



HAL
open science

**New insights into the diversity and function of
Transcription Activator-Like effectors of *Xanthomonas*
phaseoli pv. *manihotis*, the causal agent of Cassava
Bacterial Blight**

Carlos Andrés Zarate Chaves

► **To cite this version:**

Carlos Andrés Zarate Chaves. New insights into the diversity and function of Transcription Activator-Like effectors of *Xanthomonas phaseoli* pv. *manihotis*, the causal agent of Cassava Bacterial Blight. *Phytopathology and phytopharmacy*. Université Montpellier, 2021. English. NNT : 2021MONTG034 . tel-03666640

HAL Id: tel-03666640

<https://theses.hal.science/tel-03666640>

Submitted on 12 May 2022

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Mécanismes des Interactions parasites pathogènes et symbiotiques

École doctorale GAIA

Unité de recherche PHIM – Plant Health Institute of Montpellier

Titre de la thèse

New insights into the diversity and function of Transcription
Activator-Like effectors of *Xanthomonas phaseoli* pv.
manihotis, the causal agent of Cassava Bacterial Blight

Présentée par Carlos Andrés ZÁRATE CHAVES
Le 16 juillet 2021

Sous la direction de Boris SZUREK
et Adriana BERNAL

Devant le jury composé de

Allia DELLAGI, Professeure, AgroParisTech

Laurent NOEL, Directeur de Recherche, CNRS

Barbara DE CONINCK, Professeure associée, Université Catholique de Louvain

Ricardo OLIVA, Chargé de Recherche, IRR

Adriana BERNAL, Professeure associée, Université de los Andes

Boris SZUREK, Directeur de Recherche, IRD

Rapportrice

Rapporteur

Examinatrice

Examineur

Co-directrice de thèse

Directeur de thèse



UNIVERSITÉ
DE MONTPELLIER

Acknowledgments

I thank my tutors Adriana, Boris, and Camilo, who guided me through this journey. They created a friendly environment where ideas, hypothesis, results, and projections were deeply discussed, and taught me the best way to construct knowledge. I have only gratitude and admiration for them. I am also thankful for all the feedback received from the members of my thesis committee Silvia Restrepo, Adam Bogdanove, Thomas Kroj, Pierre Zcernic, and Gilles Béna; their ideas and support were highly valuable for this work. Likewise, scientific discussions and collaborations with Ralf Koebnik, Sébastien Cunnac, Emmanuel Wicker, Lionel Gagnevin, Mathilde Hutin, and Florence Auguy from the XPLAIN team were always prolific and led to new ways to see things. I am also grateful to my team colleagues Coline, Marlène, Dario, Carolina, Leidy, and Edilene, who were always there to give me a hand, to discuss ideas, and to support me in several ways. I would like to thank my former master's colleagues, César, Juan Luis, Ruben, Diana, Vivian, Natalia, and Leidy, because they were an active part of my doctoral formation through scientific discussions that led to ideas on how to resolve some of the problems that I faced during the PhD.

I thank my friends for supporting me during the tough moments and finding ways to cheer me up, they were an important part of my life during this cycle. Elvira, Natalia, Lorinda, Helena, Adrian, Lucille, Teerarat, Kader, Juliette, Matthieu, Anne Sophie, you made my life easier, thank you. Special thanks to Coline and Elvira, who were there during the most difficult moments, and helped me to carry on, to learn from experiences, and discovering new ways to see life.

El primer agradecimiento personal es para mis padres, quienes dieron lo mejor de sí para ponerme en este lugar. Todos sus esfuerzos, regaños, amor, cariño y consejos dieron forma a la persona que hoy termina esta etapa. A mi hermana agradezco su apoyo incondicional, su interés por lo que hago y su amor hacia mí. A mi familia, que siempre ha creído en mis capacidades y me ha apoyado y alentado a ser mejor persona y profesional.

Este trabajo es tan mío como de Lina Gómez, quien invirtió un gran esfuerzo, amor y tiempo en esta empresa. Agradezco infinitamente haber contado con su apoyo durante este ciclo, porque compartimos y trabajamos por muchos sueños, incluido este proyecto. Agradezco cada una de sus palabras de aliento, sus emotivas reacciones que siempre me llevaron a sentirme en el lugar correcto, su apoyo incondicional para crear y mantener el mejor ambiente, y su forma única de soñar, creer en los sueños y trabajar por ellos. El resultado no habría sido el mismo sin su presencia en mi vida, porque de ella aprendí miles de cosas que facilitaron mi crecimiento personal y profesional. Gracias por ser ese motor que nunca se apagó. También agradezco a la familia Gómez Molano por su bonita forma de hacerme sentir el soporte y cariño de una gran familia. Blanca, Gustavo, Luis Eduardo, Lina, Johana y Camilo, ustedes fueron a menudo el refugio y la fuente de alegría y coraje que se necesitó para sobrellevar las distancias. Solo tengo cariño y agradecimiento para ustedes.

Resumé

Le manioc (*Manihot esculenta* Crantz) nourrit environ 800 millions de personnes et soutient l'économie des agriculteurs à faible revenu sous les tropiques. Les stress biotiques notamment causés par des virus et bactéries sont responsables d'importantes pertes économiques. *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) est l'agent causal de la bactériose vasculaire du manioc (CBB) qui est une des bactérioses les plus dévastatrices de cette culture. La CBB est une maladie vasculaire et systémique pouvant entraîner de fortes pertes de rendement. *Xpm* dépend notamment de ses effecteurs de type III pour contourner l'immunité des plantes et coloniser les tissus de l'hôte. Ces effecteurs comprennent les *Xanthomonas outer proteins* et les Transcription Activator-Like effectors (TALE) qui agissent comme de véritables facteurs de transcription eucaryotes. Une fois injectés dans la cellule cible, les TALE migrent vers le noyau de l'hôte et ciblent des motifs d'ADN spécifiques (*effector binding element*, EBE) via des motifs répétés situés dans la région centrale de la protéine. Une fois fixé sur le promoteur du gène cible, les TALE interagissent avec les facteurs de transcription de l'hôte via sa région C-terminale, recrutant ainsi l'ARN polymérase de type II pour initier la transcription du gène. Les TALE jouent un rôle clé dans la pathogenèse de *Xpm*, notamment via l'activation transcriptionnelle de gènes de susceptibilité (*S*) tels que le gène *MeSWEET10a* qui code pour un transporteur de sucrose. S'il est connu que TAL20_{xam668} et TAL14_{xam668} et leurs homologues sont des facteurs de virulence majeurs de *Xpm* car favorisent le développement de lésions aqueuses et la croissance bactérienne *in planta*, on connaît mal la diversité des TALE parmi les populations de *Xpm*, la distribution des TALE de virulence majeurs, et les autres déterminants de la maladie. Dans cette thèse, nous avons traité ces questions selon trois chapitres : i) diversité et fonction des TALEs, ii) gènes *S* alternatifs, et iii) développement d'un outil de type CRISPRi pour éteindre l'expression des TALE chez *Xanthomonas*. L'étude de la diversité et de la fonction des TALEs a révélé 13 nouveaux variants chez *Xpm*, dont un avec 22 répétitions qui active *MeSWEET10a*. De plus, nous avons identifié deux cas potentiels de convergence fonctionnelle. La pertinence de ces résultats réside dans la description de nouveaux variants de TALE majeurs et de « hubs » de sensibilité potentiels. L'étude des gènes *S* alternatifs a conduit à l'identification de *MeSWEET10e* qui code un nouveau transporteur de sucre SWEET manipulé par *Xpm*. *MeSWEET10e* semble toutefois induit indirectement par des variants de TALE15 mais le mécanisme reste à découvrir. Ces résultats sont utiles pour la génération de variétés de manioc durablement résistantes à *Xpm*. Enfin, ayant rencontré des difficultés pour muter les TALE dans certaines souches de *Xpm*, nous avons mis en place une plateforme d'extinction de l'expression génique basée sur la technologie CRISPRi, qui pourrait être extrapolée à d'autres familles de gènes et à d'autres espèces de *Xanthomonas*. Trois ARN guides (sgRNA) ont été conçus pour éteindre les TALE de *Xpm*, et nous avons montré que la plupart des TALE d'une souche sont bien éteints par un seul sgRNA, provoquant une diminution des symptômes et des titres bactériens *in planta*. En outre, des TALE protégés contre les sgRNAs sont bien exprimés dans les souches exprimant les sgRNA, ce qui permet des tests de complémentation. L'utilisation de cet outil chez *X. cassavae*, un pathogène non vasculaire du manioc, montre une réduction forte des symptômes révélant l'implication des TALEs dans la virulence de ce pathogène orphelin. Notre étude fournit une plateforme CRISPRi utile pour évaluer le rôle des TALEs lors de la colonisation de l'hôte, et globalement, pour l'analyse fonctionnelle de familles multigéniques chez cette espèce.

Abstract

Cassava (*Manihot esculenta* Crantz) is a staple crop that feeds about 800 million people and supports the economy of low-income farmers in tropics. Due to its outstanding performance, cassava grows under a wide range of temperatures, soil pH, salinity degrees, and water availability, which is highly desirable for climate change adaptation. Biotic stresses affecting cassava include pests and pathogens, being viral and bacterial pathogens the most economically relevant. The cassava mosaic disease and the cassava brown streak disease are the two main viral cassava diseases, which cause important yield losses. Among bacterial pathogens, *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), causal agent of the Cassava Bacterial Blight (CBB), is the most relevant one. CBB is a vascular disease characterized by water-soaked angular leafspots, leaf blight, wilting, stem rotting, and dieback. *Xpm* relies on its effectors to subvert plant immunity and establish successful infections. These effectors include *Xanthomonas* outer proteins and Transcription Activator-Like effectors (TALEs). TALEs are bacterial effectors that, once translocated to the host cell, migrate to the nucleus and search for specific DNA motifs (effector binding element - EBE) by means of their repetitive central region. Once located onto the target EBE, the C-terminal region interacts with host transcription factors and recruits the RNA polymerase to initiate transcription of downstream genes. TALEs play key roles in *Xpm* pathogenesis, notably through the transcriptional activation of susceptibility (*S*) genes like the sugar transporter *MeSWEET10a*. Although it is known that TAL20_{Xam668} and TAL14_{Xam668} and their homologs are key factors that enhance water-soaking and *in-planta* bacterial growth (known as major virulence TALEs), little is known about TALE diversity among *Xpm* populations, the distribution of the major virulence TALEs, and other mechanisms used by *Xpm* to promote disease. In this thesis, we addressed these matters in three chapters: i) diversity and function of TALEs, ii) search for alternative *S* genes, and iii) development of a CRISPRi-based tool to study TALE biology. The study of TALE diversity and function revealed 13 novel variants in *Xpm*, including a variant with 22 repeats that activates the *S* gene *MeSWEET10a*. Moreover, we identified two potential cases of functional convergence. The relevance of these results lies in the description of new variants of major virulence TALEs and potential TALE-targeted hubs. The study of alternative *S* genes led to the identification of another SWEET sugar transporter, *MeSWEET10e*, that is transcriptionally activated by *Xpm*, potentially as a secondary target of TALE15 variants. Phylogenies indicate that *Xpm* TALEs and their linked capabilities to activate *MeSWEETs* come from at least two different ancestor genes (convergent evolution). These findings will be useful in the conception of *Xpm*-resistant cassava varieties. Finally, as we faced difficulties to mutate *TALEs* in some *Xpm* strains, we set a silencing platform based on CRISPR interference (CRISPRi), which could be extrapolated to other gene families and other *Xanthomonas* species. Three single guide RNAs (sgRNAs) were designed to silence *Xpm* *TALEs*, and our data show that most *TALEs* in a strain can be silenced by one individual sgRNA. CRISPRi silencing resulted in decreased symptoms and lower *in-planta* bacterial titers. Furthermore, plasmid-borne *TALEs* protected against sgRNAs were expressed in the strains undergoing silencing and allowed complementation. Implementation of this platform in *Xanthomonas cassavae*, a non-vascular pathogen of cassava, demonstrated for the first time the role of TALEs in this orphan pathosystem and suggested the role of *MeSWEET10a* as a target of *Xc*. Our study provides a useful CRISPRi platform to assess the role of *TALEs* during host colonization, and overall, for the functional analysis of multigene families in this species.

Abbreviations

<i>Abbreviation</i>	<i>Definition</i>
<i>ABA</i>	Abscisic acid
<i>ABI3</i>	Abscisic acid insensitive3
<i>AD</i>	Activation domain
<i>AFLP</i>	Amplified fragment length polymorphism
<i>ATG</i>	Autophagy-Related Protein 8
<i>AUDPC</i>	Area under the disease progress curve
<i>BAC</i>	Bacterial artificial chromosome
<i>bHLH</i>	Basic helix-loop-helix
<i>bZIP</i>	Basic leucine zipper domain
<i>Cas9</i>	CRISPR-associated protein 9
<i>CBB</i>	Cassava bacterial blight
<i>CBL</i>	Calcineurin B-like proteins
<i>CBN</i>	Cassava bacterial necrosis
<i>CBSD</i>	Cassava brown streak disease
<i>CBSV</i>	Cassava brown streak virus
<i>CFSD</i>	Cassava frogskin disease
<i>CIPK</i>	CBL-Interacting Protein Kinases
<i>CMD</i>	Cassava Mosaic Disease
<i>CRISPR</i>	Clustered regularly interspaced short palindromic repeats
<i>DAMP</i>	Damage-associated molecular pattern
<i>DHR</i>	Double homologous recombination
<i>DNA</i>	Deoxyribonucleic acid
<i>EACMV</i>	East African cassava mosaic virus
<i>EBE</i>	Effector binding element
<i>ELISA</i>	Enzyme-linked immunosorbent assay
<i>ERF</i>	Ethylene response factors
<i>EST</i>	Expressed sequenced tags
<i>ETI</i>	Effector-triggered immunity
<i>ETS</i>	Effector-triggered susceptibility
<i>F1</i>	Filial 1 hybrid
<i>GUS</i>	β -glucuronidase
<i>HR</i>	Hypersensitive response
<i>Hsf</i>	Heat Stress TFs
<i>ID</i>	Identifier
<i>IRD</i>	Institut de Recherche pour le Développement
<i>LOB</i>	Lateral organ boundaries
<i>LRR</i>	Leucine-rich repeat

<i>MAP</i>	Months after planting
<i>MAS</i>	Marker-assisted selection
<i>MBP</i>	Melatonin biosynthesis pathway
<i>MDR</i>	Mean disease rating
<i>MLVA</i>	Multilocus variable-number tandem repeat (VNTR) analysis
<i>NAC</i>	NAM/ATAF/CUC2
<i>NADPH</i>	Nicotinamide adenine dinucleotide phosphate (reduced form)
<i>NBS</i>	Nucleotide binding site
<i>NBS-LRRs</i>	Nucleotide binding site-Leucine-rich repeat
<i>NF-Y</i>	Nuclear Factor Y
<i>NGS</i>	Next generation sequencing
<i>NHEJ</i>	Non-homologous end joining
<i>NLR</i>	Nucleotide-binding LRR
<i>NLS</i>	Nuclear localization signal
<i>NPR1</i>	Natriuretic Peptide Receptor 1
<i>OD</i>	Optical density
<i>PAL</i>	Phenylalanine ammonia lyase
<i>PAM</i>	Protospacer adjacent motif
<i>PAMP</i>	Pathogen-associated molecular pattern
<i>PCR</i>	Polymerase chain reaction
<i>PL</i>	Pectate lyase
<i>PR</i>	Pathogenesis-related
<i>PTI</i>	PAMP-triggered immunity
<i>QTL</i>	Quantitative trait loci
<i>RAPD</i>	Random amplified polymorphic DNA
<i>RAV</i>	Related to ABI3/VP1
<i>RBS</i>	Ribosome binding site
<i>Rep-PCR</i>	Repetitive element palindromic PCR
<i>RFLP</i>	Restriction fragment length polymorphism
<i>RLK</i>	Receptor-like kinases
<i>RLP</i>	Receptor-like proteins
<i>RNA</i>	Ribonucleic acid
<i>ROS</i>	Reactive oxygen species
<i>RT</i>	Retrotranscriptase
<i>RVD</i>	Repeat Variable Di-residue
<i>SA</i>	Salicylic acid
<i>SLCMV</i>	Sri Lankan cassava mosaic virus
<i>SMRT</i>	Single molecule, Real time
<i>SNP</i>	Single nucleotide polymorphism
<i>SSR</i>	Simple sequence repeats

SWEET	Sugars Will Eventually be Exported Transporters
TO	Primary transformants
T1SS	Type-I secretion system
T2SS	Type-II secretion system
T3E	Type-III effectors
T3SS	Type-III secretion system
T4SS	Type-IV secretion system
T6SS	Type-VI secretion system
TALE	Transcription Activator-Like effector
TALEN	Transcription Activator-Like effector Nuclease
TCP	TB1/CIN/PCF
TDF	Transcript-derived fragments
TF	Transcription factor
TFB	Transcription factor binding
TIR	Toll/interleukin-1 receptor
UCBSV	Ugandan cassava brown streak virus
USA	United States of America
UTR	Untranslated region
VNTR	Variable number of tandem repeats
VP1	Viviparous1
VPg	Viral genome-linked protein
WHY	Whirly TF
Xam	<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i> (f.k.a., <i>Xanthomonas axonopodis</i> pv. <i>manihotis</i>)
Xau	<i>Xanthomonas citri</i> pv. <i>aurantifolii</i>
Xc	<i>Xanthomonas cassavae</i>
Xca	<i>Xanthomonas campestris</i> pv. <i>campestris</i>
Xcc	<i>Xanthomonas citri</i> pv. <i>citri</i>
Xcf	<i>Xanthomonas citri</i> pv. <i>fuscans</i>
Xcm	<i>Xanthomonas citri</i> pv. <i>malvacearum</i>
Xev	<i>Xanthomonas euvesicatoria</i>
Xg	<i>Xanthomonas gardneri</i>
Xoc	<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>
Xoo	<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>
Xpm	<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i> (f.k.a., <i>Xanthomonas axonopodis</i> pv. <i>manihotis</i>)
Xpp	<i>Xanthomonas phaseoli</i> pv. <i>phaseoli</i>
Xtt	<i>Xanthomonas translucens</i> pv. <i>translucens</i>
Xtu	<i>Xanthomonas translucens</i> pv. <i>undulosa</i>
YPG	Yeast peptone glucose medium

Résumé détaillé des activités réalisées au cours de la thèse

Introduction

Le manioc (*Manihot esculenta* Crantz) est une plante pérenne ligneuse de la famille des Euphorbiacées, caractérisée par de grosses racines tubéreuses amylicées appréciées pour leurs valeurs nutritionnelles et énergétiques. Grâce à un certain nombre de mécanismes qui optimisent son utilisation de l'eau, cette plante est hautement tolérante à la sécheresse ce qui la rend particulièrement adaptée à des sols où d'autres cultures échoueraient. Sa grande tolérance à différents stress abiotiques comme notamment la température, le pH et la salinité, font du manioc une culture qui pourrait jouer un rôle important dans un contexte de changement climatique. Le manioc est largement cultivé comme culture vivrière de base dans les régions subtropicales d'Afrique, d'Asie et d'Amérique latine, contribuant à la sécurité alimentaire et économique de ses pays. Les plants de manioc se propagent à partir de boutures de tiges, ce qui permet d'obtenir un matériel de plantation peu coûteux et facilement disponible. Cette culture peut produire jusqu'à 250 kcal/ha/jour, soit plus du double du rendement calorique du blé et du sorgho, et bien plus que celui du maïs, de la patate douce et du riz. La culture du manioc nourrit environ 800 millions de personnes, dont un peu plus de la moitié reçoit plus de 100 calories par jour sous les tropiques. La majorité de cette culture est exploitée par des agriculteurs à faible revenu sur de petites exploitations. On estime que plus de 70% de la production est consacrée à l'alimentation humaine, tandis que les 30% restants sont utilisés pour l'alimentation du bétail et l'industrie de l'amidon et des biocarburants.

Le manioc est menacé par plusieurs ravageurs comme les aleurodes, les cochenilles et les acariens, ainsi que par des agents pathogènes viraux, bactériens et fongiques. Bien que les ravageurs puissent causer d'importantes pertes de rendement, certaines pratiques culturales et des agents de contrôle biologique peuvent atténuer leur impact. La maladie de la mosaïque du manioc (CMD) et la maladie de la striure brune du manioc (CBSD) sont les deux principales maladies virales du manioc dans l'hémisphère Est, causant d'importantes pertes de rendement. Les pathogènes fongiques comme *Cercospora* spp et *Colletrotrichum* spp, bien qu'étant capables d'infecter le manioc, sont considérés comme ayant des impacts mineurs sur la culture, affectant marginalement le rendement des racines et la qualité des boutures, généralement lors d'infections mixtes avec d'autres pathogènes. Parmi les pathogènes bactériens, *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), l'agent causal de la bactériose du manioc (CBB), est considéré comme la menace d'origine bactérienne la plus importante pour cette culture.

Le CBB se caractérise par une série de symptômes qui affectent principalement les feuilles, les pétioles et les tiges, entraînant fréquemment la mort de la plante. *Xpm* pénètre dans les feuilles par les stomates et les blessures, et colonise initialement l'apoplaste du mésophylle. À mesure que la maladie progresse, les bactéries accèdent aux vaisseaux du xylème à partir du mésophylle et se déplacent vers la tige par les pétioles. La colonisation des vaisseaux dans la tige permet à *Xpm* de se déplacer de manière systémique vers le haut et vers le bas. Lorsque l'infection atteint la partie

supérieure de la plante, où les tissus de la tige sont plus verts et moins lignifiés, la pourriture de la tige entraîne un dépérissement caractérisé par le flétrissement de l'apex de la pousse. De 1970 à 1975, en République démocratique du Congo où la diversité des cultures de manioc était à l'époque faible et où la dépendance alimentaire et économique vis-à-vis du manioc était considérable, la propagation du CBB a été à l'origine d'une importante période de famine. Les pertes de rendement peuvent atteindre les 100%, avec une médiane d'environ 50% dans les variétés sensibles. A ce jour, *Xpm* a été signalé dans 49 pays situés dans toutes les régions tropicales où le manioc est cultivé, ce qui souligne l'endémicité du pathogène et de la maladie.

Xpm est une gamma protéobactérie à Gram négatif appartenant à la famille des Xanthomonadaceae. Comme les autres espèces de Xanthomonas, elle possède de 13 à 23 effecteurs Xop, avec un noyau de 9 effecteurs conservés : HpaA, HrpF, XopE1, XopV, Hpa2, XopAK, XopL, XopN et XopAE (alias HpaF). Il a été démontré que ces effecteurs jouent un rôle clé dans l'Effector Triggered Susceptibility (ETS). D'autres effecteurs de type III, les effecteurs TALE (Transcription Activator-Like effectors) sont invariablement présents dans les souches de *Xpm*. Un criblage de 65 souches par RFLP a montré que toutes contenaient entre un et cinq TALEs et que la plupart étaient portés par des plasmides. Il a été démontré que les TALEs sont importants pour la pathogénie de *Xpm*. Les TALEs ont une structure particulière qui a évolué pour se lier spécifiquement au promoteur de gènes de l'hôte pour y recruter l'ARN polymérase de type II et initier la transcription. La structure de la protéine peut être divisée en trois parties : la région N-terminale, la région centrale répétée et la région C-terminale. La région N-terminale contient le signal de translocation vers l'hôte. La région C-terminale quant à elle comprend des signaux de localisation nucléaire permettant l'importation du TALE dans le noyau de la cellule hôte, une région de liaison à des facteurs généraux de la transcription, et un domaine d'activation de la transcription de type eucaryote. Enfin, la région centrale contient de une à 33 répétitions modulaires qui codent généralement pour 33 à 35 acides aminés chacune. La séquence en acides aminés des répétitions varie principalement au niveau des résidus 12 et 13 de chaque répétition, désignés comme " Repeat Variable Diresidue " (RVD). Chaque répétition forme un faisceau de deux hélices gauches, les résidus hypervariables étant situés à l'extrémité de la boucle qui relie les deux hélices, ce qui permet l'exposition du RVD aux bases de l'ADN. L'ensemble de ces répétitions modulaires se replie en une superhélice droite qui se loge le long du sillon majeur de l'ADN et permet l'interaction directe entre les RVD et les nucléotides d'une manière spécifique. Alors, la séquence RVD régit la reconnaissance de la séquence d'ADN. Les séquences d'ADN cibles des TALE, appelées Effector Binding Element (EBE), sont généralement situées dans la région du promoteur du gène ciblé, près de la boîte TATA. Après liaison à l'EBE, les domaines de la région C-terminale interagissent avec des facteurs de transcription hôtes, comme les sous-unités TFIIA γ et l'ARN polymérase pour induire la transcription du gène en aval.

Chez *Xpm*, il a été découvert que deux TALEs jouent un rôle important dans la pathogénie, via l'activation transcriptionnelle de gènes de susceptibilité. TAL20_{Xam668} (ou ses homologues) est nécessaire à l'apparition des symptômes des taches aqueuses sur les feuilles de manioc : la mutation de TAL20_{Xam668} en abolit le développement et entraîne une diminution des titres bactériens dans la nervure médiane et l'apoplasme de la feuille. Le gène *MeSWEET10a* a été identifié comme seule cible de TAL20_{Xam668}, dont l'activation transcriptionnelle est médiée par la reconnaissance directe de l'EBE TATAAACGCTTCTCGCCATC dans la région promotrice du gène. *MeSWEET10a* (*Manes.04G123400*) code pour un membre du clade III de la famille de transporteurs de sucres

SWEET (Sugar Will be Eventually Exported Transporter), exportant à la fois le glucose et le saccharose. De même, il a été démontré que les homologues de TALE14 sont essentiels à la multiplication *in planta* de *Xpm* et la colonisation des tissus cibles. L'absence/mutation de ce facteur de virulence affecte en effet la croissance de l'agent pathogène mais pas l'apparition de taches aqueuses sur les feuilles. A ce jour, une liste de dix gènes de sensibilité candidats ciblés par TAL14 potentiellement à l'origine de ce phénotype a été établie.

Justification

L'endémicité du CBB dans les pays tropicaux et le fait qu'il s'agisse principalement d'une culture vivrière cultivée à petite échelle rendent d'autant plus essentielle la compréhension des facteurs clés de la pathogénie et de la résistance dans le but d'appliquer ces connaissances à la sélection végétale. Les déterminants moléculaires de la résistance à la CBB chez le manioc ne sont pas bien compris, mais nous savons que cette résistance est polygénique et additive. La résistance de la plante procède de plusieurs mécanismes moléculaires et cellulaires spécifiques aux tissus qui sont activés en série relativement au mouvement du pathogène. Nos connaissances sur les déterminants de la sensibilité au pathogène sont encore plus rares, et à ce jour nous ne connaissons qu'un seul de ces facteurs : la protéine de transport de saccharose MeSWEET10a.

La diversité génétique du pathogène a été largement évaluée par différentes méthodes moléculaires, et plusieurs populations d'Amérique du Sud et d'Afrique ont été caractérisées. Les questions de recherche abordées par ces études nécessitant des méthodes visant à évaluer principalement la diversité génétique neutre, nous connaissons moins bien la diversité des gènes soumis à pression de sélection positive ou négative, tels que les gènes effecteurs par exemple. Quelques études ont montré que les TALEs jouent un rôle clé dans la pathogénie : TALE14 et TALE20 contribuent de manière significative à la fitness de *Xpm in planta*. Cependant, la diversité et la prévalence de ces effecteurs ne sont pas bien caractérisées dans les populations de *Xpm*. D'un point de vue fonctionnel, nos connaissances sur les TALE de *Xpm* peuvent se résumer comme suit: deux homologues de TALE14 et une version de TALE20 sont des facteurs de virulence clés ; la sensibilité médiée par TALE20 est liée à l'activation transcriptionnelle de *MeSWEET10a*.

Il a déjà été démontré que l'empêchement du fonctionnement d'un TALE était une excellente stratégie de résistance par perte de sensibilité des plantes. La compréhension non seulement des mécanismes sous-jacents, mais aussi de la diversité de l'agent pathogène et de ses stratégies de virulence constituent la clé de voute d'un déploiement de résistance(s) réussi. Cependant, dans le cas de *Xpm*, certaines questions restent ouvertes. Par exemple : quelle est la distribution de ces TALEs dits de virulence majeure dans les populations de *Xpm* ? Les mécanismes moléculaires décrits sont-ils les seuls déterminants clés de la sensibilité médiée par les TALEs de ce pathogène ? Ces questions structurent la problématique centrale que j'ai tenté de résoudre dans le cadre de ma thèse.

Cette étude vise à explorer la biologie des TALEs dans le cadre de l'interaction *Xpm*-manioc. Un premier chapitre est consacré à la diversité des TALEs et à leurs fonctions potentielles dans une population bien caractérisée de *Xpm*, où de nouveaux variants de TALEs avec des fonctions de virulence clés ont été découverts. Le deuxième chapitre traite de la découverte d'un nouveau gène de sensibilité activé par un TALE. Ce gène appartient au clade III des transporteurs de sucre de la famille SWEET et est ciblé comme une alternative au gène *MeSWEET10a*. Le troisième chapitre

décrit une nouvelle plateforme dédiée à l'extinction de gènes de *Xanthomonas* basée sur la technologie d'interférence CRISPR. L'extinction de plusieurs membres de la famille TALE a été réalisée avec succès. La plateforme a été établie chez *Xpm*, et a permis de confirmer l'importance des TALEs dans la pathogénie de *Xanthomonas*. Nous avons utilisé les TALEs comme proxy pour l'étude des gènes de la famille SWEET. En appliquant cet outil au pathogène non vasculaire du manioc *Xanthomonas cassavae*, nous avons découvert que *MeSWEET10a* était potentiellement induit par ce pathovar et pourrait donc agir comme gène de sensibilité dans ce pathosystème peu étudié.

Chapitre 1 : Diversité et fonction des TALEs dans une population colombienne de *Xpm*

Les informations disponibles sur les TALEs de *Xpm* et leurs gènes cibles dans le manioc sont rares, mais elles ont augmenté au cours de ces dernières années. Nous avons cherché à caractériser la diversité des TALEs dans les souches colombiennes de *Xpm* et à identifier les gènes candidats ciblés par ces derniers. Nous avons sélectionné dix-huit souches de *Xpm* à l'échelle du pays sur la base de la diversité génétique neutre afin d'analyser la diversité des TALEs. L'analyse RFLP a montré que les souches de *Xpm* portent des TALomes avec une distribution de taille bimodale, et le regroupement par affinité des TALEs séquencés a condensé cette variabilité en cinq groupes principaux. Nous rapportons l'identification de 13 nouveaux variants de TALEs dans *Xpm*, dont un variant de 22 répétitions qui active le gène de sensibilité *MeSWEET10a*, une cible précédemment signalée de TAL20_{Xam668}. Des analyses transcriptomiques et de prédictions bioinformatiques d'EBE ont permis de sélectionner plusieurs gènes candidats potentiellement ciblés par les TALEs, ainsi que deux cas probables de convergence fonctionnelle. Cette étude ouvre la voie à la caractérisation de nouvelles cibles potentielles des TALEs dans l'interaction *Xpm*-manioc, qui pourraient avoir un rôle déterminant dans l'infection.

Chapitre 2 : *MeSWEET10e*, un gène alternatif de sensibilité du manioc activé transcriptionnellement par les TALEs de *Xpm*

Accroître l'exportation de saccharose via l'activation transcriptionnelle de certains membres de la famille SWEET est une stratégie bien connue des *Xanthomonas* pour améliorer leur virulence et leur fitness *in planta*. Les TALEs jouent un rôle clé dans cette régulation positive des SWEETs chez le riz, le coton et le manioc. *MeSWEET10a* est un gène S transcriptionnellement activé par TAL20_{Xam668} et ses homologues. La formation de symptômes et la croissance bactérienne *in planta* dépendent fortement de cette induction médiée par ces TALEs. Dans cette étude, nous décrivons *MeSWEET10e*, un gène SWEET du clade-III comme un nouveau facteur de sensibilité du manioc à *Xpm*, potentiellement activé par plusieurs variants de TALE15. L'analyse RNAseq et de prédiction des EBE *in silico* suggèrent que l'activation de *MeSWEET10e* est indirecte. L'activation transcriptionnelle directe de *MeSWEET10e* via des TALEs artificiels démontre que la fonction physiologique de ce gène est équivalente au rôle de *MeSWEET10a* dans la pathogénèse de *Xpm*. Les analyses phylogénétiques indiquent que les effecteurs TAL de *Xpm* et leur capacité à activer les *MeSWEETs* proviennent d'au moins deux gènes ancêtres différents, ce qui constitue un cas d'évolution convergente par l'exploitation de paralogues fonctionnellement redondants. Ces résultats sont cruciaux pour la conception de variétés de manioc améliorées avec une résistance accrue à *Xpm*.

Chapitre 3 : CRISPRi, une stratégie innovante pour l'étude fonctionnelle de familles de gènes chez les bactéries phytopathogènes

La technologie CRISPR a amélioré l'édition du génome en augmentant considérablement son efficacité et sa spécificité. Cependant, les bactéries ne sont pas facilement mutées à l'aide de cette technologie. Une variante technologique appelée CRISPRi utilise une version modifiée de la protéine Cas9 qui est incapable de couper l'ADN (dead Cas9 - dCas9), mais qui conserve une activité de liaison spécifique à la séquence ADN. La fixation de dCas9 dans le promoteur ou les premiers nucléotides du brin non transcrit de la séquence codante interfère avec la transcription, ce qui entraîne l'extinction des gènes ciblés. Cet outil peut être extrêmement utile pour l'analyse fonctionnelle de familles multigéniques puisqu'un seul ARN guide (sgRNA) peut réduire au silence plusieurs membres à la fois. Dans cette étude, nous avons mis en place une plateforme CRISPRi pour *Xanthomonas* et testé son potentiel pour étudier une famille de gènes déterminante dans la pathogenèse de bactéries de ce genre : les TALEs. L'interférence CRISPR a été mise en œuvre via la conception de sgRNAs ciblant les éléments promoteurs -10 et -35, la séquence de liaison au ribosome et une région située dans les 50 premiers nucléotides de la séquence codante des TALEs de plusieurs espèces de *Xanthomonas*. En utilisant *Xpm* comme étude de cas, nous avons montré que l'expression de la plupart, sinon la totalité des TALEs d'une souche est éteinte par un seul sgRNA, ce qui entraîne une diminution des symptômes et des titres bactériens *in planta*. En outre, nous avons pu compléter les souches « silencées » via l'introgression de plasmides contenant des TALEs dépourvus de la séquence ciblée par le sgRNA. La mise en œuvre de cette ressource technologique chez *Xanthomonas cassavae*, un pathogène du manioc, suggère fortement que les symptômes de tâches aqueuses typiques de cette maladie sont liés à l'action des TALE. Des études de prédiction de cibles suggèrent que *MeSWEET10a* qui code pour un transporteur de sucre puisse agir comme un gène de susceptibilité dans cet autre pathosystème. Notre étude fournit une plateforme CRISPRi utile pour évaluer le rôle des TALEs pendant la colonisation de l'hôte, et plus généralement, pour l'analyse fonctionnelle de familles multigéniques chez *Xanthomonas*.

Discussion et conclusions

Dans cette thèse, trois axes principaux ont été étudiés : La diversité des TALEs parmi les populations de *Xpm*, les mécanismes de sensibilité alternatifs utilisés par *Xpm*, et le développement d'une nouvelle plateforme de « silencing » pour l'étude de familles de gènes, dont les TALEs, chez *Xanthomonas*. Grâce à l'évaluation de la diversité des TALEs parmi les isolats colombiens de *Xpm*, nous avons caractérisé 13 nouveaux variants TALE, dont un variant TALE22 capable d'activer le gène de sensibilité *MeSWEET10a*. Cette étude a également révélé que : après TALE20 et TALE14, les variants TALE22 et TALE13 sont largement distribués parmi les isolats, ce qui suggère que ces variants pourraient avoir des rôles importants dans la pathogenèse. La caractérisation du TALome a permis de découvrir des souches de *Xpm* très virulentes mais qui ne possèdent pas TALE20 ou ses variants connues pour activer transcriptionnellement *MeSWEET10a*. Les analyses phylogénétiques indiquent que les variants TALE provenant de ces souches sont considérablement différents des TALEs présents dans des isolats plus récemment obtenus dans le monde. Par conséquent, les données suggèrent que les TALEs de *Xpm* proviennent d'au moins deux gènes ancêtres différents,

et que les stratégies de promotion de la virulence associées à leur fonction biologique sont apparues indépendamment.

Après avoir découvert ce groupe de souches de *Xpm* qui ne dépendent pas de l'activation de *MeSWEET10a* pour provoquer la maladie, les analyses transcriptomiques ont montré que ces isolats sont capables de réguler positivement la transcription d'un deuxième transporteur de sucre du clade III de la famille *SWEET* : *MeSWEET10e*. Les données suggèrent que cette activation est médiée par les TALEs de manière indirecte, potentiellement par l'activation transcriptionnelle directe d'un régulateur principal. Bien que le mécanisme ne soit pas bien compris, nous avons démontré que *MeSWEET10e* peut agir comme un gène de sensibilité et que son rôle physiologique est équivalent au rôle joué par *MeSWEET10a*. Un pipeline *in silico* proposé en conjonction avec des validations *in planta* pourrait potentiellement identifier les éléments manquants qui conduisent à l'activation transcriptionnelle de ce gène de sensibilité.

Les difficultés rencontrées au cours de cette thèse, lors de la tentative de la double recombinaison homologue pour obtenir des mutants de TALEs dans certaines souches *Xpm*, nous ont conduit à la mise en place d'un outil de « silencing » basé sur CRISPRi. La plateforme de « silencing » est basée sur la protéine dCas9 de *Streptococcus pyogenes* et les résultats montrent qu'un seul sgRNA suffit pour éteindre l'expression d'au moins quatre TALEs dans une souche de *Xpm*. Les analyses *in silico* ont démontré que le RBS dans la région 5'UTR des TALEs est hautement conservé parmi les espèces de *Xanthomonas* permettant l'application du système conçu pour *Xpm* à d'autres espèces comme *Xoo*, *Xcc* et *Xc*. Chez ces pathogènes, un seul sgRNA a été suffisant pour supprimer l'expression d'au moins huit TALEs. En outre, la mise en œuvre de cette technologie chez *Xc*, couplée au séquençage complet du génome d'une souche, à la prédiction *in silico* de cibles et à des essais de virulence, indique que les TALEs sont importants pour la pathogenèse de *Xc*, potentiellement via l'activation du gène *MeSWEET10a*. Bien que certains travaux soient encore nécessaires, cet outil ouvre de multiples possibilités pour faciliter l'étude de la fonction des gènes dans le genre *Xanthomonas*. Une optimisation et d'autres technologies connexes sont prévues après cette validation de l'utilité de CRISPRi chez ces phytopathogènes.

Table of contents

ACKNOWLEDGMENTS	I
RESUMÉ	II
ABSTRACT	III
ABBREVIATIONS	IV
RÉSUMÉ DÉTAILLÉ DES ACTIVITÉS RÉALISÉES AU COURS DE LA THÈSE	VII
INTRODUCTION.....	VII
JUSTIFICATION	IX
CHAPITRE 1 : DIVERSITÉ ET FONCTION DES TALEs DANS UNE POPULATION COLOMBIENNE DE <i>Xpm</i>	X
CHAPITRE 2 : <i>MeSWEET10E</i> , UN GÈNE ALTERNATIF DE SENSIBILITÉ DU MANIOC ACTIVÉ TRANSCRIPTIONNELLEMENT PAR LES TALEs DE <i>Xpm</i>	X
CHAPITRE 3 : CRISPRi, UNE STRATÉGIE INNOVANTE POUR L'ÉTUDE FONCTIONNELLE DE FAMILLES DE GÈNES CHEZ LES BACTÉRIES PHYTOPATHOGÈNES.....	XI
DISCUSSION ET CONCLUSIONS.....	XI
INTRODUCTION	1
CASSAVA CROP OVERVIEW	1
<i>Impacts of weeds, pests, and diseases on cassava</i>	2
<i>Cassava, diseases, and climate change</i>	3
<i>Disease resistance improvement in cassava</i>	4
CASSAVA BACTERIAL BLIGHT	5
<i>Disease description and epidemiology</i>	5
Cassava bacterial blight symptoms.....	5
Disease cycle.....	6
Ecology of the pathogen.....	6
Distribution of the pathogen	7
<i>Disease impacts and management</i>	9
Prevalence, incidence, and losses associated to CBB	9
Economic and social importance of the pathogen	12
Control strategies	13
<i>Pathogen description</i>	16
Pathogen identification	16
Pathogen classification	16
Host range	17
Diversity of the pathogen	18
Diversity assessment tools	18
Diagnostic tools	19
Disease phenotyping tools under controlled conditions	19
<i>Virulence mechanisms of the pathogen</i>	20
Genomics of the pathogen	20
Secretion systems.....	21
Type 3 effectors.....	21
Toxins	22
Signaling and metabolic routes potentially involved in regulation of pathogenicity.....	22
<i>Host genetics and interaction</i>	23
Tools for <i>Xpm</i> -Cassava interaction research	23
Host susceptibility to <i>Xpm</i>	23
Structural and cellular features of cassava resistance against <i>Xpm</i>	24

Mapping resistance to <i>Xpm</i>	24
Genetic and molecular aspects of cassava defense against <i>Xpm</i> and <i>Xc</i>	28
Resistance engineering in cassava.....	30
TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS (TALES)	31
<i>TALE overview</i>	31
TALE architecture and mechanistic function	31
Artificial or designer TALES	32
Evolution of TALES	32
<i>Pathogenicity functions of TALES</i>	34
TALE targets and consequences of interactions	34
Major TALES and susceptibility genes.....	34
Avirulence TALES and executor genes	39
Avirulence TALES and resistance genes	40
Interference or truncated TALES.....	42
TALE-related loss-of-susceptibility alleles.....	43
<i>Xpm TALES and their role in cassava</i>	45
TALE20: TAL20 _{Xam668} and other variants	46
TALE14: TALE1 _{Xam} , TAL14 _{Xam668} , and TAL14 _{ClO151}	46
RATIONALE	50
CHAPTER 1: TALE DIVERSITY AND FUNCTION IN A COLOMBIAN <i>XANTHOMONAS PHASEOLI</i> PV. <i>MANIHOTIS</i> (<i>XPM</i>) POPULATION.....	52
CHAPTER 2: <i>MESWEET10E</i>, AN ALTERNATIVE CASSAVA SUSCEPTIBILITY GENE TRANSCRIPTIONALLY ACTIVATED BY TALES OF THE CASSAVA PATHOGEN <i>XANTHOMONAS PHASEOLI</i> PV. <i>MANIHOTIS</i>	80
CHAPTER 3: CRISPR INTERFERENCE AS A TOOL TO STUDY TALE BIOLOGY IN <i>XANTHOMONAS</i>.....	119
GENERAL DISCUSSION	155
<i>XPM</i> TALE DIVERSITY.....	155
SUSCEPTIBILITY GENES.....	157
CRISPRi TOOL	161
CONCLUSIONS AND PERSPECTIVES	164
REFERENCES.....	165
ANNEXES	191

Table of figures

FIGURE 1 ETIOLOGY, ECOLOGY, AND DISTRIBUTION OF <i>XPM</i> AND <i>Xc</i>	8
FIGURE 2 GRAPHICAL META-ANALYSIS OF INCIDENCE, SEVERITY AND LOSSES REPORTED FOR CBB.	9
FIGURE 3 TAXONOMIC POSITION OF <i>XPM</i> AND <i>Xc</i>	17
FIGURE 4 ROLES OF XOPS, HOST DETERMINANTS OF SUSCEPTIBILITY AND DEFENSE RESPONSES IN THE CASSAVA- <i>XPM</i> INTERACTION.	30
FIGURE 5 GEO-LOCALIZATION OF REGIONS WHERE <i>XPM</i> STRAINS USED FOR EXPERIMENTAL ASSESSMENTS HAVE BEEN ISOLATED.	156
FIGURE 6 THEORETICAL TALE-MEDIATED ACTIVATION OF MeSWEETS.	160
FIGURE 7 THEORETICAL OPTIMIZATION POINTS AND TECHNOLOGICAL SHIFTS FORESEEN FOR CRISPR-BASED PLATFORMS IN <i>XANTHOMONAS</i>	162

Introduction

Cassava crop overview

Cassava (*Manihot esculenta* Crantz), also known as manioc and yuca, among other names, is a woody perennial plant from the *Euphorbiaceae* family, characterized by deeply, palmately lobed leaves, stems of up to 5-m long, and large starchy tuberous roots that, on average, account for half of the plant weight (FAO, 2013; Tewe et al., 2004). Its tuberous roots have elevated contents of water (from 45% to 85%) and carbohydrates (25% to 35%), and contain significant levels of calcium and vitamin C, which make them valuable as food (Montagnac et al., 2009). Cassava leaves and petioles are also consumed as food in some regions of Africa and Asia. They have a high content of proteins, B1, B2 and C vitamins, phosphorous, magnesium, potassium, and calcium. However, leaves also contain high concentrations of toxic cyanogenic glucosides and antinutritional factors such as high fiber content, tannins, polyphenols and phytic acid, which require a pre-treatment to allow consumption of reasonable quantities (Latif and Müller, 2015). The plant is highly tolerant to drought stress thanks to its efficient water use through control over stomata, reduction in leaf size and amount as a response to limited water availability, and a potential hydraulic lift capacity (Byju and Suja, 2020; El-Sharkawy, 2006). Furthermore, the fact that most of the nutrients taken from soil are concentrated in the aerial part results in maintenance of soil fertility by decaying leaves. Although the plant is perennial, it is generally cultivated in cycles with a duration that ranges from six months to two years, being eight to ten months the best harvesting time for human consumption of the root. Longer cycles result generally in more starch load and larger roots but affects palatability, and therefore these are mainly used for starch industry (FAO, 2013).

Cassava is widely cultivated in the sub-tropical countries of Africa, Asia, and Latin America as a staple food crop that makes part of the food and economic security of these countries. Cassava can grow in poor soils with pH ranging from 4 to 9, at elevations of up to 2300 meters above sea level, in annual mean temperatures ranging from 18-25°C, in regions with precipitation rates between 500 to 5000 mm per year (reviewed by (Byju and Suja, 2020)). This crop is propagated from stem cuttings, which results in low-cost and highly available planting material. This crop can produce up to 250kcal/ha/day, which is more than double the caloric yield of wheat and sorghum, and it is significantly higher than maize, sweet potato, and rice (El-Sharkawy, 2012). This relevant feature, the high adaptability of cassava to poor soils and environmental conditions, the flexibility of harvesting time, the low-input and management needs, its resistance to biotic stresses, and its nutrition value make cassava a farmer-friendly and the most important root crop. Cassava crop feeds about 800 million people, with more than 500 million people receiving more than 100cal/day in the tropics (Byju and Suja, 2020; FAO, 2013). The majority of this crop is planted and harvested by low-income farmers on small holdings, usually as an intercrop system with maize, rice, legumes, melons, banana or oil palm (FAO, 2013). Although the use of cassava roots varies from region to region, it is estimated that over 70% is devoted to human food, while the remaining 30% is used for livestock feed and starch and biofuel industry (Pushpalatha and Gangadharan, 2020). However, in Asia, cassava starch is mainly used to produce bioethanol from , which is pushing up the growth of cassava-based industry, turning it into an emerging commercial crop (Byju and Suja, 2020).

At a global scale, cassava production has been steadily growing from the 1980s, with an increasing yield percentage superior to rice, wheat, and potato, and close to that of maize (Pushpalatha and Gangadharan, 2020). For 2019, the global harvested area dedicated to cassava was 27,500 ha and the related production estimated on 303.6 MT, from which 78.5% of the harvested area is located in Africa and produced 192.1 MT of cassava (FAO, 2020). Efforts to improve crop yields through technical amelioration and breeding have given excellent results in the past three decades, growing from a global average of less than 5 T/ha to more than 11 T/ha. Despite this steady pace, after 2012 the global yield growth reached a plateau and now it is far from the ~32 T/ha maximum reached, and even farther from the potential yield of 80–100 T/ha (Bredeson et al., 2016; Byju and Suja, 2020). The leading cause of this gap is thought to be the unbalanced mineral nutrition (Byju and Suja, 2020), although biotic stresses as pests and pathogens are important threats to this crop (Bart and Taylor, 2017; Lin et al., 2019).

Impacts of weeds, pests, and diseases on cassava

Despite a different perception by most farmers, weeds are a major concern for the crop and can cause yield losses of up to 95% (Harris et al., 2015; Khanthavong et al., 2016). The initial low growth of cassava and the wide spacing between planted cuttings are two factors that favor the field colonization by narrow-leaf grassy weeds and broad-leaf weeds, which will easily outcompete cassava for nutrients and water (FAO, 2013). Several cultural practices are suggested to mitigate the impact of weeds, like promoting cassava growth with appropriately-applied fertilizers, increase the density of planted cuttings, planting during the dry season under drip irrigation, mulching, intercropping with fast-growing crops, and weeding. However, for big land extensions or for periods with low labor availability, herbicides can be applied, specially, at the pre-emergence setting (FAO, 2013). Fighting weeds is also important to have healthy crops, since some of them can act as reservoirs for cassava pests and pathogens, like bacteria and fungi.

There are several pest and pathogens that affect cassava development during the crop cycle and the roots in the post-harvest stage. Around 200 arthropod pests have been observed to attack cassava, but whiteflies, mealybugs, and mites are the most destructive insects for the crop. Although these pests can cause important yield losses, cultural practices and biological control agents are used to mitigate their impact (FAO, 2013). Whiteflies include the species *Bemisia tabacci*, which vectors viral agents causing the cassava mosaic disease (CMD) and the cassava brown streak disease (CBSD), the two main viral diseases of cassava in the Eastern hemisphere. CMD has been reported in Africa and Asia, and it seems to be expanding across the Asian continent (Lin et al., 2019; McCallum et al., 2017). The cassava mosaic geminiviruses are responsible for CMD, which is characterized by chlorotic, mosaic patterning of leaves on curly and twisted leaflets. This abnormal leaf development deeply affects the plant performance and the root production. In a similar way, the Ipomovirus Cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) are responsible for CBSD, which is characterized by the brown streaking of cassava stems, leaf chlorosis, reduced root yield, and necrosis within the root. The distribution of this pathogen is restricted to the sub-Saharan Africa, but it causes important yield losses of up to 70% (Lin et al., 2019).

Cassava bacterial blight (CBB), caused by *Xanthomonas phaseoli* pv. *manihotis*, is an important bacterial disease that has been responsible for famines and high losses, but its impacts have been buffered in the last decades mainly through cultural practices (López Carrascal and Bernal, 2012).

As this study is focused on this cassava pathogen, more details can be found in the section describing CBB. Cassava frogskin disease (CFSD) and cassava witches' broom (CWB) are two emerging diseases, and little is known about their biology. It is hypothesized that CFSD can be caused by a mix of viruses (Carvajal-Yepes et al., 2014) or by a bacterial phytoplasma (de Souza et al., 2013; Oliveira et al., 2013), while CWB is known to be caused by a bacterial phytoplasma. CFSD affects roots and hampers the normal development of the starchy root (Freitas et al., 2020); it has been reported in Central and South America, and recently in Japan (Koinuma et al., 2018). CWB affects the plant sprouting pattern and leaf development, resulting in many clustered, poorly developed shoots along the stem; it has been reported to cause important yield losses in Africa, Asia, and America (Alvarez et al., 2013). Fungal pathogens are often associated to damage of the roots in the post-harvest stage. Also, *Cercospora* spp can cause leaf spot disease and *Colletotrichum* spp can cause anthracnose. These pathogens are considered as having minor impacts on the crop, marginally affecting root yield and quality of the cuttings, usually in mixed infections with other pathogens (Hillocks and Wydra, 2002).

Cassava, diseases, and climate change

As the climate changes, variations in light, CO₂, temperature, water availability will potentially compromise the performance of the main crops and will reshape the landscape of the regions where they are currently adapted. Likewise, biotic factors such as pollinators, pests, and pathogens, will also be forced to adapt and this will potentially create new niches and disease foci (Chavarriaga Aguirre et al., 2016). In line with this, a large-scale agronomical response is necessary to guarantee the food security and efforts to counterbalance the negative effects of climate change on the current agronomical production might include selection for adapted varieties with comparable or better performance. Estimations indicate that, among other important crops, cassava and sorghum will be the least negatively affected by these changes (-3.7% to +17.5% in a country-based suitability score, in comparison to $-16\% \pm 8.8$ for beans, -14.7 ± 8.2 for potato, $-2.5\% \pm 4.9$ for banana, and $-2.66\% \pm 6.45$ for sorghum) (Jarvis et al., 2012). This favorable response expected for cassava crops is based in several advantageous features of the plant, which render this a climate-smart crop. Cassava can tolerate a wide range of temperatures (16°C to 38°C), and some authors claim that its photosynthetic rate augments even up to 40°C (Pushpalatha and Gangadharan, 2020; Ravi et al., 2008), which is highly beneficial in view of the expected temperature variations. In the same trend, water availability and humidity will be affected by higher temperatures, which will increase drought and salinity. Cassava plants have exceptional biological mechanisms to cope up with water stress by reducing transpiration through leaf folding and drooping, reduction of shoot biomass, and partial stomatal closure, while maintaining high photosynthetic rates and having little impact on harvest index, a unique characteristic among the most important staple crops (Connor and Palta, 1981; El-Sharkaway and Cock, 1984; El-Sharkawy, 2007; Pushpalatha and Gangadharan, 2020). The expected augmentation of CO₂ levels, in combination with the rest of variables, is estimated to be more advantageous for cassava than for other crops. As it already has good mechanisms to balance water availability and photosynthesis rates, an increase in available CO₂ and slightly higher temperatures would lead to improved growth and root yields even under water stress (Cruz et al., 2016). All these advantages indicate that cassava will agronomically perform better than most of the current staple crops, and it can be extended to new growing areas in the future (Pushpalatha and Gangadharan, 2020). This potential extension and the adaptation and environmental interactions of its pathogens in the new niches might represent important threats for the crop health. Moreover, cassava response to biotic stressors is characterized by a large dependence on environmental conditions

(genotype by environment interactions) (Ceballos et al., 2020; Dixon et al., 2002; Zinsou et al., 2005). Luckily, the history of the crop shows that cassava can be improved through breeding and engineering (even if it is difficult) to select advantageous phenotypes and to overcome biotic stresses (Ceballos et al., 2020; Lin et al., 2019). The knowledge of the molecular mechanisms that determine the pathogenicity can boost this process through more targeted approaches, and therefore studies on the susceptibility and resistance factors are necessary to be ready for climate change adaptation in this crop.

Disease resistance improvement in cassava

Breeding and cassava plant selection has been in place from the decade of 1930 and has rendered excellent results in the selection of agronomical traits, especially root yield (a current maximum of 34.5 tons of fresh roots per ha) and root starch content (a current maximum of 26%), but also resistance to CMD and, to some extent, to CBB (Hahn et al., 1980). The contribution of marker-assisted selection (MAS) and genomic selection through association studies has been limited by constraints associated to the plant biology, such as the high heterozygosity, low multiplication rates (from 1 adult plant, a maximum of 10 plants can be produced through cuttings), and the long vegetative periods until the obtention of botanical seedling (Ceballos et al., 2020). However, the introgression of a recessive allele that changes the composition of the starch in the root to create an amylose-free starch, a desirable industrial trait, is a good example of the advances achieved through MAS (Aiemnaka et al., 2012). Although no transgenic cassava lines are commercially available, the transgenesis approaches have also resulted in plants with higher nutrition values and longer shelf life (Beyene et al., 2018; Narayanan et al., 2019). Disease resistance enhancing through transgenesis has also been explored for CMD, CBSD, and CBB (the latter one only at the laboratory/greenhouse scale) (Chavarriaga Aguirre et al., 2016; Díaz-Tatis et al., 2019; Lin et al., 2019). Application of the RNAi in cassava against the CBSV and UCBSV has, for example, led to obtention of potential CBSD-resistant cassava varieties, now under regulatory assessment field trials in Kenya and Uganda (Beyene et al., 2017). The test of this technology against the East African cassava mosaic virus (EACMV) (Vanderschuren et al., 2009) and Sri Lankan cassava mosaic virus (SLCMV) (Ntui et al., 2015) also resulted in plants with significantly enhanced resistance to CMD.

New biotechnological tools, like RNA interference (RNAi) and the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9 (CRISPR/Cas9), have led to a new era of genome engineering and boosted the improvement of pathogen resistance in plants (Sharma et al., 2013; Wada et al., 2020). CRISPR/Cas9-based tools have been applied in cassava to suppress known susceptibility alleles. This is for example the case of the simultaneous disruption of two alleles (*nCBP-1*, and *nCBP-2*) of the translation initiation factor 4E (*elf4E*) gene family, which is associated to the interaction with the viral genome-linked protein (VPg) of CBSD-causing viruses. Through engineered loss of susceptibility, the mutants showed delayed symptoms, reduced root symptom severity, and decreased virus titers in the storage root (Gomez et al., 2019). Genome editing approaches have not been implemented against *Xpm* so far, but the key susceptibility gene *MeSWEET10a* (described with more details in the following sections) could be a suitable target for this technology (Lin et al., 2019).

As pointed out, the knowledge of virulence factors, susceptibility, and resistance genes in a plant-pathogen interaction can be exploited in targeted approaches to obtain varieties with improved resistance. The following sections of this document are dedicated to better understand some of

these molecular factors that define the interaction between the CBB-causing pathogen *Xpm* and cassava.

Cassava bacterial blight

This introductory section addresses the state of the art for the cassava disease caused by *Xanthomonas phaseoli* pv. *manihotis*, and, when the information was available the disease caused by *Xanthomonas cassavae*. This introduction to the pathosystem is divided in five subsections starting by the etiology and epidemiology, impacts and management of the disease, and moving to the molecular characterization of the pathogen and its molecular determinants that play a role in pathogenesis. The section finishes by addressing what we know about the interaction and the host molecular determinants that play important roles in the disease. Most of these sections were put together to publish the first pathogen profile on *Xanthomonas* causing diseases in cassava. The accepted version of the article (DOI:10.1111/mpp.13094) can be found appended to this thesis document.

Disease description and epidemiology

Cassava bacterial blight symptoms

Cassava bacterial blight (CBB) which is caused by the bacterial pathogen *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), is characterized by a range of symptoms that mainly affect leaves, petioles, and stems, frequently leading to plant death (Figure 1). Early symptoms appear as brown to dark-brown water-soaked translucent angular spots on the leaf tissue browning at later stages, occasionally surrounded by a chlorotic halo. Veins around these spots discolor and affected tissues frequently produce creamy white and later yellow-to-orange exudates on the abaxial side of the leaf. Blight results from spot coalescence which creates necrotic areas that become dry and curl the leaflets, giving them the aspect of a superficial burn. As disease progresses, bacteria access the xylem vessels from the mesophyll and move towards the stem through the petioles, which become brown and collapse, causing the leaf to wilt. Vessel colonization in the stem allows *Xpm* to systemically move upwards and downwards. When infection reaches the plant upper part where stem tissues are greener and less lignified, stem rotting leads to dieback characterized by shoot apex wilting. New sprouts can grow from buds located in more basal zones, giving the plant the appearance of a candle stick. However, if these buds are also contaminated by *Xpm* they will eventually wilt. Infected fruits also show water-soaked spots, and the resulting seeds suffer cotyledon and endosperm necrosis, and seed deformation. Roots from highly susceptible cultivars can show delayed symptoms restricted to the vascular tissues, with discolored vascular strands surrounded by dry and rotten spots (Boher et al., 1995; Lozano, 1986; Lozano and Sequeira, 1974a; Maraite and Meyer, 1975).

Symptoms of cassava bacterial necrosis (CBN) caused by *Xanthomonas cassavae* (*Xc*) might be similar at a first sight (Figure 1H), but the outcome of the infection is not as devastating (Mostade and Butare, 1978). Infections caused by *Xc* initially produce rounded water-soaked spots surrounded by a yellow chlorotic halo and radial necrosis of the veins (Maraite and Perreaux, 1979). As the disease progresses, bacteria colonize the adjacent mesophyll tissues, lesions expand and yellow exudates can be observed in them (Maraite, 1993) Leaves wilt, collapse and dry, but the pathogen is not able to colonize vascular tissues and there is no formation of secondary spots that turn into

extended blight areas (Lozano and Sequeira, 1974a; Maraite, 1993; Maraite and Perreux, 1979; Van Den Mooter et al., 1987; Verdier et al., 1994).

Disease cycle

Xpm has an initial epiphytic phase where it is able to colonize the surface of cassava aerial tissues. When conditions are favorable, especially humidity, multiplication of the pathogen increases around trichomes and penetration of the outer layer takes place (Daniel and Boher, 1985a). *Xpm* accesses the plant through wounds and natural openings of the leaf, potentially primarily via adaxial stomates (Kemp et al., 2004). Although it is common for other vascular *Xanthomonas* to enter through hydathodes, their role in the initial stages of infection for *Xpm* is not known. Internal tissue colonization is accompanied by the production of highly hygroscopic bacterial exopolysaccharides (EPS), which are exudated from lesions and then hydrated and carried by raindrops. Wind and rain-mediated splashing of these bacterial suspensions is considered the main natural way of horizontal pathogen transmission from one plant to another (Lozano and Sequeira, 1974b). Once *Xpm* hits new cassava leaves, the disease cycle restarts. Since this crop is propagated from cuttings, presence of the pathogen on the propagative material and working tools is the major factor for disease spreading (Lozano and Sequeira, 1974b). Plants sprouted from contaminated cuttings quickly develop the disease and are a major source for secondary infections in the field. Bacteria can also be transmitted inside the sexual seed (Daniel and Boher, 1985b; Elango and Lozano, 1980). Stems can get infected by the dispersal of wind-driven sand or hail (Maraite, 1993).

Ecology of the pathogen

Xpm shows two different lifestyles in the field, an epiphytic phase (Daniel and Boher, 1985b; Elango and Lozano, 1981) and a biotrophic parasitic phase that starts when environmental conditions facilitate pathogen growth and entrance into the mesophyll apoplast of cassava leaves (Verdier et al., 1990). The epiphytic stage plays an important role ensuring the natural persistence of the pathogen intercropping. Epiphytic populations of *Xpm* have been found on asymptomatic leaves in fields where CBB has been reported. Bacterial titers vary importantly according to environmental conditions, humidity and temperature being two of the most important factors. The optimal temperature for CBB development is 30°C, whereas that for CBN is 25°C (Maraite and Perreux, 1979). There is a marked increase of epiphytic populations and a shift to the parasitic phase during the rainy season. Conversely, symptoms of CBB are less frequent and epiphytic populations decrease during the dry season (Daniel and Boher, 1985a, 1985b). *Xpm* has also been detected in several weeds that occur naturally in cassava fields in South America, suggesting that the pathogen can epiphytically survive on them during intercropping (Elango and Lozano, 1981). Artificial inoculation of several cassava crop-associated weeds (non-host interactions in all cases) highlighted different degrees of bacterial survival, and pathogen viability for up to 54 days (Fanou et al., 2017; Marcano and Trujillo, 1984).

Infected plant debris and soil are also suggested to play a role in *Xpm* field persistence. However, survival of *Xpm* in a free-living state in the soil has been experimentally proven to be limited to up to three weeks (Fanou et al., 2017). In contrast, *Xpm* was shown to survive in slow-decaying dry debris for more than two months and one year under field and controlled conditions respectively (Daniel and Boher, 1985b; Fanou et al., 2017). Insect-mediated dissemination of *Xpm* has been reported. The African grasshopper pest *Zonocerus variegatus* recovered from diseased cassava

fields harbored a significant number of infective bacterial cells (Daniel and Boher, 1985b; Zandjanakou-Tachin et al., 2007). The coreid bug *Pseudotheraptus devastans* has also been reported as facilitator of infection through the generation of punctures through which the bacterium gains access to internal tissue (Maraite and Meyer, 1975).

Distribution of the pathogen

Botanical, geographical and domestication origin of cassava is a controversial topic. A widely-accepted theory suggests that the cultivated cassava species *Manihot esculenta* subsp. *esculenta* arose from the wild *Manihot esculenta* subsp. *flabellifolia* in the Amazon basin, and its domestication began 5000 to 7000 years b.c. in the southern Amazon border region (Allem, 2002; Olsen and Schaal, 1999). Cassava was introduced to Africa by Portuguese traders during the 16th century and was adopted as a staple crop across several countries (Jones, 1959). In Asia it is believed to be introduced from Mexico to the Philippines during the 17th century. In agreement with the geographical origin of the crop, genetic diversity of *Xpm* populations in South America has been shown to be high (Bart et al., 2012; Verdier et al., 2004).

CBB was first reported in Brazil in 1912 by Bondar and then identified in different South American countries during and after the 1970s. The disease and the pathogen were detected in Africa for the first time in Nigeria in 1972, and then systematically detected in several Sub-Saharan countries (Hillocks and Wydra, 2002; Persley, 1976). To date, *Xpm* has been reported in 49 countries located all over in the tropics (CABI, 2020a; Taylor et al., 2017). On the other hand, *Xc* has only been reported in the East African countries of Democratic Republic of the Congo, Rwanda, Malawi, Tanzania, Uganda, Kenya, and Burundi (Boher and Verdier, 1994; CABI, 2020b; Maraite, 1993; Mostade and Butare, 1978). Figure 1 shows cassava production areas and the distribution of both pathogens.

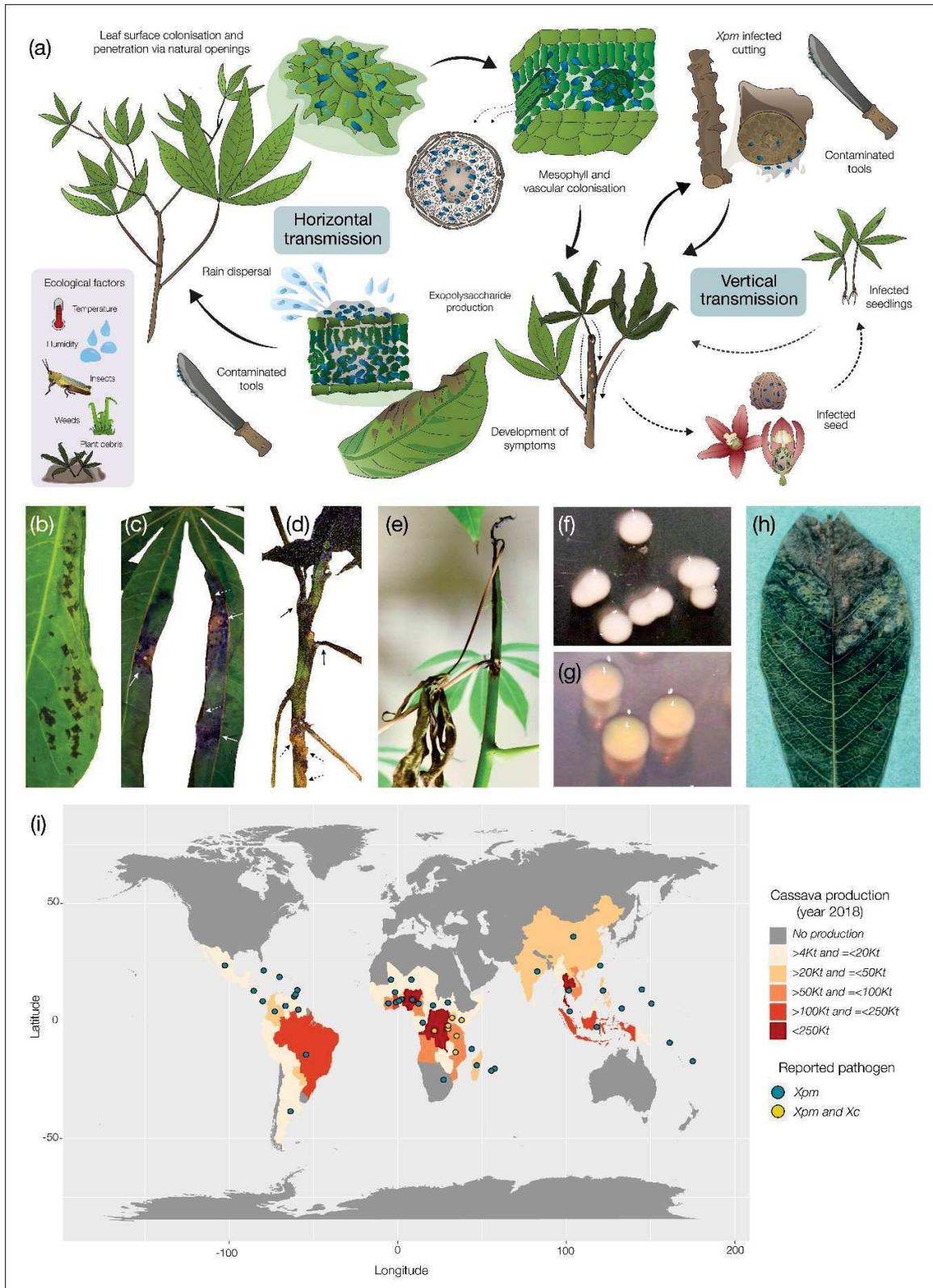


Figure 1 Etiology, ecology, and distribution of *Xpm* and *Xc*.

(a) Cassava bacterial blight disease cycle. Solid arrows indicate processes that are relevant for cassava cropping systems. Dashed arrows indicate processes that have been detected and are scientifically interesting but are not relevant for cassava cropping systems. The inset indicates environmental and ecological factors that affect the spread or development of disease. (b) angular leaf spots caused by *Xpm*. (c) blight (solid white arrows) and leaflet curling (dashed white arrows) caused by *Xpm*. (d) collapsed petioles from wilted leaves (solid black arrows) and gum exudation (dashed black arrows) from stems caused by *Xpm* infection. (e) shoot apex wilting and dieback caused by *Xpm*. (f) typical colonies of *Xpm* on YPG medium. (g) typical colonies of *Xc* on YPG medium. (h) leaf spots caused by *Xc*. (i) world-wide cassava production (FAO, 2020) by country (missing data for South Africa, Guam, and Palau) for 2018 and distribution of *Xpm* and *Xc* (CABI, 2020a).

Disease impacts and management

Prevalence, incidence, and losses associated to CBB

Historically CBB is reminiscent of a famine period from 1970 to 1975 in the Democratic Republic of the Congo, where cassava crop diversity was low and food and economic dependence on cassava was considerable (Hillocks and Wydra, 2002; Lozano, 1986; Maraite and Meyer, 1975). Several studies have addressed the distribution and effects of the pathogen on the crop. Figure 2 (and Tables 1, 2, and 3) summarizes data from several reports where incidence, severity and/or yield losses were assessed, mainly in farms or field trials. Incidence varies according to environmental factors, but most of the reports show ranges between 30% and 90%, with an incidence peak between 60% and 70%. Disease severity reports show systemic symptoms in all the surveys (values greater than 2 in 1 - 5 scales), even for measurements performed within the three first months after planting (MAP). Severity seems higher between the third and the sixth MAP, and there is a slight decrease in the following months, but complete dieback and plant stunting have been recorded in all the surveys performed after the third MAP.

Yield losses are more difficult to measure accurately mainly due to two reasons: i) cassava fields are not only affected by CBB in most of the cases, and ii) as for other diseases, outcomes are dependent on environmental conditions, i.e. plants can counterbalance the negative effects of CBB when favorable growth conditions are given (Harris et al., 2015; Zinsou et al., 2005). However, available metadata suggest that fresh root yield losses (Figure 2) can reach up to 100 %, with a median of about 50% in susceptible varieties; while it reaches up to 76% in resistant varieties, but the median is below 25%. This is an indication of the high potential of *Xpm* to cause important losses.

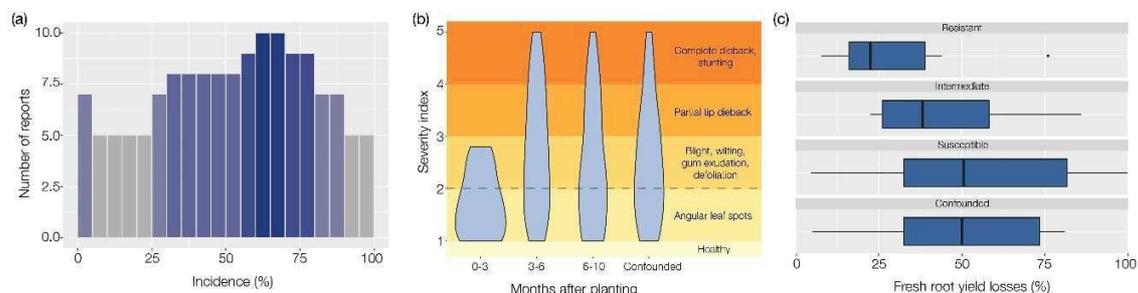


Figure 2 Graphical meta-analysis of incidence, severity and losses reported for CBB.

(a) Histogram of the incidence ranges reported by 11 studies (Table 1). The color scale correlates with the frequency of the reported range. (b) Violin plots showing severity ranges recorded by 15 studies (Table 2) according to measurement timepoints after planting. The confounded variable groups data from reports that did not include timepoint information. A harmonized severity grading scale is presented at the right side of the plot. The dashed line indicates that values above 2 reflect systemic disease. (c) Boxplots showing fresh root yield losses ranges described by 7 reports (some studies reported several ranges - Table 3) according to the designated resistance status of the plants. The confounded variable groups data from reports that did not include the designated resistance status of the evaluated plants.

Table 1. Incidence data extracted from several studies on CBB. Incidence was defined in these studies as the percentage of symptomatic plants (No. of symptomatic plants*100/No. of evaluated plants) recorded in cassava fields that were naturally infected by *Xpm* (no artificial inoculations). Incidence ranges (minimum and maximum), timepoint of the incidence evaluation, and country where the study was performed were extracted from each report. Min. = minimum, Max. = maximum, MAP = months after planting, nd = missing data.

Study reference	Min. (%)	Max. (%)	Timepoint (MAP)	Country
(Arene, 1976)	22.8	22.8	4	Nigeria
(Fokunang et al., 2000)	12.5	100	3	Nigeria
(Fokunang et al., 2000)	25.3	100	6	Nigeria
(Restrepo et al., 2000a)	0	17.6	4	Colombia
(Restrepo et al., 2000a)	0	100	7	Colombia
(Jorge et al., 2001)	0	2.2	4	Colombia
(Jorge et al., 2001)	94.4	100	7	Colombia
(Wydra and Verdier, 2002)	27	86	nd	Benin
(Wydra and Verdier, 2002)	0	8	nd	Ghana
(Banito et al., 2007)	27.4	90.5	nd	Togo
(Night et al., 2011)	2	2	nd	Rwanda
(Abaca et al., 2013)	73.8	100	nd	Uganda
(Affery et al., 2016)	2.2	66.7	nd	Ivory coast
(Affery et al., 2016)	1.00	89.00	4	Ivory coast
(Affery et al., 2017)	64.7	79.4	2	Ivory coast
(Dania and Ojeyemi, 2019)	33.9	76.1	nd	Nigeria
(Mbah et al., 2019)	58.3	67.5	9	Nigeria

Table 2. Severity data extracted from several studies on CBB. Severity was measured in these studies using similar scales based on symptom recording in cassava fields that were naturally infected by *Xpm* (no artificial inoculations). Severity ranges (minimum and maximum), timepoint of the severity evaluation, country where the study was performed, and grading scales were extracted from each report. Transformation column indicates if the values presented as minimum and maximum were mathematically transformed to obtain values of the same type and scale. Min. = minimum, Max. = maximum, MAP = months after planting, nd = missing data. Grading scales showed slight modifications, but all of them reflected the same damage extension with the corresponding grading

numbers. Most of the scales were grouped within the Scale A: 1, no disease symptoms; 2, only angular leaf spots on leaves; 3, leaf blight, leaf wilt, defoliation, gum exudation in petioles and stems; 4, extensive leaf blight, wilt, defoliation, and partial stem die-back; 5, complete defoliation, stem die-back and stunting in most of the plants. Only one study used a similar grading system with different scale numbers, Scale B: 0 = no symptoms; 1 = symptoms on leaves only - blight; 2 = presence of necrotic lesions on the stem or petiole; 3 = most severe symptoms on leaves and/or the presence of necrotic lesions with gum exudation; 4 = complete loss of leaves with apical death or death of the plant.

Reference	Min.	Max.	Timepoint (MAP)	Country	Scale	Transformation
(Umemura and Kawano, 1983)	1	2.5	2	Colombia	A	No
(Umemura and Kawano, 1983)	1.5	5	4	Colombia	A	No
(Umemura and Kawano, 1983)	1.5	5	6	Colombia	A	No
(Joseph and Elango, 1991)	1	3.452	nd	Trinidad	A	No
(Lamptey et al., 1998)	1	4.8	6	Ghana	A	No
(Fokunang et al., 2000)	1.5	2.8	3	Nigeria	A	No
(Fokunang et al., 2000)	2	4	6	Nigeria	A	No
(Restrepo et al., 2000a)	1	2.9	4	Colombia	A	No
(Restrepo et al., 2000a)	1	5	7	Colombia	A	No
(Jorge et al., 2001)	1	4.5	4	Colombia	A	No
(Jorge et al., 2001)	1.6	4.5	7	Colombia	A	No
(Dixon et al., 2002)	1	4	6	Nigeria	A	No
(Wydra and Verdier, 2002)	1	5	nd	Benin	A	No
(Wydra and Verdier, 2002)	1	2	nd	Benin	A	No
(Onyeka et al., 2004)	1	3.19	nd	Nigeria	A	No
(Restrepo et al., 2004)	1	4.5	7	Colombia	A	No
(Banito et al., 2007)	1.35	2.4	nd	Togo	A	Yes
(Affery et al., 2016)	1.01	1.09	2	Ivory Coast	A	Yes
(Affery et al., 2016)	1.07	1.8	3	Ivory Coast	A	Yes
(Affery et al., 2016)	1.23	2.43	4	Ivory Coast	A	Yes
(Affery et al., 2016)	1.11	1.69	5	Ivory Coast	A	Yes
(Affery et al., 2016)	1.09	1.51	6	Ivory Coast	A	Yes
(Affery et al., 2016)	1.05	1.22	7	Ivory Coast	A	Yes
(Affery et al., 2016)	1.02	1.2	8	Ivory Coast	A	Yes
(Affery et al., 2016)	1	1.22	9	Ivory Coast	A	Yes
(Affery et al., 2016)	1	1.19	10	Ivory Coast	A	Yes
(Oliveira et al., 2016)	1.96	3.85	10	Brazil	B	Yes
(Affery et al., 2017)	2.03	2.41	9	Ivory Coast	A	Yes
(Dania and Ojeyemi, 2019)	1.5	4.5	nd	Nigeria	A	No
(Mbah et al., 2019)	1.97	2.21	9	Nigeria	A	No

Table 3. Losses data extracted from several studies on CBB. Only losses reported as the percentage of fresh root weight lost were considered. All the studies were carried out in cassava fields naturally or artificially infected by *Xpm*. Loss ranges (minimum and maximum), resistance status of the tested varieties, and country or region where the study was performed were extracted from each report. Min. = minimum, Max. = maximum.

Reference	Inoculum	Min. (%)	Max. (%)	Status	Country/Region
(CIAT, 1973)	Artificial	7.6	14.4	Resistant	Colombia
(CIAT, 1973)	Artificial	22.5	27.3	Intermediate	Colombia
(CIAT, 1973)	Artificial	34.2	57.0	Susceptible	Colombia
(Otim-Nape, 1980)	Natural	90.0	100.0	Susceptible	Uganda
(Umemura and Kawano, 1983)	Natural	44.0	76.0	Resistant	Colombia
(Umemura and Kawano, 1983)	Natural	49.0	86.0	Intermediate	Colombia
(Umemura and Kawano, 1983)	Natural	51.0	92.0	Susceptible	Colombia
(Lozano, 1986)	Natural	75.0	75.0	Confounded	Zaire
(Lozano, 1986)	Natural	80.0	80.0	Confounded	Central Africa
(Lozano, 1986)	Natural	50.0	50.0	Confounded	Brazil
(Lozano, 1986)	Natural	5.0	40.0	Confounded	South America
(Lozano, 1989)	Artificial	29.4	72.0	Confounded	South America
(Fanou, 1999)	Natural	21.0	24.0	Resistant	Benin-Nigeria
(Fanou, 1999)	Natural	32.0	50.0	Susceptible	Benin-Nigeria
(Fanou, 1999)	Natural	5.5	5.5	Susceptible	Benin-Nigeria
(Fanou et al., 2018)	Natural	50.0	50.0	Confounded	Nigeria
(Fanou et al., 2018)	Natural	10.0	30.0	Confounded	Taiwan

Economic and social importance of the pathogen

In a recent report of the Evans School Policy Analysis and Research (Harris et al., 2015) for the Bill and Melinda Gates Foundation, economic impacts of CBB and postharvest physiological deterioration were weighed in the context of main constraints for cassava production. Despite its well-known potential to cause important losses, the current impacts of CBB on the crop are masked by the lack of recent surveys due to limited research on the disease, underestimations due to the lack of farmers' training for CBB identification, and the influence of environmental factors like drought and coinfections on the onset and severity of the disease. A two-year follow-up of farmers practices and crop outcomes in Uganda and Kenya revealed that despite the high concern of farmers about the impacts of diseases like CBB and cassava anthracnose disease (CAD), the factors impacting yield losses the most were soil fertility and weed management (Fermont et al., 2009). This latter could result from the cumulative experience of agricultural control measures applied to cassava farming since the 1980s, which may have buffered the impact of CBB (Lozano, 1986).

Conversely, CBB dynamics seem to be altered by social behaviors linked mainly to cutting trading, which favors the distribution of the pathogen over short and long distances (Restrepo et al., 2004; Restrepo and Verdier, 1997; Trujillo et al., 2014b, 2014a). In a case of study in Colombia, disease incidence is positively correlated with agrochemicals use, land ownership and propagative material

sharing. Land ownership limitations due to internal conflict and inequity in cassava farmer communities hampers the long-term establishment of crops and force to inter-cycle renewal of propagative material, which leads to the use of any available sources of stakes that frequently lack of phytosanitary controls (Pérez et al., In prep.).

Control strategies

Efficient CBB control is based on three main pillars: sanitary controls, cultural practices, and deployment of tolerant or resistant varieties. Control practices have been deployed based on our knowledge about *Xpm* and its ecology. Sanitary controls of propagative materials and seeds are large-scale measures that aim at stopping pathogen dispersal. In general these measures comprise the deployment of disease-free materials, treatments to eliminate the pathogen, establish quarantine, and *Xpm* detection tools (Chavarriaga Aguirre et al., 2016; Frison and Feliu, 1991).

Cultural practices for CBB management include crop rotation, intercropping, fallowing, removal or burying of crop debris, weed management, delayed planting (at the end of the rainy season), and use of clean propagative materials. Crop rotation and fallowing aim at depleting the pathogen inoculum sources through time; they are deemed as effective at buffering the impact of CBB in successive crop cycles (Lozano, 1986; Persley, 1978). Since weeds are reservoirs of the pathogen, their management during the entire crop cycle, as well as debris treatments before planting, are strongly recommended. Fanou and collaborators (Fanou et al., 2017) showed that *Xpm* survival is markedly reduced when debris is covered with soil or buried (less than 30 days vs more than 120 days for non-buried debris under dry conditions). Moreover, Fanou and Wydra (Wydra and Fanou, 2015) showed that removal of symptomatic leaves can reduce CBB severity and improve the quality of the crop as source of propagative material; however, this practice did not show effects on yields in their study. Planting during the second half of the rainy season, disease incidence and severity decrease, while yields are maintained or even improved generally (Ambe, 1993; Umemura and Kawano, 1983; Zinsou et al., 2004).

Considering that cassava is devoted to sustaining mainly low-income farmers, deployment of tolerant or resistant varieties is considered the main solution for CBB control. Clonal selection and breeding for cassava disease resistance started in the 1940s and were mainly focused on CMD. Resistance against CBB is essentially due to genes introgressed from wild *Manihot* species, like the ceara rubber tree *M. glaziovii* (Nassar, 2007). Several studies addressed the performance of diverse cassava varieties against CBB (Dixon et al., 2002; Fokunang et al., 2000; Lamptey et al., 1998; Lozano and Laberry, 1982; Restrepo et al., 2000a; Umemura and Kawano, 1983; Wydra et al., 2004, 2007; Zinsou et al., 2005) (see Table 2). In summary, three main aspects condition the successful deployment of CBB resistant varieties: i) interactions of cassava with environmental factors profusely affect the fitness and resistance of some cultivars; ii) the need for co-selection of agronomical traits of value for farmers and the local market; and iii) the resistance against CBB is polygenic, additively inherited (Umemura and Kawano, 1983), pathotype-specific (Wydra et al., 2004) and its molecular basis needs to be elucidated.

Regarding the first aspect, several studies (Dixon et al., 2002; Restrepo et al., 2000a; Wydra et al., 2007; Zinsou et al., 2005) demonstrated that the environment has a significantly greater influence on CBB disease incidence than the genotypic component, and that the interaction of both factors can mask differences between genotypes. For instance, the highly resistant variety TMS30572

performed as resistant in two edaphoclimatic zones of Benin, while it was moderately resistant in four other climatic zones (Zinsou et al., 2005). Therefore researchers draw attention to test resistance in relevant conditions, with high disease pressure, and in parallel in greenhouse settings where pathogenicity assays are optimal (Restrepo et al., 2000a; Zinsou et al., 2005). Regarding the second aspect, surveys have shown that smallholders prioritize higher yields, taste, and good cooking qualities over diseases tolerance when selecting materials, which is highly relevant when engineering disease resistance (Harris et al., 2015). The latter aspect - cellular, genetic, and molecular bases of the resistance - is addressed later in this document.

Despite the above-mentioned caveats and to summarize the attempts to find high-performing cassava varieties (see Table 4), the outstanding resistance against CBB of genotype TMS30572 has been widely demonstrated in several trials in African countries. This genotype comes from the CBB-resistant parent 58308, a low-productive interspecific hybrid with *M. glaziovii*, and the susceptible, but high yielding Brazilian variety Branca Caterina de Santa (<https://seedtracker.org/cassava/> (Nassar, 2007; Umanah, 1977)). Although performance of cassava varieties tested in South American country fields has been highly variable, genotypes CMC40, MECU82, MCOL1916, MPAN19, MPAN12B, MBRA685, MBRA886, MBRA902, MNGA2, CM523-7, and CM6438-14 were considered resistant against different *Xpm* strains in a set of studies (Table 4).

Table 4. Result compilation of cassava resistance assessments reported in literature. Test period: the time span from planting to the last symptom reading. nd: not described. Setting: environmental conditions of the test, inoculation technique, and *Xpm* strains.

Test period	Setting	No. of tested varieties	Resistant varieties	Reference
5 years (4 cycles)	Field (one ecozone)/Naturally occurring disease coupled to greenhouse/leaf clipping (strain 1060)	1800	Four resistant varieties: M Ecu 82, M Col 1916, M pan 19, and M Pan 12B.	(Lozano and Laberry, 1982)
6 years	Field (two ecozones)/Naturally occurring disease.	~ 1400	Resistance degrees among tested plants is not explicit. CM 523-7 is the only variety explicitly deemed as resistant.	(Umemura and Kawano, 1983)
2 cycles	Field (two locations)/Naturally occurring disease.	12	Three resistant varieties: TMS 30001, TMS 30572 and TMS 91934.	(Akparobi et al., 1998)

2 years (2 cycles)	Field (two ecozones)/Naturally occurring disease coupled to Greenhouse/stem puncture (several undefined <i>Xpm</i> strains).	19	Three resistant varieties: M NGA 2, M BRA 685, and CM 6438-14.	(Restrepo et al., 2000a)
3 months	Greenhouse/Leaf inoculation and stem puncture (26 undefined <i>Xpm</i> strains)	17	Five resistant varieties: MBRA886, MBRA685, MBRA902, MNGA2, and CM6438-14.	(Restrepo et al., 2000b)
3 years	Field trials (three ecozones)/Naturally occurring disease	9	Two stable resistant varieties: 63397 and 4(2)1425.	(Dixon et al., 2002)
6 weeks	Glasshouse/Stem puncture or leaf infiltration (four <i>Xpm</i> strains: GSPB2506, GSPB2507, GSPB2511, and Uganda 12).	113	Three resistant varieties: TMS30572, CM8820-40, and CM8873-69.	(Wydra et al., 2004)
2 years	Field trials (three ecozones)/Naturally occurring disease and pray inoculation (three <i>Xpm</i> strains: GSPB 2506, GSPB 2510, and Save 10)	37	Five stable resistant varieties: CAP94030, BEN86040, RB89509, RB92132, and TMS30572.	(Zinsou et al., 2005)
2 years	Field (five ecozones)/Spray inoculation (<i>Xpm</i> strain X27).	22	Six resistant varieties: CVTM4, TMS92/0429, Main27, TMS30572, TMS91/02316, and TMS4(2)1425.	(Wydra et al., 2007)
2 months	Greenhouse/Stem puncture (4 <i>Xpm</i> strains: GSPB2506, GSPB2507,	24	Eleven resistant varieties: TMS92/0343, TMS92/0429, TMS92/0326, TMS92/0067, TMS92/0057, Toma289,	(Banito et al., 2010)

	GSPB251, and Uganda 12).		Toma378, TMS4(2)1425, TMS91/02316, CVTM4, and Gbazékouté.	
2 months	<i>In vitro</i> plants/leaf clipping (undefined <i>Xpm</i> strain)	21	Four resistant varieties: TME419, 30572,98/0505, Kibaha.	(Mbaringong et al., 2017)
3 months	Greenhouse/Stem puncture (undefined <i>Xpm</i> strain).	11	Four resistant varieties: TME419, 30572,98/0505, Kibaha.	(Mbaringong et al., 2017)

Pathogen description

Pathogen identification

Xpm and *Xc* are gammaproteobacteria that belong to the Xanthomonadaceae family. *Xpm* is a Gram negative rod with a polar flagellum that forms shiny, slimy, convex and circular colonies with entire margins, when sugars as glucose or sucrose are present in the medium, without any pigmentation (Figure 1F) (Lozano and Sequeira, 1974a; Maraite and Meyer, 1975; Van Den Mooter et al., 1987). The characteristic white color is due to the absence of xanthomonadin pigment, a rare trait also observed in *Xanthomonas citri* pv. *mangiferaeindicae* and *Xanthomonas campestris* pv. *viticola* (Midha and Patil, 2014). Colonies of the Gram negative rod *Xc* are slimy, convex, circular with entire margins and a deep yellow color (Figure 1G). According to Van Den Mooter and coworkers (Van Den Mooter et al., 1987), these two pathogens can be differentiated by colony color, growth on D-saccharic acid (positive for *Xc*, but not for *Xpm*), hydrolysis of Tween 60 and growth on DL-glyceric acid (positive for *Xpm*, but not for *Xc*).

Pathogen classification

First described by Bondar in 1912 the causal agent of CBB was initially named *Bacillus manihotis* (Arthaud-Berthet). After several taxonomical reclassifications (Burkholder, 1942; Dedal et al., 1980; Lozano and Booth, 1974; Maraite and Meyer, 1975; Vauterin et al., 1995), the most recent classification based on a polyphasic taxonomic approach (a seven-gene multilocus sequence analysis, average nucleotide identity and biochemical analyses) assigned this pathogen to the species *X. phaseoli*, with the definite nomenclature *X. phaseoli* pv. *manihotis* (Constantin et al., 2016). As to the causative pathogen of CBN, it was first described by Wiehe and Dowson in 1953 in Malawi and named *Xanthomonas cassavae* n. sp. (Wiehe and Dowson, 1953), which was further confirmed upon DNA-DNA hybridization later on (Vauterin et al., 1995). Figure 3 shows the taxonomic position of the two cassava pathogens among most known *Xanthomonas* species.

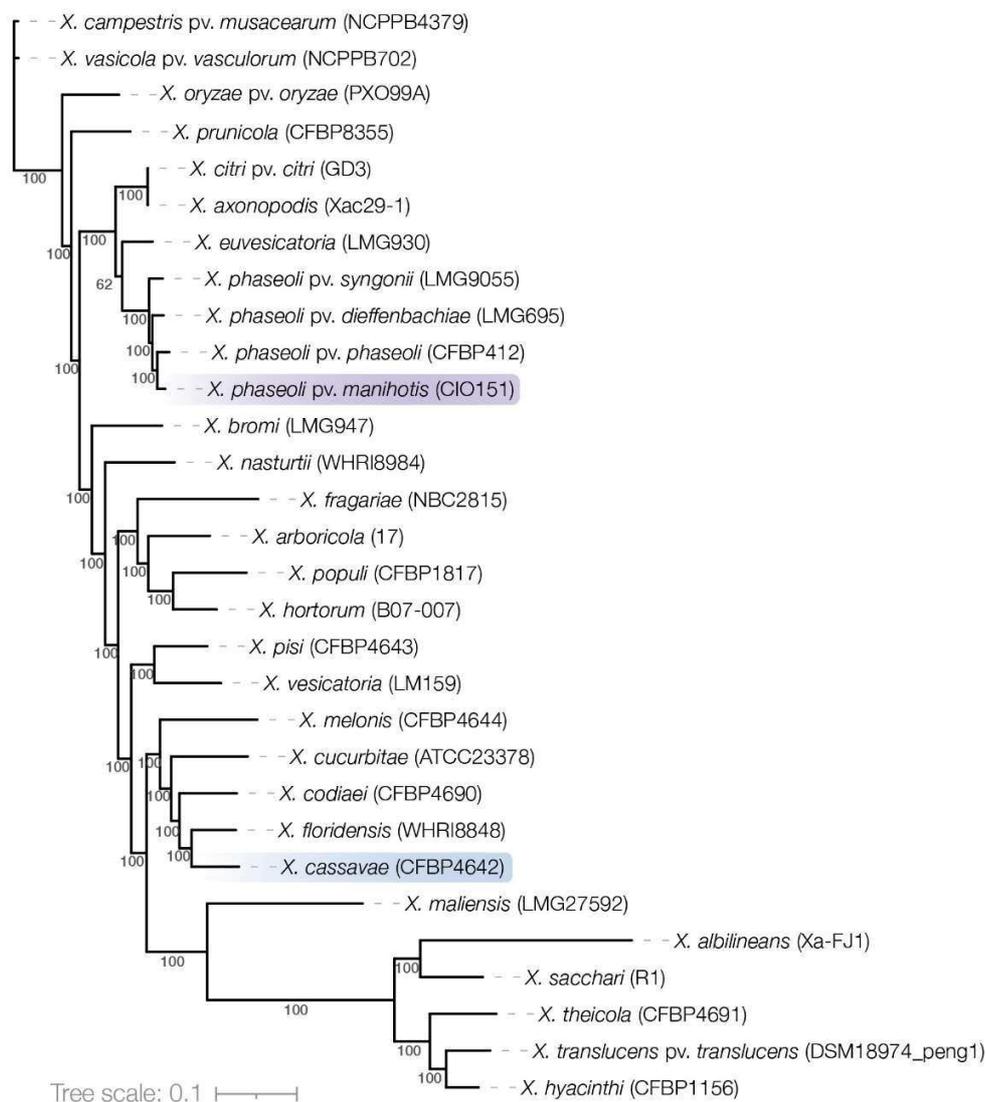


Figure 3 Taxonomic position of *Xpm* and *Xc*.

Phylogeny of 30 representative *Xanthomonas*, including 27 species, and 4 pathovars of *X. phaseoli*. Strain code is indicated in parentheses. *Xpm* and *Xc* are highlighted in purple and blue, respectively. Phylogeny was constructed from RefSeq complete genomes using the bioinformatic workflow PhaME (Shakya et al., 2020). Core genome alignments resulted in 155,507 aligned SNP positions that covered coding and non-coding regions. Trees were reconstructed with the GTRGAMMAI model of RAxML and consensus tree was calculated from 100 bootstraps; results higher than 80 are shown above branches.

Host range

Alternative hosts can play a key role in pathogen persistence in the environment, acting as reservoirs during long periods. To date, five Euphorbiaceous plants have been reported to be symptomatically affected by *Xpm*, supporting its multiplication and release to the environment: the three cassava wild relative species *Manihot glaziovii*, *Manihot palmata*, and *Manihot aipi* (Lozano and Sequeira, 1974a), *Euphorbia pulcherrima* and *Pedilanthus tithymaloides* (Dedal et al., 1980). On the other hand, the limited distribution of *Xc* in Africa and its apparent absence in South America (origin of cassava) suggests that coevolution with cassava is rather short and that this plant may not be the main host of this bacterium (Hayward, 1993). No alternative host has been reported for this pathogen so far.

Diversity of the pathogen

As reviewed previously (López Carrascal and Bernal, 2012), the first studies on African isolates highlighted a rather clonal populations context. Diversity seemed to be increasing from 1980s to 2010s, probably as a result of the introduction of new varieties to African countries. More recently, high genetic relatedness of isolates from different cassava-growing regions in Kenya using rep-PCR has been reported, suggesting that despite the high prevalence of the disease, genetic diversity of *Xpm* is still low in some African countries (Chege et al., 2017).

Of contrast, *Xpm* diversity is higher in South America, in agreement with the hypothesis of the origin of the pathogen. Studies on *Xpm* populations isolated in several regions in Colombia showed no differentiated geographic structure, traces of pathogen migration, yearly changes of pathogen population structure in the same field and indications of plant host pressure on the pathogen (reviewed by (López Carrascal and Bernal, 2012)). Recent studies highlighted two contrasting scenarios. Most haplotypes of *Xpm* were not structured geographically in northern regions of Colombia when using Amplified fragment length polymorphisms (AFLPs) (Trujillo et al., 2014b) or Multilocus variable-number tandem repeat (VNTR) analysis (MLVA) (Rache et al., 2019), suggesting an important role of pathogen migration through contaminated propagative material. However, a center of origin of several *Xpm* haplotypes was discovered in one of the studied locations, which acts as a potential source of new founder pathogens (Trujillo et al., 2014b). In contrast, AFLP and MLVA-based analysis showed that *Xpm* populations of the eastern plains were structured geographically, but still showing some degree of genetic flow between distant regions. Authors consider that these contrasting scenarios reflect different agricultural practices, crop intensiveness and distribution of lands dedicated to the crop (Trujillo et al., 2014a).

At a larger geographical scale, 65 *Xpm* strains, mainly from Africa and South America were sequenced by Illumina (Bart et al., 2012). Phylogenies based on more than 12,000 chromosomal SNPs show that African, Colombian and some Brazilian strains share a common ancestor, in line with the idea of the introduction of the pathogen to Africa as a consequence of the slave trade and or Portuguese missions. Authors highlight that *Xpm* populations are evolving independently, but also show genetic flow between geographically distant places, which should be taken into consideration for surveillance and CBB control purposes.

Diversity assessment tools

Xpm diversity has been analyzed by different methods during the last three decades, including RFLP (Berthier et al., 1993; Restrepo et al., 2004; Restrepo & Verdier, 1997; Verdier et al., 1993; Verdier

et al., 2001; Verdier et al., 1994), Repetitive element palindromic PCR (rep-PCR) (Chege et al., 2017; Restrepo et al., 2000c), Random Amplified Polymorphic DNA (RAPDs) (Ogunjobi et al., 2007, 2006), and AFLPs (Gonzalez et al., 2002; Ogunjobi et al., 2010; Restrepo et al., 2000c; Trujillo et al., 2014b). Recent studies have shown that a slightly better discriminatory power is achieved using MLVA (Rache et al., 2019; Trujillo et al., 2014a). An MLVA scheme with 15 VNTRs amplified by multiplex-PCR was developed to assess the pathogen diversity at a local scale and this scheme was tested in the Northern Coast of Colombia. This method has several advantages : it does not require DNA extraction, it can be easily formatted for high-throughput settings, the amplification of up to four loci is parallelized in a unique PCR reaction, and outputs show high reproducibility and portability (Rache et al., 2019).

Diagnostic tools

Early detection and diagnostic tools relied on *in vitro* isolation of the pathogen coupled with phage typing, cassava reinoculation and/or serological assays on cassava extracts. With the popularization of more straightforward molecular methods, standard (Verdier et al., 1998) and nested PCR (Ojeda and Verdier, 2000), dot blots (Verdier & Mosquera, 1999) and an ELISA assay were developed (Verdier et al., 2001). Three PCR-based methods were reported for *Xpm* detection recently, a first one using an optimized version of a previous assay (Verdier et al., 1998) resulting in improved *Xpm* detection (Cerqueira-Melo et al., 2019). A second approach is based on a multiplexed nested PCR including a broadly conserved fragment of *Xpm* Transcription Activator-Like effector (TALE) genes and a semi-specific region of *rpoB*, resulting in a wider detection potential for *Xpm* strains (Bernal-Galeano et al., 2018). A third strategy consisted of a duplex PCR amplifying highly conserved chromosomal regions in *Xpm* and *Xc* to detect and distinguish both cassava pathogens proved to be highly selective and sensitive (Flores et al., 2019).

Disease phenotyping tools under controlled conditions

Generally, *Xpm* must first overcome plant defenses in the leaf apoplast, before accessing to the vessels and migrating towards the stem. Defense mechanisms in the mesophyll are different from those exerted by the plant in the leaf and stem vessels, adding barriers of a different nature that are surpassed only by successful pathogens (Kpémoua et al., 1996; Restrepo et al., 2000a; Wydra et al., 2007). This multilayer tissue-specific resistance might be the explanation for the quantitative and additive nature of the genetic factors reigning resistance in cassava, which increases the complexity of disease phenotyping. Moreover, the propagation of cassava through vegetative cuttings results in developmentally unsynchronized plants with variations in their physiology that could affect the outcomes of phenotyping methods (Mutka et al., 2016). Hence, disease phenotyping tools in cassava and their relationship to the real-world plant performance are not trivial subjects.

Bacterial virulence is usually assessed upon infiltration of leaves of adult 2- to 4-month old plants, through bacterial growth analysis, bacterial movement through the leaf, and symptom development (compilation of protocols in (Cohn et al., 2015)). Bacterial growth and movement within the host quantitatively reflect the ability to grow locally and to migrate through the xylem, while symptom development is a qualitative estimation of *Xpm* capacity to induce water soaking lesion, which might improve bacterial fitness. Symptom development can also be quantified by measuring lesion areas around a perforated hole or upon leaf clipping or leaf spraying (Pacumbaba, 1987; Verdier et al., 1994; Zandjanakou-Tachin et al., 2007). More recently, image-based phenotyping of *Xpm* with

luminescent reporters efficiently allowed to quantitatively track the disease in time and space (Mutka et al., 2016).

Leaf inoculation may be questionable for resistance phenotyping since discrepancies were observed between leaf lesion measurements and scoring upon stem inoculation (Muñoz-Bodnar et al., 2015; Restrepo et al., 2000b). However, others have successfully detected symptom differences when infiltrating leaves of susceptible and resistant cultivars at low densities (i.e., from 10^2 to 10^5 ufc/mL) (Flood et al., 1995; Wydra et al., 2007, 2004).

In the stem, the most discriminant methodology consists of inoculating one internode of the apical region by puncturing bacteria with a sharp tool and evaluating disease development over a 30-day span according to a standardized severity scale. Measurements from several replicates are mathematically integrated to obtain a dimensionless quantity (area under the disease progression curve, AUDPC) that can be comparable even among experimental sets with similar conditions. Resistance and susceptibility is established from scores (in a 0-5 scale, >4 is deemed as susceptible and <3 is deemed resistant), or by setting thresholds when performing AUDPC analyses, this latter being the preferred method to assess resistance (Jorge & Verdier, 2002; Restrepo, Duque, et al., 2000b).

To avoid the recurrent problem of space limitation when phenotyping cassava (plants are large), two studies recently reported the use of cassava plants grown *in-vitro*. A comparative resistance screening between plants grown in pots and *in vitro* showed that the latter did not wilt although some symptoms developed, yet indicating that the AUDPC methodology can be applied to assign resistance/susceptibility categories (Mbaringong et al., 2017). In another study Mora and collaborators (2019) developed methods to quantify AUDPC values and perform bacterial growth analysis, highlighting contrasted phenotypes between susceptible and resistant cassava varieties in terms of disease progression and bacterial growth (Mora et al., 2019). These technical advances, in combination with traditional and more recent tools open a new era in cassava phenotyping research.

Virulence mechanisms of the pathogen

Genomics of the pathogen

The advent of next generation sequencing technologies allowed the sequencing of the genome of 66 *Xpm* strains originating from Africa and South America mainly (Arrieta-Ortiz et al., 2013; Bart et al., 2012). Assemblies resulted into fragmented draft genomes that were used for genomic and phylogenetic analyses, highlighting two main clusters. Grouping of South American and African strains in the same cluster agrees with the hypothesis formulated years ago on the possible introduction of *Xpm* from America to Africa. The authors also identified the invariable occurrence of the well-known Pathogen-Associated Molecular Patterns (PAMPs) *fliC* and *ax21*, the presence of 14 to 22 *Xanthomonas* outer proteins (Xops) per strain, with a core set of nine of them, and reported that all strains harbor at least one TALE (Bart et al., 2012). The first manually annotated and high-quality draft genome for *Xpm* was obtained through 454 pyrosequencing technology of the Colombian strain CIO151 (Arrieta-Ortiz et al., 2013). Analysis of the genomic sequence confirmed the presence of fully functional pathogenicity mechanisms such as type II, III, IV and VI secretion systems; an exopolysaccharide production cluster; core and accessory type three effectors, and chemotaxis, type IV pili, flagella, siderophore biosynthesis, and putative polyketide synthesis genes.

This study also explored for the first time the possibility of using a MLVA scheme for *Xpm* diversity studies, and report on 16 potential VNTR loci. Third-generation long read sequencing has resulted in several high-quality non-fragmented xanthomonads genomes (Booher et al., 2015; Cox et al., 2017; Denancé et al., 2018; Gochez et al., 2018; Peng et al., 2016; Ruh et al., 2017; Showmaker et al., 2017; Tran et al., 2018; Wilkins et al., 2015). A first attempt to sequence *Xpm* with Pacific Biosciences technology highlighted a high error rate on reads precluding to obtain a high-quality genome (Bart et al., 2012). Regarding *Xc*, a draft genome of strain CFBP4642 (*a.k.a* NCPPB101, ICMP204, LMG673) isolated in Malawi in 1951, highlighted the presence of a canonical T3SS and some T3Es including TALEs (Bolot et al., 2013). Presently our team is producing high-quality reference genomes for both *Xpm* and *Xc* pathogens by using long-read sequencing technologies.

Secretion systems

The type I secretion system (T1SS) allows the active transport of hydrolases (e.g. proteases, phosphatases, glucanases, nucleases, lipases) and toxins from the cytoplasm to the extracellular matrix (reviewed by (Delepelaire, 2004)). The canonical type II secretion system (T2SS) allows the secretion of periplasmic enzymes (e.g. cellulases, pectin methylesterases, cellobiosidases, and polygalacturonases) to the extracellular matrix (reviewed by (Cianciotto and White, 2017)). The presence of canonical versions of these two systems (T1SS and T2SS) in *Xpm* was reported by Arrieta-Ortiz and coworkers, highlighting two slightly different clusters (*xps* and *xcs*) for the T2SS in CIO151 genome (Arrieta-Ortiz et al., 2013). Type IV secretion systems (T4SS) translocate proteins and DNA-protein complexes to surrounding eukaryotic host or other prokaryotic cells, playing an important role for competition against other Gram negative bacteria (reviewed by (Sgro et al., 2019)). Likewise, the type VI secretion system (T6SS) is a contractile apparatus that injects toxic effectors to accompanying eukaryotic and prokaryotic cells (reviewed by (Bayer-Santos et al., 2019)). *Xpm* carries at least one class of T4SS (Arrieta-Ortiz et al., 2013; Sgro et al., 2019), and one T6SS cluster encoded by fifteen genes in strain CIO151, but their functionality remains to be evidenced (Arrieta-Ortiz et al., 2013).

The type III secretion system (T3SS) is a translocation machinery that allows the injection of effector proteins into the host. Delivered effectors play a major role in counteracting the host defenses and hijacking the cellular mechanisms of the invaded cell (reviewed by (Timilsina et al., 2020)). In *Xpm*, the T3SS is encoded by 26 genes of the *hrp2* family cluster (Arrieta-Ortiz et al., 2013), and its role in pathogenesis has been demonstrated through the mutagenesis of *hrpX*, a key regulator of the T3SS (Medina et al., 2018).

Type 3 effectors

Xanthomonas T3Es can be classified in two major groups considering their molecular structure, function, and interactors/targets. The first group forms a heterogeneous group of so-called *Xanthomonas* outer proteins (Xops) i.e., effectors with a wide range of enzymatic activities, whose targets and associated physiological effects mainly take place in the host cytoplasm and their mechanism of action rely on protein-protein interactions. The second group is only composed by Transcription Activator-Like effectors (TALEs), which are modular proteins that share an unusual architecture combined with eukaryotic motifs that allow them to act as *bona fide* transcription factors inside the host nucleus. Most Xops are involved in disturbing plant defense through alteration of PAMP-DAMP triggered immunity (PTI/DTI) or effector-triggered immunity (ETI)

pathways, while TALEs act as TFs that re-shape the host cellular metabolism through the activation of susceptibility (*S*) genes to create better niches for bacteria (Timilsina et al., 2020).

Xpm strains harbor from 13 to 23 Xop effectors, of which 9 are conserved among all the reported *Xpm* genomes and termed as *Xpm* core effectors: HpaA, HrpF, XopE1, XopV, Hpa2, XopAK, XopL, XopN, and XopAE (a.k.a HpaF), the latter five being almost completely monomorphic among the surveyed strains (Bart et al., 2012). AvrBs2, XopAO1, XopZ and XopX are important for virulence; strains mutated in these T3Es do not multiply properly *in planta*, and vascular colonization and/or symptom formation is impacted for some of them. XopK plays a dual role by increasing the developmental rate of symptoms but limiting the spread of the pathogen, while XopN and XopQ seem to have a redundant pathogenicity function. Expression in heterologous systems showed that XopR, AvrBs2 and XopAO1 interfere with PTI, while XopAO1 and XopE4 interfere with ETI (Medina et al., 2018; Mutka et al., 2016).

The crucial role of TALEs in *Xpm* pathogenicity was known for long and finally reported recently (Castiblanco et al., 2013; Cohn et al., 2016, 2014). Several diagnostics and diversity tools were unintentionally based on TALEs detection, the results highlighting their high conservation within *Xpm* (reviewed by (Verdier et al., 2004)). The screening of 65 strains through RFLP showed that all contained between one and five TALEs with 12.5, 13.5, 14.5, 19.5, 20.5 or 21.5 repeats, and most of them were plasmid-borne (Bart et al., 2012). However, our understanding of *Xpm* TALE diversity is rather poor since limited to only seven genes from three strains (Castiblanco et al., 2013; Cohn et al., 2016, 2014). Since these effectors are the main axis of this thesis, a dedicated introductory section addresses the TALEs and what we know about them in cassava. For detailed information, see section *Xpm* TALEs and their role in cassava.

Toxins

Upon bacterial entry into the leaves, the first symptoms to appear are water-soaked angular spots. Generally, discrete spots begin to coalesce, and eventually, surrounding, and distal areas start to blight. This characteristic blight phenotype may be due to the diffusion into the laminar tissues around the infection foci of a small molecule that acts as a toxin. Perreaux and collaborators isolated from *Xpm* cultures a small organic acid, the 3-methyl thiopropionic acid, which induces the blight symptom when infiltrated alone into cassava leaves (Perreaux et al., 1982). Bacterial metabolism leads to a transamination coupled to a decarboxilation of methionine, resulting in the formation of 3-methyl thiopropionic acid (Ewbank and Maraite, 1990). Concentrations of this toxin rise up in leaves along with bacterial multiplication and reach a maximum just before the onset of the blight symptom (Perreaux et al., 1986). The exact mechanism of action of this small acid is not known. However evidence for the role of this compound as a toxin is debated, basically because concentrations of free methionine and the potential toxin in leaves are very low, and *in vitro* assays could have been biased by the acid nature of the compound (Cooper et al., 2001).

Signaling and metabolic routes potentially involved in regulation of pathogenicity

Recently, an RNAseq of a Quorum Sensing-insensitive mutant of *Xpm* vs. wild type grown *in vitro* was reported (Botero et al., 2020). The authors concluded that the QS system controls several subsequent signaling routes including a number of phosphorylation sensor and transduction pathways, some of which share a c-di-GMP phosphodiesterase activity (HD-GYP domain). Metabolic

routes that were affected by QS included the already reported xanthan biosynthesis route and the newly reported NAD(P)⁺ balance, and fatty acid elongation, which should be further studied.

Host genetics and interaction

Tools for *Xpm*-Cassava interaction research

Development of adapted durable resistance must consider the local pathogen population structure and the host response variability. Pathotyping schemes can condense this complex interaction data to cluster pathogens according to virulence (Restrepo, Duque, et al., 2000b; Trujillo, Ochoa, et al., 2014; Verdier et al., 2004) and classify cassava genotypes based on susceptibility (Wydra et al., 2004). Inoculation of *in vitro* plants (Mora et al., 2019) and image-based phenotyping methods (Mutka et al., 2016; Velez et al., 2020) are two important advances that will greatly facilitate the study of this pathosystem by leveraging high throughput applications, and evidencing the plant response in a more comprehensive way.

Cassava genomic resources include the read collection from 58 Illumina-sequenced cassava or cassava relatives accessions (Bredeson et al., 2016), four high-quality sequenced genomes from cultivars AM560-2 (an MCOL-1505-derived partially inbred cultivar) (Bredeson et al., 2016; Prochnik et al., 2012), KU50 (Wang et al., 2014), TME3, and the highly CBB-susceptible cultivar TMS60444 which is amenable to genetic transformation (Kuon et al., 2019). Also it includes at least four high density genetic maps (International Cassava Genetic Map Consortium (ICGMC), 2015; I. Rabbi et al., 2014; I. Y. Rabbi et al., 2014; Soto et al., 2015), one of them with an integrated physical map that includes immunity-related genes (Soto et al., 2015); and Bacterial Artificial Chromosome (BAC) libraries from cultivars TMS30001 and MECU72 (Tomkins et al., 2004). Transcriptomic resources for cassava challenged by *Xpm* include expressed sequenced tags (ESTs), simple sequence repeats (SSR) (Lopez et al., 2004; López et al., 2007), a microarray (López Carrascal et al., 2005), transcript-derived fragments (TDFs) (Santaella et al., 2004), and differentially expressed genes (DEGs) catalogs (Cohn et al., 2014; Gómez-Cano et al., 2019; Muñoz-Bodnar et al., 2014). The website www.cassavagenome.org provides tools incorporating the data from some of the above-described resources such as transcriptomic and SSR data displayed on the AM560-2 genome. As mentioned earlier, daTALbase allows the research of potential virulence targets of TALEs through prediction of EBE sequences in the host promoterome, and cross-referencing to host transcriptomics and polymorphism data (Pérez-Quintero et al., 2017).

Host susceptibility to *Xpm*

The susceptibility (*S*) gene *MeSWEET10a* is a Clade III member of the well-known SWEET family of phloem-loading efflux carriers (reviewed by (Chen, 2014)), *MeSWEET10a* is transcriptionally activated by the TALE TAL20_{xam668}, and the resulting protein was found to export glucose and sucrose using *Xenopus* oocytes (Cohn et al., 2014). Increased susceptibility is probably related to increase of glucose accumulation in the apoplast, where it supports bacterial growth (Cohn et al., 2014), and/or by increasing osmotic water influx into the intercellular spaces that facilitate bacterial movement (El Kasmi et al., 2018; Schwartz et al., 2017). Cassava has 27 putative *SWEET* genes that are scattered in the genome (except for three clusters with 5, 3 and 2 *SWEET* genes), and account for at least five clade III members that could potentially also act as *S* genes, as shown in rice (Streubel et al., 2013).

Mora and coworkers compared the expression of *MeSWEET10a* during a compatible and an incompatible interaction, highlighting its upregulation in the susceptible cultivar TMS60444 but not in the resistant cultivar CM6438-14 fifty hours after inoculation of a strain carrying *TAL20_{Xam668}*. Inspection of the *MeSWEET10a* promoter in both varieties showed conservation of the EBE, suggesting the immune response to stop susceptibility pathways at early stages of the infection (Mora et al., 2019).

Structural and cellular features of cassava resistance against *Xpm*

A mechanism associated with anti-*Xpm* defense in cassava is the occlusion of adaxial stomata with wax, as a way to reduce pathogen entry into the leaf. Interestingly, the number of occluded stomata seems higher in resistant varieties (Cooper et al., 2001; Zinsou et al., 2006). Once in the substomatal spaces, *Xpm* reaches the mesophyll which colonization is accompanied by the formation of a fibrillar matrix made of bacterial exopolysaccharides and loosening of plant cell walls through enzymatic lysis. These modifications allow the pathogen to become vascular by accessing the xylem vessels and migrating to other parts of the plant (Boher et al., 1997, 1995; Kemp et al., 2004). Little is known about the resistant reactions taking place in the leaf mesophyll, but it has been suggested that these responses may be overcome by *Xpm* (Kpémoua et al., 1996; Verdier et al., 1994). Cassava resistance against *Xpm* is mediated by several cellular and molecular mechanisms, and the promptness of this response is the key difference between a susceptible and a resistant variety (Kpémoua et al., 1996). Resistance in the stem is characterized by the initial secretion of bactericidal phenolic compounds into intercellular spaces and vessel lumen, followed by a deposition of lignin, suberin and callose in the paramural spaces. Later, pectic-rich tyloses are formed to occlude vessels and to secrete more phenolic compounds, while bacteria are trapped in lysis pockets through lignification and suberization (Boher et al., 1995). Moreover, resistant genotypes are able to metabolically cope up better with infection, with the maintenance of average stomatal resistance, water potential, and proline levels, factors that are markedly altered in compatible interactions (Restrepo Rubio et al., 2017).

Mapping resistance to *Xpm*

Due to the quantitative nature of CBB resistance, and the complex interaction between the host genetic background and the environment, the search for resistance sources against *Xpm* in cassava is mainly restricted to the identification of quantitative trait loci (QTLs) (Jorge et al., 2001, 2000; López Carrascal et al., 2007; Soto-Sedano et al., 2017; Tappiban et al., 2018; Wydra et al., 2004). Identified QTLs were strain-specific in all cases, explaining up to 61% of the resistance variance (reviewed by (López & Bernal, 2012)). Recently, thanks to a high density cassava genetic map (Soto et al., 2015), Soto-Sedano and coworkers associated five novel resistance QTLs, accounting for 16% to 22% of the variance phenotype and which colocalized with 29 genes potentially involved in defense (Soto-Sedano et al., 2017). Furthermore, three out of these QTLs showed to be significantly influenced by environmental conditions, especially humidity. More recently, by using an SSR-based genetic map derived from the F1 offspring of an Asian cassava cultivar cross, 10 QTLs were associated to CBB resistance, explaining up to 26.5 % of the genotypic variance, and five colocalized genes were shown to be differentially upregulated in resistant genotypes (Tappiban et al., 2018). Table 5 summarizes all the cassava QTLs (more than 100) reported to be involved in CBB resistant.

Table 5. Quantitative Trait Loci (QTLs) associated to CBB resistance. QTL name: name of the QTL associated to CBB resistance. ExpVar: percentage of phenotypic variance explained by a given QTL. Linkage group: linkage group where a given QTL can be found. Linkage groups depend on the map used for associations. Strain: strain used to inoculate the assessed plants (in most of the cases, QTLs are strain-specific). Natural inoculum: the study was performed in the field and the quantified disease resulted from infections with naturally occurring strains. Undefined strain: the study states that only one strain was used to inoculate the assessed plants, but authors do not provide the strain identifier. Technique: disease quantitation technique used to associate phenotypes to markers. AUDPC: Area Under the Disease Progression Curve, 100-DI: percentage of diseased plants (100-disease incidence), MDR: Mean Disease Rating. Marker: peak marker associated to a given QTL. Map reference: cassava map used to associate QTLs. The maps developed by Fregene and coworkers (Fregene et al., 1997) and Soto and coworkers (Soto et al., 2015) derived from the female parental TMS 30572 and the male parental CM 2177–2; the map developed by Wydra and coworkers (Wydra et al., 2004) derived from five F1 male individuals (CM7857-4, CM7857-10, CM7857-51, CM7857-77, and CM7857-115) and the female recurrent parent TMS30572; the map developed by Tappiban and coworkers (Tappiban et al., 2018) derived from the parentals Huay Bong 60 (HB60) and Hanatee (HN). Study reference: the reference to the study where QTLs were found.

QTL name	ExpVar (%)	Strain	Technique	Linkage group	Marker	Map reference	Study reference
XM1	11	CIO-84	AUDPC	D	GY222	(Fregene et al., 1997)	(Jorge et al., 2000)
XM2	10			L	GY141		
XM9	10.7			C	GY54		
XM10	11.2			G	GY93		
XM3	13	CIO-1		N	AD4b		
XM4	9			N	rGY145		
XM11	11.6			C	rGY18		
XM11	15.5			C	rGY26		
XM5	13	CIO-136		X	CBB1		
XM12	12.4			C	rNI1.C2		
XM6	10	CIO-295		B	rGY147		
XM1	20	ORST X-27		D	GY181		
XM7	12			D	GY179		
XM8	12			D	O19		
XM2	19			L	GY141		
XMF3	6.1	Natural inoculum	100-DI	F	CDY107	(Fregene et al., 1997)	(Jorge et al., 2001)
XMF3	3.6			F	GY37		
XMF3	5.2			F	GY186		
XMF5	7.7			M	W19b		
XMF5	8.1			M	GY105		
XMF1	4.3			D	GY219		
XM8	9.3			D	O19		
XMF2	8.7			E	rI4a		

XMF2	13.1			E	CDY123a		
XMF3	8			F	CDY107		
XMF4	10.8			I	rAC1		
XMF5	10.2			M	nrGY67		
XMF5	11.6			M	rGY87		
XMF5	8.9			M	W19b		
XMF5	9.6			M	GY105		
XMF1	6.5			D	GY219		
XM8	9.1			D	O19		
XMF2	9.9			E	rI4a		
XMF3	18.2			F	CDY107		
XMF3	6			F	GY37		
XMF3	7.2			F	GY186		
XMF5	12.9			M	nrGY67		
XMF5	11.7			M	rGY87		
XMF5	9.8			M	W19b		
XMF5	9.6			M	GY105		
XMF6	4.8			D	GY219		
XMF7	11.2			E	rF19b		
XMF3	7.1			F	CDY107		
XMF3	6.5			F	GY37		
XMF3	5.6			F	GY194		
XMF3	6.9			F	GY186		
XMF5	5.8			M	W19b		
XMF5	6.1			M	GY105		
XMF6	10.8			B	rGY72		
XMF6	3.2			D	GY219		
XMF2	9.9			E	rI4a		
XMF2	7.2		MDR	E	GY118		
XMF2	13.3			E	CDY123a		
XMF7	5.3			E	GY205		
XMF3	10.7			F	CDY107		
XMF3	4.3			F	GY37		
XMF3	5.4			F	GY194		
XMF3	4.8			F	GY186		
XMF5	12			M	nrGY67		
XMF5	9.5			M	rGY87		
XMF5	6.7			M	W19b		
XMF5	9.1			M	GY105		
XMF6	10.6			B	rGY72		
XMF6	4.6			D	GY219		

XMF2	11.1			E	rI4a		
XMF2	10.6			E	GY118		
XMF2	9.5			E	CDY123a		
XMF7	6.4			E	GY205		
XMF3	16.4			F	CDY107		
XMF3	8.3			F	GY37		
XMF3	6.4			F	GY194		
XMF3	8.4			F	GY186		
XMF5	11.1			M	nrGY67		
XMF5	10.6			M	rGY87		
XMF5	10.4			M	W19b		
XMF5	9.8			M	GY105		
None	21.4	CIO121		A	rGY75	(Fregene et al., 1997)	(López Carrascal et al., 2007)
None	61.6	CIO151		U	B39P22		
QTL1	16	GSPB2507		O	SSRY6	(Wydra et al., 2004)	(Wydra et al., 2004)
QTL1	18	GSPB2511		O	SSRY6		
QTL2	19.7			O	SSRY170		
QTL3	18.1	Uganda 12		N	SSRY55		
QTL4	23.2			C	SSRY93		
QTL5	21.8	GSPB2507		A	GY12		
QTL5	33.3	Uganda 12		A	GY12		
None	17.4			Not Linked	SSRY17		
None	12.8	GSPB2507		Not Linked	SSRY104		
QTL6	19.6	GSPB2511	AUDPC	No name assigned	SSRY83		
None	19.3	Uganda 12		Not Linked	SSRY104		
QTL6	18.2	GSPB2506		No name assigned	SSRY83		
None	23.4			Not Linked	SSRY104		
None	18.4	GSPB2507		Not Linked	SSRY7		
None	21.8	GSPB2511		Not Linked	SSRY84		
None	22.6	GSPB2506		Not Linked	SSRY157		
QAR681D-14	18.1	UA681		LG14	MB_9956		

QLV318RD-19	17.3	UA318		LG19	MB_38006	(Soto et al., 2015)	(Soto-Sedano et al., 2017)
QLV681RD-6	22.1	UA681		LG6	MB_50599		
QGH318-19	18.8	UA318		LG19	MB_49863		
QGH681-2.2	15.8	UA681		LG2.2	MB_23160		
CBB14_7dai_1	66	Undefined strain	MDR	LG3	SSYR180	(Tappiban et al., 2018)	(Tappiban et al., 2018)
CBB14_7dai_2	8.4			LG7	CA172		
CBB14_7dai_3	15.1			LG12	SSRY192		
CBB14_7dai_3	15.1			LG12	EME489		
CBB14_7dai_4	20.9			LG13	SSRY82		
CBB14_7dai_4	20.9			LG13	s10963n39		
CBB14_7dai_5	16			LG14	s08265n64		
CBB14_7dai_5	16			LG14	MeES485		
CBB14_10dai_1	21.3			LG7	CA567		
CBB14_10dai_1	21.3			LG7	NS898		
CBB14_10dai_2	13.8			LG12	s11341n43		
CBB14_10dai_3	32.4			LG15	NS982		
CBB14_10dai_3	32.4			LG15	NS847		
CBB14_10dai_4	25.7			LG17	CA358		
CBB14_12dai	26.5			LG5	NS1021		

Genetic and molecular aspects of cassava defense against *Xpm* and *Xc*

Most of our knowledge about the molecular aspects of the cassava-*Xpm* interaction comes from studies of resistant varieties (see Figure 4). However, challenging susceptible genotypes with *Xpm* also induces defense-related genes, but expression is delayed when compared to resistant varieties (López Carrascal et al., 2005; Santaella et al., 2004). Several studies have identified Pathogenesis-Related (PR) proteins (Li et al., 2017; Román et al., 2014; Yoodee et al., 2018) and various microRNA families potentially involved in plant defense (Pérez-Quintero et al., 2012). The role of Transcription Factor (TF) families in response to *Xpm* has been addressed by coexpression network analyses and functional studies by TF families. Specific cassava Heat Stress TFs (MeHSfs) (Wei et al., 2017, 2018b), related to ABI3 AND VP1 (MeRAVs) (Wei et al., 2018a), Nuclear Factor Y (MeNF-Ys) (He et al., 2019), Whirly (MeWHYs) (W. Liu et al., 2018), NAM/ATAF/CUC2 (NACs) (Gómez-Cano et al., 2019), basic leucine zipper (bZIPs) (Gómez-Cano et al., 2019; Li et al., 2017), WRKYs (Gómez-Cano et al., 2019; W. Liu et al., 2018; Wei et al., 2017; Yan et al., 2017; Yoodee et al., 2018) and TB1/CIN/PCF (TCPs) TFs (Gómez-Cano et al., 2019) were shown to play important roles in cassava immunity against *Xpm* through activation of defenses. For example, *MeWHY1/2/3* (through interaction with MeWRKY75), MeNF-YA1/3, MeNF-YB11/16, and MeNF-YC11/12 are crucial for upregulation of defenses against *Xpm* (He et al., 2019; W. Liu et al., 2018).

Salicylic acid (SA), calcium signaling, and melatonin synthesis also play important roles in anti-*Xpm* defense. The upregulation of *MeHsf3* in response to *Xpm* infection triggers the activation of the SA pathway through transcriptional activation of *MeEDS1*, also triggering defense by activating the non-SA-related *MePR4* gene (coding for a defense-related protein containing SCP domain) (Wei et

al., 2018b). Several Calcineurin B-like proteins (CBLs) and CBL-interacting protein kinases (MeCIPKs) which are respectively transducers of CIPKs signaling and sensors of calcium signaling, have been found upregulated during the interaction. Among them, MeCIPK23 physically interacts with MeCBL1 and MeCBL9 to modulate defense against the pathogen, potentially as a response to calcium signaling (Yan et al., 2018). During infection *Xpm* also upregulates the melatonin biosynthesis pathway (MBP), leading to melatonin accumulation, activation of defense- and ROS-related genes, and increased callose deposition (Wei et al., 2016). MeRAV1 and MeRAV2 TFs directly activate three genes of the MBP (*MeTDC2*, *MeT5H*, and *MeASMT1*), while MeWRKY79 and MeHsf20 increase the transcription rate of *MeASMT2* (a second N-acetylserotonin O-methyltransferase gene involved in the MBP), all resulting in increased defense responses against *Xpm* (Wei et al., 2018a, 2017).

Autophagy regulation is altered during *Xpm* infection, which significantly impacts immunity. Four members of the cassava Autophagy-Related Protein 8 (*MeATG8*) family, including *MeATG8f* which colocalizes with a resistance QTL (Tappiban et al., 2018) and which product was reported to interact with MeWRKY20 to induce defense against *Xpm* (Yan et al., 2017). Furthermore, *MeATG8b* and *MeATG8e* have antagonistic roles against the two cassava Glyceraldehyde-3-phosphate dehydrogenases MeGAPC4 and MeGAPC6, whose downregulation results in higher autophagic activity, higher H₂O₂ levels, and increased callose deposition (Zeng et al., 2018). Little attention has been paid to the cassava-*Xc* interaction, but analysis of an incompatible interaction showed that, unlike for *Xpm*, Phenylalanine Ammonia Lyase (PAL) transcription and activity, and cell wall-bound peroxidase activity are markedly increased in response to the pathogen (Pereira et al., 2000, 1999).

Pathogen perception relies on Receptor-Like Kinases (RLKs), Receptor-Like Proteins (RLPs), and Nucleotide Binding Site-Leucine-Rich Repeat (NBS-LRRs) proteins. Cassava contains at least 253 RLK- (Soto et al., 2015) and 327 NBS-LRR-encoding genes. Most *NBS-LRRs* are grouped in 39 clusters, with two superclusters with 43 and 19 of the genes located on chromosomes 16 and 17, respectively (Lozano et al., 2015). Mapping for resistance QTLs allowed the identification of the two resistance gene candidates *RXam1* and *RXam2*, respectively coding for a RLK and an NBS-LRR (López et al., 2003). *RXam1* shows similarity to the anti-*Xanthomonas oryzae* pv. *oryzae* rice resistance gene *Xa21*, and was recently shown to be involved in defense against *Xpm* strain CIO136 (Díaz Tatis et al., 2018). *RXam2* is a typical non-TIR NB-LRR located in a QTL explaining 61% of the resistance variance to *Xpm* strain CIO151. Recently, it was demonstrated that this gene confers partial and broad-spectrum resistance to different *Xpm* strains (Díaz-Tatis et al., *In prep*).

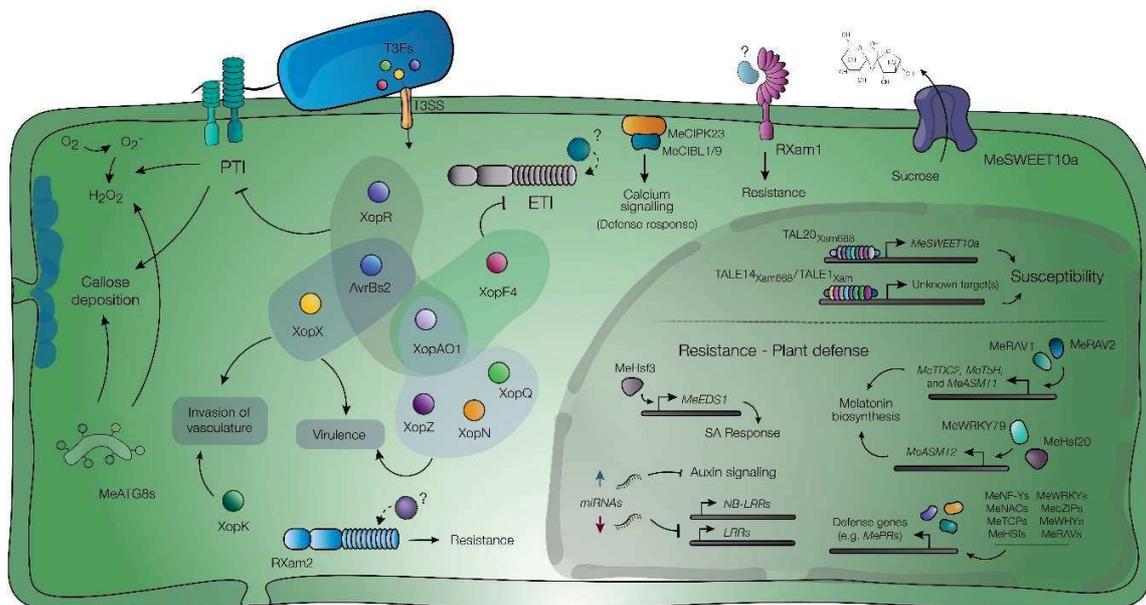


Figure 4 Roles of Xops, host determinants of susceptibility and defense responses in the cassava-*Xpm* interaction.

Xop effectors (colored circles) are injected by *Xpm* into the plant cytoplasm, where they interfere with host cellular processes to repress. T3Es are grouped (shading) according to their predicted involvement in PAMP-triggered immunity (PTI), effector-triggered immunity (ETI), and/or other virulence-related phenotypes (see Type 3 effectors section). The upper section of the depicted nucleus outlines the contribution of TAL effectors to cassava susceptibility, in which transcriptional activation of the *S* gene *MeSWEET10a* leads to an over-accumulation of sugar transporters (purple, anchored to the membrane) and sugar export (see Type 3 effectors section). The RXam1 RLK (lilac, in the membrane) and the RXam2 NBS-LRR (blue, in the cytoplasm) are associated to defense response triggered against *Xpm*, but their matching elicitors are unknown. Calcium signaling and autophagy elements are represented in the cytoplasm as part of the anti-*Xpm* defense mechanisms. The roles of miRNA and transcription factors are schematized in the nucleus (see Genetic and molecular aspects of cassava defense against *Xpm* and *Xc* section). H₂O₂ and O₂⁻ represent reactive oxygen species (ROS) accumulation. SA: salicylic acid.

Resistance engineering in cassava

Stable transformation of cassava is mainly achieved through bacterial-mediated transformation of friable embryogenic calli and many transgenic cassava lines are reported in the literature (reviewed by (Chavarriga Aguirre et al., 2016). However, the incipient efforts to engineer CBB resistance reflect the limited information on genetic determinants for resistance to *Xpm*. Stable transformants of the cassava cultivar 60444 overexpressing and silencing *RXam1* were generated to study the function of this resistance gene candidate (Chavarriga Aguirre et al., 2016; Díaz Tatis et al., 2018), as well as overexpressing the pepper *R* gene *Bs2*, which did not allow resistance against *Xpm* (Díaz-Tatis et al., 2019).

Since TALEs recognize *S* genes through interaction with DNA, genome editing is a nice tool to disturb or favor the TAL-DNA interaction and induce resistance or prevent susceptibility (reviewed by (Schornack et al., 2013). This approach was applied recently to help rice against *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) by editing five EBEs in the promoter of three *SWEET* genes targeted by *Xoo* in elite varieties, resulting in a robust and broad-spectrum resistance in the field (Oliva et al., 2019). This strategy could also be applied in cassava to engineer resistance against *Xpm*, given the requirement of *MeSWEET10a* for successful disease development.

Transcription Activator-Like effectors (TALEs)

As TALEs are the main axis of this thesis, this introductory section addresses the general characteristics of this effector family, their known roles in the *Xanthomonas* pathosystems, and a data compilation of what we know about TALEs in *Xpm* and their molecular role in cassava.

TALE overview

TALE architecture and mechanistic function

TALEs have a particular structure evolved to selectively bind DNA promoters in the host and recruit the RNA polymerase complex in order to initiate transcription of downstream genes (Boch and Bonas, 2010; Bogdanove et al., 2010). The protein structure can be divided in three different regions: N-terminal region, central repeat region and C-terminal region. The N-terminal region contains the signal for T3SS-mediated translocation (Rossier et al., 1999; Szurek et al., 2002), presumably within the first 63 aminoacids (Schreiber et al., 2015), and the sequence preceding the central repeat region contains cryptic repeats, whose aminoacidic identities differ from the central repeats, but which are predicted to have structural similarities and to engage in the DNA interaction (generally with a thymine) (Boch and Bonas, 2010; Bogdanove et al., 2010; Doyle et al., 2013; Schreiber and Bonas, 2014). The C-terminal region harbors three different functional elements: a transcription factor binding (TFB) region, one or several nuclear localization signals (NLS), and a carboxyl-terminal transcription activation domain (AD). The TFB region is a 135 to 145 aa domain that directly interacts with transcription factor IIA subunits, which is crucial for transcriptional induction of the downstream genes (Hui et al., 2019a; Yuan et al., 2016). The nuclear localization signals allow the interaction of the TALE protein with nucleocytoplasmic transporters (importins), which transfer the TALE from the cytoplasm to the nucleus (Hui et al., 2019b; Szurek et al., 2001; Yang and Gabriel, 1995). The carboxy-terminal transcription activation domain (AD) is a 38-aminoacid domain characterized by several acidic amino-acid residues interspersed with bulky hydrophobic residues, which is characteristic of eukaryotic acidic transcription activation domains (Szurek et al., 2001; Zhu et al., 1998).

The central region contains from one to 33 modular repeats, with a high pairwise nucleotide identity (ranging from 60% to 100%, but usually over 80%) that usually code for 33 to 35 amino acids per repeat (Deng et al., 2014; Schandry et al., 2018). The amino acid sequence of the repeats varies primarily in the identity of residues at position 12 and 13 in each repeat, which have been designated as “repeat variable diresidue” (RVD). The last repeat of the central region is characterized by being truncated at the 20th amino acid, but still containing the RVD. Each repeat forms a left-handed, two-helix bundle with the hypervariable residues located at the end of the loop that connects the two helices, which allows the exposition of the RVD to DNA bases. The set of these

modular repeats folds in a right-handed superhelix that gets accommodated along the major groove of DNA and allows the direct interaction between RVDs and nucleotides in a specific manner (D. Deng et al., 2012; Mak et al., 2012, 2013). The interaction code for RVD-DNA base was cracked in 2009 (Boch et al., 2009; Moscou and Bogdanove, 2009), showing that the RVD sequence governed the DNA sequence recognition.

From a mechanistic standpoint, TALE proteins have two different conformational stages that allow the search and recognition of their target. TALEs can associate with the DNA molecules to start their search through non-specific DNA-protein interactions mediated by the N-terminal regions. In the search conformation, the TALE protein moves in a one dimensional-sliding along the molecule for a rapid scan, with a loose central repeat region super helical structure that avoids strong interactions between the repeats and the phosphate backbone. Steric and electrostatic clashes during the sliding would cause destabilization of the complex, but TALE proteins are able to detach the DNA helix and shortly reattach to it (hopping). On the other hand, a proper target sequence would allow energetically favorable interactions between the superhelix and the DNA bases, leading to a contraction of the repeat region, the interaction of the RVDs with the DNA bases, and an arrest of the advance of the TALE on the DNA helix. The target recognition phase is characterized by an extended time onto the given sequence, which would allow the rest of the protein domains to take action (Cuculis et al., 2015; Wicky et al., 2013). The target sequences for TALEs are termed Effector Binding Elements (EBEs) (Antony et al., 2010), which are usually located on the promoter region of the targeted gene, near the TATA boxes. After recognition of the EBE, the domains in the C-terminal region interact with host transcription factors, like TFIIA γ subunits, and the RNA polymerase II to induce transcription of the downstream genes (Yuan et al., 2016).

Artificial or designer TALEs

The modularity nature of TALEs has leveraged the development of tools to design and construct artificial TALEs (arTALEs or dTALEs) through the sequential combination of the typical domains (Bogdanove and Voytas, 2011; Booher and Bogdanove, 2014; Briggs et al., 2012; Cermak et al., 2011; Reyon et al., 2012; Sanjana et al., 2012; Schmid-Burgk et al., 2013; Wang et al., 2016; Yang et al., 2013). These constructs have several biotechnological uses in targeted genome editing when combined with nuclease domains as FokI (Han et al., 2020; Sanjana et al., 2012; Streubel et al., 2013). In phytopathology, arTALEs/dTALEs have been mainly used to validate the roles of the TALE-targeted genes, by creating TALEs that target alternative EBEs in the promoter sequence of a candidate gene. When the observed phenotypes are equivalent after the transcriptional activation of one candidate gene with two or more arTALEs/dTALEs, then it could be said that the candidate, and not the alternative targets of the natural TALEs, is responsible for the phenotype.

Evolution of TALEs

TALE evolution is a subject that is not well understood yet, however, diversity studies and the more recent experimental evolution approaches have shed some light about this subject. At a genomic level, the distribution and number of members of this gene family in a genome seems to depend on the pathovar; for example, strains from pathovars containing large TALomes tend to carry TALE genes exclusively in their chromosomes (Schandry et al., 2018). On the other hand, the wide distribution of plasmids carrying TALEs among *Xanthomonas* species reflects the importance of the horizontal gene transfer (HGT) in TALE evolution. In this regard, it was shown that TALE gene

acquisition and HGT events involving these genes is clearly associated to the adaptation of *Xanthomonas citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, and *X. phaseoli* pv. *phaseoli* to common bean. Some insertion events have led to insertion of TALE genes into the bacterial chromosome of these strains, but there has not been a related expansion of the gene family into the chromosome. By considering this feature, it is tempting to hypothesize that larger chromosomal TALomes are an indicator of the coevolution time from the acquisition of TALEs (Ruh et al., 2017).

The associated elements found in the flanking sequences of TALEs indicate that dispersion of these genes in the genomic context is driven mainly by two mechanisms: transposons and site-specific recombinases. In the first group, which is representative of most of the TALE genes from *X. citri*, *X. axonopodis*, *X. campestris* pv. *vesicatoria*, *X. campestris*, *X. phaseoli*, and Asian strains from *X. oryzae* pv. *oryzae*, TALEs are flanked by inverted repeats, which act as substrates for Tn3-like systems (like *TnXax1*), to form mobile insertion cassettes. The abundance of these cassettes, the presence of remaining uncoupled repetitive elements, and the evidenced chromosomal insertion and/or duplication of long plasmid-borne regions harboring TALEs flanked by repetitive regions, indicate that transposition of these cassettes is an important driver of dispersion and diversification (Ferreira et al., 2015; Ruh et al., 2017). The second group comprises only TALEs from *Xanthomonas oryzae* pv. *oryzicola*, which are not flanked by transposon-related inverted repeats, but highly conserved spacer regions with putative hairpin structures. The structure of the chromosomal regions containing TALE genes in this group is typical from integrons, where genes flanked by conserved sequences with secondary structures are sequentially integrated in the locus, potentially through a site-specific recombinase (Erkes et al., 2017).

Some of the mechanisms that have been proposed to affect TALE gene evolution are point mutations, and deletion/duplications of repeats. Point mutations can create non-synonym changes on the RVD residues and then affect the DNA specificity of the TALE. Deletion/duplication of repeats are thought to arise from replication errors due to the high degree of conservation of the repeats that could cause loop formation into the replication fork (Erkes et al., 2017), or by slipped strand mispairing (Schandry et al., 2016). Recombination has been evidenced as one of the most common pathways on TALE gene evolution (Denancé et al., 2018; Erkes et al., 2017; Ruh et al., 2017). Recombination of the N- and C-terminal-coding sequences is commonly observed in *X. citri* pv. *fuscans*, *X. citri* pv. *aurantifolii*, *X. phaseoli* pv. *phaseoli*, *X. phaseoli* pv. *manihotis*, and *X. campestris* TALEs, which have allowed to propose that in the latter case, TALEs come from only two ancestral genes (Denancé et al., 2018; Ruh et al., 2017). The analysis of repeats and their position in the TALE gene has led to hypothesize that recombinational events often implicated a breakpoint in the N- or C-terminal-encoding region and another one in the central repeat region, being the first and last repeats less prone to recombining (Erkes et al., 2017). As a special case of recombination, segmental gene conversion, where an existing repeat (or a gene segment including more repeats) of the same gene, other gene or a pseudogene is recombined into the repeat region, has been proposed as an explanation for observations on *Ralstonia solanacearum* TALEs (Schandry et al., 2016).

Studies on the pathosystems comprised by *X. citri* strains and several citrus hosts have revealed that the pathogen complex has evolved different TALE alleles to maximize the binding affinity to different alleles of key target genes, leading to some degree of host specialization. Approaches in experimental evolution of this pathogen complex showed that TALEs indeed evolve by recombination of repeats present in other TALE genes (segmental gene conversion), by deletion of

complete repeats, and point mutations causing non-synonymous mutations in the RVDs, while duplication was not observed (Teper and Wang, 2021). Although authors claim that mismatch positions between the TALE protein and the EBE are not a determinant factor of adaptation, it is apparent that all the mutational/recombinational events that led to novel functional variants reverted the mismatches represented by the N-terminal-proximal repeats more than the distal ones.

Pathogenicity functions of TALEs

TALE targets and consequences of interactions

Transcriptional induction mediated by TALE family members plays a significant role for pathogenesis and virulence in several *Xanthomonas* pathosystems. This biological activity can result in three possible scenarios: i) the upregulation of a gene (susceptibility or S gene) whose activity results advantageous for bacteria and increases the susceptibility of the plant to the disease; ii) the upregulation of a gene whose activity results disadvantageous for bacteria and confers resistance to the plant; iii) the upregulation of a gene whose activity does not affect the bacterial fitness, nor the susceptibility to disease. In the following subsections, examples of the first and the second scenarios are exhaustively documented.

Major TALEs and susceptibility genes

Disruption or knocking out of some TALEs significantly affects the virulence of the pathogen or can even abolish pathogenicity. In those cases, the TALEs involved in the compatible interaction are termed major TALEs, and the TALE-targeted gene is known as a susceptibility (S) gene (Antony et al., 2010; Cernadas et al., 2014; Cohn et al., 2014; Cox et al., 2017; Hu et al., 2014; Kay et al., 2007; Perez-Quintero and Szurek, 2019; Schwartz et al., 2017; Sugio et al., 2007; Tran et al., 2018; Yang et al., 2006; Yu et al., 2011). Most of the S genes have been described for the rice-*Xanthomonas oryzae* pv. *oryzae* (*Xoo*) and rice-*Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) interactions, but there are also reports for the *Xanthomonas* pathosystems that involve other hosts such as citrus, cotton, pepper, tomato, wheat, and cassava. Table 6 lists the known S genes exploited by TALEs.

Table 6. TALEs and targeted susceptibility genes exploited by *Xanthomonas* species in rice, citrus, cotton, pepper, tomato, wheat, and cassava.

Pathogen	Host	TALE	Target gene	Characterization of the target gene	Reference
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	Rice	PthXo1	<i>OsSWEET11</i> (<i>Xa13</i>)	Clade III SWEET transporter	(Römer et al., 2010; Yang et al., 2006)
		PthXo2A PthXo2B PthXo2C	<i>OsSWEET13</i> (<i>Xa25</i>)		(Oliva et al., 2019; Zhou et al., 2015)
		AvrXa7 PthXo3 TalC TalF (Tal5)	<i>OsSWEET14</i> (<i>Xa41</i>)		(Antony et al., 2010; Römer et al., 2010;

					Streubel et al., 2013; Yu et al., 2011)
		PthXo6 TalB	<i>OsTFX1</i> (<i>OsZIP73</i>)	Transcription factor (Master regulator?)	(Sugio et al., 2007; Tran et al., 2018)
		TalB	<i>OsERF#123</i>	Transcription factor (Master regulator?)	(Tran et al., 2018)
		PthXo7	<i>OsTFIIAγ1</i>	Transcription factor	(Sugio et al., 2007)
<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>	Rice	Tal2g	<i>OsSULTR3;6</i>	Sulfate transporter	(Cernadas et al., 2014)
		Tal7	<i>Os09g29100</i>	Cyclin-D4-1	(Cai et al., 2017)
		Unknown	<i>OsERF#123</i>	Transcription factor (Master regulator?)	(Tran et al., 2018)
<i>Xanthomonas citri</i> pv. <i>citri</i>	Citrus	PthA PthA4 PthAw PthA*	<i>CsSWEET1</i>	Clade I SWEET transporter	(Hu et al., 2014)
		PthA and 13 allelic variants	<i>CsLOB1</i>	Transcription factor (master regulator)	(Duan et al., 2018; Hu et al., 2014; Teper and Wang, 2021)
		PthA2 PthA4	<i>CsLOB2</i>	Transcription factor (master regulator)	(Pereira et al., 2014; Zhang et al., 2017)
<i>Xanthomonas citri</i> pv. <i>aurantifolii</i>	Citrus	PthB PthC	<i>CsLOB1</i>	Transcription factor (master regulator)	(Hu et al., 2014)
		PthC1	<i>CsLOB2</i>	Transcription factor (master regulator)	(Pereira et al., 2014; Zhang et al., 2017)
<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i>	Cassava	TAL20 _{Xam668}	<i>MeSWEET10a</i>	Clade III SWEET transporter	(Cohn et al., 2014)
		TAL14 _{Xam668}	Unknown	Unknown	(Cohn et al., 2016, 2014)
	Cotton	Avrb6	<i>GhSWEET10</i>	Clade III SWEET transporter	(Cox et al., 2017)

<i>Xanthomonas citri</i> pv. <i>malvacearum</i>		Tal2	Unknown	Unknown	(Haq et al., 2020)
<i>Xanthomonas translucens</i> pv. <i>undulosa</i>	Wheat	Tal8 _{XT4699}	TaNCED_5BS	9-cis-epoxycarotenoid dioxygenase	(Peng et al., 2019)
		Tal2 Tal4b	Unknown	Unknown	(Falahi Charkhabi et al., 2017)
<i>Xanthomonas euvesicatoria</i>	Pepper/ Tobacco	AvrBs3	<i>Upa20</i>	Transcription factor (master regulator)	(Kay et al., 2007)
<i>Xanthomonas gardneri</i>	Tomato	AvrHah1	<i>bHLH3</i> <i>bHLH6</i>	Transcription factor (master regulator)	(Schwartz et al., 2017)

Most of the characterized S genes belong to the clade III of the sugar transporter SWEET family. These genes encode for glucose-sucrose transporters that, under non-pathogenic conditions, participate in the apoplastic phloem loading through exportation of photoassimilated sugars (especially sucrose) into the apoplast surrounding the phloem, next to the sieve element/companion cells (Chen, 2014). However, it is thought that TALE-mediated transcriptional activation of these genes triggers the sugar exportation to the mesophyll apoplast, where the pathogen uses these sugars as a carbon source. A second consequence of this sugar imbalance in the mesophyll apoplast is water accumulation, which hypothetically facilitates the movement of the pathogen and the entrance of more bacteria to the soaked tissue (Aung et al., 2018; Chen, 2014; El Kasmi et al., 2018; Schwartz et al., 2017). In rice, three different SWEET paralogues from the clade III have been described as S genes for the bacterial blight caused by *Xanthomonas oryzae* pv. *oryzae*. The *OsSWEET11* gene (a.k.a *Xa13*) is transcriptionally induced by the TALE PthXo1, the *OsSWEET13* gene (a.k.a *Xa25*) by TALEs PthXo2A, PthXo2B, and PthXo2C, and *OsSWEET14* (a.k.a *Xa41*) gene by TALEs AvrXa7, PthXo3, TalC or TalF (a.k.a Tal5). In the latter case, all the mentioned TALEs except for TalC use overlapped EBEs, which depicts a case of convergent evolution. In most cases, disruption of the rice SWEET-activating TALEs leads to decreased virulence and bacterial titers (reviewed by (Hutin et al., 2015a)). In cotton, the *GhSWEET10* gene has been shown to be activated by the major TALE Avr6 present in strain XcmH1005, and its activation was not only associated to sucrose export and water-soaked lesion formation, but the release of bacteria from the leaf interior to the leaf surface during infections, but without affecting the bacterial titers reached in the mesophyll (Cox et al., 2017; Yang Yinong et al., 1994). In cassava, transcriptional activation of the *MeSWEET10a* gene by the TAL20_{Xam668} effector leads to glucose and sucrose export to the apoplast, water-soaked symptom development, and an increase in apoplastic bacterial growth (when compared to the mutant lacking this TALE) (Cohn et al., 2014). This interaction is described more in detail in the section **TAL20_{Xam668} and other variants** below. In citrus, TALEs PthA4, PthAw, and PthA* can mediate the transcriptional activation of the clade I sugar transporter *CsSWEET1* gene, but this activation does not result in measurable phenotypic improvements, which leads to think that this transporter does not play a role in pathogenesis (Hu et al., 2014).

A second more heterogeneous category of S genes targeted by TALEs are transcription factors and master regulators. The transcriptional induction of these genes results in the transcriptional alteration of secondary targets, whose coordinated role induces host susceptibility. The first characterized TALE-targeted master regulator was the *upa20* gene, which encodes an acidic bHLH transcription factor for cell enlargement that stimulates cell growth and, in the context of the disease, causes hypertrophy that allows for the development of leaf symptoms and potentially enhances the epiphytic survival of the pathogen (Kay et al., 2007; Wichmann and Bergelson, 2004). In tomato, two more members of the bHLH transcription factor family have been described as TALE targets and S genes: bHLH3 and bHLH6. The major TALE AvrHah1, translocated to the host cells by *Xanthomonas gardneri*, upregulates both master regulators, whose downstream targets include a pectate lyase. This secondary S gene was demonstrated to be important for water-soaked symptom formation, hypothetically through an increase of the cell wall hygroscopicity due to the depolymerizing action of this enzyme (Schwartz et al., 2017). The *CsLOB1* gene belongs to the lateral organ boundaries (LOB) domain family, whose members, under non-pathogenic conditions, act as transcription factors to regulate plant growth and development (Majer and Hochholdinger, 2011). However, some strains of *Xanthomonas citri* are able to exploit this gene through transcriptional activation to increase the susceptibility of the plant. PthA4 and several of its homologs bind the promoter sequence of *CsLOB1* and induce the high-level transcription of this factor, which leads to hypertrophy, hyperplasia and water soaking symptoms (typical canker symptoms) (Hu et al., 2014; Teper and Wang, 2021). Although the physiological mechanism undertaken by the secondary targets is not clear, ectopic induction (in a non-pathogenic context) of *CsLOB1* is enough to develop pustules in grapefruit leaves, but it does not induce water soaking. Two pectate lyases and one zinc finger C3HC4-type RING finger protein, another transcription factor, might be involved in the pustule formation as secondary targets (Duan et al., 2018).

In rice, the gene *OsTFX1* (a.k.a *OsbZIP73*) encodes a bZIP transcription factor that, under non-pathogenic conditions, is involved in cold adaptation and response to cold stress (Liu et al., 2019). TALEs, PthXo6 and TalB, from Asian and African *Xanthomonas oryzae* pv. *oryzae* directly upregulate this transcription factor during pathogenesis, and the physiological consequences of this transcriptional activation (so far unknown) support the virulence of the pathogen (Sugio et al., 2007; Tran et al., 2018). Likewise, TalB from the African Xoo strain MAI1 and TALEs from several *Xanthomonas oryzae* pv. *oryzicola* strains upregulate the *OsERF#123* gene, a protein that contains an AP2 domain, whose upregulation confers susceptibility to the rice host in an unknown way (Tran et al., 2018). Transcription factors of the AP2/ERF family are implicated in the response to several stresses and are able to respond to abscisic acid and ethylene hormone signaling (reviewed by (Xie et al., 2019)), with a subgroup implicated in defense response, suggesting that this transcription factor could act as a hub for modulating the defense against the pathogen (Tran et al., 2018). Finally, in rice, the *OsTFIIA γ 1* gene encodes for one of the gamma subunits of general transcription factor A required for RNA polymerase II-dependent transcription, which stabilizes the TATA box-binding protein TFIID complex and is crucial for transcription. PthXo7, a TALE translocated by the PXO99^A strain, upregulates the expression of this gene, whose product is crucial for TALE function (Sugio et al., 2007). As described earlier, C-terminal regions of TALEs contain domains that interact with transcription factors to initiate the transcriptional activation, and TFIIA γ units have been described as the TAL-effector interacting transcription factors in rice, citrus, pepper and tomato (Huang et al., 2017; Yuan et al., 2016). Therefore, the susceptibility induced by the TALE-mediated upregulation

of this transcription factor does not rely on secondary target activation, but in an increase of the available pool of TFIIA γ subunits that will serve to initiate transcription. In line with this, *Xanthomonas oryzae* pv. *oryzae* strains carrying PthXo7 (naturally or transformed) are able to cause disease in the context of plants carrying the recessive allele *xa5*, a protein version of TFIIA γ 5 that weakly interacts with TALEs (more details in the following section), since the role of TFIIA γ 1 allows for the rest of the TALEs to carry out their functions (Han et al., 2020; Ma et al., 2018).

Some other TALE-targeted genes have been described as key virulence factors exploited by the pathogen. The rice sulfate transporter-encoding gene *OsSULTR3;6* has been shown to be upregulated in rice through the transcriptional action of the *Xoc* TALE Tal2g. Disruption of this molecular interaction leads to smaller water-soaked lesions and less exudated bacteria on the lesion surface. Although the exact physiological implications of the upregulation of this gene in the pathogenesis is not understood, it is hypothesized that it alters redox status or osmotic equilibrium of the cell to interfere with defense signaling and to enhance water-soaking (Cernadas et al., 2014). In the same pathosystem, TALE Tal7 is able to transcriptionally activate the rice gene *Os09g29100*, which encodes a Cyclin-D4-1 protein. Although the relevance of this S gene has not been demonstrated for this pathosystem, its heterologous expression into *Xoo* strains carrying the *avrXa7* avirulence gene disrupts the effector-triggered immunity process triggered by the presence of this protein in a host plant carrying the *Xa7* allele (more details about this gene in the next subsection). In non-pathogenic scenarios, Cyclin-D4-1 proteins are involved in cell cycle progression regulation through the interaction with cyclin-dependent kinases (CDK), but in the context of the disease, it is speculated that upregulation of this gene could improve *Xoc* proliferation through promotion of nutrient production by parenchyma cells and/or by alteration of the CDK-related signal induction of defense (Cai et al., 2017).

In citrus, upregulation of the *CsLOB1* homolog *CsLOB2* was described as PthA2-, PthA4-, and PthC1-dependent, but the role of this transcriptional activation in pathogenicity was initially not tested (Pereira et al., 2014). However, although activation was not demonstrated as a consequence of a natural TALE, but from artificial TALEs, *CsLOB2* and *CsLOB3* were shown to be functionally redundant to *CsLOB1* in citrus and to trigger cell hypertrophy and hyperplastic lesion formation (Zhang et al., 2017). Strains of the pathogen *Xanthomonas translucens* pv. *undulosa* that contain the TALE Tal8 exhibit longer leaf lesions when compared to strains that do not carry homologs of this TALE, and its expression in strains lacking this factor dramatically increases the lesion lengths and bacterial migration from the inoculation point. Tal8 is able to transcriptionally induce two alleles of the wheat gene 9-cis-epoxycarotenoid dioxygenase 3 (*TaNCED_5BS* and *TaNCED_5DS*) and a member of the apetala-2/ethylene response factor (AP-2/ERF) family designated as *TaERF_1BL*. However, only induction of *TaNCED_5BS/5DS* has a measurable role in virulence and therefore can be considered a S gene for wheat. The enzymatic product of this gene carries over the first step in the biosynthetic pathway of the abscisic acid (ABA), whose characterized hormonal effects included stomatal closure, gas and water transpiration reductions, and repression of NPR1 (a master regulator of the salicylic acid-controlled signaling pathway). It is then hypothesized that Tal8 manipulates the phytohormone balance to create a favorable aqueous apoplasmic environment and to suppress the ability of wheat to perceive SA (Peng et al., 2019).

Finally, there are studies that have identified a number of major TALEs, but to date, the targeted S genes remain unknown. This is the case of TAL14_{xam668} from the cassava pathogen *Xanthomonas*

phaseoli pv. *manihotis*, whose deletion results in lower bacterial titers in cassava leaves (Cohn et al., 2016, 2014) (this is described in more detail in the subsection **TALE14: TALE1_{Xam}, TAL14_{Xam668}, and TAL14_{CIO151}**); Tal2 and Tal4b from the wheat pathogen *Xanthomonas translucens* pv. *undulosa*, whose individual disruption results in shorter lesions after inoculation of *Triticum aestivum* cv. Chinese Spring plant leaves (Falahi Charkhabi et al., 2017); and Tal2 from the cotton pathogen *Xanthomonas citri* pv. *malvacearum* (strain Xss-V2–18), whose disruption is followed by lower bacterial titers (when compared to the wild type) and abolishment of water-soaking in cotton leaves (Haq et al., 2020).

Avirulence TALEs and executor genes

Evolution has led to the rise of molecular traps that take advantage of the targeted transcriptional activation caused by TALEs. In these traps, EBEs are located in the promoters of genes, whose expression triggers defense and/or results disadvantageous for bacteria and confers resistance to the plant. These genes are collectively called executor (E) genes (see table 7). So far, only a few examples have been found in nature. TALEs AvrBs3 and AvrHah1 from *Xanthomonas euvesicatoria* and *Xanthomonas gardneri*, respectively, target the pepper *Bs3* gene. Although these two TALE genes have a different number of repeats (17.5 repeats for AvrBs3 vs 13.5 repeats for AvrHah1), they share 10 RVDs and bind overlapped EBEs in the promoter of *Bs3* (Römer et al., 2007; Schornack et al., 2008). *Bs3* encodes a YUCCA-like flavin-dependent monooxygenase whose enzymatic action is thought to trigger a hypersensitive response (HR) through the conversion of NADPH into H₂O₂. Although the exact mechanism by which *Bs3* triggers HR is unknown, its enzymatic action is accompanied by increased levels of salicylic and pipecolic acids, suggesting that it exploits the established immune signaling pathways to trigger HR (Krönauer et al., 2019). Similarly, AvrBs4 from *X. euvesicatoria* upregulates the tightly regulated *Bs4C* gene in pepper, leading to a HR (Strauß et al., 2012). Although little is known about the nature and function of the encoded protein, several *Solanaceus* species and grapevine have homologs for this gene (Strauß et al., 2012), and its expression in rice and *Nicotiana benthamiana* leads to cell death. The molecular characterization of the encoded peptide indicates that it is a transmembrane protein located in the endoplasmic reticulum (ER) (Wang et al., 2018).

In rice, four E genes have been described, namely, *Xa27* (Gu et al., 2005), *Xa10* (Gu et al., 2008), *Xa23* (Wang et al., 2015), and *Xa7* (Chen et al., 2021), which are transcriptionally activated by TALEs AvrXa27, AvrXa10, and AvrXa23, AvrXa7/PthXo3, respectively. As *Bs4C*, *Xa27* is tightly regulated in rice and its encoded protein has a key N-terminal signal-anchor-like sequence required for its defensive role. Little is known about its physiological role, but this protein is translocated to the apoplast, and its overexpression leads to thickened vascular bundle elements (Wu et al., 2008). *Xa10* is a small protein of 126 amino acids that anchors to the ER membrane as hexamers, potentially through its four transmembrane helices. *Xa10*-mediated HR is triggered by the Ca²⁺ depletion and the disruption of the ER, which activates programmed cell death (Tian et al., 2014). *Xa23* is also a short protein with 50% identity to *Xa10* and is then thought to be a transmembrane protein. It was shown to trigger HR in rice, tobacco, and tomato, but the mechanism that activates the programmed cell death is so far unknown (Wang et al., 2015). By contrast, *Xa7* encodes a 113 amino-acid protein which differs significantly from the previously described E genes (but contains putative transmembrane domains) and has no homologs in other plants. This gene is targeted by two known TALEs, AvrXa7 and PthXo3, and its activation correlates with elevated levels of PR proteins (as

defense markers) and triggers HR in rice and tobacco (as heterologous system) (Chen et al., 2021). This E gene has two interesting characteristics: its expression depends on environmental temperature, which creates a heat-tolerant resistance (Chen et al., 2021; Cohen et al., 2017; Webb et al., 2010), and it has been demonstrated to be highly durable. This latter characteristic is presumably due to the fact that AvrXa7 has a dual role: in virulence by targeting the S gene *OsSWEET14* and in avirulence by targeting the E gene *Xa7* (Vera Cruz et al., 2000).

In sweet orange (*C. sinensis*) and lemon (*C. limon*), the targeted transcriptional enhancing activity of the short TALE PthA4AT (7.5 repeats) triggers HR. This TALE is a rare allele of PthA4, a major TALE for canker-causing *Xanthomonas citri*, and it is thought to bind the promoter of an unknown E gene. This unknown factor is potentially also present in *Nicotiana benthamiana* since *Agrobacterium*-mediated PthA4AT expression results in HR (Roeschlin et al., 2019).

Table 7. TALEs and targeted executor genes activated by *Xanthomonas* species in citrus, pepper, and rice.

Pathogen	Host	TALE	Target E gene	Characterization of the target E gene	Reference
<i>Xanthomonas euvesicatoria</i>	Pepper	AvrBs3	<i>Bs3</i>	YUCCA-like flavin-dependent monooxygenase	(Römer et al., 2007)
<i>Xanthomonas gardneri</i>	Pepper	AvrHah1			(Schornack et al., 2008)
<i>Xanthomonas euvesicatoria</i>	Pepper	AvrBs4	<i>Bs4C</i>	Transmembrane protein located in the ER	(Strauß et al., 2012)
<i>Xanthomonas citri</i> pv. <i>citri</i>	Citrus	PthA4AT	Unknown	Unknown	(Roeschlin et al., 2019)
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	Rice	AvrXa27	<i>Xa27</i>	Protein exported to the apoplast	(Gu et al., 2005)
		AvrXa10	<i>Xa10</i>	Transmembrane protein located in the ER	(Gu et al., 2008)
		AvrXa23	<i>Xa23</i>	Putative transmembrane protein	(Wang et al., 2015)
		AvrXa7 PthXo3	<i>Xa7</i>	Putative transmembrane protein	(Chen et al., 2021)

Avirulence TALEs and resistance genes

Evolution has also led to the onset of classical resistance (R) genes that directly recognize TALE structures through protein-protein interactions. Although these TALE-recognizing genes are not widely distributed among plants, they have been found in both monocots and dicots and have canonical LRR-RLK or NBS-LRR structures (see table 8). The tomato *Bs4* gene encodes a TIR-NBS-LRR

protein that locates to the cytoplasm and is able to recognize TALEs as AvrBs4, Hax3, and Hax4 (Kay et al., 2005; Schornack et al., 2004). Although the physical protein-protein interaction has not been demonstrated, it is known that the N-terminal region plus the first 3.5 repeats are necessary for TALE recognition and defense response activation. This defense response is maintained when Bs4 is heterologously expressed on potato (*Solanum tuberosum*) and several *Nicotiana* species, and it is hypothesized to result from the SA pathway activation through *EDS1* (Schornack et al., 2004).

In rice, Xa3/Xa26 encodes an R gene with the typical structure of an LRR-RLK which triggers HR in the presence of the cognate TALE AvrXa3 (Li et al., 2004; Sun et al., 2004; Xiang et al., 2006). Although the exact mechanism by which AvrXa3 is recognized and triggers the defense response is unknown, the intricate Xa3/Xa26-mediated defense response involves thiamine synthesis through *OsDR8* gene (Wang et al., 2006), the WRKY13 and WRKY45-2 transcription factors as master regulators of secondary genes (Cheng et al., 2015; Tao et al., 2009; Xiao et al., 2013), the nucleic acid-binding-like protein C3H12 (H. Deng et al., 2012), the XA3/XA26-type kinase OsSERK2 (Chen et al., 2014), and the triosephosphate isomerase OsTPI1.1 (Y. Liu et al., 2018). Xa1 was the first NBS-LRR gene described as an R gene against a bacterial disease in rice (Yoshimura et al., 1998). This NBS-LRR contains an N-terminal zfBED domain and six highly conserved tandem repeats in the C-terminal region, which are thought to be responsible for TALE recognition (Read et al., 2020; Yoshimura et al., 1998). Several *Xoo* TALEs have been tested against Xa1, and so far, all the tested effectors have been recognized by Xa1, resulting in HR. Little is known about the mechanism behind HR triggering, but it has been demonstrated that this R protein acts inside the nucleus, potentially after the import of TALEs through importins (Ji et al., 2016). As it is closely-related to Xa1, Xo1 is an NBS-LRR protein characterized by an N-terminal region that includes a zfBED domain and five highly conserved tandem repeats located in the C-terminal region (Read et al., 2020). This gene triggers HR in the presence of TALEs from *Xoo* and *Xoc*, and it has been demonstrated that recognition is abolished in the absence of the N-terminal region and the first 3.5 repeats (as for Xa1 and Bs4) (Triplett et al., 2016). As explained below in more detail, the Xa1- and Xo1-TALE interaction, and the consequential HR, is somehow abolished by the action of truncated TALEs/interference TALEs (Ji et al., 2016; Read et al., 2016).

Table 8. TALEs and interactive R proteins described in tomato and rice.

Pathogen	Host	Interacting TALE	R protein	Characterization of the R protein	Reference
<i>Xanthomonas euvesicatoria</i>	Tomato	AvrBs4	Bs4	NBS-LRR	(Schornack et al., 2004)
<i>Xanthomonas campestris</i> pv. <i>campestris</i>	Tomato	Hax3 Hax4			(Kay et al., 2005)
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	Rice	AvrXa3	Xa3/Xa26	LRR-RLK	(Sun et al., 2004; Xiang et al., 2006)
		All tested <i>Xoo</i> TALEs	Xa1	NBS-LRR	(Ji et al., 2016; Yoshimura et al., 1998)

		All tested <i>Xoo</i> and <i>Xoc</i> TALEs	Xo1	NBS-LRR	(Read et al., 2016; Triplett et al., 2016)
--	--	--	-----	---------	--

Interference or truncated TALEs

Among this effector family, there are some TALEs characterized by having truncations and lacking some functional domains, but which have important roles in pathogenicity that do not involve transcriptional activation. These type of TALEs have been only observed in *Xanthomonas oryzae* pv. *oryzae* and *Xanthomonas oryzae* pv. *oryzicola* (see Table 9) and are known as truncTALEs due to their truncated architecture, or iTALEs, due to their interfering role in TALE recognition by the R proteins of the host. TruncTALEs or iTALEs are structurally different from the traditional TALEs, since they carry deletions on the N-terminal region (after the first 306 nucleotides, therefore preserving the exporting signal), and deletions on the C-terminal region that result in preservation of some NLSs but the removal of the activation domain. Although all the truncTALEs/iTALEs carry central repeat domains with different RVD composition, only the first 3.5 repeats of this domain seem to be important for the action mechanism of these effectors, potentially for the protein-protein interaction (Ji et al., 2016; Triplett et al., 2016). This specific architecture plays an important role for the nuclear localization of the truncTALE/iTALE, and the interaction with Xa1 or Xo1. In the absence of truncTALEs/iTALEs, the NLRs Xa1 or Xo1 recognize TALEs and triggers the ETI, and conversely, the injection of truncTALEs/iTALEs into the host cell by the pathogen avoid the Xa-1 or Xo-1 mediated recognition of traditional TALEs and the subsequent ETS triggering (Ji et al., 2016; Read et al., 2016). Therefore, these effectors potentially act as decoys that carry out dominant suppression of the NLR function (Zuluaga et al., 2017).

Table 9. TruncTALEs/iTALEs identified within *Xoo* and *Xoc* genomes. The asterisk indicates that the effect of this truncTALE/iTALE on virulence was tested in planta.

TruncTALE/ iTALE ID	Pathovar	Strain	Virulence function	Ref
Tal3a*	<i>Xoo</i>	PXO99 ^A	Interference with Xa-1 mediated TALE recognition.	(Ji et al., 2016)
Tal3*	<i>Xoo</i>	PXO86		
Tal12*	<i>Xoc</i>	BXOR1		
Tal3b*	<i>Xoo</i>	PXO99 ^A		
Tal11h*	<i>Xoc</i>	BXOR1		
Tal5e*	<i>Xoc</i>	RS105		
Tal2h*	<i>Xoc</i>	BLS256	Interference with Xo-1 mediated TALE recognition.	(Read et al., 2016)
Tal15	<i>Xoo</i>	K74	The particular architecture indicates that they have the same role of other truncTALEs/iTALEs.	(Ji et al., 2016; Read et al., 2016)
Tal6*	<i>Xoo</i>	PXO86		
Tal6	<i>Xoo</i>	T7174		
Tal1	<i>Xoo</i>	K74		
Tal8f	<i>Xoc</i>	CFBP7341		
Tal8f	<i>Xoc</i>	CFBP7331		
Tal3b	<i>Xoo</i>	JSB2-24		

Tal5i	<i>Xoc</i>	B8-12		
Tal5g	<i>Xoc</i>	BLS279		
Tal5i	<i>Xoc</i>	L8		
Tal2j	<i>Xoc</i>	CFBP2286		

TALE-related loss-of-susceptibility alleles

Since most of the key virulence functions exerted by TALEs are linked to specific DNA binding and transcriptional activation, the interference with these two crucial steps to create susceptibility results in improved resistance of the host. This avoidance of interaction with TALEs is naturally mediated by alleles of *S* genes whose EBE was changed or lost by mutations that maintain the protein function in the host, but prevent the interaction with TALEs (see table 10). In literature, there are two examples of rice and citrus alleles of the key TALE interactor TFIIA γ 5. In rice, the *xa5* allele encodes a TFIIA γ 5 where the 39th amino acid is changed from valine to glutamine and acts as a resistant factor to the bacterial leaf blight (Iyer and McCouch, 2004; Jiang et al., 2006). This mutation in the N-terminal domain of the protein decreases the association between TFIIA γ 5 and the transcription factor binding domain of the TALE, which is crucial for effective transcriptional activation (Han et al., 2020; Hui et al., 2019a; Yuan et al., 2016). Likewise, a screening of germplasm of citrus varieties allowed the identification of a different allele of *TFIIA γ 5* in *Atalantia buxifolia*, which shows resistance to *Xanthomonas* induced citrus canker. The molecular characterization showed 3D structural differences due to mutations in the 81th, 87th, and 90th aminoacids, and its expression was associated with low levels of TALE-mediated transcriptional activation (Tang et al., 2021).

As described earlier, most of loss-of-susceptibility (LoS) alleles rely on the natural variation found in the EBEs of major TALEs. This is the case of the recently found *AbLOB1* allele, which has two SNPs in the EBE of the major TALE PthA4 that significantly alter the transcriptional rate induced by PthA4 and results in increased resistance to canker (Tang et al., 2021). In rice, there are three examples of loss-of-susceptibility alleles that disturb recognition of *OsSWEET* genes. *xa13*, *xa25*, and *xa41* are recessive alleles of the susceptibility *OsSWEET11* (Chu et al., 2006; Römer et al., 2010; Yang et al., 2006; Yuan et al., 2009), *OsSWEET13* (Liu et al., 2011; Zhou et al., 2015), and *OsSWEET14* (Antony et al., 2010; Hutin et al., 2015b; Streubel et al., 2013; Yu et al., 2011) genes (respectively), whose promoter sequences naturally lack a matching EBE for several major TALEs from *Xoo*. Thus, these alleles confer improved resistance to *Xoo* strains that rely on the exploitation of only one *SWEET* gene through the variation in the corresponding EBE.

Table 10. Natural loss of susceptibility (LoS) alleles found in rice and citrus that disrupt TALE-gene interactions and compromise their physiological effect.

Pathogen	Host	TALE lacking EBE	Target gene	Characterization of the LoS gene	Reference
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	rice	None	<i>xa5</i>	Recessive allele of TFIIA γ 5	(Hui et al., 2019b; Iyer and McCouch,

					2004; Jiang et al., 2006)
		PthXo1	<i>xa13</i>	Recessive allele of <i>OsSWEET11</i>	(Chu et al., 2006; Römer et al., 2010; Yang et al., 2006; Yuan et al., 2009)
		PthXo2	<i>xa25</i>	Recessive allele of <i>OsSWEET13</i>	(Liu et al., 2011; Zhou et al., 2015)
		AvrXa7 PthXo3 TalC TalF (Tal5)	<i>xa41</i>	Recessive allele of <i>OsSWEET14</i>	(Antony et al., 2010; Hutin et al., 2015b; Streubel et al., 2013; Yu et al., 2011)
<i>Xanthomonas citri</i> pv. <i>citri</i>	citrus	None	<i>AbTFIIAγ5</i>	<i>Atalantia buxifolia</i> allele of <i>TFIIAγ5</i>	(Tang et al., 2021)
		PthA4	<i>AbLOB1</i>	<i>Atalantia buxifolia</i> allele of <i>LOB1</i>	

Based on this knowledge and the fact that TALEs are key virulence factors for most of the *Xanthomonas* species affecting crops, two major crops have been edited to remove EBEs for all the major TALEs that target S genes, such as SWEETs and LOB1, conferring improved resistance through loss of susceptibility. One of the first proof-of-concepts in rice (Blanvillain-Baufumé et al., 2017; Li et al., 2012) achieved the engineering of Kitaake rice plants with modified *OsSWEET14* promoter sequences that lacked EBEs for TalC, TalF (a.k.a Tal5), and AvrXa7. From a systematic strategy based on TALENs, this led to rice plants with enhanced resistance to Xoo strains relying on AvrXa7 and TalF, while showed that TalC probably targets an alternative S gene so far unknown (Blanvillain-Baufumé et al., 2017). The most recent and successful approach comprised the simultaneous edition of the rice *OsSWEET11*, *13*, and *14* promoters to eliminate the transcriptional activation mediated by variants of PthXo1, PthXo2, PthXo3, TalC, AvrXa7, and TalF. This genomic edition was carried out by CRISPR-Cas9 rice in the highly Xoo-susceptible line Kitaake and the elite mega varieties IR64 and Ciherang-Sub1, which resulted in resistance to most of the tested Xoo strains, without affecting the

agronomical traits in most of the engineered plants (Oliva et al., 2019). In a smaller approach for basic research, TALEN-based edition of the rice *OsTFIIA γ 5* resulted in improved resistance to bacterial blight, but the resistance level was not as good as the one conferred by the *xa5* allele (Han et al., 2020). Citrus has also been edited with the aid of the CRISPR-Cas9 system to enhance canker resistance through EBE to create artificial loss-of-susceptibility alleles. A total deletion of the PthA4 EBE on the CsLOB1 promoter of the sweet orange (*Citrus sinensis* Osbeck) resulted in morphologically normal plants with improved resistant to canker, which did not developed pustules nor allowed the pathogen to grow after inoculation (Peng et al., 2017). In a similar CRISPR-Cas9-based strategy, the coding sequences of the *CsLOB1* alleles of the grapefruit Duncan (*Citrus paradisi* Macf.) were mutated to create knock-out alleles. This strategy was based on the knowledge that *CsLOB1* homologs (*CsLOB2* and *CsLOB4*) play redundant roles and resulted in phenotypically normal plants that were highly resistant to *Xcc* canker (Jia et al., 2017). Further improvements on the CRISPR-Cas9 system through codon optimization allowed the obtention of homozygous canker-resistant Pummelo (*Citrus maxima*) plants in the T0 generation. In this case, the promoter of the pummelo *LOB1* gene was modified by creating deletions inside the PthA4 EBE, which resulted in homozygous and biallelic mutants with high resistance to canker (Jia and Wang, 2020).

Xpm TALEs and their role in cassava

As in other *Xanthomonas* species, TALEs are involved in *Xpm* pathogenicity (Bart et al., 2012; Cohn et al., 2016, 2014). According to southern blot analysis, all the tested *Xpm* strains carry *TALE* genes and most of them are located in plasmids. The number of *TALE* gene copies ranges from 1 to 5 TALEs, but most of the tested strains carry multiple *TALE* orthologs. A weak positive correlation has been observed between the number of *TALE* genes (assessed by southern blot) per strain and its virulence, measured as the area under the disease progression curve (AUDPC) (Bart et al., 2012). The repetitive nature of these genes has precluded their assembly from the short-read sequencing projects. Before the development of this thesis, there were only seven TALE variants for which the RVD sequence is publicly available (Bart et al., 2012; Castiblanco et al., 2013), and only one of these is deposited as nucleotide sequence in the GeneBank (Castiblanco et al., 2013). Table 11 shows the RVD sequence of these TALE variants. The pathogenicity role of these variants (except for TAL21_{ClO151}) was tested by multiple approaches, showing that TALE1_{xam}, TAL14_{xam668}, TAL14_{ClO151}, and TAL20_{xam668} play critical roles in pathogenicity, while no significant role was observed for TAL13_{xam668}, TAL15_{xam668}, TAL22_{xam668}. The role on pathogenesis of each TALE involved in *Xpm* virulence in cassava will be addressed in the following subsections.

Table 11. *Xpm* TALEs described in literature.

TALE gene ID	Repeat number	RVD sequence	Reference
TALE1 _{xam}	13.5	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NN-NI-NG	(Castiblanco et al., 2013)
TAL13 _{xam668}	12.5	NI-NS-NN-HD-NG-HD-NI-NG-HD-NN-NI-NI-NG	(Bart et al., 2012)
TAL14 _{xam668}	13.5	NI-NG-NI-NN-NG-HD-NS-NS-NN-NG-HD-NN-NI-NG	
TAL15 _{xam668}	14.5	NI-NG-NI-NN-HD-HD-NS-NS-NS-HD-HD-NS-HD-NG-NG	
TAL20 _{xam668}	19.5	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NS-HD-NN-HD-HD-HD-NI-NG-NG	

<i>TAL22</i> _{Xam668}	21.5	NI-NG-HD-NG-NG-NG-HD-HD-NG-NG-HD-NG-HD-HD-NG-NG-HD-NG-NG-NG	
<i>TAL14</i> _{CIO151}	13.5	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NN-NI-NG	
<i>TAL21</i> _{CIO151}	20.5	NI-NG-HD-NG-HD-N*-NG-NG-HD-HD-HD-NG-N*-NG-HD-NG-NG-NG-HD-NG-NG	

TALE20: *TAL20*_{Xam668} and other variants

*TAL20*_{Xam668} is one of the two TAL genes in *Xpm* associated with virulence and watersoaking symptoms. It was isolated from a cosmid library derived from the highly pathogenic Indonesian strain Xam668. This TALE gene is located in the smallest plasmid of the Xam668 genome. The key role of this TALE in pathogenesis was first described by Cohn and coworkers through insertional mutation of the gene and virulence testing. Infiltration of leaves of the cassava variety 60444 with the mutant Xam668Δ*TAL20*_{Xam668} showed that this TALE was involved in water soaking symptom formation. The expression of this TALE variant into strain CIO151, which causes weak water soaking symptoms, increased the capacity of this strain to cause water soaking, but it did not increase bacterial titers. The role of this TALE was also tested on the ability of the pathogen to grow in the leaf midvein and the apoplast, and results show that disrupting this TALE gene leads to significantly lower bacterial titers in both tissues. A transcriptional analysis coupled to TALE-targeted gene predictions identified the *MeSWEET10a* gene as the only target of *TAL20*_{Xam668}, and further experiments demonstrated that this transcriptional activation is direct through recognition of the EBE TATAAACGCTTCTCGCCCATC. The target gene *MeSWEET10a* (*Manes.04G123400*) encodes a clade III SWEET sugar transporter family member, which was demonstrated to export both glucose and sucrose. Transcriptional activation of this target gene by means of artificial TALEs (arTALEs) proved that the physiological consequences of its activation are important for bacterial fitness during infection and constitutes the only cassava susceptibility gene exploited by *Xpm* known so far (Cohn et al., 2014).

The activation of the susceptibility gene *MeSWEET10a* was tested upon infection with different *Xpm* strains, and results show that most of the tested strains induce the overexpression of the target gene, produce important water soaking symptoms, and carry TALE variants with 20 repeats (as assessed by southern blot). Conversely, the two tested strains (CIO151 and CFBP1851) lacking TALE20 variants did not transcriptionally induce this gene (Cohn et al., 2014). Mora and coworkers tested the transcriptional activation of the susceptibility gene *MeSWEET10a* in the leaves of the 60444 cassava variety 48 hpi with four highly virulent *Xpm* strains, including Xam668. Results show that the four strains are able to transcriptionally induce the gene, but the TALome of each strain was not available at the moment of the study release. Moreover, *TAL20*_{Xam668}-mediated transcriptional activation of *MeSWEET10a* is highly significant even after 24 hpi in the 60444 cultivar leaves, while there is no activation of this susceptibility gene in the leaves of the resistant variety CM6438-14 at the same timepoints (Mora et al., 2019).

TALE14: *TALE1*_{Xam}, *TAL14*_{Xam668}, and *TAL14*_{CIO151}

*TALE1*_{Xam} was the first sequenced *Xpm* TALE gene associated with virulence and was isolated from a previously cloned region of the p44 plasmid (present in the Colombian *Xpm* strain CFBP1851). The genomic context of this variant shows that the gene is flanked by imperfect repeats of 1300 bp, two

200-bp-long tandem repeats, and eight palindromic repeats, followed downstream by a putative resolvase, all of them suggesting recombination in the locus. The role of this TALE in *Xpm* pathogenesis was tested on the non-pathogenic and spontaneous TALE-mutant strain CFBP1851Δp44 (a.k.a. ORST4), which has an 8-kb deletion on the p44 plasmid that included *TALE1_{Xam}*. Results showed that expression of *TALE1_{Xam}* turns CFBP1851Δp44 into a pathogenic strain able to cause water soaking symptoms, reach significant titers (as the wild-type strain CFBP1851) in the local tissue, and migrate to the leaf vessels. Although the transcriptional activation activity of this TALE was demonstrated by the increase of GUS activity in the presence of the effector and the corresponding binding sequence in the promoter of GUS (in *Nicotiana benthamiana*), the target genes for this TALE were not assessed in this first study (Castiblanco et al., 2013).

TAL14_{Xam668} is a TALE variant slightly different from *TALE1_{Xam}* (the fifth RVD codes for NI in *TALE1_{Xam}*, while in TAL14_{Xam668} codes for NG), which was isolated from a cosmid library from the Indonesian highly virulent *Xpm* strain Xam668 (Bart et al., 2012). Its role in pathogenesis was described by Cohn and coworkers, who showed that TAL14_{Xam668} deletion resulted in lower bacterial titers in the midvein and the apoplast at 6 days post inoculation (dpi), when compared to the wild-type strain (Xam668) (Cohn et al., 2014). TAL14_{CIO151} is a TALE variant (from the Colombian *Xpm* strain CIO151) that has the same RVD sequence of *TALE1_{Xam}*, but, due to the lack of the nucleotide sequence for TAL14_{CIO151}, it is not possible to know if it is the exact homolog or not. Cohn and coworkers corroborated the role of this TALE (*TALE1_{Xam}*/TAL14_{CIO151}) by insertional mutagenesis in the strain CIO151, which resulted in decreased bacterial growth. Moreover, TAL14_{Xam668} and TAL14_{CIO151} were shown to be functionally interchangeable, since they complemented the function of each other in the context of the corresponding TALE mutants (Cohn et al., 2016). This indicates that despite the RVD difference, both variants are able to target one or more cassava susceptibility gene(s). A transcriptomics analysis by RNAseq and subsequent confirmation by RT-qPCR (in the cassava variety 60444) showed that TAL14_{Xam668} directly targets at least twenty-six genes (upregulated at 48 hpi), and ten of them are also directly targeted by the TAL14_{CIO151} variant. Table 12 shows the description of genes transcriptionally activated by the TALE14 variants. In spite of narrowing the target gene set, the quantity of potential targets did not allow the functional validation of each candidate target gene, nor the discovery of the susceptibility gene(s) targeted by the TALE14 variants (Cohn et al., 2016).

Table 12. TAL14-induced genes described for the *Xpm*-cassava interaction.

Gene ID	Annotation	Targeting variant	Reference
<i>cassava4.1_030094m.g</i> / <i>Manes.15G090000</i>	PTHR24326:SF257 - HOMEBOX-LEUCINE ZIPPER PROTEIN ATHB-52	TAL14 _{Xam668}	(Cohn et al., 2014)
<i>cassava4.1_026646m.g</i> / <i>Manes.03G183800</i>	PTHR13683:SF268 - EUKARYOTIC ASPARTYL PROTEASE FAMILY PROTEIN	TAL14 _{Xam668}	(Cohn et al., 2016, 2014)
<i>cassava4.1_026121m.g</i> / <i>Manes.15G026800</i>	PTHR33312:SF5 - MEMBRANE-ASSOCIATED KINASE REGULATOR 1-RELATED	TAL14 _{Xam668}	
<i>cassava4.1_031361m.g</i> / <i>Manes.04G012000</i>	PTHR33142:SF1 - CYCLIN-DEPENDENT PROTEIN KINASE INHIBITOR SIM-RELATED	TAL14 _{Xam668}	
<i>cassava4.1_023036m.g</i> /	No data	TAL14 _{Xam668}	

<i>cassava4.1_020499m.g</i> /	No data	TAL14 _{Xam668}	
<i>cassava4.1_024404m.g</i> /Manes.11G151300	PTHR11802//PTHR11802:SF95 - SERINE PROTEASE FAMILY S10 SERINE CARBOXYPEPTIDASE // SUBFAMILY NOT NAMED	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_007568m.g</i> /Manes.03G152600	PTHR31683:SF20 - PECTATE LYASE 1- RELATED	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_034150m.g</i> /	No data	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_023665m.g</i> /Manes.13G013900	PTHR31623:SF15 - ACETYL- COA:BENZYLALCOHOL ACETYLTRANSFERASE-LIKE PROTEIN- RELATED	TAL14 _{Xam668}	(Cohn et al., 2016)
<i>cassava4.1_001042m.g</i> /Manes.S044400	PTHR12526//PTHR12526:SF341 - GLYCOSYLTRANSFERASE // SUBFAMILY NOT NAMED	TAL14 _{Xam668}	
<i>cassava4.1_012090m.g</i> /Manes.11G050300	PTHR10209//PTHR10209:SF182 - OXIDOREDUCTASE, 2OG-FE II OXYGENASE FAMILY PROTEIN // SUBFAMILY NOT NAMED	TAL14 _{Xam668}	
<i>cassava4.1_022534m.g</i> /Manes.01G000600	PFO0892 - EAMA-LIKE TRANSPORTER FAMILY (EAMA)	TAL14 _{Xam668}	
<i>cassava4.1_016646m.g</i> /Manes.05G017000	PF14009 - DOMAIN OF UNKNOWN FUNCTION (DUF4228) (DUF4228)	TAL14 _{Xam668}	
<i>cassava4.1_025591m.g</i> /Manes.08G044900	PTHR33493:SF2 - LATE EMBRYOGENESIS ABUNDANT PROTEIN 4-5	TAL14 _{Xam668}	
<i>cassava4.1_015102m.g</i> /Manes.01G268500	PTHR10809:SF43 - VESICLE-ASSOCIATED PROTEIN 1-3	TAL14 _{Xam668}	
<i>cassava4.1_024542m.g</i> /Manes.05G186300	PTHR12482:SF5 - PROTEIN C09D4.4, ISOFORM C	TAL14 _{Xam668}	
<i>cassava4.1_022805m.g</i> /Manes.02G112100	PTHR33070:SF3 - EXPRESSED PROTEIN	TAL14 _{Xam668}	
<i>cassava4.1_020743m.g</i> /	No data	TAL14 _{Xam668}	
<i>cassava4.1_022871m.g</i> /	No data	TAL14 _{Xam668}	
<i>cassava4.1_011524m.g</i> /Manes.17G043200	PTHR10639//PTHR10639:SF8 - CLATHRIN LIGHT CHAIN // SUBFAMILY NOT NAMED	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_007516m.g</i> /Manes.15G048700	PTHR31683:SF20 - PECTATE LYASE 1- RELATED	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_033289m.g</i> /Manes.16G104900	PTHR10795//PTHR10795:SF362 - PROPROTEIN CONVERTASE SUBTILISIN/KEXIN // SUBFAMILY NOT NAMED	TAL14 _{Xam668} / TAL14 _{CI0151}	
<i>cassava4.1_011345m.g</i> /Manes.17G107100	K00134 - GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (GAPDH, GAPA)	TAL14 _{Xam668} / TAL14 _{CI0151}	

<i>cassava4.1_026299m.g</i> <i>/Manes.12G142100</i>	PTHR35771 - FAMILY NOT NAMED	TAL14 _{xam668} / TAL14 _{CI0151}	
<i>cassava4.1_019005m.g</i> <i>/Manes.09G123900</i>	PTHR35122 - FAMILY NOT NAMED	TAL14 _{xam668} / TAL14 _{CI0151}	
<i>cassava4.1_009347m.g</i> <i>/Manes.02G162000</i>	PF00892 - EAMA-LIKE TRANSPORTER FAMILY (EAMA)	TAL14 _{xam668} / TAL14 _{CI0151}	

Rationale

As pointed out in the introduction of this thesis, *Xpm* is considered one of the most relevant plant pathogen (Mansfield et al., 2012), and the impact of CBB can be devastating under favorable conditions for the pathogen. Prophylactic cultural practices and the use of resistant varieties have prevented the burst of new epidemics at a large scale (López Carrascal and Bernal, 2012). However, the endemicity of the disease and the fact that cassava is a staple crop mainly cultured in small-scale farming in tropical countries are critical reasons to understand the key factors of pathogenicity and resistance, and ultimately apply this knowledge to plant breeding.

The molecular determinants of resistance to CBB in cassava are not well understood. Until now, we know that cassava resistance is polygenic and additive (Umemura and Kawano, 1983), and complete resistance to *Xpm* has never been documented. Plant resistance to the pathogen arises from several molecular and cellular mechanisms, which are tissue specific and are serially activated with the movement of the pathogen (Boher et al., 1995; Kpémoua et al., 1996). It is believed that the promptness of the plant response is the main driver for resistance, which could be linked to the presence of specific genes. At the molecular level, the leaf tissue response to *Xpm* relies on several transcription factors that activate melatonin synthesis (Wei et al., 2016), the salicylic acid pathway (Wei et al., 2018), the increase of the autophagy rate (Yan et al., 2017), besides the canonical PTI-related responses such as reactive oxygen species accumulation, pathogen-related protein expression, and increased callose deposition (Medina et al., 2018). The role of an RLK and an NB-LRR has been described as part of the cassava defense response, but their action is strain-specific, and the elicitors remain unknown (Díaz Tatis et al., 2018; López Carrascal et al., 2003). Knowledge on the susceptibility determinants is even more scarce, and to-date we only know one of these factors: the sucrose exporting protein MeSWEET10a (Cohn et al., 2014).

On the side of the pathogen, the diversity has been extensively assessed through different molecular methods, and several populations from South America and Africa have been characterized (Bart et al., 2012; Chege et al., 2017; Rache et al., 2019; Restrepo et al., 2000; Trujillo et al., 2014a, 2014b). These studies have shown that the genetic pool available in South American populations is larger than the one found in African populations, which correlates with the possible origin of the pathogen and its more recent introduction in Africa. Likewise, these studies highlight the role of trade of cassava cuttings as a key factor that contributes to pathogen diversity in cassava (Bart et al., 2012; Trujillo et al., 2014b). Since the research questions addressed by these studies required molecular methods to assess neutral genomic diversity, there was less data about diversity of genes under positive or negative evolutionary pressure, such as effector genes. However, the sequencing of 65 genomes of strains from African and South American *Xpm* populations established the core effectome of the pathovar and showed that all the tested strains carried *TALEs* (Bart et al., 2012).

TALEs play key roles in the pathogenicity of most *Xanthomonas* species, and *Xpm* *TALEs* are not the exception. Although little is known about the distribution, diversity, and function of *TALE* effectors in this pathovar, *TALE14* variants (*TALE1_{Xam}*, *TAL14_{Xam668}*) and *TAL20_{Xam668}* significantly contribute to pathogen fitness in the plant (Castiblanco et al., 2013; Cohn et al., 2016, 2014). Knockout of *TALE14* variants resulted in reduced bacterial titers *in planta*, but the susceptibility gene(s) targeted by these *TALEs* and their function remain unknown (Castiblanco et al., 2013; Cohn et al., 2014). Furthermore,

it was demonstrated that the homolog *TAL14_{CI0151}* plays similar molecular roles, with slightly different target repertoire in the host (Cohn et al., 2016). On the other hand, the knockout of *TAL20_{Xam668}* in strain Xam668 resulted in a mutant which is not able to cause water soaked lesions and shows reduced titers *in planta*, as well. It was then demonstrated that these disease symptoms were dependent on the transcriptional activation of the susceptibility gene *MeSWEET10a*, and the increased number of the encoded sugar transporter, and this activation was mediated by *TAL20_{Xam668}* (Cohn et al., 2014). In summary, what we know about the roles of TALEs in the interaction between *Xpm* and cassava is that TALE14 and TALE20 are key virulence factors, and the TALE20-mediated susceptibility is related to *MeSWEET10a* transcriptional activation.

However, some questions remain open. For example: what is the distribution of these major TALEs among the *Xpm* populations? What is the function/target of TALE14 ? Are there more key susceptibility determinants targeted by TALEs of this pathogen? Most of these matters reside in the center of this thesis as the main biological questions. Although we know that no strain of *Xpm* without TALEs has ever been reported in the literature and that some variants are key for pathogenicity, we do not know the distribution of these variants, the combination patterns of TALE orthologs, and we know even less about other potential molecular functions. From this standpoint, we ignore if the described mechanisms where *Xpm* exploits the sugar transporter MeSWEET10a through TALE-mediated transcriptional activation is a common feature for *Xpm* strains, if there are alternative targets whose activation converge to the same phenotype, or if there are other important plant factors activated (or repressed) by TALEs in this pathosystem. These data are crucial to engineer cassava resistance, especially if it is based on TALE-effector biology. It has already been shown that TALE biology is an excellent opportunity to engineer loss of susceptibility in plants (Jia et al., 2017; Oliva et al., 2019).

This thesis explores the TALE biology in the framework of the *Xpm*-cassava interaction. A first chapter was devoted to address the TALE diversity and potential functions on a well-characterized *Xpm* population, where new TALE variants with key virulence functions are described. The second chapter addresses the discovery of a novel susceptibility gene belonging to the clade III sugar transporters from the *SWEET* family, that is targeted as an alternative to the *MeSWEET10a* gene. The third chapter describes a novel silencing platform for *Xanthomonas* based on the CRISPR interference technology and tested for silencing of several members of the TAL effector gene family at once. The platform was developed in *Xpm*, and we confirmed previous findings on TALE-related pathogenicity. We used TALEs as a proxy for family gene study, and by applying this tool to the cassava non-vascular pathogen *Xanthomonas cassavae*, we predict that *MeSWEET10a* would also be exploited as a susceptibility gene through TALE-mediated activation.

Chapter 1: TALE diversity and function in a Colombian *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) population

TALEs are key virulence factors for most phytopathogenic *Xanthomonas* species (Schandry et al., 2018). *Xpm* is the most relevant bacterial pathogen in cassava, and the role of TALEs is crucial for a successful host infection (Castiblanco et al., 2013; Cohn et al., 2014; López Carrascal and Bernal, 2012). TALEs seem to be present in all the *Xpm* strains that have been studied so far, and strains usually contain between 2 and 6 members of these genes (Bart et al., 2012; Cohn et al., 2014). Despite the abundance of TALEs in *Xpm* populations, we only know the sequence of seven variants coming from two strains (Castiblanco et al., 2013; Cohn et al., 2014). The particular architecture of these genes has hampered their sequencing by next generation sequencing (NGS), while sequencing through sanger usually required the cloning of each gene as a preliminary step. These constraints created a gap on the knowledge of *Xpm* TALE diversity, and functional studies are restricted to the analysis of a few strains. These studies showed that TALE14 and TALE20 can be considered as major virulence factors, since their absence results in lower bacterial titers and/or disruption of water soaked symptom (Castiblanco et al., 2013; Cohn et al., 2016, 2014). However, there is a lack of knowledge about the distribution of these TALEs and the functions that they fulfill.

This chapter addresses several research questions related to TALE diversity: what is the distribution of TALE homologs across *Xpm* strains from populations isolated from different regions (at a country scale)? are the two TALEs described as major virulence factors present in all the pathogenic strains? Are there novel TALE paralogs with potential new functions? How functionally close are the TALE variants present in the populations? Do they functionally converge? Answering these questions would start filling the gap on TALE and TALE-mediated functional diversity in *Xpm*, which is crucial for a successful plant breeding scheme based on plant-pathogen molecular interactions knowledge.

We selected Colombian populations of *Xpm* because these populations have previously been studied for genetic neutral diversity (and some of them for other types of studies) (Trujillo et al., 2014b, 2014a), and they were isolated from fields located in two regions where the cassava cultivation is intensive and an important source of income for smallholder farmers. We used several approaches to characterize the pathogen populations and their TALE contents, which resulted in the description of several TALEs, including a novel variant from a strain lacking TALE20. Based on this, we moved forward to transcriptomics to characterize the effect on the plant of strains with divergent TALomes. This molecular characterization led to the discovery of TALE22D, which possibly derives from TALE20 and plays the same molecular role as TAL20_{xam668}. Finally, *in silico* analyses based on the transcriptomics data and target predictions indicate that there are potentially interesting targets which could play a role in pathogenesis. This resulted in the description of one case of functional convergence for two TALEs, the first described in this pathosystem.

This work was performed in collaboration with several researchers, graduate students, and undergraduate students. My main role was to perform the wet-lab experiments and to analyze and exploit data generated by the bioinformaticians that collaborated with us. This work allowed me to

interact with graduate (Daniela Osorio, Rubén Mora), and undergraduate students on biology and microbiology (María Camila Buitrago, Luisa Jaimes) and guide them mainly to extract TALomes. The experimental set for the RNAseq was based on the bulk sequencing data that we obtained from *Xpm* strains, and then executed by a Rubén Mora (a master student by then). This experiment allowed us to successfully test and validate the use of *in vitro*-propagated cassava plants for transcriptomics. The two bioinformaticians involved in this project, Alexis Deereper and Álvaro Pérez, helped with the processing and analysis of NGS data, as with target prediction analyses. Camilo López, Silvia Restrepo, Adriana Bernal and Boris Szurek provided significant scientific guidance and helped me to improve my writing skills. The paper was submitted in December 2020, and published in February 2021, in the scientific journal *Microorganisms* (IF=4.167). This publication makes part of a special issue devoted to integration of science on *Xanthomonas* and *Xylella* for plant disease management.



Article

TAL Effector Repertoires of Strains of *Xanthomonas phaseoli* pv. *manihotis* in Commercial Cassava Crops Reveal High Diversity at the Country Scale

Carlos A. Zárate-Chaves ^{1,2}, Daniela Osorio-Rodríguez ², Rubén E. Mora ³, Álvaro L. Pérez-Quintero ¹ , Alexis Dereeper ¹, Silvia Restrepo ⁴ , Camilo E. López ³, Boris Szurek ¹ and Adriana Bernal ^{2,*}

- ¹ PHIM, CIRAD, INRAe, IRD, Montpellier SupAgro, University of Montpellier, 34090 Montpellier, France; carlos.zarate-chaves@ird.fr (C.A.Z.-C.); alperezqui@gmail.com (Á.L.P.-Q.); alexis.dereeper@ird.fr (A.D.); boris.szurek@ird.fr (B.S.)
- ² Laboratorio de Interacciones Moleculares de Microorganismos Agrícolas, Departamento de Ciencias Básicas, Universidad de los Andes, Bogotá 111711, Colombia; d.osorio23@uniandes.edu.co
- ³ Manihot Biotec, Departamento de Biología, Universidad Nacional de Colombia, Bogotá 111321, Colombia; remoram@unal.edu.co (R.E.M.); celopezc@unal.edu.co (C.E.L.)
- ⁴ Laboratorio de Micología y Fitopatología de la Universidad de los Andes (LAMFU), Departamento de Ciencias Básicas, Universidad de los Andes, Bogotá 111711, Colombia; srestrep@uniandes.edu.co
- * Correspondence: abernal@uniandes.edu.co; Tel.: +57-1-339-4949



Citation: Zárate-Chaves, C.A.; Osorio-Rodríguez, D.; Mora, R.E.; Pérez-Quintero, Á.L.; Dereeper, A.; Restrepo, S.; López, C.E.; Szurek, B.; Bernal, A. TAL Effector Repertoires of Strains of *Xanthomonas phaseoli* pv. *manihotis* in Commercial Cassava Crops Reveal High Diversity at the Country Scale. *Microorganisms* **2021**, *9*, 315. <https://doi.org/10.3390/microorganisms9020315>

Academic Editor: Christopher B. Blackwood
 Received: 15 November 2020
 Accepted: 24 December 2020
 Published: 4 February 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Transcription activator-like effectors (TALEs) play a significant role for pathogenesis in several xanthomonad pathosystems. *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), the causal agent of Cassava Bacterial Blight (CBB), uses TALEs to manipulate host metabolism. Information about *Xpm* TALEs and their target genes in cassava is scarce, but has been growing in the last few years. We aimed to characterize the TALE diversity in Colombian strains of *Xpm* and to screen for TALE-targeted gene candidates. We selected eighteen *Xpm* strains based on neutral genetic diversity at a country scale to depict the TALE diversity among isolates from cassava productive regions. RFLP analysis showed that *Xpm* strains carry TALomes with a bimodal size distribution, and affinity-based clustering of the sequenced TALEs condensed this variability mainly into five clusters. We report on the identification of 13 novel variants of TALEs in *Xpm*, as well as a functional variant with 22 repeats that activates the susceptibility gene *MeSWEET10a*, a previously reported target of TAL20_{Xam668}. Transcriptomics and EBE prediction analyses resulted in the selection of several TALE-targeted candidate genes and two potential cases of functional convergence. This study provides new bases for assessing novel potential TALE targets in the *Xpm*–cassava interaction, which could be important factors that define the fate of the infection.

Keywords: cassava bacterial blight; SWEET; host target genes; susceptibility; transcriptomics; EBE prediction

1. Introduction

Transcription activator-like effectors (TALEs) play a significant role for pathogenesis and virulence in several xanthomonads. Disruption or knocking out of some TALE encoding genes significantly affects the virulence of the pathogen or even abolishes pathogenicity [1–12]. This family of effectors shares a particular structure evolved to selectively bind plant host promoters and recruit the RNA polymerase complex in order to initiate transcription of downstream genes [13,14]. TALE proteins include a signal for translocation through the type III secretion system (T3SS) [15], nuclear localization signals (NLS) [16] and an acidic transcriptional activation domain (AAD) to promote transcription in the host cell [17]. The central region of the effector contains modular tandem repeats mainly composed of 33–35 amino acids [18,19] whose sequence varies primarily at residues 12th and 13th (so-called “repeat variable diresidue”, RVD). RVDs interact with DNA bases

with a nucleotide recognition preference [20–24], ruling the DNA sequence recognition. Sequences targeted by TALEs are usually located within or near the promoter of target genes, and are termed Effector Binding Elements (EBEs) [1]. After secretion, TALEs interact with host transcription factors like TFIIA γ subunits, to induce polymerase II–dependent transcription [25].

Many host genes targeted by TALEs have been described for the rice-*Xanthomonas oryzae* pathovars *oryzae* (*Xoo*) and *oryzicola* (*Xoc*), but data are also available for wheat [26], citrus, cotton, pepper, tomato, and cassava-interacting *Xanthomonas* (reviewed by [4]). TALE targets whose activation during infection is promoting host disease are defined as susceptibility (*S*) genes, many of which are nutrient transporters. For example, TALE-mediated activation of SWEET sugar transporters in rice, cotton and cassava leads to promotion of virulence potentially through nutrient hijacking [10,27,28], while activation of sulfate transporter *OsSULTR3;6* in rice may alter redox status or osmotic equilibrium to interfere with defense signaling and induce water-soaking [3]. In wheat, induction by TAL8 of *Xanthomonas translucens* pv. *undulosa* of a gene involved in abscisic acid regulation alters water management in favor of the pathogen [26]. Other *S* genes include plant transcription factors, which control more complex host cellular processes through multiple indirect targets [2,6,9,10,12]. In pepper, AvrBs3-dependent activation of *upa20*, a basic helix-loop-helix (bHLH) family member, induces hypertrophy [12], while AvrHah1-dependent activation of *bHLH3* and *bHLH6* upregulates cell wall-degrading enzymes [10]. In citrus, PthA4-mediated (and its functional equivalents) activation of the transcription factor *CsLOB1*, results in pustule development, which may be dependent on activation of DNA-interacting secondary targets [2,29].

Xanthomonas phaseoli pv. *manihotis* (*Xpm*), previously known as *Xanthomonas axonopodis* pv. *manihotis* [30], is a Gram-negative vascular pathogen responsible for Cassava Bacterial Blight (CBB). CBB has been reported in all the continents where cassava is grown [31]. As other xanthomonads, *Xpm* uses a suite of effectors to manipulate physiological processes in host plant cells. *Xpm* has up to 24 effectors that are translocated to the host cytoplasm through the T3SS (termed type III effectors), including members of the TALE family [32]. Population diversity studies using TALE1_{Xam} (a.k.a *pthB*) as a probe showed that all isolates from Latin-American and African *Xpm* collections carry TALEs [33–37]. Illumina-based genomic sequencing of more than 60 *Xpm* strains later confirmed that all strains contain at least one TALE [32,38]. Full *Xpm* TALE sequence is available for only eight effectors [8,39,40], and only two complete *Xpm* TALomes (whole set of TALEs present in a strain) have been reported so far [8,39].

Information about *Xpm* TALEs and their target genes in cassava has been growing in the last years. The first TALE characterized in *Xpm* is TALE1_{Xam} (contains 14 repeats), which has an important role in pathogenicity [40], but whose targets remain unknown [8,39–41]. Cohn and coworkers [8] deciphered the contribution of each of the five TALEs of strain Xam668, showing that mutant strains for TAL14_{Xam668} and TAL20_{Xam668} significantly affected pathogen fitness characterized by reduction of bacterial growth and/or symptom formation. Importantly they also demonstrated that TAL20_{Xam668} induces the *S* gene *MeSWEET10a*, which codes for a clade-III sugar transporter from the SWEET family.

Studying cassava pathogens is highly relevant since this crop is one of the most important starchy root crops for food security in America, Africa and Asia [42]. The annual worldwide production of cassava is estimated to 270 million tons and serves as a staple food for more than 800 million people, mainly in tropical countries. This and its tolerance to drought, make it an important food security crop [43,44]. CBB is the most devastating bacterial disease of cassava, causing losses that range from 12% to 100% [31] depending on environmental conditions [45,46], with incidences ranging between 25% and 100% [46,47]. CBB symptoms include angular leaf spotting, blight, wilt, gum exudation, vascular necrosis, stem canker and dieback (reviewed by [48]). In Colombia, CBB was first described in 1971 [35]. More recently, Trujillo and coworkers described the genetic structure of *Xpm* populations in the most prominent cassava production regions, showing high susceptibility

of the most common cultivars grown and a complex pathogen population structure in some of these regions [49,50].

Science-based crop resistance improvement and disease tackling require a compilation of knowledge to understand the cellular and molecular bases of plant-pathogen interactions. As *Xanthomonas* pathogenicity notably relies on TALEs, knowledge on diversity and function of these effectors, in conjunction with resistance (R) proteins, are a cornerstone for resistance breeding in the affected crops [51–54]. Among the mechanisms of resistance associated with TALEs [55], recessive resistance is mediated by loss-of-susceptibility (LoS) alleles, where alternative alleles of an *S* gene that does not possess the EBE on the promoter sequence prevent the recognition and upregulation of the *S* gene (reviewed by [53]). As previously stated, only one *S* gene has been described for cassava (*MeSWEET10a*) as a direct target of TALE20_{Xam668} [8]. No LoS alleles, executor genes or TALE-recognizing nucleotide-binding leucine-rich repeat (NLR) receptors in cassava have been reported so far. In this work we aimed to characterize the TALE diversity of one of the best characterized *Xpm* populations in the world. We sequenced the repeat regions of 46 TALEs from 18 *Xpm* strains, including the complete TALomes from seven strains. We also aimed to determine functional diversity of the potential TALE-targeted genes and to propose candidates that could contribute to a better understanding of this pathosystem. The TALE-affinity analyses, host target prediction, and transcriptomics on cassava plants inoculated with the pathogen resulted in potential novel targets that might be relevant for pathogenesis. This study expands our understanding of the complex interactions between *Xpm* and cassava by picturing the TALE diversity in the context of this pathosystem.

2. Materials and Methods

2.1. Bacterial Strains, Plant Material, Culture Conditions and Media

All *Xanthomonas phaseoli* pv. *manihotis* strains were previously collected in diversity studies of this pathogen from our group (Laboratorio de Micología y Fitopatología de la Universidad de los Andes, Universidad de los Andes, Colombia). All the selected strains were pathogenic to the highly CBB-susceptible variety 60444. Isolates from two geographically distant cassava productive regions in Colombia were selected through maximization of differences of neutral genetic diversity. Additionally, two strains, one per region, isolated at least one decade before the collection dates of most isolates, were selected as reference strains. Bacteria were streaked on YPG (yeast extract 5 gL⁻¹, peptone 5 gL⁻¹, glucose 5 gL⁻¹, agar-agar 15 gL⁻¹) solid media [35] and incubated for 48 h at 28 °C or grown in Phi broth (yeast extract 1 gL⁻¹, peptone 10 gL⁻¹, casaminoacids 1 gL⁻¹) [35] at 28 °C, under constant shaking at 220 rpm for 24 h. Cassava cuttings from the cultivar 60444 were planted on individual peat pots and grown under greenhouse conditions (27 °C ± 5 °C; photoperiod 12:12, relative humidity >60%). Bacterial inoculations were performed on 3-month old plants. In vitro-grown plants from the same cultivar were grown as described elsewhere [56].

2.2. Aggressiveness Assays and Bacterial Growth Curves

Bacterial aggressiveness was quantified as the leaf lesion area formed by *Xpm* upon inoculation of 3-month-old cassava plants. Bacteria were cultured in liquid media and washed with 10 mM MgCl₂ sterile solution. Cell density was adjusted to an OD₆₀₀ of 0.2 (c.a. 2 × 10⁸ cfu/mL). Adjusted bacterial suspensions were inoculated by placing a 10-μL drop of inoculum over a 2-mm (Ø) hole made with a cork borer through the leaf tissue. Plants were kept in a greenhouse under the same conditions described above. Lesions were individually photographed at 15 days post-inoculation (dpi) in a stereoscope and areas were measured using Image-J software (version 1.48) [57]. Each treatment was inoculated once in three individual plants, and the aggressiveness assay was replicated three times.

Bacterial growth in planta was assessed as described elsewhere [58] except that inoculations were performed on the variety 60444 and 0, 5, and 10 dpi were evaluated. Each

treatment was inoculated once in three different plants and the bacterial growth assessment was replicated two times.

2.3. Genomic DNA Extraction, Restriction Fragment Length Polymorphism (RFLP), and Southern Blot Analysis

A single colony of each *Xpm* strain was grown on Phi broth overnight. DNA was extracted from 2-mL bacterial pellets using the Genelute™ Bacterial Genomic DNA kit (Sigma-Aldrich, St. Louis, MO, USA) according to manufacturer's instructions. Genomic DNA concentration was determined by gel quantification and spectrophotometry on a NanoDrop™ 1000 (Thermo Fisher Scientific, Waltham, MA, USA). Ten micrograms of each genomic DNA were digested with 120 units of *Bam*HI-HF (NEB, Ipswich, MA, USA) sequentially added (3 times 40 units) over a 24-h incubation period. *Bam*HI-digested DNA was precipitated using sodium acetate and ethanol [59] and then resuspended in TE solution. Electrophoresis of digested DNA was performed on a 0.8% agarose gel on 0.5× TBE (Thermo Fisher Scientific, Waltham, MA, USA); migration was performed at a constant power of 5 Watts for 20 h, in a cold room and with periodic buffer renewal. DNA blotting was performed according to Roche's instruction manual for Digoxigenin application to filter hybridization [60]. The probe was synthesized from a conserved portion of the TALE1_{Xam} (from nucleotide 2909 to 3409) DNA contained in pF3, a pBluescript derivative with a 5.4-kb fragment from CFBP1851 plasmid p44 containing the full TALE1_{Xam} gene [35]. The probe was labeled with digoxigenin by means of the Random Primed DNA Labeling kit (Roche Diagnostics GmbH, Mannheim, Germany). Luminescent detection was achieved using chloro-5-substituted adamantyl-1, 2-dioxetane phosphate (CSPD) according to manufacturer's instructions. Chemiluminescent images were acquired on the ChemiDoc XRS + System (BioRad, Hercules, CA, USA) with the Chemi Hi Resolution protocol and signal accumulation mode.

2.4. Isolation and Sequencing of TALEs

The isolation of TALE genes was achieved by two different approaches. The first one consisted of PCR amplification of the central repeat region using GoTaq polymerase (Promega, Madison, WI, USA) and subsequent cloning into pGEM®-T easy (Promega, Madison, WI, USA). Amplification was performed using 200 nM of primers 427 Fw (5'-CGGTGGAGGCAGTGCATG-3') and 428 Rv (5'-ATCAGGGCGAGATAACTGGGC-3'); 1× GoTaq Green buffer, 1.5 mM MgCl₂; 100 μM dNTPs, 20 mM betaine, 0.025 GoTaq polymerase units/μL and a total amount of 40 ng of template DNA in a final volume of 20 μL. PCR conditions were as follows: Initial denaturation at 94 °C for 3 min; then 25 cycles of 94 °C for 40 s, 60 °C for 40 s, and 72 °C for 3 min and 30 s. Cloning into pGEM®-T easy was performed according to manufacturer's instructions. Ligation products were electroporated into *E. coli* DH5α cells. Screen of clones harboring inserts was performed using blue/white screening on LB agar using 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and a PCR-based confirmation (as described above).

The second approach was based on direct cloning of TALE genes into pBlueScript. RFLP data was used to calculate the position of each TALE in the electrophoresed *Bam*HI-digested DNA. For each strain, 70 μg of DNA were digested. Electrophoresis was carried out in a 0.8% agarose gel on 1× TAE; migration was performed at a constant power of 5 Watts for 5 h at room temperature. Gels were stained using a 3× GelRed (Biotium, Fremont, CA, USA) pool. Gels were quickly visualized in a ChemiDoc XRS+ System (BioRad, Hercules, CA, USA), regions where the expected TALE bands were excised and recovered using the Zymoclean™ Gel DNA Recovery Kit (ZymoResearch, Irvine, CA, USA) according to manufacturer's instructions. pBlueScript was digested with *Bam*HI-HF (NEB, Ipswich, MA, USA) and dephosphorylated with Antarctic Phosphatase (NEB, Ipswich, MA, USA) according to manufacturer's instructions. Eluted DNA was ligated with the prepared vector with T4 DNA ligase (Thermo Fisher Scientific, Waltham, MA, USA) at 22 °C overnight. Transformation of ligations, screening and confirmation were performed as described earlier.

Sequencing of isolated TALE genes was carried out by standard Sanger chemistry (Macrogen Inc., Seoul, Korea) and the assembly of reported fragments was performed using the Geneious software (version R11). In the case of TALE genes obtained by PCR, they were only considered accurate if they were found in at least two rounds of independent amplifications and cloning, or if they were also cloned directly from digested DNA.

2.5. TALE Characterization and Clustering

Assembled TALE sequences were treated and classified in variants using an in-house script on R (version 3.6.1). TALEs were clustered by predicted DNA binding specificity using FuncTAL software (version 1.1) [61]; default parameters were used. TALEs were also clustered by nucleotide repeat composition using DisTAL software (version 1.1) [61]; default parameters were used.

2.6. In-Vitro Plant Inoculation and Cassava RNA Extraction

Xpm strains were cultured in liquid media and prepared, as described earlier. Inocula were adjusted to an OD₆₀₀ of 0.02 (ca. 2×10^7 cfu/mL) in a 10-mM MgCl₂ solution. In-vitro propagated plants were inoculated with bacterial suspensions or mock solution (10-mM MgCl₂) with a swab on axial and abaxial surfaces of punctured leaves (nine needle punctures per leaf); each treatment was inoculated on three leaves per plant and on three different plants. Tissue surrounding inoculated punctures was collected at 50 hpi using a 3-mm diameter cork borer. Total RNA was extracted with the Invitrap Spin plant RNA minikit (STRATEC, Birkenfeld, Germany), using the RP buffer per manufacturer's instructions. Total RNAs were treated with RNase-free DNase I (Thermo Fisher Scientific, Waltham, MA, USA) and quality was controlled with the RNA 6000 Nano Kit (Agilent, Santa Clara, CA, USA).

2.7. RNA Sequencing and Transcriptome Analysis

Barcoded, paired-end (150-bp inserts) libraries were constructed with the TruSeq RNA Library Prep Kit (Illumina, Vancouver, British Columbia, Canada), per manufacturer's instructions. RNA libraries were pooled and sequenced on four flow cells of the NextSeq500 System. RNAseq analyses were performed using the Kallisto pseudo-mapper [62] and EdgeR [63] for differential expression analysis by comparing against the mock-inoculated treatment on R (version 3.6.1). GO-term enrichment analyses were performed using the topGO package [64].

2.8. TALE Target Prediction and Candidate Analysis

TALE targets were predicted using four different software: TALVEZ [65], TAL Effector Nucleotide Targeter 2.0 (TALENT 2.0) [66], TALgetterLong [67] and PrediTALE [68]. The *Manihot esculenta* promoterome (1-kb sequences preceding annotated translational start sites) was extracted from Phytozome's cassava genome version 6.1 [69], by means of the Biomart tool, and it was used as input for target prediction. All the algorithms were run using the default parameters. Output data were merged and compared on R (version 3.6.1) using an in-house script.

2.9. Semi-Quantitative and Quantitative RT-PCR

Xpm strains were cultured in liquid media and prepared as described earlier. Bacterial inoculum was adjusted to an OD₆₀₀ of 0.5 (ca. 5×10^8 ufc/mL) in a 10-mM MgCl₂ solution. Bacterial suspensions or the mock solution (10-mM MgCl₂) were infiltrated into leaves of 3-month old adult plants grown from stakes, by means of a needleless syringe; each treatment was inoculated on one leaflet per plant and on three different individuals. Infiltrated tissue was collected at 50 hpi in sterile tubes containing RNA-free glass beads, samples were frozen with liquid nitrogen and ground by vortexing. Total RNA was extracted as described earlier. cDNA synthesis was performed with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Waltham, MA, USA),

per manufacturer's instructions. Semi-quantitative RT-PCRs were performed with a 20- μ L reaction mix per tube containing $1 \times$ of $5 \times$ GoTaq Green Buffer, 1.5 mM of $MgCl_2$, 100 μ M of a dNTP mix, 0.2 μ M of each primer, 5 ng/ μ L of cDNA, and 0.025 U/ μ L of GoTaq polymerase (Promega, Madison, WI, USA). Amplification was programmed as follows: one step at 95 °C for 3 min, followed by a variable number of cycles (20 to 28, as needed) of 95 °C for 30 s, 60 °C for 40 s, and 72 °C for 30 s. Amplicons were resolved in 1% agarose gels on $0.5 \times$ TBE. RT-qPCRs were performed on a 7500 Fast & 7500 Real-Time PCR System. Each well contained 10 μ L of the following reaction mix: $1 \times$ of $2 \times$ SsoFastTM EvaGreen Supermix with Low ROX (BioRad, Hercules, CA USA), 300 nM of each primer, and 20 ng/ μ L of cDNA. PCR cycling was as follows: one step at 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s, and 60 °C for 30 s; data was acquired during the second step of each cycle. The melting curve was evaluated from 65 °C to 95 °C. Primers for reference and candidate genes are summarized in Table 1.

Table 1. Primers used for RT-qPCR assays.

Target Gene ^a	Type	Sequence (5'→3')	Reference
Candidate gene: Manes.06G123400 (<i>MeSWEET10a</i>)	Fw	TCCTCACCTTGACTGCGGTG	[8]
	Rv	AGCACCATCTGGACAATCCCA	
Candidate gene: Manes.04G033900 (Dof domain, zinc finger)	Fw	AAAGTGCCCAAGAGGTGGTG	This study
	Rv	GCCTTTCACCTGAAGCTGGG	
Candidate gene: Manes.13G045100 (Clavata3/ESR CLE-related protein)	Fw	CCACGACGAAC TTCACCCA	This study
	Rv	CGCTGGGAAC TTCATGAGCT	
Candidate gene: Manes.11G151300 (Serine carboxypeptidase)	Fw	GCCCCAACTGTTAGATTTGTGG	This study
	Rv	GGTGACCAGCTTCATACACCTT	
Candidate gene: Manes.15G052000 (Beta-glucosidase)	Fw	TTGAAGATATGCTCAACGACACG	This study
	Rv	CGTCTGCTCCGTTCTGATA	
Reference gene: Manes.08G061700.1 (Tubulin beta-6 chain)	Fw	GGAAAGATGAGACCAAGGA	[56]
	Rv	ACCAGTATACCAGTGCAAGAAG	

^a Gene ID for Phytozome's Cassava genome annotation version 6.1, main functional annotation is indicated in parentheses.

2.10. Statistical Analysis and Packages

All the statistical analyses were performed in R (version 3.6.1). Aggressiveness data was analyzed using an in-house script through a linear mixed model—packages lme4 [59] and multcomp [60]—for log-transformed data, where strain was used as predictor variable with fixed effects, and replicates in time were included as the random effect. All the statistical comparisons were performed at a 0.95 significance level. Correlation analyses were performed by using ggpubr [70]. RT-qPCR data were analyzed using pcr package [71], where a two-tailed *t*-test (alpha 0.05) was applied to the normalized expression values (normalized by using tubulin gene as reference) of each target gene in the inoculated treatment versus the mock-inoculated treatment. Boxplots, barplots, dotplots, and heatmaps were created with ggplot2 [72], gplots [73] package. EBEs were depicted in the promoter contexts by using the gggenes package [74] or Geneious software (version R11). Distance-based trees were modified on Geneious software (version R11).

3. Results

3.1. Aggressiveness of Selected *Xpm* Strains Is Homogeneous on Cassava Cultivar 60444

Xpm Colombian strains were selected based on three criteria: year of collection, the edaphoclimatic zone (ECZ) of CBB-endemic cassava-growing lands [35], and previous AFLP- and MLVA-based diversity information [49,50]. From the two geographically distant ECZs, which harbor genetically differentiated *Xpm* populations [50], a total of 18 strains

representative of the Colombian *Xpm* diversity were selected (Table 2 and Supplementary Figure S1). It is also worth noting that strains UA318 and UA681 were previously used in mapping studies for QTL of resistance in cassava against CBB [75]. Also, a subset of them was sequenced by Illumina (strains CIO151, CFBP1851, UA226, and UA306) [32,38], and/or pathotyped on a set of cassava varieties (strains UA226, UA306, UA318, UA531 and UA681) [50].

Table 2. General characteristics of the *Xpm* strains selected in this study.

Strain	Region of Origin (ECZ ^a)	Location of Origin	Year of Isolation	AFLP-Based Haplotype ^b	Estimated Number of TALEs ^c
CFBP1851	Caribbean coast (ECZ 1)	nd	1974	3 *	2
CIO151	Eastern plains (ECZ 2)	Meta	1995	3 *	5
UA226	Caribbean coast (ECZ1)	Chinú	2008	5 *	3
UA306	Caribbean coast (ECZ 1)	Palmitos	2008	7 *	4
UA318	Caribbean coast (ECZ 1)	Ciénaga de oro	2008	3 *	4
UA522	Caribbean coast (ECZ 1)	Chinú	2009	4 *	5
UA531	Caribbean coast (ECZ 1)	Chinú	2009	5 *	3
UA681	Caribbean coast (ECZ 1)	Chinú	2009	8 *	5
UA1061	Caribbean coast (ECZ 1)	Chinú	2009	5 *	2
UA1069	Caribbean coast (ECZ 1)	Chinú	2009	5 *	2
UA1183	Eastern plains (ECZ 2)	Villavicencio	2011	1 **	5
UA1211	Eastern plains (ECZ 2)	Granada	2011	2 **	5
UA1235	Eastern plains (ECZ 2)	Villavicencio	2011	3 **	5
UA1245	Eastern plains (ECZ 2)	Villavicencio	2011	3 **	5
UA1357	Eastern plains (ECZ 2)	Orocué	2012	8 **	4
UA1381	Eastern plains (ECZ 2)	Orocué	2012	6 **	4
UA1396	Eastern plains (ECZ 2)	Orocué	2012	7 **	5
UA1399	Eastern plains (ECZ 2)	Orocué	2012	7 **	5

^a Previously described in [35]. ^b Numbers were arbitrarily assigned to clades of distance trees from data reported by [49] (*) or [50] (**). Since analyses were performed separately, clade numbers from [49,50] are not shared. ^c Number of TALE bands determined by RFLP. nd, no data.

Aggressiveness of the strains was evaluated based on leaf lesion area on cassava adult plants from the cultivar 60444 at 15 dpi (Figure 1). All the strains were pathogenic and caused symptomatic infections, with water-soaked angular lesions and necrosis in some parts of the tissues. The biological variation among treatments was considerably high. However, lesion areas caused by isolates from the Eastern plain region show narrower inter-quartile ranges (IQRs) for most of the treatments when compared to Caribbean coast strains, which would point to a lower variability in aggressiveness for these strains. The reference strain from the Caribbean coast, CFBP1851, consistently caused the smallest lesions on 60444, while largest lesions were observed with strains UA318 and UA1061. However, due to the considerable variance, little could be inferred from this analysis, except that *Xpm* aggressiveness on 60444 seems to achieve homogeneous levels after 15 dpi and cannot be correlated with the AFLP- and VNTR-based genetic neutral diversity used as a criterion to select strains.

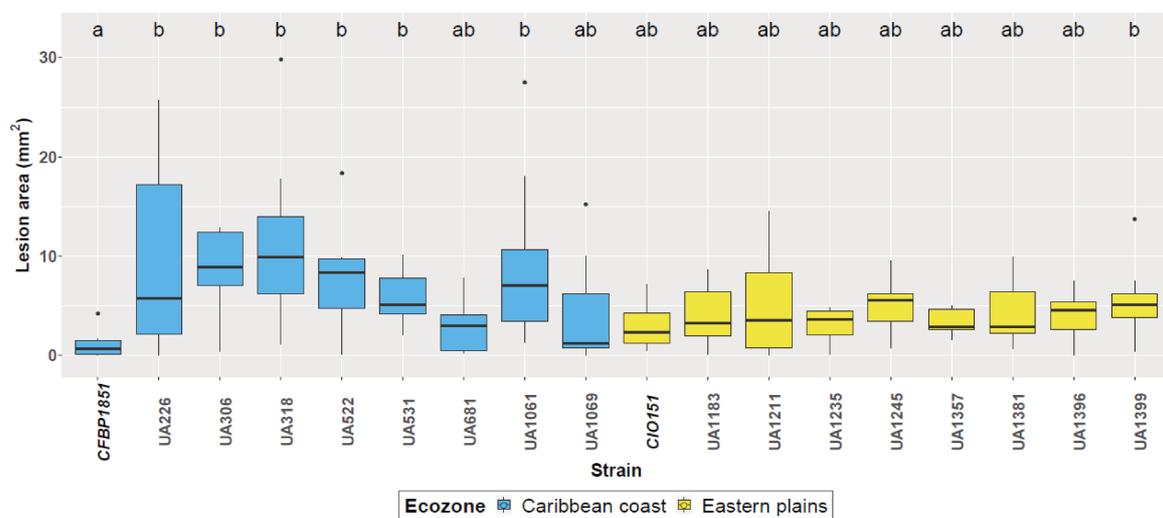


Figure 1. Aggressiveness of selected *Xpm* strains. Lesion areas in squared millimeters are displayed for each strain. Strains used as reference for each region (CFBP1851 and CIO151) are bolded and italicized in the x-axis. Boxes for strains isolated in the Caribbean coast region are colored in blue, while boxes for strains isolated in the Eastern plains are colored in dark yellow. Boxplots were constructed with measurements from three biological replicates with three replicates each. A linear mixed model was fitted to the log₁₀ transformed data, with replicates of the experiment as random effect, which showed to be near to zero. Different letters indicate significant differences according to a post-hoc Tukey's test ($\alpha = 0.05$).

3.2. RFLP Analysis Highlights TALome Patterns with Two to Six TALEs Per Strain and Restricted Size Ranges

RFLP analysis was performed using a 504-bp probe corresponding to a segment of the C-terminal region of the TALE1_{Xam} gene (a.k.a *pthB*). Figure 2A shows TALome patterns, where each band observed in the RFLP blot is represented by a letter or a star. Among the 18 tested strains, TALomes are consistently composed by effectors with 13, 14, 15, 20 or 22 apparent repeats, while bands corresponding to effectors sizes ranging from 16 to 19 repeats are completely absent (bimodal size distribution). RFLP analysis was followed by cloning and sequencing of most of the TALE genes present in the studied *Xpm* strains. From the 73 bands observed in the blot (Supplementary Figure S2), we were able to sequence 46 genes. TALE genes will be named according to their predicted number of repeats (e.g., TALE20 has 19.5 repeats). Variants are indicated in Figure 2A through letter codes according to RVD sequence classes of TALEs with the same repeat number. Gene names included the suffix *Xpm*. However, for the sake of readability of this document, suffix will be obviated. Eighteen RVD-based variants were found.

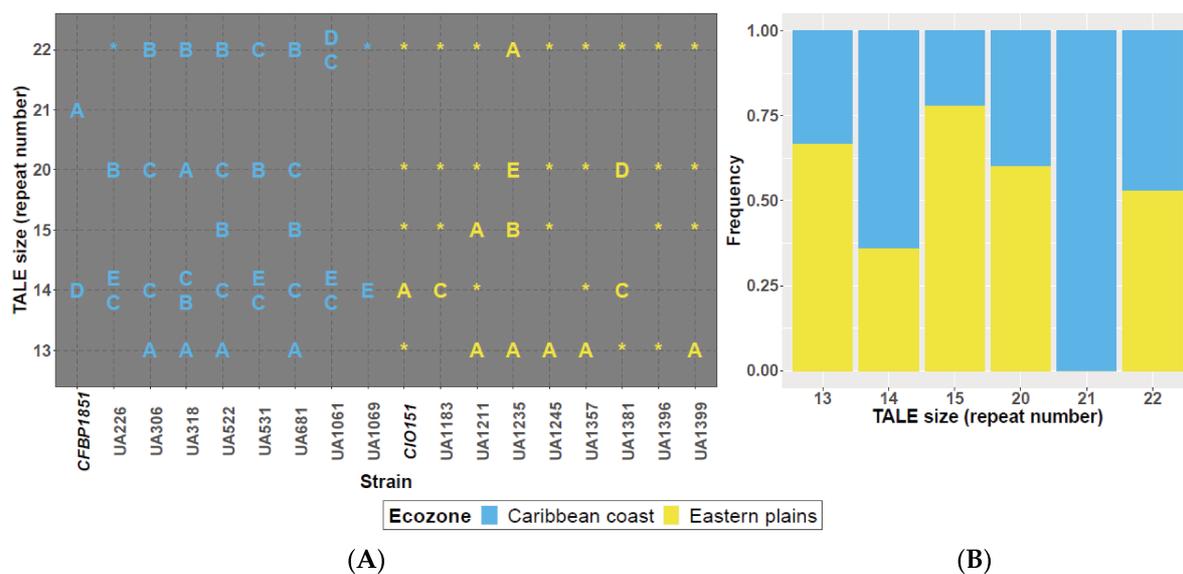


Figure 2. RFLP analysis, sequenced TALE gene variants and distribution. **(A)** TALome representation based on RFLP analysis and sequencing. Each band observed in the RFLP blot indicates the presence of at least one TALE gene. Cloned and sequenced TALE gene variants were first classified according to their size in terms of repeat number and then a letter was assigned to each variant according to the translated RVD sequence. Different letters in a row indicate different RVD strings for each TALE size. TALE genes that were only detected based on the RFLP analysis are indicated with a star. Reference strains of each ecozone are shown bolded and italicized in the *x*-axis. **(B)** TALE size prevalence according to regions. Frequencies are based on the size and number of bands observed on the RFLP analysis.

As shown in Figure 2, eighteen different TALE variants were detected in the Colombian strains. Thirteen of them are novel variants, whereas five are variants that were previously reported in *Xpm* (namely, TALE14C, TALE14D, TALE15B, and TALE22A [32,40]). Variants TALE21A, TALE14A, and TALE14D were only observed in the reference strains, collected in the 90s, a possible indication of TALE selection through time. TALE22 variants are the most commonly distributed (present in 17 out of 18 strains), followed by TALE20 and TALE14 variants (15 and 14 out of 18 strains, respectively). Interestingly, distribution of sizes according to the source edaphoclimatic zone (Figure 2A,B) indicates a bias on the TALome content for TALE13, TALE14, and TALE15 variants. TALE14 variants are more widespread among isolates from the Caribbean coast, while TALE13 and TALE15 variants are more frequently found among strains from Eastern plains. A statistically relevant correlation between the absence of TALE14 and the presence of TALE15 variants was found by Spearman correlation test ($Rho = -0.535$, p -value = 0.022), but it was not significant for TALE13 vs. TALE14 variants ($Rho = -0.378$, p -value = 0.122). Correlations between TALE presence and aggressiveness data were also tested, but no significant associations were found (data not shown). At RVD sequence level, TALE13 has a unique widespread variant, while TALE15 shows two variants differentiated by only one RVD (change of the nucleotides that code for a NN to a NS in the RVD of the twelfth repeat; NN12NS). TALE14, TALE20, and TALE22 showed at least four variants each by the coding sequence of at least one RVD and a maximum of 11, 6, and 13 RVDs, respectively. Table 3 details the RVD sequences of sequenced TALE variants.

Table 3. Translated RVD strings of the TALE gene variants sequenced in this study.

Variant	RVD Sequence	Occurrences	Strains
TALE13A	NI-NS-NN-HD-NG-HD-NI-NG-HD- NN-NI-NI-NG	9	UA306, UA318, UA522, UA681, UA1211, UA1235, UA1245, UA1357, UA1399
TALE14A	NI-NG-NI-NN-HD-HD-NG-NG-HD- NG-NG-HD-NG-NG	1	CIO151
TALE14B	NI-NG-NI-NN-HD-HD-NS-NS-NS- HD-HD-NS-HD-NG	1	UA318
TALE14C	NI-NG-NI-NN-NG-HD-NS-NS-NN- NG-HD-NN-NI-NG	9	UA226, UA306, UA318, UA522, UA531, UA681, UA1061, UA1183, UA1381
TALE14D	NI-NG-NI-NN-NI-HD-NS-NS-NN- NG-HD-NN-NI-NG	1	CFBP1851
TALE14E	NI-NS-NN-HD-NG-HD-NI-NS-NN- HD-NN-NI-NI-NG	4	UA226, UA531, UA1061, UA1069
TALE15A	NI-NG-NI-NN-HD-HD-NS-NS-NS- HD-HD-NN-HD-NG-NG	1	UA1211
TALE15B	NI-NG-NI-NN-HD-HD-NS-NS-NS- HD-HD-NS-HD-NG-NG	3	UA522, UA681, UA1235
TALE20A	NI-NG-NI-NN-HD-HD-NN-HD- NN-NG-HD-NS-HD-NN-HD-NG- HD-NI-NG-NG	1	UA318
TALE20B	NI-NG-NI-NN-NG-HD-NN-HD- NN-NG-HD-NG-HD-NN-HD-NG- HD-NN-NG-NG	2	UA226, UA531
TALE20C	NI-NG-NI-NN-NG-HD-NN-HD- NN-NG-HD-NS-HD-NN-HD-NG- HD-NI-NG-NG	3	UA306, UA522, UA681
TALE20D	NI-NG-NI-NN-NG-HD-NS-NS-NN- NG-HD-NG-HD-NN-HD-HD-NS- NI-NG-NG	1	UA1381
TALE20E	NI-NG-NI-NN-NG-HD-NS-NS-NN- NG-HD-NS-HD-NN-HD-NN-HD- NI-NS-NG	1	UA1235
TALE21A	NI-NG-HD-NG-HD-N*-NG-NG-HD- HD-NG-NG-N*-NG-HD-NG-NG- NG-HD-NG-NG	1	CFBP1851
TALE22A	NI-NG-HD-NG-NG-NG-HD-HD- NG-NG-HD-NG-HD-HD-NG-NG- HD-NG-NG-HD-NG-NG	1	UA1235
TALE22B	NI-NG-HD-NG-NG-NG-HD-HD- NG-NG-NG-NG-HD-HD-NG-NG- HD-NG-NG-HD-NG-NG	4	UA306, UA318, UA522, UA681
TALE22C	NI-NG-HD-NG-NG-NG-HD-HD- NG-NG-NG-NN-HD-NN-NG-NG- HD-NG-NG-HD-NG-NG	2	UA531, UA1061
TALE22D	NI-NG-NI-NG-NI-NN-NG-HD-NN- HD-NN-NG-HD-NG-HD-NN-HD- NG-HD-NN-NG-NG	1	UA1061

3.3. Five TALE Clusters with Similar Predicted DNA-Binding Affinities

To investigate potential DNA-binding affinities and redundant functions for TALE variants found among these isolates, we used the program FuncTAL. FuncTAL clusters TALEs based on correlations between potential target DNA sequences predicted from each RVD string [61]. Figure 3 shows the dendrogram that resulted from the analysis of the 18 TALE variants reported in Table 3. The color code used to represent RVD types in the analyzed TALEs highlights interesting general features: (i) the first repeat invariably contains NI as RVD, and most of the effectors initiate the repeat sequence with a NI-NG-NI-NN repeat block; (ii) the last half-repeat is consistently NG, and the repeat region of

the majority of variants ends with an NG-NG repeat block; (iii) NG is the most frequent RVD in *Xpm* TALEs while N* is only present in one effector extracted from reference strain CFBP1851, and (iv) repeats with NS tend to be generally restricted to central positions along the repeat array.

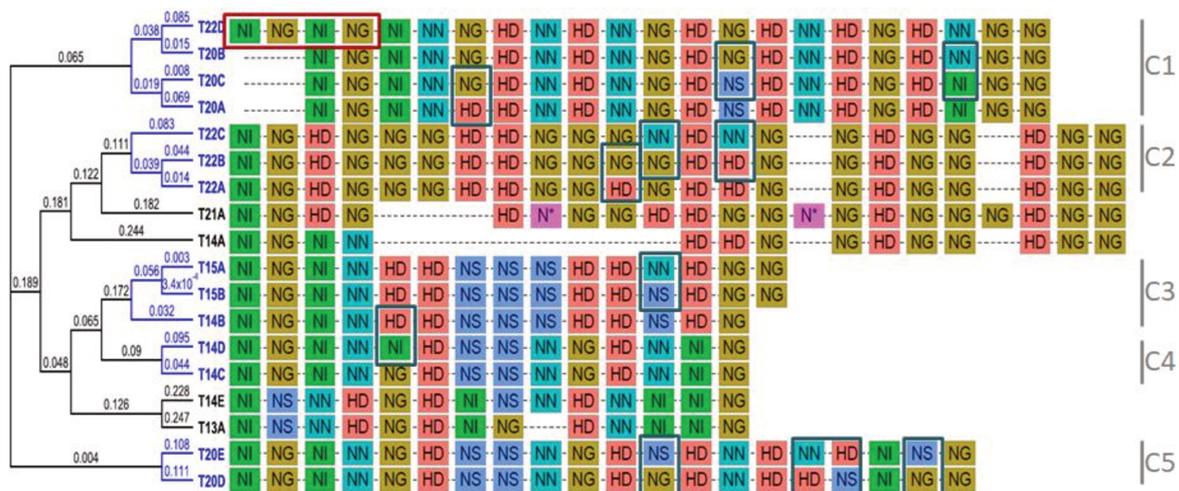


Figure 3. RVD sequence clustering based on predicted DNA-binding affinity. FuncTAL was used to estimate distances based on the predicted TALE affinity. The dendrogram on the left shows the estimated relationship, where numbers represent the distance between the position weight matrices calculated for each TALE. In branch labels, T stands for TALE. Blue branches limit each cluster, and black terminal branches indicate that the corresponding effector was not clustered (orphans). Central alignments depict RVD strings for each TALE variant, where each box represents a repeat with its RVD. Boxes are colored according to RVD type. Dotted lines show gaps inserted by the alignment procedure. Labels on the right indicate tags of clusters as defined by blue branches. The red square shows a possible duplication of the first two repeats, blue squares show repeat swaps between TALE variants.

Blue terminal branches on the dendrogram highlight five binding affinity clusters (a maximum of 3 individual RVD mismatches between each pair of compared RVD strings), which cover most of the variants for TALE15, TALE20 and TALE22 (Figure 3). In general, TALE variants in the same affinity group have the same number of RVDs. However, two out of the five clusters include variants with different number of repeats, and the alignment of RVDs shows that this is due to duplication/deletion of whole repeats at the beginning or end of the repeat region. Among them, TALE22D variant shares the whole RVD sequence with TALE20B variant, which indicates that this novel TALE might also promote transcription of the S gene *MeSWEET10a*. It is also worth noting that variants TALE20D and TALE20E exhibit a considerable number of differences relative to A, B and C variants, and instead show similarities to TALE15 and TALE14 variants. On the other hand, black terminal branches point to TALE variants whose predicted affinities are considerably different to the rest and cannot be grouped, suggesting that they do not share EBEs with TALEs from any of the clusters.

In parallel, a repeat-based phylogeny was constructed using the software DisTAL [61] (Supplementary Figure S3). Results show nearly the same clustering patterns as observed with FuncTAL, but TALE20, TALE21, and TALE22 variants are grouped into one big clade, while TALE15 and TALE14 variants (except for TALE14E) constitute another big clade. The nucleotide versions of TALE13A and TALE14E variants form a big clade that seem to be at the origin of the two above-mentioned clusters.

3.4. Expression Profiles of Plants Inoculated with UA681 and UA1061 Are Highly Similar

To gain insights into the cassava genes targeted by the TALomes of Colombian *Xpm*, we performed RNA-seq experiments of cassava leaves challenged with strains UA681 and UA1061. These strains were selected because their TALomes are fully sequenced (Figure 2),

and they collectively contain eight different TALE variants, allowing to cover four out of the five binding affinity clusters (see Figure 3), as well as the “orphan” variants TALE13A and TALE14E. Moreover, strain UA1061 lacks a TALE20 variant but it is as aggressive as others that contain TALE20, therefore, we aimed at investigating alternative explanations for this aggressiveness that could provide insight into the pathogenesis of *Xpm*. To this end, a mock solution (10-mM MgCl₂) or *Xpm* strains were inoculated into two-month old cassava plants propagated in-vitro, and tissue was harvested at 50 hpi. As illustrated in Figure 4A, expression profiles are remarkably similar and some of the most up- (Log₂FoldChange ≥ 2 and *p*-value < 0.001) and down-regulated (Log₂FoldChange ≤ −2 and *p*-value < 0.001) genes are shared by both treatments (up-regulated genes *Manes.06G123400*, *Manes.17G063800*, *Manes.07G120000*; downregulated genes *Manes.17G042600*, *Manes.05G151100*, *Manes.18G029800*). However, treatment with UA681 significantly affected the transcription level of 588 genes, while UA1061 significantly affected that of 1021 genes. When analyzed as a whole, most of the differentially expressed genes (DEGs) shared by both treatments show the same transcriptomic profile (271 out of 273 genes, Figure 4B). The top three upregulated DEGs common to both treatments are *Manes.06G123400*, *Manes.17G063800* and *Manes.07G120000* which respectively encode the *S* gene *MeSWEET10a*, an oxidoreductase 2OG-Fe II oxygenase family protein, and a wall-associated receptor kinase galacturonan-binding. The top three common downregulated DEGs encode a tyrosinase (*Manes.17G042600*), a cytochrome P450 71B21-related protein (*Manes.05G151100*) and a protein that belongs to the mitochondrial calcium uniporter family (*Manes.18G029800*). The Gene Ontology (GO) term enrichment analysis (Fisher’s test, $\alpha = 0.01$) performed on shared DEGs (Figure 4C) shows that, as defined by cell compartment, most of the differential transcriptional activity accounts for proteins that localize to the cell membrane or the apoplast. Moreover, the molecular function and biological process classifications point towards a response that involves a significant number of genes altering cell redox status, glucan metabolism, and nitrogen transport.

3.5. Potential TALE Targets Are Involved in the Manipulation of Host Cell Redox Status and Nutrient Transport

Correlation of inoculated host gene expression profiles with EBE prediction is a powerful approach to find genes targeted by TALEs. The promoterome (defined as the 1-kb regions that flank annotated translational start sites) of cassava inbred line AM560-2 [76] was used as input to predict the EBEs of each of the 18 TALE variants discovered in this study. We used four different on-line available software: TALVEZ [77], TAL Effector Nucleotide Targeter 2.0 (TALENT) [78], TALgetter [65], and PrediTALe [66]. The output of each program was restricted to the first 400 results and the rank of each predicted EBE was used as a homogenized score to enable comparisons. We cross-referenced the DUGs with the set of predictions for the TALomes of each inoculated strain. Results for UA1061 show that from the 320 DUGs, only 36 were predicted to have at least one EBE on the promoter region, while for UA681 the ratio was 69 out of 396.

The study of common DEGs could shed light on potential hubs used by the pathogen to manipulate the host. When predictions for the two TALomes and transcriptomics from plants inoculated with both strains were compared, we detected six candidate genes (Figure 5A, first six genes from the top to the bottom of the figure) with high-quality EBE predictions whose transcription could be upregulated by TALEs from clusters 1, 2, and 4 (see threshold in Figure 5A). Table 4 details the annotation and the prediction quality obtained for the proposed candidates. In the analysis of individual TALomes (Supplementary Figure S4), a clear signature prediction/expression for probable candidates was observed for 19 additional genes in the case of UA681, and 7 additional genes for UA1061 (see Table 4). Distance to the translation start site and orientation of EBEs are two factors known to affect transcriptional activation rate driven by TALEs [3,67]. Cernadas and coworkers [3] found that EBEs in real targets ranked consistently on the first 200 predictions, and their distance to annotated transcriptional start site (TSS) generally ranged between 152 bp upstream and 63 bp downstream; being this latter parameter the most predictive

descriptor. We applied a second filter by using information of EBE rank among predictors (best prediction) and distance to TSS (when available) and translational start site (TLS) to further classify potential real EBEs. Figure 5B shows EBE predictions in the promoter context of the retained candidates. Twenty candidate genes (see Table 4) out of the 34 and their predicted EBEs fell into the predictor ranges, and only three of them are located on the minus strand (reverse orientation). Among these candidates, we found potential novel targets that were not found in previous studies [8,39,41] and could be directly linked to defense regulation, like two different Abscisic Acid Receptor PYL4 encoding genes (*Manes.10G007900*, *Manes.07G135200*), a Rho GTPase-Activating Protein Ren1 encoding gene (*Manes.18G060400*), and an ammonium transporter (*Manes.08G141800*) among others.

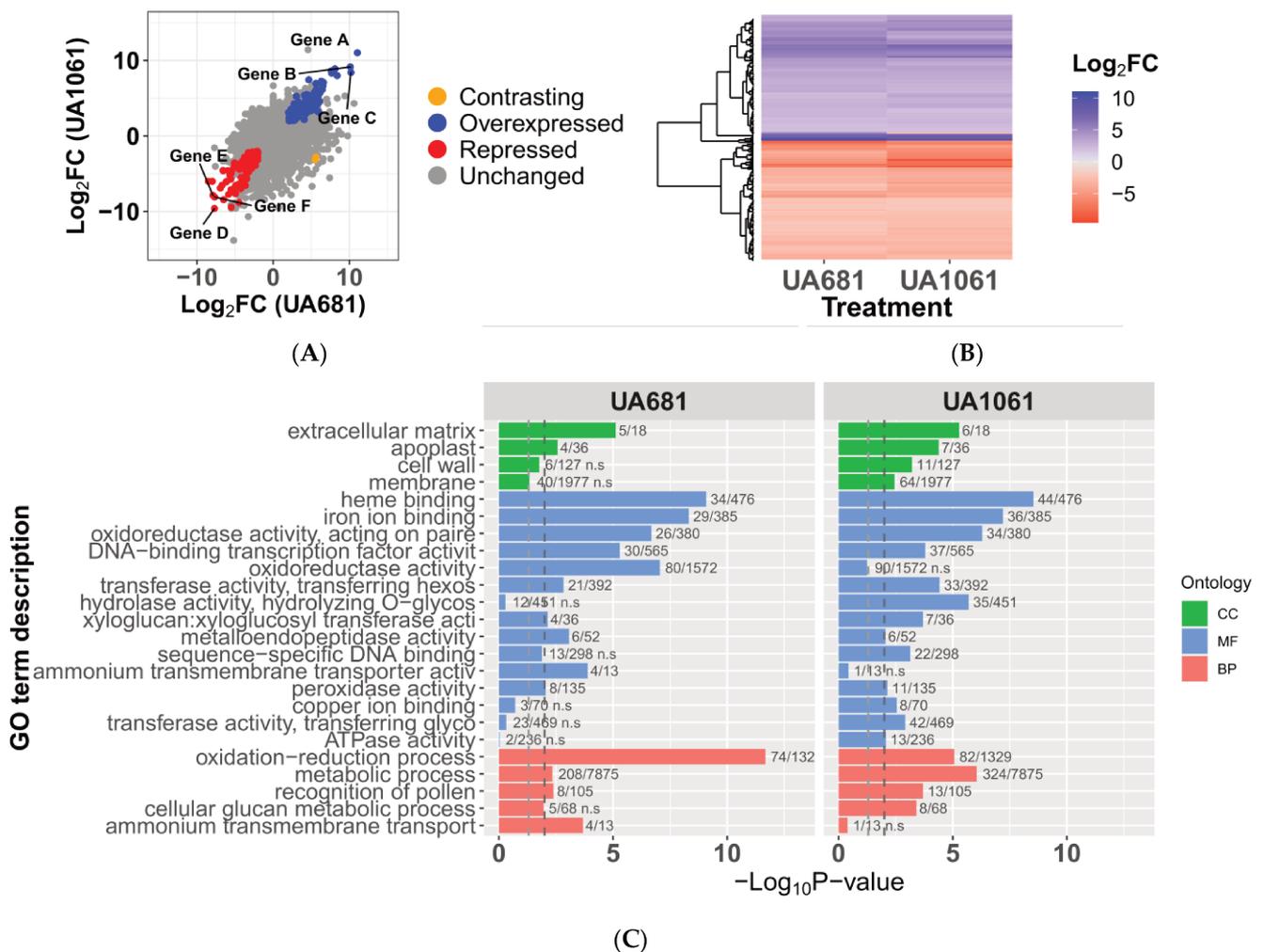


Figure 4. Overview of RNAseq data and GO-term enrichment for differentially expressed genes. **(A)** Comparison of cassava gene expression profiles observed at 50 hpi of 60444 plants inoculated with *Xpm* strains UA681 and UA1061; mock-inoculated plants were used to determine the differentially expressed genes in each case. Plot shows the shared differentially upregulated genes (DUGs) in blue and differentially down-regulated genes (DDGs) in red, while a gene that has a contrasting transcriptional behavior between the two bacterial treatments is marked in yellow. Marked genes correspond to the three shared DUGs and DDGs with most extreme response magnitudes: A: *Manes.06G123400*; B: *Manes.17G063800*; C: *Manes.07G120000*; D: *Manes.17G042600*; E: *Manes.05G151100*; F: *Manes.18G029800*. **(B)** Heatmap of the shared DUGs and DDGs transcriptional behavior. **(C)** GO-term enrichment analysis showing the cellular component (CC), molecular function (MF) and biological process (BP) categories that are significantly enriched in the differentially expressed genes in UA681- or UA1061-treated plants. The corresponding GO term IDs are summarized in the Supplementary Table S3. Discontinuous lines show the *p*-value thresholds corresponding to 0.05 (blue) and 0.01 (dark gray).

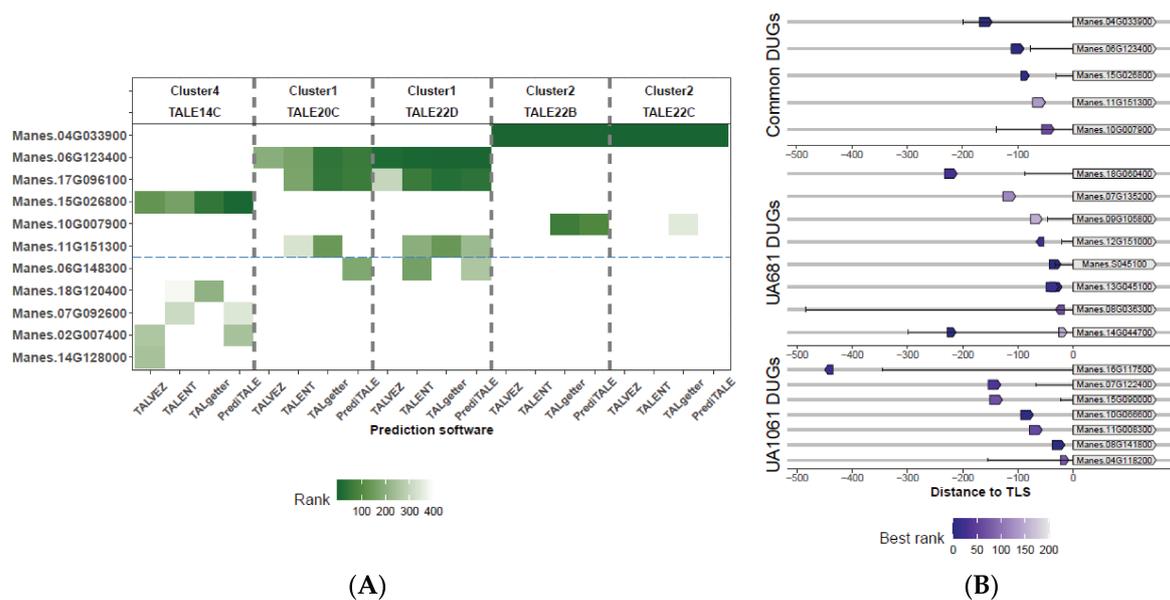


Figure 5. Differentially upregulated genes with predicted EBEs and their promoter context. **(A)** Shared DUGs with EBE predictions for TALE variants or TALE clusters shared by both strains. The heatmap shows the best candidates and the EBE-prediction quality from each software. Top labels indicate the TALE variant predicted to bind each gene and the DNA-binding affinity cluster where it belongs. Rank indicates the position of the prediction (based on the score) among the set of the top 400 predictions. Genes over the dashed line are considered as the best optioned candidates (see text). **(B)** Promoter context and predicted EBE position for candidates retained after the second filter (see text above). Genes are represented by gray arrows, while EBEs are represented by short arrows with a color code according to prediction quality. Direction of short arrows indicates if EBEs are located in the plus (pointing to right) or the minus (pointing to left) strands. Transcription start sites (as annotated on Phytozome’s cassava genome v6.1) are marked as vertical black lines connected to the translational start site with a solid black line to represent the 5’-UTR.

Since predictions were based on the well-annotated cassava genome from inbred line AM560-2, we blasted the candidate genes against the recently published genome from variety 60444 [68] to extract the corresponding promoters and check for EBE presence. Among the 34 candidates, exact EBE matches were found for 29 of them, and imperfect matches (from 1 to 3 mismatches) for three more genes. Predicted EBEs for *Manes.S013800* (coding for a peptidase of plants and bacteria) and *Manes.04G096600* (coding for a Xyloglucan:Xyloglucosyl Transferase) were not found in the 60444 promoters. All the EBE predictions, except for two cases (*Manes.09G134100* and *Manes.04G109300*), were found in similar positions relative to TLS, with a maximum shift of 30 bp (Supplementary Table S1).

To explore the effect of potential candidates on transcriptome modification from a functional point of view, a GO term enrichment analysis for each transcriptome was run keeping only DUGs that contained EBEs for the corresponding TALome (Supplementary Figure S5). For strain UA681, there is a significant (Fisher’s test, $\alpha = 0.01$) term enrichment mainly for molecular functions and biological processes related to oxidoreductase activity, indicating that TALEs might be involved in host cell redox status manipulation. However, the analysis does not show the same landscape for plants inoculated with UA1061, where none of the categories (cell compartment, molecular function, and biological process) seemed to be significantly enriched. Nevertheless, the ammonium transport is shown as a differential characteristic possibly linked to TALE activity from UA1061 TALome.

Table 4. Candidate gene targets for TALEs present in strains UA681 and UA1061 with good to excellent EBE predictions.

	Target Gene ID ^a	Annotations ^a	EBE Prediction Quality ^c						Retained After Second Filter? ^e	Responsible for Activation (Ref)		
			T13A (NC ^b)	T14E (NC ^b)	T14C (C4 ^b)	T15B (C3 ^b)	T20C (C1 ^b)	T22D (C1 ^b)			T22B (C2 ^b)	T22C (C2 ^b)
Common DUGs	Manes.04G033900	Dof Domain, Zinc Finger Protein						E	E	Yes	TAL22 [8]	
	Manes.06G123400	Bidirectional Sugar Transporter Sweet10					VG	E		Yes	TAL20 [8]	
	Manes.17G096100	No Data					VG	VG		No		
	Manes.15G026800	Membrane-Associated Kinase Regulator			E					Yes	TAL14 [8]	
	Manes.10G007900	Abscisic Acid Receptor PYL4							G	F	Yes	
	Manes.11G151300	Serine Carboxypeptidase S10					F	G			Yes	TAL14 [8]
DUGs for UA681	Manes.15G041200	12s Seed Storage Protein							VG	No		
	Manes.18G060400	Rho GTPase-Activating Protein Ren1							E	Yes		
	Manes.13G045100	Clavata3/ESR (CLE)-Related Protein							E	Yes	TAL22 [8]	
	Manes.S045100	Class IV Chitinase					E			Yes		
	Manes.04G109300	Oxidoreductase, 2OG-Fe (II) Oxygenase Family Protein				F	VG			No		
	Manes.06G159100	Ring Finger Domain (ZF-Ring_2)/Wall-Associated Receptor Kinase C-Terminal	F			G					No	
	Manes.09G134100	mlo Protein				E				No		
	Manes.14G044700	Protein Containing an AP2 Domain			F	E					Yes	TAL15 [8]
	Manes.04G096600	Xyloglucan:Xyloglucosyl Transferase			VG ^d						No	
	Manes.17G038600	No Data			E						No	
Manes.05G118500	No Data	G			F	F				No		

^a Gene identifiers and annotations were extracted from Phytozome's cassava genome version 6.1. ^b T stands for TALE. The bracketed alphanumeric code corresponds to the cluster number derived from FuncTAL analysis (Section 3.3 and Figure 3). NC, non-clustered; C1 to C4, cluster 1 through 4. ^c The quality of EBE prediction for a given candidate was calculated as the sum of the four determined ranks (one per software). Totals were then categorized as E (Excellent, combined rank = 1 to 400), VG (Very Good, combined rank = 401 to 800), G (Good, combined rank = 801 to 1200), and F (Fair, combined rank = 1201 to 1600). ^d These EBE predictions were not found in the corresponding 60444 promoter sequence. ^e EBE distance to TSS/TLS between 152 bp and 63 bp.

Table 4. Cont.

Target Gene ID ^a	Annotations ^a	EBE Prediction Quality ^c								Retained After Second Filter? ^e	Responsible for Activation (Ref)
		T13A (NC ^b)	T14E (NC ^b)	T14C (C4 ^b)	T15B (C3 ^b)	T20C (C1 ^b)	T22D (C1 ^b)	T22B (C2 ^b)	T22C (C2 ^b)		
Manes.02G120700	Protein Containing a Tetratricopeptide Repeat (TPR_16)	G								No	
Manes.12G151000	No Data	VG								Yes	
Manes.09G105800	Protein Containing a Myb-Like DNA-Binding Domain and a Myb-CC Type Transfactor, LHEQLE Motif					G				Yes	
Manes.18G074600	No Data					G				No	
Manes.S013800	Basic Secretory Protein Family/Peptidase of Plants and Bacteria					G ^d				No	
Manes.06G057400	Equilibrative Nucleoside Transporter							G		No	
Manes.07G135200	Abscisic Acid Receptor PYL4							G		Yes	
Manes.04G053400	Galactolipase/Phospholipase A (1)					F		G		No	
Manes.08G036300	Calcium-Dependent Protein Kinase					G				Yes	
Manes.07G122400	No Data								VG	Yes	
Manes.08G141800	Ammonium Transporter 1							E		Yes	
Manes.11G008300	Member Of 'GDYG' Family of Lipolytic Enzymes							VG		Yes	
DUGs for UA1061	Manes.10G066600	2-Hydroxyisoflavanone Dehydratase						E		Yes	
	Manes.16G117500	No Data			VG					Yes	
	Manes.13G136600	Coniferyl-Alcohol Glucosyltransferase			E					No	
	Manes.04G118200	Protein of Unknown Function (Duf642)		G ^d						Yes	
	Manes.15G090000	Homeobox-Leucine Zipper Protein ATHB-52						G		Yes	TAL14 [8]

^a Gene identifiers and annotations were extracted from Phytozome's cassava genome version 6.1. ^b T stands for TALE. The bracketed alphanumeric code corresponds to the cluster number derived from FuncTAL analysis (Section 3.3 and Figure 3). NC, non-clustered; C1 to C4, cluster 1 through 4. ^c The quality of EBE prediction for a given candidate was calculated as the sum of the four determined ranks (one per software). Totals were then categorized as E (Excellent, combined rank = 1 to 400), VG (Very Good, combined rank = 401 to 800), G (Good, combined rank = 801 to 1200), and F (Fair, combined rank = 1201 to 1600). ^d These EBE predictions were not found in the corresponding 60444 promoter sequence. ^e EBE distance to TSS/TLS between 152 bp and 63 bp.

3.6. Candidate Gene Expression Profiles Are in Agreement with Trends Observed in RNAseq and Previous Reports

We confirmed the expression profile of some candidates through RT-qPCR with RNA extracted (at 50 hpi) from adult plants inoculated with UA681. The candidate genes were *Manes.13G045100*, *Manes.04G033900*, *Manes.11G151300*, *Manes.06G123400*. The four candidates have strong predictions; but first two genes have fold-changes near 2 (threshold to define upregulated genes), while the latter two have large fold-changes (6.19 and 11.06, respectively). Relative overexpression is statistically significant (Figure 6A), and expression patterns are well correlated between RNAseq data and RT-qPCR data (Figure 6B). Cohn and colleagues [8,39] demonstrated the TALE-dependent overexpression of 28 cassava genes in the Xam668-cassava infection context. RVD sequences of the TALEs of the UA681 strain are very similar to the ones present in Xam668: TAL13, TAL14 and TAL15 have the same RVD sequence, while TAL20 differs in 4 central RVDs and TAL22 shows the HD11NG change. Cross-reference of validated targets with the set of common and UA681 DUGs shows that 19 of these validated candidates are present in this set of transcriptionally activated genes, but only seven of them had predicted EBEs. Previously unreported candidates found in this study might be specific targets of TAL20C and TAL22B variants. The last column of Table 4 shows the validated candidates and which TALE is potentially responsible for their transcriptional activation.

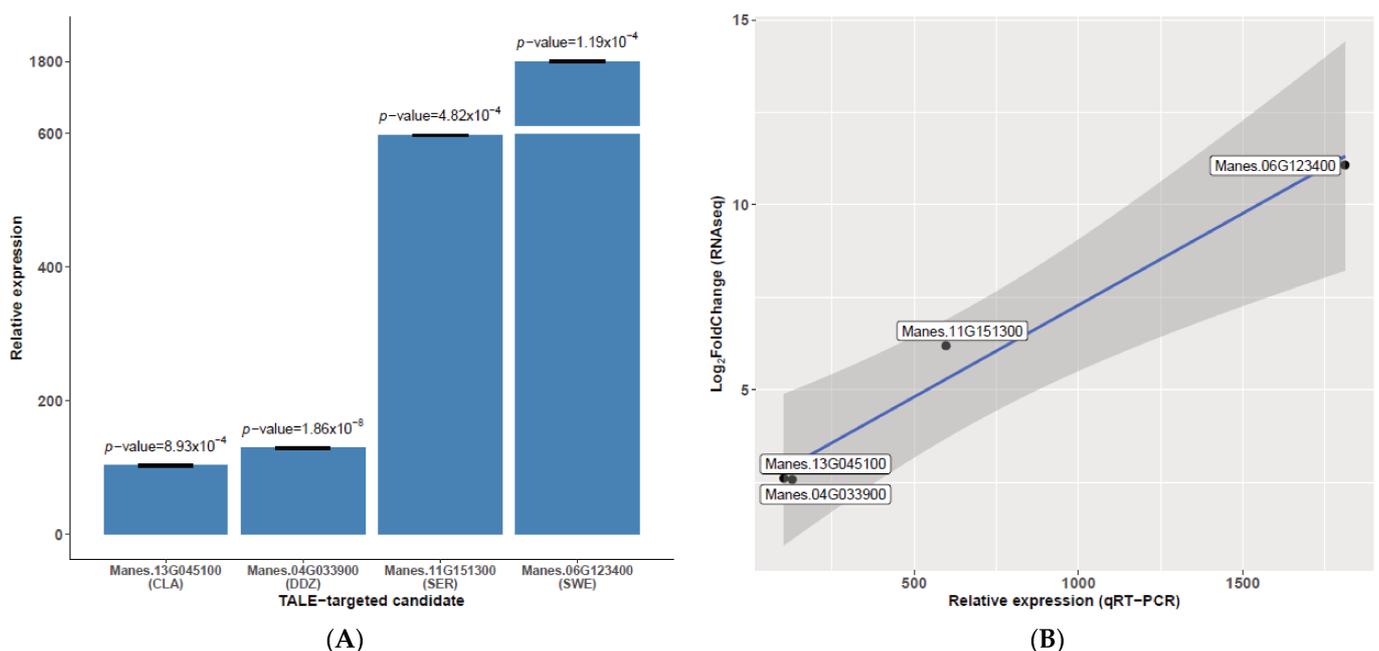


Figure 6. RT-qPCR validation of expression profile for selected genes. (A) Relative expression of four DUGs in UA681-inoculated vs. mock-inoculated plants. Main bars show the average of three independent replicates, and error bars show standard deviation. Note that y-axis has been broken to allow visualization of less expressed genes. p-values correspond to a two-tailed *t*-test ($\alpha = 0.05$) comparing normalized expression in the UA681-inoculated and mock-inoculated treatments. X axis shows gene identifiers and a three-letter code to denote each gene. (B) Relative expression tested by RT-qPCR plotted against RNAseq profiles for the selected candidates. White boxes show the gene identifiers.

3.7. Potential Functional Convergence within TALome Members and Among Expanded TALome

Functional convergence among TALEs is defined as two or more distantly related TALEs targeting the same gene, using the same, overlapped or completely different EBEs. This feature can be assessed at two different levels: (i) intra-TALome, for genes that are targeted by at least two TALEs expressed by the same bacterium (functional redundancy)

and (ii) among the TALE variants present in the bacterial population. We did not find any cases of intra-TALome functional convergence. However, analyses extended to all the TALEs isolated in this study allowed us to track potential cases of functional convergence. Due to marked similarities between the affinity of clusters 1 and 5 (both harboring TALE20 variants), these convergence analyses did not take them as individual clusters. This analysis considered only DUGs (from UA681 or UA1061 treatments) with EBE for two or more TALEs from different clusters and resulted in eight candidates (Supplementary Table S2). The gene *Manes.04G033900*, which encodes a Dof Domain, Zinc Finger Protein, showed the strongest evidence for convergence. This gene was upregulated in both treatments and has been already validated as an actual target for TALE22_{Xam668} [8]. Theoretically, this gene is targeted by TALEs from cluster 2 (TALE22A, TALE22B and TALE22C) and TALE21A. Likewise, a second validated target for TALE22_{Xam668}, upregulated in both RNAseq treatments and with EBE predictions for TALE21A, *Manes.13G045100* (coding for a Clavata3/ESR (CLE)-Related Protein), was found among the potential candidates. EBE predictions for TALE21A are almost totally overlapped with EBE predictions for cluster 2, sharing 20 base pairs in both cases (Figure 7A). Since predictions are well-supported and these genes were already validated as real TALE22 targets, we decided to test activation of these genes by inoculating cassava leaves with strain CFBP1851, a strain that lacks TALE22 variants but contains TALE21A. RNA was extracted after 72 hpi (since aggressiveness of this strain is considerably low. See Figure 1) and semiquantitative RT-PCR was performed on the derived cDNA. Figure 7B shows results for both genes. Strain CFBP1851 was able to differentially upregulate gene *Manes.13G045100* when compared to the mock-inoculated treatment, indicating that this is probably a case of TALE convergence (still, activation by a non-TALE factor cannot be ruled out with this experiment). However, transcripts from gene *Manes.04G033900* were not detected by the RT-PCR in plants treated with CFBP1851, indicating that TALEs from this strain are not able to activate this target.

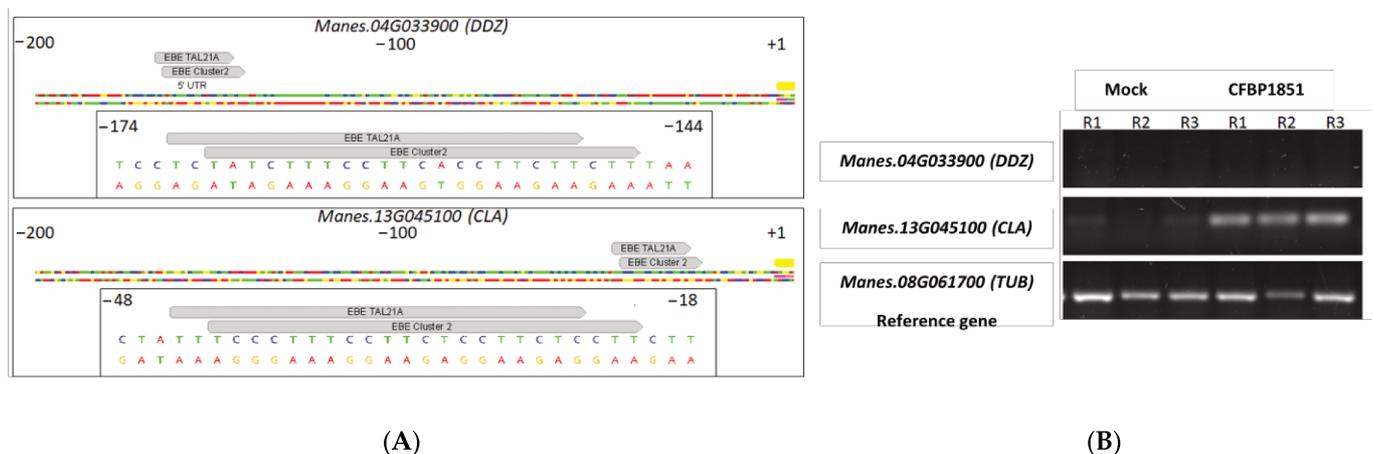


Figure 7. EBE predictions and semiquantitative RT-PCR in the context of a convergent activation of two candidates by two different TALEs. **(A)** Overlapping EBEs for TALE21A and TALE variants of the cluster 2 in the context of promoters for *Manes.04G033900* and *Manes.13G045100* genes. Each case is displayed as an overview and a zoom on the region of interest. EBEs are shown as gray arrows pointing to the right. CDS is represented as a yellow band, while 5-UTR region (only annotated for *Manes.04G033900*) is shown as a gray band. TSS or TSL was centered on the nucleotide number 1000, so coordinates are a reflect of this setting. **(B)** Semiquantitative RT-PCR for genes *Manes.04G033900* and *Manes.13G045100* on cDNA from cassava leaves inoculated with CFBP1851 or with a 10-mM MgCl₂ solution (Mock). R1, R2 and R3 are independent replicates of the inoculation. Tubuline (TUB—*Manes.08G061700*) was used as the reference gene for normalization.

4. Discussion

This study started with a selection of *Xpm* strains based on neutral genetic diversity at a country scale that allowed us to depict the TALE variation among strains isolated from different cassava productive regions over different years in Colombia. Southern blot analysis showed that *Xpm* strains carry TALomes with a bimodal size distribution, and sequence analysis showed thirteen novel TALE variants. Sequence clustering by predicted DNA-binding affinity allowed us to condense this variability into five clusters and four unrelated TALE variants. A complementary transcriptomic approach allowed us to move from diversity to function and to screen host genes that might act as CBB susceptibility determinants. Transcriptomics showed that redox processes and nutrient transport are upregulated during infection, and some of the involved genes are potentially targeted by TALEs. Importantly, our results also suggest that TALEs different from TALE20 might activate the transcription of *MeSWEET10a* and potentially confer aggressiveness to those strains lacking TALE20 in *Xpm* populations. This study sets new bases for assessing unidentified TALE targets in the *Xpm*-cassava interaction, which could be important factors that define the fate of the infection.

TALEs play major roles in the pathogenesis of several xanthomonads (reviewed by [4]), mainly by conferring bacteria the ability to create a favorable niche in the plant vascular system. In some cases, *Xanthomonas* spp. pathogenicity is drastically altered by the activity of a given TALE [69,79]. Bart and coworkers [32] found a positive trend between the number of TALEs per TALome vs. aggressiveness and virulence assessed in 18 strains of *Xpm*. In our study, we profiled the aggressiveness of 18 *Xpm* strains and found that only CFBP1851 shows reduced capacity to cause lesions. Interestingly, this strain does not carry TALE20 or any TALE20 variant that could activate the *S* gene *MeSWEET10a* [8], which could explain reduced aggressiveness. Strains UA1061 and UA1069 also lack this TALE variant, but their aggressiveness is average. However, UA1061 possess two variants of TALE22, one of them with an RVD sequence similar to that of the TALE20B variant plus a duplication of the first two repeats (Figures 2 and 3). This new TALE22 variant (TALE22D) could explain the fact that strain UA1061, and potentially UA1069, do not have a decrease in aggressiveness even if they lack TALE20 variants. The DisTAL analysis (Supplementary Figure S3) indicates that TALE22D shares ancestry with TALE20B, and, based on the transcriptomic profile of UA1069-inoculated plants and the EBE predictions, this TALE seems to be able to induce *MeSWEET10a*. However, despite the reduced ability of CFBP1851 to cause symptoms, it still induces water-soaked lesions, indicating that there might be mechanisms that do not rely on *MeSWEET10a* transcriptional activation to promote water soaking, yet less efficiently.

As seen in Southern blot analysis, *Xpm* TALomes are characterized by a limited and bimodal range of variant sizes. This bimodal size distribution could be explained by a recent acquisition of TALEs from two different ancestors or the fixation of two useful alleles that then diverged. Our data favors the second hypothesis since the repeat-based phylogeny (Supplementary Figure S3) shows two clades that group TALEs by size, both emerging from the TALE13 variants. This is probably why there are shared repeat block patterns between the two TALE size groups (Figure 3). Ferreira and collaborators [80] created a TALE-based phylogeny with 122 coding sequences from *Xanthomonas* spp. TALEs. *Xanthomonas phaseoli* pv. *phaseoli* and *Xanthomonas citri* pv. *fuscans*, two pathogens closely related to *Xpm*, show unimodal and continuous TALE size distribution that ranges from 18 to 23 repeats [81]. N-, C-terminal and TALE flanking region analysis suggests that these TALE variants come from a unique common ancestor [81]. *X. campestris* (*Xc*) exhibits a more complex TALE size distribution which however seems to be even from 11 to 22 repeats. Yet, analysis of N-, C-terminal and repeats indicate that *Xc* TALEs emerged from two different ancestors [82], where smaller TALEs (variants with 12, 14 and 15 repeats) have structurally different repeats from larger TALEs with 21 and 22 repeats. Since N- and C-terminal sequences are not available in our study and repeat structure is highly homogeneous, we could not perform such analyses.

TALE sequence diversity was assessed by potential DNA-binding affinity using FuncTAL [61], resulting in five DNA affinity groups and four orphan (non-clustered) effectors. Among them, there are several examples of deletion/duplication, like the case of TALE22D from cluster 1 (discussed earlier), TALE14B in the cluster 3 and TALE13A when compared to TALE14E (see Figure 3). Repeat recombination patterns are also observed; for example, TALE14A variant shows the most common starting RVD block NI-NG-NI-NN, but the remaining block of ten RVDs can be found in TALE22A and B variants. Repeat swaps (non-synonymous point mutations affecting the RVD) seem to be a key variability driver for TALE diversity in *Xpm*, since all the clusters show at least two variants with one to four swaps (see Figure 3). These evolutionary mechanisms for TALEs have been well documented [83,84], and several examples of deletion/duplication, repeat swap and recombination can be found among *Xanthomonas* spp. TALEs [6,81,82].

To elucidate the effects of this functional diversity, we selected two strains whose TALomes covered most of the TALE diversity. Strains UA681 and UA1061 were inoculated in 60,444 varieties to obtain a transcriptomic snapshot of the infection. Briefly, plant response to infection with either of the strains shows a marked alteration of oxidoreductase activity, DNA-transcription factor activity, interaction with metals such as iron and copper, and carbon catabolism. Muñoz-Bodnar and collaborators compared the transcriptomic profiles of cassava plants from the variety MCOL2215 inoculated with two *Xpm* strains with contrasting virulence [41]. In that study, *Xpm* transcriptomic profiles did not significantly differ and showed a transcriptional alteration of genes related to catabolic processes, cellular glucan metabolism, transcription regulation and response to oxidative stress, among other categories. A transcriptomic study in *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) showed altered expression of genes involved in catalytic and oxidoreductase activities, which suggest an active reactive oxygen species detoxification and cellular redox status control [3]. Taken together, this indicates that alteration of the redox environment and catabolic processes are common to the compatible interactions between *Xanthomonas* spp. and their hosts, highlighting their relevance in pathogenesis.

The analysis of the DUGs coupled to EBE predictions resulted in several candidates that could be important for *Xpm* pathogenesis. Some of these candidate genes were previously validated as actual targets of *Xpm* TALEs (see Table 4), which supports our findings. The sugar transporter encoding *MeSWEET10a* is a DUG highly overexpressed in both treatments (UA681 and UA1061), and it has been shown to play a key role in symptom formation and bacterial growth *in planta* [8]. Our data provide indirect evidence that the new TALE20-derived TALE22D variant is functional and capable of targeting this important *S* gene. In this regard, we explored if the addition of these two repeats at the beginning would negatively affect the predicted DNA-binding affinity on the already known EBE (EBE for TAL20: TATAAACGCTTCTCGCCCATC). Results show that the two nucleotides located right before the EBE for TAL20 are the same first two nucleotides of the EBE (EBE for TAL22D: TATATAAACGCTTCTCGCCCATC), allowing then a perfect match for the new variant. We foresee that these changes might affect TALE-driven induction dynamics, but it is difficult to establish if these will result beneficial for the molecular interaction. Repeats located near to the N-terminal region contribute more to the specific DNA recognition and are more important to the overall affinity of the TALE-DNA interaction [85,86], but, paradoxically, increasing the number of repeats results in lower specificity due to more off-target activation [87]. What implications this change could have in the intensity of binding induction of the gene or in the adaptation of the pathogen to mutations in the promoter needs to be determined. A comparison of *MeSWEET10a* induction in plants treated with strains expressing these two TALE variants would indicate if this mutation has a significant impact on transcriptional activity of the *S* gene, and the search for this novel variant among populations would shed light on evolutionary dynamics of major TALEs.

The ammonium transporter encoding gene *Manes.08G141800* is one of the most interesting candidates targeted by this new TALE variant. Grewal and coworkers demonstrated

that rice response after a few hours of being inoculated with *Xanthomonas oryzae* pv. *oryzae*, a vascular pathogen responsible for the bacterial leaf blight in rice, encompasses the transcriptional alteration of several transporters, including ammonium transporters. Activity of these transporters change ion fluxes and affect the membrane potential, disturbing uptake of other channels and altering activation of defense response. In this regard, they point out that ammonium transporter activity affects K^+ uptake by cells, which would lead to modified cation balances [88]. It is possible that these ammonium transporters have an analogous effect to that hypothesized for the sulfate transporter coding gene *OsSULTR3;6* induced by Tal2g in the *Xoc*-rice pathosystem [3]. The expression of these genes might allow the bacteria to alter antioxidant capacity through ion leaking, which would result in interference with redox signaling and defense or in the promotion of water soaking through alteration of osmotic equilibrium.

Functional redundancy by convergence has been demonstrated for TALEs in other pathosystems. In those cases, important susceptibility genes are targeted by two or more TALEs that have potentially evolved from different ancestors and converge on EBEs present in the same promoter region [1,7,53,89–91]. Our data suggested that candidate genes *Manes.04G033900* and *Manes.13G045100* are targeted by TALE22 variants (which was already validated by Cohn and coworkers [8]) and also by the TALE21A variant found in CFBP1851, through overlapped EBEs. Our RT-PCR results indicate that *Manes.13G045100* is actually upregulated *in planta* after inoculation with CFBP1851 strain (Figure 7B), showing that this is a potential case of functional convergence and a possible relevant target for the pathogen. However, even if at a first glance TALE21A variant has important structural differences when compared to TALE22 variants, DisTAL analysis indicates that variant TALE21A is very close to TALE22 variants, which would reflect the history of close evolving effectors that maintain affinities. *Manes.13G045100* encodes for a CLAVATA3/ESR (CLE) peptide, a protein family of ligands involved in definition of cell fate with a key role in cellular division control in the shoot apical meristem [92,93]. This ligand family has been shown to be important in symbiotic processes through the regulation of nodule formation, and in pathogenic interactions where phytopathogens hijack the CLE-mediated developmental signaling pathway to improve fitness and colonization rate (reviewed by [94]). Nematodes *Globodera rostochiensis* [95] and *Heterodera glycines* [96] secrete effectors that mimic CLE functions which allow them to reprogram cell growth at the point of infection. To our knowledge, there are no reports of phytopathogenic bacteria using this pathway as a virulence factor, but there are other examples of cell fate-related transcriptional factors, *UPA20* and *UPA7* [12,97], which are also transcriptionally activated by TALEs (AvrBs3). Taking all this into consideration, validation and further studies on this TALE target could shed light on new host susceptibility pathways.

5. Conclusions

This work describes for the first time the TALE variability on *Xpm* populations. TALome sequences allowed us to make some evolutionary and functional inferences on TALE biology for the *Xpm*-cassava pathosystem. Functional clustering of TALEs unveiled traces of recombination, repeat swaps and repeat duplication/deletion. Transcriptomics indicate that new TALE20-derived TALE22D variant preserves the ability to activate the *S* gene *MeSWEET10a*. RNAseq with two *Xpm* strains showed that common up- and down-regulated genes have virtually the same behavior in terms of expression magnitude and regulation, and differentially expressed genes fall consistently in oxidoreductase and catabolic activity categories. Our results pinpoint new potential TALE targets that could have a role in pathogenesis. Further studies are needed to validate the TALE-dependent activation of these targets through the use of mutant strains inactivated in these TALEs and characterize their role during infection using artificial TALEs.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2076-2607/9/2/315/s1>, Figure S1: Two edaphoclimatic zones (ECZs) where cassava is grown in Colombia and strain origins. The blue region corresponds to the Caribbean coast ECZ delimited as two departments:

Córdoba and Sucre. The yellow region corresponds to the Eastern plains ECZ delimited as two departments: Casanare and Meta. Black dots show municipalities where strains were collected. The list of strains from each ECZ includes the collection date in parentheses. Figure S2: Southern blots for the characterized strains. The upper panel shows lanes for strains from the Caribbean coast and strains isolated from the department Meta (Eastern plains). The lower panel shows lanes for the strains from the department Casanare (Eastern plains). The ID of each strain is marked above each lane. MWM = molecular weight marker; TALE1_{Xam} = pF3 vector containing TALE1_{Xam}. This vector was also digested with *Bam*HI and loaded into each gel as a positive control and size indicator. The molecular weight marker is presented as seen in the stained gel to estimate sizes. Figure S3: Repeat-based phylogeny created by DisTAL. Unrooted cladogram showing the phylogenetic relationship among nucleotide variants according to nucleotide composition of repeats after masking positions coding for RVDs. Labels indicate the variant classification according to the RVD string (T stands for TALE) and the strain where the gene was isolated. Colors are related to the clusters defined in Figure 3: blue = Cluster 1, magenta = Cluster 2, red = Cluster 3, cyan = Cluster 4, yellow = Cluster 5; orphans are all in black. Values on branches correspond to distances obtained from pair-wise alignments for repeat-based coded TALEs calculated by DisTAL. Figure S4: TALE-targeted candidate prediction for A) UA681 and B) UA1061 DUGs. These heatmaps show the best optioned candidates and their EBE-prediction quality from each software used. Top labels indicate the TALE variant that is predicted to bind each gene and the DNA-binding affinity cluster where it belongs. Rank indicates the position of the prediction (which is a reflect of the score) among the set of the top 400 predictions. Genes over the dashed horizontal line are considered as the best optioned candidates (see text). Figure S5: Selective GO-term enrichment analysis showing the cellular component (CC), molecular function (MF) and biological process (BP) categories that are significantly enriched with DEGs carrying EBEs on their promoters in UA681- or UA1061-treated plants. Discontinuous lines show the *p*-value thresholds corresponding to 0.05 (blue) and 0.01 (dark gray). Table S1. EBE position comparison between the promoter context of genes reported for varieties AM560-2 and 60444 genomes. Table S2. Potential cases of functional convergence among DUGs with predicted EBEs. Candidates are sorted from the top to the bottom according to the EBE prediction quality. Table S3. ID and short description of GO Terms displayed on Figure 4 and Figure S5. The order of GO Terms matches the order displayed on the *y*-axis of both figures.

Author Contributions: D.O.-R. performed strain selection, isolation, and sequencing of a number of TALEs; R.E.M. carried out the RNAseq experiments in in-vitro plants; Á.L.P.-Q. analyzed the RNAseq data and provided advice on TALVEZ and QueTAL suite use; A.D. processed the RNAseq data and cross-referenced this information with prediction data for further analyses. C.A.Z.-C. performed aggressiveness assessments, isolation, and sequencing of TALEs, data processing and analyses, and wrote the initial draft of the article; A.B. conceived the project, secured funding, directed experiments, and helped draft the manuscript in collaboration with C.E.L. and B.S.; S.R., C.E.L., B.S. and A.B. wrote and reviewed sections of the manuscripts. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Colombian Ministry of Sciences (Former Colciencias), through project No. 315-2013; by Universidad de los Andes and its Basic Sciences funding through multiple research projects (seed project NUMBER assigned to CAZ and Project No. INV-2019-84-1855 assigned to AB); the Research Program on Roots, Tubers and Bananas, and the Agropolis Foundation (project #1403-073). CAZ was supported by a doctoral fellowship awarded by the Insitut de Recherche pour le Développement.

Data Availability Statement: Publicly available datasets were analyzed in this study. TALE nucleotide sequences can be found at GenBank under accession numbers MW413757-MW413802. RNA-Seq data can be found in the SRA database under the BioProject No. PRJNA688032, BioSample No. SAMN17167922 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA688032>).

Acknowledgments: The *Xanthomonas* strains were collected and investigated under the contract for access to genetic resources from the Colombian Environmental Ministry Number 261 or 2019 with Universidad de los Andes. We acknowledge the contribution of undergraduate students from Universidad de los Andes Luisa Jaimes Niño and María Camila Buitrago during their internships. We also thank Adolfo Amezcua for statistical advice.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Antony, G.; Zhou, J.; Huang, S.; Li, T.; Liu, B.; White, F.; Yang, B. Rice *xa13* Recessive Resistance to Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene *Os-11N3*. *Plant Cell* **2010**, *22*, 3864–3876. [[CrossRef](#)] [[PubMed](#)]
2. Hu, Y.; Zhang, J.; Jia, H.; Sosso, D.; Li, T.; Frommer, W.B.; Yang, B.; White, F.F.; Wang, N.; Jones, J.B. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E521–E529. [[CrossRef](#)] [[PubMed](#)]
3. Cernadas, R.A.; Doyle, E.L.; Niño-Liu, D.O.; Wilkins, K.E.; Bancroft, T.; Wang, L.; Schmidt, C.L.; Caldo, R.; Yang, B.; White, F.F.; 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. [[CrossRef](#)] [[PubMed](#)]
4. Perez-Quintero, A.L.; Szurek, B. A Decade Decoded: Spies and Hackers in the History of TAL Effectors Research. *Annu. Rev. Phytopathol.* **2019**, *57*, 459–481. [[CrossRef](#)] [[PubMed](#)]
5. Yang, B.; Sugio, A.; White, F.F. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10503–10508. [[CrossRef](#)] [[PubMed](#)]
6. Tran, T.T.; Pérez-Quintero, A.L.; Wonni, I.; Carpenter, S.C.D.; Yu, Y.; Wang, L.; Leach, J.E.; Verdier, V.; Cunnac, S.; Bogdanove, A.J.; et al. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **2018**, *14*, e1007092. [[CrossRef](#)]
7. Yu, Y.; Streubel, J.; Balzergue, S.; Champion, A.; Boch, J.; Koebnik, R.; Feng, J.; Verdier, V.; Szurek, B. 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*, 1102–1113. [[CrossRef](#)]
8. Cohn, M.; Bart, R.S.; Shybut, M.; Dahlbeck, D.; Gomez, M.; Morbitzer, R.; Hou, B.-H.; Frommer, W.B.; Lahaye, T.; Staskawicz, B.J. *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*, 1186–1198. [[CrossRef](#)]
9. Sugio, A.; Yang, B.; Zhu, T.; White, F.F. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIA γ 1* and *OsTFX1* during bacterial blight of rice. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10720–10725. [[CrossRef](#)]
10. Schwartz, A.R.; Morbitzer, R.; Lahaye, T.; Staskawicz, B.J. TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E897–E903. [[CrossRef](#)]
11. Cox, K.L.; Meng, F.; Wilkins, K.E.; Li, F.; Wang, P.; Booher, N.J.; Carpenter, S.C.D.; Chen, L.-Q.; Zheng, H.; Gao, X.; et al. TAL effector driven induction of a *SWEET* gene confers susceptibility to bacterial blight of cotton. *Nat. Commun.* **2017**, *8*, 15588. [[CrossRef](#)] [[PubMed](#)]
12. 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* **2007**, *318*, 648–651. [[CrossRef](#)] [[PubMed](#)]
13. Boch, J.; Bonas, U. *Xanthomonas* AvrBs3 Family-Type III Effectors: Discovery and Function. *Annu. Rev. Phytopathol.* **2010**, *48*, 419–436. [[CrossRef](#)] [[PubMed](#)]
14. Bogdanove, A.J.; Schornack, S.; Lahaye, T. TAL effectors: Finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* **2010**, *13*, 394–401. [[CrossRef](#)]
15. Szurek, B.; Rossier, O.; Hause, G.; Bonas, U. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol. Microbiol.* **2002**, *46*, 13–23. [[CrossRef](#)]
16. Yang, Y.; Gabriel, D.W. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. *Mol. Plant-Microbe Interact.* **1995**, *8*, 627–631. [[CrossRef](#)]
17. Zhu, W.; Yang, B.; Chittoor, J.M.; Johnson, L.B.; White, F.F. AvrXa10 Contains an Acidic Transcriptional Activation Domain in the Functionally Conserved C Terminus. *Mol. Plant-Microbe Interact.* **1998**, *11*, 824–832. [[CrossRef](#)]
18. Deng, D.; Yan, C.; Wu, J.; Pan, X.; Yan, N. Revisiting the TALE repeat. *Protein Cell* **2014**, *5*, 297–306. [[CrossRef](#)]
19. Schandry, N.; Jacobs, J.M.; Szurek, B.; Perez-Quintero, A.L. A cautionary TALE: How plant breeding may have favoured expanded TALE repertoires in *Xanthomonas*. *Mol. Plant Pathol.* **2018**, *19*, 1297–1301. [[CrossRef](#)]
20. Mak, A.N.-S.; Bradley, P.; Bogdanove, A.J.; Stoddard, B.L. TAL effectors: Function, structure, engineering and applications. *Curr. Opin. Struct. Biol.* **2013**, *23*, 93–99. [[CrossRef](#)]
21. Mak, A.N.-S.; Bradley, P.; Cernadas, R.A.; Bogdanove, A.J.; Stoddard, B.L. The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. *Science* **2012**, *335*, 716–719. [[CrossRef](#)] [[PubMed](#)]
22. Deng, D.; Yan, C.; Pan, X.; Mahfouz, M.; Wang, J.; Zhu, J.-K.; Shi, Y.; Yan, N. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **2012**, *335*, 720–723. [[CrossRef](#)] [[PubMed](#)]
23. Moscou, M.J.; Bogdanove, A.J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **2009**, *326*, 1501. [[CrossRef](#)] [[PubMed](#)]
24. Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **2009**, *326*, 1509–1512. [[CrossRef](#)] [[PubMed](#)]
25. Yuan, M.; Ke, Y.; Huang, R.; Ma, L.; Yang, Z.; Chu, Z.; Xiao, J.; Li, X.; Wang, S. A host basal transcription factor is a key component for infection of rice by TALE-carrying bacteria. *eLife* **2016**, *5*, e19605. [[CrossRef](#)]
26. Peng, Z.; Hu, Y.; Zhang, J.; Hugueta-Tapia, J.C.; Block, A.K.; Park, S.; Sapkota, S.; Liu, Z.; Liu, S.; White, F.F. *Xanthomonas translucens* commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 20938–20946. [[CrossRef](#)]

27. Chen, L.-Q.; Hou, B.-H.; Lalonde, S.; Takanaga, H.; Hartung, M.L.; Qu, X.-Q.; Guo, W.-J.; Kim, J.-G.; Underwood, W.; Chaudhuri, B.; et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **2010**, *468*, 527–532. [[CrossRef](#)]
28. El Kasmi, F.; Horvath, D.; Lahaye, T. Microbial effectors and the role of water and sugar in the infection battle ground. *Curr. Opin. Plant Biol.* **2018**, *44*, 98–107. [[CrossRef](#)]
29. Duan, S.; Jia, H.; Pang, Z.; Teper, D.; White, F.; Jones, J.; Zhou, C.; Wang, N. Functional characterization of the citrus canker susceptibility gene *CsLOB1*. *Mol. Plant Pathol.* **2018**, *19*, 1908–1916. [[CrossRef](#)]
30. Constantin, E.C.; Cleenwerck, I.; Maes, M.; Baeyen, S.; Van Malderghem, C.; De Vos, P.; Cottyn, B. 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. [[CrossRef](#)]
31. CABI. *Xanthomonas axonopodis* pv. *manihotis* (cassava Bacterial Blight). Available online: www.cabi.org/isc (accessed on 1 October 2020).
32. Bart, R.; Cohn, M.; Kassen, A.; McCallum, E.J.; Shybut, M.; Petriello, A.; Krasileva, K.; Dahlbeck, D.; Medina, C.; Alicai, T.; et al. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1972–E1979. [[CrossRef](#)]
33. Restrepo, S.; Valle, T.L.; Duque, M.C.; Verdier, V. Assessing genetic variability among Brazilian strains of *Xanthomonas axonopodis* pv. *manihotis* through restriction fragment length polymorphism and amplified fragment length polymorphism analyses. *Can. J. Microbiol.* **1999**, *45*, 754–763. [[CrossRef](#)]
34. Restrepo, S.; Vélez, C.M.; Verdier, V. Measuring the Genetic Diversity of *Xanthomonas axonopodis* pv. *manihotis* Within Different Fields in Colombia. *Phytopathology* **2000**, *90*, 683–690. [[CrossRef](#)]
35. Restrepo, S.; Verdier, V. Geographical Differentiation of the Population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Appl. Environ. Microbiol.* **1997**, *63*, 4427–4434. [[CrossRef](#)] [[PubMed](#)]
36. Verdier, V.; Restrepo, S.; Mosquera, G.; Duque, M.C.; Gerstl, A.; Laberry, R. Genetic and pathogenic variation of *Xanthomonas axonopodis* pv. *manihotis* in Venezuela. *Plant Pathol.* **2001**, *47*, 601–608. [[CrossRef](#)]
37. Gonzalez, C.; Restrepo, S.; Tohme, J.; Verdier, V. Characterization of pathogenic and nonpathogenic strains of *Xanthomonas axonopodis* pv. *manihotis* by PCR-based DNA fingerprinting techniques. *Fems Microbiol. Lett.* **2002**, *215*, 23–31. [[CrossRef](#)]
38. Arrieta-Ortiz, M.L.; Rodríguez-R, L.M.; Pérez-Quintero, Á.L.; Poulin, L.; Díaz, A.C.; Arias Rojas, N.; Trujillo, C.; Restrepo Benavides, M.; Bart, R.; Boch, J.; et al. Genomic survey of pathogenicity determinants and VNTR markers in the cassava bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis* strain CIO151. *PLoS ONE* **2013**, *8*, e79704. [[CrossRef](#)]
39. Cohn, M.; Morbitzer, R.; Lahaye, T.; Staskawicz, B.J. Comparison of gene activation by two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility genes in cassava. *Mol. Plant Pathol.* **2016**, 875–889. [[CrossRef](#)] [[PubMed](#)]
40. Castiblanco, L.F.; Gil, J.; Rojas, A.; Osorio, D.; Gutiérrez, S.; Muñoz-Bodnar, A.; Perez-Quintero, A.L.; Koebnik, R.; Szurek, B.; López, C.; et al. TALE1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. *Mol. Plant Pathol.* **2012**, *14*, 84–95. [[CrossRef](#)] [[PubMed](#)]
41. Munoz-Bodnar, A.; Pérez-Quintero, A.; Gomez-Cano, F.; Gil, J.; Michelmore, R.; Bernal, A.; Szurek, B.; Lopez, C. RNAseq analysis of cassava reveals similar plant responses upon infection with pathogenic and non-pathogenic strains of *Xanthomonas axonopodis* pv. *manihotis*. *Plant Cell Rep.* **2014**, *33*, 1901–1912. [[CrossRef](#)]
42. Aristizábal, J.; Sánchez, T. *Guía Técnica para Producción y Análisis de Almidón de Yuca*; FAO: Roma, Italy, 2007.
43. FAO. *Save and Grow: Cassava, a Guide to Sustainable Production Intensification*; FAO: Roma, Italy, 2013; ISBN 978-92-5-107641-5.
44. Parmar, A.; Sturm, B.; Hensel, O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. *Food Secur.* **2017**, *9*, 907–927. [[CrossRef](#)]
45. Lozano, J.C. Cassava Bacterial Blight: A Manageable Disease. *Plant Dis.* **1986**, *70*, 1089. [[CrossRef](#)]
46. Wydra, K.; Verdier, V. Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. *Agric. Ecosyst. Environ.* **2002**, *93*, 211–226. [[CrossRef](#)]
47. Abaca, A.; Kiryowa, M.; Awori, E.; Andema, A.; Dradiku, F.; Moja, A.; Mukalazi, J. Cassava Pests and Diseases' Prevalence and Performance as Revealed by Adaptive Trial Sites in North Western Agro-Ecological Zone of Uganda. *J. Agric. Sci.* **2013**, *6*. [[CrossRef](#)]
48. López, C.E.; Bernal, A.J. Cassava Bacterial Blight: Using Genomics for the Elucidation and Management of an Old Problem. *Trop. Plant Biol.* **2012**, *5*, 117–126. [[CrossRef](#)]
49. Trujillo, C.; Arias-Rojas, N.; Poulin, L.; Medina, C.; Tapiero, A.; Restrepo, S.; Koebnik, R.; Bernal, A. Population typing of the causal agent of cassava bacterial blight in the Eastern Plains of Colombia using two types of molecular markers. *BMC Microbiol.* **2014**, *14*, 161. [[CrossRef](#)]
50. Trujillo, C.; Ochoa, J.C.; Mideros, M.F.; Restrepo, S.; Lopez, C.; Bernal, A. A complex population structure of the cassava pathogen *Xanthomonas axonopodis* pv. *manihotis* in recent years in the Caribbean Region of Colombia. *Microb. Ecol.* **2014**, *68*, 155–167. [[CrossRef](#)]
51. Blanvillain-Baufumé, S.; Reschke, M.; Solé, M.; Auguy, F.; Doucoure, H.; Szurek, B.; Meynard, D.; Portefaix, M.; Cunnac, S.; Guiderdoni, E.; 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*, 306–317. [[CrossRef](#)]

52. Oliva, R.; Ji, C.; Atienza-Grande, G.; Huguet-Tapia, J.C.; Perez-Quintero, A.; Li, T.; Eom, J.-S.; Li, C.; Nguyen, H.; Liu, B.; et al. Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **2019**, *37*, 1344–1350. [[CrossRef](#)]
53. Hutin, M.; Pérez-Quintero, A.L.; Lopez, C.; Szurek, B. MorTAL Kombat: The story of defense against TAL effectors through loss-of-susceptibility. *Front. Plant Sci.* **2015**, *6*, 535. [[CrossRef](#)]
54. Shantharaj, D.; Römer, P.; Figueiredo, J.F.L.; Minsavage, G.V.; Krönauer, C.; Stall, R.E.; Moore, G.A.; Fisher, L.C.; Hu, Y.; Horvath, D.M.; et al. An engineered promoter driving expression of a microbial avirulence gene confers recognition of TAL effectors and reduces growth of diverse *Xanthomonas* strains in citrus. *Mol. Plant Pathol.* **2017**, *18*, 976–989. [[CrossRef](#)] [[PubMed](#)]
55. Zhang, J.; Yin, Z.; White, F. TAL effectors and the executor R genes. *Front. Plant Sci.* **2015**, *6*, 641. [[CrossRef](#)] [[PubMed](#)]
56. Mora, R.; Rodriguez, M.; Gayosso, L.; López, C. Using in vitro plants to study the cassava response to *Xanthomonas phaseoli* pv. *manihotis* infection. *Trop. Plant Pathol.* **2019**. [[CrossRef](#)]
57. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 Years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)]
58. Medina, C.A.; Reyes, P.A.; Trujillo, C.A.; Gonzalez, J.L.; Bejarano, D.A.; Montenegro, N.A.; Jacobs, J.M.; Joe, A.; Restrepo, S.; Alfano, J.R.; et al. The role of type III effectors from *Xanthomonas axonopodis* pv. *manihotis* in virulence and suppression of plant immunity. *Mol. Plant Pathol.* **2018**, *19*, 593–606. [[CrossRef](#)]
59. Sambrook, J.R.D.W. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2001.
60. Eisel, D.; Seth, O.; Grünewald-Janho, S.; Kruchen, B.; Rieger, B. *DIG Application Manual for Filter Hybridization*; Roche Applied Science: Mannheim, Germany, 2008.
61. Pérez-Quintero, A.L.; Lamy, L.; Gordon, J.L.; Escalon, A.; Cunnac, S.; Szurek, B.; Gagnevin, L. QueTAL: A suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Front. Plant Sci.* **2015**, *6*, 545. [[CrossRef](#)]
62. Bray, N.L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **2016**, *34*, 525–527. [[CrossRef](#)]
63. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2009**, *26*, 139–140. [[CrossRef](#)]
64. Rahnenfuhrer, J.; Alexa, A. TopGO: Enrichment Analysis for Gene Ontology; R Package Version 2.38.1. 2019. Available online: <https://bioconductor.org/packages/release/bioc/manuals/topGO/man/topGO.pdf> (accessed on 1 October 2020).
65. 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*, e1002962. [[CrossRef](#)]
66. Erkes, A.; Mücke, S.; Reschke, M.; Boch, J.; Grau, J. PrediTAL: A novel model learned from quantitative data allows for new perspectives on TALE targeting. *PLoS Comput. Biol.* **2019**, *15*, e1007206. [[CrossRef](#)]
67. Wang, L.; Rinaldi, F.C.; Singh, P.; Doyle, E.L.; Dubrow, Z.E.; Tran, T.T.; Pérez-Quintero, A.L.; Szurek, B.; Bogdanove, A.J. TAL Effectors Drive Transcription Bidirectionally in Plants. *Mol. Plant* **2017**, *10*, 285–296. [[CrossRef](#)] [[PubMed](#)]
68. Kuon, J.-E.; Qi, W.; Schläpfer, P.; Hirsch-Hoffmann, M.; von Bieberstein, P.R.; Patrignani, A.; Poveda, L.; Grob, S.; Keller, M.; Shimizu-Inatsugi, R.; et al. Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. *BMC Biol.* **2019**, *17*, 75. [[CrossRef](#)] [[PubMed](#)]
69. Lin, H.-C.; Chang, Y.-A.; Chang, H. A *pthA* homolog from a variant of *Xanthomonas axonopodis* pv. *citri* enhances virulence without inducing canker symptom. *Eur. J. Plant Pathol.* **2013**, *137*, 677–688. [[CrossRef](#)]
70. Kassambara, A. ggpubr: “ggplot2” Based Publication Ready Plots; R Package Version 0.2.4. 2019. Available online: <https://cran.r-project.org/web/packages/ggpubr/index.html> (accessed on 1 October 2020).
71. Ahmed, M.; Kim, D.R. pcr: An R package for quality assessment, analysis and testing of qPCR data. *PeerJ* **2018**, *6*, e4473. [[CrossRef](#)] [[PubMed](#)]
72. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4.
73. Warnes, G.R.; Bolker, B.; Bonebakker, L.; Gentleman, R.; Huber, W.; Liaw, A.; Lumley, T.; Maechler, M.; Magnusson, A.; Moeller, S.; et al. gplots: Various R Programming Tools for Plotting Data; R Package Version 3.0.1.2. 2020. Available online: <https://cran.r-project.org/web/packages/gplots/index.html> (accessed on 1 October 2020).
74. Wilkins, D. gggenes: Draw Gene Arrow Maps in “ggplot2”; R Package Version 0.4.0. 2019. Available online: <http://cran.wustl.edu/web/packages/gggenes/gggenes.pdf> (accessed on 1 October 2020).
75. Soto-Sedano, J.C.; Mora-Moreno, R.E.; Mathew, B.; León, J.; Gómez-Cano, F.A.; Ballvora, A.; López-Carrascal, C.E. Major Novel QTL for Resistance to Cassava Bacterial Blight Identified through a Multi-Environmental Analysis. *Front. Plant Sci.* **2017**, *8*, 1169. [[CrossRef](#)] [[PubMed](#)]
76. Bredeson, J.V.; Lyons, J.B.; Prochnik, S.E.; Wu, G.A.; Ha, C.M.; Edsinger-Gonzales, E.; Grimwood, J.; Schmutz, J.; Rabbi, I.Y.; Egesi, C.; et al. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* **2016**, *34*, 562–570. [[CrossRef](#)] [[PubMed](#)]
77. Pérez-Quintero, A.L.; Rodriguez-R, L.M.; Dereeper, A.; López, C.; Koebnik, R.; Szurek, B.; Cunnac, S. 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. [[CrossRef](#)]
78. Doyle, E.L.; Booher, N.J.; Standage, D.S.; Voytas, D.F.; Brendel, V.P.; VanDyk, J.K.; Bogdanove, A.J. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res.* **2012**, *40*, W117–W122. [[CrossRef](#)]

79. Ye, G.; Hong, N.; Zou, L.-F.; Zou, H.-S.; Zakria, M.; Wang, G.-P.; Chen, G.-Y. tale-Based Genetic Diversity of Chinese Isolates of the Citrus Canker Pathogen *Xanthomonas citri* subsp. *citri*. *Plant Dis.* **2013**, *97*, 1187–1194. [[CrossRef](#)]
80. Ferreira, R.M.; de Oliveira, A.C.P.; Moreira, L.M.; Belasque, J.; Gourbeyre, E.; Siguier, P.; Ferro, M.I.T.; Ferro, J.A.; Chandler, M.; Varani, A.M. 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*. [[CrossRef](#)]
81. Ruh, M.; Briand, M.; Bonneau, S.; Jacques, M.-A.; Chen, N.W.G. Xanthomonas adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors. *BMC Genom.* **2017**, *18*, 670. [[CrossRef](#)] [[PubMed](#)]
82. Denancé, N.; Szurek, B.; Doyle, E.L.; Lauber, E.; Fontaine-Bodin, L.; Carrère, S.; Guy, E.; Hajri, A.; Cerutti, A.; Boureau, T.; et al. Two ancestral genes shaped the *Xanthomonas campestris* TAL effector gene repertoire. *New Phytol.* **2018**, *219*, 391–407. [[CrossRef](#)] [[PubMed](#)]
83. Erkes, A.; Reschke, M.; Boch, J.; Grau, J. Evolution of Transcription Activator-Like Effectors in *Xanthomonas oryzae*. *Genome Biol. Evol.* **2017**, *9*, 1599–1615. [[CrossRef](#)] [[PubMed](#)]
84. Schandry, N.; de Lange, O.; Prior, P.; Lahaye, T. TALE-Like Effectors Are an Ancestral Feature of the *Ralstonia solanacearum* Species Complex and Converge in DNA Targeting Specificity. *Front. Plant Sci.* **2016**, *7*, 1225. [[CrossRef](#)] [[PubMed](#)]
85. Meckler, J.F.; Bhakta, M.S.; Kim, M.-S.; Ovadia, R.; Habrian, C.H.; Zykovich, A.; Yu, A.; Lockwood, S.H.; Morbitzer, R.; Elsässer, J.; et al. Quantitative analysis of TALE–DNA interactions suggests polarity effects. *Nucleic Acids Res.* **2013**, *41*, 4118–4128. [[CrossRef](#)]
86. Wan, H.; Hu, J.; Li, K.; Tian, X.; Chang, S. Molecular Dynamics Simulations of DNA-Free and DNA-Bound TAL Effectors. *PLoS ONE* **2013**, *8*, e76045. [[CrossRef](#)]
87. Guilinger, J.P.; Pattanayak, V.; Reyon, D.; Tsai, S.Q.; Sander, J.D.; Joung, J.K.; Liu, D.R. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods* **2014**, *11*, 429–435. [[CrossRef](#)]
88. Grewal, R.K.; Gupta, S.; Das, S. *Xanthomonas oryzae* pv. *oryzae* triggers immediate transcriptomic modulations in rice. *BMC Genom.* **2012**, *13*, 49. [[CrossRef](#)]
89. Yang, B.; White, F.F. Diverse Members of the AvrBs3/PthA Family of Type III Effectors Are Major Virulence Determinants in Bacterial Blight Disease of Rice. *Mol. Plant-Microbe Interact.* **2004**, *17*, 1192–1200. [[CrossRef](#)]
90. 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*, 808–819. [[CrossRef](#)]
91. Zhou, J.; Peng, Z.; Long, J.; Soso, D.; Liu, B.; Eom, J.-S.; Huang, S.; Liu, S.; Vera Cruz, C.; Frommer, W.B.; et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* **2015**, *82*, 632–643. [[CrossRef](#)] [[PubMed](#)]
92. Betsuyaku, S.; Sawa, S.; Yamada, M. The Function of the CLE Peptides in Plant Development and Plant-Microbe Interactions. *Arab. Book/Am. Soc. Plant Biol.* **2011**, *9*, e0149. [[CrossRef](#)] [[PubMed](#)]
93. Strabala, T.J.; O'donnell, P.J.; Smit, A.-M.; Ampomah-Dwamena, C.; Martin, E.J.; Netzler, N.; Nieuwenhuizen, N.J.; Quinn, B.D.; Foote, H.C.C.; Hudson, K.R. Gain-of-function phenotypes of many CLAVATA3/ESR genes, including four new family members, correlate with tandem variations in the conserved CLAVATA3/ESR domain. *Plant Physiol.* **2006**, *140*, 1331–1344. [[CrossRef](#)] [[PubMed](#)]
94. Yamaguchi, Y.L.; Ishida, T.; Sawa, S. CLE peptides and their signaling pathways in plant development. *J. Exp. Bot.* **2016**, *67*, 4813–4826. [[CrossRef](#)]
95. Guo, Y.; Ni, J.; Denver, R.; Wang, X.; Clark, S.E. Mechanisms of molecular mimicry of plant CLE peptide ligands by the parasitic nematode *Globodera rostochiensis*. *Plant Physiol.* **2011**, *157*, 476–484. [[CrossRef](#)]
96. Wang, X.; Mitchum, M.G.; Gao, B.; Li, C.; Diab, H.; Baum, T.J.; Hussey, R.S.; Davis, E.L. A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of *Arabidopsis thaliana*. *Mol. Plant Pathol.* **2005**, *6*, 187–191. [[CrossRef](#)]
97. Marois, E.; Van den Ackerveken, G.; Bonas, U. The *Xanthomonas* Type III Effector Protein AvrBs3 Modulates Plant Gene Expression and Induces Cell Hypertrophy in the Susceptible Host. *Mol. Plant-Microbe Interact.* **2002**, *15*, 637–646. [[CrossRef](#)]

Chapter 2: *MeSWEET10e*, an alternative cassava susceptibility gene transcriptionally activated by TALEs of the cassava pathogen *Xanthomonas phaseoli* pv. *manihotis*

As seen in the previous chapter, TALE variability in *Xpm* strains is considerable, but this diversity becomes reduced at the light of predicted binding affinity. Hence, most of the variants (TALEs with the same number of repeats, but different RVD sequence) are tightly clustered in groups whose members usually share the same number of repeats and potentially common recent ancestors and function. However, it was also shown that evolution has resulted in TALEs, like TALE22D and TALE14B, that show gain or loss of repeats, without affecting the binding affinity of the effector, and potentially preserving its role in the interaction. This work confirmed the wide distribution of TALE20 and TALE14 variants, described as major TALEs in the *Xpm*-cassava interaction (Castiblanco et al., 2013; Cohn et al., 2016, 2014), but also pointed out to a wide distribution of other TALEs like TALE22 and TALE13. Before sequencing the TALome of *Xpm* strain UA1061, it was interesting to observe that strains with apparent “exotic” TALomes (i.e., absence of TALE20 or TALE14 based on Southern blot analysis) could be highly aggressive and virulent, suggesting that other mechanisms (TALE-associated or not) may be involved to compensate the absence of these major virulence factors. In the case of strain UA1061, we evidenced TALE22D to be a variant of TALE20 with two extra repeats located at the beginning of the central repeat region that do not affect the binding affinity to the *S* gene *MeSWEET10a*. However, in line with the second research question of this thesis (“are the described molecular mechanisms the only key susceptibility determinants targeted by TALEs of this pathogen?”), we focused our interest on highly aggressive *Xpm* strains that harbored exotic TALomes in order to shed light on potential alternative virulence strategies in *Xpm*.

This work was initiated by a PCR screening of *Xpm* strains deposited in the *Xanthomonas* collection from the Institute de Recherche pour le Développement, from which we selected some strains carrying “exotic” TALomes that contained only TALEs with 13, 14, 15, 16, and 21 RVDs (No TALE20/22). One of these strains, CIAT1241, was particularly interesting. Despite of containing TALEs with no more than 15 RVDs, it appeared as aggressive as *Xpm* strains naturally armed with TALE14 and TALE20. In this chapter, we used the SMRT technology to compare the genome sequences of five *Xpm* strains, including the high-virulence strain CIAT1241 and the low-virulence strain CIAT1205, and cloned their TALomes which we characterized. Transcriptomic analysis of cassava leaves infected with strain CIAT1241 revealed it to upregulate the clade III SWEET transporter-encoding gene *MeSWEET10e*. Interestingly we report that other *Xpm* strains with “exotic” TALomes (also sequenced using SMRT) are also capable of activating this gene, and they all have in common TALE15 variants with close predicted DNA affinities. Unfortunately, we were not able to find the responsible for the induction of *MeSWEET10e*, but we found evidence that suggests

that *MeSWEET10e* upregulation is mediated by an intermediate host transcriptional factor. Furthermore, the use of arTALEs allowed to demonstrate that *MeSWEET10e* and *MeSWEET10a* upregulation have similar physiological effects during a compatible interaction, both leading to the development of water-soaked symptoms. Phylogenetic analyses indicate that *Xpm* TAL effectors and their linked capabilities to activate *MeSWEETs* come from at least two different ancestor genes, constituting a case of convergent evolution through exploitation of functionally-redundant paralogs. This information is relevant for the pathosystem and the potential use of TALE biology as a driver for resistance engineering, as already seen in the pathosystem *Xanthomonas oryzae* pv. *oryzae*-rice (Oliva et al., 2019).

This work was performed in collaboration with several researchers and graduate students. My main role was to perform the wet-lab experiments and to analyze and exploit data generated by the bioinformaticians that collaborated with us. DNA preparation for sequencing was done under the guidance of Lionel Gagnevin (researcher), while assemblies were performed under the guidance of the graduate student David Ayala (master's in computational biology) at Universidad de los Andes. The experimental set for the RNAseq was designed in conjunction with Camilo López (researcher), Adriana Bernal (researcher), Boris Szurek (researcher) and Ruben Mora (graduate student), who then carried out the experiment. Alexis Deereper (researcher in bioinformatics) processed and analyzed the Illumina data derived from the RNAseq experiment and uploaded the information to the cassavagenome.org platform. Phylogenetic and ortholog analyses were performed under the guide of Natalia Vargas, a post-doctoral researcher at Universidad de los Andes. Camilo López, Adriana Bernal and Boris Szurek provided significant scientific guidance during the experimental planning, execution, and document writing. Since we still need to perform some experiments, we envision to publish these data at the end of this year, in the frame of a short post-doctoral stay at IRD. We aimed at publishing the present manuscript to MPP or MPMI-like journals.

MeSWEET10e*, an alternative cassava susceptibility gene transcriptionally activated by TALEs of *Xanthomonas phaseoli* pv. *manihotis

Carlos Andrés Zárate-Chaves^{1,2}, Rubén Mora³, Alexis Dereeper¹, Emilie Thomas¹, Camilo López³, Adriana Bernal^{2*}, Boris Szurek^{1*}

¹PHIM, Univ Montpellier, IRD, CIRAD, INRAe, Institut Agro, Montpellier, France.

²Laboratorio de interacciones moleculares en microorganismos agrícolas (LIMMA), Universidad de los Andes, Bogotá, Colombia

³Manihot Biotec, Departamento de Biología, Universidad Nacional de Colombia, Bogotá, Colombia.

*Corresponding author: boris.szurek@ird.fr, abernal@uniandes.edu.co

Abstract

Transcriptional activation of members of the Sugar Will be Eventually Exported Transporter (SWEET) family is a conserved strategy used by *Xanthomonas* and mediated by Transcription Activator-Like (TAL) effectors to promote pathogen virulence in rice, cotton, and cassava. *MeSWEET10a* is a major cassava bacterial blight (CBB) susceptibility (*S*) gene that is transcriptionally activated by TAL20_{Xam668} and its homologs, resulting in increased water soaking and *in planta* bacterial growth. However, not all strains of *Xpm* isolated from cassava fields possess TAL20_{Xam668}, suggesting alternative virulence functions used by these strains. In this study, we report on the identification of *MeSWEET10e* as a novel cassava *S* gene encoding an alternative clade-III sugar transporter activated by *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*). RNAseq experiments confirmed by quantitative RT-PCR analysis show that *Xpm* strain CIAT1241, which is devoid of the well-conserved TAL20_{Xam668}, is unable to induce *MeSWEET10a* but induces *MeSWEET10e*. Further validation of the latter as a novel CBB *S* gene came from the use of arTALE-based activation of *MeSWEET10e* in mutant strain Xam668ΔTAL20 hence demonstrating that the physiological function of this gene is equivalent to the role of *MeSWEET10a* in *Xpm* pathogenesis. Phylogenetic analyses indicate that *Xpm* TALEs and their linked capabilities to activate *MeSWEETs* come from at least two different ancestor genes, constituting a case of convergent evolution through exploitation of functionally-redundant paralogs. These findings are crucial in the conception of improved cassava varieties with enhanced *Xpm* resistance based on key molecular interactions.

Keywords

Cassava bacterial blight, convergent evolution, indirect transcriptional activation, *MeSWEET10e*

Introduction

Sugars are the main product of plant photosynthesis and carbohydrate assimilation, transport, and distribution are key factors impacting plant interactions with environmental biotic factors. After production in the mesophyll cells, sugars, mainly in the form of sucrose, must be mobilized to the phloem to be subsequently distributed to heterotrophic tissues. The movement of sucrose between the mesophyll cells and to the phloem parenchymal cells is achieved through plasmodesmata, while the sucrose movement from the parenchymal cells to the sieve element (SE)/companion cell (CC)

complex of the phloem can be accomplished through apoplasmic loading and/or symplasmic loading. In the apoplasmic loading, sucrose is exported from the parenchymal cells to the apoplast through the action of the Sugar Will be Eventually Exported Transporter (SWEET) family of transporters, and then CCs uptake the apoplasmic free sucrose to load it in the SEs through the H⁺-coupled sucrose transporters (SUTs) (reviewed by (Chen et al. 2010)). The SWEET family is defined as transmembrane sugar transporters usually carrying seven transmembrane helices, which are organized as two triple helix bundles connected by an inversion linker helix (Feng and Frommer 2015). Phylogenetically, in plants, *SWEETs* are distributed in four clades (I to IV), and, at least in *Arabidopsis thaliana*, the phylogenetic relationship correlates with the sugar type that is exported through the transporter. *SWEETs* from clades I, II, and IV are predominantly hexose transporters, while clade III *SWEETs* transport sucrose and some hexoses (Feng and Frommer 2015). Apart from sucrose loading in the phloem, clade III *SWEETs* are implicated in several physiological processes as nectar secretion (Lin et al. 2014), seed filling (Chen et al. 2015; Fei et al. 2021), pollen development (Sun et al. 2013), flowering (Andrés et al. 2020), and water-stress adaptation (Durand et al. 2016). Given the role of *SWEETs*, a tight regulation is crucial for an efficient carbohydrate partitioning, but so far, little is known about this aspect (Anjali et al. 2020). To date, only one transcription factor (TF), OsDOF11, has been demonstrated to coordinately control rice apoplasmic loading through a direct binding to *OsSUT1*, *OsSWEET11*, and *OsSWEET14* promoter regions (Wu et al. 2018). Nevertheless, it is widely known that bacterial, fungal, and oomycete phytopathogens deregulate *SWEETs*, supposedly to access carbon during the colonization of the host, turning these important genes into *S* genes (recently reviewed by (Breia et al. 2021)).

Pathogenicity of several *Xanthomonas* species rely on *SWEET* gene exploitation through transcriptional deregulation. Alteration of these *S* genes is achieved through transcriptional induction mediated by the bacterial Transcription Activator-Like (TAL) effectors (Perez-Quintero and Szurek 2019). TALEs have a particular structure that allows them to be imported to the host nucleus and recognize specific DNA sequences to modify transcription patterns (Schreiber and Bonas 2014; Doyle et al. 2013; Bogdanove et al. 2010; Boch and Bonas 2010). The N-terminal portion of the protein contains a short signal for translocation to the host cells and cryptic repeats that allow non-specific interaction with DNA (Szurek et al. 2002; Rossier et al. 1999; Doyle et al. 2013; Schreiber and Bonas 2014). The central portion of the protein is characterized by the presence of tandem, near-identical repeats that form a superhelix that wraps around the DNA. Each repeat is folded in 34- to 35-aminoacid scaffolds that allow residues 12 and 13 to interact with the DNA nitrogenous bases in a specific manner. The amino acids present in these two positions are referred as repeat variable di-residue (RVD), and their steric effects and other non-covalent interactions favor interactions with specific nitrogenous bases. Thus, the RVDs exposed by the repeats work as a code to find a specific nucleotide pattern where the superhelix stabilizes and the C-terminal region has the sufficient time to be positioned and carry out its role in transcriptional alteration (Moscou and Bogdanove 2009; Boch et al. 2009; Cuculis et al. 2015; Wicky et al. 2013). The C-terminal region is characterized by the presence of nuclear localization signal(s), which allow the import of the translocated effector to the nucleus. It also has a TF binding (TFB) region and a eukaryotic activation domain (AD) (Szurek et al. 2001; Hui et al. 2019b; Yang and Gabriel 1995; Zhu et al. 1998; Yuan et al. 2016; Hui et al. 2019a). The position of the protein on the effector binding element (EBE) on the target DNA (Antony et al. 2010), the interaction with the host basal TF subunits, and the presence of the AD recruit the RNA polymerase II, and transcription of the downstream region of the EBE is

started. As for *SWEET*s and other target genes (TGs), EBEs are often located in the proximity of the natural transcription start site, which ensures an effective transcriptional activation of the TG.

Strains of *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), the causing agent of rice bacterial leaf blight, possess several TALEs that cause individual upregulation of three *SWEET* genes belonging to clade III. The transcriptional activation of these three *OsSWEET*s results in increased rice susceptibility measured as larger lesions, water-soaked symptoms, and higher bacterial populations *in planta* (Römer et al. 2010; Yang et al. 2006; Oliva et al. 2019; Zhou et al. 2015; Antony et al. 2010; Yu et al. 2011). *OsSWEET14* (a clade-III *SWEET*) is targeted by several TALEs (*AvrXa7*, *PthXo3*, *TalC*, and *TalF*) present in Asian and African *Xoo* strains. This is a fine example of convergent evolution, and shows how important this *S* gene is for the *Xoo*-rice pathogenic interaction (Oliva et al. 2019; Tran et al. 2018). *OsSWEET11* (another clade-III *SWEET*) transcriptional activation is mediated only (as far as we know) by *PthXo1* (Römer et al. 2010; Yang et al. 2006), a TALE variant present in some Asian *Xoo* strains, while *OsSWEET13* activation is mediated by *PthXo2* variants (A, B, and C) (Zhou et al. 2015), which are usually present in Asian *Xoo* strains that also carry *PthXo3* or *AvrXa7* (Oliva et al. 2019). Comparably, upland cotton (*Gossypium hirsutum*) *SWEET10D* is transcriptionally activated during the compatible interactions with *Xanthomonas citri* pv. *malvacearum* through the action of *Avrb6*, a TALE. Mutation of this effector leads to a decrease in virulence for the pathogen measured as lower bacterial titers, and the loss of the capacity to induce water soaking in leaf lesions (Cox et al. 2017). Although *Xanthomonas citri* pv. *citri* can cause upregulation of the *Citrus sinensis* *SWEET1* (a clade-I *SWEET*) through the action of the *PthA4* variants, the physiological consequences of this activation do not seem to impact pathogenesis (Hu et al. 2014).

Xanthomonas phaseoli pv. *manihotis* (*Xpm*) is a cassava vascular pathogen that initially colonizes the mesophyll apoplast and then migrates to the leaf vessels to reach the stem. During mesophyll colonization, *Xpm* induces the formation of water-soaked angular leaf spots, which is also dependent on the transcriptional activation of the clade-III *MeSWEET10a* gene (Cohn et al. 2014). Different variants of TALE20 (*TAL20_{Xam668}*, *TALE20C_{XpmUA681}*, and *TALE20B_{XpmUA531}*) and the variant *TALE22D_{XpmUA1061}* promote the activation of cassava *MeSWEET10a*, whose product has been characterized as a glucose/sucrose transporter (Cohn et al. 2014; Zárate-Chaves et al. 2021; Mora et al. 2019). Lack or mutation of *MeSWEET10a*-inducing TALEs among *Xpm* strains results in lower bacterial titers and abolishment of water-soaked lesions (Cohn et al. 2014). Water soaking seems to be an important symptomatic feature for some plant pathogenic *Pseudomonas* and *Xanthomonas* species, since it improves the bacterial movement in the plant tissue and facilitates the entrance of epiphytic bacteria to the affected tissue (Schwartz et al. 2017; El Kasmi et al. 2018). Water soaking can be the consequence of the induction of cell-wall degrading enzymes that in turn results in increased hygroscopicity of the leaf tissue, a strategy exhibited by the tomato pathogen *Xanthomonas gardneri* and its TALE *AvrHah1* (Schwartz et al. 2017). Theoretically, osmotic/ionic unbalances that favor osmosis towards the apoplast (Aung et al. 2018; Cernadas et al. 2014) and the presence of hygroscopic molecules, like the xanthan gum, in the apoplast can also drive water soaking formation. Therefore, *SWEET*-mediated sugar export may not only support *Xanthomonas* growth accompanied by more bacterial exopolysaccharides in the apoplast, but it could also tune the osmotic potential and help the establishment of an aqueous environment (Aung et al. 2018).

Cassava has 28 *SWEET* genes, and 13 of them belong to clade III (Cao et al. 2019) but only *MeSWEET10a* was reported so far to be targeted by *Xpm* (Cohn et al. 2014). In this study, we

perform the characterization of a few strains that lack TALE20/22 *MeSWEET10a*-inducing variants, whilst they retain full capacity to cause water-soaked lesions. This being in disagreement with the statement that *Xpm* strains without TALE20/22 variants are poorly virulent, we reasoned that these strains may either carry other TALEs with *MeSWEET10a*-inducing capacity that were not known so far, or that they induce alternative clade-III *SWEETs* or even other unknown *S* genes. Combining TALome sequencing of these exotic strains and transcriptomic analysis led us to discover *MeSWEET10e* as a novel CBB *S* gene and draw new hypothesis about a potential TALE-mediated indirect activation of this gene.

Results

***Xpm* strain CIAT1241 causes water-soaked symptoms without TALE20 homologs.**

In a screen to assess sequence and functional diversity of TALomes among natural populations of *Xpm* (Zárate-Chaves et al. 2021), we performed Southern blot analyses to identify strains with exotic TALome patterns (Figure 1b). *Xpm* strain CIAT1241, isolated from Argentina in 1984, displayed TALEs with apparent 13, 14 and 15 RVDs and does not have any of the larger ones (20 or 22 RVDs), which are known to target *MeSWEET10a*. Despite this, CIAT1241 retains full aggressiveness and is able to cause symptoms comparable to those produced by *MeSWEET10a*-inducing strains UA681 and UA1061 (Figure 1A and (Zárate-Chaves et al. 2021)). Another interesting strain is CIAT1205 which was isolated from symptomatic leaves of cassava in New Zealand in 1966. CIAT1205 is comparable to *Xpm* strain CFBP1851 which was investigated previously (Zárate-Chaves et al. 2021). Both strains show reduced aggressiveness and carry only two TALEs with 14 and 21 RVDs, as shown by Southern blot. To determine the TALomes of strains CIAT1241 and CIAT1205, we sequenced the genomes of both strains using the SMRT technology (Pacific Biosciences, Menlo Park, CA). Table 1 shows the extracted TALE sequences in the form of RVD string. Interestingly, CIAT1241 and CIAT1205 share TALE14D, indicating that this gene is not responsible *per se* for the full aggressiveness phenotype of CIAT1241. Remarkably, *TALE15E_{XpmCIAT1241}* is present in two copies in the genome of CIAT1241, one of them is located on a plasmid together with *TALE14D_{XpmCIAT1241}* and *TALE15C_{XpmCIAT1241}*. The second copy is chromosomal, laying in a region flanked by repetitions, next to *XopE4* and several transposon-related genes, suggestive of a previous event of insertion (Supplementary Figure S1).

CIAT1205 carries the same TALome as CFBP1851 and both TALEs variants (*TALE14D* and *TALE21A*) are located on the same plasmid. The chromosome alignment shows, however that the strains are slightly different due to translocation and inversion of two syntenic segments of the genome (Supplementary Figure S2). In summary, TALomes from strain CIAT1205 (low virulence) and strain CIAT1241 (high virulence) share the TALE14D variant, and CIAT1241 strain has three unique TALEs (*TALE13B_{XpmCIAT1241}*, *TALE15C_{XpmCIAT1241}*, and *TALE15E_{XpmCIAT1241}*) that may support the full virulence of the strain in the absence of *MeSWEET10a*-inducing TALE20-22 variants.

Transcriptomic analysis of cassava leaves infected with *Xpm* strain CIAT1241 reveals *MeSWEET10e* as an alternative susceptibility gene candidate of the clade III *SWEET* family

To gain insights into the molecular mechanisms underlying the virulence of strain CIAT1241, an RNAseq experiment was performed to profile the transcriptomic responses of *in-vitro* propagated cassava plants of the 60444 variety at 50 hours post-inoculation (hpi). Comparisons were made between mock-inoculated plants and those inoculated with strain CIAT1241. The analysis of the

relative expression of genes showed that 131 cassava genes were upregulated ($\text{Log}_2\text{FoldChange} \geq 2$ and $p\text{-value} \leq 0.001$) in response to treatment with CIAT1241. Strikingly, the most upregulated gene ($\text{Log}_2\text{FoldChange} = 12.68$, $p\text{-value} = 2.9 \times 10^{-21}$) corresponds to *Manes.14G047800*, annotated as a bidirectional sugar transporter SWEET encoding gene designated as *MeSWEET10e* (Cao et al. 2019). Figure 2A shows the results of the transcriptomic analysis overall and identifies the expression profile of the 28 members of the cassava SWEET family (Cao et al. 2019). The known S gene *MeSWEET10a* was also shown to be upregulated by CIAT1241 in this experiment, but to a lower extent than that of *MeSWEET10e* ($\text{Log}_2\text{FoldChange} = 5.03$, $p\text{-value} = 6.9 \times 10^{-5}$). Semiquantitative RT-PCR analysis of *MeSWEET10a* and *MeSWEET10e* further confirmed the expression pattern revealed by the RNAseq, showing that CIAT1241 but not UA681 (which was used as positive control for *MeSWEET10a* induction (Zárate-Chaves et al. 2021)), is targeting *MeSWEET10e* (Figure 2B). Interestingly, no induction of either *MeSWEET10a* or *MeSWEET10e* was observed when strain CIAT1205 was tested. This is in line with the absence of *MeSWEET10a*-inducing TALE20/22 variants in this strain, and suggests that the TALE14D variant, which is present in both CIAT1241 and CIAT1205, is not involved in the activation of *MeSWEET10e*.

The amino acid sequence of the orthologs of cassava SWEET proteins from *Arabidopsis thaliana*, rice (*Oryza sativa*), and clade-III SWEETs from upland cotton were aligned and clustered by Neighbor-Joining to assess relationships with *MeSWEET10e* (Figure 3). *MeSWEET10e*, which is also a clade-III SWEET protein, tightly clusters with 4 other *MeSWEET10* paralogs as well as with their orthologs in *A. thaliana* and cotton. The *MeSWEET10* paralogs are physically encoded in two different clusters of the cassava genome (Figure 4). The first cluster is located in the chromosome 6 and contains *MeSWEET10a*, *b*, and *c* and *MeSWEET12a*, and *b*, while the second cluster is located in the chromosome 14 and is composed by *MeSWEET10d*, *e* and *MeSWEET12c*. Both *Xpm*-targeted *MeSWEETs* are composed by six exons and they share 67% of identity.

Cross-referencing transcriptomics data and TALE target prediction has allowed the discovery of most TALE-targeted genes with key roles in pathogenicity. Hence, the potential EBEs for TALEs from CIAT1241 were predicted in the promoterome of the AM560-2 accession, by means of three available tools: TALVEZ (Pérez-Quintero et al. 2013), TALE-NT 2.0 (Doyle et al. 2012), and prediTAL (Erkes et al. 2019). Predictions were pooled and classified according to their rank and redundancy among predictions generated by the three tools (supplementary Table 1). As seen in Figure 4, the promoter of *MeSWEET10e* has two EBE candidates predicted for TALE15C_{XpmCIAT1241}, and TALE15E_{XpmCIAT1241}, which overlapped in an AT rich region c.a. 500-bp away from the start codon. No other binding sites were predicted within this region for any of the other TALEs in this strain.

***MeSWEET10e* is activated by other *Xpm* strains carrying TALE15C_{XpmCIAT1241} functional homologs.**

Since predictions indicated that both TALE15 variants might be involved in *MeSWEET10e* transcriptional activation, we further screened our collection of *Xpm* strains to identify other strains lacking TALE20 homologs but having TALE15 variants. From 30 strains that were analyzed through PCR-based profiling of TALEs repertoires (data not shown), we were able to identify three that also lacked TALE20: CIAT1135, CIAT1211, and ORST2. CIAT1135 was isolated in China in 1975, CIAT1211 was isolated in New Zealand in 1980, while ORST2 was isolated in Venezuela in 1971. Interestingly, quantitative RT-PCR of 60444-adult-plant leaves inoculated with these strains demonstrated that, like for CIAT1241, strains CIAT1211 and ORST2 also upregulate *MeSWEET10e* but no *MeSWEET10a*

at 50 hpi. As for CIAT1135, no induction of any of these genes was detected, as previously observed for CIAT1205 (Figure 5).

Sequencing of the identified strains was performed using the SMRT technology, and their TALomes were retrieved. Table 1 shows the RVD sequences of TALEs from these three *Xpm* strains. ORST2 was shown to possess a TALome almost identical to the one of CIAT1241, except for the lack of the TALE13B variant. To clarify the diversity, functionality and ancestry of the TALE variants, we used the tools DisTAL and FuncTAL of the QueTAL suite (Pérez-Quintero et al. 2015), and reconstructed phylogenies based on the N- and C-terminal-coding regions of the TALEs. As controls for the functional analysis, we included the RVD strings of TAL20_{Xam668}, TALE22D_{XpmUA1061}, and TALE20C_{XpmUA681}, which are validated *MeSWEET10a* activators (Cohn et al. 2014; Zárate-Chaves et al. 2021). For the sake of comparability between FuncTAL, DisTAL, and the phylogenies, we included TAL14_{Xam668} and TALE14C_{XpmUA681} (both having the same RVD sequence), since the nucleotide sequence for TAL14_{Xam668} is not publicly available, but they have been reported to activate important *S* genes in cassava (Cohn et al. 2014, 2016). Figure 6 shows the relationship of TALEs according to their predicted affinity (as determined by FuncTAL), which indicates that there are three potential functional clusters. Clusters 3 and 2 group the control sequences and described DNA affinity groups already defined in our previous work (Zárate-Chaves et al. 2021). Cluster 1 shows a potential functional relationship among TALE15C, TALE15D, and TALE16A variants, while TALE15E variant seems to be functionally closer to TALE21A variant.

Phylogenetic analyses based on repeat sequences and the N-terminal-coding sequence (Supplementary Figure S3A and B) agree with the functional clustering produced by FuncTAL. TALEs from the FuncTAL clusters 2 and 3 grouped together and were always separated from the clade that harbors TALE15C, TALE15D, TALE15E, and TALE16A. In line with our previous assumptions, the *MeSWEET10e*-upregulation capacity have a strong correlation with the clades that include TALE15C/D/E variants. In contrast, the phylogeny based on C-terminal-coding sequences (Supplementary Figure S3C) is not congruent with the previously-described clustering.

TALE15C and TALE15D variants seem to be responsible for *MeSWEET10e* activation, probably not as a direct target.

The capacity to transcriptionally activate *MeSWEET10e* was analyzed as a function of the TALome members, and the functional groups set by FuncTAL. Figure 7A shows that the common functional factor among CIAT1241, CIAT1211 and ORST2 is the presence of TALE15 variants from the FuncTAL cluster 1. EBE predictions indicate that variants TALE15C, TALE15D, TALE16A (FuncTAL cluster 1), and TALE15E are able to bind the promoter region of *MeSWEET10e* on overlapped EBEs (Figure 7B). Since the EBEs seem to be far away from the translational start site (TXS) and the annotated transcriptional start site (TSS) corresponds to the TXS, we used the RNAseq data to reconstruct the potential transcript induced 50 hpi in the mock- and the CIAT1241-inoculated treatment. Unluckily, the read abundance for the mock-inoculated treatment was extremely low, and they all mapped to the last exons (as expected after a polyT-based cDNA synthesis). However, reads from the CIAT1241 library mapped to *MeSWEET10e* and its promoter region revealed that the potential TSS is 40-bp away from the TXS (Figure 7B) and there are no reads supporting transcription from predicted EBEs.

None of the CIAT1241 TALEs are able to individually activate *MeSWEET10e*.

We used a gain-of-function approach to look for the CIAT1241 TALE responsible for *MeSWEET10e* transcriptional activation, based on the well-founded assumption that *MeSWEET10* overexpression would increase or restore the ability of a mutant defective in causing water soaking (Cohn et al. 2014). First, we cloned the four TALE variants in pSKX1, under the control of the *TALE1_{Xam}* promoter region; we included a TALE20 construct as a positive control. These constructs and the empty vector were transformed into the low-virulence mutant CFBP1851- Δ TALE14D- Δ TALE21A, a TALE-free *Xpm* strain, and then inoculated on 60444 leaves of adult plants. Figure 8A shows that none of the transformants was able to induce water soaking at 96 hpi, except for the control transformant expressing *TALE20*. Since wild type CFBP1851 strain has a low virulence in itself (even with its complete TALome) (Zárate-Chaves et al. 2021), the Xam668 Δ TAL20_{Xam668} mutant used by Cohn and coworkers was transformed with the same set of plasmids. As observed in Figure 8B, mixes of the transformants were infiltrated on 60444-adult-plant leaves, but results came to the same conclusion: none of the individual variants of TALE15, nor the mix of transformants carrying the CIAT1241 TALome were able to activate *MeSWEET10e*.

ArTALEs rescue the water-soaking phenotype in the absence of TALE20 through *MeSWEET10e* transcriptional activation.

As it is not clear if *MeSWEET10e* is a direct TALE target and if its promoter includes functional EBEs, we designed arTALEs to test if a direct TALE-mediated transcriptional activation of this gene would restore the water soaking in a defective mutant. We designed two arTALEs, arTALES10E.1 and arTALES10E.2 (see Table 3), which anchor 52-bp and 75-bp away from the TXS, respectively. The Xam668 Δ TAL20_{Xam668} mutant was transformed with these *arTALEs* and the EV and then infiltrated on 60444-adult-plant leaves. Figure 9 shows that both arTALEs drive the upregulation of *MeSWEET10e* at 50 hpi, and such transcriptional induction is accompanied by the restoration of the capacity to cause water-soaked symptoms in the absence of TALE20 functional homologs.

Discussion and conclusions

The devastating cassava pathogen *Xpm* heavily relies on its type-three effectors to cause disease and specially on TALEs to exploit *S* genes, such as the sugar transporter encoding gene *MeSWEET10a* (Bart et al. 2012; Medina et al. 2018; Cohn et al. 2014, 2016; Zárate-Chaves et al. 2021). In this study, we describe a second sugar transporter gene of the SWEET family, *MeSWEET10e*, that is transcriptionally induced by three *Xpm* strains that lack *MeSWEET10a*-activating TALEs. Convergent evolution through targeting functionally redundant *S* genes has been described for other *Xanthomonas* pathogens as *Xoo* in rice and *Xanthomonas citri* pv. *citri* in citrus. In the case of rice, the TALE-mediated upregulation of three different clade-III *OsSWEETs* leads to increased levels of sucrose in the apoplast that support the pathogen growth and water soaking (Römer et al. 2010; Yang et al. 2006; Oliva et al. 2019; Zhou et al. 2015; Antony et al. 2010; Yu et al. 2011). Furthermore, Streubel and coworkers demonstrated that, in rice, arTALE-mediated transcriptional activation of the clade-III *OsSWEET* paralogs leads to enhanced water-soaked symptoms and *in-planta* growth (Streubel et al. 2013). In citrus, the key master regulator *CsLOB1* is targeted by several homologs of the TALE PthA4, whose upregulation in leaves leads to pustule formation and canker. As in rice, *Xcc* TALEs seem not only to target *CsLOB1*, but also the paralog *CsLOB2* (Pereira et al. 2014), and arTALE-mediated upregulation of *CsLOB2* and *CsLOB3* results in the same disease phenotypes due to functional redundancy of these master regulators (Zhang et al. 2017). Here, we report that arTALE-

mediated direct upregulation of *MeSWEET10e* can restore water-soaked symptoms in a mutant of a strain that activates *MeSWEET10a*, showing the potential of this transporter gene as S factor. Hence, upregulation of *MeSWEET10e* would constitute another case of functional convergence for enhanced susceptibility in cassava.

TALE-mediated transcriptional activation for most of the validated *SWEETs* is characterized by strong changes in the transcript abundance of the gene (Cohn et al. 2014; Cox et al. 2017; Oliva et al. 2019; Streubel et al. 2013; Antony et al. 2010), as observed here for *MeSWEET10e*. Although we could not validate the importance of TALEs in this activation, correlative evidence suggests TALE15C and/or 15E, which would be functionally redundant according to FuncTAL, as candidate factors responsible for *MeSWEET10e* upregulation. Moreover, TALE15E variant cannot be ruled out as another *MeSWEET10e*-activating TALE, since it is present in two of the three *Xpm* strains that transcriptionally activate *MeSWEET10e*, and it is predicted to have EBEs that overlap with the ones from TALE15C and TALE15D variants. Nevertheless, it is also plausible that *Xpm* uses non-TALE effectors and/or other virulence factors to cause *MeSWEET10e* upregulation, as it has been reported for other phytopathogens that do not express TALEs, like *Pseudomonas syringae* pv. *tomato*, *Botrytis cinerea*, and *Fusarium oxysporum*, among others (reviewed by (Devanna et al. 2021)). The TALE dependency of this activation should be further studied through loss-of-function mutants.

It is interesting that *TALE15E* is found in two copies, one chromosomal and one plasmid-borne, in the genome of CIAT1241. Transposon-mediated integration of TALEs has been well documented (Ferreira et al. 2015; Ruh et al. 2017). Reconstruction of the association between TALEs and adaptation of *Xanthomonas* to common bean showed that acquisition of *tal23A* through horizontal gene transfer from *X. citri* pv. *fuscans* (*Xcf*) to *X. phaseoli* pv. *phaseoli* (*Xpp*) was followed by a transposon-mediated insertion of this gene into the chromosome. The high occurrence of this gene and its homologs in the common bean pathogens (even in phylogenetically distant species) suggests a key role and it is then tempting to suggest that insertion of *tal23A* facilitated the adaptation of *Xpp* to its host (Ruh et al. 2017). In this case, the *TALE15E* insertion could reflect a recent duplication of the gene and/or, in a scenario where important target activation and function is TALE-dose dependent, the fact that having more copies of this gene could result in higher fitness of the pathogen.

Functional and phylogenetic relationships suggest that *TALE15C/D/E* and *TALE16A* (the potential activators of *MeSWEET10e*) evolved from a *TALE* ancestor that was different from the one that gave rise to *TALE14*, *TALE20*, and *TALE22* variants. However, this would have resulted in functional convergence through upregulation of two different functionally-redundant target genes (*MeSWEET10a* and *MeSWEET10e*). Although *TALE16A* is, in theory, functionally related to *TALE15C* and *TALE15D*, strain CIAT1135 is not able to activate *MeSWEET10e*. However, this gene seems to have been evolved to have an uncommon feature among *Xpm* TALEs: the last repeat and the last half repeat have HD as RVD (see Figure 6). The last half repeat is suggested by some researchers to stabilize the TALE structure when interacting with DNA (Wan et al. 2016), nevertheless, Zheng and coworkers demonstrated that it is dispensable for gene activation from arTALEs (Zheng et al. 2014). Furthermore, the HD-encoding last repeat seems to be common among functional TALEs from other species like *Xanthomonas campestris* pv. *campestris* (*Xca*) (Denancé et al. 2018) and *Xpp* (Ruh et al. 2017). Validation of EBEs and target genes for TALE15 variants would clarify the affinity and activity of these exotic TALEs.

It is also worth noting that the three *MeSWEET10e*-upregulating strains with exotic TALomes were isolated more than 35 years ago, and most of the identified TALE variants are not frequently found within the more recently-collected *Xpm* isolates (Zárte-Chaves et al. 2021). Although, we do not have enough data to draw conclusions, we could hypothesize that, in a scenario without sampling biases, this would reflect the extinction of populations that used *MeSWEET10e* activation as a strategy to promote disease. The decrease in these populations would have been facilitated by the gradual increase in the distribution of *MeSWEET10a-activating* TALEs, selected as a more parsimonious mechanism to upregulate clade-III *MeSWEET* genes. It would be interesting to expand the screening of this strategy (and associated TALomes) among *Xpm* isolates to test if this feature can be tracked to modern isolates.

Transcript reconstruction from RNAseq data reads and the distance between the predicted EBEs and the transcription start point to indirect TALE-mediated upregulation of *MeSWEET10e*. Although indirect TALE-mediated activation of *SWEET* genes has not been reported for *Xanthomonas*, it is known that the TALE-mediated upregulation of master regulators (bHLH3, bHLH6, upa20, and CsLOB1) induces the concerted activation of downstream genes that code for proteins with key enzymatic action (Duan et al. 2018; Kay et al. 2009; Schwartz et al. 2017). Plant SWEETs have been shown to be commonly targeted and deregulated by phytopathogenic bacteria and fungi, which implies that TALEs are not the exclusive mechanism to upregulate these genes (Devanna et al. 2021). Although little is known about these mechanisms, bZIP11 has been shown to be involved in *SWEET* control in *Arabidopsis thaliana*, and manipulated by *Pseudomonas syringae* pv. *tomato* to potentially shift the regulation of some of these sugar transporters in benefit of the pathogen (Prior et al. 2021). Upregulation of the *MeSWEET10e* was tested through individual or combined infiltration of mutant strains carrying CIAT1241 TALEs (gain-of-function mutants). However, none of the treatments induced the transcriptional activation of the candidate target. As we did not test all the possible TALE combinations and we do not dispose of loss-of-function mutants, we cannot rule out that *MeSWEET10e* activation arises from a coordinated biological activity between multiple TALEs and/or non-TALE factors (other translocated or secreted effectors, for example), or that it is not mediated by TALEs. As SWEETs, extracellular invertases contribute to carbohydrate metabolism in plants through sucrose hydrolysis, and some members are associated to defense/stress responses. In tomato, the circadian transcriptional regulation of the extracellular invertase LIN6 is jointly mediated by two TFs (Circadian Clock Associated 1 -CCA1- and Late Elongated Hypocotyl -LHY), where co-occurrence and cooperative promotive activity is necessary for such regulation (Proels and Roitsch 2009). Therefore, it is not unreasonable to think the *MeSWEET10e* activation as a result of coordinated TALE (or non-TALE factors) action. Comparative genomics between genomes of strains with contrasting capacities to induce *MeSWEET10e* and the obtention of loss-of-function mutants would hopefully provide insights into this activation.

Other scenarios could be also speculated, like the potential of TALEs as trans-activating enhancers. In this case, TALEs would bind to EBEs located far away from the target gene, and then would come in close contact with the promoter region due to the tridimensional DNA structure. This has not been previously reported for TALEs, but enhancers are very common elements that control gene expression in eukaryotes (Weber et al. 2016). Alternatively, we cannot rule out the possibility that the expression vector used for this experiment (pSKX1) does compromise the delivery of the TALEs and their virulence. Although, this expression vector has been widely used for TALE biology studies

in *Xoo* and *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) (Tran et al. 2018; Triplett et al. 2016; Hutin et al. 2016, 2015), recent observations in our group points to a strong vector effect on *Xpm* (data not shown).

In conclusion, here we demonstrate that some *Xpm* strains are able to activate *MeSWEET10e* and this gene acts as an *S* gene. Although the factor responsible for this activation is unknown, evidence point to TALE15C, TALE15D, and TALE15E variants, probably upregulating it as an indirect target. Phylogenetic analyses show that TALEs that activate the already known *S* gene *MeSWEET10a* stem from a different ancestor than the TALEs that potentially activate the novel *S* gene *MeSWEET10e*, possibly constituting a case of convergent evolution through functional redundance. Further studies are required to reveal the mechanisms behind this transcriptional activation.

Materials and methods

Bacterial strains, plant material, culture conditions and media

Most of the *Xanthomonas phaseoli* pv. *manihotis* strains are part of the *Xanthomonas* collection of the Plant Health Institute of Montpellier. Strains UA681 and UA1061 are included in the *Xanthomonas* collection of the Laboratorio de Interacciones Moleculares de Microorganismos Agrícolas, Universidad de los Andes. The strain Xam668 and the derived mutant XamUA668 Δ TAL20_{Xam668} were kindly provided by Brian Staskawicz. Bacteria were streaked on YPG (yeast extract 5 gL⁻¹, peptone 5 gL⁻¹, glucose 5 gL⁻¹, agar-agar 15 gL⁻¹) solid media and incubated for 48 hours at 28°C or grown in Phi broth (yeast extract 1 gL⁻¹, peptone 10 gL⁻¹, casaminoacids 1 gL⁻¹) at 28°C, under constant shaking at 220 rpm for 24 hours. For DNA extraction, bacteria were grown on YP solid media, whose composition is the same of YPG but without glucose. Cassava cuttings from the cultivar 60444 were planted on individual peat pots and grown under greenhouse conditions (27°C \pm 5°C; photoperiod 12:12, relative humidity > 60%). *In vitro*-grown plants from the same cultivar were grown as described elsewhere (Mora et al. 2019).

Aggressiveness assays and evaluation of the development of water-soaked symptoms

Bacterial aggressiveness was quantified as the leaf lesion area formed by *Xpm* upon inoculation of 3-month-old cassava plants. Bacteria were cultured in liquid media and washed with 10 mM MgCl₂ sterile solution. Cell density was adjusted to an OD₆₀₀ of 0.2 (c.a. 2 \times 10⁸ cfu/mL). Adjusted bacterial suspensions were inoculated by placing a 10- μ L drop of inoculum over a 2-mm (\varnothing) hole made with a cork borer through the leaf tissue. Plants were kept in a greenhouse under the same conditions described above. Lesions were individually photographed at 15 days post-inoculation (dpi) in a stereoscope and areas were measured using Image-J software (version 1.48) (Schneider et al. 2012). Each treatment was inoculated once in three individual plants, and the aggressiveness assay was replicated three times.

To monitor water-soaked symptom development, wild-type bacteria and transformants were grown as described earlier in liquid media, supplemented accordingly with gentamycin (20 μ g/ml). Leaves of 6-week-old cassava plants were infiltrated with the adjusted bacterial suspensions (OD₆₀₀ = 0.2) with a needleless syringe through the abaxial area of the leaflet. For mixed inoculations, the OD₆₀₀ of each transformant suspension was adjusted before mixing to obtain a final concentration for each transformant of 2 \times 10⁸ cfu/mL. Plants were kept in a greenhouse under the same conditions described above. Infiltrating points were individually photographed at 96 hours post-inoculation

(hpi) under a stereoscope. Each treatment was infiltrated twice in three individual plants, and each assay was replicated two times.

Southern blot analysis

Southern blots were performed as described in our previous work (Zárate-Chaves et al. 2021). Briefly, total bacterial DNA was digested with BamHI, electrophoresed, and crosslinked to a positively charged Nylon membrane. The probe was synthesized from a conserved portion of the *TALE1_{Xam}* (from nucleotide 2909 to 3409) DNA and labeled with digoxigenin. The chloro-5-substituted adamantyl-1, 2-dioxetane phosphate (CSPD) reagent was used to detect the probe on a ChemiDoc XRS+ System (BioRad, USA).

Isolation and sequencing of TALEs

TALE genes were cloned directly from the target strain by direct cloning, as described in our previous work (Zárate-Chaves et al. 2021). Briefly, TALE genes were purified from agarose gels where the BamHI-digested total bacterial DNA was electrophoresed, and then cloned into BamHI-digested pBlueScript. Sequencing of isolated TALE genes was carried out by standard Sanger chemistry (Macrogen Inc., Korea) and the assembly of reported fragments was performed using the Geneious software (version R11).

Genome sequencing by PacBio SMRT technology and assembly

Total DNA was extracted from bacterial patches grown overnight on YP solid medium. For each strain, two 20- μ L loops of bacteria were taken from the patch and resuspended in 1.5 mL of 0.5 M NaCl by vortexing. Samples were centrifuged at 4500 rpm for 4 minutes and supernatants were discarded. Washed pellets were used as the starting material for the Genomic-tips 100/G (Qiagen, Germany). Total DNA was extracted according to the manufacturer's instructions, including the treatment with RNase A. DNA was resuspended in TE overnight at 4°C, and its quality and quantity were measured by NanoDrop One (Thermo Fisher Scientific, USA), Qubit 1.0 (Thermo Fisher Scientific, USA) and 2100 Bioanalyzer (Agilent, USA), per manufacturer's instructions. Barcoded DNA libraries spanning 8 to 12 kilobases were prepared at the Plateforme GENTYANE UMR INRA (France) and then sequenced by multiplexing six libraries (strains) per SMRT cell, in a PacBio Sequel sequencer (PacBio, USA). Genomes were assembled with CANU version 1.8 (Koren et al. 2017), circularized with Circlator version 1.5.5 (Hunt et al. 2015), and polished three times with Arrow (included in the GenomicConsensus package of PacBio).

TALE gene annotation and TALE-based functional and phylogenetical analysis

TALE genes were annotated using AnnoTALE version 1.2 (Grau et al. 2016). Genome annotation was performed via RASTtk (Brettin et al. 2015) service included in the PATRIC platform (Davis et al. 2020). Assembled TALE sequences were treated and classified in variants using an in-house script on R (version 3.6.1). TALEs were clustered by predicted DNA binding specificity using FuncTAL software (version 1.1) (Pérez-Quintero et al. 2015); default parameters were used. TALEs were also clustered by nucleotide repeat composition using DisTAL software (version 1.1) (Pérez-Quintero et al. 2015); default parameters were used. The nucleotide sequences coding for the N- and C-terminal regions were independently aligned using MAFFT version 7.450 (Katoh and Standley 2013) with the automatic algorithm selection of the Geneious software plugin. Alignments were then used to find

the best nucleotide substitution model by using ModelTest version 3.7 (Posada and Crandall 1998). Maximum Likelihood phylogenetic trees were created with the Geneious software plugin of PHYML version 3.3 (Guindon et al. 2010), under the HKY85 model with 100 bootstraps. Trees were visualized and edited in Geneious software (version R11).

***In-vitro* propagated cassava plant inoculation, RNA extraction, and RNA sequencing**

In-vitro-propagated cassava plants were inoculated with the *Xpm* strain CIAT1241 or a 10-mM MgCl₂ solution (mock inoculation) as described in our previous work to sequence the mRNA of the plant at 50 hpi (Zárate-Chaves et al. 2021). Briefly, 50 hours after inoculation, total RNA was extracted with the Invitrap Spin plant RNA minikit (STRATEC, Germany), using the RP buffer per manufacturer's instructions. The quality of the DNase I-treated samples was assessed with the RNA 6000 Nano Kit (Agilent, USA). The barcoded, paired-end TruSeq RNA libraries were pooled and sequenced in four flow cells of the NextSeq500 System. RNAseq analyses were performed using the Kallisto pseudo-mapper (Bray et al. 2016) on the AM560-2 genome version 6 (Bredeson et al. 2016) and EdgeR (Robinson et al. 2009) for differential expression analysis by comparing against the mock-inoculated treatment on R (version 3.6.1).

Retro-transcriptase (RT) semi-quantitative and quantitative PCR

Xpm strains were cultured in liquid media and prepared as described earlier. Bacterial inoculum was adjusted to an OD₆₀₀ of 0.5 (c.a. 5×10⁸ cfu/mL) in a 10-mM MgCl₂ solution. Bacterial suspensions or the mock inoculation solution (10-mM MgCl₂) were infiltrated into leaves of 6-week-old adult plants grown from cuttings, by means of a needleless syringe; each treatment was inoculated on one leaflet per plant and on three different individuals. Infiltrated tissue was collected at 50 hpi in sterile tubes containing RNA-free glass beads, samples were frozen with liquid nitrogen and ground by vortexing. Total RNA was extracted as described earlier. Total RNA was treated with TURBO™ DNase (Thermo Fisher Scientific, USA) per manufacturer's instructions. cDNA synthesis was performed with the SuperScript™ III Reverse Transcriptase (Thermo Fisher Scientific, USA) and an Oligo(dT)12-18 Primer (Thermo Fisher Scientific, USA), per manufacturer's instructions.

Semi-quantitative RT-PCRs were performed with a 20-μL reaction mix per tube containing 1X of 5X GoTaq Green Buffer, 1.5 mM of MgCl₂, 100 μM of a dNTP mix, 0.2 μM of each primer, 5 ng/μL of cDNA, and 0.025 U/μL of GoTaq polymerase (Promega, USA). Amplification was programmed as follows: one step at 95°C for 3 min, followed by a variable number of cycles (20 to 28, as needed) of 95°C for 30 sec, 60°C for 40 sec, and 72°C for 30 sec. Amplicons were resolved in 1% agarose gels on 0.5X TBE. RT-qPCRs were performed in 7-μL reactions of the Eurogentec Takyon™ SYBR® 2X qPCR Mastermix Blue (Sycamore Life Sciences, USA) containing 0.3 μM of each primer. PCR cycling was as follows: one step at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, and 60°C for 30 sec; data was acquired during the second step of each cycle. The melting curve was evaluated from 65°C to 95°C. Primers for gene *Manes.06G123400* encoding the sucrose transporter MeSWEET10a, *Manes.14G047800* encoding the sucrose transporter MeSWEET10e, and the reference gene *Manes.01G193700* encoding a serine/threonine-PP2A catalytic subunit (STPP) are summarized in Table 2.

TALE target prediction and artificial TALE construction and transformation

TALE targets were predicted using three different software: TALVEZ (Pérez-Quintero et al. 2013), TALE Nucleotide Targeter 2.0 (TALENT 2.0) (Doyle et al. 2012), and PrediTAL (Erkes et al. 2019). The *Manihot esculenta* promoterome (1-kb sequences preceding annotated translational start sites) was extracted from cassava genome version 6.1 from Phytozome (Bredeson et al. 2016), by means of the Biomart tool, and it was used as input for target prediction. All the algorithms were run using the default parameters. Output data were merged and compared on R (version 3.6.1) using an in-house script. arTALEs were designed by using the TALE Targeter tool of the TALE Nucleotide Targeter 2.0 suite, with the default parameters on the 500 bp preceding the start codon of the target gene (Doyle et al. 2012). Targets were predicted for the obtained candidates by means of TALVEZ, as described earlier. Selected arTALEs showed the target gene as first prediction, and the score of the second prediction was at least 5 units lower. Assembly was achieved through Golden TAL Technology, by using the arTALE construction kit described by (Geißler et al. 2011), using pSKX1 as expression vector. Constructed arTALEs are shown in Table 3.

Ortholog analysis

The *MeSWEETs* were retrieved from the Phytozome platform version 12 by searching for the PANTHER gene family PTHR10791, which corresponds to the RAG1-ACTIVATING PROTEIN 1, in the version 6.0 of the AM560-2 genome (Bredeson et al. 2016). The search resulted in 28 genes and 31 transcripts, whose amino acid sequence was downloaded. The tables of orthologs identified by Phytozome (via inParanoid 4.1 software) were retrieved for each gene, and the amino acid sequences of *MeSWEET* orthologs in rice (*O. sativa* version 7.0) and *Arabidopsis thaliana* TAIR10 were extracted through the Phytozome's tool BioMart. Additionally, the amino acid sequence of the *GhSWEET10D* (*Gh_D12G1898*) was retrieved from the CottonGen portal, from the *Gossypium hirsutum* genome version 1.1. Alignment of *MeSWEETs*, *AtSWEETs*, *OsSWEETs*, and *GhSWEET10D* was performed in ClustalW version 2.1 (Larkin et al. 2007) and then used as input to create a Neighbor-Joining tree with 1000 bootstraps in the Geneious tree builder. The resulting tree was visualized and edited in Geneious software (version R11).

Plots and synteny analysis

Volcano plots, RVD string representations, and bar plots were created using ggplot2 package (Wickham 2016) through an in-house script on R (version 3.6.1). Venn diagrams were created on R using the ggVennDiagram package (Gao and Yu 2021). Gene clusters, genes, and EBEs were plotted using Geneious software (version R11). Read mapping to the AM560-2 *MeSWEET* clusters and display were performed on the Geneious software (version R11) with the aid of the plugins BBDuk (version 37.64) and BBMerge (version 37.64). Synteny analyses and visualization were carried out in the Mauve (Darling et al. 2010) plugin of the Geneious software (version R11).

Acknowledgements

This research was funded by Research Program on Roots, Tubers and Bananas, and the Agropolis Foundation (project #1403-073). CAZ was supported by a doctoral fellowship awarded by the Institut de Recherche pour le Développement and received funding from the Biological Sciences Faculty from Universidad de los Andes. We acknowledge the advice of Lionel Gagnevin for DNA extraction and SMRT sequencing and Natalia Vargas for ortholog analysis.

References

- Andrés, F., Kinoshita, A., Kalluri, N., Fernández, V., Falavigna, V. S., Cruz, T. M. D., Jang, S., Chiba, Y., Seo, M., Mettler-Altmann, T., Huettel, B., and Coupland, G. 2020. The sugar transporter SWEET10 acts downstream of FLOWERING LOCUS T during floral transition of *Arabidopsis thaliana*. *BMC Plant Biol.* 20:53
- Anjali, A., Fatima, U., Manu, M. S., Ramasamy, S., and Senthil-Kumar, M. 2020. Structure and regulation of SWEET transporters in plants: An update. *Plant Physiol. Biochem.* 156:1–6
- Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., and Yang, B. 2010. Rice *xa13* Recessive Resistance to Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene *Os-11N3*. *Plant Cell.* 22:3864–3876
- Aung, K., Jiang, Y., and He, S. Y. 2018. The role of water in plant–microbe interactions. *Plant J.* 93:771–780
- Bart, R., Cohn, M., Kassen, A., McCallum, E. J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L. M., Neto, J. R., Verdier, V., Santana, M. A., Kositcharoenkul, N., Vanderschuren, H., Grisse, W., Bernal, A., and Staskawicz, B. J. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc. Natl. Acad. Sci. U. S. A.* 109:E1972-9
- Boch, J., and Bonas, U. 2010. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Annu. Rev. Phytopathol.* 48:419–436
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* (80-.). 326:1509–1512
- Bogdanove, A. J., Schornack, S., and Lahaye, T. 2010. TAL effectors: finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* 13:394–401
- Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34:525–527
- Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., Grimwood, J., Schmutz, J., Rabbi, I. Y., Egesi, C., Nauluvula, P., Lebot, V., Ndunguru, J., Mkamilo, G., Bart, R. S., Setter, T. L., Gleadow, R. M., Kulakow, P., Ferguson, M. E., Rounsley, S., and Rokhsar, D. S. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* 34:562–570
- Breia, R., Conde, A., Badim, H., Fortes, A. M., Gerós, H., and Granell, A. 2021. Plant SWEETs: from sugar transport to plant–pathogen interaction and more unexpected physiological roles. *Plant Physiol.*
- Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason 3rd, J. A., Stevens, R., Vonstein, V., Wattam, A. R., and Xia, F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365
- Cao, Y., Liu, W., Zhao, Q., Long, H., Li, Z., Liu, M., Zhou, X., and Zhang, L. 2019. Integrative analysis

- reveals evolutionary patterns and potential functions of SWEET transporters in Euphorbiaceae. *Int. J. Biol. Macromol.* 139:1–11
- Cernadas, R. A., Doyle, E. L., Niño-Liu, D. O., Wilkins, K. E., Bancroft, T., Wang, L., Schmidt, C. L., Caldo, R., Yang, B., White, F. F., Nettleton, D., Wise, R. P., and Bogdanove, A. J. 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 Pathog.* 10:e1003972
- Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., Guo, W.-J., Kim, J.-G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F. F., Somerville, S. C., Mudgett, M. B., and Frommer, W. B. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature.* 468:527–532
- Chen, L.-Q., Lin, I. W., Qu, X.-Q., Sosso, D., McFarlane, H. E., Londoño, A., Samuels, A. L., and Frommer, W. B. 2015. A Cascade of Sequentially Expressed Sucrose Transporters in the Seed Coat and Endosperm Provides Nutrition for the Arabidopsis Embryo. *Plant Cell.* 27:607–619
- Cohn, M., Bart, R. S., Shybut, M., Dahlbeck, D., Gomez, M., Morbitzer, R., Hou, B.-H., Frommer, W. B., Lahaye, T., and Staskawicz, B. J. 2014. *Xanthomonas axonopodis* Virulence Is Promoted by a Transcription Activator-Like Effector-Mediated Induction of a SWEET Sugar Transporter in Cassava. *Mol. Plant-Microbe Interact.* 27:1186–1198
- Cohn, M., Morbitzer, R., Lahaye, T., and Staskawicz, B. J. 2016. Comparison of gene activation by two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility genes in cassava. *Mol. Plant Pathol.* :875–889
- Cox, K. L., Meng, F., Wilkins, K. E., Li, F., Wang, P., Booher, N. J., Carpenter, S. C. D., Chen, L.-Q., Zheng, H., Gao, X., Zheng, Y., Fei, Z., Yu, J. Z., Isakeit, T., Wheeler, T., Frommer, W. B., He, P., Bogdanove, A. J., and Shan, L. 2017. TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. *Nat. Commun.* 8:15588
- Cuculis, L., Abil, Z., Zhao, H., and Schroeder, C. M. 2015. Direct observation of TALE protein dynamics reveals a two-state search mechanism. *Nat. Commun.* 6:7277
- Darling, A. E., Mau, B., and Perna, N. T. 2010. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS One.* 5:e11147
- Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., Chlenski, P., Conrad, N., Dickerman, A., Dietrich, E. M., Gabbard, J. L., Gerdes, S., Guard, A., Kenyon, R. W., Machi, D., Mao, C., Murphy-Olson, D., Nguyen, M., Nordberg, E. K., Olsen, G. J., Olson, R. D., Overbeek, J. C., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomas, C., VanOeffelen, M., Vonstein, V., Warren, A. S., Xia, F., Xie, D., Yoo, H., and Stevens, R. 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res.* 48:D606–D612
- Denancé, N., Szurek, B., Doyle, E. L., Lauber, E., Fontaine-Bodin, L., Carrère, S., Guy, E., Hajri, A., Cerutti, A., Boureau, T., Poussier, S., Arlat, M., Bogdanove, A. J., and Noël, L. D. 2018. Two ancestral genes shaped the *Xanthomonas campestris* TAL effector gene repertoire. *New Phytol.* 219:391–407
- Devanna, B. N., Jaswal, R., Singh, P. K., Kapoor, R., Jain, P., Kumar, G., Sharma, Y., Samantaray, S., and Sharma, T. R. 2021. Role of transporters in plant disease resistance. *Physiol. Plant.*

171:849–867

- Doyle, E. L., Booher, N. J., Standage, D. S., Voytas, D. F., Brendel, V. P., VanDyk, J. K., and Bogdanove, A. J. 2012. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.* 40:W117–W122
- Doyle, E. L., Hummel, A. W., Demorest, Z. L., Starker, C. G., Voytas, D. F., Bradley, P., and Bogdanove, A. J. 2013. 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
- Duan, S., Jia, H., Pang, Z., Teper, D., White, F., Jones, J., Zhou, C., and Wang, N. 2018. Functional characterization of the citrus canker susceptibility gene *CsLOB1*. *Mol. Plant Pathol.* 19:1908–1916
- Durand, M., Porcheron, B., Hennion, N., Maurousset, L., Lemoine, R., and Pourtau, N. 2016. Water Deficit Enhances C Export to the Roots in *Arabidopsis thaliana*; Plants with Contribution of Sucrose Transporters in Both Shoot and Roots. *Plant Physiol.* 170:1460 LP – 1479
- Erkes, A., Mücke, S., Reschke, M., Boch, J., and Grau, J. 2019. PrediTAL: A novel model learned from quantitative data allows for new perspectives on TALE targeting. *PLOS Comput. Biol.* 15:e1007206
- Fei, H., Yang, Z., Lu, Q., Wen, X., Zhang, Y., Zhang, A., and Lu, C. 2021. OsSWEET14 cooperates with OsSWEET11 to contribute to grain filling in rice. *Plant Sci.* 306:110851
- Feng, L., and Frommer, W. B. 2015. Structure and function of SemiSWEET and SWEET sugar transporters. *Trends Biochem. Sci.* 40:480–486
- Ferreira, R. M., de Oliveira, A. C. P., Moreira, L. M., Belasque, J., Gourbeyre, E., Siguier, P., Ferro, M. I. T., Ferro, J. A., Chandler, M., and Varani, A. M. 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
- Gao, C.-H., and Yu, G. 2021. ggVennDiagram: A “ggplot2” Implement of Venn Diagram. Available at: <https://cran.r-project.org/package=ggVennDiagram> [Accessed April 28, 2021].
- Geißler, R., Scholze, H., Hahn, S., Streubel, J., Bonas, U., Behrens, S.-E., and Boch, J. 2011. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One.* 6:e19509
- Grau, J., Reschke, M., Erkes, A., Streubel, J., Morgan, R. D., Wilson, G. G., Koebnik, R., and Boch, J. 2016. AnnoTALE: bioinformatics tools for identification, annotation and nomenclature of TALEs from *Xanthomonas* genomic sequences. *Sci. Rep.* 6:21077
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* 59:307–321
- Hu, M., Hu, W., Xia, Z., Zhou, X., and Wang, W. 2016. Validation of Reference Genes for Relative Quantitative Gene Expression Studies in Cassava (*Manihot esculenta* Crantz) by Using Quantitative Real-Time PCR. *Front. Plant Sci.* 7:680

- Hu, Y., Zhang, J., Jia, H., Sosso, D., Li, T., Frommer, W. B., Yang, B., White, F. F., Wang, N., and Jones, J. B. 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc. Natl. Acad. Sci. U. S. A.* 111:E521–E529
- Hui, S., Liu, H., Zhang, M., Chen, D., Li, Q., Tian, J., Xiao, J., Li, X., Wang, S., and Yuan, M. 2019a. The host basal transcription factor IIA subunits coordinate for facilitating infection of TALEs-carrying bacterial pathogens in rice. *Plant Sci.* 284:48–56
- Hui, S., Shi, Y., Tian, J., Wang, L., Li, Y., Wang, S., and Yuan, M. 2019b. TALE-carrying bacterial pathogens trap host nuclear import receptors for facilitation of infection of rice. *Mol. Plant Pathol.* 20:519–532
- Hunt, M., Silva, N. De, Otto, T. D., Parkhill, J., Keane, J. A., and Harris, S. R. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 16:294
- Hutin, M., Césari, S., Chalvon, V., Michel, C., Tran, T. T., Boch, J., Koebnik, R., Szurek, B., and Kroj, T. 2016. Ectopic activation of the rice NLR heteropair RGA4/RGA5 confers resistance to bacterial blight and bacterial leaf streak diseases. *Plant J.* 88:43–55
- Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R., and Szurek, B. 2015. A knowledge-based molecular screen uncovers a broad-spectrum OsSWEET14 resistance allele to bacterial blight from wild rice. *Plant J.* 84:694–703
- El Kasmi, F., Horvath, D., and Lahaye, T. 2018. Microbial effectors and the role of water and sugar in the infection battle ground. *Curr. Opin. Plant Biol.* 44:98–107
- Katoh, K., and Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780
- Kay, S., Hahn, S., Marois, E., Wieduwild, R., and Bonas, U. 2009. Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3 Δ rep16. *Plant J.* 59:859–871
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27:722–736
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Thompson, J. D., Gibson, T. J., and Higgins, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.* 23:2947–2948
- Lin, I. W., Sosso, D., Chen, L.-Q., Gase, K., Kim, S.-G., Kessler, D., Klinkenberg, P. M., Gorder, M. K., Hou, B.-H., Qu, X.-Q., Carter, C. J., Baldwin, I. T., and Frommer, W. B. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature.* 508:546–549
- Medina, C. A., Reyes, P. A., Trujillo, C. A., Gonzalez, J. L., Bejarano, D. A., Montenegro, N. A., Jacobs, J. M., Joe, A., Restrepo, S., Alfano, J. R., and Bernal, A. 2018. The role of type III effectors from *Xanthomonas axonopodis* pv. *manihotis* in virulence and suppression of plant immunity. *Mol. Plant Pathol.* 19:593–606
- Mora, R., Rodriguez, M., Gayosso, L., and López Carrascal, C. E. 2019. Using in vitro plants to study

- the cassava response to *Xanthomonas phaseoli* pv. *manihotis* infection. *Trop. Plant Pathol.*
- Moscou, M. J., and Bogdanove, A. J. 2009. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* (80-). 326:1501 LP – 1501
- Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J. C., Perez-Quintero, A., Li, T., Eom, J.-S., Li, C., Nguyen, H., Liu, B., Auguy, F., Sciallano, C., Luu, V. T., Dossa, G. S., Cunnac, S., Schmidt, S. M., Slamet-Loedin, I. H., Vera Cruz, C., Szurek, B., Frommer, W. B., White, F. F., and Yang, B. 2019. Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* 37:1344–1350
- Pereira, A., Carazzolle, M., Abe, V., de Oliveira, M., Domingues, M., Silva, J., Cernadas, R., and Benedetti, C. 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-Quintero, A. L., Lamy, L., Gordon, J. L., Escalon, A., Cunnac, S., Szurek, B., and Gagnevin, L. 2015. QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Front. Plant Sci.* 6:545
- Pérez-Quintero, A. L., Rodriguez-R, L. M., Dereeper, A., López Carrascal, C. E., Koebnik, R., Szurek, B., and 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
- Perez-Quintero, A. L., and Szurek, B. 2019. A Decade Decoded: Spies and Hackers in the History of TAL Effectors Research. *Annu. Rev. Phytopathol.* 57:459–481
- Posada, D., and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 14:817–818
- Prior, M. J., Selvanayagam, J., Kim, J.-G., Tomar, M., Jonikas, M., Mudgett, M. B., Smeekens, S., Hanson, J., and Frommer, W. B. 2021. Arabidopsis bZIP11 Is a Susceptibility Factor During *Pseudomonas syringae* Infection. *Mol. Plant-Microbe Interact.* 34:439–447
- Proels, R. K., and Roitsch, T. 2009. Extracellular invertase LIN6 of tomato: a pivotal enzyme for integration of metabolic, hormonal, and stress signals is regulated by a diurnal rhythm. *J. Exp. Bot.* 60:1555–1567
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2009. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 26:139–140
- Römer, P., Recht, S., Strauß, T., Elsaesser, J., Schornack, S., Boch, J., Wang, S., and 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 Phytol.* 187:1048–1057
- Rossier, O., Wengelnik, K., Hahn, K., and Bonas, U. 1999. The *Xanthomonas* Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. *Proc. Natl. Acad. Sci. U. S. A.* 96:9368–9373
- Ruh, M., Briand, M., Bonneau, S., Jacques, M.-A., and Chen, N. W. G. 2017. *Xanthomonas* adaptation to common bean is associated with horizontal transfers of genes encoding TAL

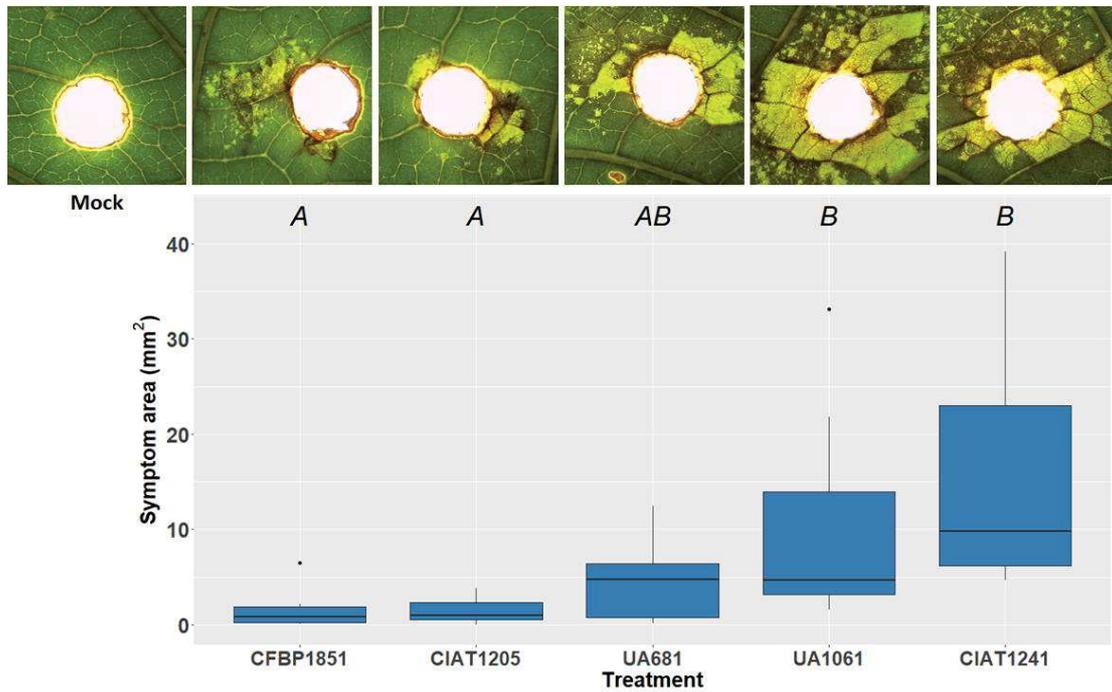
- effectors. *BMC Genomics*. 18:670
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. 9:671–675
- Schreiber, T., and Bonas, U. 2014. Repeat 1 of TAL effectors affects target specificity for the base at position zero. *Nucleic Acids Res.* 42:7160–7169
- Schwartz, A. R., Morbitzer, R., Lahaye, T., and Staskawicz, B. J. 2017. TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. *Proc. Natl. Acad. Sci.* 114:E897 LP-E903
- Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J., and Szurek, B. 2013. Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* 200:808–819
- Sun, M.-X., Huang, X.-Y., Yang, J., Guan, Y.-F., and Yang, Z.-N. 2013. Arabidopsis RPG1 is important for primexine deposition and functions redundantly with RPG2 for plant fertility at the late reproductive stage. *Plant Reprod.* 26:83–91
- Szurek, B., Marois, E., Bonas, U., and Van den Ackerveken, G. 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 J.* 26:523–534
- Szurek, B., Rossier, O., Hause, G., and Bonas, U. 2002. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol. Microbiol.* 46:13–23
- Tran, T. T., Pérez-Quintero, A. L., Wonni, I., Carpenter, S. C. D., Yu, Y., Wang, L., Leach, J. E., Verdier, V., Cunnac, S., Bogdanove, A. J., Koebnik, R., Hutin, M., and Szurek, B. 2018. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLOS Pathog.* 14:e1007092
- Triplett, L. R., Cohen, S. P., Heffelfinger, C., Schmidt, C. L., Huerta, A. I., Tekete, C., Verdier, V., Bogdanove, A. J., and Leach, J. E. 2016. A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. *Plant J.* 87:472–483
- Wan, H., Chang, S., Hu, J., Tian, X., and Wang, M. 2016. Potential Role of the Last Half Repeat in TAL Effectors Revealed by a Molecular Simulation Study Z. Liang, ed. *Biomed Res. Int.* 2016:8036450
- Weber, B., Zicola, J., Oka, R., and Stam, M. 2016. Plant Enhancers: A Call for Discovery. *Trends Plant Sci.* 21:974–987
- Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wicky, B. I. M., Stenta, M., and Dal Peraro, M. 2013. TAL Effectors Specificity Stems from Negative Discrimination. *PLoS One.* 8:e80261
- Wu, Y., Lee, S.-K., Yoo, Y., Wei, J., Kwon, S.-Y., Lee, S.-W., Jeon, J.-S., and An, G. 2018. Rice Transcription Factor OsDOF11 Modulates Sugar Transport by Promoting Expression of Sucrose Transporter and *SWEET* Genes. *Mol. Plant.* 11:833–845
- Yang, B., Sugio, A., and White, F. F. 2006. *Os8N3* is a host disease-susceptibility gene for bacterial

blight of rice. Proc. Natl. Acad. Sci. U. S. A. 103:10503–10508

- Yang, Y., and Gabriel, D. W. 1995. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. Mol. Plant. Microbe. Interact. 8:627–631
- Yu, Y., Streubel, J., Balzergue, S., Champion, A., Boch, J., Koebnik, R., Feng, J., Verdier, V., and 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. Mol. Plant-Microbe Interact. 24:1102–1113
- Yuan, M., Ke, Y., Huang, R., Ma, L., Yang, Z., Chu, Z., Xiao, J., Li, X., and Wang, S. 2016. A host basal transcription factor is a key component for infection of rice by TALE-carrying bacteria T. Nürnberg, ed. Elife. 5:e19605
- Zárate-Chaves, C. A., Osorio-Rodríguez, D., Mora, R. E., Pérez-Quintero, Á. L., Dereeper, A., Restrepo, S., López, C. E., Szurek, B., and Bernal, A. 2021. TAL Effector Repertoires of Strains of *Xanthomonas phaseoli* pv. *manihotis* in Commercial Cassava Crops Reveal High Diversity at the Country Scale. Microorganisms. 9
- Zhang, J., Huguet -Tapia, J. C., Hu, Y., Jones, J., Wang, N., Liu, S., and White, F. F. 2017. Homologues of CsLOB1 in citrus function as disease susceptibility genes in citrus canker. Mol. Plant Pathol. 18:798–810
- Zheng, C.-K., Wang, C.-L., Zhang, X.-P., Wang, F.-J., Qin, T.-F., and Zhao, K.-J. 2014. The last half-repeat of transcription activator-like effector (TALE) is dispensable and thereby TALE-based technology can be simplified. Mol. Plant Pathol. 15:690–697
- Zhou, J., Peng, Z., Long, J., Sosso, D., Liu, B., Eom, J.-S., Huang, S., Liu, S., Vera Cruz, C., Frommer, W. B., White, F. F., and Yang, B. 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. Plant J. 82:632–643
- Zhu, W., Yang, B., Chittoor, J. M., Johnson, L. B., and White, F. F. 1998. AvrXa10 Contains an Acidic Transcriptional Activation Domain in the Functionally Conserved C Terminus. Mol. Plant-Microbe Interact. 11:824–832

Figures

A



B

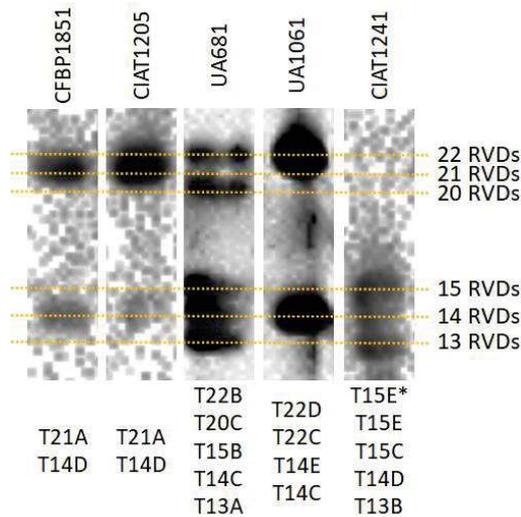
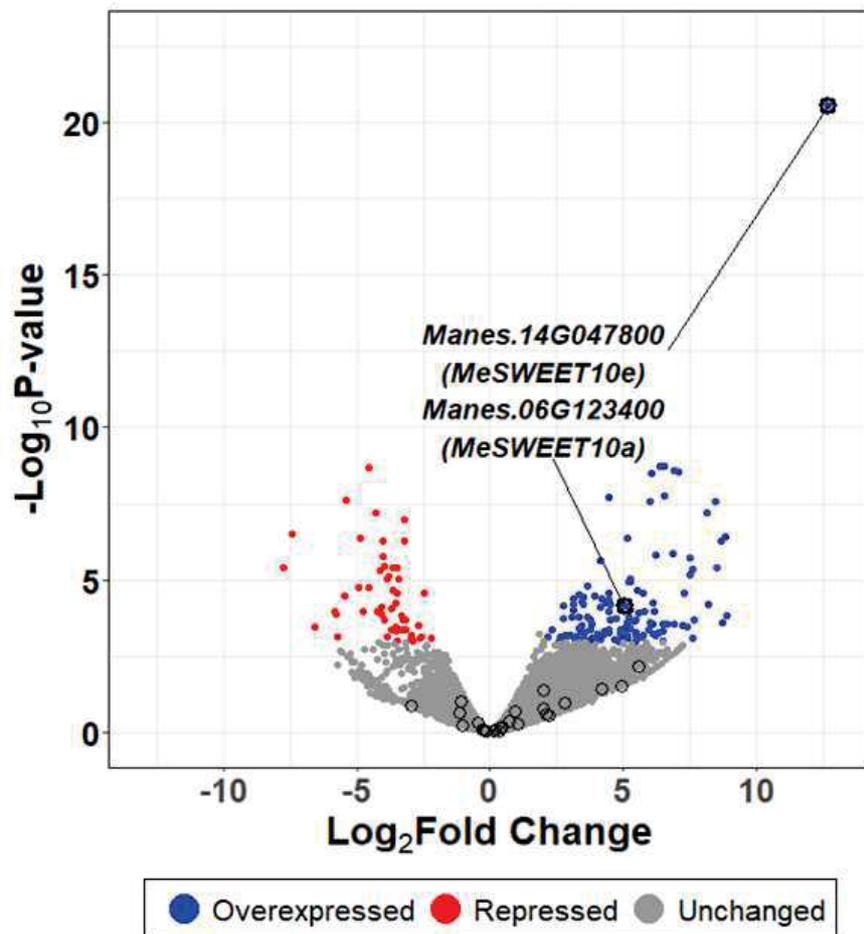


Figure 1. Aggressiveness and TALE repertoires of a panel of *Xpm* strains. A. Boxplot of the aggressiveness of *Xpm* strains CFBP1851, CIAT1205, UA681, UA1061 and CIAT1241 quantified as lesion areas. Italicized letters indicate groups with significant differences according to Tukey ($\alpha=0.05$). Representative photographs of lesions induced by the corresponding strain are shown at the top. B. Southern blot analysis of *Xpm* genomic DNA digested with *Bam*HI and using a 506-bp probe amplified from the C-terminal-coding region of *TALE1_{Xam}*. Results come from two different membranes; each lane was therefore aligned to build the figure. Dotted yellow lines show the position of bands with the predicted RVD number indicated on the right. The TALEs present in each strain are detailed at the bottom, based on genomic sequences (Table 1).

A



B

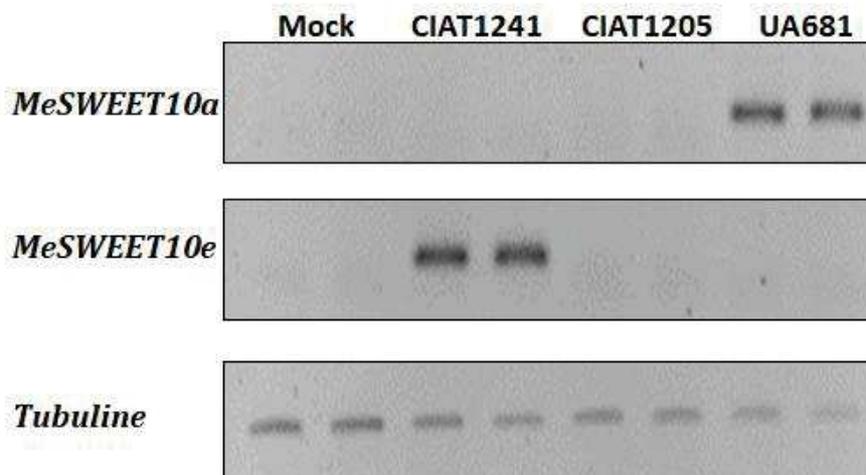


Figure 2. *Xpm* strain CIAT1241 induces the expression of *MeSWEET10e* but not *MeSWEET10a*. A. Volcano plot summarizing the relative expression of each cassava gene at 50 hpi when comparing

the treatments for cassava leaves inoculated with *Xpm* strain CIAT1241 versus mock-inoculated plants. Differentially expressed genes are colored in blue ($\text{Log}_2\text{FoldChange} \geq 2$ and $p\text{-value} \leq 0.001$) or red ($\text{Log}_2\text{FoldChange} \leq -2$ and $p\text{-value} \leq 0.001$). Black circles indicate the 28 *MeSWEET* paralogs of which *MeSWEET10a* and *MeSWEET10e* are highlighted. B. Semiquantitative RT-PCR analysis of *MeSWEET10a* and *MeSWEET10e* at 50 hours post inoculation of susceptible cassava 60444 with *Xpm* strains CIAT1241, CIAT1205 and UA681. This experiment was reproduced three times with similar results.

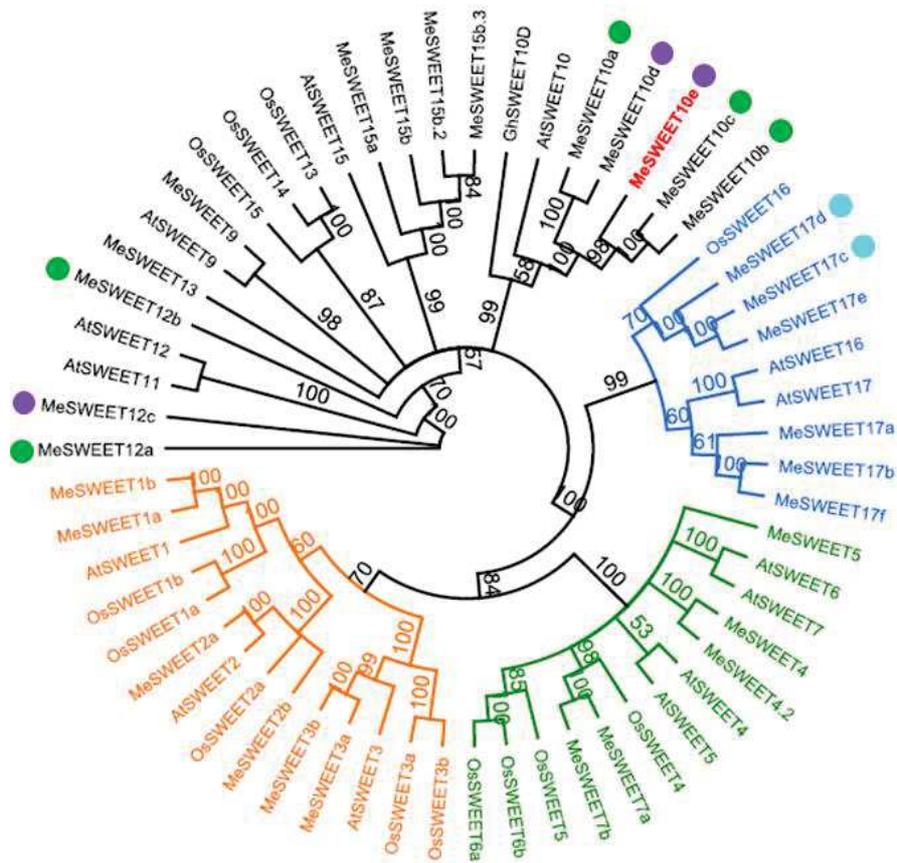


Figure 3. Ortholog analysis of cassava SWEET proteins. Neighbor-Joining analysis with 1000 bootstraps showing the relationship between cassava (*MeSWEET*), *A. thaliana* (*AtSWEET*), rice (*OsSWEET*) SWEET proteins, and cotton *GhSWEET10D*. Branch colors reflect the four clades: clade I in orange, clade II in green, clade III in black, and clade IV in blue, while branch numbers correspond to bootstrap values. *MeSWEET10e* is labeled in bold, and red. Colored dots highlight *MeSWEETs* in genomic clusters as identified in the AM560-2 cassava genome: green for the cluster located in chromosome 6, cyan for the cluster located in chromosome 8, and purple for the cluster located in chromosome 14.

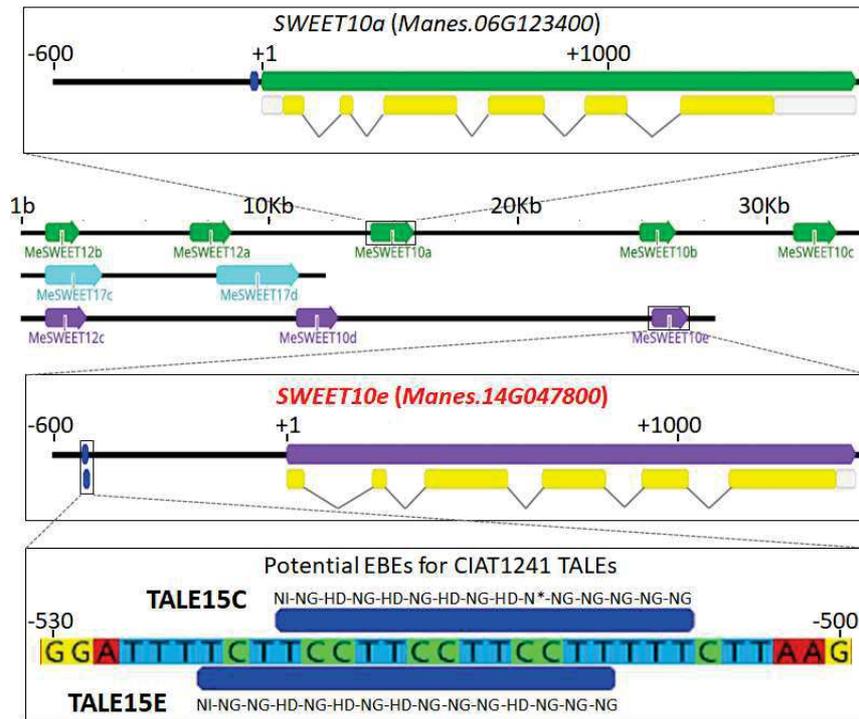


Figure 4. Genomic localization of *MeSWEET10e* and its predicted EBEs. Genomic *MeSWEET* clusters with *MeSWEETs* represented as colored arrows. The color code is the same as for Figure 3. The upper and middle boxes depict *MeSWEET10a* and *MeSWEET10e* predicted gene structure (exons as yellow rectangles, untranslated regions as gray rectangles) and EBEs (blue rectangles). The lower box highlights the predicted EBEs for TALE15C_{XpmCIAT1241} and TALE15E_{XpmCIAT1241} in the promoter region of *MeSWEET10e*.

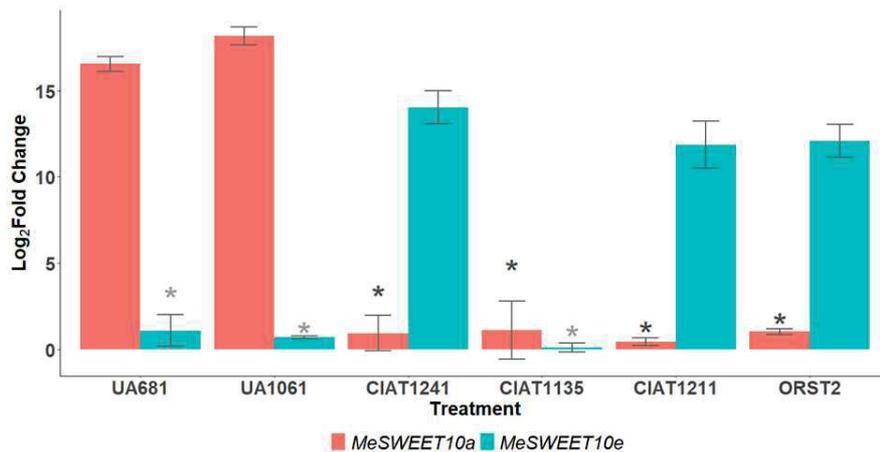


Figure 5. *MeSWEET10a* and *MeSWEET10e* induction 50hpi with a set of *Xpm* strains lacking TALE20 variants. Relative fold change was normalized against a mock-inoculated sample, using *STPP* gene as calibrator. Significance of mean differences were assessed through a t-test ($\alpha=0.05$) against the corresponding comparative treatment: gray stars indicate significant differences for *MeSWEET10a* relative fold change of a given treatment against the UA681 treatment, black stars indicate

significant differences for *MeSWEET10e* relative fold change of a given treatment against the CIAT1241 treatment. Error bars represent the standard deviation of three replicates.

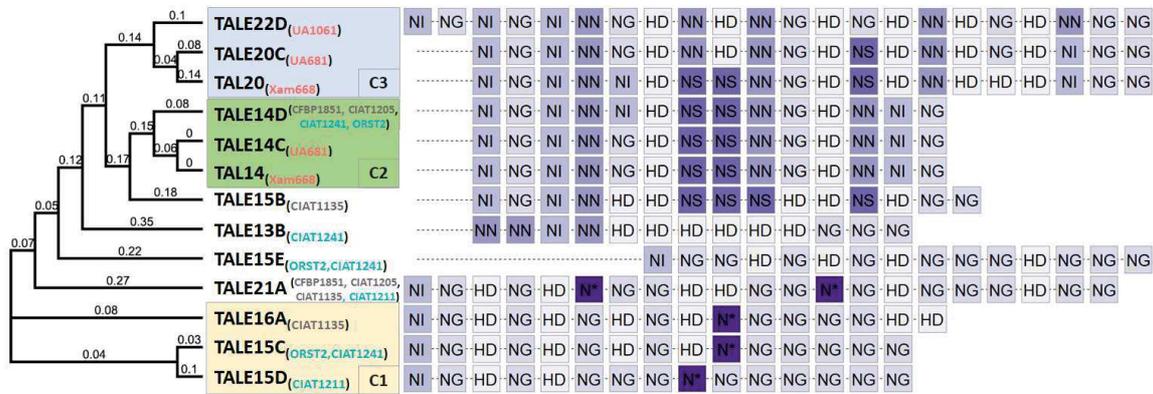


Figure 6. Functional relationships and shared RVD strings among TALE variants found in this study. Left dendrogram depicts the distance-based clusters defined by FuncTAL. Values on the branches correspond to distances. Labels indicate the variant as bolded black text, and the strain where the given TALE variant is present between brackets and subscript. The name of each strain is color coded as in Figure 5, coral for strains activating *MeSWEET10a*, light blue for strains activating *MeSWEET10e*, and gray for strains that do not activate either of these genes. Blue, green, and yellow boxes around labels highlight the binding affinity clusters C1, C2 and C3, respectively. The cut-off value for defining these clusters was based on cluster 3 (0.28), whose functional redundancy has been validated (Cohn et al. 2014; Zárate-Chaves et al. 2021). The right part of the plot shows the corresponding alignment of RVD strings based on the FuncTAL clustering. Each box represents a repeat, and the text inside indicates their corresponding RVD. RVD types are color coded in a purple palette to facilitate visualization of patterns.

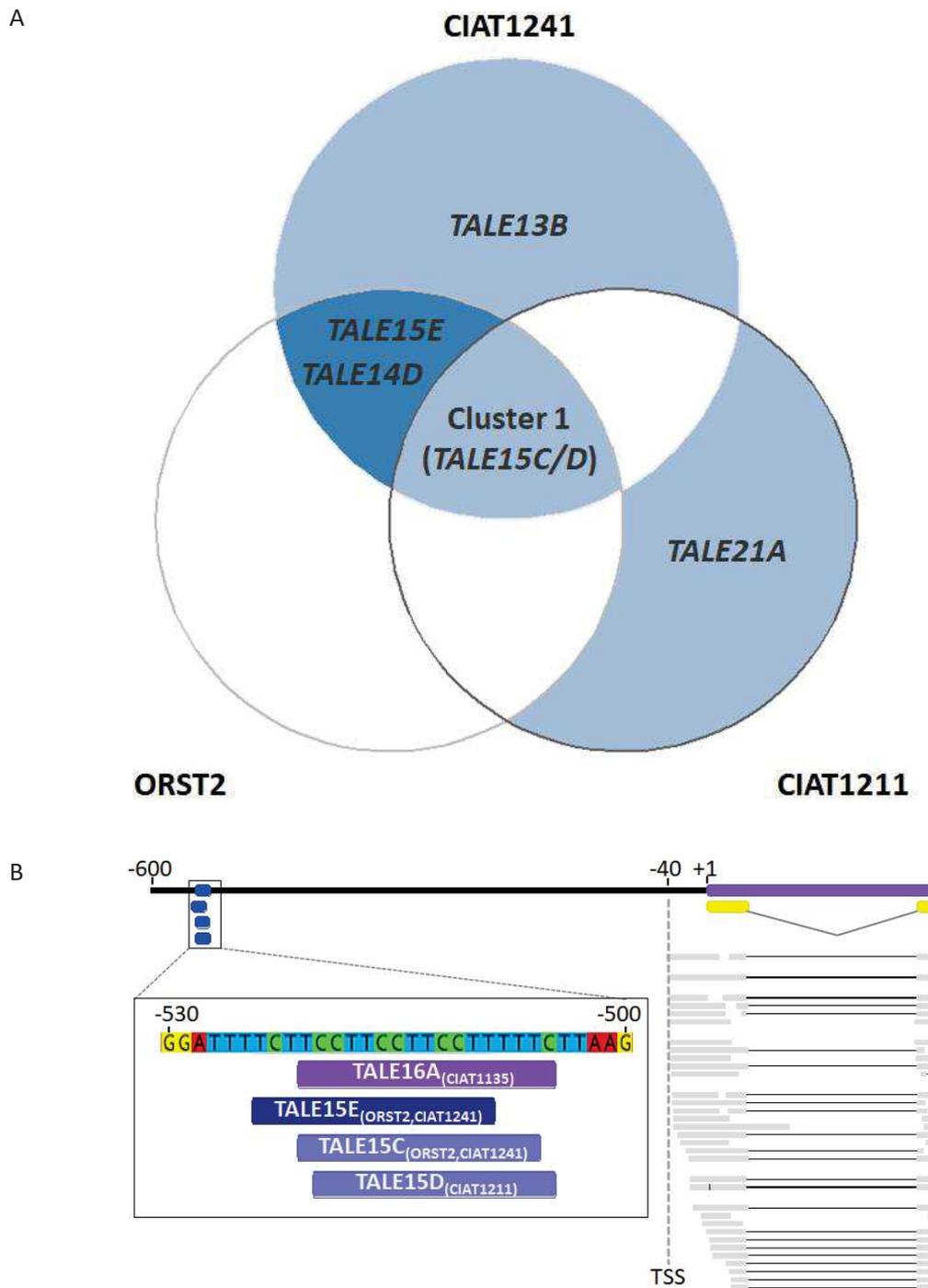


Figure 7. Comparison of functional clusters and TALE variants present in *MeSWEEP10e*-upregulating strains, predicted EBEs, and predicted transcript start. A. Venn's diagram of TALE variants present in strains (CIAT1241, CIAT1211 and ORST2) that activate *MeSWEEP10e* after inoculation. Variants TALE15C and TALE15D were considered as functionally redundant according to FuncTAL results. B. EBEs (blue and purple rectangles on the zoom box) predicted in the promoter region of *MeSWEEP10e* for FuncTAL cluster 1 TALE variants. RNAseq reads were assembled and mapped to

the *MeSWEET10e* sequence and its promoter region. Only the first two exons are shown (yellow rectangles) with the mapped reads as gray rectangles linked by a black line. The dotted gray line indicates that the potential TSS is located at the -40 bp.

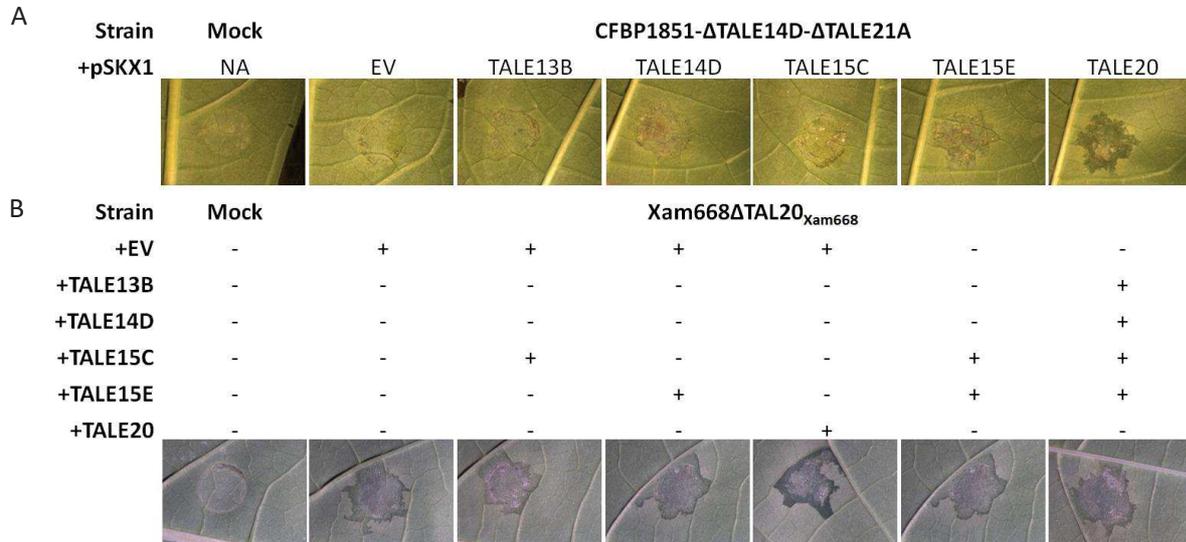


Figure 8. Search for *SWEET10e*-inducing candidate TALEs from *Xpm* strain CIAT1241. A. Leaves of cassava 60444 adult plants were infiltrated with the TALE-free mutant strain CFBP1851- Δ TALE14D- Δ TALE21A carrying an empty vector (EV) or expressing *TALE13B*, *TALE14D*, *TALE15C* or *TALE15E* from *Xpm* strain CIAT1241, or *TALE20C* from strain UA522 as positive control. B. Leaves of cassava 60444 adult plants were infiltrated with individual or combinations of mutant strains Xam668 Δ TAL20_{Xam668} complemented with an empty vector (EV), or expressing *SWEET10e*-inducing CIAT1241 *TALE* candidates, or *TALE20* used as positive control. NA: not applicable. Infiltration points of all panels were photographed at 96 hpi.

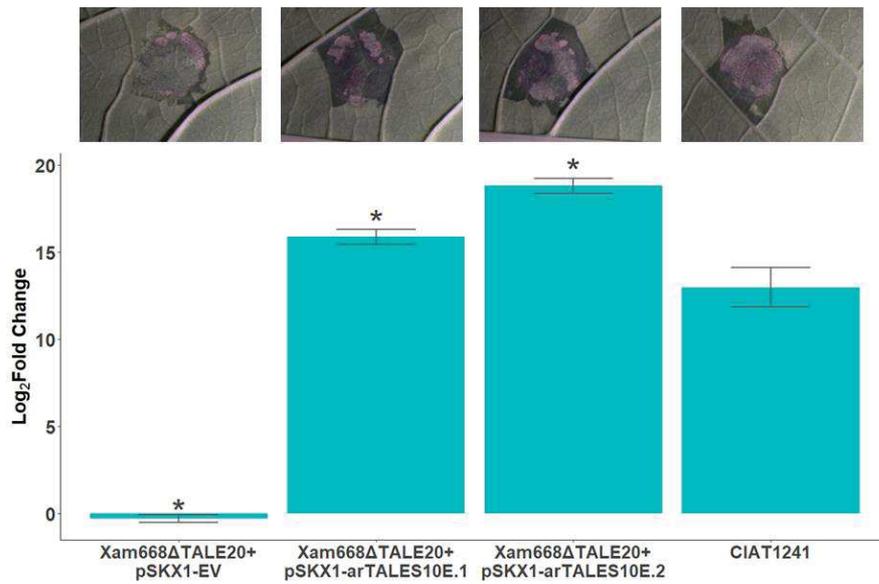


Figure 9. ArTALEs targeting *MeSWEET10e* restore the virulence of mutant strain Xam668Δ*TAL20*_{xam668}. Water-soaked symptoms and *MeSWEET10e* induction at, respectively, 96 and 50 hours post infiltration of cassava leaves with mutant strain Xam668Δ*TAL20*_{xam668} carrying an empty vector (EV) or expressing *MeSWEET10e*-targeting *arTALEs*. Stars indicate significant differences between the induction of the given treatment and infection with CIAT1241 according to a t-test ($\alpha=0.05$). Error bars represent the standard deviation of three replicates.

Tables

Table 1. TALomes from *Xpm* strains sequenced in this study. The variant name shown here has been shortened, but it must be followed by the subscript *Xpm* + the name of the strain. ^{chr} indicates that there is one copy of this TALE inserted into the chromosome. The rest of these genes were found in plasmids.

Strain	Variant	RVDs
CIAT1205	TALE14D	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NN-NI-NG
	TALE21A	NI-NG-HD-NG-HD-N*-NG-NG-HD-HD-NG-NG-N*-NG-HD-NG-NG-NG-HD-NG-NG
CIAT1241	TALE13B	NN-NN-NI-NN-HD-HD-HD-HD-HD-NG-NG-NG
	TALE14D	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NN-NI-NG
	TALE15C	NI-NG-HD-NG-HD-NG-HD-NG-HD-N*-NG-NG-NG-NG-NG
	TALE15E	NI-NG-NG-HD-NG-HD-NG-HD-NG-NG-NG-HD-NG-NG-NG
	TALE15E ^{chr}	NI-NG-NG-HD-NG-HD-NG-HD-NG-NG-NG-HD-NG-NG-NG
CIAT1135	TALE15B	NI-NG-NI-NN-HD-HD-NS-NS-NS-HD-HD-NS-HD-NG-NG
	TALE16A	NI-NG-HD-NG-HD-NG-HD-NG-HD-N*-NG-NG-NG-NG-HD-HD
	TALE21A	NI-NG-HD-NG-HD-N*-NG-NG-HD-HD-NG-NG-N*-NG-HD-NG-NG-NG-HD-NG-NG
CIAT1211	TALE15D	NI-NG-HD-NG-HD-NG-NG-NG-N*-NG-NG-NG-NG-NG-NG
	TALE21A	NI-NG-HD-NG-HD-N*-NG-NG-HD-HD-NG-NG-N*-NG-HD-NG-NG-NG-HD-NG-NG
ORST2	TALE14D	NI-NG-NI-NN-NI-HD-NS-NS-NN-NG-HD-NN-NI-NG
	TALE15C	NI-NG-HD-NG-HD-NG-HD-NG-HD-N*-NG-NG-NG-NG-NG
	TALE15E	NI-NG-NG-HD-NG-HD-NG-HD-NG-NG-NG-HD-NG-NG-NG

Table 2. Primers used in this study. Primers conceived for this study were designed with primer3 (<http://primer3.ut.ee/>) with set parameters to obtain melting temperatures of 60±2°C.

ID	Sequence (5'→3')	Reference
MeSWEET10a-Fw	TCCTCACCTTGACTGCGGTG	(Cohn et al. 2014)
MeSWEET10a-Rv	AGCACCATCTGGACAATCCCA	(Cohn et al. 2014)
MeSWEET10e-Fw	CGTTTTCGTTGCTATTCCAAAT	This study
MeSWEET10e-Rv	TCAAGTGTTGGCTTCTCAA	This study
STPP-Fw	GCTTGTCATGGAAGGGTACAA	(Hu et al. 2016)
STPP-Rv	TTCCACATCGGTAGCAATAG	(Hu et al. 2016)

Table 3. arTALEs targeting *MeSWEET10e*. Both were expressed from pSKX1 and both target the positive strand in the *MeSWEET10e* promoter region.

arTALE ID	RVD sequence	Target sequence	Start	End
arTALES10E.1	NH-HD-NI-NI-NI-NI-HD-NI-NG-NI-NH-NH-HD-NI-NH-NI-NN-NG	TGCAAAACATAGGCAGAGT	-52	-33

arTALES10E.2	NH-NH-NG-NG-HD-NI-HD-NI- NI-HD-HD-NG-NH-NG-NI-NN- NG-NG	TGGTTCACAACCTGTAGTT	-75	-56
--------------	---	---------------------	-----	-----

Supporting Information

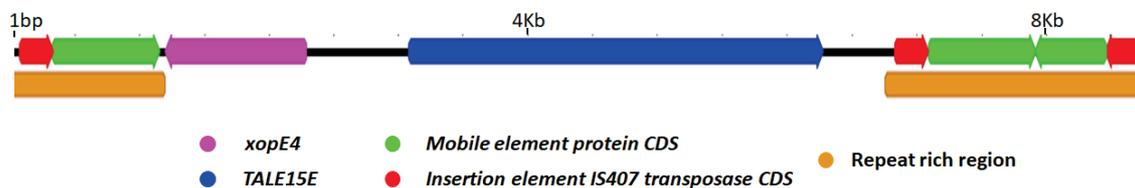


Figure S1. Graphical representation of the 8.8-Kb chromosome insertion where *TALE15E*_{*Xpm*CIAT1241} is found. Genes are represented as colored arrows. Ochre rectangles indicate repeat regions.

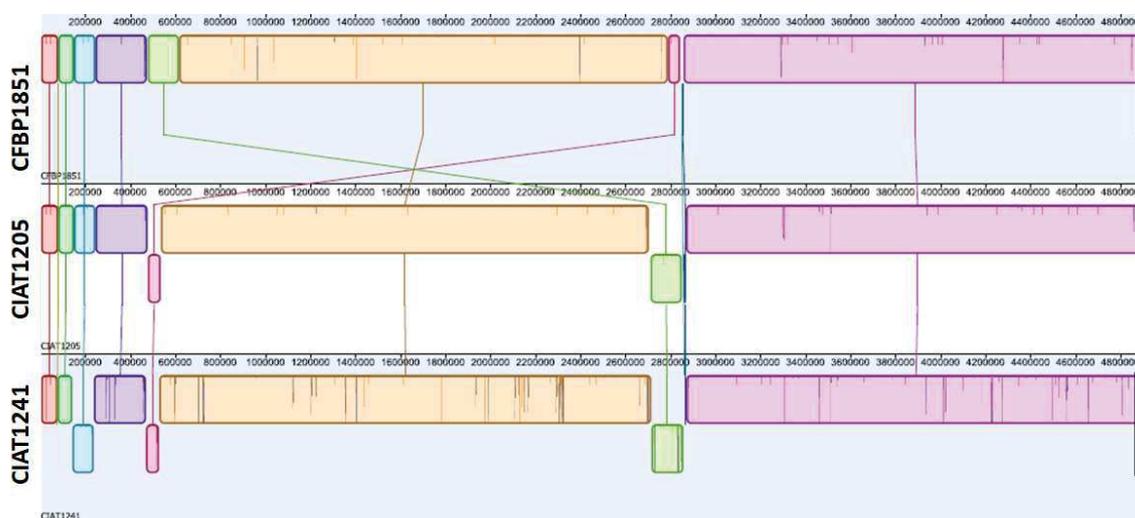
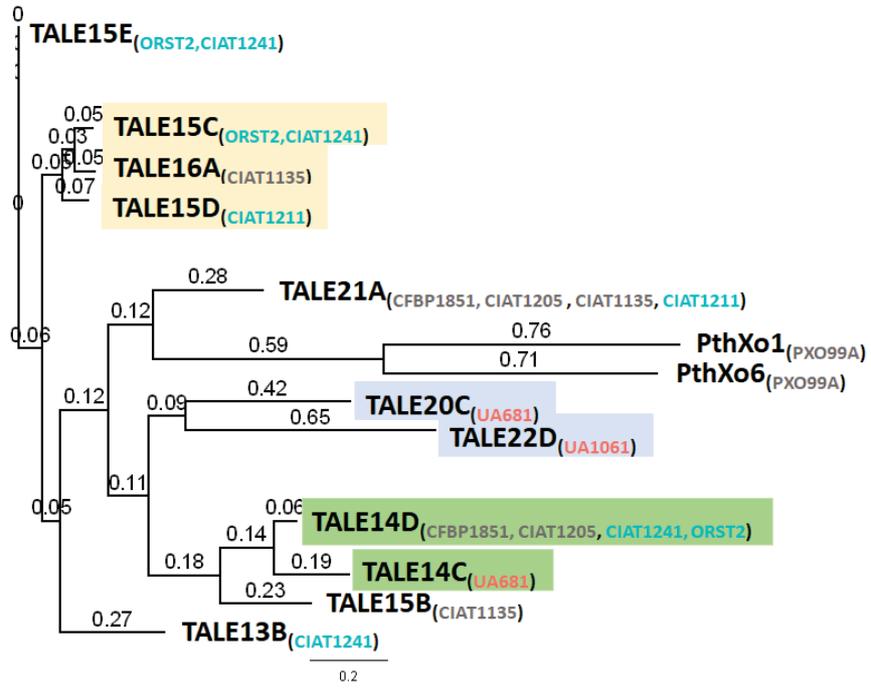


Figure S2. Synteny analysis for chromosomes of *Xpm* strains CFBP1851, CIAT1205, and CIAT1241. Each colored block represents a syntenic block connected by a colored line across the three chromosomes. Blocks shifted below the other blocks indicate inversions against the reference.

A



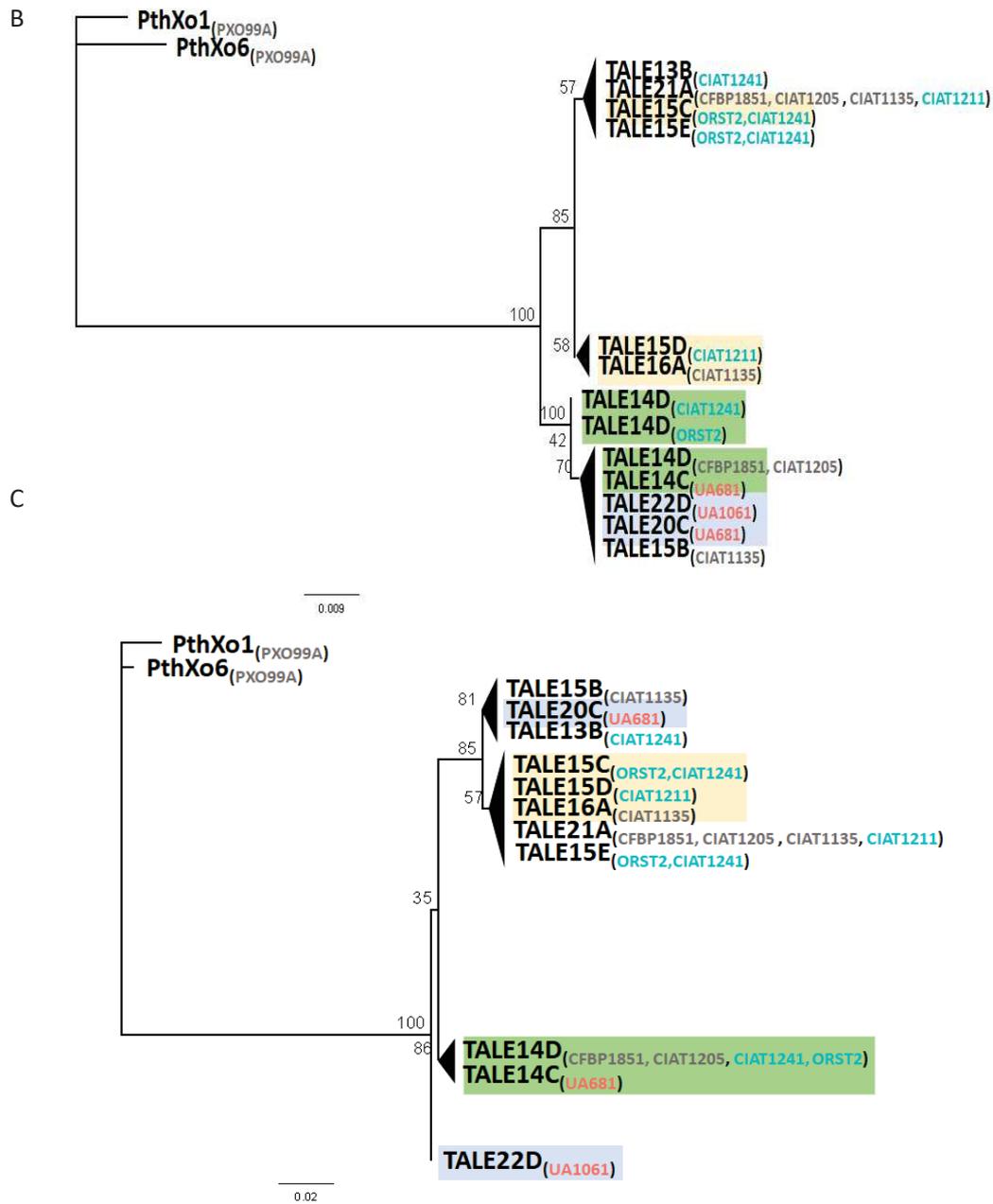


Figure S3. Phylogenetic relationships among TALE variants according to their repeat composition, N- and C-terminal regions. A. Phylogeny of TALEs according to repeat composition as inferred by DisTAL. Numbers on branches correspond to distances. Labels indicate the variant as bolded black text, and the strain where the given TALE variant is present between brackets and subscript. The name of each strain is color-coded as in Figure 5, coral for strains activating *MeSWEET10a*, light blue for strains activating *MeSWEET10e*, and gray for strains that do not activate either of these genes. Yellow, green, and blue boxes around labels reflect the FuncTAL clustering. N-terminal-based (B) and C-terminal-based (C) phylogeny of TALEs inferred with the HKY model with 100 bootstrap replications. Numbers on branches correspond to bootstrap values. TALE variant labeling is as

explained above. In all the cases, *PthXo1* and *PthXo6* from *Xoo* strain PX099^A were used as an outgroup.

Table S1. Differentially upregulated genes and their EBE predictions. The quality of EBE prediction for a given candidate was calculated as the sum of the three determined ranks (one per software). Totals were then categorized as E (Excellent, combined rank = 1 to 300), VG (Very Good, combined rank = 301 to 600), G (Good, combined rank = 601 to 900), and F (Fair, combined rank = 901 to 1200). The Log₂Fold Change value corresponds to the RNAseq-based differential expression analysis.

Gene ID	EBE predictions				Log ₂ Fold Change
	TALE13B	TALE14D	TALE15C	TALE15E	
<i>Manes.14G047800</i>			Fair	Good	12.68
<i>Manes.04G109300</i>		Very good			8.88
<i>Manes.10G008800</i>	Excellent	Fair	Fair	Good	8.85
<i>Manes.03G157000</i>		Good	Fair	Fair	8.74
<i>Manes.05G168000</i>		Good		Very good	8.67
<i>Manes.09G145800</i>		Good			8.49
<i>Manes.01G141900</i>	Excellent	Very good	Very good	Good	8.44
<i>Manes.14G130000</i>	Fair	Fair		Excellent	8.19
<i>Manes.10G025900</i>	Very good	Fair	Good	Fair	8.11
<i>Manes.01G067400</i>	Good	Good		Fair	7.67
<i>Manes.01G101800</i>	Very good		Fair	Fair	7.6
<i>Manes.02G111500</i>		Fair	Fair		7.59
<i>Manes.07G120000</i>	Fair	Fair	Fair	Fair	7.52
<i>Manes.07G058500</i>	Good	Very good	Good	Good	7.51
<i>Manes.12G001500</i>		Good	Fair	Good	7.39
<i>Manes.12G117300</i>	Very good	Very good	Fair	Good	7.29
<i>Manes.06G148300</i>	Very good	Good	Fair	Fair	7.21
<i>Manes.03G164700</i>	Good	Very good			7.05
<i>Manes.02G172200</i>		Very good			6.9
<i>Manes.11G151300</i>		Excellent		Fair	6.89
<i>Manes.13G117300</i>	Good	Good	Fair		6.84
<i>Manes.01G170600</i>	Excellent	Fair		Fair	6.62
<i>Manes.09G164700</i>		Good	Fair	Fair	6.56
<i>Manes.04G153000</i>	Fair	Fair	Fair	Good	6.53
<i>Manes.09G056300</i>		Very good	Good	Good	6.48
<i>Manes.01G030800</i>			Fair		6.48
<i>Manes.12G029800</i>			Fair	Good	6.48
<i>Manes.12G123200</i>		Good			6.48
<i>Manes.04G118200</i>	Good	Excellent	Fair		6.39
<i>Manes.01G228700</i>	Excellent	Fair	Fair	Fair	6.34
<i>Manes.14G098200</i>	Very good	Fair	Fair		6.24
<i>Manes.05G167500</i>	Very good	Very good	Fair		6.2

<i>Manes.01G085600</i>	Good	Good	Fair	Good	6.16
<i>Manes.01G232400</i>	Very good		Fair	Good	6.13
<i>Manes.06G118700</i>		Good		Fair	6.13
<i>Manes.05G078000</i>	Good		Fair		6.12
<i>Manes.05G167700</i>				Good	6.05
<i>Manes.14G060800</i>	Very good	Fair		Fair	6.02
<i>Manes.12G135500</i>	Good	Fair			5.98
<i>Manes.02G058200</i>	Good	Excellent	Good	Good	5.96
<i>Manes.02G161100</i>	Excellent	Very good		Fair	5.8
<i>Manes.14G096700</i>	Excellent	Very good	Fair	Good	5.74
<i>Manes.04G136100</i>		Fair			5.73
<i>Manes.02G118300</i>	Excellent	Fair	Fair	Good	5.69
<i>Manes.14G145500</i>	Very good	Very good	Fair	Fair	5.62
<i>Manes.12G071700</i>	Excellent	Excellent	Fair	Good	5.56
<i>Manes.03G105000</i>	Good	Very good	Fair	Fair	5.5
<i>Manes.18G139500</i>		Good	Good		5.48
<i>Manes.13G114400</i>	Good	Very good	Excellent	Very good	5.48
<i>Manes.01G196000</i>	Good	Fair	Fair	Good	5.46
<i>Manes.07G092900</i>	Fair	Very good		Good	5.45
<i>Manes.S062500</i>		Good			5.45
<i>Manes.12G157700</i>		Fair	Fair	Very good	5.4
<i>Manes.01G067500</i>		Very good	Fair	Fair	5.26
<i>Manes.05G170500</i>	Fair	Excellent	Good	Good	5.26
<i>Manes.15G071000</i>	Fair	Good	Fair	Very good	5.24
<i>Manes.05G012900</i>		Fair	Fair		5.24
<i>Manes.17G048200</i>	Excellent	Good	Excellent	Good	5.19
<i>Manes.12G079500</i>	Excellent	Good			5.18
<i>Manes.03G044100</i>	Very good	Excellent		Good	5.17
<i>Manes.05G023400</i>	Excellent	Fair	Fair	Fair	5.16
<i>Manes.04G147900</i>		Very good			5.15
<i>Manes.03G152600</i>	Good	Excellent	Fair	Fair	5.14
<i>Manes.13G012700</i>	Excellent	Fair		Fair	5.13
<i>Manes.11G159900</i>	Good		Fair	Fair	5.12
<i>Manes.03G059400</i>		Good	Fair		5.11
<i>Manes.14G028400</i>	Very good	Fair	Fair	Good	5.05
<i>Manes.11G072700</i>		Good			5.05
<i>Manes.06G123400</i>	Very good	Good		Good	5.03
<i>Manes.09G113800</i>		Very good		Fair	4.96
<i>Manes.09G105800</i>		Good			4.95
<i>Manes.10G103600</i>	Good	Very good	Fair	Fair	4.93
<i>Manes.02G042500</i>	Very good	Good	Fair	Good	4.93

<i>Manes.16G055200</i>			Very good	Good	4.89
<i>Manes.09G149000</i>	Very good	Very good	Fair	Good	4.88
<i>Manes.08G082800</i>		Very good	Fair	Fair	4.83
<i>Manes.10G030200</i>		Fair	Fair	Good	4.72
<i>Manes.17G076200</i>		Good		Good	4.68
<i>Manes.13G051600</i>	Good	Very good	Fair	Fair	4.67
<i>Manes.16G118800</i>	Very good	Fair	Fair	Good	4.53
<i>Manes.13G143800</i>		Excellent		Fair	4.53
<i>Manes.05G167900</i>	Excellent	Very good		Fair	4.48
<i>Manes.02G164500</i>	Excellent	Fair	Very good	Good	4.48
<i>Manes.09G175800</i>	Very good		Good	Good	4.48
<i>Manes.05G091500</i>		Fair		Good	4.47
<i>Manes.17G038600</i>		Excellent	Very good	Fair	4.44
<i>Manes.05G018200</i>	Fair	Good	Fair	Fair	4.36
<i>Manes.09G113500</i>	Good	Very good	Fair	Very good	4.32
<i>Manes.01G121300</i>	Excellent	Very good	Fair		4.29
<i>Manes.S002300</i>	Good	Very good			4.18
<i>Manes.04G113100</i>		Fair		Fair	4.16
<i>Manes.14G029100</i>	Excellent	Fair	Fair	Good	4.14
<i>Manes.13G028500</i>		Very good	Fair		4.11
<i>Manes.03G048200</i>	Fair	Very good	Fair	Good	4.04
<i>Manes.16G116900</i>	Good	Very good	Fair	Good	3.96
<i>Manes.06G051700</i>	Excellent	Very good	Fair	Fair	3.93
<i>Manes.11G165400</i>	Excellent	Very good	Fair	Fair	3.81
<i>Manes.05G086500</i>	Good	Very good	Good	Good	3.79
<i>Manes.01G129800</i>	Good	Fair	Fair		3.78
<i>Manes.13G043900</i>	Very good	Good		Fair	3.76
<i>Manes.06G164800</i>	Excellent	Good	Excellent	Good	3.75
<i>Manes.15G090000</i>	Good	Very good	Good	Good	3.71
<i>Manes.14G173500</i>		Very good	Fair	Good	3.66
<i>Manes.12G109200</i>		Good	Good	Good	3.56
<i>Manes.03G091000</i>	Good	Very good	Very good	Good	3.54
<i>Manes.12G027000</i>	Good		Fair		3.52
<i>Manes.03G043500</i>	Good	Fair			3.51
<i>Manes.06G094300</i>	Excellent	Fair	Good	Good	3.48
<i>Manes.13G031100</i>	Very good	Fair		Fair	3.45
<i>Manes.07G033400</i>	Excellent	Good	Fair	Good	3.42
<i>Manes.15G122900</i>	Good	Good	Fair	Good	3.38
<i>Manes.07G033000</i>	Excellent	Fair	Fair		3.36
<i>Manes.15G113300</i>		Very good	Good	Good	3.36
<i>Manes.14G152100</i>		Very good			3.33

<i>Manes.07G033300</i>	Excellent	Fair	Fair		3.24
<i>Manes.16G013800</i>	Good	Good	Excellent	Fair	3.13
<i>Manes.04G164400</i>	Excellent	Very good			3.12
<i>Manes.15G176200</i>		Good			3.11
<i>Manes.05G055600</i>		Very good			3.1
<i>Manes.15G134900</i>	Very good	Excellent	Fair	Fair	3.09
<i>Manes.15G013100</i>	Fair	Good	Fair	Excellent	3.04
<i>Manes.11G019100</i>	Very good	Fair	Good	Excellent	2.96
<i>Manes.11G007000</i>	Very good	Fair			2.78
<i>Manes.17G103900</i>	Fair	Very good			2.74
<i>Manes.17G013400</i>	Excellent	Good	Good	Good	2.73
<i>Manes.16G026000</i>	Very good	Good	Fair	Good	2.66
<i>Manes.02G075400</i>	Fair	Excellent	Good		2.61
<i>Manes.05G041000</i>	Fair	Fair		Fair	2.33
<i>Manes.01G043200</i>	Excellent	Very good			2.17

Chapter 3: CRISPR interference as a tool to study TALE biology in *Xanthomonas*

As seen in previous chapters, *Xpm* strains, and in general most of the *Xanthomonas* species, carry multiple *TALE*s copies (Schandry et al., 2018). *Xpm* *TALE*s are mainly plasmid-borne or integrated into the chromosome but deriving from plasmidic regions. Typically, the identity of the nucleotide stretches coding for the N-terminal and C-terminal regions of *TALE*s is high, which is in agreement with their prevalent evolutionary mechanisms (Erkes et al., 2017). Then, factors such as the presence of multiple paralogs, the occurrence of multiple copies due to their plasmid-borne nature, and the high degree of conservation, encumber directed *TALE* mutation strategies in this pathovar. Additionally, we have applied classical mutagenesis approaches (insertional and double homologous recombination-based mutation) to several strains of this pathovar, and results indicate that *TALE* mutants are difficult to obtain in some cases. This, and the fact that there are no reports of *Xpm* strains depleted of *TALE*s motivated us to find an alternative tool for *TALE* functional studies in *Xpm*.

The CRISPR/Cas system has revolutionized genome editing by increasing accuracy and simplicity. At a glance, this system uses a protein (CAS) that forms a complex with a short RNA, which will guide the complex to the target, and after positioning, will induce a double strand breaking (DSB) in the DNA. In organisms that possess the molecular machinery for the non-homologous end joining (NHEJ) repair system, mutations arise from DNA repairing errors (Kondrateva et al., 2021). Unfortunately, NHEJ systems are only sporadic in bacteria, and they are absent in *Xanthomonas* (Sharda et al., 2020). This means that CRISPR-induced DSB are deleterious for this genus. However, the use of a Cas9 variant that does not cut the DNA, a dead variant, can be useful to silence expression through interference with gene transcription (Qi et al., 2013). This kind of silencing platforms has been established for other bacteria, mainly human pathogens, but, to the best of our knowledge, it has never been tested in *Xanthomonas*.

In this chapter, we addressed the following research question: is the CRISPR interference (CRISPRi) platform implementation feasible and useful for the study of gene families such as *TALE*s, in the context of the plant-pathogen interaction? The expected positive answer to this question will provide our research field with a powerful tool to study single or multiple genes that share a certain homology degree, like gene families. We envisioned that such platform would simplify the study of the molecular pathogenicity determinants, and that it can be tuned with little effort to meet researchers needs.

We selected the *Xpm*-cassava pathosystem to design and implement this platform, since we have a good expertise in the biology of this model, and we were eager to have a tool to study the function of *TALE*s in this pathosystem (see chapter 2). After setting the platform, we expanded its application to other *Xanthomonas* pathogens, where *TALE*s are essential for pathogenicity. Our results demonstrate that the CRISPRi platform based on the *Streptococcus pyogenes*-derived dCas9 is functional in *Xanthomonas*, as it proved to silence up to 9 *TALE*s in one strain by using just one sgRNA. Moreover, we provided a system for gene complementation, allowing for WT-like levels of virulence when major virulence *TALE*s are expressed. Finally, the application of this platform to the poorly studied cassava non-vascular pathogen *Xanthomonas cassavae* resulted in the discovery that

TALEs are collectively essential for the virulence of this pathogen and that the susceptibility gene *MeSWEET10a* may be induced by one of them.

This work results of a concerted effort between a previous PhD student César Medina and post-doctoral researcher Jonathan Jacobs, today researchers, and several other collaborators. César and Jonathan set up the first tests and constructed the vectors for the system. They also demonstrated that a plasmid-borne copy of dCas9 was able to silence GFP when expressed together with an anti-*gfp* sgRNA in *Xpm*. The conception of sgRNAs to target *TALEs* in *Xpm* was guided by Ralf Koebnik and supervised by Adriana Bernal and Boris Szurek. All the experimental work reported here was carried out by myself, while there was a constant guidance and feedback from my tutors. Finally, the result of this work has allowed me to interact with other researchers in the field, and this will probably be followed by a collaborative publication of the tool. The target journal for this publication is *eLife*.

1 CRISPRi to study the function of gene families in plant pathogenic bacteria

2
3 Carlos Andrés Zárate-Chaves¹, César Medina², Emilie Thomas¹, Jonathan Jacobs^{3,4}, Camilo López⁵,
4 Ralf Koebnik¹, Adriana Bernal^{6†}, Boris Szurek^{1*†}

5 ¹ PHIM, Univ Montpellier, IRD, CIRAD, INRAe, Institut Agro, Montpellier, France

6 ² United States Department of Agriculture-Agricultural Research Service, Plant Germplasm
7 Introduction and Testing Research, Prosser, WA, USA.

8 ³ Department of Plant Pathology, The Ohio State University, Columbus, OH, USA

9 ⁴ Infectious Disease Institute, The Ohio State University, Columbus, OH, USA

10 ⁵ Manihot Biotec, Departamento de Biología, Universidad Nacional de Colombia, Bogotá, Colombia

11 ⁶ Laboratorio de interacciones moleculares en microorganismos de interés agronómico (LIMMA),
12 Universidad de los Andes, Bogotá, Colombia

13 * To whom correspondence should be addressed.

14 † These two authors contributed equally to this work.

16 **Abstract**

17 CRISPR technology has improved genome editing by dramatically increasing efficiency and
18 specificity. However, bacteria are not easily mutated using this technology. A technological variation
19 called CRISPRi employs a Cas9 modified version that is unable to cut DNA (dead Cas9 – dCas9) but
20 maintains sequence-specific DNA binding activity. dCas9 interferes with gene transcription when
21 targeted against the promoter or the first nucleotides of the non-template strand of the coding
22 sequence, thus resulting in gene silencing. This tool can be extremely useful for the functional
23 analysis of multigene families since a single guide RNA (sgRNA) may silence several members at
24 once. In this study, we set a CRISPRi platform for *Xanthomonas* and tested its potential to study a
25 gene family and pathogenesis determinants in this genus. We silenced the genes of the
26 Transcription Activator-Like Effector (TALE) family, which encode major virulence factors in several
27 *Xanthomonas* species. CRISPRi was implemented through the design of sgRNAs targeting the -10
28 and -35 promoter elements, the Ribosome Binding Site, and a region within the first 50 nucleotides
29 of the coding sequence of *TALEs* of several *Xanthomonas* species. Using *Xanthomonas phaseoli* pv.
30 *manihotis* (*Xpm*) as a case study, we showed that most if not all *TALEs* in a strain are silenced by one
31 individual sgRNA, resulting in decreased symptoms and *in planta* bacterial titers. Furthermore,
32 plasmid-borne *TALEs* lacking the sgRNA-targeted sequence can be expressed and used for
33 complementation. Implementation of this technological resource in *Xanthomonas cassavae*, a
34 pathogen of cassava, strongly suggest that water soaking symptoms are linked to the TALE-
35 dependent activation of the sugar transporter encoding-gene *MeSWEET10a*, a susceptibility gene
36 exploited by the distantly-related cassava pathogen *Xpm*. Our study provides a useful CRISPRi
37 platform to assess the role of *TALEs* during host colonization, and overall, for the functional analysis
38 of multigene families in bacteria.

39 **Introduction**

40 The CRISPR–Cas (Clustered Regularly Interspaced Short Palindromic Repeats–CRISPR associated
41 protein) system is an adaptative immunity mechanism present in several bacteria and archaea that
42 protects cells from phage attack through nucleic acid restriction (Barrangou et al., 2007). In general
43 terms, a small RNA transcribed from the CRISPR locus (crRNA) interacts with a transactivating CRISPR

44 RNA (tracrRNA) which serves as a binding scaffold, and the Cas endonuclease to form an active
45 complex. The Cas nuclease on the complex scans DNA for a short nucleotide motif called
46 protospacer-adjacent motif (PAM), which is positioned between two protein domains present in the
47 nuclease lobe. This interaction unfolds the target DNA and the crRNA in the complex is exposed to
48 the test its complementarity. Once the Cas nuclease interacts with the correct PAM, and the crRNA
49 hybridizes with the target sequence, the enzyme cleaves the target nucleic acid (Brouns et al., 2008;
50 Gasiunas et al., 2012). CRISPR-Cas systems are classified into two classes, 6 types and 33 subtypes
51 (Makarova et al., 2020). Class 1 systems use multiprotein complexes for nucleic acid restriction,
52 while class 2 uses a single-protein nuclease. Class 2 CRISPR-Cas has gained attention for
53 biotechnological applications due to its simplicity, and type-II Cas9 (specially the one from
54 *Streptococcus pyogenes*) variants and type-V Cas12 variants, which code for RNA-guided DNA
55 endonucleases, are widely used (Anzalone et al., 2020). Besides the successful implementation of
56 this technology for genome editing in eukaryotes and some prokaryotes, modification, or addition
57 of protein domains to the Cas9 nuclease has increased the application range for these systems.
58 Fusion to deaminases, transcriptional repressors/activators, histone modification enzymes, and
59 fluorescent proteins allows base editing, gene expression control, epigenetic regulation and
60 imaging, among others (Liu et al., 2021).

61 Until recently, CRISPR-based genome editing in eukaryotes was achieved through double-strand
62 break (DSB) of the target DNA, which was then repaired by the host. When a homologous sequence
63 with the desired DNA alteration was provided, the homology-directed repair (HDR) pathway
64 resulted in recombination with the altered DNA. In the absence of a donor homologous template,
65 the non-homologous end joining (NHEJ) DNA repair mechanism, which is error-prone, resulted in
66 insertions or deletions (indels) in the repaired position. However, new technological variants allow
67 for more accurate genome editing approaches involving or not these repair mechanisms (Anzalone
68 et al., 2020). Most bacteria lack NHEJ systems, and Cas9 activity results in lethality due to unresolved
69 chromosome breaks (Cui and Bikard, 2016). The use of catalytically impaired variants of the Cas9
70 (dCas9 and nCas9) protein allowed the development of technologies for DSB-free gene editing and
71 gene silencing suitable for bacteria. nCas9, a catalytically impaired variant of Cas9 inducing nick
72 formation, can be fused to deaminase domains to induce cytosine-to-thymine or adenine-to-
73 guanine substitutions (Gaudelli et al., 2017), or be fused to a reverse transcriptase to create
74 template-based insertions *in situ* (Anzalone et al., 2019). The dCas9, a catalytically dead variant of
75 dCAS9, *per se* preserves its PAM specificity and the complex formed with the single guide RNA
76 (sgRNA, a construct where the crRNA and the tracrRNA are fused through a short nucleotide loop
77 for biotechnological purposes) can bind specifically to the target DNA, leading to sgRNA-guided
78 protein complex position. This interaction can interfere with transcription initiation or elongation,
79 resulting in the silencing of the target gene in a process known as CRISPR interference (CRISPRi)
80 (Bikard et al., 2013; Qi et al., 2013). CRISPRi disturbs transcription by physically interfering with RNA
81 polymerase and transcription factors. When the sgRNA targets any of the strands of the promoter
82 region, the dCAS9/sgRNA complex sterically prevents the association of the trans-acting
83 transcription factors and their corresponding DNA elements. When the sgRNA targets the non-
84 template strand of the coding region located downstream the promoter, RNA polymerase
85 progression is blocked by the association of the dCAS9/sgRNA complex and the DNA.

86 Gene families are groups of genes that derive from a common ancestor, and which generally have
87 similar biochemical functions. According to the Innovation–Amplification–Divergence (IAD) model
88 (Bergthorsson et al., 2007), when a gene evolves side functions that can be advantageous for fitness,
89 a selective pressure acts on gene duplicates (usually created by segmental duplication in bacteria),
90 allowing divergency of one of the copies to improve the new function and resulting in two members
91 of the same gene family. Horizontal gene transfer mechanisms also contribute to paralog and
92 homolog acquisition (Copley, 2020). Bacterial genomes contain thousands of gene families mainly
93 composed of one or two genes, but some gene families can have hundreds of members in one
94 genome, like transcriptional regulators and ABC transporters (Collins et al., 2011).

95 Transcription Activator-Like Effectors (TALEs) are virulence factors present in most of the
96 *Xanthomonas* species (Schandry et al., 2018), which are delivered to plant host cytoplasm through
97 the type three secretion system. The particular structure of these effectors allows them to
98 selectively bind host DNA promoters to induce transcription of downstream genes (Boch and Bonas,
99 2010). TALE proteins are composed by an N-terminal region containing the signal for T3SS-mediated
100 translocation (Szurek et al., 2002), while the C-terminal region holds two to three nuclear
101 localization signals (NLSs) (Yang and Gabriel, 1995) and an acidic transcriptional activation domain
102 (AAD) responsible for RNA polymerase recruitment (Zhu et al., 1998). The central region is
103 characterized by the arrangement of a variable number of tandem modular repeats of 33-35 amino
104 acids (Deng et al., 2014; Schandry et al., 2018). Each repeat has a two-residue hotspot at positions
105 12 and 13 referred to as repeat variable diresidue (RVD), which determines the binding affinity of
106 the repeat for a DNA base in the target sequence (Boch et al., 2009; Moscou and Bogdanove, 2009).
107 The central region wraps around the DNA in a right-handed superhelix (Deng et al., 2012; Mak et
108 al., 2013, 2012), and the RVD sequence of the repeats governs the target DNA recognition. Usually,
109 this target sequence known as the Effector Binding Element (EBE) (Antony et al., 2010), is located in
110 the vicinity of the transcription start site and often close the TATA-box. Upon DNA-binding, TALEs
111 interact with host transcription factors to induce polymerase II–dependent transcription of the
112 downstream gene (Yuan et al., 2016). The number of *TALE* genes in a single genome (TALome) can
113 vary from 0 to 29, depending on the pathovar and strain. Most pathovars of *Xanthomonas* with
114 TALEs harbor TALomes with 1 to 5 paralogs, others like *Xanthomonas oryzae* pv. *leersiae*,
115 *Xanthomonas citri* pv. *malvacearum* (*Xcm*), and *Xanthomonas citri* pv. *mangiferaeindicae* contain
116 10-12 paralogs, while *Xanthomonas oryzae* displays up to 19 and 29 *TALEs* in their genomes in the
117 case of *oryzae* (*Xoo*) and *oryzicola* (*Xoc*) pathovars, respectively (Schandry et al., 2018). With the
118 exception of *Xoo* and *Xoc* where all *TALEs* are chromosomal (Erkes et al., 2017), *TALEs* are often
119 present in plasmids or plasmid-borne chromosomal insertions (Bart et al., 2012; Cox et al., 2017;
120 Denancé et al., 2018; Ruh et al., 2017), where transposition and horizontal gene transfer play key
121 roles in *TALE* distribution and evolution (Denancé et al., 2018; Ferreira et al., 2015).

122 The *TALE* gene family plays important roles in the virulence and avirulence of many *Xanthomonas*
123 species by transcriptional activation of host genes that can confer advantages to the pathogen,
124 known as susceptibility (*S*) genes, or genes that act like molecular traps and induce cell death, known
125 as effector (*E*) genes. Most of the *S* genes discovered so far can be grouped into two main functional
126 categories, transcriptional master regulators and nutrient secretion systems (reviewed by (Xue et
127 al., 2021)). Disruption of major virulence *TALEs*, i.e., that activate *S* genes, results in a dramatic
128 reduction of pathogen aggressiveness and/or virulence, as reported in several pathosystems such

129 as *Xoc*- and *Xoo*-rice, *Xcm*-cotton, *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*)-cassava,
130 *Xanthomonas citri* pv. *citri* (*Xcc*)-citrus, *Xanthomonas gardneri* (*Xg*)-tomato, *Xanthomonas*
131 *campestris* pv. *vesicatoria* (*Xcv*)-pepper, *Xanthomonas translucens* pv. *undulosa* (*Xtu*)-wheat, and
132 *Ralstonia solanacearum*-tomato (Perez-Quintero and Szurek, 2019; Xue et al., 2021). Based on our
133 knowledge about TALE function, it has been possible to breed naturally occurring or engineered
134 loss-of-susceptibility alleles leading to plants with highly improved pathogen resistance (Oliva et al.,
135 2019; Peng et al., 2017). Nonetheless our knowledge about the role(s) of *TALE*s in *Xanthomonas*
136 virulence is poor, taking in account the high number of species or pathovars where they are
137 conserved and for which no data is available. This is in part due to the difficulty of performing
138 targeted gene deletion since, as a gene family, *TALE*s maintain a high degree of similarity at the
139 nucleotide level, they are clustered and/or they are located in multi-copy plasmids.

140 In this study, we aimed at validating a CRISPRi technology platform allowing for the functional
141 analysis of gene families in *Xanthomonas*. We used *TALE*s as a model gene family to demonstrate
142 the feasibility and relevance of knocking-down a full gene family in the context of a plant-pathogen
143 interaction. We first show that sgRNAs targeting regulatory functional elements conserved in the
144 promoter and 5'-UTR regions of *TALE*s of the cassava pathogen *Xpm* successfully silenced *TALE*
145 expression. We then demonstrate that complementation with a designer major virulence *TALE*
146 expressed from a sgRNA-protected plasmid-borne construct allows to restore WT-like molecular
147 and physiological phenotypes. To highlight its applicability in other pathosystems, the tool was next
148 applied to strains of *Xoo*, *X. citri*, and *X. cassavae* (*Xc*), and full *TALE* repertoires knockdown was
149 evidenced as shown upon Western-blot analysis. Pathogenicity assays with *Xoo* and *Xc* confirmed
150 the crucial role of the *TALE* gene family in *Xoo* and evidenced its role in *Xc* virulence on cassava.
151 Overall, we demonstrate that the *Streptococcus pyogenes*-derived dCas9 is a praiseful strategy for
152 gene(s) knockdown in *Xanthomonas* plant-pathogenic bacteria.

153 **Results**

154 **A silencing platform based on *S. pyogenes*-derived dCas9 adapted to *Xpm***

155 Seventy-nine *TALE* sequences (including the promoter region) obtained by SMRT long-read
156 sequencing from 20 *Xpm* world-wide isolates (Zárate-Chaves et al., *In prep*) were aligned to guide
157 the sgRNA design process. The consensus sequence for the promoter and first 60 nucleotides of the
158 coding sequence of *Xpm TALE*s were scanned for potential X(20)NGG sgRNAs (see Supplementary
159 Figure S1). Three were selected that were directed against the non-template strand: one is targeting
160 the region comprising the -10 and -35 elements (from here on -10-35), a second one is targeting the
161 ribosome binding site (from here on RBS), and a third one is directed against the region spanning
162 nucleotides +31 to +50 of the coding sequence (from here on CDS). Table 1 shows the characteristics
163 of the selected sgRNAs, highlighting that none of them are predicted to have off-targets (Hsu et al.,
164 2013). A construct comprising the sgRNA, dCas9 and Gentamycin resistance expression cassette was
165 cloned into a transposable Tn7 element (Choi and Schweizer, 2006) for targeted transposition within
166 the intergenic region located downstream of the *glmS* gene of *Xpm*. Figure 1 describes the
167 constructs used for the silencing and complementation strategy. Two control constructs were
168 created: one carrying only the resistance marker and the *dCas9* gene (from here on dCas9), and a
169 second one carrying the resistance marker, *dCas9*, and a promoterless *GUS* gene instead of a sgRNA

170 scaffold (from here on GUS). None of these controls is expected to induce silencing, since they lack
171 sgRNA expression.

172 **Silencing of *Xpm* TALE genes results in decreased pathogen virulence**

173 Since strain Xam668 was previously used to characterize TALE function through insertional
174 mutagenesis (Cohn et al., 2014), we selected this strain to test the silencing platform. After
175 transposition of the CRISPRi units and derivative controls into the Xam668 genome, the resulting
176 transformants were first tested by Western-blot analysis using an anti-TALE antibody to determine
177 TALE protein levels (Figure 2A). As expected, WT strain Xam668 showed five bands corresponding
178 to TAL13_{Xam668}, TAL14_{Xam668}, TAL15_{Xam668}, TAL20_{Xam668}, and TAL22_{Xam668} similarly to what was observed
179 with the negative controls (dCas9 alone, and GUS). Interestingly Xam668 derivative strains
180 expressing sgRNAs directed against -10-35 and RBS targets generally accumulated less TALE
181 proteins, except for TAL13. In contrast, no significant changes could be evidenced when derivative
182 strains expressing the anti-CDS sgRNA were analyzed, suggesting that interfering with transcription
183 initiation is more effective than with elongation in this assay. Strains were next tested for
184 pathogenicity *in planta*. In contrast to WT Xam668, CRISPRi strains carrying -10-35 and RBS sgRNAs
185 were unable to induce water-soaking symptoms at 96 hours post inoculation (hpi) of the susceptible
186 cassava cultivar 60444, similarly to what was observed upon inoculation of the strain
187 Xam668ΔTAL20 which is mutated in a *MeSWEET10a*-inducing major virulence TALE (Cohn et al.,
188 2014). As expected, negative dCAS9 and GUS controls and the strain carrying the CDS sgRNA also
189 showed normal symptom formation. Together, these data indicate that interfering with *TALE*
190 transcription initiation is an effective strategy to knock-down several TALEs with a single construct
191 in *Xpm*, thus resulting in strong virulence decrease.

192 **An arTALE protected against sgRNAs and inducing the CBB susceptibility gene *MeSWEET10a*** 193 **restores virulence**

194 As *MeSWEET10a* TALE-mediated activation is a key feature of *Xpm* pathogenesis that results in
195 water soaking symptom formation and higher bacterial titers in cassava (Cohn et al., 2014), we used
196 an artificial TALE (arTALE) capable of activating *MeSWEET10a* (Cohn et al., 2014). Unlike natural *Xpm*
197 TALEs, the N-terminal-coding sequence (from *Hax3*) and the lac promoter (Geißler et al., 2011) of
198 this arTALE cannot be targeted by any of the tested sgRNAs. The set of Xam668 WT and CRISPRi
199 derivative strains were transformed with the pME6010-plac-arTALE or the empty vector (EV) and
200 tested for *TALE* expression, the capacity to form water soaking symptoms, *in planta* bacterial
201 growth, and the transcriptional induction of *MeSWEET10a*. As expected, the arTALE is
202 homogeneously expressed in the context of the CRISPRi strains (Figure 3A), confirming that it is
203 protected against silencing. In consequence, pathogenicity assays show that arTALE expression
204 (Figure 3B) restores virulence, as judged by the resurgence of water soaking symptoms. Figure 3C
205 shows the bacterial titers of some of the transformed CRISPRi strains and the WT at 7 days post
206 inoculation (dpi) in the cassava variety 60444. Bacterial titers for -10-35 and RBS strains are
207 significantly lower ($p < 0.05$) than those for the WT, the controls, and the CDS CRISPRi strain. In line
208 with virulence restoration, arTALE expression improved the bacterial titers for the -10-35 and RBS
209 CRISPRi strains, reaching titers statistically equivalent to the ones observed for the rest of the tested
210 strains. *MeSWEET10a* activation patterns (Figure 3D) indicate that TAL20_{Xam668} silencing is correlated
211 with the loss of the *MeSWEET10a*-upregulating capacity of Xam668, while arTALE expression

212 restores the induction of this *S* gene. Overall, results indicate that expressing a major *TALE* from a
213 non-sgRNA-targeted plasmid restores most of the aggressiveness and virulence defects that
214 resulted from *TALE* silencing. Interestingly, silencing most of the *Xpm* *TALEs* did not prevent bacteria
215 to reach significant titers *in planta*. Altogether, these data evidence that the silencing platform
216 implemented in *Xpm* is useful for the study of gene families.

217 **The *TALE* silencing platform can be used in other *Xanthomonas* species**

218 In order to test if the silencing platform that we developed for *Xpm* *TALEs* can be applied to other
219 *Xanthomonas* species, we aimed at knocking-down the *TALE* repertoires of *X. oryzae* pv. *oryzae*, *X.*
220 *citri* and *X. cassavae* that, respectively, cause bacterial leaf blight (BLB) of rice, Asian citrus canker,
221 and Cassava Bacterial Necrosis (CBN). To that end we selected the reference *Xoo* strains PXO99^A and
222 MAI1 (two representatives of the Asian and African lineages respectively), *Xcc* IAPAR306 and *Xc*
223 CFBP4642. Alignments of the upstream region of the *TALE* genes of each strain with the designed
224 sgRNAs showed that 35 out of 38 *TALEs* investigated were predicted to be targeted by the RBS
225 sgRNA (Table 2 and Supplementary Figure S2). Interestingly the whole TALomes of *Xoo* MAI1, *Xcc*
226 IAPAR306, and *Xc* CFBP4642 are predicted to be silenced, while 3 out of 19 *TALEs* in the case of
227 PXO99^A would be protected because of one mismatch in the seed sequence (Table 2). Figure 4 shows
228 the results for the Western Blot analyses to test for the reduction of protein accumulation in the
229 strains transformed with the CRISPRi construct with the *TALE* RBS sgRNA. For strains *Xoo* MAI1, *Xcc*
230 IAPAR306, and *Xc* CFBP4642, a reduction in protein expression confirms the knockdown of *TALE*
231 genes, as predicted by the bioinformatic analyses. Unfortunately, Western blot results obtained so
232 far for strain PXO99^A are not optimal (data not shown) and prevented us from drawing conclusions.

233 As silencing of the MAI1 major virulence *TALEs* *talC* and *talF* was confirmed and *pthXo1* silencing in
234 PXO99^A is predicted, water soaking symptoms and lesion length were measured as an indication of
235 pathogen virulence in rice. Figure 4B shows that lesion lengths caused by the *Xoo* MAI1 strain
236 transformed with the CRISPRi unit containing the RBS sgRNA significantly decreased, when
237 compared to those caused upon inoculation of the WT and GUS control strains. In agreement with
238 this, water soaking symptoms at 96 hpi were significantly reduced in this CRISPRi strain when
239 compared to the WT and the GUS control (Figure 4C). In the case of PXO99^A, reduction in lesion
240 lengths was also statistically significant, but less dramatic than MAI1. Taken together, the silencing
241 platform is suitable for other *Xanthomonas* species, although some pathogens would need
242 adjustment of target specificity.

243 **Silencing the *Xc* TALome reveals its major role in pathogen virulence**

244 Information about *Xc* pathogenesis is scarce, and molecular interaction data between the cassava
245 host and this pathogen is, to our best knowledge, inexistent. Although the draft sequence of the
246 reference strain CFBP4642 showed that *Xc* contains *TALEs* (Bolot et al., 2013), nothing is known
247 about their involvement in pathogenesis. Recently, Zárte-Chaves and coworkers resequenced the
248 genome of this strain through long-read SMRT sequencing technology, leading to the obtention of
249 *Xc* CFBP4642 TALome (Zárte-Chaves et al., *In prep*). Transformation of the CRISPRi unit containing
250 the *TALE* RBS sgRNA in strain CFBP4642 strikingly resulted in silencing of the whole TALome (Figure
251 4A). Surprisingly, although sequencing revealed the existence of 6 *TALEs* in *Xc* (and a potential
252 pseudogenized *TALE*, see Supplementary Figure S2), there are only four bands in the corresponding
253 molecular weight range. CRISPRi *Xc* strains infiltration in a panel of three cassava varieties (Figure

254 5A) indicates that water soaking symptom development is compromised in the RBS CRISPRi strain,
255 but not in the GUS control strain. This finding prompted us to search for the cassava genes
256 targeted by *Xc* TALEs using the program Talvez that allows to identify possible EBEs in plant
257 promoter collections (promoteromes) (Pérez-Quintero et al., 2013) (Supplementary Table S1).
258 Surprisingly, TALE23 from *Xc* CFBP4642 is predicted to target *MeSWEET10a*, which encodes for a
259 sugar transporter and acts as a susceptibility gene during the infection of cassava by *Xpm*.
260 *MeSWEET10a* is induced by TAL20_{Xam668} upon binding to a 20-nt EBE that overlaps with the one
261 predicted to be recognized by TALE23 (Figure 5B). In conclusion, the use of the silencing platform in
262 *Xc* shows that TALEs are crucial for water soaking during Cassava Bacterial Necrosis. Moreover, the
263 search for cassava genes targeted by these TALEs indicates that the sugar transporter *MeSWEET10a*
264 may be involved in symptom formation.

265 **Discussion and conclusions**

266 A first generation of CRISPR/Cas tools allowed the precise gene mutations in eukaryotes by inducing
267 DSB and taking advantage either of natural repair of errors or the recombination machinery present
268 in cells. Since NHEJ mechanisms are sporadically distributed among bacterial genomes (Sharda et
269 al., 2020), the suitability of this tool for precise genome editing is very limited. However, the CRISPRi,
270 where the targeted gene is silenced by a catalytically dead Cas9 variant, has been successfully used
271 in bacteria for engineering biotechnological desirable characteristics, identifying drug targets, or
272 studying the role of genes and pathways (reviewed by (Ding et al., 2020; Todor et al., 2021)). CRISPRi
273 platforms have mainly been established for human and animal pathogens like *Vibrio cholerae*,
274 *Yersinia pestis*, *Leptospira* spp., *Burkholderia* spp., among others (Afonina et al., 2020; Caro et al.,
275 2019; Ellis et al., 2021; Fernandes et al., 2019; Hogan et al., 2019; Lai et al., 2020; Takacs et al., 2021;
276 Tan et al., 2018; Wang et al., 2019). Only a few examples of CRISPRi application to plant-associated
277 microorganisms are reported in literature: *P. fluorescens* (Tan et al., 2018) and *Paenibacillus sonchi*
278 (Brito et al., 2020). Here, we demonstrate not only the usefulness of this platform in *Xanthomonas*
279 species, without dCAS9-associated toxicity and with appropriate silencing rates, but also the
280 feasibility of targeting several genes sharing the sgRNA-targeted sequence. The involvement of
281 TALEs as major virulence factors was corroborated in four different pathosystems (rice-*Xoo* MAI1,
282 rice-*Xoo* PXO99^A, cassava-*Xpm* Xam668 and cassava-*Xc* CFBP4642) and we showed, for the first time,
283 the potential contribution of TALE-mediated *MeSWEET10a* activation within *Xc* pathogenesis.

284 As previously reported, blocking transcription elongation is position-dependent, being more
285 stringent when the target sequence is closer to the start codon (Qi et al., 2013). Our results show
286 that *TALE* silencing by CRISPRi in the -10-35 and RBS strains is sufficient to significantly reduce the
287 *TALE*-mediated induction of *MeSWEET10a*, with the concomitant reduction of water soaking
288 symptoms in the plant. On the other hand, targeting positions +31 to +50 of the CDS results in
289 uneven silencing of targets, which does not significantly affect virulence of the bacterium. However,
290 as we do not yet possess the nucleotide sequence of Xam668 TALEs, it cannot be ruled out that, as
291 other *Xpm* TALEs (see Supplementary Figure S1), unsilenced Xam668 TALEs would carry SNPs
292 preventing sgRNA binding with this CDS region.

293 The suitability of this CRISPRi platform was tested with non-optimized sgRNAs (since they were
294 designed for *Xpm*) in other important *Xanthomonas* pathogens: *Xc*, the causal agent of CBN, *Xcc*,
295 the causal agent of Asian citrus canker, and *Xoo*, the causal agent of BLB of rice. Results for *Xc*

296 CFBP4642, *Xcc* IAPAR306, and *Xoo* MAI1, pathogens with a relatively low number of *TALE* genes (7,
297 4, and 9 respectively) confirm that a single sgRNA is enough for multi-*TALE* knock down. Silencing of
298 the 19 *TALE* paralogs of the Asian *Xoo* PXO99^A strain could not be assessed so far through
299 quantitative Western blots, but virulence assays with CRISPRi strains are encouraging (see Figure 4).
300 We foresee that the silencing platform is useful but requires strain- or pathovar-specific
301 optimization (discussed below) in some cases. Interestingly, the implementation of the silencing
302 platform in *Xc* CFBP4642 strain showed that *Xc* TALEs may play a key role for water soaking symptom
303 formation. It is hypothesized that the water soaked lesions enhance bacterial fitness by facilitating
304 the entrance and movement of *Xanthomonas* to the apoplast, and a few TALEs from various
305 *Xanthomonas* species are involved in the development of this type of symptom (Cohn et al., 2014;
306 Cox et al., 2017; Schwartz et al., 2017; Streubel et al., 2013). TALE-mediated *SWEET* gene activation
307 during a compatible interaction results in water soaking development in cassava (Cohn et al., 2014),
308 rice (Streubel et al., 2013) and cotton (Cox et al., 2017), possibly through osmotic changes caused
309 by sugar accumulation in the apoplast and the increase of hygroscopic exopolysaccharides
310 associated to the enhanced pathogen growth. Alternatively, in tomato leaves, the transcriptional
311 activation of a pectate lyase (an indirect target of the *X. gardneri* TALE AvrHah1) is responsible for
312 the water soaking symptom formation, hypothetically by increasing the hygroscopicity of the cell
313 wall (Schwartz et al., 2017). Although phenotyping and *in silico* predictions indicate that *Xc* activates
314 *MeSWEET10a* resulting in water soaking lesions, more work is needed to confirm the role of TALEs
315 in *Xc* pathogenesis. Confirmation of this activation would represent a case of convergent evolution
316 between two distantly-related pathogens, with different lifestyles (vascular vs non-vascular) that
317 exploit the same susceptibility factor.

318 The silencing platform presented in this study allowed the efficient silencing of up to 9 genes from
319 the TALE family with one sgRNA. Impacts of this accomplishment are highly significant, since the
320 obtention of an equivalent mutant through DHR methodologies takes considerable effort, time, and
321 several *in-vitro* passages that could attenuate pathogenicity. In this regard, Ji and coworkers deleted
322 the nine TALE clusters present in the *Xoo* strain PXO99^A through eight rounds of sequential
323 mutation, by using differential flanking sequences for DHR of clusters containing several *TALEs*. This
324 strain (named PH) was used to test TALE-mediated avirulence (HR screenings) and intermediate
325 mutants led to the discovery of iTALEs (Ji et al., 2016). Likewise, a series of three studies spanning
326 four years of research led to the obtention of a *Pseudomonas syringae* DC3000 mutant
327 (DC3000D28E) where the 28 type-three effectors (T3Es) were successively deleted. These mutants
328 allowed to discover the substantial contribution of T3Es like HopQ1-1 and AvrPtoB in the *Ps* DC3000
329 pathogenesis in *Nicotiana benthamiana* (Cunnac et al., 2011; Kvitko et al., 2009; Wei et al., 2007).
330 Implementation of CRISPRi with multiplexed sgRNAs has been reported for up to 10 different targets
331 at the same time through the use of an arrangement of crRNAs and spacers controlled by a unique
332 strong promoter carrying an enhancing element (Ellis et al., 2021). For the T3Es in *Xanthomonas*,
333 *Burkholderia*, and *Ralstonia*, it would be also interesting to target Plant-Inducible Promoter (PIP)
334 boxes. PIP boxes are cis-regulatory elements that allow transcriptional induction after HrpX binding
335 (Cunnac et al., 2004; Fenselau and Bonas, 1995). *hrpX* and *hrpG* constitute a two-component
336 regulatory system that controls the expression of pathogenicity genes *in planta* (Fenselau and
337 Bonas, 1995; Koebnik et al., 2006; Wengelnik et al., 1996). Targeting PIP boxes would theoretically
338 interfere with HrpG DNA binding and transcriptional activity, potentially blocking upregulation of
339 downstream T3Es.

340 dCAS9 and sgRNA expression patterns can be controlled by stronger and inducible promoters, which
341 is a key factor when silencing, for example, genes that control cell growth (Caro et al., 2019; Tan et
342 al., 2018). Some authors have used the tetracycline-inducible tet promoter to control dCAS9
343 expression, allowing fine-tuning of gene silencing in a tetracycline dose-dependent manner (Caro et
344 al., 2019; Ellis et al., 2021), or more sophisticated systems like the nisA promoter that is activated
345 upon the nisin recognition by the NisKR two-component system included in the CRISPRi silencing
346 platform (Afonina et al., 2020). Nevertheless, the optimization of the system should be in
347 accordance with the experimental set and goals, and the possibilities to modulate the expression of
348 the system *in vivo*. For example, for *Xanthomonas*, *Burkholderia*, and *Ralstonia* species, PIP boxes
349 could be used to control *dCAS9* expression, ensuring that silencing takes place only during host
350 infection.

351 One of the strains presented in this study, the strain expressing the dCAS9, could also be used to
352 easily screen sgRNAs expressed from replicative plasmids, without the need of chromosomal
353 insertion. This system has proven to be useful in *Burkholderia*, where a chromosomally-inserted
354 *dCAS9* is under the control of the *E. coli* rhamnose-inducible promoter while sgRNAs are expressed
355 from the pBBR1-derived pSCrhaB2 plasmid (Hogan et al., 2019). While this approach is useful to
356 check the efficient silencing of the sgRNA-targeted gene *in vitro*, the inclusion of the sgRNA scaffold
357 into the chromosomal CRISPR unit ensures sgRNA expression *in planta*, even if its expression results
358 in a loss of fitness for the pathogen in the host.

359 In conclusion, our study demonstrates the usefulness of the CRISPRi platform to study of large gene
360 families in the plant pathogenic bacterium *Xanthomonas*, and its strong potential to dissect the
361 function of genes involved in bacterial pathogenesis. Hopefully, further optimization and adaptation
362 of the platform (e.g., *dCAS9* codon optimization, tighter control of *dCAS9* expression, adjustment of
363 gene dosing, sgRNA multiplexing, etc.), along with other dCAS9 based technologies like base editors
364 (Banno et al., 2018), will increase the technology adoption and widen the array of molecular tools
365 for the study of *Xanthomonas*.

366 **Materials and Methods**

367 **Bacterial strains, media, and culture conditions**

368 The bacterial strains used in this study were *Escherichia coli* DH10 β , *X. phaseoli* pv. *manihotis* (*Xpm*)
369 strain Xam668, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strains MAI1 and PXO99^A, *Xanthomonas citri*
370 pv. *citri* (*Xcc*) strain IAPAR306, and *Xanthomonas cassavae* (*Xc*) strain CFBP4642. Wild-type *Xpm*
371 strains were streaked on YPG (yeast extract 5 gL⁻¹, peptone 5 gL⁻¹, glucose 5 gL⁻¹, agar-agar 15 gL⁻¹)
372 solid media, while transformants were streaked on YPG supplemented with gentamycin and/or
373 tetracycline and incubated at 28°C for 48 hours. When needed, they were grown in Phi broth (yeast
374 extract 1 gL⁻¹, peptone 10 gL⁻¹, casaminoacids 1 gL⁻¹), supplemented with gentamycin or tetracycline
375 as appropriate, at 28°C, under constant shaking at 220 rpm overnight. *Xoo* strains MAI1 and PXO99^A,
376 *Xcc* strain IAPAR306, and *Xc* strain CFBP4642 were grown in PSA (peptone 10 gL⁻¹, sucrose 10 gL⁻¹,
377 glutamic acid 1 gL⁻¹, and agar 15 gL⁻¹, pH 7.5) and the derived transformants on PSA supplemented
378 with gentamycin. When needed, they were grown in PSA broth (peptone 10 gL⁻¹, sucrose 10 gL⁻¹,
379 glutamic acid 1 gL⁻¹, pH 7.5) at 28°C, under constant shaking at 22 rpm overnight.

380 **Design of CRISPRi sgRNAs**

381 Sequences of the promoter and N-terminal-coding region (first 870 nucleotides) of *TALEs* from *Xpm*
382 were aligned with Muscle (Edgar, 2004). Putative promoter -10 element was predicted using the
383 Virtual Footprint software version 3.0 (Münch et al., 2005) with the position weight matrix for Sig70
384 (-10) from *Escherichia coli*. Putative promoter -35 element was located 14±2 bp upstream from the
385 putative -10 element trying to match the canonical *E. coli* TTGACA motif. The ribosome-binding site
386 (RBS) was located in the 5'-UTR region by identifying the *Xanthomonas* canonical pattern GGAG
387 eight nucleotides upstream of the start codon (Barak et al., 2016). The consensus sequence was
388 scanned for potential PAMs on the non-template strand by using the Geneious suite for CRISPR with
389 the model proposed by Doench and coworkers for in-target activity (Doench et al., 2014) and the
390 off-target scoring method proposed by Hsu and coworkers (Hsu et al., 2013). Three potential crRNAs
391 were selected based on the targeted regions, a high score for in-target activity, and the absence of
392 significant off-targets in the *Xpm* genome. Nucleotides were added to each crRNA to obtain a 5'
393 overhang AAC and a 3' overhang GTTTT in order to insert them into the CRISPR scaffold using *BsaI*.
394 Designed crRNAs were synthesized as single-stranded oligonucleotides (Table 1). Alignment of
395 *TALEs* from *Xoo* MAI1, *Xoo* PXO99^A, *Xcc* IAPAR306, and *Xc* CFBP4642 was performed as described
396 earlier in this paragraph for *Xpm*, and crRNA guides were then mapped to alignments to detect
397 mismatches. Criteria to predict the effectiveness of the targeted sequence were adapted from the
398 ones proposed by Hsu and coworkers (Hsu et al., 2013): the PAM should be adequate in sequence
399 and position, the global sequence similarity between the sgRNA and the target sequence should
400 have fewer than three mismatches, these mismatches should not be consecutive, and there should
401 be no more than two mismatches within the 12 nucleotides of the PAM-proximal region.

402 **CRISPRi plasmid construction and chromosomal insertion in *Xpm***

403 Both strands of the crRNAs were synthesized as single-stranded oligonucleotides, then 100 µM were
404 phosphorylated with the T4 Polynucleotide Kinase (NEB, USA) according to manufacturer's
405 instructions. NaCl (Sigma Aldrich, Germany) was added to a final concentration of 0.05 M and
406 phosphorylated sgRNAs were heated for 5 minutes at 95°C and allowed to cool down during 12
407 hours for annealing. The resulting solution was diluted 10 times. pENTR/D TOPO pgRNA *BsaI*:*BsaI*
408 (Supplementary Figure S3) is a vector containing the J23119 promoter, the Cas9-handle and a
409 *Streptococcus pyogenes* tracr terminator, which allows directional cloning of the crRNAs between
410 the promoter and the handle by using *BsaI* to produce the sgRNA scaffold. pENTR/D TOPO pgRNA
411 *BsaI*:*BsaI* was digested with *BsaI* HF (NEB, USA) and each phosphorylated double-stranded sgRNA
412 was ligated using the T4 DNA ligase (NEB, USA), according to the manufacturer's instructions. Each
413 ligation product was transformed by heat shock in competent DH10β *E. coli* cells. Screening was
414 performed by PCR on boiled colonies, using one of the oligonucleotides synthesized for each sgRNA
415 and the M13-Rv primer (Table 3). PCRs were performed in a 20-µL reaction mix containing 1X of 5X
416 GoTaq Green Buffer, 200 µM of a dNTP mix, 0.3 µM of each primer, 0.05 U/µL of GoTaq G2
417 polymerase (Promega, USA), and 1 µL of the bacterial boiling. PCR-positive colonies were grown in
418 LB broth supplemented with kanamycin to prep the plasmid DNA with the Monarch® Plasmid
419 Miniprep Kit (NEB, USA) following manufacturer's instructions. Clones were sequenced by Sanger.

420 pUC18-mini-Tn7T-Gm-Gw-dCAS9 is a non-replicative vector for *Xanthomonas* containing the border
421 sequences recognized by the Tn7 transposon, which will provide the cassette to be transposed to
422 the bacterial chromosome (Supplementary Figure S4). Between these borders, the plasmid contains
423 the *aacC1* gene, the *dCAS9* gene under the lac promoter control, and the *cat* and *ccdB* genes flanked

424 by attR1 and attR2 sequences. Each sgRNA scaffold previously cloned into pENTR/D TOPO pgRNA
425 BsaI:BsaI was transferred to the pUC18-mini-Tn7T-Gm-Gw-dCAS9 through an LR reaction using the
426 Gateway™ LR Clonase™ II Enzyme mix (Thermo Fischer Scientific, USA), according to manufacturer's
427 instructions. The pENTR-GUS control provided by Thermo Fischer Scientific was used to create a
428 control where the sgRNA scaffold was replaced by the promoterless *GUS* gene. Each LR product was
429 transformed in competent DH10β *E. coli* cells. Colony PCR-screening was performed using one of
430 the sgRNA oligonucleotides and the MiniTn7T-Rv primer (Table 3). PCR-positive colonies were grown
431 in LB broth supplemented with gentamycin and plasmids were purified and sequenced as described
432 earlier.

433 Transposition of the CRISPRi units (*aacC1*, *dCAS9*, and the sgRNA scaffold) was achieved by
434 electroporation of the corresponding plasmid and the helper plasmid pTNS1, which carries the
435 transposon machinery (Figure S4). Competent Xam668 cells were electroporated with pUC18-mini-
436 Tn7T-Gm-Gw-dCAS9 containing each sgRNA scaffold, pUC18-mini-Tn7T-Gm-Gw-dCAS9 containing
437 the promoterless *GUS* gene in the place of the sgRNA scaffold, and pUC18-mini-Tn7T-Gm-dCAS9, a
438 version of the plasmid where sequences flanked by attR1 and attR2 sites were eliminated, resulting
439 in a CRISPRi unit containing only *aacC1* and *dCAS9*. Gentamycin resistant Xam668 colonies were
440 selected and screened by PCR. Screening was performed by PCR on boiled colonies, using the
441 Scaffold-Fw and pglmS-Rv primers (Table 3). The expected band reflects the insertion of the CRISPRi
442 unit in the intergenic region located downstream of the *glms* gene. The GUS control was screened
443 with the GUS-Fw primer (Table 3) instead of Scaffold-Fw, and the scaffold-less control was screened
444 for *dCAS9* presence with primers dCAS9-Nterm-Fw and dCAS9-Nterm-Rv (Table 3). PCRs were
445 performed as described earlier. PCR-positive colonies were grown in Phi broth supplemented with
446 gentamycin and preserved in 20% glycerol stocks. Amplification products were sequenced by
447 Sanger.

448 **ArTALE and TALE-expressing vector construction and transformation**

449 The *MeSWEET10a*-activating arTALE used in this study was designed by Cohn and coworkers (Cohn
450 et al., 2014). Assembly was achieved through Golden TAL Technology, by using the arTALE
451 construction kit described by (Geißler et al., 2011) in the pSKX1 vector. The arTALE was transferred
452 to the expression vector pME6010 by using HindIII HF (NEB, USA) and BglII (NEB, USA), according to
453 manufacturer's instructions. Transformation of Xam668 CRISPRi strains was performed by
454 electroporation and selection on YPG plates supplemented with gentamycin and tetracycline.

455 **Plant material, culture conditions, and pathogen inoculation**

456 Cassava cuttings from cultivars 60444, CM6438-14, MTA18, and MCOL2215 were planted on
457 individual peat pots and grown under greenhouse conditions (27°C ± 5°C, photoperiod 12:12,
458 relative humidity > 70%). The japonica rice cultivar Kitaake, highly susceptible to PXO99^A and MAI1,
459 was planted in trays and grown under greenhouse conditions (27°C ± 2°C, photoperiod 12:12,
460 relative humidity 80% during day and 60% during night).

461 Water soaking symptom formation was assessed after Xam668 WT and CRISPRi strains infiltration
462 on leaves of 6-week-old cassava plants from the 60444 variety. Bacteria were cultured in Phi and
463 washed once with 10 mM MgCl₂ sterile solution. Cell density was adjusted to an OD₆₀₀ of 0.2 (c.a.
464 2×10⁸ cfu/mL). Adjusted bacterial suspensions were inoculated with a needleless syringe by

465 infiltrating an abaxial area of the leaflet per treatment. Plants were kept in a greenhouse under the
466 same conditions described above. Infiltrating points were individually photographed at 96 hours
467 post-inoculation (hpi) under a stereoscope. Each treatment was inoculated twice in four individual
468 plants, and the assay was replicated three times. Water soaking symptom formation for *Xc*
469 *CFBP4642* WT and CRISPRi strains was assessed using the described protocol with some changes: *Xc*
470 strains were cultured in PSA broth; the cassava varieties used for infiltration were CM6438-14,
471 MTAI8, and MCOL2215; photographs were taken at 5 days post-inoculation (dpi); each treatment
472 was inoculated twice in three individual plants; and the assay was carried out once.

473 Bacterial growth was also assessed by infiltration in 60444 plants. Bacterial suspensions were
474 adjusted to an OD₆₀₀ of 0.002, c.a. 2×10^6 cfu/mL, and infiltrated with a needleless syringe on the
475 abaxial surface. Two areas of the same treatment were infiltrated per leaflet, which correspond to
476 timepoints 0 and 7 days post-inoculation (DPI). A 4-mm leaf disc was taken from each infiltrated
477 area and ground with three glass beads in 1 mL of sterile 10 mM MgCl₂ in a TissueLyser II (Qiagen,
478 Germany). Serial dilutions of the grounded leaf discs were plated on YPG solid media supplemented
479 with cycloheximide and tetracycline. Three plants were used by experimental set, and the
480 experiment was repeated two times.

481 Water soaking symptom formation and lesion length were tested in rice for *Xoo* strains MAI1 and
482 PXO99^A. An isolated colony of each strain was streaked in PSA, allowed to grow for 48 hours, and
483 then resuspended in sterile ultrapure water. Cell density was adjusted to an OD₆₀₀ of 0.2 (c.a. 2×10^8
484 cfu/mL). For symptom formation, the adjusted bacterial suspensions and the mock solution were
485 inoculated with a needleless syringe by infiltrating three separate points on the abaxial area of the
486 second leaf of 3-week-old plants. For lesion length, the adjusted bacterial suspensions (OD₆₀₀=0.2)
487 and the mock (sterile water) were inoculated by leaf clipping of the second leaf in 4-week-old plants.
488 Plants were kept in a greenhouse under the same conditions described above. Infiltrated points
489 were individually photographed at 96 hpi under a stereoscope. Lesion length was measured at 15
490 dpi. Assays were replicated twice.

491 **TALE protein expression assays**

492 A single colony of the wildtype Xam668 (WT), CRISPRi-derived Xam668 strains, or complemented
493 Xam668 strains was grown overnight in Phi broth; WT and CRISPRi strains were grown in broth
494 without antibiotics, while complemented CRISPRi strains were grown in broth supplemented with
495 tetracycline. Cells were pelleted and washed twice with phosphate saline buffer (Euromedex,
496 France), and OD concentration was adjusted to 1.0. Three milliliters of each suspension were
497 pelleted by centrifugation and supernatant was removed. One milliliter of 1X Laemmli buffer
498 containing 700 mM of β -mercaptoethanol (Sigma Aldrich, Germany) was added to each tube and
499 heated to 95°C for 5 minutes, and then immediately put on ice.

500 For quantitative Western blots, 15 μ L of each bacterial lysate were loaded into a Mini-PROTEAN TGX
501 Stain-Free Precast Gel (Bio-Rad, USA) and run in tris-glycine-SDS buffer (Euromedex, France) at
502 constant 200 volts for 1.5 hours in a Mini-PROTEAN Tetra Cell (Bio-Rad, USA). Protein loading was
503 imaged in an E-box gel documentation imaging (Vilber, France) after three minutes of UV activation.
504 Blot transfer was performed with a Trans-Blot® Turbo™ Transfer Pack (Bio-Rad, USA) per
505 manufacturer's instruction, in a Trans-Blot Turbo System (Bio-Rad, USA) by using the integrated
506 protocol MIXED MW. After transfer, membranes were blocked with a 5% Bovine Serum Albumin –

507 BSA – (Sigma Aldrich, Germany) solution on Tris-Buffered Saline added with Tween-20 (TBS-T). TBS-
508 T was used for subsequent washes, and rabbit anti-TALE (Ac33450) antibody and HRP-conjugated
509 goat anti-rabbit (Thermo Fisher Scientific, USA) hybridizations were performed in 1% BSA TBS-T.
510 Luminescence was developed with the Pierce™ ECL Plus Western Blotting Substrate (Thermo Fisher
511 Scientific, USA) and images were acquired in a Typhoon™ FLA 9500 (General Electric Healthcare Life
512 Sciences, USA) with the ECL2 protocol. After imaging, membranes were soaked in Ponceau S solution
513 (Sigma-Aldrich, Germany) and washed with ultrapure water until the protein bands were visible.

514 **Infiltration, RNA extraction and retro-transcriptase quantitative PCR (RT-qPCR)**

515 *Xpm* strains were cultured in Phi and washed as described earlier. Bacterial inoculum was adjusted
516 to an OD₆₀₀ of 0.5 (c.a. 5×10^8 ufc/mL) in a 10-mM MgCl₂ solution. Bacterial suspensions and the mock
517 solution (10-mM MgCl₂) were infiltrated into leaves by means of a needleless syringe; each
518 treatment was inoculated on one leaflet per plant and on three different individuals. Infiltrated
519 tissue was collected at 50 hpi in sterile tubes containing RNA-free glass beads, samples were frozen
520 with liquid nitrogen and ground by vortexing. Total RNA was extracted with the Invitrap Spin plant
521 RNA minikit (STRATEC, Germany), using the RP buffer per manufacturer's instructions. Total RNAs
522 were treated with TURBO™ DNase (Thermo Fisher Scientific, USA) per manufacturer's instructions.
523 cDNA synthesis was performed with the SuperScript™ III Reverse Transcriptase (Thermo Fisher
524 Scientific, USA) and an Oligo(dT)12-18 Primer (Thermo Fisher Scientific, USA), per manufacturer's
525 instructions. RT-qPCRs were performed in 7-μL reactions of the Eurogentec Takyon™ SYBR® 2X qPCR
526 Mastermix Blue (Sycamore Life Sciences, USA) containing 0.3 μM of each primer. PCR cycling was as
527 follows: one step at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, and 60°C for 30 sec;
528 data was acquired during the second step of each cycle. The melting curve was evaluated from 65°C
529 to 95°C. Primers for gene *Manes.06G123400* encoding the sucrose transporter MeSWEET10a and
530 the reference gene *Manes.01G193700* encoding a serine/threonine-PP2A catalytic subunit (STPP)
531 are summarized in Table 3.

532 **TALE target prediction and candidate analysis**

533 TALE targets were predicted using TALVEZ (Pérez-Quintero et al., 2013). The *Manihot esculenta*
534 promoterome (1-kb sequences preceding annotated translational start sites) was extracted from
535 Phytozome's cassava genome version 6.1 (Bredeson et al., 2016), by means of the Biomart tool, and
536 it was used as input for target prediction. The algorithm was run using the default parameters.
537 Output data were analyzed in R (version 3.6.1) using an in-house script.

538 **Acknowledgements**

539 This research was funded by the Laboratorio Mixto Internacional (LMI) Biolnca and the Research
540 Program on Roots, Tubers and Bananas, and the Agropolis Foundation (project #1403-073). CAZ was
541 supported by a doctoral fellowship awarded by the Insitut de Recherche pour le Développement
542 and received funding from the Biological Sciences Faculty from Universidad de los Andes. We
543 acknowledge the advice of Lionel Gagnevin for DNA extraction and SMRT sequencing.

544 **References**

545 Afonina I, Ong J, Chua J, Lu T, Kline KA. 2020. Multiplex CRISPRi System Enables the Study of Stage-
546 Specific Biofilm Genetic Requirements in <span class="named-content genus-

- 547 species" id="named-content-1">Enterococcus faecalis" /span>
548 *MBio* **11**:e01101-20. doi:10.1128/mBio.01101-20
- 549 Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. 2010. Rice *xa13* Recessive Resistance to
550 Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene *Os-11N3*. *Plant*
551 *Cell* **22**:3864–3876. doi:10.1105/tpc.110.078964
- 552 Anzalone A V, Koblan LW, Liu DR. 2020. Genome editing with CRISPR–Cas nucleases, base editors,
553 transposases and prime editors. *Nat Biotechnol* **38**:824–844. doi:10.1038/s41587-020-0561-9
- 554 Anzalone A V, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA,
555 Raguram A, Liu DR. 2019. Search-and-replace genome editing without double-strand breaks
556 or donor DNA. *Nature* **576**:149–157. doi:10.1038/s41586-019-1711-4
- 557 Banno S, Nishida K, Arazoe T, Mitsunobu H, Kondo A. 2018. Deaminase-mediated multiplex
558 genome editing in *Escherichia coli*. *Nat Microbiol* **3**:423–429. doi:10.1038/s41564-017-0102-6
- 559 Barak JD, Vancheva T, Lefeuvre P, Jones JB, Timilsina S, Minsavage G V, Vallad GE, Koebnik R. 2016.
560 Whole-Genome Sequences of *Xanthomonas euvesicatoria* Strains Clarify Taxonomy and
561 Reveal a Stepwise Erosion of Type 3 Effectors . *Front Plant Sci* .
- 562 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.
563 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (80-)*
564 **315**:1709 LP – 1712. doi:10.1126/science.1138140
- 565 Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C,
566 Alicai T, Kumar L, Moreira LM, Neto JR, Verdier V, Santana MA, Kositcharoenkul N,
567 Vanderschuren H, Gruissem W, Bernal A, Staskawicz BJ. 2012. High-throughput genomic
568 sequencing of cassava bacterial blight strains identifies conserved effectors to target for
569 durable resistance. *Proc Natl Acad Sci U S A* **109**:E1972-9. doi:10.1073/pnas.1211014109
- 570 Bergthorsson U, Andersson DI, Roth JR. 2007. Ohno's dilemma: Evolution of new genes
571 under continuous selection. *Proc Natl Acad Sci* **104**:17004 LP – 17009.
572 doi:10.1073/pnas.0707158104
- 573 Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. 2013. Programmable repression
574 and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic*
575 *Acids Res* **41**:7429–7437. doi:10.1093/nar/gkt520
- 576 Boch J, Bonas U. 2010. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into
577 the plant cell. *Annu Rev Phytopathol* **48**:419–436. doi:10.1146/annurev-phyto-080508-
578 081936
- 579 Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. 2009.
580 Breaking the code of DNA binding specificity of TAL-type III effectors. *Science (80-)*
581 **326**:1509–1512. doi:10.1126/science.1178811
- 582 Bolot S, Munoz Bodnar A, Cunnac S, Ortiz E, Szurek B, Noël LD, Arlat M, Jacques M-A, Gagnevin L,
583 Portier P, Fischer-Le Saux M, Carrere S, Koebnik R. 2013. Draft Genome Sequence of the
584 *Xanthomonas cassavae* Type Strain CFBP 4642. *Genome Announc* **1**:e00679-13.
585 doi:10.1128/genomeA.00679-13
- 586 Bredeson J V, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, Grimwood J, Schmutz J,

- 587 Rabbi IY, Egesi C, Nauluvula P, Lebot V, Ndunguru J, Mkamilo G, Bart RS, Setter TL, Gleadow
588 RM, Kulakow P, Ferguson ME, Rounsley S, Rokhsar DS. 2016. Sequencing wild and cultivated
589 cassava and related species reveals extensive interspecific hybridization and genetic
590 diversity. *Nat Biotechnol* **34**:562–570. doi:10.1038/nbt.3535
- 591 Brito LF, Schultenkämper K, Passaglia LMP, Wendisch VF. 2020. CRISPR interference-based gene
592 repression in the plant growth promoter *Paenibacillus sonchi* genomovar *Riograndensis*
593 SBR5. *Appl Microbiol Biotechnol* **104**:5095–5106. doi:10.1007/s00253-020-10571-6
- 594 Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova
595 KS, Koonin E V, van der Oost J. 2008. Small CRISPR RNAs guide antiviral defense in
596 prokaryotes. *Science* **321**:960–964. doi:10.1126/science.1159689
- 597 Caro F, Place NM, Mekalanos JJ. 2019. Analysis of lipoprotein transport depletion in
598 Vibrio cholerae using CRISPRi. *Proc Natl Acad Sci* **116**:17013 LP –
599 17022. doi:10.1073/pnas.1906158116
- 600 Choi K-H, Schweizer HP. 2006. mini-Tn7 insertion in bacteria with single attTn7 sites: example
601 *Pseudomonas aeruginosa*. *Nat Protoc* **1**:153–161. doi:10.1038/nprot.2006.24
- 602 Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T,
603 Staskawicz BJ. 2014. *Xanthomonas axonopodis* Virulence Is Promoted by a Transcription
604 Activator-Like Effector–Mediated Induction of a SWEET Sugar Transporter in Cassava. *Mol*
605 *Plant-Microbe Interact* **27**:1186–1198. doi:10.1094/MPMI-06-14-0161-R
- 606 Cohn M, Morbitzer R, Lahaye T, Staskawicz BJ. 2016. Comparison of gene activation by two TAL
607 effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility
608 genes in cassava. *Mol Plant Pathol* 875–889. doi:10.1111/mpp.12337
- 609 Collins RE, Merz H, Higgs PG. 2011. Origin and evolution of gene families in Bacteria and Archaea.
610 *BMC Bioinformatics* **12 Suppl 9**:S14–S14. doi:10.1186/1471-2105-12-S9-S14
- 611 Copley SD. 2020. Evolution of new enzymes by gene duplication and divergence. *FEBS J* **287**:1262–
612 1283. doi:https://doi.org/10.1111/febs.15299
- 613 Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, Carpenter SCD, Chen L-Q, Zheng H, Gao X,
614 Zheng Y, Fei Z, Yu JZ, Isakeit T, Wheeler T, Frommer WB, He P, Bogdanove AJ, Shan L. 2017.
615 TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of
616 cotton. *Nat Commun* **8**:15588. doi:10.1038/ncomms15588
- 617 Cui L, Bikard D. 2016. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*.
618 *Nucleic Acids Res* **44**:4243–4251. doi:10.1093/nar/gkw223
- 619 Cunnac S, Boucher C, Genin S. 2004. Characterization of the cis-Acting
620 Regulatory Element Controlling HrpB-Mediated Activation of the Type III Secretion System
621 and Effector Genes in Ralstonia solanacearum; *J Bacteriol* **186**:2309 LP
622 – 2318. doi:10.1128/JB.186.8.2309-2318.2004
- 623 Cunnac S, Chakravarthy S, Kvitko BH, Russell AB, Martin GB, Collmer A. 2011. Genetic disassembly
624 and combinatorial reassembly identify a minimal functional repertoire of type III effectors in
625 Pseudomonas syringae; *Proc Natl Acad Sci* **108**:2975 LP – 2980.
626 doi:10.1073/pnas.1013031108

- 627 Denancé N, Szurek B, Doyle EL, Lauber E, Fontaine-Bodin L, Carrère S, Guy E, Hajri A, Cerutti A,
628 Boureau T, Poussier S, Arlat M, Bogdanove AJ, Noël LD. 2018. Two ancestral genes shaped
629 the *Xanthomonas campestris* TAL effector gene repertoire. *New Phytol* **219**:391–407.
630 doi:10.1111/nph.15148
- 631 Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J-K, Shi Y, Yan N. 2012. Structural basis for
632 sequence-specific recognition of DNA by TAL effectors. *Science* **335**:720–723.
633 doi:10.1126/science.1215670
- 634 Deng D, Yan C, Wu J, Pan X, Yan N. 2014. Revisiting the TALE repeat. *Protein Cell* **5**:297–306.
635 doi:10.1007/s13238-014-0035-2
- 636 Ding W, Zhang Y, Shi S. 2020. Development and Application of CRISPR/Cas in Microbial
637 Biotechnology . *Front Bioeng Biotechnol* .
- 638 Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier
639 RJ, Root DE. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene
640 inactivation. *Nat Biotechnol* **32**:1262–1267. doi:10.1038/nbt.3026
- 641 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
642 *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340
- 643 Ellis NA, Kim B, Tung J, Machner MP. 2021. A multiplex CRISPR interference tool for virulence gene
644 interrogation in *Legionella pneumophila*. *Commun Biol* **4**:157. doi:10.1038/s42003-021-
645 01672-7
- 646 Erkes A, Reschke M, Boch J, Grau J. 2017. Evolution of Transcription Activator-Like Effectors in
647 *Xanthomonas oryzae*. *Genome Biol Evol* **9**:1599–1615. doi:10.1093/gbe/evx108
- 648 Fenselau S, Bonas U. 1995. Sequence and expression analysis of the hrpB pathogenicity operon of
649 *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to
650 components of the Hrp, Ysc, Spa, and Fli secretion systems. *Mol Plant Microbe Interact*
651 **8**:845–854. doi:10.1094/mpmi-8-0845
- 652 Fernandes LG V, Guaman LP, Vasconcellos SA, Heinemann MB, Picardeau M, Nascimento ALTO.
653 2019. Gene silencing based on RNA-guided catalytically inactive Cas9 (dCas9): a new tool for
654 genetic engineering in *Leptospira*. *Sci Rep* **9**:1839. doi:10.1038/s41598-018-37949-x
- 655 Ferreira RM, de Oliveira ACP, Moreira LM, Belasque J, Gourbeyre E, Siguier P, Ferro MIT, Ferro JA,
656 Chandler M, Varani AM. 2015. A TALE of Transposition: Tn3-Like Transposons Play a Major
657 Role in the Spread of Pathogenicity Determinants of *Xanthomonas citri* and Other
658 *Xanthomonads*. *MBio* **6**. doi:10.1128/mBio.02505-14
- 659 Gasiunas G, Barrangou R, Horvath P, Siksnys V. 2012. Cas9-crRNA ribonucleoprotein complex
660 mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci*
661 **109**:E2579 LP-E2586. doi:10.1073/pnas.1208507109
- 662 Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. 2017. Programmable
663 base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**:464–471.
664 doi:10.1038/nature24644
- 665 Geißler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens S-E, Boch J. 2011. Transcriptional
666 activators of human genes with programmable DNA-specificity. *PLoS One* **6**:e19509.

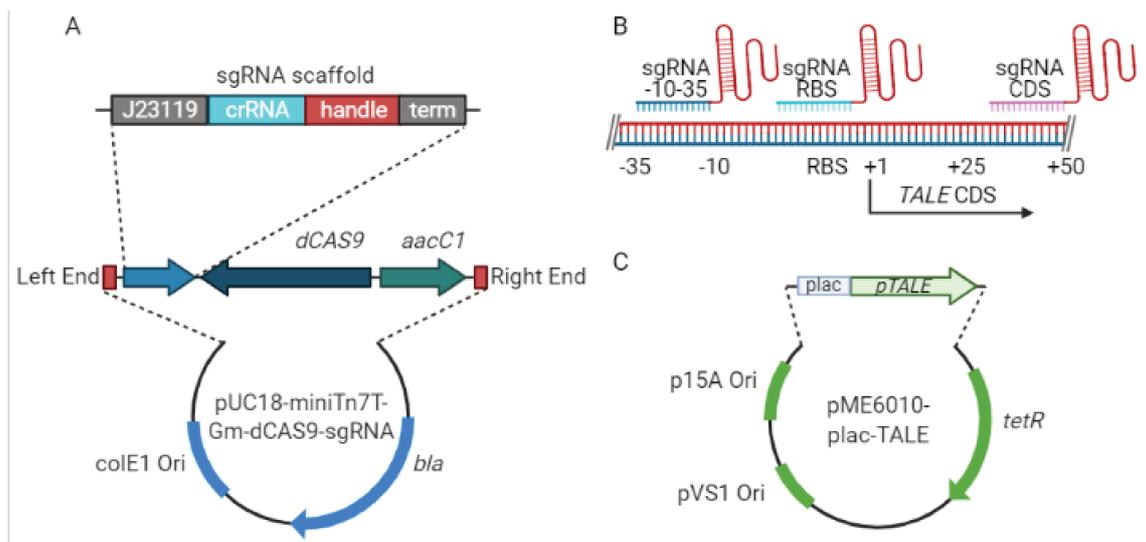
- 667 doi:10.1371/journal.pone.0019509
- 668 Hogan AM, Rahman ASMZ, Lightly TJ, Cardona ST. 2019. A Broad-Host-Range CRISPRi Toolkit for
669 Silencing Gene Expression in Burkholderia. *ACS Synth Biol* **8**:2372–2384.
670 doi:10.1021/acssynbio.9b00232
- 671 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O,
672 Cradick TJ, Marraffini LA, Bao G, Zhang F. 2013. DNA targeting specificity of RNA-guided Cas9
673 nucleases. *Nat Biotechnol* **31**:827–832. doi:10.1038/nbt.2647
- 674 Hu M, Hu W, Xia Z, Zhou X, Wang W. 2016. Validation of Reference Genes for Relative Quantitative
675 Gene Expression Studies in Cassava (*Manihot esculenta* Crantz) by Using Quantitative Real-
676 Time PCR. *Front Plant Sci* **7**:680. doi:10.3389/fpls.2016.00680
- 677 Ji Z, Ji C, Liu B, Zou L, Chen G, Yang B. 2016. Interfering TAL effectors of *Xanthomonas oryzae*
678 neutralize R-gene-mediated plant disease resistance. *Nat Commun* **7**:13435.
679 doi:10.1038/ncomms13435
- 680 Koebnik R, Krüger A, Thieme F, Urban A, Bonas U. 2006. Specific binding of the *Xanthomonas*
681 *campestris* pv. *vesicatoria* AraC-type transcriptional activator HrpX to plant-inducible
682 promoter boxes. *J Bacteriol* **188**:7652–7660. doi:10.1128/JB.00795-06
- 683 Kvitko BH, Park DH, Velásquez AC, Wei C-F, Russell AB, Martin GB, Schneider DJ, Collmer A. 2009.
684 Deletions in the Repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 Type III Secretion
685 Effector Genes Reveal Functional Overlap among Effectors. *PLOS Pathog* **5**:e1000388.
- 686 Lai Y, Babunovic GH, Cui L, Dedon PC, Doench JG, Fortune SM, Lu TK. 2020. Illuminating Host-
687 Mycobacterial Interactions with Genome-wide CRISPR Knockout and CRISPRi Screens. *Cell*
688 *Syst* **11**:239-251.e7. doi:https://doi.org/10.1016/j.cels.2020.08.010
- 689 Liu R, Liang L, Freed EF, Gill RT. 2021. Directed Evolution of CRISPR/Cas Systems for Precise Gene
690 Editing. *Trends Biotechnol* **39**:262–273. doi:10.1016/j.tibtech.2020.07.005
- 691 Mak AN-S, Bradley P, Bogdanove AJ, Stoddard BL. 2013. TAL effectors: function, structure,
692 engineering and applications. *Curr Opin Struct Biol* **23**:93–99.
693 doi:https://doi.org/10.1016/j.sbi.2012.11.001
- 694 Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. 2012. The Crystal Structure of TAL
695 Effector PthXo1 Bound to Its DNA Target. *Science (80-)* **335**:716–719.
696 doi:10.1126/science.1216211
- 697 Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D,
698 Haft DH, Horvath P, Moineau S, Mojica FJM, Scott D, Shah SA, Siksnyš V, Terns MP, Venclovas
699 Č, White MF, Yakunin AF, Yan W, Zhang F, Garrett RA, Backofen R, van der Oost J, Barrangou
700 R, Koonin E V. 2020. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and
701 derived variants. *Nat Rev Microbiol* **18**:67–83. doi:10.1038/s41579-019-0299-x
- 702 Moscou MJ, Bogdanove AJ. 2009. A Simple Cipher Governs DNA Recognition by TAL Effectors.
703 *Science (80-)* **326**:1501 LP – 1501. doi:10.1126/science.1178817
- 704 Münch R, Hiller K, Grote A, Scheer M, Klein J, Schobert M, Jahn D. 2005. Virtual Footprint and
705 PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*
706 **21**:4187–4189. doi:10.1093/bioinformatics/bti635

- 707 Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom J-S, Li C, Nguyen H,
708 Liu B, Auguy F, Sciallano C, Luu VT, Dossa GS, Cunnac S, Schmidt SM, Slamet-Loedin IH, Vera
709 Cruz C, Szurek B, Frommer WB, White FF, Yang B. 2019. Broad-spectrum resistance to
710 bacterial blight in rice using genome editing. *Nat Biotechnol* **37**:1344–1350.
711 doi:10.1038/s41587-019-0267-z
- 712 Peng A, Chen S, Lei T, Xu L, He Y, Wu L, Yao L, Zou X. 2017. Engineering canker-resistant plants
713 through CRISPR/Cas9-targeted editing of the susceptibility gene CsLOB1 promoter in citrus.
714 *Plant Biotechnol J* **15**:1509–1519. doi:10.1111/pbi.12733
- 715 Pérez-Quintero AL, Rodriguez-R LM, Dereeper A, López Carrascal CE, Koebnik R, Szurek B, Cunnac
716 S. 2013. An Improved Method for TAL Effectors DNA-Binding Sites Prediction Reveals
717 Functional Convergence in TAL Repertoires of *Xanthomonas oryzae* Strains. *PLoS One*
718 **8**:e68464. doi:10.1371/journal.pone.0068464
- 719 Perez-Quintero AL, Szurek B. 2019. A Decade Decoded: Spies and Hackers in the History of TAL
720 Effectors Research. *Annu Rev Phytopathol* **57**:459–481. doi:10.1146/annurev-phyto-082718-
721 100026
- 722 Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. 2013. Repurposing
723 CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*
724 **152**:1173–1183. doi:10.1016/j.cell.2013.02.022
- 725 Ruh M, Briand M, Bonneau S, Jacques M-A, Chen NWG. 2017. *Xanthomonas* adaptation to
726 common bean is associated with horizontal transfers of genes encoding TAL effectors. *BMC*
727 *Genomics* **18**:670. doi:10.1186/s12864-017-4087-6
- 728 Schandry N, Jacobs JM, Szurek B, Perez-Quintero AL. 2018. A cautionary TALE: how plant breeding
729 may have favoured expanded TALE repertoires in *Xanthomonas*. *Mol Plant Pathol* **19**:1297–
730 1301. doi:10.1111/mpp.12670
- 731 Schwartz AR, Morbitzer R, Lahaye T, Staskawicz BJ. 2017. TALE-induced bHLH transcription factors
732 that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. *Proc*
733 *Natl Acad Sci* **114**:E897 LP-E903. doi:10.1073/pnas.1620407114
- 734 Sharda M, Badrinarayanan A, Seshasayee ASN. 2020. Evolutionary and Comparative Analysis of
735 Bacterial Nonhomologous End Joining Repair. *Genome Biol Evol* **12**:2450–2466.
736 doi:10.1093/gbe/evaa223
- 737 Streubel J, Pesce C, Hutin M, Koebnik R, Boch J, Szurek B. 2013. Five phylogenetically close rice
738 *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*.
739 *New Phytol* **200**:808–819. doi:10.1111/nph.12411
- 740 Szurek B, Rossier O, Hause G, Bonas U. 2002. Type III-dependent translocation of the
741 *Xanthomonas* AvrBs3 protein into the plant cell. *Mol Microbiol* **46**:13–23. doi:10.1046/j.1365-
742 2958.2002.03139.x
- 743 Takacs CN, Scott M, Chang Y, Kloos ZA, Irnov I, Rosa PA, Liu J, Jacobs-Wagner C. 2021. A CRISPR
744 Interference Platform for Selective Downregulation of Gene Expression in *Borrelia burgdorferi*
745 *class="named-content genus-species" id="named-content-*
746 *1" >Borrelia burgdorferi*. *Appl Environ Microbiol* **87**:e02519-20.
747 doi:10.1128/AEM.02519-20

- 748 Tan SZ, Reisch CR, Prather KLJ. 2018. A Robust CRISPR Interference Gene Repression System in
 749 <span class="named-content genus-species" id="named-content-
 750 1">&Pseudomonas; *J Bacteriol* **200**:e00575-17. doi:10.1128/JB.00575-17
- 751 Todor H, Silvis MR, Osadnik H, Gross CA. 2021. Bacterial CRISPR screens for gene function. *Curr*
 752 *Opin Microbiol* **59**:102–109. doi:https://doi.org/10.1016/j.mib.2020.11.005
- 753 Wang T, Wang M, Zhang Q, Cao S, Li X, Qi Z, Tan Y, You Y, Bi Y, Song Y, Yang R, Du Z. 2019.
 754 Reversible Gene Expression Control in <span class="named-content genus-
 755 species" id="named-content-1">&Yersinia pestis; by Using
 756 an Optimized CRISPR Interference System. *Appl Environ Microbiol* **85**:e00097-19.
 757 doi:10.1128/AEM.00097-19
- 758 Wei C-F, Kvitko BH, Shimizu R, Crabill E, Alfano JR, Lin N-C, Martin GB, Huang H-C, Collmer A. 2007.
 759 A *Pseudomonas syringae* pv. tomato DC3000 mutant lacking the type III effector HopQ1-1 is
 760 able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J* **51**:32–46.
 761 doi:https://doi.org/10.1111/j.1365-313X.2007.03126.x
- 762 Wengelnik K, Van den Ackerveken G, Bonas U. 1996. HrpG, a key hrp regulatory protein of
 763 *Xanthomonas campestris* pv. vesicatoria is homologous to two-component response
 764 regulators. *Mol Plant Microbe Interact* **9**:704–712. doi:10.1094/mpmi-9-0704
- 765 Xue J, Lu Z, Liu W, Wang S, Lu D, Wang X, He X. 2021. The genetic arms race between plant and
 766 *Xanthomonas*: lessons learned from TALE biology. *Sci China Life Sci* **64**:51–65.
 767 doi:10.1007/s11427-020-1699-4
- 768 Yang Y, Gabriel DW. 1995. *Xanthomonas* avirulence/pathogenicity gene family encodes functional
 769 plant nuclear targeting signals. *Mol Plant Microbe Interact* **8**:627–631. doi:10.1094/mpmi-8-
 770 0627
- 771 Yuan M, Ke Y, Huang R, Ma L, Yang Z, Chu Z, Xiao J, Li X, Wang S. 2016. A host basal transcription
 772 factor is a key component for infection of rice by TALE-carrying bacteria. *Elife* **5**:e19605.
 773 doi:10.7554/eLife.19605
- 774 Zhu W, Yang B, Chittoor JM, Johnson LB, White FF. 1998. AvrXa10 Contains an Acidic
 775 Transcriptional Activation Domain in the Functionally Conserved C Terminus. *Mol Plant-
 776 Microbe Interact* **11**:824–832. doi:10.1094/MPMI.1998.11.8.824

777

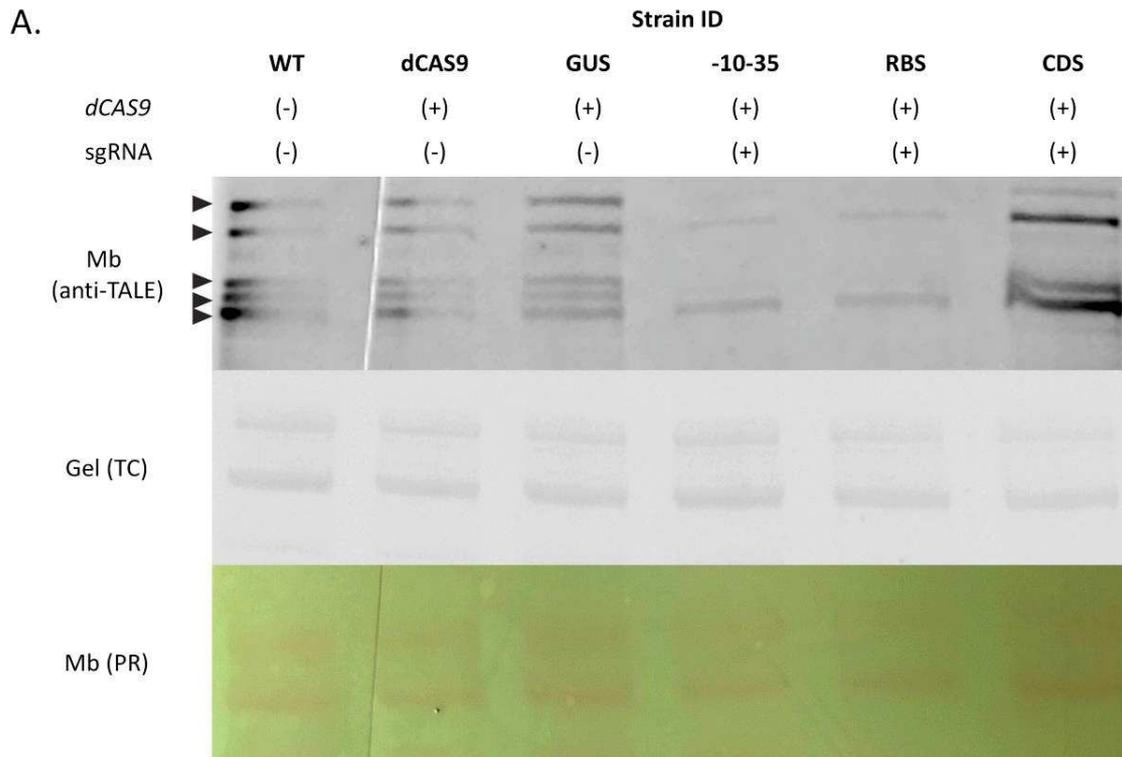
778 **Figures**



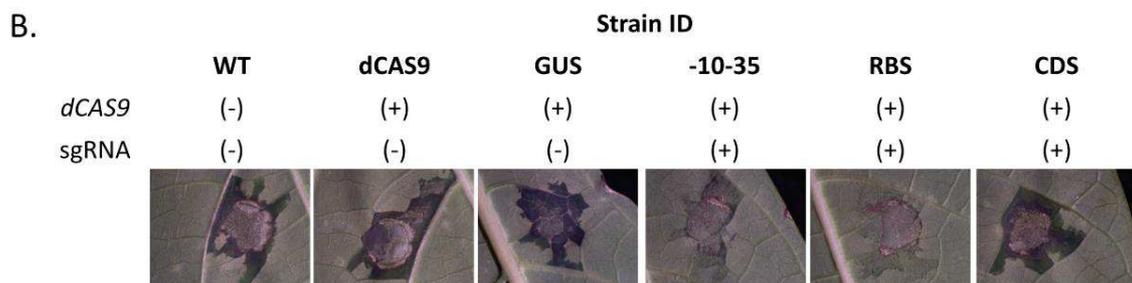
779

780 **Figure 1. Elements of the CRISPRi platform in *Xanthomonas* and complementation plasmid.** A.
 781 Scheme showing the structure of the non-replicative plasmid pUC18-mini-Tn7T-Gm-dCas9-sgRNA
 782 which is carrying the transposable CRISPRi unit. Arrows represent genes or expression cassettes and
 783 their directionality. Replication origin is depicted as a blue box. The left and the right border delimit
 784 the transposable unit (red rectangles). The CRISPRi unit comprises the sgRNA scaffold, the *S.*
 785 *pyogenes* *dCas9* gene (*dCas9*), and the gentamycin resistance gene (*aacC1*) marker. The sgRNA
 786 scaffold is expressed under the highly-active and constitutive J23119 promoter, the crRNA is fused
 787 to the *S. pyogenes* CRISPR handle (sgRNA) and the tracrRNA terminator of *S. pyogenes*. ColE1 Ori:
 788 origin of replication; *bla*: ampicillin resistance cassette. B. Schematic representation of sgRNAs
 789 targeting *TALE* genes and their elements. Positions of the -35 and -10 promoter elements and the
 790 ribosome binding sites (RBS) of *TALE*s are indicated under DNA strands. Positive numeration
 791 indicates the position on the *TALE* coding sequence C. Replicative vector pME6010 carrying a
 792 modified *TALE* that is protected against sgRNAs (*pTALE*). The *pTALE* gene is under the control of the
 793 lac promoter (light blue box) and its RBS and CDS sequences were modified to avoid sgRNA-
 794 mediated silencing. The plasmid carries two replication origins and a tetracycline resistance gene
 795 (*tetR*) as selection marker. The graphical representation was created with BioRender.com.

796

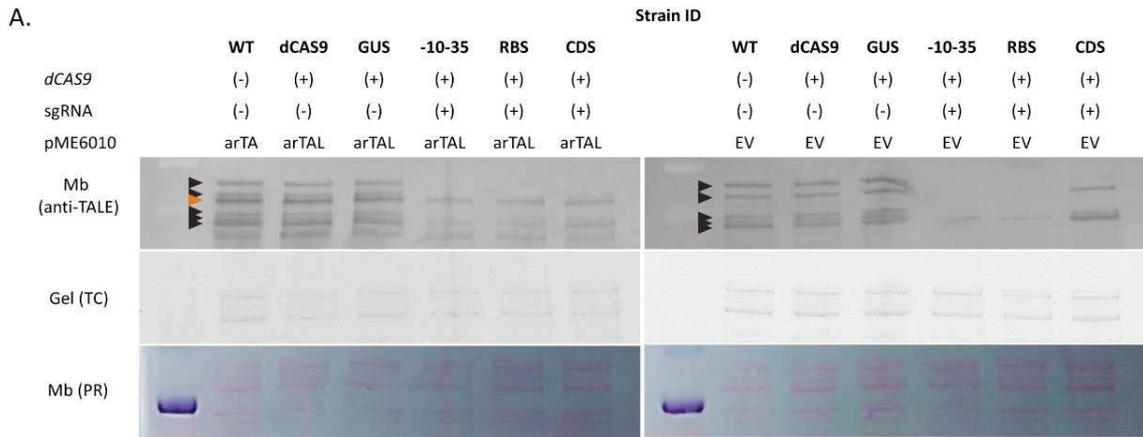


797

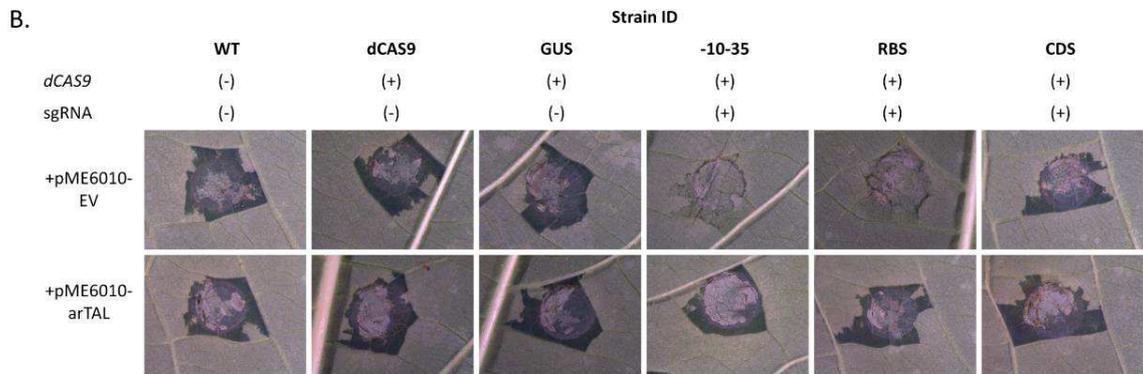


798

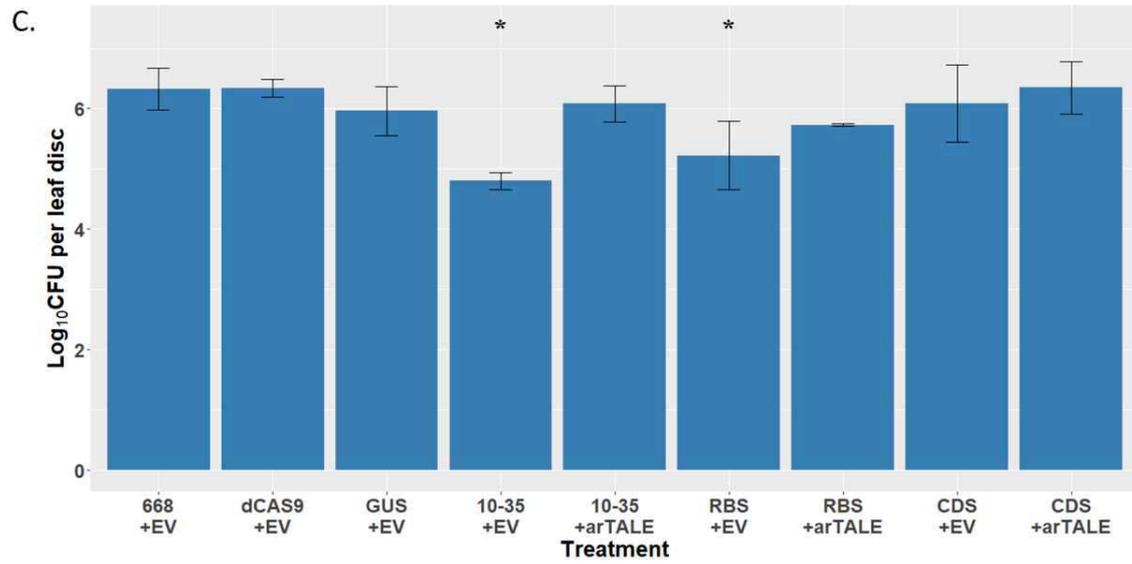
799 Figure 2. Knocking-down *TALE* genes in *Xpm* abolishes pathogen virulence. A. Western-blot analysis
 800 of total protein extracts of wildtype Xam668 (WT) and derivative strains transformed with a CRISPRi
 801 unit without sgRNA (dCas9), where the sgRNA insertion is replaced by a promoterless *GUS* (GUS), or
 802 expressing the -10-35 (-10-35), RBS (RBS) or CDS (CDS) anti-*TALE* sgRNAs, resolved by SDS-PAGE
 803 using an anti-*TALE* antibody. Gray arrowheads point to TAL13_{Xam668}, TAL14_{Xam668}, TAL15_{Xam668},
 804 TAL20_{Xam668}, and TAL22_{Xam668} bands (from smallest to largest). Protein loading controls are shown
 805 through UV activation of the gel trihalo compound (TC) and Ponceau S red (PR) staining of the
 806 membrane. B. Water soaking symptoms observed at 5 dpi on cassava 60444 leaves infiltrated with
 807 strains of Xam668 harboring the indicated constructs.



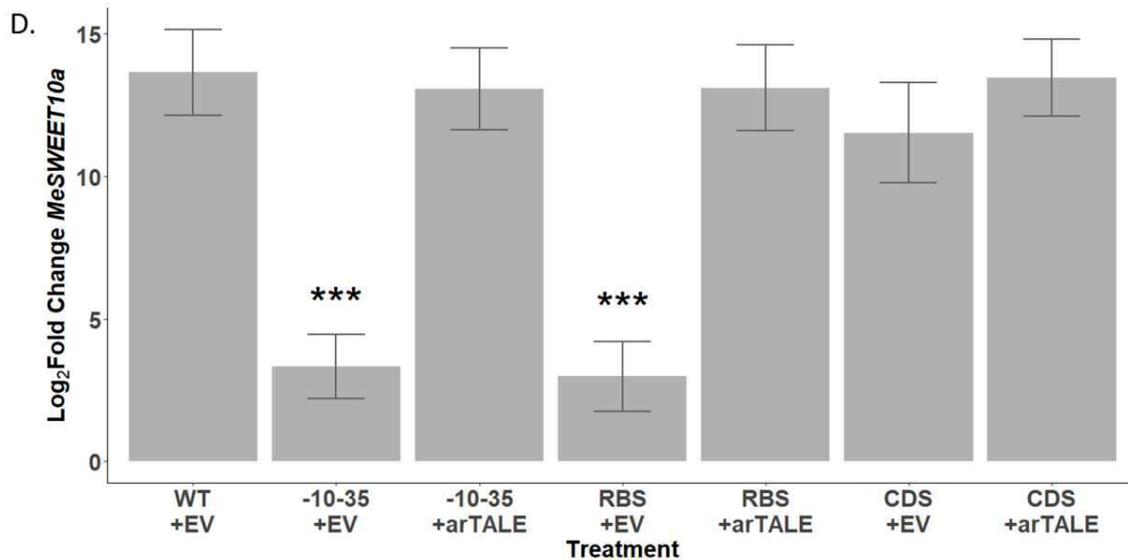
808



809

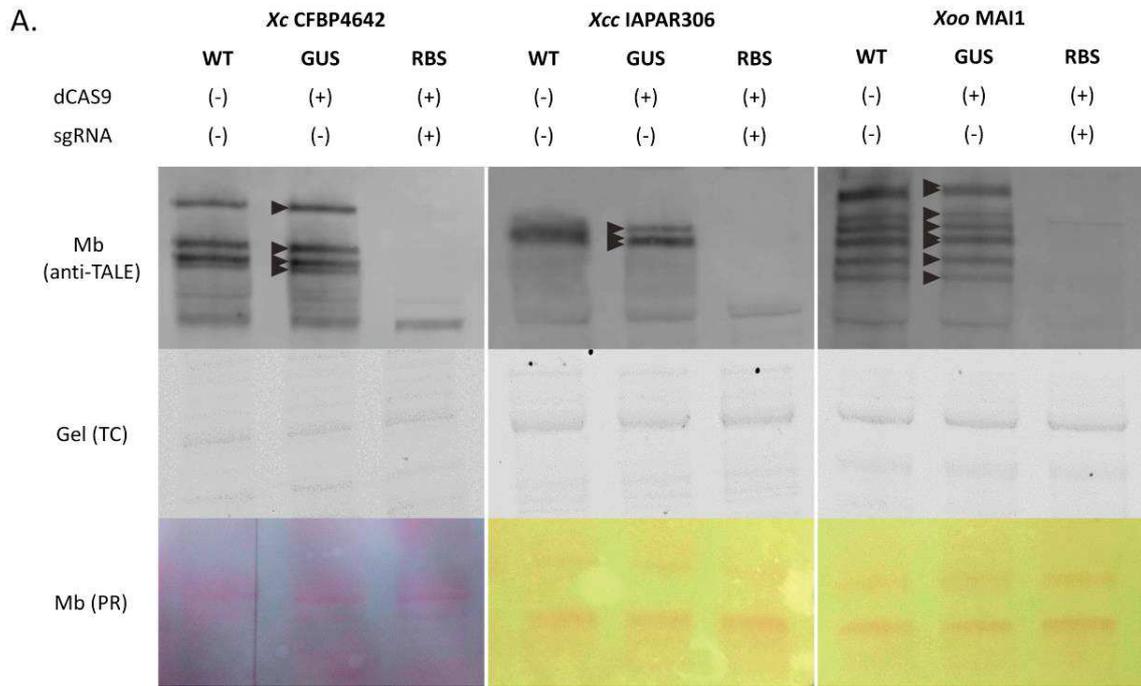


810

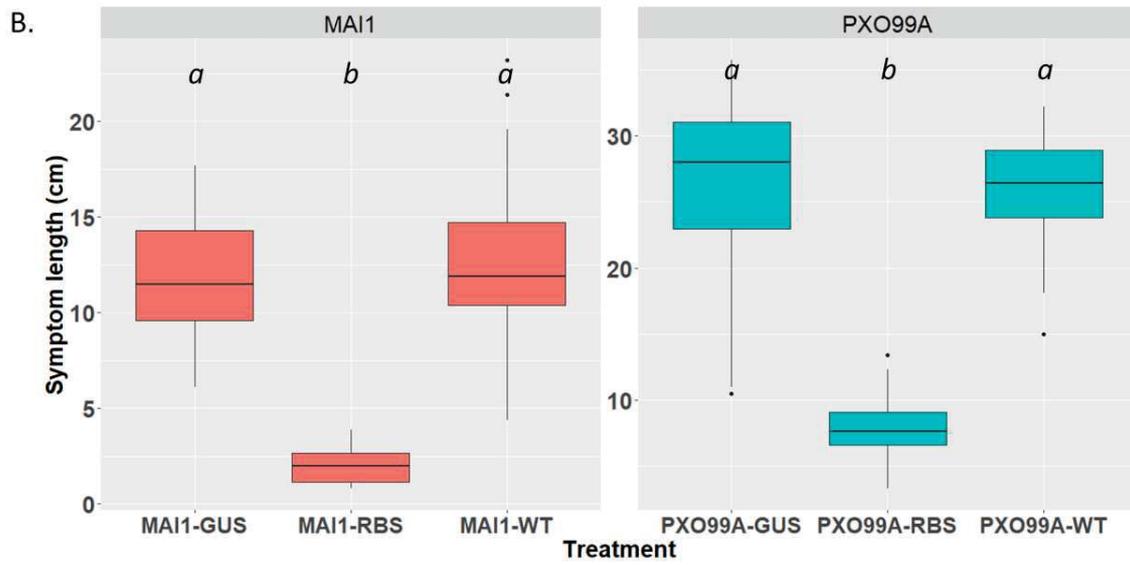


811

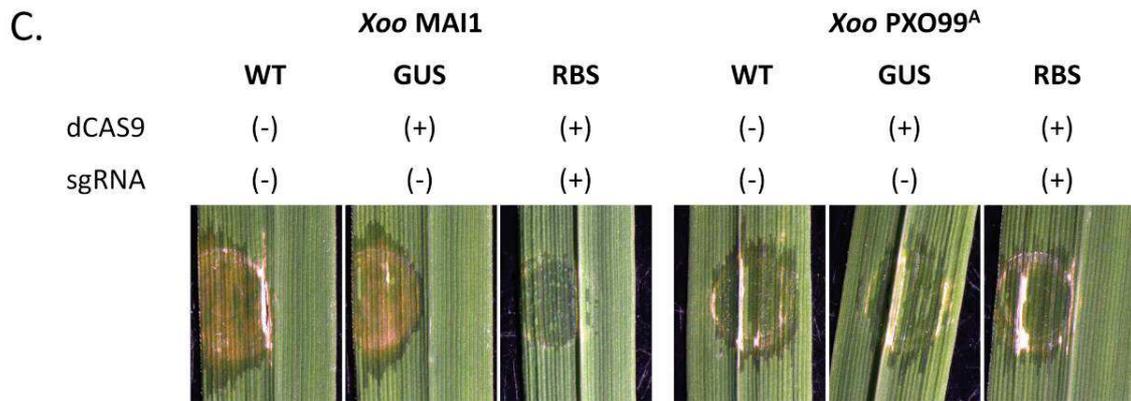
812 Figure 3. CRISPRi strains of *Xpm* knocked down for *TALEs* can be complemented in protein
813 expression and virulence with an arTALE resistant to the sgRNAs. A. Western blots showing *TALE*
814 expression in the WT and the CRISPRi strains for *TALE* genes transformed with pME6010-arTALE or
815 pME6010-EV. First panel shows luminescent signals (anti-TALE) after development of the Western
816 blot membrane. Gray arrowheads point to TAL13_{Xam668}, TAL14_{Xam668}, TAL15_{Xam668}, TAL20_{Xam668}, and
817 TAL22_{Xam668} bands (from smallest to largest), while orange arrowhead points to the arTALE band.
818 Protein loading controls are shown through UV activation of the gel trihalo compound (TC) and
819 Ponceau S red (PR) staining of the membrane. B. Water soaking symptoms at 96 hpi in cassava
820 cultivar 60444. Photographs are representative of three experimental replicates with similar results.
821 C. Bacterial titers at 7 dpi on the cassava variety 60444 infiltrated with the transformed WT and
822 CRISPRi strains. Bacterial titers significantly different from the ones obtained for the WT strain are
823 marked with an asterisk (Tukey's test, $\alpha=0.05$). Bars represent standard deviations of three
824 biological replicates. The experiment was repeated twice, with similar results. D. *MeSWEET10a*
825 relative expression at 50 hpi after infiltration with the WT strain and the CRISPRi strains transformed
826 with the pME6010-EV or pME6010-arTALE. The fold change was calculated against a mock-
827 inoculated treatment, and the reference gene was *Manes.01G193700*, a Serine/threonine-PP2A
828 catalytic subunit encoding gene (STPP). Bars represent transformed standard deviations from three
829 biological replicates, and the three stars indicate that the relative expression is significantly different
830 (t-student test, $\alpha = 0.001$) when compared to the WT strain transformed with the EV.



831

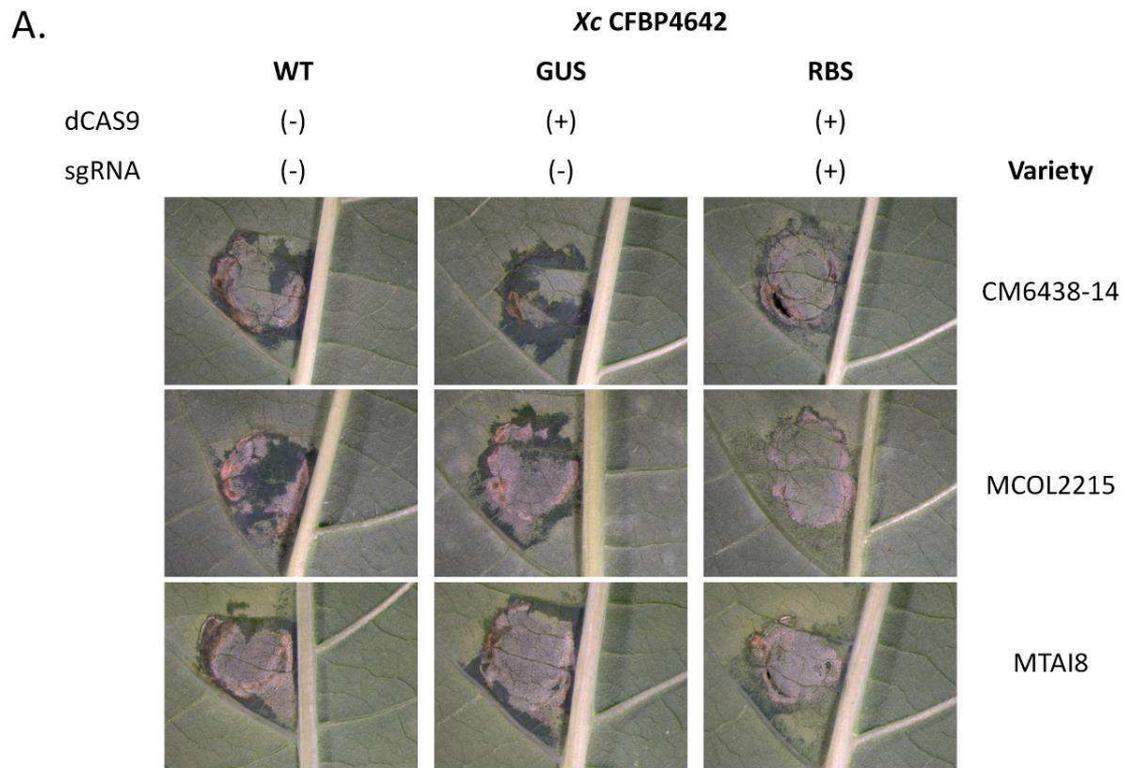


832



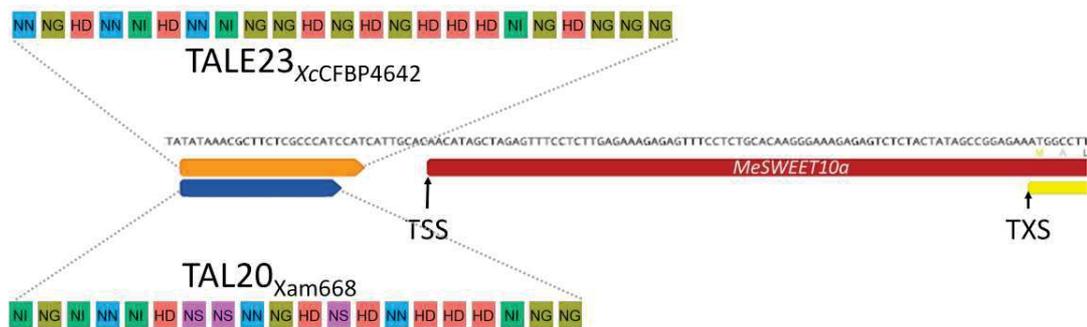
833

834 Figure 4. *TALE* silencing in other *Xanthomonas* species. A. Western blots showing *TALE* expression
835 in the CRISPR or WT strains of *Xc*, *Xcc*, and *Xoo*. The control CRISPRi strains contained the CRISPR
836 unit where sgRNA was replaced by a promoterless GUS gene, while the test CRISPRi strains
837 expressed the RBS sgRNA that were designed for *Xpm*. First panel shows luminescent signals (anti-
838 *TALE*) after development of the Western blot membrane. Gray arrowheads point to *TALE* bands.
839 Protein loading controls are shown through UV activation of the gel trihalo compound (TC) and
840 Ponceau S red (PR) staining of the membrane. The lower bands are common to the Western blots
841 of *Xanthomonas* total protein extracts when using this antibody. In any case the molecular weight
842 would correspond to a *TALE* protein (>98 kDa). B. Lesion length caused by the WT and CRISPRi strains
843 measured 15 days after inoculation by leaf clipping of Kitaake plants (n=15). Different italicized
844 letters indicate statistically significant differences according to Dunn's test ($\alpha=0.05$) among each
845 group (MAI1 or PXO99^A). The experiment was repeated twice with similar results. C. Representative
846 photos of infiltration of the WT and the CRISPRi strains for both *Xoo* strains in Kitaake rice. Water
847 soaking symptoms were photographed at 96 hpi.



848

B.



849

850 Figure 5. Knocking-down of TALE impairs virulence in *X. cassavae*. A. Representative photos of water
 851 soaking symptoms observed at 5 dpi on leaves of three cassava varieties infiltrated with *Xc*
 852 CFBP4642 WT and CRISPRi strains. The control CRISPRi strain contained the CRISPR unit where
 853 sgRNA was replaced by a promoterless GUS gene, while the test CRISPRi strain expressed the RBS
 854 sgRNA that were designed for *Xpm*. Cassava varieties are identified at the right side of each panel.
 855 The experiment was performed once. B. Schematic representation highlighting the predicted
 856 evolutionary convergence of *Xc* TALE23_{XcCFBP4642} and *Xpm* TAL20_{Xam668}. Both TALEs bind to the same
 857 EBE in the promoter sequence of the *S* gene *Manes.06123400* (*MeSWEET10a*). The nucleotide
 858 sequence is depicted in gray. EBEs are represented as orange (TALE23) and blue (TAL20) arrows,
 859 while red and yellow boxes correspond to the predicted transcript and coding sequences of

860 *MeSWEET10a*. TSS indicates transcriptional start site, TXS indicates translational start site. Both
 861 TALEs are depicted as a collection of color-coded boxes that reflect the RVD sequence of each
 862 protein. The graphical representation was created using Geneious software version 11.

863 **Tables**

864 Table 1. Characteristics of crRNAs used to construct the sgRNAs of this study. All of them were 20
 865 nucleotides in length and were complementary to the Watson strand (for -10-35) or the non-
 866 template strand. ^aPosition is relative to the first nucleotide of the coding sequence, which is
 867 considered as +1. ^bCalculated in Geneious software (version 11) with the algorithm proposed by
 868 Doench and coworkers (Doench et al., 2014). ^cCalculated in Geneious software (version 11) with
 869 the algorithm proposed by Hsu and coworkers (Hsu et al., 2013).

crRNA	Targeted region ^a	Sequence	PAM	Activity score ^b	Off-target score ^c
-10-35	-49:-30	GACCAGAGATCTTTTAGTCT	TGG	0.029	100.00%
RBS	-19:+1	CTATAAAGAGGTATGCCTGA	TGG	0.583	100.00%
CDS	+31:+50	CCTGCCGACGAACCTTTGGC	CGG	0.553	100.00%

870

871 Table 2. Characterization of RBS anti-*Xpm* sgRNA specificity against *TALEs* from other *Xanthomonas*
 872 species. ^aMismatches found in the seed sequence (5 base pairs adjacent to the PAM). ^bMismatches
 873 found in the first 12 nucleotides. ^cMismatches found in the whole crRNA sequence.

<i>Xanthomonas</i> spp.	TALE ID	M5 ^a	M12 ^b	M20 ^c	Target potential
<i>Xcc</i> IAPAR306	<i>pthA1, pthA2, pthA3, pthA4</i>	0	0	2	High
<i>Xc</i> CFBP4642	<i>tale14_{XcCFBP4642}, tale15A_{XcCFBP4642}, tale15B_{XcCFBP4642}, tale16_{XcCFBP4642}, tale17_{XcCFBP4642}, tale23_{XcCFBP4642}</i>	0	0	2	High
<i>Xoo</i> MAI1	<i>talA, talB, talE, talF, talG, talI</i>	0	0	1	High
	<i>talC, talD, talH</i>	0	0	2	High
<i>Xoo</i> PXO99 ^A	<i>tal2a, tal4, tal5a, tal5b (pthXo6), tal6a, tal6b, tal7a, tal7b, tal8a, tal8b, tal9a, tal9b, tal9d, tal9e, pthXo1, pthXo7</i>	0	0	2	High
	<i>tal9c (avrXa27), tal3a, tal3b</i>	1	1	3	Low

874

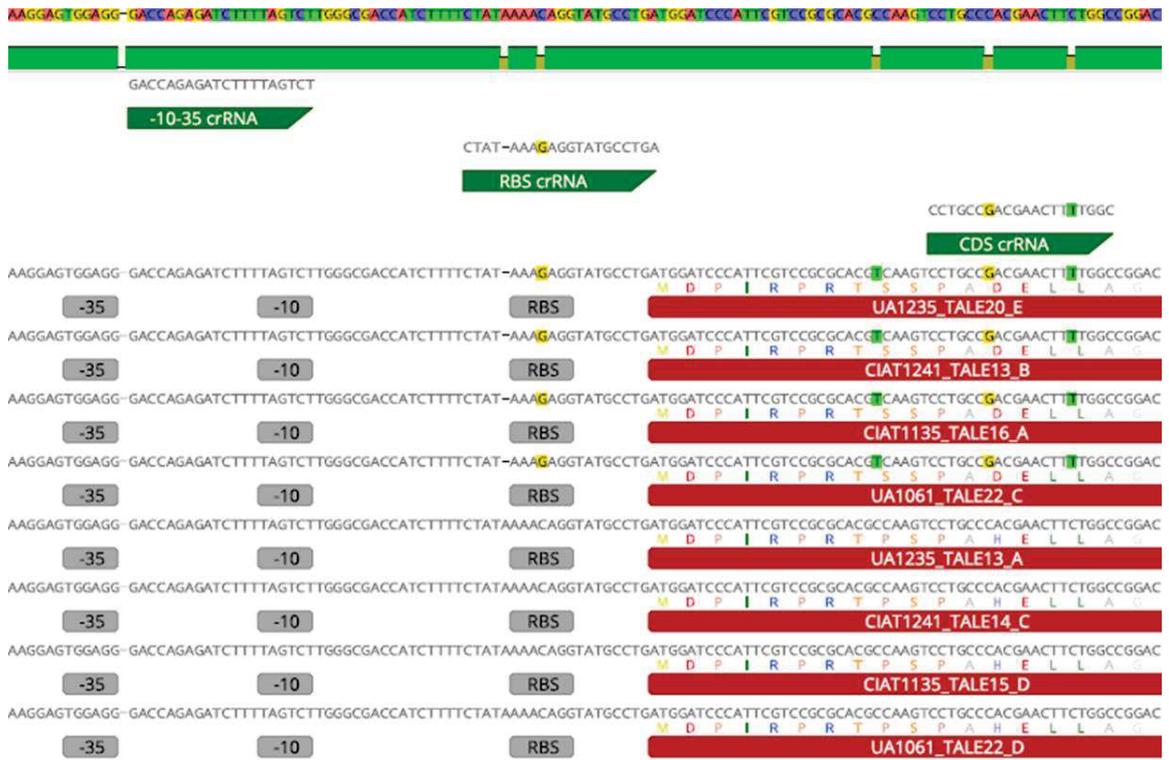
875 Table 3. Primers used in this study. Primers were designed with the primer3 software
 876 (<http://primer3.ut.ee/>) with set parameters to obtain melting temperatures of 60±2°C.

ID	Sequence (5'→3')	Reference
-10-35-sgRNA-Fw	AAACGACCAGAGATCTTTTAGTCTG	This study
RBS-sgRNA-Fw	AAACCTATAAAGAGGTATGCCTGAG	This study
CDS-sgRNA-Fw	AAACCCTGCCGACGAACCTTTGGCG	This study
M13-Rv	CAGGAAACAGCTATGAC	Universal primer
MiniTn7T-Rv	CTTATCTGGTTGGCCTGCAA	This study
Scaffold-Fw	CAGCTTTTGACAGCTAGCTCAGT	This study
pglM5-Rv	GACTCGTGCATACGTCCTCC	This study

GUS-Fw	TGCGAGGTACGGTAGGAGTT	This study
dCAS9-Nterm-Fw	GGGGAGCTAGCCAAGAAGAA	This study
dCAS9-Nterm-Rv	ACAAACGTCCCAACCAGTA	This study
MeSWEET10a-Fw	TCCTCACCTTGACTGCGGTG	(Cohn et al., 2014)
MeSWEET10a-Rv	AGCACCATCTGGACAATCCCA	(Cohn et al., 2014)
STPP-Fw	GCTTGTCATGGAAGGGTACAA	(Hu et al., 2016)
STPP-Rv	TTCCCACATCGGTAGCAATAG	(Hu et al., 2016)

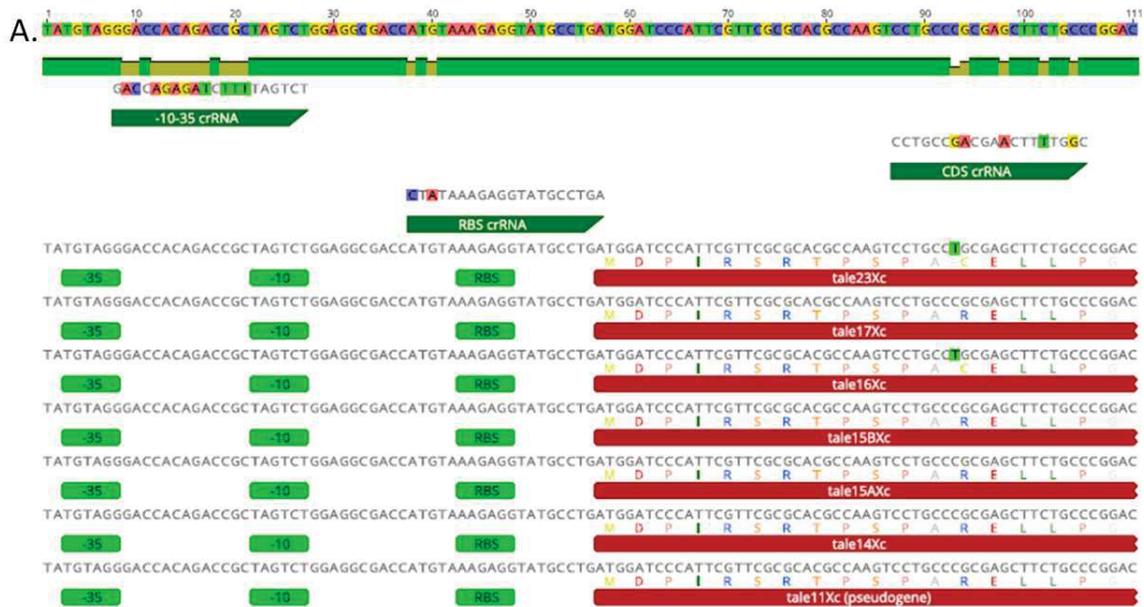
877

878 **Supplementary figures**

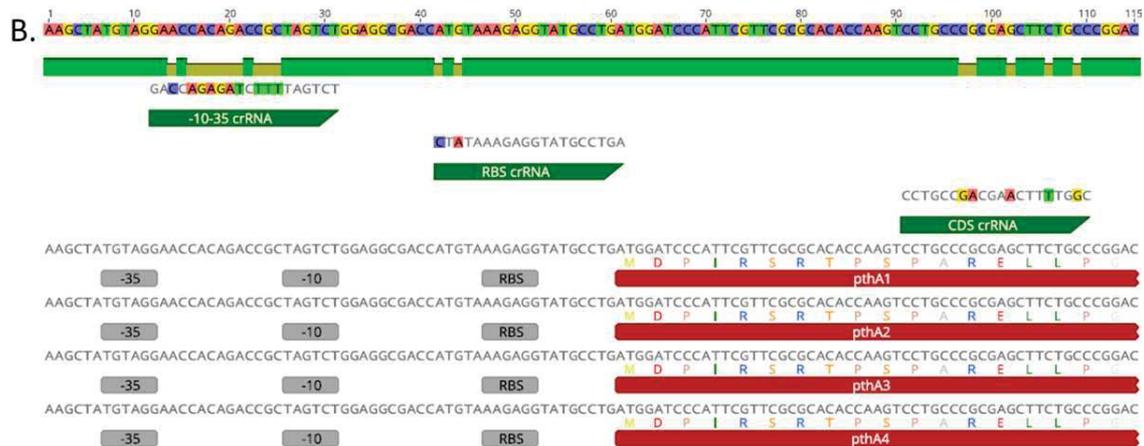


879

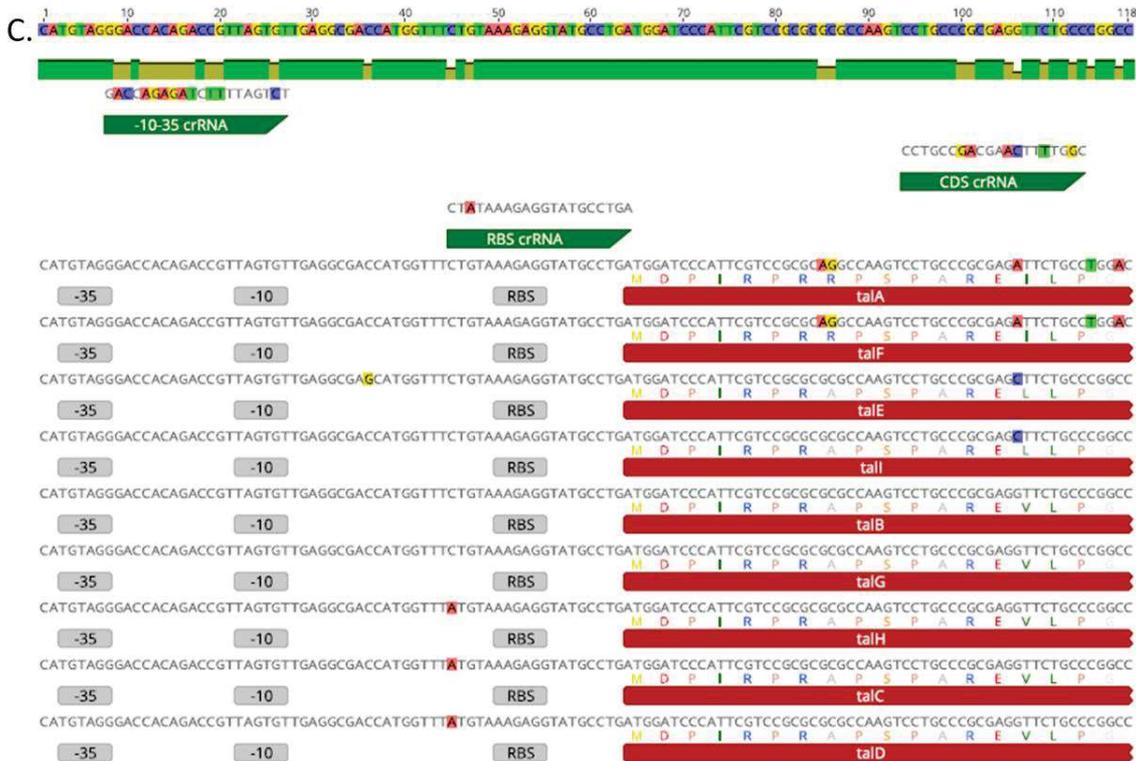
880 Figure S1. Alignment of eight representative *Xpm* TALE gene variants and designed crRNAs. Green
 881 arrows indicate the crRNA (-10-35, RBS, and CDS), with the arrowhead pointing to the PAM. Red
 882 arrows indicate the first 55 nucleotides of the TALE coding sequences; translations are indicated
 883 above each arrow. Gray boxes indicate the position of putative -35 and -10 promoter elements, and
 884 the ribosome binding site (RBS). The nucleotide sequence of each TALE gene is shown above each
 885 translation, mismatches to a global consensus are highlighted in colors.



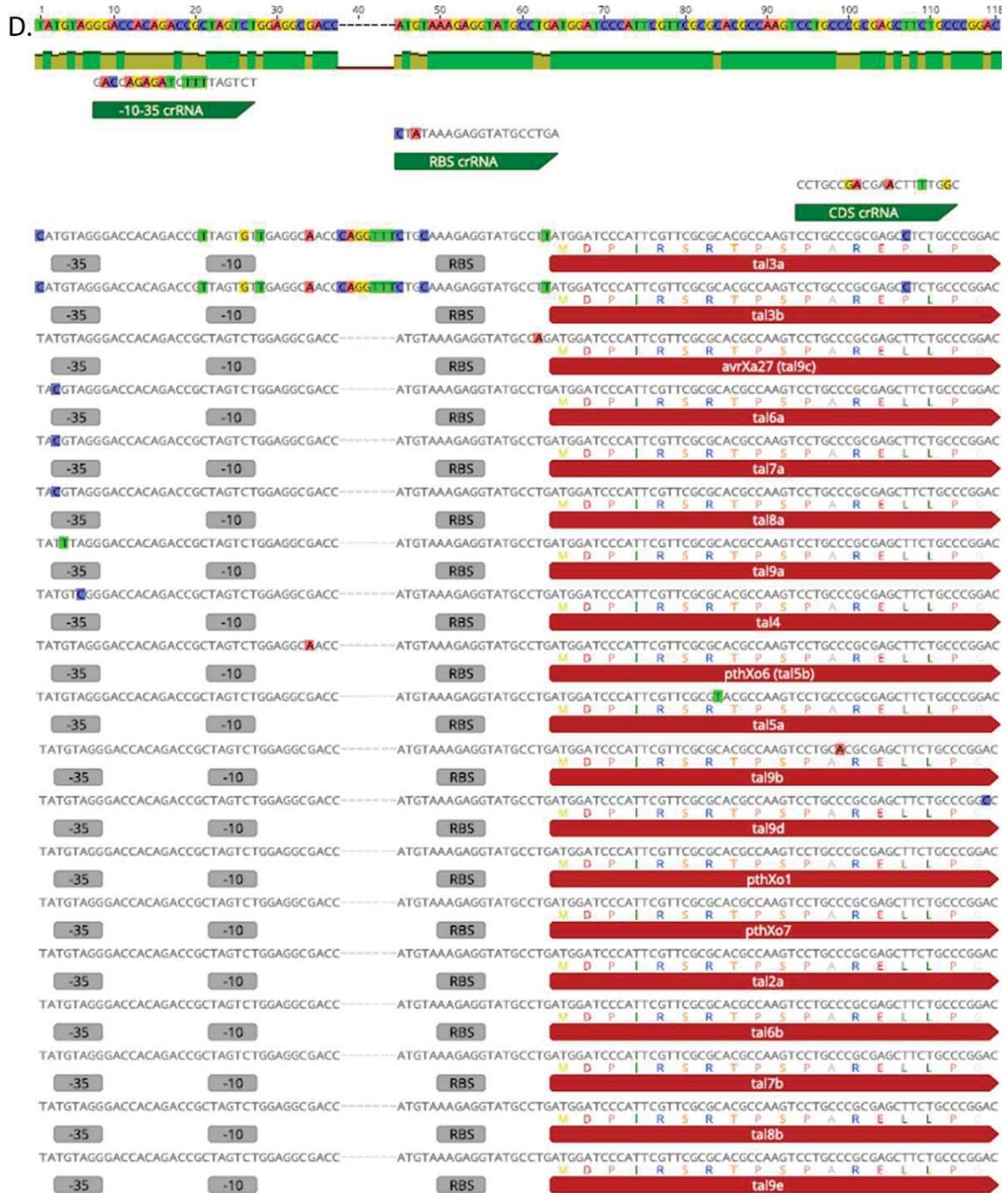
886



887



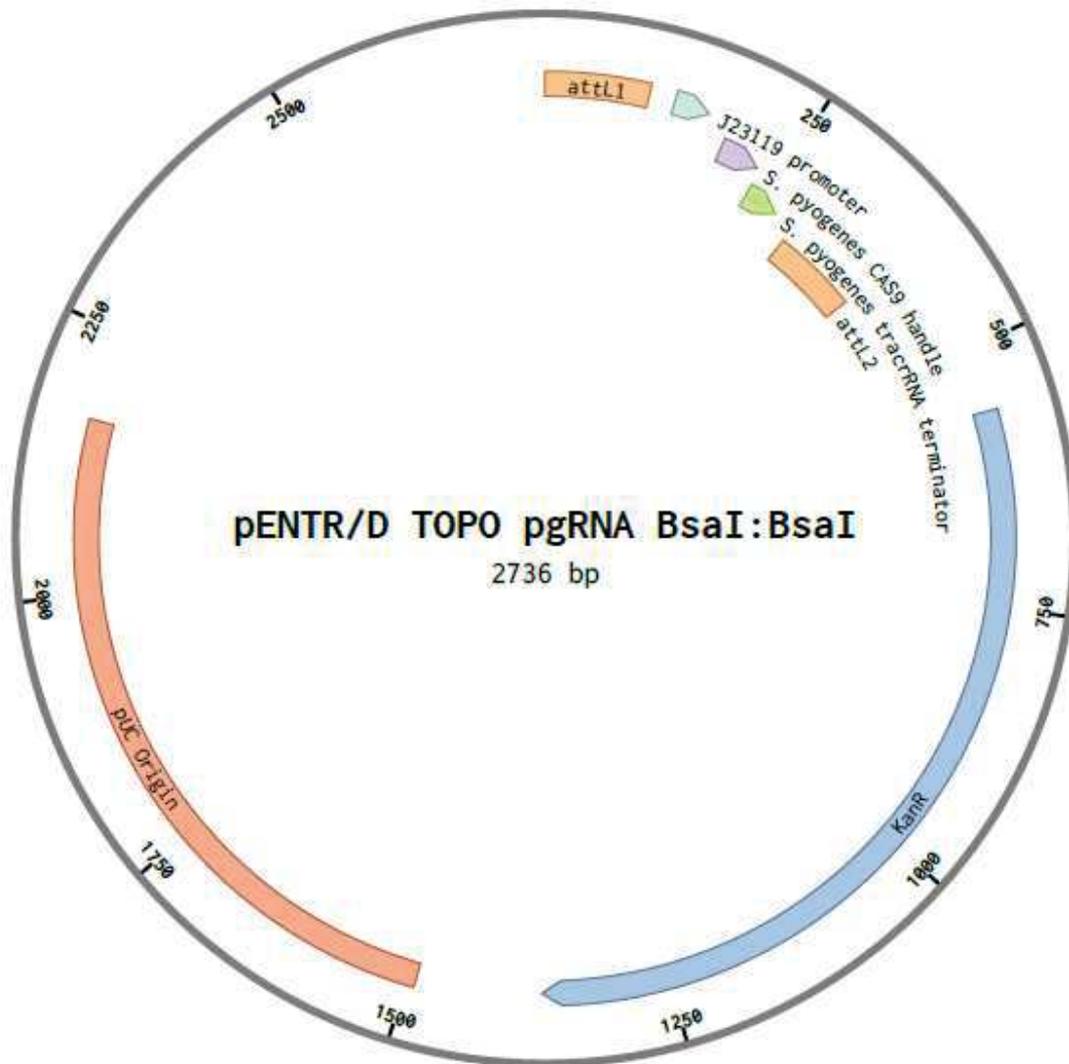
888



889

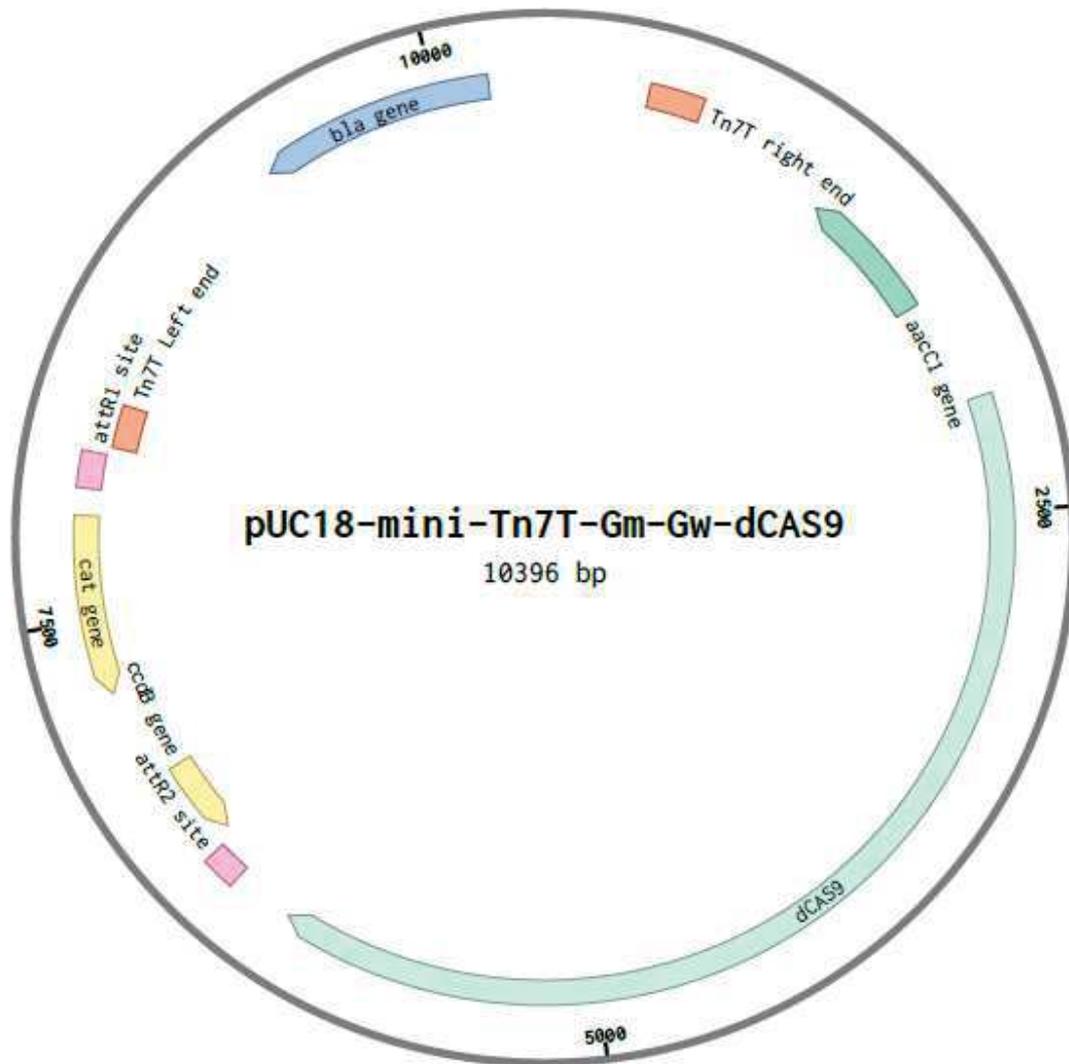
890 Figure S2. Alignment of *TALE* genes from the four tested *Xanthomonas* pathogens and crRNAs
 891 designed for *Xpm*. Green arrows indicate the crRNA (-10-35, RBS, and CDS), with the arrowhead
 892 pointing to the PAM. Red arrows indicate the first 55 nucleotides of the *TALE* coding sequences;
 893 translations are indicated above each arrow. Gray boxes indicate the position of putative -35 and -
 894 10 promoter elements, and the ribosome binding site (RBS). The nucleotide sequence of each *TALE*
 895 gene is shown above each translation, mismatches to a global consensus are highlighted in colors.

896 A. TALE genes from *Xc* CFBP4642. B. TALE genes from *Xcc* IAPAR306. C. TALE genes from *Xoo* MAI1.
897 D. TALE genes from *Xoo* PXO99^A.



898

899 Figure S3. pENTR/D TOPOpgRNA BsaI:BsaI plasmid map. pUC origin is shown as a red box, *attL1* and
900 *attL2* regions are shown as orange boxes, and the *kanR* gene is shown as a blue arrow. The J23119
901 promoter (light blue arrow), the dCAS9 handle (violet arrow) and the terminator (light green arrow)
902 are located between the *attL* regions. The sequence located between the promoter and the handle
903 allow the directional cloning of the crRNAs by using *BsaI* restriction enzyme. The plasmid map was
904 generated with Benchling.com.



905

906 Figure S4. pUC18-min-Tn7T-Gm-Gw-dCAS9 plasmid map. *attR1* and *attR2* sites are depicted as pink
 907 boxes, while Tn7T left and right borders are showed as red boxes. Left and right borders contain the
 908 *aacC1* gene (green arrow), the *dCAS9* gene (blue arrow), and the region comprised between the two
 909 *attR* sites. The *attR* sites flank the *ccdB* and *cat* genes (yellow arrows), which will be exchanged by
 910 the sgRNA scaffold. The plasmid backbone also carries the *bla* gene (dark blue arrow), which codes
 911 for a beta-lactamase.

912 **Supplementary tables**

913 Table S1. Predicted targets for *Xc* CFBP4642 TALEs that were also validated for Xam668-cassava
 914 interaction in previous studies (column Ref). The first six columns show the TALVEZ results for *Xc*
 915 TALEs, while the following four show results for the equivalent assessment for *Xpm* TALEs. TALE IDs
 916 were simplified for readability, gene ID corresponds to the cassava gene annotation version 6
 917 provided by Phytozome (<https://phytozome.jgi.doe.gov/>), rank indicates the position of the
 918 prediction according to the score, score corresponds to TALVEZ score of the TALE-EBE relationship,
 919 dist to TXS is the distance from the EBE to the translational start site (start codon).

TALVEZ output for Xc CFBP4642 TALEs						TALVEZ output for Xam668 TALEs				Annotation	Ref
TALE ID	Gene ID	Rank	Score	Distance to TXS	EBE sequence	TALE ID	Score	Distance to TXS	EBE sequence		
TALE15 A	Manes.16G10490 0	83	12.6	492	CCCTCTCCCATCTCCC	TAL14	6.3	208	TATGACCTGATCACT	Proprotein convertase subtilisin/kexin	(Cohn et al., 2016)
TALE16	Manes.01G05890 0	40	16.1	995	CCCCTCCCCCTCCCTC	TAL15	9.3	128	TATAACCCACCCCCA	Ring finger domain-containing	(Cohn et al., 2014)
TALE16	Manes.01G05890 0	140	14.6	991	TCCCCCTCCCTCCCAA	TAL15	9.3	128	TATAACCCACCCCCA		
TALE17	Manes.06G13440 0	11	13.4	866	TGCATTTCTAAGCTCGT	TAL13	nd	nd	nd	Cytoplasmic dynein light chain	(Cohn et al., 2014)
TALE23	Manes.06G12340 0	43	15.2	110	TATAAACGCTTCTCGCCATCCAT	TAL20	12.9	110	TATAAACGCTTCTCGCCAT C	SWEET transporter	(Cohn et al., 2014)
TALE23	Manes.15G02680 0	150	13.3	369	CACCCACGATCTAACCCACCTTT	TAL14	10.8	94	TATAATCATATCGAT	Membrane-associated kinase regulator 1-related	(Cohn et al., 2014)

920

General discussion

Xpm TALE diversity

As evidenced in chapter one, there is a considerable TALE diversity among *Xpm* populations with an expanding trend in terms of variants, but a preservation of the predicted affinity towards a reduced group of theoretical targets. This implies that evolution is constantly acting on *Xpm* TALEs, but they might be under a positive selection constrained by DNA affinity and target function. TALome composition, TALE genomic context, and the phylogenetic relatedness between *Xpm* and other pathovars like the common bean pathogens *X. phaseoli* pv. *phaseoli* (*Xpp*) and *X. citri* pv. *fuscans* (*Xcf*) suggest that evolutionary history of these TALEs and *Xpm* ones might be related. In their study, Ruh and coworkers (Ruh et al., 2017) describe the TALomes of *Xpp* and *Xcf*, highlighting trends similar to the ones found here: an expanding number of variants that preserve affinity to a small group of targets, a high prevalence of plasmidic TALE genes, and some transposon-mediated TALE insertions into the chromosome. According to their study, a first ancestor TALE was vertically inherited upon the split between the *X. oryzae* species and the *X. axonopodis* species complex, while there was an important second event of TALE divergence in the *Xpm*, *Xpp* and *Xcf* clade after the split from *X. euvesicatoria* species. Horizontal Gene Transfer (HGT) is crucial to TALE evolution and dispersion, and it is suggested that TALEs were transferred from the ancestors of *X. axonopodis* to the *X. citri* species through this mechanism (Ruh et al., 2017). Likewise, Denancé and coworkers described the distribution and phylogenetic relationships of TALEs in the *Xanthomonas campestris* species, with an emphasis on *Xanthomonas campestris* pv. *campestris* (*Xca*). *Xca* TALEs seem to be predominantly vertically inherited, and evidence suggests that TALomes in this pathovar emerged from recombination of two ancestral TALE genes. *Xca* TALEs are scattered among the genome and can be both plasmid-borne and chromosomal. Moreover, the transposon-associated elements in the vicinity of TALE genes are indicators of an evolutionary history characterized by TALE mobilization (Denancé et al., 2018).

It would be interesting to infer the evolutionary history of TALEs in *Xpm*. To that end we are sequencing the genome of several *Xpm* strains that were isolated from geographically distant places, different farming contexts and time periods (see Figure 5). Through long read sequencing technologies, it will be possible not only to compare TALomes, but their genomic contexts, and to make inferences to compare neutral genetic markers to non-neutral markers such as effectors. Besides the evolutionary history of TALEs in *Xpm*, two other related challenges will be of major interest in the future: assessing i) the *Xpm* population genetic structure and evolution of TALEs at the level of an epidemic; and ii) the evolutionary dynamic of TALEs in response to host adaptation (recombination, loss or gain of gene, etc.). For the first goal, our team has several suitable resources and tools, like a vast collection of temporally and geographically diverse *Xpm* strains, including population sets (>15 isolates/plot) sampled from devastated fields in various agro-ecological contexts; the recently developed tool to assess *Xpm* neutral genetic diversity through Multilocus variable-number tandem repeat (VNTR) analysis (MLVA) (Rache et al., 2019); and the PCR-based TALE assessment tool developed by Escalone and coworkers (Escalon, 2013) that I adapted and optimized for *Xpm* during this thesis. In the rice pathogen *Xoo*, the comparative analysis of the TALome structure of strains from populations sampled in infected fields and their MLVA patterns

has shown that, in this pathogen, TALomes tend to evolve faster than the studied VNTR loci. Interestingly, it was reported that several strains with different TALome patterns could co-exist within one field, and that signs of loss or gains of *TAL*E could be evidenced in a population level, thus highlighting the strong evolutionary dynamic of TALomes in *Xoo*. This may highlight the adaptation through the loss of genes that have avirulence effects or whose impacts on the disease are negligible (Diallo A et al., *In prep*).

It is believed that *Xpm* origin and adaptation to cassava is geographically linked to the origin of cassava *per se* (Bart et al., 2012; Verdier et al., 2004). The high neutral genetic diversity seen in *Xpm* populations isolated from South America would indicate a longer period of coevolution between the host and the pathogen, and the reduced neutral diversity observed among for instance African isolates support these observations (Verdier et al., 2004). To gain insights into the origin of *Xpm* and adaptation to cassava as a host, we have performed samplings within the Amazonian region (the theoretical geographic origin of cassava) and collected several *Xpm* strains from cassava fields in the middle of the primary forest (Szurek et al., unpublished, see Figure 5). These fields have two important characteristics: the propagative material exchange with regions out of the Amazonia is potentially reduced (therefore the linked pathogen exchange may also be impacted), and the permaculture schemes and high crop diversity within and around cassava plants differs greatly from conventional farming systems. These factors may be advantageous for our studies, since they may help to analyze characteristics of a well-established endemic population of *Xpm*, including its TALome.

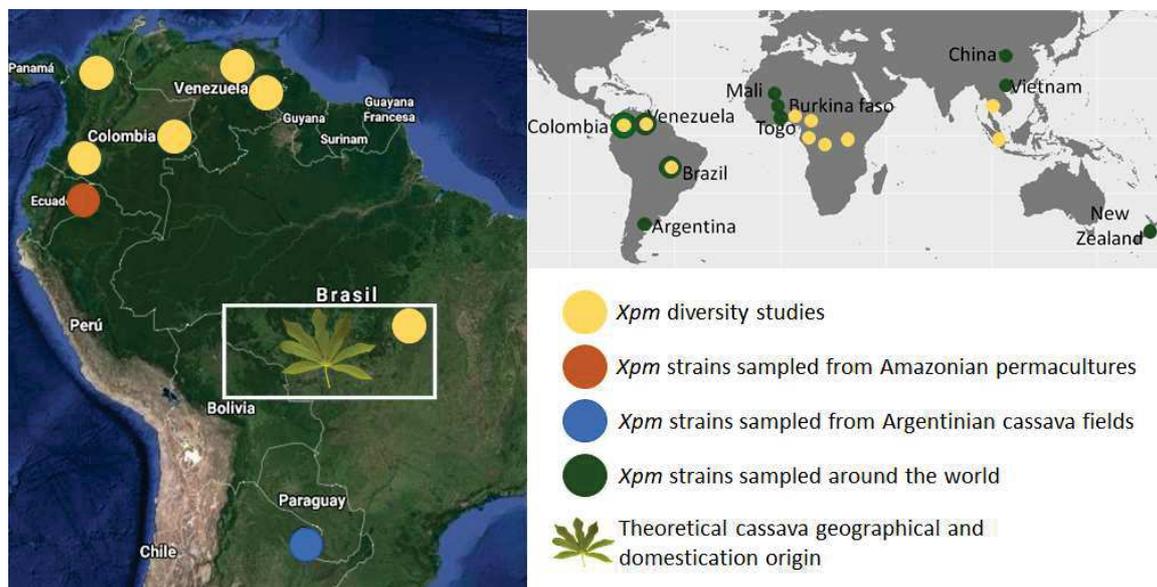


Figure 5 Geo-localization of regions where *Xpm* strains used for experimental assessments have been isolated.

Left panel shows part of the South American continent. The hypothetical cassava origin in the Amazon basin is indicated in the center of the map. Golden circles show the regions of origin of *Xpm* isolates used for diversity studies in literature. The orange circle shows the Amazonian region where our group has recently sampled *Xpm*. The blue circle indicates the region where several *Xpm* strains were isolated from cassava crops in a recent survey performed by our team in the frame of this

thesis. The right panel shows a world map segment indicating the geographical origin of *Xpm* strains assessed for diversity in literature (golden circles) and the *Xpm* strains that we profiled for future diversity (including TALE diversity) studies.

In the second chapter of this thesis, we demonstrated that although the functional diversity among TALEs is reduced, there are alternative and convergent mechanisms to promote disease in cassava. Our data suggest that some TALE15 variants are responsible for activation of the potential *S* gene *MeSWEET10e*, potentially as a consequence of direct upregulation of a transcription factor. Phylogenetically, the TALEs associated to strains that use these redundant mechanisms (activation of *MeSWEET10a* or *MeSWEET10e*) seem to come from different sources, but more work is needed to understand where these TALEs came from, and if and how they are distributed in modern *Xpm* populations. This first report of evolutionary convergence among *Xpm* TALEs raises important pragmatical questions as to the distribution of these TALEs within *Xpm* populations and the occurrence of the two underlying strategies in the field. From the information that we have so far, TALEs activating *MeSWEET10e* are exclusively present in *Xpm* strains that were isolated c.a. 40 to 60 years ago, but their distribution is not geographically restricted (Argentina, China, and Venezuela). It is tempting to conjecture (ignoring the role of other pathogenicity factors and potential sampling biases) that the wide distribution of *MeSWEET10a*-upregulating TALEs among more recently isolated strains would reflect a positive selection towards this mechanism. The activation efficiency of the *S* gene, where one TALE directly binds the promoter of the *S* gene to activate it, could be the reason for this selection. Nevertheless, other factors affecting population dynamics or sampling biases may explain the relative reduced number of strains with *MeSWEET10e*-activating TALEs among recently isolated strains. In an attempt to track the occurrence of this alternative virulence strategy, I carried out in 2019 samplings in cassava fields located in the North of Argentina, the country where CIAT1241 was isolated in 1984 (see Figure 5). Although the occurrence of CBB in the Argentinian cassava fields was considerably low, we managed (in collaboration with Martín Domínguez from the Instituto Nacional de Tecnología Agropecuaria) to isolate over 60 isolates of *Xpm* strains from symptomatic cassava plants. An objective for the team in the near future, will be to assess the TALomes of these isolates to check if *MeSWEET10e*-activating TALEs are present in the surveyed fields. Overall, this report illustrates well the importance of sugar transport in host susceptibility to *Xanthomonas* species overall and to *Xpm* in particular. It is worth noting that the results presented here indicate that CBB is the second example, after bacterial leaf blight in rice, where several clade-III SWEETs are mobilized to promote pathogenicity. In line with this, we cannot rule out that other clade-III SWEETs than *MeSWEET10a* and *MeSWEET10e* may be manipulated by *Xpm*, another interesting subject to address in this pathosystem.

Susceptibility genes

Xc and *Xpm* provide an interesting system to compare vascular and non-vascular lifestyles, but also host adaptation, since *Xc* is phylogenetically distant from *Xpm* (see Figure 3). Due to its restricted distribution, it is believed that *Xc* recently adapted to cassava in Africa and that this plant is not its primary or main host (Hayward, 1993). This study constitutes the first report of *Xc* TALE sequences and suggests the potential TALE-mediated exploitation of an *S* gene that is also targeted by *Xpm*. Interestingly and in contrast to *Xpm*, all *Xc* TALE genes are inserted and scattered into the chromosome. Although more work is needed to test these hypotheses, this organization suggest that *Xc* TALEs were vertically inherited from the ancestors that gave rise to the clades of *X. oryzae*,

X. axonopodis, and the *campestris/cassavae/maliensis/arboricola/hortorum/gardneri* clade. Knowing this, it is striking that TALEs from *Xc* and *Xpm* have evolved totally independently since there is no evidence of HGT (data not shown) to target the same susceptibility gene (*MeSWEET10a*) and by using the exact same EBE (experimental confirmation of *MeSWEET10a* induction by *Xc* is still pending). This could be explained by several hypothesis: an ancestral capacity to bind plant TATA boxes that was already present in the genes that gave rise to TALomes in *Xpm* and *Xc*, a biological factor related to the host genome that positively selects TALEs with this specific affinity, or by mere chance. This scenario presents a unique opportunity to study evolutionary dynamics and constraints to TALE selection imposed by the host biology, and how TALEs can help “distantly-related” pathogens to adapt to the same host.

Under the light of the hypothesis that *Xc* started coevolving with cassava after the introduction of this crop to Africa in the XVI century (Hayward, 1993; Jones, 1959), it is tempting to speculate that host adaptation was promoted by the evolved TALE variants targeting *MeSWEET10a*. This fact would also imply that such TALEs became fixed (if they did not exist before) in less than 500 years of coevolution, demonstrating the versatility and high impacts of TALEs as pathogenicity determinants. Regarding this, Teper and Wang demonstrated through an experimental evolution approach that only nine serial disease cycles (colonization of a new susceptible host by a pathogen derived from a diseased plant) are necessary to rearrange the central repeat region of TALEs to overcome mismatching interactions between TALEs and S gene EBEs (Teper and Wang, 2021). Clearly, evolution of completely new TALE variants targeting relevant genes in the context of adaptation might take much more time. Hence, the *Xc*-cassava pathosystem represents a good opportunity to understand a recent adaptation of *Xanthomonas* to a dicot host as a non-vascular pathogen and TALEs implication. Adaptation has been extensively studied for *Xanthomonas citri* pv. *citri*, suggesting that the Fabaceae colonizers *X. axonopodis clitoriae* and *X. axonopodis cajani* gave rise to *X. citri*, and a host jump was implicated in the onset of *Xcc*. It was estimated that this pathovar was originated after the last glacial maximum, and diversification burst of the pathovar coincides with the spread of Citrus cultivars c.a. 2500-3500 years ago (Patané et al., 2019).

Provided that further experiments validate *MeSWEET10a* as an S gene for *Xc*, it is worth noting that the common activation of a *SWEET* gene by *Xc* and *Xpm* suggests that it is not a lifestyle determinant. Although more studies are needed to clearly demonstrate the involvement of *MeSWEET10a* in *Xc*-induced susceptibility, this would be the first report of a TALE of a non-vascular pathogen targeting a clade-III *SWEET* gene. One could speculate that the role of *SWEETs* is more important from the pathogen penetration through stomata (a feature shared by *Xc* and *Xpm*) and during mesophyll colonization, allowing for the early step multiplication of bacteria before reaching the main vessels of the leaf. Although it is clear that susceptibility genes shape the pathogen behavior (for example, *CsLOB1* and its linked canker-like phenotype that promotes the dissemination of the pathogen), a single genetic determinant has recently been found to determine the lifestyle (vascular vs non-vascular) of plant pathogenic bacteria. *CsbA*, a cell wall-degrading cellobiohydrolase, seems to be responsible for the capacity of *Xanthomonas* pathogens to leave the parenchyma and enter the vessels. Though it is not the sole factor responsible for this lifestyle, its presence can turn the non-vascular wheat pathogen *Xanthomonas translucens* pv. *undulosa* into a pathogen that migrates from the inoculation site to distal tissues through the wheat leaf vessels (Gluck-Thaler et al., 2020). In line with these findings, our preliminary results show that *Xc* CFBP4642 lacks a functional *cbsA*,

while *Xpm* strains have a unique and complete chromosomal version of this gene (data not shown). Although deeper studies are needed, especially in *Xc*, understanding the TALE biology of these two distant pathogens could ease the identification of TALE-related susceptibility hubs, and provide valuable loss-of-*S* alleles to promote broad-spectrum resistance against both CBB and CBN (even if disease co-occurrence has never been documented so far).

Assessment of *Xpm* diversity based on *TALE* profiles resulted in the discovery of strains that exploit a second clade-III *SWEET* gene to cause susceptibility (see Figure 6). This result is relevant for the design of resistant varieties based on TALE biology with the aim of deploying loss-of-susceptibility alleles through genome editing. A proof of concept for this has been carried out in rice and citrus to fight bacterial leaf blight caused by *Xoo* (Oliva et al., 2019) and citrus canker caused by *Xcc* (Jia et al., 2017; Peng et al., 2017), respectively. In both cases a detailed knowledge about the diversity of the pathogenicity mechanisms is at the center of the strategy aiming at the safe the deployment of durable resistance. This is exemplified by the fact that, although the genome-edited *Xoo*-resistant mega varieties seem resistant to most of the *Xoo* diversity representatives, African *Xoo* strains harboring the *OsSWEET14*-activating TaC effector use different strategies to circumvent the loss of susceptibility generated by edition, and consequently still cause disease albeit moderately (Blanvillain-Baufumé et al., 2017; Oliva et al., 2019). Currently, researchers are working to understand this more complex virulence strategy in order to engineer the appropriate factors that promote disease in rice. What is known is that *OsSWEET13* and *OsSWEET14* may be both implicated in host susceptibility, since simultaneous knockout of these genes results in resistance against the African strain AXO1947, yet *OsSWEET13* transcriptional activation is not evident (Eom et al., 2019). In the case of the alternative *S* gene *MeSWEET10e*, our data suggest that transcriptional activation is indirect, potentially through the action of a TALE-upregulated transcription factor. It has been demonstrated that several phytopathogens are able to upregulate host *SWEETs* during infection, but the underlying mechanisms are mainly unknown (Devanna et al., 2021). Recently, Prior and coworkers described the involvement of the transcription factor bZIP11 in the transcriptional regulation of several nutrient transporters, including *SWEETs*, in *Arabidopsis thaliana*. bZIP11 is somehow manipulated by the bacterial pathogen *Pseudomonas syringae* pv. *tomato* strain DC3000 through its type III effectors, potentially increasing the mRNA transcript abundance during infection. This upregulation leads to a potential shift from a non-secretory cell to a secretory cell with an increased number of transporters, which in turn cause the apoplastic accumulation of nutrients like sugars and some amino acids. Since TALE-based manipulation of TF transcription rates has already been documented and characterized as susceptibility genes for other *Xanthomonas* species (Hu et al., 2014; Kay et al., 2007; Peng et al., 2019; Schwartz et al., 2017), it would not be surprising that *Xpm* can manipulate TFs involved in carbon metabolism. Further analyses (see Figure 6) would narrow down the candidate TFs through cross-referencing expression data of tissues (Wilson et al., 2017) where *MeSWEET10e* is upregulated under normal conditions. Additionally, our group, in collaboration with Paul Chavarriaga (CIAT, Colombia), is generating transgenic cassava plants where *MeSWEET10a* and *MeSWEET10e* are knocked-out. They will be useful in the further characterization of this cassava susceptibility factor.

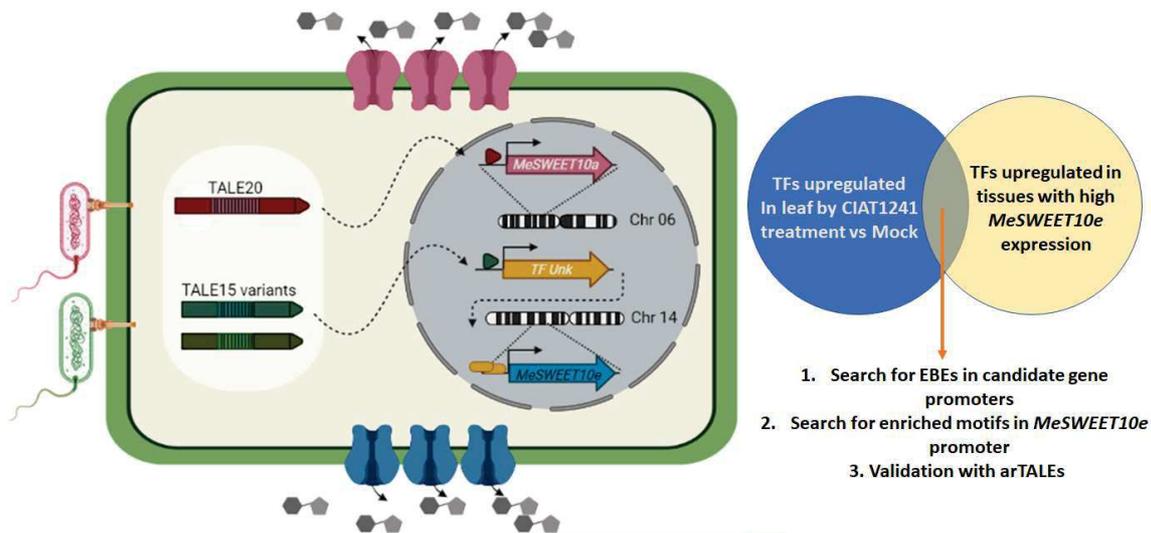


Figure 6 Theoretical TALE-mediated activation of MeSWEETs.

Model of the potential mechanisms of transcriptional activation of *MeSWEET10e* and *MeSWEET10a* by *Xpm* TALEs, and a proposed strategy to find the missing elements on the activation chain. The figure of the left was created with BioRender.

Another remarkable feature arising from the diversity studies is the wide distribution and conservation of other TALE variants like TALE22 and TALE13. However, previous reports highlighted them as dispensable factors for *Xpm* virulence (Cohn et al., 2014). TALE-related virulence and aggressiveness have been traditionally investigated during the process of colonization of the mesophyll apoplast and following the movement of the pathogen in the leaf (Castiblanco et al., 2013; Cohn et al., 2016, 2014), obviating the implications of lacking TALE variants during stem vessel colonization and systemic movement. Mutka and coworkers have established new methodologies to measure these pathogenic capabilities through luminescence real-time monitoring (Mutka et al., 2016), while the classical method to calculate Area Under Disease Progress Curve (AUDPC) is another resource that has not been used to account TALE impacts on cassava systemic disease. Measuring these impacts is important for this pathosystem, since plant defenses (and the associated ways used by bacteria to bypass them) in petioles and stems are substantially different from the poorly understood cellular defense in the mesophyll (Boher et al., 1995; Kpémoua et al., 1996). It is appealing to hypothesize that TALEs could have roles that would also increase the fitness of bacteria during stem colonization, since they are not only restricted to the vessels, but also interact from intercellular spaces with the parenchymal cells of xylem and phloem (Kpémoua et al., 1996). In this setting, several physiological changes were noted, such as a shift in cellular growth patterns to create tyloses and occlude vessels, or the deposition of large amounts of pectic elements to block and reinforce cellular structures (Kpémoua et al., 1996). Hence, it is not unreasonable to think that TALEs could undermine these responses through transcriptional reprogramming. Therefore, more work is needed on plant-pathogen interaction during this disease phase, which potentially would explain the conservation of some TALE variants.

CRISPRi tool

Implementation of the silencing platform based on CRISPRi for *Xanthomonas* species represents a powerful tool for the study of the biology of this pathogen. When comparing the time/effort invested to achieve gene knock-outs by double homologous recombination (DHR), it is clear that CRISPRi is easier, faster and leads to similar results. Moreover, the possibility to easily multiplexing sgRNAs to simultaneously target several genes has been already documented in analogous platforms in bacteria (Afonina et al., 2020; Dong et al., 2017; Myrbråten et al., 2019), reaching silencing levels with physiological consequences near to the ones seen in deletional mutants. Application of this technology in high throughput experimental designs (settings that cannot be established by traditional DHR or directed insertional mutagenesis) has allowed the discovery of genes with key roles in, for example, pathogenic interactions (Todor et al., 2021). This is the case of the interaction between *Mycobacterium bovis* BCG (Bacillus Calmette-Guérin) and human phagocytic cells, where a high-throughput hybrid platform (CRISPR and CRISPRi) allowed the identification of more than 100 genes associated with phagocytic cell survival. Among these, components of the type I interferon and aryl hydrocarbon receptor signaling pathways were identified as potential pharmacological targets to improve the tuberculosis treatments (Lai et al., 2020). Although the design conceived in this thesis was intended to study plant-pathogen interactions through the silencing of gene families, the fact that we validated the function of the *Streptococcus pyogenes* dCas9 in several pathovars of *Xanthomonas* imply that optimization and alternative designs based on this protein are feasible (see Figure 7).

In view of the recent advances of the CRISPR technology, it is predictable that base editors and prime editing will be developed for *Xanthomonas* (see Figure 7). Base editors use a modified version of Cas9 that nicks the DNA after recognition (nCas9) and unwinds local DNA. As the nCas9 is fused to a cytidine deaminase domain or an adenosine deaminase domain, nCas9 localization and DNA unwind allow the fused domain to convert cytosine to uracil, or adenine to guanine, respectively. These conversions are then resolved by the DNA repair systems, which lead to fixation of the targeted point mutation (Kondrateva et al., 2021). Recently, a system based on a dCas9 fused to a *Petromyza marinus* cytidine deaminase was used in the phytopathogens *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. This resulted in efficient mutation of several targets through creation of premature stop codons and low off-target rates, which are thought to be more related to the base editor than to the dCas9 activity (Rodrigues et al., 2021). Abdullah and coworkers reviewed the base editor approaches successfully applied in bacteria (Abdullah et al., 2020). Prime editing is a novel technique that allows targeted DNA editing that not depends on errors of the DNA repair methods. In this technology, a nCas9 is fused to a reverse transcriptase that uses a prime editing guide RNA (pegRNA) with a dual purpose: to guide the complex to the target, and to serve as reverse-transcriptase template. The portion of the pegRNA that provides the template can be designed to obtain the desired insertion at the position of the DNA nick. Therefore, after positioning and DNA nicking, the complex synthesizes a single stranded DNA flap, resulting in an edited and an unedited strand. Equilibration between these two strands occurs and it is followed by the cleavage of the unedited flap, ligation, and DNA reparation (Anzalone et al., 2019). This system has been successfully applied in mice and plants (reviewed by (Abdullah et al., 2020)), but so far, it has not been tested in bacteria.

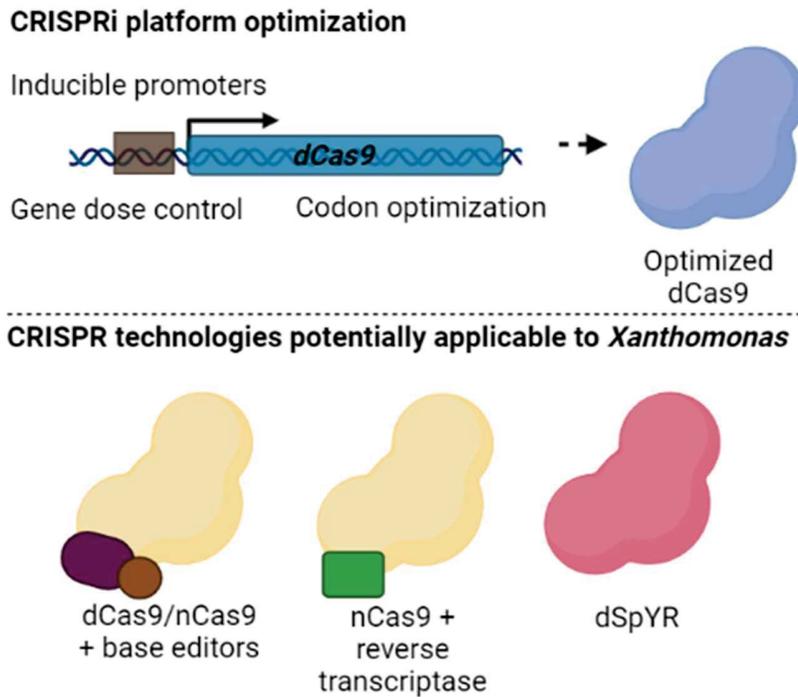


Figure 7 Theoretical optimization points and technological shifts foreseen for CRISPR-based platforms in *Xanthomonas*.

The upper panel shows three proposed ways to optimize the CRISPRi platform present in this study. Inducible promoters would allow the study of genes that play key roles in pathogenicity and central metabolism. Codon optimization and gene dose control would potentially maximize the on-target effects, and control some of the “bad-seed” problems that could be encountered. The lower panel shows the CRISPR technologies that could be applied in the future to *Xanthomonas* species. The use of base editors and prime editing would allow targeted mutation of genes to create knock-downs or intended insertions, while the use of a dead version of the PAM-less SpYR Cas9 protein would increase the possibilities to silence genes in *Xanthomonas*. The figure was created with BioRender.

The implementation of this tool in *Xc* led to unveil the potential use of *MeSWEET10a* as a key susceptibility factor during pathogenesis. This is a clear demonstration of the effectiveness of this platform in the study of gene families and, specifically, TALEs in the pathogen-host interaction. *In silico* testing of the sgRNAs developed for *Xpm* in the genomes of at least one member *X. oryzae*, *X. axonopodis*, and the *campestris/cassavae/maliensis/arboricola/hortorum/gardneri* clades indicates that the TALE RBS sequence is highly conserved among these pathovars, indicating that this is a feature from the ancestor gene that gave rise to TALEs in the *Xanthomonas* genus. Although more comprehensive analyses are needed, if true, this would be a significant discovery that would allow the silencing of most of the TALE variants in any *Xanthomonas* species. In this scenario, the CRISPRi tool developed for TALEs could be used as a screening methodology to assess TALE roles in the pathogenicity of poorly-studied pathovars. Furthermore, other genes, different from TALEs, implicated in the pathogenicity of *Xanthomonas* could also be studied using this silencing methodology. For example, Medina and coworkers assessed the involvement of the Xops of *Xpm* in

the cassava immunity disturbance during compatible interactions. Results highlight that some of the Xops, like XopN and XopQ, have redundant functions, and related phenotypes can only be observed in the double knock-out mutants (Medina et al., 2018). By using the proposed approach, silencing these two genes would be achieved after just one transformation event, which significantly decreases times and efforts, while avoid the *in vitro* serial passages (at least 8 for a double mutant) required to obtain these mutants.

Conclusions and perspectives

In this thesis three main axes were studied: *TALE* diversity among *Xpm* populations, alternative susceptibility mechanisms used by *Xpm*, and development of a new silencing platform to study gene families, including *TALEs*, in *Xanthomonas*. Through the assessment of TALome diversity among Colombian *Xpm* isolates, we characterized 13 new *TALE* variants, including *TALE22D* which represents a new variant with 22 RVDs capable of activating the susceptibility gene *MeSWEET10a*. This survey also revealed that, in addition to *TALE14* and *TALE20*, the variants *TALE13* and *TALE22* are widely distributed among isolates, suggesting that they could have important roles in pathogenesis. Further characterization of the TALomes resulted in the discovery of *Xpm* strains that do not carry *MeSWEET10a*-inducing *TALE20* variants yet retaining full virulence capacity. Phylogenetic analyses indicate that *TALE* variants coming from these strains are considerably different from those present in isolates that were collected more recently around the world. Therefore, data suggest that *Xpm TALEs* come from at least two different ancestor genes, and the virulence-promoting strategies associated to their biological function arose independently.

After discovering three *Xpm* strains that do not rely on *MeSWEET10a* activation to cause disease, transcriptomics analyses showed that these isolates are able to upregulate a second clade-III sugar transporter of the *SWEET* family: *MeSWEET10e*. Our data suggest that this activation is mediated by *TALEs* in an indirect fashion, potentially through direct transcriptional activation of a master regulator. Although the molecular mechanism underlying its activation is not well understood so far, we demonstrated that *MeSWEET10e* can act as a susceptibility gene similarly to *MeSWEET10a*. A proposed *in silico* pipeline in conjunction with *in planta* validations would potentially identify the missing factors that lead to transcriptional activation of this new *S* gene.

The difficulties faced during the development of this thesis, when DHR was attempted to obtain *TALE* mutants in some *Xpm* strains, led us to the implementation of a silencing tool based on CRISPRi. The silencing platform is based in the *Streptococcus pyogenes* dCas9 and results show that only one sgRNA is enough to silence up to five *TALEs* in one *Xpm* strain. *In silico* analyses demonstrated that the RBS within the 5'UTR region of *TALEs* is highly conserved among the *Xanthomonas* species and allowed the application of the system designed for *Xpm* to other species like *Xoo*, *Xcc*, and *Xc*. In these pathogens, one sgRNA was enough to silence up to nine *TALEs*. Furthermore, implementation of this technology in *Xc* coupled to genome sequencing, *in silico* target prediction, and virulence assays indicate that *TALEs* are important for *Xc* pathogenesis, potentially through the activation of the known *S* gene *MeSWEET10a*. Although some improvement is necessary, this tool opens multiple possibilities to facilitate the study of gene function in the *Xanthomonas* genus. Optimization and other related technologies are foreseen after this validation of the utility of CRISPRi in these phytopathogens.

Finally, we hope that the knowledge generated in this thesis will help to advance in the fight against pathogens through engineered crop protection.

References

- Abaca A, Kiryowa M, Awori E, Andema A, Dradiku F, Moja A, Mukalazi J. 2013. Cassava Pests and Diseases' Prevalence and Performance as Revealed by Adaptive Trial Sites in North Western Agro-Ecological Zone of Uganda. *J Agric Sci* **6**. doi:10.5539/jas.v6n1p116
- Abdullah, Jiang Z, Hong X, Zhang S, Yao R, Xiao Y. 2020. CRISPR base editing and prime editing: DSB and template-free editing systems for bacteria and plants. *Synth Syst Biotechnol* **5**:277–292. doi:https://doi.org/10.1016/j.synbio.2020.08.003
- Affery AM, Abo K, Issa W, Seydou T, Kassi M, Camara B, Kone D. 2017. Assessment of the NECO's Effectiveness against Cassava Bacterial Blight Caused by *Xanthomonas axonopodis* pv. *manihotis* In Côte D'Ivoire. *Int J Agric Res* **12**:190–198. doi:10.3923/ijar.2017.190.198
- Affery AM, Abo K, Seydou T, Boni N, Kone D. 2016. Geographical Distribution and Incidence of Cassava Bacterial Blight (*Manihot esculenta* Crantz) Caused by *Xanthomonas axonopodis* pv. *manihotis* in Two Agro-ecological Zones of Cote d'Ivoire. *Plant Pathol J* **16**:1–11. doi:10.3923/ppj.2017.1.11
- Afonina I, Ong J, Chua J, Lu T, Kline KA. 2020. Multiplex CRISPRi System Enables the Study of Stage-Specific Biofilm Genetic Requirements in *Enterococcus faecalis*. *MBio* **11**:e01101-20. doi:10.1128/mBio.01101-20
- Aiemnaka P, Wongkaew A, Chanthaworn J, Nagashima SK, Boonma S, Authapun J, Jenweerawat S, Kongsila P, Kittipadukul P, Nakasathien S, Sreewongchai T, Wannarat W, Vichukit V, López-Lavalle LAB, Ceballos H, Rojanaridpiched C, Phumichai C. 2012. Molecular Characterization of a Spontaneous Waxy Starch Mutation in Cassava. *Crop Sci* **52**:2121–2130. doi:https://doi.org/10.2135/cropsci2012.01.0058
- Akparobi SO, Togun AO, Ekanayake IJ. 1998. Assessment Of Cassava Genotypes For Resistance To Cassava Mosaic Disease, Cassava Bacterial Blight And Cassava Green Mite At A Lowland And Mid-Altitude Site In Nigeria. *African Crop Sci J* **6**:385–396.
- Allem AC. 2002. The origins and taxonomy of cassava. In: Hillocks RJ, Thresh JM, Bellotti A, editors. *Cassava: Biology, Production and Utilization*. Wallingford, UK: New York: CABI Publishing. hlm. pp. 1–16.
- Alvarez E, Pardo J, Mejia JF, Bertaccini A, Thanh N, Hoat T. 2013. Detection and identification of *Candidatus Phytoplasma asteris*'-related phytoplasmas associated with a witches' broom disease of cassava in Vietnam. *Phytopathogenic Mollicutes* **3**:77. doi:10.5958/j.2249-4677.3.2.018
- Ambe JT. 1993. The effect of planting dates on three cassava diseases in Cameroon. *Int J Pest Manag* **39**:309–311. doi:10.1080/09670879309371811
- 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*. *Plant Cell* **22**:3864–3876. doi:10.1105/tpc.110.078964
- Anzalone A V, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR. 2019. Search-and-replace genome editing without double-strand breaks

- or donor DNA. *Nature* **576**:149–157. doi:10.1038/s41586-019-1711-4
- Arene OB. 1976. Influence of shade and intercropping on the incidence of cassava bacterial blight. Proceedings of an Interdisciplinary Workshop Held at IITA, Ibadan, Nigeria. pp. 28–29.
- Arrieta-Ortiz ML, Rodríguez-R LM, Pérez-Quintero ÁL, Poulin L, Díaz AC, Arias Rojas N, Trujillo C, Restrepo Benavides M, Bart R, Boch J, Boureau T, Darrasse A, David P, Dugé de Bernonville T, Fontanilla P, Gagnevin L, Guérin F, Jacques M-A, Lauber E, Lefeuvre P, Medina C, Medina E, Montenegro N, Muñoz Bodnar A, Noël LD, Ortiz Quiñones JF, Osorio D, Pardo C, Patil PB, Poussier S, Pruvost O, Robène-Soustrade I, Ryan RP, Tabima J, Urrego Morales OG, Vernière C, Carrere S, Verdier V, Szurek B, Restrepo S, López Carrascal CE, Koebnik R, Bernal A. 2013. Genomic survey of pathogenicity determinants and VNTR markers in the cassava bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis* strain CIO151. *PLoS One* **8**:e79704. doi:10.1371/journal.pone.0079704
- Aung K, Jiang Y, He SY. 2018. The role of water in plant–microbe interactions. *Plant J* **93**:771–780. doi:https://doi.org/10.1111/tpj.13795
- Banito A, Kpémoua KE, Wydra K. 2010. Screening of Cassava Genotypes for Resistance to Bacterial Blight Using Strain x Genotype Interactions. *J Plant Pathol* **92**:181–186.
- Banito A, Verdier V, Kpemoua K, Wydra K. 2007. Assessment of major cassava diseases in Togo in relation to agronomic and environmental characteristics in a systems approach. *African J Agric Res* **2**:418–428.
- Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, Kumar L, Moreira LM, Neto JR, Verdier V, Santana MA, Kositcharoenkul N, Vanderschuren H, Gruissem W, Bernal A, Staskawicz BJ. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc Natl Acad Sci U S A* **109**:E1972-9. doi:10.1073/pnas.1211014109
- Bart RS, Taylor NJ. 2017. New opportunities and challenges to engineer disease resistance in cassava, a staple food of African small-holder farmers. *PLOS Pathog* **13**:e1006287.
- Bayer-Santos E, Ceseti L de M, Farah CS, Alvarez-Martinez CE. 2019. Distribution, Function and Regulation of Type 6 Secretion Systems of Xanthomonadales . *Front Microbiol* .
- Bernal-Galeano V, Ochoa JC, Trujillo C, Rache L, Bernal A, López Carrascal CE. 2018. Development of a multiplex nested PCR method for detection of *Xanthomonas axonopodis* pv. *manihotis* in Cassava. *Trop Plant Pathol* **43**:341–350. doi:10.1007/s40858-018-0214-4
- Berthier Y, Verdier V, Guesdon JL, Chevrier D, Denis JB, Decoux G, Lemattre M. 1993. Characterization of *Xanthomonas campestris* Pathovars by rRNA Gene Restriction Patterns. *Appl Environ Microbiol* **59**:851–859.
- Beyene G, Chauhan RD, Ilyas M, Wagaba H, Fauquet CM, Miano D, Alicai T, Taylor NJ. 2017. A Virus-Derived Stacked RNAi Construct Confers Robust Resistance to Cassava Brown Streak Disease. *Front Plant Sci* **7**:2052. doi:10.3389/fpls.2016.02052
- Beyene G, Solomon FR, Chauhan RD, Gaitán-Solis E, Narayanan N, Gehan J, Siritunga D, Stevens RL, Jifon J, Van Eck J, Linsler E, Gehan M, Ilyas M, Fregene M, Sayre RT, Anderson P, Taylor NJ, Cahoon EB. 2018. Provitamin A biofortification of cassava enhances shelf life but reduces dry matter content of storage roots due to altered carbon partitioning into starch. *Plant*

Biotechnol J **16**:1186–1200. doi:<https://doi.org/10.1111/pbi.12862>

- Blanvillain-Baufumé S, Reschke M, Solé M, Auguy F, Doucoure H, Szurek B, Meynard D, Portefaix M, Cunnac S, Guiderdoni E, Boch J, Koebnik R. 2017. Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for *SWEET14*-inducing TAL effectors. *Plant Biotechnol J* **15**:306–317. doi:10.1111/pbi.12613
- Boch J, Bonas U. 2010. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Annu Rev Phytopathol* **48**:419–436. doi:10.1146/annurev-phyto-080508-081936
- 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 (80-)* **326**:1509–1512. doi:10.1126/science.1178811
- Bogdanove AJ, Schornack S, Lahaye T. 2010. TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* **13**:394–401. doi:<https://doi.org/10.1016/j.pbi.2010.04.010>
- Bogdanove AJ, Voytas DF. 2011. TAL Effectors: Customizable Proteins for DNA Targeting. *Science (80-)* **333**:1843 LP – 1846. doi:10.1126/science.1204094
- Boher B, Kpemoua K, Nicole M, Luisetti J, Geiger JP. 1995. Ultrastructure of interactions between cassava and *Xanthomonas campestris* pv. *manihotis*: cytochemistry of cellulose and pectin degradation in a susceptible cultivar. *Phytopathology* **85**.
- Boher B, Nicole M, Potin M, Geiger JP. 1997. Extracellular Polysaccharides from *Xanthomonas axonopodis* pv. *manihotis* Interact with Cassava Cell Walls During Pathogenesis. *Mol Plant-Microbe Interact* **10**:803–811. doi:10.1094/MPMI.1997.10.7.803
- Boher B, Verdier V. 1994. Cassava bacterial blight in Africa: The state of knowledge and implications for designing control strategies. *African Crop Sci J* **2**:505–509.
- Bolot S, Munoz Bodnar A, Cunnac S, Ortiz E, Szurek B, Noël LD, Arlat M, Jacques M-A, Gagnevin L, Portier P, Fischer-Le Saux M, Carrere S, Koebnik R. 2013. Draft Genome Sequence of the *Xanthomonas cassavae* Type Strain CFBP 4642. *Genome Announc* **1**:e00679-13. doi:10.1128/genomeA.00679-13
- Booher NJ, Bogdanove AJ. 2014. Tools for TAL effector design and target prediction. *Methods* **69**:121–127. doi:<https://doi.org/10.1016/j.ymeth.2014.06.006>
- 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. *Microb genomics* **1**:e000032. doi:10.1099/mgen.0.000032
- Botero D, Monk J, Rodríguez Cubillos MJ, Rodríguez Cubillos A, Restrepo M, Bernal-Galeano V, Reyes A, González Barrios A, Palsson BØ, Restrepo S, Bernal A. 2020. Genome-Scale Metabolic Model of *Xanthomonas phaseoli* pv. *manihotis*: An Approach to Elucidate Pathogenicity at the Metabolic Level. *Front Genet*.
- Bredeson J V, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, Grimwood J, Schmutz J, Rabbi IY, Egesi C, Nauluvula P, Lebot V, Ndunguru J, Mkamilo G, Bart RS, Setter TL, Gleadow RM, Kulakow P, Ferguson ME, Rounsley S, Rokhsar DS. 2016. Sequencing wild and cultivated

- cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat Biotechnol* **34**:562–570. doi:10.1038/nbt.3535
- Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, Church GM. 2012. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res* **40**:e117–e117. doi:10.1093/nar/gks624
- Burkholder WH. 1942. Three bacterial plant pathogens: *Phytomonas earyophylli* sp.n., *Phytomonas alliiicola* sp.n., and *Phytomonas manihotis* (Arthaud-Berthet et Sondar) Viégas. *Phytopathology* **32**:141–149.
- Byju G, Suja G. 2020. Chapter Five - Mineral nutrition of cassava In: Sparks DLBT-A in A, editor. Academic Press. pp. 169–235. doi:https://doi.org/10.1016/bs.agron.2019.08.005
- CABI. 2020a. *Xanthomonas axonopodis* pv. *manihotis* (cassava bacterial blight). *Invasive Species Compend*. <https://www.cabi.org/isc/datasheet/56952>
- CABI. 2020b. *Xanthomonas cassavae* (cassava leaf spot). *Invasive Species Compend*. <https://www.cabi.org/isc/datasheet/56927>
- Cai L, Cao Y, Xu Z, Ma W, Zakria M, Zou L, Cheng Z, Chen G. 2017. A Transcription Activator-Like Effector Tal7 of *Xanthomonas oryzae* pv. *oryzicola* Activates Rice Gene Os09g29100 to Suppress Rice Immunity. *Sci Rep* **7**:5089. doi:10.1038/s41598-017-04800-8
- Carvajal-Yepes M, Olaya C, Lozano I, Cuervo M, Castaño M, Cuellar WJ. 2014. Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. *Virus Res* **186**:76–86. doi:https://doi.org/10.1016/j.virusres.2013.12.011
- Castiblanco LF, Gil J, Rojas A, Osorio D, Gutiérrez S, Muñoz-Bodnar A, Perez-Quintero AL, Koebnik R, Szurek B, López Carrascal CE, Restrepo S, Verdier V, Bernal AJ. 2013. TALE1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. *Mol Plant Pathol* **14**:84–95. doi:10.1111/j.1364-3703.2012.00830.x
- Ceballos H, Rojanaridpiched C, Phumichai C, Becerra LA, Kittipadakul P, Iglesias C, Gracen VE. 2020. Excellence in Cassava Breeding: Perspectives for the Future. *Crop Breeding, Genet Genomics* **2**:e200008. doi:10.20900/cbagg20200008
- 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 Res* **39**:e82–e82. doi:10.1093/nar/gkr218
- Cernadas RA, Doyle EL, Niño-Liu DO, Wilkins KE, Bancroft T, Wang L, Schmidt CL, Caldo R, Yang B, White FF, Nettleton D, Wise RP, Bogdanove AJ. 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 Pathog* **10**:e1003972. doi:10.1371/journal.ppat.1003972
- Cerqueira-Melo RC, Dórea-Bragança CA, Nogueira-Pestana K, Alves da Silva HS, Fortes-Ferreira C, Oliveira SA. 2019. Improvement of the specific detection of *Xanthomonas phaseoli* pv. *manihotis* based on the pthB gene. *Acta Sci Agron*.
- Chavarriaga Aguirre P, Brand A, Medina A, Prías M, Escobar R, Martínez J, Díaz Tatis P, López Carrascal CE, Roca WM, Tohme J. 2016. The potential of using biotechnology to improve

- cassava: a review. *Vitr Cell Dev Biol - Plant* **52**:461–478. doi:10.1007/s11627-016-9776-3
- Chege M, Wamunyokoli F, Kamau J, Nyaboga E. 2017. Phenotypic and genotypic diversity of *Xanthomonas axonopodis* pv. *manihotis* causing bacterial blight disease of cassava in Kenya. *J Appl Biol Biotechnol* **5**:38–44. doi:10.7324/JABB.2017.50206
- Chen L-Q. 2014. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol* **201**:1150–1155. doi:10.1111/nph.12445
- Chen X, Liu P, Mei L, He X, Chen L, Liu H, Shen S, Ji Z, Zheng X, Zhang Y, Gao Z, Zeng D, Qian Q, Ma B. 2021. Xa7, a new executor R gene that confers durable and broad-spectrum resistance to bacterial blight disease in rice. *Plant Commun* 100143. doi:https://doi.org/10.1016/j.xplc.2021.100143
- Chen X, Zuo S, Schwessinger B, Chern M, Canlas PE, Ruan D, Zhou X, Wang J, Daudi A, Petzold CJ, Heazlewood JL, Ronald PC. 2014. An XA21-associated kinase (OsSERK2) regulates immunity mediated by the XA21 and XA3 immune receptors. *Mol Plant* **7**:874–892. doi:10.1093/mp/ssu003
- Cheng H, Liu H, Deng Y, Xiao J, Li X, Wang S. 2015. The WRKY45-2 WRKY13 WRKY42 Transcriptional Regulatory Cascade Is Required for Rice Resistance to Fungal Pathogen. *Plant Physiol* **167**:1087 LP – 1099. doi:10.1104/pp.114.256016
- Chu Z, Fu B, Yang H, Xu C, Li Z, Sanchez A, Park YJ, Bennetzen JL, Zhang Q, Wang S. 2006. Targeting xa13, a recessive gene for bacterial blight resistance in rice. *Theor Appl Genet* **112**:455–461. doi:10.1007/s00122-005-0145-6
- Cianciotto NP, White RC. 2017. Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. *Infect Immun* **85**:e00014-17. doi:10.1128/IAI.00014-17
- CIAT. 1973. CIAT Annual Report. Cali, Colombia.
- Cohen SP, Liu H, Argueso CT, Pereira A, Vera Cruz C, Verdier V, Leach JE. 2017. RNA-Seq analysis reveals insight into enhanced rice Xa7-mediated bacterial blight resistance at high temperature. *PLoS One* **12**:e0187625.
- Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, 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. *Mol Plant-Microbe Interact* **27**:1186–1198. doi:10.1094/MPMI-06-14-0161-R
- Cohn M, Morbitzer R, Lahaye T, Staskawicz BJ. 2016. Comparison of gene activation by two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility genes in cassava. *Mol Plant Pathol* 875–889. doi:10.1111/mpp.12337
- Cohn M, Shybut M, Dahlbeck D, Staskawicz B. 2015. Assays to Assess Virulence of *Xanthomonas axonopodis* pv. *manihotis* on Cassava. *Bio-protocol* **5**:e1522. doi:10.21769/BioProtoc.1522
- Connor DJ, Palta J. 1981. Response of cassava to water shortage III. Stomatal control of plant water status. *F Crop Res* **4**:297–311. doi:https://doi.org/10.1016/0378-4290(81)90080-0
- 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.

doi:10.1111/ppa.12461

- Cooper RM, Kemp B, Day R, Gomez-Vasquez R, Beeching JR. 2001. Pathogenicity and Resistance in *Xanthomonas* Blight of Cassava In: De Boer SH, editor. *Plant Pathogenic Bacteria*. Dordrecht: Springer. pp. 319–323.
- Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, Carpenter SCD, Chen L-Q, Zheng H, Gao X, Zheng Y, Fei Z, Yu JZ, Isakeit T, Wheeler T, Frommer WB, He P, Bogdanove AJ, Shan L. 2017. TAL effector driven induction of a *SWEET* gene confers susceptibility to bacterial blight of cotton. *Nat Commun* **8**:15588. doi:10.1038/ncomms15588
- Cruz JL, Alves AAC, LeCain DR, Ellis DD, Morgan JA. 2016. Elevated CO₂ concentrations alleviate the inhibitory effect of drought on physiology and growth of cassava plants. *Sci Hortic (Amsterdam)* **210**:122–129. doi:https://doi.org/10.1016/j.scienta.2016.07.012
- Cuculis L, Abil Z, Zhao H, Schroeder CM. 2015. Direct observation of TALE protein dynamics reveals a two-state search mechanism. *Nat Commun* **6**:7277. doi:10.1038/ncomms8277
- Dania VO, Ojeyemi TD. 2019. Prevalence and polymerase chain reaction detection of *Xanthomonas axonopodis* pv. *manihotis* causal agent of cassava bacterial blight disease in Osun state, Southwestern Nigeria. *Niger J Biotechnol* **36**:159. doi:10.4314/njb.v36i1.21
- Daniel JF, Boher B. 1985a. Epiphytic phase of *Xanthomonas campestris* pathovar *manihotis* on aerial parts of cassava. *Agronomie* **5**:111–116.
- Daniel JF, Boher B. 1985b. Etude des modes de survie de l'agent causal de la bacteriose vasculaire du manioc, *Xanthomonas campestris* pathovar *manihotis*. *Agron Sci des Prod Veg l'environnement* **5**:339–346.
- de Souza AN, da Silva FN, Bedendo IP, Carvalho CM. 2013. A Phytoplasma Belonging to a 16SrIII-A Subgroup and dsRNA Virus Associated with Cassava Frogskin Disease in Brazil. *Plant Dis* **98**:771–779. doi:10.1094/PDIS-04-13-0440-RE
- Dedal OI, Palomar MK, Napiere CM. 1980. Host Range of *Xanthomonas manihotis* Starr. *Ann Trop Res* **2**:149–155.
- Delepelaire P. 2004. Type I secretion in gram-negative bacteria. *Biochim Biophys Acta - Mol Cell Res* **1694**:149–161. doi:https://doi.org/10.1016/j.bbamcr.2004.05.001
- Denancé N, Szurek B, Doyle EL, Lauber E, Fontaine-Bodin L, Carrère S, Guy E, Hajri A, Cerutti A, Boureau T, Poussier S, Arlat M, Bogdanove AJ, Noël LD. 2018. Two ancestral genes shaped the *Xanthomonas campestris* TAL effector gene repertoire. *New Phytol* **219**:391–407. doi:10.1111/nph.15148
- Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J-K, Shi Y, Yan N. 2012. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**:720–723. doi:10.1126/science.1215670
- Deng D, Yan C, Wu J, Pan X, Yan N. 2014. Revisiting the TALE repeat. *Protein Cell* **5**:297–306. doi:10.1007/s13238-014-0035-2
- Deng H, 卿邓汉, Liu H, 波刘红, Li X, 花李香, Xiao J, 华肖景, Wang S, 平王石. 2012. A CCCH-Type Zinc Finger Nucleic Acid-Binding Protein Quantitatively Confers Resistance against Rice

- Bacterial Blight Disease. *Plant Physiol* **158**:876–889. doi:10.1104/pp.111.191379
- Devanna BN, Jaswal R, Singh PK, Kapoor R, Jain P, Kumar G, Sharma Y, Samantaray S, Sharma TR. 2021. Role of transporters in plant disease resistance. *Physiol Plant* **171**:849–867. doi:https://doi.org/10.1111/ppl.13377
- Díaz-Tatis PA, Ochoa JC, García L, Chavarriaga P, Bernal AJ, López Carrascal CE. 2019. Interfamily transfer of Bs2 from pepper to cassava (*Manihot esculenta* Crantz). *Trop Plant Pathol* **44**:225–237. doi:10.1007/s40858-019-00279-y
- Díaz Tatis PA, Herrera Corzo M, Ochoa Cabezas JC, Medina Cipagauta A, Prías MA, Verdier V, Chavarriaga Aguirre P, López Carrascal CE. 2018. The overexpression of *RXam1*, a cassava gene coding for an RLK, confers disease resistance to *Xanthomonas axonopodis* pv. *manihotis*. *Planta* **247**:1031–1042. doi:10.1007/s00425-018-2863-4
- Dixon AGO, Ngeve JM, Nukenine EN. 2002. Genotype × environment Effects on Severity of Cassava Bacterial Blight Disease caused by *Xanthomonas axonopodis* pv. *manihotis*. *Eur J Plant Pathol* **108**:763–770. doi:10.1023/A:1020876019227
- Dong X, Jin Y, Ming D, Li B, Dong H, Wang L, Wang T, Wang D. 2017. CRISPR/dCas9-mediated inhibition of gene expression in *Staphylococcus aureus*. *J Microbiol Methods* **139**:79–86. doi:https://doi.org/10.1016/j.mimet.2017.05.008
- Doyle EL, Hummel AW, Demorest ZL, Starker CG, Voytas DF, Bradley P, Bogdanove AJ. 2013. 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.
- Duan S, Jia H, Pang Z, Teper D, White F, Jones J, Zhou C, Wang N. 2018. Functional characterization of the citrus canker susceptibility gene *CsLOB1*. *Mol Plant Pathol* **19**:1908–1916. doi:10.1111/mpp.12667
- El-Sharkaway MA, Cock JH. 1984. Water Use Efficiency of Cassava. I. Effects of Air Humidity and Water Stress on Stomatal Conductance and Gas Exchange1. *Crop Sci* **24**:cropsci1984.0011183X002400030017x. doi:https://doi.org/10.2135/cropsci1984.0011183X002400030017x
- El-Sharkawy M. 2012. Stress-Tolerant Cassava: The Role of Integrative Ecophysiology-Breeding Research in Crop Improvement. *Open J Soil Sci* **02**. doi:10.4236/ojss.2012.22022
- El-Sharkawy MA. 2007. Physiological characteristics of cassava tolerance to prolonged drought in the tropics: implications for breeding cultivars adapted to seasonally dry and semiarid environments. *Brazilian J Plant Physiol*.
- El-Sharkawy MA. 2006. International research on cassava photosynthesis, productivity, ecophysiology, and responses to environmental stresses in the tropics. *Photosynthetica* **44**:481–512. doi:10.1007/s11099-006-0063-0
- El Kasmi F, Horvath D, Lahaye T. 2018. Microbial effectors and the role of water and sugar in the infection battle ground. *Curr Opin Plant Biol* **44**:98–107. doi:https://doi.org/10.1016/j.pbi.2018.02.011
- Elango F, Lozano JC. 1981. Epiphytic survival of *Xanthomonas manihotis* on common weeds in

- Colombia 5th Int. Conf. Plant Path. Bact. Cali, Colombia. pp. 203–209.
- Elango F, Lozano JC. 1980. Transmission of *Xanthomonas manihotis* in Seed of Cassava (*Manihot esculenta*). *Plant Dis* **64**:784–786.
- Eom J-S, Luo D, Atienza-Grande G, Yang J, Ji C, Thi Luu V, Huguet-Tapia JC, Char SN, Liu B, Nguyen H, Schmidt SM, Szurek B, Vera Cruz C, White FF, Oliva R, Yang B, Frommer WB. 2019. Diagnostic kit for rice blight resistance. *Nat Biotechnol* **37**:1372–1379. doi:10.1038/s41587-019-0268-y
- Erkes A, Reschke M, Boch J, Grau J. 2017. Evolution of Transcription Activator-Like Effectors in *Xanthomonas oryzae*. *Genome Biol Evol* **9**:1599–1615. doi:10.1093/gbe/evx108
- Escalon A. 2013. Evolution et spécialisation du pouvoir pathogène de *Xanthomonas citri* pv. *citri* : rôle des effecteurs de type 3. Saint-Denis: Université de la Réunion.
- Ewbank E, Maraite H. 1990. Conversion of methionine to phytotoxic 3-methylthiopropionic acid by *Xanthomonas campestris* pv. *manihotis*. *Microbiology* **136**:1185–1189. doi:https://doi.org/10.1099/00221287-136-7-1185
- Falahi Charkhabi N, Booher NJ, Peng Z, Wang L, Rahimian H, Shams-Bakhsh M, Liu Z, Liu S, White FF, Bogdanove AJ. 2017. Complete Genome Sequencing and Targeted Mutagenesis Reveal Virulence Contributions of Tal2 and Tal4b of *Xanthomonas translucens* pv. *undulosa* ICMP11055 in Bacterial Leaf Streak of Wheat. *Front Microbiol* .
- Fanou A. 1999. Epidemiological and ecological investigations on cassava bacterial blight and development of integrated methods for its control in Africa. Cuvillier.
- Fanou A, Zinsou V, Wydra K. 2018. Cassava Bacterial Blight: A Devastating Disease of Cassava. doi:10.5772/intechopen.71527
- Fanou A, Zinsou V, Wydra K. 2017. Survival of *Xanthomonas axonopodis* pv. *manihotis* in weed species and in cassava debris: implication in the epidemiology of cassava bacterial blight. *Int J Adv Res* **5**:2098–2112. doi:10.21474/IJAR01/4057
- FAO. 2020. FAOSTAT Crops. <http://www.fao.org/faostat/en/#data/QC>
- FAO. 2013. SAVE AND GROW: Cassava, a guide to sustainable production intensification. Roma, Italia: FAO.
- Fermont AM, van Asten PJA, Tittonell P, van Wijk MT, Giller KE. 2009. Closing the cassava yield gap: An analysis from smallholder farms in East Africa. *F Crop Res* **112**:24–36. doi:https://doi.org/10.1016/j.fcr.2009.01.009
- Ferreira RM, de Oliveira ACP, Moreira LM, Belasque J, 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**. doi:10.1128/mBio.02505-14
- Flood J, Cooper RM, Deshappriya N. 1995. Resistance of cassava (*Manihot esculenta*) to *Xanthomonas* blight in vitro and in planta. *Aspects of Applied Biology* (United Kingdom).
- Flores C, Zarate C, Triplett L, Maillot-Lebon V, Moufid Y, Kanté M, Bragard C, Verdier V, Gagnevin L, Szurek B, Robène I. 2019. Development of a duplex-PCR for differential diagnosis of

- Xanthomonas phaseoli* pv. *manihotis* and *Xanthomonas cassavae* in cassava (*Manihot esculenta*). *Physiol Mol Plant Pathol* **105**:34–46.
doi:<https://doi.org/10.1016/j.pmpp.2018.07.005>
- Fokunang CN, Akem CN, Dixon AGO, Ikotun T. 2000. Evaluation of a cassava germplasm collection for reaction to three major diseases and the effect on yield. *Genet Resour Crop Evol* **47**:63–71. doi:10.1023/A:1008797225343
- Fregene M, Angel F, Gomez R, Rodriguez F, Chavarriaga P, Roca W, Tohme J, Bonierbale M. 1997. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor Appl Genet* **95**:431–441. doi:10.1007/s001220050580
- Freitas EL, Brito AC, de Oliveira SAS, de Oliveira EJ. 2020. Early diagnosis of cassava frog skin disease in powdered tissue samples using near-infrared spectroscopy. *Eur J Plant Pathol* **156**:547–558. doi:10.1007/s10658-019-01904-x
- Frison E, Feliu E. 1991. FAO/IBPGR Technical Guidelines for the Safe Movement of Cassava Germplasm. doi:10.13140/RG.2.1.2878.2569
- Gluck-Thaler E, Cerutti A, Perez-Quintero AL, Butchacas J, Roman-Reyna V, Madhavan VN, Shantharaj D, Merfa M V, Pesce C, Jauneau A, Vancheva T, Lang JM, Allen C, Verdier V, Gagnevin L, Szurek B, Beckham GT, De La Fuente L, Patel HK, Sonti R V, Bragard C, Leach JE, Noël LD, Slot JC, Koebnik R, Jacobs JM. 2020. Repeated gain and loss of a single gene modulates the evolution of vascular plant pathogen lifestyles. *Sci Adv* **6**:eabc4516. doi:10.1126/sciadv.abc4516
- Gochez AM, Huguet-Tapia JC, Minsavage G V, Shantaraj D, Jalan N, Strauß A, Lahaye T, Wang N, Canteros BI, Jones JB, Potnis N. 2018. Pacbio sequencing of copper-tolerant *Xanthomonas citri* reveals presence of a chimeric plasmid structure and provides insights into reassortment and shuffling of transcription activator-like effectors among *X. citri* strains. *BMC Genomics* **19**:16. doi:10.1186/s12864-017-4408-9
- Gómez-Cano F, Soto J, Restrepo S, Bernal A, López-Kleine L, López Carrascal CE. 2019. Gene co-expression network for *Xanthomonas*-challenged cassava reveals key regulatory elements of immunity processes. *Eur J Plant Pathol* **153**:1083–1104. doi:10.1007/s10658-018-01628-4
- Gomez MA, Lin ZD, Moll T, Chauhan RD, Hayden L, Renninger K, Beyene G, Taylor NJ, Carrington JC, Staskawicz BJ, Bart RS. 2019. Simultaneous CRISPR/Cas9-mediated editing of cassava eIF4E isoforms nCBP-1 and nCBP-2 reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnol J* **17**:421–434. doi:<https://doi.org/10.1111/pbi.12987>
- Gonzalez C, Restrepo S, Tohme J, Verdier V. 2002. Characterization of pathogenic and nonpathogenic strains of *Xanthomonas axonopodis* pv. *manihotis* by PCR-based DNA fingerprinting techniques. *Fems Microbiol Lett* **215**:23–31.
- Gu K, Sangha JS, Li Y, Yin Z. 2008. High-resolution genetic mapping of bacterial blight resistance gene Xa10. *Theor Appl Genet* **116**:155–163. doi:10.1007/s00122-007-0655-5
- Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang G-L, White FF, Yin Z. 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**:1122–1125. doi:http://www.nature.com/nature/journal/v435/n7045/supinfo/nature03630_S1.html

- Hahn SK, Howland AK, Terry ER. 1980. Correlated resistance of cassava to mosaic and bacterial blight diseases. *Euphytica* **29**:305–311. doi:10.1007/BF00025127
- Han J, Xia Z, Liu P, Li C, Wang Y, Guo L, Jiang G, Zhai W. 2020. TALEN-based editing of TFIIA5 changes rice response to *Xanthomonas oryzae* pv. *Oryzae*. *Sci Rep* **10**:2036. doi:10.1038/s41598-020-59052-w
- Haq F, Xie S, Huang K, Shah SMA, Ma W, Cai L, Xu X, Xu Z, Wang S, Zou L, Zhu B, Chen G. 2020. Identification of a virulence tal gene in the cotton pathogen, *Xanthomonas citri* pv. *malvacearum* strain Xss-V(2)-18. *BMC Microbiol* **20**:91. doi:10.1186/s12866-020-01783-x
- Harris KP, Martin A, Novak S, Kim S-H, Reynolds T, Leigh A. 2015. Cassava Bacterial Blight and Postharvest Physiological Deterioration, Production Losses and Control Strategies. [https://evans.uw.edu/wp-content/uploads/files//EPAR_UW_Request_298_CBB & PPD Production Losses and Control Strategies_3.23.15_0.pdf](https://evans.uw.edu/wp-content/uploads/files//EPAR_UW_Request_298_CBB_&PPD_Production_Losses_and_Control_Strategies_3.23.15_0.pdf)
- Hayward AC. 1993. The hosts of *Xanthomonas* In: Swings JG, Civerolo EL, editors. *Xanthomonas*. Dordrecht: Springer. pp. 1–119.
- He X, Liu G, Li B, Xie Y, Wei Y, Shang S, Tian L, Shi H. 2019. Functional analysis of the heterotrimeric NF-Y transcription factor complex in cassava disease resistance. *Ann Bot* **124**:1185–1197. doi:10.1093/aob/mcz115
- Hillocks RJ, Wydra K. 2002. Bacterial, fungal and nematode diseasesCassava: Biology, Production and Utilization. Wallingford, UK ; New York: CABI Publishing. pp. 261–280.
- 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. *Proc Natl Acad Sci U S A* **111**:E521–E529. doi:10.1073/pnas.1313271111
- Huang R, Hui S, Zhang M, Li P, Xiao J, Li X, Yuan M, Wang S. 2017. A Conserved Basal Transcription Factor Is Required for the Function of Diverse TAL Effectors in Multiple Plant Hosts . *Front Plant Sci* .
- Hui S, Liu H, Zhang M, Chen D, Li Q, Tian J, Xiao J, Li X, Wang S, Yuan M. 2019a. The host basal transcription factor IIA subunits coordinate for facilitating infection of TALEs-carrying bacterial pathogens in rice. *Plant Sci* **284**:48–56. doi:https://doi.org/10.1016/j.plantsci.2019.04.004
- Hui S, Shi Y, Tian J, Wang L, Li Y, Wang S, Yuan M. 2019b. TALE-carrying bacterial pathogens trap host nuclear import receptors for facilitation of infection of rice. *Mol Plant Pathol* **20**:519–532. doi:https://doi.org/10.1111/mpp.12772
- Hutin M, Pérez-Quintero AL, López Carrascal CE, Szurek B. 2015a. MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. *Front Plant Sci* **6**:535. doi:10.3389/fpls.2015.00535
- Hutin M, Sabot F, Ghesquière A, Koebnik R, Szurek B. 2015b. A knowledge-based molecular screen uncovers a broad-spectrum OsSWEET14 resistance allele to bacterial blight from wild rice. *Plant J* **84**:694–703. doi:https://doi.org/10.1111/tpj.13042
- International Cassava Genetic Map Consortium (ICGMC). 2015. High-Resolution Linkage Map and Chromosome-Scale Genome Assembly for Cassava (*Manihot esculenta* Crantz) from 10

- Populations. *G3 Genes/Genomes/Genetics* **5**:133 LP – 144. doi:10.1534/g3.114.015008
- Iyer AS, McCouch SR. 2004. The Rice Bacterial Blight Resistance Gene xa5 Encodes a Novel Form of Disease Resistance. *Mol Plant-Microbe Interact* **17**:1348–1354. doi:10.1094/MPMI.2004.17.12.1348
- Jarvis A, Ramirez-Villegas J, Herrera Campo BV, Navarro-Racines C. 2012. Is Cassava the Answer to African Climate Change Adaptation? *Trop Plant Biol* **5**:9–29. doi:10.1007/s12042-012-9096-7
- Ji Z, Ji C, Liu B, Zou L, Chen G, Yang B. 2016. Interfering TAL effectors of *Xanthomonas oryzae* neutralize R-gene-mediated plant disease resistance. *Nat Commun* **7**:13435. doi:10.1038/ncomms13435
- Jia H, Wang N. 2020. Generation of homozygous canker-resistant citrus in the T0 generation using CRISPR-SpCas9p. *Plant Biotechnol J* **18**:1990–1992. doi:https://doi.org/10.1111/pbi.13375
- Jia H, Zhang Y, Orbović V, Xu J, White FF, Jones JB, Wang N. 2017. Genome editing of the disease susceptibility gene CsLOB1 in citrus confers resistance to citrus canker. *Plant Biotechnol J* **15**:817–823. doi:https://doi.org/10.1111/pbi.12677
- Jiang G-H, Xia Z-H, Zhou Y-L, Wan J, Li D-Y, Chen R-S, Zhai W-X, Zhu L-H. 2006. Testifying the rice bacterial blight resistance gene xa5 by genetic complementation and further analyzing xa5 (Xa5) in comparison with its homolog TFIIAγ1. *Mol Genet Genomics* **275**:354–366. doi:10.1007/s00438-005-0091-7
- Jones W. 1959. *Manioc in Africa*. Stanford, CA: Stanford University Press.
- Jorge V, Fregene M, Duque MC, Bonierbale MW, Tohme J, Verdier V. 2000. Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz). *Theor Appl Genet* **101**:865–872. doi:10.1007/s001220051554
- Jorge V, Fregene M, Vélez CM, Duque MC, Tohme J, Verdier V. 2001. QTL analysis of field resistance to *Xanthomonas axonopodis* pv. *manihotis* in cassava. *Theor Appl Genet* **102**:564–571. doi:10.1007/s001220051683
- Jorge V, Verdier V. 2002. Qualitative and quantitative evaluation of cassava bacterial blight resistance in F1 progeny of a cross between elite cassava clones. *Euphytica* **123**:41–48. doi:10.1023/A:1014400823817
- Joseph J, Elango F. 1991. The Status of Cassava Bacterial Blight Caused by *Xanthomonas campestris* pv. *manihotis* in Trinidad. *J Phytopathol* **133**:320–326. doi:10.1111/j.1439-0434.1991.tb00167.x
- 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. *Mol Plant-Microbe Interact* **18**:838–848. doi:10.1094/MPMI-18-0838
- 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 (80-)* **318**:648–651. doi:10.1126/science.1144956
- Kemp BP, Horne J, Bryant A, Cooper RM. 2004. *Xanthomonas axonopodis* pv. *manihotis* gumD gene is essential for EPS production and pathogenicity and enhances epiphytic survival on cassava (*Manihot esculenta*). *Physiol Mol Plant Pathol* **64**:209–218.

doi:<https://doi.org/10.1016/j.pmpp.2004.08.007>

- Khanthavong P, Oudthachit S, Souvannalat A, Matsumoto N. 2016. Effect of weed biomass on cassava yield related to weeding times. *Adv Plants Agric Res* **5**.
- Koinuma H, Miyazaki A, Wakaki R, Fujimoto Y, Iwabuchi N, Nijo T, Kitazawa Y, Shigaki T, Maejima K, Yamaji Y, Namba S. 2018. First report of 'Candidatus Phytoplasma pruni' infecting cassava in Japan. *J Gen Plant Pathol* **84**:300–304. doi:[10.1007/s10327-018-0787-2](https://doi.org/10.1007/s10327-018-0787-2)
- Kondrateva E, Demchenko A, Lavrov A, Smirnikhina S. 2021. An overview of currently available molecular Cas-tools for precise genome modification. *Gene* **769**:145225. doi:<https://doi.org/10.1016/j.gene.2020.145225>
- Kpémoua K, Boher B, Nicole M, Calatayud P, Geiger JP. 1996. Cytochemistry of defense responses in cassava infected by *Xanthomonas campestris* pv. *manihotis*. *Can J Microbiol* **42**:1131–1143. doi:[10.1139/m96-145](https://doi.org/10.1139/m96-145)
- Krönauer C, Kilian J, Strauß T, Stahl M, Lahaye T. 2019. Cell Death Triggered by the YUCCA-like Bs3 Protein Coincides with Accumulation of Salicylic Acid and Pipecolic Acid But Not of Indole-3-Acetic Acid. *Plant Physiol* **180**:1647 LP – 1659. doi:[10.1104/pp.18.01576](https://doi.org/10.1104/pp.18.01576)
- Kuon J-E, Qi W, Schläpfer P, Hirsch-Hoffmann M, von Bieberstein PR, Patrignani A, Poveda L, Grob S, Keller M, Shimizu-Inatsugi R, Grossniklaus U, Vanderschuren H, Gruitsem W. 2019. Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. *BMC Biol* **17**:75. doi:[10.1186/s12915-019-0697-6](https://doi.org/10.1186/s12915-019-0697-6)
- Lai Y, Babunovic GH, Cui L, Dedon PC, Doench JG, Fortune SM, Lu TK. 2020. Illuminating Host-Mycobacterial Interactions with Genome-wide CRISPR Knockout and CRISPRi Screens. *Cell Syst* **11**:239-251.e7. doi:<https://doi.org/10.1016/j.cels.2020.08.010>
- Lamphey JNL, Okoli OO, Frimpong-Manso PP. 1998. Incidence and severity of African cassava mosaic disease (ACMD) and cassava bacterial blight (CBB) on some local and exotic cassava varieties in different ecological zones of Ghana. *Ghana J Agric Sci* **31**:35–43.
- Latif S, Müller J. 2015. Potential of cassava leaves in human nutrition: A review. *Trends Food Sci Technol* **44**:147–158. doi:<https://doi.org/10.1016/j.tifs.2015.04.006>
- Li P, Long J, Huang Y, Zhang Y, Wang J. 2004. AvrXa3: A novel member of avrBs3 gene family from *Xanthomonas oryzae* pv. *oryzae* has a dual function. *Prog Nat Sci* **14**:774–780. doi:[10.1080/10020070412331344311](https://doi.org/10.1080/10020070412331344311)
- Li T, Liu B, Spalding MH, Weeks DP, Yang B. 2012. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol* **30**:390–392. doi:[10.1038/nbt.2199](https://doi.org/10.1038/nbt.2199)
- Li X, Fan S, Hu W, Liu G, Wei Y, He C, Shi H. 2017. Two Cassava Basic Leucine Zipper (bZIP) Transcription Factors (MebZIP3 and MebZIP5) Confer Disease Resistance against Cassava Bacterial Blight. *Front Plant Sci* .
- Lin ZJD, Taylor NJ, Bart R. 2019. Engineering Disease-Resistant Cassava. *Cold Spring Harb Perspect Biol* . doi:[10.1101/cshperspect.a034595](https://doi.org/10.1101/cshperspect.a034595)
- Liu C, Schläppi MR, Mao B, Wang W, Wang A, Chu C. 2019. The bZIP73 transcription factor controls rice cold tolerance at the reproductive stage. *Plant Biotechnol J* **17**:1834–1849. doi:<https://doi.org/10.1111/pbi.13104>

- Liu Q, Yuan M, Zhou YAN, 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 Environ* **34**:1958–1969. doi:<https://doi.org/10.1111/j.1365-3040.2011.02391.x>
- Liu W, Yan Y, Zeng H, Li X, Wei Y, Liu G, He C, Shi H. 2018. Functional characterization of WHY–WRKY75 transcriptional module in plant response to cassava bacterial blight. *Tree Physiol* **38**:1502–1512. doi:[10.1093/treephys/tpy053](https://doi.org/10.1093/treephys/tpy053)
- Liu Y, Cao Y, Zhang Q, Li X, Wang S. 2018. A Cytosolic Triosephosphate Isomerase Is a Key Component in XA3/XA26-Mediated Resistance. *Plant Physiol* **178**:923–935. doi:[10.1104/pp.18.00348](https://doi.org/10.1104/pp.18.00348)
- López Carrascal CE, Bernal AJ. 2012. Cassava Bacterial Blight: Using Genomics for the Elucidation and Management of an Old Problem. *Trop Plant Biol* **5**:117–126. doi:[10.1007/s12042-011-9092-3](https://doi.org/10.1007/s12042-011-9092-3)
- López Carrascal CE, Jorge V, Piégu B, Mba C, Cortes D, Restrepo S, Soto M, Laudíe M, Berger C, Cooke R, Delseny M, Tohme J, Verdier V. 2004. A unigene catalogue of 5700 expressed genes in cassava. *Plant Mol Biol* **56**:541–554. doi:[10.1007/s11103-004-0123-4](https://doi.org/10.1007/s11103-004-0123-4)
- López Carrascal CE, Quesada-Ocampo LM, Bohórquez A, Duque MC, Vargas J, Tohme J, Verdier V. 2007. Mapping EST-derived SSRs and ESTs involved in resistance to bacterial blight in *Manihot esculenta*. *Genome* **50**:1078–1088. doi:[10.1139/G07-087](https://doi.org/10.1139/G07-087)
- López Carrascal CE, Soto M, Restrepo S, Piégu B, Cooke R, Delseny M, Tohme J, Verdier V. 2005. Gene expression profile in response to *Xanthomonas axonopodis* pv. *manihotis* infection in cassava using a cDNA microarray. *Plant Mol Biol* **57**:393–410. doi:[10.1007/s11103-004-7819-3](https://doi.org/10.1007/s11103-004-7819-3)
- López Carrascal CE, Zuluaga AP, Cooke R, Delseny M, Tohme J, Verdier V. 2003. Isolation of Resistance Gene Candidates (RGCs) and characterization of an RGC cluster in cassava. *Mol Genet Genomics* **269**:658–671. doi:[10.1007/s00438-003-0868-5](https://doi.org/10.1007/s00438-003-0868-5)
- Lozano JC. 1989. Outbreaks of cassava diseases and losses induced. *Fitopatol Bras* **14**:7–11.
- Lozano JC. 1986. Cassava Bacterial Blight: A Manageable Disease. *Plant Dis*, Plant disease **70**:1089. doi:[10.1094/pd-70-1089](https://doi.org/10.1094/pd-70-1089)
- Lozano JC, Booth RH. 1974. Diseases of Cassava (*Manihot esculenta* Crantz). *PANS Pest Artic News Summ* **20**:30–54. doi:[10.1080/09670877409412334](https://doi.org/10.1080/09670877409412334)
- Lozano JC, Laberry R. 1982. Screening for Resistance to Cassava Bacterial Blight. *Plant Dis* **66**:316–318.
- Lozano JC, Sequeira L. 1974a. Bacterial Blight of Cassava in Colombia: Etiology. *Phytopathology* **64**:74–82.
- Lozano JC, Sequeira L. 1974b. Bacterial Blight of Cassava in Colombia: Epidemiology and Control. *Phytopathology* **64**:83–88. doi:[10.1094/Phyto-64-83](https://doi.org/10.1094/Phyto-64-83).
- Lozano R, Hamblin MT, Prochnik S, Jannink J-L. 2015. Identification and distribution of the *NBS-LRR* gene family in the Cassava genome. *BMC Genomics* **16**:360. doi:[10.1186/s12864-015-1554-9](https://doi.org/10.1186/s12864-015-1554-9)
- Ma W, Zou L, Zhiyuan JI, Xiameng XU, Zhengyin XU, Yang Y, Alfano JR, Chen G. 2018. *Xanthomonas*

- oryzae pv. oryzae TALE proteins recruit OsTFIIA γ 1 to compensate for the absence of OsTFIIA γ 5 in bacterial blight in rice. *Mol Plant Pathol* **19**:2248–2262. doi:<https://doi.org/10.1111/mpp.12696>
- Majer C, Hochholdinger F. 2011. Defining the boundaries: structure and function of LOB domain proteins. *Trends Plant Sci* **16**:47–52. doi:10.1016/j.tplants.2010.09.009
- Mak AN-S, Bradley P, Bogdanove AJ, Stoddard BL. 2013. TAL effectors: function, structure, engineering and applications. *Curr Opin Struct Biol* **23**:93–99. doi:<https://doi.org/10.1016/j.sbi.2012.11.001>
- Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. 2012. The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. *Science (80-)* **335**:716–719. doi:10.1126/science.1216211
- Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer S V, Machado MA, Toth I, Salmond G, Foster GD. 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol* **13**:614–629. doi:10.1111/j.1364-3703.2012.00804.x
- Maraite H. 1993. *Xanthomonas campestris* pathovars on cassava: cause of bacterial blight and bacterial necrosis In: Swings JG, Civerolo EL, editors. *Xanthomonas*. London, UK: Springer Science + Business Media Dordrecht. pp. 18–25.
- Maraite H, Meyer J. 1975. *Xanthomonas manihotis* (Arthaud-Berthet) Starr, Causal Agent of Bacterial Wilt, Blight and Leaf Spots of Cassava in Zaire. *Int J Pest Manag - INT J PEST Manag* **21**:27–37. doi:10.1080/09670877509411484
- Maraite H, Perreux D. 1979. Comparative symptom development in cassava after infection by *Xanthomonas manihotis* or *X. cassavae* under controlled conditions. In: Terry ER, Persley GJ, Cook SCA, editors. *Cassava Bacterial Blight in Africa; Past, Present and Future. Report of an Interdisciplinary Workshop, IITA, Ibadan, Nigeria, 1978*. London, UK: Centre for Overseas Pest Research. pp. 17–24.
- Marcano M, Trujillo G. 1984. Papel de las malezas en relacion a la perpetuacion del anublo bacteria no de la yuca. *Rev la Fac Agron* **13**:167–181.
- Mbah EU, Nwankwo BC, Njoku DN, Gore MA. 2019. Genotypic evaluation of twenty-eight high- and low-cyanide cassava in low-land tropics, southeast Nigeria. *Heliyon* **5**:e01855. doi:<https://doi.org/10.1016/j.heliyon.2019.e01855>
- Mbaringong G, Nyaboga E, Ondu V, Kanduma E. 2017. Evaluation of Selected Cassava (*Manihot esculenta* Crantz) Cultivars Grown in Kenya for Resistance to Bacterial Blight Disease. *World J Agric Res* **5**:94–101. doi:10.12691/wjar-5-2-5
- McCallum EJ, Anjanappa RB, Grisse W. 2017. Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). *Curr Opin Plant Biol* **38**:50–58. doi:<https://doi.org/10.1016/j.pbi.2017.04.008>
- Medina CA, Reyes PA, Trujillo CA, Gonzalez JL, Bejarano DA, Montenegro NA, Jacobs JM, Joe A, Restrepo S, Alfano JR, Bernal A. 2018. The role of type III effectors from *Xanthomonas axonopodis* pv. *manihotis* in virulence and suppression of plant immunity. *Mol Plant Pathol* **19**:593–606. doi:10.1111/mpp.12545

- Midha S, Patil PB. 2014. Genomic insights into the evolutionary origin of *Xanthomonas axonopodis* pv. *citri* and its ecological relatives. *Appl Environ Microbiol* **80**:6266–6279. doi:10.1128/AEM.01654-14
- Montagnac JA, Davis CR, Tanumihardjo SA. 2009. Nutritional Value of Cassava for Use as a Staple Food and Recent Advances for Improvement. *Compr Rev food Sci food Saf* **8**:181–194. doi:10.1111/j.1541-4337.2009.00077.x
- Mora R, Rodriguez M, Gayosso L, López Carrascal CE. 2019. Using in vitro plants to study the cassava response to *Xanthomonas phaseoli* pv. *manihotis* infection. *Trop Plant Pathol*. doi:10.1007/s40858-019-00296-x
- Moscou MJ, Bogdanove AJ. 2009. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science (80-)* **326**:1501 LP – 1501. doi:10.1126/science.1178817
- Mostade JM, Butare I. 1978. Symptomatologie et epidemiologie de la necrose bacterienne du manioc causee par *Xanthomonas cassavae* au Rwanda In: Maraite H, Meyer JA, editors. Disease of Tropical Food Crops. Proceedings of an International Symposium. U.c.L., Louvain-la-Neuve, Belgium: Universite Catholique de Louvain, Louvain-la-Neuve, Belgium. pp. 95–117.
- Muñoz-Bodnar A, Cruz-Gómez LM, Bernal A, Szurek B, López Carrascal CE. 2015. Comparing Inoculation Methods to Evaluate the Growth of *Xanthomonas axonopodis* pv. *manihotis* on Cassava Plants. *Acta Biológica Colomb*.
- Muñoz-Bodnar A, Perez-Quintero AL, Gomez-Cano F, Gil J, Michelmore R, Bernal A, Szurek B, López Carrascal CE. 2014. RNAseq analysis of cassava reveals similar plant responses upon infection with pathogenic and non-pathogenic strains of *Xanthomonas axonopodis* pv. *manihotis*. *Plant Cell Rep* **33**:1901–1912. doi:10.1007/s00299-014-1667-7
- Mutka AM, Fentress SJ, Sher JW, Berry JC, Pretz C, Nusinow DA, Bart R. 2016. Quantitative, Image-Based Phenotyping Methods Provide Insight into Spatial and Temporal Dimensions of Plant Disease. *Plant Physiol* **172**:650 LP – 660. doi:10.1104/pp.16.00984
- Myrbråten IS, Wiull K, Salehian Z, Håvarstein LS, Straume D, Mathiesen G, Kjos M. 2019. CRISPR Interference for Rapid Knockdown of Essential Cell Cycle Genes in *Lactobacillus plantarum*; *mSphere* **4**:e00007-19. doi:10.1128/mSphere.00007-19
- Narayanan N, Beyene G, Chauhan RD, Gaitán-Solís E, Gehan J, Butts P, Siritunga D, Okwuonu I, Woll A, Jiménez-Aguilar DM, Boy E, Grusak MA, Anderson P, Taylor NJ. 2019. Biofortification of field-grown cassava by engineering expression of an iron transporter and ferritin. *Nat Biotechnol* **37**:144–151. doi:10.1038/s41587-018-0002-1
- Nassar N. 2007. Cassava improvement: Challenges and impacts. *J Agric Sci - J AGR SCI* **145**. doi:10.1017/S0021859606006575
- Night G, Asiimwe P, Gashaka G, Nkezabahizi D, Legg JP, Okao-Okuja G, Obonyo R, Nyirahorana C, Mukakanyana C, Mukase F, Munyabarenzi I, Mutumwinka M. 2011. Occurrence and distribution of cassava pests and diseases in Rwanda. *Agric Ecosyst Environ* **140**:492–497. doi:https://doi.org/10.1016/j.agee.2011.01.014
- Ntui VO, Kong K, Khan RS, Igawa T, Janavi GJ, Rabindran R, Nakamura I, Mii M. 2015. Resistance to Sri Lankan Cassava Mosaic Virus (SLCMV) in Genetically Engineered Cassava cv. KU50 through RNA Silencing. *PLoS One* **10**:e0120551.

- Ogunjobi A, Dixon A, Fagade O. 2007. Molecular genetic study of cassava bacterial blight casual agent in Nigeria using random amplified polymorphic DNA. *Electron J Environ Agric Food Chem* **6**.
- Ogunjobi A, Fagade O, Dixon A. 2006. Molecular variation in population structure of *Xanthomonas axonopodis* pv *manihotis* in the south eastern Nigeria. *African J Biotechnol* **5**:1868–1872. doi:10.4314/ajb.v5i20.55891
- Ogunjobi AA, Fagade OE, Dixon AGO. 2010. Comparative analysis of genetic variation among *Xanthomonas axonopodis* pv *manihotis* isolated from the western states of Nigeria using RAPD and AFLP. *Indian J Microbiol* **50**:132–138. doi:10.1007/s12088-010-0037-6
- Ojeda S, Verdier V. 2000. Detecting *Xanthomonas axonopodis* pv. *manihotis* in cassava true seeds by nested polymerase chain reaction assay. *Can J Plant Pathol* **22**:241–247. doi:10.1080/07060660009500470
- Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom J-S, Li C, Nguyen H, Liu B, Auguy F, Sciallano C, Luu VT, Dossa GS, Cunnac S, Schmidt SM, Slamet-Loedin IH, Vera Cruz C, Szurek B, Frommer WB, White FF, Yang B. 2019. Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat Biotechnol* **37**:1344–1350. doi:10.1038/s41587-019-0267-z
- Oliveira EJ de, Aud FF, Morales CFG, Oliveira SAS de, Santos V da S. 2016. Non-hierarchical clustering of *Manihot esculenta* Crantz germplasm based on quantitative traits. *Rev Ciência Agronômica*.
- Oliveira SAS, Abreu EFM, Araújo TS, Oliveira EJ, Andrade EC, Garcia JMP, Álvarez E. 2013. First Report of a 16SrIII-L Phytoplasma Associated with Frogskin Disease in Cassava (*Manihot esculenta* Crantz) in Brazil. *Plant Dis* **98**:153. doi:10.1094/PDIS-05-13-0499-PDN
- Olsen KM, Schaal BA. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proc Natl Acad Sci* **96**:5586 LP – 5591. doi:10.1073/pnas.96.10.5586
- Onyeka T, Dixon A, Bandyopadhy R, Okechukwu R. 2004. Distribution and current status of bacterial blight and fungal diseases of cassava in Nigeria.
- Otim-Nape GW. 1980. Cassava Bacterial Blight in Uganda. *Trop Pest Manag* **26**:274–277. doi:10.1080/09670878009414412
- Pacumbaba RP. 1987. A Screening Method for Detecting Resistance Against Cassava Bacterial Blight Disease. *J Phytopathol* **119**:1–6. doi:10.1111/j.1439-0434.1987.tb04378.x
- Patané JSL, Martins J, Rangel LT, Belasque J, Digiampietri LA, Facincani AP, Ferreira RM, Jaciani FJ, Zhang Y, Varani AM, Almeida NF, Wang N, Ferro JA, Moreira LM, Setubal JC. 2019. Origin and diversification of *Xanthomonas citri* subsp. *citri* pathotypes revealed by inclusive phylogenomic, dating, and biogeographic analyses. *BMC Genomics* **20**:700. doi:10.1186/s12864-019-6007-4
- Peng A, Chen S, Lei T, Xu L, He Y, Wu L, Yao L, Zou X. 2017. Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. *Plant Biotechnol J* **15**:1509–1519. doi:10.1111/pbi.12733
- 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. doi:10.1186/s12864-015-2348-9
- Peng Z, Hu Y, Zhang J, Huguet-Tapia JC, Block AK, Park S, Sapkota S, Liu Z, Liu S, White FF. 2019. *Xanthomonas translucens* commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. *Proc Natl Acad Sci* **116**:20938 LP – 20946. doi:10.1073/pnas.1911660116
- Pereira A, Carazzolle M, Abe V, de Oliveira M, Domingues M, Silva J, Cernadas R, Benedetti C. 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.
- Pereira LF, Goodwin PH, Erickson L. 2000. Peroxidase Activity During Susceptible and Resistant Interactions Between Cassava (*Manihot esculenta*) and *Xanthomonas axonopodis* pv. *manihotis* and *Xanthomonas cassavae*. *J Phytopathol* **148**:575–578. doi:10.1111/j.1439-0434.2000.00548.x
- Pereira LF, Goodwin PH, Erickson L. 1999. The Role of a Phenylalanine Ammonia Lyase Gene during Cassava Bacterial Blight and Cassava Bacterial Necrosis. *J Plant Res* **112**:51–60. doi:10.1007/PL00013858
- Pérez-Quintero AL, Lamy L, Zarate CA, Cunnac S, Doyle E, Bogdanove A, Szurek B, Dereeper A. 2017. daTALbase: A Database for Genomic and Transcriptomic Data Related to TAL Effectors. *Mol Plant-Microbe Interact* **31**:471–480. doi:10.1094/MPMI-06-17-0153-FI
- Pérez-Quintero ÁL, Quintero A, Urrego O, Vanegas P, López Carrascal CE. 2012. Bioinformatic identification of cassava miRNAs differentially expressed in response to infection by *Xanthomonas axonopodis* pv. *manihotis*. *BMC Plant Biol* **12**:29. doi:10.1186/1471-2229-12-29
- Perez-Quintero AL, Szurek B. 2019. A Decade Decoded: Spies and Hackers in the History of TAL Effectors Research. *Annu Rev Phytopathol* **57**:459–481. doi:10.1146/annurev-phyto-082718-100026
- Perreaux D, Maraite H, Meyer JA. 1986. Detection of 3(methylthio) propionic acid in cassava leaves infected by *Xanthomonas campestris* pv. *manihotis*. *Physiol Mol Plant Pathol* **28**:323–328. doi:https://doi.org/10.1016/S0048-4059(86)80074-1
- Perreaux D, Maraite H, Meyer JA. 1982. Identification of 3(methylthio) propionic acid as a blight-inducing toxin produced by *Xanthomonas campestris* pv. *manihotis* in vitro. *Physiol Plant Pathol* **20**:313–319. doi:https://doi.org/10.1016/0048-4059(82)90056-X
- Persley GJ. 1978. Studies on the epidemiology and ecology of cassava bacterial blight In: Terry ER, Persley GJ, Cook SCA, editors. Workshop on Cassava Bacterial Blight in Africa, Past, Present and Future. Ibadan, Nigeria: Centre for Overseas Pest Research. pp. 5–7.
- Persley GJ. 1976. Distribution and importance of cassava bacterial blight in Africa In: Persley JG, Terry ER, MacIntyre R, editors. Cassava Bacterial Blight, Report on an Interdisciplinary Workshop. pp. 9–14.
- Prochnik S, Marri PR, Desany B, Rabinowicz PD, Kodira C, Mohiuddin M, Rodriguez F, Fauquet C, Tohme J, Harkins T, Rokhsar DS, Rounsley S. 2012. The Cassava Genome: Current Progress, Future Directions. *Trop Plant Biol* **5**:88–94.

- Pushpalatha R, Gangadharan B. 2020. Is Cassava (*Manihot esculenta* Crantz) a Climate “Smart” Crop? A Review in the Context of Bridging Future Food Demand Gap. *Trop Plant Biol* **13**:201–211. doi:10.1007/s12042-020-09255-2
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**:1173–1183. doi:10.1016/j.cell.2013.02.022
- Rabbi I, Hamblin M, Gedil M, Kulakow P, Ferguson M, Ikpan AS, Ly D, Jannink J-L. 2014. Genetic Mapping Using Genotyping-by-Sequencing in the Clonally Propagated Cassava. *Crop Sci* **54**:1384–1396. doi:10.2135/cropsci2013.07.0482
- Rabbi IY, Hamblin MT, Kumar PL, Gedil MA, Ikpan AS, Jannink J-L, Kulakow PA. 2014. High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res* **186**:87–96. doi:https://doi.org/10.1016/j.virusres.2013.12.028
- Rache L, Blondin L, Flores C, Trujillo C, Szurek B, Restrepo S, Koebnik R, Bernal A, Vernière C. 2019. An Optimized Microsatellite Scheme for Assessing Populations of *Xanthomonas phaseoli* pv. *manihotis*. *Phytopathology*TM **109**:859–869. doi:10.1094/PHYTO-06-18-0210-R
- Ravi V, Ravindran CS, Ramesh V. 2008. The impact of climate change on photosynthesis and productivity of cassava and sweet potato: Effect of rise in temperature, CO₂ and UV-B Radiation: an Overview. *J Root Crop* **34**:95–107.
- Read AC, Moscou MJ, Zimin A V, Peretea G, Meyer RS, Purugganan MD, Leach JE, Triplett LR, Salzberg SL, Bogdanove AJ. 2020. Genome assembly and characterization of a complex zBED-NLR gene-containing disease resistance locus in Carolina Gold Select rice with Nanopore sequencing. *PLOS Genet* **16**:e1008571.
- Read AC, Rinaldi F, Hutin M, He Y, Triplett L, Bogdanove AJ. 2016. Suppression of Xo1-mediated disease resistance in rice by a truncated, non-DNA-binding TAL effector of *Xanthomonas oryzae*. *Front Plant Sci* **7**.
- Restrepo Rubio JS, López Carrascal CE, Melgarejo LM. 2017. Physiological behavior of cassava plants (*Manihot esculenta* Crantz) in response to infection by *Xanthomonas axonopodis* pv. *manihotis* under greenhouse conditions. *Physiol Mol Plant Pathol* **100**:136–141. doi:https://doi.org/10.1016/j.pmpp.2017.09.004
- Restrepo S, Duque MC, Verdier V. 2000a. Resistance spectrum of selected *Manihot esculenta* genotypes under field conditions. *F Crop Res* **65**:69–77. doi:https://doi.org/10.1016/S0378-4290(99)00072-6
- Restrepo S, Duque MC, Verdier V. 2000b. Characterization of pathotypes among isolates of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Plant Pathol* **49**:680–687. doi:10.1046/j.1365-3059.2000.00513.x
- Restrepo S, Velez CM, Duque MC, Verdier V. 2004. Genetic Structure and Population Dynamics of *Xanthomonas axonopodis* pv. *manihotis* in Colombia from 1995 to 1999. *Appl Environ Microbiol* **70**:255 LP – 261. doi:10.1128/AEM.70.1.255-261.2004
- Restrepo S, Vélez CM, Verdier V. 2000c. Measuring the Genetic Diversity of *Xanthomonas axonopodis* pv. *manihotis* Within Different Fields in Colombia. *Phytopathology* **90**:683–690.

doi:10.1094/PHYTO.2000.90.7.683

- Restrepo S, Verdier V. 1997. Geographical Differentiation of the Population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Appl Environ Microbiol* **63**:4427–4434.
- Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* **30**:460–465. doi:10.1038/nbt.2170
- Rodrigues SD, Karimi M, Impens L, Van Lerberge E, Coussens G, Aesaert S, Rombaut D, Holtappels D, Ibrahim HMM, Van Montagu M, Wagemans J, Jacobs TB, De Coninck B, Pauwels L. 2021. Efficient CRISPR-mediated base editing in *Agrobacterium* spp. *Proc Natl Acad Sci* **118**:e2013338118. doi:10.1073/pnas.2013338118
- Roeschlin RA, Uviedo F, García L, Molina MC, Favaro MA, Chiesa MA, Tasselli S, Franco-Zorrilla JM, Forment J, Gadea J, Marano MR. 2019. PthA4AT, a 7.5-repeats transcription activator-like (TAL) effector from *Xanthomonas citri* ssp. *citri*, triggers citrus canker resistance. *Mol Plant Pathol* **20**:1394–1407. doi:https://doi.org/10.1111/mpp.12844
- Román V, Bossa-Castro AM, Vásquez AX, Bernal-Galeano V, Schuster M, Bernal N, López Carrascal CE. 2014. Construction of a cassava PR protein-interacting network during *Xanthomonas axonopodis* pv. *manihotis* infection. *Plant Pathol* **63**:792–802. doi:https://doi.org/10.1111/ppa.12155
- 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 (80-)* **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 Phytol* **187**:1048–1057. doi:10.1111/j.1469-8137.2010.03217.x
- Rossier O, Wengelnik K, Hahn K, Bonas U. 1999. The *Xanthomonas* Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. *Proc Natl Acad Sci U S A* **96**:9368–9373. doi:10.1073/pnas.96.16.9368
- 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. doi:10.1186/s12864-017-4087-6
- Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. 2012. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* **7**:171–192. doi:10.1038/nprot.2011.431
- Santaella M, Suárez E, López Carrascal CE, González C, Mosquera G, Restrepo S, Tohme J, Badillo A, Verdier V. 2004. Identification of genes in cassava that are differentially expressed during infection with *Xanthomonas axonopodis* pv. *manihotis*. *Mol Plant Pathol* **5**:549–558. doi:10.1111/j.1364-3703.2004.00254.x
- Schandry N, de Lange O, Prior P, Lahaye T. 2016. TALE-Like Effectors Are an Ancestral Feature of the *Ralstonia solanacearum* Species Complex and Converge in DNA Targeting Specificity. *Front Plant Sci*.
- Schandry N, Jacobs JM, Szurek B, Perez-Quintero AL. 2018. A cautionary TALE: how plant breeding

- may have favoured expanded TALE repertoires in *Xanthomonas*. *Mol Plant Pathol* **19**:1297–1301. doi:10.1111/mpp.12670
- Schmid-Burgk JL, Schmidt T, Kaiser V, Höning K, Hornung V. 2013. A ligation-independent cloning technique for high-throughput assembly of transcription activator–like effector genes. *Nat Biotechnol* **31**:76–81. doi:10.1038/nbt.2460
- 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 J* **37**:46–60. doi:10.1046/j.1365-313X.2003.01937.x
- 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 Phytol* **179**:546–556. doi:10.1111/j.1469-8137.2008.02487.x
- Schornack S, Moscou MJ, Ward ER, Horvath DM. 2013. Engineering Plant Disease Resistance Based on TAL Effectors. *Annu Rev Phytopathol* **51**:383–406. doi:10.1146/annurev-phyto-082712-102255
- Schreiber T, Bonas U. 2014. Repeat 1 of TAL effectors affects target specificity for the base at position zero. *Nucleic Acids Res* **42**:7160–7169. doi:10.1093/nar/gku341
- Schreiber T, Sorgatz A, List F, Blüher D, Thieme S, Wilmanns M, Bonas U. 2015. Refined Requirements for Protein Regions Important for Activity of the TALE AvrBs3. *PLoS One* **10**:e0120214.
- 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. *Proc Natl Acad Sci* **114**:E897 LP-E903. doi:10.1073/pnas.1620407114
- Sgro GG, Oka GU, Souza DP, Cenens W, Bayer-Santos E, Matsuyama BY, Bueno NF, dos Santos TR, Alvarez-Martinez CE, Salinas RK, Farah CS. 2019. Bacteria-Killing Type IV Secretion Systems. *Front Microbiol*.
- Shakya M, Ahmed SA, Davenport KW, Flynn MC, Lo C-C, Chain PSG. 2020. Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life. *Sci Rep* **10**:1723. doi:10.1038/s41598-020-58356-1
- Sharda M, Badrinarayanan A, Seshasayee ASN. 2020. Evolutionary and Comparative Analysis of Bacterial Nonhomologous End Joining Repair. *Genome Biol Evol* **12**:2450–2466. doi:10.1093/gbe/evaa223
- Sharma V, Sanghera GS, Kashyap P, Sharma B, Chandel C. 2013. RNA interference: A novel tool for plant disease management. *AFRICAN J Biotechnol* **12**:2303–2312. doi:10.5897/AJB12.2791
- Showmaker KC, Arick 2nd MA, Hsu C-Y, Martin BE, Wang X, Jia J, Wubben MJ, Nichols RL, Allen TW, Peterson DG, Lu S-E. 2017. The genome of the cotton bacterial blight pathogen *Xanthomonas citri* pv. *malvacearum* strain MSCT1. *Stand Genomic Sci* **12**:42. doi:10.1186/s40793-017-0253-3
- Soto-Sedano JC, Mora-Moreno RE, Mathew B, León J, Gómez-Cano FA, Ballvora A, López Carrascal CE. 2017. Major Novel QTL for Resistance to Cassava Bacterial Blight Identified through a

- Multi-Environmental Analysis. *Front Plant Sci* **8**:1169. doi:10.3389/fpls.2017.01169
- Soto JC, Ortiz JF, Perlaza-Jiménez L, Vásquez AX, Lopez-Lavalle LAB, Mathew B, León J, Bernal AJ, Ballvora A, López Carrascal CE. 2015. A genetic map of cassava (*Manihot esculenta* Crantz) with integrated physical mapping of immunity-related genes. *BMC Genomics* **16**:190. doi:10.1186/s12864-015-1397-4
- Strauß T, van Poecke RMP, Strauß Annett, Römer P, Minsavage G V, Singh S, Wolf C, Strauß Axel, Kim S, Lee H-A, Yeom S-I, Parniske M, Stall RE, Jones JB, Choi D, Prins M, Lahaye T. 2012. RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proc Natl Acad Sci* **109**:19480–19485.
- 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 Phytol* **200**:808–819. doi:10.1111/nph.12411
- 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 *OstFIIAγ1* and *OstFX1* during bacterial blight of rice. *Proc Natl Acad Sci* **104**:10720–10725.
- Sun X, Cao Y, Yang Z, Xu C, Li X, Wang S, Zhang Q. 2004. Xa26, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant J* **37**:517–527. doi:https://doi.org/10.1046/j.1365-313X.2003.01976.x
- Szurek B, Marois E, Bonas U, Van den Ackerveken G. 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 J* **26**:523–534. doi:10.1046/j.0960-7412.2001.01046.x
- Szurek B, Rossier O, Hause G, Bonas U. 2002. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol Microbiol* **46**:13–23. doi:10.1046/j.1365-2958.2002.03139.x
- Tang X, Wang X, Huang Y, Ma L, Jiang X, Rao MJ, Xu Y, Yin P, Yuan M, Deng X, Xu Q. 2021. Natural variations of TFIIAγ gene and LOB1 promoter contribute to citrus canker disease resistance in *Atalantia buxifolia*. *PLOS Genet* **17**:e1009316.
- Tao Z, Liu H, Qiu D, Zhou Y, Li X, Xu C, Wang S. 2009. A Pair of Allelic WRKY Genes Play Opposite Roles in Rice-Bacteria Interactions. *Plant Physiol* **151**:936 LP – 948. doi:10.1104/pp.109.145623
- Tappiban P, Sraphet S, Srisawad N, Smith DR, Triwitayakorn K. 2018. Identification and expression of genes in response to cassava bacterial blight infection. *J Appl Genet* **59**:391–403. doi:10.1007/s13353-018-0457-2
- Taylor RK, Griffin RL, Jones LM, Pease B, Tsatsia F, Fanai C, Macfarlane B, Dale CJ, Davis RI. 2017. First record of *Xanthomonas axonopodis* pv. *manihotis* in Solomon Islands. *Australas Plant Dis Notes* **12**:49. doi:10.1007/s13314-017-0275-0
- Teper D, Wang N. 2021. Consequences of adaptation of TAL effectors on host susceptibility to *Xanthomonas*. *PLOS Genet* **17**:e1009310.
- Tewe OO, FAO R (Italy). PP and PD eng, Lutaladio N. 2004. Cassava for livestock feed in sub-

Saharan Africa.

- Tian D, Wang J, Zeng X, Gu K, Qiu C, Yang X, Zhou Z, Goh M, Luo Y, Murata-Hori M, White FF, Yin Z. 2014. The Rice TAL Effector–Dependent Resistance Protein XA10 Triggers Cell Death and Calcium Depletion in the Endoplasmic Reticulum. *Plant Cell* **26**:497–515. doi:10.1105/tpc.113.119255
- Timilsina S, Potnis N, Newberry EA, Liyanapathirana P, Iruegas-Bocardo F, White FF, Goss EM, Jones JB. 2020. *Xanthomonas* diversity, virulence and plant–pathogen interactions. *Nat Rev Microbiol*. doi:10.1038/s41579-020-0361-8
- Todor H, Silvis MR, Osadnik H, Gross CA. 2021. Bacterial CRISPR screens for gene function. *Curr Opin Microbiol* **59**:102–109. doi:https://doi.org/10.1016/j.mib.2020.11.005
- Tomkins J, Fregene M, Main D, Kim H, Wing R, Tohme J. 2004. Bacterial artificial chromosome (BAC) library resource for positional cloning of pest and disease resistance genes in cassava (*Manihot esculenta* Crantz). *Plant Mol Biol* **56**:555–561. doi:10.1007/s11103-004-5045-7
- Tran TT, Pérez-Quintero AL, Wonni I, Carpenter SCD, Yu Y, Wang L, Leach JE, Verdier V, Cunnac S, Bogdanove AJ, Koebnik R, Hutin M, Szurek B. 2018. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLOS Pathog* **14**:e1007092.
- Triplett LR, Cohen SP, Heffelfinger C, Schmidt CL, Huerta AI, Tekete C, Verdier V, Bogdanove AJ, Leach JE. 2016. A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. *Plant J* **87**:472–483. doi:10.1111/tpj.13212
- Trujillo C, Arias-Rojas N, Poulin L, Medina C, Tapiero A, Restrepo S, Koebnik R, Bernal A. 2014a. Population typing of the causal agent of cassava bacterial blight in the Eastern Plains of Colombia using two types of molecular markers. *Bmc Microbiol* **14**:161.
- Trujillo C, Ochoa JC, Mideros MF, Restrepo S, López Carrascal CE, Bernal A. 2014b. A complex population structure of the cassava pathogen *Xanthomonas axonopodis* pv. *manihotis* in recent years in the Caribbean Region of Colombia. *Microb Ecol* **68**:155–167. doi:10.1007/s00248-014-0411-8
- Umanah EE. 1977. Cassava Research in Nigeria Before 1972 In: Cock J, MacIntyre R, Graham M, editors. Fourth Tropical Root Crops Symposium. pp. 137–141.
- Umehura Y, Kawano K. 1983. Field Assessment and Inheritance of Resistance to Cassava Bacterial Blight1. *Crop Sci* **23**. doi:10.2135/cropsci1983.0011183X002300060025x
- Van Den Mooter M, Maraitte H, Meiresonne L, Swings J, Gillis M, Kersters K, De Ley J, Perreux D. 1987. Comparison Between *Xanthomonas campestris* pv. *manihotis* (ISPP List 1980) and *X. campestris* pv. *cassavae* (ISPP List 1980) by Means of Phenotypic, Protein Electrophoretic, DNA Hybridization and Phytopathological Techniques. *Microbiology* **133**:57–71. doi:https://doi.org/10.1099/00221287-133-1-57
- Vanderschuren H, Alder A, Zhang P, Gruissem W. 2009. Dose-dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. *Plant Mol Biol* **70**:265–272. doi:10.1007/s11103-009-9472-3

- Vauterin G, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. International journal of systematic bacteriology (USA).
- Veley KM, Okwuonu I, Jensen G, Yoder M, Taylor NJ, Meyers BC, Bart RS. 2020. Visualizing cassava bacterial blight at the molecular level using CRISPR-mediated homology-directed repair. *bioRxiv* 2020.05.14.090928. doi:10.1101/2020.05.14.090928
- Vera Cruz CM, Bai J, Oña I, Leung H, Nelson RJ, Mew T-W, 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. *Proc Natl Acad Sci* **97**:13500 LP – 13505. doi:10.1073/pnas.250271997
- Verdier V, Boher B, Maraite H, Geiger J-P. 1994. Pathological and Molecular Characterization of *Xanthomonas campestris* Strains Causing Diseases of Cassava (*Manihot esculenta*). *Appl Environ Microbiol* **60**:4478 LP – 4486.
- Verdier V, Dongo P, Boher B. 1993. Assessment of genetic diversity among strains of *Xanthomonas campestris* pv. *manihotis*. *J Gen Microbiol* **139**:2591–2601. doi:10.1099/00221287-139-11-2591
- Verdier V, Mosquera G. 1999. Specific Detection of *Xanthomonas axonopodis* pv. *manihotis* with a DNA Hybridization Probe. *J Phytopathol* **147**:417–423. doi:10.1111/j.1439-0434.1999.tb03843.x
- Verdier V, Mosquera G, Assigbétsé K. 1998. Detection of the Cassava Bacterial Blight Pathogen, *Xanthomonas axonopodis* pv. *manihotis*, by Polymerase Chain Reaction. *Plant Dis* **82**:79–83. doi:10.1094/PDIS.1998.82.1.79
- Verdier V, Ojeda S, Mosquera G. 2001a. Methods for detecting the cassava bacterial blight pathogen: A practical approach for managing the disease. *Euphytica* **120**:103–107. doi:10.1023/A:1017516007945
- Verdier V, Restrepo S, Mosquera G, Duque MC, Gerstl A, Laberry R. 2001b. Genetic and pathogenic variation of *Xanthomonas axonopodis* pv. *manihotis* in Venezuela. *Plant Pathol* **47**:601–608. doi:10.1046/j.1365-3059.1998.00271.x
- Verdier V, Restrepo S, Mosquera G, Jorge V, López Carrascal CE. 2004. Recent progress in the characterization of molecular determinants in the *Xanthomonas axonopodis* pv. *manihotis*–cassava interaction. *Plant Mol Biol* **56**:573–584. doi:10.1007/s11103-004-5044-8
- Verdier V, Schmit J, Lemattre M. 1990. Étude en microscopie électronique à balayage de l'installation de deux souches de *Xanthomonas campestris* pv *manihotis* sur feuilles de vitroplants de manioc. *Agronomie* **10**:93–102.
- Wada N, Ueta R, Osakabe Y, Osakabe K. 2020. Precision genome editing in plants: state-of-the-art in CRISPR/Cas9-based genome engineering. *BMC Plant Biol* **20**:234. doi:10.1186/s12870-020-02385-5
- Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, Qin T, Li Y, Che J, Zhang M, Yang B, Liu Y, Zhao K. 2015. XA23 Is an Executor R Protein and Confers Broad-Spectrum Disease Resistance in Rice. *Mol Plant* **8**:290–302. doi:10.1016/j.molp.2014.10.010
- Wang F, Zhang H, Gao J, Chen F, Chen S, Zhang C, Peng G. 2016. Rapid and accurate synthesis of

- TALE genes from synthetic oligonucleotides. *Biotechniques* **60**:299–305. doi:10.2144/000114422
- Wang G, Ding X, Yuan M, Qiu D, Li X, Xu C, Wang S. 2006. Dual Function of Rice OsDR8 Gene in Disease Resistance and Thiamine Accumulation. *Plant Mol Biol* **60**:437–449. doi:10.1007/s11103-005-4770-x
- Wang J, Zeng X, Tian D, Yang X, Wang L, Yin Z. 2018. The pepper Bs4C proteins are localized to the endoplasmic reticulum (ER) membrane and confer disease resistance to bacterial blight in transgenic rice. *Mol Plant Pathol* **19**:2025–2035. doi:10.1111/mpp.12684
- Wang W, Feng B, Xiao J, Xia Z, Zhou X, Li P, Zhang W, Wang Y, Møller BL, Zhang P, Luo M-C, Xiao G, Liu J, Yang J, Chen S, Rabinowicz PD, Chen X, Zhang H-B, Ceballos H, Lou Q, Zou M, Carvalho LJC, Zeng C, Xia J, Sun S, Fu Y, Wang H, Lu C, Ruan M, Zhou S, Wu Z, Liu H, Kannangara RM, Jørgensen K, Neale RL, Bonde M, Heinz N, Zhu W, Wang S, Zhang Y, Pan K, Wen M, Ma P-A, Li Z, Hu M, Liao W, Hu W, Zhang S, Pei J, Guo A, Guo J, Zhang J, Zhang Z, Ye J, Ou W, Ma Y, Liu X, Tallon LJ, Galens K, Ott S, Huang J, Xue J, An F, Yao Q, Lu X, Fregene M, López-Lavalle LAB, Wu J, You FM, Chen M, Hu S, Wu G, Zhong S, Ling P, Chen Y, Wang Q, Liu G, Liu B, Li K, Peng M. 2014. Cassava genome from a wild ancestor to cultivated varieties. *Nat Commun* **5**:5110. doi:10.1038/ncomms6110
- Webb KM, Oña I, Bai J, Garrett KA, Mew T, Vera Cruz CM, Leach JE. 2010. A benefit of high temperature: increased effectiveness of a rice bacterial blight disease resistance gene. *New Phytol* **185**:568–576. doi:https://doi.org/10.1111/j.1469-8137.2009.03076.x
- Wei Y, Chang Y, Zeng H, Liu G, He C, Shi H. 2018a. RAV transcription factors are essential for disease resistance against cassava bacterial blight via activation of melatonin biosynthesis genes. *J Pineal Res* **64**:e12454. doi:10.1111/jpi.12454
- Wei Y, Hu W, Wang Q, Liu W, Wu C, Zeng H, Yan Y, Li X, He C, Shi H. 2016. Comprehensive transcriptional and functional analyses of melatonin synthesis genes in cassava reveal their novel role in hypersensitive-like cell death. *Sci Rep* **6**:35029. doi:10.1038/srep35029
- Wei Y, Liu G, Bai Y, Xia F, He C, Shi H. 2017. Two transcriptional activators of N-acetylserotonin O-methyltransferase 2 and melatonin biosynthesis in cassava. *J Exp Bot* **68**. doi:10.1093/jxb/erx305
- Wei Y, Liu G, Chang Y, He C, Shi H. 2018b. Heat shock transcription factor 3 regulates plant immune response through modulation of salicylic acid accumulation and signaling in cassava. *Mol Plant Pathol* **19**. doi:10.1111/mpp.12691
- Wichmann G, Bergelson J. 2004. Effector genes of *Xanthomonas axonopodis* pv. *vesicatoria* promote transmission and enhance other fitness traits in the field. *Genetics* **166**:693–706. doi:10.1534/genetics.166.2.693
- Wicky BIM, Stenta M, Dal Peraro M. 2013. TAL Effectors Specificity Stems from Negative Discrimination. *PLoS One* **8**:e80261. doi:10.1371/journal.pone.0080261
- Wiehe PO, Dowson WJ. 1953. A bacterial disease of Cassava (*Manihot utilissima*) in Nyasaland. *Emp J Exp Agric* **21**:141–143.
- 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. *Front Plant Sci* **6**:536. doi:10.3389/fpls.2015.00536
- Wilson MC, Mutka AM, Hummel AW, Berry J, Chauhan RD, Vijayaraghavan A, Taylor NJ, Voytas DF, Chitwood DH, Bart RS. 2017. Gene expression atlas for the food security crop cassava. *New Phytol* **213**:1632–1641. doi:https://doi.org/10.1111/nph.14443
- 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 Physiol* **148**:1497 LP – 1509. doi:10.1104/pp.108.123356
- Wydra K, Banito A, Kpémoua KE. 2007. Characterization of resistance of cassava genotypes to bacterial blight by evaluation of leaf and systemic symptoms in relation to yield in different ecozones. *Euphytica* **155**:337–348. doi:10.1007/s10681-006-9335-9
- Wydra K, Fanou A. 2015. Removal of symptomatic cassava leaves as cultural practice to control cassava bacterial blight. *Int J Plant Pathol* **3**:117–124.
- Wydra K, Verdier V. 2002. Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. *Agric Ecosyst Environ* **93**:211–226. doi:https://doi.org/10.1016/S0167-8809(01)00349-8
- Wydra K, Zinsou V, Jorge V, Verdier V. 2004. Identification of Pathotypes of *Xanthomonas axonopodis* pv. *manihotis* in Africa and Detection of Quantitative Trait Loci and Markers for Resistance to Bacterial Blight of Cassava. *Phytopathology*TM **94**:1084–1093. doi:10.1094/PHYTO.2004.94.10.1084
- Xiang Y, Cao Y, Xu C, Li X, Wang S. 2006. Xa3, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as Xa26. *Theor Appl Genet* **113**:1347–1355. doi:10.1007/s00122-006-0388-x
- Xiao Jun, Cheng H, Li X, Xiao Jinghua, Xu C, Wang S. 2013. Rice WRKY13 Regulates Cross Talk between Abiotic and Biotic Stress Signaling Pathways by Selective Binding to Different cis-Elements. *Plant Physiol* **163**:1868 LP – 1882. doi:10.1104/pp.113.226019
- Xie Z, Nolan TM, Jiang H, Yin Y. 2019. AP2/ERF Transcription Factor Regulatory Networks in Hormone and Abiotic Stress Responses in Arabidopsis. *Front Plant Sci* **10**:228. doi:10.3389/fpls.2019.00228
- Yan Y, He X, Hu W, Liu G, Wang P, He C, Shi H. 2018. Functional analysis of MeCIPK23 and MeCBL1/9 in cassava defense response against *Xanthomonas axonopodis* pv. *manihotis*. *Plant Cell Rep* **37**:887–900. doi:10.1007/s00299-018-2276-7
- Yan Y, Wang P, He C, Shi H. 2017. MeWRKY20 and its interacting and activating autophagy-related protein 8 (MeATG8) regulate plant disease resistance in cassava. *Biochem Biophys Res Commun* **494**:20–26. doi:https://doi.org/10.1016/j.bbrc.2017.10.091
- Yang B, Sugio A, White FF. 2006. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc Natl Acad Sci U S A* **103**:10503–10508. doi:10.1073/pnas.0604088103
- Yang J, Yuan P, Wen D, Sheng Y, Zhu S, Yu Y, Gao X, Wei W. 2013. ULtiMATE System for Rapid Assembly of Customized TAL Effectors. *PLoS One* **8**:e75649.

- Yang Y, Gabriel DW. 1995. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. *Mol Plant Microbe Interact* **8**:627–631. doi:10.1094/mpmi-8-0627
- Yang Yinong Y, De Feyter R, Gabriel DW. 1994. Host-specific symptoms and increased release of *Xanthomonas citri* and *X.campestris* pv. *malvacearum* from leaves are determined by the 102-bp tandem repeats of *pthA* and *avrb6*, respectively. *Mol Plant-Microbe Interact* **7**:345–355. doi:10.1094/mpmi-7-0345
- Yoodee S, Kobayashi Y, Songnuan W, Boonchird C, Thitamadee S, Kobayashi I, Narangajavana J. 2018. Phytohormone priming elevates the accumulation of defense-related gene transcripts and enhances bacterial blight disease resistance in cassava. *Plant Physiol Biochem* **122**:65–77. doi:https://doi.org/10.1016/j.plaphy.2017.11.016
- Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang ZX, Kono I, Kurata N, Yano M, Iwata N, Sasaki T. 1998. Expression of *Xa1*, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. *Proc Natl Acad Sci U S A* **95**:1663–1668. doi:10.1073/pnas.95.4.1663
- Yu Y, Streubel J, Balergue 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. *Mol Plant-Microbe Interact* **24**:1102–1113. doi:10.1094/mpmi-11-10-0254
- 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 Cell Physiol* **50**:947–955. doi:10.1093/pcp/pcp046
- Yuan M, Ke Y, Huang R, Ma L, Yang Z, Chu Z, Xiao J, Li X, Wang S. 2016. A host basal transcription factor is a key component for infection of rice by TALE-carrying bacteria. *Elife* **5**:e19605. doi:10.7554/eLife.19605
- 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. doi:10.1111/j.1439-0434.2007.01210.x
- Zeng H, Xie Y, Liu G, Lin D, He C, Shi H. 2018. Molecular identification of GAPDHs in cassava highlights the antagonism of MeGAPCs and MeATG8s in plant disease resistance against cassava bacterial blight. *Plant Mol Biol* **97**:201–214. doi:10.1007/s11103-018-0733-x
- Zhang J, Huguet -Tapia JC, Hu Y, Jones J, Wang N, Liu S, White FF. 2017. Homologues of CsLOB1 in citrus function as disease susceptibility genes in citrus canker. *Mol Plant Pathol* **18**:798–810. doi:https://doi.org/10.1111/mpp.12441
- Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom J-S, Huang S, Liu S, Vera Cruz C, Frommer WB, White FF, Yang B. 2015. Gene targeting by the TAL effector *PthXo2* reveals cryptic resistance gene for bacterial blight of rice. *Plant J* **82**:632–643. doi:10.1111/tpj.12838
- Zhu W, Yang B, Chittoor JM, Johnson LB, White FF. 1998. *AvrXa10* Contains an Acidic Transcriptional Activation Domain in the Functionally Conserved C Terminus. *Mol Plant-Microbe Interact* **11**:824–832. doi:10.1094/MPMI.1998.11.8.824
- Zinsou V, Wydra K, Ahohuendo B, Hau B. 2005. Genotype × environment interactions in symptom

development and yield of cassava genotypes with artificial and natural cassava bacterial blight infections. *Eur J Plant Pathol* **111**:217–233. doi:10.1007/s10658-004-2877-6

Zinsou V, Wydra K, Ahohuendo B, Hau B. 2004. Effect of soil amendments, intercropping and planting time in combination on the severity of cassava bacterial blight and yield in two ecozones of West Africa. *Plant Pathol* **53**:585–595. doi:10.1111/j.0032-0862.2004.01056.x

Zinsou V, Wydra K, Ahohuendo B, Schreiber L. 2006. Leaf Waxes of Cassava (*Manihot esculenta* Crantz) in Relation to Ecozone and Resistance to *Xanthomonas* Blight. *Euphytica* **149**:189–198. doi:10.1007/s10681-005-9066-3

Zuluaga P, Szurek B, Koebnik R, Kroj T, Morel J-B. 2017. Effector Mimics and Integrated Decoys, the Never-Ending Arms Race between Rice and *Xanthomonas oryzae* . *Front Plant Sci* .

Annexes

The following three annexes correspond to three articles that were in the frame of this thesis. The first article corresponds to the first pathogen profile of *Xanthomonas* causing disease in cassava. This document is accepted and under typesetting to be published in *Molecular Plant Pathology* (IF=4.379), therefore we annex the submitted version. This article is product of an exhaustive bibliographic research and derived from the wording of the introduction of this thesis. Adriana Bernal, Camilo López, Boris Szurek, and Valérie Verdier are all plant pathology specialist with substantial experience in the field. Diana Gómez de la Cruz, a PhD student in molecular plant pathology at The Sainsbury Laboratory, synthesized the information of several sections and produced the artworks of the article.

The second annex corresponds to an article entitled “daTALbase: A Database for Genomic and Transcriptomic Data Related to TAL Effectors”, published in the *Molecular Plant-Microbe Interactions* journal (IF=3.696). I contributed to this work by providing results from analyses performed in *Xpm* that assessed the TALome diversity and affinity. I also reviewed and edited the manuscript before submission. This work resulted in a web-based interface that allow exploring TALE biology in several pathosystems (<http://bioinfo-web.mpl.ird.fr/cgi-bin2/datalbase/home.cgi>).

The last annex corresponds to an article entitled “Development of a duplex-PCR for differential diagnosis of *Xanthomonas phaseoli* pv. *manihotis* and *Xanthomonas cassavae* in cassava (*Manihot esculenta*)”, which was published in the *Physiological and Molecular Plant Pathology* journal (IF=1.646). I contributed to this work by carrying out some experiments to demonstrate the specificity of the PCR-based tool. This work resulted in a tool that allows the identification of *Xpm* and *Xc* from cassava infected tissue.

1 **Cassava diseases caused by *Xanthomonas phaseoli* pv. *manihotis* and *Xanthomonas cassavae***

2 Carlos A. Zárate-Chaves¹, Diana Gómez de la Cruz², Valérie Verdier¹, Camilo E. López³, Adriana
3 Bernal^{4*}, Boris Szurek^{1*}

4 ¹PHIM, Univ Montpellier, CIRAD, INRAe, IRD, Montpellier SupAgro, Montpellier, France

5 ²The Sainsbury Laboratory, University of East Anglia, NR4 7UH, Norwich, United Kingdom

6 ³Manihot Biotec, Departamento de Biología, Universidad Nacional de Colombia, 111321, Bogotá,
7 Colombia

8 ⁴Laboratorio de Interacciones Moleculares de Microorganismos Agrícolas, Departamento de
9 Ciencias Básicas, Universidad de los Andes, 111711, Bogotá, Colombia

10 *Corresponding author: boris.szurek@ird.fr, abernal@uniandes.edu.co

11

12 **Running header:** Cassava diseases caused by *Xanthomonas*

13

14 **Keywords:** cassava bacterial blight, cassava bacterial necrosis, quantitative resistance, cassava,
15 *Xanthomonas*.

16

17 **Word count:** 8796

18

19 **Summary**

20 *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) and *Xanthomonas cassavae* (*Xc*) are two bacterial
21 pathogens attacking cassava. Cassava bacterial blight (CBB) is a systemic disease caused by *Xpm*,
22 which might have dramatic effects on plant growth and crop production. Cassava bacterial necrosis
23 (CBN) is a non-vascular disease caused by *Xc* with foliar symptoms similar to CBB, but its impacts on
24 the plant vigor and the crop are limited. In this review, we describe the epidemiology and ecology
25 of the two pathogens, the impacts and management of the diseases, and the main research
26 achievements for each pathosystem. Since *Xc* data is sparse, our main focus is on *Xpm* and CBB.

27

28 **Introduction**

29 Cassava is an important staple crop as a source of food and income for hundreds of millions of
30 people in tropical countries. This major crop is threatened by several pests and pathogens, which
31 significantly affect productivity. Among bacterial pathogens, the vascular and systemic
32 *Xanthomonas phaseoli* pv. *manihotis* has received considerable attention due to its devastating
33 potential in the tropics and scientific importance worldwide (Mansfield *et al.*, 2012). This pathogen
34 profile describes the main hallmarks of this pathosystem and updates several reviews and original
35 articles on the subject. Although more bacterial pathogens infect cassava, information in the
36 literature is scarce. In this regard, we also present a compilation of information for a second
37 xanthomonad infecting cassava, the non-vascular pathogen *Xanthomonas cassavae*.

38

39 **Disease description and epidemiology**

40 ***Cassava bacterial blight symptoms***

41 Cassava bacterial blight (CBB) which is caused by the bacterial pathogen *Xanthomonas phaseoli* pv.
42 *manihotis* (*Xpm*), is characterized by a range of symptoms that mainly affect leaves, petioles, and

43 stems, frequently leading to plant death (Figure 1). Early symptoms appear as brown to dark-brown
44 water-soaked translucent angular spots on the leaf tissue browning at later stages, occasionally
45 surrounded by a chlorotic halo. Veins around these spots discolor and affected tissues frequently
46 produce creamy white and later yellow-to-orange exudates on the lower side of the leaf. Blight
47 results from spot coalescence which creates necrotic areas that become dry and curl the leaflets,
48 giving them the aspect of a superficial burn. As disease progresses, bacteria access the xylem vessels
49 from the mesophyll and move towards the stem through the petioles, which become brown and
50 collapse, causing the leaf to wilt. Vessel colonization in the stem allows *Xpm* to systemically move
51 upwards and downwards. When infection reaches the plant upper part where stem tissues are
52 greener and less lignified, stem rotting leads to dieback characterized by shoot apex wilting. New
53 sprouts can grow from buds located in more basal zones, giving the plant the appearance of a candle
54 stick. However, if these buds are also contaminated by *Xpm* they will eventually wilt. Infected fruits
55 also show water-soaked spots, and the resulting seeds suffer cotyledon and endosperm necrosis,
56 and seed deformation. Roots from highly susceptible cultivars can show delayed symptoms
57 restricted to the vascular tissues, with discolored vascular strands surrounded by dry and rotten
58 spots (Boher *et al.*, 1995; Lozano, 1986; Lozano and Sequeira, 1974b; Maraite and Meyer, 1975).

59 Symptoms of cassava bacterial necrosis (CBN) caused by *Xanthomonas cassavae* (*Xc*) might be
60 similar at a first sight (Figure 1H), but the outcome of the infection is not as devastating (Mostade
61 and Butare, 1978). Infections caused by *Xc* initially produce rounded water-soaked spots surrounded
62 by a yellow chlorotic halo and radial necrosis of the veins (Maraite and Perreux, 1979). As the
63 disease progresses, bacteria colonize the adjacent mesophyll tissues, lesions expand and yellow
64 exudates can be observed in them (Maraite, 1993) Leaves wilt, collapse and dry, but the pathogen
65 is not able to colonize vascular tissues and there is no formation of secondary spots that turn into
66 extended blight areas (Lozano and Sequeira, 1974b; Maraite, 1993; Maraite and Perreux, 1979;

67 Van Den Mooter *et al.*, 1987; Verdier *et al.*, 1994). Stem bark can be locally colonized, potentially
68 through wind-induced wounds in the stem, resulting in cortical necrotic areas. These areas can
69 expand and cause girdling of the stem, which in turn can initially induce turgidity loss of proximal
70 petioles and then wilting (Maraite and Weyns, 1979; Mostade and Butare, 1978).

71 ***Disease cycle***

72 *Xpm* has an initial epiphytic phase where it is able to colonize the surface of cassava aerial tissues.
73 When conditions are favorable, especially humidity, multiplication of the pathogen increases
74 around trichomes and penetration of the outer layer takes place (Daniel and Boher, 1985a). *Xpm*
75 accesses the plant through wounds and natural openings of the leaf, potentially primarily via adaxial
76 stomates (Kemp *et al.*, 2004). Although it is common for other vascular *Xanthomonas* to enter
77 through hydathodes, their role in the initial stages of infection for *Xpm* is not known. Internal tissue
78 colonization is accompanied by the production of highly hygroscopic bacterial exopolysaccharides
79 (EPS), which are exudated from lesions and then hydrated and carried by raindrops. Wind and rain-
80 mediated splashing of these bacterial suspensions is considered the main natural way of horizontal
81 pathogen transmission from one plant to another (Lozano and Sequeira, 1974a). Once *Xpm* hits new
82 cassava leaves, the disease cycle restarts. Since this crop is propagated from cuttings, presence of
83 the pathogen on the propagative material and working tools is the major factor for disease
84 spreading (Lozano and Sequeira, 1974a). Plants sprouted from contaminated cuttings quickly
85 develop the disease and are a major source for secondary infections in the field. Bacteria can also
86 be transmitted inside the sexual seed (Daniel and Boher, 1985b; Elango and Lozano, 1980). Stems
87 can get infected by the dispersal of wind-driven sand or hail (Maraite, 1993).

88

89 ***Ecology of the pathogen***

90 *Xpm* shows two different lifestyles in the field, an epiphytic phase (Daniel and Boher, 1985b; Elango
91 and Lozano, 1981) and a biotrophic parasitic phase that starts when environmental conditions
92 facilitate pathogen growth and entrance into the mesophyll apoplast of cassava leaves (Verdier *et*
93 *al.*, 1990). The epiphytic stage plays an important role ensuring the natural persistence of the
94 pathogen intercropping. Epiphytic populations of *Xpm* have been found on asymptomatic leaves in
95 fields where CBB has been reported. Bacterial titers vary importantly according to environmental
96 conditions, humidity and temperature being two of the most important factors. The optimal
97 temperature for CBB development is 30°C, whereas that for CBN is 25°C (Maraite and Perreux,
98 1979). There is a marked increase of epiphytic populations and a shift to the parasitic phase during
99 the rainy season. Conversely, symptoms of CBB are less frequent and epiphytic populations
100 decrease during the dry season (Daniel and Boher, 1985a; Daniel and Boher, 1985b). *Xpm* has also
101 been detected in several weeds that occur naturally in cassava fields in South America, suggesting
102 that the pathogen can epiphytically survive on them during intercropping (Elango and Lozano,
103 1981). Artificial inoculation of several cassava crop-associated weeds (non-host interactions in all
104 cases) highlighted different degrees of bacterial survival, and pathogen viability for up to 54 days
105 (Fanou *et al.*, 2017; Marcano and Trujillo, 1984).

106 Infected plant debris and soil are also suggested to play a role in *Xpm* field persistence. However,
107 survival of *Xpm* in a free-living state in the soil has been experimentally proven to be limited to up
108 to three weeks (Fanou *et al.*, 2017). In contrast, *Xpm* was shown to survive in slow-decaying dry
109 debris for more than two months and one year under field and controlled conditions respectively
110 (Daniel and Boher, 1985b; Fanou *et al.*, 2017). Insect-mediated dissemination of *Xpm* has been
111 reported. The African grasshopper pest *Zonocerus variegatus* recovered from diseased cassava
112 fields harbored a significant number of infective bacterial cells (Daniel and Boher, 1985b;
113 Zandjanakou-Tachin *et al.*, 2007). The coreid bug *Pseudothrips devastans* has also been reported

114 as facilitator of infection through the generation of punctures through which the bacterium gains
115 access to internal tissue (Maraite and Meyer, 1975).

116

117 ***Distribution of the pathogen***

118 Botanical, geographical and domestication origin of cassava is a controversial topic. A widely-
119 accepted theory suggests that the cultivated cassava species *Manihot esculenta* subsp. *esculenta*
120 arose from the wild *Manihot esculenta* subsp. *flabellifolia* in the Amazon basin, and its
121 domestication began 5000 to 7000 years b.c. in the southern Amazon border region (Allem, 2002;
122 Olsen and Schaal, 1999). Cassava was introduced to Africa by Portuguese traders during the 16th
123 century and was adopted as a staple crop across several countries (Jones, 1959). In Asia it is believed
124 to be introduced from Mexico to the Philippines during the 17th century. In agreement with the
125 geographical origin of the crop, genetic diversity of *Xpm* populations in South America has been
126 shown to be high (Bart et al., 2012; Verdier et al., 2004).

127 CBB was first reported in Brazil in 1912 by Bondar and then identified in different South American
128 countries during and after the 1970s. The disease and the pathogen were detected in Africa for the
129 first time in Nigeria in 1972, and then systematically detected in several Sub-Saharan countries
130 (Hillocks and Wydra, 2002; Persley, 1976). To date, *Xpm* has been reported in 49 countries located
131 all over in the tropics (CABI, 2020a; Taylor *et al.*, 2017). On the other hand, *Xc* has only been reported
132 in the East African countries of Democratic Republic of the Congo, Rwanda, Malawi, Tanzania,
133 Uganda, Kenya, and Burundi (Boher and Verdier, 1994; CABI, 2020b; Maraite, 1993; Mostade and
134 Butare, 1978). Figure 1 shows cassava production areas and the distribution of both pathogens.

135

136 **Disease impacts and management**

137 ***Prevalence, incidence, and losses associated to CBB***

138 Historically CBB is reminiscent of a famine period from 1970 to 1975 in the Democratic Republic of
139 the Congo, where cassava crop diversity was low and food and economic dependence on cassava
140 was considerable (Hillocks and Wydra, 2002; Lozano, 1986; Maraité and Meyer, 1975). Several
141 studies have addressed the distribution and effects of the pathogen on the crop. Figure 2 (and Tables
142 S1, S2, and S3) summarizes data from several reports where incidence, severity and/or yield losses
143 were assessed, mainly in farms or field trials. Incidence varies according to environmental factors,
144 but most of the reports show ranges between 30% and 90%, with an incidence peak between 60%
145 and 70%. Disease severity reports show systemic symptoms in all the surveys (values greater than 2
146 in 1 - 5 scales), even for measurements performed within the three first months after planting
147 (MAP). Severity seems higher between the third and the sixth MAP, and there is a slight decrease in
148 the following months, but complete dieback and plant stunting have been recorded in all the surveys
149 performed after the third MAP.

150 Yield losses are more difficult to measure accurately mainly due to two reasons: i) cassava fields are
151 not only affected by CBB in most of the cases, and ii) as for other diseases, outcomes are dependent
152 on environmental conditions, i.e. plants can counterbalance the negative effects of CBB when
153 favorable growth conditions are given (Harris *et al.*, 2015; Zinsou *et al.*, 2005). However, available
154 metadata suggest that fresh root yield losses (Figure 2) can reach up to 100 %, with a median of
155 about 50% in susceptible varieties; while it reaches up to 76% in resistant varieties, but the median
156 is below 25%. This is an indication of the high potential of *Xpm* to cause important losses.

157

158 ***Economic and social importance of the pathogen***

159 In a recent report of the Evans School Policy Analysis and Research (Harris *et al.*, 2015) for the Bill
160 and Melinda Gates Foundation, economic impacts of CBB and postharvest physiological
161 deterioration were weighed in the context of main constraints for cassava production. Despite its

162 well-known potential to cause important losses, the current impacts of CBB on the crop are masked
163 by the lack of recent surveys due to limited research on the disease, underestimations due to the
164 lack of farmers' training for CBB identification, and the influence of environmental factors like
165 drought and coinfections on the onset and severity of the disease. A two-year follow-up of farmers
166 practices and crop outcomes in Uganda and Kenya revealed that despite the high concern of farmers
167 about the impacts of diseases like CBB and cassava anthracnose disease (CAD), the factors impacting
168 yield losses the most were soil fertility and weed management (Fermont *et al.*, 2009). This latter
169 could result from the cumulative experience of agricultural control measures applied to cassava
170 farming since the 1980s, which may have buffered the impact of CBB (Lozano, 1986).

171 Conversely, CBB dynamics seem to be altered by social behaviors linked mainly to cutting trading,
172 which favors the distribution of the pathogen over short and long distances (Restrepo *et al.*, 2004;
173 Restrepo and Verdier, 1997; Trujillo *et al.*, 2014b; Trujillo *et al.*, 2014a). In a case of study in
174 Colombia, disease incidence is positively correlated with agrochemicals use, land ownership and
175 propagative material sharing. Land ownership limitations due to internal conflict and inequity in
176 cassava farmer communities hampers the long-term establishment of crops and force to inter-cycle
177 renewal of propagative material, which leads to the use of any available sources of stakes that
178 frequently lack of phytosanitary controls (Pérez *et al.*, In prep.).

179

180 ***Control strategies***

181 Efficient CBB control is based on three main pillars: sanitary controls, cultural practices, and
182 deployment of tolerant or resistant varieties. Control practices have been deployed based on our
183 knowledge about *Xpm* and its ecology. Sanitary controls of propagative materials and seeds are
184 large-scale measures that aim at stopping pathogen dispersal. In general these measures comprise

185 the deployment of disease-free materials, treatments to eliminate the pathogen, establish
186 quarantine, and *Xpm* detection tools (Chavarriaga Aguirre *et al.*, 2016; Frison and Feliu, 1991).

187 Cultural practices for CBB management include crop rotation, intercropping, fallowing, removal or
188 burying of crop debris, weed management, delayed planting (at the end of the rainy season), and
189 use of clean propagative materials. Crop rotation and fallowing aim at depleting the pathogen
190 inoculum sources through time; they are deemed as effective at buffering the impact of CBB in
191 successive crop cycles (Lozano, 1986; Persley, 1978). Since weeds are reservoirs of the pathogen,
192 their management during the entire crop cycle, as well as debris treatments before planting, are
193 strongly recommended. Fanou and collaborators (Fanou *et al.*, 2017) showed that *Xpm* survival is
194 markedly reduced when debris is covered with soil or buried (less than 30 days vs more than 120
195 days for non-buried debris under dry conditions). Moreover, Fanou and Wydra (Wydra and Fanou,
196 2015) showed that removal of symptomatic leaves can reduce CBB severity and improve the quality
197 of the crop as source of propagative material; however, this practice did not show effects on yields
198 in their study. Planting during the second half of the rainy season, disease incidence and severity
199 decrease, while yields are maintained or even improved generally (Ambe, 1993; Umemura and
200 Kawano, 1983; Zinsou *et al.*, 2004).

201

202 Considering that cassava is devoted to sustaining mainly low-income farmers, deployment of
203 tolerant or resistant varieties is considered the main solution for CBB control. Clonal selection and
204 breeding for cassava disease resistance started in the 1940s and were mainly focused on CMD.
205 Resistance against CBB is essentially due to genes introgressed from wild *Manihot* species, like the
206 Ceara rubber tree *M. glaziovii* (Nassar, 2007). Several studies addressed the performance of diverse
207 cassava varieties against CBB (Dixon *et al.*, 2002; Fokunang *et al.*, 2000; Lamptey *et al.*, 1998; Lozano
208 and Laberry, 1982; Restrepo *et al.*, 2000b; Umemura and Kawano, 1983; Wydra *et al.*, 2004; Wydra

209 *et al.*, 2007; Zinsou *et al.*, 2005) (see Table S2). In summary, three main aspects condition the
210 successful deployment of CBB resistant varieties: i) interactions of cassava with environmental
211 factors profusely affect the fitness and resistance of some cultivars; ii) the need for co-selection of
212 agronomical traits of value for farmers and the local market; and iii) the resistance against CBB is
213 polygenic, additively inherited (Umemura and Kawano, 1983), pathotype-specific (Wydra *et al.*,
214 2004) and its molecular basis needs to be elucidated.

215 Regarding the first aspect, several studies (Dixon *et al.*, 2002; Restrepo *et al.*, 2000b; Wydra *et al.*,
216 2007; Zinsou *et al.*, 2005) demonstrated that the environment has a significantly greater influence
217 on CBB disease incidence than the genotypic component, and that the interaction of both factors
218 can mask differences between genotypes. For instance, the highly resistant variety TMS30572
219 performed as resistant in two edaphoclimatic zones of Benin, while it was moderately resistant in
220 four other climatic zones (Zinsou *et al.*, 2005). Therefore researchers draw attention to test
221 resistance in relevant conditions, with high disease pressure, and in parallel in greenhouse settings
222 where pathogenicity assays are optimal (Restrepo *et al.*, 2000b; Zinsou *et al.*, 2005). Regarding the
223 second aspect, surveys have shown that smallholders prioritize higher yields, taste, and good
224 cooking qualities over diseases tolerance when selecting materials, which is highly relevant when
225 engineering disease resistance (Harris *et al.*, 2015). The latter aspect - cellular, genetic, and
226 molecular bases of the resistance - is addressed later in this document.

227 Despite the above-mentioned caveats and to summarize the attempts to find high-performing
228 cassava varieties (see Table S5), the outstanding resistance against CBB of genotype TMS30572 has
229 been widely demonstrated in several trials in African countries. This genotype comes from the CBB-
230 resistant parent 58308, a low-productive interspecific hybrid with *M. glaziovii*, and the susceptible,
231 but high yielding Brazilian variety Branca Caterina de Santa (<https://seedtracker.org/cassava/>
232 (Nassar, 2007; Umanah, 1977)). Although performance of cassava varieties tested in South American

233 country fields has been highly variable, genotypes CMC40, MECU82, MCOL1916, MPAN19,
234 MPAN12B, MBRA685, MBRA886, MBRA902, MNGA2, CM523-7, and CM6438-14 were considered
235 resistant against different *Xpm* strains in a set of studies (Table S5).

236

237 **Pathogen description**

238 ***Pathogen identification***

239 *Xpm* and *Xc* are gammaproteobacteria that belong to the Xanthomonadaceae family. *Xpm* is a Gram
240 negative rod with a polar flagellum that forms shiny, slimy, convex and circular colonies with entire
241 margins, when sugars as glucose or sucrose are present in the medium, without any pigmentation
242 (Figure 1F) (Lozano and Sequeira, 1974b; Maraite and Meyer, 1975; Van Den Mooter *et al.*, 1987).
243 The characteristic white color is due to the absence of xanthomonadin pigment, a rare trait also
244 observed in *Xanthomonas citri* pv. *mangiferaeindicae* and *Xanthomonas campestris* pv. *viticola*
245 (Midha and Patil, 2014). Colonies of the Gram negative rod *Xc* are slimy, convex, circular with entire
246 margins and a deep yellow color (Figure 1G). Optimal temperature for *in-vitro* growth of *Xpm* is
247 around 30°C, while for *Xc* is around 28°C (Maraite and Weyns, 1979). According to Van Den Mooter
248 and coworkers (Van Den Mooter *et al.*, 1987), these two pathogens can be differentiated by colony
249 color, growth on D-saccharic acid (positive for *Xc*, but not for *Xpm*), hydrolysis of Tween 60 and
250 growth on DL-glyceric acid (positive for *Xpm*, but not for *Xc*).

251

252 ***Pathogen classification***

253 First described by Bondar in 1912 the causal agent of CBB was initially named *Bacillus manihotis*
254 (Arthaud-Berthet). After several taxonomical reclassifications (Burkholder, 1942; Dedal *et al.*, 1980;
255 Lozano and Booth, 1974; Maraite and Meyer, 1975; Vauterin *et al.*, 1995), the most recent
256 classification based on a polyphasic taxonomic approach (a seven-gene multilocus sequence

257 analysis, average nucleotide identity and biochemical analyses) assigned this pathogen to the
258 species *X. phaseoli*, with the definite nomenclature *X. phaseoli* pv. *manihotis* (Constantin *et al.*,
259 2016). As to the causative pathogen of CBN, it was first described by Wiehe and Dowson in 1953 in
260 Malawi and named *Xanthomonas cassavae* n. sp. (Wiehe and Dowson, 1953), which was further
261 confirmed upon DNA-DNA hybridization later on (Vauterin *et al.*, 1995). Figure 3 shows the
262 taxonomic position of the two cassava pathogens among most known *Xanthomonas* species.

263

264 **Host range**

265 Alternative hosts can play a key role in pathogen persistence in the environment, acting as reservoirs
266 during long periods. To date, five Euphorbiaceous plants have been reported to be symptomatically
267 affected by *Xpm*, supporting its multiplication and release to the environment: the three cassava
268 wild relative species *Manihot glaziovii*, *Manihot palmata*, and *Manihot aipi* (Lozano and Sequeira,
269 1974b), *Euphorbia pulcherrima* and *Pedilanthus tithymaloides* (Dedal *et al.*, 1980). On the other
270 hand, the limited distribution of *Xc* in Africa and its apparent absence in South America (origin of
271 cassava) suggests that coevolution with cassava is rather short and that this plant may not be the
272 main host of this bacterium (Hayward, 1993). No alternative host has been reported for this
273 pathogen so far.

274

275 **Diversity of the pathogen**

276 *Xpm* diversity has been analyzed by different methods during the last three decades, including RFLP
277 (Berthier *et al.*, 1993; Restrepo *et al.*, 2004; Restrepo & Verdier, 1997; Verdier *et al.*, 1993; Verdier
278 *et al.*, 2001; Verdier *et al.*, 1994), rep-PCR (Chege *et al.*, 2017; Restrepo *et al.*, 2000c), RAPDs
279 (Ogunjobi *et al.*, 2006; Ogunjobi *et al.*, 2007), and AFLPs (Gonzalez *et al.*, 2002; Ogunjobi *et al.*, 2010;
280 Restrepo *et al.*, 2000c; Trujillo *et al.*, 2014b). Recent studies have shown that a slightly better

281 discriminatory power is achieved using MLVA (Rache *et al.*, 2019; Trujillo *et al.*, 2014a). An MLVA
282 scheme with 15 VNTRs amplified by multiplex-PCR was developed to assess the pathogen diversity
283 at a local scale and this scheme was tested in the Northern Coast of Colombia. This method has
284 several advantages : it does not require DNA extraction, it can be easily formatted for high-
285 throughput settings, the amplification of up to four loci is parallelized in a unique PCR reaction, and
286 outputs show high reproducibility and portability (Rache *et al.*, 2019).

287

288 As reviewed previously (López Carrascal and Bernal, 2012), the first studies on African isolates
289 highlighted a rather clonal populations context. Diversity seemed to be increasing from 1980s to
290 2010s, probably as a result of the introduction of new varieties to African countries. More recently,
291 high genetic relatedness of isolates from different cassava-growing regions in Kenya using rep-PCR
292 has been reported, suggesting that despite the high prevalence of the disease, genetic diversity of
293 *Xpm* is still low in some African countries (Chege *et al.*, 2017).

294 Of contrast, *Xpm* diversity is higher in South America, in agreement with the hypothesis of the origin
295 of the pathogen. Studies on *Xpm* populations isolated in several regions in Colombia showed no
296 differentiated geographic structure, traces of pathogen migration, yearly changes of pathogen
297 population structure in the same field and indications of plant host pressure on the pathogen
298 (reviewed by (López Carrascal and Bernal, 2012)). Recent studies highlighted two contrasting
299 scenarios. Most haplotypes of *Xpm* were not structured geographically in northern regions of
300 Colombia when using AFLPs (Trujillo *et al.*, 2014b) or MLVA (Rache *et al.*, 2019), suggesting an
301 important role of pathogen migration through contaminated propagative material. However, a
302 center of origin of several *Xpm* haplotypes was discovered in one of the studied locations, which
303 acts as a potential source of new founder pathogens (Trujillo *et al.*, 2014b). In contrast, AFLP and
304 MLVA-based analysis showed that *Xpm* populations of the eastern plains were structured

305 geographically, but still showing some degree of genetic flow between distant regions. Authors
306 consider that these contrasting scenarios reflect different agricultural practices, crop intensiveness
307 and distribution of lands dedicated to the crop (Trujillo *et al.*, 2014a).

308 At a larger geographical scale, 65 *Xpm* strains, mainly from Africa and South America were
309 sequenced by Illumina (Bart *et al.*, 2012). Phylogenies based on more than 12,000 chromosomal
310 SNPs show that African, Colombian and some Brazilian strains share a common ancestor, in line with
311 the idea of the introduction of the pathogen to Africa as a consequence of the slave trade and or
312 Portuguese missions. Authors highlight that *Xpm* populations are evolving independently, but also
313 show genetic flow between geographically distant places, which should be taken into consideration
314 for surveillance and CBB control purposes.

315

316 ***Diagnostic tools***

317 Early detection and diagnostic tools relied on *in vitro* isolation of the pathogen coupled with phage
318 typing, cassava reinoculation and/or serological assays on cassava extracts. With the popularization
319 of more straightforward molecular methods, standard (Verdier *et al.*, 1998) and nested PCR (Ojeda
320 and Verdier, 2000), dot blots (Verdier & Mosquera, 1999) and an ELISA assay were developed
321 (Verdier *et al.*, 2001). Three PCR-based methods were reported for *Xpm* detection recently, a first
322 one using an optimized version of a previous assay (Verdier *et al.*, 1998) resulting in improved *Xpm*
323 detection (Cerqueira-Melo *et al.*, 2019). A second approach is based on a multiplexed nested PCR
324 including a broadly conserved fragment of *Xpm* TAL effector genes and a semi-specific region of
325 *rpoB*, resulting in a wider detection potential for *Xpm* strains (Bernal-Galeano *et al.*, 2018). A third
326 strategy consisted of a duplex PCR amplifying highly conserved chromosomal regions in *Xpm* and *Xc*
327 to detect and distinguish both cassava pathogens proved to be highly selective and sensitive (Flores
328 *et al.*, 2019).

329

330 **Disease phenotyping tools under controlled conditions**

331 Generally, *Xpm* must first overcome plant defenses in the leaf apoplast, before accessing to the
332 vessels and migrating towards the stem. Defense mechanisms in the mesophyll are different from
333 those exerted by the plant in the leaf and stem vessels, adding barriers of a different nature that are
334 surpassed only by successful pathogens (Kpémoua *et al.*, 1996; Restrepo *et al.*, 2000b; Wydra *et al.*,
335 2007). This multilayer tissue-specific resistance might be the explanation for the quantitative and
336 additive nature of the genetic factors reigning resistance in cassava, which increases the complexity
337 of disease phenotyping. Moreover, the propagation of cassava through vegetative cuttings results
338 in developmentally unsynchronized plants with variations in their physiology that could affect the
339 outcomes of phenotyping methods (Mutka *et al.*, 2016). Hence, disease phenotyping tools in
340 cassava and their relationship to the real-world plant performance are not trivial subjects.

341 Bacterial virulence is usually assessed upon infiltration of leaves of adult 2- to 4-month old plants,
342 through bacterial growth analysis, bacterial movement through the leaf, and symptom development
343 (compilation of protocols in (Cohn *et al.*, 2015)). Bacterial growth and movement within the host
344 quantitatively reflect the ability to grow locally and to migrate through the xylem, while symptom
345 development is a qualitative estimation of *Xpm* capacity to induce water soaking lesion, which might
346 improve bacterial fitness. Symptom development can also be quantified by measuring lesion areas
347 around a perforated hole or upon leaf clipping or leaf spraying (Pacumbaba, 1987; Verdier *et al.*,
348 1994; Zandjanakou-Tachin *et al.*, 2007). More recently, image-based phenotyping of *Xpm* with
349 luminescent reporters efficiently allowed to quantitatively track the disease in time and space
350 (Mutka *et al.*, 2016).

351 Leaf inoculation may be questionable for resistance phenotyping since discrepancies were observed
352 between leaf lesion measurements and scoring upon stem inoculation (Muñoz-Bodnar *et al.*, 2015;

353 Restrepo *et al.*, 2000a). However, others have successfully detected symptom differences when
354 infiltrating leaves of susceptible and resistant cultivars at low densities (i.e., from 10^2 to 10^5 ufc/mL)
355 (Flood *et al.*, 1995; Wydra *et al.*, 2004; Wydra *et al.*, 2007).

356 In the stem, the most discriminant methodology consists of inoculating one internode of the apical
357 region by puncturing bacteria with a sharp tool and evaluating disease development over a 30-day
358 span according to a standardized severity scale. Measurements from several replicates are
359 mathematically integrated to obtain a dimensionless quantity (area under the disease progression
360 curve, AUDPC) that can be comparable even among experimental sets with similar conditions.
361 Resistance and susceptibility is established from scores (in a 0-5 scale, >4 is deemed as susceptible
362 and <3 is deemed resistant), or by setting thresholds when performing AUDPC analyses, this latter
363 being the preferred method to assess resistance (Jorge & Verdier, 2002; Restrepo, Duque, et al.,
364 2000b).

365 To avoid the recurrent problem of space limitation when phenotyping cassava (plants are large),
366 two studies recently reported the use of cassava plants grown *in-vitro*. A comparative resistance
367 screening between plants grown in pots and *in vitro* showed that the latter did not wilt although
368 some symptoms developed, yet indicating that the AUDPC methodology can be applied to assign
369 resistance/susceptibility categories (Mbaringong *et al.*, 2017). In another study Mora and
370 collaborators (2019) developed methods to quantify AUDPC values and perform bacterial growth
371 analysis, highlighting contrasted phenotypes between susceptible and resistant cassava varieties in
372 terms of disease progression and bacterial growth (Mora *et al.*, 2019). These technical advances, in
373 combination with traditional and more recent tools open a new era in cassava phenotyping
374 research.

375

376 **Virulence mechanisms of the pathogen**

377 **Genomics of the pathogen**

378 The advent of next generation sequencing technologies allowed the sequencing of the genome of
379 66 *Xpm* strains originating from Africa and South America mainly (Arrieta-Ortiz *et al.*, 2013; Bart *et*
380 *al.*, 2012). Assemblies resulted into fragmented draft genomes that were used for genomic and
381 phylogenetic analyses, highlighting two main clusters. Grouping of South American and African
382 strains in the same cluster agrees with the hypothesis formulated years ago on the possible
383 introduction of *Xpm* from America to Africa. The authors also identified the invariable occurrence
384 of the well-known Pathogen-Associated Molecular Patterns (PAMPs) *fliC* and *ax21*, the presence of
385 14 to 22 Xanthomonas outer proteins (Xops) per strain, with a core set of nine of them, and reported
386 that all strains harbor at least one TAL effector (Bart *et al.*, 2012). The first manually annotated and
387 high-quality draft genome for *Xpm* was obtained through 454 pyrosequencing technology of the
388 Colombian strain CIO151 (Arrieta-Ortiz *et al.*, 2013). Analysis of the genomic sequence confirmed
389 the presence of fully functional pathogenicity mechanisms such as type II, III, IV and VI secretion
390 systems; an exopolysaccharide production cluster; core and accessory type three effectors, and
391 chemotaxis, type IV pili, flagella, siderophore biosynthesis, and putative polyketide synthesis genes.
392 This study also explored for the first time the possibility of using a MLVA scheme for *Xpm* diversity
393 studies, and report on 16 potential VNTR loci. Third-generation long read sequencing has resulted
394 in several high-quality non-fragmented xanthomonads genomes (Booher *et al.*, 2015; Cox *et al.*,
395 2017; Denancé *et al.*, 2018; Gochez *et al.*, 2018; Peng *et al.*, 2016; Ruh *et al.*, 2017; Showmaker *et*
396 *al.*, 2017; Tran *et al.*, 2018; Wilkins *et al.*, 2015). A first attempt to sequence *Xpm* with Pacific
397 Biosciences technology highlighted a high error rate on reads precluding to obtain a high-quality
398 genome (Bart *et al.*, 2012). Regarding *Xc*, a draft genome of strain CFBP4642 (*a.k.a* NCPPB101,
399 ICMP204, LMG673) isolated in Malawi in 1951, highlighted the presence of a canonical T3SS and

400 some T3Es including TAL effectors (Bolot *et al.*, 2013). Presently our team is producing high-quality
401 reference genomes for both *Xpm* and *Xc* pathogens by using long-read sequencing technologies.

402

403 ***Secretion systems***

404 The type I secretion system (T1SS) allows the active transport of hydrolases (e.g. proteases,
405 phosphatases, glucanases, nucleases, lipases) and toxins from the cytoplasm to the extracellular
406 matrix (reviewed by (Delepelaire, 2004)). The canonical type II secretion system (T2SS) allows the
407 secretion of periplasmic enzymes (e.g. cellulases, pectin methylesterases, cellobiosidases, and
408 polygalacturonases) to the extracellular matrix (reviewed by (Cianciotto and White, 2017)). The
409 presence of canonical versions of these two systems (T1SS and T2SS) in *Xpm* was reported by
410 Arrieta-Ortiz and coworkers, highlighting two slightly different clusters (*xps* and *xcs*) for the T2SS in
411 CIO151 genome (Arrieta-Ortiz *et al.*, 2013). Type IV secretion systems (T4SS) translocate proteins
412 and DNA-protein complexes to surrounding eukaryotic host or other prokaryotic cells, playing an
413 important role for competition against other Gram negative bacteria (reviewed by (Sgro *et al.*,
414 2019)). Likewise, the type VI secretion system (T6SS) is a contractile apparatus that injects toxic
415 effectors to accompanying eukaryotic and prokaryotic cells (reviewed by (Bayer-Santos *et al.*,
416 2019)). *Xpm* carries at least one class of T4SS (Arrieta-Ortiz *et al.*, 2013; Sgro *et al.*, 2019), and one
417 T6SS cluster encoded by fifteen genes in strain CIO151, but their functionality remains to be
418 evidenced (Arrieta-Ortiz *et al.*, 2013).

419 The type III secretion system (T3SS) is a translocation machinery that allows the injection of effector
420 proteins into the host. Delivered effectors play a major role in counteracting the host defenses and
421 hijacking the cellular mechanisms of the invaded cell (reviewed by (Timilsina *et al.*, 2020)). In *Xpm*,
422 the T3SS is encoded by 26 genes of the *hrp2* family cluster (Arrieta-Ortiz *et al.*, 2013), and its role in

423 pathogenesis has been demonstrated through the mutagenesis of *hrpX*, a key regulator of the T3SS
424 (Medina *et al.*, 2018).

425

426 ***Type 3 effectors***

427 Xanthomonas T3Es can be classified in two major groups considering their molecular structure,
428 function, and interactors/targets. The first group forms a heterogeneous group of so-called
429 Xanthomonas outer proteins (Xops) i.e., effectors with a wide range of enzymatic activities, whose
430 targets and associated physiological effects mainly take place in the host cytoplasm and their
431 mechanism of action rely on protein-protein interactions. The second group is only composed by
432 Transcription Activator-Like (TAL) effectors, which are modular proteins that share an unusual
433 architecture combined with eukaryotic motifs that allow them to act as *bona fide* transcription
434 factors inside the host nucleus. Most Xops are involved in disturbing plant defense through
435 alteration of PAMP-DAMP triggered immunity (PTI/DTI) or effector-triggered immunity (ETI)
436 pathways, while TAL effectors act as TFs that re-shape the host cellular metabolism through the
437 activation of susceptibility (*S*) genes to create better niches for bacteria (Timilsina *et al.*, 2020).

438 *Xpm* strains harbor from 13 to 23 Xop effectors, of which 9 are conserved among all the reported
439 *Xpm* genomes and termed as *Xpm* core effectors: HpaA, HrpF, XopE1, XopV, Hpa2, XopAK, XopL,
440 XopN, and XopAE (a.k.a HpaF), the latter five being almost completely monomorphic among the
441 surveyed strains (Bart *et al.*, 2012). *AvrBs2*, *XopAO1*, *XopZ* and *XopX* are important for virulence;
442 strains mutated in these T3Es do not multiply properly *in planta*, and vascular colonization and/or
443 symptom formation is impacted for some of them. *XopK* plays a dual role by increasing the
444 developmental rate of symptoms but limiting the spread of the pathogen, while *XopN* and *XopQ*
445 seem to have a redundant pathogenicity function. Expression in heterologous systems showed that

446 XopR, AvrBs2 and XopAO1 interfere with PTI, while XopAO1 and XopE4 interfere with ETI (Medina
447 *et al.*, 2018; Mutka *et al.*, 2016).

448 The crucial role of TAL effectors in *Xpm* pathogenicity was known for long and finally reported
449 recently (Castiblanco *et al.*, 2013; Cohn *et al.*, 2014; Cohn *et al.*, 2016). Several diagnostics and
450 diversity tools were unintentionally based on TAL effectors detection, the results highlighting their
451 high conservation within *Xpm* (reviewed by (Verdier *et al.*, 2004)). The screening of 65 strains
452 through RFLP showed that all contained between one and five TAL effectors with 12.5, 13.5, 14.5,
453 19.5, 20.5 or 21.5 repeats, and most of them were plasmid-borne (Bart *et al.*, 2012). However, our
454 understanding of *Xpm* TAL effector diversity is rather poor since limited to only seven genes from
455 three strains (Castiblanco *et al.*, 2013; Cohn *et al.*, 2014; Cohn *et al.*, 2016).

456 Castiblanco and coworkers (Castiblanco *et al.*, 2013) described the genomic context, gene structure
457 and function of *TALE1_{xam}* (a.k.a *pthB*), showing that *Xpm* Δ *TALE1_{xam}* is affected in bacterial growth
458 and movement *in planta*. A combination of transcriptomics and *in silico* prediction for TAL effector
459 binding sites led to the identification of a heat shock transcription factor as a potential target of
460 *TALE1_{xam}* in cassava (Muñoz-Bodnar *et al.*, 2014). Cohn and coworkers (Cohn *et al.*, 2014) analyzed
461 the respective role of the five TAL effectors present in the strain Xam668. This study showed that
462 mutation of *TAL14_{xam668}* (which differs by only one RVD from *TALE1_{xam}*) and *TAL20_{xam668}*, both
463 significantly affect pathogen growth *in planta*. Moreover, disruption of *TAL20_{xam668}* leads to the
464 abolishment of water-soaked symptoms. *TAL20_{xam668}* transcriptionally activates *MeSWEET10a*
465 which encodes a clade-III sugar transporter from the *SWEET* family that acts as a susceptibility (*S*)
466 gene, as reported in other pathosystems (reviewed by (Perez-Quintero and Szurek, 2019)). Despite
467 of being a major virulence factor in *Xpm*, the identification and validation of the targeted *S* gene(s)
468 have been hindered because *TALE14* induces transcription of an elevated number of genes (Cohn *et*
469 *al.*, 2016). Pérez-Quintero and coworkers (Pérez-Quintero *et al.*, 2017) developed a web-based

470 platform called daTALbase, where available information on *Xpm* TAL effectors and genomics and
471 transcriptomics resources of the host allows the *in-silico* research of TALE-targeted genes. Figure 4
472 summarizes our current understanding about *Xpm* T3Es roles *in planta*.

473 Isolation and sequencing of TAL effector genes remain challenging due to their repetitive
474 architecture and the existence of multiple copies and/or variants in the genome. Third generation
475 sequencing technologies like Pacific Biosciences Single Molecule Real Time (PacBio SMRT) and
476 Nanopore have been widely used to sequence genomes in other *Xanthomonas* species and
477 eventually recovered high-quality TALome sequences (Booher *et al.*, 2015; Cox *et al.*, 2017; Denancé
478 *et al.*, 2018; Gochez *et al.*, 2018; Peng *et al.*, 2016; Ruh *et al.*, 2017; Showmaker *et al.*, 2017; Tran *et*
479 *al.*, 2018). Targeted cloning techniques have also been optimized to isolate and study TAL effector
480 function, diversity and evolution systematically (Li *et al.*, 2019; Yu *et al.*, 2015). No *Xpm* genomes
481 generated by these techniques has been reported to date, the only two TALomes available so far
482 were isolated from cosmid genomic libraries (Bart *et al.*, 2012). Our team has sequenced several
483 *Xpm* genomes by Pacific Biosciences SMRT technology and cloned more than 50 TAL effectors to
484 assess their diversity and function among *Xpm* strains (Zárate-Chaves *et al.*, 2021). This information
485 will allow TAL effector evolutionary analyses in *Xpm* and searches for other *S* genes in cassava.

486

487 **Toxins**

488 Upon bacterial entry into the leaves, the first symptoms to appear are water-soaked angular spots.
489 Generally, discrete spots begin to coalesce, and eventually, surrounding, and distal areas start to
490 blight. This characteristic blight phenotype may be due to the diffusion into the laminar tissues
491 around the infection foci of a small molecule that acts as a toxin. Perreaux and collaborators isolated
492 from *Xpm* cultures a small organic acid, the 3-methyl thiopropionic acid, which induces the blight
493 symptom when infiltrated alone into cassava leaves (Perreaux *et al.*, 1982). Bacterial metabolism

494 leads to a transamination coupled to a decarboxilation of methionine, resulting in the formation of
495 3-methyl thiopropionic acid (Ewbank and Maraite, 1990). Concentrations of this toxin rise up in
496 leaves along with bacterial multiplication and reach a maximum just before the onset of the blight
497 symptom (Perreaux *et al.*, 1986). The exact mechanism of action of this small acid is not known.
498 However evidence for the role of this compound as a toxin is debated, basically because
499 concentrations of free methionine and the potential toxin in leaves are very low, and *in vitro* assays
500 could have been biased by the acid nature of the compound (Cooper *et al.*, 2001).

501

502 ***Signaling and metabolic routes potentially involved in regulation of pathogenicity***

503 Recently, an RNAseq of a Quorum Sensing-insensitive mutant of *Xpm* vs. wild type grown *in vitro*
504 was reported (Botero *et al.*, 2020). The authors concluded that the QS system controls several
505 subsequent signaling routes including a number of phosphorylation sensor and transduction
506 pathways, some of which share a c-di-GMP phosphodiesterase activity (HD-GYP domain). Metabolic
507 routes that were affected by QS included the already reported xanthan biosynthesis route and the
508 newly reported NAD(P)⁺ balance, and fatty acid elongation, which should be further studied.

509

510 **Host genetics and interaction**

511

512 ***Tools for Xpm-Cassava interaction research***

513 Development of adapted durable resistance must consider the local pathogen population structure
514 and the host response variability. Pathotyping schemes can condense this complex interaction data
515 to cluster pathogens according to virulence (Restrepo, Duque, et al., 2000b; Trujillo, Ochoa, et al.,
516 2014; Verdier et al., 2004) and classify cassava genotypes based on susceptibility (Wydra *et al.*,
517 2004). Inoculation of *in vitro* plants (Mora *et al.*, 2019) and image-based phenotyping methods

518 (Mutka *et al.*, 2016; Veley *et al.*, 2020) are two important advances that will greatly facilitate the
519 study of this pathosystem by leveraging high throughput applications, and evidencing the plant
520 response in a more comprehensive way.

521 Cassava genomic resources include the read collection from 58 Illumina-sequenced cassava or
522 cassava relatives accessions (Bredeson *et al.*, 2016), four high-quality sequenced genomes from
523 cultivars AM560-2 (an MCOL-1505-derived partially inbred cultivar) (Bredeson *et al.*, 2016; Prochnik
524 *et al.*, 2012), KU50 (Wang *et al.*, 2014), TME3, and the highly CBB-susceptible cultivar TMS60444
525 which is amenable to genetic transformation (Kuon *et al.*, 2019). Also it includes at least four high
526 density genetic maps (International Cassava Genetic Map Consortium (ICGMC), 2015; I., Rabbi *et al.*,
527 2014; I., Y., Rabbi *et al.*, 2014; Soto *et al.*, 2015), one of them with an integrated physical map that
528 includes immunity-related genes (Soto *et al.*, 2015); and Bacterial Artificial Chromosome (BAC)
529 libraries from cultivars TMS30001 and MECU72 (Tomkins *et al.*, 2004). Transcriptomic resources for
530 cassava challenged by *Xpm* include expressed sequenced tags (ESTs), simple sequence repeats (SSR)
531 (Lopez *et al.*, 2004; López *et al.*, 2007), a microarray (López Carrascal *et al.*, 2005), transcript-derived
532 fragments (TDFs) (Santaella *et al.*, 2004), and differentially expressed genes (DEGs) catalogs (Cohn
533 *et al.*, 2014; Gómez-Cano *et al.*, 2019; Muñoz-Bodnar *et al.*, 2014). The website
534 www.cassavagenome.org provides tools incorporating the data from some of the above-described
535 resources such as transcriptomic and SSR data displayed on the AM560-2 genome. As mentioned
536 earlier, daTALbase allows the research of potential virulence targets of TAL effectors through
537 prediction of EBE sequences in the host promoterome, and cross-referencing to host
538 transcriptomics and polymorphism data (Pérez-Quintero *et al.*, 2017).

539

540 ***Host susceptibility to Xpm***

541 The susceptibility (*S*) gene *MeSWEET10a* is a Clade III member of the well-known SWEET family of
542 phloem-loading efflux carriers (reviewed by (Chen, 2014)), *MeSWEET10a* is transcriptionally
543 activated by the TAL effector TAL20_{Xam668}, and the resulting protein was found to export glucose and
544 sucrose using *Xenopus* oocytes (Cohn *et al.*, 2014). Increased susceptibility is probably related to
545 increase of glucose accumulation in the apoplast, where it supports bacterial growth (Cohn *et al.*,
546 2014), and/or by increasing osmotic water influx into the intercellular spaces that facilitate bacterial
547 movement (El Kasmi *et al.*, 2018; Schwartz *et al.*, 2017). Cassava has 27 putative *SWEET* genes that
548 are scattered in the genome (except for three clusters with 5, 3 and 2 *SWEET* genes), and account
549 for at least five clade III members that could potentially also act as *S* genes, as shown in rice (Streubel
550 *et al.*, 2013).

551 Mora and coworkers compared the expression of *MeSWEET10a* during a compatible and an
552 incompatible interaction, highlighting its upregulation in the susceptible cultivar TMS60444 but not
553 in the resistant cultivar CM6438-14 fifty hours after inoculation of a strain carrying *TAL20_{Xam668}*.
554 Inspection of the *MeSWEET10a* promoter in both varieties showed conservation of the EBE,
555 suggesting the immune response to stop susceptibility pathways at early stages of the infection
556 (Mora *et al.*, 2019).

557

558 ***Structural and cellular features of cassava resistance against Xpm***

559 A mechanism associated with anti-*Xpm* defense in cassava is the occlusion of adaxial stomata with
560 wax, as a way to reduce pathogen entry into the leaf. Interestingly, the number of occluded stomata
561 seems higher in resistant varieties (Cooper *et al.*, 2001; Zinsou *et al.*, 2006). Once in the substomatal
562 spaces, *Xpm* reaches the mesophyll which colonization is accompanied by the formation of a fibrillar
563 matrix made of bacterial exopolysaccharides and loosening of plant cell walls through enzymatic
564 lysis. These modifications allow the pathogen to become vascular by accessing the xylem vessels

565 and migrating to other parts of the plant (Boher *et al.*, 1995; Boher *et al.*, 1997; Kemp *et al.*, 2004).
566 Little is known about the resistant reactions taking place in the leaf mesophyll, but it has been
567 suggested that these responses may be overcome by *Xpm* (Kpémoua *et al.*, 1996; Verdier *et al.*,
568 1994). Cassava resistance against *Xpm* is mediated by several cellular and molecular mechanisms,
569 and the promptness of this response is the key difference between a susceptible and a resistant
570 variety (Kpémoua *et al.*, 1996). Resistance in the stem is characterized by the initial secretion of
571 bactericidal phenolic compounds into intercellular spaces and vessel lumen, followed by a
572 deposition of lignin, suberin and callose in the paramural spaces. Later, pectic-rich tyloses are
573 formed to occlude vessels and to secrete more phenolic compounds, while bacteria are trapped in
574 lysis pockets through lignification and suberization (Boher *et al.*, 1995). Moreover, resistant
575 genotypes are able to metabolically cope up better with infection, with the maintenance of average
576 stomatal resistance, water potential, and proline levels, factors that are markedly altered in
577 compatible interactions (Restrepo Rubio *et al.*, 2017).

578

579 **Mapping resistance to *Xpm***

580 Due to the quantitative nature of CBB resistance, and the complex interaction between the host
581 genetic background and the environment, the search for resistance sources against *Xpm* in cassava
582 is mainly restricted to the identification of quantitative trait loci (QTLs) (Jorge *et al.*, 2000; Jorge *et*
583 *al.*, 2001; López Carrascal *et al.*, 2007; Soto-Sedano *et al.*, 2017; Tappiban *et al.*, 2018; Wydra *et al.*,
584 2004). Identified QTLs were strain-specific in all cases, explaining up to 61% of the resistance
585 variance (reviewed by (López & Bernal, 2012). Recently, thanks to a high density cassava genetic
586 map (Soto *et al.*, 2015), Soto-Sedano and coworkers associated five novel resistance QTLs,
587 accounting for 16% to 22% of the variance phenotype and which colocalized with 29 genes
588 potentially involved in defense (Soto-Sedano *et al.*, 2017). Furthermore, three out of these QTLs

589 showed to be significantly influenced by environmental conditions, especially humidity. More
590 recently, by using an SSR-based genetic map derived from the F1 offspring of an Asian cassava
591 cultivar cross, 10 QTLs were associated to CBB resistance, explaining up to 26.5 % of the genotypic
592 variance, and five colocalized genes were shown to be differentially upregulated in resistant
593 genotypes (Tappiban *et al.*, 2018). Table S6 summarizes all the cassava QTLs (more than 100)
594 reported to be involved in CBB resistant.

595

596 ***Genetic and molecular aspects of cassava defense against Xpm and Xc***

597 Most of our knowledge about the molecular aspects of the cassava-*Xpm* interaction comes from
598 studies of resistant varieties (see Figure 4). However, challenging susceptible genotypes with *Xpm*
599 also induces defense-related genes, but expression is delayed when compared to resistant varieties
600 (López Carrascal *et al.*, 2005; Santaella *et al.*, 2004). Several studies have identified Pathogenesis-
601 Related (PR) proteins (Li *et al.*, 2017; Román *et al.*, 2014; Yoodée *et al.*, 2018) and various microRNA
602 families potentially involved in plant defense (Pérez-Quintero *et al.*, 2012). The role of Transcription
603 Factor (TF) families in response to *Xpm* has been addressed by coexpression network analyses and
604 functional studies by TF families. Specific cassava Heat Stress TFs (MeHSfs) (Wei *et al.*, 2017; Wei *et*
605 *al.*, 2018b), related to ABI3 AND VP1 (MeRAVs) (Wei *et al.*, 2018a), Nuclear Factor Y (MeNF-Ys) (He
606 *et al.*, 2019), Whirly (MeWHYs) (Liu *et al.*, 2018), NAM/ATAF/CUC2 (NACs) (Gómez-Cano *et al.*,
607 2019), basic leucine zipper (bZIPs) (Gómez-Cano *et al.*, 2019; Li *et al.*, 2017), WRKYs (Gómez-Cano
608 *et al.*, 2019; Liu *et al.*, 2018; Wei *et al.*, 2017; Yan *et al.*, 2017; Yoodée *et al.*, 2018) and TB1/CIN)/PCF
609 (TCPs) TFs (Gómez-Cano *et al.*, 2019) were shown to play important roles in cassava immunity
610 against *Xpm* through activation of defenses. For example, *MeWHY1/2/3* (through interaction with
611 *MeWRKY75*), *MeNF-YA1/3*, *MeNF-YB11/16*, and *MeNF-YC11/12* are crucial for upregulation of
612 defenses against *Xpm* (He *et al.*, 2019; Liu *et al.*, 2018).

613 Salicylic acid (SA), calcium signaling, and melatonin synthesis also play important roles in anti-*Xpm*
614 defense. The upregulation of *MeHsf3* in response to *Xpm* infection triggers the activation of the SA
615 pathway through transcriptional activation of *MeEDS1*, also triggering defense by activating the
616 non-SA-related *MePR4* gene (coding for a defense-related protein containing SCP domain) (Wei *et al.*,
617 2018b). Several Calcineurin B-like proteins (CBLs) and CBL-interacting protein kinases (MeCIPKs)
618 which are respectively transducers of CIPKs signaling and sensors of calcium signaling, have been
619 found upregulated during the interaction. Among them, MeCIPK23 physically interacts with MeCBL1
620 and MeCBL9 to modulate defense against the pathogen, potentially as a response to calcium
621 signaling (Yan *et al.*, 2018). During infection *Xpm* also upregulates the melatonin biosynthesis
622 pathway (MBP), leading to melatonin accumulation, activation of defense- and ROS-related genes,
623 and increased callose deposition (Wei *et al.*, 2016). MeRAV1 and MeRAV2 TFs directly activate three
624 genes of the MBP (*MeTDC2*, *MeT5H*, and *MeASMT1*), while MeWRKY79 and MeHsf20 increase the
625 transcription rate of *MeASMT2* (a second N-acetylserotonin O-methyltransferase gene involved in
626 the MBP), all resulting in increased defense responses against *Xpm* (Wei *et al.*, 2017; Wei *et al.*,
627 2018a).

628 Autophagy regulation is altered during *Xpm* infection, which significantly impacts immunity. Four
629 members of the cassava Autophagy-Related Protein 8 (*MeATG8*) family, including *MeATG8f* which
630 colocalizes with a resistance QTL (Tappiban *et al.*, 2018) and which product was reported to interact
631 with MeWRKY20 to induce defense against *Xpm* (Yan *et al.*, 2017). Furthermore, *MeATG8b* and
632 *Me8ATG8e* have antagonistic roles against the two cassava Glyceraldehyde-3-phosphate
633 dehydrogenases MeGAPC4 and MeGAPC6, whose downregulation results in higher autophagic
634 activity, higher H₂O₂ levels, and increased callose deposition (Zeng *et al.*, 2018). Little attention has
635 been paid to the cassava-*Xc* interaction, but analysis of an incompatible interaction showed that,
636 unlike for *Xpm*, Phenylalanine Ammonia Lyase (PAL) transcription and activity, and cell wall-bound

637 peroxidase activity are markedly increased in response to the pathogen (Pereira *et al.*, 1999; Pereira
638 *et al.*, 2000).

639 Pathogen perception relies on Receptor-Like Kinases (RLKs), Receptor-Like Proteins (RLPs), and
640 Nucleotide Binding Site-Leucine-Rich Repeat (NBS-LRRs) proteins. Cassava contains at least 253 RLK-
641 (Soto *et al.*, 2015) and 327 NBS-LRR-encoding genes. Most *NBS-LRRs* are grouped in 39 clusters, with
642 two superclusters with 43 and 19 of the genes located on chromosomes 16 and 17, respectively (
643 Lozano *et al.*, 2015). Mapping for resistance QTLs allowed the identification of the two resistance
644 gene candidates *RXam1* and *RXam2*, respectively coding for a RLK and an NBS-LRR (López *et al.*,
645 2003). *RXam1* shows similarity to the anti-*Xanthomonas oryzae* pv. *oryzae* rice resistance gene
646 *Xa21*, and was recently shown to be involved in defense against *Xpm* strain CIO136 (Díaz Tatis *et al.*,
647 2018). *RXam2* is a typical non-TIR NB-LRR located in a QTL explaining 61% of the resistance variance
648 to *Xpm* strain CIO151. Recently, it was demonstrated that this gene confers partial and broad-
649 spectrum resistance to different *Xpm* strains (Díaz-Tatis *et al.*, *In prep*).

650

651 ***Resistance engineering in cassava***

652 Stable transformation of cassava is mainly achieved through bacterial-mediated transformation of
653 friable embryogenic calli and many transgenic cassava lines are reported in the literature (reviewed
654 by (Chavarriaga Aguirre *et al.*, 2016). However, the incipient efforts to engineer CBB resistance
655 reflect the limited information on genetic determinants for resistance to *Xpm*. Stable transformants
656 of the cassava cultivar 60444 overexpressing and silencing *RXam1* were generated to study the
657 function of this resistance gene candidate (Chavarriaga Aguirre *et al.*, 2016; Díaz Tatis *et al.*, 2018),
658 as well as overexpressing the pepper *R* gene *Bs2*, which did not allow resistance against *Xpm* (Díaz-
659 Tatis *et al.*, 2019).

660 Since TAL effectors recognize *S* genes through interaction with DNA, genome editing is a nice tool
661 to disturb or favor the TAL-DNA interaction and induce resistance or prevent susceptibility
662 (reviewed by (Schornack *et al.*, 2013). This approach was applied recently to help rice against
663 *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) by editing five EBEs in the promoter of three *SWEET* genes
664 targeted by *Xoo* in elite varieties, resulting in a robust and broad-spectrum resistance in the field
665 (Oliva *et al.*, 2019). This strategy could also be applied in cassava to engineer resistance against *Xpm*,
666 given the requirement of *MeSWEET10a* for successful disease development.

667

668 **Conclusion and future directions**

669 Whilst the application of measures based on the ecology of the pathogen have helped mitigating
670 the impact of CBB, *Xpm* remains a devastating pathogen causing important yield losses. Yet, genetic
671 resistance remains the best option to control the disease. To be effective and durable breeding
672 efforts must consider among others the local diversity of the pathogen populations, the impact of
673 other biotic and abiotic stresses, the farming practices, and growers' preferences. Despite cassava-
674 *Xpm* was considered an orphan pathosystem, research efforts lead by several institutions and
675 researchers have resulted in robust knowledge and applied tools either for diagnosis or pathogen
676 diversity assessment, that constitute important contributions to better understand some
677 epidemiological aspects of the disease. Moreover, new image-based techniques coupled to
678 fluorescent or luminescent markers, the use of plants grown *in vitro* allowing for high-throughput
679 phenotyping, and the third-generation sequencing data will certainly help to progress on genome-
680 wide association studies. Quantitative resistance studies reveal that several genetic factors involved
681 in resistance are still awaiting to be more deeply characterized. On the other side, functional analysis
682 of *Xpm* TAL effectors has allowed to unmask a *SWEET* sugar transporter as a major susceptibility
683 determinant. This opens a new perspective for CBB control by loss of susceptibility, as was

684 successfully implemented in rice against the bacterial leaf blight, resulting in plants with reduced
685 susceptibility to *Xoo*. More research is however needed to pinpoint the function of other major
686 virulence TAL effectors and their targets. By exploiting this marked dependence, genome editing is
687 a powerful tool to decrease susceptibility and improve resistance against this pathogen, but parallel
688 efforts need to be done to understand the durability and farmer's acceptability of these strategies,
689 considering the high diversity of the pathogen and its genomic repertoires. Further research should
690 be deployed to better understand *Xc* lifestyle and its phylogenetic relationship to *Xpm*. It could be
691 an excellent opportunity and contribution to unveil the mechanisms that differentiate vascular from
692 nonvascular diseases in plants.

693

694 **Acknowledgements**

695 We thank the Research Program on Roots, Tubers and Bananas, and the Agropolis Foundation
696 (project #1403-073) for funding support, and the Institut de Recherche pour le Développement for
697 the doctoral fellowship awarded to C.A.Z.C. We acknowledge the IRD i-Trop HPC (South Green
698 Platform) for providing HPC resources that have contributed to the results reported within this
699 paper. URL: <https://bioinfo.ird.fr/>- <http://www.southgreen.fr>. We want to thank Ndomassi Tando
700 and Julie Orjuela for their support in bioinformatics. Likewise, we are thankful to C. Bragard and H.
701 Maraite (UCL, Belgium) for providing pictures of CBN.

702

703 **References**

- 704 **Allem, A.C.** (2002) The origins and taxonomy of cassava. In *Cassava: Biology, Production and*
705 *Utilization.* (Hillocks, R.J., Thresh, J.M., and Bellotti, A., eds), pp. 1–16. Wallingford, UK: New
706 York: CABI Publishing. hlm.
- 707 **Ambe, J.T.** (1993) The effect of planting dates on three cassava diseases in Cameroon. *Int. J. Pest*
708 *Manag.* **39**, 309–311.
- 709 **Arrieta-Ortiz, M.L., Rodríguez-R, L.M., Pérez-Quintero, Á.L., et al.** (2013) Genomic survey of
710 pathogenicity determinants and VNTR markers in the cassava bacterial pathogen
711 *Xanthomonas axonopodis* pv. *manihotis* strain CIO151. *PLoS One* **8**, e79704.
- 712 **Bart, R., Cohn, M., Kassen, A., et al.** (2012) High-throughput genomic sequencing of cassava
713 bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc.*
714 *Natl. Acad. Sci. U. S. A.* **109**, E1972-9.
- 715 **Bayer-Santos, E., Ceseti, L. de M., Farah, C.S. and Alvarez-Martinez, C.E.** (2019) Distribution,
716 Function and Regulation of Type 6 Secretion Systems of Xanthomonadales . *Front.*
717 *Microbiol.* **10**, 1635.
- 718 **Bernal-Galeano, V., Ochoa, J.C., Trujillo, C., Rache, L., Bernal, A. and López Carrascal, C.E.** (2018)
719 Development of a multiplex nested PCR method for detection of *Xanthomonas axonopodis*
720 pv. *manihotis* in Cassava. *Trop. Plant Pathol.* **43**, 341–350.
- 721 **Berthier, Y., Verdier, V., Guesdon, J.L., Chevrier, D., Denis, J.B., Decoux, G. and Lemattre, M.**
722 (1993) Characterization of *Xanthomonas campestris* Pathovars by rRNA Gene Restriction
723 Patterns. *Appl. Environ. Microbiol.* **59**, 851–859.
- 724 **Boher, B., Kpemoua, K., Nicole, M., Luisetti, J. and Geiger, J.P.** (1995) Ultrastructure of
725 interactions between cassava and *Xanthomonas campestris* pv. *manihotis*: cytochemistry of
726 cellulose and pectin degradation in a susceptible cultivar. *Phytopathology* **85**.

727 **Boher, B., Nicole, M., Potin, M. and Geiger, J.P.** (1997) Extracellular Polysaccharides from
728 *Xanthomonas axonopodis* pv. *manihotis* Interact with Cassava Cell Walls During
729 Pathogenesis. *Mol. Plant-Microbe Interact.* **10**, 803–811.

730 **Boher, B. and Verdier, V.** (1994) Cassava bacterial blight in Africa: The state of knowledge and
731 implications for designing control strategies. *African Crop Sci. J.* **2**, 505–509.

732 **Bolot, S., Munoz Bodnar, A., Cunnac, S., et al.** (2013) Draft Genome Sequence of the
733 *Xanthomonas cassavae* Type Strain CFBP 4642. *Genome Announc.* **1**, e00679-13.

734 **Boher, N.J., Carpenter, S.C.D., Sebra, R.P., Wang, L., Salzberg, S.L., Leach, J.E. and Bogdanove,**
735 **A.J.** (2015) Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a
736 dynamic structure and complex TAL (transcription activator-like) effector gene relationships.
737 *Microb. genomics* **1**, e000032.

738 **Botero, D., Monk, J., Rodríguez Cubillos, M.J., et al.** (2020) Genome-Scale Metabolic Model of
739 *Xanthomonas phaseoli* pv. *manihotis*: An Approach to Elucidate Pathogenicity at the
740 Metabolic Level. *Front. Genet.* **11**, 837.

741 **Bredeson, J. V., Lyons, J.B., Prochnik, S.E., et al.** (2016) Sequencing wild and cultivated cassava and
742 related species reveals extensive interspecific hybridization and genetic diversity. *Nat.*
743 *Biotechnol.* **34**, 562–570.

744 **Burkholder, W.H.** (1942) Three bacterial plant pathogens: *Phytomonas earyophylli* sp.n.,
745 *Phytomonas alliicola* sp.n., and *Phytomonas manihotis* (Arthaud-Berthet et Sondar) Viégas.
746 *Phytopathology* **32**, 141–149.

747 **CABI** (2020a) *Xanthomonas axonopodis* pv. *manihotis* (cassava bacterial blight). *Invasive Species*
748 *Compend.*

749 **CABI** (2020b) *Xanthomonas cassavae* (cassava leaf spot). *Invasive Species Compend.*

750 **Castiblanco, L.F., Gil, J., Rojas, A., et al.** (2013) TALE1 from *Xanthomonas axonopodis* pv.

751 *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity
752 in cassava plants. *Mol. Plant Pathol.* **14**, 84–95.

753 **Cerqueira-Melo, R.C., Dórea-Bragança, C.A., Nogueira-Pestana, K., Alves da Silva, H.S., Fortes-**
754 **Ferreira, C. and Oliveira, S.A..** (2019) Improvement of the specific detection of *Xanthomonas*
755 *phaseoli* pv. *manihotis* based on the pthB gene. *Acta Sci. Agron.* **41**.

756 **Chavarriga Aguirre, P., Brand, A., Medina, A., et al.** (2016) The potential of using biotechnology
757 to improve cassava: a review. *Vitr. Cell. Dev. Biol. - Plant* **52**, 461–478.

758 **Chege, M., Wamunyokoli, F., Kamau, J. and Nyaboga, E.** (2017) Phenotypic and genotypic
759 diversity of *Xanthomonas axonopodis* pv. *manihotis* causing bacterial blight disease of
760 cassava in Kenya. *J. Appl. Biol. Biotechnol.* **5**, 38–44.

761 **Chen, L.-Q.** (2014) SWEET sugar transporters for phloem transport and pathogen nutrition. *New*
762 *Phytol.* **201**, 1150–1155.

763 **Cianciotto, N.P. and White, R.C.** (2017) Expanding Role of Type II Secretion in Bacterial
764 Pathogenesis and Beyond. *Infect. Immun.* **85**, e00014-17.

765 **Cohn, M., Bart, R.S., Shybut, M., et al.** (2014) *Xanthomonas axonopodis* Virulence Is Promoted by
766 a Transcription Activator-Like Effector–Mediated Induction of a SWEET Sugar Transporter in
767 Cassava. *Mol. Plant-Microbe Interact.* **27**, 1186–1198.

768 **Cohn, M., Morbitzer, R., Lahaye, T. and Staskawicz, B.J.** (2016) Comparison of gene activation by
769 two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host
770 susceptibility genes in cassava. *Mol. Plant Pathol.*, 875–889.

771 **Cohn, M., Shybut, M., Dahlbeck, D. and Staskawicz, B.** (2015) Assays to Assess Virulence of
772 *Xanthomonas axonopodis* pv. *manihotis* on Cassava. *Bio-protocol* **5**, e1522.

773 **Constantin, E.C., Cleenwerck, I., Maes, M., Baeyen, S., Malderghem, C. Van, Vos, P. De and**
774 **Cottyn, B.** (2016) Genetic characterization of strains named as *Xanthomonas axonopodis* pv.

775 *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant*
776 *Pathol.* **65**, 792–806.

777 **Cooper, R.M., Kemp, B., Day, R., Gomez-Vasquez, R. and Beeching, J.R.** (2001) Pathogenicity and
778 Resistance in *Xanthomonas* Blight of Cassava. In *Plant Pathogenic Bacteria*. (Boer, S.H. De,
779 ed), pp. 319–323. Dordrecht: Springer.

780 **Cox, K.L., Meng, F., Wilkins, K.E., et al.** (2017) TAL effector driven induction of a *SWEET* gene
781 confers susceptibility to bacterial blight of cotton. *Nat. Commun.* **8**, 15588.

782 **Daniel, J.F. and Boher, B.** (1985a) Epiphytic phase of *Xanthomonas campestris* pathovar *manihotis*
783 on aerial parts of cassava. *Agronomie* **5**, 111–116.

784 **Daniel, J.F. and Boher, B.** (1985b) Etude des modes de survie de l'agent causal de la bacteriose
785 vasculaire du manioc, *Xanthomonas campestris* pathovar *manihotis*. *Agron. Sci. des Prod.*
786 *Veg. l'environnement.* **5**, 339–346.

787 **Dedal, O.I., Palomar, M.K. and Napiere, C.M.** (1980) Host Range of *Xanthomonas manihotis* Starr.
788 *Ann. Trop. Res.* **2**, 149–155.

789 **Delepelaire, P.** (2004) Type I secretion in gram-negative bacteria. *Biochim. Biophys. Acta - Mol.*
790 *Cell Res.* **1694**, 149–161.

791 **Denancé, N., Szurek, B., Doyle, E.L., et al.** (2018) Two ancestral genes shaped the *Xanthomonas*
792 *campestris* TAL effector gene repertoire. *New Phytol.* **219**, 391–407.

793 **Díaz-Tatis, P.A., Ochoa, J.C., García, L., Chavarriaga, P., Bernal, A.J. and López Carrascal, C.E.**
794 (2019) Interfamily transfer of Bs2 from pepper to cassava (*Manihot esculenta* Crantz). *Trop.*
795 *Plant Pathol.* **44**, 225–237.

796 **Díaz Tatis, P.A., Herrera Corzo, M., Ochoa Cabezas, J.C., Medina Cipagauta, A., Prías, M.A.,**
797 **Verdier, V., Chavarriaga Aguirre, P. and López Carrascal, C.E.** (2018) The overexpression of
798 *RXam1*, a cassava gene coding for an RLK, confers disease resistance to *Xanthomonas*

799 *axonopodis* pv. *manihotis*. *Planta* **247**, 1031–1042.

800 **Dixon, A.G.O., Ngeve, J.M. and Nukenine, E.N.** (2002) Genotype × environment Effects on Severity
801 of Cassava Bacterial Blight Disease caused by *Xanthomonas axonopodis* pv. *manihotis*. *Eur. J.*
802 *Plant Pathol.* **108**, 763–770.

803 **Elango, F. and Lozano, J.C.** (1981) Epiphytic survival of *Xanthomonas manihotis* on common weeds
804 in Colombia. In 5th Int. Conf. Plant Path. Bact. Cali, Colombia., pp. 203–209.

805 **Elango, F. and Lozano, J.C.** (1980) Transmission of *Xanthomonas manihotis* in Seed of Cassava
806 (*Manihot esculenta*). *Plant Dis.* **64**, 784–786.

807 **Ewbank, E. and Maraite, H.** (1990) Conversion of methionine to phytotoxic 3-methylthiopropionic
808 acid by *Xanthomonas campestris* pv. *manihotis*. *Microbiology* **136**, 1185–1189.

809 **Fanou, A., Zinsou, V. and Wydra, K.** (2017) Survival of *Xanthomonas axonopodis* pv. *manihotis* in
810 weed species and in cassava debris: implication in the epidemiology of cassava bacterial
811 blight. *Int. J. Adv. Res.* **5**, 2098–2112.

812 **FAO** (2020) FAOSTAT Crops.

813 **Fermont, A.M., Asten, P.J.A. van, Tiftonell, P., Wijk, M.T. van and Giller, K.E.** (2009) Closing the
814 cassava yield gap: An analysis from smallholder farms in East Africa. *F. Crop. Res.* **112**, 24–36.

815 **Flood, J., Cooper, R.M. and Deshappriya, N.** (1995) Resistance of cassava (*Manihot esculenta*) to
816 *Xanthomonas* blight in vitro and in planta UN, F.A.O. of the, ed. , 277–284.

817 **Flores, C., Zarate, C., Triplett, L., et al.** (2019) Development of a duplex-PCR for differential
818 diagnosis of *Xanthomonas phaseoli* pv. *manihotis* and *Xanthomonas cassavae* in cassava
819 (*Manihot esculenta*). *Physiol. Mol. Plant Pathol.* **105**, 34–46.

820 **Fokunang, C.N., Akem, C.N., Dixon, A.G.O. and Ikotun, T.** (2000) Evaluation of a cassava
821 germplasm collection for reaction to three major diseases and the effect on yield. *Genet.*
822 *Resour. Crop Evol.* **47**, 63–71.

823 **Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W., Tohme, J. and**
824 **Bonierbale, M.** (1997) A molecular genetic map of cassava (*Manihot esculenta* Crantz).
825 *Theor. Appl. Genet.* **95**, 431–441.

826 **Frison, E. and Feliu, E.** (1991) FAO/IBPGR Technical Guidelines for the Safe Movement of Cassava
827 Germplasm,.

828 **Gochez, A.M., Huguet-Tapia, J.C., Minsavage, G. V, et al.** (2018) Pacbio sequencing of copper-
829 tolerant *Xanthomonas citri* reveals presence of a chimeric plasmid structure and provides
830 insights into reassortment and shuffling of transcription activator-like effectors among *X. citri*
831 strains. *BMC Genomics* **19**, 16.

832 **Gómez-Cano, F., Soto, J., Restrepo, S., Bernal, A., López-Kleine, L. and López Carrascal, C.E.**
833 (2019) Gene co-expression network for *Xanthomonas*-challenged cassava reveals key
834 regulatory elements of immunity processes. *Eur. J. Plant Pathol.* **153**, 1083–1104.

835 **Gonzalez, C., Restrepo, S., Tohme, J. and Verdier, V.** (2002) Characterization of pathogenic and
836 nonpathogenic strains of *Xanthomonas axonopodis* pv. *manihotis* by PCR-based DNA
837 fingerprinting techniques. *Fems Microbiol. Lett.* **215**, 23–31.

838 **Harris, K.P., Martin, A., Novak, S., Kim, S.-H., Reynolds, T. and Leigh, A.** (2015) Cassava Bacterial
839 Blight and Postharvest Physiological Deterioration, Production Losses and Control Strategies.
840 , 1–34.

841 **Hayward, A.C.** (1993) The hosts of *Xanthomonas*. In *Xanthomonas*. (Swings, J.G. and Civerolo, E.L.,
842 eds), pp. 1–119. Dordrecht: Springer.

843 **He, X., Liu, G., Li, B., Xie, Y., Wei, Y., Shang, S., Tian, L. and Shi, H.** (2019) Functional analysis of the
844 heterotrimeric NF-Y transcription factor complex in cassava disease resistance. *Ann. Bot.* **124**,
845 1185–1197.

846 **Hillocks, R.J. and Wydra, K.** (2002) Bacterial, fungal and nematode diseases. In *Cassava: biology,*

847 production and utilization., pp. 261–280. Wallingford, UK ; New York: CABI Publishing.

848 **International Cassava Genetic Map Consortium (ICGMC)** (2015) High-Resolution Linkage Map and
849 Chromosome-Scale Genome Assembly for Cassava (*Manihot esculenta* Crantz) from 10
850 Populations. *G3 Genes/Genomes/Genetics* **5**, 133 LP – 144.

851 **Jones, W.** (1959) Manioc in Africa, Stanford, CA: Stanford University Press.

852 **Jorge, V., Fregene, M., Duque, M.C., Bonierbale, M.W., Tohme, J. and Verdier, V.** (2000) Genetic
853 mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz).
854 *Theor. Appl. Genet.* **101**, 865–872.

855 **Jorge, V., Fregene, M., Vélez, C.M., Duque, M.C., Tohme, J. and Verdier, V.** (2001) QTL analysis of
856 field resistance to *Xanthomonas axonopodis* pv. *manihotis* in cassava. *Theor. Appl. Genet.*
857 **102**, 564–571.

858 **Jorge, V. and Verdier, V.** (2002) Qualitative and quantitative evaluation of cassava bacterial blight
859 resistance in F1 progeny of a cross between elite cassava clones. *Euphytica* **123**, 41–48.

860 **Kasmi, F. El, Horvath, D. and Lahaye, T.** (2018) Microbial effectors and the role of water and sugar
861 in the infection battle ground. *Curr. Opin. Plant Biol.* **44**, 98–107.

862 **Kemp, B.P., Horne, J., Bryant, A. and Cooper, R.M.** (2004) *Xanthomonas axonopodis* pv. *manihotis*
863 *gumD* gene is essential for EPS production and pathogenicity and enhances epiphytic survival
864 on cassava (*Manihot esculenta*). *Physiol. Mol. Plant Pathol.* **64**, 209–218.

865 **Kpémoua, K., Boher, B., Nicole, M., Calatayud, P. and Geiger, J.P.** (1996) Cytochemistry of
866 defense responses in cassava infected by *Xanthomonas campestris* pv. *manihotis*. *Can. J.*
867 *Microbiol.* **42**, 1131–1143.

868 **Kuon, J.-E., Qi, W., Schläpfer, P., et al.** (2019) Haplotype-resolved genomes of geminivirus-
869 resistant and geminivirus-susceptible African cassava cultivars. *BMC Biol.* **17**, 75.

870 **Lamphey, J.N.L., Okoli, O.O. and Frimpong-Manso, P.P.** (1998) Incidence and severity of African

871 cassava mosaic disease (ACMD) and cassava bacterial blight (CBB) on some local and exotic
872 cassava varieties in different ecological zones of Ghana. *Ghana J. Agric. Sci.* **31**, 35–43.

873 **Li, C., Ji, C., Huguet-Tapia, J.C., White, F.F., Dong, H. and Yang, B.** (2019) An efficient method to
874 clone TAL effector genes from *Xanthomonas oryzae* using Gibson assembly. *Mol. Plant*
875 *Pathol.* **20**, 1453–1462.

876 **Li, X., Fan, S., Hu, W., Liu, G., Wei, Y., He, C. and Shi, H.** (2017) Two Cassava Basic Leucine Zipper
877 (bZIP) Transcription Factors (MebZIP3 and MebZIP5) Confer Disease Resistance against
878 Cassava Bacterial Blight. *Front. Plant Sci.* **8**, 2110.

879 **Liu, W., Yan, Y., Zeng, H., Li, X., Wei, Y., Liu, G., He, C. and Shi, H.** (2018) Functional
880 characterization of WHY–WRKY75 transcriptional module in plant response to cassava
881 bacterial blight. *Tree Physiol.* **38**, 1502–1512.

882 **López Carrascal, C.E. and Bernal, A.J.** (2012) Cassava Bacterial Blight: Using Genomics for the
883 Elucidation and Management of an Old Problem. *Trop. Plant Biol.* **5**, 117–126.

884 **López Carrascal, C.E., Jorge, V., Piégu, B., et al.** (2004) A unigene catalogue of 5700 expressed
885 genes in cassava. *Plant Mol. Biol.* **56**, 541–554.

886 **López Carrascal, C.E., Quesada-Ocampo, L.M., Bohórquez, A., Duque, M.C., Vargas, J., Tohme, J.**
887 **and Verdier, V.** (2007) Mapping EST-derived SSRs and ESTs involved in resistance to bacterial
888 blight in *Manihot esculenta*. *Genome* **50**, 1078–1088.

889 **López Carrascal, C.E., Soto, M., Restrepo, S., Piégu, B., Cooke, R., Delseny, M., Tohme, J. and**
890 **Verdier, V.** (2005) Gene expression profile in response to *Xanthomonas axonopodis* pv.
891 *manihotis* infection in cassava using a cDNA microarray. *Plant Mol. Biol.* **57**, 393–410.

892 **López Carrascal, C.E., Zuluaga, A.P., Cooke, R., Delseny, M., Tohme, J. and Verdier, V.** (2003)
893 Isolation of Resistance Gene Candidates (RGCs) and characterization of an RGC cluster in
894 cassava. *Mol. Genet. Genomics* **269**, 658–671.

895 **Lozano, J.C.** (1986) Cassava Bacterial Blight: A Manageable Disease. *Plant Dis.* **70**, 1089.

896 **Lozano, J.C. and Booth, R.H.** (1974) Diseases of Cassava (*Manihot esculenta* Crantz). *PANS Pest*
897 *Artic. News Summ.* **20**, 30–54.

898 **Lozano, J.C. and Laberry, R.** (1982) Screening for Resistance to Cassava Bacterial Blight. *Plant Dis.*
899 **66**, 316–318.

900 **Lozano, J.C. and Sequeira, L.** (1974a) Bacterial Blight of Cassava in Colombia: Epidemiology and
901 Control. *Phytopathology* **64**, 83–88.

902 **Lozano, J.C. and Sequeira, L.** (1974b) Bacterial Blight of Cassava in Colombia: Etiology.
903 *Phytopathology* **64**, 74–82.

904 **Lozano, R., Hamblin, M.T., Prochnik, S. and Jannink, J.-L.** (2015) Identification and distribution of
905 the *NBS-LRR* gene family in the Cassava genome. *BMC Genomics* **16**, 360.

906 **Mansfield, J., Genin, S., Magori, S., et al.** (2012) Top 10 plant pathogenic bacteria in molecular
907 plant pathology. *Mol. Plant Pathol.* **13**, 614–629.

908 **Maraite, H.** (1993) *Xanthomonas campestris* pathovars on cassava: cause of bacterial blight and
909 bacterial necrosis. In *Xanthomonas*. (Swings, J.G. and Civerolo, E.L., eds), pp. 18–25. London,
910 UK: Springer Science + Business Media Dordrecht.

911 **Maraite, H. and Meyer, J.** (1975) *Xanthomonas manihotis* (Arthaud-Berthet) Starr, Causal Agent of
912 Bacterial Wilt, Blight and Leaf Spots of Cassava in Zaire. *Int. J. Pest Manag. - INT J PEST*
913 *Manag.* **21**, 27–37.

914 **Maraite, H. and Perreaux, D.** (1979) Comparative symptom development in cassava after infection
915 by *Xanthomonas manihotis* or *X. cassavae* under controlled conditions. In *Cassava Bacterial*
916 *Blight in Africa; Past, Present and Future. Report of an interdisciplinary workshop, IITA,*
917 *Ibadan, Nigeria, 1978.* (Terry, E.R., Persley, G.J., and Cook, S.C.A., eds), pp. 17–24. London,
918 UK: Centre for Overseas Pest Research.

919 **Maraite, H. and Weyns, J.** (1979) Distinctive physiological, biochemical and pathogenic
920 characteristics of *Xanthomonas manihotis* and *X. cassavae*. In Diseases of
921 tropical food crops. Proceedings of an International Symposium, U.C.L., 1978. Louvain-la-
922 Neuve, Belgium: Université Catholique de Louvain. (Maraite, H. and Meyer, J., eds), pp. 103–
923 117.

924 **Marcano, M. and Trujillo, G.** (1984) Papel de las malezas en relacion a la perpetuacion del anublo
925 bacteria no de la yuca. *Rev. la Fac. Agron.* **13**, 167–181.

926 **Mbaringong, G., Nyaboga, E., Ondu, V. and Kanduma, E.** (2017) Evaluation of Selected Cassava
927 (*Manihot esculenta* Crantz) Cultivars Grown in Kenya for Resistance to Bacterial Blight
928 Disease. *World J. Agric. Res.* **5**, 94–101.

929 **Medina, C.A., Reyes, P.A., Trujillo, C.A., et al.** (2018) The role of type III effectors from
930 *Xanthomonas axonopodis* pv. *manihotis* in virulence and suppression of plant immunity. *Mol.*
931 *Plant Pathol.* **19**, 593–606.

932 **Midha, S. and Patil, P.B.** (2014) Genomic insights into the evolutionary origin of *Xanthomonas*
933 *axonopodis* pv. *citri* and its ecological relatives. *Appl. Environ. Microbiol.* **80**, 6266–6279.

934 **Mooter, M. Van Den, Maraite, H., Meiresonne, L., Swings, J., Gillis, M., Kersters, K., Ley, J. De**
935 **and Perreaux, D.** (1987) Comparison Between *Xanthomonas campestris* pv. *manihotis* (ISPP
936 List 1980) and *X. campestris* pv. *cassavae* (ISPP List 1980) by Means of Phenotypic, Protein
937 Electrophoretic, DNA Hybridization and Phytopathological Techniques. *Microbiology* **133**, 57–
938 71.

939 **Mora, R., Rodriguez, M., Gayosso, L. and López Carrascal, C.E.** (2019) Using in vitro plants to study
940 the cassava response to *Xanthomonas phaseoli* pv. *manihotis* infection. *Trop. Plant Pathol.*

941 **Mostade, J.M. and Butare, I.** (1978) Symptomatology et epidemiologie de la necrose bacterienne
942 du manioc causee par *Xanthomonas cassavae* au Rwanda. In Disease of Tropical Food Crops.

943 Proceedings of an International Symposium. (Maraite, H. and Meyer, J.A., eds), pp. 95–117.
944 U.c.L., Louvain-la-Neuve, Belgium: Universite Catholique de Louvain, Louvain-la-Neuve,
945 Belgium.

946 **Muñoz-Bodnar, A., Cruz-Gómez, L.M., Bernal, A., Szurek, B. and López Carrascal, C.E.** (2015)
947 Comparing Inoculation Methods to Evaluate the Growth of *Xanthomonas axonopodis* pv.
948 *manihotis* on Cassava Plants. *Acta Biológica Colomb.* **20**, 47–55.

949 **Muñoz-Bodnar, A., Perez-Quintero, A.L., Gomez-Cano, F., Gil, J., Michelmore, R., Bernal, A.,**
950 **Szurek, B. and López Carrascal, C.E.** (2014) RNAseq analysis of cassava reveals similar plant
951 responses upon infection with pathogenic and non-pathogenic strains of *Xanthomonas*
952 *axonopodis* pv. *manihotis*. *Plant Cell Rep.* **33**, 1901–1912.

953 **Mutka, A.M., Fentress, S.J., Sher, J.W., Berry, J.C., Pretz, C., Nusinow, D.A. and Bart, R.** (2016)
954 Quantitative, Image-Based Phenotyping Methods Provide Insight into Spatial and Temporal
955 Dimensions of Plant Disease. *Plant Physiol.* **172**, 650 LP – 660.

956 **Nassar, N.** (2007) Cassava improvement: Challenges and impacts. *J. Agric. Sci. - J AGR SCI* **145**.

957 **Ogunjobi, A., Dixon, A. and Fagade, O.** (2007) Molecular genetic study of cassava bacterial blight
958 casual agent in Nigeria using random amplified polymorphic DNA. *Electron. J. Environ. Agric.*
959 *Food Chem.* **6**.

960 **Ogunjobi, A., Fagade, O. and Dixon, A.** (2006) Molecular variation in population structure of
961 *Xanthomonas axonopodis* pv *manihotis* in the south eastern Nigeria. *African J. Biotechnol.* **5**,
962 1868–1872.

963 **Ogunjobi, A.A., Fagade, O.E. and Dixon, A.G.O.** (2010) Comparative analysis of genetic variation
964 among *Xanthomonas axonopodis* pv *manihotis* isolated from the western states of Nigeria
965 using RAPD and AFLP. *Indian J. Microbiol.* **50**, 132–138.

966 **Ojeda, S. and Verdier, V.** (2000) Detecting *Xanthomonas axonopodis* pv. *manihotis* in cassava true

967 seeds by nested polymerase chain reaction assay. *Can. J. Plant Pathol.* **22**, 241–247.

968 **Oliva, R., Ji, C., Atienza-Grande, G., et al.** (2019) Broad-spectrum resistance to bacterial blight in
969 rice using genome editing. *Nat. Biotechnol.* **37**, 1344–1350.

970 **Olsen, K.M. and Schaal, B.A.** (1999) Evidence on the origin of cassava: Phylogeography of *Manihot*
971 *esculenta*. *Proc. Natl. Acad. Sci.* **96**, 5586 LP – 5591.

972 **Pacumbaba, R.P.** (1987) A Screening Method for Detecting Resistance Against Cassava Bacterial
973 Blight Disease. *J. Phytopathol.* **119**, 1–6.

974 **Peng, Z., Hu, Y., Xie, J., Potnis, N., Akhunova, A., Jones, J., Liu, Z., White, F.F. and Liu, S.** (2016)
975 Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector
976 gene analysis of *Xanthomonas translucens*. *BMC Genomics* **17**, 21.

977 **Pereira, L.F., Goodwin, P.H. and Erickson, L.** (2000) Peroxidase Activity During Susceptible and
978 Resistant Interactions Between Cassava (*Manihot esculenta*) and *Xanthomonas axonopodis*
979 *pv. manihotis* and *Xanthomonas cassavae*. *J. Phytopathol.* **148**, 575–578.

980 **Pereira, L.F., Goodwin, P.H. and Erickson, L.** (1999) The Role of a Phenylalanine Ammonia Lyase
981 Gene during Cassava Bacterial Blight and Cassava Bacterial Necrosis. *J. Plant Res.* **112**, 51–60.

982 **Pérez-Quintero, A.L., Lamy, L., Zarate, C.A., Cunnac, S., Doyle, E., Bogdanove, A., Szurek, B. and**
983 **Dereeper, A.** (2017) daTALbase: A Database for Genomic and Transcriptomic Data Related to
984 TAL Effectors. *Mol. Plant-Microbe Interact.* **31**, 471–480.

985 **Pérez-Quintero, Á.L., Quintero, A., Urrego, O., Vanegas, P. and López Carrascal, C.E.** (2012)
986 Bioinformatic identification of cassava miRNAs differentially expressed in response to
987 infection by *Xanthomonas axonopodis* *pv. manihotis*. *BMC Plant Biol.* **12**, 29.

988 **Perez-Quintero, A.L. and Szurek, B.** (2019) A Decade Decoded: Spies and Hackers in the History of
989 TAL Effectors Research. *Annu. Rev. Phytopathol.* **57**, 459–481.

990 **Perreaux, D., Maraite, H. and Meyer, J.A.** (1986) Detection of 3(methylthio) propionic acid in

991 cassava leaves infected by *Xanthomonas campestris* pv. *manihotis*. *Physiol. Mol. Plant Pathol.*
992 **28**, 323–328.

993 **Perreaux, D., Maraite, H. and Meyer, J.A.** (1982) Identification of 3(methylthio) propionic acid as a
994 blight-inducing toxin produced by *Xanthomonas campestris* pv. *manihotis* in vitro. *Physiol.*
995 *Plant Pathol.* **20**, 313–319.

996 **Persley, G.J.** (1976) Distribution and importance of cassava bacterial blight in Africa. In Cassava
997 bacterial blight, report on an interdisciplinary workshop. (Persley, J.G., Terry, E.R., and
998 MacIntyre, R., eds), pp. 9–14.

999 **Persley, G.J.** (1978) Studies on the epidemiology and ecology of cassava bacterial blight. In
1000 Workshop on Cassava Bacterial Blight in Africa, Past, Present and Future. (Terry, E.R., Persley,
1001 G.J., and Cook, S.C.A., eds), pp. 5–7. Ibadan, Nigeria: Centre for Overseas Pest Research.

1002 **Prochnik, S., Marri, P.R., Desany, B., et al.** (2012) The Cassava Genome: Current Progress, Future
1003 Directions. *Trop. Plant Biol.* **5**, 88–94.

1004 **Rabbi, I., Hamblin, M., Gedil, M., Kulakow, P., Ferguson, M., Ikpan, A.S., Ly, D. and Jannink, J.-L.**
1005 (2014) Genetic Mapping Using Genotyping-by-Sequencing in the Clonally Propagated
1006 Cassava. *Crop Sci.* **54**, 1384–1396.

1007 **Rabbi, I.Y., Hamblin, M.T., Kumar, P.L., Gedil, M.A., Ikpan, A.S., Jannink, J.-L. and Kulakow, P.A.**
1008 (2014) High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava
1009 using genotyping-by-sequencing and its implications for breeding. *Virus Res.* **186**, 87–96.

1010 **Rache, L., Blondin, L., Flores, C., Trujillo, C., Szurek, B., Restrepo, S., Koebnik, R., Bernal, A. and**
1011 **Vernière, C.** (2019) An Optimized Microsatellite Scheme for Assessing Populations of
1012 *Xanthomonas phaseoli* pv. *manihotis*. *Phytopathology*TM **109**, 859–869.

1013 **Restrepo Rubio, J.S., López Carrascal, C.E. and Melgarejo, L.M.** (2017) Physiological behavior of
1014 cassava plants (*Manihot esculenta* Crantz) in response to infection by *Xanthomonas*

1015 *axonopodis* pv. *manihotis* under greenhouse conditions. *Physiol. Mol. Plant Pathol.* **100**, 136–
1016 141.

1017 **Restrepo, S., Duque, M.C. and Verdier, V.** (2000a) Characterization of pathotypes among isolates
1018 of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Plant Pathol.* **49**, 680–687.

1019 **Restrepo, S., Duque, M.C. and Verdier, V.** (2000b) Resistance spectrum of selected *Manihot*
1020 *esculenta* genotypes under field conditions. *F. Crop. Res.* **65**, 69–77.

1021 **Restrepo, S., Velez, C.M., Duque, M.C. and Verdier, V.** (2004) Genetic Structure and Population
1022 Dynamics of *Xanthomonas axonopodis* pv. *manihotis* in Colombia from 1995 to 1999. *Appl.*
1023 *Environ. Microbiol.* **70**, 255 LP – 261.

1024 **Restrepo, S., Vélez, C.M. and Verdier, V.** (2000c) Measuring the Genetic Diversity of *Xanthomonas*
1025 *axonopodis* pv. *manihotis* Within Different Fields in Colombia. *Phytopathology* **90**, 683–690.

1026 **Restrepo, S. and Verdier, V.** (1997) Geographical Differentiation of the Population of
1027 *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Appl. Environ. Microbiol.* **63**, 4427–
1028 4434.

1029 **Román, V., Bossa-Castro, A.M., Vásquez, A.X., Bernal-Galeano, V., Schuster, M., Bernal, N. and**
1030 **López Carrascal, C.E.** (2014) Construction of a cassava PR protein-interacting network during
1031 *Xanthomonas axonopodis* pv. *manihotis* infection. *Plant Pathol.* **63**, 792–802.

1032 **Ruh, M., Briand, M., Bonneau, S., Jacques, M.-A. and Chen, N.W.G.** (2017) *Xanthomonas*
1033 adaptation to common bean is associated with horizontal transfers of genes encoding TAL
1034 effectors. *BMC Genomics* **18**, 670.

1035 **Santaella, M., Suárez, E., López Carrascal, C.E., González, C., Mosquera, G., Restrepo, S., Tohme,**
1036 **J., Badillo, A. and Verdier, V.** (2004) Identification of genes in cassava that are differentially
1037 expressed during infection with *Xanthomonas axonopodis* pv. *manihotis*. *Mol. Plant Pathol.*
1038 **5**, 549–558.

1039 **Schornack, S., Moscou, M.J., Ward, E.R. and Horvath, D.M.** (2013) Engineering Plant Disease
1040 Resistance Based on TAL Effectors. *Annu. Rev. Phytopathol.* **51**, 383–406.

1041 **Schwartz, A.R., Morbitzer, R., Lahaye, T. and Staskawicz, B.J.** (2017) TALE-induced bHLH
1042 transcription factors that activate a pectate lyase contribute to water soaking in bacterial
1043 spot of tomato. *Proc. Natl. Acad. Sci.* **114**, E897 LP-E903.

1044 **Sgro, G.G., Oka, G.U., Souza, D.P., et al.** (2019) Bacteria-Killing Type IV Secretion Systems. *Front.*
1045 *Microbiol.* **10**, 1078.

1046 **Shakya, M., Ahmed, S.A., Davenport, K.W., Flynn, M.C., Lo, C.-C. and Chain, P.S.G.** (2020)
1047 Standardized phylogenetic and molecular evolutionary analysis applied to species across the
1048 microbial tree of life. *Sci. Rep.* **10**, 1723.

1049 **Showmaker, K.C., Arick 2nd, M.A., Hsu, C.-Y., et al.** (2017) The genome of the cotton bacterial
1050 blight pathogen *Xanthomonas citri* pv. *malvacearum* strain MSCT1. *Stand. Genomic Sci.* **12**,
1051 42.

1052 **Soto-Sedano, J.C., Mora-Moreno, R.E., Mathew, B., León, J., Gómez-Cano, F.A., Ballvora, A. and**
1053 **López Carrascal, C.E.** (2017) Major Novel QTL for Resistance to Cassava Bacterial Blight
1054 Identified through a Multi-Environmental Analysis. *Front. Plant Sci.* **8**, 1169.

1055 **Soto, J.C., Ortiz, J.F., Perlaza-Jiménez, L., et al.** (2015) A genetic map of cassava (*Manihot*
1056 *esculenta* Crantz) with integrated physical mapping of immunity-related genes. *BMC*
1057 *Genomics* **16**, 190.

1058 **Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J. and Szurek, B.** (2013) Five phylogenetically
1059 close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae*
1060 pv. *oryzae*. *New Phytol.* **200**, 808–819.

1061 **Tappiban, P., Sraphet, S., Srisawad, N., Smith, D.R. and Triwitayakorn, K.** (2018) Identification
1062 and expression of genes in response to cassava bacterial blight infection. *J. Appl. Genet.* **59**,

1063 391–403.

1064 **Taylor, R.K., Griffin, R.L., Jones, L.M., Pease, B., Tsatsia, F., Fanai, C., Macfarlane, B., Dale, C.J.**
1065 **and Davis, R.I.** (2017) First record of *Xanthomonas axonopodis* pv. *manihotis* in Solomon
1066 Islands. *Australas. Plant Dis. Notes* **12**, 49.

1067 **Timilsina, S., Potnis, N., Newberry, E.A., Liyanapathirana, P., Iruegas-Bocardo, F., White, F.F.,**
1068 **Goss, E.M. and Jones, J.B.** (2020) *Xanthomonas* diversity, virulence and plant–pathogen
1069 interactions. *Nat. Rev. Microbiol.*

1070 **Tomkins, J., Fregene, M., Main, D., Kim, H., Wing, R. and Tohme, J.** (2004) Bacterial artificial
1071 chromosome (BAC) library resource for positional cloning of pest and disease resistance
1072 genes in cassava (*Manihot esculenta* Crantz). *Plant Mol. Biol.* **56**, 555–561.

1073 **Tran, T.T., Pérez-Quintero, A.L., Wonni, I., et al.** (2018) Functional analysis of African
1074 *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf
1075 blight of rice. *PLOS Pathog.* **14**, e1007092.

1076 **Trujillo, C., Arias-Rojas, N., Poulin, L., Medina, C., Tapiero, A., Restrepo, S., Koebnik, R. and**
1077 **Bernal, A.** (2014a) Population typing of the causal agent of cassava bacterial blight in the
1078 Eastern Plains of Colombia using two types of molecular markers. *Bmc Microbiol.* **14**, 161.

1079 **Trujillo, C., Ochoa, J.C., Mideros, M.F., Restrepo, S., López Carrascal, C.E. and Bernal, A.** (2014b) A
1080 complex population structure of the cassava pathogen *Xanthomonas axonopodis* pv.
1081 *manihotis* in recent years in the Caribbean Region of Colombia. *Microb. Ecol.* **68**, 155–167.

1082 **Umanah, E.E.** (1977) Cassava Research in Nigeria Before 1972. In Fourth Tropical Root Crops
1083 Symposium. (Cock, J., MacIntyre, R., and Graham, M., eds), pp. 137–141.

1084 **Umemura, Y. and Kawano, K.** (1983) Field Assessment and Inheritance of Resistance to Cassava
1085 Bacterial Blight1. *Crop Sci.* **23**.

1086 **Vauterin, G., Hoste, B., Kersters, K. and Swings, J.** (1995) Reclassification of *Xanthomonas* UN,

1087 F.A.O. of the, ed. **45**, 472–489.

1088 **Veley, K.M., Okwuonu, I., Jensen, G., Yoder, M., Taylor, N.J., Meyers, B.C. and Bart, R.S.** (2020)

1089 Visualizing cassava bacterial blight at the molecular level using CRISPR-mediated homology-

1090 directed repair. *bioRxiv*, 2020.05.14.090928.

1091 **Verdier, V., Boher, B., Maraite, H. and Geiger, J.-P.** (1994) Pathological and Molecular

1092 Characterization of *Xanthomonas campestris* Strains Causing Diseases of Cassava (*Manihot*

1093 *esculenta*). *Appl. Environ. Microbiol.* **60**, 4478 LP – 4486.

1094 **Verdier, V., Dongo, P. and Boher, B.** (1993) Assessment of genetic diversity among strains of

1095 *Xanthomonas campestris* pv. *manihotis*. *J. Gen. Microbiol.* **139**, 2591–2601.

1096 **Verdier, V. and Mosquera, G.** (1999) Specific Detection of *Xanthomonas axonopodis* pv. *manihotis*

1097 with a DNA Hybridization Probe. *J. Phytopathol.* **147**, 417–423.

1098 **Verdier, V., Mosquera, G. and Assigbétsé, K.** (1998) Detection of the Cassava Bacterial Blight

1099 Pathogen, *Xanthomonas axonopodis* pv. *manihotis*, by Polymerase Chain Reaction. *Plant Dis.*

1100 **82**, 79–83.

1101 **Verdier, V., Ojeda, S. and Mosquera, G.** (2001a) Methods for detecting the cassava bacterial blight

1102 pathogen: A practical approach for managing the disease. *Euphytica* **120**, 103–107.

1103 **Verdier, V., Restrepo, S., Mosquera, G., Duque, M.C., Gerstl, A. and Laberry, R.** (2001b) Genetic

1104 and pathogenic variation of *Xanthomonas axonopodis* pv. *manihotis* in Venezuela. *Plant*

1105 *Pathol.* **47**, 601–608.

1106 **Verdier, V., Restrepo, S., Mosquera, G., Jorge, V. and López Carrascal, C.E.** (2004) Recent progress

1107 in the characterization of molecular determinants in the *Xanthomonas axonopodis* pv.

1108 *manihotis*–cassava interaction. *Plant Mol. Biol.* **56**, 573–584.

1109 **Verdier, V., Schmit, J. and Lemattre, M.** (1990) Étude en microscopie électronique à balayage de

1110 l’installation de deux souches de *Xanthomonas campestris* pv *manihotis* sur feuilles de

1111 vitroplants de manioc. *Agronomie* **10**, 93–102.

1112 **Wang, W., Feng, B., Xiao, J., et al.** (2014) Cassava genome from a wild ancestor to cultivated
1113 varieties. *Nat. Commun.* **5**, 5110.

1114 **Wei, Y., Chang, Y., Zeng, H., Liu, G., He, C. and Shi, H.** (2018a) RAV transcription factors are
1115 essential for disease resistance against cassava bacterial blight via activation of melatonin
1116 biosynthesis genes. *J. Pineal Res.* **64**, e12454.

1117 **Wei, Y., Hu, W., Wang, Q., et al.** (2016) Comprehensive transcriptional and functional analyses of
1118 melatonin synthesis genes in cassava reveal their novel role in hypersensitive-like cell death.
1119 *Sci. Rep.* **6**, 35029.

1120 **Wei, Y., Liu, G., Bai, Y., Xia, F., He, C. and Shi, H.** (2017) Two transcriptional activators of N-
1121 acetylserotonin O-methyltransferase 2 and melatonin biosynthesis in cassava. *J. Exp. Bot.* **68**.

1122 **Wei, Y., Liu, G., Chang, Y., He, C. and Shi, H.** (2018b) Heat shock transcription factor 3 regulates
1123 plant immune response through modulation of salicylic acid accumulation and signaling in
1124 cassava. *Mol. Plant Pathol.* **19**.

1125 **Wiehe, P.O. and Dowson, W.J.** (1953) A bacterial disease of Cassava (*Manihot utilissima*) in
1126 Nyasaland. *Emp. J. Exp. Agric.* **21**, 141–143.

1127 **Wilkins, K.E., Booher, N.J., Wang, L. and Bogdanove, A.J.** (2015) TAL effectors and activation of
1128 predicted host targets distinguish Asian from African strains of the rice pathogen
1129 *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of
1130 five TAL effectors. *Front. Plant Sci.* **6**, 536.

1131 **Wydra, K., Banito, A. and Kpémoua, K.E.** (2007) Characterization of resistance of cassava
1132 genotypes to bacterial blight by evaluation of leaf and systemic symptoms in relation to yield
1133 in different ecozones. *Euphytica* **155**, 337–348.

1134 **Wydra, K. and Fanou, A.** (2015) Removal of symptomatic cassava leaves as cultural practice to

1135 control cassava bacterial blight. *Int. J. Plant Pathol.* **3**, 117–124.

1136 **Wydra, K., Zinsou, V., Jorge, V. and Verdier, V.** (2004) Identification of Pathotypes of
1137 *Xanthomonas axonopodis* pv. *manihotis* in Africa and Detection of Quantitative Trait Loci and
1138 Markers for Resistance to Bacterial Blight of Cassava. *Phytopathology*TM **94**, 1084–1093.

1139 **Yan, Y., He, X., Hu, W., Liu, G., Wang, P., He, C. and Shi, H.** (2018) Functional analysis of MeCIPK23
1140 and MeCBL1/9 in cassava defense response against *Xanthomonas axonopodis* pv. *manihotis*.
1141 *Plant Cell Rep.* **37**, 887–900.

1142 **Yan, Y., Wang, P., He, C. and Shi, H.** (2017) MeWRKY20 and its interacting and activating
1143 autophagy-related protein 8 (MeATG8) regulate plant disease resistance in cassava. *Biochem.*
1144 *Biophys. Res. Commun.* **494**, 20–26.

1145 **Yoodee, S., Kobayashi, Y., Songnuan, W., Boonchird, C., Thitamadee, S., Kobayashi, I. and**
1146 **Narangajavana, J.** (2018) Phytohormone priming elevates the accumulation of defense-
1147 related gene transcripts and enhances bacterial blight disease resistance in cassava. *Plant*
1148 *Physiol. Biochem.* **122**, 65–77.

1149 **Yu, Y.-H., Lu, Y., He, Y.-Q., Huang, S. and Tang, J.-L.** (2015) Rapid and efficient genome-wide
1150 characterization of *Xanthomonas* TAL effector genes. *Sci. Rep.* **5**, 13162.

1151 **Zandjanakou-Tachin, M., Fanou, A., Gall, P. Le and Wydra, K.** (2007) Detection, Survival and
1152 Transmission of *Xanthomonas axonopodis* pv. *manihotis* and *X. axonopodis* pv. *vignicola*,
1153 Causal Agents of Cassava and Cowpea Bacterial Blight, respectively, in/by Insect Vectors. *J.*
1154 *Phytopathol.* **155**, 159–169.

1155 **Zárate-Chaves, C.A., Osorio-Rodríguez, D., Mora, R.E., Pérez-Quintero, Á.L., Dereeper, A.,**
1156 **Restrepo, S., López, C.E., Szurek, B. and Bernal, A.** (2021) TAL Effector Repertoires of Strains
1157 of *Xanthomonas phaseoli* pv. *manihotis* in Commercial Cassava Crops Reveal High Diversity at
1158 the Country Scale. *Microorganisms* **9**.

1159 **Zeng, H., Xie, Y., Liu, G., Lin, D., He, C. and Shi, H.** (2018) Molecular identification of GAPDHs in
1160 cassava highlights the antagonism of MeGAPCs and MeATG8s in plant disease resistance
1161 against cassava bacterial blight. *Plant Mol. Biol.* **97**, 201–214.

1162 **Zinsou, V., Wydra, K., Ahohuendo, B. and Hau, B.** (2004) Effect of soil amendments, intercropping
1163 and planting time in combination on the severity of cassava bacterial blight and yield in two
1164 ecozones of West Africa. *Plant Pathol.* **53**, 585–595.

1165 **Zinsou, V., Wydra, K., Ahohuendo, B. and Hau, B.** (2005) Genotype × environment interactions in
1166 symptom development and yield of cassava genotypes with artificial and natural cassava
1167 bacterial blight infections. *Eur. J. Plant Pathol.* **111**, 217–233.

1168 **Zinsou, V., Wydra, K., Ahohuendo, B. and Schreiber, L.** (2006) Leaf Waxes of Cassava (*Manihot*
1169 *esculenta* Crantz) in Relation to Ecozone and Resistance to *Xanthomonas* Blight. *Euphytica*
1170 **149**, 189–198.

1171

1172

1173 **Supporting Information legends**

1174

1175 **Supplementary Table S1.** Incidence data extracted from several studies on CBB. Incidence was
1176 defined in these studies as the percentage of symptomatic plants (No. of symptomatic
1177 plants*100/No. of evaluated plants) recorded in cassava fields that were naturally infected by *Xpm*
1178 (no artificial inoculations). Incidence ranges (minimum and maximum), timepoint of the incidence
1179 evaluation, and country where the study was performed were extracted from each report. Min. =
1180 minimum, Max. = maximum, MAP = months after planting, nd = missing data.

1181

1182 **Supplementary Table S2.** Severity data extracted from several studies on CBB. Severity was
1183 measured in these studies using similar scales based on symptom recording in cassava fields that
1184 were naturally infected by *Xpm* (no artificial inoculations). Severity ranges (minimum and
1185 maximum), timepoint of the severity evaluation, country where the study was performed, and
1186 grading scales were extracted from each report. Transformation column indicates if the values
1187 presented as minimum and maximum were mathematically transformed (see text after table) to
1188 obtain values of the same type and scale. Min. = minimum, Max. = maximum, MAP = months after
1189 planting, nd = missing data. Grading scales showed slight modifications, but all of them reflected the
1190 same damage extension with the corresponding grading numbers. Most of the scales were grouped
1191 within the Scale A: 1, no disease symptoms; 2, only angular leaf spots on leaves; 3, leaf blight, leaf
1192 wilt, defoliation, gum exudation in petioles and stems; 4, extensive leaf blight, wilt, defoliation, and
1193 partial stem die-back; 5, complete defoliation, stem die-back and stunting in most of the plants.
1194 Only one study used a similar grading system with different scale numbers, Scale B: 0 = no
1195 symptoms; 1 = symptoms on leaves only - blight; 2 = presence of necrotic lesions on the stem or
1196 petiole; 3 = most severe symptoms on leaves and/or the presence of necrotic lesions with gum
1197 exudation; 4 = complete loss of leaves with apical death or death of the plant.

1198 **Supplementary Table S3.** Losses data extracted from several studies on CBB. Only losses reported
1199 as the percentage of fresh root weight lost were considered. All the studies were carried out in
1200 cassava fields naturally or artificially infected by *Xpm*. Loss ranges (minimum and maximum),
1201 resistance status of the tested varieties, and country or region where the study was performed were
1202 extracted from each report. Min. = minimum, Max. = maximum.

1203

1204 **Supplementary Table S4.** Reference genomes used for phylogeny reconstruction. Phylogeny was
1205 constructed using the nucleotide concatenated sequence of five housekeeping genes (*atpD*, *dnaK*,
1206 *efp*, *glnA*, *gyrB*, and *rpoD*).

1207

1208 **Supplementary Table S5.** Result compilation of cassava resistance assessments reported in
1209 literature.

1210

1211 **Supplementary Table S6.** Quantitative Trait Loci (QTLs) associated to CBB resistance. QTL name:
1212 name of the QTL associated to CBB resistance. ExpVar: percentage of phenotypic variance explained
1213 by a given QTL. Linkage group: linkage group where a given QTL can be found. Linkage groups
1214 depend on the map used for associations. Strain: strain used to inoculate the assessed plants (in
1215 most of the cases, QTLs are strain-specific). Natural inoculum: the study was performed in the field
1216 and the quantified disease resulted from infections with naturally occurring strains. Undefined
1217 strain: the study states that only one strain was used to inoculate the assessed plants, but authors
1218 do not provide the strain identifier. Technique: disease quantitation technique used to associate
1219 phenotypes to markers. AUDPC: Area Under the Disease Progression Curve, 100-DI: percentage of
1220 diseased plants (100-disease incidence), MDR: Mean Disease Rating. Marker: peak marker
1221 associated to a given QTL. Map reference: cassava map used to associate QTLs. The maps developed

1222 by Fregene and coworkers (Fregene *et al.*, 1997) and Soto and coworkers (Soto *et al.*, 2015) derived
1223 from the female parental TMS 30572 and the male parental CM 2177–2; the map developed by
1224 Wydra and coworkers (Wydra *et al.*, 2004) derived from five F1 male individuals (CM7857-4,
1225 CM7857-10, CM7857-51, CM7857-77, and CM7857-115) and the female recurrent parent
1226 TMS30572; the map developed by Tappiban and coworkers (Tappiban *et al.*, 2018) derived from the
1227 parentals Huay Bong 60 (HB60) and Hanatee (HN). Study reference: the reference to the study
1228 where QTLs were found.
1229

1230 **Figure legends**

1231

1232 **Figure 1.** Etiology, ecology, and distribution of *Xpm* and *Xc*. (a) cassava bacterial blight disease cycle.
1233 Solid arrows indicate processes that are relevant for cassava cropping systems. Dashed arrows
1234 indicate processes that have been detected and are scientifically interesting but are not relevant for
1235 cassava cropping systems. The inset indicates environmental and ecological factors that affect the
1236 spread or development of disease. (b) angular leaf spots caused by *Xpm*. (c) blight (solid white
1237 arrows) and leaflet curling (dashed white arrows) caused by *Xpm*. (d) collapsed petioles from wilted
1238 leaves (solid black arrows) and gum exudation (dashed black arrows) from stems caused by *Xpm*
1239 infection. (e) shoot apex wilting and dieback caused by *Xpm*. (f) typical colonies of *Xpm* on LPGA
1240 medium. (g) typical colonies of *Xc* on LPGA medium. (h) leaf spots caused by *Xc*. (i) world-wide
1241 cassava production (FAO, 2020) by country (missing data for South Africa, Guam, and Palau) for 2018
1242 and distribution of *Xpm* and *Xc* (CABI, 2020a).

1243

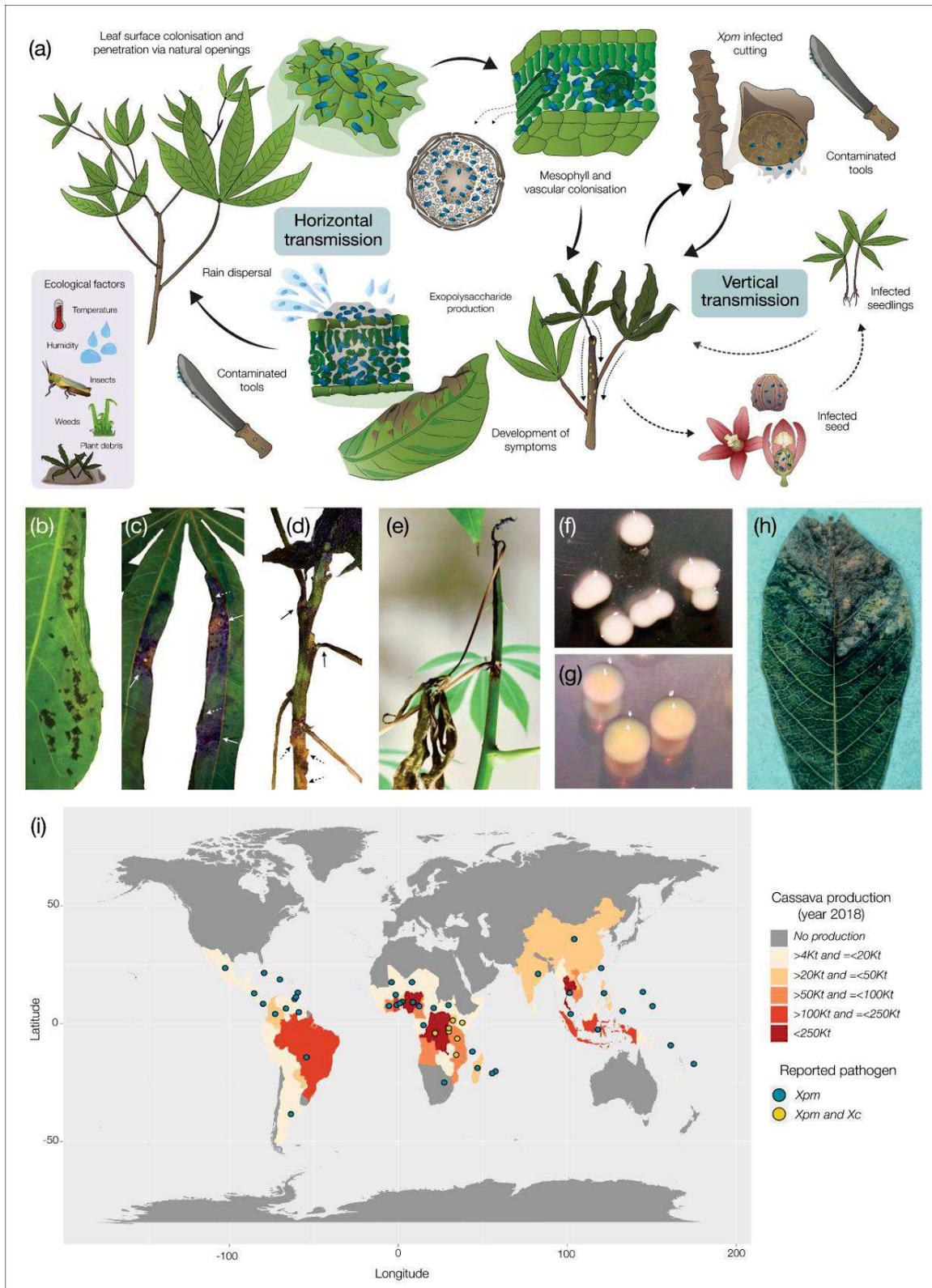
1244 **Figure 2.** Graphical meta-analysis of incidence, severity and losses reported for CBB. (a) Histogram
1245 of the incidence ranges reported by 11 studies (Table S1). The color scale correlates with the
1246 frequency of the reported range. (b) Violin plots showing severity ranges recorded by 15 studies
1247 (Table S2) according to measurement timepoints after planting. The confounded variable groups
1248 data from reports that did not include timepoint information. A harmonized severity grading scale
1249 is presented at the right side of the plot. The dashed line indicates that values above 2 reflect
1250 systemic disease. (c) Boxplots showing fresh root yield losses ranges described by 7 reports (some
1251 studies reported several ranges - Table S3) according to the designated resistance status of the
1252 plants. The confounded variable groups data from reports that did not include the designated
1253 resistance status of the evaluated plants.

1254

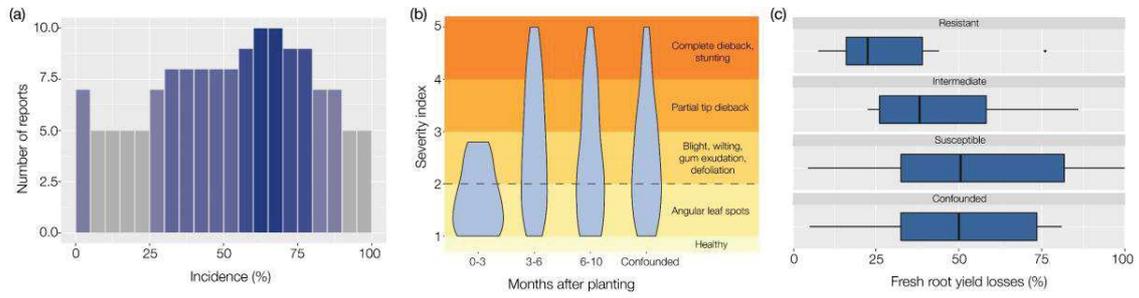
1255 **Figure 3.** Taxonomic position of *Xpm* and *Xc*. Phylogeny of 30 representative *Xanthomonas*,
1256 including 27 species, and 4 pathovars of *X. phaseoli* (Table S4). Strain code is indicated in
1257 parentheses. *Xpm* and *Xc* are highlighted in purple and blue, respectively. Phylogeny was
1258 constructed from RefSeq complete genomes using the bioinformatic workflow PhaME (Shakya *et*
1259 *al.*, 2020). Core genome alignments resulted in 155,507 aligned SNP positions that covered coding
1260 and non-coding regions. Trees were reconstructed with the GTRGAMMAI model of RAxML and
1261 consensus tree was calculated from 100 bootstraps; results higher than 80 are shown above
1262 branches.

1263

1264 **Figure 4.** Roles of Xops, host determinants of susceptibility and defense responses in the cassava-
1265 *Xpm* interaction. Xop effectors (colored circles) are injected by *Xpm* into the plant cytoplasm, where
1266 they interfere with host cellular processes to repress. T3Es are grouped (shading) according to their
1267 predicted involvement in PAMP-triggered immunity (PTI), effector-triggered immunity (ETI), and/or
1268 other virulence-related phenotypes (see Type 3 effectors section). The upper section of the depicted
1269 nucleus outlines the contribution of TAL effectors to cassava susceptibility, in which transcriptional
1270 activation of the *S* gene *MeSWEET10a* leads to an over-accumulation of sugar transporters (purple,
1271 anchored to the membrane) and sugar export (see Type 3 effectors section). The RXam1 RLK (lilac,
1272 in the membrane) and the RXam2 NBS-LRR (blue, in the cytoplasm) are associated to defense
1273 response triggered against *Xpm*, but their matching elicitors are unknown. Calcium signaling and
1274 autophagy elements are represented in the cytoplasm as part of the anti-*Xpm* defense mechanisms.
1275 The roles of miRNA and transcription factors are schematized in the nucleus (see Genetic and
1276 molecular aspects of cassava defense against *Xpm* and *Xc* section). H₂O₂ and O₂⁻ represent reactive
1277 oxygen species (ROS) accumulation. SA: salicylic acid.

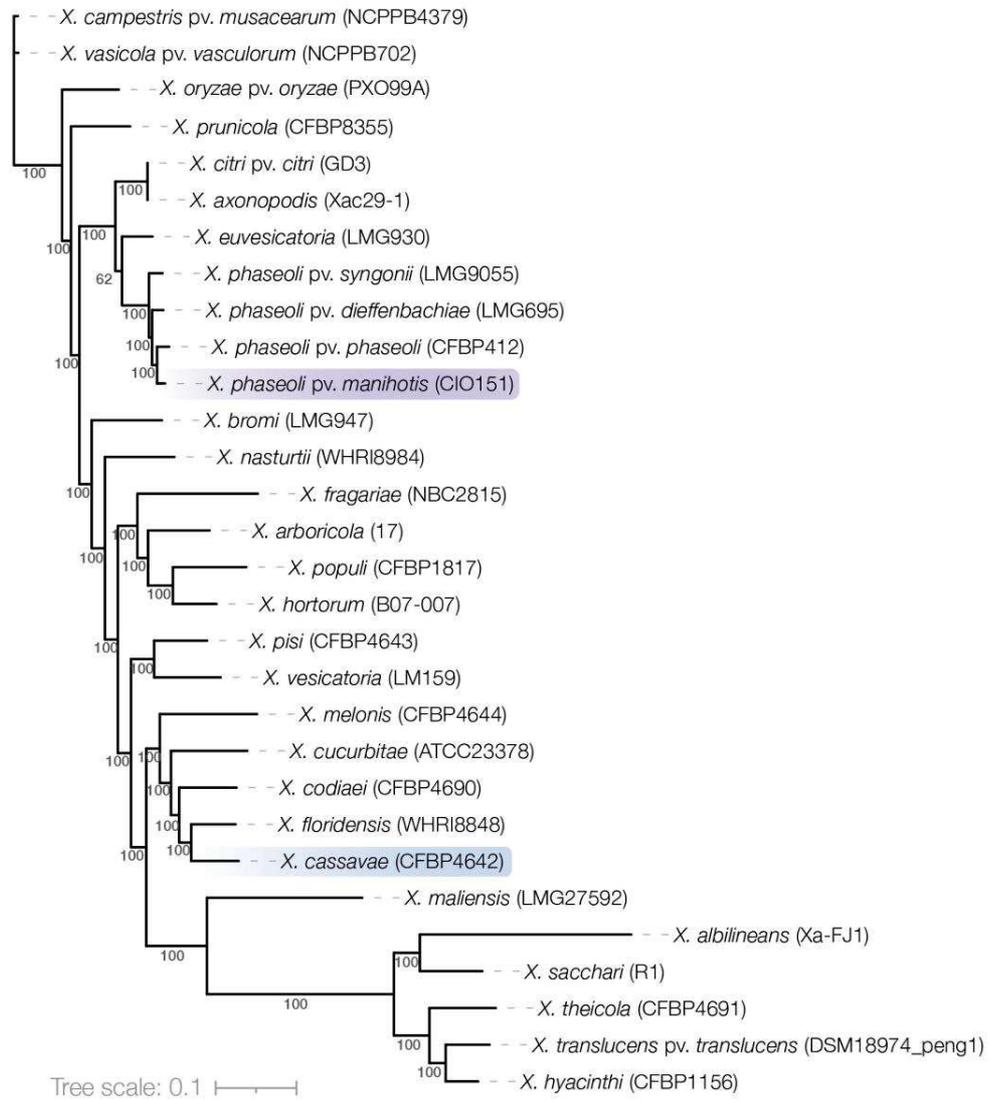


1280 Figure 1.



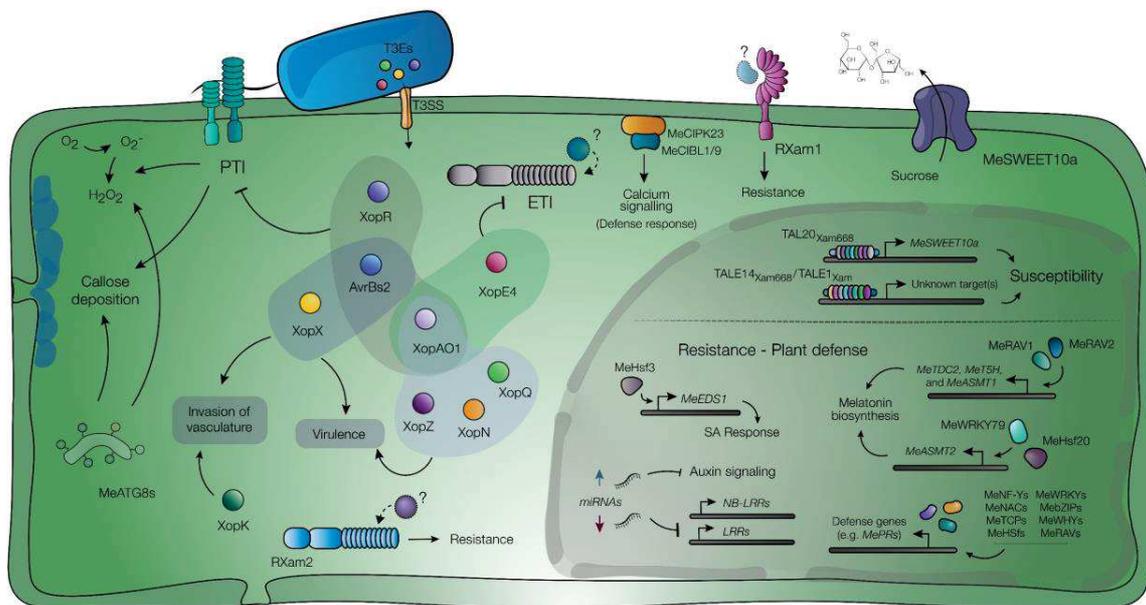
1281

1282 Figure 2.



1283

1284 Figure 3.



1285

1286 Figure 4.

daTALbase: A Database for Genomic and Transcriptomic Data Related to TAL Effectors

Alvaro L. Pérez-Quintero,^{1,2} Léo Lamy,¹ Carlos A. Zarate,¹ Sébastien Cunnac,¹ Erin Doyle,³ Adam Bogdanove,^{3,4} Boris Szurek,¹ and Alexis Dereeper^{1,†}

¹IRD, Cirad, Université Montpellier, IPME, Montpellier (34000), France; ²Institut de Biologie de l'École Normale Supérieure, École Normale Supérieure, CNRS, INSERM, PSL Research University, 75005 Paris, France; ³Department of Biology, Doane University, 1014 Boswell Avenue, Crete, NE 68333, U.S.A.; and ⁴Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, 334 Plant Science Building, Ithaca, NY 14853, U.S.A.

Accepted 15 November 2017.

Transcription activator-like effectors (TALEs) are proteins found in the genus *Xanthomonas* of phytopathogenic bacteria. These proteins enter the nucleus of cells in the host plant and can induce the expression of susceptibility genes (*S* genes), triggering disease. TALEs bind the promoter region of *S* genes following a specific code, which allows the prediction of binding sites based on TALEs amino acid sequences. New candidate *S* genes can then be discovered by finding the intersection between genes induced in the presence of TALEs and genes containing predicted effector binding elements. By contrasting differential expression data and binding site predictions across different datasets, patterns of TALE diversification or convergence may be unveiled, but this requires the seamless integration of different genomic and transcriptomic data. With this in mind, we present daTALbase, a curated relational database that integrates TALE-related data including bacterial TALE sequences, plant promoter sequences, predicted TALE binding sites, transcriptomic data of host plants in response to TALE-harboring bacteria, and other associated data. The database can be explored to uncover new candidate *S* genes as well as to study variation in TALE repertoires and their corresponding targets. The first version of the database here presented includes data for *Oryza* sp.–*Xanthomonas* pv. *oryzae* interactions. Future versions of the database will incorporate information for other pathosystems involving TALEs.

The plant-pathogenic bacteria of genus *Xanthomonas* cause devastating diseases on a wide range of hosts and impact the yield of important crops such as rice, cassava, cotton, wheat, banana, mango, citrus, and cabbage, both quantitatively and qualitatively (Hayward 1993). In rice, the two closely related pathovars *Xanthomonas oryzae* pv. *oryzae* and *Xanthomonas oryzae* pv. *oryzicola* are responsible for two diseases, bacterial leaf blight and bacterial leaf streak, respectively. *X. oryzae* pv. *oryzae* is a vascular pathogen that enters leaves via hydathodes and colonizes the xylem parenchyma, while *X. oryzae* pv. *oryzicola* is an intercellular pathogen that enters through stomata and colonizes the mesophyll

apoplast (Niño-Liu et al. 2006; White and Yang 2009). Yield losses caused by these pathogens can amount up to 50% for *X. oryzae* pv. *oryzae* and 30% for *X. oryzae* pv. *oryzicola*. These diseases are therefore important constraints for rice production worldwide (Niño-Liu et al. 2006).

To colonize their host, *Xanthomonas* species, like other bacteria, rely on a type III secretion system specialized in the injection of virulence factors (also called type III effectors [T3Es]) into the host cell. They notably rely on a family of T3Es known as transcription activator-like effectors (TALEs), which act as plant transcription factors to reprogram the host transcriptome upon translocation into the plant cell and localization to the nucleus (Boch and Bonas 2010). To induce host genes, TALEs are able to directly bind DNA through their central repeat region according to the so-called TALE code (Boch et al. 2009; Moscou and Bogdanove 2009). Each repeat forms a hairpin structure made by two α -helices connected by a loop. Upon binding to DNA, the repeats form a superhelix wrapped around the DNA major groove with the loops from each repeat on the inner side of the helix, directly interacting with the DNA (Deng et al. 2012; Mak et al. 2013). The specificity of interaction with DNA is determined by amino acids located within the loop of each repeat at positions 12 and 13, which are usually highly variable and are, thus, designated RVD, for repeat variable di-residues. Within the RVD, amino acid 12 helps stabilize the loop, while amino acid 13 can interact directly or indirectly with the nitrogenous bases through hydrogen bonds and van der Waals forces (Deng et al. 2012; Mak et al. 2012).

TALE-mediated induction of a subset of genes, referred to as susceptibility genes (*S* genes), can promote host colonization and disease. To date, several of them have been described (particularly in rice), and *S* genes with similar function are often targeted by multiple TALEs in a redundant and convergent manner (Boch et al. 2014; Pérez-Quintero et al. 2013). *S* genes targeted by *X. oryzae* pv. *oryzae* TALEs include sugar transporters of the *SWEET* family (Boch et al. 2014; Chen et al. 2010) as well as multiple types of transcription factors (Sugio et al. 2007). In contrast, OsSULTR3;6, a putative sulfate transporter, is, so far, the only *S* gene identified as target for *X. oryzae* pv. *oryzicola* TALEs (Cernadas et al. 2014). Proposed common targets for *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* include the small RNA 2'-*O*-methyltransferase *HEN1* and a flavanone 3-hydroxylase (F3H), which have been shown to be induced by TALEs from both pathovars (Moscou and Bogdanove 2009; Pérez-Quintero et al. 2013), but no phenotype has yet been associated to their induction. Importantly, plants have evolved different mechanisms to detect or neutralize TALEs. They

[†]Corresponding author: Alexis Dereeper: E-mail: alexis.dereeper@ird.fr

Funding: This work was funded by the Fondation Agropolis grant number 1403-073, the Agence Nationale de la Recherche grant number ANR-14-CE19-443-0002, and the United States National Science Foundation, grant number IOS-1444511.

include loss-of-susceptibility alleles such as *xa13*, *xa25*, and *xa41*, in which TALE binding to the promoters of *S* genes is precluded by target sequence polymorphism (Chu et al. 2006; Hutin et al. 2015a and b; Liu et al. 2011). Other forms of resistance also entail direct recognition of TALE structures (potentially Xo1 and Xa1) and subsequent defense response activation (Ji et al. 2016; Read et al. 2016; Triplett et al. 2016) or so-called executor *E* gene (*Xa7*, *Xa10*, *Xa23*, *Xa27*) induction (Zhang et al. 2015).

Because the mechanism of action of TALEs is relatively well-understood, they have become an ideal probe to investigate physiological processes governing plant susceptibility to bacteria. Binding sites for TALEs can be predicted in the host genomes, using various available softwares (Doyle et al. 2012; Grau et al. 2013; Pérez-Quintero et al. 2013; Rogers et al. 2015), and these predictions can be contrasted with transcriptomic data to identify genes that are likely to be targets of TALEs, i.e., genes that contain a predicted binding site (effector-binding element [EBE]) in their promoters and that are shown to be induced in presence of a bacteria harboring the TALE (Noël et al. 2013). These candidate targets can then be tested experimentally for either a role in disease (Cernadas et al. 2014), resistance (Strauss et al. 2012), or both.

In recent years, genomic and transcriptomic resources for the rice-*X. oryzae* system have expanded considerably; transcriptomic profiles for plants infected with various *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae* strains are becoming increasingly available (Wilkins et al. 2015). SMRT (single molecule, real time) sequencing technologies now allow the straightforward assembly of TALE repetitive regions and several finished *X. oryzae* genomes with full TALome sequences (i.e., TAL effector repertoires) have also been released (Grau et al. 2016; Quibod et al. 2016; Wilkins et al. 2015). Likewise, recent sequencing projects have made available multiple sets of genomic sequences from rice, including fully assembled *de novo* genomes (Chen et al. 2013; Wang et al. 2014) and rich single nucleotide polymorphism (SNP) data encompassing more than 3,000 rice cultivars (The 3,000 Rice Genomes Project 2014). For TALE research, this data holds the promise of not only helping discover new *S* genes but, also, of bringing important insight into the coevolution of the interacting organisms.

While there are currently multiple tools available to predict TALE binding sites (Doyle et al. 2012; Grau et al. 2013; Pérez-Quintero et al. 2013; Rogers et al. 2015) as well as tools for analyzing genomic data (The 3,000 Rice Genomes Project 2014) and pathogen-specific transcriptomic data. (Dash et al. 2012), the type of data produced by these tools is often heterogeneous and comparisons among them are often burdensome and time-consuming. Seeing the need for an accessible way to interrogate these types of data, we here present daTALbase, a relational database that integrates publicly available TALE-related genomic and transcriptomic data. This database will easily allow users to explore TALE sequences from *X. oryzae*, their predicted targets in available *Oryza* sp. genomes, target expression in transcriptomic data, target genomic variation, and more. Future versions of the database will integrate data for other pathosystems.

RESULTS

Description of the database.

The database consists mainly of five types of information: i) TALE sequences, ii) predicted targets for these sequences in promoters of annotated genes in available genomes, iii) orthology relations among genes in the available genomes, iv) genetic variants in the predicted binding sites in promoters, and v) transcriptomic data.

daTALbase v.1 includes a total of 528 TALE sequences from two *X. oryzae* pathovars, *X. oryzae* pv. *oryzae* (30 strains, 270

effectors) and *X. oryzae* pv. *oryzicola* (10 strains, 258 effectors) (Fig. 1A). RVD sequences for these available TALEs were used to predict EBEs on available assembled and annotated *Oryza* genomes (13 genomes in total) (Fig. 1B). A total of 3,405,793 putative EBEs were incorporated into the database. Among them, 259,000 were predicted in the reference *O. sativa* Nipponbare genome, corresponding to 39,811 potential target genes. More precisely, we found 8,472 genes targeted by a single TALE and 9,872 possible “hub” genes that were predicted to be targeted by at least 10 TALEs.

Distribution of EBEs along the *O. sativa* Nipponbare genome is reported in Figure 2 and reveals that predicted EBEs are distributed continuously and homogeneously along the chromosomes for both strains *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* taken together. All information regarding the scoring and location of the EBEs was included in the database. Orthology relations among annotated genes in the available genomes were also predicted to facilitate comparisons between different species or cultivars. We identified 50,015 orthologous gene sets containing 392,183 genes, accounting for between 66% (21,194 of 32,037 genes of *O. brachyantha*) and 92.7% (35,447 of 38,245 genes of *O. sativa* DJ123) of all predicted proteins. Additionally, we identified SNPs and indels in the predicted EBEs from publicly available data. In total, 112,202 SNPs and 13,605 indels from the 3,000 Rice Genomes dataset, 2,280 SNPs from the high density rice array (HDRA) were incorporated into daTALbase.

Published RNA-seq and microarray experiments comparing rice plants inoculated with various *X. oryzae* strains and compared with control conditions have been integrated into daTALbase. These included nine microarray experiments and one RNA-seq experiment, and represented experimental treatments involving 14 of the *X. oryzae* strains included in the database (Fig. 1A).

Querying the web interface.

daTALbase has been made available online. The interface is organized in five main tabs representing the main types of data integrated in the database: TALE sequences, EBE predictions, orthology relations among genes, transcriptomic data, and SNP/indel data. An additional tab “My gene lists” allows the user to compare lists generated in the other tabs, mainly to contrast EBE predictions with transcriptomic data. daTALbase also provides links to external sources for further exploration of data, including Talvez (Pérez-Quintero et al. 2013), QueTAL (Pérez-Quintero et al. 2015), GEO datasets (Edgar et al. 2002), as well as the Rice genome browser (JBrowse) of the South Green bioinformatics platform. In each tab, the data can be filtered according to relevant fields (i.e., strain for TALE sequences) and the results can be exported as Excel files (.xlsx).

daTALbase can be used for multiple types of queries, depending on the interests of the researcher, including, for example: What are the genes predicted as targets for all TALEs from a certain strain and are these genes induced? Is a certain target conserved across different *Oryza* genomes?

The interface was organized to be as intuitive as possible, so that users can perform these types of queries. The different tabs are connected to each other and allow researchers to easily find relationships among the different types of data as depicted in Figure 3. For example, users can select TALEs of a strain of interest using the filters available in the “TAL effector” tab (Fig. 3A) and, from there, they could find the predicted targets available in the database for any desired genome (Fig. 3D) or use the external link to do their own predictions using Talvez (Fig. 3B). Users can also use the external link to the QueTAL suite to draw phylogenetic relationships among TALEs of interest (Fig. 3C).

From the “TAL targets in plants” tab (Fig. 3D), users can see results for TALEs chosen in the “TAL effector” tab or they can

search for EBEs predicted in any genes of interest. For the displayed set of predicted targets, users can then check whether there is expression data available (Fig. 3H), they can display the genomic region of the EBE in a genome browser (Fig. 3F), they

can save a list of predicted target genes to compare with selected experiments in the “RNA-seq/microarray” tab (Fig. 3I), or they can search for associated SNP data in the available datasets (Fig. 3E). Detailed information of genomic variation is

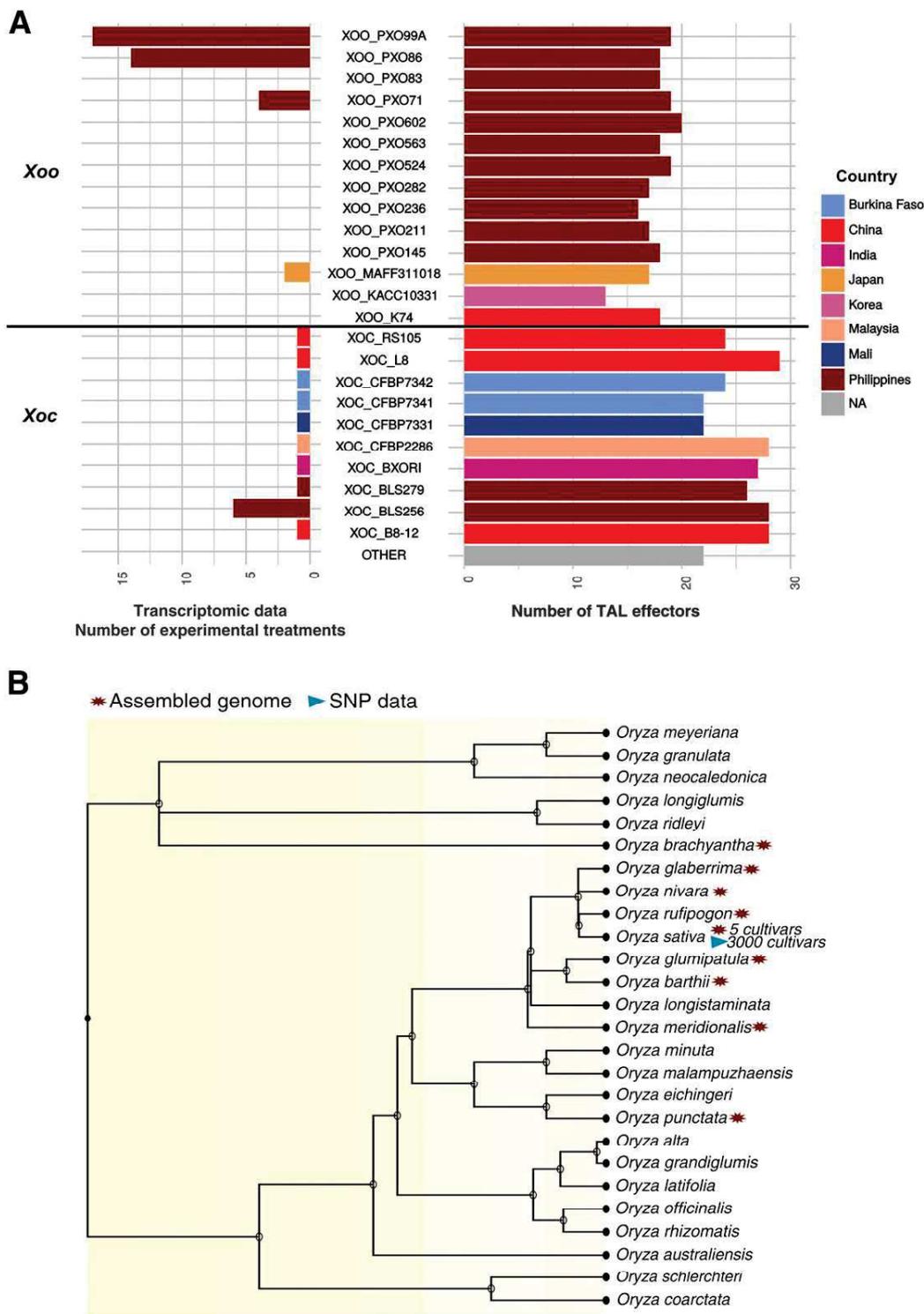


Fig. 1. A, Transcription activator-like effector (TALE) sequences and transcriptomic data included in daTALbase v1. On the left, transcriptomic experimental data included in the database associated to *Xanthomonas oryzae* strains. Each treatment represents a transcriptomic set from rice inoculated with the corresponding strain under a unique set of conditions (e.g., time postinoculation, rice variety). On the right, number of TALEs included in the database for each *X. oryzae* strain. Only strains with fully sequenced genomes are shown. Individual TALE sequences not coming from full genomes were added up as “other”. Bars are shaded according to the country of origin of each strain. Xoo = *X. oryzae* pv. *oryzae*, Xoc = *X. oryzae* pv. *oryzicola*. **B**, *Oryza* species phylogenetic tree adapted from Timetree (Hedges and Kumar 2009). Stars indicate independently assembled sequenced genomes, both draft and complete, for which predicted effector binding element data and orthology relations are available in daTALbase. The triangle indicates cultivars with available single nucleotide polymorphism data.

shown in the “SNPs/Indels” tab, these EBE variants can be of particular interest when looking for loss-of-susceptibility alleles. To assess the predicted impact of EBE variants on TALE binding, users can choose the option “Re-evaluate mutated EBEs prediction using Talvez”, which allows running TALE binding predictions on the different variants and compare their prediction scores.

In the “orthologs” tab (Fig. 3G), users can look for genes similar to any gene of interest in the available genomes and, then, look for predicted EBEs in these orthologs. Finally, the RNA-seq/microarray tab (Fig. 3H) allows users, in addition to obtaining data for previously selected genes, to explore differentially expressed genes in any of the available experiments. Users can, for example, select experiments showing genes induced in the presence of their strain of interest, save this list, and then compare it to predicted EBEs for TALEs from a said strain, using the previously described tab. For any set of genes, this tab also displays bar plots showing expression values in the relevant experiments, one gene at a time. Other possible interactions with the data are displayed in Figure 3.

Examples of usage and analysis of results from the database.

If users are interested in a specific strain of *X. oryzae*, they can use the database to identify candidate targets for all TALEs from this strain. For example, we can study the candidate targets for TALEs from the strain *X. oryzae* pv. *oryzicola* BLS256. Using daTALbase (“TAL effectors” tab, filtering according to strain), we can see that this strain has 28 TALEs, whose

predicted EBEs could be identified in the Nipponbare genome by using the link to the “TAL targets in plants” tab (2,722 genes in total, with rank less than 100). We can then explore all the experimental data available for this strain (six treatments) and identify 2,525 differentially expressed genes in the presence of this strain. Intersection between predictions and expression data represents 182 candidate target genes (Fig. 4A), which includes previously identified targets for TALEs from this strain (Cernadas et al. 2014).

This analysis can be made for each of the strains for which there is available experimental data. This reveals 747 candidate target genes for 315 TALEs from 14 strains. A hierarchical clustering based on induction of target genes reveals that strains in the database can be grouped into three main groups: i) Asian *X. oryzae* pv. *oryzae*, ii) African and Indian *X. oryzae* pv. *oryzicola*, and iii) east Asian *X. oryzae* pv. *oryzicola* (Fig. 4B), suggesting that strains from related populations have similar TALE repertoires and activate similar sets of genes, as has been previously suggested for *X. oryzae* pv. *oryzicola* (Wilkins et al. 2015) and *X. oryzae* pv. *oryzae* (Quibod et al. 2016).

Notably, some genes were identified as targets of multiple strains including both *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae*. These included “LOC_Os01g40290”, an expressed protein with unknown function, predicted as a target for TALEs from 12 of the 14 strains analyzed and differentially expressed in 41 conditions in the available transcriptome data. Other common targets included *OsSULTR3;6* (LOC_Os01g52130), a *S* gene involved in sulfate transport previously reported as a

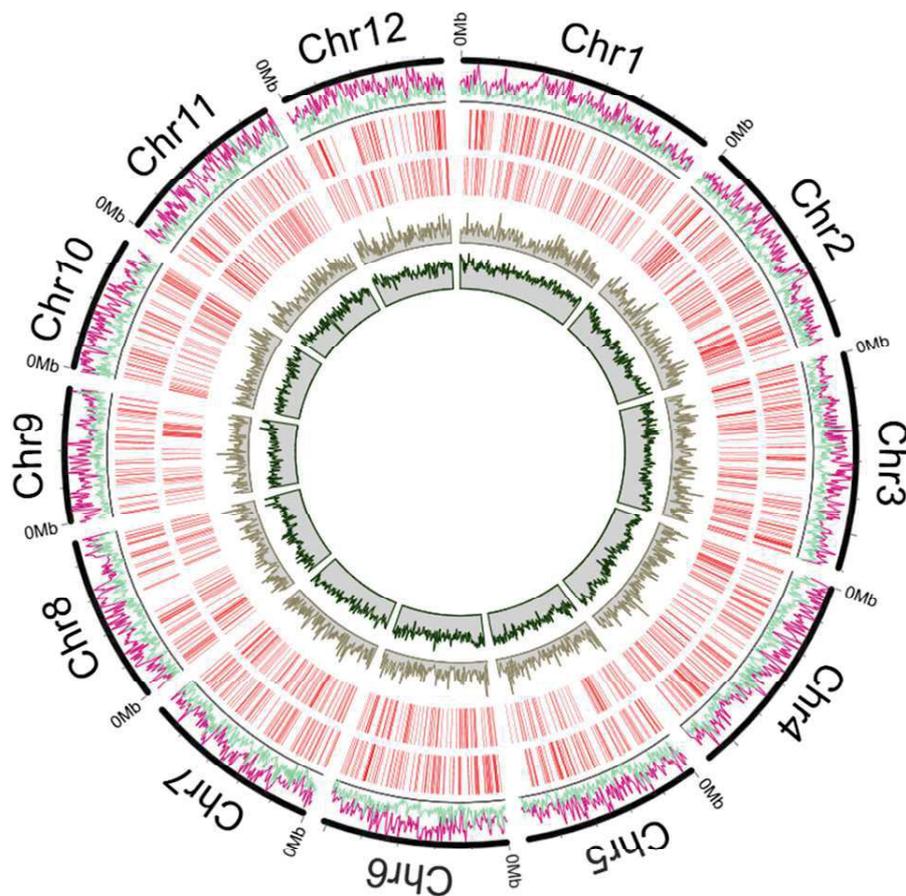


Fig. 2. Distribution of genomic features along the rice genome *O. sativa* cv. Nipponbare MSU7 (200-kb sliding windows). 1) Gene density, 2) total effector binding elements (EBEs) predicted with rank less than 500, 3) target genes for EBEs predicted with rank less than 10 for *Xanthomonas oryzae* pv. *oryzicola*, and 4) *X. oryzae* pv. *oryzae* transcription activator-like effectors, 5) single nucleotide polymorphisms and indels from The 3,000 Rice Genome Project (2014) located in predicted EBEs.

common target for *X. oryzae* pv. *oryzicola* (Cernadas et al. 2014), and *OsHEN1* (LOC_Os07g06970), a common target for both *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae*, involved in the stability of small RNAs but with a yet-unknown function in the rice-*X. oryzae* interaction (Moscou and Bogdanov 2009).

Users can also use the database to explore commonalities of target genes. For instance, it has been recently reported that TALEs can induce gene expression bidirectionally (Streubel et al. 2017; Wang et al. 2017), that is, binding to either strand in the promoter region of a gene can drive transcription of the downstream gene. With this in mind, we can look at the

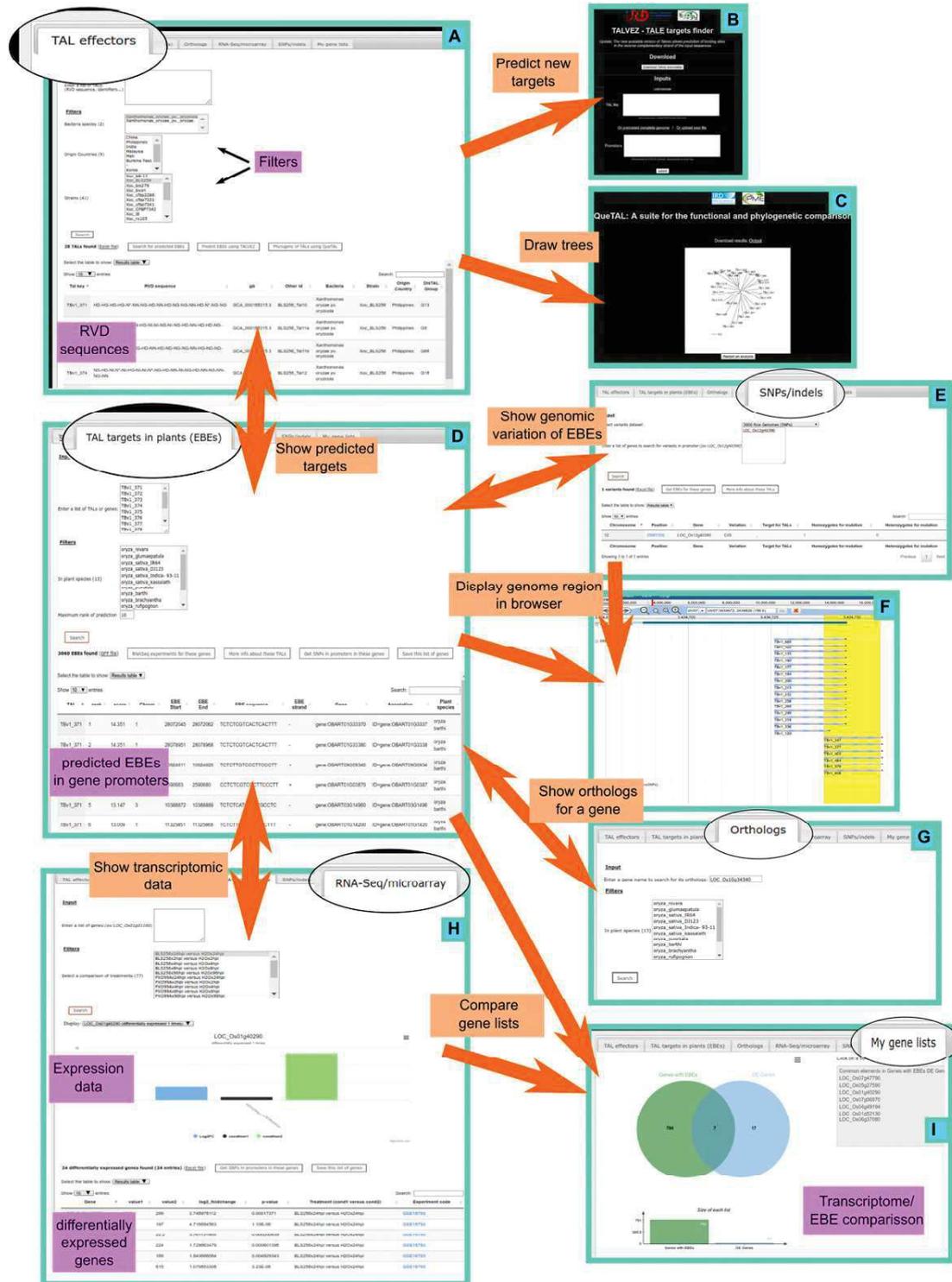


Fig. 3. Navigation process and links using the daTALbase interface, shown in screenshots of different tabs or links accessible from the web interface. Thick arrows indicate links between different tabs, bidirectional arrows indicate that queries can be made in both directions between the linked tabs. **A**, Transcription activator-like (TAL) effector tab, **B**, link to Talvez prediction, **C**, link to QueTAL phylogeny, **D**, TAL targets in plants effector binding elements (EBEs) tab, **E**, single nucleotide polymorphisms (SNPs)/Indels tab, **F**, link to Jbrowse displaying EBEs and SNPs, **G**, orthologs tab, **H**, RNA-Seq/microarray tab, and **I**, My gene lists tab.

frequency at which candidate target genes contain EBEs in the forward (same orientation as the gene) or reverse strand of the promoter. This suggests that binding in the forward strand is more common (almost twice) than binding in the reverse

strand but that, nonetheless, a large number of targets might be induced through “antisense” transcription (Fig. 4B). It’s also possible that this is the result of unknown biases in the target predictions.

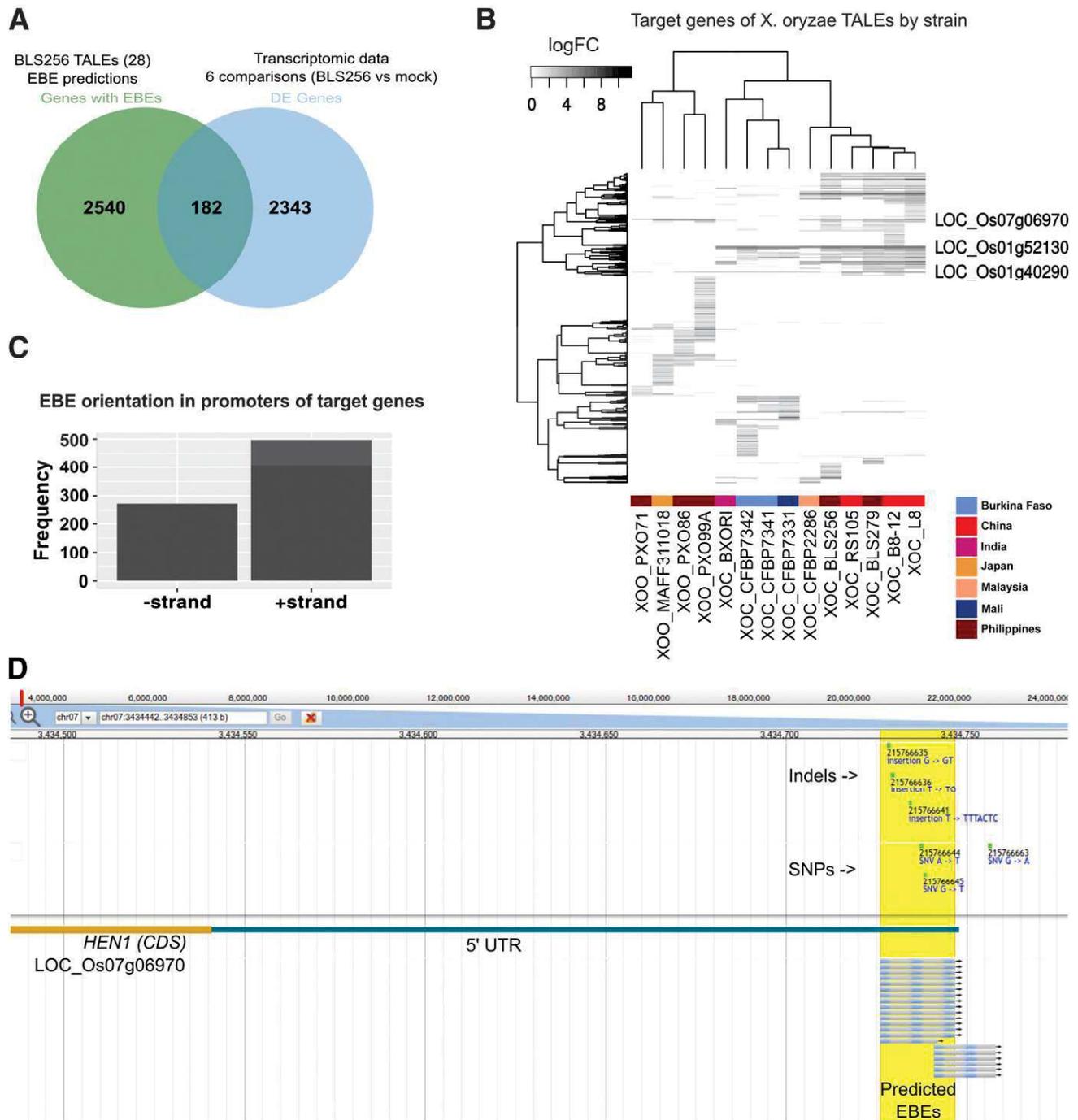


Fig. 4. A, Venn diagram showing the intersection between genes containing predicted effector binding elements (EBEs) for transcription activator-like effectors (TALEs) from *Xanthomonas oryzae* pv. *oryzicola* BLS256 and genes differentially expressed in transcriptomic data comparing plants inoculated with *X. oryzae* pv. *oryzicola* BLS256 against a control, as identified using daTALbase. A gene is considered as differentially expressed if it is identified as such in any of the six experimental treatments evaluated. **B**, Target genes were identified for TALEs from 14 *X. oryzae* strains, i.e., genes containing predicted EBEs for the corresponding strain and induced by said strain in transcriptomic data as shown for BLS256 in A. The heatmap shows the highest log fold change for each gene in the treatments evaluated. On the left, hierarchical clustering showing grouping of target genes (top) and hierarchical clustering showing grouping of strains used (bottom); bottom right, country of origin of each strain is shown. Three genes shown to be common targets for various strains are highlighted. **C**, For the target genes identified in B, the bar graph shows the frequency at which EBEs were identified in the forward strand (same orientation as the gene) or the reverse strand in the promoter of each gene. **D**, Screenshot of a JBrowse session showing the 5' untranslated region (UTR) of *OsHEN1* in the *O. sativa* cv. Nipponbare genome. The region targeted by several TALEs from *X. oryzae* pv. *oryzicola* or *X. oryzae* pv. *oryzae* is highlighted. Tracks on the top indicate single nucleotide polymorphisms (SNPs) and indels detected in The 3,000 Rice Genome Project dataset, and tracks at the bottom indicate predicted EBEs.

Finally, a user can also query the database to look further into genomic variation in the predicted EBEs for these genes. For example, we can look for possible orthologs of *HEN1*, thus identifying 11 orthologs in the 13 *Oryza* genomes included in the database (no orthologs were identified in *O. punctata* or *O. sativa* cv. *kassalath* under the parameters used). When looking for predicted EBEs for these orthologs, it can be seen that EBEs for TALEs of both *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* are greatly conserved across the different *Oryza* species, with some variation in the *O. glaberrima* and *O. barthii* genomes (Table 1). Likewise, we can look at the available genetic variants for this region, which reveals three SNPs and three insertion or deletion events identified in the 3,000 accessions. A researcher could then perform wet-lab experiments to associate the variation found in orthologous EBEs with possible phenotypes in the presence of strains harboring *HEN1*-inducing TALEs. This could help in the search for loss-of-susceptibility alleles as a source of resistance against *Xanthomonas* spp.

DISCUSSION

Data curation and future improvement.

daTALbase is conceived to be a constantly expanding and curated database for TALE-related data. The current version only integrates data related to the rice-*X. oryzae* system because a wealth of transcriptomic and genomic resources is available for this system. We are currently in the process of integrating additional rice transcriptomic data and TALE sequences generated in our laboratory related to African strains of *X. oryzae* pv. *oryzae* that await publication (T. T. Tran, A. L. Perez-Quintero, M. Hutin, and B. Szurek *in preparation*) and adding recently released rice genomes (Li et al. 2017), and we plan to add more data as it becomes available. New data can also be integrated upon request.

Future versions of the database will incorporate data related to *Xanthomonas* pathogens of beans, cabbage, citrus, wheat, and cassava that are currently being generated in collaboration with partners from the CropTAL project and the International Center for Tropical Agriculture (CIAT) cassava website. The working version of the cassava database integrates publicly available data corresponding to seven TALEs sequences from *Xanthomonas axnopodis* pv. *manihotis* (Bart et al. 2012; Castiblanco et al.

2013), their predicted targets on the cassava genome (v 6.1) (Bredeson et al. 2016), and two sets of RNA-seq data (Cohn et al. 2016, 2014; Muñoz-Bodnar et al. 2014). We expect to expand this database to include newly sequenced TALEs upon their release. Integrating other hosts will be of special interest to study convergence and evolution of targets, considering how some targets like the *SWEET* family of genes are being found to be important for different pathosystems (Cohn et al. 2014; Cox et al. 2017; Hu et al. 2014).

We also envision improving on the methods used for curating the data, including the possibility of adding EBE predictions using other available software (Doyle et al. 2012; Grau et al. 2013; Rogers et al. 2015) and improving the existing predictions using different sets of parameters. Likewise, we plan on improving the strategy to identify orthologs to make sure it is suitable for the inclusion of phylogenetically distant genomes. Finally, we hope daTALbase will constitute both a reference and an analysis tool for the community of TALE researchers and we encourage feedback for its improvement and curation.

MATERIALS AND METHODS

Data collection, contents, and features.

TALE sequences. TALE sequences have been retrieved from the National Center for Biotechnology Information protein databases from the two *X. oryzae* pathovars *X. oryzae* pv. *oryzae* (30 strains, 270 effectors) and *X. oryzae* pv. *oryzicola* (10 strains, 258 effectors) (Fig. 1A). Of these sequences, 487 were extracted from complete genome sequences. Each TALE was assigned an identifying number for the database in the format TBv1_001 (TBv1 indicates daTALbase version 1). For each TALE, associated information was registered including: published identifiers (e.g., PthXo1, Tal2g), gene bank database identifier of the TALE nucleotide sequence or the corresponding genome sequence, RVD sequence, the *X. oryzae* strain in which it was found, and its country of origin. TALEs with identical sequences found in different strains are considered as different entries in the database. RVD sequences were extracted using in-house perl scripts. TALEs were also assigned to groups according to similarities in their repeat sequences, as determined using the program DisTAL (Pérez-Quintero et al. 2013).

Table 1. Predicted effector binding elements (EBEs) for two transcription activator-like effectors (TALEs) from *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* in the promoters of orthologs of *HEN1* as identified using daTALbase

TALE name/genome	Target gene	Talvez rank	Talvez score	EBE sequence
XocBLS256_Tal1c				
<i>Oryza barthii</i>	OBART07G04060	1	15.51	TCCCCTCGCTTCCCTT
<i>O. glaberrima</i> CG14	ORGLA07G0034600	1	15.522	TCCCCTCGCTTCCCTT
<i>O. glumaepatula</i>	OGLUM07G03470	1	15.471	TCCCCTCGCTTCCCTT
<i>O. meridionalis</i>	OMERI07G02610	1	15.406	TCCCCTCGCTTCCCTT
<i>O. nivara</i>	ONIVA07G02640	1	15.493	TCCCCTCGCTTCCCTT
<i>O. rufipogon</i>	ORUFIO7G03760	1	15.462	TCCCCTCGCTTCCCTT
<i>O. sativa</i> DJ123	Maker-scaffold_131-snap-gene-2.32	1	15.533	TCCCCTCGCTTCCCTT
<i>O. sativa indica</i> 93-11	BGIOSGA024768	1	15.471	TCCCCTCGCTTCCCTT
<i>O. sativa</i> IR65	Maker-scaffold_17-snap-gene-8.38	1	17.377	TCCCCTCGCTTCCCTT
<i>O. sativa</i> Nipponbare	LOC_Os07g06970	1	15.434	TCCCCTCGCTTCCCTT
XooPXO99A_Tal9a				
<i>O. barthii</i>	OBART07G04060	144	11.893	TCCCTTCCCTAAATCCCCAACT
<i>O. glaberrima</i> CG14	ORGLA07G0034600	123	11.899	TCCCTTCCCTAAATCCCCAACT
<i>O. glumaepatula</i>	OGLUM07G03470	1	17.395	TCCCTTCCCTAAATCCCCAACT
<i>O. meridionalis</i>	OMERI07G02610	1	17.364	TCCCTTCCCTAAATCCCCAACT
<i>O. nivara</i>	ONIVA07G02640	1	17.406	TCCCTTCCCTAAATCCCCAACT
<i>O. rufipogon</i>	ORUFIO7G03760	5	14.858	TCCCTTCCCTAAATCCCCAACT
<i>O. sativa</i> DJ123	Maker-scaffold_131-snap-gene-2.32	1	17.426	TCCCTTCCCTAAATCCCCAACT
<i>O. sativa indica</i> 93-11	BGIOSGA024768	1	17.395	TCCCTTCCCTAAATCCCCAACT
<i>O. sativa</i> IR64	Maker-scaffold_17-snap-gene-8.38	1	15.493	TCCCTTCCCTAAATCCCCAACT
<i>O. sativa</i> Nipponbare	LOC_Os07g06970	1	17.377	TCCCTTCCCTAAATCCCCAACT

TALE targets (EBEs) in different *Oryza* genomes. Genomes included in the database are the reference *O. sativa* cv. Nipponbare (assembly and annotation version MSU7) from the Rice Genome Annotation Project (Kawahara et al. 2013). Ten rice genomes were obtained from the Ensembl genome database release 35: *O. barthii* (ABRL000000000), *O. brachyantha* (v1.4b), *O. glaberrima* cv. CG14 (AGI1.1), *O. glumaepatula* (*O. glumipatula*) (ALNW000000000), *O. meridionalis* (ALNW000000000), *O. nivara* (AWHD000000000), *O. punctata* (AVCL000000000), *O. rufipogon* (PRJEB4137), and *O. sativa* cv. 93-11 (ASM465v1). *O. sativa* cvs. DJ123 and IR64 (versions CSHL 1.0) were obtained from the Schatz lab (Schatz et al. 2014) and the *O. sativa* cv. Kasalath genome (v. NIAS-RAP-1.0) was obtained from rap-db (Ohyang et al. 2006).

Predictions were made using the Talvez software (Pérez-Quintero et al. 2013). This prediction tool uses the TALE-DNA code to convert the RVD sequence in a positional weight matrix. Then, the program uses the matrix to scan all the possible EBEs in the host genome sequence and gives a rank and a score for each putative EBE. For each of the genomes used, promoter sequences (1,000 bp upstream) were extracted from all annotated genes and Talvez was used to find EBEs on both strands of the promoter to reflect their bidirectional binding, allowing 500 hits per TALE, a minimum score of 7, and using updated RVD-DNA specificities that reflect recent experimental data for TALE-binding, including predictions for all possible RVD combinations (Yang et al. 2014) and the contribution of strong versus weak RVDs (Streubel et al. 2012) as employed in the program FuncTAL (Pérez-Quintero et al. 2015).

Transcriptomic data. Published RNA-seq and microarray experiments comparing rice plants inoculated with various *X. oryzae* strains and compared with control conditions have been integrated into daTALbase. Nine microarray experiments

(GSE16793, GSE19239, GSE19844, GSE33411, GSE34192, GSE36093, GSE36272, GSE43050, GSE8216) and one RNA-seq experiment (GSE67588) were used to feed the database. Microarray data were obtained from the PlexDB database, mean MAS-normalized values were downloaded from the database and differential expression was assessed using the limma package, as described by Pérez-Quintero et al. (2013) and Smyth (2005). RNA-seq data were obtained from GEO datasets (Edgar et al. 2002) and were processed as reported in (Wilkins et al. 2015).

Annotation of probes and RNA-seq mappings were based on the reference genome sequence of Nipponbare (MSU7 annotation). For all experiments, only genes considered as significantly induced or repressed, when the *P* value was <0.05, were kept and stored in the database. In total, 104,346 entries are recorded in a dedicated table for expression information, representing 14,071 differentially expressed genes potentially involved in the molecular basis of diseases caused by *X. oryzae*.

Orthology information. To allow comparisons between the available GFF-file *Oryza* genomes, the annotated proteome for each species and cultivar was obtained from the corresponding assembly, and the reconstruction of orthology groups was based on the commonly used approach combining an “all against all” BLASTP of whole proteomes and the clustering of blast results by the OrthoMCL suite (Li et al. 2003) (default parameters). Future versions of the database will include orthology with available genomes from other genera to allow the study of TALE target convergence in a wide scale.

Genetic variants (SNPs and indels) in predicted EBEs. The 3,000 Rice Genomes Project (2014) provides a considerable genetic resource recording millions of SNPs and indels. Another important resource with genotype information for 700,000 SNPs from a diverse set of rice accessions is the HDRA (McCouch

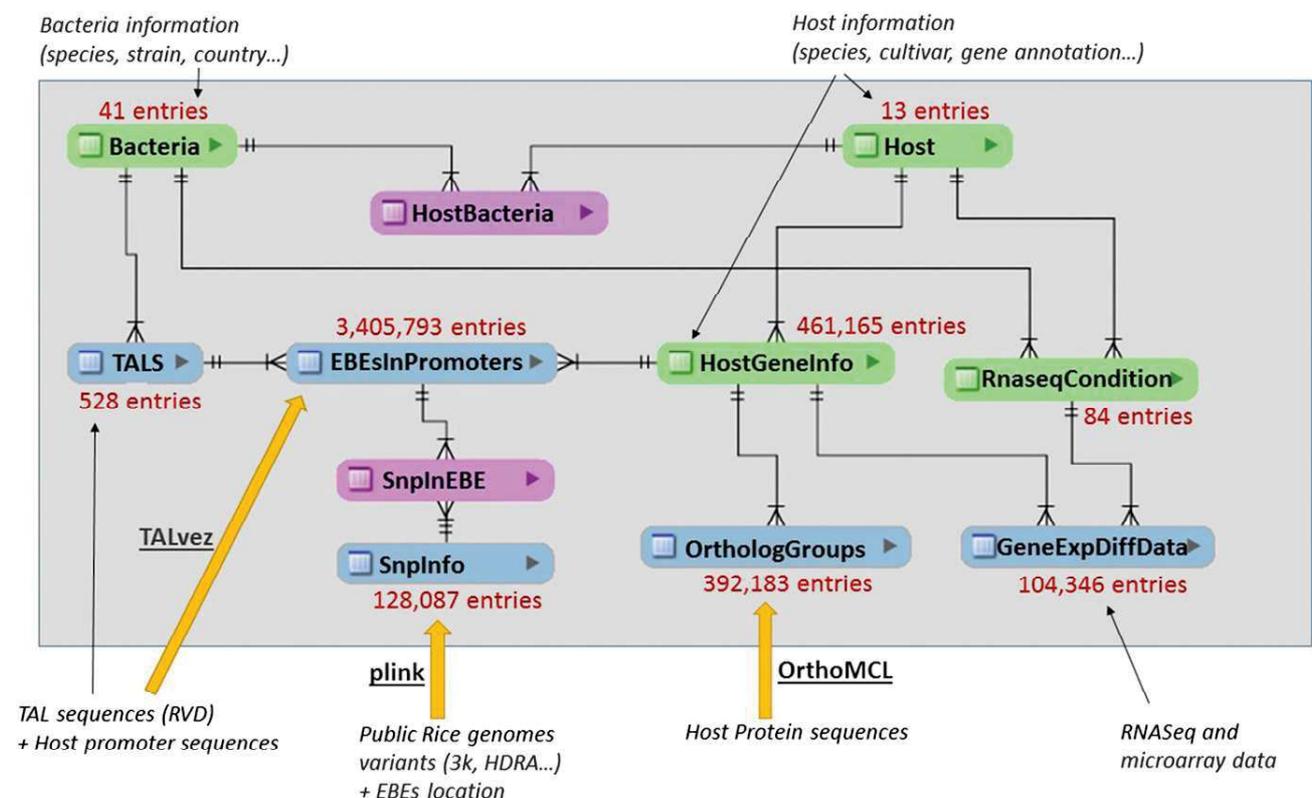


Fig. 5. daTALbase architecture and process for constructing the daTALbase database. Each square represents a table, number of entries in version 1 of daTALbase are also shown.

et al. 2016) with more than 1,600 genotyped accessions. We used this data to search for variations within the predicted EBEs in the *O. sativa* Nipponbare reference.

Variants overlapping a predicted EBE were extracted from these resources, using PLINK 1.9 (Chang et al. 2015) and EBE coordinates. In addition, EBEs and associated polymorphisms were also integrated as specific new tracks into the Rice genome browser (JBrowse) of the South Green bioinformatics platform.

System architecture and implementation.

The daTALbase database system combines both a relational MySQL database and JSON flat files. The web interface is implemented in Perl CGI scripts running on an Apache web server. The interactivity and smooth navigation is allowed by JavaScript and various libraries, such as jQuery user interface and Highcharts application programming interface, to manage graphical layouts, JVenn (Bardou et al. 2014) to handle Venn diagram representation, or DataTables plugin for jQuery to facilitate the manipulation of output tables. To manage the access to private data, the application is also equipped with login authentication to keep private entries password protected.

The database is normalized and consists mainly of nine tables, which approximately correspond either to the information reported by the different tabs of the web interface (TALS, EBEsInPromoters, OrthologGroups, GeneExpDiffData, SnpInfo) derived from genome wide analyses or to sparser information that can be applied for filtering (Bacteria, Host, HostGeneInfo, RnaseqCondition), plus two additional association tables. Tables and processes associated with the data are summarized in Figure 5. Finally, the application also includes a series of Perl scripts facilitating the extraction, conversion, and integration of new data.

Data availability.

The instance for rice hosted at the French Research Institute for Development (IRD) may be accessed at the daTALbase website. The source code of the application, including Perl, CGI, as well as SQL scripts for populating the database, is available for download and installation at GitHub South Green. Full portable copies of the current release of the database including the data presented here is available upon request.

ACKNOWLEDGMENTS

We thank the South Green bioinformatics platform and the French Research Institute for Development (IRD) bioinformatics “i-trop” for hosting the database and providing computational resources. A. Pérez-Quintero was supported by doctoral fellowship awarded by the Erasmus Mundus Action 2 PANACEA, PRECIOSA program of the European Community. C. A. Zarate is supported by the Allocations de recherche pour une thèse au Sud (ARTS) program (IRD). This project was supported by a grant from Agence Nationale de la Recherche (ANR-14-CE19-443-0002) and from Fondation Agropolis (number 1403-073) and from the United States National Science Foundation (IOS-1444511) to A. Bogdanove and E. Doyle. We also acknowledge L.-A. Becerra and A. Gkanogiannis for their collaboration in the deployment of the cassava instance of daTALbase at the International Center for Tropical Agriculture.

LITERATURE CITED

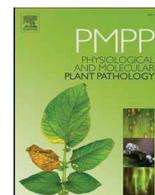
Bardou, P., Mariette, J., Escudie, F., Djemiel, C., and Klopp, C. 2014. jvenn: An interactive Venn diagram viewer. *BMC Bioinformatics* 15:293.
Bart, R., Cohn, M., Kassen, A., McCallum, E. J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L. M., Rodrigues Neto, J., Verdier, V., Santana, M. A., Kositcharoenkul, N., Vanderschuren, H., Gruissem, W., Bernal, A., and Staskawicz, B. J. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc. Natl. Acad. Sci. U.S.A.* 109:E1972-E1979.

Boch, J., and Bonas, U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: Discovery and function. *Annu. Rev. Phytopathol.* 48:419-436.
Boch, J., Bonas, U., and Lahaye, T. 2014. TAL effectors—Pathogen strategies and plant resistance engineering. *New Phytol.* 204:823-832.
Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326:1509-1512.
Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edinger-Gonzales, E., Grimwood, J., Schmutz, J., Rabbi, I. Y., Egesi, C., Nauluvula, P., Lebot, V., Ndunguru, J., Mkamilo, G., Bart, R. S., Setter, T. L., Gleadow, R. M., Kulakow, P., Ferguson, M. E., Rounsley, S., and Rokhsar, D. S. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* 34:562-570.
Castiblanco, L. F., Gil, J., Rojas, A., Osorio, D., Gutiérrez, S., Muñoz-Bodnar, A., Perez-Quintero, A. L., Koebnik, R., Szurek, B., López, C., Restrepo, S., Verdier, V., and Bernal, A. J. 2013. TALE1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. *Mol. Plant Pathol.* 14:84-95.
Cernadas, R. A., Doyle, E. L., Niño-Liu, D. O., Wilkins, K. E., Bancroft, T., Wang, L., Schmidt, C. L., Caldo, R., Yang, B., White, F. F., Nettleton, D., Wise, R. P., and Bogdanove, A. J. 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 Pathog.* 10:e1003972.
Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. 2015. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4:7.
Chen, J., Huang, Q., Gao, D., Wang, J., Lang, Y., Liu, T., Li, B., Bai, Z., Luis Goicoechea, J., Liang, C., Chen, C., Zhang, W., Sun, S., Liao, Y., Zhang, X., Yang, L., Song, C., Wang, M., Shi, J., Liu, G., Liu, J., Zhou, H., Zhou, W., Yu, Q., An, N., Chen, Y., Cai, Q., Wang, B., Liu, B., Min, J., Huang, Y., Wu, H., Li, Z., Zhang, Y., Yin, Y., Song, W., Jiang, J., Jackson, S. A., Wing, R. A., Wang, J., and Chen, M. 2013. Whole-genome sequencing of *O. brachyantha* reveals mechanisms underlying *Oryza* genome evolution. *Nat. Commun.* 4:1595.
Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., Guo, W.-J., Kim, J.-G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F. F., Somerville, S. C., Mudgett, M. B., and Frommer, W. B. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468:527-532.
Chu, Z., Yuan, M., Yao, J., Ge, X., Yuan, B., Xu, C., Li, X., Fu, B., Li, Z., Bennetzen, J. L., Zhang, Q., and Wang, S. 2006. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes Dev.* 20:1250-1255.
Cohn, M., Bart, R. S., Shybut, M., Dahlbeck, D., Gomez, M., Morbitzer, R., Hou, B.-H., Frommer, W. B., Lahaye, T., and Staskawicz, B. J. 2014. *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. *Mol. Plant-Microbe Interact* 27:1186-1198.
Cohn, M., Morbitzer, R., Lahaye, T., and Staskawicz, B. J. 2016. Comparison of gene activation by two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility genes in cassava. *Mol. Plant Pathol.* 17:875-889.
Cox, K.L., Meng, F., Wilkins, K.E., Li, F., Wang, P., Booher, N.J., Carpenter, S.C.D., Chen, L.-Q., Zheng, H., Gao, X., Zheng, Y., Fei, Z., Yu, J.Z., Isakeit, T., Wheeler, T., Frommer, W.B., He, P., Bogdanove, A.J., Shan, L., 2017. TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. *Nat. Commun.* 8:15588.
Dash, S., Van Hemert, J., Hong, L., Wise, R. P., and Dickerson, J. A. 2012. PLEXdb: Gene expression resources for plants and plant pathogens. *Nucleic Acids Res.* 40:D1194-D1201.
Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. 2012. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* 335:720-723.
Doyle, E. L., Booher, N. J., Standage, D. S., Voytas, D. F., Brendel, V. P., Vandyk, J. K., and Bogdanove, A. J. 2012. TAL effector-nucleotide targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res.* 40:W117-W122.
Edgar, R., Domrachev, M., and Lash, A. E. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30:207-210.
Grau, J., Reschke, M., Erkes, A., Streubel, J., Morgan, R. D., Wilson, G. G., Koebnik, R., and Boch, J. 2016. AnnoTALE: Bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. *Sci. Rep.* 6:21077.
Grau, J., Wolf, A., Reschke, M., Bonas, U., Posch, S., and Boch, J. 2013. Computational predictions provide insights into the biology of TAL effector target sites. *PLOS Comput. Biol.* 9:e1002962.

- Hayward, A. C. 1993. The hosts of *Xanthomonas*. Pages 1-119 in: Xanthomonas. J. G. Swings, and E. L. Civerolo, eds. Springer, Dordrecht, The Netherlands.
- Hedges, S. B., and Kumar, S. 2009. The Timetree of Life. Oxford University Press, Oxford.
- Hu, Y., Zhang, J., Jia, H., Sosso, D., Li, T., Frommer, W. B., Yang, B., White, F. F., Wang, N., and Jones, J. B. 2014. *Lateral organ boundaries 1* is a disease susceptibility gene for citrus bacterial canker disease. Proc. Natl. Acad. Sci. U.S.A. 111:E521-E529.
- Hutin, M., Pérez-Quintero, A. L., Lopez, C., and Szurek, B. 2015a. MorTAL Kombat: The story of defense against TAL effectors through loss-of-susceptibility. Front. Plant Sci. 6:535.
- Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R., and Szurek, B. 2015b. A knowledge-based molecular screen uncovers a broad-spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. Plant J. 84:694-703.
- Ji, Z., Ji, C., Liu, B., Zou, L., Chen, G., and Yang, B. 2016. Interfering TAL effectors of *Xanthomonas oryzae* neutralize R-gene-mediated plant disease resistance. Nat. Commun. 7:13435.
- Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., Schwartz, D. C., Tanaka, T., Wu, J., Zhou, S., Childs, K. L., Davidson, R. M., Lin, H., Quesada-Ocampo, L., Vaillancourt, B., Sakai, H., Lee, S. S., Kim, J., Numa, H., Itoh, T., Buell, C. R., and Matsumoto, T. 2013. Improvement of the *O. sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6:4.
- Li, L., Stoeckert, C. J., Jr., and Roos, D. S. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Res. 13:2178-2189.
- Li, L.-F., Li, Y.-L., Jia, Y., Caicedo, A. L., and Olsen, K. M. 2017. Signatures of adaptation in the weedy rice genome. Nat. Genet. 49:811-814.
- Liu, Q., Yuan, M., Zhou, Y., Li, X., Xiao, J., and Wang, S. 2011. A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. Plant Cell Environ. 34:1958-1969.
- Mak, A. N.-S., Bradley, P., Bogdanove, A. J., and Stoddard, B. L. 2013. TAL effectors: Function, structure, engineering and applications. Curr. Opin. Struct. Biol. 23:93-99.
- Mak, A. N.-S., Bradley, P., Cernadas, R. A., Bogdanove, A. J., and Stoddard, B. L. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335:716-719.
- McCouch, S. R., Wright, M. H., Tung, C.-W., Maron, L. G., McNally, K. L., Fitzgerald, M., Singh, N., DeClerck, G., Agosto-Perez, F., Korniliev, P., Greenberg, A. J., Naredo, M. E. B., Mercado, S. M. Q., Harrington, S. E., Shi, Y., Branchini, D. A., Kuser-Falcão, P. R., Leung, H., Ebana, K., Yano, M., Eizenga, G., McClung, A., Mezey, J., 2016. Open access resources for genome-wide association mapping in rice. Nat. Commun. 7:10532.
- Moscou, M. J., and Bogdanove, A. J. 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326:1501-1501.
- Muñoz-Bodnar, A., Perez-Quintero, A. L., Gomez-Cano, F., Gil, J., Micheltore, R., Bernal, A., Szurek, B., and Lopez, C. 2014. RNAseq analysis of cassava reveals similar plant responses upon infection with pathogenic and non-pathogenic strains of *Xanthomonas axonopodis* pv. *manihotis*. Plant Cell Rep. 33:1901-1912.
- Niño-Liu, D. O., Ronald, P. C., and Bogdanove, A. J. 2006. *Xanthomonas oryzae* pathovars: Model pathogens of a model crop. Mol. Plant Pathol. 7:303-324.
- Noël, L. D., Denancé, N., and Szurek, B. 2013. Predicting promoters targeted by TAL effectors in plant genomes: From dream to reality. Front. Plant Sci. 4:333.
- Ohyang, H., Tanaka, T., Sakai, H., Shigemoto, Y., Yamaguchi, K., Habara, T., Fujii, Y., Antonio, B. A., Nagamura, Y., Imanishi, T., Ikeo, K., Itoh, T., Gojobori, T., and Sasaki, T. 2006. The Rice Annotation Project Database (RAP-DB): Hub for *Oryza sativa* ssp. *japonica* genome information. Nucleic Acids Res. 34:D741-D744.
- Pérez-Quintero, A. L., Lamy, L., Gordon, J. L., Escalon, A., Cunnac, S., Szurek, B., and Gagnevin, L. 2015. QueTAL: A suite of tools to classify and compare TAL effectors functionally and phylogenetically. Front. Plant Sci. 6:545.
- Pérez-Quintero, A. L., Rodríguez-R, L. M., Dereper, A., López, C., Koebnik, R., Szurek, B., and 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.
- Quibod, I. L., Perez-Quintero, A., Booher, N. J., Dossa, G. S., Grande, G., Szurek, B., Vera Cruz, C., Bogdanove, A. J., and Oliva, R. 2016. Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. Sci. Rep. 6:34137.
- Read, A. C., Rinaldi, F. C., Hutin, M., He, Y.-Q., Triplett, L. R., and Bogdanove, A. J. 2016. Suppression of *Xo1*-mediated disease resistance in rice by a truncated, non-DNA-binding TAL effector of *Xanthomonas oryzae*. Front. Plant Sci. 7:1516.
- Rogers, J. M., Barrera, L. A., Reyon, D., Sander, J. D., Kellis, M., Joung, J. K., and Bulky, M. L. 2015. Context influences on TALE-DNA binding revealed by quantitative profiling. Nat. Commun. 6:7440.
- Schatz, M. C., Maron, L. G., Stein, J. C., Hernandez Wences, A., Gurtowski, J., Biggers, E., Lee, H., Kramer, M., Antoniou, E., Ghiban, E., Wright, M. H., Chia, J., Ware, D., McCouch, S. R., and McCombie, W. R., 2014. Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of *aus* and *indica*. Genome Biol. 15:506.
- Smyth, G. K. 2005. LIMMA: Linear models for microarray data. Pages 397-420 in: Bioinformatics and Computational Biology Solutions Using R and Bioconductor, Statistics for Biology and Health. R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, S. Dudoit, eds. Springer, New York.
- Strauss, T., van Poecke, R. M. P., Strauss, A., Römer, P., Minsavage, G. V., Singh, S., Wolf, C., Strauss, A., Kim, S., Lee, H.-A., Yeom, S.-I., Parniske, M., Stall, R. E., Jones, J. B., Choi, D., Prins, M., and Lahaye, T. 2012. RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. Proc. Natl. Acad. Sci. U.S.A. 109:19480-19485.
- Streubel, J., Baum, H., Grau, J., Stuttmann, J., and Boch, J. 2017. Dissection of TALE-dependent gene activation reveals that they induce transcription cooperatively and in both orientations. PLoS One 12:e0173580.
- Streubel, J., Blücher, C., Landgraf, A., and Boch, J. 2012. TAL effector RVD specificities and efficiencies. Nat. Biotechnol. 30:593-595.
- Sugio, A., Yang, B., Zhu, T., and White, F. F. 2007. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIAγ1* and *OsTFX1* during bacterial blight of rice. Proc. Natl. Acad. Sci. U.S.A. 104:10720-10725.
- The 3,000 Rice Genomes Project. 2014. The 3,000 rice genomes project. GigaScience 3:7.
- Triplett, L. R., Cohen, S. P., Heffelfinger, C., Schmidt, C. L., Huerta, A. I., Tekete, C., Verdier, V., Bogdanove, A. J., and Leach, J. E. 2016. A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. Plant J. 87:472-483.
- Wang, L., Rinaldi, F. C., Singh, P., Doyle, E. L., Dubrow, Z. E., Tran, T. T., Pérez-Quintero, A. L., Szurek, B., and Bogdanove, A. J. 2017. TAL effectors drive transcription bidirectionally in plants. Mol. Plant 10:285-296.
- Wang, M., Yu, Y., Haberer, G., Marri, P. R., Fan, C., Goicoechea, J. L., Zuccolo, A., Song, X., Kudrna, D., Ammiraju, J. S. S., Cossu, R. M., Maldonado, C., Chen, J., Lee, S., Sisneros, N., de Baynast, K., Golser, W., Wissotski, M., Kim, W., Sanchez, P., Ndjondjop, M.-N., Sanni, K., Long, M., Carney, J., Panaud, O., Wicker, T., Machado, C. A., Chen, M., Mayer, K. F. X., Rounsley, S., and Wing, R. A. 2014. The genome sequence of African rice (*O. glaberrima*) and evidence for independent domestication. Nat. Genet. 46:982-988.
- White, F. F., and Yang, B. 2009. Host and pathogen factors controlling the rice-*Xanthomonas oryzae* interaction. Plant Physiol. 150:1677-1686.
- Wilkins, K. E., Booher, N. J., Wang, L., and Bogdanove, A. J. 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. Front. Plant Sci. 6:536.
- Yang, J., Zhang, Y., Yuan, P., Zhou, Y., Cai, C., Ren, Q., Wen, D., Chu, C., Qi, H., and Wei, W. 2014. Complete decoding of TAL effectors for DNA recognition. Cell Res. 24:628-631.
- Zhang, J., Yin, Z., and White, F. 2015. TAL effectors and the executor R genes. Front. Plant Sci. 6:641.

AUTHOR-RECOMMENDED INTERNET RESOURCES

- ComputerScience and Quantitative Biology Schatz Lab Cold Spring Harbor Laboratory and Johns Hopkins University:
<http://schatzlab.cshl.edu/data/rice>
- CropTAL project: <https://umr-pvbmt.cirad.fr/principaux-projets/croptal>
 daTALbase homepage: <http://bioinfo-web.mpl.ird.fr/cgi-bin2/datalbase/home.cgi>
 daTALbase for Rice: <http://bioinfo-web.mpl.ird.fr/cgi-bin2/datalbase/index.cgi>
 Ensembl genome database: <http://www.ensembl.org>
- GitHub South Green bioinformatics platform:
<https://github.com/SouthGreenPlatform/daTALbase>
- Highsoft AS Highcharts application programming interface:
<http://api.highcharts.com/highcharts>
- The International Center for Tropical Agriculture (CIAT)cassava website:
<http://ciat.cgiar.org/what-we-do/ breeding-better-crops/rooting-for-cassava>
 jBrowse: http://jbrowse.southgreen.fr/?data=oryza_sativa_japonica_v7
 jQuery user interface: <https://jqueryui.com>
 DataTables plugin for jQuery: <https://datatables.net>
 PlexDB (the plant expression database): <http://www.plexdb.org>
 The Rice Annotation Project database rap-db:
<http://rapdb.dna.affrc.go.jp/download/irgsp1.html>
 South Green bioinformatics platform: <http://www.southgreen.fr>



Development of a duplex-PCR for differential diagnosis of *Xanthomonas phaseoli* pv. *manihotis* and *Xanthomonas cassavae* in cassava (*Manihot esculenta*)[☆]



Carolina Flores^a, Carlos Zarate^a, Lindsay Triplett^b, Véronique Maillot-Lebon^c, Yassine Moufid^a, Moussa Kanté^{a,d,e}, Claude Bragard^f, Valérie Verdier^a, Lionel Gagnevin^g, Boris Szurek^a, Isabelle Robène^{c,*}

^a IRD, Cirad, Univ Montpellier, IPME, Montpellier, France

^b Department of Plant Pathology and Ecology, The Connecticut Agricultural Experiment Station, New Haven, CT, USA

^c CIRAD-UMR-PVBMT, Saint-Pierre, La Réunion, France

^d Université de Ségou, Faculté d'Agronomie et de Médecine Animale-(FAMA), Ségou, Mali

^e Laboratoire de Biologie Moléculaire Appliquée (LBMA) de l'Université des Sciences Techniques et Technologues de Bamako, Mali

^f Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

^g CIRAD-UMR IPME, Montpellier, France

ARTICLE INFO

Keywords:

Cassava

Diagnostic PCR

Xanthomonas phaseoli pv. *manihotis*

Xanthomonas cassavae

ABSTRACT

Cassava Bacterial Blight and Cassava Bacterial Necrosis are two bacterial diseases affecting cassava, respectively caused by *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) and *Xanthomonas cassavae* (*Xc*). Since both pathogens may be present on leaves and cause similar symptoms, we developed a new molecular diagnostic tool to detect and distinguish between *Xpm* and *Xc*. Based on genome sequences from the target species as well as from non-target species, *in silico* analysis was performed to select candidate primers through a novel strategy targeting specificity. Experimental validation enabled to establish a duplex-PCR diagnostic tool that will be useful for future surveillance programs and better disease control.

1. Introduction

Cassava (*Manihot esculenta*, Crantz) is a staple crop for one billion people in the tropics. It is equally important as a source of feed and industrial applications, and is also a source of energy [1]. Cassava Bacterial Blight (CBB) is the predominant and most economically significant bacterial disease of cassava [2]. Its causal agent, formerly known as *Xanthomonas axonopodis* pv. *manihotis*, was recently reassessed taxonomically [3], and the accepted nomenclature is now *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*). *Xpm* is a bacterial vascular pathogen that penetrates the host through hydathodes, stomata and/or wounds, causing angular leaf spots, leaf wilting, stem gum exudates, vascular necrosis of the stem, and shoot dieback [4]. Transmission of CBB is mainly driven by contaminated farming tools and rain splashing, while infected stakes allow long-distance dissemination. Other factors influencing *Xpm* epidemiology include its epiphytic behavior, survival on plant debris and reservoir plants [5,6].

Cassava Bacterial Necrosis (CBN), which is caused by *Xanthomonas cassavae* (*Xc*) has so far only been reported in Africa [7,8]. This contrasts with CBB, which was first reported in Latin America in the early 1900s and is now present throughout the tropics [9]. *Xc* provokes foliar symptoms somehow similar to those caused by *Xpm*, except that no vascular colonization has been observed for *Xc*. Despite many common biochemical characteristics, *Xpm* and *Xc* differ in the presence of xanthomonadin, a yellow pigment typical of xanthomonads, which is conserved in *Xc*. In contrast, the cluster of genes responsible for xanthomonadin biosynthesis is mutated in *Xpm* resulting in a white phenotype [10].

The major control strategies against CBB include the use of healthy propagative plant material, removal of infected plant material, regular surveillance, monitoring of regional outbreaks, determining highly affected areas, and predicting zones under immediate threat. These strategies rely on the ability to efficiently detect and identify the pathogen in plant material. The diversity of the pathogen has to be taken

[☆] This article is part of a Special Issue entitled 'Crop Pathology in Africa' published at the journal Physiological and Molecular Plant Pathology 105C, 2019.

* Corresponding author.

E-mail address: Isabelle.robene@cirad.fr (I. Robène).

<https://doi.org/10.1016/j.pmpp.2018.07.005>

Received 28 November 2017; Received in revised form 21 July 2018; Accepted 23 July 2018

Available online 24 July 2018

0885-5765/ © 2018 Elsevier Ltd. All rights reserved.

into account for the development of such tools. Several studies based on biochemical, serological, pathological and molecular approaches clearly distinguished *Xc* from *Xpm* strains and showed intra-species variability [11–14]. Nevertheless, for *Xpm*, genetic variability was shown to be pronounced in strains from South America, but was limited in strains from other regions such as in Africa [11,14]. Yet other studies suggest that genetic variability of *Xpm* in this continent might have been underestimated [4]. Recently, a full-genome SNP-based phylogeny of 65 geographically and temporally diverse *Xpm* genomes showed a strong clustering by geographic origin with only a few outliers [15].

Systematic diagnostic studies began in the 1990s, following different approaches from the use of selective growth medium to molecular characterization. The development of a semi-selective agar medium greatly improved the recovery of *Xpm* from plant material and soil [16]. Nevertheless, culture-based methods are not specific enough and require complementary identification methods. A monoclonal antibody was also used for an ELISA assay to detect *Xpm* from plant material, but some false negative results, as well as cross-reactions with non-target strains, were obtained [17]. Different molecular tests were also developed for the detection of *Xpm*: a dot-blot assay [18], a PCR assay [19] and a nested-PCR assay [20], all targeting a gene encoding a major *Xpm* pathogenicity factor. No molecular test has been developed until now for the detection of *Xc*.

Next generation sequencing technologies have produced a wealth of *Xanthomonas* spp. genomes, including draft sequences of *Xpm* strains [10,15] and of a single strain of *Xc* [21]. Based on these resources we searched *in silico* for appropriate targets and tested them against a collection of *Xpm* and *Xc* strains. This allowed us to develop a novel duplex-PCR diagnostic tool that successfully detects and distinguishes *Xpm* and *Xc* from symptomatic cassava plant material.

2. Materials and methods

2.1. Bacterial strains, culture conditions and DNA extraction

The bacterial strains used in this study include 50 strains of *Xpm*, 23 strains of *Xc* and 32 strains belonging to other genera, species or pathovars, including saprophytic strains isolated from cassava (Table 1). Strains were stored at -80°C in 10% glycerol for long-term storage, and routinely grown on YPGA medium (5 g/l yeast extract, 5 g/l peptone, 5 g/l glucose, 15 g/l agar) at 28°C for 48–72 h. Bacterial cultures were also grown in liquid medium Φ (1 g/l yeast extract, 1 g/l acid-hydrolyzed casein) in 5 ml tubes, prior to DNA extraction or preparation of bacterial suspensions. For this, a single colony was picked on YPGA medium with a sterile toothpick and dropped into Φ liquid medium. The tubes were incubated at 28°C for 48 h in Ecotron shaker (INFORS HT) at a speed of 160 RPM.

Genomic DNA was extracted from 1 ml liquid culture using the Wizard[®] genomic DNA purification kit (Promega) according to supplier's instructions. DNA concentration was measured with a NanoDrop ND-1000 spectrophotometer and adjusted to 1 ng/ μl . To prepare boiled bacterial suspensions, 1 ml of liquid culture was adjusted spectrophotometrically to approximately 10^8 CFU/ml ($\text{O.D}_{600} = 0.2$) using a WPA biowave CO8000 Cell Density Meter, heated at 95°C for 10 min in a water bath (F25, Jubalo) and chilled on ice.

2.2. Pathogenicity tests

The pathogenicity of *Xpm* (CIX1088, CIAT1120, CIAT1135, CIO151, CIAT1111, CI4, CIAT1206, ORST2, K419b, CFBP 2603) and *Xc* strains (UPB008, UPB029, UPB033, UPB043, UPB046, UPB047, UPB054, UPB059, UPB146) was evaluated under greenhouse conditions at 28°C and a relative humidity of 80% on 6-week-old cassava plants of the variety SG107-15. Strains grown for 48 h in Φ medium were used to prepare bacterial suspensions with sterile pure water and adjusted spectrophotometrically to $\text{O.D}_{600} = 0.2$. About 500 μl of bacterial

suspensions were infiltrated into the abaxial side of the leaf by means of a needleless syringe. For each strain, three infiltrations were performed in each lobe of the leaf. This experiment was repeated three times. The observations of symptoms were performed up to 7 days after inoculation.

2.3. Design and screening of primers

To develop diagnostic primers that differentiate between *Xpm* and *Xc*, unique sequences were identified using in-house primer pipeline named Uniqprimer version 0.5.0 (L. Triplett and J.E. Leach, in preparation) that was scripted to complement the speed and utility of the nucmer sequence aligner in the MUMmer 3.0 package [22]. The pipeline runs a series of nucmer alignments (default conditions except for the options $-\text{minmatch } 300$, $-\text{maxgap } 1$) to align whole or draft genomic sequence of a diagnostic target with genomes of non-target strains and additional target strains, and parses the output to identify coordinates of regions aligning to target genomes but not to non-target genomes. Sequences are retrieved into a multifasta file that is the template for Primer3 primer design [23]. Primer3 outputs are then aligned to the target and non-target genomes using the Primersearch program from the EMBOSS package [24] to ensure amplification specificity (cutoffs = 0% mismatch with target sequence, 20% mismatch with non-target sequence). The source code, instructions, and further validation of this pipeline will be made publicly available through two publications in preparation.

The *Xpm* primers were designed from the complete draft sequence of the template genome UG23, and aligned against the genome sequences of *X. cassavae* strain CFBP4642, *Xanthomonas oryzae* PXO99A, *Xanthomonas vasicola* NCPPB206, *Xanthomonas citri* pv. *citri* Aw12879, *Xanthomonas campestris* pv. *campestris* ATCC33913, *Xanthomonas phaseoli* pv. *fuscans* 4834-R and *Xanthomonas axonopodis* pv. *citrumelo* F1 (NCBI accession numbers ATMC00000000.1, NC_010717, AKBM00000000, CP003778.1, NC_003902, FO681494.1, and CP002914, respectively). The draft genome of *Xc* strain CFBP4642 was used as a template for the selection of *Xc* primers. The specified non-target genomes were those of *Xpm* strain UG23 (NCBI accession AKEV00000000), and reference strains of *X. oryzae* pv. *oryzicola*, *X. campestris* pv. *campestris* and *X. vasicola* named above. Six primer pairs for *Xpm* and eight for *Xc* were selected (Table 2).

2.4. In silico screening

Blastn searches performed on the NCBI nucleotide collection (nr/nt) database with the primers and target sequences showed no significant sequence similarities with non-*Xanthomonas* bacteria (data not shown). The primers and target sequences were then used in BLASTn queries against the 901 genome sequences of different strains belonging to the *Xanthomonadales* order that were available in the NCBI database (16/08/17), including 66 *Xpm* and one *Xc* genome sequences (Microbial genome Blast, Databases = All genomes/Draft genomes, Organism: *Xanthomonadales* (taxid:135614), Program selection: blastn).

Some types of primer-template mismatches are known to have no significant effect on PCR efficiency. We relied on previous studies [25,26] to predict PCR amplifications based on different criteria, mainly the nature, the number and the location of the mismatches. We only took into account template/primer duplexes including up to two mismatches internal or located at primer 5' terminal base. No mismatch located in the primer 3' terminal base was tolerated and for the 3 bases before the 3' terminal base, we did not consider the mismatches known to have a significant effect on PCR yield (A:G, G:A, and C:C and A:A).

2.5. Experimental screening

The primer pairs were tested using boiled culture suspensions of a subset of 10 *Xpm* (CIX1088, CIAT1120, CIAT1135, CIO151, CIAT1111,

Table 1*Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), *Xanthomonas cassavae* (*Xc*) and other bacterial strains used in this study.

Strain ID ^a	Other known ID	Species	Origin	Date of isolation	Host	Pathogenicity ^b	References
CIAT1047		<i>Xpm</i>	Africa	1970	<i>Manihot esculenta</i>	+	[19]
CIX1179	T121	<i>Xpm</i>	Africa	NA ^d	<i>M. esculenta</i>	+	Verdier, unpublished
INTA1, INTA5		<i>Xpm</i>	Argentina	NA	<i>M. esculenta</i>	+	[20]
CIX1088^e	INTA4	<i>Xpm</i>	Argentina	NA	<i>M. esculenta</i>	+	[20]
AFNC1360 ^f		<i>Xpm</i>	Benin or Nigeria	NA	<i>M. esculenta</i>	+	[16]
CIAT1120	Orst3, CFBP1852	<i>Xpm</i>	Brazil	1974	<i>M. esculenta</i>	+	[12]
CIAT1135		<i>Xpm</i>	Taiwan	1975	<i>M. esculenta</i>	+	[14]
CIO151^e		<i>Xpm</i>	Colombia	1995	<i>M. esculenta</i>	+	[14]
CIAT1180		<i>Xpm</i>	Venezuela	1977	<i>M. esculenta</i>	+	[14]
CIAT1202		<i>Xpm</i>	Colombia	1981	<i>M. esculenta</i>	+	[14]
CIAT1111^e	CFBP1851	<i>Xpm</i>	Colombia	1974	<i>M. esculenta</i>	+	[14]
CI4		<i>Xpm</i>	Ivory Coast	2015	<i>M. esculenta</i>	+	[31]
CIX2481	Xam5M	<i>Xpm</i>	Mali	2016	<i>M. esculenta</i>	+	Kanté & Szurek, unpublished
CIAT1211, CIAT1205		<i>Xpm</i>	New Zealand	NA	<i>M. esculenta</i>	+	Verdier, unpublished
CIAT1206		<i>Xpm</i>	New Zealand	NA	<i>M. esculenta</i>	+	Verdier, unpublished
CIAT1241		<i>Xpm</i>	New Zealand	NA	<i>M. esculenta</i>	+	[20]
CIX1111	T139	<i>Xpm</i>	Togo	NA	<i>M. esculenta</i>	+	Verdier, unpublished
ORST2	CIAT1061, CFBP1850	<i>Xpm</i>	Venezuela	1971	<i>M. esculenta</i>	+	[13]
CIO238, CIO240		<i>Xpm</i>	Venezuela	1998	<i>M. esculenta</i>	+	[25]
008Xam		<i>Xpm</i>	Venezuela	1998	<i>M. esculenta</i>	+	Hernandez, unpublished
K19a, K419b , K19c, K19d, K19e		<i>Xpm</i>	Vietnam	2016	<i>M. esculenta</i>	+	Tran & Szurek, unpublished
CIAT1069		<i>Xpm</i>	NA	NA	<i>M. esculenta</i>	+	Verdier, unpublished
CIX1102		<i>Xpm</i>	NA	NA	<i>M. esculenta</i>	+	Verdier, unpublished
CIX1098		<i>Xpm</i>	NA	NA	<i>M. esculenta</i>	+	Verdier, unpublished
UPB003	LMG766	<i>Xpm</i>	DR Congo	1973	<i>M. esculenta</i>	+	[12]
UPB006	LMG767	<i>Xpm</i>	DR Congo	1973	<i>M. esculenta</i>	+	[12]
UPB009	NCPPB3058, LMG768	<i>Xpm</i>	DR Congo	1973	<i>M. esculenta</i>	+	[12]
UPB010	NCPPB3059, LMG769	<i>Xpm</i>	Zaire	1973	<i>M. esculenta</i>	+	[12]
UPB025	NCPPB3060, LMG771	<i>Xpm</i>	Nigeria	1976	<i>M. esculenta</i>	+	[12]
UPB026		<i>Xpm</i>	Nigeria	NA	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB027	C631, LMG629	<i>Xpm</i>	Cameroun	1977	<i>M. esculenta</i>	+	[12]
UPB034	671	<i>Xpm</i>	Cameroun	1977	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB058	810	<i>Xpm</i>	Nigeria	NA	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB071	NCPPB1161, LMG775	<i>Xpm</i>	Mauritius	1946	<i>M. esculenta</i>	+	[12]
UPB090	914	<i>Xpm</i>	Uganda	1979	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB091	915a	<i>Xpm</i>	Uganda	1979	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB092	917	<i>Xpm</i>	Uganda	1979	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB093	919	<i>Xpm</i>	Uganda	1979	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB094	920	<i>Xpm</i>	Uganda	1979	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB174	VQM1883	<i>Xpm</i>	Java	1981	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB176	VQM1891	<i>Xpm</i>	Sumatra	1981	<i>M. esculenta</i>	NA	Verdier, unpublished
UPB178	VQM1923	<i>Xpm</i>	Sumatra	NA	<i>M. esculenta</i>	NA	Verdier, unpublished
CFBP 2603	LMG776, NCPPB2443	<i>Xpm</i>	Colombia	1970	<i>M. esculenta</i>	+	[12]
UPB008	2283	<i>Xc</i>	DR Congo	1978	<i>M. esculenta</i>	+	[13]
UPB029	662	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	Verdier, unpublished
UPB030	662a	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB031	662 b	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[17]
UPB032	668a	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB033	668 d	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB035	689 b	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB038	LMG672, NCPPB3062	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB039	704c	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB041	LMG5265	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB043	723	<i>Xc</i>	Rwanda		<i>M. esculenta</i>	+	[13]
UPB044	LMG5267	<i>Xc</i>	Rwanda	1977	<i>M. esculenta</i>	+	[13]
UPB045	LMG5268	<i>Xc</i>	Rwanda	1978	<i>M. esculenta</i>	+	[13]
UPB046	732 b	<i>Xc</i>	Rwanda	1978	<i>M. esculenta</i>	+	[13]
UPB047	733a	<i>Xc</i>	Rwanda	1978	<i>M. esculenta</i>	+	[13]
UPB049	735a	<i>Xc</i>	Rwanda	1978	<i>M. esculenta</i>	+	[13]
UPB053	LMG5270	<i>Xc</i>	Rwanda	1978	<i>M. esculenta</i>	+	[13]
UPB054	LMG673, CIAT1148	<i>Xc</i>	Malawi	1953	<i>M. esculenta</i>	+	[13]

(continued on next page)

Table 1 (continued)

Strain ID ^a	Other known ID	Species	Origin	Date of isolation	Host	Pathogenicity ^b	References
UPB059	LMG764	<i>Xc</i>	Tanzanie	1978	<i>M. esculenta</i>	+	[13]
UPB146	LMG5271	<i>Xc</i>	Kenya	1979	<i>M. esculenta</i>	+	[13]
CIX157	LMG 5264	<i>Xc</i>	Malawi	1951	<i>M. esculenta</i>	+	[20]
CIX739	CFBP4642, LMG673	<i>Xc</i>	Malawi	1951	<i>M. esculenta</i>	+	[13]
CIX196	CFBP4642	<i>Xc</i>	Malawi	1951	<i>M. esculenta</i>	+	[22]
CPX05		Saprophytic strain	Venezuela	2015	<i>M. esculenta</i>	-	Flores & Szurek, unpublished
LN149, LN154, LN160, LN161, LN163, LN164, LN168, LN170		Saprophytic strains	Réunion, France	2016	<i>M. esculenta</i>	-	Flores & Robène, unpublished
LMG 876		<i>Xanthomonas translucens</i> pv. <i>translucens</i>	USA	1933	<i>Hordeum vulgare</i>	NA	
CFBP2053		<i>X. translucens</i> pv. <i>translucens</i>	Switzerland	1973	<i>Dactylis glomerata</i>	NA	
LMG695		<i>Xanthomonas phaseoli</i> pv. <i>dieffenbachiae</i>	Brazil	1965	<i>Anthurium</i> sp.	NA	
CIX227		<i>X. phaseoli</i> pv. <i>dieffenbachiae</i>	NA	NA	NA	NA	
LMG9055		<i>X. phaseoli</i> pv. <i>syngonii</i>	USA	1984	<i>Syngonium</i> <i>podophyllum</i>	NA	
LMG8014		<i>X. phaseoli</i> pv. <i>phaseoli</i>	Romania	1966	<i>Phaseolus vulgaris</i>	NA	
CIX208		<i>X. phaseoli</i>	NA	NA	NA	NA	
CIX228		<i>X. phaseoli</i> pv. <i>phaseoli</i>	NA	NA	NA	NA	
LMG539		<i>Xanthomonas axonopodis</i> pv. <i>axonopodis</i>	Colombia	1949	<i>Axonopus scoparius</i>	NA	
CIX230		<i>Xanthomonas citri</i> pv. <i>vignicola</i>	NA	NA	NA	NA	
LMG7429		<i>X. citri</i> pv. <i>malvacearum</i>	Uganda	1965	<i>Gossypium</i> sp.	NA	
CFBP2525		<i>X. citri</i> pv. <i>citri</i>	New Zeland	1956	<i>Citrus limon</i>	NA	
CFBP2868		<i>Xanthomonas fuscans</i> pv. <i>aurantifolii</i> (type B)	Argentina	NA	<i>Citrus limon</i>	NA	
CFBP2903		<i>X. fuscans</i> pv. <i>aurantifolii</i> (type B)	Argentina	NA	<i>Citrus limon</i>	NA	
CFBP2905		<i>X. fuscans</i> pv. <i>aurantifolii</i> (type C)	Argentina	NA	<i>Citrus limon</i>	NA	
LMG901		<i>X. axonopodis</i> pv. <i>vasculorum</i>	Mauritius	1960	<i>Saccharum officinarum</i>	NA	
CIX202		<i>Xanthomonas begoniae</i>	NA	NA	NA	NA	
CFBP2532		<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	India	1965	<i>Oryza sativa</i>	NA	
CFBP4641		<i>Xanthomonas sacchari</i>	Guadeloupe, France	NA	<i>Saccharum officinarum</i>	NA	
LMG2804		<i>Pectobacterium</i> <i>chrysanthemii</i>	USA	NA	<i>Chrysanthemum</i> × <i>morifolium</i>	NA	
Run 215		<i>Ralstonia solanacearum</i> (phylotype I)	Cameroon	NA	<i>Gaylussacia</i> sp.	NA	
Run 17		<i>Ralstonia solanacearum</i> (phylotype II)	Martinique, France	NA	<i>Heliconia</i> sp.	NA	
JX63B		<i>Pseudomonas fluorescens</i>	Réunion, France	2001	<i>Anthurium</i> sp.	NA	

^a Prefixes of strain identifiers indicate the origin of the strain: CIAT, Xanthomonas collection of the Centro Internacional de Agricultura, Cali, Colombia; INTA, Collection from the Instituto Nacional de Tecnología Agropecuaria, Bella Vista, Argentina; CIO, ORSTOM collection; UPB, Collection from Unité de Phytopathologie (ELI-UCL), Louvain-La-Neuve, Belgium; CFBP, Collection Française de Bactéries associées aux Plantes, Angers, France; LMG, Laboratorium voor Microbiologie Gent culture collection, Universiteit Gent, Belgium; Run, JX and LN, Bacterial strains Collection, Cirad, La Réunion, France; CIX, CPX and other strains, IRD bacterial collections, Montpellier, France.

^b Bold font refers to results from pathogenicity tests performed in this study on cassava variety SG107-15, normal font refers to results from unpublished data or cited references. +, pathogenic strain; -, non pathogenic strain.

^c Strains marked in bold were used for the screening of the different primers sets.

^d Not available.

^e Strains which genome sequence is available {Bart, 2012#2840;Arrieta-Ortiz, 2013#1744}.

CI4, CIAT1206, ORST2, K419b, CFBP 2603) and 9 *Xc* (UPB008, UPB029, UPB033, UPB043, UPB046, UPB047, UPB054, UPB059, UPB146) strains (Table 1). The selected strains were temporally and especially geographically diverse in order to cover as much as possible the genetic diversity of the bacteria, as shown by Bart et al. for *Xpm* [15].

We also tested a set of eight non-target strains including strains that were positive in the *in silico* evaluation (LMG 876, CFBP2053, LMG695,

LMG9055, LMG8014, CFBP2868, CFBP2903, CFBP2905) (Tables 1 and 4).

2.6. PCR assays

Simplex PCR reactions were performed in a final volume of 10 µl containing 0.5 µM of each forward and reverse primers, 2 mM dNTPs, 2 mM MgCl₂, 1,25 U GoTaq[®] DNA polymerase (Promega), buffer 1X and

Table 2
Primer pairs used in this study and *In silico* analysis of their specificity.

Targeted species	Primer name	Sequence (5' → 3')	Amplicon size (bp)	Locus tag ^a (position) and function of the target sequence	Number of matching target genomes	Non-target species detected by <i>in silico</i> PCR (No. of genomes ^b)	<i>in silico</i> PCR inclusivity (%)	<i>in silico</i> PCR exclusivity (%)
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam1F Xam1R	TATAGGGCAGCGCTACGAGT ATCGGCTATGCTGAACCACT	203	AKEV01000028 (1349..1551) lipase	66 ^c /66	<i>Xanthomonas fuscans</i> pv. <i>aurantifolli</i> (1)	100	99.99
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam2F Xam2R	AAGGCAAACCACTCGTGTTT ATGCTGGGAATCAACCTCAG	201	AKEV01000044 (13822..14022) hypothetical protein	66/66	<i>X. phaseoli</i> pv. <i>syngonii</i> (2), <i>X. phaseoli</i> pv. <i>dieffenbachiae</i> (2), <i>X. phaseoli</i> pv. <i>phaseoli</i> (9)	100	98.56
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam3F Xam3R	ACATGGGACCAAACGAAAAG GCTTTAGGATGCCTGAGCAC	204	AKEV01000080 (20701..20904) helicase	66/66	<i>X. phaseoli</i> pv. <i>syngonii</i> (2), <i>X. phaseoli</i> pv. <i>dieffenbachiae</i> (2), <i>X. phaseoli</i> pv. <i>phaseoli</i> (9),	100	98.56
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam4F Xam4R	GCTATGTCGACTGGCTCGAT GACCGTTGACCAAAGGTAA	200	AKEV01000001 (6372..6571) hypothetical protein	66/66	<i>X. phaseoli</i> pv. <i>syngonii</i> (2), <i>X. translucens</i> pv. <i>graminis</i> (8) <i>X. translucens</i> pv. <i>poae</i> (2), <i>X. translucens</i> pv. <i>translucens</i> (1), <i>X. campestris</i> pv. <i>campestris</i> (1), <i>X. arboricola</i> pv. <i>corylina</i> (1)	100	98.34
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam5F Xam5R	GAATTTGCCGCGTTATAGA TCCAGCGCTAATACCTGGAC	205	AKEV01000001 (5829..6033) hypothetical protein	66/66	<i>X. phaseoli</i> pv. <i>syngonii</i> (2), <i>X. translucens</i> pv. <i>graminis</i> (8) <i>X. translucens</i> pv. <i>poae</i> (2), <i>X. translucens</i> pv. <i>translucens</i> (1), <i>X. campestris</i> pv. <i>campestris</i> (1)	100	98.45
<i>X. phaseoli</i> pv. <i>manihotis</i>	Xam6F Xam6R	ACAGCTACGACCTGGACCAC GTCCGTGATATCGGGGTAGA	203	AKEV01000020 (583..785) putative fic-family protein	2/66	–	3	100
<i>Xanthomonas cassavae</i>	XC1F XC1R	GACCACAAAGGTGGTCTCGT CAGGCGGTGATACTGACGTA	318	ATMC01000058 (124346..124663) hypothetical protein	1/1	–	NA ^d	100
<i>Xanthomonas cassavae</i>	XC2F XC2R	GCAGGTGGTGCTCAGTGTC AGGGAATCATGCAACGAAGA	314	ATMC01000058 (132021..132334) hypothetical protein	1/1	–	NA	100
<i>Xanthomonas cassavae</i>	XC3F XC3R	AAGTACACCAGCGTGTCTGC ACGAACAACCTGCTGCACAAC	310	ATMC01000001 (17386..17695) methyltransferase	1/1	<i>Xanthomonas</i> sp. (1), <i>Xanthomonas pisi</i> (1)	NA	99.78
<i>Xanthomonas cassavae</i>	XC4F XC4R	TATTGCCCAAGTTGTTGCAG ACTGCAGTCGCTCAGTCGT	300	ATMC01000001 (18194..18493) methyltransferase	1/1	<i>Xanthomonas</i> sp. (1)	NA	99.99
<i>Xanthomonas cassavae</i>	XC5F XC5R	TTACCTTTGCCGAGACCAAC ATCGACGACTTCCTCAGTGC	304	ATMC01000010 (2049..2352) and ATMC01000002 (2049..2352)	1/1	–	NA	100
<i>Xanthomonas cassavae</i>	XC6F XC6R	CATCTGCGATCACCATAGGC GTGGATCTCCAGGACTGAA	304	ATMC01000058 (148699..149002)	1/1	–	NA	100
<i>Xanthomonas cassavae</i>	XC7F XC7R	AGTGGCTGTGAAAAGGGTA TGCATTAGGATGGTGTCCAA	306	ATMC01000058 153960..154265	1/1	–	NA	100
<i>Xanthomonas cassavae</i>	XC8F XC8R	TTGGCTTGCTGTTGTAICTC CAAGGCCTGATTCTGACAC	302	ATMC01000003 7065..7364 Hypothetical protein	1/1	–	NA	100

^a Locus tags and positions are given for the strains UG23 (*X. phaseoli* pv. *manihotis*) and CFBP 4642 (*X. cassavae*).

^b Among 901 *Xanthomonadales* genomes (taxid:135614).

^c Strain UG28 was not detected using "Microbial nBlast" on the database "draft genomes" but the target sequence was found in the raw data in the non aligned reads, using "nBlast" on the database SRA "SRR522435".

^d NA = not applicable.

Table 3
Field Cassava samples and duplex- PCR responses.

	Organ	Origine	Date	Symptoms	diagnostic ^b	Duplex-PCR responses											
						on serially diluted lesion macerates ^a										on DNA extracts	
						Xpm1					Xc1					Xpm1	Xc1
						10 ¹	10 ²	10 ³	10 ⁴	10 ⁵	10 ¹	10 ²	10 ³	10 ⁴	10 ⁵		
1	Stems	Taiwan	1978	Brown spots	<i>Xpm</i> ^c (HMB60 = LMG774)	-	-	-	+	+	-	-	-	-	-	NA ^d	NA
2	Leaves	Kenya	1979	angular watery spots	<i>Xpm</i> (HMB148 = LMG783)	-	+	++	+	-	-	-	-	-	-	+	-
3	Leaves	Nigeria-Cameroon	1977	angular watery spots	<i>Xpm</i> (HMB27)	-	-	++	+	-	-	-	-	-	-	-	-
4	Leaves	Nigeria	1977	angular watery spots	<i>Xpm</i> (HMB36 = LMG772)	-	-	++	+	+	-	-	-	-	-	NA	NA
5	Leaves	Tanzania	1979	angular watery spots	<i>Xc</i> (HMB59 = UPB059 = LMG674)	-	-	-	-	-	-	-	-	-	-	-	+
6	Leaves	Rwanda	1977	angular watery spots	<i>Xc</i> (HMB44 = LMG 5267)	-	-	-	-	-	-	-	-	-	-	-	+
7	Stems	Rwanda	1977	Brown spots	<i>Xc</i> (HMB29 = LMG 671)	-	-	-	-	-	-	-	-	-	-	NA	NA
8	Leaves	Brazil	1976	angular watery spots	<i>Xpm</i> (UPB080)	-	-	-	+	+	-	-	-	-	-	+	-
9	Leaves	Cameroon	1977	angular watery spots	<i>Xpm</i> (UPB034)	-	+	++	+	+	-	-	-	-	-	NA	NA
10	Leaves	Nigeria	1977	angular watery spots	<i>Xpm</i> (UPB058)	-	-	++	+	-	-	-	-	-	-	NA	NA
11	Leaves	Ivory Coast	1981	angular watery spots	<i>Xpm</i> (HMB203 = LMG5249)	-	+	++	+	-	-	-	-	-	-	NA	NA
12	Leaves	Uganda	1980	angular watery spots	<i>Xpm</i> (UPB090, UPB091, UPB092, UPB093, UPB094)	-	-	++	+	-	-	-	-	-	-	NA	NA
13	Leaves	Uganda	1979	angular watery spots	<i>Xpm</i> (HMB81 = LMG780)	-	-	++	+	-	-	-	-	-	-	NA	NA
14	Leaves	Uganda	1979	angular watery spots	<i>Xpm</i> (HMB93 = LMG782)	-	-	-	++	-	-	-	-	-	-	NA	NA

^a Dilutions from 10¹ to 10⁵-fold.

^b The diagnosis relied on both symptomatology and characterization of bacterial colonies isolated from these samples: morphological, physiological, biochemical features and pathogenicity tests, and for some strains, protein gel electrophoregrams and DNA:DNA hybridizations (Van der Mooter et al., 1987). The characters within brackets refers to the collection numbers of the isolated strains.

^c *Xpm*: *Xanthomonas phaseoli* pv. *manihotis* and *Xc*: *Xanthomonas cassavae*.

^d Not available.

Table 4
PCR screening of the unique *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) and *X.cassavae* (*Xc*) primer candidates.

Strains	Species	Origin	Xpm (F/R)					Xc (F/R)									
			1	2	3	4	5	6	1	2	3	4	5	6	7	8	
CIX1088	<i>X. phaseoli</i> pv. <i>manihotis</i>	Argentina	+ ^a	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CIAT1120	<i>X. phaseoli</i> pv. <i>manihotis</i>	Brazil	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CIAT1135	<i>X. phaseoli</i> pv. <i>manihotis</i>	China	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CIO151	<i>X. phaseoli</i> pv. <i>manihotis</i>	Colombia	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CIAT1111	<i>X. phaseoli</i> pv. <i>manihotis</i>	Colombia	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CI4	<i>X. phaseoli</i> pv. <i>manihotis</i>	Ivory Coast	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
CIAT1206	<i>X. phaseoli</i> pv. <i>manihotis</i>	New Zealand	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
ORST2	<i>X. phaseoli</i> pv. <i>manihotis</i>	Venezuela	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
K419b	<i>X. phaseoli</i> pv. <i>manihotis</i>	Vietnam	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
CFBP2603	<i>X. phaseoli</i> pv. <i>manihotis</i>	Colombia	+	+	+	+	+	-	-	-	x	-	-	-	-	-	-
UPB008	<i>X.cassavae</i>	DR Congo	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+
UPB029	<i>X.cassavae</i>	Rwanda	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
UPB033	<i>X.cassavae</i>	Rwanda	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
UPB043	<i>X.cassavae</i>	Rwanda	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
UPB046	<i>X.cassavae</i>	Rwanda	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
UPB047	<i>X.cassavae</i>	Rwanda	-	-	-	-	-	-	+	+	-	+	+	+	+	+	-
UPB054	<i>X.cassavae</i>	Malawi	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+
UPB059	<i>X.cassavae</i>	Tanzania	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
UPB146	<i>X.cassavae</i>	Kenya	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-
LMG 695	<i>X. phaseoli</i> pv. <i>dieffenbachiae</i>	Brazil	-	+	+	-	-	x	-	-	x	-	-	-	-	-	-
LMG 9055	<i>X. phaseoli</i> pv. <i>syngonii</i>	USA	-	+	+	+	+	-	-	-	x	-	-	-	-	-	-
LMG 8014	<i>X. phaseoli</i> pv. <i>phaseoli</i>	Romania	-	+	+	-	-	x	-	-	x	-	-	-	-	-	-
CFBP 2868	<i>X. fuscans</i> pv. <i>aurantifolii</i> (type B)	Argentina	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CFBP 2903	<i>X. fuscans</i> pv. <i>aurantifolii</i> (type B)	Argentina	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CFBP 2905	<i>X. fuscans</i> pv. <i>aurantifolii</i> (type C)	Argentina	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LMG 876	<i>X. translucens</i> pv. <i>translucens</i>	USA	-	-	-	-	-	x	-	-	-	-	-	-	-	-	-
CFBP 2053	<i>X. translucens</i> pv. <i>graminis</i>	Switzerland	-	-	-	-	-	x	-	-	-	-	-	-	-	-	-

^a +: amplification, -: no amplification, x: amplification but not the expected size for the amplicons.

1 µl of template DNA.

For the duplex-PCR, preliminary results showed the *Xpm* target to be preferentially amplified. Optimization experiments were therefore conducted in order to homogeneously amplify both targets. For this, various concentrations of *Xpm*1 F/R and *Xc*1 F/R primers were tested: 0.5 µM *Xpm*1/0.5 µM *Xc*1, 0.25 µM *Xpm*1/0.5 µM *Xc*1 and 0.125 µM *Xpm*1/0.5 µM *Xc*1 in the same final mix as for the simplex PCR. The duplex-PCR assay was tested on gDNA (1 ng/µl) or boiled suspensions of *Xpm* strain CIO151 and *Xc* strain CIX739 mixed at ratios 1:1 and 1:2. This experiment was repeated thrice. PCR amplifications were carried out in a thermal cycler (Applied Biosystems 2720). After 2 min of denaturation at 95 °C, 35 cycles of 45 s of denaturation at 95 °C, 45 s of annealing at 61 °C and 1.5 min of extension at 72 °C were performed, with a final extension at 72 °C for 5 min. Aliquots containing the PCR amplification products were subjected to electrophoresis in 1.5% agarose gels at 50 V for 1 h and visualized by ultraviolet light after staining with ethidium bromide.

2.7. Evaluation of specificity and sensitivity of the different PCR assays

To determine the specificity, we evaluated both the criteria of inclusivity, i.e. the ability of the different primers to detect all strains of the target organism, and of exclusivity, i.e. no amplicon acquisition of the expected size from non-target strains.

The specificity of the simplex PCRs directed by the primer pairs *Xpm*1 F/R and *Xc*1 F/R and of the duplex-PCR (combining the two sets of primers) was evaluated on the collection of *Xpm*, *Xc* and non-target strains (Table 1), using 1 µl of genomic DNA (1 ng/µl) and/or boiled liquid cultures as templates. The primer pairs were also tested on an historical DNA collection including 39 *Xpm* and *Xc* strains (20 ng/µl) collected in the 1970s [27] (Supplemental Table S1).

The sensitivity of the duplex-PCR was assayed for both bacterial pure cultures and plant extracts. *Xpm* strain CIO151 and *Xc* strain

CIX739 were grown overnight in Φ liquid medium, were suspended at $O. D_{600} = 0.2$ and serially diluted in Tris buffer pH 7.2. Negative control received only Tris buffer. The different suspensions were boiled as described and tested with the duplex-PCR assay. Before boiling, bacterial cells concentration was checked by plating 100 µl of the 10^4 , 10^5 and 10^6 -fold dilutions on YPGA (10 replicates per dilution). For determination of the sensitivity in a plant matrix, the dilution series were added to plant homogenates. Cassava leaves were ground using mortar and pestle in Tris buffer pH 7.2, 2% polyvinylpyrrolidone (PVP) (20 ml buffer per 0.2 g of leaf material), and total DNA was extracted from 2 ml samples using the DNeasy[®] Plant Mini Kit (Qiagen), according to the supplier's instructions and was tested using the duplex-PCR assay. This experiment was performed three time independently (including the preparation of the initial bacterial suspensions).

2.8. Detection of *Xpm* and *Xc* from plant samples

To determine the ability of the primers to detect the presence of *Xpm* and *Xc* in infected plant tissues, cassava leaves from the variety SG107-15 were artificially inoculated with *Xpm* strain CIO151 or *Xc* strain CIX739 as described above. Leaves were collected 7 dpi and angular lesions were cut from the leaves and macerated for 2 min in 0.5 ml of sterile distilled water [19]. Ten-fold dilution series were performed in distilled water down to 10^8 -fold, and 1 µl of each dilution was used as a template in simplex and duplex-PCR assays. Total DNA was also extracted from CBB and CBN symptoms obtained after inoculation of different *Xpm* (CIO151 and CIX1088) and *Xc* (UPB059, UPB146, UPB029, UPB043, UPB046) strains using the DNeasy[®] Plant Mini Kit (Qiagen) as described above, and 1 µl of DNA extract was used as a template in duplex-PCR assays.

The presence of *Xpm* and *Xc* was also tested from diseased leaf samples collected in the 1970s and 80s in different countries, all maintained at the Earth and Life Institute (Université catholique de

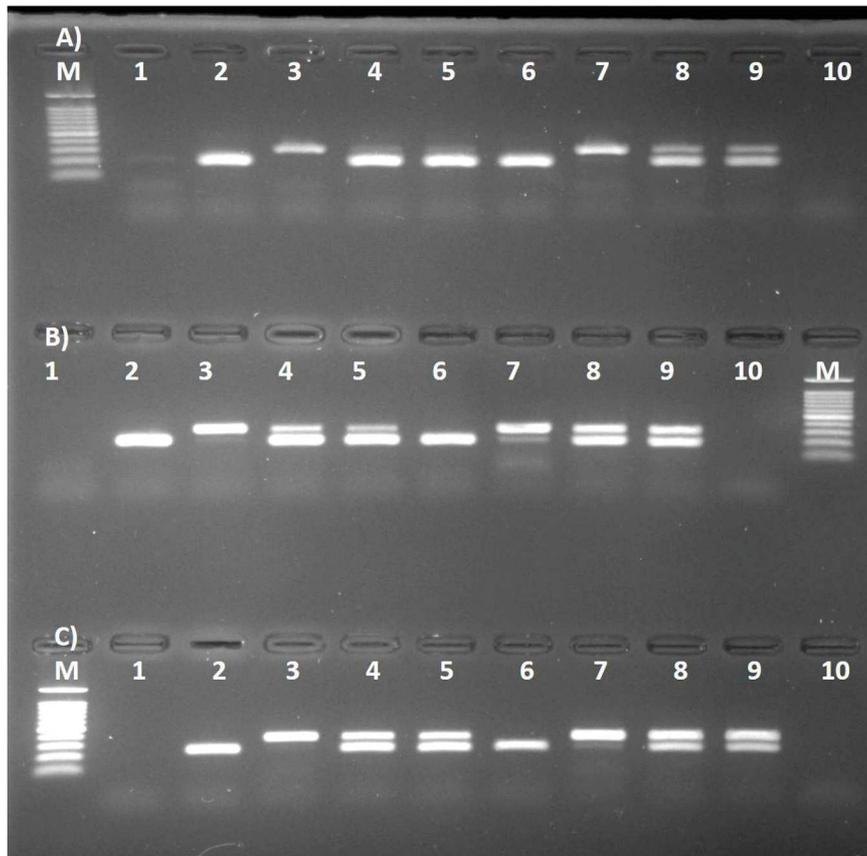


Fig. 1. Optimization of the duplex-PCR assay.

The duplex-PCR was performed using various concentrations of Xpm1 F/R and Xc1 F/R primers. A, B and C: agarose gel electrophoresis of PCR products. A, Xpm1 0.5 μM, Xc1 0.5 μM. B, Xpm1 0.25 μM, Xc1 0.5 μM. C, Xpm1 0.125 μM, Xc1 0.5 μM. Lane 1, water. Lanes 2 and 3, gDNA (1 ng/μl) of *Xanthomonas cassavae* (Xc) strain CIX739 and *Xanthomonas phaseoli* pv. *manihotis* (Xpm) strain CIO151, respectively. Lanes 4 and 5, mix of gDNA (1 ng/μl) of CIX739 and CIO151 at ratios of 1:1 and 2:1, respectively. Lanes 6 and 7, boiled suspensions (OD = 0.2) of CIX739 and CIO151, respectively. Lanes 8 and 9, mix of boiled suspensions (OD = 0.2) of CIX739 and CIO151 at ratios of 1:1 and 2:1, respectively. M: Size marker 100 bp (Promega).

Louvain, Belgium), at room temperature (20–25 °C), under dry conditions (Table 3). At the time of sample collection, bacterial colonies had been isolated from these samples and characterized by morphological, physiological, biochemical features and pathogenicity tests. Some isolates were also characterized by protein gel electrophoregrams and DNA:DNA hybridizations [28]. Typical symptomatic lesions were sampled and serially diluted down to 10⁵-fold as described above, or total DNA was extracted from lesions using the DNeasy® Plant Mini Kit (Qiagen). Duplex-PCR was performed using 1 μl of each dilution or 1 μl DNA extract.

3. Results

3.1. Search for unique primers differentiating Xpm and Xc

Comparative genomic analysis of target and non-target genomes using the Uniqprimer pipeline 0.5.0 successfully generated candidate primer pairs which are *a priori* specific to each of the two pathogens (Table 2). The screening of these primer pairs was performed based on both *in silico* and experimental tests. The pathogenicity of the subset of Xpm and Xc strains (n = 19) used for PCR screening was previously confirmed by inoculation assays (Fig. S1).

In silico analyses showed 100% of inclusivity for all Xpm primer pairs that were selected except for Xpm6. Indeed, while primer pairs Xpm1 to 5 matched with the 66 Xpm genome sequences available at NCBI (Table 2), only two genomes were putatively detected by primer pair Xpm6. When tested experimentally, pairs of primers Xpm1 to 5 all resulted in amplification of products of the expected sizes for all 10 Xpm target strains that originate from diverse geographic areas. In contrast, only one Xpm strain among 10 was tested positive using the pair Xpm6 (Table 4).

In silico analyses showed that Xpm1 primers displayed the best exclusivity value, with only one strain of *X. fuscans* pv. *aurantifolii* being

potentially detected among 901 Xanthomonadales genomes. This was not the case for other primer pairs which resulted in amplification from other xanthomonads, including strains of the phylogenetically-close pathovars *syngonii*, *dieffenbachiae* and *phaseoli* of the species *X. phaseoli* (Table 2).

When tested experimentally on a set of non-target strains, Xpm1 primers showed the expected exclusivity with only strains of *X. fuscans* pv. *aurantifolii* being amplified (n = 2). As predicted by *in silico* analyses, the other primer pairs generated amplicons of the expected sizes from non-target strains DNA, and particularly with *X. phaseoli* pathovars for Xpm2 and Xpm3 primers. Importantly, Xpm pairs of primers led to no amplification when tested on Xc strains (Table 4).

In silico analyses were less informative for Xc primers, because only one Xc genome sequence was available at the time of this study and formed the basis of all Xc primers. Nevertheless, these results showed high exclusivity value for most of Xc primers with only a few non-target strains potentially detected (for Xc3 and Xc4 primer pairs) among the 901 Xanthomonadales genomes (Table 2).

Most primer pairs generated an amplicon of the expected size when tested on the nine Xc target strains of our collection (Table 4). Exceptions are pairs of primers Xc3 and Xc8 for which no amplification was obtained in 1 and 7 samples, respectively. The experimental results obtained on the set of non-target strains reflected partially *in silico* predictions, with high specificity of the different primer pairs, but yielded an unexpected nonspecific amplification (600 bp amplicon) for *X. phaseoli* pathovars and Xpm strain CFBP2603 when using primer pair Xc3. Importantly, no amplification of the expected size was observed when Xpm strains were used as templates using any of the tested Xc pairs of primers.

In conclusion, the Xpm1 and Xc1 primer pairs were chosen as candidates for the diagnosis of Xpm and Xc. Also, the difference in amplicon sizes obtained for Xpm1 and Xc1 (203 bp and 314 bp, respectively) was compatible with the development of a duplex-PCR

Table 5

PCR results obtained for Xpm1 and Xc1 primer pairs on the collection of *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*), *Xanthomonas cassavae* (*Xc*) and other bacterial strains.

<i>Xanthomonas phaseoli</i> pv. <i>manihotis</i>	Simplex-PCR			
	Xpm1 Boiled samples	gDNA 1 ng/ μl	Xc1 Boiled samples	DNA 1 ng/μl
CIAT1047, CIX1179, INTA1, INTA5, CIX1088, AFCN1360, CIAT1120, CIAT1135, CIO151, CIAT1180, CIAT1202, CIAT1111, CI4, CIX2481, CIAT1211, CIAT1206, CIAT1205, CIAT1241, CIX1111, ORST2, CIO238, CIO240, 008Xam, K419a, K419b, K419c, K419d, K419e, CIAT1069, CIX1102, CIX1098, CFBP 2603	+	+	–	–
	Duplex-PCR			
	Xpm1 Boiled samples	Xc1		
UPB003, UPB006, UPB009, UPB010, UPB025, UPB026, UPB027, UPB034, UPB058, UPB071, UPB090, UPB091, UPB092, UPB093, UPB094, UPB174, UPB176, UPB178	+	–		
<i>Xanthomonas cassavae</i>	Simplex-PCR			
	Xpm1 Boiled samples	DNA 1 ng/μl	Xc1 Boiled samples	DNA 1 ng/μl
CIX157, CIX739, CIX196	–	–	+	+
	Duplex-PCR			
	Xpm1 Boiled samples	Xc1		
UPB008, UPB029, UPB030, UPB031, UPB032, UPB033, UPB035, UPB038, UPB039, UPB041, UPB043, UPB044, UPB045, UPB046, UPB047, UPB049, UPB053, UPB054, UPB059, UPB146	–	+		
Other bacterial strains	Duplex-PCR			
	Xpm1 Boiled samples	Xc1		
LMG539, LMG695, LMG876, LMG901, LMG2804, LMG7429, LMG8014, LMG9055, CFBP2053, CFBP2525, CFBP2532, CFBP2905, CFBP4641, CIX22, CIX202, CIX208, CIX228, CIX230, CPX05, Run 215, Run 17, JX 63 B, LN149, LN154, LN160, LN161, LN163, LN164, LN168, LN170	–	–		
CFBP 2868, CFBP 2903	+	–		

assay.

3.2. Optimization of the duplex-PCR assay

We compared the effect of various concentrations of Xpm1 and Xc1 primer sets on the ability of the duplex-PCR to amplify uniformly both Xpm1 and Xc1 targets. The best results were obtained by reducing the concentration of Xpm1 primers in the mix PCR (Fig. 1C). Particularly, concentrations of 0.125 μM for the Xpm1 primers and 0.5 μM for the Xc1 primers generated a comparable amplification of both amplicons.

3.3. Specificity of the simplex and duplex-PCR assays

The specificity of the simplex PCR assays with primer pairs Xpm1 and Xc1, and of the duplex-PCR combining the two sets of primers was tested on the collection of 50 *Xpm* strains, 23 *Xc* strains, and 32 non-

target strains, using 1 μl of genomic DNA (1 ng/μl) or boiled samples as templates. Our results evidenced 100% of inclusivity for both primer pairs, all target strains being tested positive with their respective primer pairs (Table 5 and Fig. 2). This result was consistent whether the template was purified bacterial genomic DNA or boiled cultures.

We also tested our sets of primers using 39 DNA samples from the IRD historical collection as templates (Table S1). All led to the production of an amplicon of the expected size when using the corresponding primer pair.

Exclusivity values were also very high for both primer pairs, as tested on the collection of non-target strains (n = 32), which includes other *Xanthomonas* species or pathogens, saprophytic strains isolated from cassava and other genera of phytopathogenic bacteria. This list also included strains of *Xpm* (n = 50) when testing *Xc* primers, and strains of *Xc* when testing *Xpm* primers (n = 23). As a result, 100% exclusivity was obtained for Xc1. The Xpm1 primers did not amplify a

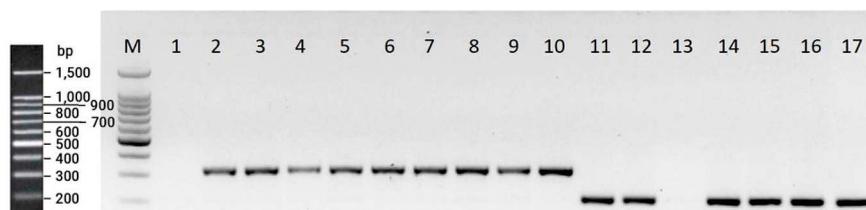


Fig. 2. Agarose gel electrophoresis (negative picture) of the duplex-PCR products using the Xpm1 and Xc1 primer pairs. Strains of *Xanthomonas cassavae* (*Xc*) and *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) were tested for amplification using Xpm1 F/R (203 bp) and Xc1 F/R (318 bp). Lanes M, DNA size marker 100 bp (Promega); lanes 1 and 13, water negative control; lanes 2–10, *Xc* strains UPB041, UPB146, UPB008, UPB029, UPB030, UPB031, UPB032, UPB033 and CIX739; lanes 11–12 and 14–17, *Xam* strains UPB003, UPB006, UPB025, UPB026, UPB027 and CIO151.

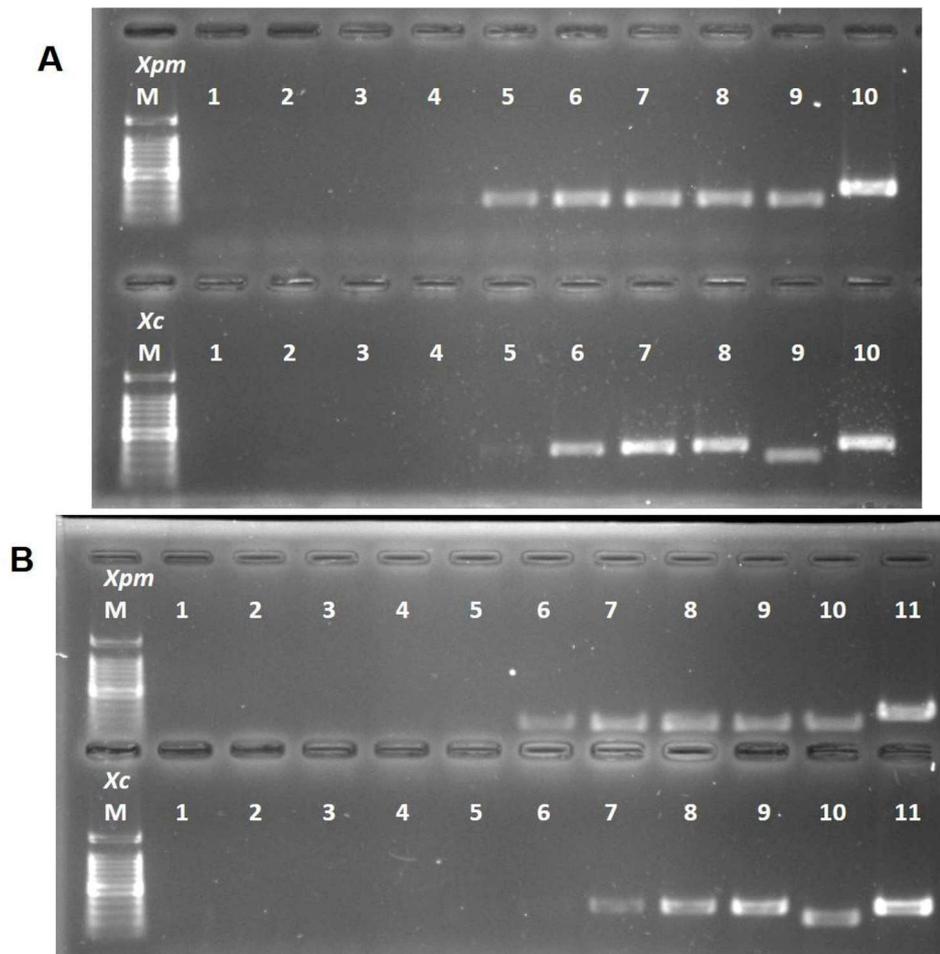


Fig. 3. Duplex-PCR assay performed on serial dilutions of *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) and *Xanthomonas cassavae* (*Xc*) template DNA in absence or presence of plant matrix.

A. Ten-fold diluted boiled suspensions of *Xpm* strain CIO151 (top) and *Xc* strain CIX739 (bottom) were used as templates for duplex-PCR. M: Size marker 100 bp (Promega). Lane 1, water; lanes 2–8: bacterial suspensions ranging from 10^2 CFU/ml to 10^8 CFU/ml for *Xpm* strain CIO151 and 10^1 CFU/ml to 10^7 CFU/ml for *Xc* strain CIX739; lane 9, CIO151 and lane 10, CIX739 gDNA used as positive controls.

B. Total DNA extracted from cassava leaf homogenates spiked with ten-fold diluted suspensions of *Xpm* strain CIO151 (top) and *Xc* strain CIX739 (bottom) were used as templates for duplex-PCR. M: Size marker 100 bp (Promega). Lane 1, water; lane 2, Plant SG-107-5 gDNA, lanes 3–9, bacterial suspensions in plant matrix ranging from 10^1 CFU/ml to 10^7 CFU/ml for *Xpm* strain CIO151, and 10^0 CFU/ml to 10^6 CFU/ml for *Xc* strain CIX739; lane 10, CIO151 gDNA and lane 10, CIX739 gDNA used as positive controls.

product from any of the non-target strains with the exception of two *X. fuscans* pv. *aurantifolii* B strains, leading to a high exclusivity value of 96.4% (Table 5).

3.4. Sensitivity of the duplex-PCR assay

The concentrations of the CIO151 and CIX739 bacterial dilution series used to determine the limit of detection were estimated by spread plate method. The suspensions adjusted to $OD_{600} = 0.2$ corresponded to $\approx 10^8$ CFU/ml ($1.02 \pm 0.26 \times 10^8$ CFU/ml) and $\approx 10^7$ CFU/ml ($1.79 \pm 0.65 \times 10^7$ CFU/ml) for CIO151 and CIX739, respectively.

The duplex-PCR assay tested on pure bacterial suspensions could reliably detect *Xpm* strain CIO151 to a concentration of 10^5 CFU/ml (3/3 positive replicates) which corresponds to a quantity of 100 bacterial cells per reaction. At 10^4 CFU/ml, weak amplification was obtained for only one among the three replicates. For *Xc* strain CIX739, the sensitivity of the assay was slightly lower with only two replicates among three tested positive for 10^5 CFU/ml and a weak amplification in only one among the three replicates at 10^4 CFU/ml (Fig. 3A).

When tested with total DNA extracted from a plant matrix spiked with different concentrations of the bacteria, the reliable threshold of detection was lowered to 10^3 CFU/ml for *Xpm* (3/3 positives replicates) which corresponds here to 20 bacterial cells per reaction after the step of DNA extraction (Fig. 3B). We did not take into account the signals obtained for the lower bacterial concentrations because they were very weak and non-reproducible. For *Xc*, the reliable threshold of detection was 10^4 CFU/ml (3/3 positive replicates).

3.5. Detection of *Xpm* and *Xc* in plant tissues

To address the question whether the duplex-PCR diagnostic tool was efficient in detecting *Xpm* and *Xc* within cassava plant tissues, we artificially inoculated *Xpm* strain CIO151 and *Xc* strain CIX739 into leaves of cassava plants grown under greenhouse conditions. Our results demonstrated consistent detection of *Xpm*. Nevertheless, detection was only possible with macerated lesions that were diluted prior to the PCR assay, for dilutions ranging from 10^2 to 10^5 -fold dilutions (Fig. 4A). As expected, no amplification was demonstrated when using *Xc*1 primers on the same set of samples (Fig. 4B). *Xpm* detection from total DNA extracted from CBB lesions was also demonstrated (Fig. 5).

Conversely, no amplification occurred with macerated lesions from plants inoculated with *Xc* strain regardless of the dilution factor. However, detection of *Xc* could be achieved when the duplex-PCR was performed with total DNA extracted from CBN lesions (Fig. 5).

When performing the duplex-PCR on cassava leaves collected in the 1970s - 80s, *Xpm* was detected in several samples showing characteristic CBB symptoms and previously declared as infected by *Xpm* as is the case of samples 1–4 and 8–14 (Table 3). As with tissues inoculated artificially, dilution of the tissue macerates was necessary prior to performing the duplex-PCR assay (dilutions from 10^2 to 10^5 -fold), with optimal results at 10^3 -fold dilution for the majority of samples. No positive signals could be obtained using this dilution method with CBN symptoms (samples 5–7). However, like for artificially contaminated lesions, *Xc* detection could be achieved after a step of DNA extraction (Table 3).

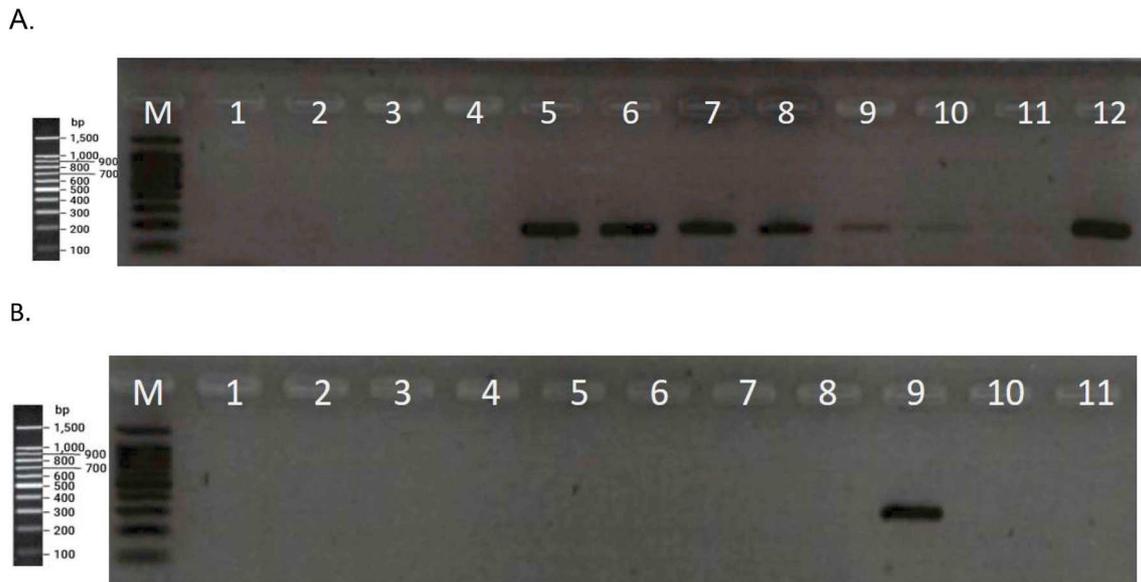


Fig. 4. Duplex-PCR detection of *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) from artificially inoculated cassava leaf tissues.

Macerated lesions of cassava leaves that were infiltrated with *Xpm* strain CIO151 were used as template in the duplex-PCR assay using *Xpm*1 (A) and *Xc*1 (B) primers. A. Agarose gel electrophoresis of *Xpm*1 PCR products (negative picture). Lane 1, water; lane 2, macerate obtained from untreated leaf; lane 3, undiluted macerated lesion; lanes 4 to 11, serial dilutions of macerated lesion material ranging from 10^1 -fold to 1×10^8 -fold and lane 12, gDNA of *Xam* strain CIO151. B. Agarose gel electrophoresis of *Xc*1 PCR products (negative picture). Lane 1, undiluted macerated lesion material; lanes 2–8 and 10, serial dilutions of macerated lesion material ranging from 10^1 -fold to 1×10^8 -fold; lane 9, gDNA of *Xc* strain CIX196 (positive control) and lane 11, water. Lane M, DNA size marker 100 bp (Promega).

4. Discussion

This paper describes the development of a duplex-PCR able to differentially detect *X. phaseoli* pv. *manihotis* and *X. cassavae*, both pathogenic on cassava. CBB is a major threat in the main cassava producing areas in the tropics, while until now, CBN has only been reported in Africa [7]. Nevertheless, the foliar symptoms caused by these two pathogens are very similar, and it is essential to be able to distinguish them in order to deploy adapted control strategies.

Tools based on PCR and N-PCR have been developed, but these only target *Xpm*, and are based on the amplification of a fragment of *TALE1_{Xam}* gene that encodes a major pathogenicity factor widely conserved in *Xpm* [14,19]. This gene is highly prone to horizontal transfer and recombinational events [29]. Therefore using *TALE1_{Xam}* for diagnosis purposes is not optimal because of the pitfall that some strains don't have it [30] and the risk that some non-target strains harbor it, thereby leading to false negative and positive results, respectively.

For the screening of specific primers targeting *Xpm* and *Xc* we used the genome alignment-based computational pipeline designated Uniqprimer 0.5.0 (L. Triplett and J. E. Leach, in preparation), which allows to compare multiple genomes to find unique regions. It has been successfully used for the search of specific loop-mediated isothermal amplification (LAMP) primers for the detection of *X. translucens* pathovars [31]. By combining rapid global alignments before primer

design with stringent screening of the primers, primers specific to the target genome sequences are more efficiently produced than manual or targeted alignment techniques. However, as seen with the initial list of candidate *Xc* primers produced in this study, the proportion of primers that will be robust in the field may be dependent on the quantity and diversity of genomes available for primer design.

Screening of the different primer pairs was performed by exploiting genome sequences available in NCBI databases, particularly for *Xpm*. Our *in silico* predictions were globally consistent with the generated experimental data. Inclusivity results predicted *in silico* for the different *Xpm* primers were in agreement with the results obtained in our PCR assays. *Xc* inclusivity could not be tested *in silico* due to limited genomic information, so we experimentally tested the tool on a collection of 23 strains. The cross reactions of *Xpm*1 primers with *X. fuscans* pv. *aurantifolii* strains and of *Xpm*2 and *Xpm*3 primer pairs with the phylogenetically related strains belonging to the species *X. phaseoli* were also experimentally confirmed. Some *in silico* predictions were not experimentally confirmed (e.g. amplification of *X. translucens* species with primers *Xpm*4 and *Xpm*5). However, because of a lack of availability in some instances, the strains we tested *in vitro* were not necessarily the same isolates that those present in the NCBI databases. These two approaches were complementary and allowed us to develop a duplex-PCR assay with primer set *Xpm*1/*Xc*1 which displayed a high specificity when tested on a broad collection of target and non-target strains.

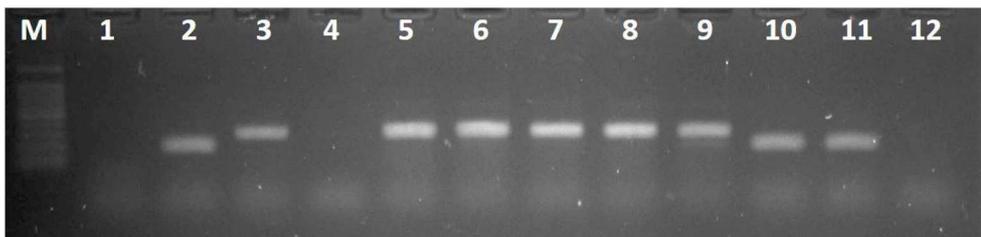


Fig. 5. Duplex-PCR detection of *Xanthomonas cassavae* (*Xc*) and *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) from artificially inoculated cassava leaf tissues.

Total DNA extracted from cassava leaves that were infiltrated with different *Xc* or *Xpm* strains were used as template in the duplex-PCR assay using *Xc*1 and *Xpm*1 primers. Agarose gel electrophoresis of PCR products. Lanes 1 and 12, PCR mix without DNA (negative control); lane 2, gDNA of strain CIO151 (1 ng/μl); lane 3, gDNA of *Xc* strain CIX739; lane 4, total DNA from healthy cassava leaf (negative control); lanes 5 to 9, total DNA extracted from leaves inoculated with *Xc* strains UPB059, UPB146, UPB029, UPB043, UPB046, respectively; lanes 10 and 11, total DNA from leaves inoculated with *Xpm* strains CIO151 and CIX1088, respectively; lane M, DNA size marker 100 bp (Promega).

Indeed, the simplex or the duplex-PCR assays led to the amplification of DNA fragments of the expected sizes with all *Xpm* and *Xc* tested strains. No cross-reactions were obtained except for strains of *X. fuscans* pv. *aurantifolii* when tested with *Xpm1* primers. Nevertheless, it is unlikely that these cross-reactions will interfere in the diagnosis of *Xpm* because most xanthomonads have a highly restricted host range [32] and the ability to induce symptoms on cassava plants seems to be specific to *Xpm* and *Xc*. The most important result is that no amplification was obtained with strains sharing the same ecological niche, such as saprophytic strains isolated from cassava leaves.

The duplex-PCR assay was able to reliably detect *Xpm* from pure culture down to a concentration of 10^5 CFU/ml. For *Xc*, the PCR sensitivity was slightly lower and the limit of detection was between 10^5 CFU/ml and 10^6 CFU/ml. The detection limit in a plant matrix was lowered to 10^3 CFU/ml for *Xpm*, and 10^4 CFU/ml for *Xc*, which corresponds to 20 and 200 CFU per reaction, respectively. This gain of sensitivity compared to pure culture is explained by the step of DNA extraction which both concentrates the DNA target and eliminates PCR-inhibiting plant substances. These results are consistent with the data found in the literature for conventional PCR [33,34]. Consistently, this DNA extraction step greatly improved the diagnosis from artificially or naturally infected plants, particularly for *Xc* for which detection was seen only from total DNA plant extracts. Recently, Zou et al. [35] provided a review on quick and inexpensive approaches to simplify nucleic acid preparation, which could guide the further development of a simplified protocol adapted to high volume of samples, especially in low-resource settings. The level of sensitivity of the current protocol is enough for confirming the presence of the different pathogens in symptomatic lesions. For certification purposes, more sensitive molecular tools like real-time quantitative PCR assays should be next developed.

The duplex-PCR assay described here is a reliable and specific procedure that can be used for epidemiological monitoring of the cassava-pathogenic *Xanthomonas* strains, useful for efficient surveillance programs of these diseases, allowing for the monitoring of crop health and for deploying the control measures that are necessary to prevent severe disease epidemics.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author's contributions

IR and BS conceived and designed the research work. CF, CZ, LT, MK, YM and VML performed the experiments. IR, BS and CF analyzed the data. IR, CF, LG and BS wrote the manuscript. VV, CB, LG and LT critically reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgement

This work was supported by the Agropolis Foundation grants PAIX (#1403-073) and E-SPACE (1504-004), and by the Agricultural Fund for Rural Development (EAFRD) of the European Union. CF was supported by a doctoral fellowship awarded by Fundayacucho and IRD, and CZ was supported by a doctoral fellowship awarded by IRD. We also wish to thank A. Rieux (CIRAD) and H. Maraite (UCL) for fruitful discussions and advices.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.pmp.2018.07.005>.

References

- [1] S. Latif, J. Müller, Cassava – how to explore the “all-sufficient”, *Rural Times* 21 (2014) 30–31.
- [2] E.J. McCallum, R.B. Anjanappa, W. Gruissem, Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*), *Curr. Opin. Plant Biol.* 38 (2017) 50–58 <https://doi.org/10.1016/j.pbi.2017.04.008>.
- [3] E.C. Constantin, I. Cleenwerck, M. Maes, S. Baeyen, C. Van Malderghem, P. De Vos, B. Cottyn, Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex, *Plant Pathol.* 65 (2016) 792–806 <https://doi.org/10.1111/ppa.12461>.
- [4] V. Verdier, S. Restrepo, G. Mosquera, V. Jorge, C. Lopez, Recent progress in the characterization of molecular determinants in the *Xanthomonas axonopodis* pv. *manihotis*-cassava interaction, *Plant Mol. Biol.* 56 (2004) 573–584 <https://doi.org/10.1007/s11103-004-5044-8>.
- [5] A. Fanou, Z. Amégnikin, K. Wydra, Survival of *Xanthomonas axonopodis* pv. *manihotis* in weed species and in cassava debris: implication in the epidemiology of Cassava Bacterial Blight, *Int. J. Adv. Res.* 5 (2017) 2320–2407 <https://doi.org/10.21474/IJAR01/4057>.
- [6] V. Verdier, C. Lopez, A. Bernal, Cassava bacterial blight, caused by *Xanthomonas axonopodis* pv. *manihotis*, in: B. Ospina Patiño, H. Ceballos (Eds.), *Cassava in the Third Millennium: Modern Production, Processing, Use, and Marketing Systems*, CIAT, Cali, Colombia, 2012, pp. 200–212.
- [7] R.J. Hillocks, K. Wydra, Bacterial, fungal, and nematode disease, in: R.J. Hillocks, J.M. Thresh, A.C. Bellotti (Eds.), *Cassava: Biology, Production and Utilization*, International CAB, New York, USA, 2002, pp. 261–280.
- [8] D. Kone, S. Dao, C. Tekete, I. Doumbia, O. Koita, K. Abo, E. Wicker, V. Verdier, Confirmation of *Xanthomonas axonopodis* pv. *manihotis* causing cassava bacterial blight in Ivory Coast, *Plant Dis.* 99 (2015) 1445 <https://doi.org/10.1094/PDIS-02-15-0172-PDN>.
- [9] C.E. López, A.J. Bernal, Cassava bacterial blight: using genomics for the elucidation and management of an old problem, *Trop Plant Biol* 5 (2012) 117–126 <https://doi.org/10.1007/s12042-011-9092-3>.
- [10] M.L. Arrieta-Ortiz, R.L. Rodriguez, A. Perez-Quintero, L. Poulin, A.C. Diaz, N. Arias Rojas, C. Trujillo, M. Restrepo Benavides, R. Bart, J. Boch, T. Boureau, A. Darrasse, P. David, T. Duge de Bernonville, P. Fontanilla, L. Gagnevin, F. Guérin, M.A. Jacques, E. Lauber, P. Lefeuvre, C. Medina, E. Medina, N. Montenegro, A. Munoz Bodnar, L.D. Noël, J.F. Ortiz Quinones, D. Osorio, C. Pardo, P.B. Patil, S. Poussier, O. Pruvost, I. Robène-Soustrade, R.P. Ryan, J. Tabima, O.G. Urrego Morales, C. Vernière, S. Carrère, V. Verdier, B. Szurek, S. Restrepo, C. Lopez, R. Koebnik, A. Bernal, Genomic survey of pathogenicity determinants and VNTR markers in the cassava bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis* strain CIO151, *PLoS One* 8 (2013) e79704 <https://doi.org/10.1371/journal.pone.0079704>.
- [11] V. Verdier, P. Dongo, B. Boher, Assessment of genetic diversity among strains of *Xanthomonas campestris* pv. *manihotis*, *J. Gen. Microbiol.* 139 (1993) 2591–2601.
- [12] V. Verdier, B. Boher, H. Maraite, J.-P. Geiger, Pathological and molecular characterization of *Xanthomonas campestris* strains causing disease of cassava, *Appl. Environ. Microbiol.* 60 (1994) 4478–4486.
- [13] S. Restrepo, V. Verdier, Geographical differentiation of the population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia, *Appl. Environ. Microbiol.* 63 (1997) 4427–4434.
- [14] S. Restrepo, M. Duque, J. Tohme, V. Verdier, AFLP fingerprinting: an efficient technique for detecting genetic variation of *Xanthomonas axonopodis* pv. *manihotis*, *Microbiology* 145 (1999) 107–114.
- [15] R. Bart, M. Cohn, A. Kassen, E.J. McCallum, M. Shybut, A. Petriello, K. Krasileva, D. Dahlbeck, C. Medina, T. Alicai, L. Kumar, L.M. Moreira, J. Rodrigues Neto, V. Verdier, M.A. Santana, N. Kositcharoenkul, H. Vanderschuren, W. Gruissem, A. Bernal, B.J. Staskawicz, High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance, *Proc. Natl. Acad. Sci. U.S.A.* 109 (2012) E1972–E1979 <https://doi.org/10.1073/pnas.1208003109>.
- [16] A. Fessehaie, K. Wydra, K. Rudolph, Development of a new semiselective medium for isolating *Xanthomonas campestris* pv. *manihotis* from plant material and soil, *Phytopathology* 87 (1999) 591–597.
- [17] V. Verdier, S. Ojeda, G. Mosquera, Methods for detecting the cassava bacterial blight pathogen: a practical approach for managing the disease, *Euphytica* 120 (2001) 103–107.
- [18] V. Verdier, G. Mosquera, Specific detection of *Xanthomonas axonopodis* pv. *manihotis* with a DNA hybridization probe, *J. Phytopathol.* 147 (1999) 417–423.
- [19] L. Verdier, G. Mosquera, K. Assigbetsé, Detection of the cassava bacterial blight pathogen, *Xanthomonas axonopodis* pv. *manihotis*, by polymerase chain reaction, *Plant Dis.* 82 (1998) 79–83.
- [20] S. Ojeda, V. Verdier, Detecting *Xanthomonas axonopodis* pv. *manihotis* in cassava true seeds by nested polymerase chain reaction assay, *J. Indian Dent. Assoc.* 22 (2000) 241–247.
- [21] S. Bolot, A. Munoz Bodnar, S. Cunnac, E. Ortiz, B. Szurek, L.D. Noël, M. Arlat, M.A. Jacques, L. Gagnevin, P. Portier, M. Fischer-Le Saux, S. Carrère, R. Koebnik, Draft genome sequence of the *Xanthomonas cassavae* type strain CFBP 4642, *Genome Announc.* 1 (2013), <https://doi.org/10.1128/genomeA.00679-13>.
- [22] S. Kurtz, A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S.L. Salzberg, Versatile and open software for comparing large genomes, *Genome Biol.* 5 (2004) R12 <https://doi.org/10.1186/gb-2004-5-2-r12>.
- [23] T. Koressaar, M. Remm, Enhancements and modifications of primer design program Primer3, *Bioinformatics* 23 (2007) 1289–1291 <https://doi.org/10.1093/>

- bioinformatics/btm091.
- [24] P. Rice, I. Longden, A. Bleasby, EMBOSS: the european molecular biology open software suite, *Trends Genet.* 16 (2000) 276–277 [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2).
- [25] S. Kwok, D.E. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson, J.J. Sninsky, Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies, *Nucleic Acids Res.* 18 (1990) 999–1005.
- [26] D. Bru, F. Martin-Laurent, L. Philippot, Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example, *Appl. Environ. Microbiol.* 74 (2008) 1660–1663 <https://doi.org/10.1128/AEM.02403-07>.
- [27] V. Verdier, S. Restrepo, G. Mosquera, M.C. Duque, A. Gerstl, R. Laberry, Genetic and pathogenic variation of *Xanthomonas axonopodis* pv. *manihotis* in Venezuela, *Plant Pathol.* 47 (1998) 601–608.
- [28] M. Van Den Mooter, H. Maraite, L. Meiresonne, J. Swings, M. Gillis, K. Kersters, J. De Ley, Comparison between *Xanthomonas campestris* pv. *manihotis* (ISPP List 1980) and *X. campestris* pv. *cassavae* (ISPP List 1980) by means of phenotypic, protein electrophoretic, DNA hybridization and phytopathological techniques, *J. Gen. Microbiol.* (1987) 57–71.
- [29] L.F. Castiblanco, J. Gil, A. Rojas, D. Osorio, S. Gutierrez, A. Munoz-Bodnar, A.L. Perez-Quintero, R. Koebnik, B. Szurek, C. Lopez, S. Restrepo, V. Verdier, A.J. Bernal, TALE1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants, *Mol. Plant Pathol.* 14 (1) (2013) 84–95.
- [30] M. Cohn, R.S. Bart, M. Shybut, D. Dahlbeck, M. Gomez, R. Morbitzer, B.H. Hou, W.B. Frommer, T. Lahaye, B.J. Staskawicz, *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava, *Mol. Plant Microbe Interact.* 27 (2014) 1186–1198 <https://doi.org/10.1094/MPMI-06-14-0161-R>.
- [31] P.A. Langlois, J. Snelling, J.P. Hamilton, C. Bragard, R. Koebnik, V. Verdier, L.R. Triplett, J. Blom, N.A. Tisserat, J.E. Leach, Characterization of the *Xanthomonas translucens* complex using draft genomes, comparative genomics, phylogenetic analysis, and diagnostic LAMP assays, *Phytopathology* 107 (2017) 519–527 <https://doi.org/10.1094/PHYTO-08-16-0286-R>.
- [32] F. Leyns, M. De Cleene, J.G. Swings, J. De Ley, The host range of the genus *Xanthomonas*, *Bot. Rev.* 50 (1984) 308–356 <https://doi.org/10.1007/BF02862635>.
- [33] D.A. Cuppels, F.J. Louws, T. Ainsworth, Development and evaluation of PCR-based diagnostic assays for the bacterial speck and bacterial spot pathogens of tomato, *Plant Dis.* 90 (2006) 451–458 <https://doi.org/10.1094/PD-90-0451>.
- [34] Y. Fang, R.P. Ramasamy, Current and prospective methods for plant disease detection, *Biosensors* 5 (2015) 537–561 <https://doi.org/10.3390/bios5030537>.
- [35] Y. Zou, M.G. Mason, Y. Wang, E. Wee, C. Turni, P.J. Blackall, M. Trau, J.R. Botella, Nucleic acid purification from plants, animals and microbes in under 30 seconds, *PLoS Biol.* 15 (2017) e2003916 <https://doi.org/10.1371/journal.pbio.2003916>.