4444	<i>avrHah1</i>	EF436255	Schornack <i>et al.</i> , 2008
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	AXO1947	<i>tal1</i> ; <i>tal2</i> ; <i>tal3</i> ; <i>tal4</i> ; <i>tal5</i> ; <i>tal6</i> ; <i>tal7</i> ; <i>tal8</i>	CP013666	Huguet-Tapia <i>et al.</i> , 2016
	JXOIII	<i>avrXa3</i> ; <i>avrXa5</i>	AY129298; FJ593881	Wu <i>et al.</i> , 2007; Zou <i>et al.</i> , 2010
	PXO86	<i>avrXa10</i>	XOU50552	Hopkins <i>et al.</i> , 1992
	PXO99A	<i>AvrXa27</i> ; <i>pthXo1</i> ; <i>pthXo6</i> ; <i>pthXo7</i>	AY986494; AY495676; DQ233317; DQ231560	Yang & White, 2004; Gu <i>et al.</i> , 2005; Sugio <i>et al.</i> , 2007
<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>	B8-12	<i>tal1</i> ; <i>tal2a</i> ; <i>tal2b</i> ; <i>tal3</i> ; <i>tal4a</i> ; <i>tal4b</i> ; <i>tal5b</i> ; <i>tal5c</i> ; <i>tal5d</i> ; <i>tal5e</i> ; <i>tal11</i> ; <i>tal12</i>	CP011955	Wilkins <i>et al.</i> , 2015
	BLS256	<i>tal1a</i> ; <i>tal2a</i> ; <i>tal2b</i> ; <i>tal2c</i> ; <i>tal2d</i> ; <i>tal2e</i> ; <i>tal2f</i> ; <i>tal2g</i> ; <i>tal3a</i> ; <i>tal3b</i> ; <i>tal3c</i> ; <i>tal4b</i> ; <i>tal5a</i> ; <i>tal5b</i> ; <i>tal6</i> ; <i>tal9b</i> ; <i>tal10</i>	CP003057	
	BLS279	<i>tal1</i> ; <i>tal5a</i> ; <i>tal5c</i> ; <i>tal5e</i>	CP011956	
	BXOR1	<i>tal1</i> ; <i>tal3</i> ; <i>tal5</i> ; <i>tal6</i> ; <i>tal7</i> ; <i>tal8a</i> ; <i>tal9a</i> ; <i>tal9b</i> ; <i>tal9c</i> ; <i>tal10c</i> ; <i>tal11a</i> ; <i>tal11b</i> ; <i>tal11d</i> ; <i>tal11f</i> ; <i>tal11g</i> ; <i>tal13</i>	CP011957	
	CFBP2286	<i>tal1a</i> ; <i>tal1b</i> ; <i>tal2a</i> ; <i>tal2b</i> ; <i>tal2e</i> ; <i>tal2f</i> ; <i>tal2h</i> ; <i>tal4a</i> ; <i>tal4b</i> ; <i>tal6</i> ; <i>tal9b</i> ; <i>tal11a</i>	CP011962	
	CFBP7331	<i>tal2</i> ; <i>tal3</i> ; <i>tal4</i> ; <i>tal5</i> ; <i>tal6b</i> ; <i>tal7a</i> ; <i>tal7b</i> ; <i>tal7c</i> ; <i>tal8a</i> ; <i>tal8b</i> ; <i>tal8c</i> ; <i>tal8d</i> ; <i>tal8e</i> ; <i>tal9a</i> ; <i>tal9b</i> ; <i>tal11a</i> ; <i>tal11b</i> ; <i>tal12</i>	CP011958	
	CFBP7341	<i>tal1</i> ; <i>tal6b</i> ; <i>tal8c</i>	CP011959	
	CFBP7342	<i>tal1</i> ; <i>tal2b</i> ; <i>tal3</i> ; <i>tal4b</i> ; <i>tal6a</i> ; <i>tal6b</i> ; <i>tal7a</i> ; <i>tal8</i> ; <i>tal9</i> ; <i>tal10b</i> ; <i>tal10c</i> ; <i>tal11a</i> ; <i>tal11b</i> ; <i>tal11d</i> ; <i>tal11e</i> ; <i>tal12</i>	CP007221	
	L8	<i>tal1</i> ; <i>tal2a</i> ; <i>tal5</i> ; <i>tal9b</i> ; <i>tal9c</i>	CP011960	
	RS105	<i>tal5c</i> ; <i>tal8b</i> ; <i>tal9a</i> ; <i>tal9c</i>	CP011961	
<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i>	CFBP1851	<i>pthB</i>	HQ113297	Castiblanco <i>et al.</i> , 2013
<i>Xanthomonas phaseoli</i> pv. <i>phaseoli</i>				
<i>Xanthomonas translucens</i> pv. <i>undulosa</i>	XT4699	<i>tal1</i> ; <i>tal2</i> ; <i>tal3</i> ; <i>tal4</i> ; <i>tal5</i> ; <i>tal6</i> ; <i>tal7</i> ; <i>tal8</i>	CP008714	Peng <i>et al.</i> , 2016

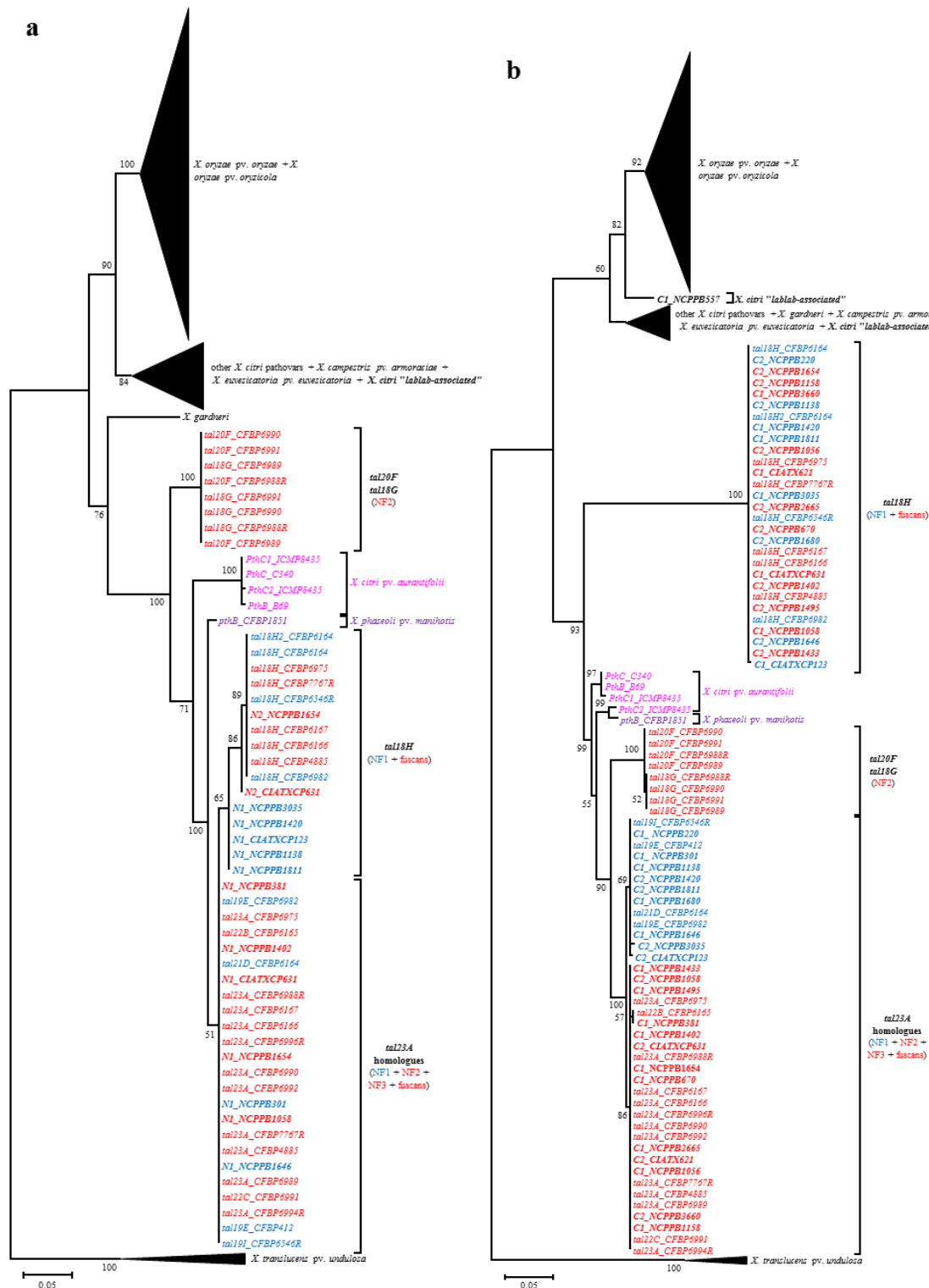
Additional file 5: Table S3. References for previously published *tal* genes (a) and whole genome (b) sequences from *Xanthomonas* spp. strains used in this study. (2/2)

(b)		Whole genomes	
Species	Strain	Accession number	Reference
<i>Xanthomonas campestris</i> pv. <i>armoraciae</i>	Xca5	CAOR00000000	Bolot <i>et al.</i> , 2013
<i>Xanthomonas citri</i> pv. <i>aurantifolii</i>	ICPB11122	ACPX00000000	Moreira <i>et al.</i> , 2010
<i>Xanthomonas citri</i> pv. <i>citri</i>	306	GCA_000007165 (AE008923; AE008924; AE008925)	da Silva <i>et al.</i> , 2002
	FDC 1083	CCVZ00000000	Gordon <i>et al.</i> , 2015
	JF90-2	CCWA00000000	
	JJ238-24	CCVX00000000	
	LC80	CCWJ00000000	
	LE20-1	CCWK00000000	
	LMG9322	JPYD00000000	
<i>Xanthomonas citri</i> pv. <i>fuscans</i>	NCPPB3660	JSEX00000000	Aritua <i>et al.</i> , 2015
	NCPPB381	JTKK00000000	
	CIATX621	JXHS00000000	
	CIATXCP631	JXLW00000000	
	NCPPB1056	JSEV00000000	
	NCPPB1058	JSEY00000000	
	NCPPB1433	JSBT00000000	
	NCPPB2665	JSBQ00000000	
	NCPPB1654	JSBR00000000	
	NCPPB670	JRRE00000000	
	NCPPB1402	JSEW00000000	
	NCPPB1158	JSBS00000000	
	NCPPB1495	JSEU00000000	
<i>Xanthomonas citri</i> pv. <i>malvacearum</i>	GSPB1386	AHIB00000000	Midha & Patil, 2014
	GSPB2388	AHIC00000000	
	X18	ATMA00000000	Cunnac <i>et al.</i> , 2013
	X20	ATMB00000000	
<i>Xanthomonas citri</i> "lablab-associated strains"	NCPPB556	JTJF00000000	Aritua <i>et al.</i> , 2015
	NCPPB557	JWTE00000000	
	NCPPB2064	JSEZ00000000	
	NCPPB1713	JWTD00000000	
<i>Xanthomonas euvesicatoria</i> pv. <i>euvesicatoria</i>	85-10	GCA_000009165 (AM039948; AM039949; AM039950; AM039951; AM039952)	Thieme <i>et al.</i> , 2005
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	AXO1947	CP013666	Huguet-Tapia <i>et al.</i> , 2016
	KACC10331	AE013598	Lee <i>et al.</i> , 2005
	MAFF311018	AP008229	Ochiai <i>et al.</i> , 2005
	PXO99A	CP000967	Salzberg <i>et al.</i> , 2008
<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>	B8-12	CP011955	Wilkins <i>et al.</i> , 2015
	BLS256	CP003057	
	BLS279	CP011956	
	BXOR1	CP011957	
	CFBP2286	GCA_001042735 (CP011962; CP011963)	
	CFBP7331	CP011958	
	CFBP7341	CP011959	
	CFBP7342	CP007221	
	L8	CP011960	
	RS105	CP011961	
<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i>	CFBP1851	AKCY00000000	Bart <i>et al.</i> , 2012
<i>Xanthomonas phaseoli</i> pv. <i>phaseoli</i>	NCPPB1646	JTCT00000000	Aritua <i>et al.</i> , 2015
	NCPPB301	JTCU00000000	
	CIATXCP123	JXLV00000000	
	NCPPB1420	JTCV00000000	
	NCPPB1811	JWTF00000000	
	NCPPB1680	JWTG00000000	
	NCPPB3035	JSFA00000000	
	NCPPB220	JWTH00000000	
<i>Xanthomonas translucens</i> pv. <i>undulosa</i>	XT4699	CP008714	Peng <i>et al.</i> , 2016

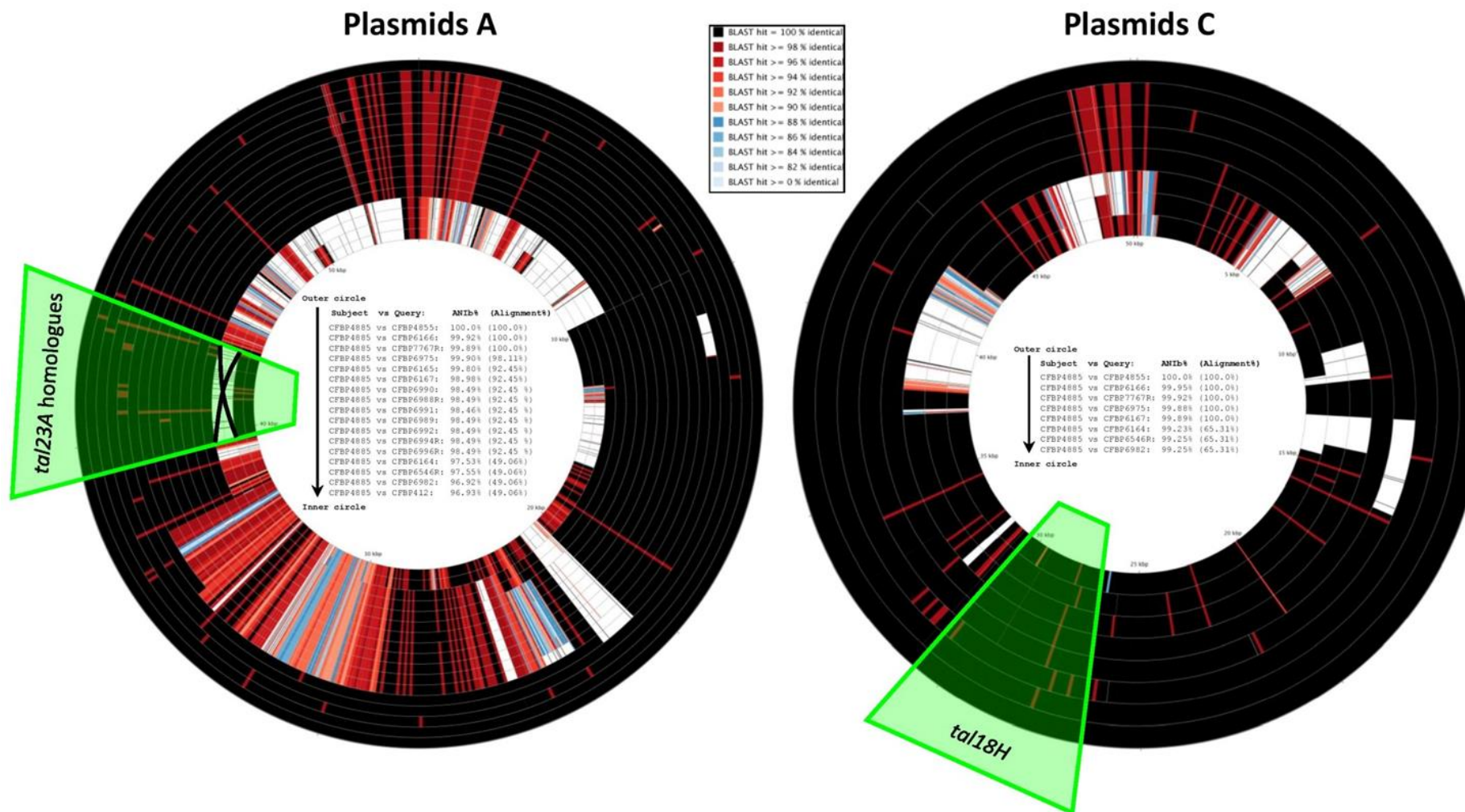
Additional file 6: Table S4. Average nucleotide identity and percentage of alignment between whole genome sequences from 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. BLAST calculation of average nucleotide identity (ANiB) and alignment percentages were estimated using Jspecies (Richter & Rosselló-Móra, 2009).

ANiB (%)	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)						<i>X. citri</i> pv. <i>fuscans</i> (NF2)				<i>X. citri</i> pv. <i>fuscans</i> (NF3)			<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)			
Species	Strain	CFBP4885	CFBP6165	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6988R	CFBP6989	CFBP6990	CFBP6991	CFBP6992	CFBP6994R	CFBP6996R	CFBP412	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	99.88	99.95	99.75	99.87	99.93	98.37	98.38	98.37	98.37	98.18	98.17	98.15	93.45	93.67	93.43	93.48
	CFBP6165	99.78	---	99.77	99.80	99.80	99.76	98.57	98.57	98.58	98.55	98.30	98.31	98.24	93.40	93.51	93.28	93.39
	CFBP6166	99.91	99.87	---	99.71	99.83	99.88	98.41	98.41	98.41	98.42	98.19	98.20	98.16	93.41	93.61	93.39	93.43
	CFBP6167	99.63	99.87	99.64	---	99.66	99.62	98.40	98.40	98.40	98.38	98.20	98.20	98.19	93.45	93.70	93.48	93.49
	CFBP6975	99.89	99.87	99.91	99.76	---	99.87	98.36	98.36	98.36	98.36	98.16	98.15	98.11	93.40	93.60	93.39	93.45
<i>X. citri</i> pv. <i>fuscans</i> (NF2)	CFBP7767R	99.84	99.82	99.83	99.65	99.73	---	98.36	98.36	98.36	98.37	98.27	98.17	98.10	93.37	93.41	93.35	93.46
	CFBP6988R	98.52	98.89	98.50	98.64	98.47	98.54	---	100.0	99.98	99.96	98.70	98.65	98.65	93.08	93.08	92.99	93.06
	CFBP6989	98.51	98.86	98.50	98.62	98.45	98.53	100.0	---	99.98	99.97	98.69	98.65	98.61	93.08	93.08	93.00	93.07
	CFBP6990	98.51	98.87	98.49	98.60	98.43	98.53	99.98	99.98	---	99.98	98.65	98.62	98.60	93.07	93.06	92.99	93.05
<i>X. citri</i> pv. <i>fuscans</i> (NF3)	CFBP6991	98.46	98.88	98.46	98.56	98.42	98.50	99.93	99.93	99.95	---	98.63	98.58	98.58	93.02	92.98	92.94	93.00
	CFBP6992	98.16	98.59	98.16	98.09	98.13	98.29	98.50	98.50	98.48	98.50	---	99.88	99.84	92.99	92.80	92.91	92.95
	CFBP6994R	98.19	98.60	98.19	98.15	98.15	98.22	98.60	98.60	98.58	98.60	99.99	---	99.93	92.97	92.96	92.88	92.94
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP6996R	98.07	98.51	98.07	98.02	98.02	98.09	98.44	98.44	98.42	98.42	99.78	99.77	---	92.86	92.85	92.78	92.83
	CFBP412	93.59	93.93	93.59	93.59	93.54	93.59	93.26	93.26	93.26	93.43	93.28	93.26	93.23	---	99.89	99.82	99.95
	CFBP6164	93.66	93.91	93.65	93.64	93.58	93.48	93.19	93.19	93.19	93.42	93.14	93.29	93.27	99.85	---	99.80	99.82
	CFBP6546R	93.69	93.93	93.69	93.68	93.63	93.71	93.22	93.22	93.22	93.51	93.40	93.37	93.33	99.87	99.92	---	99.89
	CFBP6982	93.58	93.89	93.59	93.6	93.56	93.61	93.19	93.19	93.19	93.39	93.29	93.26	93.21	99.88	99.84	99.77	---

Alignment (%)	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)						<i>X. citri</i> pv. <i>fuscans</i> (NF2)				<i>X. citri</i> pv. <i>fuscans</i> (NF3)			<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)			
Species	Strain	CFBP4885	CFBP6165	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6988R	CFBP6989	CFBP6990	CFBP6991	CFBP6992	CFBP6994R	CFBP6996R	CFBP412	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	94.89	99.96	98.42	99.51	99.47	96.26	96.26	96.26	96.34	96.44	96.44	95.72	94.75	95.01	94.77	95.19
	CFBP6165	98.12	---	98.58	99.5	98.9	98.8	96.71	96.71	96.71	96.71	97.73	97.69	97.01	95.03	95.07	94.41	94.89
	CFBP6166	98.94	94.57	---	98.47	98.88	98.98	95.7	95.7	95.7	95.76	95.82	95.8	95.1	94.04	94.39	93.76	94.55
	CFBP6167	96.14	91.88	97.27	---	96.8	96.91	93.68	93.68	93.68	93.9	94.74	94.74	93.98	93.53	94.4	93.49	93.98
	CFBP6975	98.26	94.19	98.64	97.95	---	98.42	96.02	96.02	96.02	96.16	95.98	96.08	95.34	94.54	94.85	94.27	94.91
<i>X. citri</i> pv. <i>fuscans</i> (NF2)	CFBP7767R	96.84	92.94	97.26	96.58	97.11	---	93.84	93.84	93.84	93.88	96.19	93.99	93.32	92.6	93.94	92.43	92.88
	CFBP6988R	93.53	89.92	94.02	93.98	94.59	93.84	---	100.0	100.0	100.0	95.98	96.15	95.35	92.00	92.13	91.39	92.37
	CFBP6989	93.58	90.17	94.05	93.99	94.62	93.93	100.0	---	100.0	100.0	95.9	96.06	95.23	91.93	91.91	91.16	92.24
	CFBP6990	93.5	90.14	94.03	94.07	94.62	93.87	100.0	100.0	---	100.0	96.01	96.19	95.34	92.16	92.2	91.46	92.52
<i>X. citri</i> pv. <i>fuscans</i> (NF3)	CFBP6991	91.62	87.76	92.1	92.08	92.66	91.93	97.81	97.81	97.81	---	93.81	94.01	93.24	90.11	90.34	89.49	90.51
	CFBP6992	91.86	89.09	92.32	93.74	92.69	94.44	94.15	94.15	94.15	94.15	---	97.44	96.75	90.18	91.11	89.47	90.59
	CFBP6994R	93.73	91.04	94.18	95.42	94.67	94.05	95.76	95.76	95.76	95.76	99.12	---	98.29	91.75	91.83	91.1	92.12
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP6996R	94.57	91.72	95.01	96.37	95.51	94.89	96.73	96.73	96.73	96.73	100.0	100.0	---	92.97	93.05	92.3	93.25
	CFBP412	93.96	82.91	94.32	94.77	94.89	94.34	94.06	94.06	94.06	94.08	94.43	94.51	93.9	---	99.47	98.88	99.88
	CFBP6164	92.72	81.06	93.06	94.01	93.55	93.95	92.59	92.59	92.59	92.72	93.71	92.91	92.32	97.42	---	97.16	97.73
	CFBP6546R	94.6	82.31	94.67	95.12	95.38	94.69	94.52	94.52	94.52	94.56	94.81	94.89	94.44	99.49	99.75	---	99.77
	CFBP6982	93.75	81.9	94.12	94.25	94.49	94.02	93.56	93.56	93.56	93.58	93.64	93.72	93.16	98.96	98.88	98.34	---



Additional file 7: Figure S3. Phylogenetic trees of N-ter- and C-ter-encoding regions of *tal* genes with sequences published by Aritua *et al.*, 2015. Bootstrap values greater than 50% (100 replicates) are shown and horizontal scale bars represent the number of nucleotide substitutions per site. The *tal* genes from the *X. citri* pv. *fuscans* genetic lineages fuscans, NF2 and NF3 are indicated in red, *X. phaseoli* pv. *phaseoli* NF1 lineage in blue, *X. citri* pv. *aurantifolii* in pink and *X. phaseoli* pv. *manihotis* in purple. The sequences published by Aritua *et al.*, 2015 are in bold. Both ML trees were constructed using *tal* genes from *X. translucens* pv. *undulosa* XT4699 as outgroups. **a.** ML tree constructed on a nucleotide alignment of the N-ter-encoding region of *tal* genes. **b.** ML tree constructed on a nucleotide alignment of the C-ter-encoding region of *tal* genes.

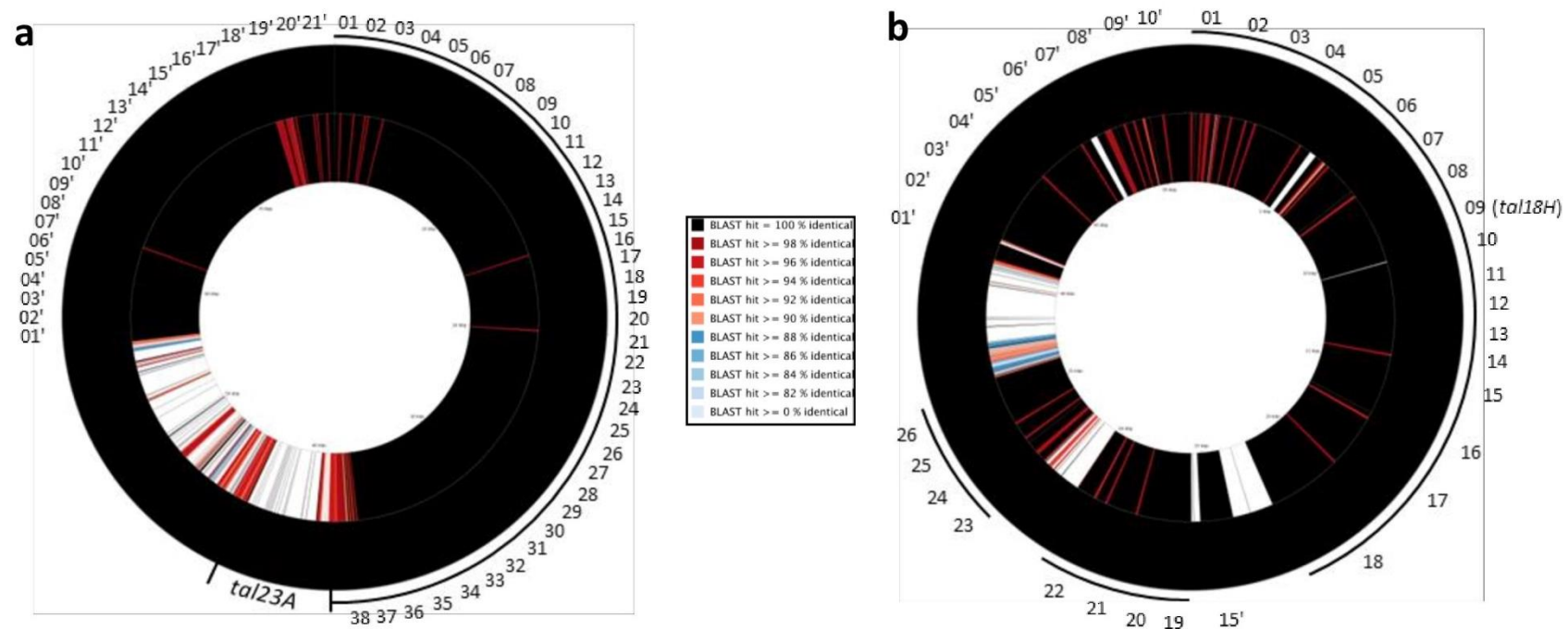


Additional file 8: Figure S4. Circular representation of plasmids A and C. Genomic sequences were compared and converted in a graphical map using CGView (Grant *et al.*, 2012) with strain CFBP4885 as reference. Colours differ according to identity percentage (see legend). Strains order, ANIb and alignment percentages are indicated in the center of each graphical map. Localisation of *tal* genes are highlighted by green zones. Absence of *tal* genes in plasmid A from strains CFBP6164 and CFBP6546R is indicated by a black cross.

Additional file 9: Table S5. Average nucleotide identity and percentage of alignment between sequences of plasmid A from 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. BLAST calculation of average nucleotide identity (ANiB) and alignment percentages were estimated using Jspecies (Richter & Rosselló-Móra, 2009).

ANiB (%)	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)						<i>X. citri</i> pv. <i>fuscans</i> (NF2)				<i>X. citri</i> pv. <i>fuscans</i> (NF3)			<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)			
Species	Strain	CFBP4885	CFBP6165	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6988R	CFBP6989	CFBP6990	CFBP6991	CFBP6992	CFBP6994R	CFBP6996R	CFBP412	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	99.8	99.92	98.98	99.9	99.89	98.49	98.49	98.49	98.46	98.49	98.49	98.49	96.93	97.53	97.55	96.92
	CFBP6165	99.8	---	99.8	98.69	99.8	99.77	98.5	98.5	98.5	98.47	98.5	98.5	98.5	96.93	97.84	97.87	96.92
	CFBP6166	99.92	99.78	---	98.13	99.9	99.89	97.37	97.37	97.37	97.27	97.37	97.35	97.37	95.9	96.32	95.94	95.89
	CFBP6167	98.72	98.58	98.72	---	98.72	98.7	99.79	99.79	99.79	99.79	99.79	99.78	99.79	97.05	99.64	99.34	97.04
	CFBP6975	99.97	99.84	99.97	98.12	---	99.94	97.58	97.58	97.58	97.54	97.58	97.58	97.58	96.33	96.67	96.63	96.32
	CFBP7767R	99.95	99.82	99.94	98.94	99.93	---	98.69	98.69	98.69	98.6	98.69	98.7	98.69	97.13	97.95	97.9	97.11
<i>X. citri</i> pv. <i>fuscans</i> (NF2)	CFBP6988R	98.54	98.39	98.54	99.81	98.54	98.51	---	99.97	99.97	99.96	99.97	99.96	99.97	97.17	99.87	99.65	97.09
	CFBP6989	98.5	98.31	98.5	99.81	98.49	98.47	99.99	---	99.99	99.96	99.99	99.98	99.98	98.12	99.84	99.81	98.07
	CFBP6990	98.17	98.05	98.17	99.81	98.17	98.14	99.99	99.99	---	99.92	99.99	99.98	99.99	96.44	99.59	99.33	96.44
	CFBP6991	98.08	97.87	98.08	99.72	98.08	98.05	99.91	99.91	99.91	---	99.91	99.9	99.9	97.56	99.78	99.53	97.55
<i>X. citri</i> pv. <i>fuscans</i> (NF3)	CFBP6992	98.0	97.79	98.0	99.81	97.96	97.98	99.99	99.99	99.99	99.96	---	99.98	99.99	97.72	99.85	99.83	97.62
	CFBP6994R	97.42	97.26	97.41	99.81	97.41	97.4	99.99	99.99	99.99	99.92	99.99	---	99.99	97.47	99.88	99.6	97.47
	CFBP6996R	98.15	97.96	98.15	99.79	98.14	98.13	99.99	99.99	99.99	99.91	99.99	99.98	---	97.76	99.42	99.42	97.74
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP412	96.52	96.22	96.53	97.31	96.53	96.48	97.12	97.12	97.12	97.11	97.12	97.12	97.12	---	98.25	98.24	99.98
	CFBP6164	96.86	96.75	96.86	99.22	96.81	96.82	99.41	99.41	99.41	99.41	99.42	99.41	99.42	98.59	---	99.75	98.61
	CFBP6546R	96.48	96.39	96.47	99.27	96.42	96.46	98.91	98.91	98.91	98.91	98.91	98.91	98.91	98.74	99.85	---	98.79
	CFBP6982	97.23	97.09	97.23	97.56	97.22	97.2	96.94	96.94	96.94	96.9	96.94	96.94	96.94	99.98	98.91	98.66	---

Alignment (%)	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)						<i>X. citri</i> pv. <i>fuscans</i> (NF2)				<i>X. citri</i> pv. <i>fuscans</i> (NF3)			<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)			
Species	Strain	CFBP4885	CFBP6165	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6988R	CFBP6989	CFBP6990	CFBP6991	CFBP6992	CFBP6994R	CFBP6996R	CFBP412	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	92.45	100.0	92.45	98.11	100.0	92.45	92.45	92.45	92.45	92.45	92.45	92.45	49.06	49.06	49.06	49.06
	CFBP6165	98.25	---	98.25	87.72	98.25	98.25	87.72	87.72	87.72	87.72	87.72	87.72	87.72	59.65	64.91	64.91	59.65
	CFBP6166	100.0	92.31	---	92.31	98.08	100.0	90.38	90.38	90.38	90.38	90.38	90.38	90.38	61.54	61.54	61.54	61.54
	CFBP6167	77.05	73.77	77.05	---	77.05	77.05	96.72	96.72	96.72	96.72	96.72	96.72	96.72	62.3	73.77	72.13	62.3
	CFBP6975	100.0	92.59	100.0	88.89	---	100.0	88.89	88.89	88.89	88.89	88.89	88.89	88.89	64.81	62.96	61.11	64.81
	CFBP7767R	100.0	92.16	100.0	90.2	100.0	---	86.27	86.27	86.27	86.27	86.27	86.27	86.27	50.98	58.82	56.86	50.98
<i>X. citri</i> pv. <i>fuscans</i> (NF2)	CFBP6988R	66.67	62.67	66.67	86.67	65.33	66.67	---	100.0	100.0	100.0	100.0	100.0	100.0	65.33	76.0	77.33	64.0
	CFBP6989	71.64	67.16	71.64	86.57	70.15	71.64	100.0	---	100.0	100.0	100.0	100.0	100.0	50.75	59.7	59.7	49.25
	CFBP6990	55.88	55.88	55.88	88.24	55.88	55.88	100.0	100.0	---	100.0	100.0	100.0	100.0	67.65	72.06	75.0	67.65
	CFBP6991	68.57	64.29	68.57	82.86	68.57	68.57	100.0	100.0	100.0	---	100.0	100.0	100.0	62.86	71.43	72.86	62.86
<i>X. citri</i> pv. <i>fuscans</i> (NF3)	CFBP6992	67.69	64.62	67.69	90.77	64.62	67.69	100.0	100.0	100.0	100.0	---	100.0	100.0	66.15	73.85	73.85	63.08
	CFBP6994R	70.77	69.23	70.77	87.69	70.77	70.77	100.0	100.0	100.0	100.0	100.0	---	100.0	66.15	69.23	72.31	66.15
	CFBP6996R	72.13	70.49	72.13	88.52	72.13	72.13	100.00	100.00	100.00	100.00	100.00	100.00	---	65.57	65.57	67.21	65.57
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP412	49.23	46.15	49.23	58.46	49.23	47.69	64.62	64.62	64.62	64.62	64.62	64.62	64.62	---	76.92	75.38	98.46
	CFBP6164	46.38	44.93	46.38	57.97	44.93	46.38	65.22	65.22	65.22	65.22	65.22	65.22	65.22	82.61	---	100.0	84.06
	CFBP6546R	56.0	54.67	56.0	66.67	54.67	56.0	76.0	76.0	76.0	76.0	76.0	76.0	76.0	81.33	98.67	---	84.0
	CFBP6982	37.66	36.36	37.66	40.26	37.66	37.66	46.75	46.75	46.75	46.75	46.75	46.75	46.75	89.61	75.32	74.03	---



Additional file 10: Figure S5. Circular representation of the regions shared by phylogenetically distant lineages in plasmids A and C. Genomic sequences were compared and converted in a graphical map using CGView (Grant *et al.*, 2012) with strain CFBP4885 as reference. Colours differ according to identity percentage (see legend). Non-redundant regions shared by both strains are surrounded by a black line. Numbers represent genes conserved in both strains (see **Additional files 11 and 13** for more details). **a.** Plasmids A from strains CFBP6988R (outer circle = reference) and CFBP6546R (inner circle). **b.** Plasmids C from strains CFBP6166 (outer circle = reference) and CFBP6982 (inner circle).

Additional file 11: Table S6. Genes localized in plasmid A shared by strains CFBP6988R and CFBP6546R. Please refer to **Additional file 10a** for graphic representation of the data.

Genes shared by both strains	Start (bp)	End (bp)	Genes duplicated in strain CFBP6988R	Start (bp)	End (bp)	Predicted function (InterProScan)
<i>plasmidA_01</i>	1	2698	<i>plasmidA_01'</i>	56767	59766	tnpA : IPR002513:Tn3 transposase DDE domain
<i>plasmidA_02</i>	2695	3272	<i>plasmidA_02'</i>	59763	60341	IPR006119:Resolvase, N-terminal
<i>plasmidA_03</i>	3750	4166	<i>plasmidA_03'</i>	60819	61235	IPR002716:PiIT protein, N-termina
<i>plasmidA_04</i>	4174	4734	<i>plasmidA_04'</i>	61243	61803	IPR012912:Plasmid pRiA4b, Orf3
<i>plasmidA_05</i>	4803	4946	<i>plasmidA_05'</i>	61872	62015	IPR002514:Transposase IS3/IS911family
<i>plasmidA_06</i>	4995	5543	<i>plasmidA_06'</i>	62064	62612	IPR000380:DNA topoisomerase, type IA
<i>plasmidA_07</i>	5695	7018	<i>plasmidA_07'</i>	62767	64087	IPR001590:Peptidase M12B, ADAM/reprolysin
<i>plasmidA_08</i>	7117	8616	<i>plasmidA_08'</i>	64186	65685	IPR011050:Pectin lyase fold/virulence factor
<i>plasmidA_09</i>	8811	9665	<i>plasmidA_09'</i>	65880	66734	IPR000380:DNA topoisomerase, type IA
<i>plasmidA_10</i>	9667	9915	<i>plasmidA_10'</i>	67057	67296	Conserved hypothetical protein
<i>plasmidA_11</i>	9988	10227	<i>plasmidA_11'</i>	67307	67615	Conserved hypothetical protein
<i>plasmidA_12</i>	10238	10546	<i>plasmidA_12'</i>	67686	69038	Conserved hypothetical protein
<i>plasmidA_13</i>	10617	11969	<i>plasmidA_13'</i>	69054	70046	IPR005498:Type IV secretion system, VirB10 / TraB / TrbI
<i>plasmidA_14</i>	11985	12977	<i>plasmidA_14'</i>	70073	70723	IPR010258:Conjugal transfer, TrbG/VirB9/CagX
<i>plasmidA_15</i>	13004	13654	<i>plasmidA_15'</i>	70790	72268	IPR007430:Bacterial virulence protein VirB8
<i>plasmidA_16</i>	13721	15119	<i>plasmidA_16'</i>	72835	73515	IPR007688:Conjugal transfer, TrbL/VirB6
<i>plasmidA_17</i>	15767	16540	<i>plasmidA_17'</i>	73490	73606	Conserved hypothetical protein
<i>plasmidA_18</i>	16559	19129	<i>plasmidA_18'</i>	73767	76182	IPR018145:CagE, TrbE, VirB component of type IV transporter system, central domain
<i>plasmidA_19</i>	19196	19462	<i>plasmidA_19'</i>	76249	76515	Conserved hypothetical protein
<i>plasmidA_20</i>	19465	19779	<i>plasmidA_20'</i>	76672	76830	IPR007039:Conjugal transfer TrbC/type IV secretion VirB2
<i>plasmidA_21</i>	19859	20863	<i>plasmidA_21'</i>	76910	77006	IPR001482:Type II secretion system protein E
<i>plasmidA_22</i>	21860	22678				Conserved hypothetical protein
<i>plasmidA_23</i>	23254	23805				Conserved hypothetical protein
<i>plasmidA_24</i>	23786	24232				IPR005468:Avidin/streptavidin
<i>plasmidA_25</i>	24614	25840				TXax1 : IPR001207:Transposase, mutator type
<i>plasmidA_26</i>	26349	26924				IPR012912:Plasmid pRiA4b, Orf3
<i>plasmidA_27</i>	26921	27412				cryptic genes
<i>plasmidA_28</i>	27409	29460				IPR003688:Ti-type conjugative transfer system, TraG-like
<i>plasmidA_29</i>	29465	32419				IPR005094:Endonuclease relaxase, MobA/VirD2
<i>plasmidA_30</i>	32706	33286				IPR008258:Lytic transglycosylase-like, catalytic
<i>plasmidA_31</i>	33289	33885				IPR000223:Peptidase S26A, signal peptidase I
<i>plasmidA_32</i>	33890	34156				IPR021263:Protein of unknown function DUF2840
<i>plasmidA_33</i>	34598	34933				cryptic genes
<i>plasmidA_34</i>	34930	35565				IPR002586:CobQ/CobB/MinD/ParA nucleotide binding domain
<i>plasmidA_35</i>	36583	37329				Conserved hypothetical protein
<i>plasmidA_36</i>	37758	38342				cryptic genes
<i>plasmidA_37</i>	38454	38798				Conserved hypothetical protein
<i>plasmidA_38</i>	56567	56677				Conserved hypothetical protein

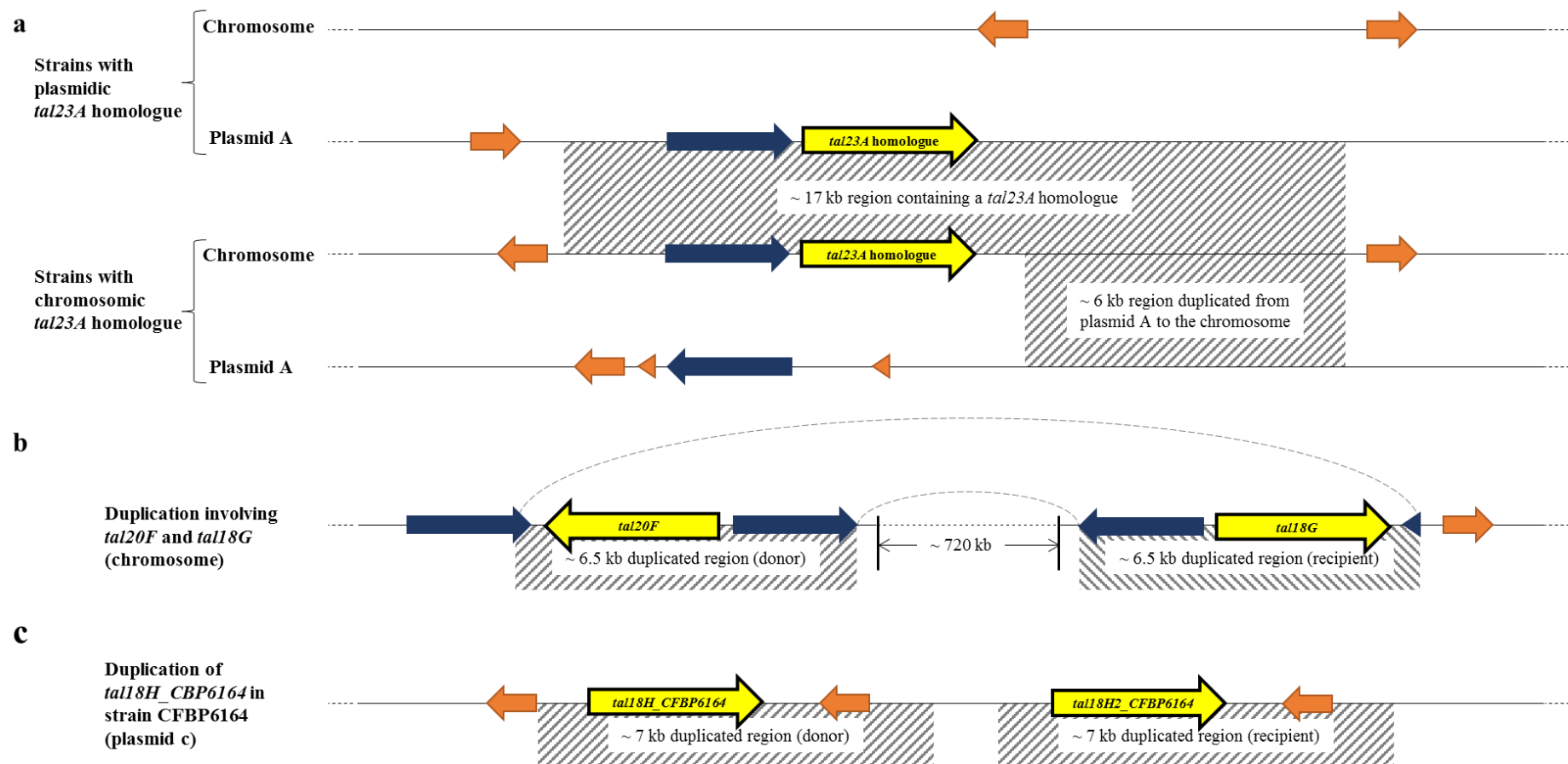
Additional file 12: Table S7. Average nucleotide identity and percentage of alignment between sequences of plasmid C in eight *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* (NF1) strains. BLAST calculation of average nucleotide identity (ANiB) and alignment percentages were estimated using (Richter & Rosselló-Móra, 2009).

ANiB (%)	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)					<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)		
Species	Strain	CFBP4885	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	99.95	99.89	99.88	99.92	99.23	99.25	99.4
	CFBP6166	99.96	---	99.88	99.89	99.92	99.24	99.25	99.43
	CFBP6167	99.93	99.94	---	99.93	99.91	99.39	99.38	99.48
	CFBP6975	99.89	99.89	99.89	---	99.86	99.26	99.27	99.34
	CFBP7767R	99.94	99.93	99.87	99.85	---	99.43	99.44	99.43
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP6164	99.53	99.54	99.53	99.53	99.49	---	99.82	99.87
	CFBP6546R	99.3	99.3	99.28	99.2	99.28	99.72	---	99.7
	CFBP6982	99.51	99.52	99.48	99.42	99.49	99.88	99.87	---

Alignment	Species	<i>X. citri</i> pv. <i>fuscans</i> (fuscans)					<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)		
Species	Strain	CFBP4885	CFBP6166	CFBP6167	CFBP6975	CFBP7767R	CFBP6164	CFBP6546R	CFBP6982
<i>X. citri</i> pv. <i>fuscans</i> (fuscans)	CFBP4885	---	100.0	100.0	93.88	100.0	65.31	65.31	65.31
	CFBP6166	100.0	---	96.0	90.0	100.0	72.0	72.0	76.0
	CFBP6167	88.14	88.14	---	74.58	88.14	67.8	67.8	69.49
	CFBP6975	96.36	96.36	94.55	---	96.36	61.82	61.82	67.27
	CFBP7767R	100.0	100.0	100.0	90.2	---	70.59	70.59	72.55
<i>X. phaseoli</i> pv. <i>phaseoli</i> (NF1)	CFBP6164	66.67	66.67	66.67	62.96	66.67	---	96.3	90.74
	CFBP6546R	62.32	62.32	63.77	59.42	62.32	94.2	---	86.96
	CFBP6982	68.42	68.42	71.93	61.4	68.42	91.23	91.23	---

Additional file 13: Table S8. Genes localized in plasmid C shared by strains CFBP6982 and CFBP6166. Please refer to **Additional file 10b** for graphic representation of the data.

Genes shared by both strains	Start (bp)	End (bp)	Genes duplicated in strain CFBP6988R	Start (bp)	End (bp)	Predicted function (InterProScan)
<i>plasmidC_01</i>	1	175	<i>plasmidC_01'</i>	41886	42206	Conserved hypothetical protein
<i>plasmidC_02</i>	385	1826	<i>plasmidC_02'</i>	42385	43878	Replication and maintenance
<i>plasmidC_03</i>	2732	3316	<i>plasmidC_03'</i>	44012	44527	IPR021104:KfrA, N-terminal DNA-binding domain
<i>plasmidC_04</i>	3329	3898	<i>plasmidC_04'</i>	44783	45367	IPR011335:Restriction endonuclease type II-like
<i>plasmidC_05</i>	3965	4570	<i>plasmidC_05'</i>	45380	45949	pin : IPR012337:Ribonuclease H-like domain
<i>plasmidC_06</i>	4586	4792	<i>plasmidC_06'</i>	46016	46621	Conserved hypothetical protein
<i>plasmidC_07</i>	4786	5052	<i>plasmidC_07'</i>	46637	46843	IPR002514:Transposase IS3/IS911family
<i>plasmidC_08</i>	5049	5846	<i>plasmidC_08'</i>	46837	47103	Conserved hypothetical protein
<i>plasmidC_09</i>	6103	8830	<i>plasmidC_09'</i>	47100	47897	IPR005042:Avirulence protein PthA-related (tal gene encoding TAL18H RVD combination)
<i>plasmidC_10</i>	9705	10025	<i>plasmidC_10'</i>	48311	51342	Conserved hypothetical protein
<i>plasmidC_11</i>	11079	11903				Conserved hypothetical protein
<i>plasmidC_12</i>	12576	12797				Conserved hypothetical protein
<i>plasmidC_13</i>	13220	13474				IPR011991: TrfB transcriptional repressor protein
<i>plasmidC_14</i>	14044	15021				IPR003812:Adenosine monophosphate-protein transferase VbhT
<i>plasmidC_15</i>	15513	16739	<i>plasmidC_15'</i>	24009	25235	IPR001207:Transposase, mutator type
<i>plasmidC_16</i>	16886	18432				Conserved hypothetical protein
<i>plasmidC_17</i>	19165	20712				IPR014128:Type IV conjugative transfer system, coupling protein TraD
<i>plasmidC_18</i>	20709	22109				IPR014129:Conjugative transfer relaxase protein Tral
<i>plasmidC_19</i>	25572	25811				cryptic genes
<i>plasmidC_20</i>	26130	26618				IPR006119:Resolvase, N-terminal
<i>plasmidC_21</i>	27197	27886				IPR001091:Restriction/modification DNA-methylase
<i>plasmidC_22</i>	27883	28884				IPR019071:Restriction endonuclease, type II, Xcyl
<i>plasmidC_23</i>	32023	32706				IPR012337:Ribonuclease H-like domain
<i>plasmidC_24</i>	32703	32861				Conserved hypothetical protein
<i>plasmidC_25</i>	33266	35770				xopC1 : IPR023214:HAD-like domain; virulence associated
<i>plasmidC_26</i>	36323	37150				TXcd1_2 : IPR012337:Ribonuclease H-like domain

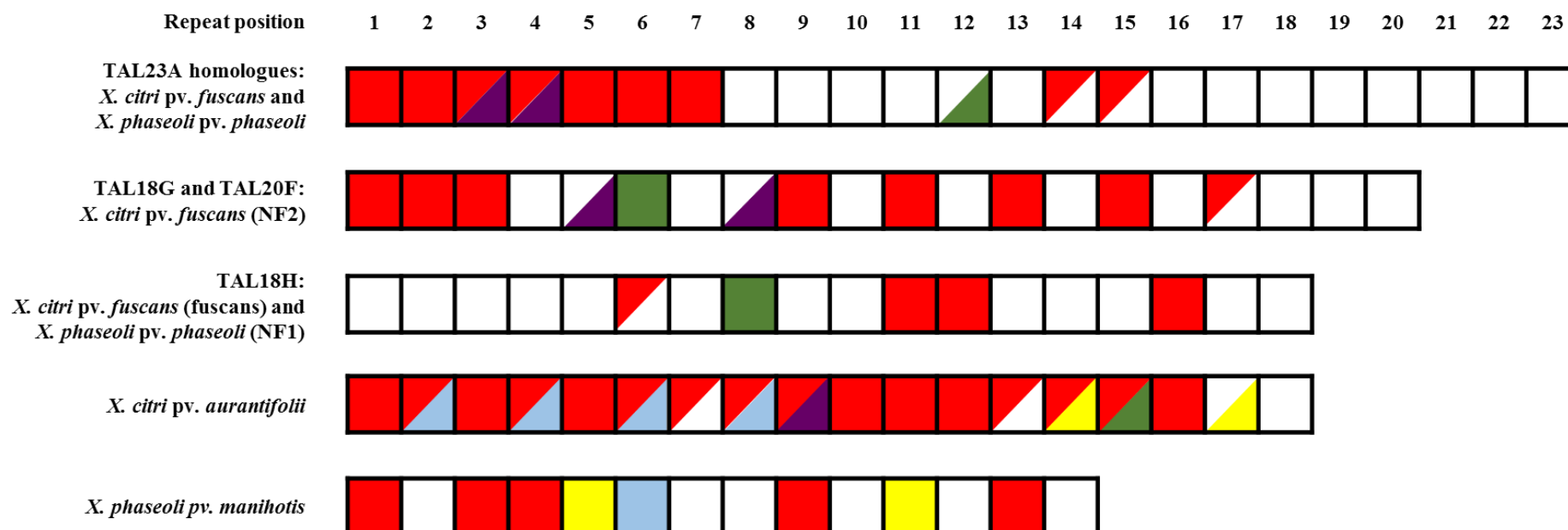


Additional file 14: Figure S6. Duplications and movements of *tal* genes associated with IS3 and Tn3-like transposons. Arrows represent *tal* genes (yellow), IS3 sequences (orange) or Tn3-like transposons (dark blue). Arrowheads correspond to truncated IS3 (orange) or Tn3-like transposon (dark blue). Striped areas correspond to conserved regions. Dashed lines represent the borders of duplication events between different regions of a single molecule. **a.** Movement of *tal23A* homologues from plasmid A to the chromosome, associated with IS3 sequences. **b.** Duplication of *tal20F* and *tal18G* associated with Tn3-like transposons. **c.** Duplication of *tal18H_CFBP6164* associated with IS3 sequences in strain CFBP6164.

Additional file 15: Table S9. Repeats found within all TAL effectors from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains. Sequences were compared after removal of RVD. *Xcf* = *X. citri* pv. *fuscans*; *Xpp* = *X. phaseoli* pv. *phaseoli*; *Xca* = *X. citri* pv. *aurantifolii*; *Xpm* = *X. phaseoli* pv. *manihotis*. A coloured line indicates that the repeat is shared between TAL effectors from different groups, one colour being associated with a single repeat and corresponding to the colour of repeats in **Additional file 16: Figure S7**.

Repeat sequence	Total number	Comment
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	259	shared between all TAL effectors from <i>Xcf</i> , <i>Xpp</i> , <i>Xca</i> and <i>Xpm</i> strains
LTPEQVVAIASGGKQALETVQRLLPVLCEQHG	78	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>)
LTPEQVVAIASGGKQALETVQRLLPVLCEQAHG	44	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>)
LTPAQVVAIASGGKQALETVQRLLPVLCEQHG	33	shared between TAL23A homologs (except TAL23A_CFBP6975 (<i>Xcf</i> -fuscans); TAL19E_CFBP6164 and TAL19E_CFBP412 (<i>Xpp</i> -NF1)); TAL18G and TAL20F (<i>Xcf</i> -NF2); PthC_C340 and PthC1_ICMP8435 (<i>Xca</i>)
LTREQVVAIASGGKQALETVQRLLPVLCEQHG	33	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>)
LTREQVVAIASGGKQALETVQRLLPVLCEQAHG	31	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>)
LTPDQVVAIASGGKQALETVQRLLPVLCEQAHG	31	shared between TAL23A homologs (<i>Xcf</i> -NF2 and NF3); TAL18G and TAL20F (<i>Xcf</i> -NF2); PthC_C340 (<i>Xca</i>)
LTPDQVVAIASGGKQALETVQRLLPVLCEQAHG	20	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
LTREQVVAIASGGKQALETVQRLLPVLCEQHG	17	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>)
LTPDQVVAIASGGKQALETVQRLLPVLCEQELG	17	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALETVQRLLPVLCEQAHG	16	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
LTPAQVVAIASGGKQALE	14	specific for TAL23A homologs (<i>Xpp</i> -NF1 and <i>Xcf</i>) except TAL22B_CFBP6165 (<i>Xcf</i> -fuscans); TAL21D_CFBP6164 and TAL19I_CFBP6546R (<i>Xpp</i> -NF1)
LTPEQVVAIASGGKQALETVQRLLPVLCEQHG	12	specific for TAL23A homologs (<i>Xcf</i>) except TAL22_CFBP6165 (<i>Xcf</i> -fuscans) and <i>Xpp</i> -NF1 strains
LTPDQVVAIASGGKQALETVQRLLPVLCKDHG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALEAVQRLLPVLCEQDHG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALETVQRLLPVLCEQELG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
MTLAQVVAIASGGKQALEAVQRLLPVLCEQELG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
MTLAQVVAIASGGKQALETVQRLLPVLCEQELG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LNPQVVAIASGGKQALE	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPAQVVAIASGGKQALETVQRLLPVLCKDHG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALETVQRLLPVLCEQELG	9	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans)
LTPAQVVAIASGGKQALETVQRLLPVLCKDHG	8	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans) except TAL18H_CFBP6982 (<i>Xpp</i> -NF1)
LTPDQVVAIASGGKQALE	8	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
MTLAQVVAIASGGKQALEAVQRLLPVLCEQDLG	8	specific for TAL18H (<i>Xpp</i> -NF1 and <i>Xcf</i> -fuscans) except TAL18H_CFBP6975 (<i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	8	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	8	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	8	specific for TAL18G and TAL20F (<i>Xcf</i> -NF2)
LTPDQVVAIASGGKQALETVQRLLPVLCEQAHG	7	shared between TAL effectors from <i>Xca</i> and <i>Xpm</i> except PthC2_ICMP8435 (<i>Xca</i>)
LTPAQVVAIASGGKQALETVQRLLPVLCEQHG	6	shared between TAL effectors from <i>Xca</i> and <i>Xpm</i> except PthC2_ICMP8435 (<i>Xca</i>)
LTPAQVVAIASGGKPALE	4	specific for TAL effectors from <i>Xca</i>
LTPDQVVAIASGGKQALETVQRLLPVLCEQELG	3	specific for TAL18H from TAL18H_CFBP4885; TAL18H_CFBP6166; TAL18H_CFBP7767R (<i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALE	2	specific for TAL23A homologs from TAL21D_CFBP6164; TAL19I_CFBP6546R (<i>Xpp</i> -NF1)
LTPDQVVAIASGGKQALETVQRLLPVLCEQHG	2	specific for TAL23A homologs from TAL119E_CFBP412; TAL19E_CFBP6982 (<i>Xpp</i> -NF1)
LTPEQVVAIASGGKQALE	1	specific for TAL23A homologs from TAL22B_CFBP6165 (<i>Xcf</i> -fuscans)
LTPDQVVAIASGGKQALERVQRLLPVLCEQHG	1	specific for PthC1_ICMP8435 (<i>Xca</i>)
LTLQVVAIASGGKQALETVQRLLPVLCEQHG	1	specific for pthB_CFBP1851 (<i>Xpm</i>)
MTLAQVVAIASGGKQALETVQRLLPVLCEQDLG	1	specific for TAL18H from TAL18H_CFBP6975 (<i>Xcf</i> -fuscans)
LTPAQVVAIASGGKQALETVQRLLPVLCEQHG	1	specific for pthB_CFBP1851 (<i>Xpm</i>)
LTPAQVVAIASGGKPALETVQRLLPVLCEQHG	1	specific for pthB_CFBP1851 (<i>Xpm</i>)
LTPAQVVAIASGGKQALETVQRLLPVLCEQAHG	1	specific for TAL23A homologs from TAL23A_CFBP6975 (<i>Xcf</i> -fuscans)
LTPDQVVAIASGGKPALETVQRLLPVLCEQHG	1	specific for PthC2_ICMP8435 (<i>Xca</i>)
LTPEQVVAIASGGKQALETVQRLLPVLCEQELG	1	specific for TAL18H from TAL18H_CFBP6167 (<i>Xcf</i> -fuscans)
LTPDQVVAIASGGKPALE	1	specific for pthB_CFBP1851 (<i>Xpm</i>)
LTPAQVVAIASGGKQALETVQRLLPVLCKDHG	1	specific for TAL18H from TAL18H_CFBP6982 (<i>Xpp</i> -NF1)
LTPAQVVAIASGGKQALETVQRLLPVLCEQHG	1	specific for PthB_B69 (<i>Xca</i>)
LTPAQVVAIASGGKQALETVQRLLPVLCEQAHG	1	specific for pthB_CFBP1851 (<i>Xpm</i>)
LTPDQVVAIASGGKPALETVQRLLPVLCEQHG	1	specific for pthB_CFBP1851 (<i>Xpm</i>)

Total 772



Additional file 16: Figure S7. Repeats conserved and shared between different TAL effectors from *X. citri* pv. *fuscans*, *X. phaseoli* pv. *phaseoli*, *X. citri* pv. *aurantifolii* and *X. phaseoli* pv. *manihotis* strains. Repeats conserved between different TAL effectors are indicated by the same colour (red, purple, green, blue or yellow). Half-coloured boxes indicate that polymorphism exist among TAL effectors homologues from different strains. Please see **Additional file 15: Table S9** for more detailed information.

Additional file 17: Table S10. Top 10 predicted targets for the nine TAL effectors with unique RVD sequences retrieved in 17 *X. citri* pv. *fuscans* and *X. phaseoli* pv. *phaseoli* strains. Target names correspond to locus tags from the *P. vulgaris* genome available at <https://phytozome.jgi.doe.gov/>. Similar targets predicted for different TAL effectors are highlighted by colours corresponding to the predicted function of the target (see the legend below).

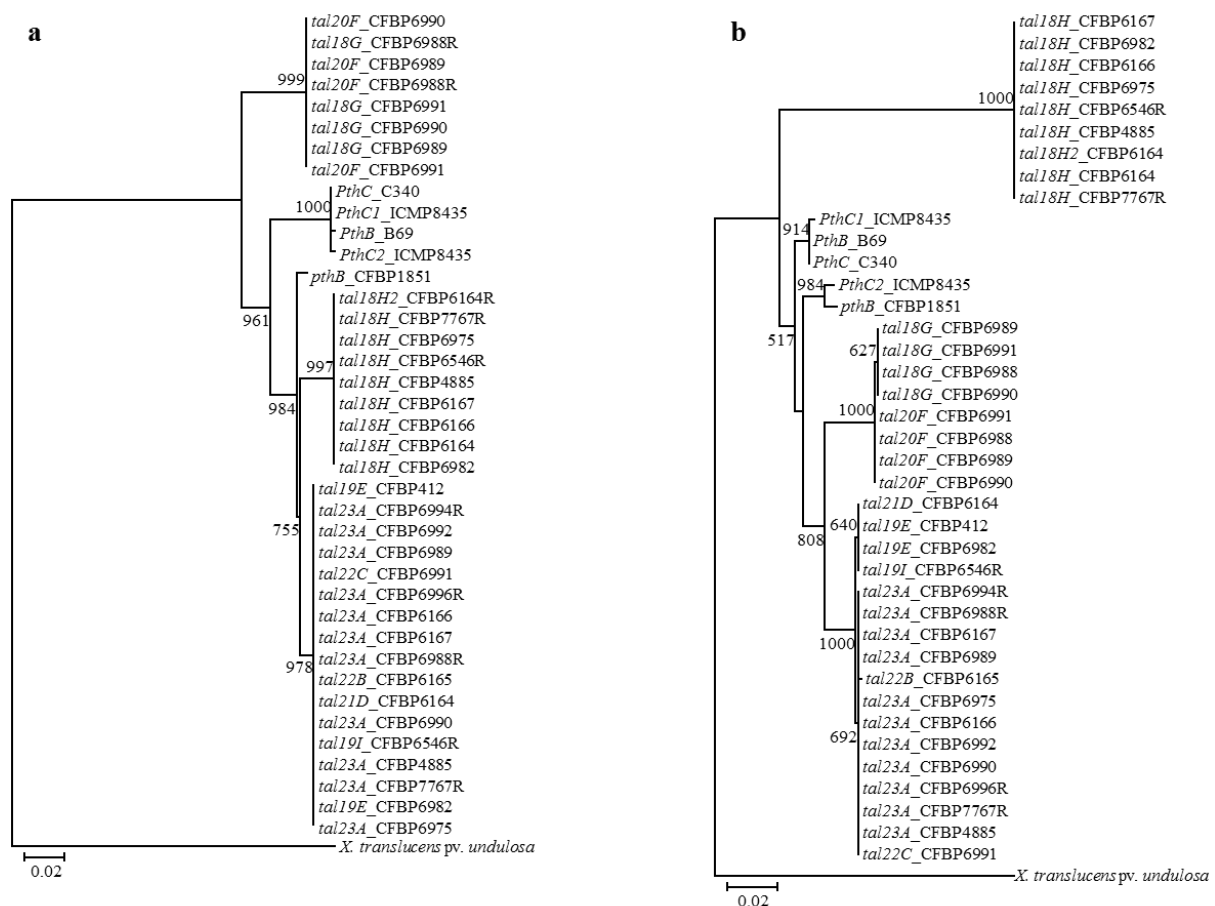
TAL RVD combination	TAL23A		TAL22B		TAL22C		TAL21D		TAL19E		TAL19I		TAL20F		TAL18G		TAL18H	
Rank in TALVEZ 3.2 predictions	Target	Score	Target	Score	Target	Score	Target	Score	Target	Score	Target	Score	Target	Score	Target	Score	Target	Score
1	Phvul.002G328604	15.698	Phvul.008G275300	13.895	Phvul.002G052200	16.652	Phvul.010G030100	14.399	Phvul.006G032900	13.938	Phvul.010G159300	13.639	Phvul.005G097300	14.785	Phvul.010G009400	15.069	Phvul.009G148800	14.799
2	Phvul.006G032900	15.194	Phvul.004G052500	13.797	Phvul.003G042100	15.884	Phvul.002G052200	14.184	Phvul.011G016600	13.933	Phvul.009G058400	13.265	Phvul.005G088500	14.308	Phvul.005G097300	14.410	Phvul.001G247600	13.488
3	Phvul.002G052200	14.350	Phvul.002G328604	13.686	Phvul.003G097300	15.877	Phvul.008G022975	13.881	Phvul.002G043100	13.509	Phvul.009G058500	13.265	Phvul.006G167500	13.991	Phvul.002G123400	14.163	Phvul.007G045600	13.442
4	Phvul.010G030100	14.151	Phvul.003G201700	13.322	Phvul.004G069900	15.648	Phvul.L001655	13.845	Phvul.010G137400	13.248	Phvul.006G023100	13.073	Phvul.006G048600	13.950	Phvul.006G048600	13.980	Phvul.003G074300	13.364
5	Phvul.005G184500	13.923	Phvul.006G032900	13.182	Phvul.002G195000	15.640	Phvul.003G097300	13.823	Phvul.011G155600	13.241	Phvul.006G032900	13.070	Phvul.003G248500	13.922	Phvul.005G058600	13.977	Phvul.001G202300	13.336
6	Phvul.003G097300	13.575	Phvul.004G025700	13.024	Phvul.004G091300	15.640	Phvul.011G031300	13.659	Phvul.002G193900	12.793	Phvul.009G022032	12.901	Phvul.008G285600	13.887	Phvul.003G292800	13.806	Phvul.001G079500	13.336
7	Phvul.002G270000	13.545	Phvul.001G039300	13.007	Phvul.005G077900	15.640	Phvul.010G159200	13.603	Phvul.001G217500	12.624	Phvul.002G052200	12.882	Phvul.011G074200	13.829	Phvul.008G022900	13.785	Phvul.003G243400	13.336
8	Phvul.008G176100	13.357	Phvul.001G236400	12.984	Phvul.001G112200	15.163	Phvul.003G201700	13.509	Phvul.003G152000	12.606	Phvul.011G115900	12.810	Phvul.003G292800	13.785	Phvul.002G258200	13.752	Phvul.004G073200	13.336
9	Phvul.008G022975	13.237	Phvul.002G103000	12.970	Phvul.008G089600	15.136	Phvul.003G193900	13.469	Phvul.003G124100	12.578	Phvul.004G018200	12.782	Phvul.010G009400	13.716	Phvul.010G029900	13.732	Phvul.008G092400	13.158
10	Phvul.003G196200	13.215	Phvul.002G302600	12.936	Phvul.001G062100	15.128	Phvul.003G196200	13.463	Phvul.007G251200	12.439	Phvul.007G172700	12.774	Phvul.011G006500	13.691	Phvul.011G214900	13.693	Phvul.003G183900	12.855

Predicted functions were coloured as follow

Vacuolar protein sorting-associated protein VPS13
Peptidase of plants and bacteria
40S ribosomal protein S29
Putative serine/threonine protein kinase
UDP-glucosyl transferase
Nucleolar ribonucleoprotein subunit
RING/Ubox like zinc-binding domain
F-box/Leucine rich repeat protein
Unknown function
SAM dependent carboxyl methyltransferase
Nucleic acid deaminase of DYW family
Oligopeptide transporter-related

Additional file 18: Table S11. Primers used in this study.

Primer name	Designed on strain	<i>tal</i> genes	Localisation	Sequence	Product size	Used in strains
Xap-F2	CFBP4885	<i>tal23A_CFBP4885</i> and <i>tal18H_CFBP4885</i>	N-ter	CATACAGAGGCTGCCTCAG	3686 bp or 3176 bp	CFBP4885
Xap-R	CFBP4885	<i>tal23A_CFBP4885</i> and <i>tal18H_CFBP4885</i>	C-ter	TCGTTGAATGCCGGGAAATC		
F7767C23	CFBP7767R	<i>tal23A_CFBP7767R</i>	C-ter	GCAGGCACGGGTTGTTACAGC	479 bp	CFBP4885; CFBP7767R; CFBP6546R; CFBP6994R
R7767C23	CFBP7767R	<i>tal23A_CFBP7767R</i>	C-ter	GGCCGCATCTTGTTCCCAGATC		
F7767C18	CFBP7767R	<i>tal18H_CFBP7767R</i>	C-ter	CGGCACGGTTTGTTACAACTCTTTC	469 bp	CFBP4885; CFBP6546R; CFBP7767R
R7767C18	CFBP7767R	<i>tal18H_CFBP7767R</i>	C-ter	CGTGTTCCTCAAGCCACGGTG		
F7767N	CFBP7767R	<i>tal18H_CFBP7767R</i>	upstream of the gene	GATGGGTTTCAGGATCGCC	691 bp	CFBP6546R; CFBP7767R
R7767N	CFBP7767R	<i>tal18H_CFBP7767R</i>	N-ter	CTATCGGACGAGCCTTCAG		
F6988N18	CFBP6988R	<i>tal18G_CFBP6988R</i>	N-ter	GGTGAAGAATTGCAGCCTCCGAC	2740 bp	CFBP6988R
R6988N	CFBP6988R	<i>tal18G_CFBP6988R</i>	upstream of the gene	GAGCGCCGAAGGCAGGC		



Additional file 19: Figure S8. Phylogenetic trees of N-ter- and C-ter-encoding regions of *tal* genes used for the Kishino-Hasegawa-Templeton test. Bootstrap values greater than 50% are shown for 1000 replicates and horizontal scale bars represent the number of nucleotide substitutions per site. a. ML tree constructed on a nucleotide alignment of the N-ter-encoding region of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains using a *tal* gene from *X. translucens* pv. *undulosa* XT4699 as outgroup. b. ML tree constructed on a nucleotide alignment of the C-ter-encoding region of *tal* genes from *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* and *X. phaseoli* pv. *manihotis* strains using a *tal* gene from *X. translucens* pv. *undulosa* XT4699 as outgroup.

III. Contenu génomique d'une diversité de souches de *X. citri* pv. *vignicola*

Phaseolus et *Vigna* sont les genres regroupant les principales légumineuses vivrières de la sous tribu des Phaseolinae. Le niébé (*Vigna unguiculata* subsp. *unguiculata* (L.) Walp., 1843) est une légumineuse cultivée dans les régions semi-arides tropicales et subtropicales essentiellement en Afrique, mais également en Amérique et en Asie. Le niébé joue un rôle important dans l'alimentation humaine où il peut être consommé de différentes manières, mais il aussi utilisé en tant que fourrage et pour l'alimentation animale. La brûlure bactérienne du niébé est une maladie majeure de cette culture car elle engendre d'importantes pertes de rendement partout où le niébé est cultivé. Cette bactériose est provoquée par *X. citri* pv. *vignicola*, voisin phylogénétique de *X. citri* pv. *fuscans* (Mhedbi-Hajri *et al.*, 2013). Une précédente analyse phylogénétique réalisée dans l'équipe sur la base de sept gènes de ménage a montré que le positionnement des cinq souches disponibles de *X. citri* pv. *vignicola* n'était pas identique pour chacun des sept gènes (Mhedbi, 2010). La souche CFBP7111 était toujours distante des quatre autres souches, tandis que la souche CFBP7113 n'était pas regroupée avec les autres souches pour quatre des sept gènes de ménage utilisés. Aucun génome de *X. citri* pv. *vignicola* n'étant disponible, nous avons séquencé le génome de trois souches représentant la diversité génétique observée : CFBP7111, CFBP7112 et CFBP7113. Les ADN ont été extraits et envoyés à séquencer par Sophie Bonneau, les génomes ont été séquencés par l'Icahn School of Medicine at Mount Sinai de New York et les données génomiques brutes obtenues ont été assemblées par Martial Briand et on fait l'objet d'une publication dont je suis le premier auteur (Ruh *et al.*, 2017b).



AMERICAN
SOCIETY FOR
MICROBIOLOGY

genomeAnnouncements™

PROKARYOTES



First Complete Genome Sequences of *Xanthomonas citri* pv. *vignicola* Strains CFBP7111, CFBP7112, and CFBP7113 Obtained Using Long-Read Technology

Mylène Ruh, Martial Briand, Sophie Bonneau, Marie-Agnès Jacques,
Nicolas W. G. Chen

IRHS, INRA, Agrocampus Ouest, Université d'Angers, SFR4207 QUASAV, Beaucozoué, France

ABSTRACT *Xanthomonas citri* pv. *vignicola* strains cause bacterial blight of the legume crop cowpea. We report whole-genome sequences of three *X. citri* pv. *vignicola* strains obtained using PacBio single-molecule real-time sequencing. Such genomic data provide new information on pathogenicity factors, such as transcription activator-like effectors.

Cowpea (*Vigna unguiculata*) is a legume crop grown in semiarid tropical and subtropical regions of Africa, America, and Asia (1). It is used as staple food in different forms (grain, green pods, roots, or leaves) and as animal feed (2). Cowpea bacterial blight (CoBB) is a major disease of cow pea that severely reduces crop yields (3) and is present wherever cowpea is grown (4). The causal agent of CoBB is *Xanthomonas citri* pv. *vignicola* (formerly *X. Axonopodis* pv. *vignicola*) (5). *X. Citri* pv. *vignicola* mainly produces symptoms on leaves, starting as tiny water-soaked spots that evolve into large necrotic areas surrounded by a yellow halo. This pathogen can also cause stem canker and water-soaking of pods, providing the opportunity for the bacteria to gain access to seeds (6). CoBB is seed-borne and can also be transmitted by aerosols and insects (7). Moreover, *X. citri* pv. *vignicola* is able to survive in infected plant debris, seeds, and soil, which constitute different inoculum sources (8, 9). Some genotypes of cowpea have been selected for resistance to CoBB (10); however, no information on the molecular basis for this resistance is available to date.

We sequenced three strains of *X. citri* pv. *vignicola* using PacBio single-molecule real-time (SMRT) sequencing (11). Genomic DNA was extracted using the Wizard genomic DNA purification kit (Promega) according to the manufacturer's recommendations. Whole genomes were sequenced using the PacBio RSII SMRT technology at the Icahn School of Medicine at Mount Sinai (New York, NY). One SMRT cell was used per strain to achieve ~100× coverage. *De novo* assembly was performed using Canu version 1.5 (12), circularization of the contigs was done using Circlator version 1.5.1 (13), and polishing was done using the Quiver algorithm (14). Annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline (15).

Each *X. citri* pv. *vignicola* genome comprised one chromosome plus zero to three circularized plasmids per strain. The genome of strain CFBP7111 comprised three plasmids, for a total assembly size of 5,710,499 bp, with a G+C content of 64.5% and a total of 5,095 predicted coding sequences (CDS). The genome of strain CFBP7112 comprised two plasmids, for a total assembly size of 5,155,996 bp, with a G+C content of 64.7% and a total of 4,460 predicted CDS. The genome of strain CFBP7113 comprised no plasmids, for a total assembly size of 5,105,205 bp, with a G+C content of 64.7% and a total of 4,446 predicted CDS. Interestingly, the content and localization of genes encoding transcription activator-like type III effectors (*tal*) differed among the three

Received 11 July 2017 Accepted 18 July 2017 Published 7 September 2017

Citation Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017. First complete genome sequences of *Xanthomonas citri* pv. *vignicola* strains CFBP7111, CFBP7112, and CFBP7113 obtained using long-read technology. *Genome Announc* 5:e00813-17. <https://doi.org/10.1128/genomeA.00813-17>.

Copyright © 2017 Ruh et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Nicolas W. G. Chen, nicolas.chen@agrocampus-ouest.fr.

genomes. Strain CFBP7111 had no *tal* gene, while strain CFBP7112 contained seven *tal* genes located on plasmids, and strain CFBP7113 bore one *tal* gene located on the chromosome. This diversity may reflect different modes of adaptation of these strains to cowpea.

Accession number(s). The three whole-genome sequences have been deposited at GenBank under accession numbers CP022263 to CP022266 for strain CFBP7111, CP022267 to CP022269 for strain CFBP7112, and CP022270 for strain CFBP7113.

ACKNOWLEDGMENTS

This work was supported by grants from the French National Research Agency (CROpTAL project, grant number ANR-14-CE19-0002-04) and Angers-Loire Metropole, France (X-HOT project).

We thank Laurent D. Noël, Boris Szurek, Lionel Gagnevin, and Ralf Koebnik for recurrent scientific exchanges on the project CROpTAL led by Laurent D. Noël.

PacBio sequencing was performed at the Icahn School of Medicine at Mount Sinai (NY, USA) under the supervision of Robert Sebra and preliminary quality control of DNA was performed at the ANAN platform from the SFR 4207 QUASAV (Angers, France).

We also thank the CIRM-CFBP (Angers, France) for strain preservation and supply.

REFERENCES

- Ehlers JD, Hall AE. 1997. Cowpea (*Vigna unguiculata* L. Walp.). Field Crops Res 53:187–204. [https://doi.org/10.1016/S0378-4290\(97\)00031-2](https://doi.org/10.1016/S0378-4290(97)00031-2).
- Singh BB, Ajeigbe HA, Tarawali SA, Fernandez-Rivera S, Abubakar M. 2003. Improving the production and utilization of cowpea as food and fodder. Field Crops Res 84:169–177. [https://doi.org/10.1016/S0378-4290\(03\)00148-5](https://doi.org/10.1016/S0378-4290(03)00148-5).
- de Lima-Primo HE, de L. Nechet K, de A. Halfeld-Vieira B, de Oliveira JR, Mizubuti ESG, de Souza GR. 2015. Epidemiological aspects of cowpea bacterial blight. Trop Plant Pathol 40:46–55. <https://doi.org/10.1007/s40858-015-0007-y>.
- Khatri-Chhetri GB, Wydra K, Rudolph K. 2003. Metabolic diversity of *Xanthomonas axonopodis* pv. *vignicola*, causal agent of cowpea bacterial blight and pustule. Eur J Plant Pathol 109:851–860. <https://doi.org/10.1023/A:1026104716650>.
- Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B. 2016. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. Plant Pathol 65:792–806. <https://doi.org/10.1111/ppa.12461>.
- Williams RJ. 1975. Diseases of cowpea (*Vigna unguiculata* (L.) Walp.) in Nigeria. Trop Pest Manag 21:253–267. <https://doi.org/10.1080/09670877509411407>.
- Zandjanakou-Tachin M, Fanou A, Le Gall P, Wydra K. 2007. Detection, survival and transmission of *Xanthomonas axonopodis* pv. *manihotis* and *X. axonopodis* pv. *vignicola*, causal agents of cassava and cowpea bacterial blight, respectively, in/by insect vectors. J Phytopathol 155: 159–169. <https://doi.org/10.1111/j.1439-0434.2007.01210.x>.
- Okechukwu R, Ekpo E. 2008. Survival of *Xanthomonas campestris* pv. *vignicola* in infested soil, cowpea seed and cowpea debris. Trop Agric Res Ext 11:45–48.
- Sikirou R, Wydra K. 2004. Persistence of *Xanthomonas axonopodis* pv. *vignicola* in weeds and crop debris and identification of *Sphenostylis stenocarpa* as a potential new host. Eur J Plant Pathol 110:939–947. <https://doi.org/10.1007/s10658-004-8949-9>.
- Duche TR, Iheukwumere CC, Omoigui L. 2015. Evaluation of selected cowpea genotypes for resistance to bacterial blight. Int J Curr Microbiol Appl Sci 4:257–270.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. <https://doi.org/10.1126/science.1162986>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity GM, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta) genomic annotation. OMICS 12:137–141. <https://doi.org/10.1089/omi.2008.0017>.

Tableau 1 : Caractéristiques génomiques des souches séquencées.

Espèce	Hôte	Souche	Année d'isolation (pays)	Nombre de contig	Taille du génome (pb)	% de GC	Nombre total de gènes	Nombre de gène <i>tal</i> (localisation)
<i>X. citri</i> pv. <i>vignicola</i>	Niébé	CFBP7111	1942 (Etats-Unis)	4	5 710 499	64,5	4 934	0
		CFBP7112	1942 (Etats-Unis)	3	5 155 996	64,7	4 344	7 + 2 ^a (Ch. et Pl.)
		CFBP7113	1966 (Soudan)	1	5 105 205	64,7	4 328	1 (Ch.)
<i>X. citri</i> pv. ' <i>dieffenbachiae</i> '	<i>Dieffenbachia</i>	CFBP3132	1950 (Etats-Unis)	3	5 378 423	64,4	4 676	4 + 2 ^a (Pl.)
<i>X. phaseoli</i> pv. <i>dieffenbachiae</i>	<i>Anthurium</i>	CFBP3133	1965 (Brésil)	1	5 083 349	64,9	4 379	0
<i>X. euvesicatoria</i> pv. <i>alfalfae</i>	Luzerne	CBFP3836R	1996* (Soudan)	2	5 150 212	64,7	4 421	9 (Ch.)
<i>X. citri</i> pv. <i>glycines</i>	Soja	CFBP2526R	1956 (Soudan)	3	5 345 826	64,6	4 688	5 + 4 ^a (Ch. et Pl.)
		CFBP7119R	1981 (Brésil)	3	5 621 420	64,4	4 980	5 + 5 ^a (Ch. et Pl.)

Pl. : Plasmidique

Ch. : Chromosomique

* date de transmission à la CFBP

^a : pseudogène de *tal*

Les souches dont le nom se termine par R sont résistantes à la Rifampicine

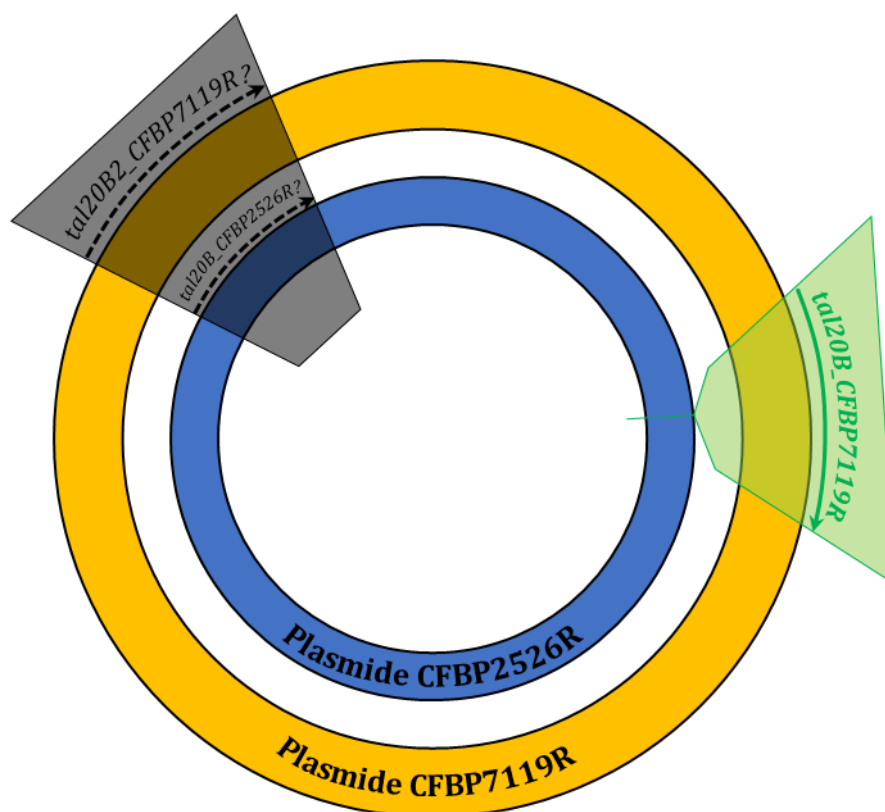


Figure 1 : Schéma du plasmide conservé au sein des souches de *X. citri* pv. *glycines* séquencées. A l'extérieur en jaune est représenté le plasmide de la souche CFBP7119R et à l'intérieur en bleu est représenté celui de la souche CFBP2526R. La région surlignée en vert indique la présence de *tal20B_CFBP7119R*, tandis que sur le plasmide de CFBP2526R elle est absente ce qui est indiqué par le trait vert. La région surlignée en noir indique la présence de pseudogènes de *tal* ayant la combinaison de RVD TAL20B chez les deux souches.

IV. Données et résultats complémentaires

En plus des trois souches de *X. citri* pv. *vignicola* responsables de la brûlure bactérienne du niébé présentées ci-dessus, nous avons séquencé d'autres souches pathogènes de légumineuses, comprenant une souche de *X. euvesicatoria* pv. *alfalfae*, provoquant des taches bactériennes sur luzerne (*Medicago sativa* L.) ; et deux souches de *X. citri* pv. *glycines*, responsable d'une bactériose sur soja (*Glycine max* L.). Egalement, nous avons séquencé une souche de *X. citri* pv. *dieffenbachiae* pathogène sur *Dieffenbachia* et phylogénétiquement proches de *X. citri* pv. *fuscans*, et une souche de *X. phaseoli* pv. *dieffenbachiae* pathogène sur *Anthurium* et phylogénétiquement proche de *X. phaseoli* pv. *phaseoli* (Constantin *et al.*, 2016). De la même façon que pour les souches infectant le haricot, des analyses préliminaires pour étudier les *tal* de ces souches et comparer les distances phylogénétiques de leurs génomes aux distances phylogénétiques de leurs *tal* ont été entreprises

Ces souches ont été séquencées en utilisant le séquençage SMRT avec les mêmes paramètres que précédemment (Ruh *et al.*, 2017b). Les données de séquençage obtenues correspondent à des données de génomiques classiques pour des souches de *Xanthomonas* (Bolot *et al.*, 2013b; Darrasse *et al.*, 2013a; Jacques *et al.*, 2013; Jalan *et al.*, 2013; Gagnevin *et al.*, 2014; Indiana *et al.*, 2014; Aritua *et al.*, 2015; Cesbron *et al.*, 2015), avec pour chaque génome une taille d'environ 5 Mb, un contenu en GC d'environ 64,5 % et environ 4500 gènes prédits. Des analyses par BLASTN ont montré qu'il existait des gènes *tal* chez presque toutes les souches séquencées. En effet, seules deux souches sur les huit séquencées, CFBP7111 de *X. citri* pv. *vignicola* et CFBP3133 de *X. phaseoli* pv. *dieffenbachiae*, ne possédaient aucune séquence correspondant à des *tal*. Pour les autres souches, entre un et sept *tal* ont été trouvés par génome (**Tableau 1**). De plus, treize pseudogènes de *tal* ont été identifiés dans ces génomes : quatre correspondaient à une séquence de *tal* avec un codon STOP prématuré (deux chez CFBP3132, un chez CFBP7119R et un chez CFBP7112), un, retrouvé chez CFBP7112, correspondait à une séquence de *tal* avec une région N-terminale tronquée, tandis que les huit autres, uniquement retrouvés dans les génomes de *X. citri* pv. *glycines*, avaient une longueur totale mais contenaient au moins une délétion d'un nucléotide dans les régions N-terminale ou C-terminale (quatre chez CFBP2526R et quatre chez CFBP7119R). En raison du nombre important de *tal* dans les génomes du pathovar glycines, les analyses de PCR-séquençage permettant de vérifier si ces délétions étaient réelles ou consécutives à des erreurs de séquençage SMRT, comme démontré pour les souches infectant le haricot commun (Ruh *et al.*, 2017a), n'ont pas pu être réalisées. De la même façon que pour les *tal* des souches infectant le haricot (Ruh *et al.*, 2017a), ces *tal* nouvellement séquencés ont été nommés en utilisant leur nombre de répétitions suivies d'une lettre pour différencier les combinaisons de RVD unique portant le même nombre de répétition.

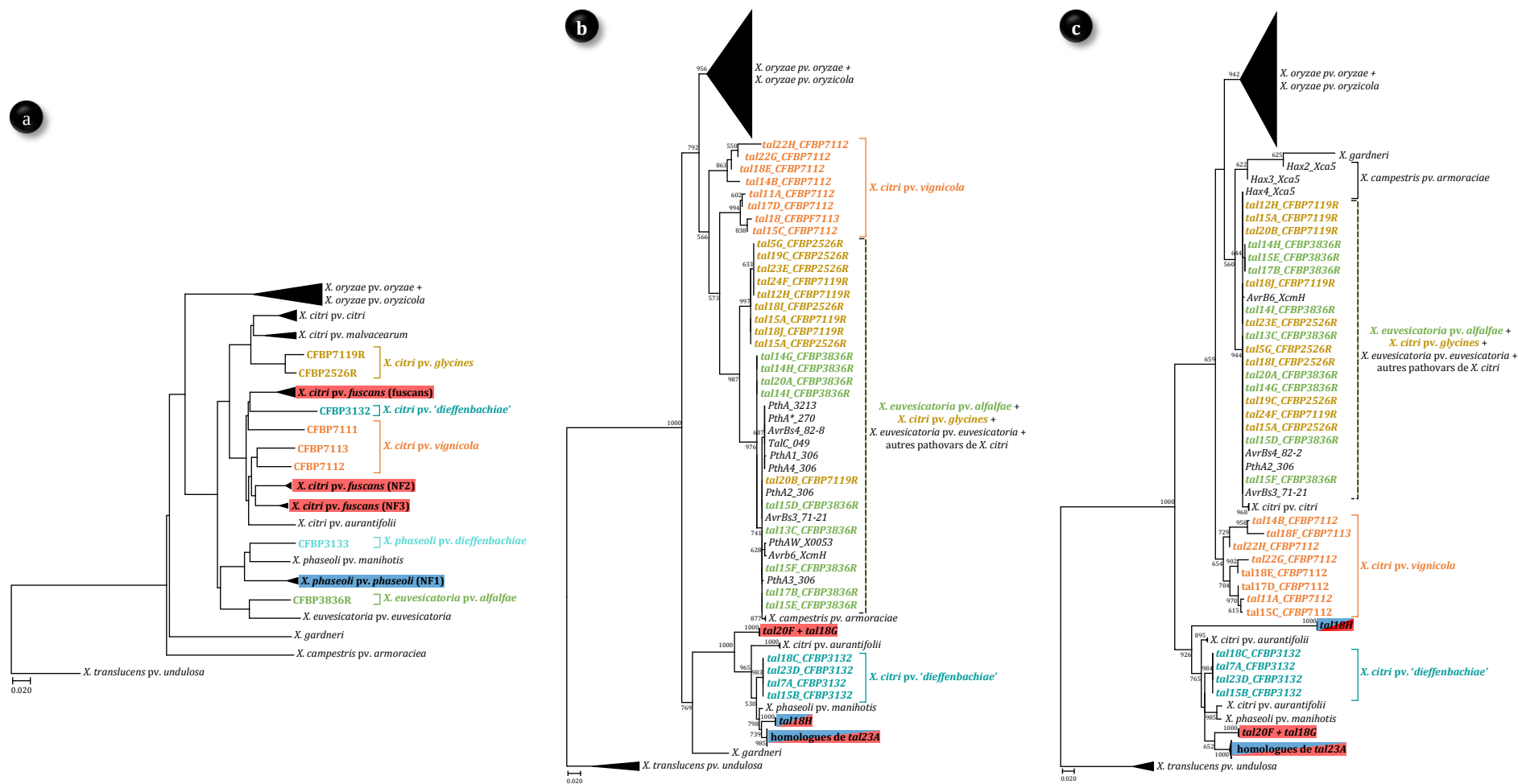


Figure 2 : Phylogénie des souches de *Xanthomonas* et des *tal*. Les souches et les *tal* de *X. phaseoli* pv. *phaseoli* (NF1) et des trois lignées NF2, NF3 et fuscans de *X. citri* pv. *fuscans* sont respectivement surlignées en bleu et en rouge. Les souches et les *tal* de *X. citri* pv. *glycines* sont écrits en jaune, ceux de *X. citri* pv. *vignicola* en orange, *X. euvesicatoria* pv. *alfalfae* en vert et *X. citri* pv. '*dieffenbachiae*' en turquoise foncé et *X. phaseoli* pv. *dieffenbachiae* en turquoise clair. (a) CVtree représentant l'évolution des souches de *Xanthomonas* construites en utilisant l'ensemble du protéome des souches. (b) Arbre ML représentant l'évolution des *tal*, construit en utilisant les séquences nucléotidiques correspondant à la région N-terminale. (c) Arbre ML représentant l'évolution des *tal* construit en utilisant les séquences nucléotidiques correspondant à la région C-terminale. Pour (b) et (c), les valeurs de bootstrap supérieures à 50% sont indiquées pour 1000 répétitions et les échelles représentent le nombre de substitutions de nucléotides par site.

La localisation des *tal* dans ces génomes variait selon le pathovar. En effet, pour *X. citri* pv. '*dieffenbachiae*', tous les *tal* de CFBP3132 étaient plasmidiques. Pour *X. citri* pv. *vignicola*, tous les *tal* étaient également plasmidiques chez la souche CFBP7112 alors que la souche CFBP7113 possédait un *tal* chromosomique et que la souche CFBP7111 ne possédait pas de *tal*. Presque tous les *tal* de la souche de *X. euvesicatoria* pv. *alfalfae* étaient localisés sur une région chromosomique de 60 kb, tandis que le dernier *tal* était également chromosomique mais sur une région plus éloignée. Pour finir, chez les deux souches de *X. citri* pv. *glycines*, quatre *tal* étaient sur le chromosome et un *tal* était sur un plasmide. De ce fait, nous avons voulu déterminer si des plasmides contenant des *tal* étaient conservés entre souches phylogénétiquement éloignées et nous avons alors comparé les séquences de tous les plasmides provenant de toutes les souches que nous avons séquencées, y compris celles infectant le haricot. Seuls deux plasmides provenant respectivement des deux souches de *X. citri* pv. *glycines* présentaient une forte identité nucléotidique moyenne, supérieure à 99,9 % sur la totalité de la longueur. De façon intéressante, il s'agissait de plasmides porteurs de séquences correspondant à des *tal* ou à des pseudogènes de *tal* dont les séquences de RVD étaient identiques. En effet, le plasmide de la souche CFBP2526R était porteur d'un pseudogène de *tal*, dont la combinaison de RVD se nommait TAL20B, et celui de la souche CFBP7119R est porteur d'un *tal* et d'un pseudogène de *tal* avec cette même combinaison de RVD (**Figure 1**).

Les séquences correspondant aux protéomes et aux *tal* de ces souches ont également fait l'objet d'une analyse phylogénétique similaire à celle présentée dans le Chapitre 2 – II. Comme attendu, la topologie du CTVTree (**Figure 2a**) était congruente avec celle d'arbres phylogénétiques de *Xanthomonas* précédemment publiés (Rademaker *et al.*, 2005; Constantin *et al.*, 2016; Jacques *et al.*, 2016; Ruh *et al.*, 2017a). Par ailleurs, les deux arbres phylogénétiques des *tal* avaient une topologie différente de celui des souches et les topologies étaient également différentes entre les arbres des séquences N-terminale et C-terminale (**Figures 2b et 2c**). En particulier, les *tal* de *X. citri* pv. '*dieffenbachiae*' étaient toujours regroupés avec ceux de *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli* et *X. phaseoli* pv. *manihotis* tandis que les *tal* de *X. euvesicatoria* pv. *alfalfae* et *X. citri* pv. *glycines* étaient toujours dans un clade contenant les *tal* de *X. euvesicatoria* et d'autres pathovars de *X. citri*. Selon les séquences N-terminale ou C-terminale, la position des *tal* de *X. citri* pv. *vignicola* était soit postérieure à celle des *tal* des pathovars de *X. oryzae* pour l'arbre N-ter, soit antérieure pour l'arbre C-ter. L'ensemble de ces différences entre l'arbre des souches et l'arbre des *tals* suggérait donc que des événements de recombinaison et / ou de transferts horizontaux de ces gènes avaient eu lieu lors de la diversification des *Xanthomonas*. En effet, les *tal* des pathovars glycines, *alfalfae* étaient regroupés avec ceux provenant de divers pathovars de *X. citri* et *X. euvesicatoria* malgré la diversité des souches qui les portent. A l'inverse, les gènes *tal* de *X. citri* pv. *vignicola* étaient éloignés de ceux

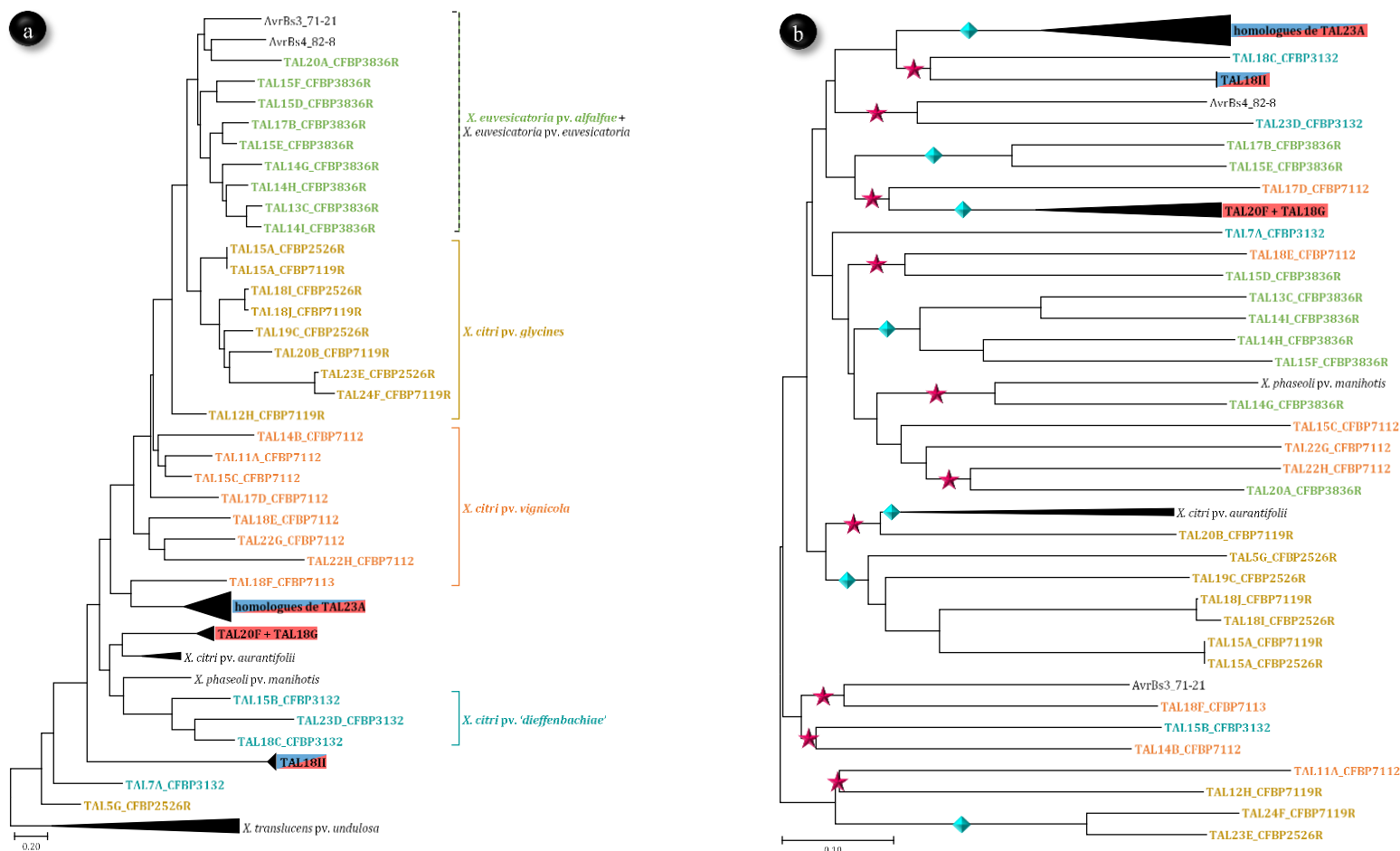


Figure 3 : Arbres représentant les distances des répétitions et des séquences de RVD d'effecteurs TAL. Les effecteurs TAL de la lignée NF1 de *X. phaseoli* pv. *phaseoli* et des trois lignées NF2, NF3 et fuscans de *X. citri* pv. *fuscans* sont respectivement surlignés en bleu et en rouge. Les effecteurs TAL de *X. citri* pv. *glycines* sont écrits en jaune, ceux de *X. citri* pv. *vignicola* en orange, *X. euvesicatoria* pv. *alfalfae* en vert et *X. citri* pv. *'dieffenbachiae'* en turquoise. **(a)** Classification phylogénétique des répétitions construite avec le programme DisTAL (Pérez-Quintero *et al.*, 2015). L'analyse a été effectuée avec les séquences protéiques des répétitions (RVD exclus) utilisant huit effecteurs TAL de la souche XT4699 de *X. translucens* pv. *undulosa* en tant qu'outgroup. **(b)** Arbre phylogénétique représentant la distance entre les séquences de RVD, construit à l'aide du programme FuncTAL (Pérez-Quintero *et al.*, 2015). Les étoiles rose indiquent que des séquences de RVD différentes situées sur cette branche sont issues d'effecteurs TAL provenant de différents pathovars. Les losanges bleus indiquent que des séquences de RVD différentes situées sur cette branche sont issues d'effecteurs TAL provenant du même pathovar.

des agents responsables de la CBB malgré la proximité phylogénétique de ces souches. Au final, cela indique également qu'il n'y a pas d'origine commune des *tals* de souches pathogènes sur légumineuses.

Pour finir, une analyse des répétitions centrales des *tal* similaire à celle présentée dans l'article présentée dans le Chapitre 2 – II a été également entreprise. Nous avons construit deux arbres phylogénétiques avec les outils Distal, pour étudier l'évolution des répétitions sans les RVD, et Functal, pour comparer les spécificités de liaison à l'ADN des effecteurs TAL en utilisant uniquement les séquences de RVD (Pérez-Quintero *et al.*, 2015). De façon intéressante, l'arbre Distal montrait à quelques exceptions près, que les répétitions étaient globalement regroupées selon les pathovars, de façon similaire aux arbres N-ter et C-ter, alors que pour l'arbre Functal cette organisation n'était plus visible (**Figures 3a et 3b**). Certaines combinaisons de RVD provenant d'un même pathovar étaient regroupées sur l'arbre Functal (losanges bleus sur la **Figure 3b**). Ces ressemblances dans les séquences de RVD de différents effecteurs TAL au sein d'un pathovar pourraient suggérer que ces effecteurs ciblent des EBE proches dans le génome d'une même plante hôte, comme cela a été montré pour les effecteurs TalC, Tal5, PthXo3 et AvrXa7 de *X. oryzae* pv. *oryzae* ciblant des EBE chevauchantes situées en amont du gène *OsSWEET14* chez le riz, ou les effecteurs PthB et PthC de *X. citri* pv. *aurantifolii* et PthA^w, PthA⁴ et PthA^{*} de *X. citri* pv. *citri* ciblant des EBE chevauchantes situées en amont du gène *CsLOB1* chez l'orange douce (Yang & White, 2004; Al-Saadi *et al.*, 2007; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013; Hu *et al.*, 2014; Hutin *et al.*, 2015). De façon intéressante, sur l'arbre Functal, il y avait aussi 10 branches présentant des séquences de RVD issues d'effecteurs TAL provenant de différents pathovars (étoiles rouges sur la **Figure 3b**), ce qui pourrait suggérer une convergence fonctionnelle de ces effecteurs TAL. Pour quatre cas, des enchaînements de RVD présentant des similarités évidentes entre effecteurs TAL provenant de pathovars différents ont été trouvés (**Figure 4**). Deux hypothèses peuvent expliquer ces similarités. Soit ces séquences de RVD ont été transmises et conservées au cours de la diversification de ces pathovars, soit ces similarités correspondent à des convergences évolutives entre ces pathovars. Ainsi, ces effecteurs TAL semblent représenter des pistes très intéressantes dans la recherche de cibles d'effecteurs TAL conservées dans différentes plantes d'intérêt agronomique.

Nom	Séquence de RVD	Pathovars (lignée génétique)
TAL17D	HD HD NG <u>N*</u> NI NN NN HD HD HD HD HD NI NG NN HD NN	<i>X. citri</i> pv. <i>vignicola</i>
TAL20F	N* HD HD NG <u>NG</u> NI NG NI NI NN <u>NI</u> HD HD HD HD HD NI HD NK NG	<i>X. citri</i> pv. <i>fuscans</i> (NF2)
TAL18E	NI <u>NG</u> NI NI NN NI HD <u>NN</u> NG NI HD NN NN HD NN NG HD NN	<i>X. citri</i> pv. <i>vignicola</i>
TAL15D	<u>NG</u> NI HD NI <u>NI</u> HD <u>NG</u> NG NI NG NG NI NG NI NG	<i>X. euvesicatoria</i> pv. <i>alfalfae</i>
pthB	NI NG NI <u>NN</u> NI HD NS NS NN NG HD NN NI NG	<i>X. phaseoli</i> pv. <i>manihotis</i>
TAL14G	NI NG NI <u>NI</u> NI NG NI NG NG NG HD HD NS NG	<i>X. euvesicatoria</i> pv. <i>alfalfae</i>
TAL15B	NN NI NI HD <u>NN</u> HD <u>NG</u> HD NN HD NI NG NG HD NG	<i>X. citri</i> pv. ' <i>dieffenbachiae</i> '
TAL14B	NG NI NN NI HD <u>NI</u> HD <u>NN</u> HD HD NI NG NG NG	<i>X. citri</i> pv. <i>vignicola</i>

Figure 4 : Séquences de RVD proches issues d'effecteurs TAL provenant de différents pathovars. Les RVD surlignés en jaune indiquent la partie de la séquence de RVD conservée entre les deux combinaisons de RVD alignées. Les RVD soulignés en rouge indiquent les RVD différents au sein des séquences de RVD conservées.

V. Perspectives et discussion

Au cours de cette thèse, nous avons donc séquencé 25 souches de *Xanthomonas*, ce qui a permis de mettre à jour 85 séquences de *tal* et de pseudogènes de *tal*. Même si le séquençage SMRT nécessite parfois la correction de certains pseudogènes, ces nouvelles données représentent une grande avancée pour l'étude de la diversité et de l'évolution des effecteurs TAL au sein des *Xanthomonas*, mais aussi pour la recherche des cibles des effecteurs TAL dans les génomes des plantes hôtes. Nous avons également montré que le transfert horizontal et la recombinaison sont les deux principales forces motrices de la diversification des *tal* et que le transfert horizontal de *tal* est probablement impliqué dans l'adaptation des *Xanthomonas* au haricot. Avec les nouvelles données présentées dans la quatrième partie de ce chapitre, il semble que des événements de recombinaison et de transferts horizontaux de *tal* se soient également produits entre *X. citri* pv. *vignicola*, *X. citri* pv. *glycines*, *X. citri* pv. '*dieffenbachiae*' et *X. euvesicatoria* pv. *alfalfae*. Il serait donc intéressant d'approfondir les analyses de ces *tal* pour mettre à jour de façon plus précise les différents types d'événements évolutifs mis en jeu ainsi que les temps de divergence des *tal*. Par ailleurs, nous avons noté que chez les *Xanthomonas* il existait deux répertoires différents de *tal*. Le premier répertoire correspond à l'accumulation d'un nombre important de *tal* chromosomiques, comme c'est le cas chez les pathovars de *X. oryzae* (Lee *et al.*, 2005; Ochiai *et al.*, 2005; Salzberg *et al.*, 2008; Wilkins *et al.*, 2015; Huguet-Tapia *et al.*, 2016). Il semblerait que ceci soit la conséquence d'une longue coévolution entre les effecteurs TAL et le génome de la plante hôte. Le second répertoire correspond à un nombre moins important de *tal* plasmidiques, comme par exemple chez les pathovars de *X. citri* (da Silva *et al.*, 2002). Il semblerait que ceci soit la conséquence d'une évolution plus courte sans coévolution entre les effecteurs TAL et le génome hôte. La localisation plasmidique des *tal* jouerait également un rôle dans la diffusion de ces gènes au sein des *Xanthomonas*, ce qui pourrait conduire à la modification rapide des répertoires d'effecteurs TAL et pourraient participer à l'acquisition de nouveaux traits pour l'adaptation à l'hôte. Avec les nouvelles données de séquençage, l'existence de deux types de répertoires pouvant correspondre à des histoires évolutives distinctes semble être confirmée ainsi que le fait que le plasmide représente un vecteur important de diffusion des *tal*. Ces histoires évolutives seraient également commune aux pathovars de *X. citri* et *X. euvesicatoria*, pour lesquels les souches possèdent des *tal* chromosomiques, des *tal* plasmidiques, voire les deux à la fois.

D'autre part, les souches pathogènes sur *Anthurium* et *Dieffenbachiae* ont une histoire évolutive liée à celle des souches infectant le haricot. En effet, la souche CFBP3132 groupe avec la lignée fuscans et la souche CFBP3133 groupe avec la lignée NF1 et les *tal* de *X. citri* pv. '*dieffenbachiae*' groupent avec les *tal* des souches responsables de CBB. Les souches de *X. citri* pv. *vignicola* infectant le niébé sont phylogénétiquement proches des souches de *X. citri* pv. *fuscans*.

Cependant, contrairement aux souches de *X. citri* pv. '*dieffenbachiae*', leurs *tal* ne groupent jamais avec ceux des souches infectant le haricot. Cela signifie donc que malgré leur proximité phylogénétique, les *tal* des souches infectant le haricot et infectant le niébé ont été acquis par le biais d'évènements indépendants. Cependant, la similarité de séquences de RVD entre TAL17D et TAL20F provenant respectivement de *X. citri* pv. *vignicola* et de la lignée NF2 de *X. citri* pv. *fuscans* (**Figure 4**) suggère une possible convergence fonctionnelle entre ces pathovars. Ainsi, les cibles de ces effecteurs TAL mériteraient une étude approfondie afin de déterminer s'il s'agit d'une cible des effecteurs TAL conservée entre le haricot et le niébé, deux légumineuses phylogénétiquement proches. Plus généralement, il serait intéressant d'étudier plus en profondeur les cibles des effecteurs des divers pathovars de *Xanthomonas* afin de trouver d'autres cibles conservées dans différentes plantes d'intérêt agronomique, qui constitueraient des pistes pour le développement de résistances durables contre les maladies induites par ces bactéries.

CHAPITRE 3 :

**TAL18H : un effecteur TAL atypique partagé
entre des lignées phylogénétiquement éloignées
de *Xcf-Xpp***

I. Introduction

L'article "*Xanthomonas* adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors" présenté dans le Chapitre 2-II a mis en évidence l'existence de *tal18H*, un gène *tal* transmis par HGT entre deux lignées phylogénétiquement distantes : les lignées fuscans de *Xcf* et NF1 de *Xpp*. L'objectif de ce chapitre est de décrire le rôle de TAL18H dans l'interaction entre les agents responsables de la CBB et le haricot commun.

L'analyse de la séquence protéique de TAL18H a permis de découvrir quelques caractéristiques peu communes voire spécifiques à cet effecteur TAL. En effet, TAL18H possède à deux reprises un RVD peu commun, le RVD HY, et les séquences des répétitions portant ce RVD commencent de façon inhabituelle, les quatre premiers acides aminés étant MTLA au lieu de LTPE ou LTPD dans la majorité des autres effecteurs TAL décrits à ce jour. Afin de déterminer la spécificité de liaison du RVD HY et l'impact du motif MTLA sur cette spécificité de liaison, une étude approfondie de ce RVD et de ce motif a été entreprise. Pour ce faire, je suis partie durant trois mois dans l'équipe de Jens Boch, à l'institut de génétique végétale de l'université Leibniz à Hanovre en Allemagne, au sein de laquelle j'ai été guidée et épaulée par Sebastian Becker, Jens Boch et Jana Streubel afin d'acquérir les compétences nécessaires à cette étude.

La caractérisation du rôle de TAL18H dans l'interaction *Xanthomonas* - haricot commun demandait la production d'un mutant de ce *tal*. Dans ce but, de multiples tentatives de construction d'un mutant de délétion de *tal18H* dans la souche CFBP6546R (*Xpp* – NF1) ont été entreprises, sous les conseils d'Armelle Darrasse, mais sont restées sans succès. Cependant, au cours de ces essais un mutant de délétion de *tal18H* spontané est apparu. Ainsi, dans le cadre du stage M1 de Justine Foucher que j'ai encadrée, et avec l'aide d'Anne Préveaux et de Céline Rousseau, des expérimentations *in planta* ont été menées afin de déterminer l'impact de TAL18H dans la pathogénie de la souche et aussi sur le transcriptome du haricot. Par la suite, des données RNAseq ont donc été produites par la plateforme Get-PlaGe du centre INRA de Castanet-Tolosan et ont été traitées par Sébastien Carrère avant d'être analysées par Justine Foucher et moi-même avec l'aide de Sandra Pelletier et Martial Briand. La comparaison de ces données de transcriptomique avec des données de prédictions des cibles de TAL18H, réalisées avec l'aide d'Alvaro L. Pérez-Quintero et de Nicolas Chen, nous a permis de sélectionner quelques gènes potentiellement ciblés par TAL18H afin d'entamer des validations fonctionnelles de ces cibles. Une partie de ces analyses sont présentées dans la deuxième partie de ce chapitre sous la forme d'une ébauche d'article dont je suis le premier auteur et qui, après des expériences et des analyses complémentaires, pourrait être soumis à la revue *New Phytologist*.

Par ailleurs, des tests de pouvoir pathogène et analyses RNAseq ont été réalisées sur deux génotypes de haricot représentatifs des deux centres de diversité majeurs du haricot : JaloEEP558 (andin) et BAT93 (mésaméricain), afin de détecter d'éventuelles résistances du haricot à TAL18H. Des graines de ces deux variétés, nous ont été préalablement fournies par Valérie Geffroy. De plus, les génomes de ces deux variétés ont été séquencés et assemblés à la plateforme Get-PlaGe du centre INRA de Castanet-Tolosan, et annotés par Sébastien Carrère.

Dans l'optique de pouvoir étudier l'ensemble des *tal* retrouvés dans les souches responsables de la CBB, j'ai également produit un second mutant de délétion dépourvu de *tal* dans la souche CFBP6165R (*Xcf* – fuscans) avec les conseils d'Armelle Darrasse et de Lionel Gagnevin. Ce mutant de délétion a été complémenté par Sophie Bonneau, avec des *ArtTAL* que j'ai construit et des *tal* naturels que Sophie Bonneau a clonés. J'ai acquis les compétences et le matériel nécessaire pour la construction des *ArtTAL* durant un séjour de trois semaines dans l'équipe de Boris Szurek et sous les conseils de Tu Tran Tuan, puis je les ai complétés lors de mon séjour dans l'équipe de Jens Boch, où j'ai construit l'ensemble des *ArtTAL* correspondant à ceux retrouvés dans les souches de *Xcf-Xpp*. Le protocole de clonage des *tal* nous a été initialement fourni par Tu Tran Tuan puis a fait l'objet de nombreuses mises au point réalisées par Sophie Bonneau, Ivanna Fuentes et Sandrine Fabre dans le cadre du projet ANR CropTAL. Ainsi, lors de son stage M2 encadré par Nicolas Chen, Justine Foucher, épaulée d'Anne Préveaux, de Sophie Bonneau et de moi-même, a entamé des expérimentations *in planta* similaires à celles précédemment décrites, afin de déterminer l'impact de tous les effecteurs TAL de *Xcf-Xpp* dans la pathogénie de la souche et sur le transcriptome du haricot.

La démarche à suivre pour compléter l'ébauche d'article ainsi que les analyses complémentaires mentionnées ci-dessus sont présentées dans la troisième partie de ce chapitre.

II. Deciphering the role of TAL18H effector in the interaction between the common bacterial blight agents and common bean

Mylène Ruh¹, Justine Foucher¹, Anne Préveaux¹, Sébastien Carrère², Martial Briand¹, Sebastian Becker³, Jens Boch³, Marie-Agnès Jacques¹ and Nicolas W.G. Chen^{1*}

¹IRHS, INRA, AGROCAMPUS OUEST, Université d'Angers, SFR4207 QUASAV, Beaucouzé, France

²CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR 2594, Castanet-Tolosan, France

³Institut für Pflanzengenetik, Leibniz Universität Hannover, Hannover, Germany

* Corresponding author

E-mail addresses of all authors:

Mylène Ruh: mylene.ruh@agrocampus-ouest.fr

Justine Foucher: justine.foucher@inra.fr

Anne Préveaux: anne.preveaux@inra.fr

Sébastien Carrère: sebastien.carrere@toulouse.inra.fr

Martial Briand: martial.briand@inra.fr

Sebastian Becker: sebastian.beckergenetik.uni-hannover.de

Jens Boch: jens.boch@genetik.uni-hannover.de

Marie-Agnès Jacques: marie-agnes.jacques@inra.fr

Nicolas W.G. Chen: nicolas.chen@agrocampus-ouest.fr

SUMMARY

Plant pathogenic bacteria from the genus *Xanthomonas* utilize transcription activator-like (TAL) effectors to activate host genes to promote infection. TAL effector specificity resides in tandem repeats of 33 to 35 amino acid from which residues 12 and 13, also called the repeat variable diresidue (RVD), are involved in the specific binding of the TAL effector to DNA. TAL18H is a TAL effector shared among strains from *Xanthomonas phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans*, the causal agents of common bacterial blight of bean. TAL18H presents atypical repeats starting with an MTLA motif, and carrying the unusual HY RVD. We characterized the binding specificity of HY and the impact of the MTLA motif in HY binding efficiency. To analyse the role of TAL18H *in planta*, we produced a *tal18H* deletion mutant and carried out comparative analyses of transcripts as well as phenotyping experiments in common bean plants inoculated with bacterial strains carrying *tal18H* or not. The HY RVD bound preferentially to adenine and the

MTLA motif increased the efficiency of this binding. TAL18H led to higher aggressiveness of the strain and impacted the transcription levels of numerous common bean genes. Most over-expressed genes corresponded to functions linked to the cell periphery, including three pectate lyases, which have been recently described as TAL effectors targets. Interestingly, numerous under-expressed genes encoded functions potentially involved in plant immunity. Together, our results show that TAL effector repeat structure can modify RVD binding efficiency to DNA, and suggest that TAL18H is able to promote infection by both inducing susceptibility genes and lowering common bean defences.

Key words: common bacterial blight (CBB) agents, common bean, pathogenicity, TAL effector target, Transcription Activator-Like (TAL) effector, transcriptome, *Xanthomonas*

INTRODUCTION

Bacterial pathogens from the genus *Xanthomonas* infect many economically important crops such as rice, soybean, cassava, citrus and bean and are responsible for significant yield losses worldwide (Hayward, 1993; Ryan *et al.*, 2011). Individually, the host range of each *Xanthomonas* strain is usually restricted to a single plant species, or several plant species belonging to the same botanical family (Vauterin *et al.*, 1995). Host specialization is a complex process apparently defined by a repertoire of genes including type III effectors (T3E) (Hajri *et al.*, 2009, 2012b; Mhedbi-Hajri *et al.*, 2011; Schwartz *et al.*, 2015; Jacques *et al.*, 2016). Among the T3E found in *Xanthomonas*, transcription activator-like (TAL) effectors are able to mimic eukaryotic transcription factors in the plant cells (Bonas *et al.*, 1989; Heuer *et al.*, 2007; Boch & Bonas, 2010; de Lange *et al.*, 2014b). These effectors have the ability to enter the plant nucleus, then wrap around DNA and bind specific sequences called Effector Binding Elements (EBE) usually located in the promoter region of targeted genes (Mak *et al.*, 2012, 2013). The central region of TAL effectors consists of a variable number of tandem repeats of 33 to 35 highly conserved amino acids, and is flanked by a conserved N-terminal domain containing a type III secretion signal, and a conserved C-terminal domain containing nuclear-localization signals and an acidic activation domain. The two amino acids on positions 12 and 13 of each repeat are highly polymorphic and referred to as the repeat variable diresidue (RVD). Each RVD binds more or less specifically to one or few nucleotides, and the RVD sequence of a given TAL effector determines its binding specificity to a given EBE. It was shown that the residue at position 12 of the repeat is involved in the linkage stability, while the residue at position 13 is responsible for DNA binding specificity (Mak *et al.*, 2012). To analyse the binding specificities and efficiencies of RVD, artificial TAL effectors (ArtTAL)

with predictable DNA specificities have been constructed based on the modular architecture of TAL effector repeats (Boch *et al.*, 2009; Cermak *et al.*, 2011; Christian *et al.*, 2012; Cong *et al.*, 2012; Streubel *et al.*, 2012). ArtTAL are also used to validate candidate host genes as bona fide virulence targets for *Xanthomonas* (Streubel *et al.*, 2013; Hutin *et al.*, 2015; Blanvillain-Baufumé *et al.*, 2017; Wang *et al.*, 2017). Breaking the TAL effector code linking each RVD to its binding specificity has enabled the development of programs able to scan the host genome for *in silico* prediction of EBE (Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Hummel *et al.*, 2012; Grau *et al.*, 2013; Pérez-Quintero *et al.*, 2013; Yang *et al.*, 2014).

A few targets of TAL effectors have been characterized and EBE are usually found upstream of susceptibility genes, which promotes bacterial infection. For example, TAL effectors from diverse *Xanthomonas* pathogens target genes encoding sugar transporters in rice, citrus, cassava or pepper, which would allow the bacteria to have access to a nutritive source favouring their multiplication and thus bacterial infection (Yang & White, 2004; Yang *et al.*, 2006; Kay *et al.*, 2009; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013; Cohn *et al.*, 2014; Hu *et al.*, 2014; Zhou *et al.*, 2015). Other TAL effectors target genes encoding a sulphate transporter, that would modify the redox potential favourably for symptoms establishment (Cernadas *et al.*, 2014); a dioxygenase, that would indirectly promote cell proliferation via an inhibition of cinnamoyl-CoA reductases activity (Abe & Benedetti, 2016); or that are involved in microRNA stability (Moscou & Bogdanove, 2009; Pérez-Quintero *et al.*, 2013) have been described. Transcription factors can also be overexpressed by TAL effector, but their secondary targets remain to be determined (Al-Saadi *et al.*, 2007; Kay *et al.*, 2007; Sugio *et al.*, 2007; Hu *et al.*, 2014; Abe & Benedetti, 2016; Schwartz *et al.*, 2017).

In some cases, EBE are found in the promoter region of resistance genes, which promotes the defence reaction of the host plant. For example, induction of *Xa27* by *AvrXa27* from *X. oryzae* pv. *oryzae* promotes resistance in some rice varieties (Gu *et al.*, 2005). Similarly, *AvrBs3* from *X. euvesicatoria* pv. *euvesicatoria* activates the expression of the resistance gene *Bs3* in pepper (Römer *et al.*, 2007). Promising strategies have been proposed for exploiting the TAL effector-EBE system to develop disease resistance in crops (Schornack *et al.*, 2013; Boch *et al.*, 2014). For example in rice, mutations of EBE by genome editing can prevent the specific binding of TAL effectors and lead to a loss of susceptibility to *X. oryzae* pv. *oryzae* (Li *et al.*, 2012b; Blanvillain-Baufumé *et al.*, 2017). Likewise, placing EBE in the promoter region of resistance genes is an efficient way to promote defence reactions against *X. oryzae* pv. *oryzae* or *X. oryzae* pv. *oryzicola* strains bearing corresponding TAL effectors (Hummel *et al.*, 2012, 2016; Wang *et al.*, 2015; Zeng *et al.*, 2015; Zhang *et al.*, 2015b; Hutin *et al.*, 2016).

Common bacterial blight of bean (CBB) is the most devastating bacterial disease of common bean (*Phaseolus vulgaris*), causing significant yield losses everywhere common bean is

cultivated (Mutlu *et al.*, 2008; Belete & Bastas, 2017). CBB is due to different genetic lineages belonging to *X. phaseoli* pv. *phaseoli* (*Xpp*) and *X. citri* pv. *fuscans* (*Xcf*) (Alavi *et al.*, 2008; EFSA PLH Panel (EFSA Panel on Plant Health), 2014; Constantin *et al.*, 2016). Although *X. phaseoli* pv. *phaseoli* and *X. citri* pv. *fuscans* strains are phylogenetically distant, they both induce similar symptoms on common bean. This pathological convergence is linked to large events of horizontal gene transfers (Aritua *et al.*, 2015). Two TAL effector-encoding genes (*tal*) *tal23A* and *tal18H* have been horizontally transferred and maintained among *Xpp* and *Xcf* (Ruh *et al.*, 2017a). In particular, the RVD sequence from TAL18H was conserved in *Xcf* and *Xpp* strains, suggesting that this TAL effector may be important for the adaptation of *Xanthomonas* to common bean, and thus for disease development.

In this study, we describe the atypical structure of TAL18H, a TAL effector bearing the uncommon HY RVD carried by atypical repeats starting with an MTLA motif. We constructed ArtTAL effectors to characterize the binding specificity of the HY RVD and to identify the role of the MTLA motif. To study the impact of TAL18H during common bean infection, we performed phenotyping experiments and comparative analyses of transcripts in common bean plants inoculated with bacterial strains carrying *tal18H* or not.

MATERIALS AND METHODS

Plant material, bacterial strains and growing conditions

Common bean seeds from genotype JaloEEP558 were individually seeded in plastic pots (7 × 7 × 8 cm) containing pre-wetted compost (NEUHAUS HUMINSUBSTRAT *N4*, NFU 44–551). Plants were grown in a controlled climatic room under 80% relative humidity at 23°C/20°C (day/night) with a photoperiod of 16 h. relative humidity was increased at 95% the day before inoculation and maintained at 95% for the duration of the pathogenicity assay. Plants were watered every two days with water supplemented with N-P-K (15-10-30) at 1.5 g/L. *Nicotiana benthamiana* plants were grown in the greenhouse under 60% relative humidity with at 23°C/19°C (day/night) with a photoperiod of 16 h.

Bacterial strains used for this study are listed in **Supplementary Material Table S1**. *Xanthomonas* strains were grown 48 h at 28°C on TSA medium (17.0 g/L pancreatic digest of casein; 3.0 g/L enzymatic digest of soya bean; 5.0 g/L NaCl; 2.5 g/L K₂HPO₄; 2.5 g/L glucose; 15 g/L agar), then 24 h at 28°C on a 1/10 dilution of TSA (TSA10) medium to obtain fresh grown bacteria. Rifamycin was added at a final concentration of 50 mg/L. All experiments with *Xanthomonas* strains were carried out under quarantine conditions at UMR1345 IRHS, Centre INRA, Beaucouzé, France. *Agrobacterium* strains were grown 48 h at 28°C on YEB medium (5.0

g/L peptone of meat; 5.0 g/L peptone of casein; 5.0 g/L sucrose; 1.0 g/L yeast extract; 15 g/L agar), then 24 h on YEB medium to obtain fresh grown bacteria. When needed, rifampicin, spectinomycin or kanamycin were added at a final concentration of 100 mg/L.

For bacterial growth control, the absorbance (λ : 600nm) of liquid cultures at initial concentration of 1×10^6 CFU.ml⁻¹ of strains CFBP6546R and CFBP6546R Δ *tal18H* in TSA10 broth in 96-well plate incubated at 28°C under shaking at 200 rpm was monitored every 30 min over a period of 46.5 h using a spectrophotometer.

Genome sequencing

Genomic DNA of strain CFBP6546R Δ *tal18H* was extracted with The Wizard® Genomic DNA Purification Kit (Promega) according to manufacturer's recommendations. The whole genome was sequenced at Icahn School of Medicine at Mount Sinai (New York, NY) using PacBio Single Molecule Real Time (SMRT) sequencing. One SMRT cell was used to achieve ~100× coverage (Eid *et al.*, 2009). De novo assembling was performed using HGAP assembler version 3.0 (Chin *et al.*, 2013; Pacific Biosciences, Menlo Park, CA). Annotation of the whole genome assembly was performed with EuGene-PP version 1.2 automated pipeline (Sallet *et al.*, 2014), using SWISS-PROT as protein and training protein databases (<http://www.uniprot.org/>).

Pathogenicity assays

Inoculations were performed by bathing the first trifoliate bean leaves of two weeks old plants for 30 seconds into bacterial suspensions calibrated at 1×10^7 CFU/mL in sterile distilled H₂O. At 2 h, 1, 5, 8, and 15 days post-inoculation (dpi), at least three trifoliate leaves by conditions were harvested. Each trifoliate leaf was ground individually (Stomacher 80; Seward, London, United Kingdom) for 2 min at maximum power in 10 mL of distilled H₂O. Samples and appropriate dilutions were plated on TSA10 medium supplemented with rifamycin and incubated for 72 h at 28°C before counting of colony-forming units (cfu). Symptoms development was monitored 8 and 15 dpi by chlorophyll fluorescence imaging (Rousseau *et al.*, 2013) at the PHENOTIC Seeds and Plants platform of the IRHS of Angers. Each leaflet was set in the dark for 30 min. Then, a first picture was taken under a modulated flash of light to measure basal fluorescence (F_0) of the tissues, followed by another picture taken under a high flash of saturating light to measure the maximum fluorescence emission level (F_m). For each pixel, the maximum quantum yield of photosystem II photochemistry ($F_v/F_m = (F_m - F_0)/F_m$) was calculated in order to cluster them according to the intensity of the symptoms (Rousseau *et al.*, 2015).

RNAseq experiments

First trifoliate leaflets of two weeks old common bean were detached from the plant and vacuum-infiltrated twice for 1 min into bacterial suspensions calibrated at 1×10^8 CFU/mL in sterile distilled H₂O. The infiltrated leaflets were maintained in survival in Petri dishes by dipping the petiole in water agar (0.7 % agar) and incubated at 28°C/25°C (day/night) with a photoperiod of 16 h. At 2 h, 1 and 2 dpi, at least, nine leaflets by conditions were sampled and distributed in three samples for microbial and transcriptional analysis. For microbial analysis, at least, two disks were cut per sample at each dpi with a 1.3 cm borer and ground (Stomacher 80; Seward, London, United Kingdom) for 2 min at maximum power in 2.5 ml of distilled H₂O. Samples and appropriate dilutions were plated on TSA10 medium supplemented with rifamycin, incubated 72 h at 28°C and cfu were counted. For transcriptional analysis, at least, ten disks were cut per sample at 1 and 2 dpi with a 1.3 cm borer. The disks were distributed in two eppendorf tubes to form biological duplicates and were liquid nitrogen-frozen and stored them at -80°C.

Frozen disks were ground in a mixer mill for 30 seconds at 25 Hz. RNA was extracted with the TRIzol® Plus RNA Purification Kit (Ambion) according to manufacturer's recommendations, except that TRIzol® was used at 60°C and that an additional DNase treatment was performed. The quantity and the quality of DNA-free total RNA was evaluated using a NanoDrop™ (Thermo Scientific) and an Experion™ chip (Bio-Rad). Preparation of libraries and paired-end (2 x 150 base pair) RNA sequencing (RNA-seq) were performed at Get-PlaGe, Centre INRA, Castanet-Tolosan, France. Libraries were prepared using the Illumina TruSeq Stranded mRNA Kit and RNA-seq was performed on an Illumina HiSeq3000 platform. Around 17 million reads per library were produced on average, and total reads were mapped on the genome sequence of common bean genotype JaloEEP558 (data not shown). Differentially expressed genes were determined using DeSeq2 with default parameters (Love *et al.*, 2014), and Benjamini-Hochberg (BH) adjusted p-value was calculated (Benjamini & Hochberg, 1995). All genes with BH-adjusted p-value less than or equal to 0.05 were taken as being differentially expressed, regardless of fold change. The functions of differentially expressed common bean genes were assessed based on BLASTP analyses on the *Arabidopsis thaliana* genome available at TAIR (<https://www.arabidopsis.org>).

GUS assays

Primers, constructions, bacterial strains and plasmids produced and used for GUS assays are described in in **Supplementary Materials Tables S1, S2 and S3**. Repeats with LTPE or MTLA motifs carrying the HY RVD were produced from LTPE repeats carrying the HD RVD by PCR with specific primers for modifying LTPE into MTLA, or HD into HY, respectively. PCR products were inserted in pUC57 plasmid, then verified by BpiI digestion coupled with SANGER sequencing. ArtTAL were constructed into pSKA2, an ENTRY vector containing a Hax3 backbone, as previously

described (Boch *et al.*, 2009). Constructs were verified by NdeI and Hind III digestion coupled with SANGER sequencing, then transformed by electroporation into *A. tumefaciens* (Supplementary Material Table S2). For β -glucuronidase (*GUS*) assays, different DNA sequences corresponding to the EBE of each ArtTAL and allowing to define HY binding specificities were inserted upstream the minimal *Bs4* promoter in front of the *GUS* gene. The whole constructs were inserted into the pGWB3 plasmid, then transformed by electroporation into *A. tumefaciens* (Nakagawa *et al.*, 2007; Streubel *et al.*, 2012). Five to seven weeks old *N. benthamiana* leaves were inoculated using a needleless syringe with a (1:1) mixture of *A. tumefaciens* ($OD_{600nm} = 0.8$) containing the pSKA2 or pGWB3 constructs suspended in AIM (Agrobacterium Infiltration Medium) buffer (0.01 M MES; 0.01 M MgCl₂; 0.15 mg/mL acetosyringone). At 2 dpi, three leaf disks per inoculated spot were cut with a 0.9 cm borer. One disk was used for qualitative GUS assay and the two other disks were used for quantitative GUS assay. For qualitative GUS assay, leaf disks were stained in GUS staining solution overnight at 37°C (10 mM sodium phosphate buffer pH 7.0; 10 mM EDTA pH 8.0; 1 mM potassium ferricyanide; 1 mM potassium ferrocyanide; 0.1% TritonTM X-100; 0.1% X-Gluc), then destained in ethanol at 96% and dried between cellophane sheets as previously described (Schornack *et al.*, 2005). For quantitative GUS assays, leaf disks were placed in liquid nitrogen before storage at -80°C. Frozen disks were ground with two grinding balls in a mixer mill for 40 seconds at 30 Hz. Proteins were extracted with 300 μ l of GUS extraction buffer (50 mM sodium phosphate buffer pH 7.0; 10 mM EDTA pH 8.0; 0.1% SDS; 0.1% TritonTM X-100; 10 mM β -mercaptoethanol). Samples were centrifuged for 5 min at 14,000 rpm and 10 μ l of supernatants were incubated in 90 μ l of MUG (4-Methylumbelliferyl- β -D-Glucuronide Trihydrate) buffer (50 mM sodium phosphate buffer pH 7.0; 10 mM EDTA pH 8.0; 0.1% SDS; 0.1% TritonTM X-100; 10 mM β -mercaptoethanol; 5 mM MUG) from 60 to 90 min at 37°C (Kay *et al.*, 2007). The GUS activity was stopped with a solution of Na₂CO₃ at 0.2 M then quantified by measuring fluorescence with excitation and emission wavelengths of 360 nm and 460 nm respectively, an integration time of 40 μ s and three excitation flashes. Proteins were quantified using Bradford assay. All assays were carried out twice and triplicate samples from different plants were combined into one data point.

RESULTS

Description of the CFBP6546R Δ tal18H mutant strain

In order to analyse the role of TAL18H, we used the *tal18H* mutant strain CFBP6546R Δ tal18H corresponding to a variant of the *Xcf* strain CFBP6546R presenting a spontaneous deletion of *tal18H_CFBP6546R* gene. The genome sequence of strain CFBP6546R Δ tal18H was obtained through PacBio SMRT sequencing and compared to the

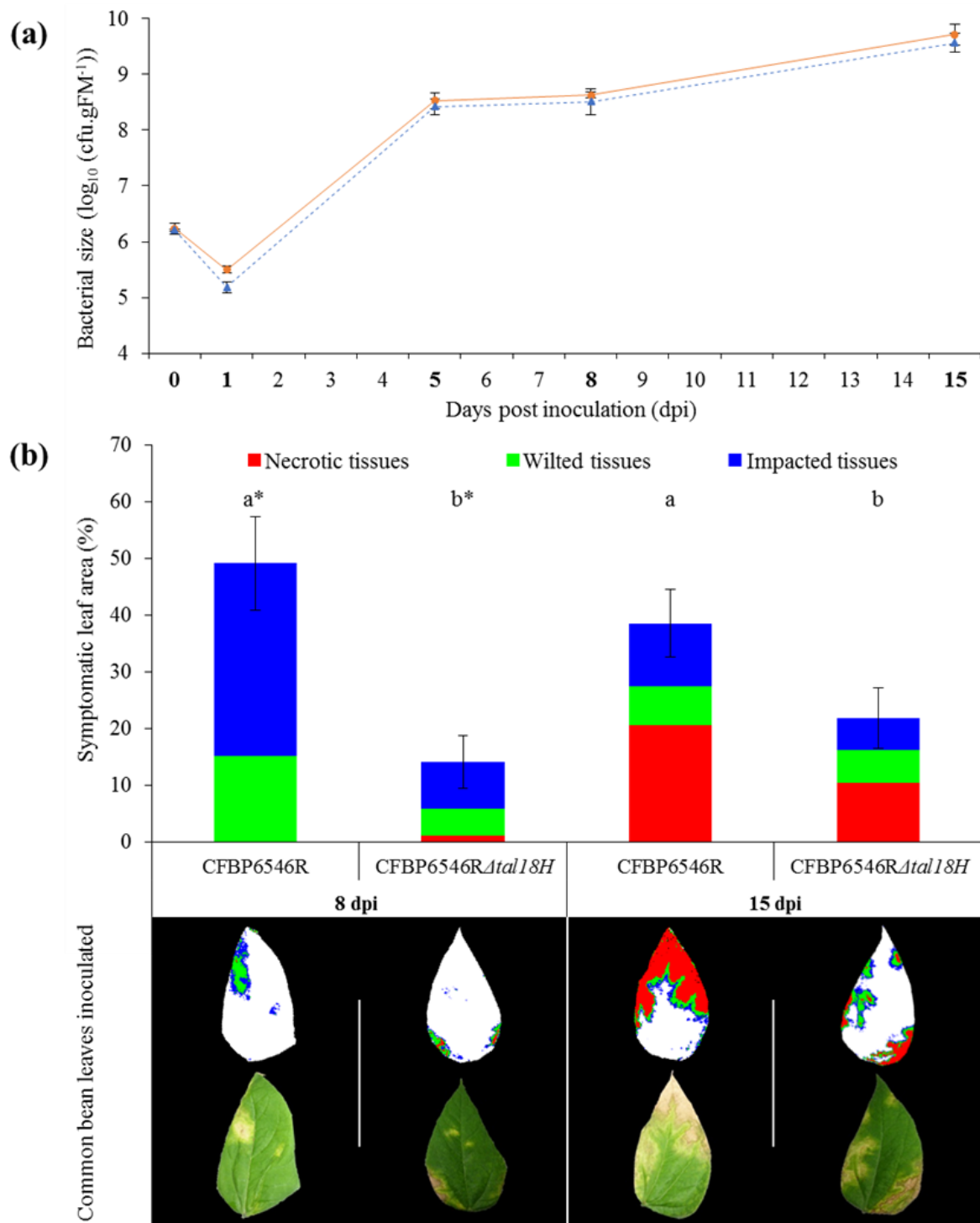


Fig. 1 Impact of *tal18H* in planta. (a) Bacterial population dynamics after inoculation of strains CFBP6546R (orange) and CFBP6546RΔ*tal18H* (blue) in common bean leaves. The bacterial population sizes are expressed in log₁₀ (cfu.gFM⁻¹) (gMF = gram of fresh material). (b) Symptomatic leaf area of common bean after inoculation with CFBP6546R and CFBP6546RΔ*tal18H* strains. The percentage of symptomatic leaf area was calculated by calculating the ratio of the pixel number associated with symptoms to the pixel number associated with the total leaf area. Under the histogram, examples of the leaves visualized by chlorophyll fluorescence imaging (top) and the visible image (bottom) in the corresponding conditions. Error bars represent the standard errors of the means. Different letters indicate that differences within sampling date are significant based on the Mann–Whitney test (p-value < 0.05, *p-value < 0.01).

sequence of the wild-type (WT) strain CFBP6546R (Ruh *et al.*, 2017a). In both strains, the whole genome sequence comprised four contigs corresponding to the chromosome plus three circularized plasmids. For strain CFBP6546R Δ *tal18H*, total assembly size was 5,223,945 bp, with a G+C content of 64.82 % and annotation predicted a total of 4,685 CDS. The average nucleotide identity between CFBP6546R and CFBP6546R Δ *tal18H* was above 99.9% across the whole genomes. In strain CFBP6546R, plasmid C contained *tal18H_CFBP6546R* gene as well as a pseudogenized paralog of *tal18H_CFBP6546R* named *tal18H_CFBP6546R** (Ruh *et al.*, 2017a). In CFBP6546R Δ *tal18H*, two duplicated regions of nearly 5,400 bp including *tal18H_CFBP6546R* and *tal18H_CFBP6546R** were deleted (**Supplementary Material Fig. S1**). Each region contained *tal18H_CFBP6546R* or *tal18H_CFBP6546R** plus four CDS encoding hypothetical proteins. These two deleted regions were flanked by insertion sequences (IS), suggesting that these deletions occurred through recombination between IS. These deletions had no major effect on the bacterial growth of either CFBP6546R or CFBP6546R Δ *tal18H* strains in liquid medium (**Supplementary Material Fig. S2**).

CFBP6546R Δ *tal18H* mutant presented reduced CBB symptoms in planta

To characterise the role of TAL18H in CBB, we characterized the pathogenicity of strain CFBP6546R Δ *tal18H*. For this, the dynamics of bacterial population sizes were evaluated over time after inoculation of the wild type and mutant strains on common bean leaves. Both bacterial dynamics were similar, with a slight decrease of the population size at 1 dpi followed by exponential growth until a maximum of around 9.6 log (cfu / gFM) at 15 dpi (**Fig. 1a**). These dynamics are typical of a compatible interaction, as previously described for *Xcf* strain CFBP4834-R on *P. vulgaris* cv. Flaver (Darsonval *et al.*, 2009). At 8 and 15 dpi, the percentage of symptomatic leaf area was quantified by chlorophyll fluorescence imaging on the same leaves than those used for bacterial dynamics (**Fig. 1b**). Interestingly, strain CFBP6546R Δ *tal18H* induced significantly less symptoms than strain CFBP6546R at both 8 and 15 dpi. Indeed, at 8 dpi, the symptomatic leaf areas were 49.89 % and 14.22 % for strains CFBP6546R and CFBP6546R Δ *tal18H*, respectively. Differences between both strains were slightly smaller at 15 dpi, with symptomatic leaf areas of 38.61 % for strain CFBP6546R and 22.94 % for strain CFBP6546R Δ *tal18H*. For both strains, severe symptoms were almost absent at 8 dpi, while they represented around half of the symptomatic areas at 15 dpi. Together these results indicate that in spite of no differences in bacterial growth, strain CFBP6546R was able to induce more symptoms than strain CFBP6546R Δ *tal18H*. These results suggested that TAL18H was involved in the aggressiveness of the CFBP6546R strain.

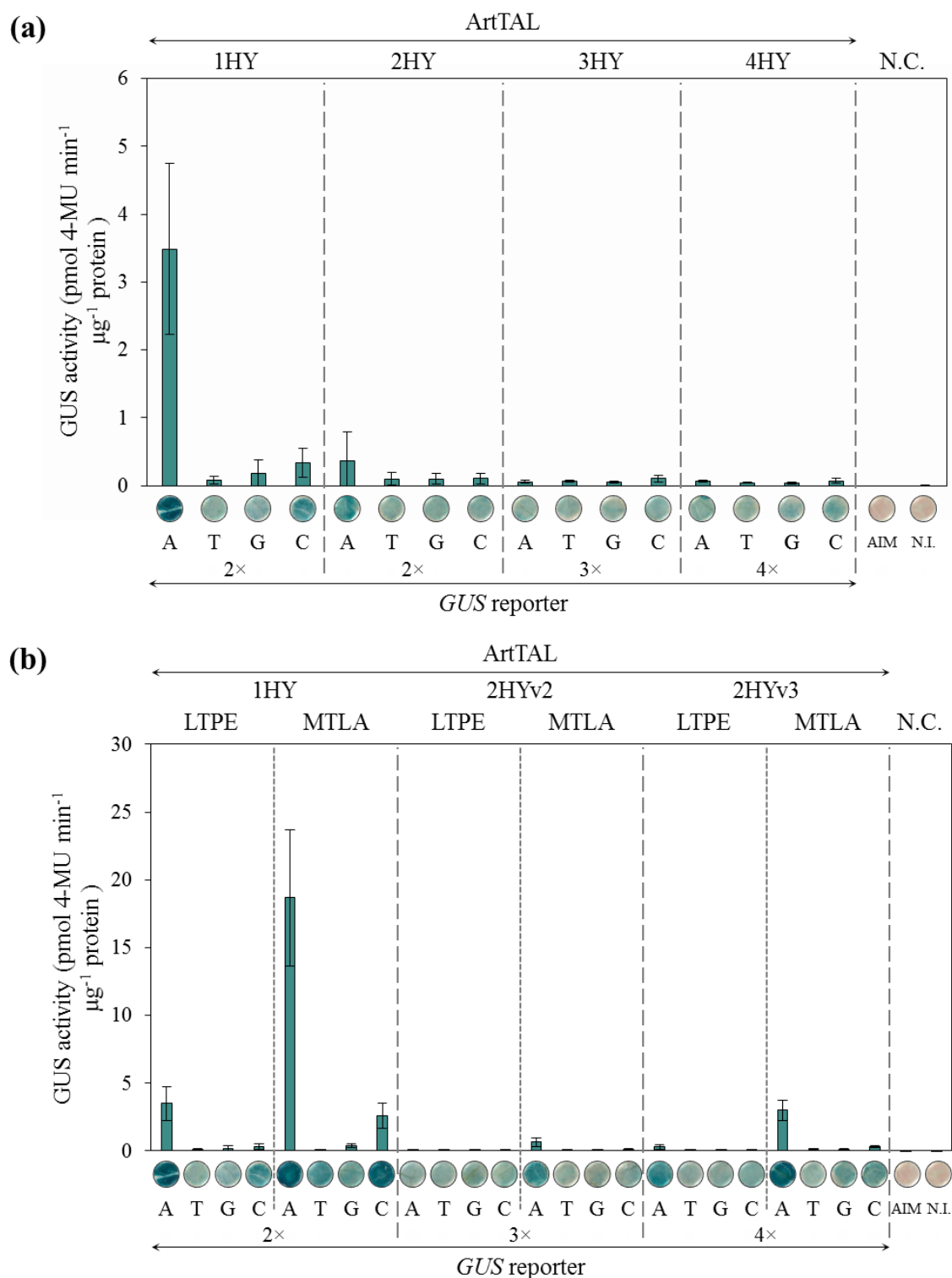


Fig. 2 Binding specificity of the HY RVD and impact of the MTLA motif. Sequences of ArtTAL and GUS reporter are described in Supplementary Material Table S1. GUS activity is expressed as pmol of 4-methylumbelliferone per minute per μg of total protein ($\text{pmol 4-MU min}^{-1} \mu\text{g}^{-1} \text{protein}$). Negative controls (N.C.) correspond to samples of non-inoculated *N. benthamiana* leaves (N.I.) or *N. benthamiana* leaves inoculated with bacterial-free AIM (AIM). **(a)** Binding specificity of the HY RVD. All HY repeat used were LTPE repeat. The activity of ArtTAL depend on the number of HY RVD. **(b)** Impact of distance between two HY RVD and of the MTLA motif. The activity of ArtTAL depend on the number of RVD between two HY RVD ($2\text{HYv}2 = 1 \text{ NS RVD between } 2 \text{ HY RVD}$ and $2\text{HYv}3 = 2 \text{ NS RDV between } 2 \text{ HY RVD}$) and is improved by the MTLA motif.

TAL18H presented an atypical repeat starting with an MTLA motif, carrying the unusual HY RVD

As described previously the repeat region of TAL18H was composed of 18 consecutive repeats and thus 18 RVD (**Supplementary Material Fig. S3**) (Ruh *et al.*, 2017a). Interestingly, at positions 10, 14 and 15, TAL18H presented atypical repeats starting with MTLA instead of the usual LTPE or LTPD motifs found in most TAL effectors reported to date (Bonas *et al.*, 1993; De Feyter *et al.*, 1993; Yang & Gabriel, 1995b; Kay *et al.*, 2005; El Yacoubi *et al.*, 2007; Boch & Bonas, 2010; Wilkins *et al.*, 2015; Huguet-Tapia *et al.*, 2016; Peng *et al.*, 2016; Ruh *et al.*, 2017a). The MTLA motif was not found in any of the TAL effectors described in *Xanthomonas* and TAL-like effectors described in other bacteria, such as RipTAL from *Ralstonia solanacearum* (de Lange *et al.*, 2013; Li *et al.*, 2013a), Bats from *Burkholderia rhizoxinica* (de Lange *et al.*, 2014b; Juillerat *et al.*, 2014; Stella *et al.*, 2014) or MOrTL from marine organisms (de Lange *et al.*, 2015). In particular, all TAL and TAL-like effectors repeats usually started with a conserved leucine or phenylalanine, but never with a methionine as it was the case here, suggesting that the MTLA repeat appeared only recently in CBB agents.

Moreover, at positions 10 and 15, MTLA repeats comprised the unusual HY RVD. This RVD was previously reported in only five other TAL effectors from *X. oryzae* pv. *oryzicola* strains, and one RipTAL from *Ralstonia solanacearum* (de Lange *et al.*, 2013; Li *et al.*, 2013a; Wilkins *et al.*, 2015). Previous large-scale study of RVD binding specificities showed that HY was not very efficient at binding any nucleotide (Yang *et al.*, 2014). However, this study was performed on three consecutive identical RVD, and used classical repeats starting with LTPE. To further characterize the binding specificity and efficiency of the HY RVD we constructed ArtTAL containing HY at different positions, and reporter constructs containing corresponding EBE within the promoter region of the *GUS* gene, as previously described (Streubel *et al.*, 2012). Different *GUS* reporter constructs were designed for testing every nucleotides (A, T, C or G) to bind with the HY RVD (**Supplementary Material Table S2**). First, we tested the activity of ArtTAL containing from one to four contiguous LTPE repeats with HY (**Fig. 2a**). When alone, HY possessed a binding specificity for adenine, and a slight *GUS* activity was observed for cytosine. These bindings were greatly impaired for two consecutive HY repeats, and were completely lost when the number of consecutive HY repeats increased. Such decrease in binding efficiency was previously observed for the NI RVD (Streubel *et al.*, 2012). We also designed two ArtTAL named 2HYv2-LTPE and 2HYv3-LTPE, both containing two HY respectively separated by one or two NS, an RVD able to bind every nucleotide (**Supplementary Material Table S2**). HY binding efficiency to adenine was abolished for two HY RVD separated with one NS RVD, and a slight binding was observed when separated by two NS RVD, although it was very weak (**Fig. 2b**).

To study the impact of the MTLA motif on HY binding specificity, we performed similar experiments after modifying the first four amino acids of the repeats containing HY from LTPE to MTLA (**Supplementary Material Table S2**). Interestingly, with the MTLA motif HY binding efficiency to adenine increased by more than fivefold compared to LTPE (**Fig. 2b**). In addition, HY was able to bind cytosine while it was not with LTPE repeats. For MTLA repeats, when two HY were separated by one NS, binding efficiency to cytosine was abolished, while binding efficiency to adenine became very weak. A rescue of HY binding efficiency to adenine was observed when the two HY were separated by two NS (**Fig. 2b**). Together, our result show that HY was able to weakly bind adenine, and that the MTLA motif increased this efficiency of HY RVD to bind adenine, but also conferred the ability to bind cytosine, albeit in a weaker way.

Comparison of RNAseq data and TAL effector target predictions revealed potential genes targeted by TAL18H

To analyse the ability of TAL18H to induce *in planta* transcription of genes, we performed RNAseq experiments on common bean leaves from cultivar JaloEEP558 two days after inoculation with strains CFBP6546R or CFBP6546R Δ *tal18H*. The genome of JaloEEP558 was previously sequenced and annotated, with a total of 32,914 predicted protein-coding genes (data not shown). A total of 5,427 common bean genes were differentially expressed when comparing the transcriptomes from leaves inoculated with CFBP6546R or CFBP6546R Δ *tal18H*. This corresponds to 16.5 % of the total genes found in JaloEEP558. Of these differentially expressed genes, 3,107 were induced after inoculation with the wild type strain compared to the mutant strain, while 2,320 genes were repressed. To analyse the gene ontologies that were the most represented among these differentially expressed genes, we used the Gene Set Enrichment tool from Genevestigator (Hruz *et al.*, 2008) on 400 orthologs from *Arabidopsis thaliana* corresponding to the most induced genes in our data. This analysis unveiled that 102 of these 400 genes were related to the cell periphery, the external encapsulating structure corresponding to the cell wall, and/or to the plasma membrane (**Supplementary Material Fig. S4a**). Similar analysis on 400 orthologs from *A. thaliana* corresponding to the most repressed genes in our data unveiled genes mostly involved in single biological processes, and more particularly 139 genes involved in catalytic activities (**Supplementary Material Fig. S4b**). Our data therefore indicate that TAL18H impacted the transcriptome of common bean by modifying the expression of genes linked to various functions, more particularly by inducing functions linked to the plant cell periphery and repressing functions linked to catalytic activities.

Among these induced genes, 153 genes had a logFC > 2, which make them good TAL18H target candidates. In order to highlight potential targets of TAL18H, we compared our RNAseq data with EBE predictions performed on the common bean promoterome consisting of the 3 kb

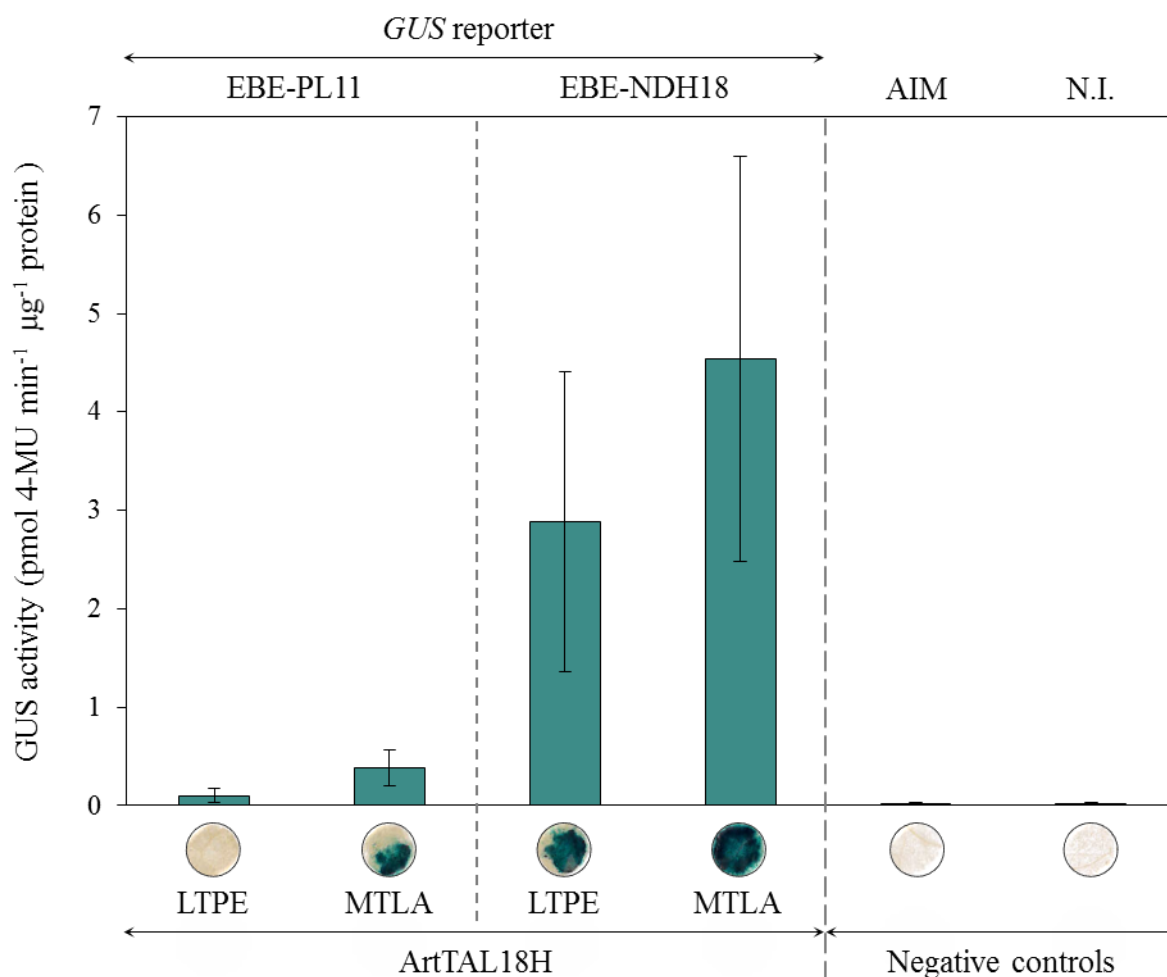


Fig. 3 Binding specificity of TAL18H and impact of the MTLA motif. Sequences of ArtTAL and GUS reporter are described in Supplementary Material Table S1. GUS activity is expressed as pmol of 4-methylumbelliferone per minute per µg of total protein (pmol 4-MU min⁻¹ µg⁻¹ protein). Negative controls (N.C.) correspond to samples of non-inoculated *N. benthamiana* leaves (N.I.) or *N. benthamiana* leaves inoculated with bacterial-free AIM (AIM). The binding specificity of TAL18H was tested on two EBE corresponding to those predicted into promoters of *PL11* and *NDH18*. Impact of the MTLA motif was tested using ArtTAL18H-MTLA.

region upstream of the transcriptional start of each annotated common bean gene, using Target Finder with default parameters (Doyle *et al.*, 2012) or TALVEZ version 3.2 using a correction matrix starting from the 14th RVD (Pérez-Quintero *et al.*, 2013). These programs did not take into account the HY RVD. It was previously shown that like HY, NI binds to adenine (Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Yang *et al.*, 2014). Therefore, predictions were done with HY replaced by NI in the RVD sequence of TAL18H. Among the 153 most induced genes, we identified 13 genes that were also predicted as potential targets of TAL18H by Target Finder, and one gene predicted by TALVEZ (**Supplementary Material Table S4**). Among the most interesting candidates for being TAL18H targets, three caught our attention for various reasons. The first gene was *PvJaloEEP55800033g0016481* which encodes a NAD(P)H dehydrogenase 18 (*NDH18*). This gene was the third in the ranking according to RNAseq data. Interestingly, three putative EBE duplicated in tandem were predicted by Target Finder in the promoter region of this gene. The second gene, *PvJaloEEP55803280g0250741*, encodes a two-component response regulator ARR8-related (*ARR8*). This gene was ranked ninth according to RNAseq data and the predicted EBE was the best ranked according to Target Finder. The third gene was *PvJaloEEP55800441g0102071* which encodes a pectate lyase 11-related (*PL11*). This gene was thirty-sixth in the ranking according to RNAseq data and the only predicted EBE according to TALVEZ.

To analyse TAL18H binding specificity to the predicted EBE of these three genes, we performed GUS assays using ArtTAL carrying the TAL18H RVD sequence with HY repeats starting with LTPE or MTLA (named ArtTAL18H-LTPE and ArtTAL18H-MTLA, respectively), and *GUS* reporter constructs containing each candidate EBE in the promoter region of the *GUS* gene. Due to technical reasons, we could not test the EBE predicted in the promoter region of *ARR8*. For the EBE located in the promoter region of *PL11*, no GUS activity was observed for ArtTAL18H-LTPE, and a very weak GUS activity was observed for ArtTAL18H-MTLA. We observed higher GUS activities for the EBE located in the promoter region of *NDH18*, suggesting that *NDH18* was a primary target of TAL18H. MTLA repeats slightly enhanced *GUS* activity, indicating that the MTLA motif favoured HY binding efficiency in the context of TAL18H structure (**Fig. 3**).

DISCUSSION

Here, we characterized the binding specificity of HY, an unusual RVD present in the RVD sequence of TAL18H, and the impact of the MTLA motif specific to this TAL effector. To understand the role of TAL18H in the interaction between CBB agents and common bean, we assessed its impact on the pathogenicity of strain CFBP6546R and on the transcriptome of common bean genotype JaloEEP558.

The maintenance of TAL18H in phylogenetically distant strains from *Xcf* and *Xpp* (Ruh *et al.*, 2017a) suggested that TAL18H was important for the pathogenicity of strains carrying it. Maintenance of the MTLA repeat, together with the HY RVD also suggested that these motifs were important for TAL18H function. Our results indicate that the MTLA motif enhanced the binding efficiency of HY to adenine, and provided a weak binding to cytosine. Additionally, consecutive HY abolished the ArtTAL binding to DNA. TAL effector binding to DNA is a complex, fine-tune mechanism involving direct interaction between plant nucleotides and the 13th residue of each TAL effector repeat (Deng *et al.*, 2012; Mak *et al.*, 2012). In some cases, hydrogen bonds can be formed between consecutive RVD (Deng *et al.*, 2012) illustrating that steric environment can be important for successful TAL effector-DNA binding. Here, the 13th residue of HY repeat was a tyrosine (Y), an aromatic residue only very rarely found in any other TAL effectors described to date (de Lange *et al.*, 2013; Wilkins *et al.*, 2015). It's tempting to speculate that the tyrosine of the HY RVD would create a steric hindrance interfering with correct binding to DNA and, in view of our results, that the MTLA motif would somehow alleviate this steric hindrance, which would increase HY binding efficiency to adenine. It would be interesting to verify whether this MTLA motif generates the same type of result for other RVD and in particular for the RVD NG because it is also present in another MTLA repeat from TAL18H.

It appears that TAL18H had no impact on bacterial growth *in planta*, yet more severe CBB symptoms were observed after inoculation with the WT strain than with the mutant strain, indicating that TAL18H enhances the aggressiveness of the CFBP6546R strain. The high bacterial population levels observed for the mutant strain can be explained by the ability of these bacteria to multiply in biofilms while performing symptomless colonization of the leaves. TAL18H impacted a large number of common bean genes, both positively and negatively. Most induced genes were related to the outer component of the plant cell, including the plasma membrane and the cell wall. Important genes could be induced for favouring the apoplastic life of *Xcf* in common bean leaves. In particular, three genes encoding pectate lyases (PL1, PL11 and PL22) were found among the most induced genes (**Supplementary Material Table S4**). Pectate lyases are involved in the degradation of the plant cell wall during cell elongation (Leng *et al.*, 2017), and can in some cases allow a more rapid progression of pathogens in the plant tissues (Marín-Rodríguez *et al.*, 2002; Xin *et al.*, 2016). In tomato, a pectate lyase was recently described as a secondary target of TAL effector AvrHah1 from *X. gardneri*, via the induction of a bHLH transcription factor (Schwartz *et al.*, 2017). The authors hypothesize that the pectate lyase activity increased the hygroscopicity of the cell wall, which may lead to the increase of symptoms of tissue maceration. Further experiments in common bean would help characterizing the contribution of pectate lyases to the water-soaking symptoms observed in CBB, as well as other genes induced by TAL18H and potentially important for CBB symptoms.

TAL18H also led to the repression of numerous genes encoding proteins with catalytic activity. Interestingly here, genes that are important for plant defence were found among the 50 most repressed genes, including genes encoding proteases, transcription factors from the WRKY family, a leucine-rich repeat-containing protein or a NPR1 protein. Proteases can be involved in diverse metabolic pathways, and many proteases have been characterised for their role in plant defence (Figueiredo *et al.*, 2014; Misas-Villamil *et al.*, 2016), such as Rcr3 in tomato (Krüger *et al.*, 2002), or CDR1 in *Arabidopsis* (Xia *et al.*, 2004). Transcription factors from the WRKY family regulate positively or negatively the activation of genes necessary for numerous aspects of plant development and defence (Ishihama & Yoshioka, 2012; Aamir *et al.*, 2017; Jiang *et al.*, 2017). Finally, NPR1 is a key protein of the salicylic acid (SA) defence pathway that is pivotal in the plant immune system (Wu *et al.*, 2012). Overall, our results show that TAL18H has a pleiotropic effect on the common bean transcriptome, and could contribute both to inducing genes favourable for symptom development and lowering genes involved in plant defences.

Combination of RNAseq, EBE predictions and GUS assays identified the NDH18-encoding gene as the putative primary target of TAL18H. The putative EBE was duplicated three times in tandem in the promoter of the NDH18-encoding gene. It would be interesting to perform additional GUS assays under the complete NDH18 promoter to study how this duplication would impact TAL18H induction ability. In *Arabidopsis*, NDH18 belongs to the NDH-PSI transmembrane supercomplex resulting from the interaction between NAD(P)H dehydrogenase (NDH) complex and the photosystem I (PSI) complex (Peng *et al.*, 2009; Ifuku *et al.*, 2011). Signals originating from photosynthetic reactions can lead to redox and pH changes as well as accumulation or reactive oxygen species that can accumulate in response to biotic or abiotic stress. How *NDH18* could be involved in the enhancement of CBB symptoms still needs more work to be elucidated.

Together, our results highlight that the repeat structure outside of the RVD can be important for TAL effectors binding specificity, and show that TAL effectors can display pleiotropic effects *in planta*.

ACKNOWLEDGEMENTS

This work was supported by grants from the French National Research Agency (CROpTAL project, grant number ANR-14-CE19-0002-04) and Angers-Loire Metropole, France (X-HOT project). The authors thank Sandra Pelletier for helping in RNAseq analysis, Céline Rousseau and Phenotic platform for analysis of chlorophyll fluorescence imaging, Alvaro L. Pérez-Quintero for helping in TAL effector target predictions and Daniel Sochard for *in planta* experiments. The authors also thank the French Network on *Xanthomonads* (FNX) (<https://www.reseau->

xantho.org/) and particularly Laurent D. Noël, Boris Szurek, Lionel Gagnevin and Ralf Koebnik for recurrent scientific exchanges on *Xanthomonas*. PacBio sequencing was performed at Icahn School of Medicine at Mount Sinai (NY, USA) under the supervision of Robert Sebra and preliminary quality control of DNA was performed at the ANAN platform from the SFR 4207 QUASAV (Angers, France). We thank the CIRM-CFBP (Angers, France) for strain preservation and supply. RNAseq was performed at Get-PlaGe, Centre INRA (Castanet-Tolosan, France) under the supervision of Olivier Bouchez and Maarten Person.

AUTHOR CONTRIBUTION

MR, MAJ and NWGC designed the study and wrote the manuscript. MR, JF, AP, SC, MB and NWGC contributed to phenotyping and RNAseq analyses. MR, SB and JB contributed in the RVD binding specificity analyses. All authors have read and approved the final version of the manuscript.

REFERENCES

- Aamir M, Singh VK, Meena M, Upadhyay RS, Gupta VK, Singh S. 2017.** Structural and Functional Insights into WRKY3 and WRKY4 Transcription Factors to Unravel the WRKY–DNA (W-Box) Complex Interaction in Tomato (*Solanum lycopersicum* L.). A Computational Approach. *Frontiers in Plant Science* **8**: 1–24.
- Abe VY, Benedetti CE. 2016.** Additive roles of PthAs in bacterial growth and pathogenicity associated with nucleotide polymorphisms in effector-binding elements of citrus canker susceptibility genes. *Molecular Plant Pathology* **17**: 1223–1236.
- Al-Saadi A, Reddy JD, Duan YP, Brunings AM, Yuan Q, Gabriel DW. 2007.** All five host-range variants of *Xanthomonas citri* carry one pthA homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. *Molecular plant-microbe interactions : MPMI* **20**: 934–943.
- Alavi SM, Sanjari S, Durand F, Brin C, Manceau C, 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 h. *Applied and Environmental Microbiology* **74**: 3295–3301.
- Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. 2010.** Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant cell* **22**: 3864–76.
- Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. 2015.** Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Frontiers in Microbiology* **6**: 1–18.
- Belete T, Bastas K. 2017.** Common Bacterial Blight (*Xanthomonas axonopodis* pv. *phaseoli*) of Beans with Special Focus on Ethiopian Condition. *Journal of Plant Pathology & Microbiology* **8**.
- Benjamini Y, Hochberg Y. 1995.** Controlling the false discovery rate: a practical and powerful approach to multiple

- testing. *Journal of the Royal Statistical Society B* **57**: 289–300.
- Blanvillain-Baufumé S, Reschke M, Solé M, Auguy F, Doucoure H, Szurek B, Meynard D, Portefaix M, Cunnac S, Guiderdoni E, et al. 2017.** Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for SWEET14-inducing TAL effectors. *Plant biotechnology journal* **15**: 306–317.
- Boch J, Bonas U. 2010.** *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* **48**: 419–36.
- Boch J, Bonas U, Lahaye T. 2014.** TAL effectors - pathogen strategies and plant resistance engineering. *New Phytologist* **204**: 823–832.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. 2009.** Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509–12.
- Bonas U, Conrads-Strauch J, Balbo I. 1993.** Resistance in tomato to *Xanthomonas campestris* pv. *vesicatoria* is determined by alleles of the pepper-specific avirulence gene *avrBs3*. *Molecular & general genetics : MGG* **238**: 261–269.
- Bonas U, Stall RE, Staskawicz B. 1989.** Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular & general genetics : MGG* **218**: 127–136.
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia N V, Bogdanove AJ, Voytas DF. 2011.** Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**: 1–11.
- Cernadas RA, Doyle EL, Niño-Liu DO, Wilkins KE, Bancroft T, Wang L, Schmidt CL, Caldo R, Yang B, White FF, et al. 2014.** Code-Assisted Discovery of TAL Effector Targets in Bacterial Leaf Streak of Rice Reveals Contrast with Bacterial Blight and a Novel Susceptibility Gene. *PLoS Pathogens* **10**: e1003972.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al. 2013.** Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* **10**: 563–569.
- Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. 2012.** Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues. *PLoS ONE* **7**.
- Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou BH, Frommer WB, Lahaye T, Staskawicz BJ. 2014.** *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. *Molecular plant-microbe interactions : MPMI* **27**: 1186–1198.
- Cong L, Zhou R, Kuo Y, Cunniff M, Zhang F. 2012.** Comprehensive Interrogation of Natural TALE DNA Binding Modules and Transcriptional Repressor Domains. *Nature Communications* **3**.
- Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B. 2016.** Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathology* **65**: 792–806.
- Darsonval A, Darrasse A, Durand K, Bureau C, Cesbron S, Jacques M-A. 2009.** Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Molecular plant-microbe interactions : MPMI* **22**: 747–757.
- Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J, Shi Y. 2012.** Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science* **335**: 720–723.
- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. 2012.** TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Research* **40**: 117–122.
- EFSA PLH Panel (EFSA Panel on Plant Health). 2014.** Scientific Opinion on the pest categorisation of *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *EFSA Journal* **12**: 3856.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, et al. 2009.** Single

- Polymerase Molecules. *Science* **323**: 133–138.
- De Feyter R, Yang Y, Gabriel DW. 1993.** Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. *malvacearum* avr genes. *Molecular plant-microbe interactions : MPMI* **6**: 225–237.
- Figueiredo A, Monteiro F, Sebastiana M. 2014.** Subtilisin-like proteases in plant-pathogen recognition and immune priming: a perspective. *Frontiers in Plant Science* **5**: 1–4.
- Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J. 2013.** Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Computational Biology* **9**.
- Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang G-L, White FF, et al. 2005.** R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122–1125.
- Hajri A, Brin C, Hunault G, Lardeux F, Lemaire C, Manceau C, Boureau T, 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.
- Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Duffy B, Manceau C. 2012.** Type Three Effector Gene Distribution and Sequence Analysis Provide New Insights into the Pathogenicity of Plant-Pathogenic *Xanthomonas arboricola*. *Applied and environmental microbiology* **78**: 371–384.
- Hayward AC. 1993.** The hosts of *Xanthomonas*. *Xanthomonas*. 1–119.
- Heuer H, Yin YN, Xue QY, Smalla K, Guo JH. 2007.** Repeat domain diversity of avrBs3-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Applied and Environmental Microbiology* **73**: 4379–4384.
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. 2008.** Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics* **2008**: 1–5.
- Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, Yang B, White FF, Wang N, Jones JB. 2014.** Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E521–9.
- Huguet-Tapia JC, Peng Z, Yang B, Yin Z, Liu S, White FF. 2016.** Complete Genome Sequence of the African Strain AXO1947 of *Xanthomonas oryzae* pv. *oryzae*. *Genome announcements* **4**: 10–11.
- Hummel AW, Doyle EL, Bogdanove AJ. 2012.** Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**: 883–893.
- Hummel AW, Wilkins KE, Wang L, Cernadas RA, Bogdanove AJ. 2016.** A transcription activator-like effector from *Xanthomonas oryzae* pv. *oryzicola* elicits dose-dependent resistance in rice. *Molecular plant pathology* **18**: 55–56.
- Hutin M, Césari S, Chalvon V, Michel C, Tran TT, Boch J, Koebnik R, Szurek B, Kroj T. 2016.** Ectopic activation of the rice NLR heteropair RGA4/RGA5 confers resistance to bacterial blight and bacterial leaf streak diseases. *Plant Journal* **88**: 43–55.
- Hutin M, Pérez-Quintero AL, Lopez C, Szurek B. 2015.** MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. *Frontiers in plant science* **6**: 535.
- Ifuku K, Endo T, Shikanai T, Aro E-M. 2011.** Structure of the chloroplast NADH dehydrogenase-like complex: Nomenclature for nuclear-encoded subunits. *Plant and Cell Physiology* **52**: 1560–1568.
- Ishihama N, Yoshioka H. 2012.** Post-translational regulation of WRKY transcription factors in plant immunity. *Current Opinion in Plant Biology* **15**: 431–437.
- Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al. 2016.** Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas*: French Network on *Xanthomonads* (FNX). *Annual Review of Phytopathology* **54**: 163–187.

- 593 **Jiang Y, Guo L, Ma X, Zhao X, Jiao B, Li C, Luo K. 2017.** The WRKY transcription factors PtrWRKY18 and PtrWRKY35
594 promote Melampsora resistance in Populus. *Tree Physiology* **37**: 665–675.
- 595 **Juillerat A, Bertonati C, Dubois G, Guyot V, Thomas S, Valton J, Beurdeley M, Silva GH, Daboussi F, Duchateau P.**
596 **2014.** BurrH: a new modular DNA binding protein for genome engineering. *Scientific reports* **4**: 3831.
- 597 **Kay S, Boch J, Bonas U. 2005.** Characterization of AvrBs3-like effectors from a Brassicaceae pathogen reveals virulence
598 and avirulence activities and a protein with a novel repeat architecture. *Molecular plant-microbe interactions : MPMI* **18**:
599 838–848.
- 600 **Kay S, Hahn S, Marois E, Hause G, Bonas U. 2007.** A bacterial effector acts as a plant transcription factor and induces
601 a cell size regulator. *Science* **318**: 648–651.
- 602 **Kay S, Hahn S, Marois E, Wieduwild R, Bonas U. 2009.** Detailed analysis of the DNA recognition motifs of the
603 Xanthomonas type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant Journal* **59**: 859–871.
- 604 **Krüger J, Thomas CM, Golstein C, Dixon MS, Smoker M, Tang S, Mulder L, Jones JDG. 2002.** A Tomato Cysteine
605 Protease Required for Cf-2-Dependent Disease Resistance and Suppression of Autonecrosis. *Science* **296**: 744–747.
- 606 **de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinowski J, Heuer H, Strauß A, Lahaye T. 2013.**
607 Breaking the DNA-binding code of Ralstonia solanacearum TAL effectors provides new possibilities to generate plant
608 resistance genes against bacterial wilt disease. *New Phytologist* **199**: 773–786.
- 609 **de Lange O de, Wolf C, Dietze J, Elsaesser J, Morbitzer R, Lahaye T. 2014.** Programmable DNA-binding proteins from
610 Burkholderia provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Research* **42**: 7436–7449.
- 611 **de Lange O, Wolf C, Thiel P, Krüger J, Kleusch C, Kohlbacher O, Lahaye T. 2015.** DNA-binding proteins from marine
612 bacteria expand the known sequence diversity of TALE-like repeats. *Nucleic Acids Research* **43**: 10065–10080.
- 613 **Leng Y, Yang Y, Ren D, Huang L, Dai L, Wang Y, Chen L, Tu Z, Gao Y, Li X, et al. 2017.** A rice PECTATE LYASE-LIKE
614 gene is required for plant growth and leaf senescence. *Plant Physiology* **174**: 1151–1166.
- 615 **Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M, Sharakuu A, Mahjoub A, Wang G, Khan S, et al. 2013.**
616 Characterization and DNA-binding specificities of Ralstonia TAL-like effectors. *Molecular Plant* **6**: 1318–1330.
- 617 **Li C, Wei J, Lin Y, Chen H. 2012.** Gene silencing using the recessive rice bacterial blight resistance gene xa13 as a new
618 paradigm in plant breeding. *Plant Cell Reports* **31**: 851–862.
- 619 **Love MI, Huber W, Anders S. 2014.** Moderated estimation of fold change and dispersion for RNA-seq data with
620 DESeq2. *Genome Biology* **15**: 1–21.
- 621 **Mak ANS, Bradley P, Bogdanove AJ, Stoddard BL. 2013.** TAL effectors: Function, structure, engineering and
622 applications. *Current Opinion in Structural Biology* **23**: 93–99.
- 623 **Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. 2012.** The Crystal Structure of TAL Effector PthXo1
624 Bound to Its DNA Target. *Science* **335**: 716–719.
- 625 **Mallory AC, Vaucheret H. 2006.** Functions of microRNAs and related small RNAs in plants. *Nature Genetics* **38**: S31–
626 S36.
- 627 **Marín-Rodríguez MC, Orchard J, Seymour GB. 2002.** Pectate lyases, cell wall degradation and fruit softening. *Journal*
628 *of Experimental Botany* **53**: 2115–2119.
- 629 **Mhedbi-Hajri N, Darrasse A, Pigné S, Durand K, Fouteau S, Barbe V, Manceau C, Lemaire C, Jacques M-A. 2011.**
630 Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. *BMC evolutionary biology* **11**: 67.
- 631 **Misas-Villamil JC, van der Hoorn RAL, Doeblemann G. 2016.** Papain-like cysteine proteases as hubs in plant
632 immunity. *New Phytologist* **212**: 902–907.
- 633 **Moscou MJ, Bogdanove AJ. 2009.** A simple cipher governs DNA recognition by TAL effectors. *Science* **326**: 1501.
- 634 **Mutlu N, Vidaver AK, Coyne DP, Steadman JR, Lambrecht PA, Reiser J. 2008.** Differential Pathogenicity of
635 Xanthomonas campestris pv. phaseoli and X. fuscans subsp. fuscans Strains on Bean Genotypes with Common Blight
636 Resistance. *Plant Disease* **92**: 546–554.

- 637 **Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T.**
 638 **2007.** Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for
 639 plant transformation. *Journal of Bioscience and Bioengineering* **104**: 34–41.
- 640 **Peng L, Fukao Y, Fujiwara M, Takami T, Shikanai T. 2009.** Efficient Operation of NAD(P)H Dehydrogenase Requires
 641 Supercomplex Formation with Photosystem I via Minor LHCI in Arabidopsis. *The Plant Cell* **21**: 3623–3640.
- 642 **Peng Z, Hu Y, Xie J, Potnis N, Akhunova A, Jones J, Liu Z, White FF, Liu S. 2016.** Long read and single molecule DNA
 643 sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics* **17**:
 644 21.
- 645 **Pérez-Quintero AL, Rodriguez-R LM, Dereeper A, López C, Koebnik R, Szurek B, Cunnac S. 2013.** An Improved
 646 Method for TAL Effectors DNA-Binding Sites Prediction Reveals Functional Convergence in TAL Repertoires of
 647 *Xanthomonas oryzae* Strains. *PLoS ONE* **8**: e68464.
- 648 **Rademaker JLW, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J, de Bruijn FJ. 2005.** A comprehensive
 649 species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* **95**: 1098–1111.
- 650 **Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T. 2007.** Plant pathogen recognition mediated by promoter
 651 activation of the pepper Bs3 resistance gene. *Science* **318**: 645–648.
- 652 **Rousseau C, Belin E, Bove E, Rousseau D, Fabre F, Berruyer R, Guillaumès J, Manceau C, Jacques M-A, Boureau T.**
 653 **2013.** High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis.
 654 *Plant methods* **9**: 17.
- 655 **Rousseau C, Hunault G, Gaillard S, Bourbeillon J, Montiel G, Simier P, Campion C, Jacques M-A, Belin E, Boureau**
 656 **T. 2015.** Phenoplant: a web resource for the exploration of large chlorophyll fluorescence image datasets. *Plant methods*
 657 **11**.
- 658 **Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017.** *Xanthomonas* adaptation to common bean is associated
 659 with horizontal transfers of genes encoding TAL effectors. *BMC Genomics* **18**: 670.
- 660 **Ryan RP, Vorhölter F-J, Potnis N, Jones JB, Van Sluys M-A, Bogdanove AJ, Dow JM. 2011.** Pathogenomics of
 661 *Xanthomonas*: understanding bacterium–plant interactions. *Nature Reviews Microbiology* **9**: 344–355.
- 662 **Sallet E, Gouzy J, Schiex T. 2014.** EuGene-PP: A next-generation automated annotation pipeline for prokaryotic
 663 genomes. *Bioinformatics* **30**: 2659–2661.
- 664 **Schornack S, Moscou MJ, Ward ER, Horvath DM. 2013.** Engineering Plant Disease Resistance Based on TAL Effectors.
 665 *Annual Review of Phytopathology* **51**: 383–406.
- 666 **Schornack S, Peter K, Bonas U, Lahaye T. 2005.** Expression levels of avrBs3-like genes affect recognition specificity
 667 in tomato Bs4- but not in pepper Bs3-mediated perception. *Molecular plant-microbe interactions : MPMI* **18**: 1215–1225.
- 668 **Schwartz AR, Morbitzer R, Lahaye T, Staskawicz BJ. 2017.** TALE-induced bHLH transcription factors that activate a
 669 pectate lyase contribute to water soaking in bacterial spot of tomato. *Proceedings of the National Academy of Sciences of*
 670 *the United States of America* **114**: E897–E903.
- 671 **Schwartz AR, Potnis N, Timilsina S, Wilson M, Patané J, Martins J, Minsavage G V., Dahlbeck D, Akhunova A,**
 672 **Almeida N, et al. 2015.** Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in
 673 effector repertoires and identifies determinants of host specificity. *Frontiers in Microbiology* **6**: 1–17.
- 674 **Stella S, Molina R, Bertonatti C, Juillerrat A, Montoya G. 2014.** Expression, purification, crystallization and
 675 preliminary X-ray diffraction analysis of the novel modular DNA-binding protein BurrH in its apo form and in complex
 676 with its target DNA. *Acta Crystallographica Section F Structural Biology Communications* **70**: 87–91.
- 677 **Streubel J, Blücher C, Landgraf A, Boch J. 2012.** TAL effector RVD specificities and efficiencies. *Nature Biotechnology*
 678 **30**: 593–595.
- 679 **Streubel J, Pesce C, Hutin M, Koebnik R, Boch J, Szurek B. 2013.** Five phylogenetically close rice SWEET genes confer
 680 TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist* **200**: 808–819.

- Sugio A, Yang B, Zhu T, White FF. 2007.** Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes OsTFIIAgamma1 and OsTFX1 during bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 10720–5.
- Vauterin L, Hoste B, Kersters K, Swings J. 1995.** Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**: 472–489.
- Wang L, Rinaldi FC, Singh P, Doyle EL, Dubrow ZE, Tu T, Pérez-quintero AL, Szurek B, Bogdanove AJ. 2017.** TAL effectors drive transcription bidirectionally in plants. *Molecular Plant* **10**: 285–296.
- Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, Qin T, Li Y, Che J, Zhang M, et al. 2015.** XA23 Is an executor R protein and confers broad-spectrum disease resistance in rice. *Molecular Plant* **8**: 290–302.
- Wilkins KE, Booher NJ, Wang L, Bogdanove AJ. 2015.** TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. *Frontiers in Plant Science* **6**: 1–15.
- Wu Y, Zhang D, Chu JY, Boyle P, Wang Y, Brindle ID, De Luca V, Després C. 2012.** The Arabidopsis NPR1 Protein Is a Receptor for the Plant Defense Hormone Salicylic Acid. *Cell Reports* **1**: 639–647.
- Xia Y, Suzuki H, Borevitz J, Blount J, Guo Z, Patel K, Dixon RA, Lamb C. 2004.** An Extracellular Aspartic Protease Functions in Arabidopsis Disease Resistance Signaling. *EMBO Journal* **23**: 980–988.
- Xin X-F, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, Chang JH, Zipfel C, He SY. 2016.** Bacteria establish an aqueous living space in plants crucial for virulence. *Nature* **539**: 524–529.
- El Yacoubi B, Brunings AM, Yuan Q, Shankar S, Gabriel DW. 2007.** In planta horizontal transfer of a major pathogenicity effector gene. *Applied and Environmental Microbiology* **73**: 1612–1621.
- Yang Y, Gabriel DW. 1995.** Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. These include : Intragenic Recombination of a Single Plant Pathogen Gene Provides a Mechanism for the Evolution of New Host Specifici. *Journal of Bacteriology* **177**: 4963–4968.
- Yang B, Sugio A, White FF. 2006.** Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 10503–8.
- Yang B, White FF. 2004.** Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular plant-microbe interactions : MPMI* **17**: 1192–200.
- Yang J, Zhang Y, Yuan P, Zhou Y, Cai C, Ren Q, Wen D, Chu C, Qi H, Wei W. 2014.** Complete decoding of TAL effectors for DNA recognition. *Cell research* **24**: 628–31.
- Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B. 2011.** Colonization of rice leaf blades by an african strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. *Molecular Plant-Microbe Interactions* **24**: 1102–13.
- Zeng X, Tian D, Gu K, Zhou Z, Yang X, Luo Y, White FF, Yin Z. 2015.** Genetic engineering of the Xa10 promoter for broad-spectrum and durable resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant Biotechnology Journal* **13**: 993–1001.
- Zhang J, Yin Z, White F. 2015.** TAL effectors and the executor R genes. *Frontiers in Plant Science* **6**: 641.
- Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom JS, Huang S, Liu S, Vera Cruz C, Frommer WB, et al. 2015.** Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant Journal* **82**: 632–643.

Supplementary Material Table S1. Bacterial strains and plasmids used in this study.

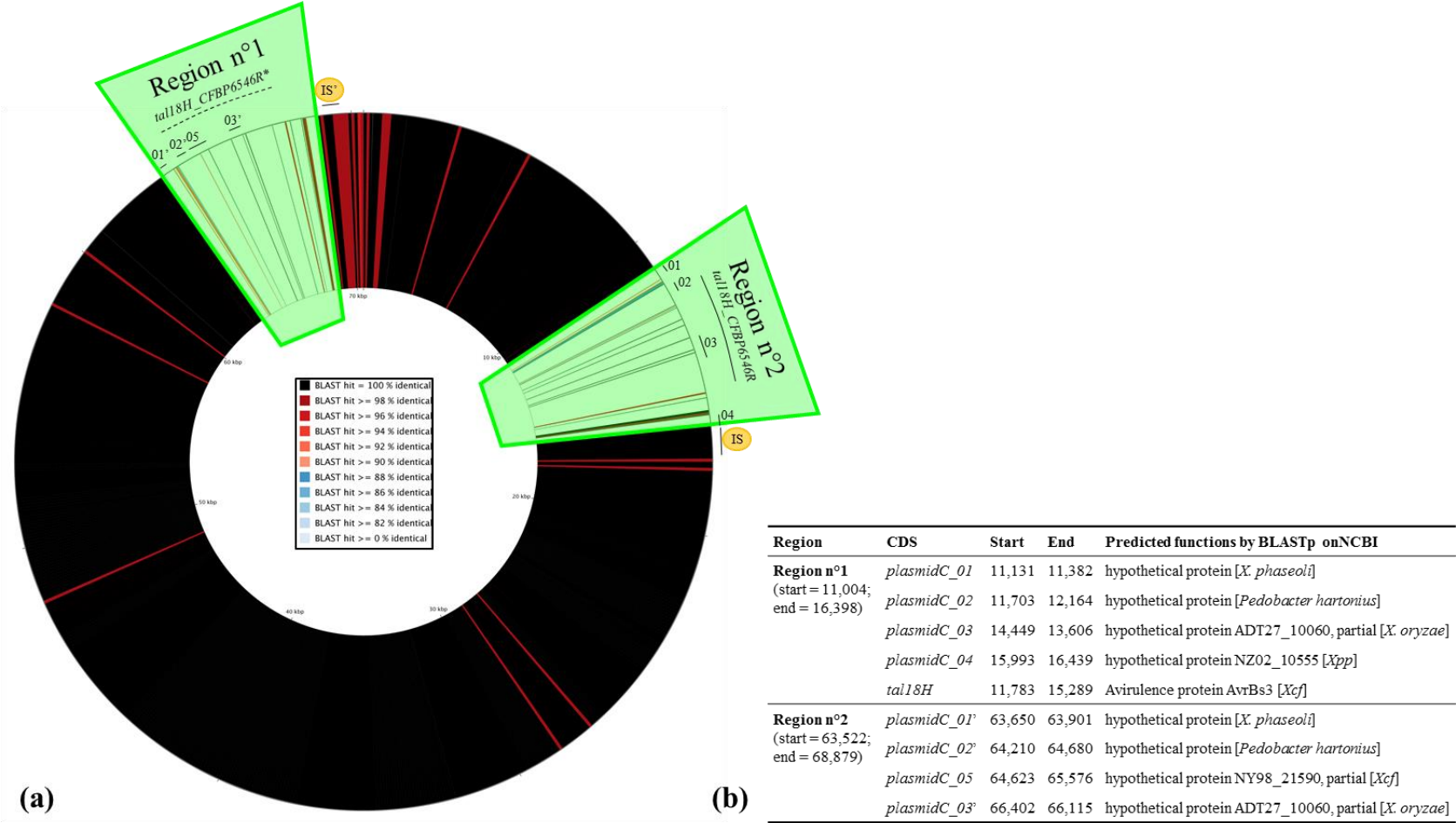
Strain or Plasmid	Characteristics	Reference
<i>X. phaseoli</i> pv. <i>phaseoli</i>		
CFBP6546R	WT strain; rifamycin resistant	Ruh <i>et al.</i> , 2017
CFBP6546R Δ <i>tal18H</i>	<i>tal18H</i> deletion mutant; rifamycin resistant	This study
<i>A. tumefaciens</i>		
GV3101	WT strain; rifampicin resistant	Van Larebeke <i>et al.</i> , 1974
GV3101(1HY-LTPE)	GV3101 derivative containing pSKA2-1HY-LTPE	This study
GV3101(2HY-LTPE)	GV3101 derivative containing pSKA2-2HY-LTPE	This study
GV3101(3HY-LTPE)	GV3101 derivative containing pSKA2-3HY-LTPE	This study
GV3101(4HY-LTPE)	GV3101 derivative containing pSKA2-4HY-LTPE	This study
GV3101(2HYv2-LTPE)	GV3101 derivative containing pSKA2-2v2HY-LTPE	This study
GV3101(2HYv3-LTPE)	GV3101 derivative containing pSKA2-2v3HY-LTPE	This study
GV3101(1HY-MTLA)	GV3101 derivative containing pSKA2-1HY-MTLA	This study
GV3101(2HYv2-MTLA)	GV3101 derivative containing pSKA2-2v2HY-MTLA	This study
GV3101(2HYv3-MTLA)	GV3101 derivative containing pSKA2-2v3HY-MTLA	This study
GV3101(<i>ArtTAL18H</i> -LTPE)	GV3101 derivative containing pSKA2- <i>ArtTAL18H</i> -LTPE	This study
GV3101(<i>ArtTAL18H</i> -MTLA)	GV3101 derivative containing pSKA2- <i>ArtTAL18H</i> -MTLA	This study
GV3101(<i>GUS</i> -2A)	GV3101 derivative containing pGWB3- <i>GUS</i> -2A	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -2T)	GV3101 derivative containing pGWB3- <i>GUS</i> -2T	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -2G)	GV3101 derivative containing pGWB3- <i>GUS</i> -2G	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -2C)	GV3101 derivative containing pGWB3- <i>GUS</i> -2C	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -3A)	GV3101 derivative containing pGWB3- <i>GUS</i> -3A	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -3T)	GV3101 derivative containing pGWB3- <i>GUS</i> -3T	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -3G)	GV3101 derivative containing pGWB3- <i>GUS</i> -3G	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -3C)	GV3101 derivative containing pGWB3- <i>GUS</i> -3C	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -4A)	GV3101 derivative containing pGWB3- <i>GUS</i> -4A	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -4T)	GV3101 derivative containing pGWB3- <i>GUS</i> -4T	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -4G)	GV3101 derivative containing pGWB3- <i>GUS</i> -4G	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -4C)	GV3101 derivative containing pGWB3- <i>GUS</i> -4C	Streubel <i>et al.</i> , 2012
GV3101(<i>GUS</i> -EBE-PL11)	GV3101 derivative containing pGWB3- <i>GUS</i> -EBE-PL11	This study
GV3101(<i>GUS</i> -EBE-NDH18)	GV3101 derivative containing pGWB3- <i>GUS</i> -EBE-NDH18	This study
Plasmids		
pUC57-HD1 to HD6	Vectors containing repeat with HD RVD at 6 positions; carbenicillin resistant	Boch <i>et al.</i> , 2009
pUC57-NI1 to NI6	Vectors containing repeat with HD RVD at 6 positions; carbenicillin resistant	Boch <i>et al.</i> , 2009
pUC57-NG1 to NG6	Vectors containing repeat with NG RVD at 6 positions; carbenicillin resistant	Boch <i>et al.</i> , 2009
pUC57-NS1 to NS6	Vectors containing repeat with NS RVD at 6 positions; carbenicillin resistant	Boch <i>et al.</i> , 2009
pUC57-HY1 to HY6	Vectors containing repeat with HY RVD at 6 positions; carbenicillin resistant	This study
pUC57-HY2- and HY6-MTLA	Vectors containing repeat with HY RVD and MTLA motif at positions n°2 and n°6; carbenicillin resistant	This study
pSKA2	ENTRY vector containing a Hax3 backbone; spectinomycin resistant	Boch <i>et al.</i> , 2009
pGWB3	Vector containing a GUS gene; kanamycin resistant	Nakagawa <i>et al.</i> , 2007

Supplementary Material Table S2. Constructions produced and used for GUS assays in this study.

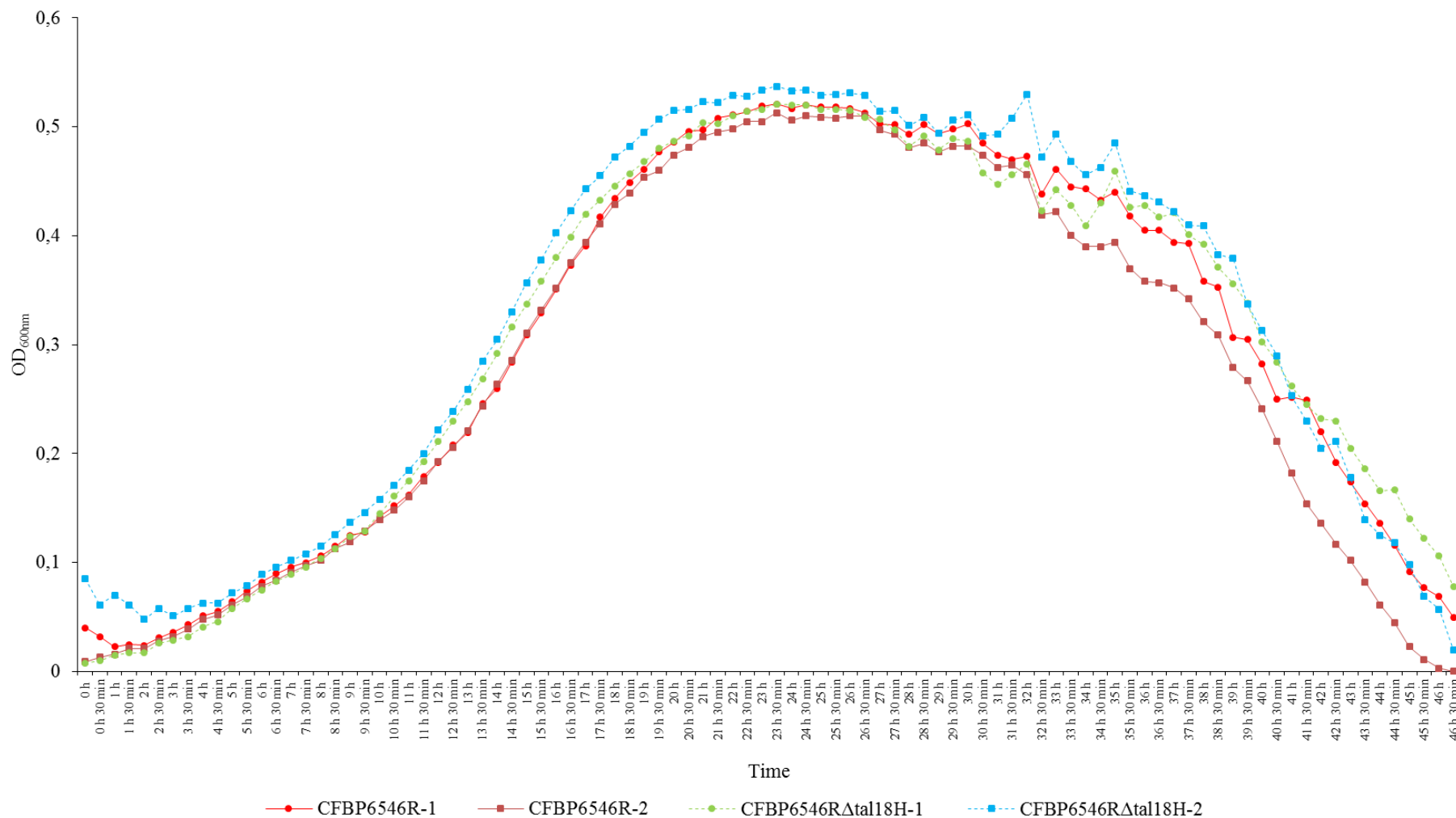
Construction	Sequence	Combination
Repeat	Repeat sequence	
HY-LTPE	LTPEQVVAIAS HY GGKQALETVQRLLPVLCQAHG	
HY-MTLA	MTLAQVVAIAS HY GGKQALETVQRLLPVLCQAHG	
other	LTPEQVVAIAS XX GGKQALETVQRLLPVLCQAHG	
ArtTAL	RVD sequence	GUS reporter
1HY-LTPE	NI-HD-NI-HD-HD-HD- HY -NS-NS-HD-NI-NG	GUS-2X
2HY-LTPE	NI-HD-NI-HD-HD-HD- HY-HY -NS-HD-NI-NG	GUS-2X
3HY-LTPE	NI-HD-NI-HD-HD-HD- HY-HY-HY -HD-NI-NG	GUS-3X
4HY-LTPE	NI-HD-NI-HD-HD-HD- HY-HY-HY-HY -NI-NG	GUS-4X
2HYv2-LTPE	NI-HD-NI-HD-HD-HD- HY -NS- HY -HD-NI-NG	GUS-3X
2HYv3-LTPE	NI-HD-NI-HD-HD-HD- HY -NS-NS- HY -NI-NG	GUS-4X
1HY-MTLA	NI-HD-NI-HD-HD-HD- HY -NS-NS-HD-NI-NG	GUS-2X
2HYv2-MTLA	NI-HD-NI-HD-HD-HD- HY -NS- HY -HD-NI-NG	GUS-3X
2HYv3-MTLA	NI-HD-NI-HD-HD-HD- HY -NS-NS- HY -NI-NG	GUS-4X
ArtTAL18H-LTPE	NI-NG-HD-NG-HD-NI-NG-NI-HD- HY -NN-N*-HD-NG- HY -NN-HD-NG	GUS-EBE
ArtTAL18H-MTLA	NI-NG-HD-NG-HD-NI-NG-NI-HD- HY -NN-N*-HD- NG-HY -NN-HD-NG	GUS-EBE
GUS reporter	EBE sequence	ArtTAL
GUS-2A	A C A C C C A A A C A T	1HY; 2HY
GUS-2T	A C A C C C T T A C A T	1HY; 2HY
GUS-2G	A C A C C C G G A C A T	1HY; 2HY
GUS-2C	A C A C C C C C A C A T	1HY; 2HY
GUS-3A	A C A C C C A A A C A T	3HY; 2HYv2
GUS-3T	A C A C C C T T T C A T	3HY; 2HYv2
GUS-3G	A C A C C C G G G C A T	3HY; 2HYv2
GUS-3C	A C A C C C C C C C A T	3HY; 2HYv2
GUS-4A	A C A C C C A A A A A T	4HY; 2HYv3
GUS-4T	A C A C C C T T T T A T	4HY; 2HYv3
GUS-4G	A C A C C C G G G G A T	4HY; 2HYv3
GUS-4C	A C A C C C C C C C A T	4HY; 2HYv3
GUS-EBE-PL11	A T C T C A T A C C A T C C T C T T	ArtTAL18H
GUS-EBE-NDH18	G T C T C A T A A A G A C T A A C T	ArtTAL18H

Supplementary Material Table S3. Primers used in this study.

Primer name	Sequence	Product size	Interest
rep-allele-links	ATGGCCACCACCTGCTCCGG	3000 bp	Production of ArtTAL repeats carrying the HY RVD
rep_mut_HY_rechts	CGCCAGCCACTATGGCGCAAGCAGGCGCTGGAGA		
MTLA2for	TTGAAGACTTATGACCCTGGCGCAGGTGGTGGCCATCGCC	3000 bp	Production of ArtTAL repeats at position n°2 with a MTLA motif and carrying the HY RVD
MTLA2rev	AGGGCCCGTCGACTGCAGAG		
MTLA6for	CTTCCATGGCATGACCCTGGCGCAGGTGGTGGCCATCGCC	3000 bp	Production of ArtTAL repeats at position n°6 with a MTLA motif and carrying the HY RVD
MTLA6rev	TCTTCAAAGGGCCCGTCGAC		
mutsite1_MTLA_for	GGCCCATGGCATGAAAGTCTTCAAAGGGAT	3000 bp	Production of ArtTAL repeats at position n°1 compatible with repeat at position n°2 with a MTLA motif
mutsite1_MTLA_rev	TGGCACAGCACCGGCAACAG		
MTLA_Rep1_F	TTTGAAGACTTTGACCCTGGCGCAGGTGGTGGCCATCGC	100 bp	Production of ArtTAL repeats at position n°1 with a MTLA motif and carrying the NG RVD
MTLA_Rep1_R	TTTGAAGACTTTCAGGCCATGGGCCTGGCACAGCACCG		
MTLA_CR_F	ATGACAAGTCTTCCTTGTCTGTAAGCGGA	3000 bp	Production of assembly vector compatible with repeat at position n°1 with a MTLA motif
MTLA_CR_R	ATGACAAGTCTTCCTTGTCTGTAAGCGGA		
Bs4_GG_R_AKR	TTTGGTCTCACCTTAGATTCGATTAAAAATAAAATTGTATG	100 bp	Production of GUS reporter carrying the EBE of <i>PL11</i> or <i>NDH18</i>
EBE-PL11_F	TTTGGTCTCACACCTATCTCATACCATCCTCTTTTCTTTCTTGTATATAACTTTGTCC		
EBE-NDH18_F	TTTGGTCTCACACCTGTCTCATAAAGACTAACTTTCTTTCTTGTATATAACTTTGTCC		



Supplementary Material Fig. S1. Deleted regions of CFBP6546RΔ*tal18H*. (a) Circular representation of plasmids C from CFBP6546R containing *tal18H_CFBP6546R* and CFBP6546RΔ*tal18H*. Genomic sequences of plasmids C from CFBP6546R and CFBP6546RΔ*tal18H* were compared and converted in a graphical map using CGView (Grant *et al.*, 2012) with strain CFBP6546R as reference. Colours differ according to identity percentage (see legend). Localisation of the deleted regions containing *tal18H_CFBP6546R* and *tal18H_CFBP6546R** pseudogene are highlighted by green zones. The predicted genes are represented by black lines while the sequence corresponding to *tal18H_CFBP6546R** pseudogene is represented by a black dotted line. Yellow circles indicate the presence of insertion sequences (IS). (b) Detail of the CDS predicted in the regions corresponding to the two deleted regions of CFBP6546RΔ*tal18H*.



Supplementary Material Fig. S2. Dynamics of population sizes of CFBP6546RΔtal18H. The bacterial growths of WT and mutant deletion of *tal18H* strains were controlled two times with biological duplicates and technical triplicates.

N-terminal encoding sequence

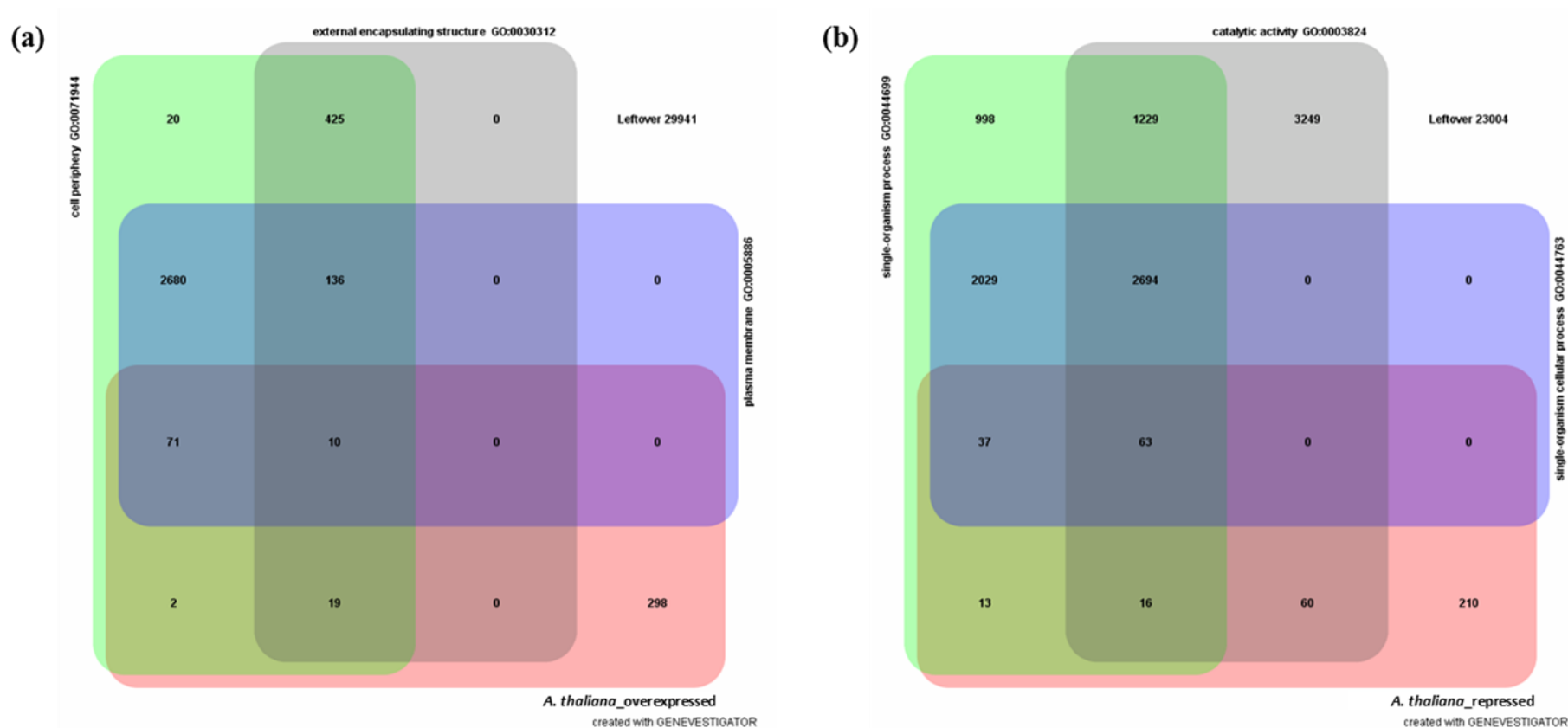
```

1      LTPDQVVAIASNIGGKQALETVQLLLPVLCEQHG
2      LTPDQVVAIASNGGGKQALETVQRLLPVLCCKDHG
3      LTPAQVVAIANHDGKQALETVQRLLPVLCCKDHG
4      LTPAQVVAIASNGGGKQALETVEQLLLPVLCCKDHG
5      LTPDQVVAIANHDGKPALETVQRLLPVLCQELG
6      LTPDQVVAIASNIGGKQALETVQRLLPVLCQELG
7      LTPDQVVAIASNGGAKQALETVQRLLPVLCQELG
8      LTPAQVVAIASNIGGKQALETVQRLLPVLCEQHG
9      LTPDQVVAIASHHDGKQALETVERLLLPVLCQELG
10     MTLAQVVAIASHYGKQALEAVQRLLPVLCQELG
11     LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
12     LTPDQVVAIASN*GGKQALETVQRLLPVLCEQHG
13     LTPDQVVAIASHHDGKQALETVERLLLPVLCQELG
14     MTLAQVVAIASNGGGKQALETVERLLLPVLCQELG
15     MTLAQVVAIASHYGKQALEAVQRLLPVLCQDLG
16     LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
17     LTPDQVVAIASHHDGKQALEAVQRLLPVLCQDHG
18     LNPDQVVAIASNGGGRQALE

```

C-terminal encoding sequence

Supplementary Material Fig. S3. Protein alignment of TAL18H repeats. The RVD are underlined. The unusual HY RVD is highlighted in yellow, the MTLA motif is highlighted in green.



Supplementary Material Fig. S4. Gene ontologies of the most overexpressed and repressed genes after inoculation with CFBP6546R vs. CFBP6546R Δ tal18H. Images were generated using Genevestigator on 400 *A. thaliana* orthologues corresponding to the most overexpressed genes in common bean (a) or on 400 *A. thaliana* orthologues corresponding to the most repressed genes in common bean (b).

Supplementary Material Table S4. Details of the genes significantly overexpressed in the presence of *tal18H*, and *in silico* predictions of TAL18H targets.

Rank*	Gene of JaloEEP558	Corresponding gene in G19833	Predicted function with Phytozome	Rank of TAL18H target prediction	
				Talvez	Target Finder
1	PVJaloEEP55800861g0145481	Phvul.007G139200.1	Protein of unknown function (DUF1635)		
2	PVJaloEEP55800679g0128221	Phvul.007G099700.1	No functional annotations		
3	PVJaloEEP55800033g0016481	Phvul.007G280100.1	NAD(P)H dehydrogenase 18		117 to 119
4	PVJaloEEP55800226g0067831	Phvul.001G041700.1	BTB/POZ domain-containing protein npy1		
5	PVJaloEEP55800203g0063461	Phvul.001G001500.1	Multi-copper oxidase type I family protein-related		
6	PVJaloEEP55800673g0127691	Phvul.003G148400.1	Allergen V5/TPX-1-related family protein-related		
7	PVJaloEEP55801425g0185111	Phvul.008G144100.1	Beta expansin 6-related		
8	PVJaloEEP55800438g0101811	Phvul.006G182700.1	No functional annotations		
9	PVJaloEEP55803280g0250741	Phvul.001G125800.1	Two-component response regulator ARR8-related		41
10	PVJaloEEP55801242g0173941	Phvul.007G046100.1	Cysteine and glycine-rich protein (CSRP)		
11	PVJaloEEP55800057g0024751	Phvul.006G197500.1	Cysteine-rich secretory protein-relate		
12	PVJaloEEP55803218g0249351	Phvul.009G012200.1	Wuschel-related homeobox 13		
13	PVJaloEEP55800889g0148081	Phvul.007G071100.1	PAR1 protein		
14	PVJaloEEP55801772g0201171	Phvul.003G207400.1	Pectate lyase 22-related		
15	PVJaloEEP55800459g0104461	Phvul.002G210700.2	Taurine-transporting ATPase		
16	PVJaloEEP55800001g0000831	NA	NA		
17	PVJaloEEP55800401g0096411	Phvul.008G270900.1	Glycosyl hydrolase family 9 protein		
18	PVJaloEEP55803316g0251431	Phvul.002G086150.1	Origin recognition complex subunit 6 (ORC6)		
19	PVJaloEEP55826286g0327781	Phvul.006G176920.1	Fruit bromelain		
20	PVJaloEEP55800545g0114001	Phvul.009G133600.1	Basic blue protein		
21	PVJaloEEP55800018g0010221	Phvul.007G236800.1	Pectate lyase 1-related		221
22	PVJaloEEP55800525g0111851	Phvul.007G274300.1	Phosphoethanolamine N-methyltransferase		
23	PVJaloEEP55801077g0162381	Phvul.011G089900.1	Glucomannan 4-beta-mannosyltransferase 2		
24	PVJaloEEP55803897g0263521	Phvul.001G186700.1	Pre-mRNA cleavage complex 2 protein Pcf11 (PCF11)		378
25	PVJaloEEP55808389g0306631	Phvul.003G143400.2	Peroxidase 16-related		
26	PVJaloEEP55800293g0079661	Phvul.001G225000.1	Nicotianamine synthase		
27	PVJaloEEP55820685g0326051	Phvul.008G235300.1	Gibberellin-regulated protein 12-related		
28	PVJaloEEP55800033g0016471	Phvul.007G280000.1	Acylaminoacyl-peptidase [EC:3.4.19.1] (APEH)		
29	PVJaloEEP55800670g0127291	Phvul.005G031700.1	Large subunit ribosomal protein L27 (RP-L27, MRPL27, rpmA)		
30	PVJaloEEP55802585g0232231	Phvul.003G247700.1	Multi-copper oxidase		
31	PVJaloEEP55800091g0037291	Phvul.007G261600.1	No functional annotations		678
32	PVJaloEEP55801592g0193431	Phvul.009G006600.1	Glyceraldehyde-3-phosphate dehydrogenase (NADP(+)) (phosphorylating) / Triosephosphate dehydrogenase (NADP+)		
33	PVJaloEEP55800033g0016461	Phvul.007G280000.1	Acylaminoacyl-peptidase [EC:3.4.19.1] (APEH)		
34	PVJaloEEP55803426g0253651	Phvul.003G086500.1	Oligopeptide transporter 3		1.185
35	PVJaloEEP55800999g0156601	Phvul.009G074300.1	AP2-like factor, ANT lineage (OVM, ANT)		
36	PVJaloEEP55800441g0102071	Phvul.004G059200.1	Pectate lyase 11-related	169	
37	PVJaloEEP55801056g0160901	Phvul.005G007700.1	Homogalacturonan degradation		
38	PVJaloEEP55800244g0070931	Phvul.008G232200.1	Expansin-A6		
39	PVJaloEEP55800273g0076371	Phvul.008G259500.1	F10K1.18 protein		
40	PVJaloEEP55800139g0049821	Phvul.001G263700.1	Immunoglobulin E-SET superfamily protein		
41	PVJaloEEP55800016g0008961	Phvul.008G255300.1	No functional annotations		
42	PVJaloEEP55808349g0306351	Phvul.004G031000.1	Pathogen-inducible salicylic acid glucosyltransferase [EC:2.4.1.-] Glc b1-2 SA (SGT1)		
43	PVJaloEEP55800382g0093781	Phvul.005G003100.1	Endoglucanase 3-related		
44	PVJaloEEP55802812g0238811	Phvul.009G221600.1	No functional annotations		
45	PVJaloEEP55800040g0018881	Phvul.002G207400.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein-related		
46	PVJaloEEP55800538g0113241	Phvul.007G186200.1	Plant protein of unknown function (DUF868)		

977

833

626

1.146

103	PVJaloEEP55800377g0093111	Phvul.002G006000.1	Glycine cleavage system H protein 3, mitochondrial
104	PVJaloEEP55805111g0281611	Phvul.008G219100.1	Amino acid transporter
105	PVJaloEEP55800848g0144211	Phvul.010G028300.1	Interleukin-1 receptor-associated kinase 1 (IRAK1)
106	PVJaloEEP55800017g0009571	Phvul.003G187700.1	Cell division protein kinase
107	PVJaloEEP55800093g0037951	Phvul.004G032800.1	Lamin-like protein
108	PVJaloEEP55800433g0100861	Phvul.008G041200.1	Gibberellin-regulated protein 4
109	PVJaloEEP55801386g0183161	Phvul.004G170100.2	Polyketide cyclase / dehydrase and lipid transport (Polyketide_cyc2)
110	PVJaloEEP55800378g0093251	NA	NA
111	PVJaloEEP55802280g0221721	Phvul.002G084100.1	Homeodomain-like superfamily protein
112	PVJaloEEP55801339g0179861	Phvul.007G086100.1	D-xylose 1-dehydrogenase (NADP(+))
113	PVJaloEEP55800845g0143861	Phvul.010G114700.1	Aspartyl protease-related
114	PVJaloEEP55800116g0044521	Phvul.007G214900.1	AGAA.5 // Domain of unknown function (DUF4228)
115	PVJaloEEP55800148g0051741	Phvul.001G188900.1	Purple acid phosphatase 21-related
116	PVJaloEEP55802511g0229681	Phvul.007G102400.1	Di-glucose binding protein with Leucine-Rich Repeat domain
117	PVJaloEEP55817868g0324541	Phvul.007G276500.1	No functional annotations
118	PVJaloEEP55800431g0100711	Phvul.007G066500.1	Dehydration-responsive element-binding protein 1A-related
119	PVJaloEEP55802633g0233641	Phvul.009G255600.1	No functional annotations
120	PVJaloEEP55800077g0032231	Phvul.002G297600.1	Endoglucanase 19-related
121	PVJaloEEP55800267g0075091	Phvul.002G084500.1	F7A19.32 protein-related // ribonuclease T2
122	PVJaloEEP55801279g0176281	Phvul.010G063300.1	Patellin-6
123	PVJaloEEP55800745g0134711	Phvul.007G167800.1	Leucine Rich Repeat (LRR_1) // Leucine rich repeat N-terminal domain (LRRNT_2)
124	PVJaloEEP55802878g0240401	Phvul.006G041100.1	Purple acid phosphatase 18
125	PVJaloEEP55802997g0243921	Phvul.001G190100.1	Ribonucleoside-diphosphate reductase subunit M2 (RRM2)
126	PVJaloEEP55801357g0181231	Phvul.010G022000.1	NINJA-family protein AFP1-related
127	PVJaloEEP55800860g0145431	Phvul.004G002600.1	Probable lipid transfer (LTP_2)
128	PVJaloEEP55802640g0233891	Phvul.007G123002.1	No functional annotations
129	PVJaloEEP55802858g0239901	Phvul.005G045500.1	Aspartyl protease family protein
130	PVJaloEEP55801236g0173471	NA	NA
131	PVJaloEEP55800333g0085641	Phvul.001G071600.1	G2/mitotic-specific cyclin-B
132	PVJaloEEP55801510g0189621	Phvul.008G290200.1	UDP-glycosyltransferase 71C4
133	PVJaloEEP55803404g0253321	Phvul.009G226900.1	Subtilisin serine protease-related
134	PVJaloEEP55800543g0113821	Phvul.003G266500.1	Aspartyl protease family protein
135	PVJaloEEP55800594g0119371	Phvul.009G142800.1	Expansin-A8
136	PVJaloEEP55800826g0142271	Phvul.009G028900.1	Allergen-like protein BRN20-related
137	PVJaloEEP55801448g0186351	Phvul.005G175700.1	Domain of unknown function (DUF4666)
138	PVJaloEEP55800116g0044681	Phvul.007G216200.1	Cotton fibre expressed protein (DUF761)
139	PVJaloEEP55803362g0252371	Phvul.007G150300.1	Adenosine kinase
140	PVJaloEEP55801300g0177501	Phvul.003G217800.1	Monocopper oxidase-like protein SKU5
141	PVJaloEEP55801510g0189631	Phvul.008G290300.1	UDP-glycosyltransferase 71C4
142	PVJaloEEP55802433g0227191	NA	NA
143	PVJaloEEP55800880g0147291	Phvul.008G251400.1	Sodium/hydrogen exchanger family (Na_H_Exchange)
144	PVJaloEEP55804744g0276761	Phvul.003G151400.1	No functional annotations
145	PVJaloEEP55803093g0246581	Phvul.008G120700.1	Expansin-A6
146	PVJaloEEP55800023g0012211	Phvul.001G171600.1	Ankyrin repeat-containing protein
147	PVJaloEEP55800997g0156411	Phvul.007G203400.1	Inositol 3-alpha-galactosyltransferase / UDP-alpha-D-galactose:myo-inositol 3-alpha-D-galactosyltransferase
148	PVJaloEEP55801977g0209971	Phvul.009G225000.1	Ethylene-responsive transcription factor ERF018
149	PVJaloEEP55802462g0228081	Phvul.004G077100.1	Dynein light chain type 1-like protein
150	PVJaloEEP55802640g0233881	Phvul.007G123002.1	No functional annotations
151	PVJaloEEP55801443g0186161	Phvul.008G119600.1	Nitrogen metabolic regulation protein NMR-related
152	PVJaloEEP55800190g0060621	Phvul.009G157900.1	Patellin-3-related
153	PVJaloEEP55800238g0069881	Phvul.007G014100.1	Ring zinc finger protein

* Rank of the most overexpressed genes (BH-adjusted p-value < 0.05 and logFC > 2) in common bean after inoculation with the WT strain with the respect to the mutant strain.

NA: not applicable.

III. Résultats complémentaires et perspectives

1. Expériences complémentaires à réaliser pour caractériser le rôle de TAL18H dans l'interaction *Xanthomonas* / haricot

Afin de confirmer que les différences observées entre la souche sauvage et le mutant de délétion *tal18H* sont directement la conséquence de la perte de *tal18H* et que *NDH18* est un gène ciblé par TAL18H, plusieurs expériences complémentaires sont à réaliser.

Tout d'abord, il faut compléter le mutant de délétion *tal18H* avec un gène codant cet effecteur TAL. Pour ce faire, nous avons complété la souche CFBP6546RΔ*tal18H* avec des vecteurs d'expression d'ArtTAL18H-LTPE et ArtTAL18H-MTLA ainsi que le vecteur vide. L'étape suivante sera de contrôler la pathogénie de ces mutants complétés sur haricot en effectuant des tests de pouvoir pathogène similaires à ceux décrits dans l'ébauche d'article. Si après inoculation des mutants complétés avec ArtTAL18H nous observons des symptômes sur haricot similaires à ceux observés après inoculation de la souche sauvage, nous pourrions confirmer que TAL18H est impliqué dans l'agressivité de la souche.

Dans le but de valider notre gène candidat, *NDH18*, comme étant une cible de TAL18H, nous pourrions également utiliser les mutants complétés pour les infiltrer sur haricot et par la suite analyser les transcrits. En effet, si après infiltration des mutants complétés avec ArtTAL18H nous observons par qRT-PCR une surexpression de *NDH18*, nous pourrions affirmer que TAL18H est bel et bien impliqué dans l'expression de ce gène. Pour valider que cette surexpression est la conséquence de l'accroche de TAL18H sur le promoteur de *NDH18*, il sera nécessaire d'effectuer des co-inoculations sur *N. benthamiana* de souches d'*A. tumefaciens* transformées avec un vecteur portant ArtTAL18H et un vecteur portant le promoteur de *NDH18* en amont du gène rapporteur *GUS*. De la même façon que pour les résultats détaillés dans l'ébauche d'article, si ArtTAL18H accroche le promoteur au niveau de l'EBE et parvient à activer l'expression du gène en aval, nous observerons une augmentation de l'activité *GUS* qui nous permettra de confirmer que TAL18H est capable d'activer l'expression de *NDH18*.

D'autre part, nous pourrions également construire d'autres ArtTAL ciblant d'autres EBE dans le promoteur de *NDH18* et réitérer les tests de pouvoir pathogènes, les co-inoculations sur *N. benthamiana* et les analyses de transcrits afin de valider définitivement que les résultats que nous avons obtenus sont directement la conséquence de la perte de *tal18H* et de l'absence de la surexpression de *NDH18* par TAL18H.

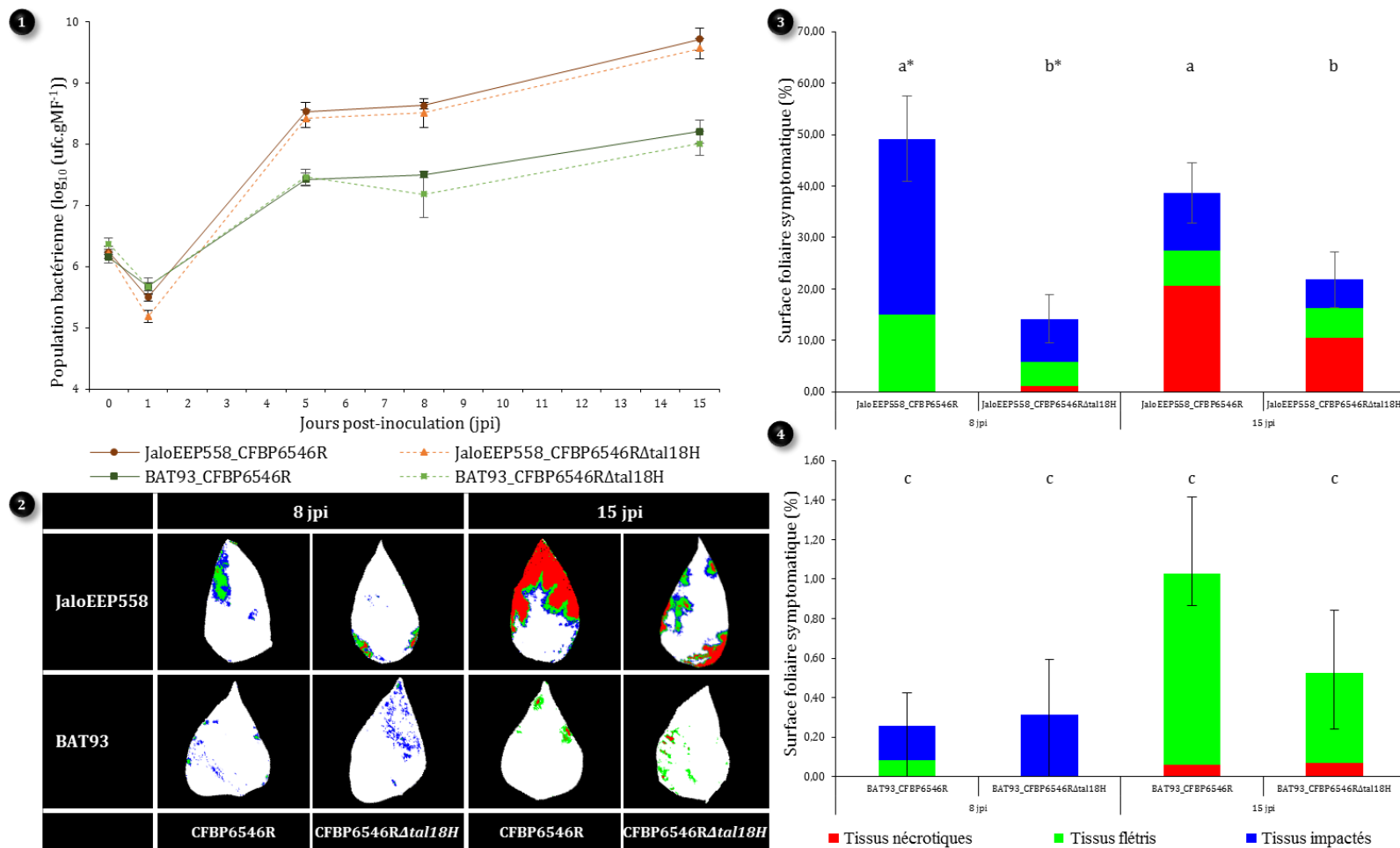


Figure 4 : Comparaison de l'impact de TAL18H sur la population bactérienne *in planta* et sur l'agressivité de la souche chez les deux variétés de haricot JaloEEP558 et BAT93. (1) Dynamique de population bactérienne après inoculation des souches CFBP6546R (ligne continue) et CFBP6546RΔtal18H (ligne pointillée) sur JaloEEP558 (en orange) et BAT93 (en vert) exprimée en $\log_{10}$ (u.f.c./gMF⁻¹) (gMF = gramme de matière fraîche). **(2)** Exemples de feuilles des deux variétés de haricot analysées par imagerie en fluorescence de chlorophylle 8 et 15 jours après inoculation des souches sauvages et mutantes. Surface foliaire symptomatique de JaloEEP558 **(3)** et BAT93 **(4)** 8 et 15 jours après inoculation des souches sauvages et mutantes. Les barres d'erreurs représentent l'erreur standard de la surface symptomatique totale. Les différentes lettres indiquent que les conditions sont significativement différentes avec une p-value < 0,05, les astérisques indiquent que c'est avec une p-value < 0,01, selon le test de Mann-Whitney.

2. Différences dans l'interaction *Xanthomonas* / haricot entre les deux variétés de haricot BAT93 et JaloEEP558

Afin de détecter d'éventuelles différences dans la réponse du haricot aux effecteurs TAL de *Xcf-Xpp*, dont notamment TAL18H, les tests de pouvoir pathogène et analyse des transcrits ont été réalisés sur deux variétés de haricot représentatives des deux centres de diversité majeurs du haricot : JaloEEP558 pour le centre andin et BAT93 pour le centre mésoaméricain. Lors des tests de pouvoir pathogène nous avons observé une différence dans l'interaction *Xanthomonas* / haricot commun. En effet, la population bactérienne *in planta* des deux souches était plus faible sur BAT93 que sur JaloEEP558 à partir du 5^{ème} jour après inoculation avec plus de $1 \log_{10}(\text{ufc.gMF}^{-1})$ d'écart. Egaleme nt, nous avons observé qu'il n'y avait pas de différence significative entre les croissances bactériennes des souches CFBP6546R et CFP6546R $\Delta tal18H$ sur un génotype donné (Figure 8). Cette différence de croissance bactérienne sur les deux variétés de haricot est associée à une forte diminution des symptômes sur BAT93. De façon plus précise, la surface foliaire symptomatique sur BAT93 ne dépassait jamais les 2 %, alors que sur JaloEEP558 elle était comprise entre 15 et 50 %. Malgré cette surface foliaire symptomatique drastiquement réduite sur BAT93, nous avons quand même observé une augmentation de la sévérité des symptômes au cours du temps (**Figure 4**). L'ensemble de ces résultats montrent donc que TAL18H n'a pas d'impact sur la croissance bactérienne *in planta* mais joue un rôle dans l'agressivité de la souche CFBP6546R chez JaloEEP558. Egaleme nt, ces résultats indiquent qu'il existe une importante différence de sensibilité entre les deux variétés de haricot que nous avons testées, JaloEEP558 étant plus sensible et BAT93 plus résistant. Il serait intéressant de tester une diversité de souches de *Xcf-Xpp* sur BAT93 afin de déterminer si la résistance observée est effective face aux différentes lignées de *Xap-Xff*, ce qui ferait de BAT93 une potentielle source de résistance pour la sélection de variétés résistantes à la CBB. Par ailleurs, des données de transcriptomique provenant des deux variétés de haricot deux jours après inoculation avec les souches sauvage et mutante ont été générées, et il serait intéressant de déterminer s'il existe une corrélation entre ces transcriptomes et les différences de résistance-sensibilité observées chez BAT93 et JaloEEP558, ainsi qu'une conservation ou une diversification des EBE ciblées par TAL18H entre ces deux variétés de haricot.

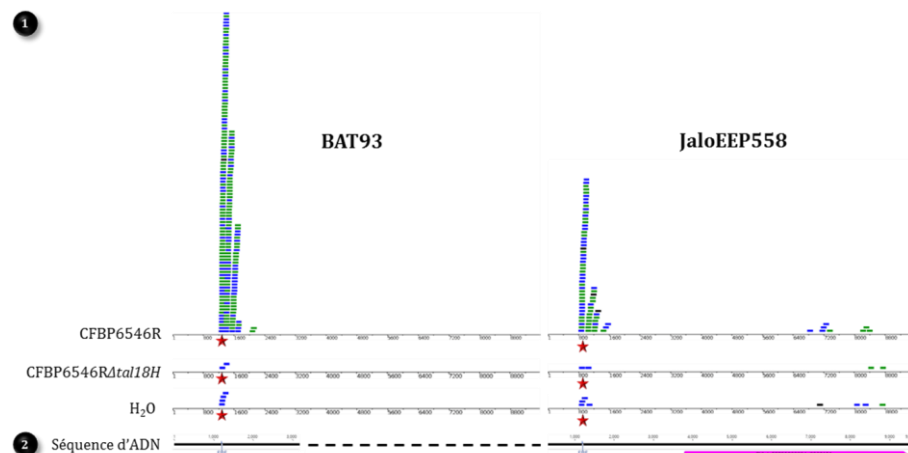


Figure 5 : Données RNAseq du petit transcrit situé au niveau d'une EBE prédite de TAL18H. (1) Cartographie des séquences RNAseq à proximité de l'EBE prédite de TAL18H. La région analysée comporte le petit transcrit situé au niveau d'une EBE (étoile rouge) et un gène en aval codant une glucuronosyltransférase. Cette analyse permet de visualiser l'accumulation des transcrits (en vert et bleu) pour chaque conditions d'inoculation. **(2)** Annotation des régions analysées. La ligne pointillée indique la fin du contig chez BAT93. Chez JaloEEP558, les séquences sont annotées comme suit : en bleu l'EBE et en rose le gène codant une glucuronosyltransférase

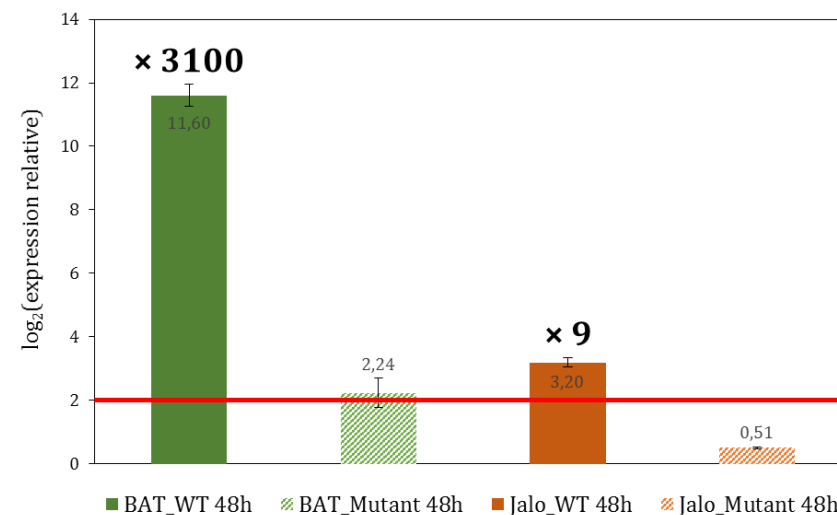


Figure 6 : Expression relative du petit transcrit situé au niveau de l'EBE prédite de TAL18H. Les feuilles des deux variétés BAT93 et JaloEEP558 ont été inoculées avec les souches CFPB6546R, CFBP6546RΔtal18H ou de l'eau. L'expression relative a été obtenue via la méthode des 2ΔΔCt en utilisant le gène de ménage *IDE* comme gène rapporteur. Le trait rouge représente la limite à partir de laquelle la surexpression est considérée comme significative soit si log₂(expression relative) > 2. Chaque résultat correspond à la moyenne de trois échantillons, avec trois répliques techniques par échantillon. Des résultats similaires ont été obtenus en utilisant les gènes *Act11* ou *EF1-α* comme gènes rapporteurs.

3. Analyse d'une potentielle cible intergénique de TAL18H

Les prédictions des cibles de TAL18H dans le génome du haricot ont révélé une cible particulièrement intéressante que nous n'avons pas mentionnée dans l'article ci-dessus car elle n'a pas été trouvée par analyses RNAseq classiques. En effet, une EBE a été retrouvée parmi les trois premières cibles prédites dans le classement de toutes les prédictions qui ont été menées pour TAL18H. Avec le RVD NI à la place du RVD HY, cette EBE était en deuxième position dans le classement des prédictions réalisées avec TALVEZ et en troisième position dans le classement des prédictions réalisées avec Target Finder. Fait intéressant, les deux bases de cette EBE prédites situées en face des RVD HY sont des cytosines et nous avons précédemment démontré que le RVD HY est capable de se lier à la cytosine lorsque la répétition le portant commence par un motif MTLA.

Cette EBE se situe à plus de 2 kb d'un gène codant pour une glucuronosyltransférase et il n'a jamais été démontré qu'un effecteur TAL est capable d'activer l'expression d'un gène aussi éloigné. Les données de transcriptomiques confirment que TAL18H n'active pas l'expression de ce gène car il n'est pas différentiellement exprimé en présence ou en absence de cet effecteur TAL. Cependant, la cartographie des séquences ARN issues du RNAseq a permis de visualiser une accumulation très importante d'un petit transcrit d'environ 600 pb situé au niveau de cette EBE chez les deux variétés de haricot que nous avons utilisé (BAT93 et JaloEEP558) deux jours après inoculation avec la souche CFPB6546R (**Figure 5**). Nous avons confirmé ce résultat par qRT-PCR (**Figure 6**). En effet, l'expression relative de ce petit transcrit était 3100 fois plus importante après inoculation de la souche sauvage CFPB6546R qu'après inoculation de l'eau chez BAT93 et neuf fois plus importante chez JaloEEP558. Egalement, les deux analyses ont montré que l'expression de ce petit transcrit n'était pas différente après infiltration du mutant de délétion *tal18H* par rapport à l'eau. Ces résultats semblent donc confirmer que TAL18H active l'expression de ce petit transcrit. Nous avons mené des expérimentations *in planta* afin de valider l'accroche de TAL18H à l'EBE prédite (**Figure 7**). Ainsi, nous avons observé que les ArtTAL18H déclenchent une activité GUS importante lorsqu'il se lie à cette EBE prédite et nous n'avons pas observé de différence entre ArtTAL18H-LTPE et ArtTAL18H-MTLA. Egalement, nous avons testé cette capacité d'activation de l'expression du gène *GUS* dans le contexte des séquences d'ADN de BAT93 et JaloEEP558 contenant cette EBE et nous avons observé une activité GUS similaire pour les deux variétés de haricot et que ce soit avec ArtTAL18H-LTPE qu'avec ArtTAL18H-MTLA. Par ailleurs, nous avons aussi testé la spécificité de liaison du RVD HY et l'impact du motif MTLA dans le contexte de cette EBE. Pour ce faire, nous avons produit quatre vecteur *GUS* dont les promoteurs contiennent respectivement une EBE où les quatre bases d'ADN font face aux RVD HY issu d'une répétition commençant par LTPE ou MTLA (**Figure 8**). Les résultats obtenus pour ces quatre EBE, nous ont

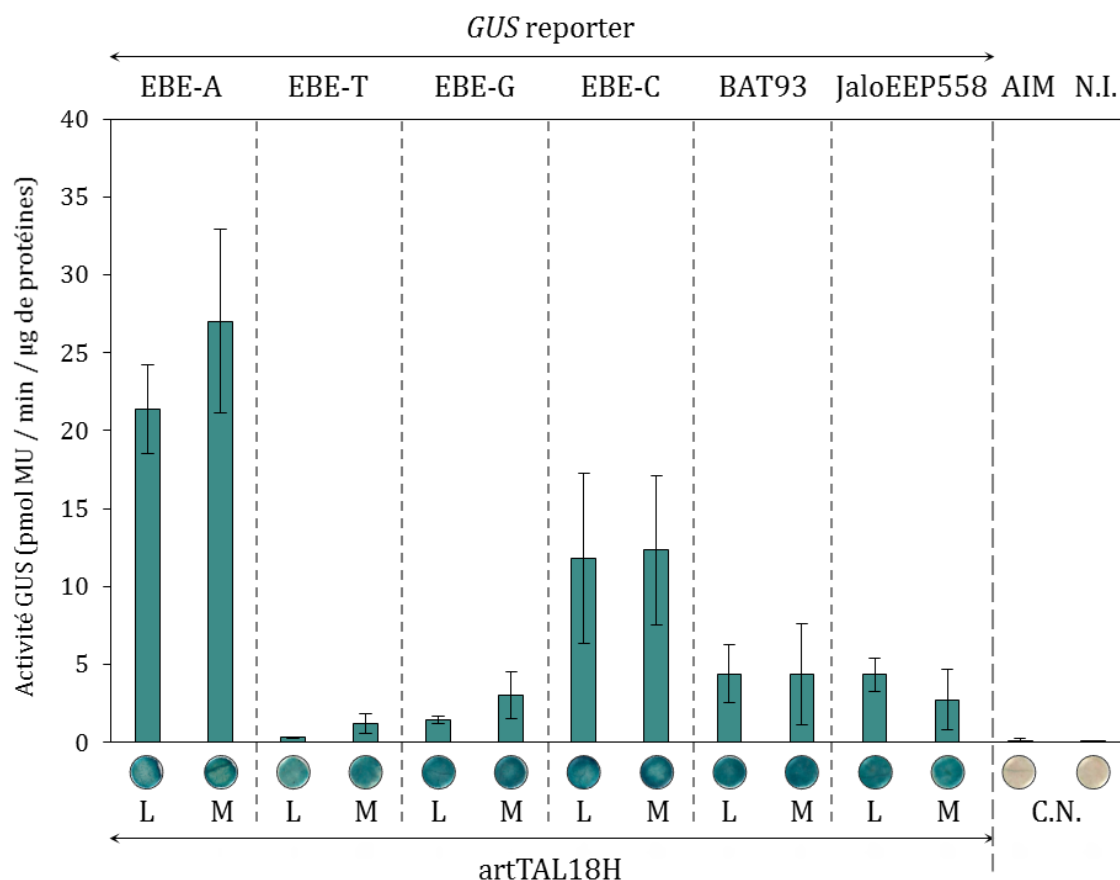


Figure 7 : Spécificité de liaison de TAL18H-LTPE et TAL18H-MTLA pour l'EBE prédite située au niveau du petit transcrit et pour ses variantes. L'activité GUS est exprimée en pmol de 4-méthylumbellifère par minute par µg de protéine totale (pmol MU / min / µg de protéine). Les contrôles négatifs (C.N.) correspondent à des échantillons de feuilles de tabac non inoculés (N.I.) ou inoculés avec de la solution d'infiltration sans bactéries (AIM). La spécificité de liaison de TAL18H a été testée sur l'EBE prédite située au niveau du petit transcrit, sur des variantes (Figure 7) et au niveau des séquences d'ADN correspondantes de BAT93 et JaloEEP558. L'impact du motif MTLA a aussi été testé en utilisant des ArtTAL contenant la séquence de RVD de TAL18H où toutes les répétitions commencent par LTPE (L) et où les répétitions associées aux RVD HY en position 10 et 15 et au RVD NG en position 14 commencent par MTLA (M).

Position :	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Séquence de RVD :		NI	NG	HD	NG	HD	NI	NG	NI	HD	HY	NN	N*	HD	NG	HY	NN	HD	NG
EBE prédite (EBE-C) :	T	A	T	C	T	C	A	T	A	C	C	G	T	C	T	C	A	C	T
Variante EBE-A :	T	A	T	C	T	C	A	T	A	C	A	G	T	C	T	A	A	C	T
Variante EBE-G :	T	A	T	C	T	C	A	T	A	C	G	G	T	C	T	G	A	C	T
Variante EBE-T :	T	A	T	C	T	C	A	T	A	C	T	G	T	C	T	T	A	C	T

Figure 8 : Liste des variants de l'EBE prédite située au niveau du petit transcrit. Les RVD issus de séquences de répétitions commençant par MTLA sont surlignés en jaune. Les bases modifiées pour former les différentes variantes de l'EBE prédite située au niveau du petit transcrit sont indiquées en rouge.

permis de confirmer que le RVD HY se lie préférentiellement à l'adénine (EBE-A), puis la cytosine (EBE-C), puis la guanine (EBE-G) et pour finir la thymine (EBE-T) et que le motif MTLA augmente l'efficacité de liaison du RVD HY à l'adénine (**Figure 7**).

Ensemble, ces résultats sont très intéressants, car il n'a jamais été démontré qu'un effecteur TAL pouvait induire une région intergénique. D'un point vu fonctionnel, il n'y a pas d'intérêt pour la bactérie de conserver un effecteur TAL liant une telle région. Or, TAL18H induit bien cette région chez les deux génotypes de haricot analysés. Si ce petit transcrit est réellement actif, il pourrait correspondre à un précurseur de micro-ARN. Les micro-ARN sont connus pour se lier par complémentarité avec les transcrits d'autres gènes, ce qui aboutit à la répression de l'expression de ces gènes (Mallory & Vaucheret, 2006). Ainsi, TAL18H ciblerait indirectement un gène afin de réprimer son expression et non de l'activer, ce qui n'a jamais été démontré. Des analyses complémentaires devront être réalisées pour confirmer ou infirmer cette hypothèse.

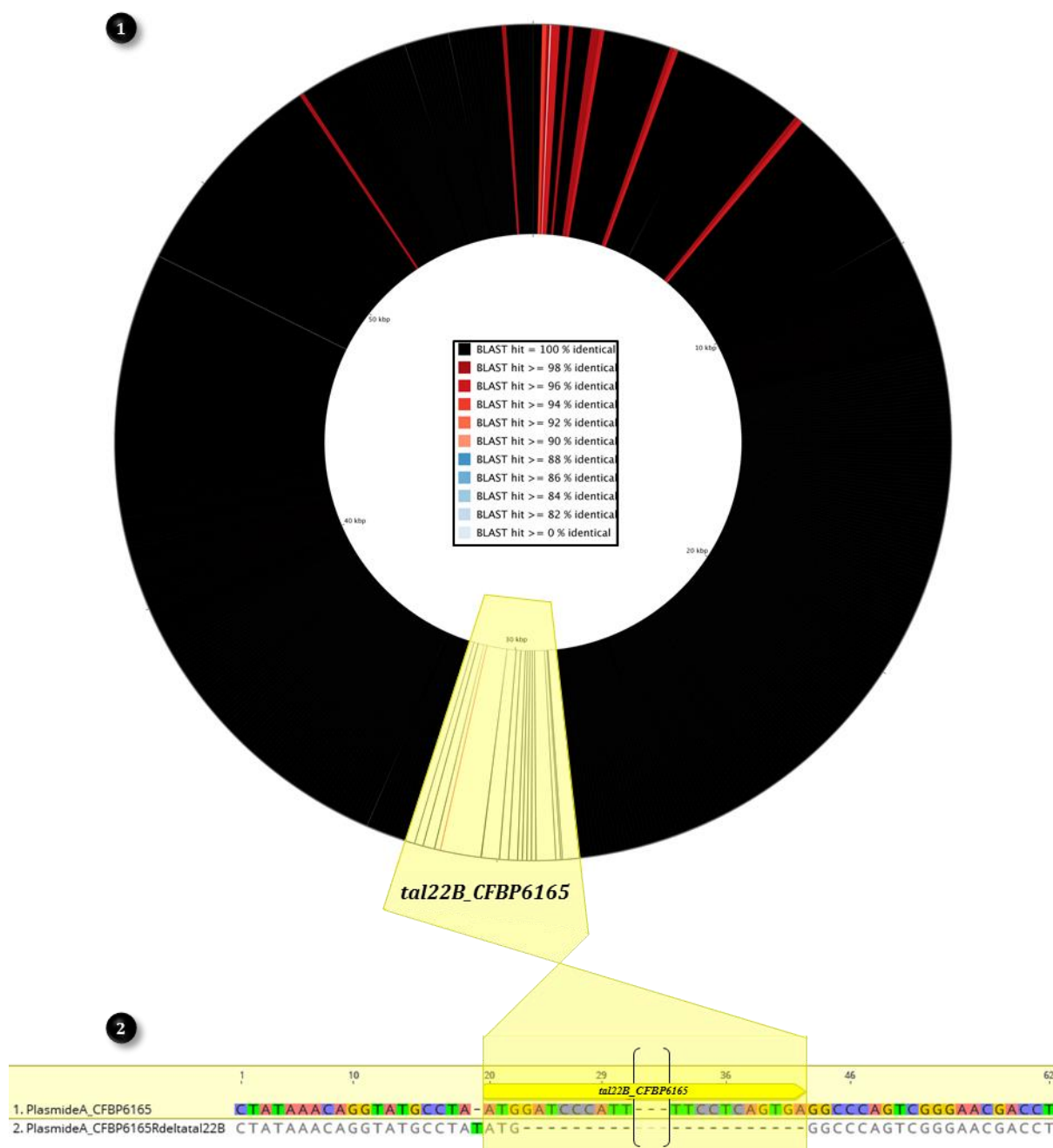


Figure 9 : Délétion du gène *tal22B_CFBP6165*. Le gène *tal22B_CFBP6165* a été éliminé du plasmide A de la souche CFBP6165R avec un protocole en deux étapes : un clonage par la technique « Golden Gate » suivie d'une délétion génomique dirigée avec un dérivatif du vecteur suicide pK18 nommé pΔ13. **(1)** Représentation circulaire des plasmides provenant respectivement de la souche sauvage CFBP6165 et de la souche mutante CFBP6165RΔ*tal22B*. Les séquences ont été comparées et converties dans une carte graphique utilisant CGView (Grant *et al.*, 2012) avec la souche CFBP6165 comme référence. Les couleurs diffèrent selon le pourcentage d'identité (voir la légende). La localisation du gène *tal22B_CFBP6165* éliminé est mise en évidence par une zone jaune. **(2)** Alignement des plasmides A provenant des souches CFBP6165 et CFBP6165RΔ*tal22B* réalisé avec Geneious et centré sur le gène *tal22B_CFBP6165*. La majeure partie de la séquence du gène *tal22B_CFBP6165* n'est pas représentée sur l'alignement et est indiquée par des crochets. Le gène *tal22B_CFBP6165* est éliminé à partir du second codon jusqu'au codon STOP inclus.

4. Production de mutants permettant l'étude de l'ensemble des effecteurs TAL des souches responsables de CBB

Afin d'étudier l'ensemble des *tal* retrouvés dans les souches responsables de CBB, nous avons produit un second mutant de délétion, CFBP6165R Δ *tal22B*, correspondant à une souche de la lignée fuscans de *X. citri* pv. *fuscans* dépourvue de *tal*. Tout comme pour le mutant de délétion *tal18H*, *tal22B_CFBP6165* étant plasmidique (plasmide A) (Ruh *et al.*, 2017a), nous avons contrôlé la délétion de ce *tal* par PCR et par séquençage SMRT PacBio pour nous assurer de l'absence de toutes les copies de *tal22B_CFBP6165* dans la souche mutante. Ainsi, la comparaison des séquences des plasmides A portant ou non *tal22B_CFBP6165* chez la souche CFBP6165R et chez la CFBP6165R Δ *tal22B* nous a permis de visualiser que *tal22B_CFBP6165* est bien délété du codon START au codon STOP (**Figure 9**).

Nous avons complété ce mutant de délétion *tal22B* avec des *ArtTAL* et *tal* clonés correspondant à toute la diversité de *tal* existant chez *Xcf-Xpp* (**Tableau 4**; Ruh *et al.*, 2017). Pour tous les mutants, nous avons constaté en milieu riche des croissances bactériennes similaires à celles observées pour les souches témoins correspondant à la souche sauvage, au mutant de délétion ou au mutant de délétion complété avec un vecteur vide (**Figure 10**). Des tests de pouvoir pathogène et expériences de RNAseq utilisant ces souches ont commencé dans l'équipe, ce qui permettra à terme de caractériser le rôle des différents TAL de *Xcf* et *Xpp* dans l'interaction *Xanthomonas*-haricot.

Tableau 4 : Listes des mutants produits pour l'étude des effecteurs TAL des souches responsables de la CBB. *Xpp* : *X. phaseoli* pv. *phaseoli* ; *Xcf* : *X. citri* pv. *fuscans* ; *ArtTAL* : *tal* artificiel ; Rif : rifamycine ; Genta : gentamycine.

Souche	Caractéristique	Gène <i>tal</i> présent	Construction portée	Résistance
CFBP6546R	Souche sauvage (<i>Xpp</i> lignée NF1)	<i>tal19I_CFBP6546R</i> et <i>tal18H_CFBP6546R</i>	aucune	Rif
CFBP6546R Δ <i>tal18H</i>	Mutant de délétion du gène <i>tal18H</i>	<i>tal19I_CFBP6546R</i>	aucune	Rif
CFBP6546R Δ <i>tal18H::arttal18H-LTPE</i>	Mutant complémenté avec l' <i>arttal tal18H-LTPE</i>	<i>tal19I_CFBP6546R</i> et <i>arttal18H-LTPE</i>	pSKX1:: <i>arttal18H-LTPE</i>	Rif et Genta
CFBP6546R Δ <i>tal18H::arttal18H-MTLA</i>	Mutant complémenté avec l' <i>arttal tal18H-MTLA</i>	<i>tal19I_CFBP6546R</i> et <i>arttal18H-MTLA</i>	pSKX1:: <i>arttal18H-MTLA</i>	Rif et Genta
CFBP6546R Δ <i>tal18H::pSKX1</i>	Mutant complémenté avec le vecteur vide pSKX1	<i>tal19I_CFBP6546R</i>	pSKX1	Rif et Genta
CFBP6165R	Souche sauvage (<i>Xcf</i> lignée <i>fuscans</i>)	<i>tal22B_CFBP6165R</i>	aucune	Rif
CFBP6165R Δ <i>tal22B</i>	Mutant de délétion du gène <i>tal22B</i>	aucun	aucune	Rif
CFBP6165R Δ <i>tal22B::arttal23A</i>	Mutant complémenté avec l' <i>arttal tal23A</i>	<i>arttal23A</i>	pSKX1:: <i>arttal23A</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal22B</i>	Mutant complémenté avec l' <i>arttal tal22B</i>	<i>arttal22B</i>	pSKX1:: <i>arttal22B</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal22C</i>	Mutant complémenté avec l' <i>arttal tal22C</i>	<i>arttal22C</i>	pSKX1:: <i>arttal22C</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal21D</i>	Mutant complémenté avec l' <i>arttal tal21D</i>	<i>arttal21D</i>	pSKX1:: <i>arttal21D</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal19E</i>	Mutant complémenté avec l' <i>arttal tal19E</i>	<i>arttal19E</i>	pSKX1:: <i>arttal19E</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal19I</i>	Mutant complémenté avec l' <i>arttal tal19I</i>	<i>arttal19I</i>	pSKX1:: <i>arttal19I</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal20F</i>	Mutant complémenté avec l' <i>arttal tal20F</i>	<i>arttal20F</i>	pSKX1:: <i>arttal20F</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal18G</i>	Mutant complémenté avec l' <i>arttal tal18G</i>	<i>arttal18G</i>	pSKX1:: <i>arttal18G</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal18H-LTPE</i>	Mutant complémenté avec l' <i>arttal tal18H-LTPE</i>	<i>arttal18H-LTPE</i>	pSKX1:: <i>arttal18H-LTPE</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::arttal18H-MTLA</i>	Mutant complémenté avec l' <i>arttal tal18H-MTLA</i>	<i>arttal18H-MTLA</i>	pSKX1:: <i>arttal18H-MTLA</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::pSKX1</i>	Mutant complémenté avec le vecteur vide pSKX1	aucun	pSKX1	Rif et Genta
CFBP6165R Δ <i>tal22B::tal23A_CFBP4885</i>	Mutant complémenté avec le gène <i>tal23A_CFBP4885</i> cloné	<i>tal23A_CFBP4885</i>	pSKX1auto:: <i>tal23A_CFBP4885</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::tal23A_CFBP6166</i>	Mutant complémenté avec le gène <i>tal23A_CFBP6166</i> cloné	<i>tal23A_CFBP6166</i>	pSKX1auto:: <i>tal23A_CFBP6166</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::tal22B_CFBP6165</i>	Mutant complémenté avec le gène <i>tal22B_CFBP6165</i> cloné	<i>tal22B_CFBP6165</i>	pSKX1auto:: <i>tal22B_CFBP6165</i>	Rif et Genta
CFBP6165R Δ <i>tal22B::pSKX1auto</i>	Mutant complémenté avec le vecteur vide pSKX1_autoligase	aucun	pSKX1autoligase	Rif et Genta

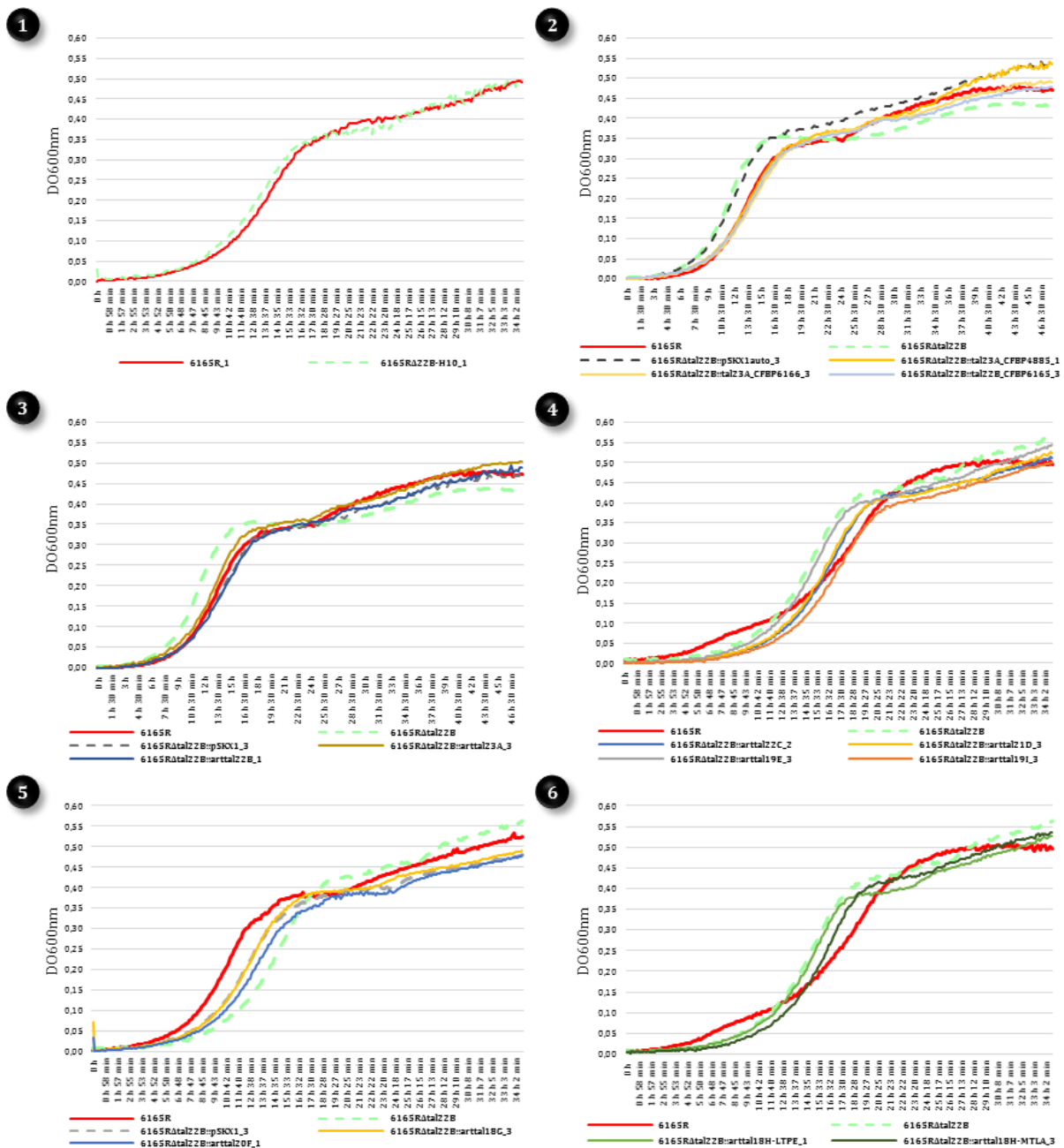


Figure 10 : Dynamique des tailles de population de l'ensemble des mutants produits pour l'étude des TALE des souches responsables de la CBB. Seuls les mutants conservés pour des expérimentations *in planta* sont présentés. **(1)** Mutant CFBP6165RΔtal22B. **(2)** Mutants CFBP6165RΔtal22B complémentés avec pSKX1auto (vecteur vide utilisé pour le clonage des *tal* naturels) et les *tal23A_CFBP4885*, *tal23A_CFBP6166* et *tal22B_CFBP6165* clonés dans pSKX1auto. **(3)** Mutants CFBP6165RΔtal22B complémentés avec pSKX1 (vecteur vide utilisé pour le clonage des *ArtTAL*) et avec les *ArtTAL23A* et *ArtTAL22B* clonés dans pSKX1. **(4)** Mutants CFBP6165RΔtal22B complémentés avec les *ArtTAL22C*, *ArtTAL21D*, *ArtTAL19E* et *ArtTAL19I* clonés dans pSKX1. **(5)** Mutants CFBP6165RΔtal22B complémentés avec pSKX1 et les *ArtTAL20F* et *ArtTAL18G* clonés dans pSKX1. **(6)** Mutants CFBP6165RΔtal22B complémentés avec les *ArtTAL18H-LTPE* et *ArtTAL18H-MTLA* clonés dans pSKX1.

DISCUSSION GÉNÉRALE

L'objectif de ma thèse était d'identifier les déterminants de la spécificité d'hôte des *Xanthomonas* pathogènes du haricot commun et déterminer le rôle des effecteurs TAL dans cette spécificité. Je vais discuter, dans ce dernier chapitre, de la méthodologie et des limites de l'étude de la spécificité d'hôte, du processus évolutif qui a conduit les souches de *Xanthomonas* à s'adapter au haricot commun, des spécificités adaptatives des effecteurs TAL et des perspectives de ces travaux dans la lutte contre la CBB.

I. Aspects méthodologiques et limites de l'étude de la spécificité d'hôte

1. Importance du jeu de données et des méthodes d'analyse

Le génome des agents responsables de la CBB contient en moyenne 4500 gènes et nous avons mis à jour 116 gènes portant des caractéristiques uniques à ces souches, dont 45 gènes qui contiennent au moins un acide aminé spécifique des quatre lignées génétiques NF1, NF2, NF3 et fuscans (Chapitre 1-II). Ainsi, 1 à 3 % du génome est potentiellement impliqué dans la spécificité des *Xanthomonas* au haricot commun. Ces chiffres sont en accord avec les résultats d'une précédente étude qui a identifié plus de 100 gènes transférés horizontalement de la lignée fuscans (*X. citri* pv. *fuscans*) vers la lignée NF1 (*X. phaseoli* pv. *phaseoli*) (Aritua *et al.*, 2015). Cependant, notre analyse et celle d'Aritua *et al.* (2015) ne comptent que neuf gènes en commun. Mis à part un gène impliqué dans le système de régulation à deux composants, reconnu comme étant impliqué dans l'adaptation à l'hôte des *Xanthomonas* (Mhedbi-Hajri *et al.*, 2011), aucun des autres gènes ne codait pour des fonctions décrites comme importantes pour la virulence. Les différences dans les jeux de données utilisés peuvent expliquer ce petit nombre de gènes en commun, puisque les deux études ne sont pas basées sur les mêmes génomes et les lignées NF2 et NF3 ne font pas partie du jeu de données d'Aritua *et al.* (2015). De plus, les objectifs et donc le type d'analyses menées étaient également différents puisque notre analyse cherchait des caractéristiques à la fois communes et spécifiques des quatre lignées génétiques de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli*, alors que l'étude d'Aritua *et al.* (2015) cherchait des caractéristiques communes mais pas forcément spécifiques aux lignées NF1 et fuscans. Notre étude a combiné plusieurs approches de génomique comparative, telles que la construction d'une matrice de présence / absence de groupes d'orthologues, la recherche de gènes monophylétiques, et l'utilisation de l'outil SkIf basé sur la recherche de k-mer partagés. Bien que la recherche de k-mer se soit avérée beaucoup plus efficace que d'autres méthodes pour trouver des gènes présentant des caractéristiques uniques aux souches responsables de la CBB, ces différentes analyses ont donné des résultats non redondants, ce qui souligne l'importance de combiner plusieurs approches pour chercher des

déterminants de la spécificité d'hôte. Enfin, ces analyses ont été menées sur des souches présentant une convergence pathologique vers le haricot, ce qui a facilité la recherche de gènes transmis par transfert horizontal, et de faire des hypothèses fortes sur l'implication de ces gènes dans la spécificité d'hôte. D'autres cas de convergence pathologique existent chez les *Xanthomonas*, comme c'est le cas entre *X. euvesicatoria* pv. *euvesicatoria*, *X. euvesicatoria* pv. *perforans*, *X. vesicatoria* et *X. gardneri* qui forment un groupe polyphylétique de souches pathogènes du poivron et de la tomate (Schwartz *et al.*, 2015), et il serait intéressant de réaliser le même type d'analyses sur ces pathovars. Au final, le jeu de donnée et la méthodologie choisis lors d'analyses de génomique comparative impactent fortement les résultats obtenus et constituent donc un verrou dans la recherche de déterminants de la spécificité d'hôte. Pour diminuer au maximum les effets de ce verrou scientifique, il serait préférable d'utiliser un nombre important de génomes représentant au mieux la diversité des souches étudiées et de combiner plusieurs approches de génomique comparative pour d'avantage consolider les résultats obtenus.

D'autre part, la méthodologie de séquençage et d'assemblage des génomes affecte le jeu de données et par conséquent les résultats d'analyses de génomique comparative. Ce qui reflète sans doute le mieux ce propos sont les différences observées dans les séquences de gènes codant des effecteurs TAL en fonction du séquençage. En effet, nous avons montré que malgré l'obtention d'un génome de très bonne qualité combinant des données de séquençage Sanger, Illumina et 454 pour la souche CFBP4834-R (Darrasse *et al.*, 2013b), les régions centrales de répétitions des deux *tal* présents dans ce génome contenaient des erreurs d'assemblage (Ruh *et al.*, 2017a). D'autre part, les séquences SMRT de *tal* provenant de souches de *X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans* comportaient des erreurs de type insertions/délétions qui ont été corrigées par séquençage de produits PCR (Ruh *et al.*, 2017a). Egalement, le nombre de *tal* peut différer en fonction de la méthodologie de séquençage. Par exemple, les souches CFBP6164 et CFBP6165 ne possèdent pas le même nombre de *tal* lorsque nous comparons leurs génomes séquencés en SMRT dans notre étude avec ceux séquencés en Illumina dans l'étude d'Aritua *et al.* (2015). Ces différences dues à la stratégie de séquençage choisie sont inévitables et constituent donc un autre verrou dans la recherche de déterminants de la spécificité d'hôte. En effet, en fonction du séquençage un bon nombre d'informations diverge ou peut être perdu. Il est donc important, lors d'analyses de génomiques comparative, d'utiliser des jeux de données homogènes en terme de stratégie de séquençage. Au final, le séquençage SMRT permet d'obtenir de longs fragments favorisant l'assemblage de séquences complexes, mais les séquences obtenues comportent de nombreuses erreurs, tandis que le séquençage Illumina donne de petits fragments mais avec un très faible taux d'erreurs. Pour minimiser les erreurs dues au séquençage, il est possible d'utiliser des génomes dont l'assemblage est basé sur le séquençage SMRT et dont les séquences sont corrigées par Illumina (Koren *et al.*, 2012).

2. Validation fonctionnelle de gènes candidats

Le Tn-seq (Transposon sequencing) est une technique de criblage de mutants à haut débit en plein essor qui pourrait constituer, en complément d'une analyse de génomique comparative, une étape préliminaire pour la validation fonctionnelle de gènes candidats impliqués dans la spécificité d'hôte. Le Tn-seq consiste à utiliser une banque de mutants d'insertion de transposon combinée à un séquençage à haut débit des sites d'insertion des transposons (van Opijnen *et al.*, 2009; van Opijnen & Camilli, 2013). Ainsi, s'agissant de bactéries phytopathogènes, seuls les mutants possédant l'arsenal suffisant pour infecter la plante devraient être amplifiés après inoculation sur plante hôte, les mutants sous-abondants sont ceux présentant l'inactivation de gènes codant des fonctions essentielles à l'interaction avec la plante hôte. Le séquençage des populations de mutants avant et après inoculation serait donc un moyen intéressant pour identifier les gènes importants pour l'établissement de la maladie. Dans le cadre d'une étude de la spécificité d'hôte des agents responsables de la CBB, il serait intéressant de réaliser des expériences de Tn-seq avec des banques de mutants provenant des quatre lignées génétiques de *X. phaseoli* pv. *phaseoli* et *X. citri* pv. *fuscans*. Les résultats obtenus seraient complémentaires aux données de génomique comparative et constitueraient une étape préliminaire à la validation fonctionnelle de gènes potentiellement impliqués dans la spécificité d'hôte.

Si la spécificité d'hôte est la conséquence de l'acquisition de plusieurs dizaines de gènes, la mutation d'un seul d'entre eux est-elle suffisante pour espérer observer un phénotype ? En effet, au regard de nos résultats et de différentes analyses de génomique comparative, il apparaît évident que la spécificité d'hôte est régie par un grand répertoire de gènes (Aritua *et al.*, 2015; Schwartz *et al.*, 2015; Hersemann *et al.*, 2017). Parmi ces gènes certaines fonctions peuvent être redondantes ou complémentaires, ce qui empêcherait d'observer un phénotype différentiel lors de l'inoculation d'une souche mutante pour un seul gène. Par exemple, parmi les protéines codées par nos 116 gènes candidats il y a quatre transporteurs Ton-B dépendants, trois ET3, trois protéines impliquées dans le système de sécrétion de type IV et six protéines impliquées dans les systèmes de régulation à deux composants. Il serait donc intéressant pour ces gènes de faire des séries de mutants, comme cela a été réalisé chez *Pseudomonas syringae* pour étudier les fonctions de différents ET3 (Cunnac *et al.*, 2009). Ces analyses pourraient être couplées à des études du transcriptome bactérien sur plantes hôtes ou non-hôtes afin d'analyser si les gènes potentiellement impliqués dans la spécificité d'hôte sont induits ou réprimés en contexte hôte ou non-hôte.

II. L'adaptation des *Xanthomonas* au haricot commun : un processus évolutif en plusieurs étapes

1. Importance du transfert horizontal

Les résultats présentés dans cette thèse retracent le processus évolutif de l'adaptation des *Xanthomonas* au haricot commun et soulignent l'importance du transfert horizontal de gènes dans ce processus. En effet, l'acquisition de matériel génétique par transfert horizontal peut potentiellement être responsable d'une modification de la gamme d'hôte ou d'un saut d'hôte souvent liées à l'émergence de nouvelles maladies. Par exemple, chez l'espèce *X. arboricola*, l'émergence des pathovars *pruni*, *juglandis* et *corylina* serait apparemment liée à l'acquisition de huit ET3 par THG (Merda *et al.*, 2016). Il a également été rapporté au Canada et en Russie que certaines souches de *X. campestris* pv. *raphani* sont capables d'infecter la tomate, alors que ce pathovar est connu pour être responsable de la tache foliaire chez les Brassicacées (Kuflu & Cuppels, 1997; Punina *et al.*, 2009). De la même façon, en Afrique du Sud, des souches de *X. vasicola* pv. *vasculorum*, connues pour infecter la canne à sucre et le maïs, auraient acquis la capacité à provoquer des brûlures sur les feuilles d'*Eucalyptus* et un dépérissement de ce nouvel hôte (Coutinho *et al.*, 2015). Il a été suggéré que ces sauts d'hôte seraient potentiellement la conséquence d'évènements de THG et / ou de recombinaison (Jacques *et al.*, 2016), qui auraient donc permis à ces souches d'acquérir, respectivement, des déterminants de la spécificité à la tomate et à l'eucalyptus. Ces modifications de la gamme d'hôte peuvent être très rapides car un faible nombre d'évènements évolutifs est suffisant. Par exemple, la bactérie commensale *Pantoea agglomerans* a acquis la capacité de causer des tumeurs sur *Gypsophila* et betterave après acquisition d'un plasmide de 10 kbp portant tous les gènes nécessaires à la formation de galles et à la spécificité d'hôte (Barash & Manulis-Sasson, 2009). Ainsi, une meilleure connaissance des populations présentes dans les compartiments non hôtes ou sauvages d'une culture donnée permettrait de caractériser les potentiels réservoirs de gènes adaptatifs représentant une menace pour la culture en question.

2. Scénarios évolutifs pour l'acquisition de la spécificité d'hôte des *Xanthomonas* au haricot commun

L'émergence d'une maladie pourrait être la conséquence de processus évolutifs complexes, dont il n'est pas toujours aisé -a posteriori- de retracer l'histoire. Par exemple, l'émergence du flétrissement bactérien du bananier et du bananier plantain en Afrique de l'Est (Tushemereirwe *et al.*, 2004) peut être expliquée par deux scénarios évolutifs. Cette maladie pourrait être due à une modification de la gamme d'hôte de *X. vasicola* pv. *musacearum*, agent

responsable du flétrissement bactérien du bananier d'Abyssinie (*Ensete ventricosum*), une espèce végétale proche du bananier (Smith *et al.*, 2008), mais elle pourrait également être la conséquence d'un saut d'hôte de *X. vasicola* pv. *holcicola* ou *X. vasicola* pv. *vasculorum*, respectivement pathogènes du sorgho et de la canne à sucre (Aritua *et al.*, 2008). La spécificité d'hôte des *Xanthomonas* au haricot commun semble elle aussi être la conséquence d'un processus évolutif complexe. En effet, dans les Chapitres 1 et 2, nous avons montré que plusieurs événements de transferts horizontaux, dont notamment des transferts de plasmides, sont impliqués dans ce processus adaptatif. L'observation de plusieurs transferts horizontaux successifs suggère que les ancêtres des quatre lignées d'agents responsables de la CBB vivaient en sympatrie ce qui leur aurait permis d'échanger du matériel génétique à plusieurs reprises. Ces événements évolutifs auraient en fait très bien pu se produire à l'île de la Réunion, puisque les quatre lignées génétiques infectant le haricot commun y sont présentes. Toutes y ont été échantillonnées en 2000. Cependant, l'observation de plusieurs vagues de transferts horizontaux plus ou moins récentes entre les lignées génétiques des agents de la CBB suggère deux scénarios évolutifs possibles pour l'acquisition de sa spécificité au haricot commun. Cette spécificité aurait pu être acquise peu à peu par la succession de plusieurs transferts horizontaux ou bien elle aurait été acquise lors d'un événement fondateur ancien suivi de plusieurs événements adaptatifs supplémentaires.

Dans la littérature, il a été rapporté que les souches infectant le haricot commun sont capables d'infecter d'autres légumineuses, telles que *Phaseolus lunatus*, *P. acutifolius*, *Vigna aconitifolia*, *V. radiata*, *V. angularis*, *V. mungo* et *V. umbellata* (Bradbury, 1986; Opio *et al.*, 1996). D'autre part, il existe des rapports contradictoires concernant la présence naturelle des agents responsables de la CBB sur le haricot écarlate (*P. coccineus*) et le niébé (*V. unguiculata*) (Bradbury, 1986; Gilbertson & Maxwell, 1992). Aujourd'hui, il est considéré que, presque certainement, tout signalement de *X. citri* pv. *fuscans* ou *X. phaseoli* pv. *phaseoli* sur niébé correspond en réalité à *X. citri* pv. *vignicola*, l'agent causal de la brûlure bactérienne du niébé (OEPP/CABI, 1986; EFSA PLH Panel (EFSA Panel on Plant Health), 2014). Parmi les souches phylogénétiquement proches des agents responsables de la CBB, telles que *X. citri* pv. *vignicola*, *X. citri* pv. *aurantifolii*, *X. citri* pv. '*dieffenbachiae*', *X. phaseoli* pv. *dieffenbachiae* et *X. phaseoli* pv. *manihotis*, seul *X. citri* pv. *vignicola* a été rapporté comme étant capable de provoquer des symptômes sur haricot (*P. vulgaris*), mais aussi sur lablab (*Dolichos lablab*) en plus de son hôte naturel, le niébé (Gena *et al.*, 2009). Ainsi, il semblerait qu'une partie du processus évolutif ayant conduit les *Xanthomonas* à s'adapter au haricot commun soit commune aux agents responsables de la CBB et à *X. citri* pv. *vignicola*, ce qui serait en accord avec un événement fondateur lié à l'acquisition de la capacité à infecter les genres *Phaseolus* et *Vigna* et *Dolichos* formant un clade au sein des Leguminosae, dont la divergence est estimée à 8 millions d'années (Lavin *et al.*, 2005). Un tel scénario serait en accord avec les résultats présentés au Chapitre 1, qui suggèrent que la capacité à infecter le haricot commun était d'abord

apparue chez les trois lignées de *X. citri* pv. *fuscans*, qui sont extrêmement proches phylogénétiquement de *X. citri* pv. *vignicola*, et que la lignée NF1 de *X. phaseoli* pv. *phaseoli* aurait acquis cette capacité seulement dans un second temps, par transferts horizontaux massifs provenant de *X. citri* pv. *fuscans*.

Les résultats correspondant à l'histoire évolutive des effecteurs TAL des souches infectant le haricot apportent d'autres éléments de réponse quant aux événements liés à la spécificité d'hôte des agents responsables de la CBB (Ruh *et al.*, 2017a). En effet, *tal23A* est commun aux quatre lignées génétiques de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli*, et aurait été transféré horizontalement de l'une des trois lignées de *X. citri* pv. *fuscans* (NF2, NF3 et *fuscans*) à la lignée NF1 de *X. phaseoli* pv. *phaseoli* puis aurait divergé au cours du temps, conduisant aujourd'hui à la présence de différents allèles portant différentes combinaisons de RVD dans les différentes souches infectant le haricot commun. D'autre part, il semblerait que *tal18H* ait été échangé plus tard entre les lignées NF1 et *fuscans* puis est resté très conservé par la suite. La conservation de ces *tal* dans des agents responsables de la CBB phylogénétiquement éloignés au cours du temps et à travers le monde entier semble confirmer leur importance pour ces souches. En effet, dans le Chapitre 3, nous avons démontré l'importance de *tal18H* dans l'agressivité de la souche et son impact sur le transcriptome du haricot commun. Par contre, la perte de TAL18H n'a pas abouti à la perte totale de virulence, et les souches des lignées NF2 et NF3, bien que virulentes, ne possèdent pas TAL18H. Ainsi, TAL18H ne paraît pas impliqué dans la spécificité d'hôte, mais son acquisition permet d'augmenter l'agressivité des agents responsables de la CBB sur haricot. Par la suite, il serait intéressant de cloner *tal18H* afin de transformer des souches appartenant aux lignées NF2 et / ou NF3 dans le but d'étudier l'impact de ce *tal* sur l'agressivité de ces souches. De manière intéressante, dans le Chapitre 2, nous avons montré que les *tal* des souches infectant le haricot sont phylogénétiquement proches de ceux de *X. citri* pv. '*dieffenbachiae*' et *X. citri* pv. *aurantifolii* et *X. phaseoli* pv. *manihotis* ce qui suggère que les *tal* de ces souches sont issus d'un transfert horizontal ancestral de *tal* chez les ancêtres des groupes 9.4 et 9.6 tels que décrits par Rademaker *et al.* (2005). A l'inverse, les *tal* de *X. citri* pv. *vignicola* sont phylogénétiquement éloignés de ceux des pathogènes du haricot ce qui suggère que *X. citri* pv. *vignicola* a perdu son/ses *tals* ancestraux, et en a acquis de nouveaux. Vu les proximités phylogénétiques qu'il existe entre *X. citri* pv. *vignicola* et les souches pathogènes du haricot d'une part, et entre leurs hôtes végétaux *V. unguiculata* et *P. vulgaris*, respectivement d'autre part, et vu les divergences dans la séquence de leurs *tals*, il serait intéressant de compléter des souches de *X. citri* pv. *fuscans* ou *X. phaseoli* pv. *phaseoli* avec des *tal* de *X. citri* pv. *vignicola*, et inversement, et de tester si ces *tal* ont un impact sur le pouvoir pathogène de ces souches et/ou la capacité de modifier la gamme d'hôte de ces souches.

Au final, l'ensemble de nos résultats sont en faveur d'un scénario évolutif correspondant à un évènement initial et fondateur de la spécificité des *Xanthomonas* aux haricots communs suivi d'acquisitions secondaires responsables d'une adaptation de plus en plus poussée de ces pathogènes à leur nouvel hôte.

III. Spécificités adaptatives des effecteurs TAL

1. L'adaptation est inhérente à la fonction des effecteurs TAL

Au cours de leur histoire évolutive, il a été démontré que les *tal* ont subi de nombreux évènements évolutifs différents comme la mutation, le transfert horizontal, mais aussi la duplication et la recombinaison, notamment favorisées par la transposition (Bonas *et al.*, 1993; Noel *et al.*, 2003; Lau *et al.*, 2014; Booher *et al.*, 2015; Ferreira *et al.*, 2015; Hutin *et al.*, 2015; Wilkins *et al.*, 2015; Erkes *et al.*, 2017). Cependant, les différents domaines fonctionnels d'un effecteur TAL ne subissent pas la même pression de sélection. En effet, les régions N- et C-terminales sont très conservées car elles confèrent des capacités à la fois essentielles et communes à tous les effecteurs TAL (sécrétion dans la cellule hôte, migration jusqu'au noyau et activation de la transcription), tandis que la région centrale de répétitions est plus variable, notamment au niveau des RVD, car elle leur confère une capacité à la fois essentielle et unique : cibler une séquence spécifique dans le génome de la plante hôte. Ainsi, l'histoire évolutive des *tal* est riche d'évènements évolutifs qui ont permis à ces effecteurs de se disséminer, d'acquérir leur capacité d'accroche à l'ADN, mais aussi de s'adapter au génome de la plante hôte afin de cibler et de lier une EBE spécifique située dans le promoteur d'un gène particulier pour en activer son expression.

Aujourd'hui, de nombreuses séquences de *tal* sont disponibles (da Silva *et al.*, 2002; Kay *et al.*, 2005; Wilkins *et al.*, 2015; Huguet-Tapia *et al.*, 2016; Peng *et al.*, 2016), ce qui a permis d'observer une grande diversité dans les séquences de RVD des effecteurs TAL et par conséquent dans les EBE potentiellement ciblées par ces effecteurs. Cependant, il apparaît clairement qu'il existe une convergence pour les gènes ciblés. En effet, de nombreux facteurs de transcription et de gènes appartenant à la famille SWEET, impliqué dans l'export de sucre, sont ciblés par de multiples effecteurs TAL dans des pathosystèmes variés tels que *X. oryzae* pv. *oryzae* chez le riz, *X. citri* pv. *citri* et pv. *aurantifolii* chez les agrumes, *X. phaseoli* pv. *manihotis* chez le manioc et *X. gardneri* et *X. euvesicatoria* sur tomate/poivron (Yang & White, 2004; Yang *et al.*, 2006; Al-Saadi *et al.*, 2007; Kay *et al.*, 2007, 2009; Sugio *et al.*, 2007; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013; Cohn *et al.*, 2014; Hu *et al.*, 2014; Zhou *et al.*, 2015; Abe & Benedetti, 2016; Schwartz *et al.*, 2017). Ainsi, il semble que ces cibles jouent un rôle récurrent dans le développement de

bactérioses causées par *Xanthomonas* et constituent donc une piste intéressante lors de la recherche de gènes ciblés par de nouveaux effecteurs TAL. Par ailleurs, il a été démontré que différents effecteurs TAL de différentes souches infectant une même plante peuvent cibler la même EBE ou des EBE chevauchantes situées dans le promoteur d'un même gène (Hutin *et al.*, 2015). En effet, quatre effecteurs TAL de différentes souches de *X. oryzae* pv. *oryzae* ciblent quatre EBE situées dans le promoteur du gène *OsSWEET14* du riz (Yang & White, 2004; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013) et cinq effecteurs TAL de différentes souches de *X. citri* pv. *citri* et *X. citri* pv. *aurantifolii* ciblent deux EBE situées dans le promoteur du gène *CsLOB1* d'agrumes (Al-Saadi *et al.*, 2007; Hu *et al.*, 2014). Ces exemples d'effecteurs TAL portant des séquences de RVD différentes mais ciblant un même gène souligne d'avantage qu'il existe de la convergence évolutive chez ces effecteurs.

2. Spécificités adaptatives de TAL18H

Nos résultats de prédictions des cibles potentielles de TAL18H combinés avec nos données de transcriptomique semblent indiquer que TAL18H ne cible pas un facteur de transcription ou un gène *SWEET*, mais plutôt une cible jamais démontrée : le gène *NDH18* codant une NAD(P)H déshydrogénase. De manière intéressante, il semble que TAL18H ait en fait la capacité de cibler différents gènes. Cette capacité a déjà été démontrée par le passé. En effet, nous savons que trois effecteurs TAL de *X. citri* pv. *citri* ont pour cible le gène *CsLOB1* et sont également capable de cibler une EBE située dans le promoteur du gène *CsSWEET11*, sans que l'on ait démontré son rôle dans le développement de la maladie (Al-Saadi *et al.*, 2007; Hu *et al.*, 2014). Dans notre cas, TAL18H aurait la capacité de cibler *NDH18*, mais aussi un petit ARN et potentiellement *PL11*, un gène codant pour une pectate lyase impliquée dans la dégradation de la paroi cellulaire végétale. Fait intéressant, il a récemment été démontré qu'un effecteur TAL de *X. gardneri* ciblait indirectement une pectate lyase (Schwartz *et al.*, 2017). Ainsi, si TAL18H est capable d'activer l'expression de *PL11*, cela constituerait alors la première preuve de convergence évolutive d'effecteurs TAL pour des gènes codant des pectates lyases, qui deviendraient alors une piste intéressante pour la recherche de cibles de TAL. Cependant, cette capacité des effecteurs TAL à cibler plusieurs séquences dans un même génome pourrait complexifier la recherche de cibles essentielles au pouvoir pathogène puisqu'à priori elles ne jouent pas toutes un rôle dans le développement de la maladie.

Nous avons également analysé la région centrale de répétitions de TAL18H, et plus particulièrement les séquences de ces répétitions. La séquence des répétitions est habituellement relativement conservée sauf au niveau des RVD qui sont hypervariables. Cependant, pour TAL18H, nous avons observé une conservation stricte de la séquence de RVD dans les souches des

lignées NF1 et fuscans étudiées, et une certaine variabilité dans les séquences de répétitions (Ruh *et al.*, 2017a). Ainsi, il semblerait que la pression de sélection ait agi contre la diversification des RVD de TAL18H ce qui confirme son importance fonctionnelle chez *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* et son implication dans l'adaptation des *Xanthomonas* au haricot commun.

D'autre part, nous nous sommes également intéressés à un RVD peu commun que possède TAL18H à deux reprises : le RVD HY (Chapitre 3-II). Nous avons caractérisé ses spécificités de liaison et avons démontré que le RVD HY cible préférentiellement l'adénine. Ainsi, TAL18H possède à deux reprises un RVD peu commun mais conservé dans les souches que nous avons étudiées. Cela suggère donc que ce RVD devrait conférer un avantage à TAL18H et donc aux souches le possédant, mais lequel ? Nous avons peut-être un début de réponse à cette question. En effet, lors de nos analyses nous avons également découvert que la séquence de répétitions portant le RVD HY différait légèrement des autres : elle commençait toujours par un motif MTLA, jamais encore décrit dans la littérature. Nous avons analysé ce motif et avons montré qu'il permettait d'augmenter l'efficacité de liaison du RVD HY. Une hypothèse serait que le motif MTLA permette un relâchement de la structure de la répétition le portant ce qui permettrait de diminuer l'encombrement stérique provoqué par la tyrosine. Au sein de TAL18H, il y aurait donc eu une coévolution entre un RVD peu commun et une nouvelle répétition permettant à ce RVD d'optimiser sa fonction. Ceci souligne également d'avantage que TAL18H est très adapté aux séquences d'ADN de haricot commun qu'il est capable de cibler.

Ces analyses confirment également une certaine direction qu'empruntent actuellement certaines équipes spécialisées dans l'étude des effecteurs TAL. En effet, les séquences de répétitions aberrantes, particulièrement celles qui diffèrent par leur taille, sont de plus en plus rapportées et intégrées dans des outils bioinformatiques dédiés à l'étude de ces effecteurs (Boch *et al.*, 2014; Booher *et al.*, 2015; Hutin *et al.*, 2015; Pérez-Quintero *et al.*, 2015; Wang *et al.*, 2015; Wilkins *et al.*, 2015; Grau *et al.*, 2016; Erkes *et al.*, 2017). Une étude approfondie de ces répétitions de longueur inhabituelle a même démontré qu'elles sont responsables d'une rupture dans le mode de liaison strict RVD-base d'ADN et permettent l'introduction d'une certaine flexibilité dans la liaison de l'effecteur TAL à l'ADN (Richter *et al.*, 2014). Nos résultats semblent être en accord avec cette flexibilité de liaison effecteur TAL-ADN puisque le motif MTLA permet une accroche faible de HY à la cytosine en plus d'augmenter l'efficacité de l'accroche à l'adénine. Ainsi, la flexibilité dans la liaison de l'effecteur TAL à l'ADN ne serait pas seulement influencée par la longueur de la répétition mais aussi par sa composition en acides aminés. Une perspective intéressante serait d'exploiter l'ensemble de ces résultats pour améliorer les outils biotechnologiques basés sur l'utilisation des TALEN. En effet, nos résultats et de récentes études montrent que la diversité naturelle des effecteurs TAL, RVD mis à part, peut être utilisée pour diversifier les forces de liaison

d'ArtTAL et des TALEN tout en conservant leurs spécificités de liaison (Tochio *et al.*, 2016; de Lange *et al.*, 2017). Nos résultats sont particulièrement intéressants pour ce point puisque les séquences riches en adénine et en thymine ne permettent pas une accroche forte des ArtTAL lorsqu'on utilise des RVD classiques, ce qui est une des limites identifiées pour l'utilisation des ArtTAL et des TALEN (Streubel *et al.*, 2012).

IV. Développement de la lutte contre la graisse commune du haricot

1. De nouveaux outils d'épidémiosurveillance ?

Par la suite, il serait intéressant d'exploiter les gènes candidats pour la spécificité d'hôte dans la lutte contre la graisse commune du haricot. En effet, une application possible serait de générer un outil exploitable en épidémiosurveillance, défini par l'Anses (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail) comme le suivi de l'évolution des maladies et des agents pathogènes et la détection de l'émergence de nouveaux agents infectieux (<https://www.anses.fr/fr>). Pour ce faire, il suffirait de dessiner des amorces PCR spécifiques sur les gènes que nous avons identifiés comme étant à la fois communs et spécifiques des agents responsables de la CBB. Ainsi, ces amorces permettraient à la fois de suivre efficacement les populations de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli* mais aussi d'anticiper l'émergence de nouvelles souches capables d'infecter le haricot qui se seraient adaptées à ce nouvel hôte après acquisition de ces gènes par transfert horizontal. Cela constituerait l'un des maillons essentiels dans la prévention des risques sanitaires liés à la CBB. Un outil d'identification moléculaire des souches infectant le haricot commun par PCR multiplex a déjà été développé (Boureau *et al.*, 2013), ce qui signifie que cette méthodologie est efficace. Cependant, cet outil n'est basé que sur deux gènes codant les ET3 *avrBst* et *xopL* et, de façon surprenante, ils ne font pas partie de notre liste de gènes portant des caractéristiques uniques aux agents responsables de la CBB. Ainsi, avec nos nouvelles données, nous pourrions éventuellement mettre à jour cet outil afin de limiter au maximum des faux positifs mais surtout des faux négatifs lors de la détection de ces agents infectieux.

2. Quelle durabilité pour les résistances basées sur les effecteurs TAL ?

Le mode d'action des effecteurs TAL représente une source de stratégies pour le développement de résistances. Par exemple, une variété résistante de riz a été développée en mutant l'EBE située dans le promoteur du gène *OsSWEET14* et ciblée par AvrXa7. Cette mutation empêche alors la fixation de l'effecteur TAL et donc la surexpression du transporteur de sucre

nécessaire au développement de la maladie (Li *et al.*, 2012a). Egalement, une autre variété de riz multi-résistante a été développée en plaçant 6 EBE différentes, préalablement synthétisées *in vitro*, en amont d'un gène exécuteur. Ceci permet de piéger 6 effecteurs TAL provenant de souches de *X. oryzae* pv. *oryzae* et *X. oryzae* pv. *oryzicola* qui induisent donc une réaction de défense de la plante (Hummel *et al.*, 2012). Ces deux stratégies permettent donc d'exploiter des effecteurs TAL dont la fonction de virulence a été démontrée mais aussi des effecteurs TAL qui, à priori, ne jouent pas un rôle majeur dans la virulence de la bactérie. Cependant, ces stratégies reposent sur de la modification du génome qui d'une part n'est pas facilement réalisable chez toutes les espèces végétales, dont notamment le haricot commun qui est réputé pour être récalcitrant (Hnatuszko-Konka *et al.*, 2014) et qui d'une autre part repose sur une législation très complexe et diversifiée à travers le monde.

Il existe une grande diversité d'effecteurs TAL à travers le genre *Xanthomonas*, mais peu de gènes ciblés codant des fonctions différentes chez les plantes, en raison d'une convergence de ces cibles. Ainsi, la stratégie consistant à muter une EBE située dans le promoteur d'un gène dont la fonction est majoritairement ciblée par des effecteurs TAL pourrait s'avérer être durable. Cependant, cette stratégie est dépendante d'une bonne connaissance de la diversité des *tal* d'un pathovar donné, de l'implication de ces effecteurs TAL dans la virulence de la bactérie et de la caractérisation de leurs cibles qui doivent de préférence converger vers un même gène *S*. De plus, la liaison effecteur TAL-ADN n'est pas stricte et peut tolérer certains mésappariements. En effet, certains RVD ont la capacité de lier plusieurs bases de l'ADN, comme le RVD NN qui préfère se lier à la guanine, mais qui peut aussi fixer l'adénine. Ces observations suggèrent que pour générer des résistances, il est plus efficace de modifier des EBE par insertions ou délétions que par substitution. En effet, il a été démontré que, chez le riz, les allèles résistants résultent souvent d'évènements d'insertion ou de délétion situés dans les EBE ciblés par les effecteurs TAL (Chu *et al.*, 2006b; Yang *et al.*, 2006; Römer *et al.*, 2010; Zhou *et al.*, 2015). De plus, il est avéré que les RVD situés en position C-terminale contribuent moins à la spécificité de l'effecteur TAL que ceux situés en position N-terminale (Kay *et al.*, 2009; Pérez-Quintero *et al.*, 2015). De ce fait, il est probablement plus intéressant pour générer des résistances, de modifier une EBE par délétion ou insertion dans la région 3' que dans la région 5'. Néanmoins, comme évoqué précédemment, il existe des répétitions aberrantes, notamment dans leur longueur, qui confèrent aux effecteurs TAL une flexibilité dans leur liaison à l'ADN (Richter *et al.*, 2014). Ainsi, un effecteur TAL porteur d'une telle répétition pourrait potentiellement être capable de lier son EBE malgré la délétion ou l'insertion d'un nucléotide. Ainsi, ce serait un élément supplémentaire à prendre en compte avant d'entreprendre la mutation d'une EBE (Becker & Boch, 2016; Tochio *et al.*, 2016).

Contrairement aux souches infectant le riz, les agents responsables de la CBB possèdent peu de *tal*. Ainsi, comme chez les souches de *X. oryzae* seuls un ou deux effecteurs TAL jouent un rôle majeur dans la virulence de la bactérie (Yang & White, 2004), nous nous attendions à ce que les effecteurs TAL de *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli*, qui sont conservés, impactent fortement la virulence des bactéries. Or, pour TAL18H nous avons observé une diminution de l'agressivité des souches, mais celles-ci sont toujours capables de produire des symptômes après la délétion des effecteurs TAL. Ainsi, la question de la durabilité des résistances se pose d'autant plus. En effet, produire des variétés de haricot commun résistantes en piégeant ces effecteurs TAL au niveau du promoteur de gènes exécuteurs alors que ces effecteurs TAL ne sont pas indispensables pour la maladie engendrerait probablement une diversification rapide de leurs séquences de RVD en raison de la forte pression de sélection exercée par les gènes exécuteurs. Ceci est en accord avec une étude qui a montré que les gènes *R* reconnaissant des effecteurs TAL dont le rôle dans la virulence est majeur constituent une source de résistance plus durable, car la perte ou la modification de ces effecteurs représenterait un coût important en terme de fitness pour le pathogène (Vera Cruz *et al.*, 2000). D'autre part, l'existence de TAL jouant un rôle mineur dans l'agressivité pose la question de leurs intérêts pour la bactérie. Pour répondre à cette question, une première hypothèse serait que ces effecteurs TAL pourraient jouer un rôle de réservoir génétique afin de favoriser la génération de nouveaux effecteurs TAL par recombinaison. Une hypothèse alternative serait que ces TAL pourraient être importants pour la souche qui les porte, mais lors d'une interaction avec une autre plante que celles étudiées, c'est-à-dire une autre variété ou une autre plante hôte, d'où l'importance d'utiliser plusieurs génotypes de plantes hôtes pour étudier le rôle des effecteurs TAL.

CONCLUSION GÉNÉRALE

Ces travaux de thèse ont permis d'identifier des déterminants potentiels de la spécificité d'hôte des *Xanthomonas* au haricot commun et quatre effecteurs TAL chez les agents responsables de la CBB. De plus, ils ont permis de caractériser le rôle de TAL18H, un effecteur TAL partagé par *X. citri* pv. *fuscans* et *X. phaseoli* pv. *phaseoli*.

Nos résultats indiquent que plusieurs événements de transferts horizontaux se sont produits entre des souches de *Xanthomonas* phylogénétiquement éloignées infectant le haricot commun. Ces transferts horizontaux ont conduit à une combinaison spécifique de gènes seulement retrouvés chez les agents responsables de la CBB, ce qui suggère qu'ils sont impliqués dans l'adaptation des *Xanthomonas* au haricot commun. Notamment, deux plasmides porteurs de *tal* ont été des vecteurs lors de ces transferts. De ce fait, nos résultats indiquent également que les effecteurs TAL ont probablement contribué à l'adaptation des *Xanthomonas* au haricot commun. La caractérisation de TAL18H, partagé par les lignées NF1 et *fuscans*, semble confirmer cette implication des effecteurs TAL dans l'adaptation, puisque la mutation de cet effecteur a diminué l'agressivité de la souche. L'étude de l'impact de TAL18H sur le transcriptome du haricot a montré que les effecteurs TAL pouvaient présenter des effets pléiotropiques *in planta* et a contribué à l'identification des cibles potentielles de TAL18H, NDH18 étant la plus prometteuse. Ainsi, l'ensemble de ces données ouvre des perspectives de recherche sur les bases moléculaires de la spécificité hôte et peuvent également être utilisée dans la lutte contre la CBB, aussi bien en épidémiologie que pour la production de variétés de haricot commun résistantes.

Par ailleurs, la caractérisation du RVD HY et du motif MTLA a montré que ce RVD liait préférentiellement une adénine et que le motif MTLA permettait d'augmenter l'efficacité de liaison du RVD HY. Ainsi, nous avons confirmé que la composition d'une séquence de répétition (RDV exclus) joue un rôle important dans l'interaction effecteur TAL – ADN, ce qui peut être exploité pour améliorer des outils biotechnologiques basés sur ces effecteurs, tel que les TALEN.

D'autre part, les connaissances que nous avons acquises sur les répertoires d'effecteurs TAL des agents responsables de la CBB et de quelques autres souches de *Xanthomonas* ont fourni à la communauté scientifique des données sur la diversité des effecteurs TAL au sein du genre *Xanthomonas*, ont montré que la phylogénie des *tal* ne suivait pas la phylogénie des souches et ont confirmé que l'histoire évolutive de ces effecteurs était riche en événements évolutifs divers. Ainsi, nous avons obtenu de précieuses informations qui, par la suite, pourraient être utiles pour appréhender le potentiel évolutif des effecteurs TAL, notamment dans le contournement des stratégies de résistances basées sur ces effecteurs, ce qui permettrait de produire des résistances potentiellement plus durables contre les bactérioses induites par les bactéries du genre *Xanthomonas*.

RÉFÉRENCES BIBLIOGRPAHIQUES

- Aamir M, Singh VK, Meena M, Upadhyay RS, Gupta VK, Singh S. 2017.** Structural and Functional Insights into WRKY3 and WRKY4 Transcription Factors to Unravel the WRKY-DNA (W-Box) Complex Interaction in Tomato (*Solanum lycopersicum* L.). A Computational Approach. *Frontiers in Plant Science* **8**: 1–24.
- Abd-Alla MH, Bashandy SR, Schnell S. 2010.** Occurrence of *Xanthomonas axonopodis* pv. *phaseoli*, the causal agent of common bacterial blight disease, on seeds of common bean (*Phaseolus vulgaris* L.) in upper Egypt. *Folia Microbiologica* **55**: 47–52.
- Abe VY, Benedetti CE. 2016.** Additive roles of PthAs in bacterial growth and pathogenicity associated with nucleotide polymorphisms in effector-binding elements of citrus canker susceptibility genes. *Molecular Plant Pathology* **17**: 1223–1236.
- Ah-You N, Gagnevin L, Grimont PAD, Brisse S, Nesme X, Chiroleu F, Bui Thi Ngoc L, Jouen E, Lefeuvre P, Vernière C, et al. 2009.** Polyphasic characterization of *Xanthomonas* pathogenic to members of the Anacardiaceae and their relatedness to species of *Xanthomonas*. *International Journal of Systematic and Evolutionary Microbiology* **59**: 306–318.
- Akhavan A, Bahar M, Askarian H, Lak MR, Nazemi A, Zamani Z. 2013.** Bean common bacterial blight: pathogen epiphytic life and effect of irrigation practices. *SpringerPlus* **2**: 41.
- Al-Saadi A, Reddy JD, Duan YP, Brunings AM, Yuan Q, Gabriel DW. 2007.** All five host-range variants of *Xanthomonas citri* carry one pthA homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. *Molecular plant-microbe interactions : MPMI* **20**: 934–943.
- Alavi SM, Sanjari S, Durand F, Brin C, Manceau C, 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 h. *Applied and Environmental Microbiology* **74**: 3295–3301.
- Albuquerque P, Caridade CMR, Marcal ARS, Cruz J, Cruz L, Santos CL, Mendes M V., Tavares F. 2011.** Identification of *Xanthomonas fragariae*, *Xanthomonas axonopodis* pv. *phaseoli*, and *Xanthomonas fuscans* subsp. *fuscans* with novel markers and using a dot blot platform coupled with automatic data Analysis. *Applied and Environmental Microbiology* **77**: 5619–5628.
- Alexander RP, Zhulin IB. 2007.** Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 2885–2890.
- Alexandre G. 2010.** Coupling metabolism and chemotaxis-dependent behaviours by energy taxis receptors. *Microbiology* **156**: 2283–2293.
- Alfano JR, Collmer A. 1996.** Bacterial Pathogens in Plants : Life up against the Wall. *The Plant cell* **8**: 1683–1698.
- Alfano JR, Collmer A. 2004.** TYPE III SECRETION SYSTEM EFFECTOR PROTEINS : Double Agents in Bacterial Disease and Plant Defense. *Annual Review of Phytopathology* **42**: 385–414.
- Alladassi BME, Nkalubo ST, MukanKusi C, Mwale ES, Gibson P, Edema R, Urrea CA, Kelly JD, Rubaihayo PR. 2017.** Inheritance of resistance to common bacterial blight in four selected common bean (*Phaseolus vulgaris* L.) genotypes. *Journal of Plant Breeding and Crop Science* **9**: 71–78.
- Alteri CJ, Mobley HLT. 2016.** The versatile type VI secretion system. *Microbiology Spectrum* **4**.
- Amano A. 2010.** Bacterial adhesins to host components in periodontitis. *Periodontology 2000* **52**: 12–37.
- Anderson JC, Wan Y, Kim Y-M, Pasa-Tolic L, Metz TO, Peck SC. 2014.** Decreased abundance of type III secretion system-inducing signals in *Arabidopsis* mkp1 enhances resistance against *Pseudomonas syringae*. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 6846–51.
- Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Zeuli PLS, Papa R. 2010.** Beans in Europe: Origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theoretical and Applied Genetics* **121**: 829–843.
- Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. 2010.** Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant cell* **22**: 3864–76.
- Aragão FJL, Nogueira EOPL, Tinoco MLP, Faria JC. 2013.** Molecular characterization of the first commercial transgenic common bean immune to the Bean golden mosaic virus. *Journal of Biotechnology* **166**: 42–50.
- Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. 2015.** Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Frontiers in Microbiology* **6**: 1–18.
- Aritua V, Parkinson N, Thwaites R, Heeney J V., Jones DR, Tushemereirwe W, Crozier J, Reeder R, Stead DE, Smith J. 2008.** Characterization of the *Xanthomonas* sp. causing wilt of enset and banana and its proposed reclassification as a strain of *X. vasicola*. *Plant Pathology* **57**: 170–177.
- Arnaud-Santana E, Pena-Matos E, Coyne DP, Vidaver AK. 1991.** Longevity of *Xanthomonas campestris* pv. *phaseoli* in naturally infested dry bean (*Phaseolus vulgaris*) debris. *Plant Disease* **75**: 952–953.

- Audy P, Laroche A, Saindon G, Huang HC, Gilbertson RL. 1994.** Detection of the Bean Common Blight Bacteria, *Xanthomonas campestris* pv. *phaseoli* and *X. c. phaseoli* var. *fuscans*, Using the Polymerase Chain Reaction. *Molecular Plant Pathology* **84**: 1185–92.
- Bai Y, Michaels TE, Pauls KP. 1997.** Identification of RAPD markers linked to common bacterial blight resistance genes in *Phaseolus vulgaris* L. *Genome* **40**: 544–551.
- Baker KF, Smith SH. 1966.** Dynamics of Seed Transmission of Plant Pathogens. *Annual Review of Phytopathology* **4**: 311–332.
- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL. 2011.** Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**.
- Baptista JC, Machado MA, Homem RA, Torres PS, Vojnov AA, Do Amaral AM. 2010.** Mutation in the *xpsD* gene of *Xanthomonas axonopodis* pv. *citri* affects cellulose degradation and virulence. *Genetics and Molecular Biology* **33**: 146–153.
- Barak JD, Vancheva T, Lefeuvre P, Jones JB, Timilsina S, Minsavage G V., Vallad GE, Koebnik R. 2016.** Whole-Genome Sequences of *Xanthomonas euvesicatoria* Strains Clarify Taxonomy and Reveal a Stepwise Erosion of Type 3 Effectors. *Frontiers in Plant Science* **7**.
- Barash I, Manulis-Sasson S. 2009.** Recent evolution of bacterial pathogens: the gall-forming *Pantoea agglomerans* case. *Annual Review of Phytopathology* **47**: 133–152.
- Bardoel BW, van der Ent S, Pel MJC, Tommassen J, Pieterse CMJ, van Kessel KPM, van Strijp JAG. 2011.** *Pseudomonas* evades immune recognition of flagellin in both mammals and plants. *PLoS Pathogens* **7**.
- Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, et al. 2012.** High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E1972–E1979.
- Bartoli C, Roux F, Lamichhane JR. 2016.** Molecular mechanisms underlying the emergence of bacterial pathogens: An ecological perspective. *Molecular Plant Pathology* **17**: 303–310.
- Bäumler A, Fang FC. 2013.** Host specificity of bacterial pathogens. *Cold Spring Harbor perspectives in medicine* **3**: a010041.
- Becker S, Boch J. 2016.** Single-molecule biophysics: TALEs spin along, but not around. *Nature Chemical Biology* **12**: 5–7.
- Belas R, Zhulin IB, Yang Z. 2008.** Bacterial signaling and motility: Sure bets. *Journal of Bacteriology* **190**: 1849–1856.
- Belete T, Bastas K. 2017.** Common Bacterial Blight (*Xanthomonas axonopodis* pv. *phaseoli*) of Beans with Special Focus on Ethiopian Condition. *Journal of Plant Pathology & Microbiology* **8**.
- Belhaj K, Chaparro-garcia A, Kamoun S, Patron NJ, Nekrasov V. 2015.** Editing plant genomes with CRISPR / Cas9. *Current Opinion in Biotechnology* **32**: 76–84.
- Benjamini Y, Hochberg Y. 1995.** Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* **57**: 289–300.
- Berg HC. 2003.** The Rotatory Motor of Bacterial Flagella. *Annual Review of Biochemistry* **72**: 19–54.
- Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, Liu X, Gao D, Clevenger J, Dash S, et al. 2016.** The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature genetics* **48**: 438–446.
- Birch PRJ, Hyman LJ, Taylor R, Opio AF, Bragard C, Toth IK. 1997.** RAPD PCR-based differentiation of *Xanthomonas campestris* pv. *phaseoli* and *Xanthomonas campestris* pv. *phaseoli* var. *fuscans*. *European Journal of Plant Pathology* **103**: 809–814.
- Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Santo D, Nanni L, Attene G, Papa R. 2017.** Beans (*Phaseolus* ssp.) as a Model for Understanding Crop Evolution. *Frontiers in Plant Science* **8**: 1–21.
- Blanvillain-Baufumé S, Reschke M, Solé M, Auguy F, Doucoure H, Szurek B, Meynard D, Portefaix M, Cunnac S, Guiderdoni E, et al. 2017.** Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for SWEET14-inducing TAL effectors. *Plant biotechnology journal* **15**: 306–317.
- Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M. 2007.** Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**.
- Boch J, Bonas U. 2010.** *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* **48**: 419–36.
- Boch J, Bonas U, Lahaye T. 2014.** TAL effectors - pathogen strategies and plant resistance engineering. *New Phytologist* **204**: 823–832.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. 2009.** Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509–12.
- Bogdanove AJ. 2014.** Principles and applications of TAL effectors for plant physiology and metabolism. *Current Opinion in Plant*

Biology **19**: 99–104.

Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli S V., Patil PB, Van Sluys MA, Ryan RP, Meyer DF, Han SW, et al. 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *Journal of Bacteriology* **193**: 5450–5464.

Bogdanove AJ, Schornack S, Lahaye T. 2010. TAL effectors: Finding plant genes for disease and defense. *Current Opinion in Plant Biology* **13**: 394–401.

Bogdanove a. J, Voytas DF. 2011. TAL Effectors: Customizable Proteins for DNA Targeting. *Science* **333**: 1843–1846.

Bogdanove AJ, Wei ZM, Zhao L, Beer S V. 1996. *Erwinia amylovora* secretes harpin via a type III pathway and contains a homolog of yopN of *Yersinia* spp. *Journal of Bacteriology* **178**: 1720–1730.

Boller T, He SY. 2009. Innate Immunity in Plants : An Arms Race. *Science* **324**: 742–4.

Bolot S, Bodnar AM, Cunnac S, Ortiz E, Szurek B, Noel LD, Arlat M, Jacques M-A, Gagnevin L, Portier P, et al. 2013a. Draft Genome Sequence of the *Xanthomonas cassavae* Type Strain CFBP 4642. *Genome announcements* **1**: e00679-13.

Bolot S, Guy E, Carrere S, Barbe V, Arlat M, Noël LD. 2013b. *Xanthomonas campestris* pv. *campestris* Strain Xca5. *Genome Announcements* **1**: e00032-12.

Bonas U, Conrads-Strauch J, Balbo I. 1993. Resistance in tomato to *Xanthomonas campestris* pv. *vesicatoria* is determined by alleles of the pepper-specific avirulence gene *avrBs3*. *Molecular & general genetics : MGG* **238**: 261–269.

Bonas U, Stall RE, Staskawicz B. 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular & general genetics : MGG* **218**: 127–136.

Bonfim K, Faria JC, Nogueira EOPL, Mendes E a, Aragão FJL, Mendes É a, Aragão FJL. 2007. RNAi-Mediated Resistance to Bean golden mosaic virus in Genetically Engineered Common Bean (*Phaseolus vulgaris*). *Molecular plant-microbe interactions : MPMI* **20**: 717–726.

Boni MF, Posada D, Feldman MW. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**: 1035–1047.

Booher NJ, Carpenter SCD, Sebra RP, Wang L, Salzberg SL, Leach JE, Bogdanove AJ. 2015. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Microbial Genomics* **1**: 1–22.

Bortesi L, Fischer R. 2015. The CRISPR / Cas9 system for plant genome editing and beyond. *Biotechnology Advances* **33**: 41–52.

Boulanger A, Noël LD. 2016. *Xanthomonas* Whole Genome Sequencing: Phylogenetics, Host Specificity and Beyond. *Frontiers in Microbiology* **7**: 1–2.

Boureau T, Kerkoud M, Chhel F, Hunault G, Darrasse A, Brin C, Durand K, Hajri A, Poussier S, Manceau C, et al. 2013. A multiplex-PCR assay for identification of the quarantine plant pathogen *Xanthomonas axonopodis* pv. *phaseoli*. *Journal of Microbiological Methods* **92**: 42–50.

Bradbury JF. 1986. *Guide to plant pathogenic bacteria*. CAB International.

Briand M, Gaborieau R, Jacques M, Barret M, Boureau T, Gaillard S, Chen NWG. 2016. Sklf : a tool for rapid identification of genes or regulators of interest [version 1; not peer reviewed]. *F1000Research* **5**: (poster).

Burkholder WH. 1930. The bacterial diseases of the bean. *Memoirs. Cornell University Agricultural Experiment Station* **127**: 1–88.

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

Büttner D, Bonas U. 2010. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Reviews* **34**: 107–133.

Büttner D, He SY. 2009. Type III protein secretion in plant pathogenic bacteria. *Plant physiology* **150**: 1656–1664.

Lo Cantore P, Shanmugaiah V, Iacobellis NS. 2009. Antibacterial activity of essential oil components and their potential use in seed disinfection. *Journal of agricultural and food chemistry* **57**: 9454–61.

Cao H, Baldini RL, Rahme LG. 2011. COMMON MECHANISMS FOR PATHOGENS OF PLANTS AND ANIMALS. *Annual Review of Phytopathology* **39**: 259–84.

Cao TB, Saier MH. 2003. The general protein secretory pathway: Phylogenetic analyses leading to evolutionary conclusions. *Biochimica et Biophysica Acta - Biomembranes* **1609**: 115–125.

Carroll D. 2014. Genome engineering with targetable nucleases. *Annual Review of Biochemistry* **83**: 409–439.

Carvalho JLV, de Oliveira Santos J, Conte C, Pacheco S, Nogueira EOPL, Souza TLPO, Faria JC, Aragão FJL. 2015. Comparative analysis of nutritional compositions of transgenic RNAi-mediated virus-resistant bean (event EMB-PV051-1) with its non-transgenic counterpart. *Transgenic Research* **24**: 813–819.

- Castiblanco LF, Gil J, Rojas A, Osorio D, Gutiérrez S, Muñoz-Bodnar A, Perez-Quintero AL, Koebnik R, Szurek B, López C, et al. 2013.** TALE1 from *Xanthomonas axonopodis* pv. *Manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. *Molecular Plant Pathology* **14**: 84–95.
- Castro Guerrero NA, Isidra-Arellano MC, Mendoza-Cozatl DG, Valdes-Lopez O. 2016.** Common bean: a legume model on the rise for unraveling responses and adaptations to iron, zinc and phosphate deficiencies. *Frontiers in Plant Science* **7**: 1–7.
- Celia H, Noinaj N, Zakharov SD, Bordignon E, Botos I, Santamaria M, Barnard TJ, Cramer W a, Lloubes R, Buchanan SK. 2016.** Structural insight into the role of the Ton complex in energy transduction. *Nature* **538**: 60–65.
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia N V, Bogdanove AJ, Voytas DF. 2011.** Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**: 1–11.
- Cernadas RA, Doyle EL, Niño-Liu DO, Wilkins KE, Bancroft T, Wang L, Schmidt CL, Caldo R, Yang B, White FF, et al. 2014.** Code-Assisted Discovery of TAL Effector Targets in Bacterial Leaf Streak of Rice Reveals Contrast with Bacterial Blight and a Novel Susceptibility Gene. *PLoS Pathogens* **10**: e1003972.
- Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques M-A. 2015.** Comparative Genomics of Pathogenic and Nonpathogenic Strains of *Xanthomonas arboricola* Unveil Molecular and Evolutionary Events Linked to Pathoadaptation. *Frontiers in Plant Science* **6**: 1126.
- Chacón S. MI, Pickersgill B, Debouck DG, Arias JS. 2007.** Phylogeographic analysis of the chloroplast DNA variation in wild common bean (*Phaseolus vulgaris* L.) in the Americas. *Plant Systematics and Evolution* **266**: 175–195.
- Chacón S MI, Pickersgill B, Debouck DG. 2005.** Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theoretical and Applied Genetics* **110**: 432–444.
- Chan JWYF, Goodwin PH. 1999.** Differentiation of *Xanthomonas campestris* pv. *phaseoli* from *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* by PFGE and RFLP. *European Journal of Plant Pathology* **105**: 867–878.
- Charpentier M, Oldroyd G. 2010.** How close are we to nitrogen-fixing cereals? *Current Opinion in Plant Biology* **13**: 556–564.
- Chen LQ. 2014.** SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytologist* **201**: 1150–1155.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al. 2013.** Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* **10**: 563–569.
- Chinchilla D, Bauer Z, Regenass M, Boller T, Felix G. 2006.** The Arabidopsis Receptor Kinase FLS2 Binds flg22 and Determines the Specificity of Flagellin Perception. *The Plant Cell* **18**: 465–476.
- Chisholm ST, Coaker G, Day B, Staskawicz BJ. 2006.** Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. *Cell* **124**: 803–814.
- Cho H-S, Kang JG, Lee J-H, Lee J-J, Jeon SK, Ko J-H, Kim D-S, Park K-H, Kim Y-S, Kim N-S. 2015.** Direct regulation of E-cadherin by targeted histone methylation of TALE-SET fusion protein in cancer cells. *Oncotarget* **6**: 23837–44.
- Cho HJ, Park YJ, Noh TH, Kim YT, Kim JG, Song ES, Lee DH, Lee BM. 2008.** Molecular analysis of the *hrp* gene cluster in *Xanthomonas oryzae* pathovar *oryzae* KACC10859. *Microbial Pathogenesis* **44**: 473–483.
- Christenhusz MJM, Byng JW. 2016.** The number of known plants species in the world and its annual increase. *Phytotaxa* **261**: 201–217.
- Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF. 2010.** Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**: 756–761.
- Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. 2012.** Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues. *PLoS ONE* **7**.
- Chu Z, Fu B, Yang H, Xu C, Li Z, Sanchez A, Park YJ, Bennetzen JL, Zhang Q, Wang S. 2006a.** Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. *Theoretical and Applied Genetics* **112**: 455–461.
- Chu Z, Yuan M, Yao J, Ge X, Yuan B, Xu C, Li X, Fu B, Li Z, Bennetzen JL, et al. 2006b.** Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes and Development* **20**: 1250–1255.
- Cianciotto NP. 2005.** Type II secretion: A protein secretion system for all seasons. *Trends in Microbiology* **13**: 581–588.
- Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, Grusak MA, Kamfwa K, Katuuramu DN, McClean P, et al. 2015.** A *Phaseolus vulgaris* diversity panel for andean bean improvement. *Crop Science* **55**: 2149–2160.
- Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou BH, Frommer WB, Lahaye T, Staskawicz BJ. 2014.** *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. *Molecular plant-microbe interactions : MPMI* **27**: 1186–1198.

- Cong L, Zhou R, Kuo Y, Cunniff M, Zhang F. 2012.** Comprehensive Interrogation of Natural TALE DNA Binding Modules and Transcriptional Repressor Domains. *Nature Communications* **3**.
- Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B. 2016.** Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathology* **65**: 792–806.
- Corey RR, Starr MP. 1957.** Colony types of *Xanthomonas phaseoli*. *Journal of Bacteriology* **74**: 137–140.
- Cornelis G. 2006.** The type III secretion injectisome. *Nature reviews Microbiology* **4**: 811–825.
- Costerton JW, Cheng K-J, Geesey GG, Ladd TI, Nickel JC, Dasgupta M, Marrie TJ. 1987.** BACTERIAL BIOFILMS IN NATURE AND DISEASE. *Annual Review of Microbiology* **41**: 435–464.
- Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappin-Scott HM. 1995.** Microbial Biofilms. *Annual Review of Microbiology* **49**: 711–745.
- Coutinho TA, van der Westhuizen L, Roux J, Mcfarlane SA, Venter SN. 2015.** Significant host jump of *Xanthomonas vasicola* from sugarcane to a *Eucalyptus grandis* clone in South Africa. *Plant Pathology* **64**: 576–581.
- Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, Carpenter SCD, Chen L, Zheng H, Gao X, et al. 2017.** TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. *Nature Communications* **8**: 1–14.
- Cronk Q, Ojeda I, Pennington RT. 2006.** Legume comparative genomics: Progress in phylogenetics and phylogenomics. *Current Opinion in Plant Biology* **9**: 99–103.
- Crooks G, Hon G, Chandonia J, Brenner S. 2004.** WebLogo: a sequence logo generator. *Genome Research* **14**: 1188–1190.
- Cunnac S, Bolot S, Forero Serna N, Ortiz E, Szurek B, Noël LD, Arlat M, Jacques M-A, Gagnevin L, Carrere S, et al. 2013.** High quality draft genome sequences of two *Xanthomonas citri* pv. *malvacearum* strains. *Genome announcements* **1**: e00674-13.
- Cunnac S, Lindeberg M, Collmer A. 2009.** *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. *Current Opinion in Microbiology* **12**: 53–60.
- Dangl JL, Horvath DM, Staskawicz BJ. 2013.** Pivoting the Plant Immune System from Dissection to Deployment. *Science* **341**: 746–751.
- Danhorn T, Fuqua C. 2007.** Biofilm Formation by Plant-Associated Bacteria. *Annual Review of Microbiology* **61**: 401–422.
- Darmon E, Leach DRF. 2014.** Bacterial genome instability. *Microbiology and molecular biology reviews : MMBR* **78**: 1–39.
- Darrasse A, Bolot S, Serres-Giardi L, Charbit E, Boureau T, Saux MF-L, Briand M, Arlat M, Gagnevin L, Koebnik R, et al. 2013a.** High-Quality Draft Genome Sequences of *Xanthomonas axonopodis* pv. *glycines* Strains CFBP 2526 and CFBP 7119. *Genome announcements* **1**: e01036-13.
- Darrasse A, Bureau C, Samson R, Morris CE, Jacques MA. 2007.** Contamination of bean seeds by *Xanthomonas axonopodis* pv. *phaseoli* associated with low bacterial densities in the phyllosphere under field and greenhouse conditions. *European Journal of Plant Pathology* **119**: 203–215.
- Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, Briand M, Brin C, Cociancich S, Durand K, et al. 2013b.** Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. *BMC Genomics* **14**: 761.
- Darrasse A, Darsonval A, Boureau T, Brisset MN, Durand K, Jacques MA. 2010.** Transmission of plant-pathogenic bacteria by nonhost seeds without induction of an associated defense reaction at emergence. *Applied and Environmental Microbiology* **76**: 6787–6796.
- Darsonval A, Darrasse A, Durand K, Bureau C, Cesbron S, Jacques M-A. 2009.** Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Molecular plant-microbe interactions : MPMI* **22**: 747–757.
- Darsonval A, Darrasse A, Meyer D, Demarty M, Durand K, Bureau C, Manceau C, Jacques M. 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. *Applied and environmental microbiology* **74**: 2669–2678.
- Degrave A, Siamer S, Boureau T, Barny MA. 2015.** The AvrE superfamily: Ancestral type III effectors involved in suppression of pathogen-associated molecular pattern-triggered immunity. *Molecular Plant Pathology* **16**: 899–905.
- Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J, Shi Y. 2012.** Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science* **335**: 720–723.
- Deslandes L, Rivas S. 2012.** Catch me if you can: Bacterial effectors and plant targets. *Trends in Plant Science* **17**: 644–655.
- Doorn J van, Boonekamp PM, Oudega B. 1994.** Partial Characterization of Fimbriae of *Xanthomonas campestris* pv. *campestris*. *Molecular plant-microbe interactions* **7**: 334–44.

- Doughari J. 2015.** An Overview of Plant Immunity. *Journal of Plant Pathology & Microbiology* **6**: 1–11.
- Dowson DW. 1939.** On the systematic position and generic names of the Gram negative bacterial plant pathogens. *Zentralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten*, **2** **100**: 177–193.
- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. 2012.** TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Research* **40**: 117–122.
- Doyle EL, Hummel AW, Demorest ZL, Starker CG, Voytas DF, Bradley P, Bogdanove AJ. 2013a.** TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat -1. *PLoS ONE* **8**: e82120.
- Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ. 2013b.** TAL effectors: Highly adaptable phyto-bacterial virulence factors and readily engineered DNA-targeting proteins. *Trends in Cell Biology* **23**: 390–398.
- Duncan RW, Singh SP, Gilbertson RL. 2011.** Interaction of common bacterial blight bacteria with disease resistance quantitative trait Loci in common bean. *Phytopathology* **101**: 425–435.
- Dye DW, Bradbury JF, Goto M, Hayward AC, Lelliott RA, Schroth MN. 1980.** International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Review of Plant Pathology* **59**: 153–159.
- EFSA PLH Panel (EFSA Panel on Plant Health). 2014.** Scientific Opinion on the pest categorisation of *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *EFSA Journal* **12**: 3856.
- EFSA PLH Panel (EFSA Panel on Plant Health). 2015.** Scientific Opinion on the pest categorisation of *Spodoptera littoralis*. *EFSA Journal* **13**: 3987.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, et al. 2009.** Single Polymerase Molecules. *Science* **323**: 133–138.
- EPPO/CABI. 1986.** *Xanthomonas axonopodis* pv. *phaseoli*. *EPPO quarantine pest*.
- Erkes A, Reschke M, Boch J, Grau J. 2017.** Evolution of transcription activator-like effectors in *Xanthomonas oryzae*. *Genome Biology and Evolution* **2**: 1–15.
- Escalon A, Javegny S, Vernière C, Noël LD, Vital K, Poussier S, Hajri A, Boureau T, Pruvost O, Arlat M, et al. 2013.** Variations in type III effector repertoires, pathological phenotypes and host range of *Xanthomonas citri* pv. *citri* pathotypes. *Molecular Plant Pathology* **14**: 483–496.
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE. 2000.** The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**: 199–206.
- Felix G, Duran JD, Volko S, Boller T. 1999.** Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant Journal* **18**: 265–276.
- Ferreira RM, de Oliveira ACP, Moreira LM, Belasque JJ, Gourbeyre E, Siguier P, Ferro MIT, Ferro JA, Chandler M, Varani AM. 2015.** A TALE of Transposition: Tn3-like transposons play a major role in the spread of pathogenicity determinants of *xanthomonas citri* and other *xanthomonads*. *mBio* **6**: 1–12.
- De Feyter R, Yang Y, Gabriel DW. 1993.** Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. *malvacearum* avr genes. *Molecular plant-microbe interactions : MPMI* **6**: 225–237.
- Figueiredo A, Monteiro F, Sebastiana M. 2014.** Subtilisin-like proteases in plant-pathogen recognition and immune priming: a perspective. *Frontiers in Plant Science* **5**: 1–4.
- Fininsa C. 1996.** Effect of intercropping bean with maize on bean common bacterial blight and rust diseases. *International Journal of Pest Management* **42**: 51–54.
- Fininsa C. 2003.** Relationship between common bacterial blight severity and bean yield loss in pure stand and bean-maize intercropping systems. *International Journal of Pest Management* **49**: 177–185.
- Fininsa C, Tefera T. 2001.** Effect of primary inoculum sources of bean common bacterial blight on early epidemics, seed yield and quality aspects. *International Journal of Pest Management* **47**: 221–225.
- Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015.** Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**: 4651–4668.
- Flor HH. 1971.** Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology* **9**: 275–296.
- Foyer CH, Lam H-M, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM, et al. 2016.** Neglecting legumes has compromised human health and sustainable food production. *Nature Plants* **2**: 16112.
- Fronzes R, Christie PJ, Waksman G. 2009.** The structural biology of type IV secretion systems. *Nature Review Microbiology* **7**: 703–

14.

- Furuya EY, Lowy FD. 2006.** Antimicrobial-resistant bacteria in the community setting. *Nature reviews. Microbiology* **4**: 36–45.
- Gagnevin L, Bolot S, Gordon JL, Pruvost O, Vernière C, Robène I, Arlat M, Noël LD, Carrère S, Jacques M-A, et al. 2014.** Draft Genome Sequence of *Xanthomonas axonopodis* pv. *allii* Strain CFBP 6369. *Genome announcements* **2**: e00727-14.
- Gakuubi MM, Wagacha JM, Dossaji SF, Wanzala W. 2016.** Chemical Composition and Antibacterial Activity of Essential Oils of *Tagetes minuta* (Asteraceae) against Selected Plant Pathogenic Bacteria. *International Journal of Microbiology* **2016**.
- Galán JE, Collmer A. 1999.** Type III Secretion Machines: Bacterial Devices for Protein Delivery into Host Cells. *Science* **284**: 1322–1329.
- Gaut BS. 2014.** The complex domestication history of the common bean. *Nature genetics* **46**: 663–4.
- Geißler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, Boch J. 2011.** Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**: e19509.
- Gena M, Shah R, Jadon KS, Mali BL. 2009.** Host range of *Xanthomonas axonopodis* pv. *vignicola*, the incitant of bacterial blight in cowpea. *Indian Phytopathology* **62**: 539–540.
- Gerlach RG, Hensel M. 2007.** Protein secretion systems and adhesins: The molecular armory of Gram-negative pathogens. *International Journal of Medical Microbiology* **297**: 401–415.
- Gilbertson RL, Maxwell DP. 1992.** Common bacterial blight of bean. *Plant diseases of international importance. Volume II. Diseases of vegetables and oil seed crops.*: 18–39.
- Gochez AM, Minsavage G V., Potnis N, Canteros BI, Stall RE, Jones JB. 2015.** A functional XopAG homologue in *Xanthomonas fuscans* pv. *aurantifolii* strain C limits host range. *Plant Pathology* **64**: 1207–1214.
- Goodwin PH, Sopher CR. 1994.** Water stress in leaves of *Phaseolus vulgaris* infected with *Xanthomonas campestris* pv. *phaseoli*. *Journal of Phytopathology* **140**: 219–226.
- Gordon JL, Lefeuvre P, Escalon A, Barbe V, Cruveiller S, Gagnevin L, Pruvost O. 2015.** Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. *BMC Genomics* **16**: 1098.
- Graham PH, Ranalli P. 1997.** Common bean (*Phaseolus vulgaris* L.). *Field Crops Research* **53**: 131–146.
- Grant JJR, Arantes AS, Stothard P, Carver T, Rutherford K, Berriman M, Rajandream M-A, Barrell B, Parkhill J, Elnitski L, et al. 2012.** Comparing thousands of circular genomes using the CGView Comparison Tool. *BMC genomics* **13**: 202.
- Grant SR, Fisher EJ, Chang JH, Mole BM, Dangl JL. 2006.** Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology* **60**: 425–449.
- Grau J, Reschke M, Erkes A, Streubel J, Morgan RD, Wilson GG, Koebnik R, Boch J. 2016.** AnnoTALE: bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. *Scientific Reports* **6**: 21077.
- Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J. 2013.** Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Computational Biology* **9**.
- Grennan AK. 2006.** Plant response to bacterial pathogens. Overlap between innate and gene-for-gene defense response. *Plant physiology* **142**: 809–11.
- Grimault V, Olivier V, Rolland M, Darrasse A, Jacques MA. 2014.** Seed health testing methods. 7-021: Detection of *Xanthomonas axonopodis* pv. *phaseoli* on *Phaseolus vulgaris*. *ISTA International rules for seed testing. Annexe to Chapter 7: Seed health methods 7-021. International Seed Testing Association, Basserdorf, Switzerland, 1-20*.
- Guan SH, Gris C, Cruveiller S, Pouzet C, Tasse L, Leru A, Maillard A, Médigue C, Batut J, Masson-Boivin C, et al. 2013.** Experimental evolution of nodule intracellular infection in legume symbionts. *The ISME Journal* **7**: 1367–1377.
- Gu K, Sangha JS, Li Y, Yin ZC. 2008.** High-resolution genetic mapping of bacterial blight resistance gene Xa10. *Theoretical and Applied Genetics* **116**: 155–163.
- Gu K, Tian D, Qiu C, Yin Z. 2009.** Transcription activator-like type III effector AvrXa27 depends on OsTFIIA5 for the activation of Xa27 transcription in rice that triggers disease resistance to *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology* **10**: 829–835.
- Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang G-L, White FF, et al. 2005.** R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122–1125.
- Guo W, Cui Y ping, Li Y rong, Che Y zhou, Yuan L, Zou L fang, Zou H song, Chen G you. 2012.** Identification of seven *Xanthomonas oryzae* pv. *oryzicola* genes potentially involved in pathogenesis in rice. *Microbiology* **158**: 505–518.
- Hajri A, Brin C, Hunault G, Lardeux F, Lemaire C, Manceau C, Boureau T, 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.

- Hajri A, Brin C, Zhao S, David P, Feng J, Koebnik R, Szurek B, Verdier V, Boureau T, Poussier S. 2012a.** Multilocus sequence analysis and type III effector repertoire mining provide new insights into the evolutionary history and virulence of *Xanthomonas oryzae*. *Molecular Plant Pathology* **13**: 288–302.
- Hajri A, Meyer D, Delort F, Guillaumès J, Brin C, Manceau C. 2010.** Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**: 1014–1022.
- Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Duffy B, Manceau C. 2012b.** Type Three Effector Gene Distribution and Sequence Analysis Provide New Insights into the Pathogenicity of Plant-Pathogenic *Xanthomonas arboricola*. *Applied and environmental microbiology* **78**: 371–384.
- Hammond-Kosack KE, Kanyuka K. 2007.** Resistance genes (R genes) in plants. *Encyclopedia of Life Sciences*: 1–21.
- Hane JK, Ming Y, Kamphuis LG, Nelson MN, Garg G, Atkins CA, Bayer PE, Bravo A, Bringans S, Cannon S, et al. 2017.** A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. *Plant Biotechnology Journal* **15**: 318–330.
- Hashimi H, Zíková A, Panigrahi AK, Stuart KD, Lukeš J. 2008.** TbRGG1 , an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex. *RNA* **14**: 970–980.
- Hauck P, Thilmony R, He SY. 2003.** A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible *Arabidopsis* plants. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 8577–82.
- Hausner J, Hartmann N, Jordan M, Büttner D. 2017.** The predicted lytic transglycosylase HpaH from *Xanthomonas campestris* pv. *vesicatoria* associates with the type III secretion system and promotes effector protein translocation. *Infection and Immunity* **85**: e00788-16.
- Hayat I, Ahmad A, Masud T, Ahmed A, Bashir S. 2014.** Nutritional and health perspectives of beans (*Phaseolus vulgaris* L.): an overview. *Critical reviews in food science and nutrition* **54**: 580–92.
- Hayward AC. 1993.** The hosts of *Xanthomonas*. *Xanthomonas*. 1–119.
- Hazelbauer GL. 2012.** Bacterial Chemotaxis : The Early Years of Molecular Studies. *Annual Review of Microbiology* **66**: 285–303.
- Hazelbauer GL, Falke JJ, Parkinson JS. 2008.** Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends in Biochemical Sciences* **33**: 9–19.
- He SY, Nomura K, Whittam TS. 2004.** Type III protein secretion mechanism in mammalian and plant pathogens. *Biochimica et Biophysica Acta (BBA)* **1694**: 181–206.
- He P, Shan L, Lin NC, Martin GB, Kemmerling B, Nürnberger T, Sheen J. 2006.** Specific Bacterial Suppressors of MAMP Signaling Upstream of MAPKKK in *Arabidopsis* Innate Immunity. *Cell* **125**: 563–575.
- He S, Simpson BK, Sun H, Ngadi MO, Ma Y, Huang T. 2015.** *Phaseolus vulgaris* Lectins: A Systematic Review of Characteristics and Health Implications. *Critical reviews in food science and nutrition* **19**.
- Helmstädter A. 2010.** Beans and diabetes: *Phaseolus vulgaris* preparations as antihyperglycemic agents. *Journal of medicinal food* **13**: 251–4.
- Hersemann L, Wibberg D, Blom J, Goesmann A, Widmer F, Vorholter F-J, Kolliker R. 2017.** Comparative genomics of host adaptive traits in *Xanthomonas translucens* pv. *graminis*. *BMC genomics* **18**: 35.
- Heuer H, Yin YN, Xue QY, Smalla K, Guo JH. 2007.** Repeat domain diversity of avrBs3-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Applied and Environmental Microbiology* **73**: 4379–4384.
- Hillocks RJ, Madata CS, Chirwa R, Minja EM, Msolla S. 2006.** *Phaseolus* bean improvement in Tanzania, 1959-2005. *Euphytica* **150**: 215–231.
- Hirano SS, Upper CD. 2000.** Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*-a pathogen, ice nucleus, and epiphyte. *Microbiology and molecular biology reviews : MMBR* **64**: 624–53.
- Hnatuszko-Konka K, Kowalczyk T, Gerszberg A, Wiktorek-Smagur A, Kononowicz AK. 2014.** *Phaseolus vulgaris* - Recalcitrant potential. *Biotechnology Advances* **32**: 1205–1215.
- Hopkins CM, White FF, Choi SH, Guo A, Leach JE. 1992.** Identification of a family of avirulence genes from *xanthomonas oryzae* pv. *oryzae*. *Molecular plant-microbe interactions : MPMI* **5**: 451–459.
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. 2008.** Genevestigator V3 : A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics* **2008**: 1–5.
- Hsu PD, Lander ES, Zhang F. 2014.** Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**: 1262–1278.
- Huang C-L, Pu P-H, Huang H-J, Sung H-M, Liaw H-J, Chen Y-M, Chen C-M, Huang M-B, Osada N, Gojobori T, et al. 2015.** Ecological genomics in *Xanthomonas*: the nature of genetic adaptation with homologous recombination and host shifts. *BMC genomics* **16**: 1369.

- Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, Yang B, White FF, Wang N, Jones JB. 2014.** Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E521-9.
- Huguet-Tapia JC, Peng Z, Yang B, Yin Z, Liu S, White FF. 2016.** Complete Genome Sequence of the African Strain AX01947 of *Xanthomonas oryzae* pv. *oryzae*. *Genome announcements* **4**: 10-11.
- Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN, Coles M. 2006.** The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling. *Cell* **126**: 929-940.
- Hummel AW, Doyle EL, Bogdanove AJ. 2012.** Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**: 883-893.
- Hummel AW, Wilkins KE, Wang L, Cernadas RA, Bogdanove AJ. 2016.** A transcription activator-like effector from *Xanthomonas oryzae* pv. *oryzicola* elicits dose-dependent resistance in rice. *Molecular plant pathology* **18**: 55-56.
- Hutin M. 2015.** Criblage de la diversité d'*Oryza* spp. pour l'identification de nouvelles sources de résistances dépendantes des effecteurs TAL à *X. oryzae* pv. *oryzae*, agent de la bactériose vasculaire du riz.
- Hutin M, Césari S, Chalvon V, Michel C, Tran TT, Boch J, Koebnik R, Szurek B, Kroj T. 2016.** Ectopic activation of the rice NLR heteropair RGA4/RGA5 confers resistance to bacterial blight and bacterial leaf streak diseases. *Plant Journal* **88**: 43-55.
- Hutin M, Pérez-Quintero AL, Lopez C, Szurek B. 2015.** MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. *Frontiers in plant science* **6**: 535.
- Ifuku K, Endo T, Shikanai T, Aro E-M. 2011.** Structure of the chloroplast NADH dehydrogenase-like complex: Nomenclature for nuclear-encoded subunits. *Plant and Cell Physiology* **52**: 1560-1568.
- Indiana A. 2014.** Rôles du chimiotactisme et de la mobilité flagellaire dans la fitness des *Xanthomonas*.
- Indiana A, Briand M, Arlat M, Gagnevin L, Koebnik R, Noël L, Portier P, Darrasse A, Jacques MA. 2014.** Draft genome sequence of the flagellated *Xanthomonas fuscans* subsp. *fuscans* strain CFBP 4884. *Genome announcements* **2**: e00966-14.
- Ishihama N, Yoshioka H. 2012.** Post-translational regulation of WRKY transcription factors in plant immunity. *Current Opinion in Plant Biology* **15**: 431-437.
- Islam FMA, Renjifo J, Redden RJ, Basford KE, Beebe S. 2003.** Association Between Seed Coat Polyphenolics (Tannins) and Disease Resistance in Common Bean. *Plant Foods for Human Nutrition* **58**: 285-297.
- Ištvánek J, Jaroš M, Křenek A, Řepková J. 2014.** Genome assembly and annotation for red clover (*Trifolium pratense*; Fabaceae). *American journal of botany* **101**: 327-37.
- Iyer AS, McCouch SR. 2004.** The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. *Molecular plant-microbe interactions : MPMI* **17**: 1348-1354.
- Izoré T, Job V, Dessen A. 2011.** Biogenesis, regulation, and targeting of the type III secretion system. *Structure* **19**: 603-612.
- Jacobs JM, Pesce C, Lefeuvre P, Koebnik R. 2015.** Comparative genomics of a cannabis pathogen reveals insight into the evolution of pathogenicity in *Xanthomonas*. *Frontiers in Plant Science* **6**: 431.
- Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al. 2016.** Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas* : French Network on *Xanthomonads* (FNX). *Annual Review of Phytopathology* **54**: 163-187.
- Jacques MA, Bolot S, Charbit E, Darrasse A, Briand M, Arlat M, Gagnevin L, Koebnik R, Noël LD, Portier P, et al. 2013.** High-Quality Draft Genome Sequence of *Xanthomonas alfalfae* subsp. *alfalfae* Strain CFBP 3836. *Genome Announcements* **1**: e01035-13.
- Jacques M-A, Josi K, Darrasse A, Samson R. 2005.** *Xanthomonas axonopodis* pv. *phaseoli* ar. *fuscans* Is Aggregated in Stable Biofilm Population Sizes in the Phyllosphere of Field- Grown Beans. *Applied and Environmental Microbiology* **71**: 2008-2015.
- Jalan N, Aritua V, Kumar D, Yu F, Jones JB, Graham JH, Setubal JC, Wang N. 2011.** Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. *Journal of Bacteriology* **193**: 6342-6357.
- Jalan N, Kumar D, Yu F, Jones JB, Graham JH, Wang N. 2013.** Complete Genome Sequence of *Xanthomonas citri* subsp. *citri* Strain Aw12879, a Restricted-Host-Range Citrus Canker-Causing Bacterium. *Genome Announcements* **1**: e00235-13.
- Jha G, Rajeshwari R, Sonti R V. 2007.** Functional interplay between two *Xanthomonas oryzae* pv., *oryzae* secretion systems in modulating virulence on rice. *Molecular plant-microbe interactions : MPMI* **20**: 31-40.
- Jiang Y, Guo L, Ma X, Zhao X, Jiao B, Li C, Luo K. 2017.** The WRKY transcription factors PtrWRKY18 and PtrWRKY35 promote *Melampsora* resistance in *Populus*. *Tree Physiology* **37**: 665-675.

- Johnson TL, Abendroth J, Hol WGJ, Sandkvist M. 2006.** Type II secretion: From structure to function. *FEMS Microbiology Letters* **255**: 175–186.
- Jones AL. 1999.** PHASEOLUS BEAN: Post-harvest Operations. *AGSI/FAO*.
- Jones JDG, Dangl JL. 2006.** The plant immune system. *Nature* **444**: 323–329.
- Juillerat A, Bertonati C, Dubois G, Guyot V, Thomas S, Valtan J, Beurdeley M, Silva GH, Daboussi F, Duchateau P. 2014.** BurrH: a new modular DNA binding protein for genome engineering. *Scientific reports* **4**: 3831.
- Kaiser WJ, Vakili NG. 1978.** Insect transmission of pathogenic Xanthomonads to bean and cowpea in Puerto Rico. *Phytopathology* **68**: 1057–1063.
- Katoh K, Kuma KI, Toh H, Miyata T. 2005.** MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research* **33**: 511–518.
- Katungi E, Sperling L, Karanja D, Farrow A, Beebe S. 2011.** Relative importance of common bean attributes and variety demand in the drought areas of Kenya. *Journal of Development and Agricultural Economics* **3**: 411–422.
- Kauffman P, Leben C. 1974.** Soybean bacterial blight: Flower inoculation studies. *Phytopathology* **64**: 329–331.
- Kay S, Boch J, Bonas U. 2005.** Characterization of AvrBs3-like effectors from a Brassicaceae pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Molecular plant-microbe interactions : MPMI* **18**: 838–848.
- Kay S, Hahn S, Marois E, Hause G, Bonas U. 2007.** A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**: 648–651.
- Kay S, Hahn S, Marois E, Wieduwild R, Bonas U. 2009.** Detailed analysis of the DNA recognition motifs of the Xanthomonas type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant Journal* **59**: 859–871.
- Kazan K, Lyons R. 2014.** Intervention of Phytohormone Pathways by Pathogen Effectors. *The Plant cell* **26**: 2285–2309.
- Kelemu S, Cardona C, Segura G. 2004.** Antimicrobial and insecticidal protein isolated from seeds of Clitoria ternatea, a tropical forage legume. *Plant Physiology and Biochemistry* **42**: 867–873.
- Kim J-G, Taylor KW, Hotson A, Keegan M, Schmelz EA, Mudgett MB. 2008.** XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in xanthomonas-infected tomato leaves. *THE PLANT CELL* **20**: 1915–1929.
- Kimbrel JA, Givan SA, Temple TN, Johnson KB, Chang JH. 2011.** Genome sequencing and comparative analysis of the carrot bacterial blight pathogen, Xanthomonas hortorum pv. carotae M081, for insights into pathogenicity and applications in molecular diagnostics. *Molecular Plant Pathology* **12**: 580–594.
- Kingsley MT, Gabriel DW, Marlow GC, Roberts PD. 1993.** The opsX locus of Xanthomonas campestris affects host range and biosynthesis of lipopolysaccharide and extracellular polysaccharide. *Journal of Bacteriology* **175**: 5839–5850.
- Kline KA, Fälker S, Dahlberg S, Normark S, Henriques-Normark B. 2009.** Bacterial Adhesins in Host-Microbe Interactions. *Cell Host and Microbe* **5**: 580–592.
- Koebnik R. 2005.** TonB-dependent trans-envelope signalling: The exception or the rule? *Trends in Microbiology* **13**: 343–347.
- Koonin E V, Makarova KS, Aravind L. 2001.** Horizontal gene transfer in prokaryotes: quantification and classification. *Annual Review of Microbiology* **55**: 709–42.
- Koren S, Schatz MC, Walenz BP, Martin J, Howard J, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis D, et al. 2012.** Hybrid error correction and de novo assembly of single- molecule sequencing reads Sergey. *Nature Biotechnology* **30**: 693–700.
- Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012.** Selection-driven gene loss in bacteria. *PLoS Genetics* **8**: 1–7.
- Krell T, Lacal J, Busch A, Silva-Jiménez H, Guazzaroni M-E, Ramos JL. 2010.** Bacterial sensor kinases: diversity in the recognition of environmental signals. *Annual Review of Microbiology* **64**: 539–559.
- Krüger J, Thomas CM, Golstein C, Dixon MS, Smoker M, Tang S, Mulder L, Jones JDG. 2002.** A Tomato Cysteine Protease Required for Cf-2-Dependent Disease Resistance and Suppression of Autonecrosis. *Science* **296**: 744–747.
- Kuan T-L, Minsavage G V., Schaad NW. 1986.** Aerial dispersal of Xanthomonas campestris pv. campestris from naturally infected Brassica campestris. *Plant Disease* **70**: 409–413.
- Kuflu KM, Cuppels DA. 1997.** Development of a diagnostic DNA probe for xanthomonads causing bacterial spot of peppers and tomatoes. *Applied and Environmental Microbiology* **63**: 4462–4470.
- de Lange O, Binder A, Lahaye T. 2014a.** From dead leaf, to new life: TAL effectors as tools for synthetic biology. *Plant Journal* **78**: 753–771.
- de Lange O, Schandry N, Wunderlich M, Berendzen KW, Lahaye T. 2017.** Exploiting the sequence diversity of TALE-like repeats to vary the strength of dTALE-promoter interactions. *Synthetic Biology* **2**: 1–10.

- de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinowski J, Heuer H, Strauß A, Lahaye T. 2013.** Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytologist* **199**: 773–786.
- de Lange O de, Wolf C, Dietze J, Elsaesser J, Morbitzer R, Lahaye T. 2014b.** Programmable DNA-binding proteins from Burkholderia provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Research* **42**: 7436–7449.
- de Lange O, Wolf C, Thiel P, Krüger J, Kleusch C, Kohlbacher O, Lahaye T. 2015.** DNA-binding proteins from marine bacteria expand the known sequence diversity of TALE-like repeats. *Nucleic Acids Research* **43**: 10065–10080.
- Lapin D, Van den Ackerveken G. 2013.** Susceptibility to plant disease: More than a failure of host immunity. *Trends in Plant Science* **18**: 546–554.
- Van Larebeke N, Engler G, Holsters M, Van Den Elasccker S, Zaenen I, Schilperoort RA, Schell J. 1974.** Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* **252**: 169–170.
- Lareo LR, González V F. 1988.** Acceptability and nutritional quality of common beans (*Phaseolus vulgaris* L.): A bibliography. *Coli Colombia: Centro Internacional de Agricultura Tropical*: 1–279.
- Lau C-H, Zhu H, Tay JC-K, Li Z, Tay FC, Chen C, Tan W-K, Du S, Sia V-K, Phang R-Z, et al. 2014.** Genetic rearrangements of variable di-residue (RVD)-containing repeat arrays in a baculoviral TALEN system. *Molecular therapy. Methods & clinical development* **1**: 14050.
- Lavigne J-P, Botella E, O'Callaghan D. 2006.** Les systèmes de sécrétions de type IV et leurs effecteurs. *Pathologie Biologie* **54**: 296–303.
- Lavin M, Herendeen PS, Wojciechowski MF. 2005.** Evolutionary rates analysis of leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* **54**: 575–594.
- Lazo GR, Roffey R, Gabriel DW. 1987.** Pathovars of *Xanthomonas campestris* are distinguishable by restriction fragment-length polymorphism. *International Journal of Systematic Bacteriology* **37**: 214–221.
- Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS, et al. 2005.** The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Research* **33**: 577–586.
- Leng Y, Yang Y, Ren D, Huang L, Dai L, Wang Y, Chen L, Tu Z, Gao Y, Li X, et al. 2017.** A rice PECTATE LYASE-LIKE gene is required for plant growth and leaf senescence. *Plant Physiology* **174**: 1151–1166.
- Levin HL, Moran J V. 2011.** Dynamic interactions between transposable elements and their hosts. *Nature Reviews Genetics* **12**: 615–627.
- Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M, Sharakuu A, Mahjoub A, Wang G, Khan S, et al. 2013a.** Characterization and DNA-binding specificities of *Ralstonia* TAL-like effectors. *Molecular Plant* **6**: 1318–1330.
- Li T, Huang S, Zhou J, Yang B. 2013b.** Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. *Molecular Plant* **6**: 781–789.
- Li T, Liu B, Spalding MH, Weeks DP, Yang B. 2012a.** High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature Biotechnology* **30**: 390–392.
- Li L, Stoeckert CJJ, Roos DS. 2003.** OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research* **13**: 2178–2189.
- Li J, Wang N. 2012.** The *gpsX* gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in *Xanthomonas citri* subsp. *citri*. *BMC Microbiology* **12**.
- Li C, Wei J, Lin Y, Chen H. 2012b.** Gene silencing using the recessive rice bacterial blight resistance gene *xa13* as a new paradigm in plant breeding. *Plant Cell Reports* **31**: 851–862.
- Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, et al. 2010.** De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**: 265–72.
- Librado P, Rozas J. 2009.** DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.
- Lin S, Zhao Y, Zhu Y, Gosney M, Deng X, Wang X, Lin J. 2016.** An Effective and Inducible System of TAL Effector-Mediated Transcriptional Repression in Arabidopsis. *Molecular Plant* **9**: 1546–1549.
- Lindgren PB. 1997.** The role of *hrp* genes during plant-bacterial interactions. *Annual Review of Phytopathology* **35**: 129–152.
- Lindgren PB, Peet RC, Panopoulos NJ. 1986.** Gene cluster of *Pseudomonas syringae* pv. 'phaseolicola' controls pathogenicity of bean plants and hypersensitivity of nonhost plants. *Journal of Bacteriology* **168**: 512–522.
- Lindow SE, Brandl MT. 2003.** Microbiology of the Phyllosphere. *Applied and environmental microbiology* **69**: 1875–1883.
- Liu SY, Yu K, Huffner M, Park SJ, Banik M, Pauls KP, Crosby W. 2010.** Construction of a BAC library and a physical map of a major

- QTL for CBB resistance of common bean (*Phaseolus vulgaris* L.). *Genetica* **138**: 709–716.
- Liu Q, Yuan M, Zhou Y, Li X, Xiao J, Wang S. 2011.** A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant, Cell and Environment* **34**: 1958–1969.
- Lloyd SA, Norman M, Rosqvist R, Wolf-Watz H. 2001.** Yersinia YopE is targeted for type III secretion by N-terminal, not mRNA, signals. *Molecular Microbiology* **39**: 520–531.
- Love MI, Huber W, Anders S. 2014.** Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 1–21.
- Lu H, Patil P, Van Sluys MA, White FF, Ryan RP, Dow JM, Rabinowicz P, Salzberg SL, Leach JE, Sonti R, et al. 2008.** Acquisition and evolution of plant pathogenesis-associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. *PLoS ONE* **3**.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012.** SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**: 18.
- De Lúque JJR, Creamer B. 2014.** Major constraints and trends for common bean production and commercialization ; establishing priorities for future research. *Agronomía Colombiana* **32**: 423–431.
- Ma W, Dong FFT, Stavrinides J, Guttman DS. 2006.** Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genetics* **2**: 2131–2142.
- Mageshwaran V, Walia S, Annapurna K. 2012.** Isolation and partial characterization of antibacterial lipopeptide produced by *Paenibacillus polymyxa* HKA-15 against phytopathogen *Xanthomonas campestris* pv. *phaseoli* M-5. *World Journal of Microbiology and Biotechnology* **28**: 909–917.
- Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu JK. 2012.** Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Molecular Biology* **78**: 311–321.
- Mak ANS, Bradley P, Bogdanove AJ, Stoddard BL. 2013.** TAL effectors: Function, structure, engineering and applications. *Current Opinion in Structural Biology* **23**: 93–99.
- Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. 2012.** The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. *Science* **335**: 716–719.
- Mallory AC, Vaucheret H. 2006.** Functions of microRNAs and related small RNAs in plants. *Nature Genetics* **38**: S31–S36.
- Manson MD, Armitage JP, Hoch J a, Macnab RM. 1998.** Bacterial Locomotion and Signal Transduction. *Journal of bacteriology* **180**: 1009–1022.
- Marín-Rodríguez MC, Orchard J, Seymour GB. 2002.** Pectate lyases, cell wall degradation and fruit softening. *Journal of Experimental Botany* **53**: 2115–2119.
- Marois E, Van den Ackerveken G, Bonas U. 2002.** The xanthomonas type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Molecular plant-microbe interactions : MPMI* **15**: 637–646.
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015.** RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**: 1–5.
- Martin DP, Posada D, Crandall K a, Williamson C. 2005.** A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS research and human retroviruses* **21**: 98–102.
- Martin D, Rybicki E. 2000.** RDP: detection of recombination amongst aligned sequences. *Bioinformatics (Oxford, England)* **16**: 562–563.
- Mayer L, da Silva WP, Moura AB, Vendruscolo CT. 2010.** AFLP analysis of *Xanthomonas axonopodis* and *X. arboricola* strains used in xanthan production studies reveal high levels of polymorphism. *Brazilian Journal of Microbiology* **41**: 741–748.
- Melotto M, Underwood W, Koczan J, Nomura K, He SY. 2006.** Plant Stomata Function in Innate Immunity against Bacterial Invasion. *Cell* **126**: 969–980.
- Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, Bernstein BE. 2013.** Locus-specific editing of histone modifications at endogenous enhancers using programmable TALE-LSD1 fusions. *Nature biotechnology* **31**: 1133–6.
- Mercer AC, Gaj T, Fuller RP, Barbas CF. 2012.** Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Research* **40**: 11163–11172.
- Merda D, Bonneau S, Guimbaud JF, Durand K, Brin C, Boureau T, Lemaire C, Jacques MA, Fischer-Le Saux M. 2016.** Recombination-prone bacterial strains form a reservoir from which epidemic clones emerge in agroecosystems. *Environmental Microbiology Reports* **8**: 572–581.
- Merda D, Briand M, Bosis E, Rousseau C, Portier P, Barret M, Jacques M-A, Fischer-Le Saux M. 2017.** Ancestral acquisitions, gene

flow and multiple evolutionary trajectories of the type three secretion system and effectors in *Xanthomonas* plant pathogens. *Molecular Ecology*: 1–14.

Messina V. 2014. Nutritional and health benefits of dried beans. *The American Journal of Clinical Nutrition* **100**: 437–42.

Meziadi C, Richard MMS, Derquennes A, Thareau V, Blanchet S, Gratias A, Pflieger S, Geffroy V. 2016. Development of molecular markers linked to disease resistance genes in common bean based on whole genome sequence. *Plant Science* **242**: 351–357.

Mhedbi N. 2010. Approches cumulées de phylogénie et d'écologie pour déterminer les bases génétiques de la spécificité d'hôte des bactéries phytopathogènes, cas des *Xanthomonas* spp.

Mhedbi-Hajri N, Darrasse A, Pigné S, Durand K, Fouteau S, Barbe V, Manceau C, Lemaire C, Jacques M-A. 2011. Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. *BMC evolutionary biology* **11**: 67.

Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, Saux MF Le, Manceau C, Poussier S, Pruvost O, et al. 2013. Evolutionary History of the Plant Pathogenic Bacterium *Xanthomonas axonopodis*. *PLoS ONE* **8**: e58474.

Midha S, Patil PB. 2014. Genomic insights into the evolutionary origin of *Xanthomonas axonopodis* pv. citri and its ecological relatives. *Applied and Environmental Microbiology* **80**: 6266–6279.

Midha S, Ranjan M, Sharma V, Pinnaka AK, Patil PB. 2012. Genome sequence of *xanthomonas citri* pv. mangiferaeindicae strain LMG 941. *Journal of Bacteriology* **194**: 3031.

Miklas PN, Coyne DP, Grafton KE, Mutlu N, Reiser J, Lindgren DT, Singh SP. 2003. A major QTL for common bacterial resistance derives from the common bean great northern landrace cultivar Montana No. 5. *Euphytica* **131**: 137–146.

Miklas PN, Kelly JD, Beebe SE, Blair MW. 2006. Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* **147**: 105–131.

Misas-Villamil JC, van der Hoorn RAL, Doehlemann G. 2016. Papain-like cysteine proteases as hubs in plant immunity. *New Phytologist* **212**: 902–907.

Miyanari Y. 2014. TAL effector-mediated genome visualization (TGV). *Methods* **69**: 198–204.

Miyanari Y, Ziegler-Birling C, Torres-Padilla M-E. 2013. Live visualization of chromatin dynamics with fluorescent TALEs. *Nature structural & molecular biology* **20**: 1321–4.

Mkandawire ABC, Mabagala RB, Guzmán P, Gepts P, Gilbertson RL. 2004. Genetic Diversity and Pathogenic Variation of Common Blight Bacteria (*Xanthomonas campestris* pv. phaseoli and *X. campestris* pv. phaseoli var. fuscans) Suggests Pathogen Coevolution with the Common Bean. *Phytopathology* **94**: 593–603.

Molina MA, Godoy P, Ramos-González MI, Muñoz N, Ramos JL, Espinosa-Urgel M. 2005. Role of iron and the TonB system in colonization of corn seeds and roots by *Pseudomonas putida* KT2440. *Environmental Microbiology* **7**: 443–449.

Mora Y, Díaz R, Vargas-Lagunas C, Peralta H, Guerrero G, Aguilar A, Encarnación S, Girard L, Mora J. 2014. Nitrogen-fixing rhizobial strains isolated from common bean seeds: Phylogeny, physiology, and genome analysis. *Applied and Environmental Microbiology* **80**: 5644–5654.

Moreira LM, Almeida NF, Potnis N, Digiampietri L a, Adi SS, Bortolossi JC, da Silva AC, da Silva AM, de Moraes FE, de Oliveira JC, 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.

Morris CE, Monier J-M. 2003. THE ECOLOGICAL SIGNIFICANCE OF BIOFILM FORMATION BY PLANT-ASSOCIATED BACTERIA. *Annual Review of Phytopathology* **41**: 429–453.

Moscou MJ, Bogdanove AJ. 2009. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**: 1501.

Mudgett MB. 2005. New Insights To the Function of Phytopathogenic Bacterial Type Iii Effectors in Plants. *Annual Review of Plant Biology* **56**: 509–531.

Mudgett MB, Chesnokova O, Dahlbeck D, Clark ET, Rossier O, Bonas U, Staskawicz BJ. 2000. Molecular signals required for type III secretion and translocation of the *Xanthomonas campestris* AvrBs2 protein to pepper plants. *Proceedings of the National Academy of Sciences* **97**: 13324–13329.

Mus F, Crook MB, Garcia K, Costas AG, Geddes BA, Kouri ED, Paramasivan P, Ryu M-H, Oldroyd GED, Poole PS, et al. 2016. Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental Microbiology* **82**: 3698–3710.

Mutlu N, Vidaver AK, Coyne DP, Steadman JR, Lambrecht PA, Reiser J. 2008. Differential Pathogenicity of *Xanthomonas campestris* pv. phaseoli and *X. fuscans* subsp. *fuscans* Strains on Bean Genotypes with Common Blight Resistance. *Plant Disease* **92**: 546–554.

Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T. 2007. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *Journal of*

Bioscience and Bioengineering **104**: 34–41.

Nascimento V do, Arf O, Tarsitano MAA, Garcia NFS, Penteado M de S, Sabundjian MT. 2016. Economic analysis of rice and common bean production in succession to green manure crops and mechanical soil decompression in Brazilian Savannah. *Revista Ceres* **63**: 315–322.

Navarro L, Jay F, Nomura K, He SY, Voinnet O. 2008. Suppression of the microRNA pathway by bacterial effector proteins. *Science* **321**: 964–967.

Ndakidemi PA, Dakora FD, Nkonya EM, Ringo D, Mansoor H. 2006. Yield and economic benefits of common bean (*Phaseolus vulgaris*) and soybean (*Glycine max*) inoculation in northern Tanzania. *Australian Journal of Experimental Agriculture* **46**: 571–577.

Newell PD, Boyd CD, Sondermann H, O'Toole GA. 2011. A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. *PLoS Biology* **9**: e1000587.

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. 2014. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**: 935–949.

Noel L, Thieme F, Gabler J, Buttner D, Bonas U. 2003. XopC and XopJ, two novel type III effector proteins from *Xanthomonas campestris* pv. *vesicatoria*. *Journal of Bacteriology* **185**: 7092–7102.

Noël L, Thieme F, Nennstiel D, Bonas U. 2002. Two Novel Type III-Secreted Proteins of *Xanthomonas campestris* pv. *vesicatoria* Are Encoded within the hrp Pathogenicity Island Two Novel Type III-Secreted Proteins of *Xanthomonas campestris* pv. *vesicatoria* Are Encoded within the hrp Pathogenicity Island. *Journal of Bacteriology* **184**: 1340–1348.

Noinaj N, Guillier M, Barnard, TJ, Buchanan SK. 2010. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annual Review of Microbiology* **64**: 43–60.

Ochiai H, Inoue Y, Takeya M, Sasaki A, 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. *Japan Agricultural Research Quarterly* **39**: 275–287.

OEPP/CABI. 1986. *Xanthomonas axonopodis* pv. *phaseoli*. *Organismes de Quarantaine Pour l'Europe*.

OEPP/CABI. 1996. *Xanthomonas axonopodis* pv. *citri*. *Organismes de Quarantaine Pour l'Europe 2ème édit.*

Ojanen-Reuhs T, Kalkkinen N, Westerlund-Wikström B, Van Doorn J, Haahela K, Nurmiho-Lassila EL, Wengelnik K, Bonas U, Korhonen TK. 1997. Characterization of the fimA gene encoding bundle-forming fimbriae of the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. *Journal of Bacteriology* **179**: 1280–1290.

Oldroyd GED, Murray JD, Poole PS, Downie JA. 2011. The Rules of Engagement in the Legume-Rhizobial Symbiosis. *Annual Review of Genetics* **45**: 119–144.

de Oliveira Tostes G, Paula Araujo JS, Gomes Farias AR, Rossinol Frade DÂ, Lopes Olivares F. 2014. Detection and cellular localization of *Xanthomonas campestris* pv. *viticola* in seeds of commercial 'Red Globe' grapes. *Tropical Plant Pathology* **39**: 134–140.

van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods* **6**: 767–772.

van Opijnen T, Camilli A. 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nature Reviews Microbiology* **11**: 435–442.

Opio AF, Allen DJ, Teri JM. 1996. Pathogenic variation in *Xanthomonas campestris* pv. *phaseoli*, the causal agent of common bacterial blight in *Phaseolus* beans. *Plant Pathology (United Kingdom)* **45**: 1126–1133.

Opio F, Kimani PM, Musaana SM, Buruchara R. 2002. Research on common blight and halo blight of common beans in east and central Africa. *Bean Improvement Cooperative. Annual Report (USA)*. **45**: 160–161.

Opio A., Teri J., Allen D. 1993. Studies on Seed Transmission of *Xanthomonas campestris* pv *phaseoli* in Common Beans in Uganda. *African Crop Science Journal* **1**: 59–67.

Ormeño-Orrillo E, Servín-Garcidueñas LE, Rogel MA, González V, Peralta H, Mora J, Martínez-Romero J, Martínez-Romero E. 2015. Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. *Systematic and Applied Microbiology* **38**: 287–291.

Padidam M, Sawyer S, Fauquet CM. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**: 218–225.

Pandey A, Ray SK, Sonti R V., Rajeshwari R. 2014. GltB/D mutants of *xanthomonas oryzae* pv. *oryzae* are virulence deficient. *Current Microbiology* **68**: 105–112.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **57**: 2881–2887.

- Parkinson N, Cowie C, Heeney J, Stead D. 2009.** Phylogenetic structure of *Xanthomonas* determined by comparison of gyrB sequences. *International Journal of Systematic and Evolutionary Microbiology* **59**: 264–274.
- Peng L, Fukao Y, Fujiwara M, Takami T, Shikanai T. 2009.** Efficient Operation of NAD(P)H Dehydrogenase Requires Supercomplex Formation with Photosystem I via Minor LHCl in Arabidopsis. *The Plant Cell* **21**: 3623–3640.
- Peng Z, Hu Y, Xie J, Potnis N, Akhunova A, Jones J, Liu Z, White FF, Liu S. 2016.** Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics* **17**: 21.
- Pereira ALA, Carazzolle MF, Abe VY, de Oliveira MLP, Domingues MN, Silva JC, Cernadas RA, Benedetti CE. 2014.** Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. *BMC Genomics* **15**: 157.
- Pérez-Bueno ML, Pineda M, Díaz-casado E, Barón M. 2015.** Spatial and temporal dynamics of primary and secondary metabolism in *Phaseolus vulgaris* challenged by *Pseudomonas syringae*. *Physiologia Plantarum* **153**: 161–174.
- Pérez-Quintero AL. 2017.** Bioinformatic approaches to the study of TAL effector evolution and function.
- Pérez-Quintero AL, Lamy L, Gordon JL, Escalon A, Cunnac S, Szurek B, Gagnevin L. 2015.** QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Frontiers in Plant Science* **6**: 545.
- Pérez-Quintero AL, Rodriguez-R LM, Dereeper A, López C, Koebnik R, Szurek B, Cunnac S. 2013.** An Improved Method for TAL Effectors DNA-Binding Sites Prediction Reveals Functional Convergence in TAL Repertoires of *Xanthomonas oryzae* Strains. *PLoS ONE* **8**: e68464.
- Perilla-Henao LM, Casteel CL. 2016.** Vector-borne bacterial plant pathogens: Interactions with hemipteran insects and plants. *Frontiers in Plant Science* **7**: 1163.
- Perry G, Dinatale C, Xie W, Navabi A, Reinprecht Y, Crosby W, Yu K, Shi C, Pauls KP. 2013.** A comparison of the molecular organization of genomic regions associated with resistance to common bacterial blight in two *Phaseolus vulgaris* genotypes. *Frontiers in plant science* **4**: 318.
- Pesce C, Bolot S, Berthelot E, Bragard C, Cunnac S, Saux MF-L, Portier P, Arlat M, Gagnevin L, Jacques M-A, et al. 2013.** Draft Genome Sequence of *Xanthomonas translucens* pv. *graminis* Pathotype Strain CFBP 2053. *Genome announcements* **3**: e01174-15.
- Pesce C, Bolot S, Cunnac S, Portier P, Saux MF-L, Jacques M-A, Gagnevin L, Arlat M, Noël LD, Carrère S, et al. 2015.** High-Quality Draft Genome Sequence of the *Xanthomonas translucens* pv. *cerealis* Pathotype Strain CFBP 2541. *Genome announcements* **3**: e01574-14.
- Pieretti I, Royer M, Barbe V, Carrere S, Koebnik R, Cociancich S, Couloux A, Darrasse A, Gouzy J, Jacques MA, 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**: 1–15.
- Pietrarelli L, Balestra G, Varvaro L. 2006.** Effects of simulated rain on *Pseudomonas syringae* pv. *tomato* populations on tomato plants. *Journal of Plant Pathology* **88**: 245–251.
- Poinsot V, Carpéné MA, Couderc F. 2012.** Coupled Mass Spectrometric Strategies for the Determination of Carbohydrates at Very Low Concentrations: The Case of Polysaccharides Involved in the Molecular Dialogue Between Plants and Rhizobia. The Complex World of Polysaccharides. Chapter 11.
- Portaliou AG, Tsois KC, Loos MS, Zorzini V, Economou A. 2016.** Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in Biochemical Sciences* **41**: 175–189.
- Posada D, Crandall KA. 2001.** Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of the National Academy of Sciences* **98**: 13757–13762.
- Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, Ryan RP, Sharlach M, Behlau F, Dow JM, Momol M, et al. 2011.** Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC Genomics* **12**: 146.
- Pottinger P, Reller LB, Ryan KJ. 2004.** *Sherris Medical Microbiology*.
- Punina N V., Ignatov AN, Pekhtereva ES, Kornev KP, Matveeva E V., Polityko VA, Budenkov NI, Schaad NW. 2009.** Occurrence of *Xanthomonas campestris* pv. *raphani* on tomato plants in the Russian Federation. *Acta Horticulturae* **808**: 287–290.
- Purcell A. 2013.** Paradigms: Examples from the Bacterium *Xylella fastidiosa*. *Annual Review of Phytopathology* **51**: 339–356.
- Qi J, Luo H, Hao B. 2004.** CVTree: A phylogenetic tree reconstruction tool based on whole genomes. *Nucleic Acids Research* **32**: 45–47.
- 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, et al. 2005.** Comparative and functional genomic analyses of the pathogenicity of phytopathogen. *Genome Research* **15**: 757–767.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005.** InterProScan: Protein domains identifier. *Nucleic Acids Research* **33**: 116–120.

- Rademaker JLW, Hoste B, Louws FJ, Kersters K, Swings J, Vauterin L, Vauterin P, Bruijn FJ De. 2000.** Comparison of AFLP and rep-PCR genomic fingerprinting with DNA – DNA homology studies: *Xanthomonas* as a model system. *International Journal of Systematic and Evolutionary Microbiology* **50**: 665–677.
- Rademaker JLW, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J, de Bruijn FJ. 2005.** A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* **95**: 1098–1111.
- Ramírez Rivera NG, García-Salinas C, Aragão FJL, Díaz de la Garza RI. 2016.** Metabolic engineering of folate and its precursors in Mexican common bean (*Phaseolus vulgaris* L.). *Plant Biotechnology Journal* **14**: 2021–2032.
- Ran FA, Hsu PD, Lin C, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S. 2013.** Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**: 1380–1389.
- Records AR, Gross DC. 2010.** Sensor kinases RetS and LadS regulate *Pseudomonas syringae* Type VI secretion and virulence factors. *Journal of Bacteriology* **192**: 3584–3596.
- Richter M, Rosselló-Móra R. 2009.** Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 19126–31.
- Richter A, Streubel J, Blücher C, Szurek B, Reschke M, Grau J, Boch J. 2014.** A TAL effector repeat architecture for frameshift binding. *Nature communications* **5**: 3447.
- Rinaldi FC, Doyle LA, Stoddard BL, Bogdanove AJ. 2017.** The effect of increasing numbers of repeats on TAL effector DNA binding specificity. *Nucleic Acids Research* **45**: 6960–6970.
- Robène I, Bolot S, Pruvost O, Arlat M, Noël L, Carrère S, Jacques M, Koebnik R, Gagnevin L. 2016.** High-Quality Draft Genome Sequences of Two *Xanthomonas* Pathotype Strains Infecting Aroid Plants. *Genome announcements* **4**: e00902-16.
- Rodriguez-R LM, Grajales A, Arrieta-Ortiz M, Salazar C, Restrepo S, Bernal A. 2012.** Genomes-based phylogeny of the genus *Xanthomonas*. *BMC Microbiology* **12**: 1–14.
- Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T. 2007.** Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**: 645–648.
- Römer P, Recht S, Strauss T, Elsaesser J, Schornack S, Boch J, Wang S, Lahaye T. 2010.** Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist* **187**: 1048–1057.
- Romero-Arenas O, Damián Huato MA, Rivera Tapia JA, Báez SA, Huerta LM, Huerta, Cabrera E. 2013.** The Nutritional value of Beans (*Phaseolus vulgaris* L.) and its importance for Feeding of Rural communities in Puebla-Mexico. *International Research Journal of Biological Sciences* **2**: 59–65.
- Roumagnac P, Pruvost O, Chiroleu F, Hughes G. 2004.** Spatial and Temporal Analyses of Bacterial Blight of Onion Caused by *Xanthomonas axonopodis* pv. *allii*. *Phytopathology* **94**: 138–46.
- Rousseau C, Belin E, Bove E, Rousseau D, Fabre F, Berruyer R, Guillaumès J, Manceau C, Jacques M-A, Boureau T. 2013.** High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. *Plant methods* **9**: 17.
- Rousseau C, Hunault G, Gaillard S, Bourbeillon J, Montiel G, Simier P, Campion C, Jacques M-A, Belin E, Boureau T. 2015.** Phenoplant: a web resource for the exploration of large chlorophyll fluorescence image datasets. *Plant methods* **11**.
- Roux B, Bolot S, Guy E, Denancé N, Lautier M, Jardinaud M-F, Fischer-Le Saux M, Portier P, Jacques M-A, Gagnevin L, et al. 2015.** Genomics and transcriptomics of *Xanthomonas campestris* species challenge the concept of core type III effectome. *BMC genomics* **16**: 975.
- Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017a.** *Xanthomonas* adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors. *BMC Genomics* **18**: 670.
- Ruh M, Briand M, Bonneau S, Jacques M, Chen NWG. 2017b.** First Complete Genome Sequences of *Xanthomonas citri* pv. *vignicola* Strains CFBP7111, CFBP7112, and CFBP7113 Obtained Using Long-Read Technology. *Genome Announcements* **5**: e00813-17.
- Ryan RP, Vorhölter F-J, Potnis N, Jones JB, Van Sluys M-A, Bogdanove AJ, Dow JM. 2011.** Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. *Nature Reviews Microbiology* **9**: 344–355.
- Saettler AW, Cafati CR, Weller DM. 1986.** Nonoverwintering of *Xanthomonas* Bean Blight Bacteria in Michigan. *Plant Disease* **70**: 285–287.
- Sallet E, Gouzy J, Schiex T. 2014.** EuGene-PP: A next-generation automated annotation pipeline for prokaryotic genomes. *Bioinformatics* **30**: 2659–2661.
- Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, Furutani A, Ochiai H, Delcher AL, Kelley D, et al. 2008.** Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. *BMC Genomics* **9**: 204.

- Sandkvist M. 2001.** Biology of type II secretion. *Molecular Microbiology* **40**: 271–283.
- Santalla M, Rodiño AP, De Ron AM. 2002.** Allozyme evidence supporting southwestern Europe as a secondary center of genetic diversity for the common bean. *Theoretical and Applied Genetics* **104**: 934–944.
- Santalla M, de Ron AM, de La Fuente M. 2010.** Integration of genome and phenotypic scanning gives evidence of genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe. *Theoretical and Applied Genetics* **120**: 1635–1651.
- Santini CL, Ize B, Chanal A, Müller M, Giordano G, Wu LF. 1998.** A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO Journal* **17**: 101–112.
- Santos AS, Bressan-Smith RE, Pereira MG, Rodrigues R, Ferreira CF. 2003.** Genetic linkage map of *Phaseolus vulgaris* and identification of QTLs responsible for resistance to *Xanthomonas axonopodis* pv. *phaseoli*. *Fitopatologia Brasileira* **28**: 5–10.
- Sarkar SF, Guttman DS. 2004.** Evolution of the Core Genome of *Pseudomonas syringae*, a Highly Clonal, Endemic Plant Pathogen. *Applied and Environmental Microbiology* **70**: 1999–2012.
- Sati SC, Joshi S. 2011.** Antibacterial activities of *Ginkgo biloba* L. leaf extracts. *TheScientificWorldJournal* **11**: 2237–42.
- Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, et al. 2008.** Genome structure of the legume, *Lotus japonicus*. *DNA Research* **15**: 227–239.
- Schaad NW, Dianese JC. 1981.** Cruciferous weeds as sources of inoculum of *Xanthomonas campestris* in black rot crucifers. *Phytopathology* **71**: 1215–1220.
- Schaad NW, Postnikova E, Lacy GH, Sechler A, Agarkova I, Stromberg PE, Stromberg VK, Vidaver AK. 2005.** Reclassification of *Xanthomonas campestris* pv. *citri* (ex Hasse 1915) Dye 1978 forms A, B/C/D, and E as *X. smithii* subsp. *citri* (ex Hasse) sp. nov. nom. rev. comb. nov., *X. fuscans* subsp. *aurantifolii* (ex Gabriel 1989) sp. nov. nom. rev. comb. nov., and *X. Systematic and applied microbiology* **28**: 494–518.
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, et al. 2010.** Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178–83.
- Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C, et al. 2014.** A reference genome for common bean and genome-wide analysis of dual domestications. *Nature genetics* **46**: 707–13.
- Schneider TD, Stephens RM. 1990.** Sequence logos: A new way to display consensus sequences. *Nucleic Acids Research* **18**: 6097–6100.
- Scholze H, Boch J. 2011.** TAL effectors are remote controls for gene activation. *Current Opinion in Microbiology* **14**: 47–53.
- Schornack S, Ballvora A, Gürlebeck D, Peart J, Ganai M, Baker B, Bonas U, Lahaye T. 2004.** The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *Plant Journal* **37**: 46–60.
- Schornack S, Minsavage G V., Stall RE, Jones JB, Lahaye T. 2008.** Characterization of AvrHah1, a novel AvrBs3-like effector from *Xanthomonas gardneri* with virulence and avirulence activity. *New Phytologist* **179**: 546–556.
- Schornack S, Moscou MJ, Ward ER, Horvath DM. 2013.** Engineering Plant Disease Resistance Based on TAL Effectors. *Annual Review of Phytopathology* **51**: 383–406.
- Schornack S, Peter K, Bonas U, Lahaye T. 2005.** Expression levels of avrBs3-like genes affect recognition specificity in tomato Bs4- but not in pepper Bs3-mediated perception. *Molecular plant-microbe interactions : MPMI* **18**: 1215–1225.
- Schreiber T, Bonas U. 2014.** Repeat 1 of TAL effectors affects target specificity for the base at position zero. *Nucleic Acids Research* **42**: 7160–7169.
- Schultz T, Gabrielson RL. 1986.** *Xanthomonas campestris* pv. *campestris* in western washington cricifer seed fields: occurrence and survival. *Phytopathology* **76**: 1306–1309.
- Schuster ML, Coyne DP. 1974.** Survival Mechanisms of Phytopathogenic Bacteria. *Annual Review of Phytopathology* **12**: 199–221.
- Schuster ML, Coyne DP, Hoff B. 1973.** Comparative Virulence of *Xanthomonas phaseoli* Strains from Uganda, Colombia, and Nebraska. *Plant Disease Reporter* **57**: 74–75.
- Schwartz AR, Morbitzer R, Lahaye T, Staskawicz BJ. 2017.** TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. *Proceedings of the National Academy of Sciences of the United States of America* **114**: E897–E903.
- Schwartz AR, Potnis N, Timilsina S, Wilson M, Patané J, Martins J, Minsavage G V., Dahlbeck D, Akhunova A, Almeida N, et al. 2015.** Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. *Frontiers in Microbiology* **6**: 1–17.

- Senthil-Kumar M, Mysore KS. 2013.** Nonhost Resistance Against Bacterial Pathogens: Retrospectives and Prospects. *Annual Review of Phytopathology* **51**: 407–427.
- Shade A, Jacques MA, Barret M. 2017.** Ecological patterns of seed microbiome diversity, transmission, and assembly. *Current Opinion in Microbiology* **37**: 15–22.
- Shelton A, Hunter J. 1985.** Evaluation of the potential of the flea beetle *Phyllotreta cruciferae* to transmit *Xanthomonas campestris* pv. *campestris*, causal agent of black rot of crucifers. *Canadian Journal of Plant Pathology* **7**: 308–310.
- Shi C, Chaudhary S, Yu K, Park SJ, Navabi A, McClean PE. 2011a.** Identification of candidate genes associated with CBB resistance in common bean HR45 (*Phaseolus vulgaris* L.) using cDNA-AFLP. *Molecular Biology Reports* **38**: 75–81.
- Shi C, Navabi A, Yu K. 2011b.** Association mapping of common bacterial blight resistance QTL in Ontario bean breeding populations. *BMC Plant Biology* **11**: 1–11.
- Shi C, Yu K, Xie W, Perry G, Navabi A, Pauls KP, Miklas PN, Fourie D. 2012.** Development of candidate gene markers associated to common bacterial blight resistance in common bean. *Theoretical and Applied Genetics* **125**: 1525–37.
- da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Van Sluys MA, Almeida NF, Alves LMC, et al. 2002.** Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**: 459–463.
- Simpson A, Reinach F, Arruda P, Abreu F, Acencio M, Alvarenga R, Alves L, Araya J, Baia G, Baptista C, et al. 2000.** The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature* **406**: 151–157.
- Singh SP. 2001.** Broadening the genetic base of common bean cultivars: A review. *Crop Science* **41**: 1659–1675.
- Singh SP, Muñoz CG. 1999.** Resistance to Common Bacterial Blight among *Phaseolus* Species and Common Bean Improvement. *Crop Science* **39**: 80–89.
- Singh SP, Schwartz HF. 2010.** Breeding common bean for resistance to diseases: A review. *Crop Science* **50**: 2199–2223.
- Van Sluys M, de Oliveira M, Monteiro-Vitorello C, Miyaki C, Furlan L, Camargo L, da Silva A, Moon D, Takita M, Lemos E, et al. 2003.** Comparative Analyses of the Complete Genome Sequences of Pierce's Disease and Citrus Variegated Chlorosis Strains of *Xylella fastidiosa*. *Journal of bacteriology* **185**: 1018–1026.
- Smith JM. 1992.** Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* **34**: 126–129.
- Smith JJ, Jones DR, Karamura E, Blomme G, Turyagyenda FL. 2008.** An analysis of the risk from *Xanthomonas campestris* pv. *musacearum* to banana cultivation in Eastern , Central and Southern Africa. *Bioversity International, Montpellier, France*.
- Solé M, Scheibner F, Hoffmeister AK, Hartmann N, Hause G, Rother A, Jordan M, Lautier M, Arlat M, Büttner D. 2015.** *Xanthomonas campestris* pv. *vesicatoria* secretes proteases and xylanases via the Xps type II secretion system and outer membrane vesicles. *Journal of Bacteriology* **197**: 2879–2893.
- Sory M-P, Boland A, Lambermont I, Cornelis GR. 1995.** Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. *Microbiology* **92**: 11998–12002.
- Soucy SM, Huang J, Gogarten JP. 2015.** Horizontal gene transfer: building the web of life. *Nature Reviews Genetics* **16**: 472–482.
- Spago FR, Ishii Mauro CS, Oliveira AG, Beranger JPO, Cely MVT, Stanganelli MM, Simionato AS, San Martin JAB, Andrade CGTJ, Mello JCP, et al. 2014.** *Pseudomonas aeruginosa* produces secondary metabolites that have biological activity against plant pathogenic *Xanthomonas* species. *Crop Protection* **62**: 46–54.
- Stella S, Molina R, Bertonatti C, Juillerrat A, Montoya G. 2014.** Expression, purification, crystallization and preliminary X-ray diffraction analysis of the novel modular DNA-binding protein BurrH in its apo form and in complex with its target DNA. *Acta Crystallographica Section F Structural Biology Communications* **70**: 87–91.
- Stewart V. 2014.** The HAMP signal-conversion domain: Static two-state or dynamic three-state? *Molecular Microbiology* **91**: 853–857.
- Stoodley P, Sauer K, Davies DG, Costerton JW. 2002.** Biofilms as complex differentiated communities. *Annual Review of Microbiology* **56**: 187–209.
- Strauß T, van Poecke RMP, Strauss A, Römer P, Minsavage G V, Singh S, Wolf C, Strauss A, Kim S, Lee H-A, et al. 2012.** RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 19480–5.
- Streubel J, Blücher C, Landgraf A, Boch J. 2012.** TAL effector RVD specificities and efficiencies. *Nature Biotechnology* **30**: 593–595.
- Streubel J, Pesce C, Hutin M, Koebnik R, Boch J, Szurek B. 2013.** Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist* **200**: 808–819.
- Studholme DJ, Kemen E, MacLean D, Schornack S, Aritua V, Thwaites R, Grant M, Smith J, Jones JDG. 2010.** Genome-wide sequencing data reveals virulence factors implicated in banana *Xanthomonas* wilt. *FEMS Microbiology Letters* **310**: 182–192.
- Studholme DJ, Wasukira A, Paszkiewicz K, Aritua V, Thwaites R, Smith J, Grant M. 2011.** Draft genome sequences of *xanthomonas*

- sacchari and two banana-associated xanthomonads reveal insights into the xanthomonas group 1 clade. *Genes* **2**: 1050–1065.
- Su J, Zou X, Huang L, Bai T, Liu S, Yuan M, Chou S-H, He Y-W, Wang H, He J. 2016.** DgcA, a diguanylate cyclase from *Xanthomonas oryzae* pv. *oryzae* regulates bacterial pathogenicity on rice. *Scientific reports* **6**: 25978.
- Sugio A, Yang B, Zhu T, White FF. 2007.** Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes OsTFIIAgamma1 and OsTFX1 during bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 10720–5.
- Sundin GW. 2007.** Genomic Insights into the Contribution of Phytopathogenic Bacterial Plasmids to the Evolutionary History of Their Hosts. *Annual Review of Phytopathology* **45**: 129–151.
- Sung P, Klein H. 2006.** Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature reviews. Molecular cell biology* **7**: 739–750.
- Svetleva D, Velcheva M, Bhowmik G. 2003.** Biotechnology as a useful tool in common bean (*Phaseolus vulgaris* L.) improvement. *Euphytica* **131**: 189–200.
- Szurek B, Marois E, Bonas U, Van Ackerveken G Den. 2001.** Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: Protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. *Plant Journal* **26**: 523–534.
- Szurek B, Rossier O, Hause G, Bonas U. 2002.** Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Molecular Microbiology* **46**: 13–23.
- Taghavi S, Garafola C, Monchy S, Newman L, Hoffman A, Weyens N, Barac T, Vangronsveld J, Van Der Lelie D. 2009.** Genome survey and characterization of endophytic bacteria exhibiting a beneficial effect on growth and development of poplar trees. *Applied and Environmental Microbiology* **75**: 748–757.
- Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, et al. 2014.** An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC genomics* **15**: 312.
- Tao F, Swarup S, Zhang LH. 2010.** Quorum sensing modulation of a putative glycosyltransferase gene cluster essential for *Xanthomonas campestris* biofilm formation. *Environmental Microbiology* **12**: 3159–3170.
- Tayi L, Kumar S, Nathawat R, Haque AS, Maku R V, Patel HK, Sankaranarayanan R, Sonti R V. 2017.** A mutation in an exoglucanase of *Xanthomonas oryzae* pv. *oryzae* that confers an endo mode of activity affects bacterial virulence but not the induction of immune responses in rice. *Molecular Plant Pathology*.
- Teper D, Burstein D, Salomon D, Gershovitz M, Pupko T, Sessa G. 2016.** Identification of novel *Xanthomonas euvesicatoria* type III effector proteins by a machine-learning approach. *Molecular Plant Pathology* **17**: 398–411.
- Thieme F, Koebnik R, Bekel T, Berger C, Boch J, Büttner D, Caldana C, Gaigalat L, Goesmann A, Kay S. 2005.** Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. *Journal of bacteriology* **187**: 7254–7266.
- Thomma BPHJ, Nürnberger T, Joosten MHAJ. 2011.** Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *The Plant cell* **23**: 4–15.
- Tochio N, Umehara K, Uewaki J, Flechsig H, Kondo M, Dewa T, Sakuma T, Yamamoto T, Saitoh T, Togashi Y, et al. 2016.** Non-RVD mutations that enhance the dynamics of the TAL repeat array along the superhelical axis improve TALEN genome editing efficacy. *Scientific reports* **6**.
- Todorović B, Potočnik I, Rekanović E, Stepanović M, Kostić M, Ristić M, Milijašević-Marčić S. 2016.** Toxicity of twenty-two plant essential oils against pathogenic bacteria of vegetables and mushrooms. *Journal of environmental science and health, Part B* **51**: 832–839.
- Toth IK, Hyman LJ, Taylor R, Birch PRJ. 1998.** PCR-based detection of *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* in plant material and its differentiation from *X. c. pv. phaseoli*. *Journal of Applied Microbiology* **85**: 327–336.
- Tushemereirwe W, Kangire A, Ssekiwoko F, Offord LC, Crozier J, Boa E, Rutherford M, Smith JJ. 2004.** First report of *Xanthomonas campestris* pv. *musacearum* on banana in Uganda. *Plant Pathology* **53**: 802.
- Vandemark GJ, Fourie D, Miklas PN. 2008.** Genotyping with real-time PCR reveals recessive epistasis between independent QTL conferring resistance to common bacterial blight in dry bean. *Theoretical and Applied Genetics* **117**: 513–522.
- Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe a G, Cannon S, Baek J, Rosen BD, Tar'an B, et al. 2013.** Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nature Biotechnology* **31**: 240–246.
- Vauterin L, Hoste B, Kersters K, Swings J. 1995.** Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**: 472–489.

- Vauterin L, Rademaker J, Swings J. 2000.** Synopsis on the taxonomy of the genus *Xanthomonas*. *Phytopathology* **90**: 677–682.
- Vera Cruz CM, Bai J, Ona I, Leung H, Nelson RJ, Mew TW, Leach JE. 2000.** Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 13500–5.
- Vicente JG, Holub EB. 2013.** *Xanthomonas campestris* pv. *Campestris* (cause of black rot of crucifers) in the genomic era is still a worldwide threat to brassica crops. *Molecular Plant Pathology* **14**: 2–18.
- Vinatzer BA, Monteil CL, Clarke CR. 2014.** Harnessing Population Genomics to Understand How Bacterial Pathogens Emerge, Adapt to Crop Hosts, and Disseminate. *Annual Review of Phytopathology* **52**: 19–43.
- Vladimirov N, Sourjik V. 2009.** Chemotaxis: How bacteria use memory. *Biological Chemistry* **390**: 1097–1104.
- Vorhölter F-J, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Patschkowski T, Rückert C, Schmid J, et al. 2008.** The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. *Journal of Biotechnology* **134**: 33–45.
- Vorhölter F-J, Wiggerich H-G, Scheidle H, Sidhu VK, Mrozek K, Küster H, Pühler A, Niehaus K. 2012.** Involvement of bacterial TonB-dependent signaling in the generation of an oligogalacturonide damage-associated molecular pattern from plant cell walls exposed to *Xanthomonas campestris* pv. *campestris* pectate lyases. *BMC Microbiology* **12**: 239.
- Voytas DF, Joung JK. 2009.** DNA binding made easy. *Science (New York, N.Y.)* **326**: 1491–1492.
- Wadhams GH, Armitage JP. 2004.** Making sense of it all: bacterial chemotaxis. *Nature reviews. Molecular cell biology* **5**: 1024–1037.
- Walcott RR, Gitaitis RD, Castro a C. 2003.** Role of Blossoms in Watermelon Seed Infestation by *Acidovorax avenae* subsp. *citrulli*. *Phytopathology* **93**: 528–534.
- Walker TS, Bais HP, De E, Schweizer HP, Rahme LG, Fall R, Vivanco JM. 2004.** *Pseudomonas aeruginosa*-Plant Root Interactions. Pathogenicity, Biofilm Formation, and Root Exudation. *Plant Physiology* **134**: 320–331.
- Wallden K, Rivera-Calzada A, Waksman G. 2010.** Type IV secretion systems: versatility and diversity in function. *Cellular Microbiology* **12**: 1203–1212.
- Wang L, Rinaldi FC, Singh P, Doyle EL, Dubrow ZE, Tu T, Pérez-quintero AL, Szurek B, Bogdanove AJ. 2017.** TAL effectors drive transcription bidirectionally in plants. *Molecular Plant* **10**: 285–296.
- Wang L, Vinogradov E V., Bogdanove AJ. 2013.** Requirement of the lipopolysaccharide O-chain biosynthesis gene *wxocB* for type III secretion and virulence of *Xanthomonas oryzae* pv. *Oryzicola*. *Journal of Bacteriology* **195**: 1959–1969.
- Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, Qin T, Li Y, Che J, Zhang M, et al. 2015.** *XA23* is an executor protein and confers broad-spectrum disease resistance in rice. *Molecular Plant* **8**: 290–302.
- Wasukira A, Tayebwa J, Thwaites R, Paszkiewicz K, Aritua V, Kubiriba J, Smith J, Grant M, Studholme DJ. 2012.** Genome-wide sequencing reveals two major sub-lineages in the genetically monomorphic pathogen *Xanthomonas campestris* pathovar *musacearum*. *Genes* **3**: 361–377.
- Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. 2011.** Assembly of designer tal effectors by golden gate cloning. *PLoS ONE* **6**.
- Weller DM, Saetler AW. 1980.** Evaluation of Seedborne *Xanthomonas phaseoli* and *X. phaseoli* var. *fuscans* as Primary Inocula in Bean Blights. *Phytopathology* **70**: 148–152.
- Wengelnik K, Van den Ackerveken G, Bonas U. 1996.** HrpG, a key hrp regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. *Molecular plant-microbe interactions : MPMI* **9**: 704–712.
- Wengelnik K, 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*. *Journal of Bacteriology* **178**: 3462–3469.
- Wichmann F, Vorhölter F-J, Hersemann L, Widmer F, Blom J, Niehaus K, Reinhard S, Conradin C, Kölliker R. 2013.** The noncanonical type III secretion system of *Xanthomonas translucens* pv. *graminis* is essential for forage grass infection. *Molecular Plant Pathology* **14**: 576–588.
- Wiedenbeck J, Cohan FM. 2011.** Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiology Reviews* **35**: 957–976.
- Wilkins KE, Booher NJ, Wang L, Bogdanove AJ. 2015.** TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. *Frontiers in Plant Science* **6**: 1–15.
- Williams PH. 1980.** Black rot: A continuing Threat to World Crucifers. *Plant disease* **64**: 736–742.
- Wortmann CS, Kirkby RA, Eledu CA, Allen DJ. 1998.** Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. *Centro*

Internacional de Agricultura Tropical (CIAT) **297**: 133.

Wu L, Goh ML, Sreekala C, Yin Z. 2008. XA27 depends on an amino-terminal signal-anchor-like sequence to localize to the apoplast for resistance to *Xanthomonas oryzae* pv *oryzae*. *Plant physiology* **148**: 1497–1509.

Wu XM, Li YR, Zou LF, Chen GY. 2007. Gene-for-gene relationships between rice and diverse avrBs3/pthA avirulence genes in *Xanthomonas oryzae* pv. *oryzae*. *Plant Pathology* **56**: 26–34.

Wu Y, Zhang D, Chu JY, Boyle P, Wang Y, Brindle ID, De Luca V, Després C. 2012. The Arabidopsis NPR1 Protein Is a Receptor for the Plant Defense Hormone Salicylic Acid. *Cell Reports* **1**: 639–647.

Wuichet K, Alexander RP, Zhulin IB. 2007. Comparative Genomic and Protein Sequence Analyses of a Complex System Controlling Bacterial Chemotaxis. *Methods in Enzymology* **422**: 1–24.

Wuichet K, Zhulin IB. 2003. Molecular evolution of sensory domains in cyanobacterial chemoreceptors. *Trends Microbiology* **11**: 200–203.

Xia Y, Suzuki H, Borevitz J, Blount J, Guo Z, Patel K, Dixon RA, Lamb C. 2004. An Extracellular Aspartic Protease Functions in Arabidopsis Disease Resistance Signaling. *EMBO Journal* **23**: 980–988.

Xie G, Bonner CA, Brettin T, Gottardo R, Keyhani NO, Jensen RA. 2003. Lateral gene transfer and ancient paralogy of operons containing redundant copies of tryptophan-pathway genes in *Xylella* species and in heterocystous cyanobacteria. *Genome biology* **4**: R14.

Xin X-F, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, Chang JH, Zipfel C, He SY. 2016. Bacteria establish an aqueous living space in plants crucial for virulence. *Nature* **539**: 524–529.

El Yacoubi B, Brunings AM, Yuan Q, Shankar S, Gabriel DW. 2007. In planta horizontal transfer of a major pathogenicity effector gene. *Applied and Environmental Microbiology* **73**: 1612–1621.

Yamaguchi K, Nakamura Y, Ishikawa K, Yoshimura Y, Tsuge S, Kawasaki T. 2013. Suppression of rice immunity by *Xanthomonas oryzae* type III effector Xoo2875. *Bioscience, biotechnology, and biochemistry* **77**: 796–801.

Yang Y, Gabriel DW. 1995a. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. *Molecular plant-microbe interactions : MPMI* **8**: 627–631.

Yang Y, Gabriel DW. 1995b. Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities . These include : Intragenic Recombination of a Single Plant Pathogen Gene Provides a Mechanism for the Evolution of New Host Specifi. *Journal of Bacteriology* **177**: 4963–4968.

Yang C-H, Gavilanes-Ruiz M, Okinaka Y, Vedel R, Berthuy I, Boccara M, Chen JW-T, Perna NT, Keen NT. 2002. hrp genes of *Erwinia chrysanthemi* 3937 are important virulence factors. *Molecular plant-microbe interactions : MPMI* **15**: 472–480.

Yang B, Sugio A, White FF. 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 10503–8.

Yang K, Tian Z, Chen C, Luo L, Zhao B, Wang Z, Yu L, Li Y, Sun Y, Li W, et al. 2015. Genome sequencing of adzuki bean (*Vigna angularis*) provides insight into high starch and low fat accumulation and domestication. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 13213–8.

Yang B, White FF. 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular plant-microbe interactions : MPMI* **17**: 1192–200.

Yang J, Zhang Y, Yuan P, Zhou Y, Cai C, Ren Q, Wen D, Chu C, Qi H, Wei W. 2014. Complete decoding of TAL effectors for DNA recognition. *Cell research* **24**: 628–31.

Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H, et al. 2011. The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* **480**: 520–4.

Young JM, Park DC, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**: 366–377.

Yuan M, Chu Z, Li X, Xu C, Wang S. 2009. Pathogen-induced expressional loss of function is the key factor in race-specific bacterial resistance conferred by a recessive R gene xa13 in rice. *Plant and Cell Physiology* **50**: 947–955.

Yuan T, Li X, Xiao J, Wang S. 2011. Characterization of *xanthomonas oryzae*-responsive cis-acting element in the promoter of rice race-specific susceptibility gene Xa13. *Molecular Plant* **4**: 300–309.

Yu K, Park SJ, Poysa V. 2000. Marker-assisted selection of common beans for resistance to common bacterial blight: Efficacy and economics. *Plant Breeding* **119**: 411–415.

Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B. 2011. Colonization of rice leaf blades by an african strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene.

Molecular Plant-Microbe Interactions **24**: 1102–13.

Zahran HH. 1999. Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiology and molecular biology reviews* : *MMBR* **63**: 968–89.

Zaumeyer WJ. 1930. The bacterial blight of beans caused by *Bacterium phaseoli*. *Technical Bulletin US Department of Agriculture*: 1–35.

Zeng X, Tian D, Gu K, Zhou Z, Yang X, Luo Y, White FF, Yin Z. 2015. Genetic engineering of the Xa10 promoter for broad-spectrum and durable resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant Biotechnology Journal* **13**: 993–1001.

Zerbino DR, Birney E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**: 821–829.

Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature Biotechnology* **29**: 149–154.

Zhang Y, Jalan N, Zhou X, Goss E, Jones JB, Setubal JC, Deng X, Wang N. 2015a. Positive selection is the main driving force for evolution of citrus canker-causing *Xanthomonas*. *The ISME Journal* **9**: 2128–38.

Zhang J, Yin Z, White F. 2015b. TAL effectors and the executor R genes. *Frontiers in Plant Science* **6**: 641.

Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom JS, Huang S, Liu S, Vera Cruz C, Frommer WB, et al. 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant Journal* **82**: 632–643.

Zipfel C. 2008. Pattern-recognition receptors in plant innate immunity. *Current Opinion in Immunology* **20**: 10–16.

Zipfel C. 2009. Early molecular events in PAMP-triggered immunity. *Current Opinion in Plant Biology* **12**: 414–420.

Zou HS, Zhao WX, Zhang XF, Han YC, Zou LF, Chen GY. 2010. Identification of an avirulence gene, *avrxa5*, from the rice pathogen *Xanthomonas oryzae* pv. *oryzae*. *Science China Life Sciences* **53**: 1440–1449.

RÉSUMÉ

Déterminisme de la spécificité d'hôte et rôle des effecteurs TAL dans l'interaction *Xanthomonas* – haricot

La graisse commune est la principale phyto bactériose du haricot. Cette maladie est causée par *Xanthomonas citri* pv. *fuscans* (Xcf) et *X. phaseoli* pv. *phaseoli* (Xpp). Xcf et Xpp étant phylogénétiquement distantes, leur capacité à produire les mêmes symptômes sur haricot serait le fruit d'une convergence pathologique. L'objectif de cette thèse était d'identifier les gènes candidats pour la spécificité d'hôte et d'étudier le rôle des effecteurs Transcription Activator-Like (TAL) dans l'interaction *Xanthomonas* - haricot. Par une approche de génomique comparative, nous avons identifié 116 gènes spécifiques des agents de la graisse commune, dont un grand nombre a été transféré horizontalement entre Xcf et Xpp. Ces gènes codent des protéines intervenant aux différentes étapes de l'interaction. L'obtention du génome de 17 souches par séquençage Single-Molecule Real-Time a révélé l'existence de un à trois gènes codant des effecteurs TAL par souche, pour un total de quatre gènes *tal* différents dont deux (*tal23A* et *tal18H*) ont été également transférés horizontalement entre Xcf et Xpp. L'ensemble de ces gènes constitue un répertoire spécifique de Xcf et Xpp qui pourrait expliquer la convergence pathologique entre ces pathovars. Des tests de pouvoir pathogène couplés à des analyses transcriptomiques après inoculation d'un mutant de délétion de *tal18H* sur haricot ont révélé que TAL18H était impliqué dans l'aggravation des symptômes et avait un effet pléiotrope sur le transcriptome du haricot lors de l'interaction. Les résultats de cette thèse constituent une première étape vers l'optimisation de la lutte contre la graisse commune du haricot.

Mots-clés : graisse commune du haricot, *Phaseolus vulgaris*, transfert horizontal de gène, adaptation à l'hôte

ABSTRACT

Determinism of host specificity and role of TAL effectors in *Xanthomonas* - bean interaction

Common bacterial blight is the main bacterial disease of common bean. This disease is caused by *Xanthomonas citri* pv. *fuscans* (Xcf) and *X. phaseoli* pv. *phaseoli* (Xpp). Xcf and Xpp are phylogenetically distant yet they share the ability to induce the same symptoms on common bean, which is suggestive of pathological convergence between these two pathovars. This thesis aimed at identifying candidate genes for host specificity and studying the role of Transcription Activator-Like (TAL) effectors in the *Xanthomonas* – common bean interaction. Using a comparative genomic approach, we identified 116 genes specific to common bacterial blight agents, a large number of which were horizontally transferred between Xcf and Xpp. These genes encoded proteins involved in the different steps of the interaction. Single-Molecule Real-Time sequencing of 17 Xcf and Xpp genomes unveiled one to three TAL-encoding genes per strain for a total of four different *tal* genes, two of which (*tal23A* and *tal18H*) were also horizontally transferred between Xcf and Xpp. All these genes form a repertoire specific to Xcf and Xpp that could be responsible for the pathological convergence observed between these two pathovars. Combination of pathogenicity tests and transcriptomics after inoculation of a *tal18H* deletion mutant on common bean plants revealed that TAL18H was involved in symptom development and displayed a pleiotropic effect on common bean transcriptome during the interaction. The results of this thesis constitute a stepping stone towards optimizing the monitoring of common bacterial blight of bean.

Keywords: common bacterial blight of bean, *Phaseolus vulgaris*, horizontal gene transfer, host adaptation



AGROCAMPUS OUEST • Institut supérieur des sciences agronomiques, agroalimentaires, horticoles et du paysage
65 rue de Saint-Brieuc – CS84215 – F-35042 Rennes Cedex
Tél. : 02 23 48 50 00
www.agrocampus-ouest.fr

