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

En Biologie, Interactions, Diversité Adaptative des Plantes

École doctorale GAIA

Unité de recherche PHIM

Caractérisation des populations de *Xanthomonas oryzae* pv. *oryzae* en vue d'un meilleur déploiement de la résistance en Afrique :des stratégies innovantes de lutte contre la bactériose vasculaire du riz.

Présentée par Coline Sciallano
Le 16 décembre 2022

Sous la direction de Boris Szurek et Sébastien Cunnac

Devant le jury composé de

Pedro Mestre	Rapporteur
Matthieu Arlat	Rapporteur
Rozenn Le Hir	Examinatrice
Eric Giraud	Examinateur
Boris Szurek	Direction de thèse
Sébastien Cunnac	Co-encadrant de thèse

Et présidé par Eric Giraud

UNIVERSITÉ DE MONTPELLIER

REMERCIEMENTS

Je souhaite remercier en premier lieu toute ma famille et mes amis, pour lesquels je n'ai pas été très disponible ces dernières années, merci de votre soutien, et surtout de votre patience !

Je dédie plus particulièrement ce travail à mes parents et à mon frère, je suis heureuse que vous soyez tous les trois fiers de moi. Merci de m'avoir tous soutenue comme vous l'avez fait, j'ai vu vos encouragements à chaque regard.

Je le dédie également à mes grands-parents, qui me manquent tous beaucoup, et qui j'en suis sûre, auraient été très fiers de moi, surtout mes deux grands-pères, pour qui le dur travail avait une grande valeur.

J'adresse mes remerciements à l'IRD et sa présidente directrice générale Valérie VERDIER, pour m'avoir permis de de travailler au sein de l'équipe XPLAIN depuis 2017, ainsi qu'à l'unité PHIM et son directeur adjoint Gilles Béna, pour m'avoir aidée dans le bon déroulement de mon contrat de thèse.

Je remercie les membres de mon CSI, Rozenn le HIR, Thomas KROJ, Pierre CZERNIC et Corinne AUDRAN, qui m'ont aiguillée et donné des retours très intéressants pour la poursuite de ce projet.

Un grand merci à mes précédents maîtres de stage, Tou Cheu XIONG et Eric GIRAUD. Tou Cheu, tu m'as transmis la passion de la recherche et as renforcé ma curiosité naturelle. Eric, tu m'as inculqué la rigueur scientifique, plus particulièrement dans la rédaction. Je te remercie d'avoir cru en mes capacités, ce sont tes avis et tes conseils qui m'ont donné confiance en mes compétences pour débuter cette thèse.

Je remercie sincèrement mon directeur de thèse et mes encadrants de m'avoir tant aidée, guidée, et appris au cours de ces trois années de thèse et même avant cela. Merci pour votre patience et votre dévouement.

Boris, merci de m'avoir intégrée au projet HealthyCrops, et de m'avoir donné l'opportunité de démarrer cette thèse. Ton sens de la diplomatie (franco-allemande en particulier …) et ton calme en situation de crise ont permis que ce projet fleurisse à l'IRD. Merci d'avoir été compréhensif lorsque j'en ai eu besoin, et de t'être montré disponible dans les moments difficiles.

Sébastien, merci de m'avoir fait confiance et de m'avoir accueillie dans l'équipe XPLAIN pour mon premier travail en 2017 ! Merci d'avoir toujours été présent, de toujours tenir tes engagements, et de m'avoir appris tant de choses (notamment sur R, merci de ta patience bouddhique). Tu es passionné et passionnant dans ton travail, rigoureux et autonome sur tous les plans, de la paillasse, aux traitements bioinformatiques les plus poussés... J'aimerais avoir ne serait-ce qu'un dixième de ton cerveau.

Merci à toi Florence, de m'avoir globalement tout appris au labo dès mon arrivée ! Tu es la détentrice du savoir expérimental d'XPLAIN, tout le monde le sait. Merci de gérer la vie de l'équipe et de Genetrop comme tu le fais, avec cette autorité naturelle que j'admire profondément, et ces attentions presque maternelles que tu accordes à chacun. Je n'avais encore jamais vu un labo où tout roule si bien (j'ai la secrète conviction que tout exploserait ou partirait en guérilla si tu t'en allais).

Je remercie tous les chercheurs de l'équipe XPLAIN, Lionel, Ralf, Alvaro, et Mathilde pour m'avoir apporté leur aide, tant pratique que rédactionnelle, et d'avoir participé activement à ma formation. Je remercie aussi Anne SICARD pour m'avoir appris à réaliser et interpréter les analyses MLVA (encore une fois avec une patience d'or !). Une pensée particulière à Lionel et Mathilde avec qui j'ai partagé de longues conversations pas toujours (voir peu souvent) sérieuses et toujours agréables, et aussi de belles aventures en France et ailleurs.

Je remercie évidemment tous les non-permanents du labo ! Marlène, Laurent, Chloé, Dario, Mathilde, Rayan, Kader, Hinda, Jani, Senne, Angeliki, Leonor, Nicolas, Soline, avec qui j'ai partagé de si bons moments, de rires, de cafés, de soirées au bar, de moments de danse, de repas IRDiens, et tant de discussions sur tout et rien ! Je suis contente qu'on se soit rencontrés et qu'on ait pu s'amuser tous ensemble, dans la plus grande rigueur scientifique évidemment ... Merci de m'avoir soutenue quand l'énergie ou l'envie se faisaient rares, tout cela aurait été bien plus difficile sans vous.

Une mention particulière à mon Gabiloux adoré, avec qui j'ai si souvent frôlé la rupture de cage thoracique à force de rire. Tant de blagues au ras des pâquerettes, de jeux de mots, de projections dans des situations délirantes, et surtout de personnages fictifs ! Patrick le vénère, Madame Aude, Janine de la compta, le mec aux andouillettes, Roger le kleptomane … j'en passe et j'en oublie. Un véritable rayon de soleil ! Merci infiniment d'avoir été là pour m'écouter et me remonter le moral dans mes moments difficiles, j'ai presque l'impression que je n'aurais pas pu terminer sans toi.

Je remercie bien évidemment tous mes amis, de Montpellier et d'ailleurs, que j'aime profondément et que j'ai hâte de retrouver ! Aux belles rencontres cette année, Marwan, François Mick, Océane et Fabian, tous les bons moments passés avec vous ont contribué à ce moral d'acier, même dans les périodes de rush intense. Merci !

Une pensée encore plus forte pour ma Juju et ma Roro, mon Nord et mon Sud, mes copines de toujours sans qui je dériverais au large …

Kévin, je te remercie pour tout ce j'ai partagé et appris avec toi, on a effectué nos premiers pas dans la vie d'adulte ensemble et tu m'as appris à avoir plus confiance en moi. Je ne serais pas celle que je suis aujourd'hui si tu n'avais pas été là, merci.

Enfin, je remercie bien évidemment mes colocs adorés, Lolo et Francky qui ont rendu bien plus douce ma nouvelle vie, et si agréable ce nouveau chez moi. Vous avez été (et êtes toujours) ma source quotidienne de rires, de petits moments suspendus, de matins qui chantent ou qui ronflent à l'unissons devant Arte, de soirées cocooning ou au bar, de deliveroo coupables, de sessions mannequinat de dernière minute, et de tant de petits bonheurs. Et aussi, aux fois où vous avez réussi à me ramasser à la petite cuillère, merci. Je suis heureuse de vous avoir rencontré, vous m'avez énormément appris et fait grandir (pas vieillir hein, attention !).

Enfin, je remercie tous mes lecteurs potentiels, et vous souhaite une bonne lecture !

SOMMAIRE

RÉSUMÉ 11
ABSTRACT 12
ABRÉVIATIONS 13
LISTE DES FIGURES ET TABLEAUX 16
INTRODUCTION 21
1 LE RIZ, UNE DES TROIS PREMIERES PRODUCTIONS CEREALIERES AU MONDE. 21
1.1 Origine et diversité du riz. 21
1.1.1 Le genre *Oryza*. 21
1.1.1.1 Description 21
1.1.1.2 Classification 21
1.1.1.3 Ressources génétiques 21
1.1.2 Domestication du riz 23
1.1.2.1 Syndrome de domestication et amélioration variétale 23
1.1.2.2 Origine du riz cultivé en Asie 25
1.1.2.3 Domestication indépendante du riz en Afrique 27
1.2 Culture et importance agronomique du riz. 27
1.2.1 La riziculture dans le monde. 27
1.2.1.1 Systèmes de culture 27
1.2.1.2 Production et consommation 29
1.2.1.3 Commerce 29
1.2.2 Dynamique du riz en Afrique. 31
1.2.2.1 La riziculture africaine 31
1.2.2.2 Consommation et enjeux économiques 31
1.3 Contraintes écologiques affectant la riziculture. 33
1.3.1 Contraintes d'origine abiotique 33
1.3.1.1 Sécheresse 33
1.3.1.2 Température 33
1.3.1.3 Inondations 33
1.3.1.4 Salinité 33
1.3.2 Contraintes d'origine biotique 35
1.3.2.1 Insectes 35
1.3.2.2 Virus 35
1.3.2.3 Champignons 35
1.3.2.4 Nématodes 36

1.3.2.5 Bactéries 36
1.3.2.6 Plantes adventices 36
2 LA BACTERIOSE VASCULAIRE DU RIZ 38
2.1 Distribution et impact pour l'agriculture 38
2.1.1.1 Répartition géographique 38
2.1.2 Incidence et pertes de rendement 38
2.1.3 Mesures de quarantaine 40
2.2 Diagnostic. 40
2.2.1 Symptômes 40
2.2.1.1 Blight 40
2.2.1.2 Kresek 41
2.2.1.3 Pale-yellow leaf 41
2.2.2 Outils moléculaires de diagnostic 41
2.2.2.1 Tests de culture 41
2.2.2.2 ELISA 41
2.2.2.3 Tests par amplification d'ADN 42
2.2.2.4 LAMP-PCR 42
2.3 Épidémiologie. 42
2.3.1 Colonisation de la plante 42
2.3.2 Sources d'inoculum 43
2.3.2.1 Depuis les plantes hôtes 43
2.3.2.2 Depuis le substrat 43
2.3.2.3 Depuis les graines 43
2.3.3 Dissémination 43
2.3.3.1 Vecteurs naturels 43
2.3.3.2 Pratiques culturales 45
2.3.4 Hôtes alternatifs 45
2.4 Méthodes de lutte. 45
2.4.1 Intrants 45
2.4.1.1 Composés antimicrobiens 45
2.4.1.2 Agents de biocontrôle 46
2.4.2 Prévention 46
2.4.2.1 Bonnes pratiques de culture 46
2.4.2.2 Surveillance 48
2.4.2.3 Résistances génétiques 48
3 L'AGENT PHYTOPATHOGENE XANTHOMONAS ORYZAE PV. ORYZAE 48

3.1 Classification. 48
3.1.1 Le genre *Xanthomonas*. 48
3.1.2 L'espèce *Xanthomonas oryzae* 48
3.2 Généralités 50
3.2.1 Caractéristiques bactériologiques 50
3.2.1.1 Morphologie 50
3.2.1.2 Métabolisme 50
3.2.2 Culture et étude 51
3.2.2.1 Observation de symptômes au champ 51
3.2.2.2 Isolement 51
3.2.2.3 Techniques d'infection 51
3.3 Déterminants de pathogénie. 53
3.3.1 Le Système de Sécrétion de Type 3 (T3SS) et ses effecteurs. 53
3.3.1.1 Distribution et évolution 53
3.3.1.2 Description 53
3.3.1.3 Importance et conservation chez Xanthomonas spp. 54
3.3.1.4 Les effecteurs de type 3 54
3.3.2 Les effecteurs Transcription Activator Like (TALEs). 56
3.3.2.1 Distribution et structure 56
3.3.2.2 Spécificité de la reconnaissance des cibles ADN et activation de la transcription 58
3.3.3 Autres déterminants de pathogénie 58
3.3.3.1 Autres systèmes de sécrétion 58
3.3.3.2 Molécules de surface 60
3.3.3.3 Enzymes 60
3.3.3.4 Molécules du Quorum Sensing (QS) 61
3.4 Génome et potentiel évolutif. 61
3.4.1 Données génomiques 61
3.4.1.1 Description générale du génome 61
3.4.1.2 Ressources génomiques disponibles 62
3.4.2 Plasticité du génome et potentiel évolutif 62
3.4.2.1 Principales sources d'évolution chez les microorganismes 62
3.4.2.2 Plasticité du génome chez Xanthomonas et rôle dans l'évolution de la pathogénie 63
3.4.2.3 Evolution des TALEs 65
3.5 Diversité des populations de Xoo. 66
3.5.1 Outils de caractérisation 66

3.5.1.1 Méthodes basées sur les motifs de restriction 66
3.5.1.2 Méthodes basées sur l'hybridation 66
3.5.1.3 Méthodes basées sur l'amplification d'acides nucléiques 67
3.5.1.4 Méthodes basées sur le séquençage 67
3.5.1.5 Profils de virulence sur des lignées isogéniques de riz 67
3.5.2 Distinction des lignées asiatiques et africaines de Xoo 69
3.5.2.1 Différences génétiques et phénotypiques 69
3.5.2.2 Composition des TALomes 69
4 ASPECTS MOLECULAIRES DE L'INTERACTION XOO – RIZ 70
4.1 Immunité innée : Résistances constitutives 70
4.2 L'immunité induite par les pathogènes. 70
4.2.1 La PTI, Immunité induite à large spectre non-spécifique. 70
4.2.1.1 Généralités et définitions 70
4.2.1.2 L'activation de la PTI chez le riz 72
4.2.2 Immunité spécifique induite par les Effecteurs (ETI). 73
4.2.2.1 Généralités et définitions 73
4.3 Transduction des signaux de la PTI et de l'ETI, et mise en place des réponses défensives 73
4.3.1 Voies de signalisation spécifiques à la PTI. 73
4.3.2 Signalisation des MAPK partagée par l'ETI et la PTI 75
4.4 Rôle des effecteurs TALs dans l'interaction 75
4.4.1 Activation de gènes de sensibilité 75
4.4.1.1 Définition 75
4.4.1.2 Gènes de sensibilité non SWEET 77
4.4.2 Gènes de sensibilité majeurs : les transporteurs de sucrose de la famille SWEET ... 77
4.4.2.1 Distribution, structure et fonctionnement des transporteurs SWEET 77
4.4.2.2 Rôle des SWEET dans l'homéostasie du sucre de la plante : focus sur le riz 79
4.4.2.3 Exploitation par les microorganismes 79
4.4.2.4 Les gènes SWEET ciblés par des effecteurs TAL de Xoo 81
4.4.3 Résistance induite par les TALEs (TAL-ETI) 83
4.4.3.1 Reconnaissance directe des TALEs par des récepteurs NLR 83
4.4.3.2 Gènes de résistance de type exécuteur (E) 83
4.4.3.3 Résistance par perte de sensibilité 85
4.4.3.4 Gènes de résistance dans l'interaction Xoo/riz 87
5 INTRODUCTION AU SUJET DE THESE 88
5.1 Contexte 88

5.2 Caractérisation des *Xoo* africaines 88
5.3 Décoder la dépendance des *Xoo* africaines vis-à-vis de l'utilisation des *SWEETs*. 89
5.4 Recherche des bases génétiques de la spécificité des *Xoo* africaines. 89
CHAPITRE I : L'Etude de la diversité des populations de *Xoo* Africaines révèle un clivage Est – Ouest. ... 90
1 Contexte 90
2 Stratégies 90
3 Résultats 91
4 Importance pour le domaine 92
5 Implication personnelle 92
6 Article 1 : « Molecular epidemiology by MLVA genotyping of the Rice Pathogen *Xanthomonas oryzae* pv. *oryzae* reveals two different lineages on the African continent » 93
7 Analyse de génomes Africains 125
7.1 Introduction 125
7.2 Résultats 125
7.2.1 L'assemblage des génomes Africains générés révèle la présence potentielle de plasmides. 125
7.2.2 La construction d'une phylogénie des populations mondiales de *Xoo* 125
7.3 Matériel et méthodes 127
8 Article 2 : « The complete genome resource of the *Xanthomonas oryzae* pv. *oryzae* strain CIX2779 includes the first sequence of a plasmid for an African representative of this rice pathogen » 128
CHAPITRE II : Edition du génome d'une variété élite de riz Africain pour promouvoir la résistance aux souches endémiques et émergentes de *X. oryzae* pv. *oryzae* 136
1 Contexte 136
2 Stratégies 136
3 Résultats 136
4 Importance pour le domaine 137
5 Implication personnelle 138
6 Article 3 : « Genome editing of an African rice elite variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains » 139
CHAPITRE III : comprendre les mécanismes de la virulence indépendante de l'induction de *SWEET14* chez les *Xoo* africaines 167
1 Rationale : l'énigme Africaine 167
2 Recherche des cibles requises pour le développement des *Xoo* africaines chez le riz. 169
2.1 Identification des *SWEETs* du clade III nécessaires à l'infection 169
2.1.1 Choix des lignées mutantes 169

2.1.2 Recherche d'un éventuel gain de résistance acquis chez des plantes mutantes pour *OsSWEET13* et *OsSWEET14* 169
2.1.3 Investigation d'un potentiel effet cumulatif de l'absence de *SWEETs* du clade III 171
2.1.4 Confirmation du phénotype résistant par criblage bactérien à grande échelle. 171
2.2 Recherche de voies de régulation bactériennes de *OsSWEET13* 171
2.2.1 Quantification relative de transcrits par PCR quantitative. 171
2.2.2 Investigation d'une potentielle régulation post-transcriptionnelle des SWEETs par les *Xoo* africaines. 173
3 Identification des déterminants bactériens à l'origine de la capacité des *Xoo* africaines à maintenir de la virulence sans l'induction de *SWEET14* 175
3.1 Etude d'une région génomique contenant *talC* 175
3.1.1 Découverte et implications 175
3.1.2 Un mutant ponctuel dans le gène *talC* est moins affecté que le mutant BAI3.1.1 dans sa virulence 177
3.1.3 Dans la région *talC*, des facteurs autres que *talC* contribuent à la virulence de BAI3. 177
3.1.4 Analyse génétique de l'implication de *talC* et Reg*talC* dans la virulence spécifique des *Xoo* africaines et leur dépendance aux gènes *OsSWEET* du clade III. 179
3.2 Evaluation des performances du métabolisme du sucre chez les *Xoo* africaines 181
4 Matériel et méthodes 183
4.1 Conduite des expérimentations de phytopathologie 183
4.1.1 Origine et conditions de croissance du matériel végétal utilisé 183
4.1.2 Souches bactériennes utilisées et conditions de cultures 183
4.1.3 Déroulement et suivi des tests de pathogénie 183
4.2 Techniques de biologie moléculaire 183
4.2.1 Suivi d'inductions de gènes par PCR quantitative 183
4.2.2 Reproduction de la délétion RegtalC 184
4.2.3 Complémentations talC 184
4.2.4 Séquençage de souches mutantes par la technologie Nanopore 184
4.3 Techniques Biochimiques 184
4.3.1 Quantification cinétique de l'activité GUS 184
4.3.2 Suivis d'activités métaboliques en plaques 185
7 Données supplémentaires : 186
DISCUSSION 191
1 LE CONUNDRUM AFRICAIN 191
1.1 Concernant les déterminants de pathogénie de la région *talC* 191
1.1.1 Recherche de cibles alternatives de TalC 191
1.1.2 Recherche de l'élément conditionnant la virulence apportée par Reg*talC* 191

1.1.3 Validation des déterminants par transfert dans un fond génétique de *Xoo* asiatique 192

1.1.4 Perspectives : Criblage sans a priori d'un élément génétique clé pour l'apparition du phénotype SWIIVA 193

1.2 Au sujet de la résistance des KOs *sweets* 193

1.2.1 Précisions sur l'étude du phénotype SWIIVA sur les lignées mutantes *sweet* 193

1.2.2 Potentiel priming de l'immunité du double mutant *sweet13 ;sweet14* 193

1.2.3 Formation d'hétéromères de SWEET 195

1.2.4 Concernant les profils de longueurs de lésions contrastés sur les KO *sweet* 196

1.2.5 Effet du fond génétique de la variété de riz sur le phénotype des mutants 196

2 CRÉATION DE RÉSISTANCE CONTRE LA BACTÉRIOSE VASCULAIRE DU RIZ. 197

2.1 Méthodes d'amélioration du riz 197

2.1.1 Généralités 197

2.1.2 L'édition de génome pour le gain de résistance à la BLB 197

2.2 Réglementation 198

2.3 Méthodes de maximisation de la durabilité des résistances 200

3 AU SUJET DE LA DURABILITÉ DES RÉSISTANCES CONTRE *XANTHOMONAS* 201

3.1 Émergence de nouveaux pathogènes 201

3.2 Précisions sur l'évolution des TALomes 202

4 FUTURS DÉFIS ET PERSPECTIVES 204

4.1 Changement climatique 204

4.2 Nécessité d'un suivi épidémiologique régulier des populations de *Xoo* 204

4.3 Recherche de nouvelles sources de résistance 205

REFERENCES 207

ANNEXE 1 : 237

ANNEXE 2 : Article en co-auteure "Broad-spectrum resistance to bacterial blight in rice using genome editing" 240

ANNEXE 3 : Supplemental Data "Genome editing of an African elite rice variety confers resistance against endemic and emerging Xanthomonas oryzae pv. oryzae strains" 251

ANNEXE 4 : Article en co-auteure " The Rice ILI2 Locus Is a Bidirectional Target of the African *Xanthomonas oryzae* pv. *oryzae* Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility" 306

RÉSUMÉ

La bactériose vasculaire du riz (BLB), causée par *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) est une menace majeure pour la production mondiale de riz, entraînant chaque année des pertes de rendement atteignant jusque 50% en Asie et en Afrique. La pathogénie de *Xoo* repose sur l'injection dans la cellule hôte d'effecteurs particuliers, appelés Transcription-Activator Like Effectors (TALEs). Ces effecteurs agissent comme des facteurs de transcription qui détournent la machinerie transcriptionnelle de la plante. Les TALEs se lient à des séquences spécifiques appelées Effector Binding Elements (EBE) dans le promoteur de gènes dit de sensibilité (S) et activent leur transcription, qui est essentielle au développement de la maladie. Les membres du clade III de la famille des transporteurs de sucres *SWEET* sont les principaux gènes de sensibilité à la BLB. Les *Xoo* Africaines ciblent *SWEET14* via deux TALEs distincts, soulignant l'importance de ce gène pour leur virulence. Des mutations naturelles dans les EBEs des promoteurs des gènes *SWEET* empêchent la fixation des TALEs correspondants, et annule leurinduction. Ces allèles " non-réceptifs " donnent des gènes de résistance récessifs, réduisant la formation de symptômes et la prolifération bactérienne *in planta*. La création de résistance par ingénierie génétique est un moyen efficace, économique et écologique de garantir la sécurité alimentaire et financière des pays producteurs de riz. Des lignées poly-mutantes éditées au niveau des 6 EBEs majeurs des gènes *SWEET11a*, *SWEET13* et *SWEET14* ont été généré via la technologie CRISPR-CAS. Bien que ces lignées éditées soient résistantes aux souches d'origine Asiatique, les souches Africaines de *Xoo* conservent un certain niveau de virulence.

Nous avons alors cherché à comprendre la dépendance des *Xoo* Africaines vis-à-vis de l'utilisation des gènes *SWEET*. Par l'inoculation d'une importante collection d'isolats Africains sur des lignées mutantes pour les gènes *SWEET* du clade III, nous montrons que *SWEET13* est potentiellement impliqué dans la virulence résiduelle des *Xoo* Africaines.

Nous avons également prospecté les bases génétiques spécifiant la virulence des *Xoo* Africaines. La caractérisation phénotypique de plusieurs mutants nous a permis de découvrir une région génomique de 35 kilobases, conditionnant la virulence de la souche modèle BAI3 du Burkina Faso et se révélant prometteuse dans l'explication du phénotype Africain.

Un autre volet de nos travaux a consisté en l'étude des populations de *Xoo* Africaines. Bien que les souches Asiatiques de *Xoo*, décrites pour la première fois en 1884, aient largement été étudiées, la découverte ultérieure de l'agent pathogène sur le continent Africain (1979) et l'importance plus réduite de la riziculture en Afrique font que ces populations endémiques ont été étudiées plus tardivement. Des résultats antérieurs ont montré que les *Xoo* Africaines sont très éloignées des *Xoo* Asiatiques et génétiquement plus proches de *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*). Pour mieux comprendre leur histoire évolutive et leur potentiel épidémique, ainsi que pour promouvoir le déploiement local de sources de résistance adaptées, notre objectif était de caractériser une importante collection de *Xoo* Africaines. Nos premières analyses via l'utilisation de marqueurs moléculaires ont mis en évidence la structure génétique des populations de *Xoo* sur le continent Africain et ont révélé que les isolats de l'Ouest sont clairement distincts des isolats de l'Est, ces derniers se révélant très proches des *Xoo* Asiatiques. Un deuxième volet d'analyses, par séquençages complets de génomes, a confirmé le caractère invasif de ces populations et pointé leur origine vers le Yunnan et la Thaïlande. A l'instar des souches Asiatiques, ces souches dépendent de *SWEET11a* selon le même modèle d'induction et possèdent des TALEs tronqués qui leur permettent d'échapper à la résistance médiée par le gène hôte *Xa1*.

ABSTRACT

Bacterial Leaf Blight (BLB) which is caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is a major threat to rice production worldwide, leading up to 50% yield losses in Asia and Africa during severe infections. *Xoo* pathogenicity relies on the injection into the host cell of Transcription-Activator Like Effectors (TALEs) which are effective transcription factors specialized to hijack the plant transcriptional machinery. TALEs bind to Effector Binding Elements (EBE) in the promoter of susceptibility (S) genes and activate transcription, which is essential for disease development. Clade-III members of the family of *SWEET* sugar transporters are major BLB susceptibility genes. Compared to Asian *Xoo*, African strains have fewer TALEs and specifically targets *SWEET14* as a major susceptibility gene. Besides, African *Xoo* use two distinct TAL effectors to induce *SWEET14*, highlighting the importance of this gene for bacterial proliferation. Naturally occurring polymorphism in the EBEs of *SWEET* promoters prevent TALE binding, resulting in loss of *SWEET* gene induction. These “unresponsive” alleles act as recessive resistance genes, preventing symptoms formation and reducing bacterial colonization in planta. Genetic engineering to gain resistance is an effective, economic and ecologic way to guarantee food and economic safety of rice producing countries. Poly-mutant edited lines at the level of the 5 major EBEs of the *SWEET11a*, *SWEET13* and *SWEET14* genes were generated via CRISPR-Cas9 technology. However, while our multi-edited lines were able to defeat *Xoo* strains of Asian origin, African *Xoo* strains still maintained some level of virulence.

Here, we sought to unravel African *Xoo* dependence towards the use of *SWEET* genes. By inoculating a large collection of strains on knock-out lines for clade III *SWEET* genes, we show that *SWEET13* is potentially involved in the residual virulence of African *Xoo*.

We also invastigated the bacterial genetic bases for the virulence specificity of African *Xoo*. By phenotyping several bacterial mutants, we discovered a 35 kilobases genomic region conditionning the virulence of the model strain BAI3 from Burkina Faso, which appears promising to explain African particular virulence phenotype.

Finally, while Asian strains of *Xoo*, first described in 1884, are widely studied, the later discovery of the pathogen on the African continent (1979) and the lesser importance of African rice cultivation have made these endemic strains less characterized. Previous results have shown that African difer strongly from Asian *Xoo*, and are genetically more closely related to *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*). To better understand their evolutionary history and epidemic potential, and in order to promote the deployment of locally adapted sources of resistance, our objective was to characterize a large collection of African *Xoo*. Our analyses shed light on the genetic structure of *Xoo* populations on the African contient and reveal two different lineages : West African *Xoo* and East African *Xoo*. We show that East African *Xoo* strains are closely related to Asian strains, and that they also rely on the induction of *SWEET11a* to promote disease, using similar TAL effectors. East African strains, as their Asian relatives, do countain truncated TALEs that allow them to escape from the resistance mediated by *Xa1* a rice nucleotide-binding domain leucine rich protein (NLR) that specifically recognize TALEs.

ABRÉVIATIONS

AD, Activation Domain

ADN, Acide DésoxyriboNucléique

ADP, Adenosine Di Phosphate

ANI, Average Nucleotide Identity

ARN, Acide RiboNucléique

arTALE, TALE artificiels

ATP, Adenosine Tri Phosphate

Avr, Avirulence

bhlh, Basic Helix Loop Helix

BLB, Bacterial Leaf Blight

BLS, Bacterial Leaf Streak

bzip, Basic-region leucine-Zipper

CABI, Centre for Agricultural Bioscience International

Cas9, CRISPR associated protein 9

CC, Clonal Complex

CDPK, Protéines Kinases Dépendantes du Calcium

CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats

DAMPs, Damage-Associated Molecular Patterns

EBE, Effector Binding Element

EF, Elongation Factor

EGM, éléments génétiques mobiles

EPPO, European and Mediterranean Plant Protection Organization

EPS, ExoPolySaccharides

ETI, Effector-Triggered Immunity

FAO, L'Organisation des Nations unies pour l'Alimentation et l'Agriculture

GC, Grouping Complex

HGT, transfert de gènes horizontal

HHU, Université Heinrich Heine de Düsseldorf

hpa, hrp Associated

HR, Réponse Hypersensible

hrp, Hypersensitive Response Protein

ICBN, International Code for Botanical Nomenclature

IRBB, IRRI Bacterial Blight

IRD, Institut de Recherche pour le Développement durable

IRRI, International Rice Research Institute

IS, séquences d'insertion

Itale, TALE interférent

KO, Knock-Out

L'EFSA European Food Safety Authority

LAMP-PCR, Loop-Mediated Isothermal Amplification PCR

LRR, domaine riche en répétitions leucines

LRR, Leucine Rich Repeats

MAPK, Mitogen-Activated Protein Kinases

MLVA, multiple-locus VNTR Analysis

MSTs, Monosaccharides Trasporters

NADPH, Nicotinamide Adenine Dinucleotide Phosphate

NCBI, National Center for Biotechnology Information

NILs, lignées quasi-isogéniques

NLR, Nucleotide-Binding - Leucine-Rich-Repeat Receptors

NLS, Nuclear Localization Signal

ONT, Oxford Nanopore Technologies

PAMPs, Pathogen-Associated Molecular Patterns

PCR, Polymerase Chain Reaction

PFGE, Pulsed Field Gel Electrophoresis

PR, Pathogenesis-Related proteins

PRRs, Pattern-Recognition Receptors

PSA, Peptone Sucrose Agar

PTI, Pathogen-Triggered Immunity

qPCR, quantitative PCR

RFLP, Polymorphisme de Longueur de Fragments de Restriction

RLK, Receptor Like Kinase

ROS, Reactiv Oxygen Species

RPLs, Receptor Like Proteins

RPPO, Regional Plant Protection Organization

RVD, Repeat Variable Diresidues

RVD, Repeat Variable Diresidues

RYMV, Rice Yellow Mottle Virus

SUT, Sucrose Uptake Transporters

SWEET, Sugars Will Eventually be Exported Transporter

T1SS, T2SS, T3SS, T4SS, T5SS, T6SS, Type X Secretion System

T3E, Type 3 Effector

TALE, Transcription Activator Like Effector

TALEN, arTALe fusionné à une nucléase

TFB, Transcription Factor Binding

Tn, transposon

UFC, Unit Forming Colony

USDA, Département de l'Agriculture des États-Unis

VNTR, Variable Number of Tandem Repeats

WAK, cell Wall Associated Kinase

WGS, Whole Genome Sequencing

Xcc, *Xanthomonas citri* pv. *citri*

Xo, *Xanthomonas oryzae*

Xoc, *Xanthomonas oryzae* pv. *oryzicola*

Xoo, *Xanthomonas oryzae* pv. *oryzae*

Xop, *Xanthomonas* Outer Protein

LISTE DES FIGURES ET TABLEAUX

Remarque : les figures et tableaux propres aux articles ne sont pas répertoriés ici.

Tableaux

Tableau 1 : La classification du riz au cours du temps par différents taxonomistes..... 19
Tableau 2 : classification des différentes espèces du genre Oryza selon leur formule génétique 20
Tableau 3 : Les différents gènes et QTLs de résistance à la BLB ayant été isolés depuis des riz sauvages. 22
Tableau 4 : principaux stress abiotiques contraignant le riz. 32
Tableau 5 : catégorisation en quarantaine de *Xoo* selon différentes régions et RPPOs. 39
Tableau 6 : Liste EPPO 2022 des plantes hôtes de *Xoo*. 44
Tableau 7 : Critères biochimiques d'identification de *Xoo* et *Xoc*..... 49
Tableau 8 : Les différentes interactions (bénéfiques ou pathogènes) dans lesquelles les gènes *SWEET* sont impliqués. 80
Tableau 9 : Liste des différentes EBEs trouvées dans les promoteurs SWEET du clade III. 82
Tableau 10 : Souches séquencées contenant des iTALEs..... 84
Tableau 11 : Liste des gènes présents dans la région *RegtalC* et leurs annotations respectives. 174
Tableau S1 : Amorces utilisées pour la qPCR sur les SWEET du clade III. 189
Tableau S2 : Souches sauvages et mutantes de Xoo et Xoc utilisées au cours de des études présentées dans le Chapitre III et la discussion 189

Figures

Figure 1 : Morphologie du riz. 19
Figure 2 : Classification AGP III des Angiospermes. 20
Figure 3 : carte de répartition des espèces sauvages du complexe *Oryza sativa* (génomes AA), et de Leersia. 22
Figure 4 : Syndrome de domestication du riz. 24
Figure 5 : Facteurs impliqués dans l'amélioration des variétés et des rendements de riz au cours du temps. 24
Figure 6 : Les différents groupes variétaux de *O. sativa*. 26
Figure 7 : Les différents systèmes de production rizicole 26
Figure 8 : Production et consommation de riz dans le monde. 28
Figure 9 : La part du riz dans l'apport calorique, régions sélectionnées. 28
Figure 10 : Consommation de riz dans les pays d'Afrique de l'Ouest entre 1980 et 2017. 30
Figure 11 : Autosuffisance pour la production de riz en Afrique en 2020. 30
Figure 12 : Diagramme d'interaction des différentes menaces biotiques du riz en fonction des stades de développement de l'hôte. 34
Figure 13 : Distribution géographique de *Xoo*. 37
Figure 14 : Symptômes causé par *Xoo* sur le riz. 39
Figure 15 : Phylogénie, gamme d'hôtes et maladies associées au genre *Xanthomonas*. 47
Figure 16 : Morphologie de *Xanthomonas*. 49
Figure 17 : Représentation schématique du T3SS des bactéries phytopathogènes 52
Figure 18 : Spécificité de la reconnaissance TALE/ADN. 55
Figure 19 : L'activation de la transcription eucaryote par les TALEs de *Xanthomonas* 57
Figure 20 : Répartition des différents TSS chez *Xo* 57
Figure 21 : Schéma des morphologies des différents systèmes de sécrétion. 59
Figure 22 : Principaux mécanismes d'évolution des TALEs. 64
Figure 23 : Populations de *Xo* à l'échelle mondiale. 68
Figure 24 : Heatmap de la présence / absence des TALEs majeurs chez 63 *Xoo*. 68
Figure 25 : Induction de l'ETI et de la PTI chez les plantes, et transduction du signal. 71
Figure 26 : Cibles caractérisées des TALEs de *Xanthomonas*. 74
Figure 27 : Phylogénie des membres de la famille des gènes *SWEET* chez *Arabidopsis* (At), le riz (Os) et le maïs (Zm). 76
Figure 28 : Structure protéique des SWEETs. 76
Figure 29 : Rôle des SWEET dans le loading / unloading du phloème. 78
Figure 30 : Les différents variants de TALEs majeurs retrouvés chez *Xoo*. 82
Figure 31 : Structure des gènes Exécuteurs impliqués dans la résistance à *Xoo* 86
Figure 32 : Une phylogénie basée sur 411 génomes de *Xo* mondiaux révèle l'origine des *Xoo* d'Afrique de l'Est. 126
Figure 33 : Mutagenèse de *SWEET13* et *SWEET14* par édition du génome par CRISPR-Cas9 168
Figure 34 : Gain de résistance à large spectre contre les Xoo africaines par perte de fonction *sweet13 ; sweet14.* 170
Figure 35 : Un double mutant *sweet14 ; sweet15* ne confère pas de résistance supplémentaire aux *Xoo* africaines. 170
Figure 36 : Potentielle induction de *SWEET13* par les *Xoo* africaines. 172
Figure 37 : Niveaux de transcrits des gènes *SWEET* du clade III dans les feuilles de riz non infectées. 172

Figure 38 : Investigation d'un potentielle accumulation de protéines *SWEET13* en présence de *Xoo* africaines. 174
Figure 39 : Comparaison des séquences du génome de référence BAI3 et du génome du mutant BAI3.1.1 au niveau de la région *talC*. 177
Figure 40 : La région délétée dans le mutant BAI3.1.1 a un impact dans la virulence de BAI3. 178
Figure 41 : Validation du phénotype produit par la délétion de *RegtalC*. 178
Figure 42 : La complémentation par *cbsA* seul des mutants Reg*talC* ne procure pas de gain de virulence. 180
Figure 43 : Mise en évidence d'un rôle potentiel de la région *talC* dans le phénotype de virulence Africain. 180
Figure 44 : Métabolisme différentiel des sucres chez *Xanthomonas oryzae.* 182
Figure S1 : Organigramme décisionnel pour la résolution de l'énigme Africaine. 187
Figure S2 : Vecteur de délétion pK18_Reg*talC*_SpecR. 188
Figure S3 : Vecteur de complémentation *talC*. 188
Tableau S2 : Souches sauvages et mutantes de Xoo et Xoc utilisées au cours de des études présentées dans le Chapitre III et la discussion. 189
Figure 45 : La double mutation de *SWEET13* et *SWEET14* chez Kitaake ne fournit pas de résistance à Xoc. 194
Figure 46 : Assemblages de sous-unités SWEET en homo et hétéropolymères décrits chez *A. thaliana.* 194
Figure 47 : État de la législation sur l'édition de génome dans le monde en juin 2022. 199

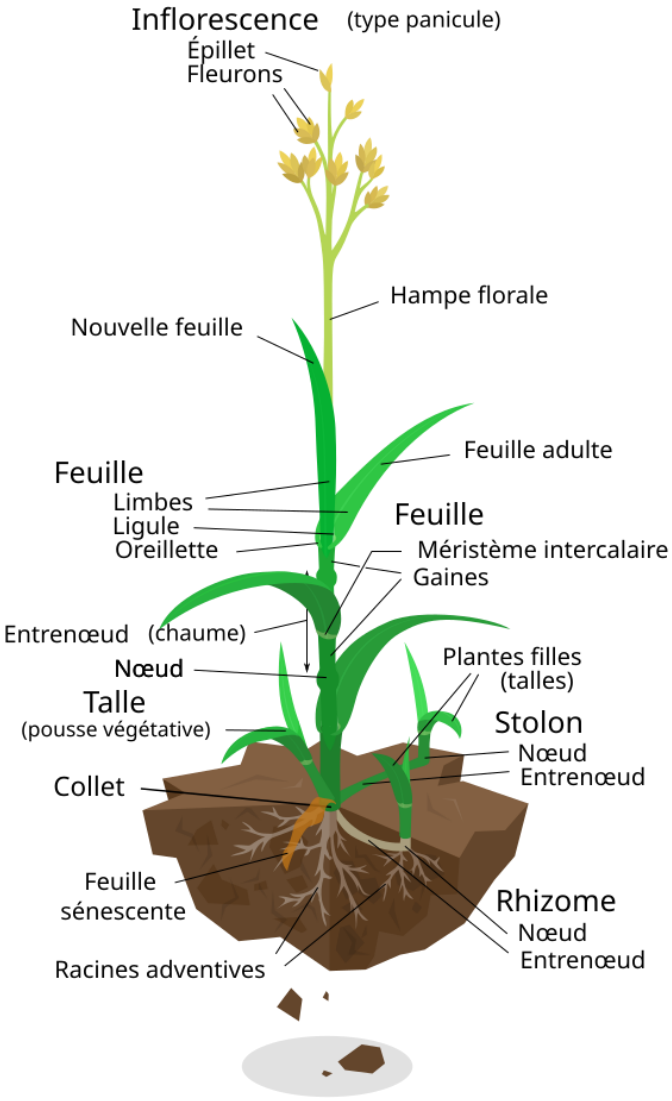
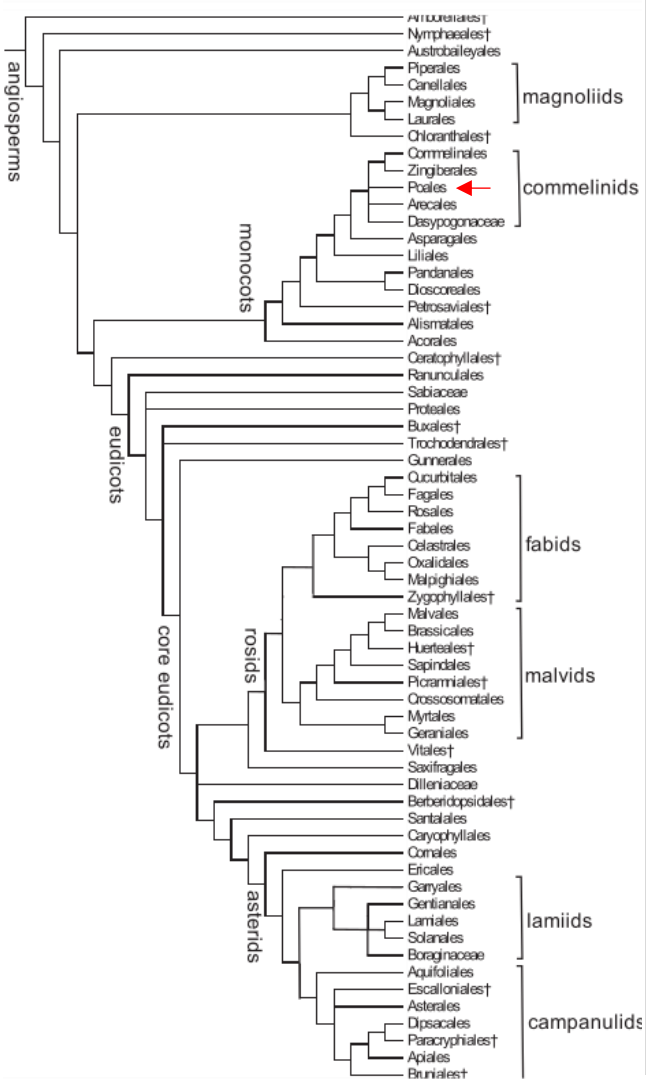


Figure 1 : Morphologie du riz. Adaptation par Spedona, dessin original de Kelvinsong.

Tableau 1 : La classification du riz au cours du temps par différents taxonomistes (Lu, 1999[1]).

a. Certaines ou toutes les espèces de cette section ne sont plus incluses dans *Oryza*.

Roschevicz (1931)	Chevalier (1932)	Ghose et al (1965)	Sharma & Shastry[b] (1965, 1972)	Tateoka[c] (1963)	Oka (1988)	Vaughan[c] (1989)
Sect. *Sativa*	Sect. *Euoryza*	Sect. *Sativa*	Sect. *Oryza*	*O. latifolia*	Sect. *Oryzae*	*O. sativa*
Sect. *Granulata*	Sect. *Padia*	Sect. *Officinalis*	Ser. *Latifoliae*	*O. sativa*	Sect. *Schlechterianae*	*O. officinalis*
Sect. *Coarctata*[a]	Sect. *Sclerophyllum*[a]	Sect. *Granulata*	Ser. *Sativae*	*O. glaberrima*	Sect. *Granulatae*	*O. ridleyi*
Sect. *Rhynchoryza*[a]	Sect. *Rhynchoryza*[a]		Sect. *Padia*	*O. ridleyi*	Sect. *Ridleyanae*	*O. meyeriana*
			Ser. *Schlechterianae*	*O. meyeriana*	Sect. *Angustifoliae*[a]	Others[e]
			Ser. *Meyerianae*	Others[d]	Sect. *Coarctatae*[a]	
			Ser. *Ridleyanae*			
			Sect. *Angustifolia*[a]			
			Ser. *Brachanthae*			
			Ser. *Perrierianae*			



Section Complex *Species*	Genome group
Oryza	
***Oryza sativa* complex**	
Oryza sativa L.	AA
O. rufipogon sensu lacto[†] (syn: *O. nivara* for the annual form *O. rufipogon sensu stricto* for the perennial form)	AA
O. glaberrima Steud.	AA
O. barthii A. Chev.	AA
O. longistaminata Chev. et Roehr.	AA
O. meridionalis Ng	AA
O. glumaepatula Steud.[‡]	AA
***O. officinalis* complex**	
O. officinalis Wall ex Watt	CC
O. minuta JS Presl. ex CB Presl.	BBCC
O. rhizomatis Vaughan	CC
O. eichingeri Peter[§]	CC
O. malapuzhaensis Krishnaswamy and Chandrasakaran	BBCC
O. punctata Kotschy ex Steud.	BB, BBCC
O. latifolia Desv.[#]	CCDD
O. alta Swallen	CCDD
O. grandiglumis (Doell.) Prod.	CCDD
O. australiensis Domin	EE[¥]
Ridleyanae Tateoka	
O. schlechteri Pilger	Unknown[**]
***O. ridleyi* complex**	
O. ridleyi Hook.	HHJJ
O. longiglumis Jansen	HHJJ
Granulata Roschev.	
***O. granulata* complex**[¶]	
O. granulata Nees et Arn ex Watt	GG
O. meyeriana (Zoll. et Mor. ex Steud.) Baill.	GG
Brachyantha B.R. Lu	
O. brachyantha Chev. Et Roehr.	FF

Figure 2 : Classification AGP III des Angiospermes. Cette classification est construite à la base de deux gènes chloroplastiques et un gène nucléaire de ribosome. Les Poales (indiquées par une flèche rouge) sont comprises dans le clade des Commelinidées. (The Angiosperm Phylogeny Group [5]).

Tableau 2 : Classification des différentes espèces du genre Oryza selon leur formule génétique (Vaughan *et al.*, 2003 [10]).

INTRODUCTION

1 LE RIZ, UNE DES TROIS PREMIERES PRODUCTIONS CEREALIERES AU MONDE.

1.1 Origine et diversité du riz.

1.1.1 Le genre *Oryza*.

1.1.1.1 Description

Le riz est une plante herbacée annuelle dont les variétés cultivées mesurent en moyenne 1 à 2 mètres avec cependant certaines variétés naines de 30 à 40 cm de haut et des variétés flottantes pouvant atteindre les 7 mètres. La plante possède des chaumes ronds, droits, lisses, creux et articulés, et des feuilles étroites et linéaires aux limbes plats et rugueux. Le riz porte des inflorescences en panicule d'épillets uniflores et autogames (moins de 10% d'allogamie chez O. *sativa*). En se développant, la plante forme plusieurs branches appelées thalles, arborant généralement chacun une panicule (Figure 1)(Chang et Bardenas, 1965[2]; Coste et Flahault, 1937 [3] ; Clément *et al.*, 2002 [4])

1.1.1.2 Classification

Les membres du genre *Oryza* sont des plantes angiospermes, monocotylédones appartenant à la famille des *Poaceae* (Figure 2) (The Angiosperm Phylogeny Group, 2009 [5]). Les *Poaceae* aussi appelées *Gramineae* sont des plantes herbacées, regroupant les espèces désignées sous le nom d'herbes et céréales. Cette famille regroupe 780 genres et plus de 12000 espèces de plantes dont le blé, le maïs, l'orge ou encore la canne à sucre. Elle est de loin la plus importante dans nos systèmes agriculturaux de par son rôle crucial dans l'alimentation humaine et animale (Xu et Zhou, 2017 [6]). La tribu des *Oryzeae* regroupe les sous-tribus *Zizaniinae* et *Oryzinae*, cette dernière comprenant les genres *Oryza* et *Leersia* (Soreng *et al*., 2015 [7]). Le genre *Oryza* a été décrit pour la première fois en 1753 par Carl Von Linné sur la base d'une unique espèce (O. *sativa Linnaeus*)[8]. *Oryza* compte deux espèces cultivées, O. *sativa* et O. *glaberrima* ainsi que 25 espèces sauvages que l'on retrouve dans les régions tropicales et subtropicales (Lu, 1999 [1]; Chen *et al*., 2019 [9]). Certaines caractéristiques génétiques et écologiques des espèces sauvages du genre ne sont pas retrouvées chez les espèces cultivées, comme la formation de rhizomes et la polyploïdie (Vaughan *et al.*, 2003 [10]). La complexité taxonomique d'*Oryza* a donné lieu à de nombreux changements dans la classification du genre, dont les remaniements majeurs sont résumés dans le Tableau 1. La nomenclature suit les règles de l'ICBN (International Code for Botanical Nomenclature) et se base sur les différences observables entre taxons et le principe d'antériorité.

1.1.1.3 Ressources génétiques

Aujourd'hui, grâce à l'utilisation de marqueurs moléculaires, les différentes espèces de riz ont pu être classées en 11 types de génomes. On distingue alors les espèces au génome diploïde (n = 12), réparties selon les formules génétiques AA, BB, CC, EE, FF et GG ; et les espèces allotétraploïdes (n = 24) de formules BBCC, CCDD, HHJJ, HHKK et KKLL (Vaughan *et al*., 2003 [10]; Stein *et al.*, 2018 [11]; Chen *et al.*, 2019 [9]) (Tableau 2). Contrairement aux génomes AA qui ne sont retrouvés que sous forme diploïde, les génomes CC et BB ont également donné naissance à plusieurs espèces polyploïdes, largement distribuées. Le génome DD, n'est retrouvé que sous forme tétraploïde en combinaison avec CC, il pourrait donc provenir soit d'une espèce éteinte, soit d'une espèce existante qui n'a pas encore été découverte. Les espèces du génome AA sont étroitement liées génétiquement. Parmi elles, on

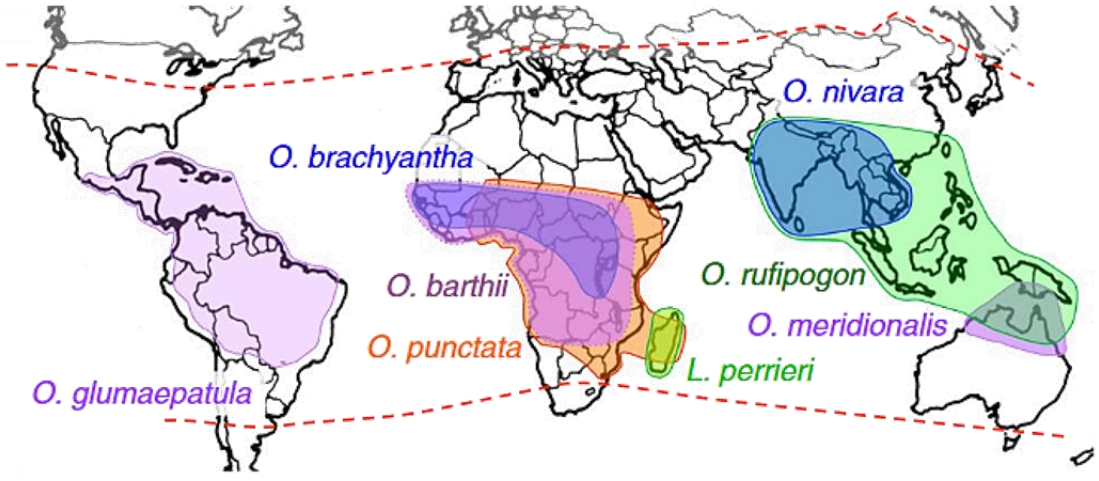


Figure 3 : carte de répartition des espèces sauvages du complexe *Oryza sativa* (génomes AA), et de Leersia. Les lignes rouges pointillées montrent les limites de la culture du riz (Stein *et al*, 2018 [11]).

Genes	Sources	Chromosome location	Dominant/ Recessive
Xa1	Kogyoku	4	Dominant
Xa2	Rantai-Emas 2	4	Dominant
Xa3/Xa26	WaseAikoku	11L	
Xa4	IR20, IR22 and IR1529-680-3	11	Dominant
Xa5	IR1545-284 and RP291-7 DV85, DV86 and DZ78	5S	Recessive
Xa6	Zenith, Malagkit Sungsong,		Dominant
Xa7	DV85, DV86 and DZ78	6	Dominant
Xa8	PI231129	7	Recessive
Xa9	Khao Lay Nhay and Sateng		Recessive
Xa10	Cas 209	11	Dominant
Xa11	IR944-102-2-3,	3S	Dominant
Xa12	Kogyoku and Java 14	4	Dominant
Xa13	BJ	8L	Recessive
Xa14	TN1	4	Dominant
Xa15	M41		Recessive
Xa16	Tetep		Dominant
Xa17	Asominori		Dominant
Xa18	Toyonishiki, Milyang 23, IR24		Dominant
Xa19	XM5		Recessive
Xa20	XM6		Recessive
Xa21	*Oryza longistaminata*	11L	Dominant
Xa22	Zhachanglong	11	Dominant
Xa23 (t)	*O. rufipogon*	11	Dominant
Xa24 (t)	DV 86	2	Recessive
Xa25 (t)/Xa25	Minghui 63	12	Recessive
Xa25 (t)	HX-3	4	Dominant
Xa26 (t)	Nep Bha Bong To		Recessive
Xa27 (t)	Arai Raj	6	Dominant
Xa28 (t)	Lota Sail		Recessive
Xa29 (t)	B5 (*Oryza officinalis*)	1	Dominant
Xa30 (t)	Y238 *O. rufipogon*	11	Dominant
Xa31 (t)	Zhachanglong	4	Recessive
Xa 32 (t)	*Oryza meyeriana*	12	Recessive
Xa32 (t)	*Oryzae ustraliensis*	11L	Dominant
Xa33 (t)	Ba7	6	Recessive
Xa33 (t)	*O. nivara* IRGC 105710	7	Dominant
Xa34 (t)	BG 122	1	Recessive
Xa35 (t)	*Oryza minuta* (Acc. No. 101133)	11	
Xa36 (t)	C4059	11L	Dominant
Xa38	*O. nivara* IRGC 81825	4L	Dominant
Xa39	FF329	11	Dominant
Xa40 (t)	IR65482-7-216-1-2	11	Dominant
Xa41 (t)	African wild and cultivated species of *O. barthii* and *O. glaberrima*		Recessive
Xa42	XM 14, a mutant of IR 24	3	Recessive
Xa43 (t)	Colombia XXI (P8)	11	Dominant
Xa44 (t)	IR73571-3B-11-3-K3 (P6)	11	Recessive
Xa45 (t)	*Oryza glaberrima* Accession IRGC 102600B	8L	Recessive

Tableau 3 : Les différents gènes et QTLs de résistance à la BLB ayant été isolés depuis des riz sauvages. Ce tableau adapté de Kesh et Kaushit, 2020 [17] répertorie les différents gènes et QTLs de résistance contre la bactériose vasculaire du riz (BLB), maladie causée par *Xanthomonas oryzae* pv.*oryzae*. Les cultivars dont les gènes ou QTLs de résistance sont issus sont indiqués. Les variétés de riz sauvages sont mises en surbrillance en vert.

retrouve les deux espèces cultivées O. *sativa* L. et O. *glaberrima* et six espèces sauvages : O. *rufipogon*, O. *barthii*, O. *nivara* (forme annuelle de O. *rufipogon*), O. *longistaminata*, O. *meridionalis*, et O. *glumaepatula* (Vaughan *et al*., 2003 [10]). De par sa culture, O. *sativa* est distribué mondialement, avec néanmoins une plus forte concentration en Asie du Sud et du Sud Est. O. *glaberrima* en revanche n'est cultivé qu'en Afrique de l'Ouest. O. *rufipogon* se trouve dans toute l'Asie et l'Océanie. Les espèces sauvages africaines O. *longistaminata* et O. *barthii* sont réparties dans toute l'Afrique pour la première et principalement en Afrique de l'Ouest pour la seconde (figure 3)(Stein *et al*., 2018 [11]). Toutes les espèces sauvages du genre *Oryza*, ainsi que les différentes variétés de riz, constituent un pool génétique précieux pouvant être utilisé dans les programmes d'amélioration variétale (Lu, 1999 [1]). Cependant, le processus de sélection par l'homme a engendré un goulot d'étranglement génétique qui a été à l'origine d'une réduction de la diversité génétique du génome des espèces cultivées (Kumagai *et al*., 2016 [12]). Des efforts internationaux ont été réalisés au début des années 1970 pour collecter du matériel génétique de riz traditionnels et sauvages (Vaughan *et al*., 2003 [10]). De nombreux gènes de résistance, notamment contre la bactériose vasculaire du riz, sont issus du génome de variétés de riz sauvages (tableau 3). Cette année encore, une région génomique conférant de la résistance contre cette maladie a été découverte chez O. *officinalis* (Chen *et al.*, 2022 [13]).

1.1.2 Domestication du riz

1.1.2.1 Syndrome de domestication et amélioration variétale

La domestication des plantes a amené à la sélection de traits adaptés aux besoins humains. Ainsi il existe une suite de caractéristiques sélectionnées de manière similaire chez différentes espèces de plantes, qui distinguent les variétés utilisées en agriculture de leurs géniteurs (Doebley *et al.*, 2006[14]). Cette « sélection convergente » en agriculture est connue sous le terme de « syndrome de domestication » (Hammer, 1984[15]). Ce syndrome, que l'on retrouve aussi chez les animaux apprivoisés par l'homme, peut être défini comme l'ensemble des traits phénotypiques associés à la transformation génétique d'un organisme en une forme domestiquée à partir d'une forme progénitrice sauvage (Allaby, 2014[16]). Les cultures destinées à l'exploitation humaine ont généralement des fruits ou des grains plus gros en comparaison à leurs précurseurs sauvages. Ce sont typiquement des plantes plus robustes dans l'ensemble, présentant souvent une dominance apicale accrue (croissance importante de la tige centrale par rapport aux tiges latérales). Aussi, dans le cas des céréales, les plantes domestiquées ne sont en général plus capables de se propager dans la nature, car la dispersion naturelle des graines a été évincée en faveur d'une récolte facilitée pour les humains, grâce à des grains plus solidement attachés. Les traits sélectionnés par l'homme chez les plantes qu'il utilise sont variés et vont de la perte de dormance des graines, à la diminution de substances amères dans les structures comestibles, en passant par la floraison synchronisée (Doebley *et al.*, 2006[14]). Chez le riz, le syndrome de domestication a conduit à une allocation élevée des ressources aux graines, des grains plus gros, mais aussi à une disparition du caractère de détachement spontané des grains. La croissance du riz cultivé est plus déterminée, avec une dominance apicale (Chen *et al.*, 2019[9]). Les connaissances sur les origines de la domestication du riz et sa diversité génétique ont progressé de manière fulgurante avec l'avènement du séquençage haut débit. Notamment, « Le projet des 3000 génomes de riz » en 2014 a constitué un effort international de reséquençage par technologie Illumina de 3000 accessions originaires des 5 continents. Ces données permettent de comprendre la diversité génétique des populations de riz à un niveau de détails plus fin, de mettre en évidence de potentiels berceaux de domestication et servent de base à la découverte à grande échelle de nouveaux allèles pour l'amélioration variétale. Des traces d'évènements indépendants de domestication de riz sauvages ont

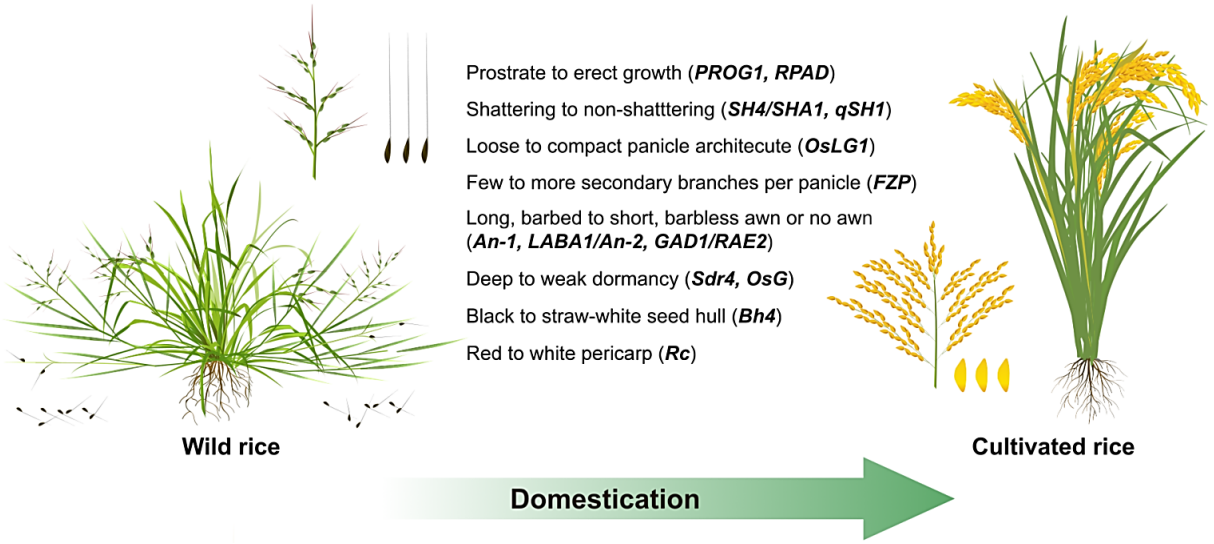


Figure 4 : Syndrome de domestication du riz. Schéma récapitulant les changements morphologiques et physiologiques ayant été sélectionnés au cours de la domestication du riz et les gènes clés sous-jacents (Chen *et al.*, 2021 [29]).

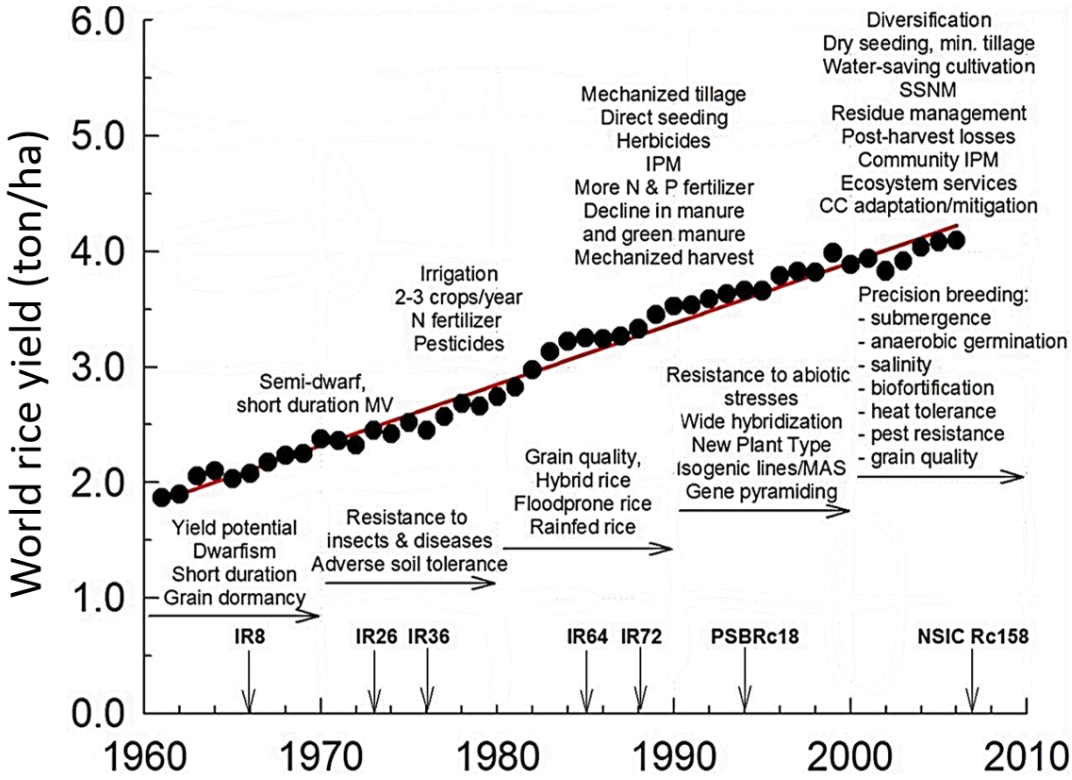


Figure 5 : Facteurs impliqués dans l'amélioration des variétés et des rendements de riz au cours du temps. (Mew et Vera Cruz [30]).

été mises en évidence en Asie (env. 9000 ans), en Afrique (env. 2000 ans), et récemment en Amérique du Sud dans une région de l'Amazonie (évènement de domestication datant d'environ 4000 ans) (Hilbert *et al.*, 2017[18] ; Cubry *et al.*, 2018 [19]). Dans les années 1950, la découverte d'un mutant naturel pour le gène *sd1* (gibbérelline oxydase), semi-nain en milieu irrigué a permis d'énormes progrès de rendement (Sasaki et al, 2002[20]) (Figure 4). L'amélioration de ces variétés s'est poursuivie par sélection d'une maturation précoce, un cycle de vie permettant leur plantation trois fois par an et une meilleure réponse aux engrais azotés, ce qui a contribué à partir des années 1960 à ce que l'on appelle la révolution verte (Muthayya *et al.*, 2014 [21]). En 2006 notamment, Xu et ses collègues [22] ont découvert un facteur de réponse à l'éthylène, *Sub1A*, conférant la tolérance à l'immersion. Grâce à l'utilisation de marqueurs facilitant la sélection, ils ont pu introgresser avec succès le gène dans une « méga variété » (variété de riz haut rendement) indienne, Swarna. Par la suite, de nombreuses variétés ont été améliorées par l'introduction de ce gène, dont la méga variété Indonésienne Ciherang, à laquelle a également été rajouté un QTL majeur de la tolérance a la germination en anaérobie, AG1, découvert dans le cultivar Khao Hlan On originaire du Myanmar (Toledo *et al.*, 2015 [23]). Les facteurs ayant contribué à l'amélioration des variétés de riz, et à l'amélioration des rendements en général au cours du temps sont présentés dans une frise en figure 5.

1.1.2.2 Origine du riz cultivé en Asie

La domestication de novo d'*O. sativa* à partir d'*O. rufipogon* s'est produite il y a environ 9000 ans dans la vallée de Yangtze dans le sud de la Chine. Le processus de domestication par l'homme a permis de générer une nouvelle sous-espèce, O. *sativa* spp *japonica*, par sélection de variants des gènes *LABA1*, *PROG1* (prostrate growth 1, croissance de type rampant) et SHATTERING 3 (*SH3*, contrôlant le phénotype d'abcission des grains), comme résumé en figure 4. Au cours des 4 000 premières années de son histoire, le riz *japonica* n'était cultivé qu'en Chine. Sa dispersion et sa diversification n'ont eu lieu qu'à partir d'une brutale période de refroidissement il y a 4200 ans (Gutaker *et al.*, 2020 [24]). Par la suite, différentes sous-populations d'*O. sativa* ont vu le jour, par l'introgression des allèles domestiqués de *japonica* dans des proto-populations de riz sauvages (Choi et Purugganan, 2018[25]). Par exemple, O. *sativa* ssp. *indica*, une autre grande sous-espèce de riz largement cultivée en Asie tropicale, est issue de l'introgression de *japonica* dans une population proto-*indica* qui était cultivée en Inde dans la vallée inférieure du Gange il y a 5 000 ans (Gutaker *et al.*, 2020 [24]; Alam *et al.*, 2021 [26]). La dissémination et diversification du riz *japonica* vers l'Asie du Sud-Est est soumise à débat. Selon une théorie dominante, la riziculture se serait répandue en Asie du Sud-Est à partir de Taïwan dans le cadre d'une émigration austronésienne il y a environ 3000 ans, d'abord dans l'archipel des Philippines, puis dans le reste de l'Asie insulaire (Alam *et al.*, 2021 [26]). Aujourd'hui, cinq différentes sous-populations d'O. *sativa* réparties dans toute l'Asie peuvent-être définies : *indica*, aus/boro, *basmati*/sadri, *japonica* tropical et *japonica* tempéré (Figure 6, The 3,000 rice genomes project, 2014 [27]). Les variétés *indica* et *japonica* peuvent être distinguées par différentes caractéristiques phénotypiques, comme la forme des grains (allongés chez *indica* et plutôt ronds chez *japonica*), la teneur en amidon, qui est plus importante chez *indica* et enfin la tolérance aux température basses, typique des variétés *japonica* (Civáň *et al.*, 2019 [28]).

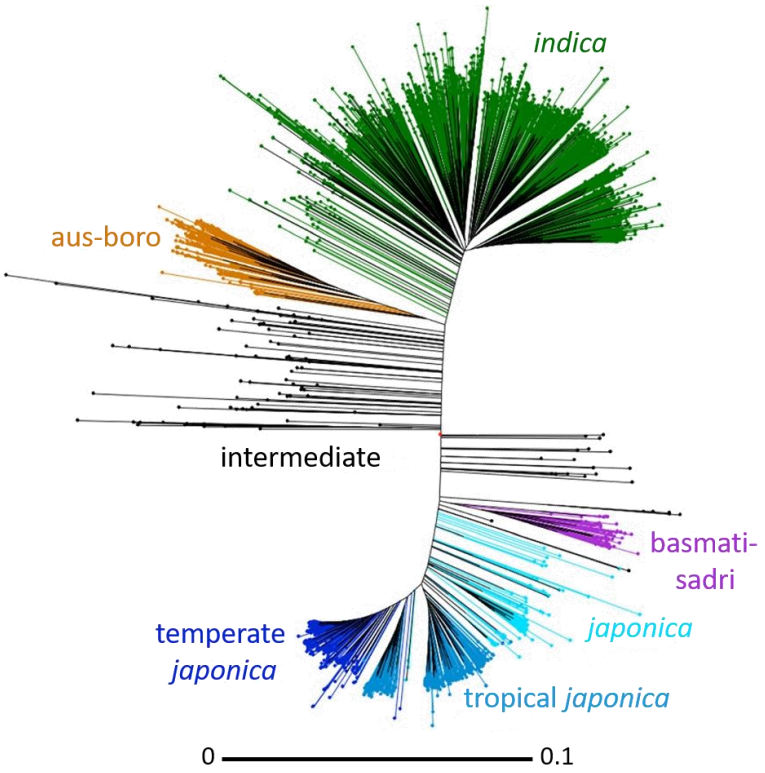


Figure 6 : Les différents groupes variétaux de *O. sativa*. Phylogénie Classification des différents groupes variétaux de O. *sativa* basée sur une analyse SNP de 3 000 accessions de riz. Les analyses ont révélé une différenciation claire des 3 000 entrées en 5 groupes : deux groupes principaux : *indica* et *japonica*, deux petits groupes (aus/boro et basmati/sadri), un groupe « intermédiaire » dont les accessions n'ont pu être classées. Adapté de « The 3,000 rice genomes project », 2014 [27].

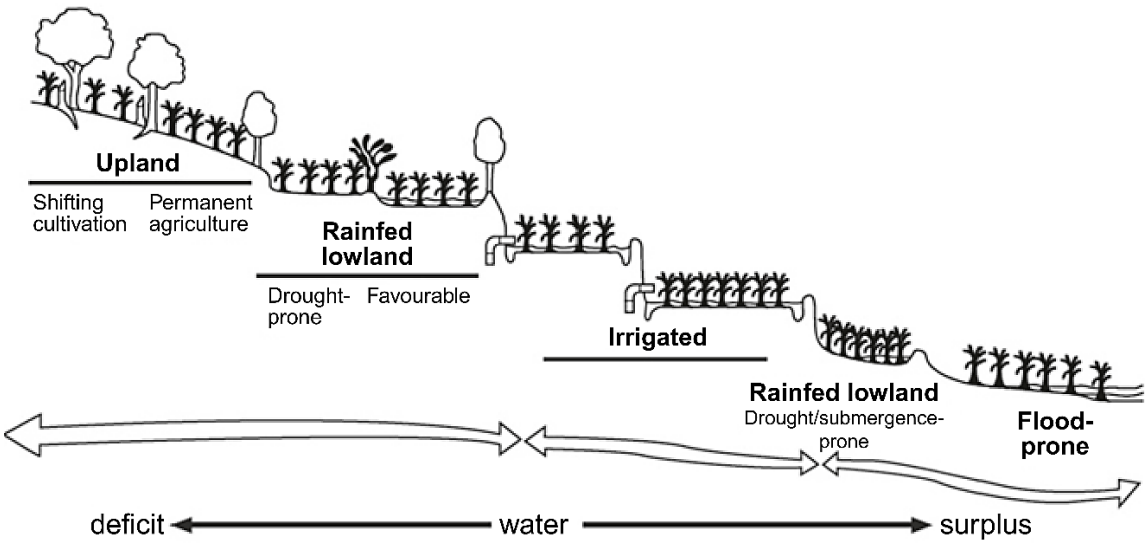


Figure 7 : Les différents systèmes de production rizicole (classification de l'IRRI, basée sur l'apport hydrique). (Schéma tiré de Prasad *et al*, 2017 [35]).

1.1.2.3 Domestication indépendante du riz en Afrique

Le riz africain, *O. glaberrima*, a été domestiqué indépendamment du riz asiatique *O. sativa* depuis l'espèce *O. barthii* (Wang *et al.*, 2014 [31]). Des études récentes, basées sur l'analyse d'indices de diversité calculés sur 163 génomes d'O. *glaberrima* et 83 génomes d'O. *barthii* collectés au Sahel et en Afrique de l'Est ont désigné le nord du delta intérieur du Niger comme centre de la domestication du riz en Afrique (Cubry *et al.*, 2018 [19]). Ces conclusions rejoignent les résultats de précédentes études ethnobotaniques et linguistiques, ainsi que les plus anciennes traces archéologiques attestant de l'existence d'une culture du riz en Afrique (Murray, 2004 [32]). Il est intéressant de retrouver certains gènes domestiqués de manière convergente entre O. *glaberrima* et O. *sativa*, comme par exemple les gènes *PROG1* et *SH3*/*SH4*. *PROG1* est impliqué dans la transition entre la croissance de type rampant et la croissance de type dressé, un changement d'architecture clé dans le processus de domestication du riz. Une version mutée de ce gène est retrouvée chez O. *sativa*, tandis qu'il est complètement absent chez O. *glaberrima*. Chez *O. sativa*, l'absence d'abscission des grains est contrôlée par des SNPs dans les gènes *SH3*/*SH4* (codant pour un facteur de transcription) et le QTL de détachement des grains sur le chromosome 1 *qSH1* (Konishi *et al.*, 2006 [33], Chen *et al.* 2019 [9]). De la même manière, il a été rapporté que le phénotype de non-abcission des grains chez O. *glaberrima* a été causée par un SNP résultant en une protéine *SH3*/*SH4* tronquée (Win *et al.*, 2017 [34]). Ces évènements, totalement indépendants, ont abouti au même phénotype de domestication (Cubry *et al.*, 2018).

1.2 Culture et importance agronomique du riz.

1.2.1 La riziculture dans le monde.

1.2.1.1 Systèmes de culture

Les nombreuses techniques de cultures du riz sont soumises aux conditions climatiques et topographiques du lieu de production. De par les besoins importants en eau de la riziculture, les systèmes de production sont principalement distingués selon l'apport hydrique. Une classification proposée en 1993 par l'International Rice Research Institute (IRRI, Philippines) permet de différencier les écosystèmes rizicoles en quatre catégories (Figure 7, Prasad *et al.*, 2017[35]) :

- La riziculture irriguée, qui assure 75% de la production mondiale et représente plus de la moitié des surfaces agricoles consacrées au riz, est caractérisée par une maîtrise totale de l'apport en eau et la possibilité de réaliser plus d'une récolte par an.

- La riziculture pluviale, dont l'apport en eau est soumis aux précipitations naturelles, ne représente que 13 % des surfaces rizicoles en Asie mais entre 60 et 75 % en Afrique et en Amérique Latine respectivement (Courtois, 2007 [36]). Elle peut être subdivisée en deux sous - catégories : la riziculture d'altitude et la riziculture de bas-fonds. La riziculture pluviale d'altitude est caractérisée par des sols ondulés en pente avec un fort ruissellement. La riziculture de bas-fonds connaît à la fois des problèmes d'inondation et de sécheresse. Environ un quart des rizières du monde sont des bas-fonds pluviaux.

- Enfin, le reste des cultures est regroupé sous l'appellation rizicultures inondables ou flottantes. Ces terrains occupent 8 % de la superficie rizicole et se trouvent aux abords des fleuves dont ils sont tributaires des crues et décrues et parfois sujets à des inondations d'eau saumâtres.

Chaque type de riziculture est confronté à différents problèmes de nutrition, de toxicité, d'adventices de maladies et de ravageurs, et de ce fait possède son propre modèle de gestion.

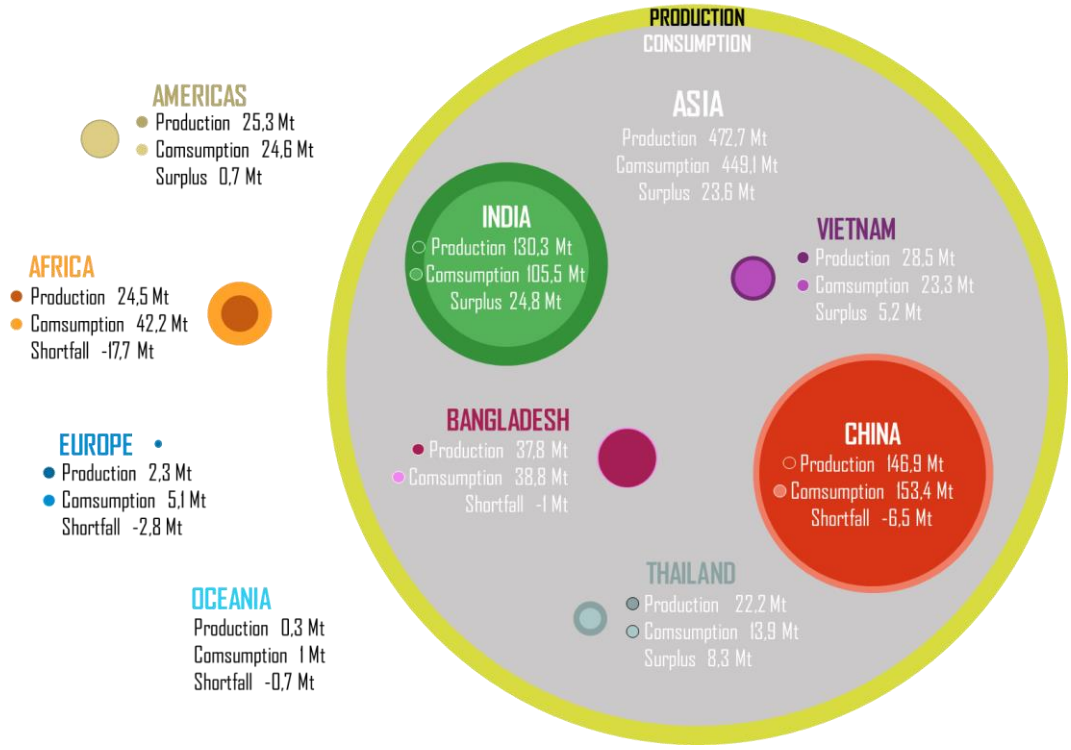


Figure 8 : Production et consommation de riz dans le monde en 2021. Réalisé depuis les données 2022 de la FAO.

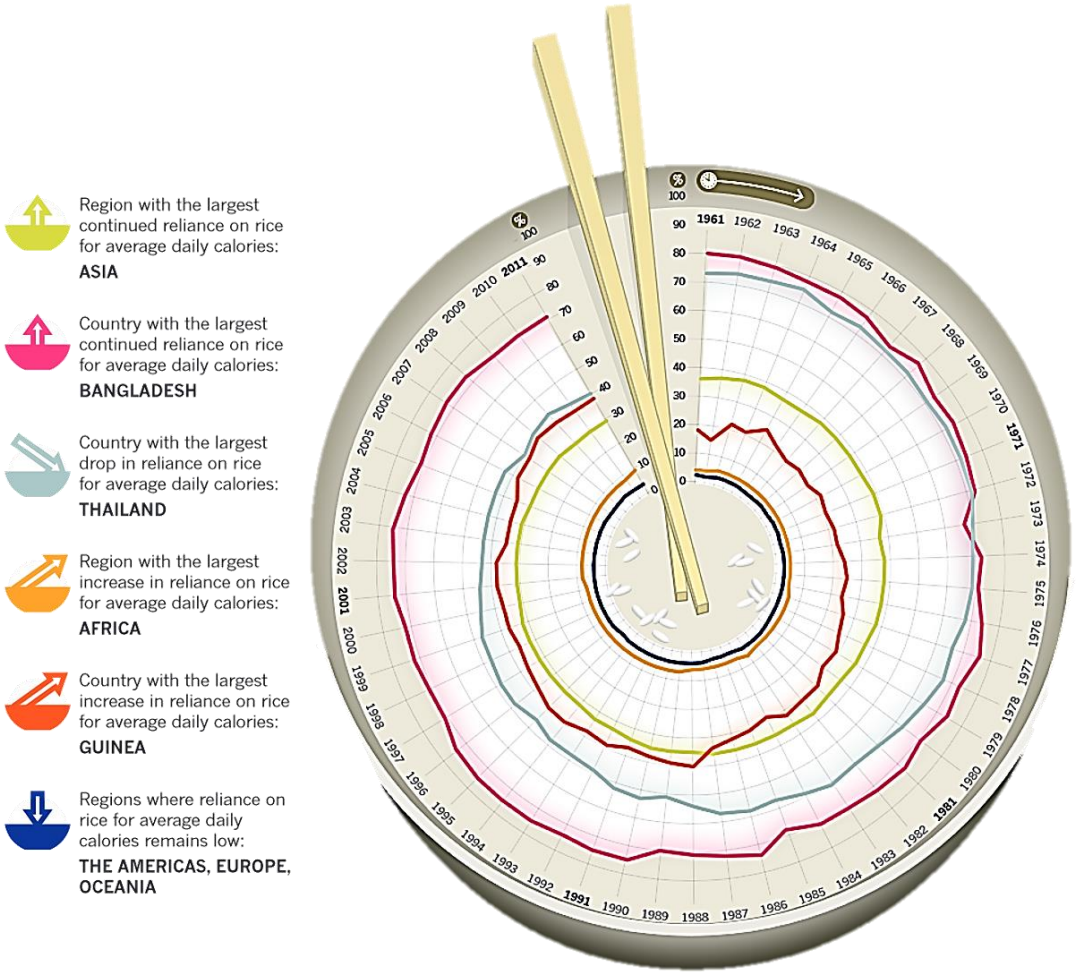


Figure 9 : La part du riz dans l'apport calorique, régions sélectionnées. (Tiré de Elert, 2014 [39])

1.2.1.2 Production et consommation

Le riz est la 3e céréale produite dans le monde après le blé et le maïs. Sa production est allée en continuelle augmentation, et est passée de 445 millions de tonnes produites en 2008/2009 à 513,5 million de tonnes pour l'année en cours (FAO 2018 [37]; USDA, 2022[38]). Bien qu'ayant noté une baisse de production de 1,2 % depuis 2010, la Chine reste le plus gros producteur mondial avec 28,9% de la production totale en 2021 c'est-à-dire 147 million de tonnes (figure 8). Les surfaces allouées à la riziculture continuent d'augmenter, bien que le perfectionnement des techniques agricoles et des variétés de riz durant la révolution verte aient permis de d'améliorer amplement les rendements. La surface moyenne pour produire une tonne de riz est passée de 0,49 hectares dans les années 1960 à 0,24 hectares dans les années 2000 (Elert, 2014 [39]). Les terres rizicoles recouvraient 166 million d'hectares en 2018, avec environ 90 % de cette surface retrouvée en Asie et 80% répartie dans huit pays, à savoir la Chine, l'Inde, l'Indonésie, le Bangladesh, les Philippines, le Vietnam, la Thaïlande et la Birmanie (Prasad *et al.*, 2017[35], FAO 2018 [37]). Les 10 % restants sont majoritairement distribués entre l'Amérique Latine et l'Afrique. Bien que faible, la part de l'Afrique dans la production mondiale est en augmentation : on passe de 3,9 % en 2015 à 4,5 % en 2022, tandis que la part de l'Asie est passée de 90,4% à 89,5 % (FAO 2018 [37]; USDA 2022[38]). La consommation domestique globale de riz est passée de 485 million de tonnes en 2018 à 514 million de tonnes en 2022. Cette augmentation est occasionnée par un accroissement de la demande en Afrique Sub-Saharienne, en Asie de l'Est, du Sud et du Sud-Est (FAO 2018 [37]; USDA 2022[38]). Cependant, cette augmentation n'est pas due à une part grandissante du riz dans l'assiette des consommateurs, mais plutôt à l'effet de la croissance démographique. Des études ont effectivement montré que la contribution du riz à l'apport calorique dans les pays asiatiques (plus gros producteurs et consommateurs de riz) était en baisse. Entre 1961 et 2007, cette part est passée de 30,2 % à 26,8 % en Chine, tandis qu'elle est passée de 79,1 % à 69,8 % au Bangladesh (Timmer 2010 [40]). Mais la baisse la plus drastique a été enregistrée en Thaïlande, où l'on est passé d'environ 72% d'apport calorique assuré par le riz en 1961 à 40% en 2011 (Elert, 2014 [39]) (figure 9).

1.2.1.3 Commerce

La majeure partie du riz est directement consommé dans le pays de production. Les plus grands producteurs se suffisent désormais à eux-mêmes et les excédents de production valorisables finissent en exportations, en aliments pour le bétail ou en semences. Les principaux exportateurs depuis 2013 sont l'Inde (21,5 million de tonnes exportées en 2021), la Thaïlande (7 million de tonnes), le Vietnam (6,6 million de tonnes), le Pakistan (4,8 million de tonnes) et les USA (2,5 million de tonnes) (FAO 2018 [37]; USDA 2022 [38]) (figure 8). Les variétés *indica* dites aromatiques *basmati* et Thaïlandaises sont les plus exportées (Courtois, 2007). Les pays en voie de développement sont les plus gros importateurs de riz. Leur part dans l'importation mondiale depuis 2013 est de 90,4 %, représentant 43.5 million de tonnes en 2017. Les plus grandes régions importatrices de riz sont le Moyen-Orient, l'Afrique subsaharienne et l'Afrique de l'Ouest (Prasad *et al.*, 2017[35]). Le prix d'achat du riz étant disparate selon les pays, les enjeux en termes de sécurité alimentaire sont importants, surtout dans les pays en état de crise géopolitique.

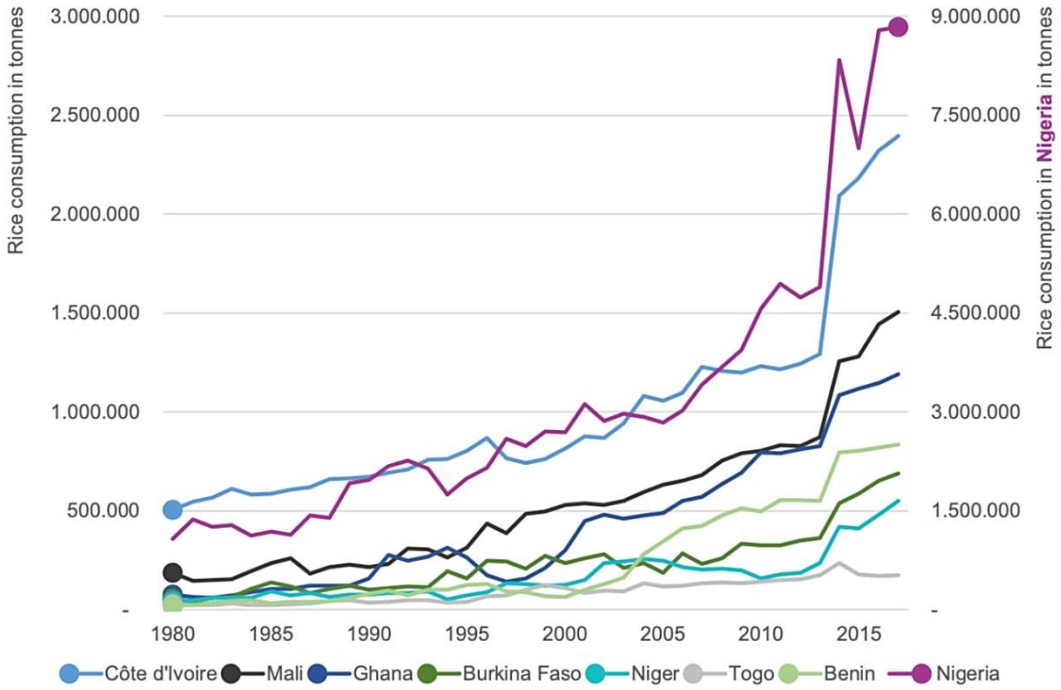


Figure 10 : Consommation de riz dans les pays d'Afrique de l'Ouest entre 1980 et 2017. (Issu de Tondel *et al*., 2020 [44]).

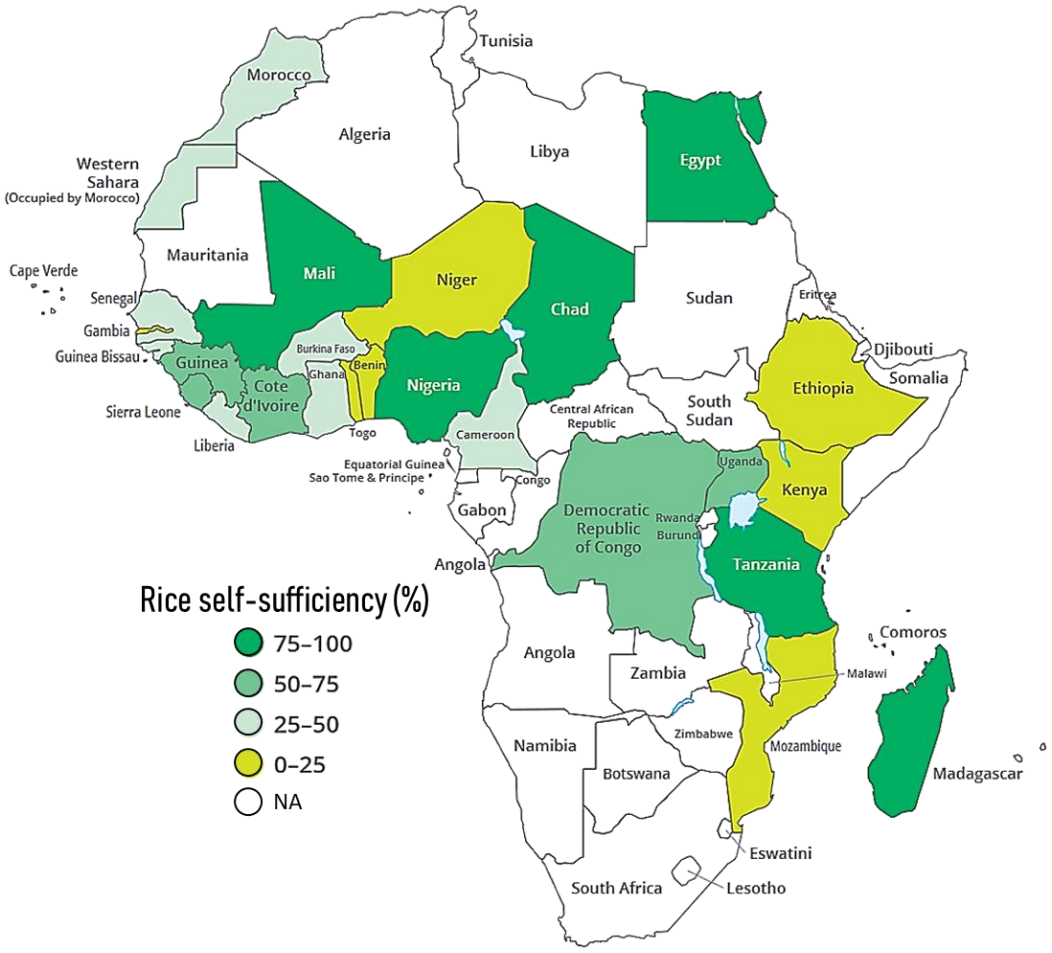


Figure 11 : Autosuffisance pour la production de riz en Afrique en 2020. (tiré du rapport AfricaRice 2020, « 50 Years of Research Development »[45]).

1.2.2 Dynamique du riz en Afrique.

1.2.2.1 La riziculture africaine

O. *glaberrima*, l'une des deux espèces cultivées en Afrique, était à l'origine la seule cultivée sur le continent. Au cours de son évolution naturelle ou sélectionnée, cette espèce a acquis des traits spécifiques aux contraintes biotiques et abiotiques africaines, notamment la compétitivité contre les mauvaises herbes, la tolérance à la sécheresse et aux maladies endémiques. L'introduction d'O. *sativa* au 16e siècle par les Portugais a conduit à une diminution notable de la production d'O. *glaberrima* en Afrique de l'Ouest. D'autres variétés asiatiques au fort potentiel de rendement sont venues s'ajouter en masse au panel de cultivars des fermiers africains entre 1870 et 1960 (Cubry *et al.*, 2018 [19]). Aujourd'hui, de nouvelles variétés hybrides d'O. *glaberrima* et O. *sativa*, notamment les récentes « NERICA » pour NEw RIce for Africa, combinent les traits avantageux des deux espèces et sont très appréciées des producteurs africains (Sié *et al.*, 2012 [41]). Chacun des 4 systèmes de culture (voir 1.2.1.1) sont utilisés en Afrique. Le plus répandu est la riziculture de bas-fonds pluviaux avec 33% des surfaces cultivées en 2013, suivie par la riziculture pluviale d'altitude (30% des surfaces), les cultures irriguées (26% des surfaces), quelques cultures inondables située dans des mangroves ou aux abords de grands fleuves (7%) et enfin, un système dérivé de la riziculture d'altitude, dont la différence se trouve dans des conditions de température plus basses (4% des surfaces) (Zenna *et al.*, 2017[42]).

1.2.2.2 Consommation et enjeux économiques

En Afrique subsaharienne, la consommation de riz n'a cessé de croître, avec un apport par habitant qui a doublé depuis 1970, atteignant dans certains pays le même niveau que les populations asiatiques (Muthayya *et al.*, 2014 [21]). En 2020, le riz a représenté en moyenne 37% des céréales consommées en Afrique, faisant de cette culture un élément de base de l'alimentation. La demande est alimentée par une croissance démographique rapide, l'urbanisation, la croissance des revenus des ménages et par la préférence des citadins pour la facilité de préparation du riz par rapport aux céréales traditionnelles (Seck *et al.* 2013 [43]). Le Nigeria est de loin le plus gros consommateur d'Afrique de l'Ouest avec presque 9 million de tonnes pour l'année 2017. Il est suivi par la Côte d'Ivoire qui a atteint les 7,4 million de tonnes la même année. Viennent ensuite le Mali, le Ghana, le Burkina Faso et le Niger (Tondel *et al.*, 2020 [44] ; Figure 10). Bien que la production de riz sur le continent africain ait augmenté de manière continuelle depuis une quarantaine d'années, passant de 1,76 % de la production mondiale en 1991-2001 à 4,6 % en 2022, la production reste inférieure à la demande. En effet, en 2017 seulement 60 % de la demande de riz fut produite localement (USDA 2022 [38], Zenna *et al.*, 2017[42]). Pour pouvoir répondre aux besoins grandissant de la population, les pays d'Afrique de l'Ouest sont devenus fortement dépendants des importations de riz. Seulement, l'approvisionnement n'est pas fiable et les prix du riz fluctuent, rendant le marché international risqué pour les pays africains. Au cours de l'année 2007-2008 par exemple, un creux dans la production mondiale de riz a conduit à une flambée des prix, ce qui donna lieu à des émeutes dans plusieurs grandes villes africaines (AfricaRice 2020 [45]). Pour surmonter cette insécurité alimentaire chronique, les dirigeants africains souhaitent atteindre l'autosuffisance et faire du riz une activité rentable pour les agriculteurs (Zenna *et al.*, 2017[42]). En 2020, plusieurs pays (dont le Tchad, l'Égypte, le Mali, Madagascar, le Nigeria et la Tanzanie) étaient autosuffisants à plus de 75 % (AfricaRice 2020 [45], Figure 11).

Tableau 4 : principaux stress abiotiques contraignant le riz. (Issu de Pathak *et al.*, 2021 [48]).

Stress	Growth stage	Details of abiotic stress	Decrease in yield (%)
Low temperature	Flowering	<13 °C for 15 days	19-29
	Reproductive and grain filling	<13 °C for 15 days	9.2-26
High temperature	Heading stage	Mean daytime temperature >33 °C	24-27
	Heading stage	Mean daytime temperature >30 °C for 15 days	20-30
	Early grain filling	Mean daytime temperature above 33°C	6.6-13.0
	Growing season	1°C increase in minimum temperature	10
Drought	Flowering	Withholding water at 60 DAT (-30±5 kPa)	23-24
	Vegetative	Withholding water for 24 days from 21 DAT	50.6
	Vegetative/ reproductive	Drying soil (dried beyond -20 kPa)	22.6
	Flowering	Sever stress	>70
	Reproductive stress	Moderate to severe stress (Withholding water-70 kPa at 0.2-m depth)	51-60
	Reproductive stress	Moderate stress	51-57
Salinity	Over growing period	1.0, 3.9, 6.5 dS m^{-1}	> 50 %
	Over growing period	7.4 dS m^{-1}	50%
	Reproductive stage	8.0 dS m^{-1}	>60%
	Reproductive stage	8.0 dS m^{-1}	50-75%
Submergence (Complete and partial)	Seedling stage	Complete submergence with 80 cm of water for 12 days	30-60%
	Vegetative to anthesis Stage	50 cm of water depth by gradual incraese of 10 cm day^{-1} rise in water level	30-70%

1.3 Contraintes écologiques affectant la riziculture.

1.3.1 Contraintes d'origine abiotique

1.3.1.1 Sécheresse

Dans les années 90 en Asie, environ 42 million d'hectares de cultures (1/3 des zones de production) étaient sujettes à des épisodes fréquents de sécheresse, conduisant à des pertes de rendement importantes (Venuprasad *et al.*, 2007 [46]). Avec l'urbanisation, l'industrialisation et les dérèglements climatiques, l'eau douce devient rare et la sécheresse menace d'être le plus grand défi pour la production agricole dans le monde. Pour les riziculture pluviales, la sécheresse ou le manque d'eau sont les défis abiotiques les plus dommageables, un stress hydrique au stade de floraison résultant en des pertes de 24 à 60% (Yang *et al.*, 2019 [47]; Venuprasad *et al.*, 2007 [46]; Pathak *et al.*, 2021 [48], Tableau 4).

1.3.1.2 Température

Selon les prévisions, l'augmentation des gaz à effet de serre élèvera la température moyenne de l'air dans le monde de 1,4 à 5,8 °C entre 1990 et 2100 (Gomez-Zavaglia *et al.*, 2020 [49]). Le stress thermique entrave le développement des plantes en perturbant les processus métaboliques. L'une des phases les plus sensibles aux températures extrêmes chez le riz est la pollinisation, car il suffit de quelques heures de stress thermique pour que les plantes deviennent stériles. Il a également été démontré que des températures élevées pendant la maturation des grains peuvent entraîner une diminution de leur remplissage et une moindre qualité lors de la mouture (Kumar *et al.*, 2021 [50]). Dans les zones de production tempérées, ainsi qu'en haute altitude, les températures basses fournissent un autre stress environnemental qui peut avoir un impact négatif sur les cultures de riz tout au long des phases de germination, de développement végétatif et de reproduction, entraînant des pertes de production entre 9 et 29 % (Pathak *et al.*, 2021 [48], Tableau 4).

1.3.1.3 Inondations

Les cultures de riz inondables aux abords des fleuves sont parfois soumises à un stress de submersion prolongé. Les dérèglements climatiques en cours et à venir vont multiplier les évènements d'inondations. Un stress immersif est plus ou moins intense selon le stade de développement des plantes et la durée de l'immersion, et peut s'avérer fatal pour les cultures. L'immersion partielle ou totale réduit considérablement les échanges gazeux dans tissus foliaires, limitant l'absorption d'oxygène ce qui contraint le métabolisme des sucres à se produire par la voie peu efficace de l'anaérobie (Pathak *et al.*, 2021 [48]). L'inondation par des eaux troubles diminue drastiquement le taux de survie des cultures, ceci étant attribué à une diminution de la lumière disponible et donc à l'impossibilité de photosynthèse (Colmer et Pedersen, 2008 [51]). Avant les travaux de Xu *et al.* en 2006 [22] (voir 1.1.2.1 «Syndrome de domestication et amélioration variétale »), la plupart des méga variétés de riz ne pouvaient survivre aux inondations et n'étaient pas non plus capables de germer en conditions anaérobie.

1.3.1.4 Salinité

La salinité des sols augmente de manière exponentielle avec l'élévation du niveau de la mer dans les zones côtières et dans les terres irriguées, où les sels solubles emportés par les eaux sont ramenés au niveau du sol. Bien que la quantité de terres touchées par la salinité (environ 900 x 10^6 hectares) ne soit pas connue avec précision, leur étendue est suffisante pour constituer une menace pour l'agriculture et la concentration en sels des sols constitue la seconde menace abiotique après la sécheresse (Flowers, 2004 [52]). Près de 40 million d'hectares de bas-fonds pluviaux en Inde, au

Threat	Pest category	Ennemy	Growth stage						
			Germinated seeds	Seedling	Tillering	Maximum tillering	Booting	Flowering	Filling	Ripening
Deadheart, whitehead	insect	Stem Borers (*Lepidoptera, Diptera*)								
Virus vectors, plant destruction and death	insect	Leafhoppers (*Cicadellidae*), planthoppers (*Delphacidae*)								
Galls, sterility	insect	Rice gall midges (Orseolia)								
Plant destruction	insect	Rice leaffolders (*Marasmia, Cnaphalocrocis*)								
Unfilled grains	insect	Grain-sucking insects (Alydidae and Pentatomidae)								
Plant destruction	insect	Rice hispa (*Dicladispa armigera*)								
Plant destruction	insect	Rice water weevil (*Lissorhoptrus oryzophilus*)								
Plant wilting and stunting	insect	Rice thrips (Stenchaetothrips bifrmis, Haplothrips aculeatus)								
Plant destruction	insect	Rice caseworm (*Nymphula depunctalis*)								
Plant destruction	insect	Whorl maggots (*Hydrellia*)								
Deadheart, whitehead, stunting	insect	Rice black bugs (*Scotinophara coarctata, Scotinophara lurida*)								
Stunting	insect	White grubs (*Holotrichia serrata*)								
Stunting, lodging	insect	Rice root aphids (*Tetraneura nigriabdominalis*)								
Stuting, death	insect	Rice root weevils (Curculionidae)								
Yellow syndrome	virus	Ragged stunted virus								
Yellow syndrome	virus	Grassy stunt virus								
Tungro	virus	Rice tungro								
Tungro	virus	Bacilliform virus								
Tungro	virus	Spherical virus								
Seedling blight	bacteria	*Burkholderia plantarii*								
Seedling-grain rot	bacteria	*Burkholderia glumae*								
Bacterial blight	bacteria	*Xoo*								
Bacterial leaf streak	bacteria	*Xoc*								
Grain discoloration	bacteria-fungus	opportunistic fungal and bacterial pathogens								
False smut	fungus	*Ustilaginnoidea virens*								
Sheath blight	fungus	*Rhizoctonia solani*								
Brown spot	fungus	*Bipolaris oryzae*								
Blast	fungus	*Magnaporthe oryzae*								
Bakanae	fungus	*Gibberella fujikuroi*								

<u>Figure 12 : Diagramme d'interaction des différentes menaces biotiques du riz en fonction des stades de développement de l'hôte.</u> Adapté et complété depuis Mew et Vera Cruz, 2018 [30], en ajoutant les données relatives aux insectes de Pathak et Khan 1994 [54]. Orange= présence avérée, gris= présence suspectée.

Bangladesh, au Myanmar et en Thaïlande, ont été ravagés par des crues salées soudaines. Les phases les plus vulnérables au stress salin dans le développement du riz sont ceux de la plantule et de la fécondation, tandis que la germination, le tallage et la maturation y sont relativement résistants (Pathak *et al.*, 2021 [48], Tableau 4).

1.3.2 Contraintes d'origine biotique

Une heatmap regroupant les principales contraintes biotiques du riz et leur période d'apparition est présentée en figure 12.

1.3.2.1 Insectes

On dénombre une centaine d'espèces d'insectes ravageurs du riz, dont une vingtaine responsable de dommages économiques significatifs, de 30% à 100% de pertes de récoltes en rizicultures pluviales d'altitude et de bas-fonds respectivement (Umeh *et al.*, 1995 [53]). On les classe selon la partie de la plante qu'ils détruisent : les foreurs de tiges, les insectes se nourrissant des feuilles, des insectes suceurs (feuilles et tiges) et enfin ceux se nourrissant au niveau des racines. Les espèces causant le plus de dommages appartiennent aux familles des Cicadelles et des Fulgores car également vectrices de virus (Pathak et Khan, 1994 [54]).

1.3.2.2 Virus

On rapporte environ 30 virus pathogènes du riz (Laha *et al*., 2017 [55]). Ils sont transmis à la plante soit par des insectes suceurs, soit par des champignons, ou bien par inoculation mécanique lors de la manutention des cultures. Après l'entrée dans la plante via leurs vecteurs, les virus du riz se propagent à l'ensemble de la plante hôte par le système vasculaire. La répartition géographique diffère selon les virus. Ainsi, on retrouve en Asie une dizaine de maladies virales, comme le virus de la rayure du riz, le virus du riz nain, le virus du rabougrissement du riz, le virus du jaunissement transitoire du riz et le tungro (RTV). En Afrique, parmi les virus signalés, on retrouve le virus de la rayure du riz et le virus de la panachure jaune du riz (RYMV) (Uehara *et al.*, 2013 [56]) Au cours des 20-25 dernières années, le RYMV est devenu une contrainte de production majeure dans les rizicultures irriguées et les rizicultures de bas-fond en Afrique, notamment au Burkina Faso, en Côte d'Ivoire, au Mali, au Niger, au Sénégal et en Sierra Leone (Séré *et al.* 2013 [57]). La maladie entraîne une décoloration jaune à orange des feuilles, un flétrissement des plantes avec un tallage réduit, des défauts de floraison avec un faible rendement des panicules, et dans les cas les plus graves la mort de l'hôte. La perte de rendement en grains due au RYMV varie de 10 à 100 % en fonction de l'âge des plantes au moment de l'infection, du niveau de sensibilité des variétés et des conditions climatiques (Laha *et al.*, 2017[55])

1.3.2.3 Champignons

Parmi les maladies fongiques du riz, les plus destructrices sont la pyriculariose, la brûlure de la gaine du riz, l'Helminthosporiose et le faux charbon du riz. Des dégâts significatifs ont également été enregistré pour des champignons pathogènes émergents responsables de la pourriture du pied et du bakanae (Laha *et al.*, 2017 [55]). La pyriculariose, causée par le champignon *Magnaporthe oryzae* est la maladie la plus dévastatrice du riz, et se retrouve dans toutes les régions productrices (Dean *et al.* 2012 [58]). Le champignon s'attaque aux feuilles, aux ramifications, mais aussi à la panicule. Les dégâts moyens engendrés par la pyriculariose vont de 10 à 30% en conditions classiques et peuvent aller à 100% de perte sur le champ en conditions épidémique (Dean *et al.* 2012 [58]). *Rhizoctonia solani*, agent de la brûlure de la gaine du riz, est répandu dans les rizières des régions tempérées et tropicales du globe,

plus particulièrement dans les cultures irriguées. Les symptômes sur la gaine apparaissent sous forme de lésions grises et aqueuses. Cette maladie est une contrainte forte pour la production de riz aux Etats Unis et en Asie (Laha *et al.*, 2017 [55]). Le faux charbon du riz est une maladie causée par *Ustilaginoidea virens* et qui affecte les ovaires immatures des épillets, les transformant en boules de spores de couleur jaune / vert que l'on appelle des pseudomorphes. Les pertes associées au faux charbon du riz vont de 3 à 81%, dépendant des variétés touchées. En plus des pertes directement liées à la destruction des grains, le champignon produit une toxine qui inhibe la germination des grains de riz, et qui est toxique pour les animaux (Koiso *et al.* 1998 [59]).

1.3.2.4 Nématodes

Environ 150 espèces de nématodes parasites du riz ont été recensées, dont certaines présentes dans la totalité des zones de production du globe (Laha *et al*., 2017 [55]). Les nématodes induisent des pertes significatives, notamment en détruisant les jeunes plants et en étant responsables d'un stress hydrique chez l'hôte qui peut devenir problématique en conditions de sécheresse. On distingue les nématodes foliaires et les nématodes racinaires, la présence des premiers étant plus facilement détectable que celle des seconds. Les nématodes racinaires, à l'origine de galles sur les racines, sont estimés responsables de 15% à 30% de perte de rendement globale sur les cultures de riz (Khan et Ahamad, 2020 [60]). En Afrique de l'Ouest, on retrouve fréquemment plus d'une dizaine d'espèces de nématodes au champ : *Hirschmanniella* et *Tylenchorhynchus* dans les cultures de bas-fond et irriguées, *Tylenchorhynchus*, *Helicotylenchus*, *Scutellonema*, *Meloidogyne*, *Heterodera* et *Pratylenchus* dans les cultures pluviales (Thio *et al.*, 2017 [61]).

1.3.2.5 Bactéries

Il existe différentes maladies bactériennes attaquant le riz. Les pathogènes s'attaquent à diverses parties de la plante : les feuilles, la tige, les panicules ou le pied. Les deux maladies bactériennes d'importance majeure pour la culture du riz sont la bactériose vasculaire et la bactériose à stries foliaires du riz, respectivement Bacterial Leaf Blight (BLB) et Bacterial Leaf Streak (BLS). Ces deux maladies sont dues à deux pathovars de la même espèce bactérienne, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) pour la première et *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) pour la seconde. Parmi les autres maladies bactériennes, on trouve la bactériose des panicules (*Burkholderia glumae*), la rayure brune (*Acidovorax avenae* subsp. *avenae*), la pourriture de la tige (*Pseudomonas fuscovaginae* et *P. syringae* pv. *syringae*) et la pourriture bactérienne du pied (*Dickeya zeae*) (Laha *et al.*, 2017 [55]).

1.3.2.6 Plantes adventices

Les plantes adventices ont un fort impact négatif sur les récoltes. Outre les contraintes que leur présence représente au moment même de la récolte, elles entrent en compétition avec le riz pour les ressources nutritives et la lumière au champ, diminuant alors les rendements et la qualité du riz. En Afrique de l'Ouest une estimation des pertes de rendement liée à ces adventices est de 8 à 15% pour le riz irrigué et de 25 à 30% pour le riz pluvial strict et de bas-fonds (Johnson, 1997 [62]). Les *Poaceae* sont les plus représentées parmi ces plantes compétitrices, suivies des *Cyperaceae*. Les principales compétitrices des cultures de riz sont elles aussi du genre *Oryza*. Il s'agit du riz herbacé ou riz rouge, *O. sativa f.spontanea*, issu de la dé-domestication du riz *O. sativa*. La dé-domestication est un procédé évolutif au cours duquel les espèces domestiquées par l'homme, qu'elles soient animales ou végétales, regagnent des traits sauvages et la capacité de se reproduire indépendamment (Dongya *et al.*, 2021 [63]). Dans le cas de O. *sativa f.spontanea*, ces traits correspondent à une large dispersion des graines, une dormance prolongée et une croissance rapide qui lui permettent d'entrer en compétition forte avec le riz cultivé (Chen *et al.*, 2019 [9]).

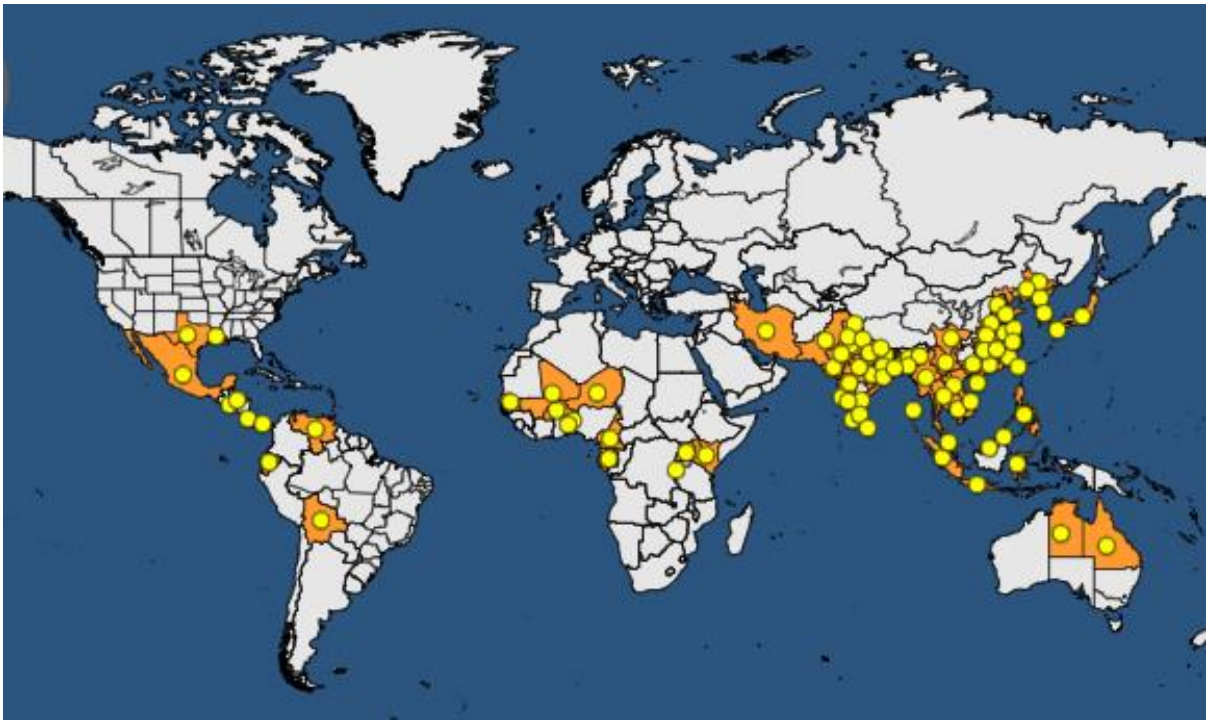


Figure 13 : Distribution géographique de *Xoo*. Les pays en orange avec une pastille jaune relèvent la présence de du pathogène (Issu du site internet de l'EPPO, 2021).

2 LA BACTERIOSE VASCULAIRE DU RIZ

2.1 Distribution et impact pour l'agriculture.

2.1.1.1 Répartition géographique

La bactériose vasculaire (BLB) fût décrite en premier lieu au Japon en 1884 dans la préfecture de Fukuoka sur l'île de Kyushu. Ce n'est qu'en 1922 que l'agent pathogène responsable de la maladie, *X. oryzae* pv. *oryzae*, ou *Xoo,* (à l'époque *Pseudomonas oryzae*) fût identifié par Ishiyama et Uyeda. Pendant longtemps confondue avec la BLB, la BLS ou bactériose à stries foliaires du riz a été distinguée en 1957 dans le sud de la Chine (Niño-Liu *et al.*, 2006). Cette maladie est due à une bactérie d'un pathovar proche de *Xoo*, appelée *Xanthomonas oryzae* pv. *orizycola* (ou *Xoc*), qui colonise les tissus du parenchyme. Rapidement, la BLB fut également décrit en Corée (Takeuchi, 1930 [64]). A partir des années 1960 et suite à l'introduction de variétés à hauts rendements mais fortement sensibles à la maladie, la BLB s'est répandu de l'Asie de l'Est vers l'Asie du Sud et du Sud-Ouest : en Malaisie, en Indonésie, au Cambodge et à Taiwan (IRRI, 1989 [65]; Naqvi, 2019 [66]). La maladie a été signalée pour la première fois en Afrique de l'Ouest par Buddenhagen *et al.* en 1979 [67] au Mali. Par la suite, elle s'est répandue au Cameroun, au Sénégal, en Mauritanie, en Gambie, au Burkina Faso, en Guinée, au Ghana, au Nigeria et dans certains pays d'Afrique de l'Est (Reckhaus, 1983 [68]; Séré *et al.* 2013 [57]). La BLB a été observée au Nord de l'Australie pour la première fois en 1973 et aux États Unis en 1977 (Aldrick *et al.*, 1973 [69]; Lozano, 1977 [70]). Aujourd'hui la bactériose vasculaire du riz sévit principalement en Asie, en Afrique et en Australie (Jeger *et al.*, 2018 [71]). Sa distribution connue la plus récente selon l'EPPO est représentée sur la carte en Figure 13 (EPPO, European and Mediterranean Plant Protection Organization, 2022).

2.1.2 Incidence et pertes de rendement

Bien que la bactériose vasculaire du riz soit dispersée aux quatre coins du globe, son impact sur les cultures n'est économiquement dommageable qu'en Asie et en Afrique, particulièrement dans les cultures irriguées ou pluviales de bas-fonds des zones tropicales. Plus l'infection a lieu à des stades précoces de la croissance de la plante, plus les symptômes sont violents (Jeger *et al.*, 2018 [71]). L'intensité de la maladie dépend également de la variété cultivée, les variétés naines à haut rendement introduites depuis les années 1960 étant bien plus sensibles à la maladie. Entrent également en jeu les conditions météorologiques (les saisons humides étant propices à l'infection) et la fertilité des sols. En effet, l'introduction d'engrais azotés se révèle favorable au développement du BLB (Naidu *et al.*, 1979 [72]). Les pertes liées au BLB sont dues à une augmentation du nombre d'épillets stériles, une diminution du rendement et de la qualité des grains et dans certains cas à la mortalité des plants (Reddy *et al.*, 1989 [73] ; Mew, 1993 [74]). Des études menées sur des champs en Corée de 2002 à 2005 ont montré que les pertes de rendement étaient plus importantes si l'infection avait lieu au stade d'initiation de la panicule. A ce stade, un champ avec une incidence de 50 % de la BLB a montré 29% de pertes de rendement (contre 17% si l'infection avait lieu au début de la maturation des grains), auxquelles s'ajoute une diminution d'environ 18% de la quantité de grains de bonne qualité, c'est-à-dire non-endommagé. La qualité gustative du riz issu des champs contaminés a également été affectée (détérioration du goût, diminution de la viscosité) (Noh *et al*., 2007 [75]).

En zones tropicales des pertes de 2 à 74% ont été rapportées (Laha *et al.*, 2017 [55]). Au Japon, dans les années 1980, les dommages imputés à la BLB étaient de 60% (Ou, 1985 [76]). Sur des variétés sensibles aux Philippines, une diminution des récoltes de 7,2 à 24,5% fut enregistrée en 1973, puis jusqu'à 50 %

Tableau 5 : catégorisation en quarantaine de *Xoo* selon différentes régions et RPPOs.

Continent ou RPPO	Nom de la RPPO	Nombre de pays membres RPPO	Pays	Liste	Année d'ajout
Afrique	NA	NA	Afrique de l'Est	A1	2001
Afrique	NA	NA	Egypte	A1	2018
Afrique	NA	NA	Maroc	Organisme de quarantaine	2018
Amérique du Sud	NA	NA	Argentine	A1	2019
Amérique du Sud	NA	NA	Brésil	A1	2018
Amérique du Sud	NA	NA	Mexique	Organisme de quarantaine	2018
Amérique du Sud	NA	NA	Paraguay	A1	1992
Amérique du Nord	NA	NA	Etats Unis d'Amérique	Organisme de quarantaine	1994
Amérique du Sud	NA	NA	Uruguay	A1	1992
Asie	NA	NA	Kazakhstan	A1	2017
Asie	NA	NA	Ouzbekistan	A1	2008
Europe	NA	NA	Azerbaijan	A1	2007
Europe	NA	NA	Turquie	A1	2016
Europe	NA	NA	Ukraine	A1	2019
COSAVE	COMITE DE SANIDAD VEGETAL DEL CONO SUR	6	Argentine, Brésil, Chili, Paraguay, Pérou, Uruguay	A2	2018
EAEU	EURASIAN ECONOMIC UNION	5	Armenie, Biélorussie, Kazakhstan, Kyrgyzstan, Russie	A1	2016
EPPO	EUROPEAN PLANT PROTECTION ORGANIZATION	52	Pays du continent Eurasiatique sans le Moyen et le Proche Orient, la Chine, l'Inde, le Japon et l'Asie du Sud Ouest	A1	1975
EU	EUOPEAN UNION	27	Pays de l'Union Européenne	A1	2019
IAPSC	INTER-AFRICAN PHYTOSANITARY COUNCIL	55	Tout le continent Africain sauf le Sahara Occidental	A2	1989

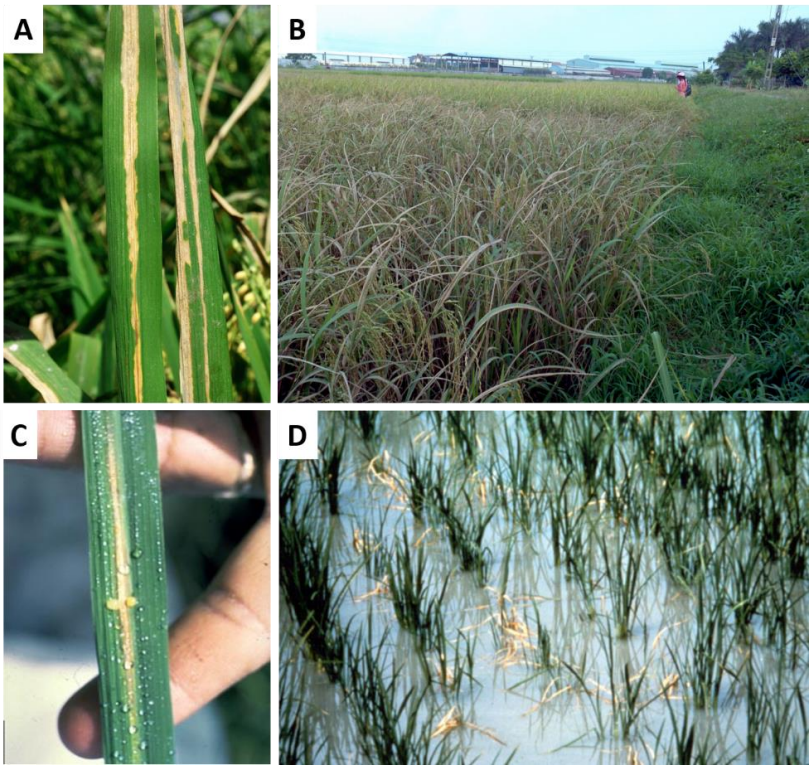


Figure 14 : Symptômes causé par *Xoo* sur le riz. A. Symptômes typiques du BLB sur des feuilles de riz (photo, IRRI Rice Knowledge Bank, 2022). B. Champ de riz infecté par *Xoo*. Les plantes malades se repèrent par leur couleur jaunâtre et leur aspect sec et vrillé (photo, courtoisie de F. Auguy, 2018). C. Gouttelettes d'exsudats bactériens visibles au matin sur de jeunes feuilles de riz infecté (photo, CABI, Plantwise Knowledge Bank, 2022). D. Nurserie atteinte de Kresek. L'attaque des jeunes plants de riz par *Xoo* après transplantation donne lieu à des symptômes violents, souvent fatals pour l'hôte. Cette manifestation de *Xoo* est appelée Kresek (Niño-Liu *et al*., 2006 [84]).

en 1993 (Exconde *et al*., 1973 [77]; Mew *et al.* 1993 [74]). De déficits à hauteur de 82% ont eu lieu en Inde et au Bangladesh dans les années 1980-1990. Adhikari et Mew en 1991 [78] ont consigné 26 % de pertes dues au BLB au Népal. En Afrique, la BLB est également devenue un problème majeur pour les cultures. En 1982 au Niger, une perte de la totalité des récoltes a été déplorée sur un cultivar chinois nain à haut rendement (Reckaus 1983 [68]). En moyenne en Afrique de l'Ouest 30 à 90 % de pertes sont incriminées au BLB (Sere *et al.* 2005 [79]; Tall *et al.*, 2019 [80]). Enfin, bien que la maladie ait été observée sur le continent Américain, elle ne cause pas de dommages notables aux récoltes (Corral *et al.* 2013 [81]).

2.1.3 Mesures de quarantaine

Les réglementations en matière de protection phytosanitaire sont régies par des organisations intergouvernementales appelées RPPOs (Regional Plant Protection Organizations) qui font office d'organe de coordination des organisations nationales de protection des végétaux (FAO 2022). L'EPPO (European and Mediterranean Plant Protection Organization) classe les organismes à risques selon 3 niveaux :

- Organismes de quarantaine : organismes constituant une menace économique pour les cultures et pas encore présents sur le territoire, ou présents mais peu répandus et faisant l'objet d'une lutte officielle (selon la définition de la FAO).
- Organismes A1 : absents de la région et recommandés pour une régulation en quarantaine EPPO.
- Organismes A2 : localement présents dans la région mais recommandés pour une régulation en quarantaine EPPO.

Xoo a été classé dans la liste A1 de l'EPPO en 1975. Bien que le pathogène ne se retrouve pas sur le territoire de l'EPPO (voir pays membres en tableau 5), il a été annoncé par la commission qu'il représentait un risque concret de persistance dans les pays méditerranéens. L'EFSA (European Food Safety Authority) a établi dans un rapport rédigé en 2018 que *Xoo* est un organisme nuisible dont l'introduction et la propagation au sein des États membres Européens via des plantes ou parties de plantes infectées est interdite (Jeger *et al.* 2018 [71]). Les recommandations de l'EPPO sont d'éviter l'importation de semences provenant de pays infestés. Cependant cette mesure est difficile à respecter et une inspection des grains avant et après importation est fortement recommandée (EPPO, CABI 2002 [82]). Les végétaux et parties de végétaux doivent être soumis à une inspection phytosanitaire sur le lieu de production avant d'être déplacés ou importés sur le territoire (EPPO, 2021 [83]). D'autres RPPOs comme le COSAVE, l'EAEU et l'IAPSC ont considéré *Xoo* comme un phytopathogène d'importance de quarantaine (EPPO, voir tableau 5).

2.2 Diagnostic.

2.2.1 Symptômes

2.2.1.1 Blight

Les symptômes les plus fréquents développés par les plants de riz infectés sont ceux de la brûlure des feuilles, ou *blight* en anglais. Ces symptômes sont visibles au stade d'apparition des talles. Ils débutent par l'apparition de taches foncées, aqueuses au niveau de la pointe et des marges de la feuille. La propagation des symptômes le long de la nervure centrale de la feuille de riz conduit à la formation de lésions nécrotiques jaunâtres qui finissent par gagner toute la feuille, causant son flétrissement (Mew, 1993[74] ; Ou, 1985 [76]) (Figure 14). Ces feuilles sont alors souvent attaquées par des champignons

saprophytes. Il arrive que les lésions démarrent loin des bords de la feuille, indiquant que l'entrée du pathogène s'est faite par une blessure (Mew *et al.*, 2018 [85]) En début de journée en zones tropicales, des gouttelettes opaques et jaunâtres peuvent être observées percolant au travers des jeunes lésions. Leur assèchement en cours de journée les transforme en petites billes, pouvant être une source d'inoculum secondaire (Mizukami and Wakimoto 1969 [86], figure 14C). La maladie peut également apparaître sur les panicules sous forme de taches brunes pâles sur les glumes, conduisant à une infertilité et une qualité médiocre des grains (Fang *et al.*, 1956 [87]).

2.2.1.2 Kresek

Le Kresek est une atteinte destructrice des jeunes plants par *Xoo*, décrite pour la première fois en 1950 par Reitsma et Shure [88]. Les symptômes du Kresek apparaissent sur les plants de riz entre 1 et 4 semaines après transplantation. Les feuilles s'enroulent complètement, deviennent jaunes ou grises ou tombent et finalement les talles se dessèchent (Reddy, 1984 [89]). S'ils ne meurent pas, les plants de riz deviennent jaunes et le tallage peut être avorté (figure 14D, voir « Pale-yellow leaf »). Cette forme de bactériose est provoquée par l'entrée de la bactérie au niveau de la partie basale de la gaine des feuilles, juste au-dessus des racines de l'hôte. Cette entrée est rendue possible par la présence de blessures, induites par la transplantation des plantes ou par des prédateurs comme le charançon du riz (Noda et al 1981 [90]). L'apparition du Kresek est favorisée par la pratique de la coupe des feuilles des jeunes plants avant le repiquage.

2.2.1.3 Pale-yellow leaf

Le jaunissement des feuilles ou « Pale yellow leaf » est un symptôme plus rare et de nature secondaire dont l'apparition n'a été rapportée qu'aux Philippines (Laha *et al.*, 2017 [55]). Ce jaunissement est visible sur les feuilles nouvellement formées de plantes ayant surmonté le blight ou le Kresek. Bien qu'aucune *Xoo* ne puisse être isolée de ces feuilles, des bactéries sont bien présentes dans la plante. Elles se concentrent au niveau des couronnes des tiges. Cette accumulation, par effet de blocage ou par la production de toxines (ces deux hypothèses n'ayant pas encore été confirmées), conduit à une chlorose des limbes des feuilles (Mew *et al.*, 2018 [85], figure 14C).

2.2.2 Outils moléculaires de diagnostic

2.2.2.1 Tests de culture

Les tests d'identification de de *X. oryzae* (ou *Xo*) reposant sur l'isolement et la culture de la bactérie sont les premiers à avoir été développés. Ils sont basés sur l'inoculation de plantes ou l'ensemencement de milieux semi-sélectifs depuis des macéras de feuilles infectées coupées, ou des « lavages » de graines suspectes. Ces tests sont longs à réaliser et peu sensibles, notamment car les bactéries se développent lentement et sont facilement recouvertes par la flore saprophyte. Par ailleurs, le nombre de bactéries présentes sur les semences est très faible (Cottyn *et al.*, 1994 [91]).

2.2.2.2 ELISA

Des anticorps monoclonaux produits en 1989 par Benedict et al [92] ont été largement utilisés pour la détection spécifique de *Xo* (Corral *et al.* 2013 [81]). Ces anticorps permettent la différenciation de *Xo*, et du pathovar *Xoo*, mais le 3e anticorps n'est que faiblement sensible au pathovar *Xoc*. Bien que hautement spécifiques, le manque de sensibilité des anticorps et leur difficulté de production rendent leurs applications limitées dans l'identification des souches, surtout à grande échelle.

2.2.2.3 Tests par amplification d'ADN

Les méthodes basées sur l'amplification d'acides nucléiques offrent une plus grande sensibilité et un temps de réponse plus court que les tests de cultures. Les premiers outils de diagnostics moléculaires pour l'étude phylogénétique des bactéries étaient basés sur l'amplification et le séquençage de l'ADN de la sous-unité 16S du ribosome (Lane, 1991 [93]). Cependant, ces approches 16S ne permettent pas l'identification des différents pathovars de *X. oryzae*, et ne permettent pas de caractériser une collection de souches d'origines diverses (Corral *et al.*, 2013 [81]). En 2010, Lang et ses collègues [94] ont mis au point des conditions et amorces adaptées à une réaction de PCR multiplexe permettant de discriminer *Xo*, *Xoo* et *Xoc* de manière spécifique. Cette technique simple à mettre en place est aujourd'hui utilisée avec succès dans différents centres de recherche, dont l'IRD.

2.2.2.4 LAMP-PCR

La LAMP-PCR ou « Loop Mediated isothermal Amplification » est une technique qui permet d'amplifier de grandes quantités d'ADN à partir d'un échantillon très peu concentré en ADN cible. En général, quatre amorces sont utilisées pour amplifier différentes régions du gène cible via une polymérase qui possède une activité de déplacement de brin supérieure aux polymérases de routine. La réaction se déroule en une heure, dans un seul tube, à une température unique comprise entre 60 et 65°C. L'amplification de la cible s'observe via des marqueurs colorés ou fluorescents (Notomi *et al.*, 2000 [95]). La LAMP est une technique moins sensible aux inhibiteurs que la PCR et elle ne nécessite que très peu d'équipements. Lang *et al.* en 2014 [96] ont mis au point un protocole de détection colorimétrique par LAMP-PCR des différents pathovars de *X. oryzae*. La réaction peut se pratiquer sur de l'ADN, des bactéries ou directement sur des échantillons de feuilles et de graines avec une sensibilité jusqu'à 104 à 105 UFC ml^{-1}. Cette technique a été utilisé avec succès par l'IRD sur le terrain lors de campagnes d'échantillonnage.

2.3 Épidémiologie.

2.3.1 Colonisation de la plante

Xoo pénètre la plante via des pores appelés hydathodes situés à l'extrémité des feuilles, ou directement via des blessures induites par des prédateurs, des évènements météorologiques violents (par collisions entre plantes), ou aux pratiques agricoles. Fang *et al.* ont pu prouver en 1956 [87] que la bactérie n'était pas capable de rentrer par les stomates, mais qu'elle profitait de l'aspiration des fluides de guttation par les hydathodes pour envahir l'hôte. Une fois entré, le pathogène se multiplie dans l'épithème (chambre de guttation en amont de l'hydathode dans laquelle débouchent les vaisseaux du xylème), puis gagne les tissus vasculaires et se répand dans toute la plante. Il a été montré par Noda et Kaku en 1999 [97] que les bactéries ne se multiplient que dans les lumens des vaisseaux du xylème et n'étaient pas retrouvées dans les autres éléments des vaisseaux. Ils ont par ailleurs pu observer que la surface des vaisseaux colonisés était dégradée, correspondant à la digestion des parois par la bactérie, libérant ainsi des composants potentiellement utilisables pour sa croissance. *Xoo* se déplace dans la feuille par les veines primaires mais progresse également dans les veines latérales. En quelques jours, les cellules bactériennes et les exopolysaccharides (EPS) remplissent les faisceaux vasculaires (Nino-liu *et al.*, 2006 [84]).

2.3.2 Sources d'inoculum

2.3.2.1 Depuis les plantes hôtes

Les sources d'inoculum majeures sont les plantes hôtes de *Xoo*, la principale étant le riz infecté. La présence de gouttelettes d'exsudats contenant des bactéries à la surface des feuilles de riz (voir paragraphe 2.2.1.1) permet la contamination de proche en proche par contact. Les débris de plantes infectées, comme les chaumes laissés au sol ou les feuilles près des aires de battage par exemple, sont des abris où la bactérie peut survivre jusqu'à la saison suivante de plantation, et constituent une source de contamination importante (Devadath & Dath, 1985 [98]). Les hôtes alternatifs de *Xoo* sont également des réservoirs infectieux non négligeables dans l'apparition du BLB. Il a notamment été montré que la bactérie pouvait hiverner en zones tropicales dans la rhizosphère d'autres *Oryzeae* des genres *Leersia* et *Zizania*, permettant au pathogène de persister toute l'année sur le terrain (Mizukami et Wakimoto, 1969 [86]).

2.3.2.2 Depuis le substrat

La survie de la bactérie dans le sol non - rhizosphérique est soumise aux conditions de température, d'humidité, de pH et la présence de prédateurs, et varie entre 1 et 3 mois (Nino-Liu *et al.* 2006 [84]). Des concentrations bactériennes élevées ont été enregistrées dans les eaux d'irrigation, bien que l'action naturelle de bactériophages puisse limiter ces populations (Ou, 1985 [76]).

2.3.2.3 Depuis les graines

Les semences collectées depuis des champs lourdement infectés sont fréquemment porteuses de la bactérie. En effet, il est facile d'isoler le pathogène depuis les glumes et l'endosperme des graines issues de panicules infectées (Fang *et al.*, 1956 [87] ; Sakthivel *et al.*, 2001 [99]). En revanche, la transmission par les graines du BLB a été sujette à controverse. Murty & Devadath en 1984 [100] n'ont pas observé l'apparition de plantules infectées après semis de graines contaminées, mais ont pu détecter la diffusion du pathogène dans le sol. Singh *et al.* en 1983 [101] ont observé une fréquence élevée de transmission depuis les graines en chambre de culture. Enfin, en 2019 et Mondal [102] et ses collègues ont observé : i) la capacité de survie de *Xoo* sur graines pendant au moins 10 mois ; ii) l'apparition de symptômes sur les jeunes plantules issues de semences infectées en pots et au champs ; iii) la détection de *Xoo* dans les graines récoltées à partir de ces plants. Bien que la transmission par les semences ne joue pas forcément un rôle important sous les tropiques là où la maladie est déjà établie, elle peut être très importante dans les pays où la maladie n'a pas encore été enregistrée (Mew, 1989 [65]).

2.3.3 Dissémination

2.3.3.1 Vecteurs naturels

Le vent, la pluie et les évènements météorologies violents comme les typhons sont autant de moyens de propagation de la maladie. Ils contribuent au contact entre plantes infectées et plantes saines, à la formation de blessures sur les plants tenant lieu d'ouvertures pour *Xoo*, mais aussi au déplacement direct de la bactérie (Corral *et al.*, 2013 [81]). Ces déplacements se font au sein du champ, ou jusqu'au champ voisin, l'amplitude du déplacement via les éclaboussures de pluie n'ayant été estimée qu'à quelques mètres (Mew *et al.* 2018 [85]). Les bactéries peuvent également être disséminées par des insectes ou des oiseaux, ou encore se répandre dans l'eau d'irrigation du champ inondé (Ou, 1985 [76]). En revanche, la durée de vie des cellules de *Xoo* libres (c'est-à-dire en dehors des feuilles) dans les eaux d'irrigation est limitée.

Tableau 6 : Liste EPPO 2022 des plantes hôtes de *Xoo*.

Host type	Species name
Major host	*Oryza sativa*
Wild/Weed	*Cenchrus ciliaris*
Wild/Weed	*Cynodon dactylon*
Wild/Weed	*Cyperus difformis*
Wild/Weed	*Cyperus rotundus*
Wild/Weed	*Echinochloa crus-galli*
Wild/Weed	*Leersia hexandra*
Wild/Weed	*Leersia japonica*
Wild/Weed	*Leersia oryzoides*
Wild/Weed	*Leersia sayanuka*
Wild/Weed	*Leptochloa chinensis*
Wild/Weed	*Leptochloa mucronata*
Wild/Weed	*Megathyrsus maximus*
Wild/Weed	*Oryza australiensis*
Wild/Weed	*Oryza glaberrima*
Wild/Weed	*Oryza longistaminata*
Wild/Weed	*Oryza rufipogon*
Wild/Weed	*Paspalum distichum*
Wild/Weed	*Paspalum scrobiculatum*
Wild/Weed	*Urochloa mutica*
Wild/Weed	*Zizania aquatica*
Wild/Weed	*Zizania latifolia*
Wild/Weed	*Zizania palustris*
Wild/Weed	*Zoysia japonica*

2.3.3.2 Pratiques culturales

La gestion des cultures par l'homme peut être elle aussi vectrice de la maladie, ne serait-ce que par le déplacement des producteurs au sein du champ, pouvant emmener les exsudats d'une feuille infectée vers un tissu foliaire sain (Corral *et al.* 2013 [81]). De la même manière, l'utilisation de matériel agricole contaminé, notamment lors de la transplantation des plantules, est propice à la propagation du BLB (CABI et EPPO, 2002 [82]). La gestion des eaux d'irrigation, pouvant traverser plusieurs champs, est elle aussi à l'origine du déplacement du pathogène sur de courtes distances. La dissémination de la maladie sur de longues distances se fait par le déplacement de matériel biologique ou agricole, et potentiellement par les semences de riz infectées (Mew *et al.*, 2018 [103]).

2.3.4 Hôtes alternatifs

En plus d'*O. sativa*, son hôte principal, *Xoo* a été recensé sur plus d'une vingtaine d'espèces de plantes. Ces dernières correspondent soit à des riz sauvages comme O. *rufipogon*, O. *longistaminata* ou O. *australiensis*, soit à des plantes herbacées de la famille des *Poaceae* non cultivées (voir tableau 6, EPPO 2022 [104]). Ces plantes hôtes sont souvent retrouvées associées aux zones rizicoles et peuvent servir de réservoir ou de relais pour la maladie.

2.4 Méthodes de lutte.

2.4.1 Intrants

2.4.1.1 Composés antimicrobiens

Le contrôle de la bactériose vasculaire du riz et de *Xanthomonas* en général via des méthodes chimiques a une efficacité limitée. Les mixtures à base de cuivre sont les plus efficaces pour les maladies bactériennes et cette technique a été utilisée dans le traitement du BLB depuis une centaine d'années. Seulement, les effets nocifs de l'épandage du cuivre dans l'environnement et l'apparition rapide de bactéries résistantes rendent ces techniques obsolètes. En effet, l'acquisition rapide de gènes de résistance au cuivre via les transferts de gènes horizontaux chez *Xanthomonas* est largement documentée (Assis *et al.*, 2021 [105] ; Huang *et al.*, 2021 [106]). Certains antibiotiques, notamment la streptomycine, ont été utilisés depuis les années 1950 pour le traitement des maladies bactériennes dans diverses cultures. Cependant, l'application de ces antibiotiques n'est pas efficace sur toutes les souches et est rapidement suivie par l'apparition d'individus résistants (An *et al.*, 2019 [107]). Depuis une vingtaine d'années, des études ont pu démontrer l'efficacité de peptides antimicrobiens dans le contrôle des maladies bactériennes des plantes, notamment contre différentes espèces de *Xanthomonas* (Cirac *et al.*, 2017 [108]; Inui Kishi *et al.*, 2018 [109]; Camo *et al.*, 2018 [110]; Li *et al.*, 2021 [111]). Bien que ces peptides soient efficaces, ils ne sont malheureusement pas très spécifiques, et ont un effet nocif à large spectre bactérien, dégradant de ce fait la flore bactérienne commensale et phytobénéfique autour des cultures. Des composés issus de microorganismes ont montré des effets prometteurs pour maîtriser *Xoo*. L'anthranilamide et la carbazomycine B isolées de deux espèces différentes de *Streptomyces* par exemple permettent de freiner respectivement la formation de biofilms et la production de xanthomonadine et d'EPS chez *Xoo* (An *et al.*, 2019 [107]). La chumacine-1/2 extraite de *Pseudomonas aeruginosa* entrave la perception des signaux du quorum sensing de la bactérie (Kanugala *et al.* 2019 [112]), ce qui va dans le sens de premiers résultats d'antagonismes obtenus par Yasmin et al en 2017 [113]. Enfin, des molécules issues de plantes comme le resvératrol et le catéchol ont donné des résultats probants *in vitro* dans le contrôle de *Xoo*, le premier inhibant la formation de biofilm et le second créant un stress oxydatif empêchant la croissance de la bactérie (Sanya *et al.*, 2022 [114]).

2.4.1.2 Agents de biocontrôle

Streptomyces toxytricini VN08-A-12 a montré des effets antagonistes sur au moins 10 races de *Xoo* isolées au Vietnam. Les résultats ont montré que VN08-A-12 était non seulement capable de réduire la longueur des lésions dues à *Xoo* chez deux cultivars de riz, mais qu'elle réduisait également de manière significative la perte de rendement liée à la maladie, tout en augmentant la productivité des plants de riz sains (Hop *et al.*, 2014 [115]). Une autre bactérie, *P. aeruginosa* a montré un antagonisme *in vitro* contre différents phytopathogènes, y compris *Xoo* et *Fusarium* spp. Utilisée comme agent de biocontrôle dans un essai au champ, cette souche a augmenté les rendements en grains et en biomasse de plus de 50% par rapport au contrôle non inoculé (Yasmin *et al.*, 2017 [113]). D'autres bactéries ont été signalées comme exerçant un biocontrôle par l'induction des défenses de la plante hôte. Par exemple, des pulvérisations foliaires répétées de *P. fluorescens* ont permis de protéger des cultures de riz contre la BLB, et ont amélioré les rendements (Vidhyasekaran *et al.* 2001 [116]).

Il a également été rapporté qu'une culture saturée en différentes amibes (*Acanthamoeba castellanii*, *Acanthamoeba lenticulata*, *Acanthamoeba polyphaga*, *Dictyostelium discoideum* et *Vermamoeba vermiformis*) présente des effets bactériostatiques et bactéricides lorsqu'elle était appliquée sur des cultures de *Xoo* (Sanya *et al.*, 2022 [114]).

Sur les deux dernières décennies, l'émergence de résistances aux composés antibactériens ainsi qu'un intérêt pour l'utilisation des bactériophages comme alternative à l'épandage de produits chimiques ont remis l'étude de ces virus au goût du jour en phytopathologie. Ubiquitaires et hautement spécifiques de leur hôte, l'utilisation de bactériophages s'est effectivement révélée efficace dans la suppression des maladies causées par *Xanthomonas* sur les tomates, les agrumes les oignons et le riz (An *et al.*, 2019 [107] ; Chae *et al.*, 2014 [117]). Bien que prometteurs, la faible stabilité des bactériophages, notamment leur sensibilité aux UVs, limite leur l'application sur le terrain.

2.4.2 Prévention

2.4.2.1 Bonnes pratiques de culture

La prévention de l'apparition et de la propagation de la BLB passe par certaines pratiques culturales. Parmi elles il y a le désherbage des champs, ayant pour but de supprimer la présence d'éventuels hôtes secondaires pouvant servir de relais et de réservoir au pathogène. Dans la même optique, un espacement plus large des plantations vise à enrayer la propagation de la maladie au sein du champ, en évitant les contacts entre les plantes. La limitation des sources d'inoculum via la destruction des débris végétaux (feuilles et chaumes), la désinfection du matériel agricole, et l'utilisation unique des eaux d'inondation via drainage font également partie des bonnes pratiques pour le contrôle de l'infection. Limiter au maximum les blessures, notamment sur les jeunes plants lors de la transplantation, est aussi une pratique favorable à la santé des cultures. L'utilisation de semences ne provenant pas de champs ayant été infectés par *Xoo* a fait ses preuves en Chine entre 1950 et 1960 (Fang et al 1956 [87]). La désinfection des semences par traitement à l'eau chaude, des antibiotiques ou des composés à base de mercure (ces deux derniers étant néanmoins interdits en Europe) est pratiquée, plus particulièrement en Asie. Enfin, la gestion raisonnée des fertilisants azotés et phosphorés permet de limiter le développement de la maladie, l'augmentation de la croissance végétative des plants lui étant favorable (Jeger *et al.*, 2018 [71]; Mew *et al.*, 2018 [103]).

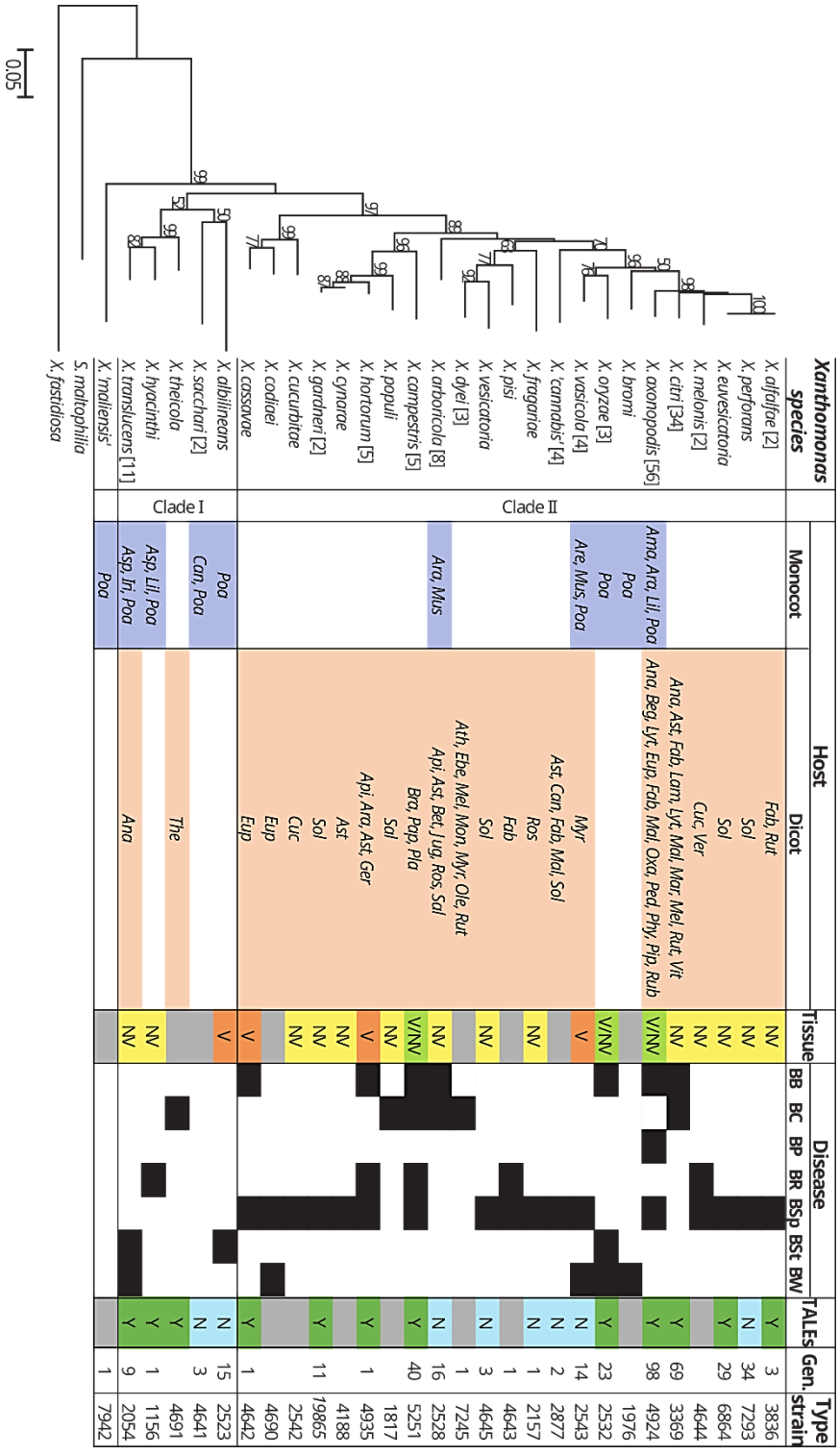


Figure 15 : Phylogénie, gamme d'hôtes et maladies associées au genre *Xanthomonas*. Phylogénie basée sur la séquence *gyrB* partielle de souches modèles représentant la diversité d'espèces du genre *Xanthomonas*. Les chiffres entre parenthèses à côté des noms d'espèces indiquent le nombre de sous-espèces ou de pathovars décrits. Les principales familles de plantes hôtes, monocotylédones et/ou dicotylédones sont iniquées. La spécificité tissulaire, lorsqu'elle est connue est mentionnée. V, vasculaire ; NV, non vasculaire. Maladies engendrées : BB, bactériose ; BC, chancre bactérien ; BP, pustule bactérienne ; BR, pourriture bactérienne ; BSp, tache bactérienne ; BSt, strie bactérienne ; BW, flétrissement bactérien. Présence de TALEs (Transcription Activator-Like Effectors) : Y, oui ; N, non. (Tiré de Jacques *et al*., 2016[119]).

2.4.2.2 Surveillance

La surveillance soutenue de l'apparition des premiers symptômes sur les jeunes plants et leur élimination constitue une étape cruciale dans la lutte contre la dispersion du pathogène. Il est difficile de prévoir l'apparition de la maladie au champ, car elle dépend notamment des variétés de riz utilisées et des conditions climatiques. Cependant, la surveillance météorologique afin de repérer l'arrivée de conditions épidémiques, ou des relevés de la concentration en bactériophages dans les eaux d'irrigation renseignant sur une augmentation des populations bactériennes au champ permettent de renforcer la prévention, à défaut de traitement efficace (Niño-Liu *et al.*, 2006 [84]).

2.4.2.3 Résistances génétiques

L'utilisation de variétés résistantes est un moyen simple, économique et écologique de protéger les cultures. La résistance génétique est à ce jour la méthode la plus efficace pour lutter contre l'infection par *Xoo* (Corral *et al.*, 2013 [81]). Il existe d'ores et déjà de nombreuses sources de résistance monogéniques caractérisées et disponibles chez le riz (voir paragraphe 4.2.3). Ces gènes ont principalement été identifiés à partir de cultivars O. *sativa* ssp. *indica*, mais aussi de variétés *japonica*, et d'espèces sauvages apparentées, notamment O. *longistaminata*, O. *rufipogon*, O. *minuta* et O. *officinalis* (Niño-Liu *et al.*, 2006 [84]). Afin de maximiser la protection fournie par ces gènes de résistance, leur déploiement doit être pensé en amont, pour ne pas favoriser la survenue d'une rupture de résistance.

3 L'AGENT PHYTOPATHOGENE *XANTHOMONAS ORYZAE* PV. *ORYZAE*

3.1 Classification.

3.1.1 Le genre *Xanthomonas*.

Le genre *Xanthomonas* fait partie de la classe des Gammaprotéobactéries, l'un des phylums les plus fournis dans le monde bactérien, regroupant plus de 250 genres (Williams *et al.*, 2010 [118]). Les *Xanthomonas* représentent 29 espèces bactériennes (voir figure 15), principalement associées aux plantes et capables d'infecter plus de 124 espèces de monocotylédones et 268 espèces de dicotylédones (Jaques *et al.* 2016 [119]). Parmi les hôtes de *Xanthomonas*, on retrouve de nombreuses plantes d'importance agronomique majeure, comme le bananier, le coton, les poivrons, les tomates, les *Brassicaceae* et de nombreux agrumes. Bien que les *Xanthomonas* soient ubiquitaires chez les végétaux, le spectre d'hôtes de chaque espèce est généralement limité à une plante en particulier (Pruvost *et al.*, 2015 [120]). Par ailleurs, des souches différentes d'une même espèce de *Xanthomonas* peuvent être à l'origine de symptômes différents chez la même plante hôte. Cette diversité au niveau intraspécifique est distinguée par l'utilisation du terme de « pathovar » afin d'identifier précisément le pathogène en question. Ainsi, le concept de pathovar ne repose pas uniquement sur la gamme d'hôtes mais également sur la spécificité d'infection des tissus et la nature des symptômes provoqués. Par exemple, *Xoo* est un pathogène vasculaire du riz, alors que *Xoc* infecte le parenchyme des feuilles de riz (Jaques *et al.* 2016 [119], Figure 15).

3.1.2 L'espèce *Xanthomonas oryzae*

Anciennement rattachés à l'espèce *X. campestris*, les différents pathovars associés au riz ont été reclassés par Swings *et al* en 1990 [121] sur la base de plusieurs critères moléculaires et biochimiques. *X. oryzae* compte quatre sous-espèces, ou pathovars : *X. oryzae* pv. *oryzae* (*Xoo*), *X. oryzae* pv. *orizycola* (*Xoc*), *X. oryzae* (*Xo*) et *X. oryzae* pv. *leersiae* (*Xol*) (Lang *et al.*, 2019). *Xoo* et *Xoc* sont deux pathovars

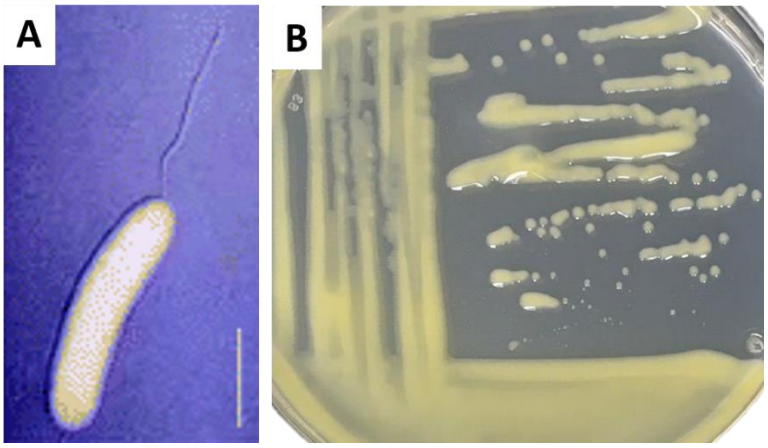


Figure 16 : Morphologie de *Xanthomonas*. A. Micrographie électronique à balayage de cellules de *Xoo* (barre d'échelle, 1.0 µm) (Photo, Niño-Liu et al., 2006 [84]). B. Colonies de *Xanthomonas* sur milieu nutritif solide.

Tableau 7 : Critères biochimiques d'identification de *Xoo* et *Xoc*. (Adapté de Mew *et al.*, 1994 [1])

Test	*Xoo*	*Xoc*
Acetoin production	-	+
Growth on L-alanine as sole carbon source	-	+
Growth in presence of 0.001% cupric nitrate	+	-
Growth on 0.2% vitamin-free casamino acids	-	+
DNA-rDNA hybridization	Both belong to second rRNA superfamily	
SDS-PAGE of cellular proteins	Both show similar banding patterns but each has unique bands	
DNA-DNA hybridization	80% homology between pathovars	
Percent G/C DNA	64.6%	65.0%
Restriction fragment length polymorphism analysis	Repetitive DNA probe from *Xoo* differentiates two pathovars; mismatch of bands between *Xoo* and *Xoc,* 84-93%; between strains of *Xoo* , 8-36%: between strains of *Xoc,* 8-69%	
Fatty acid analysis	Fatty acid profiles similar but differences in presence of 12:0 30H, I5:0 iso, 15:0 anteiso, 16:0 and 17:1 isoF fatty acids	
Serological analysis	Reacts with monoclonal antibody Xco-l but not Xccola	Reacts with monoclonal antibody Xccola but not Xco-1

pathogènes du riz, le premier étant responsable de la bactériose vasculaire du riz, et le second étant à l'origine de la bactériose à stries foliaires du riz (bacterial leaf streak of rice, BLS), moins répandue que la BLB. Les pertes de rendement associées à la strie foliaire sont estimées entre 0 et 20 %, dépendant de la variété de riz cultivée et des conditions climatiques. L'impact de la BLS sur le remplissage des grains est moindre, car l'apparition de nouvelles feuilles permet de limiter les effets de la maladie (Niño-Liu *et al.*, 2006 [84]). *X. oryzae* est une sous-espèce non pathogène découverte aux Etats-Unis et qui n'a pas pu être associée à un pathovar (Triplett, *et al.*, 2011 [122]). Enfin, *Xanthomonas oryzae* pv. *leersiae* anciennement *Xanthomonas campestris* pv. *leersiae* (*Xcl*), est un pathogène de *O. sativa* et *Leersia hexandra* Swartz une *Poaceae* commune retrouvée au sud des Etats Unis, en Amérique du Sud et en Asie (Lang *et al.*, 2019 [123]).

3.2 Généralités

3.2.1 Caractéristiques bactériologiques

3.2.1.1 Morphologie

La morphologie de *X. oryzae* est la même que le reste des espèces du genre. Ce sont des bacilles gram négatifs courts et asporulés. Les cellules mesurent de 0.4 à 0.8 µm de large sur 1.5 à 2.9 µm de long, et sont observables seules ou associées en paires ou en chaînettes. *Xanthomonas* est mobile via un unique flagelle polaire. Ces bactéries produisent un pigment jaune typique du genre, appelé xanthomonadine (Swings *et al.*, 1990 [121], figure 16A). Le mot *Xanthomonas* se compose d'ailleurs des mots grecs « xanthos », signifiant jaune et « monas », signifiant entité. Sur milieu solide, les colonies de *Xanthomonas* sont rondes, lisses, bombées, muqueuses et brillantes (figure 16B).

3.2.1.2 Métabolisme

Les bactéries du genre *Xanthomonas* sont strictement aérobies et catalase positives. Les sources de carbone sont utilisées de manière oxydative et jamais de manière fermentative. L'hydrolyse de l'amidon est lente et la croissance est visible sur D-xylose, D-glucose, D-fructose, D-galactose, cellobiose, sucrose, tréhalose, fumarate de sodium, L-lactate de sodium, L-malate de sodium, oxaloacétate de sodium et succinate de sodium. Une production d'acide est observée à partir du D-glucose, du D-fructose, du D-galactose, du cellobiose et du tréhalose. Aucune acidification du lait de Litmus n'est relevée. La croissance est inhibée dans les milieux contenant 3% de NaCl, 0,001 % de tétracycline, 0,005 % de chloramphénicol, 0,001 % de novobiocine ou 0,001 % de doxycycline, mais pas dans les milieux contenant 0,001 % de triméthoprime ou 0,005 % de métronidazole. L'esculine, le Tween 40 et le Tween 80 sont hydrolysés (Swings *et al.*, 1990 [121]).

Le développement de *X. oryzae* est optimal entre 25 et 31°C, avec une humidité relative d'environ 80%. Il semblerait que la température soit le facteur le plus limitant à l'apparition de symptômes sur la plante hôte, avec une température minimale moyenne requise de 16,7°C et maximale de 37,4°C (Diekman et Bogyo, 1992 [124]). La bactérie produit de grandes quantités d'EPS, lui donnant cet aspect brillant et muqueux sur milieu gélosé. Ces EPS sont impliqués dans la formation des gouttelettes d'exsudat bactérien apparaissant sur les jeunes feuilles infectées, fournissant une protection contre la dessiccation et aidant à la dispersion par le vent et la pluie (Niño – Liu *et al.*, 2006 [84]). Certains critères métaboliques permettant de différencier *Xoo* de *Xoc* sont répertoriés dans le tableau 7.

3.2.2 Culture et étude

3.2.2.1 Observation de symptômes au champ

L'arrivée de la BLB au champ se repère par l'apparition de plantes aux feuilles flétries et jaunâtres au stade adulte. Sur les plantules, les feuilles deviennent grisâtres et s'enroulent, puis le plant entier se dessèche et meurt. Lorsque les lésions sont anciennes, elles virent au blanc, et des tâches sombres dues à la colonisation par des champignons saprophytes se forment (IRRI Rice Knowledge Bank, 2022).

3.2.2.2 Isolement

Des sections de feuilles supposées infectées par *Xoo* sont soumises à une stérilisation de surface et macérées dans de l'eau stérile. La suspension obtenue est étalée sur un milieu solide (Peptone Sucrose Agar) semi-sélectif contenant des antibiotiques et antifongiques. Les colonies apparaissent après 4 à 5 jours d'incubation à 28 ± 2 °C (Reddy et Yin, 1989 [125]).

3.2.2.3 Techniques d'infection

Afin de confirmer un diagnostic, de typer un isolat, d'étudier les mécanismes de virulence du pathogène, ou encore de tester la résistance de lignées de riz, différentes méthodes d'inoculation de *Xoo* ont été mises au point.

Méthodes d'inoculation via les hydathodes (sur des plantes de 10 à 15 jours) :

- Par infiltration. Les plantules entières sont immergées dans une suspension bactérienne dans un récipient sous vide permettant l'infiltration dans les tissus de la plante (Mew *et al.*, 2018 [103]).
- Par pulvérisation. Une suspension bactérienne est pulvérisée sur des plantules entières (Fang *et al.*, 1956 [87]).
- Par adjonction locale d'inoculum. La guttation par les hydathodes est stimulée par le biais d'un taux d'humidité élevé, puis une suspension bactérienne est pulvérisée directement dans les gouttelettes de fluides de guttation (Fang *et al.*, 1956 [87]).

Méthode d'inoculation par piqûre à l'aiguille des feuilles de riz : des aiguilles sont plongées dans l'inoculum et utilisées pour piquer délicatement les limbes des feuilles d'une variété sensible de chaque côté de la nervure.

Infiltration directe des feuilles (sur des plantes de 10 à 15 jours) : les tissus sont infiltrés en injectant la suspension bactérienne via une seringue sans aiguille, appliquée sur la face inférieure des feuilles. Les symptômes sont observés dès 48h après inoculation. Cette méthode permet d'évaluer simplement et rapidement la virulence ou l'avirulence d'un isolat (Mew *et al.*, 2018 [103]).

Inoculation par section des feuilles (sur des plantes âgées de 3 à 5 semaines) : le découpage des feuilles (ou leaf-clipping) permet de cribler des plantes à grande échelle. Les feuilles sont coupées 2 à 3 cm en amont de la pointe avec une paire de ciseaux préalablement immergée dans l'inoculum. Les longueurs de lésions engendrées par la maladie sont mesurées dès 14 jours après l'inoculation. Le leaf-clipping se révèle être une méthode très efficace, car les cellules contournent l'étape de colonisation de l'hydathode pour directement être délivrées dans les tissus vasculaires. Cette méthode est choisie dans la plupart des études, car elle fournit des valeurs quantitatives (Kauffman *et al.*, 1973 [126]).

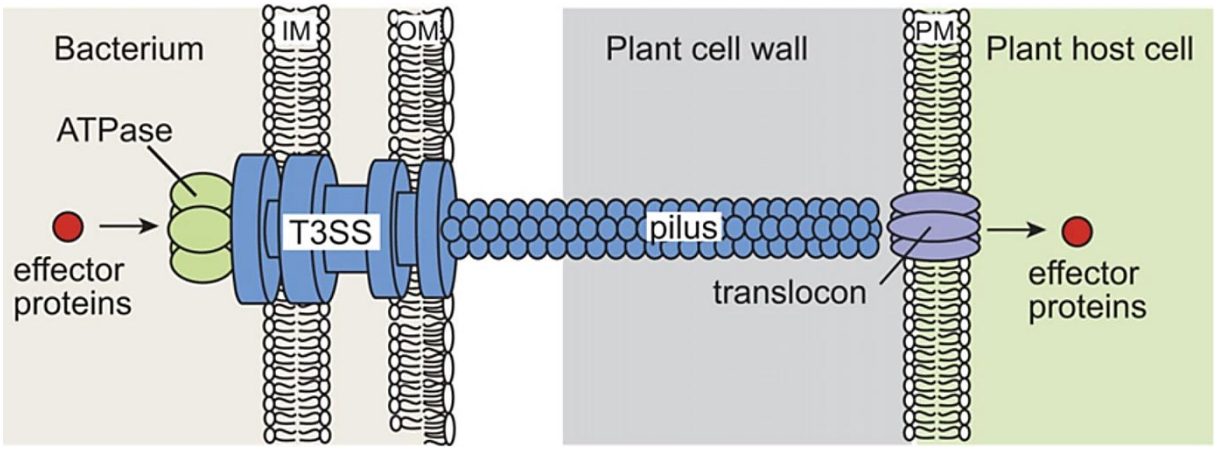


Figure 17 : Représentation schématique du T3SS des bactéries phytopathogènes. L'appareil de sécrétion traverse les deux membranes bactériennes et est associé à une ATPase cytoplasmique. Le T3SS des bactéries pathogènes de plantes possède un pilus extracellulaire qui traverse la paroi végétale de la plante. IM, membrane interne ; OM, membrane externe ; PM, membrane plasmique.

3.3 Déterminants de pathogénie.

3.3.1 Le Système de Sécrétion de Type 3 (T3SS) et ses effecteurs.

3.3.1.1 Distribution et évolution

Le système de sécrétion de type III (T3SS) est une nanostructure en forme de seringue présente dans la membrane de nombreuses bactéries Gram-négatives pathogènes et symbiotiques d'animaux, de végétaux, d'insectes et de nématodes. Ce système de sécrétion contribue au pouvoir infectieux de phytopathogènes majeurs, comme *Pseudomonas syringae*, *Ralstonia solanacearum*, *Xanthomonas* ou *Erwinia* (McCann and Guttman, 2007 [127]). Ces seringues moléculaires permettent l'injection de protéines, appelées effecteurs de type III (T3E), directement dans le cytoplasme des cellules hôtes afin de moduler l'interaction. Le T3SS semble avoir été engendré par la modification d'un appareil flagellaire ancestral. Au cours de l'évolution, sept familles distinctes de T3SS non flagellaires ont vu le jour (Hrp1, Hrp2, Rhizo, SPI-1, SPI-2, Ysc et Chlamy). Les familles Hrp1, Hrp2 sont exclusives aux phytopathogènes, tandis que Rhizo est retrouvée chez les phytosymbiotes (Alvarez-Martinez *et al.*, 2021 [128]). Les gènes associés à la structure et au fonctionnement du T3SS sont regroupés dans des îlots de pathogénicité (PAI) d'une taille d'environ 20 à 50 kpb, le plus souvent situés sur le chromosome, mais parfois présents sur des plasmides. Ils codent pour les composants du système sécrétoire lui-même, ses substrats, leurs protéines chaperonnes et des régulateurs transcriptionnels. Chez les bactéries phytopathogènes, les gènes associés au T3SS sont appelés *hrp*, pour « réponse hypersensible et pathogénicité ». Ils sont subdivisés en deux groupes : les gènes *hrc* pour *hrp*-conserved et les gènes *hpa* pour *hrp*-associated (Alvarez-Martinez *et al.*, 2021 [128]).

3.3.1.2 Description

Due à la forte homologie observée dans les composants du T3SS entre espèces bactériennes, une nomenclature universelle de ces protéines a été proposée par Hueck en 1998 [129]. Une fois assemblées, elles forment une structure composée d'un corps basal inséré dans la membrane, d'une machinerie d'export, d'une aiguille de sécrétion dans les pathogènes animaux ou d'un pilus dans les pathogènes végétaux et se termine par un pore de translocation (Notti et Stebbins, 2016 [130]).

Le corps basal traverse les membranes interne et externe de la bactérie. Il est composé de deux anneaux, SctC et SctJ, respectivements ancrés dans la membrane externe et interne via des hélices transmembranaires (figure 17). À l'intérieur de la lumière des anneaux, il est supposé que la protéine SctI forme une structure cylindrique ayant pour but de soutenir le pilus (Marlovits *et al.*, 2004 [131]).

L'appareil d'exportation est composé de 5 protéines fortement conservées, SctR, S, T, U et V, dont les fonctions ne sont pas encore bien caractérisées. SctV forme une sorte de cage à l'entrée de l'appareil d'exportation, qui faciliterait le dépliement des protéines avant leur translocation. Le complexe formé par les protéines SctR, S et T forme un pore de translocation d'une lumière d'environ 15 °A (Notti et Stebbins, 2016 [130]).

Le pilus doit franchir une barrière physique supplémentaire, la paroi cellulaire végétale, dont l'épaisseur est d'au moins 100 nm. Peut-être en raison de cette particularité, le pilus des pathogènes végétaux est environ 25 fois plus long que l'aiguille des pathogènes animaux. Par ailleurs, le passage du pilus à travers la paroi végétale nécessite l'aide d'enzymes de dégradation (Cell-Wall Degrading enzymes) (Ji et Dong, 2015 [132]). Le translocon situé à l'extrémité du pilus permet la pénétration de la membrane de la cellule hôte et est composé de plusieurs copies de protéines hydrophobes (Wagner *et al.*, 2018 [133]).

L'hydrolyse de l'ATP et la force proton-motrice (due au gradient de concentration de protons) sont toutes deux importantes pour le fonctionnement du T3SS. Les substrats sont dépliés avant d'entamer le trajet à travers le pilus propulsés par la force proton-motrice vers le translocon (Notti et Stebbins, 2016 [130]).

3.3.1.3 Importance et conservation chez Xanthomonas spp.

Les espèces du genre *Xanthomonas* peuvent être classées en deux groupes. La majorité des espèces pathogènes font partie du groupe 2, la quasi-totalité d'entre elles possédant un T3SS de type Hrp, considéré comme leur principal facteur de pathogénie. *X.albineans*, représentante pathogène majeure du groupe 1, est responsable de la brûlure des feuilles de la canne à sucre. Elle possède un T3SS de la famille SPI-1 similaire à celui présent chez Salmonella (Studholme *et al.*, 2011 [134]). Pour les souches du groupe 2, l'absence du T3SS, ou une réduction dans le répertoire de T3Es, conditionne le pouvoir pathogène de l'espèce. C'est le cas des espèces considérées commensales *X. maliensis*, *X. cannabis*, certaines espèces de *X. arboricola*, la souche *X. oryzae* XII-5A isolée aux Etats Unis, ou plus récemment des isolats de *X. melonis* retrouvés en co-occurrence avec *X. campestris* pv. *campestris* dans des feuilles symptomatiques de choux (Teper *et al.*, 2021 [135] ; Assis *et al.*, 2021 [105]; Cesbron *et al.*, 2015 [136] ; Triplett *et al.*, 2011 [122] ; Ramnarine *et al.*, 2021 [137]). Aussi, la mutation induite du T3SS -Hrp des espèces pathogènes du groupe 2 entraîne une perte de virulence, l'incapacité de coloniser l'hôte et des réactions immunitaires violentes de type HR (Hypersensitive Response) (Teper *et al.*, 2021 [135] ; Yan et Wang, 2012 [138] ; Darsonval *et al.*, 2008 [139] ; Medina *et al.*, 2018 [140] ; Mondal *et al.*, 2022 [141]). Ces résultats indiquent que les effecteurs de type 3 (au moins dans leur ensemble) ont des fonctions critiques dans la facilitation du parasitisme de l'hôte.

3.3.1.4 Les effecteurs de type 3

Plus de cinquante classes de T3Es différentes ont été recensées chez les bactéries phytopathogènes, dont une quarantaine retrouvées dans le genre *Xanthomonas* (tableau en annexe 1). Collectivement, ces T3Es permettent de supprimer les réponses défensives de l'hôte et de garantir la multiplication du pathogène dans les tissus. On note qu'il existe une certaine redondance dans le rôle des T3Es car les mutations ponctuelles de la plupart d'entre eux n'entraînent souvent pas de modification du phénotype de virulence ou ont un effet relativement faible (White *et al.*, 2009 [142]). Les protéines T3Es de *Xanthomonas* sont nommées protéines Xop (pour *Xanthomonas* outer proteins) ou protéines Avr (pour avirulence), dans le cas d'une découverte dans le cadre d'interactions incompatibles (Dangl, 1994 [143]).

De nombreux effecteurs T3 de *Xanthomonas* ont pour fonction de supprimer les réponses immunitaires déclenchées par les PAMP (Pathogen-Associated Molecular Pattern) ou la PTI. Par exemple, XopQ, XopX, XopZ et XopN sont capables d'inhiber la détection de dommages causés à la paroi par LipA, une enzyme hydrolytique de la bactérie (Sinha *et al.*, 2013 [144]). XopR, AvrBs2, XopK, XopP, XopL, XopB, AvrAC et XopS ciblent et suppriment les composants impliqués dans les réponses PTI (Alvarez-Martinez *et al.*, 2021 [128]). Chez *X. euvesicatoria* les effecteurs XopF2, XopE2, XopAP, XopAE, XopH et XopAJ empêchent la signalisation induite par flg22, un peptide hautement conservé de la flagelline bactérienne, en agissant en aval ou en parallèle des cascades de MAP kinases, inhibant ainsi le dépôt de callose dans les plantes (Popov *et al.*, 2016 [145]). XopJ de *X. campestris* perturbe le système de marquage à l'ubiquitine de l'hôte, adressant ainsi des composants de l'immunité vers le protéasome pour qu'ils y soient dégradés (Üstün *et al.*, 2013 [146]).

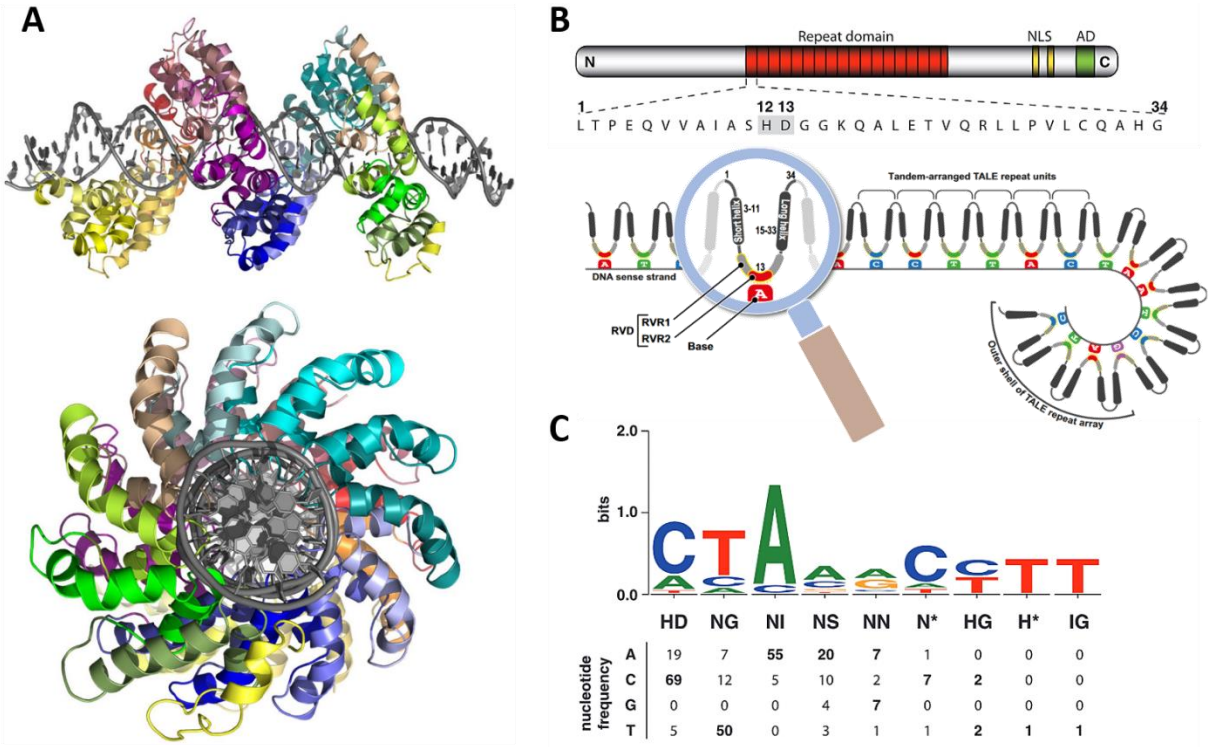


Figure 18 : Spécificité de la reconnaissance TALE/ADN. A. Structure en trois dimensions d'un TALE fixé sur son EBE. Les répétitions en tandem s'enroulent autour de l'ADN et forment une structure superhélicoïdale (Image, Mak *et al*., 2012 [154]) ; B. Base de l'interaction TALE-ADN. Les TALEs se lient à l'ADN par un nombre variable de répétitions, chaque répétition s'appariant avec une base d'ADN dans la séquence EBE correspondante (Image, Nowack *et al*., 2022 [152]). C. Code RVD/ADN. Base les plus fréquemment reconnues par les RVD indiqués (Image, Boch *et al*, 2009 [153])

Chez *X. euvesicatoria*, XopD désumoyle le facteur de transcription SlERF4 ce qui le déstabilise et réprime la transcription des gènes de défense induite par l'éthylène nécessaire à la PTI (Kim *et al.*, 2013 [147]).

Certains T3Es sont reconnus par les systèmes de surveillance de l'hôte, induisant une réponse immunitaire rapide, que l'on appelle Effector Triggered Immunity (ETI). Cependant, des résultats montrent que certains T3Es de *Xanthomonas* sont impliqués dans la suppression de l'ETI. Szczesny [148] et ses collaborateurs ont montré que l'ETI et la HR élicitée par AvrBs1 de *X. euvesicatoria* chez les lignées de poivrons portant le gène de résistance correspondant *Bs1* étaient supprimées par AvrBsT. Chez le même pathogène, XopQ perturbe une voie générale de l'ETI à la fois chez le poivron et la tomate (Teper *et al.*, 2013 [149]). Chez *Xoo*, les réponses immunitaires induites par XopQ et XopX sont supprimées par XopU, XopV, XopP, XopG, et AvrBs2 (Deb *et al.*, 2020 [150]).

Enfin, une classe d'effecteurs particuliers, la famille AvrBs3/PthA (voir aussi §3.3.2), a pour fonction l'activation de la transcription de gènes cibles variés. Ces T3Es, aussi appelés TALEs pour Transcription Activator Like Effectors vont « court-circuiter » la régulation transcriptionelle de gènes de l'hôte pour permettre à la bactérie de manipuler son environnement et garantir sa multiplication, la colonisation des tissus ou la dissémination. Une attention toute particulière est apportée à leur étude chez *Xanthomonas* où ils constituent des facteurs majeurs de pathogénie, mais ils ont également été identifiés chez *Ralstonia solanacearum* (RipTales) et *Burkholderia rhizoxinica* (Bats), une bactérie intracellulaire de champignon phytopathogène. (Alvarez-Martinez *et al.*, 2021 [128], Carter *et al.*, 2020 [151]).

3.3.2 Les effecteurs Transcription Activator Like (TALEs).

3.3.2.1 Distribution et structure

La distribution des effecteurs TAL dans le genre *Xanthomonas* est variable, étant absents de certaines espèces tandis que d'autres, comme certaines *Xoc*, codent jusqu'à 27 paralogues de ces effecteurs (figure 15).

Les TALEs forment une famille multigénique dont les différents membres présentent de 90 % à 97 % d'homologie dans leur séquence ADN. Leur structure caractéristique se décompose en quatre domaines fonctionnels (figure 18) :

- En N terminal se trouve un signal de sécrétion pour la translocation du TALE vers le cytoplasme de la cellule végétale.
- Le domaine central est une région contenant un nombre compris entre 1.5 et 33.5 de répétitions quasi-identiques de 34-35 acides aminés, avec deux positions variables au sein de chaque répétition appelées « Repeat Variable Diresidues » ou RVD. Ces acides aminés variables se trouvent en positions 12 et 13 de chaque répétition.
- La partie C terminale contient : un domaine riche en répétitions leucines (LRR) de fixation pour les facteurs de transcription (Transcription Factor Binding, TFB) ; des signaux nécessaires à l'adressage des effecteurs vers le noyau de la cellule par interaction avec les importines alpha eucaryotes (Nuclear Localization Signal, NLS) ; et un Domaine d'Activation de la transcription (AD) (Nowack *et al.*, 2022 [152]).

Les parties N-terminales et C-terminales des TALEs sont hautement conservées, et les différences entre les effecteurs en termes d'activité biologique sont majoritairement liées aux variations de la région répétitive centrale (Boch et Bonas 2009 [153]).

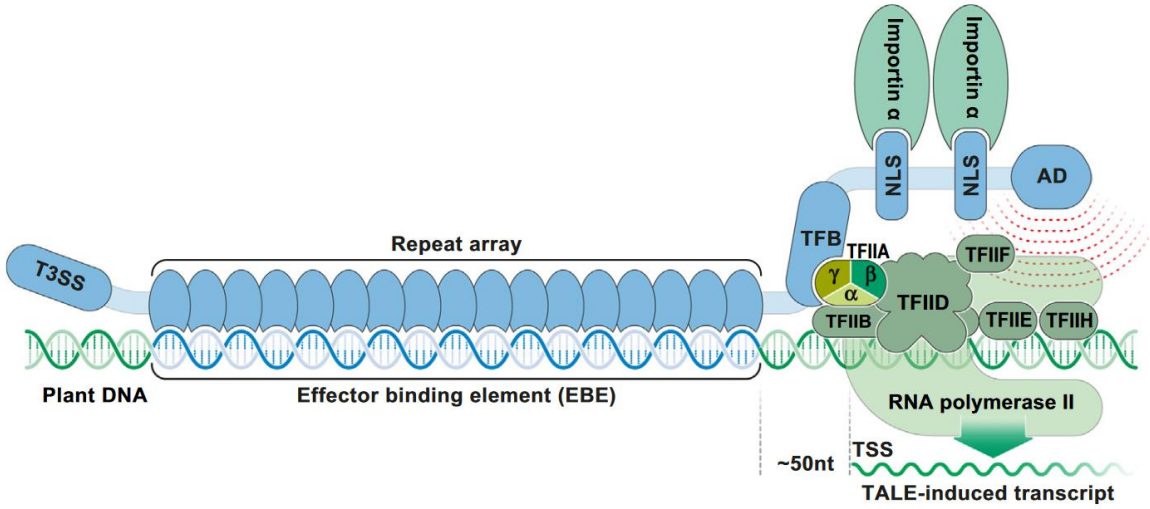


Figure 19 : L'activation de la transcription eucaryote par les TALEs de *Xanthomonas*. Le TFB assure l'interaction avec la sous-unité gamma (γ) du facteur de transcription général TFIIA. Le domaine AD stimule l'activation transcriptionnelle. Le site de départ de la transcription induite par TALE (ligne verte ondulée) commence environ 50 à 100 nucléotides (nt) en aval de l'EBE. Abréviation : NLS, signal de localisation nucléaire (Issu de Nowack *et al*., 2022 [152])

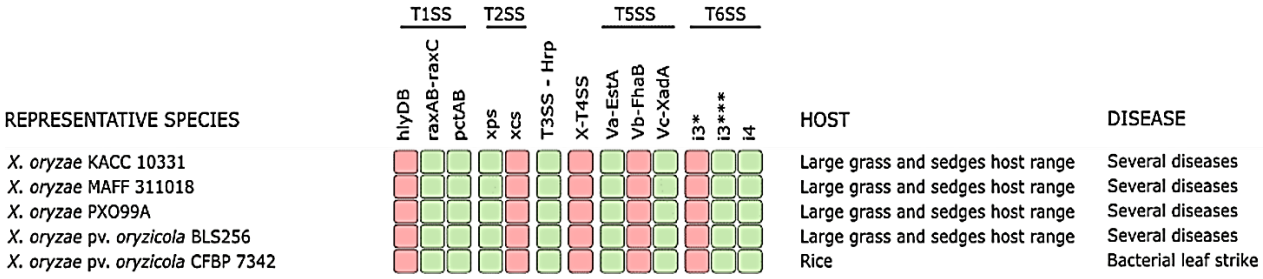


Figure 20 : Répartition des différents TSS chez *Xo*. La présence ou l'absence d'un système de sécrétion particulier est indiquée par des carrés verts et rouges, respectivement. Les espèces végétales hôtes et les maladies associées (le cas échéant) sont indiquées. (Adapté de Alvarez-Martinez *et al*., 2021 [128]).

3.3.2.2 Spécificité de la reconnaissance des cibles ADN et activation de la transcription

Une fois entré dans le noyau de la cellule hôte, le TALE va reconnaître une séquence particulière, appelée Effector Binding Element (EBE), située dans le promoteur du gène ciblé. La spécificité de la séquence reconnue par un TALE dépend des RVD présents dans chacune des répétitions du domaine central, chaque RVD pouvant se fixer à un nombre restreint de bases données (voir figure 18B ; Boch *et al.*, 2009 [155] ; Moscou et Bogdanove, 2009 [156]). Le résidu en position 12 stabilise la boucle et la projette dans le sillon majeur, où le 13e résidu, entre en contact avec le nucléotide. Un minimum de six répétitions semble être nécessaire à la fixation du TALE. En 2017, Rinaldi et ses collègues [157] ont examiné la relation entre le nombre de répétitions dans un TALE et son affinité, à la fois pour l'ADN cible et non-cible. Chaque répétition supplémentaire augmente la stabilité de la fixation du TALE à l'ADN, mais en diminue la spécificité (de manière non linéaire). Ainsi, ils ont pu mettre en évidence que la spécificité augmente, puis diminue avec l'augmentation de la longueur du TALE, atteignant un pic entre 15 et 19 répétitions. Dans la région située immédiatement avant les répétitions centrales, les effecteurs TAL contiennent plusieurs résidus chargés positivement (lysine et arginine) et quatre répétitions dites non-canoniques (répétitions -3 à 0) qui ressemblent structurellement aux répétitions centrales mais arborent des séquences différentes (Gao *et al.*, 2012 [158]). Ces dernières permettraient de stabiliser la liaison à l'ADN. À l'exception de celle de TalC, toutes les EBEs connues dans la nature débutent par un T, qui est spécifié en N-terminal, par la répétition non-canonique 0 (Yu *et al.*, 2011 [159]). Le décryptage du code de fixation à l'ADN des TALEs (voir figure 18C) a permis i) l'identification de cibles importantes dans la genèse des maladies liées à *Xanthomonas*, ii) l'identification de gènes de résistance de type « Exécuteurs ») chez les différents hôtes de *Xanthomonas* (voir 4.4.3.2 « Gènes de résistance de type exécuteur (E) », iii) la construction de TALE artificiels (ou arTALEs) permettant l'activation ciblée de gènes, iiii) la création d'arTALes fusionnés à des nucléases, appelés TALEN, pour l'édition de génomes (Scholze et Boch, 2010 [160] ; Wang *et al.*, 2017 [31]). Une fois fixé, l'effecteur va activer la transcription du gène. Pour cela, le TFB adjacent au AD du TALE, va interagir avec la sous-unité gamma (γ) du facteur de transcription général TFIIA, qui fait partie du complexe de pré-initiation de l'ARN polymérase II de l'hôte. Ce complexe est stabilisé par la présence de cyclophilines et thiorédoxines végétales (Domingues *et al.*, 2012 [161]). Enfin, le AD va stimuler la transcription eucaryote (par un mécanisme encore inconnu). Le site de début de transcription induit par le TALE (ligne verte ondulée sur la figure 19) commence environ 50 à 100 nucléotides (nt) en aval de l'EBE (Nowack *et al.*, 2022[152]). Les effecteurs TAL peuvent diriger la transcription de manière bidirectionnelle chez les plantes, à partir de l'un ou l'autre brin (Doucouré *et al.* 2022 [162]; Wang *et al.*, 2017 [31]).

3.3.3 Autres déterminants de pathogénie

3.3.3.1 Autres systèmes de sécrétion

Les autres systèmes de sécrétion rapportés chez *Xoo* sont le T1SS, le T2SS, T5SS et le T6SS (figure 20 et 21 , Alvarez-Martinez *et al.*, 2021 [128] ; Sgro *et al.*, 2019 [163]).

Le système de sécrétion de type I permet à la bactérie d'excréter des protéines dépliées dans le milieu extracellulaire. Les substrats du T1SS sont des cytotoxines de la superfamille des toxines RTX et des adhésines (Smith *et al.*, 2018 [164])

Le système de sécrétion de type II a pour fonction principale l'acquisition de nutriments pour le pathogène et de favoriser la colonisation des tissus. En effet, la majorité des substrats du T2SS caractérisés à ce jour sont des enzymes hydrolytiques qui assurent la dégradation de lipides, de protéines et de polysaccharides (voir aussi en 3.3.3.3). Chez les organismes phytopathogènes comme *Xoo*, le T2SS permet notamment l'excrétion de cellobiosidases et de cellulases impliquées dans la dégradation des parois végétales, nécessaires à la virulence de *Xoo* (Rajeshwari *et al.*, 2005). La sécrétion des substrats s'effectue en deux temps, la première étape de traversée de la membrane

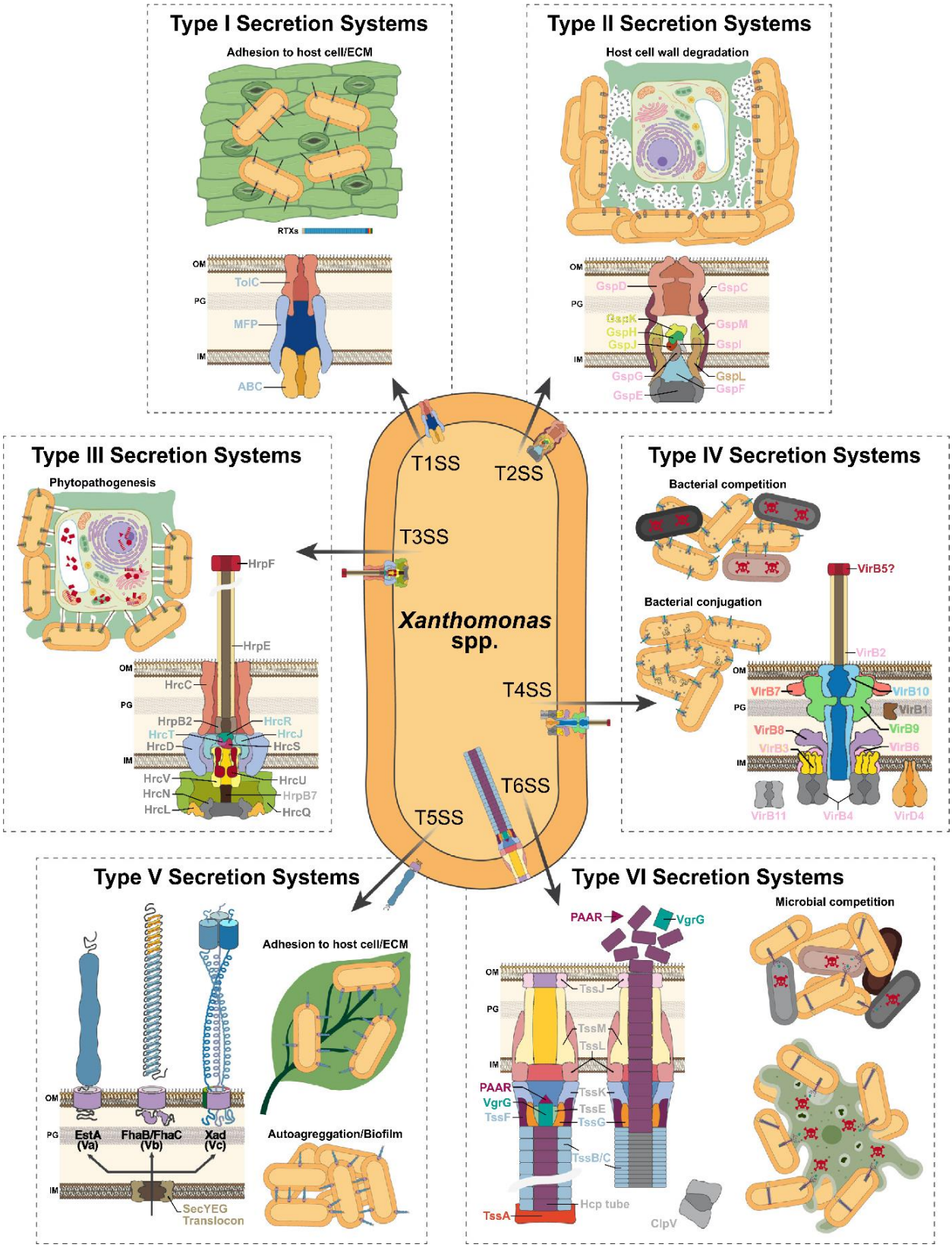


Figure 21 : Schéma des morphologies des différents systèmes de sécrétion. IM, inner membrane; PG, peptidoglycan; OM, outer membrane. (Tiré de Alvarez-Martinez *et al*., 2021 [128])

interne étant assurée par les système Sec/tat, sous forme dépliée, puis la maturation a lieu dans le périplasme avant l'export dans l'espace extracellulaire (Nivaskumar et Francetic, 2014 [165]).

Le T5SS est le système le plus simple d'un point de vu structurel. Il est constitué d'une ou deux protéines qui se localisent dans la membrane externe et ne nécessite pas d'ATP ou de gradients électrochimiques comme sources d'énergie pour le transport, d'où son appellation d'auto-transporteur (Leo *et al.*, 2012 [166]). Les effecteurs exportés sont soit libérés dans le milieu extracellulaire, soit conservés associés au transporteur, devenant ainsi exposés à la surface de la cellule bactérienne. Les T5SS de *Xoo* (types Va et Vc) exportent des substrats restants associés au transporteur, dont des protéases, des lipases et des adhésines (Alvarez-Martinez *et al.*, 2021 [128]).

Le T6SS est le dernier système de sécrétion ayant été décrit chez les Protéobactéries. Il excrète directement des protéines vers des cellules cibles procaryotes et eucaryotes par contact et est impliqué dans l'antagonisme bactérien. Les substrats du T6SS sont majoritairement des toxines, mais on trouve aussi des sidérophores (Alvarez-Martinez *et al.*, 2021 [128]). Le T6SS de *Xoo* n'ayant été caractérisé que récemment, peu d'informations concernant le type de substrats excrétés sont disponibles. Cependant, son importance dans la survie et la croissance de la bactérie a été démontrée (Choi *et al.*, 2020 [167]).

3.3.3.2 Molécules de surface

Parmi les facteurs de virulence classiques des bactéries pathogènes, on retrouve des structures nécessaires à la mobilité : les flagelles. La transcription du flagelle polaire de *Xoo* est induite au contact de facteurs de l'hôte (Kim *et al.*, 2016 [168]). Certaines données suggèrent que le chimiotactisme a un rôle important dans le déplacement vers le site d'infection, mais ne serait pas requis une fois que le pathogène est entré dans les feuilles (Shen et Ronald, 2002 [169]). Plusieurs adhésines sont impliquées dans la promotion de la virulence de *Xoo* chez le riz, facilitant l'adhésion à la surface des feuilles et les phases précoces de l'infection. Ces molécules sont pour la plupart des représentantes du T5SS, autotransporteur décrit en 3.3.3.1). Les T5SS XadA et XadB (pour *Xanthomonas* adhesin-like protein A et B) sont impliqués dans les phases précoces d'attachement et d'entrée. L'autotransporteur de type YapH et le pilus de type IV quant à eux favorisent la croissance et la migration dans les vaisseaux du xylème (Das *et al.*, 2009 [170]). Comme d'autres espèces de *Xanthomonas*, *Xoo* produit de grandes quantités d'EPS. Ces molécules fournissent une protection contre la dessiccation et les différents composés antimicrobiens rencontrés chez l'hôte et dans le milieu extérieur (notamment les antibiotiques). Elles facilitent également l'adhésion aux tissus. Une étude de 2016 par Girija et ses collègues [171] a montré que les EPS de *Xoo* étaient capables de supprimer les défenses de l'hôte induites par la détection de l'antigène O des lipopolysaccharides de la bactérie. Une souche de *Xoo* mutée dans un des *loci* impliqués dans la synthèse d'EPS (gumGxo) ne produit plus d'EPS et a une virulence réduite. Il en va de même lors de la mutation du gène *rpfC* (regulation of pathogenicity factor) impliqué dans la régulation des gènes codants pour la production d'EPS (Shen et Ronald, 2002 [169]).

3.3.3.3 Enzymes

La pectine, la cellulose et l'hémicellulose confèrent aux parois végétales leur rigidité et constituent une barrière contre les agents pathogènes. On considère depuis longtemps que les enzymes extracellulaires ont un rôle majeur dans la virulence des phytopathogènes. Via son système de sécrétion de type 2 (T2SS, voir description en 3.3.3.1), *Xoo* sécrète une variété d'enzymes pour décomposer les composants des parois cellulaires de l'hôte. Parmi ces « Cell-Wall Degrading Enzymes » on trouve deux cellulases, la xylanase XynB, la lipase/estérase LipA, la pectine méthyl

estérase Pmt et la polygalacturonase PglA. À l'exception de *pmt* et *pglA* qui ont des effets minimes, les mutations de chacun de ces gènes entraînent une réduction de la virulence de *Xoo*, indiquant leur importance dans la genèse de la maladie (Sinha *et al.*, 2013 [144]; Rajeshwari *et al.*, 2005 [172]). Parmi les cellulases excrétées via le T2SS de *Xoo*, on trouve une endoglucanase, ClsA, et une exoglucanase (ou cellobiosidase) CbsA. Les cellulases décomposent la cellulose en monosaccharides et/ou oligosaccharides simples. Il a été démontré que l'activité de la CbsA est essentielle pour la virulence de *Xoo*, et que la présence ou l'absence du gène *cbsA* chez *Xanthomonas* conditionne le mode d'infection vasculaire ou non vasculaire, respectivement (Tayi *et al.*, 2017 [173]; Gluck Thaler *et al.*, 2020 [174]; Nathawat *et al.*, 2022 [175]). Le xylane est un composant important de l'hémicellulose des parois cellulaires végétales et sa concentration est exceptionnellement élevée dans les parois du riz et d'autres graminées, représentant jusqu'à 60 % du total des glucides. Les mutations de *xynB* ou *lipA* affectent partiellement la virulence tandis que la virulence d'un double mutant *lipA-xynB* est significativement réduite (Rajeshwari *et al.*, 2005 [172]).

3.3.3.4 Molécules du Quorum Sensing (QS)

Le mécanisme de détection du quorum (QS) permet aux bactéries de détecter leur densité de population, et de synchroniser l'activité de chaque individu en un comportement communautaire coopératif. Chez les bactéries Gram-négatives, outre les signaux QS de type AHL et AI-2 bien caractérisés, des signaux de la famille des Diffusible Signal Factors (DSF) ont été signalés chez plusieurs phytopathogènes, dont *X. campestris* pv. *campestris* (*Xcc*), *Xyllela fastidiosa* et *Xoo*. Parmi les protéines régulatrices de la signalisation DSF, le système à deux composants RpfC/RpfG est impliqué dans la détection et la transduction du signal DSF (He *et al.*, 2010 [176]). Lorsque des mutants de *Xoo* pour *rpfB*, *rpfC*, *rpfF* et *rpfG* sont inoculés sur un cultivar de riz sensible, les longueurs des lésions produites sont significativement plus courtes que celles observées pour la bactérie sauvage. L'expression des gènes liés à la production d'EPS (*gumG* et *gumM*), de LPS (*xanA*, *xanB*, *wxoD*, *wxoC*), de phytase (*phyA*), de xylanase (*xynB*) et de lipase (*lipA*) est réduite de manière significative chez ces mutants (Jeong *et al.*, 2008 [177]).

3.4 Génome et potentiel évolutif.

3.4.1 Données génomiques

3.4.1.1 Description générale du génome

Le génome de *Xoo* consiste en un chromosome circulaire d'environ 4.9 Mb avec un pourcentage de GC autour de 63.7%. Il comporte 3700 à 5000 gènes codants, dont une cinquantaine pour des ARNs de transfert. Les gènes codant pour des TALEs sont en 9 à 16 exemplaires dans le génome. Les gènes *TALEs* sont répartis en clusters, mais chaque gène semble avoir son propre promoteur (pas d'ARNm polycistronique). Également, la bactérie possède une région contenant le système CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). Tous les génomes de *Xanthomonas* séquencés contiennent de nombreux éléments mobiles de type IS (Séquences d'Insertion), mais les génomes de *Xoo* contiennent le pool le plus diversifié (voir 3.4.2.2 « Plasticité du génome chez *Xanthomonas* et rôle dans l'évolution de la pathogénie » pour plus de détails). L'analyse du génome de *Xoo* laisse entrevoir sa grande plasticité, et sa rapidité d'évolution. L'exemple le plus évident est le nombre de réarrangements chromosomiques et insertions-délétions de portions de génome observées entre souches de *Xoo* (Salzberg *et al.*, 2008 [178]; Huguet-Tapia *et al.* 2016 [179]; Chen *et al.*, 2021 [180]). La présence de plasmides chez *Xoo* a été décrite pour la première fois en 1994 chez des souches

Indiennes. La caractérisation de l'un de ces plasmides, pMA36, a révélé qu'il contenait un gène de résistance à l'ampiciline et a également montré son importance dans la virulence de la souche qui le portait (XoC36) (Amuthan *et al.*, 1994 [181]). En 2015, lors du criblage de plus de 250 souches de *Xoc* en provenance de Chine, 9 souches portant des plasmides ont été découvertes. Le premier plasmide séquencé chez *Xoc* (souche GXO1), s'est montré être auto-transmissible et réplicatif dans certaines souches de *Xanthomonas* spp., et améliorer la tolérance de *Xoo* aux métaux lourds (Niu *et al.*, 2015 [182]). Plus récemment, un plasmide isolé d'une autre souche Indienne, IX-280, s'est révélé porter des gènes codant pour un système de sécrétion de type 6 ainsi que des systèmes toxines / antitoxines (Carpenter *et al.*, 2020 [183]).

3.4.1.2 Ressources génomiques disponibles

438 assemblages de génomes de *Xoo* sont disponibles sur NCBI, mais seulement 95 génomes sont complets, dont 91 avec des annotations dans Refseq. La grande majorité des génomes séquencés sont d'origine asiatique. Le premier génome obtenu fût celui de la souche coréenne KACC10331, en 2005, par Lee et ses collègues [184], suivi de près par la souche japonaise MAFF311018 (Ochiai *et al.*, 2005 [185]). Suivirent d'autres souches asiatiques, notamment des Philippines, comme les souches de référence PXO99^{A} et PXO86 (Salzberg *et al.*, 2008 [178] ; Booher *et al.*, 2015 [186]). Le premier génome complet séquencé pour une souche africaine fût celui de la souche camerounaise AXO1947 en 2016 (Huguet-Tapia *et al.*, 2016 [179]). Huit autres souches des Philippines, représentant chacune une race différente de celles de PXO86 et PXO99 furent séquencées en 2016 (Quibod *et al.*, 2016 [187]). Une campagne de séquençage de génomes provenant d'Afrique menée par l'IRD en 2018 a permis d'obtenir les génomes de huit souches maliennes ainsi que celui de la souche de référence burkinabé BAI3 (Doucouré *et al.*, 2018 [188] ; Tran *et al.*, 2018). Enfin, un effort important de séquençage a été accompli en 2019 par le consortium Healthy Crops dans le cadre des travaux de Oliva et al [189], où 26 génomes complets de souches d'origines diverses ont été séquencés : Cameroun, Burkina Faso, Mali, Niger, Bénin, Côte-d'Ivoire, Sénégal, Tanzanie, Uganda, Sri-Lanka, Népal et Philippines.

3.4.2 Plasticité du génome et potentiel évolutif

3.4.2.1 Principales sources d'évolution chez les microorganismes

La source majeure d'évolution d'un organisme est la mutation. Une mutation est une modification héréditaire de la séquence ADN, pouvant, ou non, affecter le phénotype de l'organisme qui la porte (Najafi et Pezeshki, 2013 [190]). Ce sont les mutations qui vont permettre la création de variation génétique et l'apparition de nouveaux allèles au sein d'une population. Ces mutations sont le résultat de différents processus génétiques. Tout d'abord, il y a l'apparition au cours de la réplication de l'ADN « d'erreurs » pouvant être des substitutions, délétions ou additions de nucléotides. La fréquence d'apparition de ces erreurs chez les bactéries varie généralement de 1 sur 10 millions à 1 sur 1 milliard par nucléotide et par génération. On trouve également des bactéries « ultra mutantes » avec une fréquence 100 fois plus élevée, ceci étant généralement dû à un système de réparation et de correction de l'ADN déficient (Chevallereau *et al.*, 2019 [191]).

Une autre source de mutation est la présence dans le génome de zones répétées, appelées microsatellites, consistant en des répétitions de 1 à 5 pb. Ces microsatellites favorisent les évènements de délétions, de duplications et de réarrangements génomiques (Metzgar *et al.*, 2002 [192]). Enfin, une source de mutation très importante prend la forme d'éléments génétiques mobiles au sein du génome bactérien. Les éléments génétiques mobiles (EGM) constituent une grande classe d'entités

génomiques regroupées sous le nom de « mobilome », dont les plus étudiées sont les plasmides, les îlots génomiques, les transposons (Tn), les séquences d'insertion (IS), et les séquences inversées répétées (IR) (Patel, 2016 [193]). Plus le mobilome d'une bactérie est fourni et son nombre de zones répétées important, plus son taux de mutation sera élevé.
Les bactéries ont une vitesse de reproduction rapide. De ce fait, la probabilité de réalisation d'un évènement de mutation est plus élevée. La taille de la population est alors elle aussi vectrice de potentiel évolutif, car elle influe sur la probabilité d'apparition de mutants. Ainsi, le potentiel d'évolution est plus élevé chez des organismes dont la population est importante que chez des organismes aux populations réduites. La diversité génétique d'une population peut d'ailleurs faire l'objet d'une sélection drastique dans le cas du déploiement d'un gène de résistance (ce que l'on appellera sélection directionnelle). Les mutants avantagés (c'est-à-dire toujours virulents) augmentent rapidement en fréquence au sein de la population et font perdre au gène de résistance son efficacité (McDonald et Linde, 2002 [194]).
Les échanges de gènes ou de génotypes entre des populations séparées (appelé flux de gènes ou flux de génotypes) sont eux aussi créateurs de diversité et donc instigateurs d'évolution. Au-delà des capacités de dispersion propres au pathogène, ce flux est influencé par les mouvements de matériel contaminé par le biais des pratiques culturales et du commerce réalisé par l'homme (McDonald et Linde, 2002 [194]).
Les modes de reproduction et de croisements influent sur la manière dont la diversité génétique est répartie au sein de la population et dans le génome d'un organisme. Chez les bactéries la reproduction est certes clonale, mais une forme de sexualité appelée transfert de gène horizontal existe (HGT). Ces transferts sont de trois types : conjugaison (transfert d'ADN d'une bactérie à une autre via la formation d'un pilus permettant le contact), de transduction (processus par lequel l'ADN d'une bactérie est transporté dans une autre bactérie par le biais d'un virus) et de transformation (incorporation d'ADN extracellulaire libre). Bien que des transferts de gènes horizontaux incontrôlés puissent être dangereux pour la bactérie (le système CRISPR étant la protection bactérienne la mieux étudiée pour contrer cette menace transgénique), la bactérie hôte permet ces transferts pour mieux faire face à son environnement (Dorman, 2014 [195]). La fréquence d'HGT dépend de nombreux facteurs. On sait par exemple qu'elle est plus élevée quand les bactéries sont en biofilm (Patel, 2016 [193]). Une population qui recombine fréquemment peut créer de nouvelles combinaisons d'allèles de virulence aussi rapidement que les sélectionneurs peuvent recombiner les gènes de résistance, constituant un vrai danger pour les cultures (McDonald et Linde, 2002 [194]).

3.4.2.2 Plasticité du génome chez Xanthomonas et rôle dans l'évolution de la pathogénie

La plasticité du génome joue un rôle important dans l'adaptation et la réponse à la sélection directionelle des bactéries. Des analyses métagénomiques ont montré que les antibiotiques présents dans les eaux usées ont significativement augmenté le nombre de gènes de résistance dans les communautés microbiennes, ainsi que l'abondance des EGMs dans les génomes des bactéries (Zhao *et al.*, 2021 [196]).
Les *Xanthomonas* possèdent de nombreux EGMs, comme des ISs, des Tns, des îlots génomiques et des plasmides. Les IS jouent un rôle important dans la génération de duplications, de réarrangements chromosomiques comme des inversions, des insertions ou des délétions de portions de génome, et conduisent à l'acquisition, la modification ou la perte de gènes. Sur les 19 familles d'éléments IS connues, huit apparaissent chez *Xoo*, avec pour certaines souches jusqu'à 386 IS décrites, indiquant une grande plasticité du génome de l'espèce (Ochiai *et al.*, 2005). Chez *X. arboricola*, il a été démontré que les souches pathogènes possèdent un répertoire d'IS, de prophages et d'îlots génomiques plus important que les souches non pathogènes. Les IS de *X. arboricola* sont notamment impliquées dans l'acquisition et la diversification de T3Es et de résistances , comme la tolérance aux pesticides à base

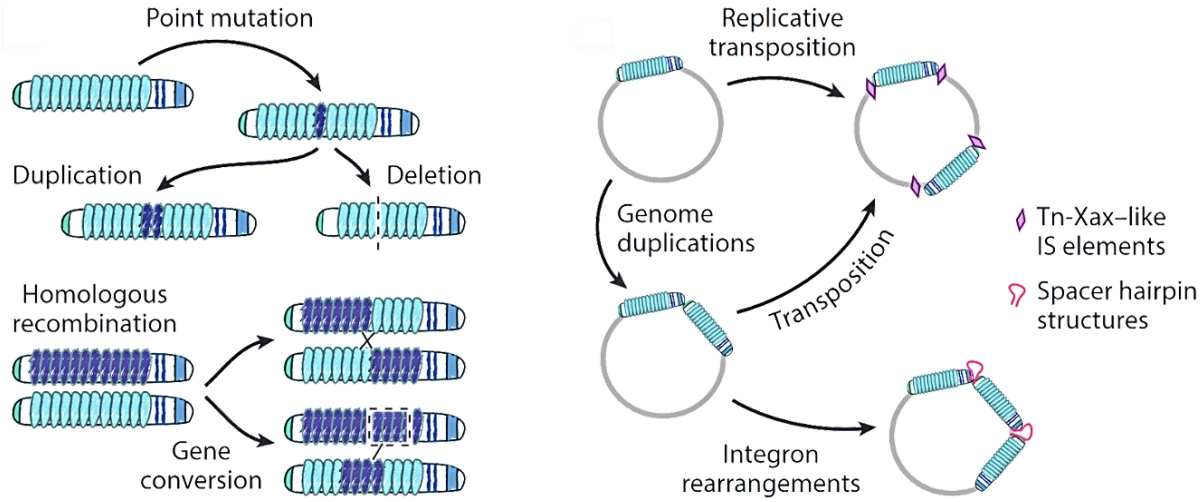


Figure 22 : Principaux mécanismes d'évolution des TALEs. IS, insertion sequence, Tn, Transposon. (Issu de Perez-Quintero et Szurek 2019 [202]).

de métaux (Assis *et al.*, 2021 [105]). Chez *Xoo*, les IS sont associées à un grand nombre de réarrangements chromosomiques.

Les plasmides sont à l'origine d'une autre source de diversité génétique pour *Xanthomonas*. Par exemple, la résistance au cuivre apparue chez *X. arboricola* est portée par un transposon situé sur un plasmide (Assis *et al.*, 2021 [105]). Certaines études mettent d'ailleurs en avant le rôle de « véhicule » des plasmides pour l'échange de gènes via l'action d'éléments génétiques mobiles, connus pour faciliter le transfert de fragments génomiques entre chromosomes et plasmides (Huang *et al.*, 2021 [106]). Ces études montrent que les plasmides portant des gènes de résistance aux métaux lourds chez *X. citri pv. citri* possèdent de multiples IS et Tn. De nombreux T3Es, dont des TALEs et autres facteurs de virulence de cette espèce sont d'ailleurs aussi trouvés associés à des transposons (Ferreira *et al.*, 2015 [197]).

Comme le montrent plusieurs études, environ 5 à 25 % du génome de *Xanthomonas* est issu de recombinaisons (Lima *et al.*, 2008 [198]). Une recombinaison à l'échelle du génome entre *X. perforans* et *X. euvesicatoria* a donné lieu à des échanges d'effecteurs et de facteurs de virulence. Des recombinaisons génétiques entre *Xanthomonas* et d'autres genres bactériens ont également été signalées. Par exemple, *X. arboricola* porte un large segment génomique (~95 kb), avec des gènes conférant une résistance au cuivre, ressemblant à des gènes de *Stenotrophomonas maltophilia* et *Pseudomonas aeruginosa* (Timilsina *et al.*, 2021 [199]).

Il semble donc exister une relation forte entre les EGM et les déterminants de pathogénie chez *Xanthomonas*. Les HGT et les mutations de gènes d'avirulence sont parmi les principaux facteurs qui influencent l'évolution de la virulence chez *Xanthomonas*.

3.4.2.3 Evolution des TALEs

Afin d'échapper à la pression de sélection des gènes de résistance de l'hôte, les différentes espèces de *Xanthomonas* voient leurs répertoires de TALE se modifier et s'enrichir. Il a d'ailleurs été montré que des souches de *Xo* ayant été exposées à des cultivars de riz domestiqués possédaient un répertoire de TALE bien plus large que leurs représentantes retrouvées sur des Poaceaes sauvages (Lang *et al.*, 2019 [123]). De même, aux Philippines, l'adaptation à des cultivars résistants portant l'allèle *Xa4* a engendré la diversification des populations de *Xoo* présentes sur le territoire, notamment via la diversification de leur arsenal de TALEs (Quibod *et al.*, 2016 [187]). La diversité acquise dans ces TALEs est principalement due à des processus de recombinaisons, d'insertions et de délétions dans la région des répétitions (figure 22). En effet, les régions répétées largement conservées chez les TALEs sont propices aux recombinaisons. Par exemple, le locus *talF* chez les souches de *Xoo* africaines BAI3 et MAI1 montre des signes de recombinaison avec d'autres TALEs de ces souches (Tran *et al*., 2018 [200]). Teper et Wang en 2021 [201] ont démontré chez *X*. *citri* pv. *citri*, en utilisant des TALEs de synthèse ne pouvant pas reconnaître les EBEs des gènes cibles de l'hôte, que des mutations adaptatives des TALEs reconstituaient la virulence des souches en moins de 30 cycles d'infections répétés. Un autre processus à l'origine de la création de nouveaux T3E est celui du « réassortiment terminal ». Il s'agit de la fusion du domaine N-terminal encodant les signaux sécrétion et de translocation de l'effecteur par le T3SS avec une autre région génomique. Si le nouvel effecteur confère un avantage sélectif, il sera conservé. Une analyse des T3Es de 23 espèces pathogènes et non-pathogènes a mis en évidence qu'en moyenne 24% des familles de T3Es possédaient un effecteur chimère de deux T3Es, ou d'un T3E et d'une autre portion génomique (McCann et Guttman, 2008 [127]). Enfin, il est intéressant de mentionner que les 19 TALEs présents chez PXO99 montrent une composition en bases différente de celle du reste du génome de la bactérie et sont adjacents à des transposases. Cela suggère qu'ils ont été acquis depuis d'autres espèces, potentiellement en copie simple, suivie d'une expansion dans le génome de *Xoo* (Salzberg *et al.*, 2008 [178]).

3.5 Diversité des populations de *Xoo*.

3.5.1 Outils de caractérisation

Les bactéries phytopathogènes se reproduisent rapidement et donnent naissance à de grandes communautés. Leur vitesse de génération de diversité, notamment par réarrangements chromosomiques, duplications de gènes, mutations et transferts de gènes horizontaux, fait que leur potentiel évolutif est bien plus important que celui de leurs hôtes (De Vries *et al.*, 2020 [203]). Il est donc crucial de surveiller attentivement l'émergence de nouvelles lignées de pathogènes afin de gérer efficacement la résistance des hôtes. Il existe de nombreuses techniques de typage, phénotypiques ou moléculaires, permettant de décrire les populations microbiennes. Pour *Xoo*, les typages phénotypiques attribuent des profils de virulence aux souches testées par inoculations sur différentes lignées de riz. Les méthodes de typage moléculaire, basées sur les acides nucléiques, peuvent quant à elles être classées en quatre types : celles utilisant des motifs de restriction, celles fonctionnant par hybridation, celles reposant sur l'amplification par PCR et enfin, le séquençage.

3.5.1.1 Méthodes basées sur les motifs de restriction

Cette classe compte l'électrophorèse en champ pulsé (« Pulsed Field Gel Electrophoresis », PFGE), la technique RFLP « Restriction Fragment Length Polymorphism » et le ribotypage. La PFGE est un type d'électrophorèse qui sépare de grandes molécules d'ADN (jusqu'à 800 kb) dans des conditions de champ pulsé, permettant la visualisation des motifs de restriction du génome bactérien entier sans utilisation de sondes. C'est une méthode de référence avec une bonne reproductibilité et une bonne capacité de typage. Cependant, elle demande beaucoup de travail et prend 2 à 4 jours (Tynkkynen *et al.*, 1999 [204]). La RFLP est une technique où l'ADN à étudier est digéré puis migré sur gel. La RFLP a constitué la première génération de typage génotypique bactérien dans les années 1980. Cette technologie présente plusieurs inconvénients dont la préparation de grandes quantités d'ADN de haute qualité et la standardisation précise des protocoles d'électrophorèse. Une variante de la RFLP est l' « Amplified Fragment Length Polymorphism » (AFLP). Cette fois, un sous-ensemble d'ADN digéré est amplifié par PCR via des amorces choisies en amont. Cette technique permet une analyse rapide de l'ensemble du génome à la recherche de polymorphismes précis (Sheeja *et al.*, 2021 [205]). Le ribotypage est basé sur la digestion de l'ADN génomique de l'organisme d'intérêt et l'hybridation par Southern Blot de sondes spécifiques aux gènes ribosomiques (cistrons d'ADNr). Il est plus approprié pour la distinction des bactéries au niveau de l'espèce et des sous-espèces qu'au niveau des souches. Cette technique est reproductible, avec un pouvoir discriminatoire élevé. Cependant, c'est une manipulation coûteuse qui nécessite un de temps de travail non-négligeable (Tynkkynen *et al.*, 1999). Une variante du ribotypage est l'ARDRA, pour « Amplified Ribosomal DNA Restriction Analysis ». Dans cette version, l'ADN ribosomique est amplifié par PCR avant d'être digéré. La résolution dépend du nombre d'endonucléases utilisées. Cette technique est rapide, facile et rentable. La complexité des profils de bandes est la principale limite de cette méthode (Vaneechoutte et Heyndrickx, 2001 [206]).

3.5.1.2 Méthodes basées sur l'hybridation

La technique d'hybridation la plus utilisée est le typage microarray. L'ADN à comparer à l'échantillon est fixé sur une matrice à laquelle sera hybridé l'ADN test. Cette technique permet la détection de la présence ou de l'absence de régions d'ADN dans le génome des échantillons testés, plusieurs souches pouvant être testées à la fois (Garaizar *et al.*, 2006 [207]). Si les microarrays sont toujours considérés comme des techniques rentables pour l'analyse de différentes communautés microbiennes, les outils de séquençage haut débit les supplantent d'année en année (Sharma *et al.*, 2020 [208]).

3.5.1.3 Méthodes basées sur l'amplification d'acides nucléiques

La RAPD (« Randomly Amplified Polymorphic DNA ») est un outil de typage qui n'utilise qu'une seule amorce de 20-25 bp. Cette amorce s'hybride de façon aléatoire dans le génome. Les produits d'amplification sont séparés sur gel pour générer une empreinte qui est utilisée pour identifier et caractériser les différentes souches. L'analyse RAPD est notamment appliquée en l'absence de séquence génomique, pour le développement d'amorces spécifiques aux souches (Sharma *et al.*, 2020 [208]). L'analyse MLVA (« Multiple-Locus VNTR Analysis ») est une méthode de typage génétique basée sur l'évaluation du nombre de séquences répétées dans plusieurs loci sélectionnés de l'ADN microbien. Ces loci, sont appelés VNTR pour «Variable Number of Tandem Repeats». La technique MLVA implique l'amplification par PCR multiplexe d'un nombre choisi de VNTRs dans le génome, ce qui conditionnera la résolution de l'analyse (son pouvoir discriminant entre populations). Poulin et ses collaborateurs ont mis au point en 2015 [209] une méthode MLVA permettant de produire des données génétiques robustes et reproductibles pour caractériser efficacement les souches de *Xo* et étudier les épidémies de BLB et de BLS.

3.5.1.4 Méthodes basées sur le séquençage

L'avantage fondamental des méthodes basées sur le séquençage par rapport aux techniques de visualisation de bandes est que les données de séquence sont sans ambiguïté et transférables entre laboratoires. Parmi les méthodes basées sur le séquençage, il y a celles qui utilisent en amont l'amplification par PCR et celles qui n'en ont pas besoin. Le « Multiloci Sequence Typing » (MLST) est une technique utilisant la PCR pour amplifier des fragments de gènes de ménage avant leur séquençage. Il permet de caractériser les allèles présents dans la population. C'est un outil plus efficace que le séquençage de l'ADN ribosomique 16S. La séquence de chaque fragment est comparée aux allèles qui ont été identifiés précédemment à ce locus, et des numéros d'allèles sont attribués pour les loci sélectionnés, discriminant ainsi différentes populations (Sharma *et al.*, 2020 [208]). Depuis les débuts de son développement dans les années 1970, le séquençage de génomes complets ou « Whole Genome Sequencing, WGS » est en compétition grandissante avec les autres méthodes de typage citées précédemment. D'ici quelques années, le WGS pourrait devenir le seul outil de diagnostic et de génétique des populations. Cette méthode permet d'inférer les relations évolutives entre isolats, est plus rapide et offre des résultats plus justes que les autres outils de typages moléculaires (Tehrani *et al.*, 2021 [210]).

3.5.1.5 Profils de virulence sur des lignées isogéniques de riz

La détermination des races (ou pathotypes) de *Xoo* est basée sur l'interaction compatible ou incompatible d'isolats de *Xoo* testés sur des cultivars spécifiques, contenant des gènes de résistance connus. Une série de lignées quasi-isogéniques (NILs) différant par la présence d'un ou plusieurs gènes de résistance contre *Xoo* (voir 4.2.1 «Gènes de l'immunité dans l'interaction *Xoo*/riz ») a été conçue dans le fond génétique du cultivar *indica* IR24 par Ogawa et al en 1991[211] à l'Institut international de recherche sur le riz (IRRI). Ces lignées appelées « IRBB » permettent de déterminer par infection le profil de virulence de l'isolat testé, afin de lui attribuer une race. Ainsi, les lignées IRBB ont permis la détection et la classification de nombreuses races de *Xoo* dans les plus grands pays rizicoles : plus de 63 races en Chine (Li *et al.*, 2009 [212]; Yang *et al.*, 2013 [213]), 11 en Inde (Mishra *et al.*, 2013 [214]), 8 pathotypes ont été signalés en Indonésie (Suryadi *et al.*, 2016 [215]), 12 races au Bangladesh (Rashid *et al.*, 2021 [216]) et quatre pathotypes dans le nord du Vietnam (Furuya *et al.*, 2012 [217]). En Afrique de l'Ouest, 9 races ont été signalées, 2 au Burkina faso, 2 au Niger et 7 au Mali (Gonzalez et al ., 2007 [218] ; Tékété *et al.*, 2020 [219]).

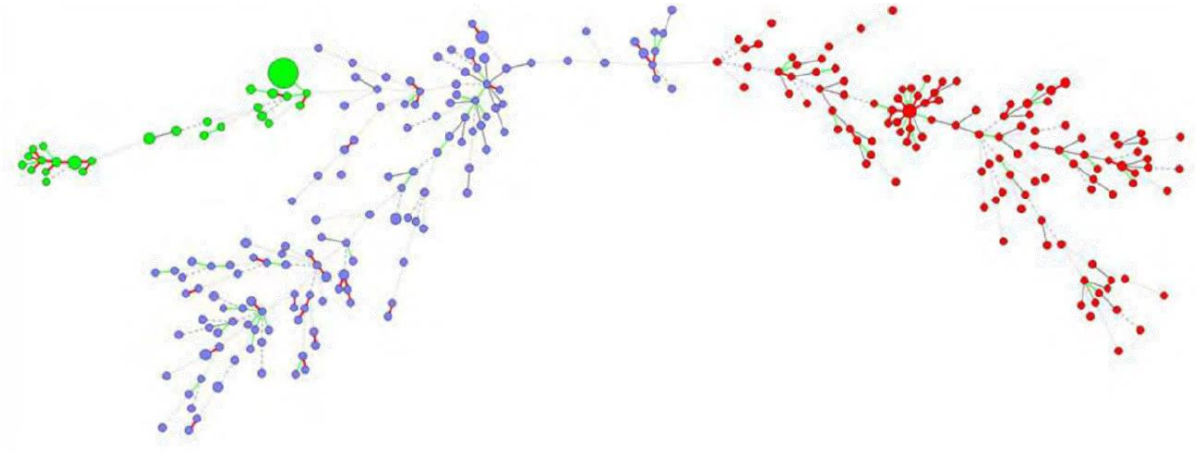


Figure 23 : Populations de *Xo* à l'échelle mondiale. Minimum spanning tree d'une collection mondiale de *Xo*. Vert : *Xoo* d'Afrique, bleu : *Xoc* et rouge : *Xoo* d'Asie. Les cercles sont proportionnels au nombre de souches par haplotype. (Poulin *et al*., 2014, unpublished).

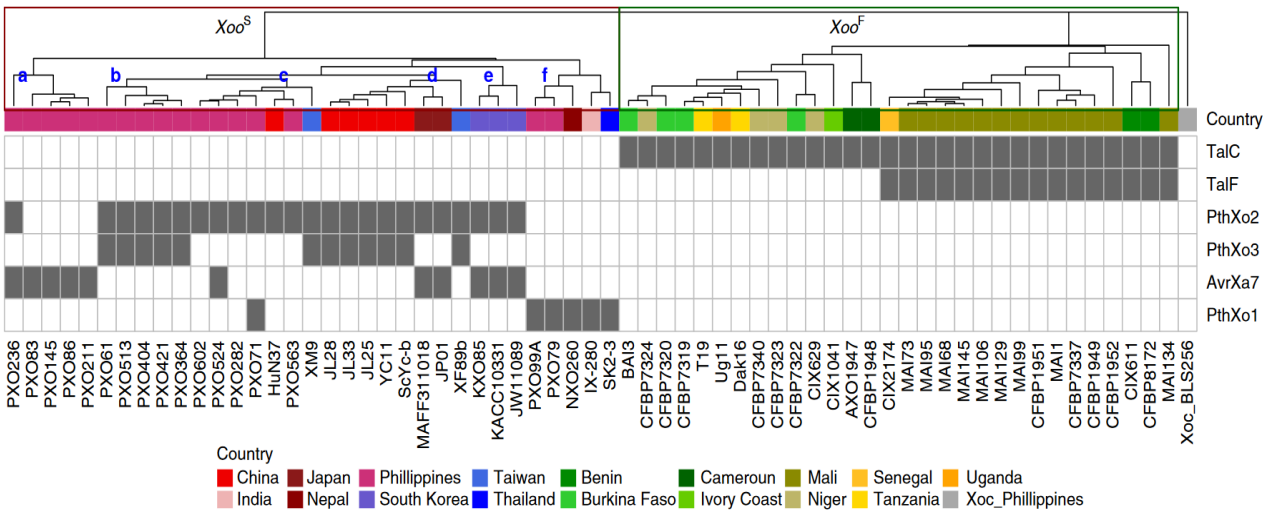


Figure 24 : Heatmap de la présence / absence des TALEs majeurs chez 63 *Xoo*. (tiré de Oliva et al., 2019 [189]). *Xoc* est utilisé comme outgroup.

3.5.2 Distinction des lignées asiatiques et africaines de *Xoo*

3.5.2.1 Différences génétiques et phénotypiques

Une analyse MLVA menée sur une collection de 338 souches de *Xoo* et *Xoc* issues d'Asie et d'Afrique, représentant une vingtaine de pays différents a révélé que les *Xoo* en provenance d'Afrique étaient plus proches des *Xoc* que des *Xoo* asiatiques (Figure 23). Ces résultats confirment de précédentes données obtenues via des profils de restriction. La richesse allélique des populations de *Xoo* africaines est également plus faible que celles des populations asiatiques, ce qui pourrait s'expliquer par une installation plus tardive du pathogène sur le territoire et / ou son essor dans les rizières du fait de changements dans les pratiques agricoles, laissant peu de temps à la diversification de celui-ci (Poulin *et al.*, 2015 [209] ; Gonzalez *et al.*, 2007 [218]). La mesure de l'identité moyenne des nucléotides (« Average Nucleotide Identity », ANI) entre des génomes de *Xoo* africaines et asiatiques met en évidence 98.2% d'identité entre les *Xoo* africaines et les *Xoc*, contre 97.8% d'identité partagée avec les *Xoo* asiatiques. Ce travail vient confirmer des données préliminaires obtenues lors du séquençage du premier génome africain complet, AXO1947, qui ne partageait que 97% d'identité avec PXO99^{A} et MAFF311018, deux souches asiatiques partageant entre elles 99% d'identité (Lang *et al.*, 2019 [123]; Huguet-Tapia *et al.*, 2016 [179]). Des phylogénies construites sur la base de SNPs sélectionnés dans les génomes complets de 63 souches de *Xoo* a clairement mis en évidence deux lignées distinctes : asiatiques (*Xoo* S) et africaines (*Xoo* F)(Oliva *et al.*, 2019 [189], Figure 24). Au niveau phénotypique, on relève quelques différences, comme une production plus faible de xanthomonadine (pigment jaune) mais plus forte d'EPS par les *Xoo* africaines par rapport aux asiatiques. Contrairement aux souches originaires d'Asie, les *Xoo* africaines à l'exception de la race A9, sont toutes contrôlés par le gène de résistance *Xa1* (Tekete *et al.*, 2020 [219]). Enfin, des données produites en 2007 par Gonzalez et al montrent que les *Xoo* africaines éllicitent une HR chez le tabac que ne déclenchent pas les *Xoo* asiatiques ou les *Xoc* (Gonzalez *et al.*, 2007 [218]).

3.5.2.2 Composition des TALomes

Lorsque l'on compare leurs TALomes, l'une des premières remarques est l'abondance moindre de TALEs chez les *Xoo* africaines. En effet, elles ne comptent que 9 TALEs, contre une quinzaine chez leurs cousines asiatiques. Les motifs RVD et les cibles de leurs TALEs diffèrent également, bien qu'elles partagent l'utilisation de certains gènes de sensibilité majeurs que sont les *SWEET*s du clade III (voir le 4.4.2.4 « Les gènes *SWEET* ciblés par des effecteurs TAL de *Xoo* ») ainsi que le facteur de transcription OsTFX1 (Yang et al 2007 [220]). Cependant, les souches africaines, visiblement moins diversifiées au niveau génétique que les souches asiatiques, ont une redondance fonctionnelle moindre quant à l'induction de ces cibles. Leurs TALE présentent également une richesse de combinaisons de répétitions moindre, comme en témoignent les distances génétiques plus courtes dans des alignements de répétitions de TALEs réalisés par Lang *et al* en 2019 [123]. En outre, des phylogénies issues des mêmes travaux et basées sur la région répétitive centrale des TALEs ont montré que les TALEs des *Xoo* africaines forment un groupe distinct, plus proche des TALEs de *Xoc*.

Une étude réalisée par Tran et ses collègues en 2018 [200] montre que les TALomes de trois *Xoo* africaines, BAI3 (Burkina Faso), MAI1 (Mali) et CFBP1947 (Cameroun), comportent 9 TALEs pouvant être répartis en 9 groupes. Ils sont hautement conservés et sont plus proches en termes de séquence des TALEs de *Xoc* que de ceux des *Xoo* asiatiques. Parmi ces 9 groupes, certains comme *talC* ne présentent aucun polymorphisme génétique. *talC*, qui induit le gène *SWEET14*, s'est montré être le seul TALE dont la mutation aboutissait à une perte de virulence forte (Yu *et al*., 2011 [159]). De plus, il

existe un autre TALE ayant *SWEET14* pour cible chez les *Xoo* africaines, *talF*, qui est retrouvé en plus de *talC* au sein de nombreuses souches (figure 24). Cette redondance et convergence fonctionnelle souligne l'importance de ce gène pour les *Xoo* africaines. Ces TALEs redondants ont potentiellement été maintenus en réponse à des variants de promoteurs *SWEET14* dans la population hôte. En effet, il existe un variant décrit, *xa41*(t), qui empêche la fixation de TalF. Autrement, *talF* pourrait être un effecteur ancestral ayant été perdu dans les souches BAI3 et CFBP1947 en raison de sa redondance avec *talC*. Sur les 9 groupes de TALEs africains, 5 groupes présentent des séquences RVDs variables. Le groupe de TalF est de ceux présentant la variabilité la plus forte. Il existe d'ailleurs un variant de ce TALE incapable d'induire *SWEET14* ayant été découvert chez BAI3.

4 ASPECTS MOLECULAIRES DE L'INTERACTION *XOO* – RIZ

4.1 Immunité innée : Résistances constitutives

Les défenses constitutives des végétaux ne nécessitent pas la perception du pathogène. Ces défenses sont d'ordre physique et chimique, et contribuent à limiter leur invasion et colonisation. La paroi des cellules végétale est la ligne de défense majeure contre les pathogènes fongiques et bactériens. Elle est constituée de différents polysaccharides, cellulose, glycanes et pectines. Les pectines forment des gels hydratés qui contribuent à "cimenter" les cellules et à réguler la teneur en eau de la paroi. De nombreuses parois cellulaires contiennent également de la lignine, un polymère de composés phénoliques très rigide, rendant les parois bien plus imperméables aux agents pathogènes. S'ajoutent à ces polymères des substances grasses, comme la cutine, la subérine et les cires (Freeman et Beattie, 2008 [221]). Les végétaux produisent de nombreux métabolites secondaires, dont la majeure partie est impliquée dans des mécanismes de défense. Ces composés appartiennent généralement à l'une des trois grandes classes chimiques suivantes : terpénoïdes, phénoliques et alcaloïdes. Parmi les composés phénoliques du riz, les flavanones et de nombreux flavonoides ont une action antifongique et antibactérienne. Une activité antibactérienne efficace contre *Xoo* a été rapportée pour quatre terpénoïdes du riz : le Linalool, le S-Limonène, le γ-Terpinene et le (E)-Nerolidol (Wang *et al.*, 2018 [222]).

4.2 L'immunité induite par les pathogènes.

4.2.1 La PTI, Immunité induite à large spectre non-spécifique.

4.2.1.1 Généralités et définitions

L'immunité non spécifique à large spectre des plantes, appelée PTI pour « Pathogen-Triggered Immunity », repose sur la détection de signatures moléculaires associées aux pathogènes dont la reconnaissance va déclencher les réactions de défenses de l'hôte. Ces motifs microbiens, que l'on nomme PAMPs pour « Pathogen-Associated Molecular Patterns » sont essentiels à la survie et la pathogénie des microorganismes, ce qui explique leur conservation (Chen et Ronald, 2011 [223]). La PTI est donc considérée comme une immunité robuste en termes évolutifs. Les PAMPs peuvent faire partie des composants de la paroi cellulaire, comme la chitine des champignons, les lipopolysaccharides ou le peptidoglycane bactérien, ou bien être des protéines spécifiques, comme la flagelline (principal composant du flagelle bactérien). En plus de cette reconnaissance directe du pathogène via les PAMPs, la plante est capable de détecter la présence de l'ennemei indirectement via les dégâts qu'il cause, notamment par la dégradation des parois cellulaires végétales. Ces molécules dérivées de l'hôte servant de signal d'alarme sont appelées DAMPs, pour « Damage-Associated Molecular Patterns ».

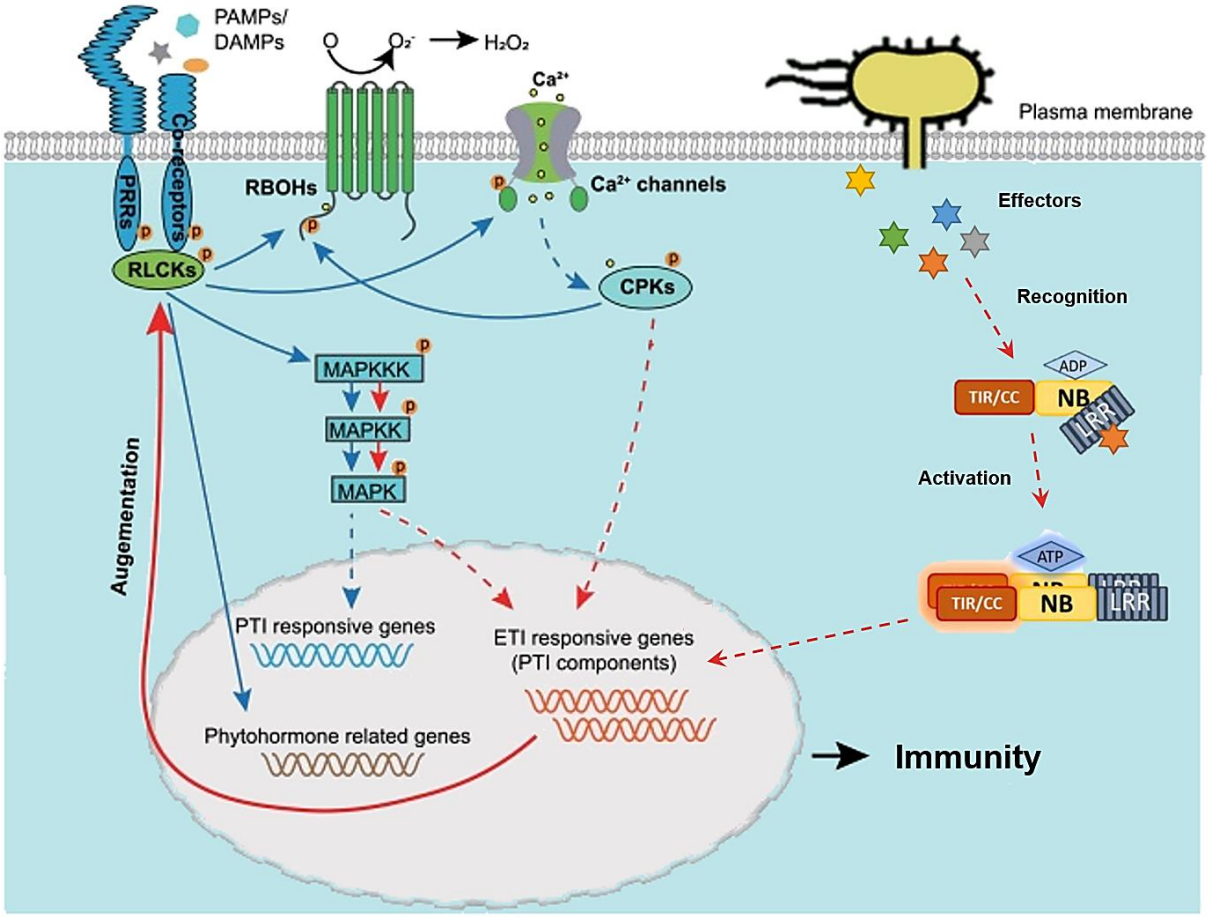


Figure 25 : Induction de l'ETI et de la PTI chez les plantes, et transduction du signal. (Adapté de Yuan *et al.*, 2021 [237])

Tous ces motifs vont être reconnus par des récepteurs spécifiques localisés à la surface des cellules végétales, les PRRs (« Pattern-Recognition Receptors »). Les PRRs sont des protéines transmembranaires possédant dans leur partie exposée vers l'apoplasme un domaine de reconnaissance pour leur ligand. Ce domaine peut-être de type LysM, EGF ou plus fréquemment de type LRR, composé de répétitions riches en leucines (« Leucine Rich Repeats », Figure 25). La présence ou l'absence dans le PRR d'un domaine kinase cytoplasmique, dont le rôle est de phosphoryler certains acides aminés des protéines avoisinantes, différencie les «Receptor Like Kinase » (RLKs) des « Receptor Like Proteins » (RLPs) respectivement. L'activation des PRRs a lieu seulement quelques minutes après la détection du ligand (Li *et al.*, 2020 [224]).

4.2.1.2 L'activation de la PTI chez le riz

Le génome du riz code pour presque 300 LRR-RLK. Parmi elles, les gènes *Xa21* et *Xa3*/*Xa26* confèrent une résistance robuste à *Xoo* (Ji *et al.*, 2018 [225]). *Xa21* est le premier gène de résistance à *Xoo* ayant été cloné, en 1995 depuis la variété *O. longistaminata* (Song *et al.*, 1995 [226]). Xa21 reconnait un peptide de 60 acides aminés sécrété par le T1SS de *Xoo*, appelée RaxX, ou Ax21 pour « Activateur de l'immunité médiée par Xa21 ». Ce peptide, qui est présent dans toutes les espèces de *Xanthomonas* séquencées, est également retrouvé chez *Xylella fastidiosa* (bactérie notamment responsable du complexe de dessèchement rapide de l'olivier) et *Stenotrophomonas maltophilia*, une bactérie ubiquitaire et pathogène opportuniste chez des patients immunodéprimés (Park *et al.*, 2010 [227] ; Pruitt *et al.*, 2015 [228]). La reconnaissance de RaxX active la voie de transduction des MAPK. La résistance médiée par Xa21 est active au stade adulte, et fonctionne contre 4 races des Philippines (Zhao *et al.*, 2009 [229] ; Yang *et al.*, 2022).

La protéine Xa3/Xa26 a une structure très similaire à celle de Xa21. Cette protéine confère une résistance a de nombreuses races de *Xoo* : 6 races des Philippines, 1 race chinoise, 2 races africaines et 6 races du Sri Lanka (Xiang *et al.*, 2006 [230]; Zhao *et al.*, 2009 [231]; Triplett *et al.*, 2015). Cependant la résistance est variable selon le fond génétique, les variétés *japonica* étant plus résistantes que les variétés *indica* (Yang *et al.*, 2022 [232]). Ceci pourrait être dû au niveau d'expression du gène, car il a été montré qu'il existe une relation dose-dépendante de la réaction de l'hôte : plus l'expression de Xa3/Xa26 est forte, plus la résistance est efficace (Zhao *et al.*, 2009 [229]). L'éliciteur d'Xa3/Xa26 a été isolé, mais son mode de reconnaissance reste méconnu (Li *et al.* 2004 [233]).

Le gène de résistance *Xa4* confère une résistance durable a de nombreuses races des Philippines. Il code pour un type spécifique de RLK associée à la paroi cellulaire appelée WAK (cell Wall Associated Kinase). L'élicitation de Xa4 par son ligand (encore inconnu), aboutit à une augmentation de la production d'enzymes de synthèse de la cellulose (CESA) et une réduction de l'expression de certaines expansines. Ces modifications vont conduire à un renforcement des parois cellulaires. L'induction d'*Xa4* a comme autre effet d'augmenter la résistance mécanique du chaume et réduit légèrement la hauteur de la plante, ce qui peut améliorer la résistance à la verse (Hu *et al.* 2017 [234]).

EF-Tu est un facteur d'élongation très abondant chez les bactéries. Chez le riz comme chez l'arabette et d'autres *Brassicaceae*, EF-Tu est un PAMP à l'origine de réactions immunitaires fortes. Bien que le récepteur associé à sa détection chez *A. thaliana* ait été trouvé (EFR pour « EF-Tu Receptor »), son équivalent chez le riz est encore inconnu (Furukawa *et al.*, 2014 [235]). Les lipopolysaccharides de *Xoo* déclenchent eux aussi la PTI chez leur hôte, par le biais d'un récepteur encore non-identifié (Desaki *et al.*, 2006 [236]). La flagelline de certains pathogènes comme *Acidovorax avenae* déclenche la PTI chez le riz. Le peptide est reconnu par OsFLS2, une LRR-RLK orthologue de la protéine FLS2, largement

caractérisée chez Arabidopsis (Gomez-Gomez et Boller, 2000 [238]). Cependant il semblerait que la flagelline de *Xoo* ait subit des changements lui permettant d'échapper au radar OsFLS2 du riz (Wang *et al.*, 2015 [239]).

4.2.2 Immunité spécifique induite par les Effecteurs (ETI).

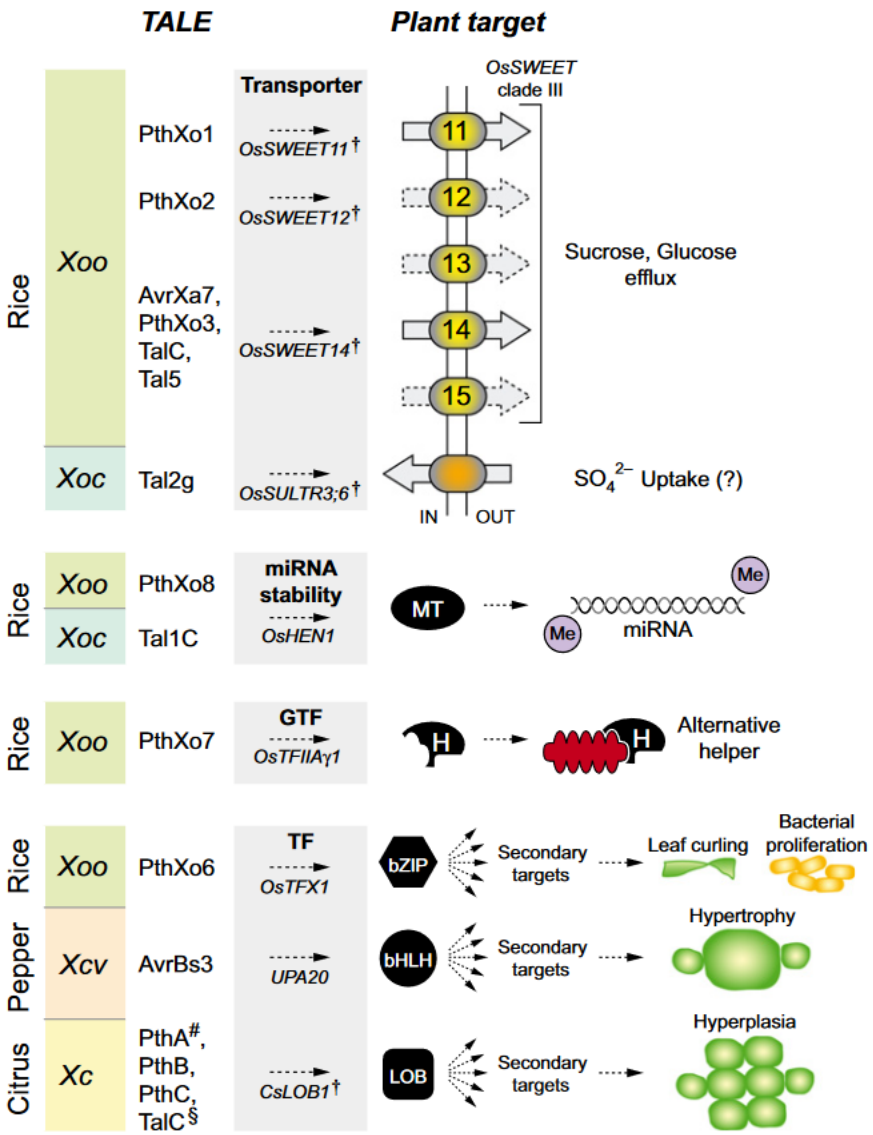
4.2.2.1 Généralités et définitions

Au cours de leur co-évolution avec leurs hôtes, les pathogènes ont déployé un arsenal d'effecteurs, leur permettant d'échapper à la PTI. C'est également le cas dans le pathosystème *Xoo*/riz comme décrit en 3.3.1.4 « Les effecteurs de type 3 ». Lorsque la PTI est franchie, l'ETI rentre alors en action. L'Effector-Triggered Immunity est une réaction immunitaire activée en réponse aux effecteurs sécrétés par le pathogène. De par sa reconnaissance d'effecteurs fortement variables d'une souche à l'autre au sein d'une même espèce, l'ETI est plus spécifique que la PTI (Chen et Ronald, 2011 [223]). Elle est liée à la PTI, car elle en partage certaines voies de signalisation et induit elle aussi l'expression de gènes de défense. De plus, de récentes données ont montré qu'il existait une synergie entre ces deux voies (Yuan *et al.*, 2021 [237]). Cependant, contrairement à la PTI, l'ETI peut entraîner l'activation d'une réponse violente, appelée Réponse Hypersensible (HR) correspondant à une mort cellulaire de type apoptotique. Ce processus rapide est supposé mettre fin à la prolifération des agents pathogènes (Li *et al.*, 2020 [224]). Les effecteurs de *Xoo* sont impliqués dans l'élicitation de trois mécanismes d'ETI différents : i) la détection par des protéines sentinelles de type NLR pour « Nucleotide-Binding - Leucine-Rich-Repeat Receptors » qui activeront des voies de défenses partagées avec la PTI, ii) le déclenchement de l'induction de gènes de résistance (dits gènes « Exécuteurs »), iii) l'interaction incompatible des TALEs majeurs avec leurs cibles. Dans ce dernier cas, l'ETI permet un gain de résistance à *Xoo* par perte de sensibilité à ses effecteurs. Il est intéressant de noter que la totalité des processus d'ETI connus en réponse à *Xoo* chez le riz impliquent les TALEs de la bactérie. Les voies d'activation de cette immunité seront donc traitées en 4.4.3 : « Résistance médiée par les TALEs, ou TAL-ETI ».

4.3 Transduction des signaux de la PTI et de l'ETI, et mise en place des réponses défensives

4.3.1 Voies de signalisation spécifiques à la PTI.

Une fois leur ligand reconnu, les PRRs vont permettre le déclenchement de plusieurs mécanismes de défenses et de signalisation. Les premiers mécanismes sont d'une part, l'activation de la voie des MAPK (« Mitogen-Activated Protein Kinases », voir 4.3.2), d'autre part, et de manière simultanée, l'induction d'un influx de calcium extracellulaire dans le cytoplasme de la cellule. Cet influx est médié par la phosphorylation de canaux membranaires Ca^{2+} spécifiques (Li *et al.*, 2020 [224], Figure 25). L'accumulation de Ca^{2+} va activer la voie des protéines kinases dépendantes du calcium, les CDPK (« Calcium Dependant protein Kinases ») (Yuan *et al.*, 2021 [237] ; Ramirez-Prado *et al.*, 2018 [240]). Ces CDPK vont induire la production d'espèces réactives de l'oxygène (ou ROS pour « Reactive Oxygen Species ») en activant une protéine membranaire de la famille des oxydases NADPH (« Nicotinamide Adenine Dinucleotide Phosphate »). La NADPH va produire de l'anion superoxyde (O_2^-), qui va à son tour engendrer d'autres ROS comme le peroxyde d'hydrogène (H_2O_2) et des oxydes de nitrate (NO) (Figure 25). Ces ROS ont, en plus de leur activité antimicrobienne directe, un rôle de messagers



<u>Figure 26 : Cibles caractérisées des TALEs de *Xanthomonas*.</u> Les cibles connues se répartissent en trois catégories : Transporteur, stabilité des miRNA, facteur de transcription. (Issu de Boch *et al.*, 2014).

secondaires de l'immunité, transmettant le signal en courte et longue distance (Lee *et al.*, 2020 [241]). Ils servent de signal de production de l'acide salyclique, de l'éthylene et du jasmonate, trois hormones clés dans l'immunité des plantes. Ces hormones vont notamment induire l'expression de protéines PR « Pathogenesis-Related », et l'accumulation de dépôts de callose (Nguyen *et al.*, 2021 [242] ; Ramirez-Prado *et al.*, 2018 [240]).

4.3.2 Signalisation des MAPK partagée par l'ETI et la PTI

Bien que les signaux et les récepteurs impliqués dans l'ETI et la PTI diffèrent, les voies de signalisation sous-jacentes permettant la transduction du signal d'alarme sont en partie partagées (Yuan *et al.*, 2021 [237]). C'est le cas des voies de signalisation des MAPK. La cascade MAPK transmet le signal d'agression via des phosphorylations successives de protéines kinases. Elle est composée au minimum de trois kinases : une MAPK, une MAPK kinase (MAPKK) et une MAPKK kinase (MAPKKK). 17 MAPK ont été identifiées dans le riz (Chen et Ronald, 2011 [223]). L'activation de la voie des MAPK a lieu dès 5 min après la perception du ligand par les PRRs, et un pic de phosphorylation est atteint dès 30 minutes après l'élicitation (Li *et al.*, 2020 [224]). La signalisation des MAPK va permettre l'induction de processus de défense robustes par l'activation massive de plusieurs familles de facteurs de transcription comme celle des Basic Helix Loop Helix (bHLH), des Basic-region leucine-Zipper (bZIP) et celle des WRKY (du nom du domaine protéique qu'ils contiennent) (Figure 25). Par exemple, parmi la centaine de facteurs de transcription WRKY présents dans le génome du riz, l'expression de WRKY53, WRKY28, WRKY62, WRKY71 et WRKY76 répond à l'infection par *Xoo* (Peng *et al.*, 2010 [243]). De même, OsTGA2, un bZIP, est lui aussi induit en réponse à la BLB (Moon *et al.*, 2018 [244]). Ces différents facteurs de transcription vont permettre la reprogrammation transcriptionnelle de nombreux gènes codant notamment des protéines exécutrices de défense, l'accumulation d'hormones de défense, le remodelage du cytosquelette, la régulation du sécrétome, la restructuration de la paroi cellulaire et la mécanique cellulaire (comme la fermeture des stomates) (Li *et al.*, 2020 [224]). Cette phase se déroule en quelques minutes à quelques heures après la détection du pathogène. Par exemple, des changements dans l'organisation des réseaux d'actine ont été détectés 15 minutes après élicitation de la PTI par de la chitine ou la flagelline (Li *et al.*, 2015 [245]). Lors des réactions de HR, on atteint un pic de mort cellulaire dans les 2 à 6h après inoculation d'une souche avirulente (c'est à une souche contenant au moins un effecteur déclenchant l'ETI) (Li *et al.*, 2020 [224]).

4.4 Rôle des effecteurs TALs dans l'interaction

4.4.1 Activation de gènes de sensibilité

4.4.1.1 Définition

Les TALEs constituent un excellent outil pour contrôler l'expression des gènes de l'hôte. Chez *Xanthomonas*, nombre d'entre eux sont connus pour cibler des gènes dont l'induction fait partie intégrante du développement de la maladie. Ces gènes contribuant à la pathogénie de *Xanthomonas* sont appelés gènes de Sensibilité (*S*). Leur induction permet par exemple de favoriser l'attraction et l'attachement de la bactérie à l'hôte, l'établissement d'un environnement favorable à la pénétration des tissus, la colonisation ou la dispersion du pathogène (Hutin *et al.*, 2015 [246]). Il existe notamment des gènes de sensibilité majeurs, c'est-à-dire des gènes dont l'induction est nécessaire à l'établissement de la maladie. Ces gènes appartiennent à une famille de transporteurs de sucres appelés SWEETs pour « Sugars Will Eventually be Exported Transporters» que nous décrirons en 4.4.2. Les autres gènes de sensibilité connus sont majoritairement des facteurs de transcription (Figure 26).

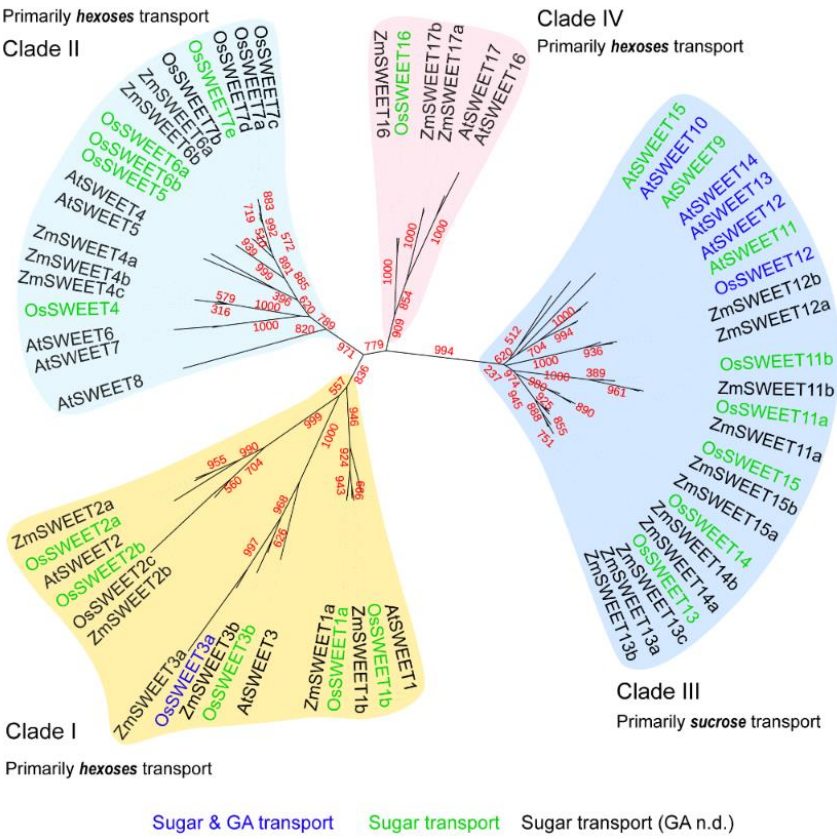


Figure 27 : Phylogénie des membres de la famille des gènes *SWEET* chez *Arabidopsis* (At), le riz (Os) et le maïs (Zm). Les gènes *SWEETs* présentés ici ont été mis en surbrillance en fonction de leur capacité de transport du saccharose/des gibérellines. La couleur bleue indique les *SWEETs* capables de transporter à la fois le saccharose/les hexoses et les gibérellines, tandis que la couleur verte montre les *SWEETs* capables uniquement de transporter le saccharose/les hexoses, enfin en noir sont représentés les *SWEETs* pour lesquels ce transport est inconnu. GA, acide gibbérellique ; n.d., non déterminé. (Tiré de Wu *et al*., 2021 [254])

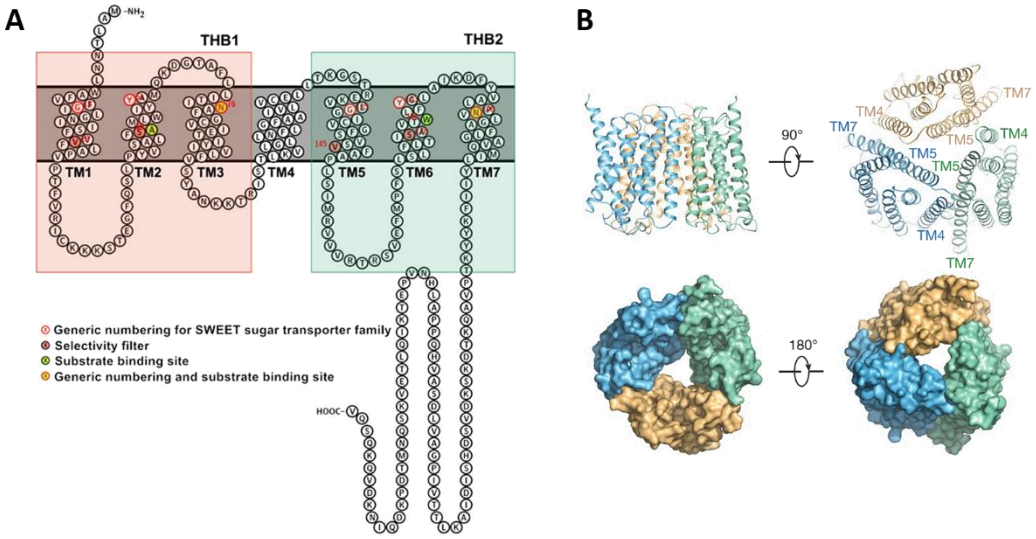


Figure 28 : Structure protéique des SWEETs. A. structure d'un monomère de SWEET. Diagramme en serpent de *AtSWEET13* avec les positions clés étiquetées (Image, Ji *et al*., 2022 [253]). B. structure tridimensionnelle d'un trimère de SWEET (Image, Tao *et al*., 2015 [256]).

4.4.1.2 Gènes de sensibilité non SWEET

Parmi les facteurs de transcription induits par des TALE lors de l'interaction de la plante hôte avec *Xanthomonas*, on retrouve des protéines de type bHLH (basic Helix – Loop Helix), dont l'induction conduit à un phénotype d'élargissement des cellules. Upa20, une bHLH induite par AvrBs3 de *Xanthomonas* campestris pv. vesicatoria conduit à une hypertrophie des cellules du poivron (Kay *et al*., 2007 [247]). bHLH3 et bHLH6, induites par AvrHah1 de *Xanthomonas* gardneri, entraînent des lésions aqueuses chez la tomate (Schwartz *et al.*, 2017 [248]). Un phénotype similaire a été observé chez les agrumes suite à l'induction du facteur de transcription CsLOB1 (Lateral Organ Boundaries) donnant lieu à l'apparition de chancres. *CsLOB1* est induit par différents TALEs retrouvés chez *Xanthomonas citri* pv. *citri* : Pth4, PthAw, PthA*, PthB, et PthC (Hu *et al.*, 2014 [249]). Chez *Xoo*, il existe différents TALEs induisant des facteurs de transcription. OsTFIIAg1, gène codant la sous-unité gamma du facteur général de transcription IIA, est la cible de PthXo7. L'induction d'OsTFIIAg1 est essentielle à la multiplication de *Xoo* sur les variétés de riz porteuses du gène de résistance *xa5* (Zhang *et al.*, 2022 [250]). *OsTFX1*, qui code un facteur de transcription bZIP, peut être induit par le TALE asiatique PthXo6 et le TALE africain TalB. TalB peut également activer un autre facteur de transcription, *OsERF#123* (Tran *et al.*, 2018 [200]). L'induction de ces trois gènes du riz favorise la croissance bactérienne et l'apparition de tâches aqueuses sur les feuilles. Tal8 de *Xanthomonas translucens* pv. *undulosa* (*Xtu*) active l'expression de *TaNCED_5BS*, un gène codant pour une enzyme régulant la vitesse de synthèse de l'acide abscissique chez le blé. L'augmentation de la concentration en acide abscissique dans les tissus infectés est supposée supprimer la voie de défense médiée par l'acide salicylique et réduire le niveau de transpiration de la plante, garantissant une humidité favorable à la multiplication de *Xtu* (Zhang *et al.*, 2022 [250]). Dans les interactions *Xoc* / riz, le transporteur de sulfate *OsSULTR3;6* est l'une des cibles de Tal2g. L'induction de ce transporteur augmente la quantité d'exsudats bactériens et favorise l'apparition de lésions (Cernadas *et al.*, 2014[251]).

4.4.2 Gènes de sensibilité majeurs : les transporteurs de sucrose de la famille *SWEET*

4.4.2.1 Distribution, structure et fonctionnement des transporteurs SWEET

Les SWEETs constituent la famille de transporteurs de sucres la plus récemment découverte (Chen *et al.*, 2010 [252]). Ils assurent le transport de mono et disaccharides, et sont présents dans tous les domaines du vivant. Chez les angiospermes, on dénombre environ une vingtaine de paralogues par espèce. Ils sont subdivisés en quatre clades, chacun présentant une spécificité de substrat. Les SWEETs des clades I et II transportent des hexoses, ceux du clade III transportent du sucrose, et ceux du clade IV, situés sur la membrane vacuolaire, transportent du fructose (Ji *et al.*, 2022 [253]). Certains SWEETs caractérisés chez Arabidopsis et le riz sont également capables de transporter des gibbérélines (Wu *et al*., 2021 [254], Figure 27). Les protéines SWEET utilisent le gradient de concentration des sucres au travers de la membrane pour transporter leurs substrats (Chen *et al.*, 2012 [255]). Ce mode de fonctionnement explique leur capacité à transporter le sucre de manière bi-directionelle. Les SWEET végétaux sont des membres de la famille MtN3/saliva (PF03083, Ji *et al*., 2022 [253]). Chaque sous-unité SWEET est divisée en deux domaines MtN3/saliva comptants chacun trois domaines transmembranaires (figure 28A). Pour pouvoir assurer le transport, au moins deux sous-unités SWEET (homologues ou hétérologues) doivent s'assembler pour former un oligomère (Tao *et al.*, 2015 [256] ; Xuan *et al.*, 2013 [257] , figure 28B). Chez le riz, on compte 23 gènes *SWEETs* (Wu *et al.*, 2021 [254] , figure 27).

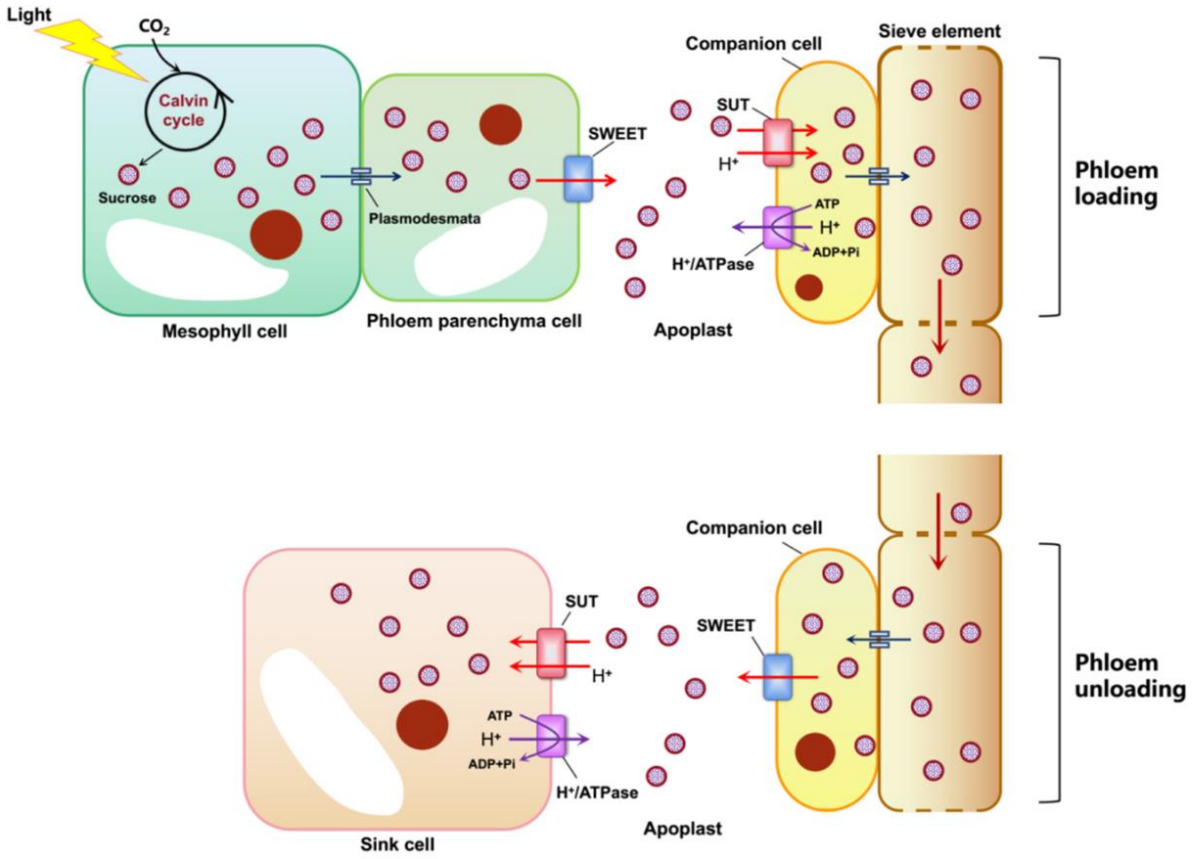


Figure 29 : Rôle des SWEET dans le chargement / déchargement du phloème. Lors du chargement du phloème, le saccharose photosynthétisé dans les cellules du mésophylle est amené vers les cellules du parenchyme du phloème par l'intermédiaire des plasmodesmes, puis transporté vers les apoplastes près du complexe cellules compagnes - éléments du tube criblé (SE-CC) par les SWEETs. Enfin, le saccharose est importé dans le SE-CC par le cotransporteur H+/saccharose SUT1/SUC2 pour être exporté vers les tissus récepteurs. Lors du déchargement, ce sont les SWEETs sur la membrane des cellules compagnes qui vont assurer l'export des sucres vers l'apoplasme (Issu de Ji *et al*., 2022 [253]).

4.4.2.2 Rôle des SWEET dans l'homéostasie du sucre de la plante : focus sur le riz

Afin d'assurer le ravitaillement en sucres dans la plante, le sucrose produit lors de la photosynthèse doit pouvoir être transporté sur de longues distances, des organes producteurs aux organes consommateurs. Pour effectuer ce trajet, le sucrose est pris en charge par le phloème. Le chargement et le déchargement du sucrose dans le phloème nécessite l'activité de différents types de transporteurs de sucres. Les premiers découverts furent les SUT, pour « Sucrose Uptake Transporters » (Riesmeier *et al.*, 1994 [258]). Ces synporteurs H^+/sucrose garantissent le chargement des cellules compagnes des tubes criblés du phloème (figure 29). Pour fonctionner, ils utilisent le gradient de protons fournit par des pompes ATPases situées sur les membranes des cellules importatrices. Le mécanisme par lequel le sucrose produit dans les cellules du mésophylle était exporté au niveau de l'apoplasme des cellules du parenchyme et des cellules compagnes était encore incompris avant la mise en évidence des SWEETs (Bezrutczyk *et al.*, 2018 [259]). Il a été découvert que le passage du sucrose des cellules du mésophylle vers celles du parenchyme suit la voie du plasmodesme, puis les SWEETs des cellules du parenchyme assurent l'export vers l'apoplaste (figure 29). Il est à noter qu'il existe sur les membranes d'autres symporteurs H^+ spécifiques des monosaccharides, les MSTs (« Monosaccharides Transporters »), dont le rôle est de ré-importer dans le cytoplasme les sucres simples présents dans l'apoplasme. En conclusion, ces prises en charge successives laissent imaginer qu'en temps normal, très peu de sucres sont vacants dans l'apoplasme. En régulant la distribution et le stockage des sucres, les SWEETs ont un rôle important dans la croissance et le développement des plantes. Chez le riz, ils sont impliqués dans divers processus physiologiques. Os*SWEET11* par exemple, est fortement exprimé dans les épillets et le pollen. Chez les plantes dont l'expression de Os*SWEET11* a été interrompue par « silencing », le développement du pollen est avorté, les rendant stériles ou semi-stériles (Yang *et al.*, 2006 [47]). Ce même gène est hautement exprimé dans les caryopses aux cotés de Os*SWEET14* et *OsSWEET15*. Les doubles mutants *Ossweet11 ;Ossweet14* et *Ossweet11*;*Ossweet15* présentent un taux de remplissage des grains fortement réduit (phénotype de « graines vides ») et une accumulation accrue d'amidon dans le péricarpe (Fei *et al.*, 2021 [260]). Les gènes *SWEETs* sont également impliqués dans la régulation du processus de sénescence. *OsSWEET5*, codant un transporteur de galactose, s'exprime dans les feuilles sénescentes. La surexpression de ce gène provoque un retard de croissance et une sénescence prématurée au stade de plantule (Zhou *et al.*, 2014 [261]).

4.4.2.3 Exploitation par les microorganismes

Nombreuses sont les interactions plante-microorganismes, qu'elles soient pathogènes ou bénéfiques, où l'on peut observer une implication des gènes de la famille *SWEET* (tableau 8). Dans le cas des systèmes plantes-pathogènes, l'observation de modifications dans l'expression de ces gènes peut s'accommoder de deux hypothèses sur le rôle des SWEETs dans l'interaction : l'hypothèse de la signalisation par les sucres, ou l'hypothèse de la nutrition. L'hypothèse de la nutrition est, semble-t-il, la plus parcimonieuse. Dans cette hypothèse, les pathogènes détourneraient le transport des sucres dans le but d'acquérir des ressources nutritives. Dans plusieurs interactions, des preuves expérimentales de l'acquisition des sucres dérivés de l'hôte par les pathogènes ont pu être produites. Par exemple lors d'attaques fongiques comme celle de l'oïdium du pois (*Erysiphe polygon*i f. sp. *pisi*, Aked *et al.*, 1993 [262]) ou dans le cas de l'interaction *Xoo* / riz (Sadoine *et al.*, 2022 [263]). Dans le cas de la seconde hypothèse, celle de la signalisation par les sucres, l'augmentation de sucres libres chez l'hôte servirait d'alarme et de carburant aux réactions immunitaires contre l'intrus. Certains travaux ont montré que l'augmentation artificielle du taux de sucres simples (hexoses), soit par l'expression transgénique d'invertases, soit par un apport exogène de solutions de sucres, augmente l'expression

Tableau 8 : Les différentes interactions (bénéfiques ou pathogènes) dans lesquelles les gènes *SWEET* sont impliqués.

Microorganisme	Hôte	Nom commun	Type d'interaction	Nom de l'interaction	Gènes SWEET impliqués	Elliciteur	Références
Pseudomonas simiae WCS417r	Arabidopsis thaliana	Arabette	bénéfique	PGPR	AtSWEET11, 12		Desrut et al., 2020
Glomus aggregatum	Glycine max	Soja	bénéfique	Mycorrhize à arbuscules	GmSWEET6, 15		Zhao et al., 2018
Rhizophagus irregularis	Glycine max	Soja	bénéfique	Mycorrhize à arbuscules	GmSWEET6, 15		Zhao et al., 2018
Sinorhizobium meliloti	Lotus japonicus	Lotier	bénéfique	Symbiose fixatrice d'azote	LjSWEET3	inconnu(s)	Sugiyama et al., 2017
Rhizophagus irregularis	Medicago truncatula	Luzerne tronquée	bénéfique	Mycorrhize à arbuscules	MtSWEET1b		An et al., 2019
Sinorhizobium meliloti	Medicago truncatula	Luzerne tronquée	bénéfique	Symbiose fixatrice d'azote	MtSWEET11		Kryvoruchko et al., 2016
Tulasnella calospora	Serapias vomeracea	Orchidée	bénéfique	Mycorrhize à arbuscules	SvNod9		Perotto et al. 2014
Rhizophagus irregularis	Solanum tuberosum	Pomme de terre	bénéfique	Mycorrhize à arbuscules	12 StSWEETs , 10 StSWEETs différents		Manck-Götzenberger et al.,
Golovinomyces cichoracearum	Arabidopsis thaliana	Arabette	pathogène	Oïdium	AtSWEET12		Chen et al., 2010
Pseudomonas syringae pv. tomato	Arabidopsis thaliana	Arabette	pathogène	Moucheture	AtSWEET4, 5, 7, 8, 10, 12, and 15		Chen et al., 2010
Pythium irregulare	Arabidopsis thaliana	Arabette	pathogène	Pourriture commune des racines	AtSWEET2, 10		Chen et al., 2015
Sclerotinia sclerotiorum	Brassica napus	Colza	pathogène	Pourriture blanche	BnSWEET9-2, 11		Jian et al., 2016
Plasmodiophora brassicae	Brassica rapa	Chou champêtre	pathogène	Hernie du chou	BrSWEET3b, 4b, 5b, 8, 9 *		Li et al., 2018
Xanthomonas citri ssp. citri	Citrus	Agrumes	pathogène	chancre bactérien	CsSWEET1	pthA4 and pthAw	Hu et al., 2014
Xanthomonas axonopodis pv. Manihotis	Manihot esculenta	Manioc	pathogène	Bactériose	MeSWEET10a	TAL20	Cohn et al., 2014
Fusarium oxysporum f.sp cubense	Musa acuminata L.	Bananier	pathogène	Fusariose	MaSWEET4c, 4d, 14h *		Miao et al., 2017
Rhizoctonia solani	Oryza sativa	Riz	pathogène	Brûlure de la gaine	OsSWEET11	inconnu(s)	Gao et al., 2018
Rhodococcus fascians	Pisum sativum L.	Pois	pathogène	Fasciation (galle du collet)	PsSW2a, 4, 5b, 7, 15a, 15b and 15d		Dhandapani et al., 2017
Botrytis cinerea	Solanum lycopersicum	Tomate	pathogène	Pourriture grise	SlSWEET15 , 21 other SlSWEETs	inconnu(s)	Asai et al., 2016
Meloidogyne incognita	Solanum lycopersicum	Tomate	pathogène	Galles	17 SlSWEETs différents	inconnu(s)	Zhao et al., 2018
Fusarium oxysporum f. sp. tuberosi	Solanum tuberosum	Pomme de terre	pathogène	Fusariose	StSWEET7a		Tamayo et al., 2022
Fusarium solani	Solanum tuberosum	Pomme de terre	pathogène	Fusariose	StSWEET7a		Tamayo et al., 2022
Puccinia graminis f. sp. tritici	Triticum aestivum	Blé	pathogène	Rouille noire	TaSWEET2b, 5a, 14a, 14g, and 14i		Gao et al., 2018
Botrytis cinerea	Vitis vinifera	Vigne	pathogène	Pourriture grise	VvSWEET4 ; VvSWEET7,15, 2a, VvSWEET10, 11, 17a,and17d		Chong et al., 2014 ; Breia et al., 2020
Colletotrichum higginsianum	Arabidopsis thaliana	Arabette	pathogène		AtSWEET11 (but not induced), 12		Gebauer et al., 2017
Fusarium oxysporum Schlecht. f. sp. batatas	Ipomoea batatas (L.) Lam	Patate douce	pathogène	Fusariose	IbSWEET10		Wang et al., 2017
Xanthomonas citri subsp. malvacearum	Gossypium hirsutum	Coton	pathogène	Bactériose	GhSWEET10	Avrb6	Cox et al. 2017
Xanthomonas campestrispv.vesicatoria	Capsicum annuum	Poivron	pathogène	Pourriture noire	UPA16	AvrBs3	Kay et al., 2009
Xanthomonas oryzae pv. oryzae	Oryza sativa	Riz	pathogène	Bactériose vasculaire	OsSWEET11, 13, 14	PthXo1, PthXo2a, PthXo2b, PthXo2c, AvrXa, PthXo3, TalC, TalF	Chu et al., 2006 ; Zhou et al., 2015 ; Yang and White, 2004 ; Antony et al., 2010 ; Streubel et al., 2013 ; Yu et al. 2011

de protéines de l'immunité (Pathogenesis Related Proteins) et la résistance à certains pathogènes comme *Magnaporthe* (Bezrutczyk *et al.*, 2018 [259]). Plus récemment, il a été observé qu'un double mutant *Atsweet11*;*Atsweet12* chez *Arabidopsis thaliana* voyait augmenter son taux d'hexoses libres et d'acide salicylique (AS), principal messager des réactions immunitaires systémiques acquises chez les végétaux (SAR), induisant l'expression de divers gènes de défense de la plante et sa résistance à *Colletotrichum higginsianum* (Gerbauer *et al.*, 2017 [264]). La capacité de certains transporteurs SWEET à transporter des gibbérélines pourrait également avoir un rôle dans l'immunité (voir 4.3.2.1). Ainsi, l'induction des gènes *SWEETs* pourrait soit être une prise de contrôle par le pathogène pour son propre bénéfice, soit constituer les prémices d'une réaction de défense de l'hôte. On peut également penser que l'augmentation de sucres libres dans l'apoplaste déclenchée par *Xoo* dans le cadre de l'hypothèse nutritive a pour effet secondaire d'amorcer les défenses de la plante, en déclenchant les cascades de signalisation et en fournissant l'énergie nécessaire au métabolisme immunitaire. La bactérie aurait donc tout intérêt utiliser rapidement les ressources acquises.

Dans le cadre du présent projet de thèse, nous privilégions l'hypothèse de la nutrition. En effet, les sucres sont naturellement peu disponibles à l'interface apoplasmique où se trouvent les pathogènes (voir 4.3.2.2). L'induction des gènes *SWEETs* leur permettrait de lever le verrou limitant leur croissance. De récents travaux par Sadoine *et al.* en 2022 [263] ont montré que lors de l'interaction *Xoo* / riz, des gènes bactériens d'internalisation et d'utilisation du sucrose nommés *sux* (« Sucrose Utilization in *Xanthomonas* ») étaient induits juste après l'induction du transporteur Os*SWEET11*a par la bactérie. Les mutants *sux* sont déficients dans leurs déplacements, leur essaimage, leur production d'EPS et leur formation de biofilms. Enfin, du côté de l'hôte, il est évident que de nombreux mécanismes sont en place pour restreindre au maximum l'accès aux sucres. Les systèmes SUT et MST, dont le rôle est de réimporter les sucres dans le milieu intracellulaire sont d'ailleurs induits lors de nombreuses infections (Bezrutczyk *et al.*, 2018 [259] ; Ji *et al.*, 2022 [253]).

4.4.2.4 Les gènes SWEET ciblés par des effecteurs TAL de Xoo

Les TALEs de *Xoo* ciblant des gènes *SWEET* du clade III sont appelés TALEs de virulence majeurs. En effet, bien qu'il y ait 23 gènes *SWEETs* chez le riz, l'induction de la plupart d'entre eux via des TALEs artificiels ne permet pas la prolifération bactérienne (Streubel *et al.*, 2013 [265]). Seuls les 6 transporteurs du clade III peuvent servir de gènes de sensibilité majeurs. Cependant, les analyses menées jusqu'à ce jour montrent que seuls les gènes Os*SWEET11*a, Os*SWEET13* et Os*SWEET14* sont retrouvés induits naturellement par *Xoo* (Streubel *et al.*, 2013 [265] ; Wu *et al.*, 2021 [254]). Le premier TALE de virulence majeur de *Xoo* découvert fut PthXo1, ciblant une séquence EBE dans le promoteur de Os*SWEET11*a (Yang *et al.*, 2006 [266]). Vinrent ensuite AvrXa7, PthXo3, TalC et TalF se liant tous trois au promoteur d'*OsSWEET14* (Antony *et al.*, 2010 [267] ; Streubel *et al.*, 2013 [265]). Enfin, PthXo2a, PthXo2b et PthXo2c peuvent reconnaître différentes séquences dans le promoteur de Os*SWEET13* (Zhou *et al.*, 2015 [268] ; Oliva *et al.*, 2019 [189]). Les différentes EBEs ciblées sont disponibles dans le tableau 9. Toutes les souches de *Xoo* virulentes isolées à partir de tissus infectés possèdent des TALEs induisant les *SWEETs* du clade III. Ces résultats soutiennent l'hypothèse selon laquelle il existe une convergence dans l'induction de ces gènes dans différentes régions du globe au cours de l'adaptation de *Xoo* au riz. L'expansion et la diversification des TALEs majeurs chez les *Xoo* asiatiques (lignée *Xoo* S, en figure 24) a été plus intense que chez les *Xoo* africaines (lignée *Xoo* F, en figure 24). En effet, les *Xoo* asiatiques sont capables d'induire 3 *OsSWEETs* différents, via treize variants des TALEs PthXo1, PthXo2a, b, c, AvrXa7 et PthXo3 (figure 30). Les souches africaines en revanche n'induisent que Os*SWEET14*, par l'utilisation de trois variants des TALE TalC et TalF (figure 30 ; Oliva *et al.*, 2019 [189]).

Tableau 9 : Liste des différentes EBEs trouvées dans les promoteurs *SWEET* du clade III. (Tiré de Oliva *et al.*, 2019 [189]).

Gène SWEET du clade III	TALe	Séquence EBE
OsSWEET11	PthXo1	GCATCTCCCCCTACTGTACACCAC
OsSWEET13	PthXo2(A)	ATAAAAGCACCACAACTCCCTT
	PthXo2(A)	ATATAAGCACCACAACTCCCTT
	PthXo2B	ATAAAGCACCACAACTCCCTTC
	PthXo2C	ATAAAGCACCACAACTCCCTTC
OsSWEET14	TalC	CATGCATGTCAGCAGCTGGTCAT
	PthXo3	ATATAAACCCCCTCCAACCAGGTGCTAAG
	AvrXa7	ATAAACCCCCTCCAACCAGGTGCTAA
	TalF	AAGCTCATCAAGCCTTCA

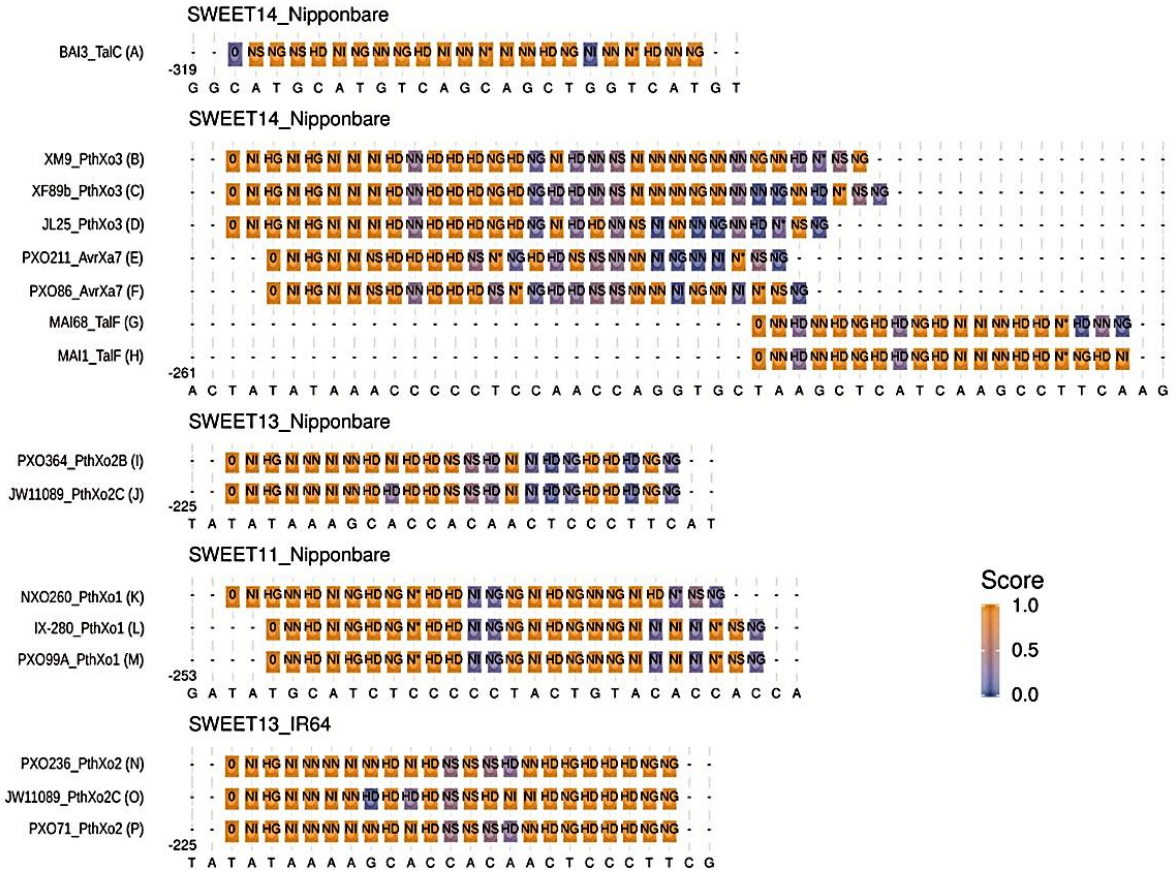


Figure 30 : Les différents variants de TALEs majeurs retrouvés chez *Xoo*. (Tiré de Oliva *et al.*, 2019 [189]).

4.4.3 Résistance induite par les TALEs (TAL-ETI)

4.4.3.1 Reconnaissance directe des TALEs par des récepteurs NLR

Le fer de lance de l'ETI est la reconnaissance des effecteurs de l'assaillant par des récepteurs spécifiques, les NLRs. Chez le riz, ces protéines sont constituées d'un domaine de répétitions riches en leucines (LRR, « Leucine-Reach Repeat ») qui assure la reconnaissance des effecteurs, d'un domaine NB («Nucleotide Binding») qui fournit l'énergie nécessaire à l'activation du NLR via la conversion d'ATP en ADP, et en N-terminal, un domaine CC « Coiled-Coil ». La reconnaissance de l'effecteur va changer la conformation du NLR, permettant le recrutement d'autres NLRs. La dimérisation ou l'oligomérisation des NLRs est indispensable à la signalisation (Nguyen *et al.*, 2021 [242], voir figure 25). La première protéine de ce type ayant un TALE pour ligand décrite fût Bs4, une NLR de la tomate, qui enclenche une HR en présence du TALE AvrBs4 de *X*. *campestris* pv. *vesicatoria* (Schornack et al., 2004 [269]). Chez le riz, 480 gènes codant des NLRs ont été prédits. Parmi eux, *Xa1* fut la première NLR impliquée dans la résistance à *Xoo* à être clonée. Non-exprimée en temps normal, elle est induite en présence du pathogène, ou lors de blessures (Jiang *et al.*, 2020 [270]). Xa1 est capable de reconnaître de nombreux TALEs de *Xoo* et *Xoc*, issus de souches africaines ou asiatiques. L'activation de Xa1 est indépendante du nombre de répétitions centrales (avec un minimum de 3.5 répétitions), ou de la composition en RVDs. Récemment, 3 allèles de *Xa1*, tous distribués sur le chromosome 4, ont été clonés avec succès (Ji *et al.*, 2020 [271]). Ces trois gènes *Xa2*/*Xa31*, *Xa14* et *Xa45*, fournissent eux aussi une résistance à large spectre selon le même mécanisme de reconnaissance. Cependant, certaines souches de *Xoo* ont mis au point des stratégies pour échapper à la reconnaissance de Xa1 et ses différents variants. Cette parade prend la forme de TALEs tronqués, appelés iTALEs pour « interfering TALEs», qui fonctionnent comme des leurres pour Xa1, laissant le champ libre aux TALEs pour opérer. Ces iTALEs ne possèdent pas de domaine d'activation de la transcription et ont une délétion de 58 acides aminés dans leur région N-terminale. Cependant, ils maintiennent leur NLS, qui semble essentiel au sabordage de Xa1 (Ji *et al.*, 2018 [225]). Ils sont présents chez les *Xoc* et les *Xoo* asiatiques, mais n'ont jusqu'alors pas étés décrits chez des *Xoo* africaines (tableau 10)(Ji *et al.*,2020 [271]). Une autre NLR capable de reconnaître les TALEs et dont l'activité est enrayée par les iTALEs est la protéine Xo1, découverte dans la variété Carolina Gold (Read *et al.*, 2016 [272]). Bien que cela reste à prouver, de nombreuses données laissent penser que Xo1 ne serait en fait qu'un autre allèle de Xa1 (Ji *et al.*,2020 [271]).

4.4.3.2 Gènes de résistance de type exécuteur (E)

Les gènes dits exécuteurs (*E*), sont des gènes dominants dont le produit est directement à l'origine de la résistance. Ces gènes ne sont pas exprimés en l'absence du pathogène, mais sont induits en présence de certains TALEs pouvant se fixer sur l'EBE contenue dans leur promoteur (Jiang *et al.*, 2020 [270]). Les gènes exécuteurs du riz codent pour des petites protéines dotées de plusieurs domaines transmembranaires. Cinq paires TALE / gènes *E* ont été clonées chez le riz : *avrXa27*/*Xa27*, *avrXa10*/*Xa10*, *avrXa23*/*Xa23*, et plus récemment, deux paires associant *Xa7*, *avrXa7*/*Xa7* et *pthXo3*/*Xa7* (Nowack *et al*., 2022 [152] ; Chen *et al*., 2021 [273]).

Les TALEs AvrXa7 et PthXo3 ciblent tous deux *OsSWEET14* comme gène *S* majeur et reconnaissent chacun une EBE sur le promoteur de *Xa7*, ce qui déclenche la réaction immunitaire. *Xa7* code pour une protéine de 113 aa, localisée sur la membrane du réticulum endoplasmique où elle s'ancre via deux domaines transmembranaires (Wang *et al*., 2021 [274]). XA7 partage peu de similitudes de séquence avec les protéines E connues et n'a aucun homologue dans les bases de données végétales actuelles, ce qui suggère un nouveau type de gène *E* (Chen *et al*., 2021 [273]). Étonnamment, aucun gène de sensibilité

Tableau 10 : Souches séquencées contenant des iTALEs. (supplemental data de Ji *et al.*, 2020 [271])

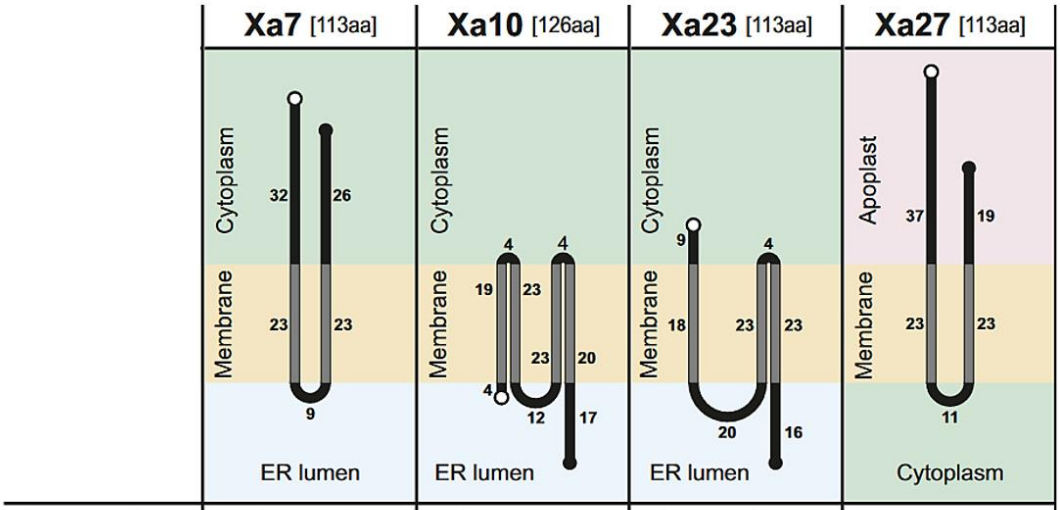
	Strain names	NCBI accession ID	Origin of strain	Country: District	Number of iTALE genes containing the 5' seq (1-690 bp)
Xoo	HuN37	CP031456.1	Asia	China: Hunan	2
	JL25	CP031457.1	Asia	China: Jilin	1
	JL28	CP031458.1	Asia	China: Jilin	1
	JL33	CP031459.1	Asia	China: Jilin	1
	K74-1	KP711411.1	Asia	China	1
	K74-15	KP711425.1	Asia	China	1
	OS198	CP031461.1	Asia	China: Guangxi	2
	XF89b	CP011532.1	Asia	China: Taiwan	1
	XM9	CP020334.1	Asia	China: Taiwan	2
	YC11	CP031464.1	Asia	China: Jiangsu	1
	YN24	CP018089.1	Asia	China: Yunan	2
	ScYc-b	CP031469.1	Asia	China: Sichuan	1
	BXO1	CP033201.1	Asia	India	2
	ICMP3125	CP031697.1	Asia	India	2
	IX-280	CP019226.1	Asia	India: Andhra Pradesh, Tanuku	2
	JP01	CP031460.1	Asia	Japan	1
	MAFF311018	AP008229.1	Asia	Japan	1
	NX0260	CP033192.1	Asia	Nepal	2
	PXO61	CP033187.1	Asia	Philippines	2
	PXO71	CP013670.1	Asia	Philippines: Palawan	2
	PXO79	CP031462.1	Asia	Philippines	2
	PXO83	CP012947.1	Asia	Philippines: Nueva Ecija	2
	PXO86	CP007166.1	Asia	Philippines	2
	PXO99A	CP000967.2	Asia	Philippines	2
	PXO142	CP031698.1	Asia	Philippines	2
	PXO145	CP013961.1	Asia	Philippines: Bontoc	2
	PXO211	CP013674.1	Asia	Philippines: Ifugao	1
	PXO236	CP013675.1	Asia	Philippines: Ifugao	1
	PXO282	CP013676.1	Asia	Philippines: Nueva Vizcaya	1
	PXO364	CP033191.1	Asia	Philippines	2
	PXO404	CP033190.1	Asia	Philippines	2
	PXO421	CP033189.1	Asia	Philippines	2
	PXO513	CP033188.1	Asia	Philippines	2
	PXO524	CP013677.1	Asia	Philippines: Laguna	2
	PXO563	CP013678.1	Asia	Philippines: Quezon	2
	PXO602	CP013679.1	Asia	Philippines: Laguna	1
	JW11089	CP033193.1	Asia	South Korea	1
	KACC10331	AE013598.1	Asia	South Korea	0
	KXO85	CP033197.1	Asia	South Korea	0
	SK2-3	CP019515.1	Asia	Thailand: Sukholthai Province	2
	AUST2013	CP033196.1	Oceanica	Australia	1
	CIAT1185	CP033194.1	South America	Colombia	2
	CFBP8172	CP033173.1	Africa	Benin: Tanguita	0
	BAI23	CP036256.1	Africa	Burkina Faso	0
	BAI3	CP025610.1	Africa	Burkina Faso	0
	CFBP7319	CP033181.1	Africa	Burkina Faso	0
	CFBP7320	CP033186.1	Africa	Burkina Faso	0
	CFBP7321	CP033180.1	Africa	Burkina Faso	0
	CFBP7322	CP033179.1	Africa	Burkina Faso	0
	CIX298	CP036378.1	Africa	Burkina Faso	0
	AXO1947	CP013666.1	Africa	Cameroon	0
	CFBP1948	CP033185.1	Africa	Cameroon	0
	CFBP1949	CP033184.1	Africa	Mali: Niono Kogoni	0
	CFBP1951	CP033183.1	Africa	Mali: Kayes	0
	CFBP1952	CP033182.1	Africa	Mali: Niono Diabali	0
	CFBP7325	CP033176.1	Africa	Mali: Sgou Niono	0
	CFBP7337	CP033175.1	Africa	Mali: Sgou Molodo	0
	MAI1	CP025609.1	Africa	Mali	0
	MAI106	CP019089.1	Africa	Mali	0
	MAI129	CP019090.1	Africa	Mali	0
	MAI134	CP019091.1	Africa	Mali	0
	MAI145	CP019092.1	Africa	Mali	0
	MAI68	CP019085.1	Africa	Mali	0
	MAI73	CP019086.1	Africa	Mali	0
	MAI95	CP019087.1	Africa	Mali	0
	MAI99	CP019088.1	Africa	Mali	0
	CFBP7323	CP033178.1	Africa	Niger	0
	CFBP7324	CP033177.1	Africa	Niger	0
	CFBP7340	CP033174.1	Africa	Niger: Bonfeba	0
	CIX2374	CP036377.1	Africa	Senegal	0
	Dak16	CP033172.1	Africa	Tanzania	0
	T19	CP033171.1	Africa	Tanzania	0
	BB151-3	CP036255.1	Africa	Uganda: Doho	0
	BB156-2	CP036254.1	Africa	Uganda: Doho	0
	Ug11	CP033170.1	Africa	Uganda	0
Xoc	O-9	CP045912.1	Asia	China, Henan	2
	GX01	CP043403.1	Asia	China, Guangxi	2
	B8-12	CP011955.1	Asia	China	1
	JSB2-24	KF447887.1	Asia	China: Jiangsu	1
	L8	CP011960.1	Asia	China	1
	RS105	CP011961.1	Asia	China	1
	BXOR1	CP011957.1	Asia	India	2
	CFBP2286	CP011962.1	Asia	Malaysia	1
	BLS256	CP003057.2	Asia	Philippines	1
	BLS279	CP011956.1	Asia	Philippines	1
	CFBP7331	CP011958.1	Africa	Mali	1
	CFBP7341	CP011959.1	Africa	Burkina Faso	1
	CFBP7342	CP007221.1	Africa	Burkina Faso	0

associé à AvrXa10, AvrXa23 ou AvrXa27 n'a été signalé dans les cultivars compatibles (Zhang *et al.*, 2015 [275]). *Xa27* a été le premier à avoir été cloné, depuis la variété sauvage *O. minuta*. Une fois exprimée, la protéine se localise au niveau de l'apoplasme sur la membrane cytoplasmique et conduit à l'inhibition de la croissance bactérienne (Figure 30). Des plantes exprimant *Xa27* de manière constitutive montrent un épaississement des parois des tissus vasculaires. La résistance conférée par XA27 est affectée par le stade de développement de la plante, atteignant un pic à 5 semaines après semis (Gu *et al.*, 2005 [276]). *Xa10* et *Xa23* sont situés proches l'un de l'autre sur chromosome 11 et produisent des protéines similaires à 50%. Toutes deux sont localisées au niveau de la membrane du réticulum endoplasmique et peuvent induire une HR dans des plantes non hôtes telles que *Nicotiana benthamiana* (Ji *et al.* 2018 [225]). XA10 confère la résistance du riz contre plusieurs races de *Xoo* des Philippines. Il a été montré que la protéine XA10 forme des hexamères sur la membrane du réticulum endoplasmique (figure 31), induisant une fuite de calcium dans le cytosol, ce qui a pour effet de déclencher la mort cellulaire (HR). Des lésions sont d'ailleurs visibles sur des plantes exprimant constitutivement *Xa10* (avec une intensité faible, aucun transformant à expression forte n'ayant pu être obtenu, Tian *et al.* 2014 [277]). *Xa23* isolé d'*O. rufipogon*, un autre riz sauvage, confère une résistance à très large spectre contre *Xoo* et n'est pas soumis à un contrôle développemental. Le spectre d'efficacité de *Xa23* s'explique par la répartition étendue de AvrXa23 chez *Xoo* (Wang *et al.* 2015 [278]).

4.4.3.3 Résistance par perte de sensibilité

Comme expliqué dans la partie 4.4.1 « Activation de gènes de sensibilité », les TALEs permettent à la bactérie de contrôler l'expression de gènes de l'hôte pour son propre bénéfice, notamment par l'utilisation de gènes de sensibilité (*S*). Il apparaît donc logique qu'au cours de la co-évolution entre *Xoo* et son hôte, des mutations spontanées apparues dans ces gènes *S* aient apporté un avantage sélectif aux variétés de riz en disposant.

C'est le cas pour chacun des trois gènes de sensibilité majeurs, *OsSWEET11* (ou *Xa13*), *OsSWEET13* (ou *Xa25*) et *OsSWEET14* (ou *Xa41*), codant pour des transporteurs de sucrose de la famille SWEET (voir 4.4.2.4). En effet, des mutations naturelles dans les séquences EBEs retrouvées sur les promoteurs de ces trois gènes ont donné lieu à l'apparition d'allèles de résistance récessifs, respectivement *xa13*, *xa25* et *xa41(t)*. Ces variations dans les EBEs empêchent la fixation des TALEs correspondants, rendant donc impossible l'induction du gène ciblé. Ce type de résistance est appelé résistance par perte de sensibilité. Le premier allèle de ce type à avoir été décrit fut *xa13* par Chu et ses collègues en 2006 [279]. Il présente une variation dans la séquence reconnue par Pthxo1, un TALE largement distribué en Asie. De même, des allèles du promoteur de *Xa25* ont été découverts dans les variétés Minghui 63, Nipponbare et Kitaake (Chen *et al.* 2002 [280]; Liu *et al.*, 2011 [281] ; Zhou *et al.* 2015 [268]). Ces différents polymorphismes, regroupés comme différents allèles *xa25*, ont conduit à l'apparition de plusieurs variants de PthXo2. Basé sur ces découvertes, Hutin et ses collègues [282] ont recherché du polymorphisme dans le promoteur de *OsSWEET14*/*Xa41* en réalisant le criblage moléculaire d'une grande collection d'accessions de riz africains. Une délétion de 18 pb dans les séquences reconnues par AvrXa7, PthXo3 et TalF (deux TALEs asiatiques et un TALE africain) a été détectée dans des accessions du riz sauvage africain *O. barthii*. Cet allèle, nommé *xa41(t)* confère une résistance robuste contre les souches asiatiques de *Xoo* dépendantes de AvrXa7 ou PthXo3 mais pas contre les *Xoo* Africaines, qui contiennent un TALE redondant à TalF : TalC (voir 3.5.2.2 « Composition des TALomes »).



	Xa7 [113aa]	Xa10 [126aa]	Xa23 [113aa]	Xa27 [113aa]
Executor *R* gene donor species	*Oryza sativa* DV85	*Oryza sativa* Cas 209	*Oryzae rufipogon*	*Oryza minuta* Acc. 101141
Distribution	*~10% of Oryza species*	*Oryza*	*Oryza*	*Oryza*
Matching TALEs	AvrXa7 (*Xoo* PXO86) PthXo3 (*Xoo* PXO61)	AvrXa10 (*Xoo* PXO86)	AvrXa23 (*Xoo* PXO99A)	AvrXa27 (*Xoo* PXO99A)

Figure 31 : Structure des gènes Exécuteurs impliqués dans la résistance à *Xoo*. Caractéristiques structurelles et fonctionnelles des gènes E spécifiques à l'interaction *Xoo*/riz. La taille de chaque protéine est indiquée entre crochets. Les représentations montrent la topologie et la localisation subcellulaire prédites des protéines E. Les cercles blancs et noirs terminaux indiquent les extrémités N et C, respectivement. (Adapté de Nowack *et al*., 2022 [152]).

Enfin, un allèle de résistance par perte de sensibilité d'un autre type a été découvert prenant la forme d'une mutation dans le gène de sensibilité TFIIAg5 codant pour la sous-unité gamma du facteur de transcription général TFIIA. La perte d'affinité des TALEs pour la protéine codée par l'allèle *xa5* diminue grandement l'induction des différentes cibles de ceux-ci (Jiang *et al.*, 2006 [283] ; Han *et al*., 2020 [284]).

4.4.3.4 Gènes de résistance dans l'interaction Xoo/riz

À ce jour, 44 gènes de résistance (*R*) à *Xoo* ont été identifiés chez le riz. Douze d'entre eux ont été clonés et caractérisés, à savoir *Xa1*, *Xa3*/*Xa26*, *Xa4*, *xa5*, *Xa10*, *xa13*, *Xa21*, *Xa23*, *xa25*, *Xa27* et *xa41*. Ces gènes *R* fournissent deux types de résistance : l'induction de résistances non-spécifiques (PTI) et des résistances spécifiques à l'action des TALEs. Certains de ces gènes sont employés depuis longtemps en agriculture dans la mise au point de variétés résistantes, comme *Xa3*/*Xa26* et *Xa4*, qui sont intégrés depuis les années 1970 dans la plupart des variétés hybrides Chinoises (Jiang *et al.*, 2020 [270]).

5 INTRODUCTION AU SUJET DE THESE

5.1 Contexte

Les membres du clade III de la famille des transporteurs SWEET sont les principaux gènes de sensibilité à *Xoo*. Du polymorphisme naturel dans les EBEs des promoteurs de ces gènes empêche la fixation des TALEs correspondants, ce qui annule leur induction par *Xoo*. Ces allèles "non-réceptifs" confèrent un phénotype de résistance récessive, prévenant la formation de symptômes et réduisant la prolifération bactérienne *in planta*. La création de résistance par ingénierie génétique est un moyen efficace, économique et écologique de garantir la sécurité alimentaire et financière des pays producteurs de riz. Dans le cadre du projet Healthy Crops (https://www.healthycrops.org) financé par la Fondation Bill et Melinda Gates, nous avons généré via la technologie CRISPR-Cas9 des lignées poly-mutantes éditées au niveau des 5 EBEs majeurs des gènes *SWEET11*, *SWEET13* et *SWEET14*. Ce projet intitulé « Transformative strategy for controlling rice disease in developing countries Phase II » est porté par Wolf Frommer en Allemagne (HHU), en collaboration avec les universités du Missouri et de Floride, l'IRRI aux Philippines, le CIAT en Colombie et notre UMR (PHIM, IRD).

Bien que les lignées éditées créées soient complètement résistantes aux *Xoo* d'origine asiatique, les souches africaines conservent un certain niveau de virulence sur ces plantes et seules des lignées doublement inactivées dans les gènes *SWEET13* et *SWEET14* sont totalement résistantes à ces souches. Ces résultats ont fait l'objet d'un article paru dans la revue *Nature Biotechnologies* en 2019, dans lequel je suis co-auteure (voir annexe 2).

Le but principal de cette thèse était de déchiffrer les mécanismes de virulence spécifiques des souches africaines de *Xoo*. Ces populations restant encore méconnues, notre premier objectif a été de caractériser leur structure et leur diversité génétique.

Afin de comprendre les résultats surprenants obtenus sur les lignées éditées du projet Healthy Crops, notre second objectif a été de décortiquer la dépendance des *Xoo* africaines vis à vis des gènes *SWEETs*.

Enfin, nous avons tenté de mettre à jour les bases génétiques à l'origine du phénotype spécifique des *Xoo* africaines.

5.2 Caractérisation des *Xoo* africaines

Pour maximiser l'efficacité du déploiement des gènes de résistance au BLB, il est crucial d'avoir une connaissance approfondie de la structure génétique et de la distribution des différentes populations de *Xoo*. Or, bien que l'Afrique Sub-Saharienne soit fortement touchée par la bactériose vasculaire du riz, les *Xoo* africaines sont bien moins connues que les souches asiatiques. On peut supposer que cet écart vient en partie de l'invasion apparemment plus tardive des cultures de riz africaines par la bactériose vasculaire mais aussi de l'importance plus faible de la riziculture sur le continent. De par ses partenariats de longue date en Afrique de l'ouest, et notamment grâce au travail de Valérie Verdier depuis 2003, l'IRD possède désormais une riche collection de souches de *Xoo*, originaire de 10 pays africains: Burkina Faso, Bénin, Sénégal, Côte-d'Ivoire, Ouganda, Niger, Mali, Cameroun et plus récemment Tanzanie et Madagascar.

Nous avons sélectionné un panel de 333 souches représentant chacun de ces pays, afin de réaliser une première analyse de diversité via une approche de typage moléculaire MLVA. Nos résultats nous ont permis d'apprécier la structure globale de nos différentes populations de *Xoo* africaines, et notamment

permis de mettre en évidence le caractère exclusif des populations récemment apparues en Afrique de l'Est (Tanzanie, Madagascar). Le séquençage et l'analyse des génomes de 41 souches de ce panel

nous a confirmé que ces souches étaient très éloignées génétiquement des souches de *Xoo* d'Afrique de l'Ouest, mais étroitement liées aux *Xoo* asiatiques.

Par ailleurs, ces données nous ont permis de découvrir une souche de *Xoo* africaine au génome particulier, dont nous avons fait une analyse plus poussée dans un article paru dans la revue Molecular Plant-Microbe Interactions.

Nous avons évalué la menace représentée par les nouvelles populations d'Afrique de l'Est pour la production de riz africain, et cherché à mettre au point de nouvelles variétés de riz afin de contrôler ces souches. Ces résultats, associés à d'autres générés par le consortium Healthy Crops, ont été à l'origine de l'écriture d'un article qui a été soumis à la revue Nature Communications.

Les résultats obtenus pour ce premier axe de réflexion et les articles en résultant sont présentés dans les Chapitres I et II.

5.3 Décoder la dépendance des *Xoo* africaines vis-à-vis de l'utilisation des *SWEETs*.

L'utilisation des transporteurs SWEET par les *Xoo* est considérée comme une condition nécessaire au développement de la maladie. Nous savons par exemple que lorsque l'induction des gènes *SWEETs* utilisés par les souches de *Xoo* asiatiques est abrogée, ces dernières n'engendrent plus de lésions (absence de développement de la maladie et baisse de la multiplication *in planta*). En revanche, les *Xoo* africaines semblent capables de maintenir la pression infectieuse malgré l'impossibilité d'induire la transcription de *SWEET14*, connu comme le seul transporteur SWEET exploité par ces souches. Cette découverte initie un questionnement sur la dépendance des souches africaines de *Xoo* vis à vis de l'utilisation des transporteurs SWEET. Sont-elles capables d'induire des *SWEETs* du clade III autre que *SWEET14* ? Ou existe-t-il une intervention post-transcriptionnelle dont le résultat serait l'accumulation de protéines SWEET ? Les souches africaines ont-elles un métabolisme des sucres plus performant ?

Afin de répondre à toutes ces interrogations, nous avons établi et testé de nombreuse hypothèses, par diverses expérimentations de phénotypages, biochimie et biologie moléculaire dont les résultats sont présentés dans le Chapitre III.

5.4 Recherche des bases génétiques de la spécificité des *Xoo* africaines.

Nous avons nommé SWIIVA (pour *SWEET* Induction Independant Virulence of African *Xoo*) l'aptitude des *Xoo* africaines à maintenir de la virulence en l'absence d'induction de gène *SWEET* du clade III. Cette capacité pourrait rester reliée à l'utilisation des SWEETs, ou leur être totalement indépendante. Ainsi, nous avons recherché les bases génétiques de SWIIVA chez *Xoo*. Plus particulièrement, nous nous sommes intéressés à une zone de 35 kb contenant le gène *talC* (TALE inducteur de *SWEET14*) que nous avons appelé *RegtalC*.

Une étude approfondie de cette région génomique est présentée dans le Chapitre III.

CHAPITRE I : L'Etude de la diversité des populations de *Xoo* Africaines révèle un clivage Est – Ouest.

Chapitre basé sur les résultats de deux articles en tant que premier auteur et une analyse de génomes.

Titre de l'article 1 : **« Molecular epidemiology by MLVA genotyping of the Rice Pathogen *Xanthomonas oryzae* pv. *oryzae* reveals two different lineages on the African continent ».** Il correspond à la première version d'un manuscrit qui sera soumis dans une revue à comité de lecture dans les prochains mois.

Titre de l'article 2 : **« The complete genome resource of the *Xanthomonas oryzae* pv. *oryzae* strain CIX2779 includes the first sequence of a plasmid for an African representative of this rice pathogen»,** accepté dans la revue ***Molecular Plant-Microbe Interactions*** le 21 octobre 2022.

1 Contexte

Aujourd'hui, la méthode de lutte la plus efficace contre la BLB est l'introduction de résistances génétiques dans les variétés cultivées. En phytopathologie, l'efficacité et la durabilité des résistances est conditionnée par l'utilisation et les schémas de déploiement dont elles font l'objet. Ainsi, pour potentialiser ces résistances, il est nécessaire d'avoir une connaissance approfondie des populations de *Xoo* présentes sur le terrain. Alors que les souches asiatiques de *Xoo*, décrites pour la première fois en 1884, sont largement étudiées, la découverte ultérieure de l'agent pathogène sur le continent Africain (1979) et l'importance plus réduite de la riziculture en Afrique expliquent probablement pourquoi les souches africaines sont moins bien caractérisées. Des résultats antérieurs ont montré que les *Xoo* africaines sont génétiquement plus proches de *X. oryzae* pv. *oryzicola* (*Xoc*). En outre, en termes de pouvoir pathogène, les *Xoo* africaines et asiatiques ne partagent aucun profil de virulence sur des lignées quasi-isogéniques (NILS, permettant la discrimination des races de *Xoo*). Notre équipe a constitué une large collection de souches de issues du continent Africain, collectées de 1979 à 2021, dans différents pays d'Afrique de l'Ouest et de l'Est. Afin de mieux comprendre l'histoire évolutive et le potentiel épidémique des *Xoo* Africaines, et pour promouvoir le déploiement de sources de résistance adaptées localement, nous avons cherché à caractériser la diversité et la structure génétique de ces populations, ainsi qu'à déterminer leurs relations phylogénétiques avec les autres populations de *Xoo* à l'échelle mondiale.

2 Stratégies

Nous avons choisi d'effectuer une analyse de la structure de 333 isolats de *Xoo* africaines d'origines variées présents en collection à l'IRD, par l'utilisation de marqueurs moléculaires MLVA développés par l'équipe en 2015. A la suite de cette première caractérisation, nous avons effectué le séquençage des génomes de 41 souches de ce panel en couplant des reads longs Oxford Nanopore Tehnologies (ONT) avec des reads Illumina. Les processus de séquençage, d'assemblage et de correction des génomes ont été réalisés à l'aide des derniers outils disponibles. Ce jeu de données a ensuite permis de réaliser des analyses phylogénétiques regroupant ces génomes, et la totalité des génomes complets de *Xoo* présents sur la base de données NCBI afin d'inférer des relations évolutives entre nos isolats africains et les *Xoo* issues d'autres parties du monde. Une analyse descriptive plus poussée du génome de la souche Nigériane NAI1 (ou CIX2779) a été réalisée.

3 Résultats

Principaux résultats de l'article 1 :

La résolution génotypique du schéma MLVA à 14 loci mis au point par Poulin et al., en 2015 est suffisante pour classer les *Xoo* Africaines en différents groupes et établir des relations entre nos populations.

Quarante et un haplotypes différents ont été attribués aux 333 souches de l'analyse, révélant une faible diversité des souches Africaines. Cette situation pourrait être la résultante d'une faible diversification, ou être imputable à la survenue de goulots d'étranglements par le passé. L'utilisation d'indices de liaison indiquent par ailleurs une forte clonalité en Afrique de l'Ouest.

L'analyse de la diversité de nos populations désigne le **Niger et le Mali comme potentiels centres** d'origine de **l'expansion de *Xoo* en Afrique de l'Ouest**.

La majorité des haplotypes isolés sur des espèces non-cultivées sont aussi retrouvés **sur des espèces cultivées et sauvages, voire même chez d'autres Poacées.** Quelques haplotypes sont exclusivement associés à des espèces sauvages de riz.

Enfin, la découverte la plus importante de cette étude est la **structuration par pays des *Xoo* africaines révélant un fort clivage entre les *Xoo* d'Afrique de l'Est et de l'Ouest**. Nous montrons que **les *Xoo* provenant de Tanzanie et de Madagascar pourraient correspondre à des membres de la lignée des *Xoo* Asiatiques**. Nos données, ainsi que des communications récentes de l'IRRI et de nos partenaires à Madagascar, suggèrent une dispersion longue distance récente médiée par l'introduction de matériel contaminé depuis l'Asie vers l'Afrique.

Principaux résultats des analyses phylogénétiques issues des données de séquençages ONT :

Nos analyses phylogénétiques se sont basées sur l'analyse de SNPs dans 411 génomes d'origines diverses, représentant les continents Africain, Asiatique, Australien et Américain. Parmi ces séquences, **41 génomes Africains ont été produits au cours de cette thèse**.

342 génomes complets ont été extraits de la base de données NCBI. Une analyse des métadonnées de ces génomes nous a permis de réaliser un état de l'art de la description génomique des populations de *Xoo* : **82 % des génomes présents sur NCBI appartiennent à des souches provenant d'Asie**.

Nos analyses phylogénétiques confirment la singularité des *Xoo* issues d'Afrique de l'Est révélée par **la précédente caractérisation MLVA**. Nous montrons que **les souches de *Xoo* en provenance de Tanzanie et de Madagascar** séquencées par ONT **groupent avec les *Xoo* Asiatiques**.

La lecture approfondie des phylogénies nous apprend que **les *Xoo* Tanzaniennes sont très proches de souches échantillonnées dans la province du Yunnan** en Chine et que **les isolats récoltés à Madagascar sont fortement apparentés à des *Xoo* provenant de Thaïlande.**

Principaux résultats de l'article 2 :

Nous avons séquencé et caractérisé le **génome de la souche Nigériane CIX2779 (ou NAI1).**

Nous montrons que cette souche possède un chromosome de **4 725 792 pb et 63.9% de GC avec une structure typique des *Xoo* Africaines**. CIX2779 contient un répertoire de **TALEs très proche de la souche Burkinabée BAI3**.

L'originalité du génome de la souche CIX2779 tient au fait qu'il abrite un **réplicon circulaire de ~40kb**. Ce plasmide, que nous avons appelé pCIX2779_1, est **présent en deux à trois copies dans la bactérie**. Les résultats d'un BLAST sur le plasmide nous montrent que les séquences les plus **proches** sont **des plasmides** issus de souches **de *X. euvesicatoria* pv. *alfalfae* et *X. vasicola* pv. *vasculorum* toutes deux originaires d'Afrique.**

Enfin, une analyse du contenu en gènes de pCIX2779_1 révèle la **présence d'une relaxase** de type MOB_F **et d'un système de sécrétion de type 4** (T4SS) apparenté au système VirB. pCIX2778_1 semble donc coder un **système de conjugaison fonctionnel**.

La majorité des autres séquences codantes sur pCIX2779_1 sont des **protéines " hypothétiques "**, ne nous permettant pas d'inférer un rôle biologique pour ce plasmide.

4 Importance pour le domaine

Nos analyses décrivent la **structure génétique et les voies de transmission potentielles de *Xoo* en Afrique**, et ont permis de révéler une **introduction récente et inédite des *Xoo* Asiatiques en Afrique de l'Est**, représentant une menace réelle pour les cultures à Madagascar et en Tanzanie.

Nos données soulignent l'**importance des réservoirs sauvages dans l'évolution et la dissémination de *Xoo***.

Les travaux de séquençage présentés ici constituent un effort continu de notre groupe visant à **améliorer l'état des connaissances et des ressources génomiques publiques concernant les *Xoo* Africaines**. Les données générées ouvriront la voie à la **conception de variétés adaptées** aux souches indigènes et émergentes de *Xoo* en Afrique.

pCIX2779_1 correspond au premier plasmide découvert chez les *Xoo* Africaines. Le génome de CIX2779 servira de base aux recherches sur la **prévalence, la diversité, le rôle biologique et l'évolution des plasmides chez les souches Africaines de *Xoo***.

5 Implication personnelle

J'ai effectué des vérifications en PCR diagnostique sur 401 entrées *Xoo* de la collection de souches de l'IRD et sélectionné 346 souches de *Xoo* sur lesquelles j'ai réalisé les différentes PCR multiplexes MLVA. J'ai effectué la totalité des analyses sur ces données. J'ai rédigé l'article 1 correspondant à ces données.

Sur la base des haplotypes découverts et de leurs métadonnées, j'ai sélectionné 41 souches Africaines pour le séquençage ONT. J'ai réalisé les extractions d'ADN génomique et les contrôles qualité nécessaires au séquençage ONT (notamment les électrophorèses en champ pulsé, PFGE), puis préparé les banques et effectué les séquençages.

La PFGE présentée dans l'article 2 est issue des contrôles qualité pré-séquençage que j'ai réalisé. J'ai mis au point un panel de phénotypage pour la caractérisation de NAI1 et effectué des inoculations par leaf-clipping et infiltrations sur riz sauvages puis enregistré les données de lésions. J'ai effectué les analyses statistiques et représentations graphiques des données quantitatives de phénotypage.

6 Article 1 : « Molecular epidemiology by MLVA genotyping of the Rice Pathogen *Xanthomonas oryzae* pv. *oryzae* reveals two different lineages on the African continent »

Molecular epidemiology of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* by Multi Locus Variable Number of Tandem Repeat Analysis reveals two different lineages on the African continent.

This study results of a collaborative project involving people from PHIM (France), ISRA (Senegal), INERA (Burkina Faso), LBMA (Mali), Fofifa (Madagascar),TARI (Tanzania), and IRRI (The Philippines).

Introduction

Rice is the staple food for more than half of the worlds' population. Increases in rice demand is predicted to reach about 14% between 2014 and 2030 [1]. In 2018, 759.6 million tons of paddy rice were produced. Major producers in Asia include China as the leader with 208 Mt of paddy rice, followed by India (170 Mt), Indonesia (75 Mt), Bangladesh (53 Mt) and Vietnam (44 Mt). Africa is the second rice producing continent, with overall production of 33.3 Mt whereby Nigeria and Madagascar are the two main producers [2]. With the present and future changes in climate conditions, rice will face various biotic and abiotic changes. Current temperatures are already close to critical levels during sensitive stages of rice plant development. High temperatures or warmer night-time temperatures can lead to a decrease in spikelet fertility [3] and it was shown in a simulation by Prabhjyot-Kaur and Hundal in 2007 [4] that a 2°C increase in temperature results in a 3 % decrease in rice yield. Rice production will face increased frequency of droughts, floods and salinity stress due to sea-level rise.

In addition to these abiotic factors, rice production already faces multiple stresses of biotic origin. They include numerous diseases, caused by fungi, bacteria, viruses or nematodes, which do not only result in serious yield losses but also affect the overall quality of the rice grains. Among the major diseases are rice blast (*Magnaporthe oryzae*), sheath blight (*Rhizoctonia solani*), brown spot (*Helminthosporium oryzae*), false smut (*Ustilaginoidea virens*), tungro virus, and bacterial leaf blight (*Xanthomonas oryzae* pv. *oryzae*) [5].

Caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) and first reported in Japan in 1884, bacterial leaf blight (BLB) of rice is a disease of major importance in tropical and sub-topical Asia and parts of Africa. Although reported in Australia, Latin America and the Caribbean, no important outbreaks have been recorded in these areas [6]. Symptoms of BLB become visible at the tillering stage. The lesions on the leaves start as watersoaked, dark green stripes. Then the bacteria move forward and the lesions expand along the veins, forming dry, yellowish stripes along the entire leaf, ultimately the entire plant can die [6]. In severe epidemics, it has been reported that the yield losses due to bacterial leaf blight range from 20 to 40% [7].*Xoo* is a rod shaped, Gram negative bacteria of 0.8-2.9 µm, motile, with a single polar flagellum of 6-8 µm [8], [9]. Bacteria enter the plant through pores at the leaf tip called hydathodes, or directly through wounds or growth cracks and colonize the epitheme. The bacteria then enter the xylem vessels and spread throughout the plant. In a couple of days, the bacterial cells and exopolysaccharides crowd the xylem vessels, resulting in the symptoms of leaf blight [8]. Sources of primary inocula commonly acknowledged are paddy water and debris from infected plants left on the ground between cultures. Although very likely, seed transmission of BLB is not considered to be a major way of dissemination, as it has been demonstrated for bacterial leaf streak of rice (BLS) caused by *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) [10]. Secondary inoculum comes from bacteria oozing from infected leaves, forming droplets easily transmitted between plants by contact, rain, cuttings or insects [6].

Similar to other bacterial diseases, bacterial leaf blight is managed traditionally through preventive cultural practices such as weeding or cleaning plant debris, the use of chemicals such as copper or Bismerthiazol, and antimicrobials like streptomycin [11], [12], [13]. However, these methods are not effective enough to control the disease and in the latter two cases, the development of resistance in *Xanthomonas* as well as the environmental and health prejudice of these approaches make them unsustainable. Biological control by PGPR bacteria (such as *Pseudomonas fluorescens* or some strains of *Bacillus)* and most importantly, host genetic resistance, are increasingly adopted as efficient management methods against BLB [13].

As phytopathogenic bacteria reproduce fast, exchange genetic material through Horizontal Gene Tranfer (HGT) and give rise to large communities, their evolutionary potential is far more important compared to their hosts [14]. It is therefore crucial to carefully monitor the emergence of novel pathogen lineages in order to manage host resistance efficiently. Population genetics measure the genetic variation within a species and or a pathovar, and is based on the evaluation of the dispersal in space and time of allele frequencies thanks to molecular typing tools and sequencing. This allow to uncover the pathogens' adaptation processes, including genetic sources of selection, dispersion patterns, repartition of virulence traits and evolution pace [14]–[16].

The earlier work of Gonzales *et al.* in 2007 [17] on a first but relatively limited set of African *X*. *oryzae* strains showed with several molecular typing approaches that African *Xoo* strains form a distinct genetic lineage within the species and tend to be more closely related to *Xoc* than to Asian *Xoo*. Subsequent studies based on DNA sequences of several housekeeping loci [18], [19] or on Average Nucleotide Identity (ANI) measurements [20] between African and Asian *Xoo* genomes confirmed these results. Finally, phylogenies constructed on the basis of SNPs selected from the complete genomes of 63 *Xoo* strains have clearly demonstrated that African and Asian *Xoo* form two distinct lineages [21].

Multi Locus Variable Number of Tandem Repeat (VNTR) Analysis (MLVA) is a genetic typing method based on the evaluation of the polymorphism in the number of repeated sequences in several VNTR loci in bacterial genomes. This technique involves multiplex PCR amplification of the selected VNTR loci. Generally speaking, the higher the number of loci, the better is the resolution of the analysis. This method is widely used to study the epidemiology of genetically monomorphic organisms. Poulin and colleagues developed in 2015 [22] an MLVA scheme able to efficiently discriminate different *Xanthomonas oryzae* populations, thereby providing a first assessment of *Xoo* and *Xoc* worldwide genetic diversity and structure. These authors notably analyzed 59 African *Xoo* strains from four West African countries and observed that strains from Niger and Burkina Faso were genetically more related while strains from Cameroon and Mali were more distinct.

Still, the knowledge concerning the genetic structure of *Xoo* populations on the African continent is incomplete. In order to better understand their evolutionary history and epidemic potential, and to promote the deployment of locally adapted sources of resistance, we sought to characterize the genetic relationships shaping African *Xoo* population using a larger collection of strains with a broader geographic origin on the continent. Here we present a first portray at the continental scale of *Xoo* strains genetic diversity contrasted as a function of geographic and temporal dimensions but also as a function of the host species of origin.

Results

Description of the set of examined African *Xoo* strains.

In total, 333 different *Xoo* strains from the IRD collection of *Xanthomonas* were selected for MLVA characterization based on minimal metadata requirements such as the country of origin and the year of isolation (Figure 1, Table S1). These isolates, originating from 9 different countries, were mainly collected in West Africa (Benin, Burkina Faso, Ivory Coast, Senegal, Niger and Mali), owed to the occurrence of BLB for a longer time in the region. Isolates were collected over 15 years, from both wild and cultivated rice but also weeds bordering the fields (Figure S1). To avoid duplicates with the pioneer study of Poulin et al [22] but also in order to compare with it, our set of strains includes a few isolates that were collected before 2012 while the majority was sampled from 2012 to 2018. East-African strains were collected from 2019 outbreaks in Madagascar and Tanzania. Because field survey activities were unequally intensive, there exist strong dissimilarities in the extent of sampling depth across countries (Figure 1, Figure S1). In particular, our collection is over-represented by strains from Mali which reflects both the high incidence of BLB in this country, and the almost continuous pathogen monitoring in this country owed to the historical partnership with national institutions that was initiated some 20 years ago. This is well highlighted in Figure S1, which shows the distribution of *Xoo* strains by year of sampling, country and sub-region. The Asian strain PXO99^{A} (CIX236) was also included in the MLVA typing as an outgroup control.

The genotypic resolution of the MLVA scheme is sufficient to uncover genetic relationships of *Xoo* in Africa and reveals low diversity of African strains.

In order to investigate the epidemiological patterns of *Xoo* in Africa, we performed molecular typing of our collection by MLVA, despite of the gaps in the representativity of strains in some countries and or region. The original MLVA scheme used in this study was developed by Poulin *et al.* in 2015 [22] and exploits 16 different VNTR loci. However, upon preliminary experiments, amplicon data for the Xo_G58 and Xo_G88 markers could not be reliably interpreted and these markers were excluded from subsequent analysis.

The genotypic resolution of this scheme on our bacterial collection was assessed by plotting the genotype accumulation curve over the 14 loci [23]. This curve describes the genotypic diversity in relation to different combinations of tandem repeats loci and is generated by randomly sampling various numbers of loci and counting the number of multilocus genotypes (MLG) differentiated. The genotypic resolution is considered good when the curve reaches a plateau. Here, we observe that 95% of our strains diversity can be discriminated with 12 loci (Figure 2), indicating that the MLVA-14 scheme is appropriate to accurately resolve the different haplotypes of our African *Xoo* collection.

Over the 333 African *Xoo* strains typed, 41 different allelic profiles or haplotypes were determined. Among the 14 loci examined to assess the population genetic structure, the most polymorphic is Xo_G55 with 14 different alleles (Table S3). The other loci scored continuously from 2 to 9 alleles. Six Grouping Complexes (GC) of haplotypes differing at three or fewer loci were identified. Each GC contains a Clonal Complex (CC) which is a group of haplotypes that differ at only one locus. The largest CC, containing 14 haplotypes, comprises 251 isolates, almost all from Mali, and represents 75.4% of the total strains studied. 6.9% of the samples studied formed singletons, i.e. isolated haplotypes with more than 4 loci of difference from any other haplotype.

Discriminant Analysis by Principal Component shows a structuration by countries of African *Xoo*

An unbiased clustering of our haplotypes via Discriminant Analysis of Principal Component (DAPC) was used because it is free of any assumption linked to a population genetic model. The optimal number of clusters to retain was determined using Kullback Information Criterion (KIC) calculated thanks to a combination of geometric approach and fast likelihood optimization developed by Beugin *et al.* in 2018 [24] (Figure S2). As shown in Table 1, DAPC clusters are generally composed of haplotypes originating from a single or a couple of countries. This suggests that African *Xoo* strains exhibit a certain degree of geographic structuring. The most interesting observation here is the singularity of the East African *Xoo*. Indeed, these strains that were sampled in 2019 in Tanzania and in Madagascar, are clearly clustering apart from West african strains (Table 1, Figure 3 and 4). Moreover they share some allelic profiles with Asian strains, and DAPC analysis groups them with our control strain PXO99^{A} from the Philippines in the DAPC6 cluster. The locus that discriminates the most powerfully East African from West African *Xoo* isolates is the locus G07 that does not amplify for East-African nor *Xoo* of Asian origin [22]. All in one, our analysis show West African and East African isolates form distinct sub-lineages and that the latter are closely related to one strain from the Philippine.

Study of the relationships between haplotypes sheds light on the distribution of *Xoo* strains in West Africa.

To further examine the genetic structure of the pathogen in Africa, we divided up manually the collection areas in 11 zones, from A to K, without taking into account the borders between states and potential trade restrictions (Figure 3A). When representing the relationships between haplotypes regarding these zones, we generally observe that zones that are geographically close cluster together and share haplotypes (Figure 3B). This is the case of haplotype 7 in DAPC2, or haplotypes 29 and 36 in DAPC1, that are shared by isolates sampled in neighboring zones C and D. This pattern could match a model of short distance dispersion of the pathogen that could be mediated by regional meteorological events. At the same time, it appears that some geographically distant zones share haplotypes or exhibit haplotypes with close links in the minimum spanning tree. As an example, the haplotype 25 of DAPC3 is found in isolates sampled in zone H but also in zone F. These cases are consistent with longer distance dispersal of the strains that could be mediated by exchanges of seed material.

When considering factors shaping the diversity of our strains, the year of sampling is undoubtedly determinant with regards to the grouping of haplotypes. A striking example is visible with strains from Niger. While Nigerien strains collected in 2004 are grouping in DAPC5, those collected in 2018 cluster with Burkinabe strains in DAPC3 (Figure 4).

Diversity and linkage disequilibrium analyses show high clonality in West Africa and point to Niger as a potential center of origin for *Xoo* expansion.

In order to investigate for the potential center(s) of diversification of *Xoo* in West Africa, we conducted diversity analyses based on the representation of the different haplotypes and alleles. Populations of interest were defined on a per-country basis. Because some populations were under-represented (N< 8), they were excluded from our genetic diversity analyses. Consequently, the focus has been on four West African countries only including Burkina Faso, Mali, Niger and Senegal. As shown in Table 2, Niger followed by Mali are the populations with the highest genotypic richness (eMLG of 6.642 and 4.818, respectively) using a rarefaction method to correct for the uneven sample size between countries, as well as one of the highest genotypic (Simpson's index of 0.898 and 0.813, respectively) and gene diversities (Nei's gene diversity index of 0.365 and 0.184, respectively). This suggests that these two

countries experienced earlier emergence of the pathogen. Of the West African countries studied, and despite a relatively low number of individuals (N=19), Niger possesses the highest allelic diversity as well. Altogether, these results suggest that Niger could correspond to a potential origin of diversification for *Xoo* in West Africa.

In contrast, Senegal displays low genotypic diversity with an eMLG of 1.851. Their Eveness (E.5) is high, 0.706 indicating that the two genotypes frequencies are quite similar. As a result, the probability of randomly choosing two different genotypes in Senegalese population is the lowest of the four countries studied (Simpson's index of 0.305). These two genotypes are very different from each other because Nei's gene diversity index (H) at all loci is rather high (with a value of 0.225) compared to other countries. This is consistent with the number of alleles observed for each locus, being on average of 2. This means that these genotypes differ from each other at almost every locus. However, these alleles are not specific to Senegal because the private allelic richness per locus is quite low, which corroborates the Eveness data. All ine one, these observations are consistent with 3 scenarios : either i) the pathogen established too recently to have time to detectably diversify, or ii) a strong selection pressure formed a bottleneck, selecting only two genotypes, or finally, iii) these values could correspond to a strong sampling bias in Senegal where only two areas were investigated and many isolates obtained from the same leaf.

Regarding linkage disequilibrium analyses, Ia and r-D index (Table 2, [25], [26]) are significantly different from 0, indicating of a strong Linkage Disequilibrium among loci. Hence, unsurprisingly, it appears that the evolution of West African populations of *Xoo* occurred mostly in a clonal manner.

Evidence supporting migration events in West Africa remains scarce.

The analysis of genetic differentiation of *Xoo* strains in West Africa was performed based on the calculation of FST and RST indices [27], [28]. Our FST values suggest high genetic differentiation between most of the West African countries studied (FST > 0.15, Table 3). As it estimates mostly migration, FST do not perform well for highly divergent populations when using microsatellites (as they have a high mutation rate). In these situations, RST is more accurate. RST is an indice representing population subdivision as inferred from allele frequencies at microsatellite loci. RST performance will depend on how well the microsatellites fit a Single Mutation Model (SMM, [29]). Indeed, RST is independent of mutation only under a strict SMM. This model takes into account the sizes of the allele (alleles of very different sizes will be more distantly related than alleles of similar sizes) [29]. The analysis of the number of repetitions obtained for the different West African haplotypes at each locus confirms that our microsatellites obey the SMM model (Table S3) and thus, legitimates the use of RST indices.

In our case RST values are very high, but as the associated p-values are all under 0.04, these indices give reliable biological information. It appears that Burkinabe strains differ highly (RST > 0.5) from Malian and Senegalese strains, and strongly (0.15 < RST < 0.25) from Nigerien strains. There is also a high differentiation between Malian and Nigerien strains. Similarly, we observe a high differentiation between Nigerien and Senegalese strains. However, there is a really weak difference between Mali and Senegal populations (RST < 0.05). Consistent with our DAPC analysis, the results of FST and RST calculation indicate that apart from this latter case, migration events and/or material exchanges between these countries are not significant.

Studying population structure at a country scale : the case of Mali

Most of our isolates consisted of Malian *Xoo*. As shown in Figure 5A, a depiction of the different haplotypes with up to five Locus-Variants (LV) allows us to trace some of the temporal pathogen diversity observed in this country. We can see two different Grouping Comlexes (GC). The larget GC consists of a Clonal Complex that corresponds to cluster DAPC1 (Table 1, Figure 4A) plus the haplotype 32, which is separated from this CC by 4 LV. This GC is relatively recent and appears to have arisen from the founder haplotype 41, which is connected to five single-locus-variants (SLV) and could be detected in Mali over seven years from 2010 to 2017 (Figure 5B). The emergence of this CC can be attributed to a different event of introduction than in the case of haplotype 1, which is separated from the CC by 5 LV.

The second GC is composed of haplotypes 4, 5, 6, 7 that differ from each other by 1 to 3 LV. This GC groups strains collected from 1979 to 2009. As these strains were collected a long time ago and since no samples are available for the period between 2004 and 2012, it is difficult to connect this GC to the largest one. Yet our analysis indicates that there were 2 events of introduction in Mali in 1979: one restricted to haplotype 8, a singleton separated by 5 LV from haplotype 7 which was collected from 1979 to 2004), and the others resulting in the second small GC.

When analyzing the network of haplotypes taking in account the nature of the host of origin (Figure 5C), our data strikingly show that haplotypes 27 and 28 are specific to the wild rice species *O. longistaminata*, while some haplotypes are found both on cultivated and wild species. It is the case for haplotypes 37, 41 and 15, which were isolated from leaves of both *O. sativa* and *O. barthii*. Finally, some genotypes are even ubiquitous, and can be found on weeds as well as on wild and cultivated rice species. This is the case for haplotype 29, which is hosted by *O. sativa*, *O. longistaminata*, and weeds such as *Leersia* (a *Poaceae* of the *Oryzeae* tribe) and *Cyperaceae* (a *Poale*).

Discussion

The overall African *Xoo* genetic diversity is relatively limited

Over the 333 strains typed, only 41 different allelic profiles (or haplotypes) were determined. Similar studies on *Xoo* populations in Asia made by Poulin et. al in 2015 [22] retrieved 71 different haplotypes for 127 Asian isolates. The allelic richness of African *Xoo* populations is lower than that of Asian populations, which could reflect a later establishment of the pathogen on the continent, leaving little time for its diversification. Alternatively, a past major population bottleneck may also explain this situation.

Unbalanced sampling across countries

Overall, the minimum spanning trees (Figures 3 and 4) drawn to detect epidemiological information, illustrates the fact that this signal is only discernible in areas that have the highest numbers of isolates. One fundamental point to consider to critically examine the generality of our observations regarding West African *Xoo* population genetics lies in the relative unbalanced counts of strains, mainly, with regards to factors such as the country of origin and the year of sampling. This is partly due to the low incidence of the disease in some countries (as in Ivory Coast or Benin), the absence of local collaborators and a lack of access to the field, or in some cases, unsuitable safety conditions for surveys. As a result, except in the case of Mali, Niger and Burkina Faso, strains from most countries are

representative of only a single year, a unique region and/or a single host type. Therefore, the number of strains that co-existed in different countries is low and diversity analysis cannot be performed with low represented countries. The samples in low represented countries should serve as spatiotemporal tags to understand global movements of the pathogen through the continent, and infer introductions in better sampled countries.

To what extent the detected genetic diversity informs our understanding of *Xoo* strains spatial distribution on the continent?

A previous African *Xoo* DAPC clustering approach [22] indicated that isolates from Burkina Faso and Niger were similar while those from both Cameroon and Mali formed more isolated populations. Here, on the basis of an extended dataset encompassing 9 countries, our DAPC analysis distinguished six groups (Table 1) and tends to deliver similar conclusions on a relative differentiation of populations in geographic units even though clear spatial boundaries are difficult to draw. In contrast, when analyzing population structure for those countries with sufficient isolates counts (Burkina Faso, Mali, Niger, Senegal), summary statistics indicate significant strong differentiation (RST > 0.15) for all pairwise comparisons but populations from neighboring Mali and Senegal. In Mali, Poulin et al [22] discerned two DAPC clusters that appeared to split strains isolated before 2010 and those isolated later. Here, the minimum spanning tree built with 258 Malian strains isolated from 1979 up to 2018 support the same conclusion with earlier strains belonging to a small GC while more contemporary isolates belong to a bigger one (Figure 5B).

Because the internal Niger river delta in Mali is believed to be the center of origin of the ancestral African *O. glaberrima* rice [30] and because Malian *Xoo* isolates were reported to be comparatively diverse [22], one could intuitively assume that this country has also been a center of diversification for African *Xoo*. While diversity indexes are consistent with previous results, one interesting outcome of our analysis is the realization that Mali ranks only second after Niger regarding diversity indices, suggesting that the center of *Xoo* diversification on the continent may also extend to this country.

Finally, it is also important to bear in mind that despite the above trends in population structure, we also observed that in several cases where strains were isolated on the same year, they tend to be more genetically related or even share the same genotype (Figure 4B, 5B). In addition, we observed geographically distant zones with isolates which have related or identical genotypes (for example haplotype 15 in DAPC5 or haplotype 25 in DAPC3; Figure 3B). These cases could indicate longer distance dispersal of the strains due to plant material exchanges.

No evidence for African *Xoo* strain specialization on alternate hosts

Because special efforts were deployed on the isolation of 30 *Xoo* strains from potential alternative host species, most notably in Mali, our study provides an unprecedented insight on whether host species tend to drive differentiation of distinct pathogen genotypes. Of the eight haplotypes that belong to strains isolated from hosts other than *O. sativa*, only two of them were not also isolated from cultivated *O. sativa* varieties (Figure 5C) and appeared to be uniquely found on *O. longistaminata*. This observation seems to contradict a possible host-specialization and further suggests that wild rice and other wild species (mostly weeds) may act as *Xoo* reservoirs that could contribute to disease dynamics across cropping seasons and should be more systematically considered in BLB epidemiological studies.

East African strains have a recent Asian origin?

Here, and as previously reported by Poulin *et al.* [22], African *Xoo* strains were found to be monomorphic for four loci of the *Xo* MLVA-16 scheme, each with a lineage specific allele (Xo_G09, 3 TRs; Xo_G15, 3 TRs; Xo_G80, 3 TRs; and Xo_83, 2 TRs). Madagascar *Xoo* strains examined here have an allele which is different from the African *Xoo* diagnostic one at all four loci whereas Tanzanian strains have different alleles at three of these loci (Table S3). Together with our evolutionary model-agnostic DAPC analysis where these East African strains clustered with the Filipino strain PXO99^{A}, this is strong evidence that they are genetically unrelated to West African *Xoo* strains and may in fact resemble Asian *Xoo*. While this hypothesis will need to be ascertained based on whole genome sequencing and pathotyping on IRBB lines for instance (see chapter 2), it nonetheless raises questions regarding the transmission mode of these strains to Africa and the potential threat they represent to rice cultivation on the continent and in East Africa specifically. Our discovery of a second lineage of *Xoo* strains, possibly related to Asian relatives, on the African continent coincide with relatively recent reports of BLB presence in Madagascar and Tanzania (Mathilde Hutin and Ricardo Oliva, personal communication). This would be consistent with long distance dispersal of Asian haplotypes by independent introduction events of infected plant material in East Africa.

Seed transmission is believed to be a major pathway for *Xoc* strains dissemination [31], [32], and bidirectional occurrences of transcontinental strain dissemination have been suggested. For example, several Chinese *Xoc* strains were found to be more related to African *Xoc* than Asian individuals [22]. Likewise, contemporary West African *Xoc* populations may have emerged as a consequence of a westward dissemination gradient following introduction of fonders Asian strains in Madagascar and East African coastal countries [12]. Contrary to *Xoc*, seed transmission of *Xoo* is not currently considered significant in the epidemiology of BLB [6]. The literature does not generally address long distance movement of *Xoo* strains possibly because of a lack of robust data supporting such events. Nonetheless, our results on the possible long distance dispersal of Asian haplotypes by introgression of material in Africa and disparate reports on the detection of *Xoo* on naturally infested seed, for example on Chinese samples [33], advocates for a reassessment of established principles regarding BLB transmission. Yet, how can new BLB epidemics be initiated from a seed-associated primary inoculum, is a key question that needs to be investigated.

Ultimately, a refined understanding of BLB etiology and *Xoo* populations dynamics in Africa would require a more balanced and nested sampling of infected samples across geographic locations and temporal frames. This will need to be confronted with data on the resistance genes make up of cultivated or natural host plants and on local cultivation practices in order to untangle the effects of selective sweeps and bottleneck events on *Xoo* populations diversity (see [14] for a general review). Our work has however classified a large collection of isolates into 41 haplotypes and this information will serve as a basis for subsequent in-depth analysis of this diversity through whole genome sequencing. Furthermore, we report on a novel East African *Xoo* lineage that deserves further scrutiny regarding its origin and virulence potential as it may poses an imminent threat to rice cultivation on the continent.

Material and methods

Bacterial strains

The list of the bacterial strains used in this study and their metadata along with their corresponding haplotypes and DAPC clusters are presented in Table S1. Each glycerol stock was streaked on PSA medium (10g/L peptone, 10g/L sucrose, 1g/L glutamic acid, 16g/L agar) in order to obtain single colonies. Plates were incubated for 3 days at 28°C. Single colonies were resuspended in sterile water and then boiled to lyse the bacteria.

MLVA genotyping

The strain collection was genotyped using the MLVA-16 scheme developed for *Xanthomonas oryzae* by Poulin *et al.* in 2015 [22]. Two loci (G58 and G88) were removed from the analyses because of amplification issues (Table S2). The multiplex PCR reactions were conducted directly on the bacterial lysates. Four reactions of multiplex PCR (corresponding to four different mixes) were performed with fluorescent primer pairs, using the QIAGEN® Multiplex PCR kit (Qiagen, Courtaboeuf, France). The conditions for amplifications for mixes 1 and 2 were 35 cycles of 94°C for 30 s, annealing temperature of 60°C for 1 min 30 s and 72°C for 1 min. For mixes 3 and 4, the conditions were 25 cycles with an annealing temperature of 64°C. PCR products were diluted in Hi-Di formamide mixed with GeneScan 500 or 600 LIZ dye Size Standard (Thermofisher Scientific). Capillary electrophoreses were performed in an ABI 3500 XL Genetic Analyzer (Applied Biosystems) at the GenSeq platform in Montpellier, France.

Amplicons sizing and allele assignment

The fluorescence peaks obtained for each locus and strain were sized using GeneMapper version 4.0 (Applied Biosystems, Life Technologies, Carlsbad, CA). The peaks corresponding to fragments sizes below the combined sizes of the flanking regions of specified VNTR were deleted. With the same software, alleles were assigned to each fragment size. When needed, the size was rounded up to the nearest integer and corresponding allele.

Genetic analyses

The sizes of the flanking regions and the patterns of the different VNTR for each locus were determined by Poulin *et al.* in 2015 [22]. The number of repeats for each allele was determined by subtracting the sizes of the corresponding flanking regions, and dividing the resulting fragment by the length of the considered VNTR pattern.

The number of Multilocus Genotypes (MLG) and associated genotypic diversities were calculated using the R package poppr v.2.8.6, which was specifically developed to facilitate analyses of clonal populations [23], [34]. The DAPC clusters and associated analysis (loading plots, scatterplots) were created using the R package adgenet [35], [36]. Inference of phylogenetic relationships between haplotypes were assessed on the software PHYLOViZ 2.0 [37] and represented by minimum spanning trees using an algorithm combining global optimal eBURST (goeBURST) and Euclidian distances. Allelic richness and private allelic richness were calculated using the software HP rare, with the rarefaction method proposed by Kalinowsky in 2005 [38]. Genetic differentiation between populations were assessed using both pairwise FST and RST statistics calculated with Arlequin software v.3.5.2.2 [39]. The Arlequin and Genepop formatted files for analysis on Arlequin and HP rare respectively were generated using the plugin Genalex version 6.5 on Microsoft Excel [40].

Bibliography

[1] S. Yuan *et al.*, « Sustainable intensification for a larger global rice bowl », *Nat Commun*, vol. 12, n° 1, Art. n° 1, déc. 2021, doi: 10.1038/s41467-021-27424-z.

[2] FAO, *FAO Rice Market Monitor (RMM), April 2018, VOLUME XXI ISSUE No. 1*. Rome, Italy: FAO, 2018. Consulté le: 18 octobre 2022. [En ligne]. Disponible sur: https://www.fao.org/publications/card/fr/c/I9243EN/

[3] R. Prasad, Y. S. Shivay, et D. Kumar, « Current Status, Challenges, and Opportunities in Rice Production », in *Rice Production Worldwide*, B. S. Chauhan, K. Jabran, et G. Mahajan, Éd. Cham: Springer International Publishing, 2017, p. 1-32. doi: 10.1007/978-3-319-47516-5_1.

[4] S. S. Hundal et Prabhjyot-Kaur, « Climatic variability and its impact on cereal productivity in Indian Punjab », *Current Science*, vol. 92, n° 4, p. 506-512, 2007.

[5] P. Singh *et al.*, « Biotic Stress Management in Rice (Oryza sativa L.) Through Conventional and Molecular Approaches », in *New Frontiers in Stress Management for Durable Agriculture*, A. Rakshit, H. B. Singh, A. K. Singh, U. S. Singh, et L. Fraceto, Éd. Singapore: Springer, 2020, p. 609-644. doi: 10.1007/978-981-15-1322-0_30.

[6] D. O. Niño-Liu, P. C. Ronald, et A. J. Bogdanove, « *Xanthomonas oryzae* pathovars: model pathogens of a model crop », *Molecular Plant Pathology*, vol. 7, n° 5, p. 303-324, 2006, doi: 10.1111/j.1364-3703.2006.00344.x.

[7] R. V. Sonti, « BACTERIAL LEAF BLIGHT OF RICE : NEW INSIGHTS FROM MOLECULAR GENETICS », *Current Science*, vol. 74, n° 3, févr. 1998.

[8] S. H. Ou, *Rice Diseases*. IRRI, 1985.

[9] S. Ishiyama, « Studies of bacterial leaf blight of rice. », vol. 45, 1922, p. 233-261.

[10] SRIVASTAVA et SHEKHAWAT, « Mode of Seed Infection and Transmission of Bacterial Leaf Streak of Rice », 1972.

[11] C. E. Morris, M. Barny, O. Berge, L. L. Kinkel, et C. Lacroix, « Frontiers for research on the ecology of plant-pathogenic bacteria: fundamentals for sustainability », *Mol Plant Pathol*, vol. 18, n° 2, p. 308-319, janv. 2017, doi: 10.1111/mpp.12508.

[12] M. Jeger *et al.*, « Pest categorisation of *Xanthomonas oryzae* pathovars *oryzae* and oryzicola », *EFSA Journal*, vol. 16, n° 1, p. e05109, 2018, doi: 10.2903/j.efsa.2018.5109.

[13] S.-Q. An *et al.*, « Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas* », *FEMS Microbiol Rev*, vol. 44, n° 1, p. 1-32, oct. 2019, doi: 10.1093/femsre/fuz024.

[14] S. de Vries, E. H. Stukenbrock, et L. E. Rose, « Rapid evolution in plant–microbe interactions – an evolutionary genomics perspective », *New Phytologist*, vol. 226, n° 5, p. 1256-1262, 2020, doi: 10.1111/nph.16458.

[15] N. J. Grünwald, S. E. Everhart, B. J. Knaus, et Z. N. Kamvar, « Best Practices for Population Genetic Analyses », *Phytopathology®*, vol. 107, n° 9, p. 1000-1010, sept. 2017, doi: 10.1094/PHYTO-12-16-0425-RVW.

[16] S. Pfeilmeier, D. L. Caly, et J. G. Malone, « Bacterial pathogenesis of plants: future challenges from a microbial perspective », *Mol Plant Pathol*, vol. 17, n° 8, p. 1298-1313, août 2016, doi: 10.1111/mpp.12427.

[17] C. Gonzalez, B. Szurek, C. Manceau, T. Mathieu, Y. Séré, et V. Verdier, « Molecular and Pathotypic Characterization of New *Xanthomonas oryzae* Strains from West Africa », *MPMI*, vol. 20, n° 5, p. 534-546, mai 2007, doi: 10.1094/MPMI-20-5-0534.

[18] L. R. Triplett *et al.*, « Genomic Analysis of *Xanthomonas oryzae* Isolates from Rice Grown in the United States Reveals Substantial Divergence from Known X. *oryzae* Pathovars », *Applied and Environmental Microbiology*, vol. 77, n° 12, p. 3930-3937, juin 2011, doi: 10.1128/AEM.00028-11.
[19] A. Hajri *et al.*, « Multilocus sequence analysis and type III effector repertoire mining provide new insights into the evolutionary history and virulence of *Xanthomonas oryzae* », *Molecular Plant Pathology*, vol. 13, n° 3, p. 288-302, 2012, doi: 10.1111/j.1364-3703.2011.00745.x.
[20] J. M. Lang *et al.*, « A Pathovar of *Xanthomonas oryzae* Infecting Wild Grasses Provides Insight Into the Evolution of Pathogenicity in Rice Agroecosystems », *Front Plant Sci*, vol. 10, p. 507, 2019, doi: 10.3389/fpls.2019.00507.
[21] R. Oliva *et al.*, « Broad-spectrum resistance to bacterial blight in rice using genome editing », *Nat Biotechnol*, vol. 37, n° 11, p. 1344-1350, 2019, doi: 10.1038/s41587-019-0267-z.
[22] L. Poulin *et al.*, « New Multilocus Variable-Number Tandem-Repeat Analysis Tool for Surveillance and Local Epidemiology of Bacterial Leaf Blight and Bacterial Leaf Streak of Rice Caused by *Xanthomonas oryzae* », *Appl Environ Microbiol*, vol. 81, n° 2, p. 688-698, janv. 2015, doi: 10.1128/AEM.02768-14.
[23] Z. N. Kamvar, J. C. Brooks, et N. J. Grünwald, « Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality », *Frontiers in Genetics*, vol. 6, 2015, Consulté le: 13 octobre 2022. [En ligne]. Disponible sur: https://www.frontiersin.org/articles/10.3389/fgene.2015.00208
[24] M.-P. Beugin, T. Gayet, D. Pontier, S. Devillard, et T. Jombart, « A fast likelihood solution to the genetic clustering problem », *Methods in Ecology and Evolution*, vol. 9, n° 4, p. 1006-1016, 2018, doi: 10.1111/2041-210X.12968.
[25] A. H. D. Brown, M. W. Feldman, et E. Nevo, « MULTILOCUS STRUCTURE OF NATURAL POPULATIONS OF HORDEUM SPONTANEUM », *Genetics*, vol. 96, n° 2, p. 523-536, oct. 1980, doi: 10.1093/genetics/96.2.523.
[26] P.-M. Agapow et A. Burt, « Indices of multilocus linkage disequilibrium », *Molecular Ecology Notes*, vol. 1, n° 1-2, p. 101-102, 2001, doi: 10.1046/j.1471-8278.2000.00014.x.
[27] S. Wright, « The genetical structure of populations », *Ann Eugen*, vol. 15, n° 4, p. 323-354, mars 1951, doi: 10.1111/j.1469-1809.1949.tb02451.x.
[28] M. Slatkin, « A Measure of Population Subdivision Based on Microsatellite Allele Frequencies », *Genetics*, vol. 139, n° 1, p. 457-462, janv. 1995.
[29] F. Balloux et N. Lugon-Moulin, « The estimation of population differentiation with microsatellite markers », *Molecular Ecology*, vol. 11, n° 2, p. 155-165, 2002, doi: 10.1046/j.0962-1083.2001.01436.x.
[30] P. Cubry *et al.*, « The Rise and Fall of African Rice Cultivation Revealed by Analysis of 246 New Genomes », *Current Biology*, vol. 28, n° 14, p. 2274-2282.e6, juill. 2018, doi: 10.1016/j.cub.2018.05.066.
[31] T. W. Mew, « Focus on Bacterial Blight of Rice », *Plant Dis.*, vol. 77, n° 1, p. 5, 1993, doi: 10.1094/PD-77-0005.
[32] G. L. Xie et T. W. Mew, « A Leaf Inoculation Method for Detection of *Xanthomonas oryzae* pv. oryzicola from Rice Seed », *Plant Dis*, vol. 82, n° 9, p. 1007-1011, sept. 1998, doi: 10.1094/PDIS.1998.82.9.1007.
[33] Y. Tian, Y. Zhao, R. Xu, F. Liu, B. Hu, et R. R. Walcott, « Simultaneous Detection of *Xanthomonas oryzae* pv. *oryzae* and X. *oryzae* pv. oryzicola in Rice Seed Using a Padlock Probe-Based

Assay », *Phytopathology®*, vol. 104, n° 10, p. 1130-1137, oct. 2014, doi: 10.1094/PHYTO-10-13-0274-R.
[34] Z. N. Kamvar, J. F. Tabima, et N. J. Grünwald, « Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction », *PeerJ*, vol. 2, p. e281, 2014, doi: 10.7717/peerj.281.
[35] T. Jombart, « adegenet: a R package for the multivariate analysis of genetic markers », *Bioinformatics*, vol. 24, n° 11, p. 1403-1405, juin 2008, doi: 10.1093/bioinformatics/btn129.
[36] T. Jombart et I. Ahmed, « adegenet 1.3-1: new tools for the analysis of genome-wide SNP data », *Bioinformatics*, vol. 27, n° 21, p. 3070-3071, nov. 2011, doi: 10.1093/bioinformatics/btr521.
[37] A. P. Francisco, C. Vaz, P. T. Monteiro, J. Melo-Cristino, M. Ramirez, et J. A. Carriço, « PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods », *BMC Bioinformatics*, vol. 13, n° 1, p. 87, mai 2012, doi: 10.1186/1471-2105-13-87.
[38] S. T. Kalinowski, « hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness », *Molecular Ecology Notes*, vol. 5, n° 1, p. 187-189, 2005, doi: 10.1111/j.1471-8286.2004.00845.x.
[39] L. Excoffier et H. E. L. Lischer, « Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows », *Mol Ecol Resour*, vol. 10, n° 3, p. 564-567, mai 2010, doi: 10.1111/j.1755-0998.2010.02847.x.
[40] R. Peakall et P. E. Smouse, « GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update », *Bioinformatics*, vol. 28, n° 19, p. 2537-2539, oct. 2012, doi: 10.1093/bioinformatics/bts460.
[41] O. Pruvost, K. Boyer, V. Ravignné, D. Richard, et C. Vernière, « Deciphering how plant pathogenic bacteria disperse and meet: Molecular epidemiology of *Xanthomonas* citri pv. citri at microgeographic scales in a tropical area of Asiatic citrus canker endemicity », *Evolutionary Applications*, vol. 12, sept. 2019, doi: 10.1111/eva.12788.
[42] R. W. Saaty, « The analytic hierarchy process—what it is and how it is used », *Mathematical Modelling*, vol. 9, n° 3, p. 161-176, janv. 1987, doi: 10.1016/0270-0255(87)90473-8.
[43] H. Akaike, « Information Theory and an Extension of the Maximum Likelihood Principle », in *Selected Papers of Hirotugu Akaike*, E. Parzen, K. Tanabe, et G. Kitagawa, Éd. New York, NY: Springer, 1998, p. 199-213. doi: 10.1007/978-1-4612-1694-0_15.
[44] G. Schwarz, « Estimating the Dimension of a Model », *The Annals of Statistics*, vol. 6, n° 2, p. 461-464, 1978.
[45] J. E. Cavanaugh, « A large-sample model selection criterion based on Kullback's symmetric divergence », *Statistics & Probability Letters*, vol. 42, n° 4, p. 333-343, mai 1999, doi: 10.1016/S0167-7152(98)00200-4.

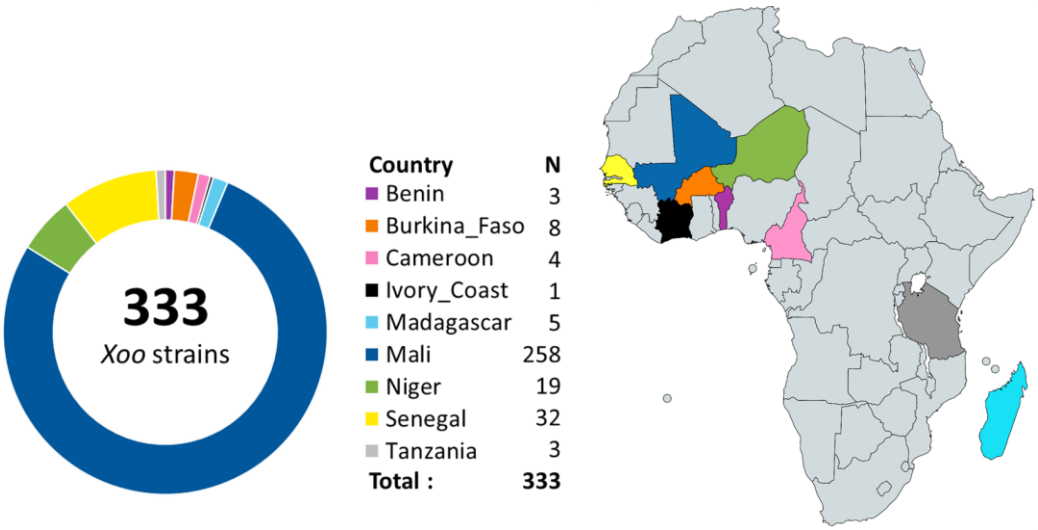


Figure 1 : African *Xoo* strains characterized in this study. In total, 333 African *Xoo* strains were considered in our analysis. The number of strains originating from each country (N) is visible on the right of the country name. On the map, the countries of origin are highlighted using the same color code.

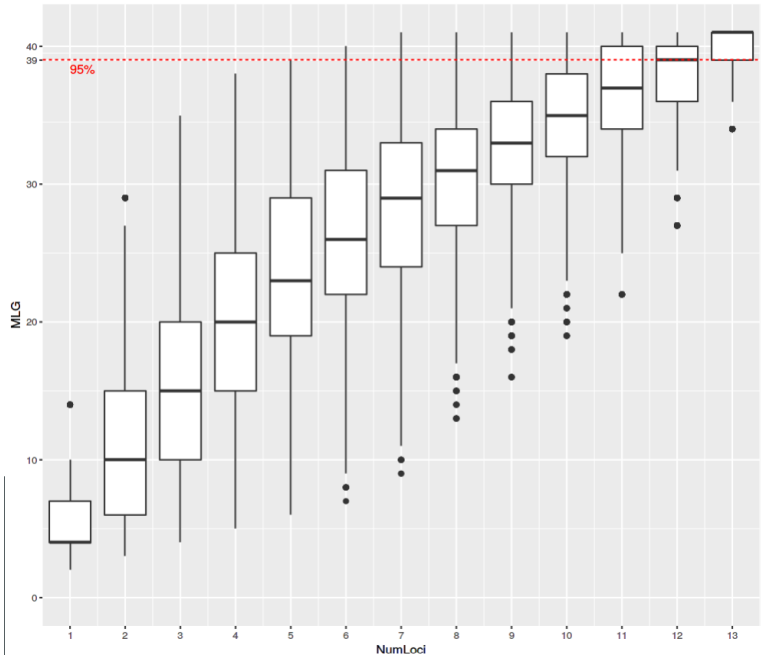


Figure 2 : Markers rarefaction analysis for African *Xoo* genotypes determination. Genotype accumulation curve for 334 *Xoo* strains (333 African *Xoo* strains and the Filipino strain PXO99^{A}) genotyped over 14 loci (42 MLG in total). The number of loci randomly sampled without replacement (up to n-1 loci) is represented on the x-axis. The y-axis shows the number of unique multilocus genotypes (MLG) observed in the data set. The red dashed line indicates the number of unique MLG incorporating 95% of our *Xoo* panel diversity.

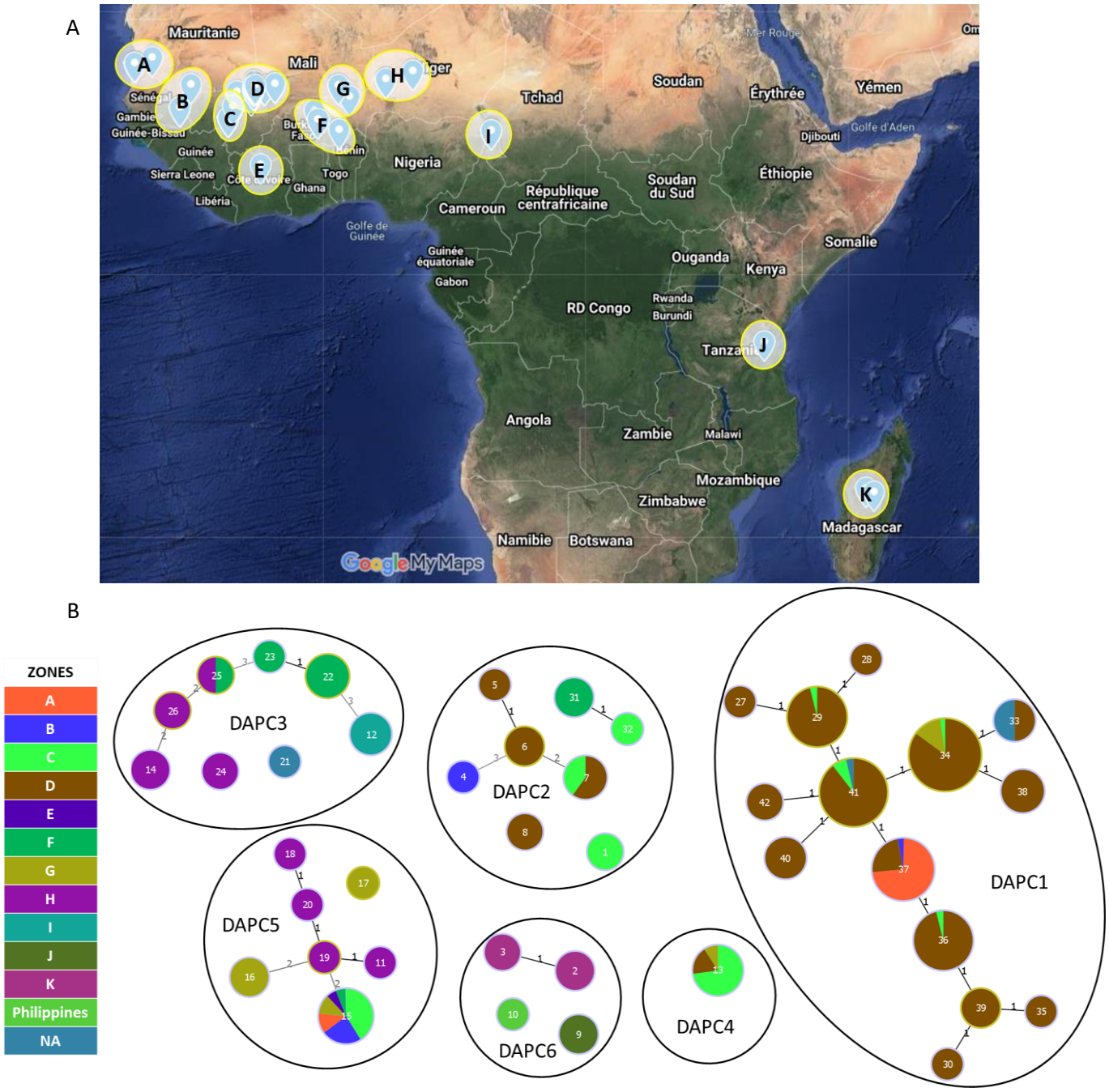


Figure 3 : Genetic diversity of *Xanthomonas oryzae* pv. *oryzae* in Africa. A. Areas of BLB sampling from 2003 to 2019. The sampling sites were segmented into zones in order to study population structure with no geographic constraints regarding frontiers. Eleven zones were determined, from A to K. **B.** Minimum spanning trees were generated using goeBURST in PHYLOVIZ. The number in each circle corresponds to the identifier of an individual haplotype. Numbers above the links between haplotypes indicate the number of polymorphic VNTR loci distinguishing haplotypes. Links marked in black connect Single Locus Variants (SLV), while those in gray connect haplotypes that differ by two or three VNTR loci. Haplotypes differing by more than three loci are not linked together. The DAPC clusters presented in Table 1 are circled in black. The colors of each circle display the geographic zone presented in Figure 3A of the isolates grouped into the considered haplotype.

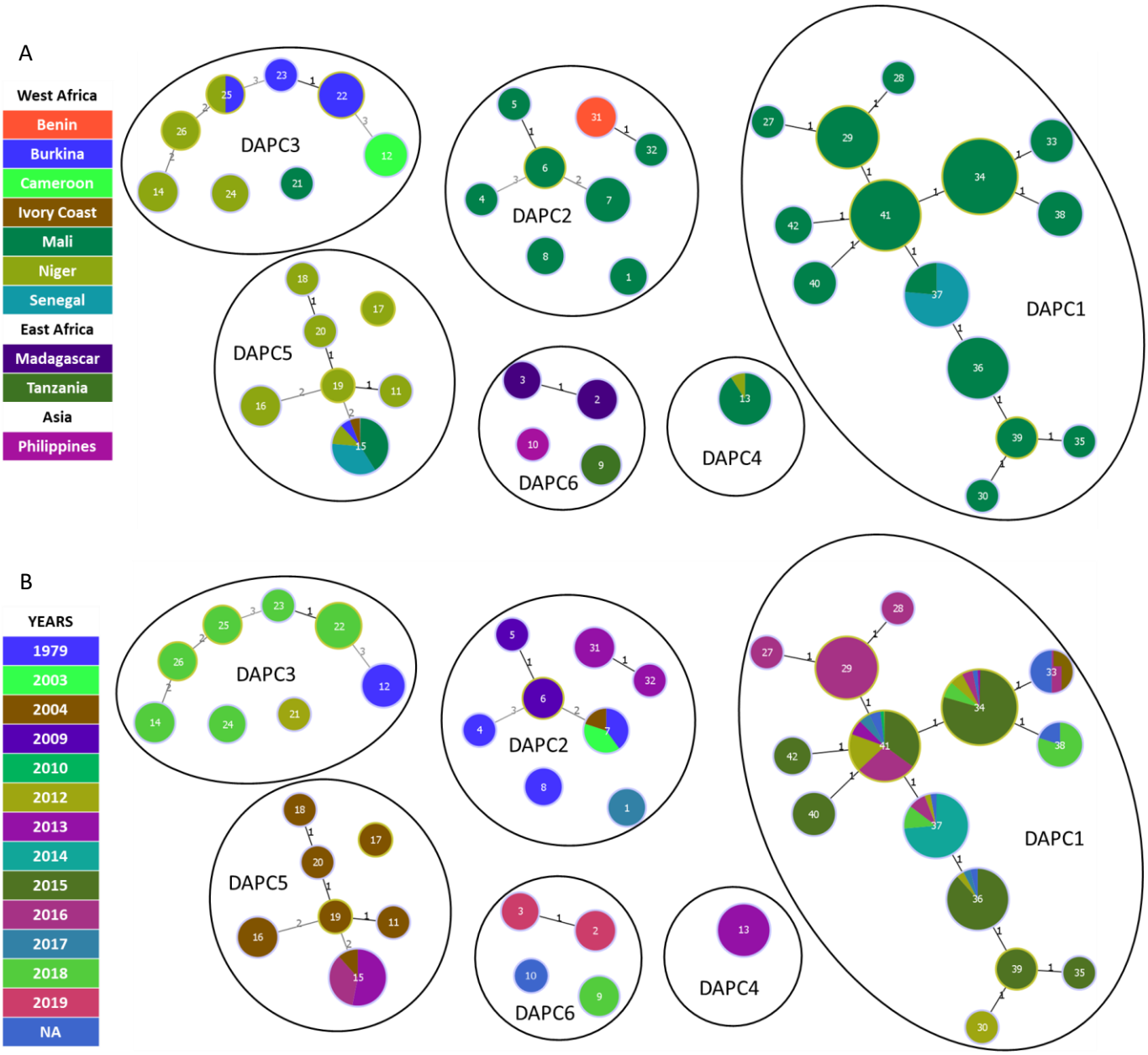
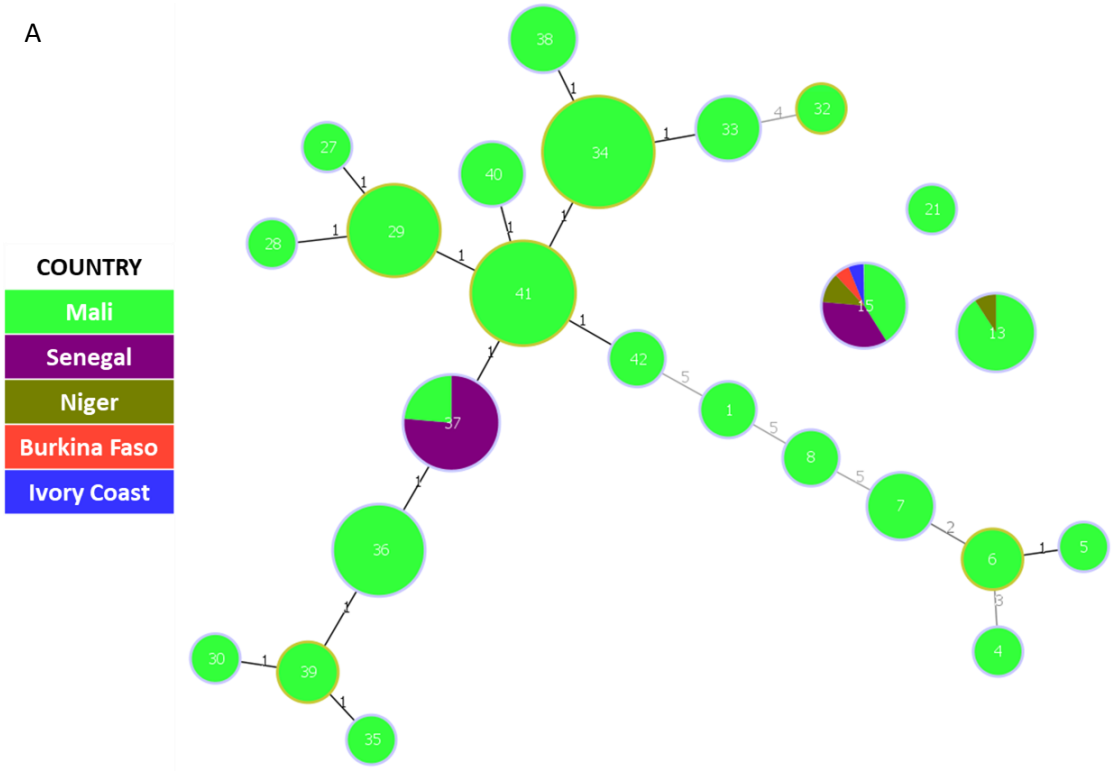
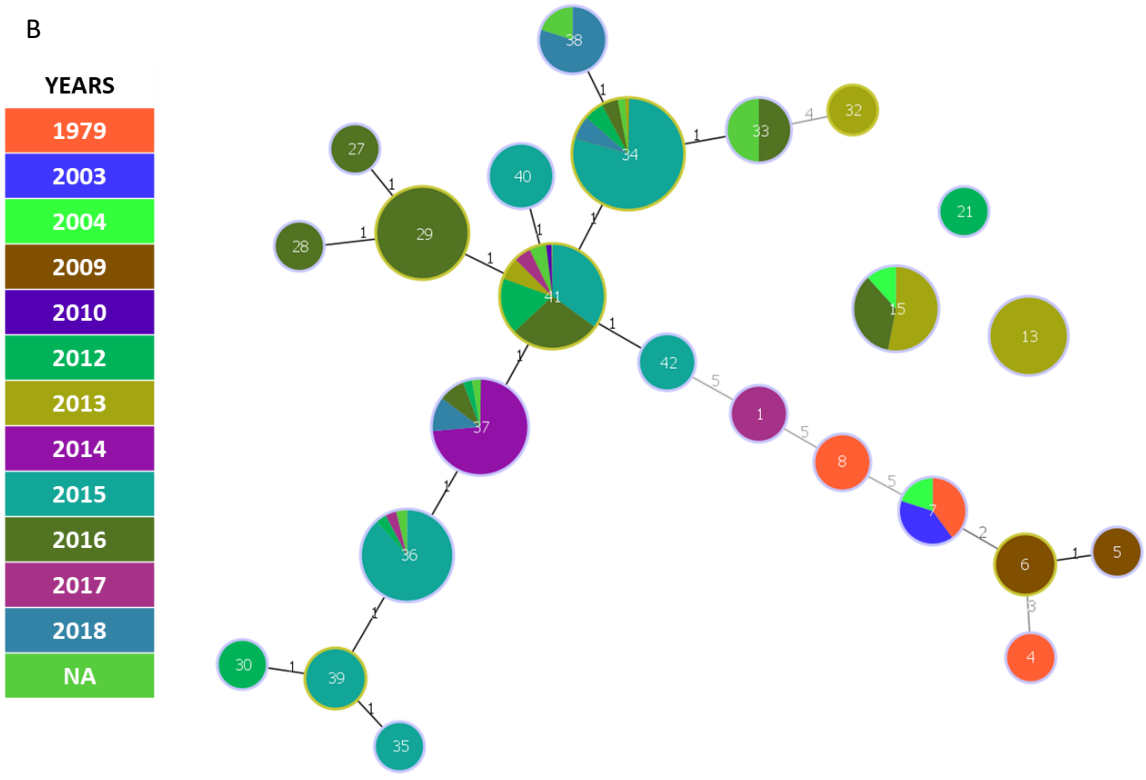


Figure 4 : Minimum spanning tree (MST) of African strains of *Xanthomonas oryzae* pv. *oryzae* based on multiple locus VNTR haplotypes. Minimum spanning trees were generated using goeBURST in PHYLOVIZ. The number in each circle corresponds to the identifier of an individual haplotype. Numbers above the links between haplotypes indicate the number of polymorphic VNTR loci distinguishing haplotypes. Links marked in black connect Single Locus Variants (SLV), while those in gray connect haplotypes that differ by two or three VNTR loci. Haplotypes differing by more than three loci are not linked together. The DAPC clusters presented in Table 1 are circled in black. The colors of each circle display information about the isolates grouped into the considered haplotype with **A**, the country of origin and **B**, the year of collection.

A


COUNTRY
Mali
Senegal
Niger
Burkina Faso
Ivory Coast
38
34
33
32
27
40
29
28
41
21
15
13
42
37
1
8
7
6
5
36
4
30
39
35

B


YEARS
1979
2003
2004
2009
2010
2012
2013
2014
2015
2016
2017
2018
NA
38
34
33
32
27
40
29
28
41
21
15
13
42
37
1
8
7
6
5
36
4
30
39
35

C

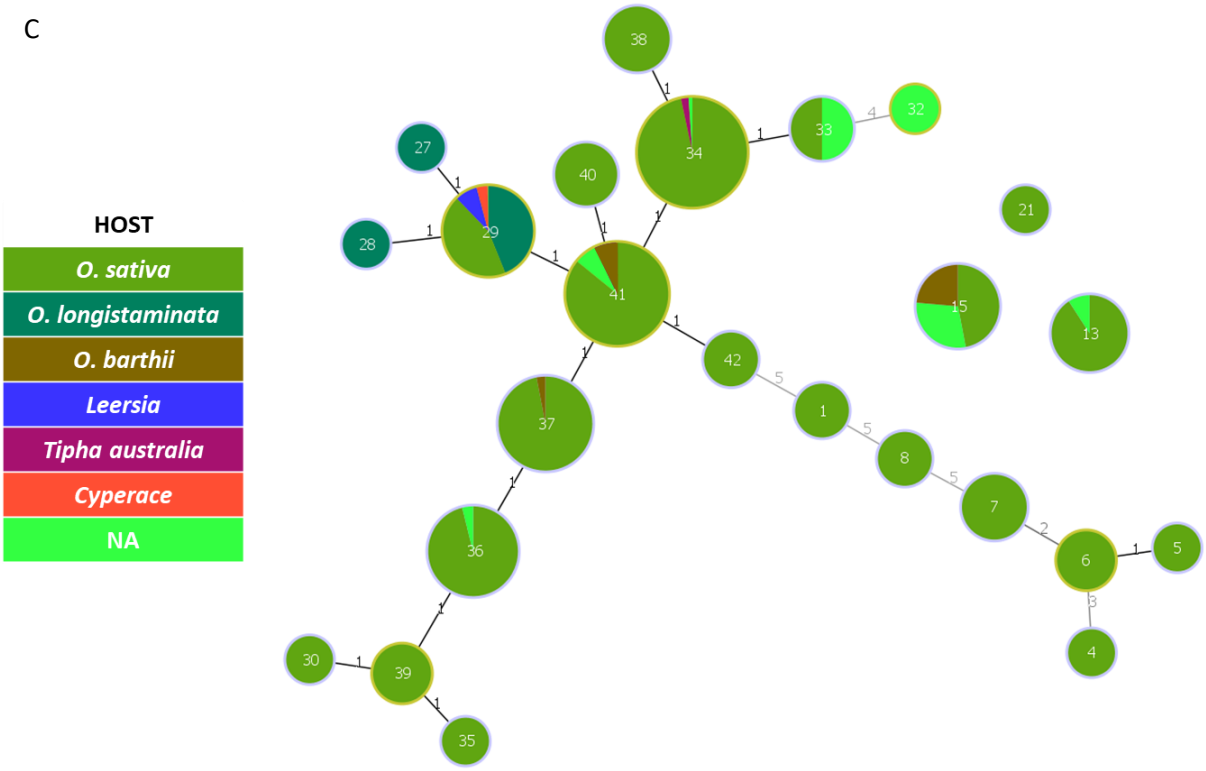


Figure 5 : Minimum spanning tree based on *Xoo* Multi Locus VNTR haplotypes identified in Mali. Minimum spanning tree was generated using goeBURST algorithm in PHYLOVIZ. The number in each circle represents an individual haplotype. Numbers above the links between haplotypes indicate the number of polymorphic TR loci distinguishing haplotypes. Links in black connect Single Locus Variants (SLV). The gray branches link haplotypes that differ in two to five TR loci, in order to highlight potential introgression events into the country. Above five TR loci of difference, no links are displayed. The colors of each circle display information about the isolates grouped into the considered haplotype : concerning countries of origin in **A** ; and year of sampling in **B** ; sampled host in **C**.

Tables :

Table 1 : Clusters of Multi Loci TR haplotypes determined by DAPC.

Country	N	DAPC1	DAPC2	DAPC3	DAPC4	DAPC5	DAPC6	MLG
Benin	3		3					1
Burkina_Faso	8			7		1		4
Cameroon	4			4				1
Ivory_Coast	1					1		1
Madagascar	5						5	2
Mali	258	225	15	1	10	7		24
Niger	19			8	1	10		12
Philippines	1						1	1
Senegal	32	26				6		2
Tanzania	3						3	1
Total N	334	251	18	20	11	25	9	/
Total MLG	/	14	8	8	1	7	4	42

N= number of individuals in each country; MLG= Number of MultiLocus Genotypes (MLG) observed in each country; Total N= total number of individuals in each group; Total MLG= total number of MLG observed in each group.

Table 2 : Analysis of the genotypic and genetic resources in West-African *Xoo* populations.

		Genotypic								
		Richness		Diversity	Evenness	Gene diversity	Allelic richness		Linkage disequilibrium*	
Population	N	MLG	eMLG	Simpson's index	E.5	H	average observed	average private	I_a	$\bar{r}D$
Burkina Faso	8	4	4	0,563	0,668	0,14	1,73	0,07	1,73	0,29
Mali	258	24	4,818	0,813	0,567	0,184	2,39	0,59	4,63	0,52
Niger	19	12	6,642	0,898	0,894	0,365	2,81	0,7	2,48	0,31
Senegal	32	2	1,851	0,305	0,706	0,225	1,89	0,05	9,00	1,00
Total	317	36	5,409	0,862	0,557	0,261	/	/	5,50	0,60

N, number of individuals in each population studied; MLG, Number of MultiLocus Genotypes observed; eMLG, number of expected MLG at the smallest sample size (here N=8) based on rarefaction; Simpson's Index, corresponds to the probability that two randomly selected genotypes inside a population are different (with 0 no genotypes are different, 1 all genotypes are different); E.5, Evenness corresponds to the genotypes abundance/frequency in each population [15], [41]; H, Nei's diversity index represents the average gene diversity per locus; Allelic richness observed, the average number of different alleles observed for each locus; Allelic richness private, the average number of different alleles observed for one sub-population for each locus; I_a, index of association; $\bar{r}D$, same but takes the number of loci into account.

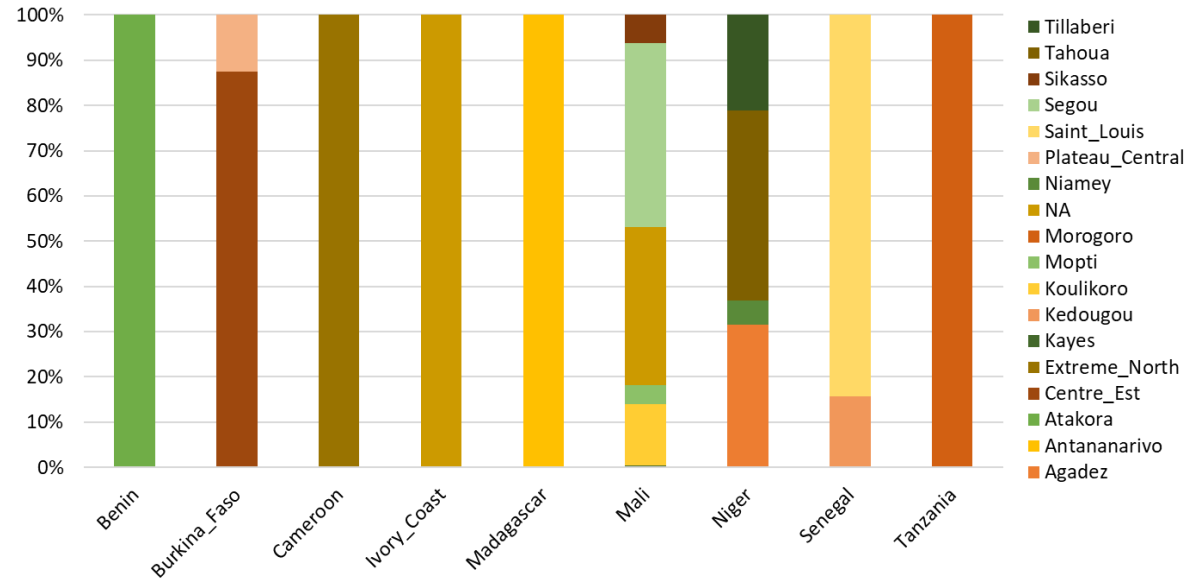
*The p values associated with I_a and $\bar{r}D$ are all < 0.01.

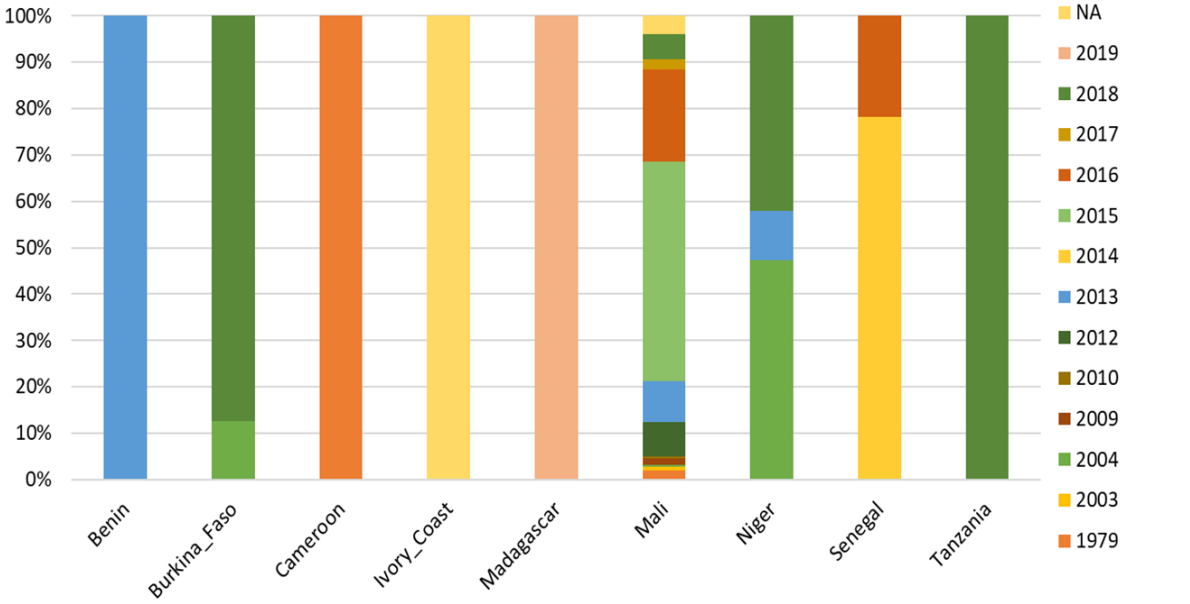
Table 3 : Genetic differentiation of populations of *Xoo* from four West African countries estimated by FST and RST.

RST\FST	Burkina Faso	Mali	Niger	Senegal
Burkina Faso	/	0,72****	0,18**	0,65****
Mali	0,78****	/	0,64****	0,18****
Niger	0,13*	0,73****	/	0,50****
Senegal	0,73****	0,003	0,66****	/

Under the diagonal, in black, are the RST values calculated for each country pairwise comparison. Above the diagonal, in grey, are the FST indices calculated for each pair compared. Associated p-values are represented with asterisks : p-values < 0,000**** ; 0,001*** ; 0,01** ; 0,1* ; 0,5.

Supplemental data





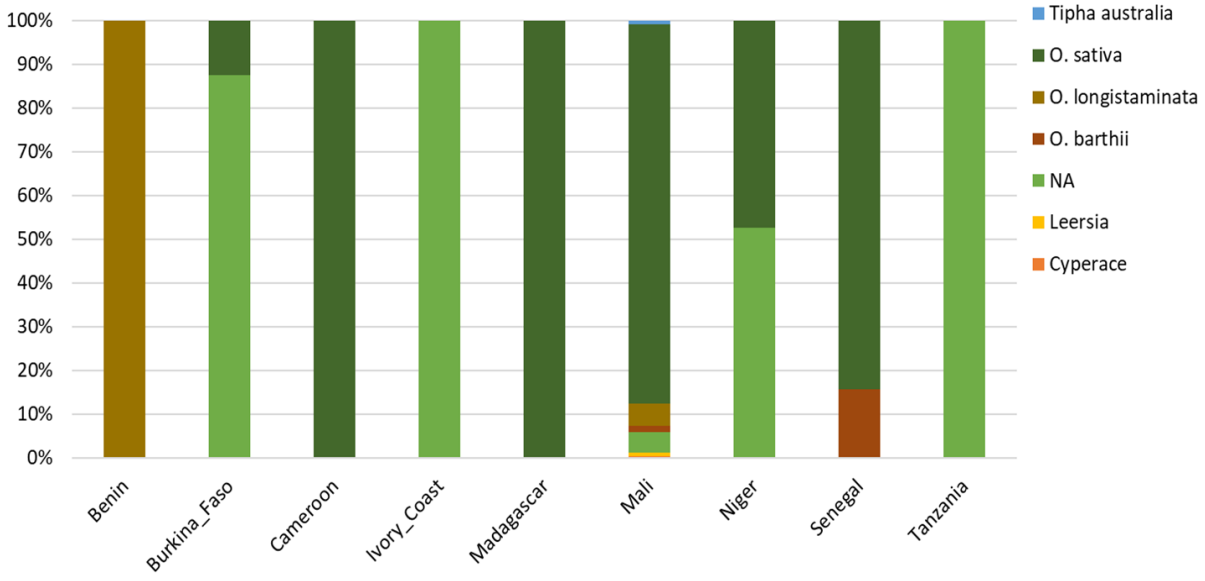


Figure S1 : Composition of the African *Xoo* panel.
On these three graphs are represented the percentages of samples collected in each country according to different categories, respectively: region, year and host of collection. The disparities in the diversity of representation of the different categories are clearly visible between the countries studied. The different categories are represented as percentages because the population sizes are disparate (see country N in Figure 1).

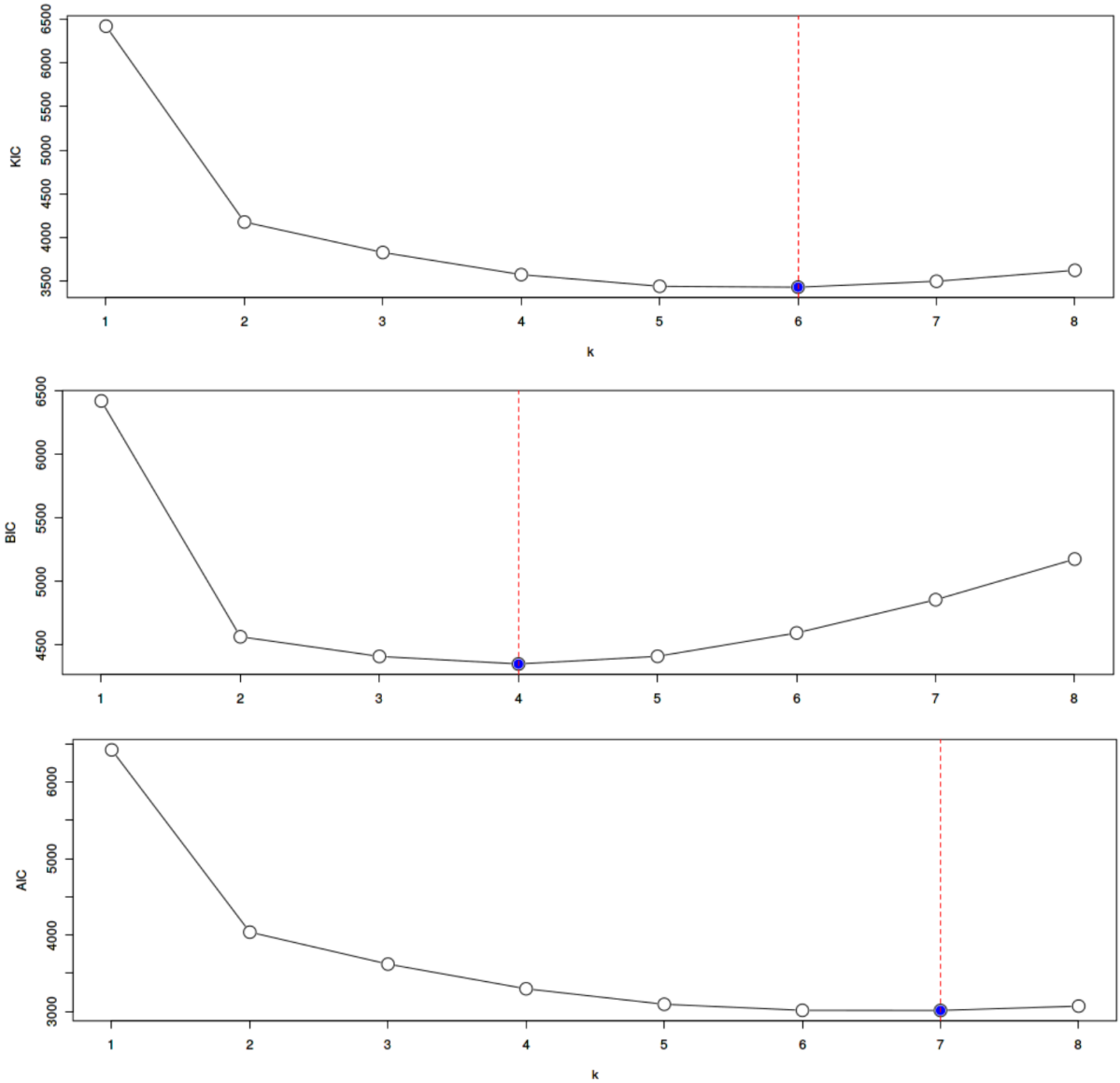


Figure S2 : Determination of the optimal number of clusters to represent African *Xoo* population structure. The optimal number of clusters was determined using model-based clustering with the combination of the Analytic Hierarchy Process (AHP; [42]). The AHP model has been created by using the information criteria Akaike's Information Criterion (AIC; [43]), Bayesian Information Criterion (BIC; [44]), and Kullback Information Criterion (KIC; [45]). The number of 6 clusters was chosen as it was the point of inflexion of the curve for both AIC and KIC.

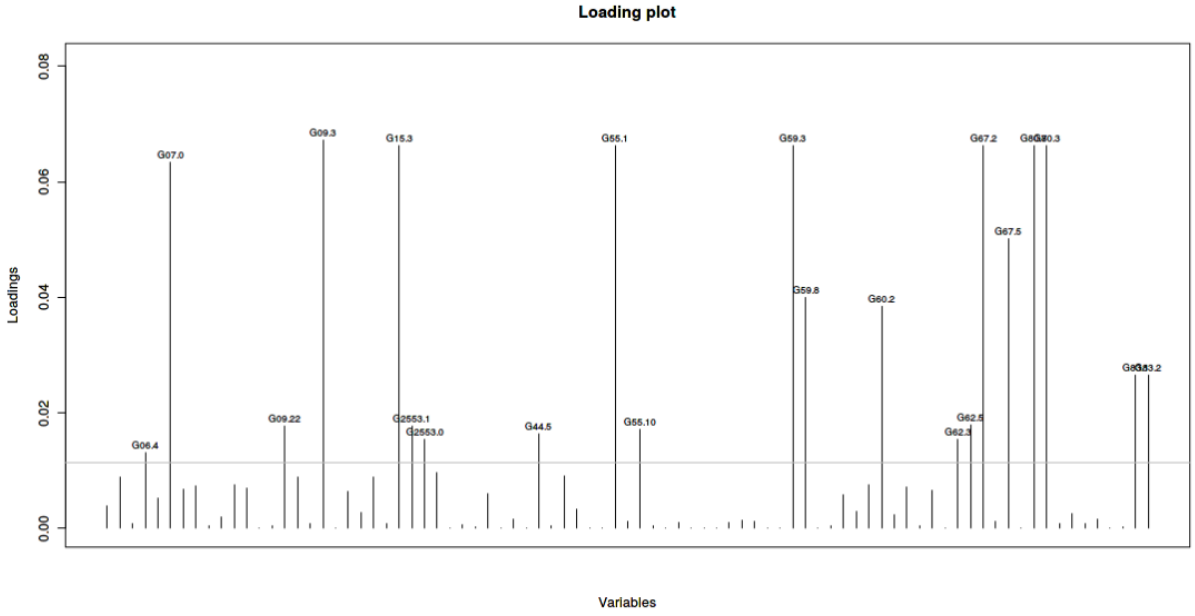


Figure S3 : Loading Plot of TR loci contributing to the Discriminant Analysis of Principal Component of *Xoo* in Africa.

The Loading Plot represents how the original variables contribute to the discriminant functions created by DAPC. Variables are plotted along the x-axis, and the contribution of those variables to the DAPC is plotted in the y-axis. Each variable, or allele, is function of a TR locus and number of repetitions at this locus. For example, the variable 3 repetitions at locus G09 is among the eight most informative variables to discriminate clusters by DAPC.

Table S1 : List of the strains used in this study and corresponding MLVA haplotypes.

CIX ID	Strain Name	Host	Country	Region	Place	Zone	Year	Haplo type	DAPC cluster
236	PXO99A	O. *sativa*	Philippines	Philippines	Philippines	NA	1980	10	6
607	Nati_Park	O. *longistaminata*	Benin	Atakora	Tanguieta	F	2013	31	2
609	Tanguieta_3	O. *longistaminata*	Benin	Atakora	Tanguieta	F	2013	31	2
611	Tanguieta_4	O. *longistaminata*	Benin	Atakora	Tanguieta	F	2013	31	2
627	Toula_13	NA	Niger	Agadez	Toula	G	2013	15	5
629	Toula_20	NA	Niger	Agadez	Toula	G	2013	13	4
633	Niono_16	O. *sativa*	Mali	Segou	Niono	D	2013	41	1
641	Niono_39	O. *sativa*	Mali	Segou	Niono	D	2013	13	4
643	Niono_41	O. *sativa*	Mali	Segou	Niono	D	2013	34	1
645	Niono_42	O. *sativa*	Mali	Segou	Niono	D	2013	13	4
649	Niono_46	O. *sativa*	Mali	Segou	Niono	D	2013	41	1
651	Niono_56	O. *sativa*	Mali	Segou	Niono	D	2013	41	1
653	Selingue_5	O. *sativa*	Mali	Sikasso	Selingue	C	2013	41	1
655	Selingue_45	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
663	Selingue_13	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
665	Selingue_18	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
669	Selingue_24	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
673	Selingue_2	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
675	Selingue_1	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
677	Selingue_22	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
681	Selingue_11	O. *sativa*	Mali	Sikasso	Selingue	C	2013	13	4
685	Selingue_10	O. *sativa*	Mali	Sikasso	Selingue	C	2013	15	5
687	Baguineda_12	O. *sativa*	Mali	Koulikoro	Baguineda	C	2013	15	5
689	Baguineda_13	NA	Mali	Koulikoro	Baguineda	C	2013	15	5
691	Baguineda_17-1	NA	Mali	Koulikoro	Baguineda	C	2013	32	2
692	Baguineda_18	NA	Mali	Koulikoro	Baguineda	C	2013	15	5
694	Baguineda_19	NA	Mali	Koulikoro	Baguineda	C	2013	15	5
696	Baguineda_20	O. *sativa*	Mali	Koulikoro	Baguineda	C	2013	15	5
698	Baguineda_31	O. *sativa*	Mali	Koulikoro	Baguineda	C	2013	15	5
767	MAI14_M20	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2009	5	2
768	MAI15_M23	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2009	6	2
770	MAI68	NA	Mali	NA	NA	D	NA	41	1
771	MAI70	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	36	1
773	MAI72	NA	Mali	NA	NA	D	NA	41	1
775	MAI74	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
776	MAI75	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
777	MAI83	NA	Mali	NA	NA	D	NA	34	1
778	MAI92	NA	Mali	NA	NA	D	NA	41	1
779	MAI93	NA	Mali	NA	NA	D	NA	36	1
781	MAI95	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	34	1
782	MAI96	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	37	1
784	MAI99	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	30	1
785	MAI100	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
787	MAI104	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
788	MAI105	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	34	1
789	MAI106	O. *sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	34	1
790	MAI107	O. *sativa*	Mali	Sikasso	Selingue	C	2012	34	1
791	MAI108	O. *sativa*	Mali	Sikasso	Selingue	C	2012	34	1
892	MAI129	O. *sativa*	Mali	NA	NA	NA	2012	41	1
893	MAI134	O. *sativa*	Mali	NA	NA	NA	2012	21	3
894	MAI145	O. *sativa*	Mali	NA	NA	NA	2012	41	1
1042	*XOO*_CII-2	NA	Ivory_Coast	NA	NA	E	2013	15	5
1250	S62-2-23	O. *sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
1253	S62-2-8	O. *sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2354	S62-2-2	O. *sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1

2355	S62-2-3	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2356	S62-2-4	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2357	S62-2-5	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2358	S62-2-6	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2359	S62-2-7	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2361	S62-2-9	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2362	S62-2-10	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2363	S62-2-11	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2364	S62-2-12	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2365	S62-2-13	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2366	S62-2-14	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2367	S62-2-15	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2368	S62-2-16	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2369	S62-2-17	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2372	S62-2-20	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2374	S62-2-22	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
2679	BAI3	*O. sativa*	Burkina_ Faso	Centre_Est	Bagre	F	2004	15	5
2711	MAI1	*O. sativa*	Mali	Segou	Niono	D	2003	7	2
2719	MAI09	*O. sativa*	Mali	Segou	Molodo	D	2004	7	2
2725	MAI15	*O. sativa*	Mali	Segou	Office_Niger_Molodo	D	2009	6	2
2779	NAI1	*O. sativa*	Niger	Niamey	Saga	G	2004	17	5
2780	NAI2	*O. sativa*	Niger	Agadez	Toula	H	2004	20	5
2781	NAI3	*O. sativa*	Niger	Tillaberi	Bonfeba	G	2004	16	5
2782	NAI4	*O. sativa*	Niger	Tillaberi	Bonfeba	G	2004	16	5
2785	NAI7	*O. sativa*	Niger	Agadez	Toula	H	2004	11	5
2786	NAI8	*O. sativa*	Niger	Tillaberi	Bonfeba	G	2004	16	5
2787	NAI9	*O. sativa*	Niger	Tillaberi	Bonfeba	G	2004	15	5
2799	CFBP_1947	*O. sativa*	Cameroo n	Extreme_North	Maroua	I	1979	12	3
2801	CFBP_1948	*O. sativa*	Cameroo n	Extreme_North	Maroua	I	1979	12	3
2802	CFBP_1948-94	*O. sativa*	Cameroo n	Extreme_North	Maroua	I	1979	12	3
2804	CFBP_1949	*O. sativa*	Mali	Segou	Niono_Kogoni	C	1979	7	2
2807	CFBP_1952	*O. sativa*	Mali	Segou	Niono_Diabali	D	1979	8	2
2951	S69-1-1	*O. sativa*	Senegal	Saint_Louis	Ndioum	A	2016	15	5
2952	S69-1-2	*O. sativa*	Senegal	Saint_Louis	Ndioum	A	2016	15	5
2969	S82-3-9	*O. barthii*	Senegal	Kedougou	Bandafassi	B	2016	37	1
2974	S82-4-1	*O. barthii*	Senegal	Kedougou	Bandafassi	B	2016	15	5
2975	S82-4-2	*O. barthii*	Senegal	Kedougou	Bandafassi	B	2016	15	5
2976	S82-4-3	*O. barthii*	Senegal	Kedougou	Bandafassi	B	2016	15	5
2977	S82-4-4	*O. barthii*	Senegal	Kedougou	Bandafassi	B	2016	15	5
4074	MAI53	*O. sativa*	Mali	NA	NA	D	2012	41	1
4077	MAI55	*O. sativa*	Mali	NA	NA	D	2010	41	1
4078	MAI71	*O. sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
4079	MAI73	*O. sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
4080	MAI101	*O. sativa*	Mali	Segou	Office_Niger_Molodo	D	2012	41	1
4084	S62-2-18	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4085	S62-2-19	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4086	S62-2-21	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4087	S62-2-24	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4089	S62-2-26	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4090	S62-2-27	*O. sativa*	Senegal	Saint_Louis	Ndiaye	A	2014	37	1
4098	MAI02	*O. sativa*	Mali	Segou	Niono	D	2003	7	2
4099	NAI5	*O. sativa*	Niger	Agadez	Toula	H	2004	18	5
4100	NAI6	*O. sativa*	Niger	Agadez	Toula	H	2004	19	5
4101	CFBP_1952-94	*O. sativa*	Mali	Segou	Niono_Diabali	D	1979	8	2
4120	MAI14	*O. sativa*	Mali	Segou	Office_Niger_Molodo	D	2009	6	2
4121	CFBP_1951	*O. sativa*	Mali	Kayes	Kayes	B	1979	4	2
4122	LMG_12464	*O. sativa*	Cameroo n	Extreme_North	Maroua	I	1979	12	3
4123	LMG_12466	*O. sativa*	Mali	Segou	Niono_Kogoni	C	1979	7	2

4218	MAI148	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4219	MAI149	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4220	MAI150	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4221	MAI151	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4222	MAI154	*Leersia*	Mali	Segou	Bewani	D	2016	29	1
4223	MAI157	*O. longistaminata*	Mali	Segou	Bewani	D	2016	28	1
4224	MAI158	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4225	MAI159	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4226	MAI160	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4227	MAI161	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4228	MAI162	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4229	MAI163	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4230	MAI164	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4231	MAI165	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4232	MAI166	*O. longistaminata*	Mali	Segou	Bewani	D	2016	27	1
4233	MAI167	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4234	MAI168	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4235	MAI170	*O. sativa*	Mali	Segou	Bewani	D	2016	29	1
4237	MAI183	*O. sativa*	Mali	Segou	Molodo_Sokoura	D	2016	29	1
4238	MAI184	*O. sativa*	Mali	Segou	Molodo_Sokoura	D	2016	29	1
4239	MAI185	*O. sativa*	Mali	Segou	Molodo_Sokoura	D	2016	29	1
4240	MAI258	*O. sativa*	Mali	Segou	Niono	D	2016	34	1
4241	MAI261	Tipha_australia	Mali	Segou	Niono	D	2016	34	1
4242	MAI262	Tipha_australia	Mali	Segou	Niono	D	2016	34	1
4243	MAI263	*O. sativa*	Mali	Segou	Niono	D	2016	33	1
4244	MAI264	*O. sativa*	Mali	Segou	Niono	D	2016	34	1
4245	MAI265	*O. sativa*	Mali	Segou	Niono	D	2016	33	1
4246	MAI266	*O. sativa*	Mali	Segou	Niono	D	2016	41	1
4247	MAI271	*O. sativa*	Mali	Segou	Molodo	D	2016	41	1
4248	MAI272	*O. sativa*	Mali	Segou	Molodo	D	2016	41	1
4249	MAI273	*O. sativa*	Mali	Segou	Molodo	D	2016	41	1
4250	MAI278	*O. sativa*	Mali	Segou	Molodo	D	2016	37	1
4251	MAI279	*O. sativa*	Mali	Segou	Molodo	D	2016	37	1
4252	MAI289	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4253	MAI290	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4254	MAI291	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4255	MAI292	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4256	MAI317	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4257	MAI318	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4258	MAI330	*O. barthii*	Mali	Segou	Bewani	D	2016	41	1
4259	MAI344	*O. sativa*	Mali	Segou	Bewani	D	2016	41	1
4260	MAI155	*Leersia*	Mali	Segou	Bewani	D	2016	29	1
4261	MAI156	*O. longistaminata*	Mali	Segou	Bewani	D	2016	29	1
4262	MAI169	*O. sativa*	Mali	Segou	Bewani	D	2016	29	1
4263	MAI274	*O. sativa*	Mali	Segou	Molodo	D	2016	41	1
4264	MAI307	*O. barthii*	Mali	Segou	Bewani	D	2016	41	1
4265	MAI370	*O. sativa*	Mali	Koulikoro	Baguineda	C	2017	1	2
4266	MAI372	*O. sativa*	Mali	Koulikoro	Baguineda	C	2017	1	2
4267	MAI392	*O. sativa*	Mali	Sikasso	Selingue	C	2017	41	1
4268	MAI393	*O. sativa*	Mali	Sikasso	Selingue	C	2017	41	1
4269	MAI394	*O. sativa*	Mali	Sikasso	Selingue	C	2017	36	1
4270	MAI395	NA	Mali	Sikasso	Selingue	C	2017	41	1
4271	MAI646	*O. sativa*	Mali	Segou	Niono	D	2018	38	1
4272	MAI647	*O. sativa*	Mali	Segou	Niono	D	2018	38	1
4273	MAI648	*O. sativa*	Mali	Segou	Niono	D	2018	37	1
4274	MAI649	*O. sativa*	Mali	Segou	Niono	D	2018	37	1
4275	MAI650	*O. sativa*	Mali	Segou	Niono	D	2018	37	1
4276	MAI651	*O. sativa*	Mali	Segou	Niono	D	2018	37	1
4277	MAI652	*O. sativa*	Mali	Segou	Niono	D	2018	38	1
4278	MAI653	*O. sativa*	Mali	Segou	Niono	D	2018	38	1
4279	MAI654	*O. sativa*	Mali	Segou	Niono	D	2018	34	1
4280	MAI655	*O. sativa*	Mali	Segou	Niono	D	2018	34	1
4281	MAI656	*O. sativa*	Mali	Segou	Niono	D	2018	34	1

4283	MAI658	*O. sativa*	Mali	Segou	Niono	D	2018	34	1
4284	MAI659	*O. sativa*	Mali	Segou	Niono	D	NA	34	1
4285	MAI660	*O. sativa*	Mali	Segou	Niono	D	2018	34	1
4286	MAI661	*O. sativa*	Mali	Segou	Niono	D	2018	34	1
4287	MAI662	NA	Mali	NA	NA	NA	NA	33	1
4288	MAI663	*O. sativa*	Mali	Segou	Niono	D	NA	38	1
4289	MAI664	NA	Mali	NA	NA	NA	NA	33	1
4290	MAI665	*O. sativa*	Mali	Segou	Niono	D	NA	37	1
4291	MAI666	O. *barthii*	Mali	Segou	Bewani	D	2016	41	1
4292	MAI667	O. *barthii*	Mali	Segou	Bewani	D	2016	41	1
4293	MAI668	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4294	MAI669	*O. sativa*	Mali	Segou	Niono	D	2016	29	1
4295	MAI670	Cyperaceae	Mali	Koulikoro	Baguineda	C	2016	29	1
4297	MAI487	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	39	1
4298	MAI488	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4299	MAI489	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4301	MAI491	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	39	1
4302	MAI492	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	35	1
4303	MAI493	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	39	1
4304	MAI494	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4305	MAI495	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4306	MAI496	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4307	MAI497	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4310	MAI500	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4311	MAI501	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4313	MAI503	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4314	MAI504	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4315	MAI505	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4316	MAI506	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4317	MAI507	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4318	MAI508	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	41	1
4319	MAI509	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4320	MAI510	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	41	1
4321	MAI511	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	41	1
4322	MAI512	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4323	MAI513	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	41	1
4324	MAI514	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4325	MAI515	*O. sativa*	Mali	Koulikoro	Niabougou	D	2015	36	1
4334	MAI524	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4335	MAI525	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4336	MAI526	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4337	MAI527	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4338	MAI528	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4339	MAI529	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4340	MAI530	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4341	MAI531	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4342	MAI532	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4343	MAI533	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4344	MAI534	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4345	MAI535	*O. sativa*	Mali	NA	RC1	D	2015	41	1
4359	MAI549	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4360	MAI550	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4361	MAI551	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4362	MAI552	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4363	MAI553	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4364	MAI554	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4365	MAI555	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4366	MAI556	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4367	MAI557	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4368	MAI558	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4369	MAI559	*O. sativa*	Mali	Mopti	Niaminami	G	2015	34	1
4370	MAI560	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4371	MAI561	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1

4372	MAI562	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4373	MAI563	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4374	MAI564	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4375	MAI565	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4376	MAI566	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4377	MAI567	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4378	MAI568	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4379	MAI569	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4380	MAI570	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4381	MAI571	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4382	MAI572	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4383	MAI573	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4384	MAI574	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4385	MAI575	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4386	MAI576	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4387	MAI577	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4388	MAI578	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4389	MAI579	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4390	MAI580	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4391	MAI581	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4392	MAI582	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4393	MAI583	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4394	MAI584	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4395	MAI585	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4396	MAI586	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4397	MAI587	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4398	MAI588	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4399	MAI589	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4400	MAI590	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4401	MAI591	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4402	MAI592	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4403	MAI593	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4404	MAI594	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4405	MAI595	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4407	MAI597	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4409	MAI599	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4410	MAI600	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4411	MAI604	*O. sativa*	Mali	NA	N10	D	2015	36	1
4413	MAI606	*O. sativa*	Mali	NA	N10	D	2015	36	1
4415	MAI608	*O. sativa*	Mali	NA	N10	D	2015	34	1
4416	MAI609	*O. sativa*	Mali	NA	N10	D	2015	34	1
4417	MAI610	*O. sativa*	Mali	NA	N10	D	2015	34	1
4418	MAI611	*O. sativa*	Mali	NA	N10	D	2015	34	1
4419	MAI612	*O. sativa*	Mali	NA	N10	D	2015	34	1
4420	MAI613	*O. sativa*	Mali	NA	N10	D	2015	42	1
4421	MAI614	*O. sativa*	Mali	NA	N10	D	2015	41	1
4422	MAI615	*O. sativa*	Mali	NA	N10	D	2015	34	1
4423	MAI616	*O. sativa*	Mali	NA	N10	D	2015	42	1
4424	MAI617	*O. sativa*	Mali	NA	N10	D	2015	34	1
4425	MAI618	*O. sativa*	Mali	NA	N10	D	2015	34	1
4426	MAI620	*O. sativa*	Mali	NA	N10	D	2015	34	1
4427	MAI621	*O. sativa*	Mali	NA	N10	D	2015	34	1
4428	MAI622	*O. sativa*	Mali	NA	N10	D	2015	34	1
4429	MAI623	*O. sativa*	Mali	NA	N10	D	2015	34	1
4430	MAI624	*O. sativa*	Mali	NA	N10	D	2015	34	1
4431	MAI625	*O. sativa*	Mali	NA	N10	D	2015	36	1
4432	MAI626	*O. sativa*	Mali	NA	N10	D	2015	36	1
4433	MAI627	*O. sativa*	Mali	NA	N10	D	2015	36	1
4434	MAI628	*O. sativa*	Mali	NA	N10	D	2015	36	1
4436	MAI630	*O. sativa*	Mali	Segou	Molodo	D	2015	34	1
4437	MAI631	*O. sativa*	Mali	Segou	Molodo	D	2015	34	1
4439	MAI633	*O. sativa*	Mali	Segou	Molodo	D	2015	41	1
4440	MAI634	*O. sativa*	Mali	Segou	Molodo	D	2015	40	1

4441	MAI635	*O. sativa*	Mali	Segou	Molodo	D	2015	41	1
4442	MAI636	*O. sativa*	Mali	Segou	Molodo	D	2015	41	1
4443	MAI637	*O. sativa*	Mali	Segou	Molodo	D	2015	40	1
4444	MAI638	*O. sativa*	Mali	Segou	Molodo	D	2015	40	1
4445	MAI639	*O. sativa*	Mali	NA	M3	D	2015	40	1
4448	MAI642	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4449	MAI643	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4450	MAI644	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4451	MAI645	*O. sativa*	Mali	NA	N3_Bis	D	2015	34	1
4452	Md5	*O. sativa*	Madagascar	Antananarivo	Ivory	K	2019	2	6
4453	Md7	*O. sativa*	Madagascar	Antananarivo	Ivory	K	2019	2	6
4454	Md9	*O. sativa*	Madagascar	Antananarivo	Ivory	K	2019	2	6
4455	Md11	*O. sativa*	Madagascar	Antananarivo	Antsirabe	K	2019	3	6
4456	Md12	*O. sativa*	Madagascar	Antananarivo	Antsirabe	K	2019	3	6
4457	Tz1	NA	Tanzania	Morogoro	DAKAWA	J	2019	9	6
4458	Tz2	NA	Tanzania	Morogoro	DAKAWA	J	2019	9	6
4462	Tz6	NA	Tanzania	Morogoro	DAKAWA	J	2019	9	6
4465	BV2.6	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	22	3
4468	BV5.4	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	22	3
4469	BV5.5	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	22	3
4470	BV5.9	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	22	3
4471	BV9.6	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	22	3
4476	B3-7	NA	Burkina_Faso	Centre_Est	Bagre	F	2018	23	3
4481	M3-10	NA	Burkina_Faso	Plateau_Central	Mogtedo	F	2018	25	3
4482	N2-4	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	24	3
4483	N2-5	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	26	3
4486	N2-8	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	26	3
4488	N8	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	14	3
4489	N8-1	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	14	3
4490	N9-1	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	24	3
4491	N16	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	14	3
4492	N3-2	NA	Niger	Tahoua	Tava_Wandiwa	H	2018	25	3

Table S2 : Allele polymorphism observed for each locus.

Mix	Locus	Observed alleles	Polymorphic for
1	Xo_G06	4	Xoo Asia / Africa / Xoc
	Xo_G07	8	
	~~Xo_G58~~	/	
	Xo_G81	5	
2	Xo_G2553	9	Xoo Asia / Xoo Africa
	~~Xo_G88~~	/	
	Xo_G44	6	
	Xo_G62	3	
3	Xo_G55	14	Xoo Africa / Xoc
	Xo_G59	7	
	Xo_G60	5	
	Xo_G67	4	
4	Xo_G09	4	Xoo Asia / Xoc
	Xo_G15	4	
	Xo_G80	2	
	Xo_G83	2	

Loci Xo_G58 and Xo_G88, from Poulin *et al.* in 2015 do not amplify.

Table S3 : Number of repetitions at each locus of the determined haplotypes.

N = Number of strains in each haplotype.

			Repetitions per locus													
Haplotype	N	Country	G06	G07	G09	G15	G2553	G44	G55	G59	G60	G62	G67	G80	G81	G83
1	2	Mali	5	17	3	3	4	5	15	6	3	4	5	3	8	2
2	3	Madagascar	5	0	22	49	1	5	1	3	2	4	2	7	7	1
3	2	Madagascar	5	0	22	50	1	5	1	3	2	4	2	7	7	1
4	1	Mali	5	10	3	3	7	6	22	8	2	4	5	3	6	2
5	1	Mali	5	9	3	3	7	3	17	7	2	4	5	3	6	2
6	3	Mali	5	9	3	3	7	3	17	8	2	4	5	3	6	2
7	5	Mali	5	9	3	3	9	3	18	8	2	4	5	3	6	2
8	2	Mali	5	9	3	3	9	4	22	6	3	4	5	3	9	2
9	3	Tanzania	10	0	21	12	0	5	1	3	2	3	2	7	6	2
10	1	Philippines	8	0	10	41	0	4	1	3	2	3	2	7	5	1
11	1	Niger	4	0	3	3	7	5	10	8	5	5	6	3	7	2
12	4	Cameroon	4	16	3	3	7	5	11	5	6	4	6	3	6	2
13	11	Mali, Niger	4	15	3	3	4	6	10	7	4	4	5	3	6	2
14	3	Niger	4	11	3	3	3	6	13	6	5	4	6	3	6	2
15	17	Mali, Senegal, Burkina Faso, Niger, Ivory Coast	4	10	3	3	7	5	11	8	5	5	6	3	7	2
16	3	Niger	4	10	3	3	7	5	10	9	5	5	6	3	7	2
17	1	Niger	4	10	3	3	7	5	10	8	4	4	6	3	6	2
18	1	Niger	4	10	3	3	8	5	10	8	5	5	6	3	7	2
19	1	Niger	4	9	3	3	7	5	10	8	5	5	6	3	7	2
20	1	Niger	4	9	3	3	8	5	10	8	5	5	6	3	7	2
21	1	Mali	4	13	12	3	3	5	26	8	4	4	3	3	4	2
22	5	Burkina Faso	4	13	3	3	7	5	11	6	5	4	6	3	6	2
23	1	Burkina Faso	4	13	3	3	8	5	11	6	5	4	6	3	6	2
24	2	Niger	4	13	3	3	9	8	15	6	5	4	5	3	6	2
25	2	Burkina Faso, Niger	4	13	3	3	11	6	12	6	5	4	6	3	6	2
26	2	Niger	4	13	3	3	4	6	13	6	5	4	6	3	6	2
27	1	Mali	6	10	3	3	3	7	18	4	3	4	5	3	8	2
28	1	Mali	6	10	3	3	3	6	17	4	3	4	5	3	8	2
29	25	Mali	6	10	3	3	3	6	18	4	3	4	5	3	8	2
30	1	Mali	6	9	3	3	3	6	21	4	3	4	5	3	8	2
31	3	Benin	6	13	3	3	5	5	17	8	3	4	5	3	7	2
32	1	Mali	6	13	3	3	6	5	17	8	3	4	5	3	7	2
33	4	Mali	6	8	3	3	3	6	17	4	3	4	5	3	7	2
34	87	Mali	6	8	3	3	3	6	17	4	3	4	5	3	8	2
35	1	Mali	6	8	3	3	3	6	22	4	3	4	5	3	8	2
36	26	Mali	6	8	3	3	3	6	20	4	3	4	5	3	8	2
37	34	Mali, Senegal	6	8	3	3	3	6	19	4	3	4	5	3	8	2
38	5	Mali	6	8	3	3	3	6	16	4	3	4	5	3	8	2
39	3	Mali	6	8	3	3	3	6	21	4	3	4	5	3	8	2
40	4	Mali	6	8	3	3	3	6	18	4	3	4	5	3	9	2
41	57	Mali	6	8	3	3	3	6	18	4	3	4	5	3	8	2
42	2	Mali	6	8	3	3	4	6	18	4	3	4	5	3	8	2

7 Analyse de génomes Africains

7.1 Introduction

Grâce aux analyses MLVA réalisées, nous avons pu avoir une vue globale de la structuration de nos populations de *Xoo* Africaines. Ce premier travail nous a permis de dégager différents groupes génétiques. Notre objectif étant la compréhension des relations évolutives liant les diverses populations de *Xoo* en Asie et en Afrique, nous avons cherché à étudier les différences observables au niveau génomique entre ces populations.

De nombreux génomes de *Xoo* sont disponibles sur la base de données publique NCBI. Cependant ce contenu est très biaisé vers les populations d'Asie. Dans un effort continu de l'équipe depuis 2018 pour une meilleure description du pool génétique Africain, nous avons effectué le séquençage de 68 souches africaines par la technologie en lectures longues de Oxford Nanopore Technologies. 41 de ces souches furent séquencées au cours de ces travaux de thèse, sur la base du panel de typage MLVA. Ces 41 souches ont été choisies de manières à couvrir la totalité des haplotypes obtenus lors du typage, si aucune des représentantes de l'haplotype n'avait encore été séquencée.

Les assemblages furent réalisés et contrôlés, puis nous les avons intégrés dans une analyse globale contenant la totalité des génomes disponibles en ligne pour créer des phylogénies entre les populations de *Xoo* à l'échelle mondiale. Une partie des résultats de ces analyses phylogénétiques a été intégrée à l'article présenté dans le Chapitre II, soumis à la revue *Nature Communications*.

7.2 Résultats

7.2.1 L'assemblage des génomes Africains générés révèle la présence potentielle de plasmides.

Les reads obtenus à l'issu du traitement des signaux de séquençages ont été assemblés, avec des profondeurs de séquençage comprises entre 50 et 450X environ (figure 32A). La plupart des souches ont pu être assemblées correctement, c'est-à-dire ont fourni des contigs circulaires complets, autour de 5.10^6 pb. Ces assemblages primaires ont été corrigés par utilisation de reads Illumina NextSeq générés pour l'occasion.

A ces chromosomes complets est parfois associé un autre contig, lui aussi circulaire, et de taille comprise entre 40 et 70 kb (figure 32A), comme observé pour CIX4420 (Mali), CIX4108 (Burkina Faso), CIX2779 et CIX2780 (Niger). La couverture de ces contigs est différente de celle du reste du génome correspondant, étant environ 3 fois plus élevée. Il semble donc que ces trois souches détiennent chacune un plasmide. Une analyse plus poussée de la souche Nigériane CIX2779 a confirmé cette théorie. Cette analyse, constituant la première description d'un génome africain de *Xoo* contenant un plasmide a donné lieu à la rédaction d'un article présenté en partie 8 de ce chapitre.

7.2.2 La construction d'une phylogénie des populations mondiales de *Xoo*

Afin de mieux comprendre l'évolution du pathogène en Afrique et de mettre au clair les origines des *Xoo* récemment apparues en Afrique de l'Est, nous avons inféré des relations phylogénétiques entre les différentes populations mondiales de *Xo* en intégrant nos séquences génomiques nouvellement créées à l'analyse.

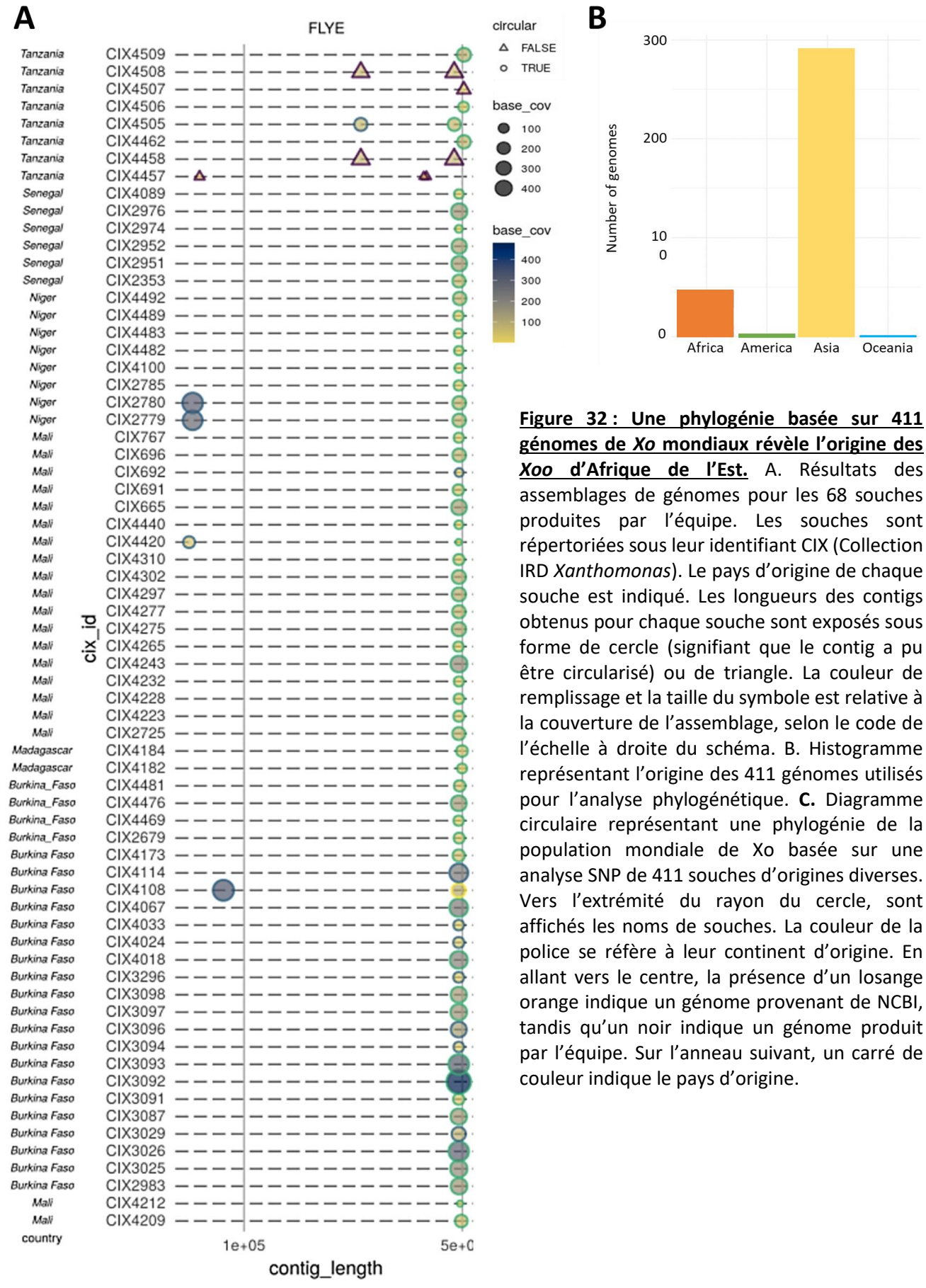
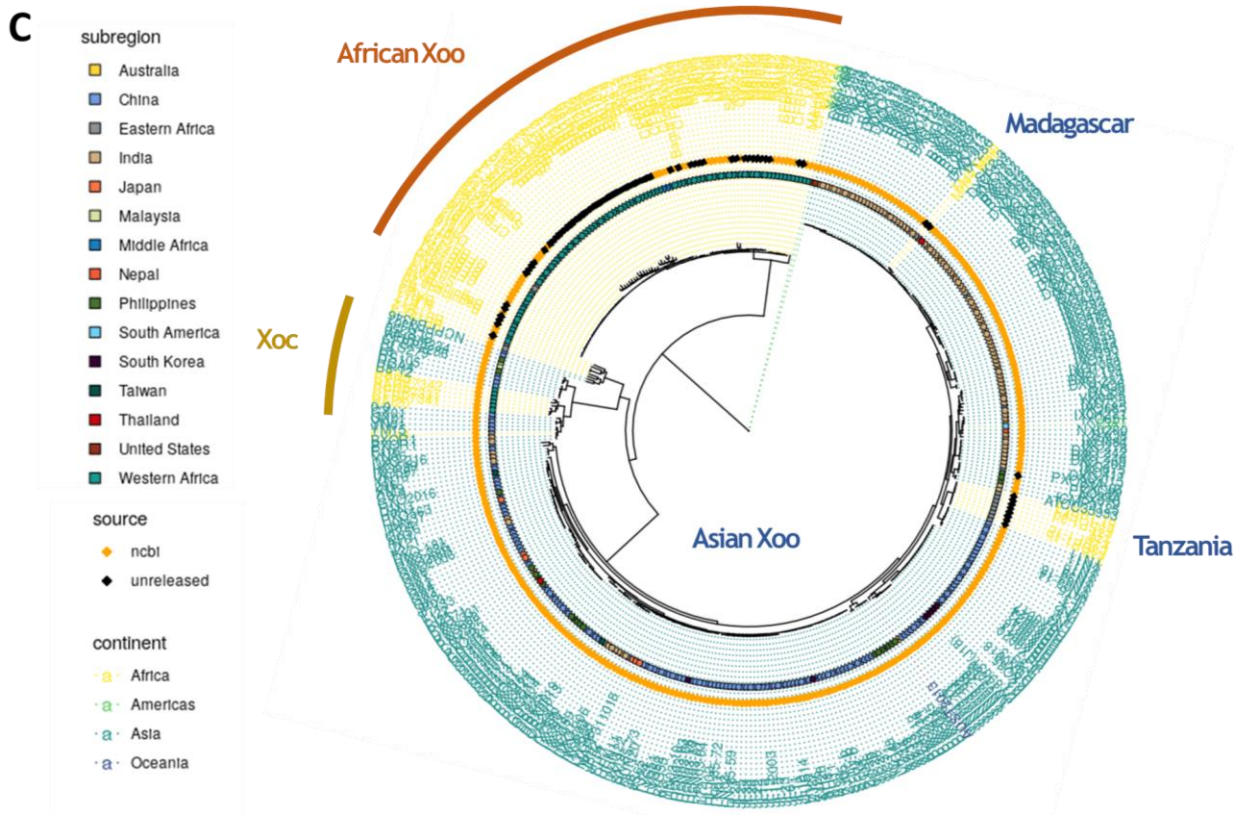


Figure 32 : Une phylogénie basée sur 411 génomes de *Xo* mondiaux révèle l'origine des *Xoo* d'Afrique de l'Est. A. Résultats des assemblages de génomes pour les 68 souches produites par l'équipe. Les souches sont répertoriées sous leur identifiant CIX (Collection IRD *Xanthomonas*). Le pays d'origine de chaque souche est indiqué. Les longueurs des contigs obtenus pour chaque souche sont exposés sous forme de cercle (signifiant que le contig a pu être circularisé) ou de triangle. La couleur de remplissage et la taille du symbole est relative à la couverture de l'assemblage, selon le code de l'échelle à droite du schéma. B. Histogramme représentant l'origine des 411 génomes utilisés pour l'analyse phylogénétique. **C.** Diagramme circulaire représentant une phylogénie de la population mondiale de Xo basée sur une analyse SNP de 411 souches d'origines diverses. Vers l'extrémité du rayon du cercle, sont affichés les noms de souches. La couleur de la police se réfère à leur continent d'origine. En allant vers le centre, la présence d'un losange orange indique un génome provenant de NCBI, tandis qu'un noir indique un génome produit par l'équipe. Sur l'anneau suivant, un carré de couleur indique le pays d'origine.



La totalité des assemblages de génomes de *Xo* présents sur la base de données NCBI ont été utilisés, représentant 343 génomes au moments de l'analyse, auxquels ont étés ajoutés les 68 génomes Africains générés par l'équipe. Sur ces 411 génomes, la grande majorité (82%) proviennent de souches isolées en Asie (figure 32B). Les souches africaines arrivent en seconde position des représentations, bien qu'elles ne constituent que 11% des entrées. Les 7% restants sont partagés entre le continent Américain et l'Océanie.

La phylogénie basée sur des SNPs générés pour ces 411 génomes a confirmé que les *Xoo* récemment prélevée en Tanzanie et à Madagascar groupaient avec les *Xoo* Asiatiques, tandis que les *Xoo* d'Afrique de l'Ouest formaient un groupe plus proche des *Xoc*, comme observé précédemment (Poulin et al., 2015 [285] ; Lang et al., 2019 [286] ; Oliva et al., 2019 [287]) (figure 32C). Trois souches prélevées en Tanzanie avant 2019 pendant une période où la BLB se faisait plus rare dans le pays ont un profil endémique des souches d'Afrique de l'Ouest. Ceci suggère une intrusion récente des populations Asiatiques sur le continent.

Une inspection plus fine cette phylogénie a montré que les souches Tanzaniennes collectées en 2019 (CIX4462, CIX4506 et CIX4509) sont très proches de souches provenant de la province du Yunnan en Chine. Les souches de Madagascar prélevées à cette période (CIX4184 et CIX4182) sont quant à elles plus proches de *Xoo* issues de Thaïlande et feront l'objet de futures caractérisations par l'équipe. Une analyse phénotypique plus poussée des *Xoo* issues de Tanzanie est présentée en dans l'article « Genome editing of an African rice elite variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains » présenté dans le Chapitre II.

7.3 Matériel et méthodes

Les extractions d'ADN, la préparation des banques, le séquençage et les assemblages ont été réalisés selon le protocole décrit dans l'article « Genome editing of an African rice elite variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains » présenté dans le Chapitre II.

Article 2 : « The complete genome resource of the *Xanthomonas oryzae* pv. *oryzae* strain CIX2779 includes the first sequence of a plasmid for an African representative of this rice pathogen »

The complete genome resource of the *Xanthomonas oryzae* pv. *oryzae* strain CIX2779 includes the first sequence of a plasmid for an African representative of this rice pathogen.

Authors
Sciallano C.[1] , Auguy F.[1] , Boulard G.[1], Szurek B.[1], Cunnac S.[1*]

Affiliations
[1]PHIM Plant Health Institute, Univ Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France
* Correspondence: sebastien.cunnac@ird.fr

Abstract
The bacterial plant pathogen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) is responsible for the foliar rice bacterial blight disease. Genetically contrasted, continent-specific, sublineages of this species can cause important damages to rice production both in Asia and Africa. We report on the genome of the CIX2779 strain of this pathogen, previously named NAI1 and originating from Niger. Oxford Nanopore long reads assembly and Illumina short reads polishing produced a genome sequence composed of a 4,725,792 bp circular chromosome and a 39,798 bp long circular plasmid designated pCIX2779_1. The chromosome structure and base-level sequence are highly related to reference strains of African *Xoo* and encode identical Transcription Activator-like effectors (TALE) for virulence. Importantly, our *in silico* analysis strongly indicates that pCIX2779_1 is a genuine conjugative plasmid, the first indigenous one sequenced from an African strain of the *X. oryzae* species.

Keywords
Oxford Nanopore Technologies; TAL effector; plant disease

Genome Announcement
The plant pathogenic bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) is the causative agent of bacterial leaf blight (BLB) of rice (*Oryza sativa*). This disease has been reported in most rice growing areas of Asia and Africa, and can severely harm yields and grain quality. The genetic diversity of the *Xoo* pathovar can be partitioned into two groups that reflect the continent of origin of the strains. In Africa, BLB has been first documented in 1979 in Mali and is being increasingly reported (Verdier et al. 2012). African *Xoo* strains notably differ from their Asian counterparts at the level of their Transcription Activator-Like Effectors (TALEs) repertoires holding up to 9 members that are unique to African strains (Tran et al. 2018; Doucouré et al. 2018).
In addition to *bona fide* chromosomes, bacterial genomes may include extra-chromosomal and autonomously replicating DNA molecules which are referred to as plasmids. They vary greatly in terms of size, sequence content or contribution to fitness. These genetic elements are often capable of shuttling between bacterial genotypes of varying relatedness and are key vehicles for horizontal gene transfer. Because they may contain antibiotic or heavy metal resistance genes, as well as virulence systems, plasmids can have a profound impact on bacterial pathogens adaptation to their hosts (Frost et al. 2005; Sundin 2007). Our knowledge on plasmids within *X. oryzae* strains is rather limited. Earlier molecular genetics work revealed that 14 out of 17 Asian *Xoo* isolates carry at least a plasmid. Furthermore, one of them contributed to symptom development (Amuthan and Mahadevan 1994). It is only recently that the first indigenous *X. oryzae* plasmid sequence was determined by Sanger sequencing from a Chinese isolate of the *oryzicola* pathovar. pXOCgx01 is a 53kb conjugative plasmid that improves tolerance to heavy metal (Niu et al. 2015). Subsequently, the first plasmid sequences from *Xoo* strains were assembled with long read sequencing of the genome of the Indian BXO1 (Kaur et al. 2019) and IX-280 (Carpenter et al. 2020) strains. However, so far, the plasmid content of African *X. oryzae* remains unexplored.
In an effort to build genomic resources enabling a better exploration of the genetic diversity of African *Xoo* strains, our group conducted long read sequencing of several strains in our collection with the MinION device of Oxford Nanopore Technologies (ONT). Here, we report on the genome sequence of the African *Xoo* strain CIX2779 which was isolated in 2004 from a diseased rice leaf collected in the Saga locality of the Niamey area in Niger and which was originally designated as NAI1 by Gonzalez *et al.* (Gonzalez et al. 2007).

Following extraction with QIAGEN Genomic-tip 100/G, the CIX2779 genomic DNA was used in multiplex library construction with the ONT rapid library preparation kit (SQK-RBK110-96) and sequenced with a MinION Mk1C device on a R10.3 (FLO-MIN111) flowcell. DNA read sequences were obtained with the Guppy basecalling software (v6.0.1+652ffd179) and a high accuracy model

(dna_r10.3_450bps_hac.cfg). To produce a circularized and polished genome assembly, those sequences served as input to the CulebrONT pipeline (v2.0.1) (Orjuela et al. 2022) with the FLYE (2.9-b1768) and Miniasm (0.3-r179) assemblers and the Medaka polisher tool. A preliminary inspection of the resulting assemblies indicated that they are structurally highly consistent across the tested assemblers (Figure 1A) and include two circular contigs. We anticipated the presence of extra-chromosomal elements in the CIX2779 genome because for this strain, in contrast to a typical African *Xoo* strain, the profile of the purified genomic DNA in pulsed field gel electrophoresis included an additional high molecular weight band of lower intensity that could correspond to relaxed circular plasmid DNA (Figure 1B). In order to improve base-level accuracy, the FLYE primary assembly sequences were further polished with polypolish (v0.5.0) as described previously (Wick and Holt 2022) with 150bp paired end Illumina NextSeq reads derived from 350bp inserts (library construction and sequencing was performed by FASTERIS, Plan-les-Ouates, Switzerland). This yielded a finished genome consisting in a 4,725,792bp circular chromosome and a 39,798bp long circular plasmid named pCIX2779_1 (Table 1). Interestingly, both ONT and Illumina reads coverages for pCIX2779_1 are about 3-4 fold those of the chromosome, suggesting that there is on average approximately three copies of this plasmid in a CIX779 cell. As a primary sequence quality control measure, gene content was estimated first with the BUSCO software (Manni et al. 2021) which returned 98.52% Complete, 0.7% Fragmented and 0.78% Missing metrics against the single-copy orthologs in the Xanthomonadales Odb10 reference dataset. This is in line with other complete *Xoo* genomes found in sequence databases. This assembly was also annotated with the the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al. 2016). Summary metrics for this annotation are also reported in Table 1 and are comparable to those found on NCBI for other complete African *Xoo* genomes. Disease assays (Figure 1C) indicated that strain CIX2779 causes typical BLB symptoms on the susceptible Kitaake variety but is avirulent on the Carolina Gold variety which contains the *Xo1* resistance gene. *Xo1* confers sublineage wide resistance to African *X. oryzae* (Triplett et al. 2016). Thus, CIX2779 has virulence properties reminiscent of African *Xoo* strains. To ascertain that the nature of the finished genome sequence was coherent with these observations, it was used to interrogate the LINbase (Tian et al. 2020) server for digital taxonomy assignment. The closest database match was the genome of the African *Xoo* strain AXO1947 (Huguet-Tapia et al. 2016) with a FastANI value of 99.98%. In addition, apart from pCIX2779_1, the CIX2279 genome displays high similarity and colinearity with the BAI3 African *Xoo* strain genome (Tran et al. 2018) (Figure 1D). Because TALEs are key virulence determinants of *Xoo* for rice colonization, we investigated the TALE coding potential in the CIX2779 genome. As summarized in Figure 1E, CIX2779 predicted TALEs are chromosomally encoded and highly similar to those of the BAI3 strain. Their sequences of Repeat Variable Diresidues (RVDs) are identical to their counterpart in BAI3 with the exception of the TalH variant whose sequence of RVD is identical to TalH from the MAI1 reference strain (Tran et al. 2018).

Next, because pCIX2779_1 would be the first example of a plasmid in African *Xo* genomes, we investigated in greater details whether its sequence exhibits canonical features of this type of genetic elements. To identify similarities with reference plasmids, we used the mash dist strategy of PLSDB (Galata et al. 2019) with default parameters. The pCIX2779_1 sequence elicited hits with 48 plasmids (Figure 1F), all of them originating from strains of the Xanthomonas genus. The best two hits corresponded to plasmids hosted by a *X. euvesicatoria* pv. *alfalfae* and a *X. vasicola* pv. *vasculorum* strain, respectively, both originating from Africa. This list also included three indigenous plasmids from Asian *X. oryzae* hosts (genbank accessions: NZ_CP019227.1, NZ_CP011963.1 and NZ_KR071788.1). To gain a better sense of the classification of pCIX2779_1, we ran the typer tool of the MOB-suite (Robertson and Nash 2018). In this analysis, pCIX2779_1 was predicted to be a conjugative plasmid because it harbors a relaxase of the MOB_F type and a mate-pair formation marker of type MPF_T (epitomized by the *vir* system of Agrobacterium). However, no origin of transfer nor replication type could be ascribed, potentially because of the inadequate coverage of the underlying databases (Robertson and Nash 2018). The detection of a mating pair formation complex which is a form of type 4 secretion system (T4SS) in pCIX2779_1 prompted us to examine secretion system coding potential more broadly. For this, the NCBI annotation predicted proteins served as input of the Macromolecular System Finder (MacSyFinder) tool together with the TXSScan models bundle (Abby et al. 2016). This analysis did not predict other secretion systems than a T4SS. The CONJscan module (Guglielmini et al. 2014) of MacSyFinder detected 12 components (all mandatory and accessory genes) for the T4SS_typeT model (VirB-like systems) which, in agreement with the typing of the MOB-suite, evaluated as the best solution. The corresponding hits were highlighted in the pCIX2779_1 map of Figure 1F. All together, these results support the hypothesis that pCIX2779_1 encodes a functional conjugation system. However, the majority of the description field for other annotated coding sequences reports on ‘hypothetical’ proteins

and beside additional elements related to conjugation or DNA metabolism, does not inform on potential other biological functions encoded on this plasmid (Figure 1F).

To date about 30 long read sequencing-assembled genomes of African *Xoo* strains are publicly available. Yet, the pCIX2779_1 sequence assembled here with Nanopore reads corresponds to the first indigenous plasmid discovered in a genome with a chromosomal genetic content typical of African strains. This sequence resource for *Xoo* CIX2779 will serve as a foundation for future questions about the prevalence, diversity and evolution of plasmids in African *Xoo* host strains, and overall expands the genomic data on the species for the design of new disease management strategies.

Data Availability

All sequence resources have been submitted to NCBI under the BioProject accession PRJNA861160 which links out to the assembly (RefSeq assembly accession: GCF_024981335.1) and primary ONT (SRA accession: SRX17276343) and Illumina (SRA accession: SRX17276344) reads sequences. The CIX2779 strain is available upon request from the corresponding author.

Funding

We gratefully acknowledge funding from the Bill and Melinda Gates Foundation to Heinrich Heine University, Düsseldorf, with sub-award to Institut de Recherche pour le Développement.

Acknowledgment

The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing High Performance Computing resources.

References

Abby, S. S., Cury, J., Guglielmini, J., Néron, B., Touchon, M., and Rocha, E. P. C. 2016. Identification of protein secretion systems in bacterial genomes. Sci Rep. 6:23080

Amuthan, G., and Mahadevan, A. 1994. Plasmid and pathogenicity in Xanthomonas oryzae pathovar oryzae, the bacterial blight pathogen of Oryza sativa. Journal of Applied Bacteriology. 76:529–538

Carpenter, S. C. D., Mishra, P., Ghoshal, C., Dash, P. K., Wang, L., Midha, S., Laha, G. S., Lore, J. S., Kositratana, W., Singh, N. K., Singh, K., Patil, P. B., Oliva, R., Patarapuwadol, S., Bogdanove, A. J., and Rai, R. 2020. An xa5 Resistance Gene-Breaking Indian Strain of the Rice Bacterial Blight Pathogen Xanthomonas oryzae pv. oryzae Is Nearly Identical to a Thai Strain. Front Microbiol. 11:579504

Doucouré, H., Pérez-Quintero, A. L., Reshetnyak, G., Tekete, C., Auguy, F., Thomas, E., Koebnik, R., Szurek, B., Koita, O., Verdier, V., and Cunnac, S. 2018. Functional and Genome Sequence-Driven Characterization of tal Effector Gene Repertoires Reveals Novel Variants With Altered Specificities in Closely Related Malian Xanthomonas oryzae pv. oryzae Strains. Front Microbiol. 9:1657

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. 2005. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 3:722–732

Galata, V., Fehlmann, T., Backes, C., and Keller, A. 2019. PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res. 47:D195–D202

Gonzalez, C., Szurek, B., Manceau, C., Mathieu, T., Séré, Y., and Verdier, V. 2007. Molecular and pathotypic characterization of new Xanthomonas oryzae strains from West Africa. Mol Plant Microbe Interact. 20:534–546

Guglielmini, J., Néron, B., Abby, S. S., Garcillán-Barcia, M. P., de la Cruz, F., and Rocha, E. P. C. 2014. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. Nucleic Acids Res. 42:5715–5727

Huguet-Tapia, J. C., Peng, Z., Yang, B., Yin, Z., Liu, S., and White, F. F. 2016. Complete Genome Sequence of the African Strain AXO1947 of Xanthomonas oryzae pv. oryzae. Genome announcements. 4:e01730-15

Kaur, A., Bansal, K., Kumar, S., Sonti, R. V., and Patil, P. B. 2019. Complete genome dynamics of a dominant-lineage strain of Xanthomonas oryzae pv. oryzae harbouring a novel plasmid encoding a type IV secretion system. Access Microbiol. 1:e000063

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol Biol Evol. 38:4647–4654

Niu, X.-N., Wei, Z.-Q., Zou, H.-F., Xie, G.-G., Wu, F., Li, K.-J., Jiang, W., Tang, J.-L., and He, Y.-Q. 2015. Complete sequence and detailed analysis of the first indigenous plasmid from Xanthomonas oryzae pv. oryzicola. BMC Microbiol. 15:233

Orjuela, J., Comte, A., Ravel, S., Charriat, F., Vi, T., Sabot, F., and Cunnac, S. 2022. CulebrONT: a streamlined long reads multi-assembler pipeline for prokaryotic and eukaryotic genomes. Peer Community Journal. 2

Robertson, J., and Nash, J. H. E. 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 4

Sundin, G. W. 2007. Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. Annu Rev Phytopathol. 45:129–151

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., and Ostell, J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44:6614–6624

Tian, L., Huang, C., Mazloom, R., Heath, L. S., and Vinatzer, B. A. 2020. LINbase: a web server for genome-based identification of prokaryotes as members of crowdsourced taxa. Nucleic Acids Res. 48:W529–W537

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

Verdier, V., Vera Cruz, C., and Leach, J. E. 2012. Controlling rice bacterial blight in Africa: Needs and prospects. Journal of Biotechnology. 159:320–328

Wick, R. R., and Holt, K. E. 2022. Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLoS Comput Biol. 18:e1009802

Table 1. Main features of the CIX2779 strain genome sequences assembly and annotation.

	Chromosome	pCIX2779_1
Genbank accessions	CP101719	CP101720
Size (bp)	4,725,792	39,798
ONT coverage	90.7x	267.9x
Illumina coverage	103.3x	381.6x
Count of polypolish corrected positions	398	8
G+C content (%)	63.9	61.0
RefSeq assembly accession	GCF_024981335.1	
Predicted genes	4,398	
Predicted CDS	4,177	
Predicted RNA genes	221	
Predicted pseudogenes	540	

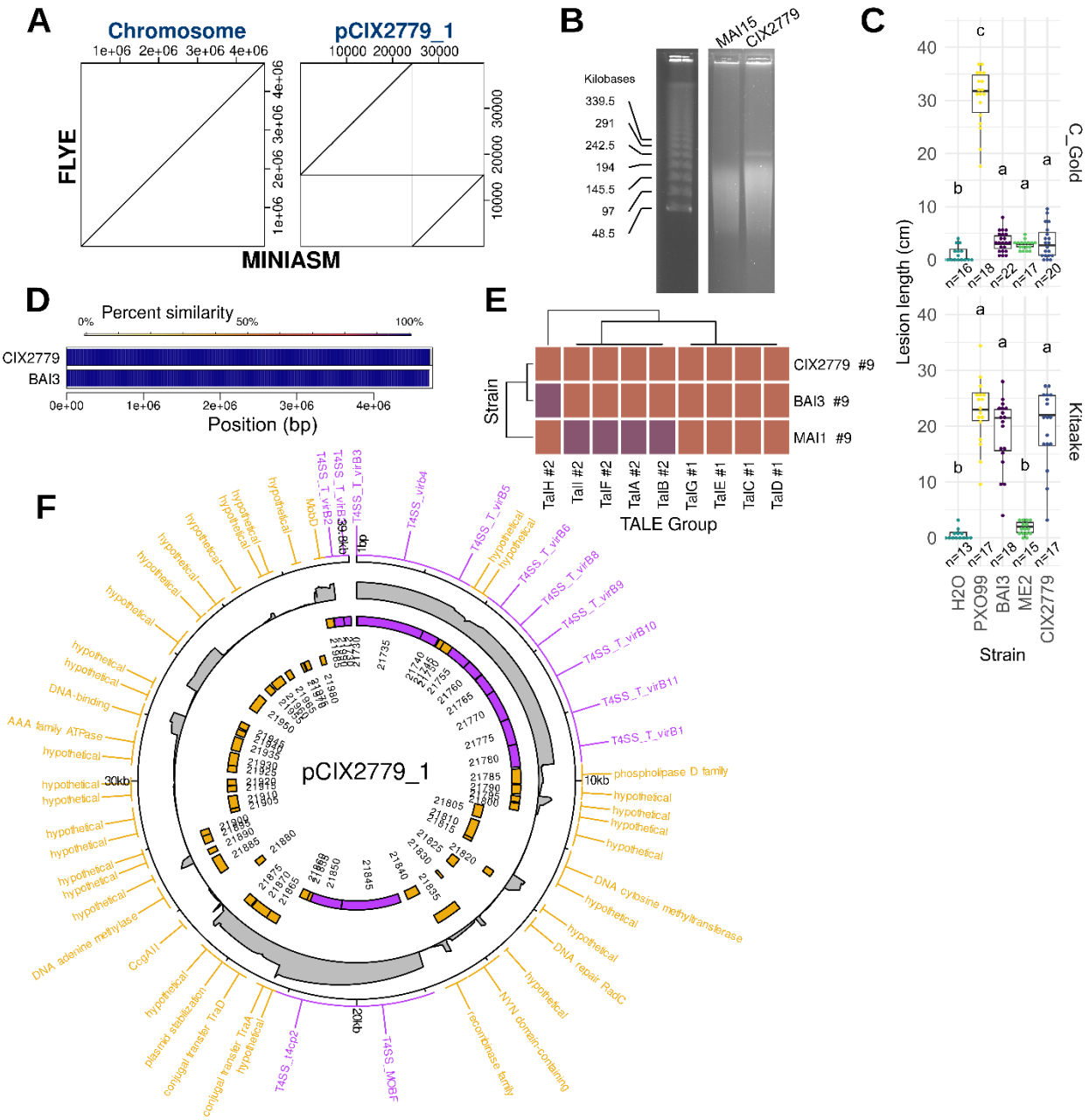

A
Chromosome
pCIX2779_1
FLYE
MINIASM
B
MAI15
CIX2779
Kilobases
339.5
291
242.5
194
145.5
97
48.5
C
Lesion length (cm)
C_Gold
Kitaake
Strain
H2O
PXO99
BAI3
ME2
CIX2779
D
Percent similarity
CIX2779
BAI3
Position (bp)
E
Strain
CIX2779 #9
BAI3 #9
MAI1 #9
TALE Group
F
pCIX2779_1
T4SS_T_virB1
T4SS_T_virB2
T4SS_T_virB3
T4SS_virB4
T4SS_T_virB5
T4SS_T_virB6
T4SS_T_virB8
T4SS_T_virB9
T4SS_T_virB10
T4SS_T_virB11
T4SS_MOBF
T4SS_f4cp2
MobD
CcgA1
DNA adenine methylase
DNA-binding
AAA family ATPase
phospholipase D family
DNA cytosine methyltransferase
DNA repair RadC
NYN domain-containing
recombinase family
plasmid stabilization
conjugal transfer TraD
conjugal transfer TraA

Figure 1. The African *Xoo* CIX2779 strain harbors a ~40kb conjugative plasmid.
A. Dotplot representations of alignments between the Flye and Miniasm versions of the CIX2779 chromosome or pCIX2779_1. Alignments were obtained with the FindSynteny and AlignSynteny functions from the DECIPHER R package.
B. Image of a stained agarose gel after pulsed field electrophoresis of CIX2779 or MAI15 (as an African *Xoo* reference) strains purified genomic DNA. Both the ladder and the samples panels derive from the same gel image (unaltered after acquisition) that was cropped to focus on the relevant areas. Samples were run for 21 hours at 4,5 $V.cm^{-2}$ with an angle of 120°, in a TBE 0,5 X, 1% agarose gel.
C. Lesion length data measured two weeks after inoculation from two replicate leaf-clipping experiments on the Kitaake or the Carolina Gold (C_Gold) varieties with water (H20), the CIX2779 strain or control strains: PXO99, the $PXO99^{A}$ Asian *Xoo* reference strain; BAI3, the African *Xoo* reference strain; ME2, a non-virulent *pthXo1 tal* gene inactivated mutant strain of PXO99A. Strains causing statistically different lesion lengths values on a host genotype, based on a Dunn's test for pairwise multiple comparisons of the ranked data (p values < 0.05), were labeled with different letters.
D. Sequence alignment of the CIX2779 and BAI3 genomes performed as in A.
E. Putative TALE proteins were predicted and processed as before (Doucouré et al. 2018) to produce this summary of TALE diversity in the CIX2779 and reference African strains genomes. Columns correspond to groups of evolutionary related sequences. Cells in the matrix with identical fill color across a column indicate that the corresponding TALEs have identical RVD sequences. Values after the hash sign indicate the count of TALEs in a strain (rows) or number of distinct variants in a group (columns).
F. Circular representation of the pCIX2779_1 plasmid with NCBI gene descriptions (golden color) or CONJscan model hits (violet color). The external line track (grey filled) represents the coverage (0 to 94) of blastn hits with the set of PLSDB related plasmid sequences. The colored rectangles on the intermediate and inner circles with NCBI locus number labels delineate predicted coding sequences spans in the forward or reverse orientation, respectively.

CHAPITRE II : Edition du génome d'une variété élite de riz Africain pour promouvoir la résistance aux souches endémiques et émergentes de *X. oryzae* pv. *oryzae*

Chapitre basé sur les résultats d'un troisième manuscrit en tant que premier auteur, soumis au journal ***Nature Communications*** le 20 octobre 2022.

Titre du manuscrit : **« Genome editing of an African rice elite variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains »**

1 Contexte

L'analyse de la diversité génétique des populations de *Xoo* Africaines que nous avons effectuée et présentée dans le Chapitre I a mis en évidence le clivage existant entre les populations d'Afrique de l'Est et celles d'Afrique de l'Ouest. Grâce à nos données de typage moléculaire et d'analyses de génomes, nous avons découvert que les récentes épidémies survenues en Tanzanie et à Madagascar sont dues à l'arrivée sur ces territoires de souches de *Xoo* issues d'Asie. Cette invasion soulève la question de la capacité des variétés de riz utilisées sur le sol Africain à contrôler ces nouvelles arrivantes. Afin d'évaluer au mieux la menace que représente ces nouvelles populations pour la riziculture en Afrique, nous avons cherché à déterminer les résistances à même de maîtriser ces envahisseurs issus d'Asie, ainsi qu'à caractériser les gènes de sensibilité majeurs ciblés au cours de l'infection. Enfin, nous avons voulu mettre au point une variété résistante à la fois aux populations endémiques et émergentes de *Xoo* en Afrique.

2 Stratégies

Afin d'évaluer quelles résistances permettaient de contrôler efficacement les populations de *Xoo* Africaines dans leur ensemble, nous avons phénotypé un panel diversifié de souches issues d'Afrique de l'Est et d'Afrique de l'Ouest sur des lignées NILS contenant différents gènes de résistance contre la BLB. Nous avons également utilisé les lignées KOs *sweet* conçues au cours du projet (Eom *et al*., 2019 [288]) en complément avec des approches de PCR quantitative et d'analyses de TALomes afin de déterminer quel(s) gène(s) *OsSWEET* du clade III et via quel effecteur(s) ces nouvelles populations promulguaient la maladie. A l'issu de cette caractérisation approfondie, nous avons développé un système hybride CRISPR/Cas9-Cpf1 pour modifier les six EBEs connues dans les promoteurs des gènes *OsSWEET* du clade III de la variété élite africaine Komboka, afin de créer une variété éditée possédant un large spectre de résistance aux souches de *Xoo* auxquelles elle pourrait être confrontée dans le cadre d'une utilisation en riziculture en Afrique.

3 Résultats

Nos campagnes de surveillance de la BLB en Afrique nous ont permis de repérer des **épidémies inédites en Tanzanie**, survenues dès 2019.

Le phénotypage de ces souches collectées en parallèle d'un panel d'isolats de *Xoo* originaires de différents pays d'Afrique de l'Ouest et de l'Est sur la **variété semi-aromatique à haut rendement Komboka**, a révélé que ce génotype était **résistant à la totalité des souches d'Afrique de l'Ouest, mais sensible aux isolats issus de Tanzanie**, ainsi qu'à notre souche contrôle issue d'Asie (PXO86).

Une recherche de résistance dans le fond génétique Komboka a révélé la présence des gènes *Xa1* et *Xa4*. En accord avec ces données, le phénotypage du même panel Africain sur des lignées isogéniques

de riz pour différents gènes de résistance à la BLB montre bien un **contrôle des *Xoo* d'Afrique de l'Ouest par IRBB1 (*Xa1*) et IRBB4 (*Xa4*)**. En revanche, aucun des gènes de résistance testés ici ne s'est montré efficace contre les souches Tanzaniennes.

Une phylogénie SNP basée sur l'analyse des génomes des souches collectées en Tanzanie révèle que **les souches CIX4462, CIX4506 et CIX4509 (collectées en 2019 et 2021)** groupent avec les souches Asiatiques de *Xoo*, et **sont très proche d'isolats provenant de la province du Yunnan en Chine**. En revanche les souches collectées en Tanzanie avant 2019 (Dak16, T19 et *Xoo*3-1) groupent bien avec les souches d'Afrique de l'Ouest comme précédemment décrit.

L'analyse des TALomes des **souches Tanzaniennes collectées après 2019** montre qu'elles **possèdent un variant du TALE PthXo1**, que nous avons nommé PthXo1B, avec deux RVDs supplémentaires à l'extrémité N-terminale et capable de se fixer sur un EBE chevauchant l'EBE ciblé par PthXo1 dans le promoteur de *OsSWEET11a*. Nous avons également découvert qu'à l'instar de nombreuses souches Asiatiques, CIX4462, CIX4506 et CIX4509 portent des TALEs tronqués (appelés **iTALes**), connus pour supprimer la résistance médiée par *Xa1*.

L'inoculation des mutants sur les lignées du kit *sweet* KO confirme bien ces données, le **KO *sweet11a* étant le seul à présenter de la résistance aux nouvelles souches Tanzaniennes**. Nos analyses de qPCR montrent bien une **induction de *OsSWEET11a*** par ces souches.

Enfin ; **l'édition des 6 EBEs connues dans les promoteurs de *SWEET11*, *SWEET13* et *SWEET14*** à l'aide d'un système d'édition hybride Cas9/Cpf1dans la variété Komboka, **a conféré avec succès de la résistance aux souches endémiques et émergentes de *Xoo*, ainsi qu'aux souches Asiatiques**.

4 Importance pour le domaine

Cette étude met en avant **l'importance d'établir des suivis épidémiologiques** réguliers, seuls capables d'alerter précocement de l'arrivée de populations pathogènes invasives.

Cette découverte montre par ailleurs que **des migrations de *Xoo* sont possibles sur de très longues distances, vraisemblablement par le biais de matériel contaminé**, mettant en avant la nécessité d'un contrôle plus rigoureux du matériel biologique lors des importations et exportations. Une nouvelle fois, cette introduction depuis la Chine **soulève la question de la transmission du pathogène par les graines**, restant jusqu'alors non-prouvée expérimentalement.

Nous montrons que **ces mouvements de populations** peuvent constituer une vraie menace pour les cultures, en révélant la présence de iTALEs chez ces nouvelles populations Tanzaniennes, jusqu'alors seulement décrits chez les *Xoo* Asiatiques, **mettant en péril les résistances utilisées au champ par les fermiers Africain** (*Xa1*).

Nous mettons en avant l'**efficacité des outils créés dans le cadre du projet pour une meilleure réactivité face à l'arrivée de nouvelles souches de *Xoo***, en déterminant avec succès le gène de sensibilité majeur utilisé par nos isolats Tanzaniens (*SWEET11a*) grâce au kit KO *sweet* produit par Eom *et al*. en 2019 [288].

Enfin, **la combinaison des systèmes CRISPR/Cas9 et Cpf1 augmente la flexibilité** du choix des ARNs guides appropriés en fonction de la richesse en G et en T dans la région à éditer, conditionnant le PAM disponible. Ce système permet également de moduler la taille des éditions (SNP pour la Cas9 et délétions courtes pour la Cpf1).

5 Implication personnelle

Parmi les données présentées dans cette étude j'ai réalisé les phénotypages des *Xoo* de Tanzanie sur les lignées éditées Komboka (anciennes et nouvelles éditions), ainsi que les tests de virulence sur les lignées du kit *sweet* KO.

J'ai réalisé les PCR quantitatives pour le suivi de l'expression des différents gènes *SWEETs* du clade III en présence de *Xoo* Tanzaniennes, sur Komboka sauvage et lignées éditées.

J'ai également préparé les librairies et réalisé le séquençage ONT des souches issues de Tanzanie ayant été utilisées pour les phylogénies SNPs et les analyses de TALomes (également présenté dans le Chapitre II).

J'ai effectué les analyses statistiques et créé les représentations graphiques (boxplots) de toutes les données quantitatives présentées dans l'article, les données étendues et les données supplémentaires (longueurs de lésions et niveaux d'ARN messager).

Enfin, j'ai préparé les figures 1, 3, 4, 5, Extended data 1, Extended data 3, Extended data 7 ainsi que les tableaux 1 et S1 et participé à la rédaction du matériel et méthodes.

6 Article 3 : « Genome editing of an African rice elite variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains »

Genome editing of an African elite rice variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains

Van Schepler-Luu[1,7,a], Coline Sciallano[3,a], Melissa Stiebner[1,a], Chonghui Ji[2,a], Gabriel Boulard[3], Amadou Diallo[3], Florence Auguy[3], Si Nian Char[2], Yugander Arra[1], Kyrylo Schenstnyi[1], Marcel Buchholzer[1], Eliza P.I. Loo[1], Atugonza L. Bilaro[4], David Lihepanyama[4], Mohammed Mkuya[5], Rosemary Murori[6], Ricardo Oliva[7,8], Sebastien Cunnac[3], Bing Yang[2,9], Boris Szurek[3,*] & Wolf B. Frommer[1,10,*]

[1] Institute for Molecular Physiology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2] Division of Plant Science and Technology, Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA

[3] Plant Health Institute of Montpellier (PHIM), Université Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France

[4] Tanzania Agricultural Research Institute (TARI)-Uyole Centre, PO Box 400 Mbeya, Tanzania

[5] International Rice Research Institute, Eastern and Southern Africa Region (IRRI ESA), BOX 30709 00100, Nairobi, Kenya

[6] International Rice Research Institute (IRRI), Africa Regional Office, Nairobi, Kenya

[7] International Rice Research Institute, Pili Drive, Los Baños, Laguna 4031, Philippines

[8] new address: World Vegetable Center, Shanhua, Tainan, Taiwan 74151

[9] Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

[10] Institute for Transformative Biomolecules, ITbM, Nagoya University, Nagoya, Japan

[a] Co-first authors, equal contribution

[*] Corresponding authors

Abstract

Bacterial leaf blight (BB) of rice, caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), threatens global food security and the livelihood of small-scale rice producers. Analyses of *Xoo* collections from Asia, Africa and the Americas demonstrated surprising continental segregation, despite robust global rice trade. Here, we report unprecedented BB outbreaks in Tanzania. The causative strains, unlike endemic *Xoo*, carry Asian-type TAL effectors targeting the sucrose transporter *SWEET11a* and suppressing *Xa1*. Phylogenomics clustered these strains with *Xoo* strains from China. African rice varieties do not carry suitable resistance genes. To protect African rice production against this emerging threat, we developed a hybrid CRISPR-Cas9/Cpf1 system to edit six TALe-binding elements in three *SWEET* promoters of the East African elite variety Komboka. The edited lines show broad-spectrum resistance against Asian and African strains of *Xoo*, including strains recently discovered in Tanzania. This strategy could help to protect global rice crops from BB pandemics.

Introduction

Rice is one of the most important staple foods for developing countries in Asia and Africa[1]. African consumers increasingly replace the traditional staples of sorghum, millet and maize with rice. Today, African farmers, 90% of which are small-scale food producers with <1 ha of land[2], produce ~60% of the local rice demand, but this demand will increase with the expected doubling of the population until 2050. Productivity is often hampered by the diseases Bacterial Leaf Blight (BB), Rice Yellow Mottle Virus (RYMV), and Rice Blast (RB)[3–5]. Breeding high-yielding varieties that are resistant to these diseases will be an important factor supporting food security in Africa. BB, caused by the bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is a devastating rice disease in many rice-growing countries. Resistance (*R*) genes for BB have been identified and used extensively to develop resistant varieties, however resistance based on single *R* genes may be rapidly overcome by new strains[6]. Regular monitoring of the virulence of current *Xoo* strains on a collection of rice tester lines carrying single *R* genes has provided guidance for stacking different *R* genes to obtain broad-spectrum resistance in Asian rice varieties[7,8].

In Africa, BB was first reported in Mali in 1979 and later in several other West and Central African countries[9,10]. Recently, East African countries like Tanzania and Uganda reported BB epidemics[11]. While the information on the spread and severity of BB in Africa is scarce, BB is not yet considered a major threat in Africa, but has established. However, climate change affects disease spread, leading to projections that the damage will become more severe in the future[10,12,13]. Other factors, such as population growth and increasing yield losses due to pathogens, as well as the distinct *Xoo* populations in Africa, make it essential to generate broad-spectrum and durable resistance in local varieties specifically for Africa.

Key aspects of the mechanisms underlying BB disease have been elucidated and provide an efficient roadmap to breeding resistance. *Xoo* secretes a series of Transcription Activator-Like effectors (TALes) into rice xylem parenchyma cells via a type III secretion system. Once inside the host cell, the TALes translocate to the nucleus and bind to effector binding elements (EBEs) in the promoters of specific host genes via a unique domain of tandemly arranged 34 amino-acid long repeats, triggering the ectopic induction of the target genes[14]. *Xoo* TALes induce one or several host SWEET sucrose uniporter-genes (*SWEET11a*, *13* and *14*), presumably causing sucrose release into the apoplasm where bacteria reside. Note that due to the discovery of a previously misannotated sucrose transporting paralog, SWEET11a was renamed (formerly SWEET11)[15]. Sucrose is consumed by the bacteria, resulting in effective reproduction[16]. Allelic EBE variants that are not recognized by TALes function as recessive gene-for-gene resistance genes. To date, six EBEs in *SWEET* promoters are known for the following TALes: PthXo1 in *SWEET11a*, PthXo2 variants in *SWEET13*, and TalC, PthXo3, AvrXa7 and TalF in *SWEET14* [17,18]. Notably, Asian and African *Xoo* strains are phylogenetically distinct and use different TALes to target *SWEETs* [18,19]. African *Xoo* strains exclusively use TalC and TalF, which both target *SWEET14*,

while Asian strains encode PthXo1, PthXo2 (variants A, B, C), PthXo3 and AvrXa7, which target *SWEET11a*, *13* and *14*. American *Xoo* strains lack TALes and hence are poorly virulent[20]. Based on this information, broad-spectrum resistance was introduced into the *Oryza sativa* ssp. *japonica* variety Kitaake and the spp. *indica* varieties IR64 and Ciherang-Sub1 by genome editing of the EBEs in all three *SWEET* genes[17,18]. The edited lines may prove to be valuable breeding materials, giving rise to lines that will benefit small-scale producers in Asia. To prepare for emerging strains with novel virulence mechanisms, the diagnostic SWEET^{R} kit was developed, enabling analyses of BB in the field and identification of suitable resistance strategies, such as which *R*-gene combinations to deploy in the next season[17].

Recent disease surveys in Tanzania, described here, have identified a BB outbreak in 2019 that has spread in subsequent years. The causative strains have features that distinguish them from other African *Xoo* strains and have virulence properties that make them a major threat to African rice production. Recently, Kenya and Nigeria have exempted SDN-1 genome-edited crops generated using site-directed nuclease without a transgene (SDN-1) from GMO regulations[21]. This exemption provides a regulatory framework for the introduction of genome-edited, BB-resistant rice into Africa. To provide small-scale rice producers with BB-resistant rice cultivars, we used an efficient transformation protocol[22] to introduce BB resistance into a rice cultivar popular in East African countries. Komboka (IRO5N-221) is a high-yielding, semi-aromatic rice variety jointly developed by IRRI (International Rice Research Institute and KALRO (Kenya Agricultural & Livestock Research Organization; BBSRC Varieties Kenya.pdf) that grows to an average height of 110 cm and has a yield potential of ~7 tons/ha, almost twice that of Basmati, a popular variety planted in Kenya and Rwanda (The Star 2020). Komboka plants mature in ~115 days, respond well to fertilization, and are suitable for irrigated lowland cultivation in Africa[23]. Komboka had been classified as moderately resistant to BB, RYMV, and RB, however, detailed resistance profiles are not available. In this work, Komboka BB-resistance genes were analyzed, and a combination of Cas9, Cpf1 and multiplexed gRNAs was used to edit all known EBEs in the *SWEET* genes, resulting in broad-spectrum resistance to not only previously known African and Asian strains, but also the newly introduced *Xoo* strains from the outbreak in Tanzania.

Results

Identification of an unprecedented disease outbreak in Tanzania

There is little information on the population genetics and dynamics of *Xoo* strains circulating in Africa. While pathogenic strains have sporadically been isolated from rice plants and characterized, there is no systematic analysis of BB across Africa. Moreover, there are many diverse rice varieties used in Africa, with no general survey of acreage growing different varieties. These factors combine to make it challenging to breed durably resistant lines. In addition, numerous international partnerships have introduced varieties and cultivation techniques to major rice production areas in different countries of

Africa. For instance, in 1975, one of the first irrigated perimeters with rice research labs and a training farm were developed in Dakawa, a major rice production zone in Tanzania, as a result of a multi-year partnership with North Korea[24]. National (NAFCO) and international institutions (USAID, Cornell University) remain active in Dakawa to build irrigation schemes and implement *The System of Rice Intensification* (SRI; http://sri.ciifad.cornell.edu/countries/tanzania/index.html#summary). In 2011, a collaboration with China resulted in 50 ha of irrigated fields managed by local farmers. There is also a Technology Demonstration Center that trains local farmers and technicians and grows examples of Chinese improved varieties, including hybrid rice[25].

The first observations of BB in Africa were made in Mali in 1979 [9]. Since then, the disease has been reported mostly in West Africa, including Senegal, Benin, Burkina Faso, Ivory Coast, Mali and Niger[26–29]. More recently, BB was also reported in the East African countries Uganda and Tanzania[11,18]. To systematically analyze these newly isolated strains and to compare them to the broader African *Xoo* landscape, 833 strains from rice fields sampled in nine African countries between 2003 and 2021 were collected and characterized using molecular diagnostic as well as pathogenicity assays (Supplementary Table 1)[26–29]. During this study, two unprecedented outbreaks were identified, one in 2019 in Dakawa and another in 2022 in Lukenge (Supplementary Table 2). These sites are located 60 km apart in the Morogoro region of Tanzania, an area that for decades has been a center of partnership initiatives aiming at increasing rice production[25]. We performed disease surveys in this area in 2019 and 2021 and identified two outbreaks on TXD 306 (SARO-5), a rice variety popular in irrigated ecologies in Tanzania. More recently BB was identified on Komboka. Leaves with typical BB symptoms were processed, and seven strains from Dakawa and 106 from Lukenge were isolated and validated as *Xoo* with diagnostic primers[30]. Notably, surveys performed in subsequent years showed increased severity and spread (Supplementary Table 2).

To identify the *R* genes best suited for BB control, 21 strains that represent the diverse geographical origins of our collection were inoculated on six Near Isogenic Lines (NILs) of rice that were harboring either *Xa1*, *Xa3*, *Xa4*, *xa5*, *Xa7* or *Xa23* and that had been reported as efficient *R* genes against some African *Xoo* strains (Table 1, Supplementary Table 1)[29]. The test panel included three strains, namely CIX4457, CIX4458 and CIX4462, that were isolated from the recently detected outbreak in Dakawa, Tanzania. NILs harboring *Xa1*, *Xa23*, *xa5* or *Xa4* were resistant to most African strains, while CIX4457, CIX4458 and CIX4462 were virulent on all NILs tested, highlighting an urgent need for a new solution to protect rice against BB outbreaks.

Resistance of Komboka to endemic African *Xoo* but not to new Tanzanian strains

Komboka, an emerging elite variety released in some countries in East Africa (Tanzania, Kenya, Uganda and Burundi), has been described as moderately resistant to *Xoo*, yet the nature of its resistance has not been elucidated. Using the African *Xoo* diversity panel, we found that most African strains were

avirulent in Komboka, except for the newly isolated strains from Tanzania (Fig. 1). By comparison, all strains were highly virulent on the susceptible variety Azucena (Extended Data Fig. 1, Table 1). Since *Xa1*, *Xa4*, *xa5* and *Xa23* confer resistance against *Xoo* (Table 1), we investigated potential *R* genes in Komboka. Mining of the IRRI QTL database (https://rbi.irri.org/resources-and-tools/qtl-profiles) revealed that Komboka contains genetic markers linked to *Xa4* but not *xa5* or *Xa23*, while there was no information on *Xa1* in the database (Extended Data Fig. 2A). We confirmed the presence of *Xa4* in Komboka by tracing an *Xa4*-associated marker and by sequencing (Extended Data Fig. 2B, C). Consistent with the broad-spectrum resistance of *Xa1* against endemic African *Xoo* strains, we found that Komboka carries a dominant *Xa1* allele that is highly similar to *Xa45*(t) (Extended Data Fig. 5D, E). The combination of *Xa4* and *Xa45(t)* in Komboka can explain the observed resistance to most African *Xoo* strains. Apparently, the two *R* genes do not protect against the strains recently isolated from the Dakawa outbreaks.

Tanzanian *Xoo* strain cluster with Asian *Xoo* via whole genome SNP-based phylogeny

To identify the mechanisms underlying the virulence of the Tanzanian strains, three representative strains: CIX4462 (Tz6) collected from Dakawa in 2019 and CIX4506 (F2SP1-1c) and CIX4509 (F4SP1-1a), collected from Lukenge in 2021, were subjected to whole-genome sequencing and SNP-based phylogenetic analyses (Supplementary Table 3). Notably, CIX4457, CIX4458 and CIX4462 clustered with Asian *Xoo* isolates (Fig. 2A). In contrast, the Tanzanian strains T19, Dak16 and Xoo3-1, which had been collected in Dakawa before the 2019 outbreak, grouped with the endemic African lineage, consistent with previous reports[18]. CIX4462, CIX4506 and CIX4509 are most closely related to strains from Yunnan province, China (Fig. 2B) and carry *iTALe* genes similar to *tal3a* (Fig. 2C). *iTALe* genes encode truncated TALes known to suppress *Xa1* resistance, but had so far only been reported in Asian *Xoo* isolates. In addition, the three strains contain a TALe (named PthXo1B) highly similar to the Asian PthXo1, with the addition of two RVD domains at the N-terminus (Fig. 2D; Supplementary Data 1). EBE prediction indicated that PthXo1B can bind to the *SWEET11a* promoter at a site overlapping with the EBE targeted by PthXo1 (Fig. 2D). Importantly, *sweet11a* knock-out mutants from the diagnostic SWEET^{R} kit were resistant to CIX4457 and CIX4505 from Tanzania, whereas *sweet13* or *sweet14* knock-out mutants remained susceptible (Extended Data Fig. 3). In accordance, *SWEET11a* was induced in Kitaake leaves when inoculated with the Tanzanian strains CIX4457 and CIX4505 (Extended Data Fig. 4). Together, our data showed that representative Tanzanian strains collected in 2019-2021 are phylogenetically related to Asian *Xoo* strains and have both *iTALe* and *PthXo1* homologs unique to Asian *Xoo* isolates[18].

Editing *SWEET* promoter sequences in Komboka to obtain resistance to all African strains

As a prerequisite for editing *SWEET* promoters in Komboka, the promoter regions of *SWEET11a*, *13* and *14* were sequenced and shown to contain EBEs for PthXo1, PthXo2A, PthXo3, AvrXa7, TalC and

TalF (Extended Data Fig. 5, Supplementary Table 4, Supplementary Data 2). To protect Komboka against endemic strains from Africa as well as introduced Asian strains, all six known EBEs in the promoters of *SWEET11a*, *13* and *14* were edited using a hybrid Cas9/Cpf1 editing system. Due to blunt cleavage, Cas9 preferentially produces SNPs, while Cpf1 (Cas12) produces staggered cuts, creating predominantly small deletions. We hypothesized that small deletions would produce more robust resistance. Due to the PAM requirement, it was not possible to design sgRNAs for Cpf1 at all EBEs; therefore, Cpf1 was combined with Cas9. Two CRISPR/Cpf1 gRNAs (cXo1 and cXo2) were designed to target PthXo1 and PthXo2A EBEs within the promoters of *SWEET11a* and *SWEET13*, respectively (Extended Data Fig. 5, 6, 7). Since the EBEs for AvrXa7, PthXo3 and TalF in the promoter of *SWEET14* are overlapping, one CRISPR-Cpf1 gRNA (cTalF) was designed to target the overlapping region of all three EBEs. Because no TTTV PAM sequence was available near the EBE for TalC, a Cas9 gRNA (gTalC) was designed to target TalC EBE. From a first round of transformation, six representative T2 lines were tested for resistance using six representative *Xoo* strains: ME2 ($PXO99^A$ mutant deficient in PthXo1), $PXO99^A$ (containing PthXo1), PXO61 (PthXo2B, PthXo3), PXO86 (AvrXa7), MAI1 (TalC, TalF) and BAI3 (TalC) (Fig. 3A, Supplementary Table 4, 5, 6). All six lines from this first round of transformation were resistant to the tested strains. One line (1.5_19) was fully resistant to all strains, while five lines were fully resistant to five strains but only moderately resistant to PXO86. This difference in resistance is most likely explained by the presence of a 4-bp deletion in the EBE for AvrXa7, while line 1.5_19 carried the same 4-bp deletion plus an additional base pair substitution (G/T) (Fig. 3B). Since the clipping assays used very high bacterial titers, it is generally assumed that moderate resistance will be sufficient for effective resistance in field conditions. However, TALes were reported to have less specific nucleotide binding at the 3'-end, hence the G/T substitution in line 1.5_19 could potentially be overcome[14]. To obtain more robust resistance to a larger population of *Xoo*, a second round of transformation was carried out. Two T2 lines (14_19 and 14_65), which contained deletions in all EBEs, including 11-bp and 5-bp deletions in the AvrXa7 EBEs, respectively, were resistant to all tested strains including PXO86 (Fig. 4). Lines 1.5_19 and 1.2_40 were also resistant to CIX4457 and CIX4505, the new isolates from Tanzania, consistent with the 12- and 9-bp deletions in the predicted PthXo1 EBE, respectively (Fig. 5). As one may have predicted, induction of *SWEET11a* by CIX4457 and CIX4505 was abolished in lines 1.2_40 and 1.5_19 (Extended Data Fig. 4). Taken together, the edited Komboka lines are fully resistant to Asian and African *Xoo* strains, including the emerging population from Dakawa and Lukenge in Tanzania.

Discussion

Here we report an outbreak of BB in East Africa. Through genetic and genomic characterization, this outbreak can be ascribed to a recent colonization event by *Xoo* strains that are most similar to Asian strains, specifically from Yunnan province.

Notably, the Asian and African strains clearly form distinct phylogenetical groups[19,31]. This came as a surprise since monitoring had indicated very low BB incidence, and only three strains were isolated between 2011 and 2018 [18]. As shown previously, those strains cluster with other African *Xoo* strains using SNP-based phylogeny analyses and carry TalC as the major virulence TALe [18]. We cannot trace the introduction of the new strains, which may have occurred through either major storms or transfer of contaminated rice. *Xoo* is well-known to be transmitted via tropical storms, and transmission within Asian countries has been implied, for example in the case of PXO99^{A}, which may have been introduced rather recently from Nepal or India to the Philippines. By contrast, rather effective continental isolation had been observed, and if the new Tanzanian strains would have been transmitted by storms from Asia, one may have expected a closer phylogenetic relationship to strains from India or Indonesia based on the geography[18,19,32]. Due to the distinct evolutionary trace and the resulting absence of coevolution, it is likely that many African rice varieties do not contain suitable *R*-genes that could protect against Asian strains. Hallmark features of the new strains identified in Tanzania are iTALes and a PthXo1 homolog; both are absent from endemic African strains. African rice lines carrying *Xa1* may thus be resistant to African strains, but will be susceptible to Asian strains, including those recently introduced to Tanzania. Surveys over several years with preliminary data from 2022 indicate that the outbreak in Dakawa is gaining momentum regarding severity and spread, thus potentially becoming a threat to, at the least, East Africa but possibly to all of Africa by the increasing occurrence of heavy storms.

Our study, which so far is based on only a few strains representative of both outbreaks in Tanzania, shows that they are very close genetically, arguing in favor of a single introduction event. However, further analysis of a greater number of strains will be required to validate their origin and relationships. To gain insight into the spread of these emerging *Xoo* populations, pathogen monitoring was further conducted in Morogoro and surrounding regions in July 2022. Strikingly, over 600 leaf samples harboring typical BB symptoms were collected from 37 fields in six different regions in Tanzania. MLVA-typing and/or whole genome SNP analysis of the strains associated with these lesions will reveal the genetic structure of the *Xoo* population in Eastern Tanzania.

We had identified *SWEETs* as essential susceptibility loci for BB [17,33,34]. Virulence appears to be critically dependent on the activation of one of the sucrose-transporting SWEET uniporters in the rice genome[18]. Another relevant difference between Asian and African strains is the fully distinct set of TALes used to target SWEETs [18]. It is thus likely that, due to the absence of coevolution, African rice varieties contain EBEs in *SWEET* promoters that can be targeted by Asian TALes.

Analysis of Komboka, a high yielding semi-aromatic rice line released in Kenya and Tanzania, shows that it is resistant to endemic African strains due to *Xa1* and *Xa4*. But Komboka is susceptible to Asian strains like PXO99^{A}, which carries both iTALes and PthXo1 that suppress Xa1 and induce *SWEET11a*, respectively. Indeed, we found Komboka samples with BB symptoms in another district in Tanzania (Mikindani, Mtwara, ~400 km from Dakawa) in 2022. Characterization of the collected leaves is

ongoing. *Xa1*-mediated resistance has been broken by 95% of the Asian *Xoo* strains[35], while *Xa4*-mediated resistance was also overcome by the new East African *Xoo* strains as well as many Asian strains reported previously[32]. Thus, *Xa1* and *Xa4* lack durability due to shifts in the *Xoo* population or the emergence or introduction of new strains. Thus new lines with broad-spectrum resistance against both Asian and African strains, including the strains found in Tanzania, is required[6,36].

We hypothesized that editing all known EBEs for both African and Asian SWEET-targeting TALes could be used to engineer full and robust resistance in an emerging elite variety for Africa. Since this strategy requires simultaneous editing of six EBEs, optimally by small deletions inside the EBE to enhance robustness, we developed a hybrid CRISPR-Cas9/Cpf1 system that targets all six EBEs in Komboka. Both CRISPR-Cas9 and CRISPR-Cpf1 systems have been widely used in genome editing. SpCas9 is the first Cas version to offer high editing efficiency across a wide range of plant species. LbCpf1 or LbCas12 is another endonuclease which is smaller than SpCas9 and requires shorter CRISPR RNA (crRNA)[37]. SpCas9 and LbCpf1 both generate double-stranded breaks, but Cas9 introduces blunted cuts 3 base pairs upstream of the PAM, while Cpf1 introduces 5-bp staggered cuts downstream of the PAM [38]. Hence, Cpf1 often produces larger mutations compared to Cas9, which mainly introduces single base pair insertions or deletions (indels). Depending on the purpose of genome editing, large mutations or small mutations are preferred. To generate knock-out mutants, single base pair indels in the first exon that result in frameshifts or premature stop codons are sufficient. To prevent the binding of TALes to their respective EBEs on *SWEET* promoters, single base pair changes are not always effective due to the degenerate code and the ability to loop out central repeats of the TALes [18]. In addition, *Xoo* evolves rapidly and, therefore, single base pair changes could be overcome quickly, as demonstrated for an *xa25*-like resistance gene[39]. Hence, to prevent rapid break of the resistance by simple modifications of the TAL effectors, a larger mutation would be preferred. Since it is still unknown how the native *SWEET* expression is affected by promoter mutations, the number of base pair indels needs to be evaluated carefully. In addition, prevention of TALe binding to EBEs depends not only on the number of deleted or inserted base pairs but also on the position of the introduced mutations within the EBE. Especially, the RVDs at the 3´-end of TALes, such as N* and NS, have less specificity in nucleotide binding, therefore, mismatches at the 3´-end of an EBE can be tolerated by the RVDs [14]. Here, the 1-bp deletion at the 5´-end of the TalC binding site is sufficient for resistance, while a 4-bp deletion at the 3´-end of the AvrXa7 binding site did not prevent binding of AvrXa7 to its EBE, indicating that mutations at the 5´-end of the TALe binding site are more effective than mutations at the 3´-end in preventing the binding of TALes to EBEs.

Cas9 cleaves target DNA adjacent to a G-rich PAM, while Cpf1 requires a T-rich PAM sequence[38,40]. Therefore, combining Cas9 and Cpf1 into a single system enables simultaneous editing in both G- and T-rich regions. Here, EBE regions used for gRNA design were comparatively short, just 23-29 bp, hence, there were a limited number of suitable gRNA design options (considering also potential

unwanted off-target effects). The BB-resistant IR64 and Ciherang-Sub1 lines developed previously for Asia were generated using four gRNAs for *SWEET11a*, *13* and *14* promoters, but did not target the TalF EBE in *SWEET14* which is targeted by African strains specifically[17,18]. The EBE in the respective promoters for TalF lacked a suitable SpCas9 NGG-PAM recognition sequence. By combining the two enzymes Cpf1 and Cas9 into a hybrid CRISPR-Cas9/Cpf1 system, it became possible to design four gRNAs (three gRNAs for Cpf1, including one for TalF; and one for to target TalC, which lacked a suitable PAM sequence for Cpf1) that target all six known EBEs relevant for BB susceptibility. The combination of CRISPR-Cas9 and Cpf1 gRNAs in a single system increases the flexibility of choosing suitable gRNAs depending on the available PAM recognition sequences and off-target frequency. An efficient transformation protocol previously established for Komboka was used here to edit all known EBEs in three *SWEET* promoters[22]. For four target sites (cXo1, cXo2c, gTalC, cTalF), we generated 15, 27, 16 and 21 variants, respectively using four gRNAs, demonstrating the successful application of the hybrid Cas9/Cpf1 in multiplex genome editing of the rice variety Komboka. Each variant represents a distinct *R* gene. Edited Komboka lines were shown to be resistant against a set of representative Asian and African *Xoo* strains, which harbor all known *SWEET*-inducing TALes.

Crosses and analyses for the reliable elimination of transgenes from the edited Komboka lines is ongoing. Ideally, the deployment of edited lines should be region-specific, based on the variation of the TALes present in the *Xoo* populations. In parallel, monitoring the evolution of *Xoo* in Africa is extremely necessary for early detection of Asian strains or hybrid African-Asian strains. In addition, the diagnostic SWEET^{R} Kit 2.0, upgraded with the newly identified *SWEET11b* susceptibility gene and containing all six translational SWEET-reporter and knockout lines, can effectively be used to detect the targeting EBEs, while TALome analysis should be performed to detect the presence of *SWEET*-inducing TALes in the current African *Xoo* population[15,17]. Kenya evaluates applications for the import of transgene-free lines on a case-by-case basis, thus it will be necessary to prepare suitable documentation for import. In parallel, it is planned to use the hybrid Cas9/Cpf1 strategy to expand the resistant germplasm to other rice varieties grown in African countries to be able to meet consumer-farmer preferences. This project aims to ultimately provide the material to small-scale producers. It will also be necessary to continue monitoring disease and *Xoo* populations. Governmental action is recommended to reduce the risk of introduction of Asian pathogens to Africa and vice versa. In parallel, improved management practices and training may also help farmers to reduce disease incidence[41].

Methods

Rice seeds and *Xoo* strains

Seeds of the rice variety Komboka (IR05N221, L17WS.06#24) were provided by the International Rice Research Institute (IRRI, The Philippines) under a Standard Material Transfer Agreement under the

Multilateral System (SMTA-MLS). *Xoo* strains were obtained from the *Xoo* strain collection of at the French National Research Institute for Sustainable Development (IRD, Montpellier, France).

Rice cultivation

Rice seed germination and plant cultivation at HHU were done as described[22]. Briefly, rice seeds were sterilized and germinated onto ½ Murashige Skoog media, supplemented with 1% sucrose. The seedlings were grown for 10 days in magenta boxes before transferring to soil. The plants were grown in greenhouses maintained at 30 °C day/ 25 °C night at a relative humidity (RH) of 50-70% and with supplemental LED light (Valoya, BX100 NS1) using a 8-hr/16-hr day/night photoperiod at 400 µmol/$m^{-2}s^{-1}$. The plants were fertilized weekly from the 2nd week and biweekly from the 6th week after germination. At IRD, plants were sown in soil complemented with 3 g of fertilizer (N: 19%; P: 5%; K: 8%) per liter and grown in greenhouses (12-h light at 200 µmol/$m^{-2}s^{-1}$, 28 °C, 80% RH; and12-h darkness at 25 °C and 60% RH).

Disease resistance scoring by leaf-clipping inoculation

Bacteria were grown on PSA media (Peptone Sucrose Agar, 1% peptone, 1% sucrose, 0.1% glutamic acid, 1.6% agar) for four days at 28 °C. Single colonies were picked and patched onto PSA and then grown for 24 h before being washed with and diluted in sterile water at OD_{600} of 0.2 (at IRD) and 0.5 (at HHU). The youngest, fully extended leaves of 4- to 6-week-old rice plants were clipped 2-3 cm from the leaf tip with scissors that had been dipped in the inoculum or sterile water. Lesion lengths were measured 14 days after inoculation. For the race typing assay, lesion length measurements < 5 cm were scored as resistant (R), 5–10 cm as moderately resistant (MR), 10–15 cm as moderately susceptible (MS), and >15 cm as susceptible (S).

Sampling of symptomatic leaf material and isolation of *Xoo* strains

Small-scale rice fields cultivated under irrigated conditions in the villages of Dakawa and Lukenge (in the Morogoro region of Tanzania) were surveyed for BB between 2019 and 2022. In each field, diseased leaves of several individual plants were sampled and processed for bacterial isolation as reported previously[29]. Colony-multiplex PCR was used to validate that isolates were *Xanthomonas oryzae* pv. *oryzae* (except for 2022 data, for which analyses have been initiated).

Tanzanian strain genome sequencing and analysis

DNA was extracted from pure bacterial cultures grown on rich media (1% peptone, 0.1% glutamic acid) with QIAGEN Genomic-tip 100/G. Multiplex libraries were prepared with the rapid library preparation kit (SQK-RBK110-96, Oxford Nanopore Technologies, ONT) and sequenced with a MinION Mk1C device on R10.3 (FLO-MIN111) flow cells. ONT electric signals were base-called with a high accuracy model (dna_r10.3_450bps_hac.cfg) and demultiplexed with the ONT Guppy base calling software (v6.0.1+652ffd179). Illumina library construction and sequencing was performed by FASTERIS (Plan-

les-Ouates, Switzerland) on an Illumina NextSeq sequencer with 350-bp inserts and 150-bp paired-end sequences. Long ONT read assembly was performed with the CulebrONT pipeline (v2.0.1)[42] and included FLYE (2.9-b1768) for primary assembly followed by RACON (v1.4.20) and MEDAKA (1.4.1) for assembly polishing. The ONT-only assemblies were subsequently polished with polypolish (v0.5.0)[43]. Core genome SNP genotyping and tree reconstruction were conducted as previously described[44], except that the raxml command included "-N 500 -m GTRGAMMAIX". This substitution model was adopted based on the output of the 'modelTest' function of the R package phangorn for model selection[45]. The X11-5A assembly (GB Acc. GCF_000212755.2) was used as an out-group, and was not displayed in Figure 2. The genomes included in the species-wide *Xanthomonas oryzae* phylogenetic tree are listed together with NCBI accessions and genome metadata in Supplementary Data 1 and Supplementary Table 3. Predictions for *TALe* genes from the genomes of the Tanzanian strains used AnnoTALE with default parameters[46]. Predictions for high-scoring EBEs of the Tanzanian TALes in the Nipponbare *SWEET11a* promoter used Talvez with default parameters[47]. Talvez target predictions on the Nipponbare SWEET11a promoter included putative TALes from Tanzanian strains (CIX4462, CIX4506, CIX4509, T19, Dak16, and Xoo3-1), Asian strains (PXO99^{A}, PXO86, PXO83, PXO71, PXO61) and an African strain (MAI1), which were included as references. Only predictions with a score above 9.05 are displayed in Figure 2.

qRT-PCR analyses of *SWEET11a* mRNA levels

Leaves of 3-week-old rice plants were infiltrated with a bacterial suspension at an OD_{600} of 0.5 or water using a needleless syringe. One leaf per plant was infiltrated, and three plants per treatment were used. Samples were collected at 24 h (Kitaake) or 48 h (wildtype Komboka and edited lines) after inoculation. Total RNA was extracted using TRI reagent (Euromedex). Following extraction, DNase I treatment was conducted using Turbo DNA-free kit (Thermo Fisher Scientific). Subsequent synthesis of complementary DNA was carried out using SuperScriptIII (Thermo Fisher Scientific) and oligo-dT primers. Quantitative PCR reactions were performed with SYBR Mesa Blue qPCR Mastermix (Eurogentec). Three technical replicates were prepared for each sample. Expression values were normalized by subtracting the values obtained for reference gene *EF-1α* (GenBank: GQ848072.1) to the studied replicates ($2^{-\Delta Ct}$ method). Mean of the technical replicates for each sample was calculated. Three independent biological replicates were analyzed. Primer sequences are listed in Supplementary Table 7).

Analysis of EBEs in *SWEET* promoters of *O. sativa* cv. Komboka

Genomic DNA was extracted from Komboka leaves using the CTAB method[48]. The gDNA fragments corresponding to promoters and first exons of *SWEET11a*, *13*, *14* were PCR-amplified using specific primer pairs. Amplicons were gel-purified and cloned into pJET1.2 (Thermo Fischer Scientific). Competent *E.coli* TOP10 competent cells were used for transformation. Plasmids were extracted from

three individual colonies and were sent for Sanger sequencing. Sanger sequencing reads representing promoter regions of *SWEET11a, 13,* and *14* from Komboka were aligned with the respective promoter regions of Kitaake (Kitaake_OsativaKitaake_499.genome; Phytozome v.13) (Supplementary Data 2).

Analyses of *Xa1* and *Xa4* sequences from *O. sativa* cv. Komboka

Full-length *Xa1* and four fragments of *Xa4* were PCR-amplified from *O. sativa* cv. Komboka gDNA using TAKARA PrimeSTART GXL polymerase and primers. Amplicons corresponding to full length *Xa1* were subjected to Sanger sequencing. *Xa4* fragments were subcloned into pUC57Gent vector (Addgene #54338)[49] and sequenced using Sanger sequencing. Individual sequence reads were mapped to their references and assembled via Geneious Prime® 2022.1.1. Neighbor-joining phylogenetic analyses were made using CLUSTAL Omega (Geneious Prime). Genomic sequences of non-Komboka derived *Xa1* and *Xa4* had previously been published[35,50]. Sequences of *Xa1* and *Xa4* genes used for phylogenetic analysis are provided in Supplementary Data 3.

Hybrid CRISPR/Cas9 and Cpf1 vector construction

A new hybrid CRISPR-Cas9/Cpf1 system was developed to edit a maximum of 6 different targets (Extended Data Fig. 7). The hybrid CRISPR-Cas9/Cpf1 combines a duplex Cas9 gRNA combined with a multiplex Cpf1 cRNA. To develop this hybrid CRISPR-Cas9/Cpf1 system, we subcloned a human-codon-optimized *LbCpf1*-coding sequence (www.addgene.org/69988/sequences)[38] and the rice Ubiquitin 1 (*OsUbi*) (LOC_Os02g06640) to create a gateway system pDEST:Cpf1. A PCR fragment of a Cas9 expression cassette, which contains a rice-codon-optimized *SpCas9*-coding sequence driven by the maize Ubiquitin 1 (*ZmUbi*) promoter[51,52], was amplified and insert into pDEST:Cpf1 by Gibson Assembly (New England BioLabs). Finally, we constructed the hybrid destination vector pDEST:Cpf1&Cas9. A Ribozyme-gRNA-Ribozyme (RGR) system was used to generate multiple cRNAs with different target sequences by flanking the cRNAs with a Hammerhead (HH) type ribozyme and a hepatitis delta virus (HDV) ribozyme[53]. Guide RNA sequences are provided in Supplementary Table 8. The promoter of the small nuclear RNA gene from rice *OsU6.1* and *OsU3*, wheat *TaU3* and maize *ZmU3* were used to drive expression of RGG-cRNAs units. Four gBlock fragments synthesized by IDT (Integrated DNA Technologies, Inc., Iowa, USA) were inserted into the vector pTLN using *Xba*I and *Xho*I to produce four intermediate pUNIT vectors: pTL-OsU6.1_CpfRNA1, pTL-OsU3_CpfRNA2, pTL-TaU3_CpfRNA3, pTL-ZmU3_CpfRNA4. A double-stranded DNA oligonucleotide for each site was produced by annealing two complementary oligonucleotides. The DNA sequence of positive clones was confirmed by Sanger sequencing. All four pUNIT vectors were transferred into another donor vector named pENTR4-ccdB using Golden Gate cloning to generate pENTR:cRNAs containing Cpf1 gRNAs cXo1, cXo2c, cTalF and cXo2d. Simultaneously, a gTalC double-stranded oligonucleotide was inserted into the *BtgZ*I-digested pENTR-gRNAs[51], and then the whole gRNA expression cassette was subcloned into pENTR:cRNAs to generated the hybrid donor plasmid pENTR:gRNA-cRNAs. The gRNA-cRNA

cassette was mobilized to the hybrid destination vector pDEST:Cpf1&Cas9 using Gateway LR Clonase (Thermo Fisher Scientific) to produce a pCam1300-CRISPR plasmid named pCpf1&Cas9:gRNAs-cRNAs (pMUGW5).

Rice transformation

Agrobacterium-mediated transformation of Komboka using immature embryos was performed as described [22]. Briefly, the *Agrobacterium tumefaciens* strain LBA4404 was transformed with pMUGW5 via electroporation. Immature rice seeds at the late milky stage were harvested for immature embryo isolation. Immature embryos were inoculated with 5 µl *Agrobacterium* suspension (OD_{600} = 0.3) and incubated in the dark for one week. The emerged shoots were removed, and immature embryos were incubated for 5 days at 28 °C under continuous light. To select positive transformants, four rounds of hygromycin selection were applied, with each round consisting of 10 days. Resistant calli were moved onto a pre-regeneration medium and incubated for 10 days. Greening calli were transferred to a regeneration medium to develop small rice plantlets. Once plantlets reached 15-cm in height, they were transferred to soil and placed in a greenhouse. After 5 months, T1 seeds were harvested.

Screening of EBE edited lines

To test whether generated T0 plants contained T-DNA insertion, the presence of *SpCas9*, *LbCpf1* and hygromycin resistance (*Hpt*) genes were checked. Leaf fragments (3 to 4 cm) were harvested for DNA extraction using a modified CTAB method for high-throughput DNA extraction. PCR was performed using GoTaq DNA Polymerase (Promega) with a melting temperature of 55 °C for *SpCas9*, *LbCpf1* and *Hpt*. For genotyping of EBE mutations, the four regions containing the six EBEs within the *SWEET11a*, *13* and *14* promoters were amplified with specific primers using Phusion HF polymerase (Thermo Fisher Scientific), and PCR amplicons were sequenced by Sanger sequencing (Microsynth seqlab). Chromatograms were analyzed manually using Benchling (www.benchling.com) to detect the mutations. Two rounds of transformation were performed. In Round 1, seven T0 plants were obtained from a single embryo, and all T0 plants contained biallelic mutations at all target sites (Supplementary Table 5, Supplementary Table 9, Supplementary Table 10). In the transformation Round 2, 18 T0 plants were T-DNA positive, from 18 independent immature embryos. Nine T0 plants were characterized further and found to contain biallelic mutations at all targeted sites (Supplementary Table 10Supplementary Table 12,). We screened 52 T1 plants from three independent events (#12, 14 and 16) and obtained 7 lines with homozygous mutations in all 6 EBEs (Supplementary Table 5, Supplementary Table 6, Supplementary Table 9).

Statistical analyses

Statistical analyses and graphical representations were prepared using R 4.0.5 (ref), on RStudio for Windows. Statistics were calculated using rstatix (https://cran.r-project.org/web/packages/rstatix). To determine if there was an effect of the treatment or the genotype on gene expression or lesion length,

the medians of the different subsamples were compared using rstatix Kruskal-Wallis test (*p*-values < 0.05). Then, mean groups were attributed to the different genotypes or treatments independently using rstatix Dunett's test (*p*-values < 0.05). Letters were attributed to each mean group. Graphical representations were realized using package ggplot2 (https://ggplot2.tidyverse.org). Data were drawn as boxplots. The boxplot is delimited by the first and the third quartile of the distribution of the studied variable. The line inside the boxplot represents the median of the variable. Finally, the two lines that start from the boxplot join the minimum and maximum theoretical values. Outliers (<7% of total values) were represented as black dots. The total number (n) of replicates for each strain/genotype condition is displayed next to each boxplot. Each observation is represented by a dot, with a color code by strain or genotype. The letters above each boxplot represent the mean groups calculated using Dunnett's test.

References

1. Odongo, P. J. *et al.* Insights into natural genetic resistance to rice yellow mottle virus and implications on breeding for durable resistance. *Front. Plant Sci.* **12**, 671355 (2021).
2. Pandey, S. *et al. Rice in the global economy: strategic research and policy issues for food security*. *IRRI Books* (International Rice Research Institute (IRRI), 2010).
3. Jiang, N. *et al.* Resistance genes and their interactions with bacterial blight/leaf streak pathogens (*Xanthomonas oryzae*) in rice (*Oryza sativa L.*)—an updated review. *Rice* **13**, 3 (2020).
4. Mutiga, S. K. *et al.* Integrated strategies for durable rice blast resistance in sub-Saharan Africa. *Plant Dis.* **105**, 2749–2770 (2021).
5. Longue, R. D. S. *et al.* Pathogenicity of rice yellow mottle virus and screening of rice accessions from the Central African Republic. *Virol. J.* **15**, 6 (2018).
6. Vera Cruz, C. M. *et al.* Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13500–5 (2000).
7. Arra, Y. *et al.* Virulence profiling of *Xanthomonas oryzae pv. oryzae* isolates, causing bacterial blight of rice in India. *Eur. J. Plant Pathol.* **149**, 171–191 (2017).
8. Arra, Y. *et al.* Incorporation of the novel bacterial blight resistance gene *Xa38* into the genetic background of elite rice variety improved Samba Mahsuri. *PloS One* **13**, e0198260 (2018).

9. Buddenhagen, I. W., Vuong, H. H. & Ba, D. D. Bacterial blight found in Africa. *Int. Rice Res. Newsl.* **4**, 11 (1979).
10. Verdier, V., Vera Cruz, C. & Leach, J. E. Controlling rice bacterial blight in Africa: Needs and prospects. *J. Biotechnol.* **159**, 320–328 (2012).
11. Duku, C., Sparks, A. H. & Zwart, S. J. Spatial modelling of rice yield losses in Tanzania due to bacterial leaf blight and leaf blast in a changing climate. *Clim. Change* **135**, 569–583 (2016).
12. Amos, O. Two genotypes of *Xanthomonas oryzae pv. oryzae* virulence identified in West Africa. *Int. J. Genet. Mol. Biol.* **5**, 28–36 (2013).
13. Séré, Y. *et al.* Rice bacterial leaf blight in West Africa: preliminary studies on disease in farmers` fields and screening released varieties for resistance to the bacteria. *Asian J. Plant Sci.* **4**, 577–579 (2005).
14. Richter, A. *et al.* A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* **5**, 3447 (2014).
15. Wu, L.-B. *et al.* OsSWEET11b, a potential sixth leaf blight susceptibility gene involved in sugar transport-dependent male fertility. *New Phytol.* **234**, 975–989 (2022).
16. Sadoine, M. *et al. Sucrose-dependence of sugar uptake, quorum sensing and virulence of the rice blight pathogen* Xanthomonas oryzae pv. oryzae. https://www.biorxiv.org/content/10.1101/2021.08.22.457195v1 (2021) doi:10.1101/2021.08.22.457195.
17. Eom, J.-S. *et al.* Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* **37**, 1372–1379 (2019).
18. Oliva, R. *et al.* Broad-spectrum resistance to bacterial blight in rice using genome-editing. *Nat. Biotechnol.* **37**, 1344–1350 (2019).
19. Tran, T. T. *et al.* Functional analysis of African *Xanthomonas oryzae pv. oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **14**, e1007092 (2018).
20. Verdier, V. *et al.* Transcription activator-like (TAL) effectors targeting *OsSWEET* genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. *New Phytol.* **196**, 1197–1207 (2012).

21. Buchholzer, M. & Frommer, W. B. An increasing number of countries regulate genome editing in crops. *New Phytol.* **online**, (2022).
22. Luu, V. T. *et al.* Efficient Agrobacterium-mediated transformation of the elite–indica rice variety Komboka. *Bio-Protoc.* **10**, e3739–e3739 (2020).
23. Kitilu, M. J. F., Nyomora, A. M. S. & Charles, J. Growth and yield performance of selected upland and lowland rainfed rice varieties grown in farmers and researchers managed fields at Ifakara, Tanzania. *Afr. J. Agric. Res.* **14**, 197–208 (2019).
24. van der Tycho, T. A. & Yonho, K. On the success and failure of North Korean development aid in Africa. in (ed. Yonho, K.) 31–42 (George Washington University, 2022).
25. Makundi, H. *Diffusing Chinese rice technology in rural Tanzania: Lessons from the Dakawa agro-technology demonstration center*. *SAIS-CARI Working Papers* https://ideas.repec.org/p/zbw/cariwp/201712.html (2017).
26. Afolabi, O. *et al.* First report of bacterial leaf blight of rice caused by *Xanthomonas oryzae pv. oryzae* in Benin. *Plant Dis.* 515 (2016).
27. Gonzalez, C. *et al.* Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. *Mol. Plant-Microbe Interact. MPMI* **20**, 534–546 (2007).
28. Tall, H. *et al.* Confirmation report of bacterial blight caused by *Xanthomonas oryzae pv. oryzae* on rice in Senegal. *Plant Dis.* **104**, 968–968 (2020).
29. Tekete, C. *et al.* Characterization of new races of *Xanthomonas oryzae pv. oryzae* in Mali informs resistance gene deployment. *Phytopathology* **110**, 267–277 (2020).
30. Lang, J. M. *et al.* Genomics-based diagnostic marker development for *Xanthomonas oryzae pv. oryzae* and *X. oryzae pv. oryzicola*. *Plant Dis.* **94**, 311–319 (2010).
31. Lang, J. M. *et al.* A pathovar of *Xanthomonas oryzae* infecting wild grasses provides insight into the evolution of pathogenicity in rice agroecosystems. *Front. Plant Sci.* **10**, (2019).
32. Quibod, I. L. *et al.* Effector diversification contributes to *Xanthomonas oryzae pv. oryzae* phenotypic adaptation in a semi-Isolated environment. *Sci. Rep.* **6**, 34137 (2016).
33. Chen, L. Q. *et al.* Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**, 207–211 (2012).

34. Streubel, J. *et al.* Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **200**, 808–19 (2013).
35. Ji, C. *et al.* Xa1 allelic R genes activate rice blight resistance suppressed by interfering TAL effectors. *Plant Commun.* **1**, 100087 (2020).
36. Quibod, I. L. *et al.* The Green Revolution shaped the population structure of the rice pathogen Xanthomonas oryzae pv. oryzae. *ISME J.* **14**, 492–505 (2020).
37. Liu, Y. *et al.* Engineering cell signaling using tunable CRISPR–Cpf1-based transcription factors. *Nat. Commun.* **8**, 2095 (2017).
38. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
39. Xu, Z. *et al.* Engineering broad-spectrum bacterial blight resistance by simultaneously disrupting variable TALE-binding elements of multiple susceptibility genes in rice. *Mol. Plant* **12**, 1434–1446 (2019).
40. Zhang, H. *et al.* The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* **12**, 797–807 (2014).
41. Mew, T. W. *et al. Rice diseases: Biology and selected management practices.* (International Rice Research Institute, 2018).
42. Orjuela, J. *et al.* CulebrONT: a streamlined long reads multi-assembler pipeline for prokaryotic and eukaryotic genomes. *bioRxiv* 2021.07.19.452922 (2022) doi:10.1101/2021.07.19.452922.
43. Wick, R. R. & Holt, K. E. Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLoS Comput. Biol.* **18**, e1009802 (2022).
44. Doucouré, H. *et al.* Functional and genome sequence-driven characterization of TAL effector Gene repertoires reveals novel variants with altered specificities in closely related Malian *Xanthomonas oryzae pv. oryzae* strains. *Front. Microbiol.* **9**, 1657 (2018).
45. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinforma. Oxf. Engl.* **27**, 592–593 (2011).
46. Grau, J. *et al.* AnnoTALE: bioinformatics tools for identification, annotation, and nomenclature of TALEs from Xanthomonas genomic sequences. *Sci. Rep.* **6**, 21077 (2016).

47. Pérez-Quintero, A. L. *et al.* An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLOS ONE* **8**, e68464 (2013).
48. Li, H. *et al.* A high-throughput, high-quality plant genomic DNA extraction protocol. *Genet. Mol. Res. GMR* **12**, 4526–4539 (2013).
49. Binder, A. *et al.* A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. *PLOS ONE* **9**, e88218 (2014).
50. Hu, K. *et al.* Improvement of multiple agronomic traits by a disease resistance gene via cell wall reinforcement. *Nat. Plants* **3**, 1–9 (2017).
51. Char, S. N. *et al.* An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnol. J.* **15**, 257–268 (2017).
52. Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H. & Yang, B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* **42**, 10903–10914 (2014).
53. Gao, Y. & Zhao, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* **56**, 343–349 (2014).

Acknowledgments

We thank Paula Emmerich Maldonado (HHU) and Britta Killing (HHU) for excellent technical assistance. We thank Dr. Marietta Wolter for advice on using the QX200 Droplet Digital PCR (ddPCR) System (Biorad) at the Institute for Neuropathology, University Hospital Düsseldorf. The authors acknowledge the ISO 9001-certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing high-performance computing resources. WBF and YA acknowledge support by the Alexander von Humboldt Foundation. This work was made possible by funding from the Bill and Melinda Gates Foundation to Heinrich Heine University, Düsseldorf, with subawards to Iowa State University/University of Missouri, University of Florida, Institut de Recherche pour le Développement, and the International Rice Research Institute.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data supporting the results are available in the main text or supplementary materials. All data that support the findings of this study are available from the corresponding author upon request. Sequencing data for strains from this study have been deposited in the NCBI Sequence Read Archive (SRA) database under the accession codes will be provided at the time of publication. Source data will be deposited before publication.

Contributions

V.S.L., M.S., Y.A., K.S., M.B., and E.P.I.L.: *R*-genes in Komboka, editing, genotyping; C.S., G.B. and F.A.: Race-typing, characterization of the resistance of genome-edited lines; F.A., C.S., A.D. and S.C.: strain isolation, whole genome sequencing and analyses; C.J., S.N.C., and B.Y.: hybrid Cas9/Cpf1 system; A.L.B., D.L., M.M., and R.M.: survey in Tanzania, disease monitoring and sampling; R.O., B.Y., B.S., and W.B.F.: conceived and designed experiments; V.S.L., M.S., Y.A., E.P.I.L., C.S., S.C, R.S., B.Y., B.S., and W.B.F. analyzed the data; R.O., M.B., B.Y., B.S. and W.B.F. wrote the manuscript.

Corresponding authors

Correspondence to Boris Szurek and Wolf B. Frommer

Ethics declarations

Competing interests

The Healthy Crops team has filed for patents, but is charitable and non-profit and aims at helping small scale rice farmers in Asia and Africa by reducing yield losses caused by pathogens . The authors declare no competing interests.

Main figures and tables

Table 1: Efficiency of resistance genes in IRBB rice lines towards a diversity panel of 21 African *Xoo* strains. Qualitative scores of the lesion lengths produced upon leaf-clip inoculation of a diversity panel consisting of 21 *Xanthomonas oryzae* pv. *oryzae* strains from West and East Africa on IRBB near-isogenic-lines (NILs) carrying different *Xanthomonas* resistance genes **(*Xa*)** and on the variety Azucena used as susceptible control (IRBB1: *Xa1*; IRBB3: *Xa3*; IRBB4: *Xa4*; IRBB5: *xa5*; IRBB7: *Xa7*; IRBB23: *Xa23*; IR24: susceptible control). Resistance or susceptibility of rice plants to *Xoo* was determined by lesion length measured 14 days after inoculation: Resistant (R) < 5 cm; moderately resistant (MR) = 5 to 10 cm; moderately susceptible (MS) = 10 to 15 cm; and susceptible (S) >15 cm.

CIX Number	Strain Name	Country	Azucena	IR24	IRBB1	IRBB3	IRBB4	IRBB5	IRBB7	IRBB23
607	Nati Park	Benin	S	**R**	**R**	**R**	**R**	**R**	**R**	**R**
705	Karfiguela 13	Burkina Faso	S	**R**	**R**	**R**	**R**	**R**	**R**	**R**
4482	N2-4	Niger	S	**R**	**R**	**R**	**R**	**R**	**R**	**R**
609	Tanguieta3	Benin	S	MR	**R**	MR	**R**	**R**	MR	**R**
629	Toula20	Niger	S	MR	**R**	MR	**R**	**R**	MR	**R**
4127	BAI250	Burkina Faso	S	MR	**R**	MR	MR	**R**	**R**	**R**
2374	S62-2-22	Senegal	S	MR	**R**	MS	**R**	**R**	MR	**R**
4083	CII-1	Ivory Coast	S	MS	MR	MS	**R**	**R**	MR	**R**
1042	CII-2	Ivory Coast	S	MS	MR	MS	MR	**R**	MR	**R**
2787	NAI9	Niger	S	MS	MR	MS	MR	MR	MS	**R**
1917	AXO1947	Cameroon	S	S	**R**	S	**R**	**R**	MS	**R**
894	MAI145	Mali	S	S	**R**	S	MR	MS	S	**R**
4092	BAI3	Burkina Faso	S	MS	MR	S	MR	MR	MS	**R**
2801	CFBP1948	Cameroon	S	MS	MR	S	MR	MR	S	**R**
4099	NAI5	Niger	S	S	MR	S	MR	MR	S	**R**
4079	MAI73	Mali	S	S	**R**	S	MS	MR	S	**R**
2976	S82-4-3	Senegal	S	MS	MR	S	MS	MR	MS	**R**
4517	MAI132	Mali	S	S	R	S	S	MS	S	**R**
4457	Tz1	Tanzania	S	S	S	S	S	S	S	S
4458	Tz2	Tanzania	S	S	S	S	S	S	S	S
4462	Tz6	Tanzania	S	S	S	S	S	S	S	S

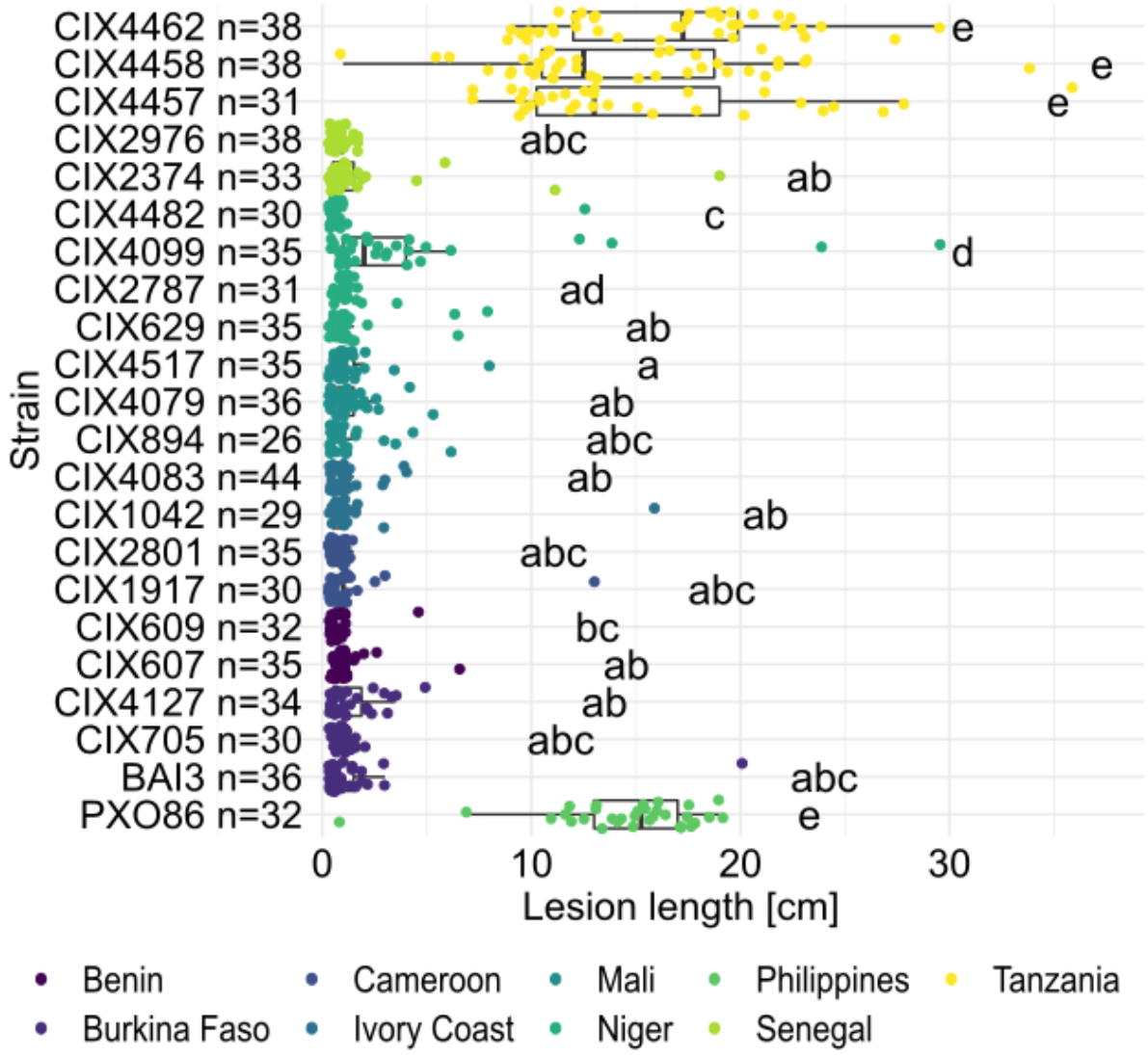


Figure 1. Resistance spectrum of wild-type Komboka rice against African *Xoo* strains. Leaf-clip inoculation of wild-type Komboka rice plants with a panel of 21 *Xoo* strains originating from 8 African countries along with Asian reference strain PXO86 from the Philippines. Lesion length (in cm) was measured 14 days after *Xoo* inoculation. Data from three independent experiments are represented. Letters in the boxplot represent significant differences.

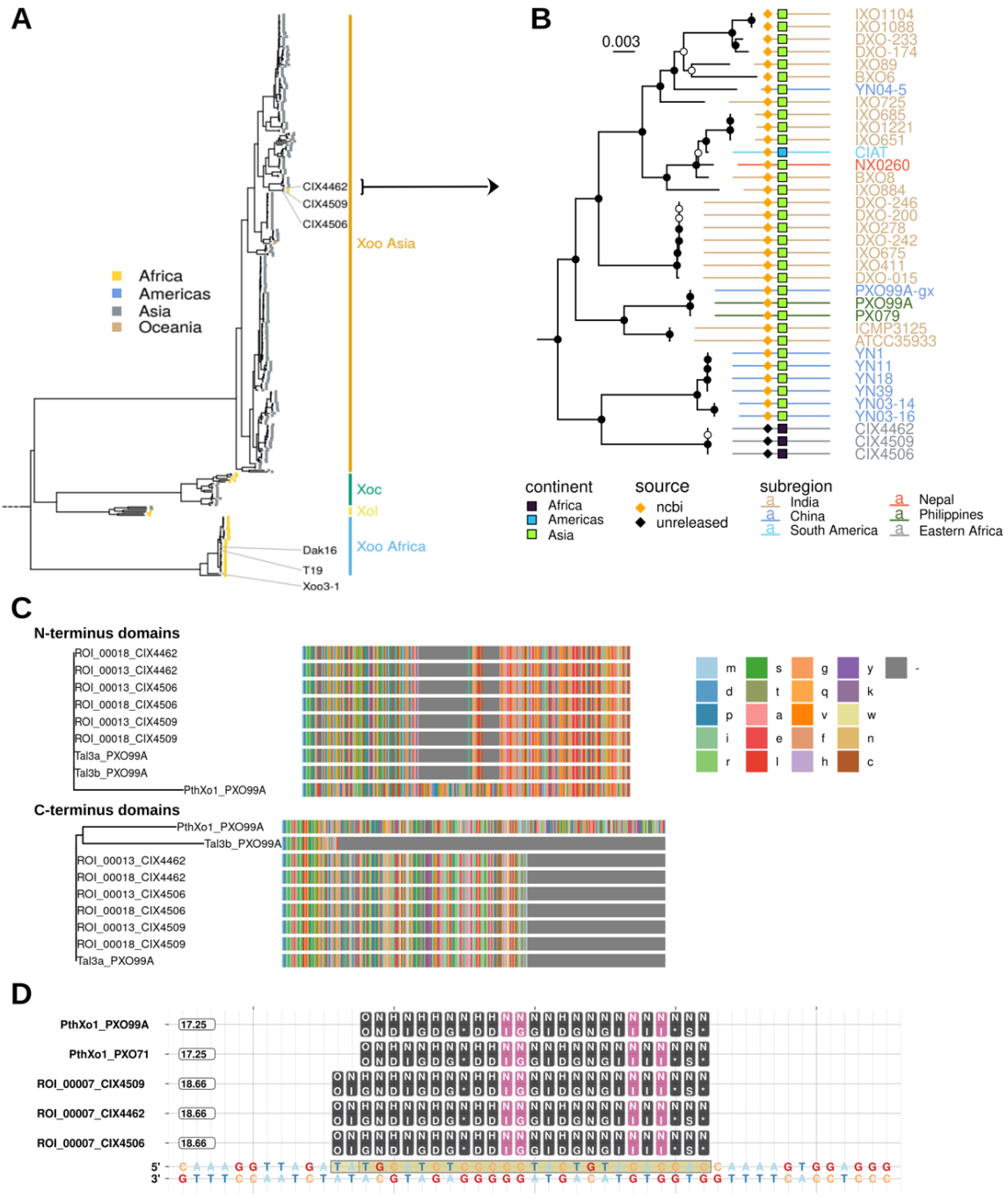


Figure 2. Analysis of the genomes of Tanzanian *Xoo* isolates. A. Core genome *Xanthomonas oryzae* phylogenetic tree. Only the names of Tanzanian isolates are indicated. Abbreviations: *Xoo*, *X. oryzae* pv. *oryzae*; *Xoc*, *X. oryzae* pv. *oryzicola*; *Xol*, *X. oryzae* pv. *leersiae*. **B.** Close-up view of a portion of the tree in A, focusing on the newly isolated Tanzanian strains and neighboring clades. Black-filled nodes have a bootstrap support value equal to or above 80%. The scale bar reflects branch length in mean number of nucleotide substitutions per site. **C.** Miniature multiple alignment of the terminal domain sequences of PXO99^{A} iTALes and the putative iTALes from the three newly isolated Tanzanian strains. The PXO99^{A} PthXo1 domains act as 'standard' TALe references, with lowercase letters representing amino acids. Gaps are colored in gray. **D.** Talvez EBE predictions on the *SWEET11a*

promoter. Repeat Variable Di-residue (RVD) sequences (rounded boxes) are aligned along their predicted matching nucleotide along the promoter sequence (the lowest row). Black-filled RVDs match their target nucleotide in the Talvez RVD-nucleotide association matrix with the best possible score for this RVD. Those in violet match with an intermediate score. Values in the rounded boxes near the TALe names correspond to Talvez prediction scores.

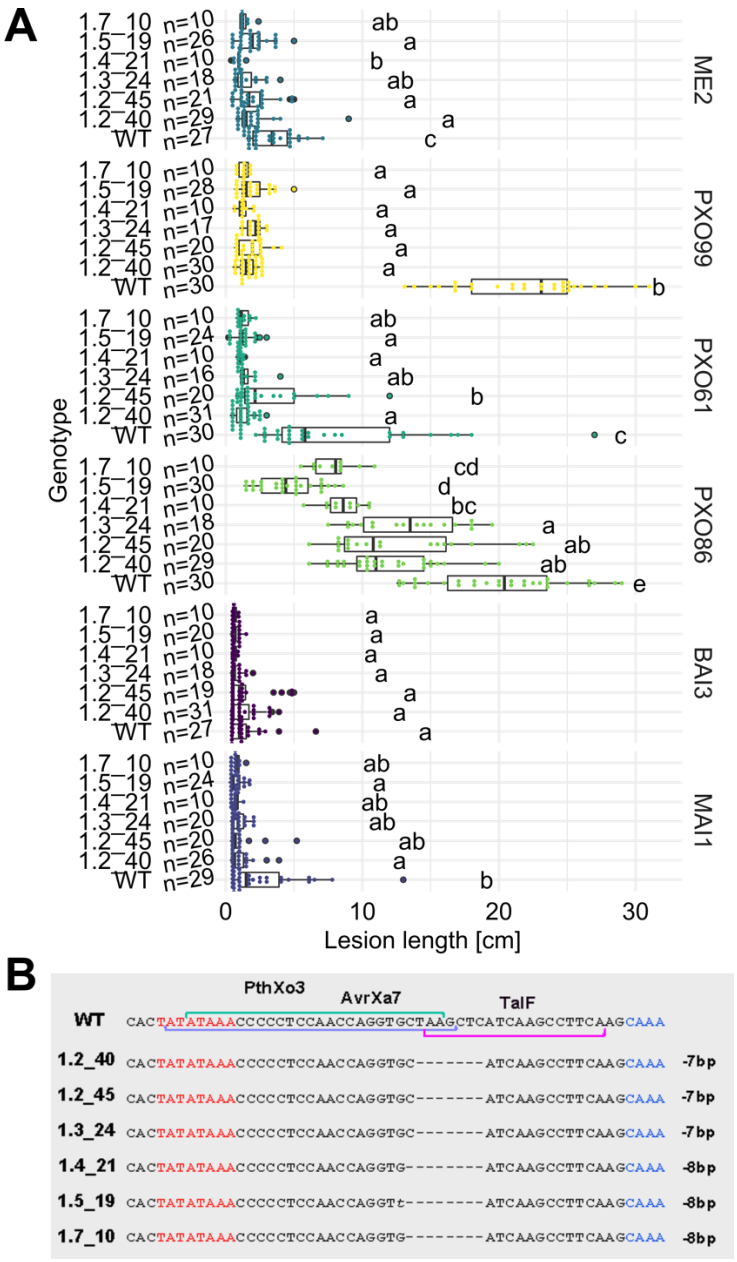


Figure 3. Resistance of EBE-edited Komboka lines against six representative *Xoo* strains. A. Reactions of WT and six edited Komboka lines to infection by *Xoo* strains (PXO99^{A}, PXO61, PXO86, BAI3 and MAI1) harboring PthXo1, PthXo2 PthXo3, AvrXa7, TalC and/or TalF. ME2 is a PXO99^{A} mutant strain with an inactivated *pthXo1* and served as a negative control. The Komboka lines were determined to be edited in all 6 EBE sites by sequencing. **B**. Details of the EBEs targeted by PthXo3, AvrXa7 and TalF and their mutations in six edited lines.

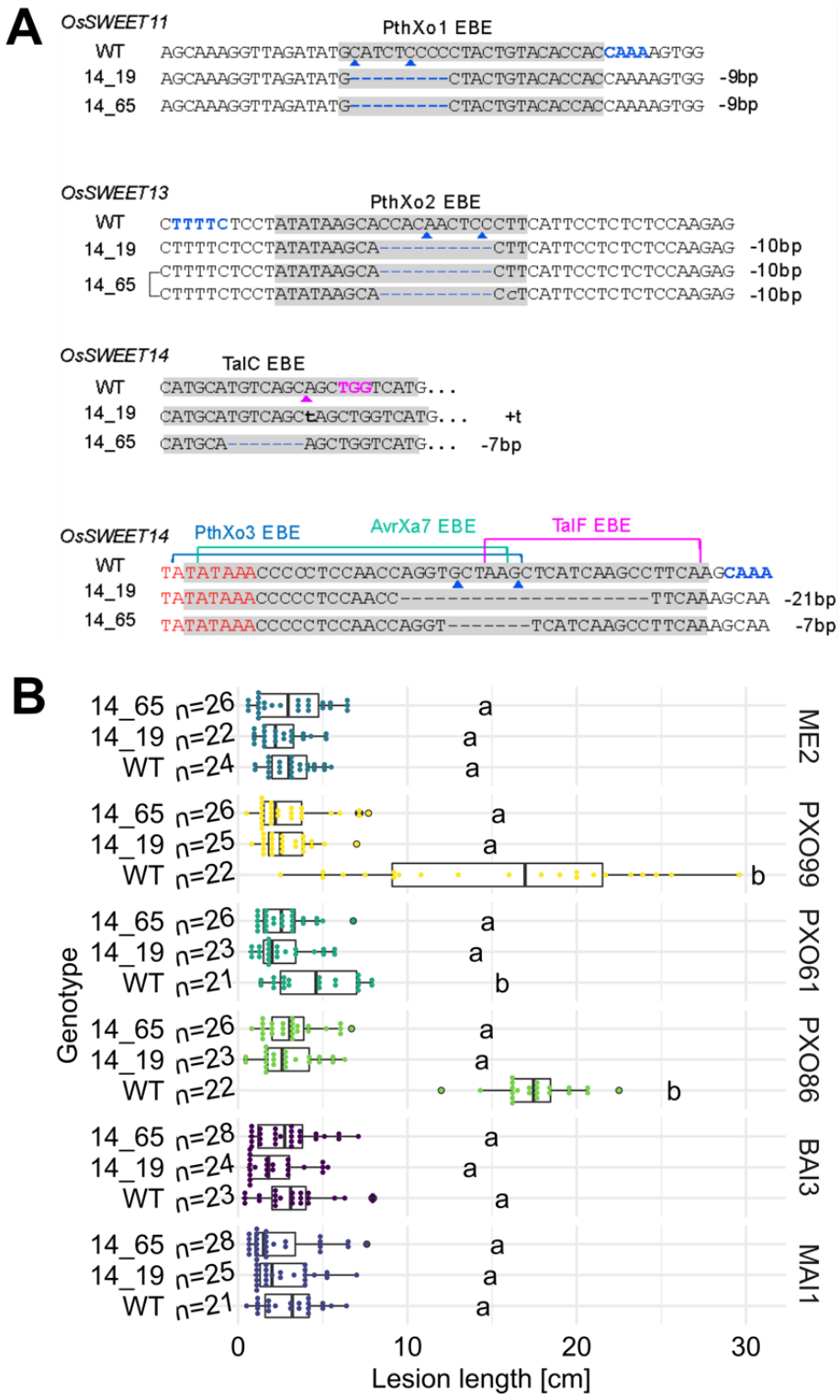


Figure 4. Genotypes and phenotypes of two EBE-edited lines generated during the second transformation experiment. A. Mutations at the EBEs for PthXo1, PthXo2, TalC, TalF, PthXo3, and AvrXa7 in two Komboka edited lines, 14_19 and 14_65. **B**. Reactions of wild-type Komboka and the two edited lines to the infections by PthXo1-, PthXo2-, PthXo3-, AvrXa7-, TalC-, and TalF-harboring *Xoo* strains (PXO99^{A}, PXO61, PXO86, BAI3 and MAI1). ME2 is a PXO99^{A} mutant that is inactivated in *pthXo1* and served as a negative control.

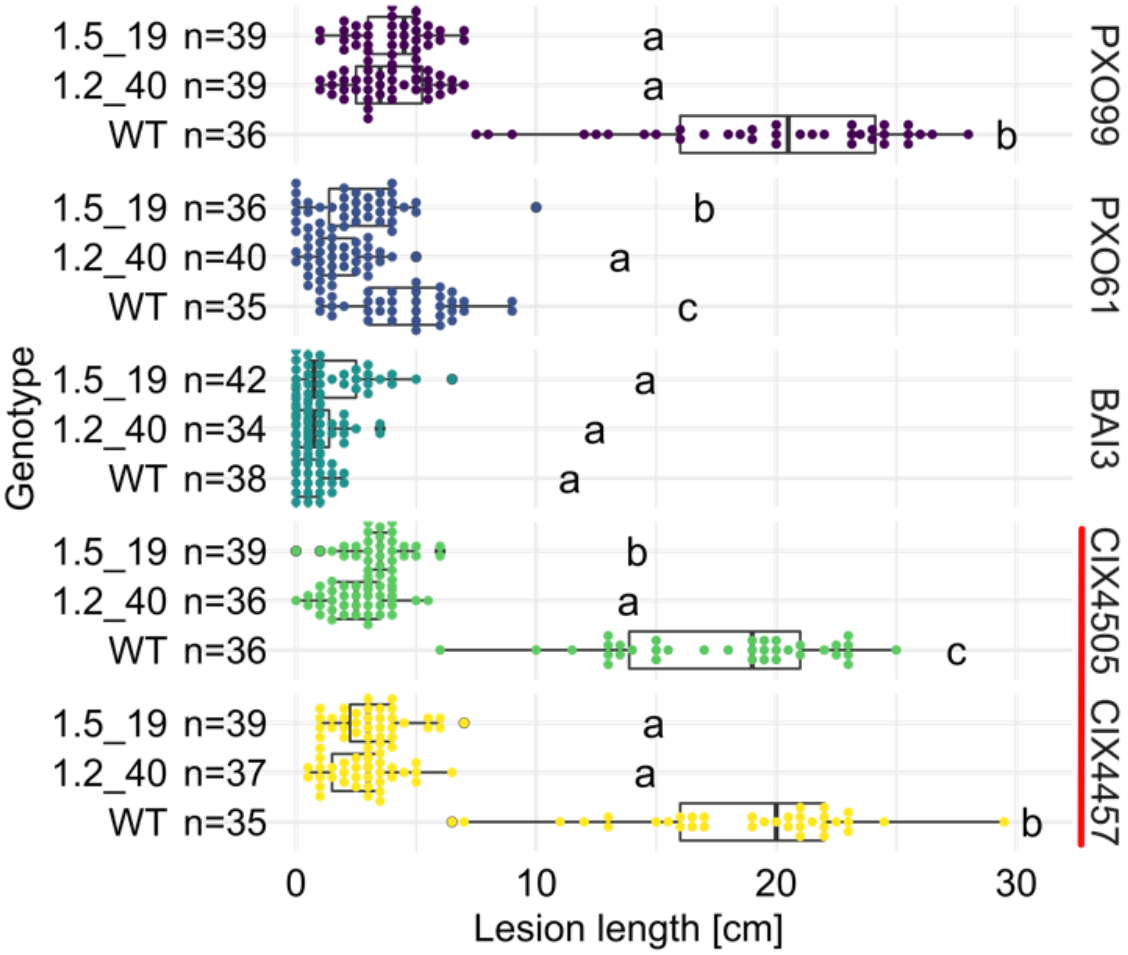


Figure 5. Komboka *SWEET* promoter edited lines are fully resistant against Tanzanian strains of *Xoo*. Lesion lengths (in cm) measured 14 days after leaf-clip inoculation of wild-type Komboka and the two multi-edited lines 1.2_40 and 1.5_19 with *Xoo* strains PXO99 (PthXo1), BAI3 (TalC), PXO61 (PthXo2 & PthXo3), and the newly isolated Tanzanian strains CIX4505 and CIX4457 (highlighted by a red bar). Four independent experiments are represented.

CHAPITRE III : comprendre les mécanismes de la virulence indépendante de l'induction de *SWEET14* chez les *Xoo* africaines

1 Rationale : l'énigme Africaine

La totalité des *Xoo* pathogènes séquencées à ce jour possèdent au moins un TALE capable d'induire un gène *SWEET* du clade III chez le riz. L'utilisation des transporteurs SWEET par les *Xoo* est par ailleurs considérée comme une condition nécessaire au développement de la maladie (Streubel *et al.*, 2013 [289]). Nous savons par exemple que lorsque l'induction des gènes *SWEET* utilisés par les *Xoo* asiatiques est abrogée, ces souches n'engendrent plus de lésions (absence de développement de la maladie) (Chu *et al.*, 2006 [290] ; Hutin *et al.*, 2015 [282]; Zhou *et al.*, 2015 [268]). *SWEET14* est connu pour être le seul SWEET induit par les *Xoo* africaines (Tran *et al.*, 2018 [291] ; Yu *et al.*, 2011 [292]). L'importance de l'activation de ce gène pour la prolifération bactérienne peut être déduite de la redondance fonctionnelle observée dans l'arsenal de TALEs africains. En effet, il existe deux TALEs ciblant *SWEET14* chez les souches africaines : TalC et TalF (Streubel *et al.*, 2013 [289] ; Yu *et al.*, 2011 [292]). TalC est conservé chez toutes les *Xoo* d'Afrique de l'Ouest ayant été étudiées, tandis que TalF n'est pas présent chez toutes les souches (Oliva *et al.*, 2019 [287]). Un variant de TalF, incapable d'induire *SWEET14*, a été découvert chez les souches BAI3 et MAI68, rendant ces souches uniquement dépendantes de TalC (Tran *et al.*, 2018 [291] ; Doucouré *et al.*, 2018 [293]). Une stratégie efficace pour le gain de résistance contre les *Xoo* asiatiques consiste en l'édition des EBEs des TALEs ciblant les gènes *SWEETs* du clade III (Oliva *et al.*, 2019 [287]). Bien que les lignées éditées dans l'EBE de TalF / Avrxa7 (TALE Asiatique) soient résistantes aux souches tributaires de TalF ou de AvrXa7, l'édition de l'EBE TalC ne confère pas la résistance attendue contre les souches dépendantes de celui-ci, malgré l'absence d'induction de *OsSWEET14* et d'à priori tout autre *OsSWEET* du clade III (Doucouré *et al.*, 2018 [293] ; Oliva *et al.*, 2019 [287]). Les *Xoo* africaines auraient donc la capacité à maintenir de la virulence en l'absence d'induction de gènes *SWEET* du clade III, un phénotype que nous avons nommé SWIIVA, pour « *SWEET* Induction Independant Virulence of African *Xoo* ». Dans ce chapitre, nous avons cherché à découvrir quels étaient les éléments conditionnant SWIIVA. Dans ce but, nous avons initié un questionnement sur la dépendance des *Xoo* africaines vis à vis de l'utilisation des transporteurs SWEET. Sont-elles capables d'induire des *SWEETs* du clade III autres que *SWEET14* sans que nous l'ayons détecté auparavant ? Ou existe-t-il une intervention post-transcriptionelle dont le résultat serait l'accumulation de protéines SWEET ? Les souches africaines ont elles un métabolisme des sucres plus performant ? Parallèlement à ces questions, nous avons cherché à mettre en évidence les bases génétiques de SWIIVA chez la bactérie, en nous intéressant plus particulièrement à une région génomique découverte chez un mutant d'une souche originaire du Burkina Faso. Nos différentes hypothèses et leur organisation logique sont présentées en figure S1.

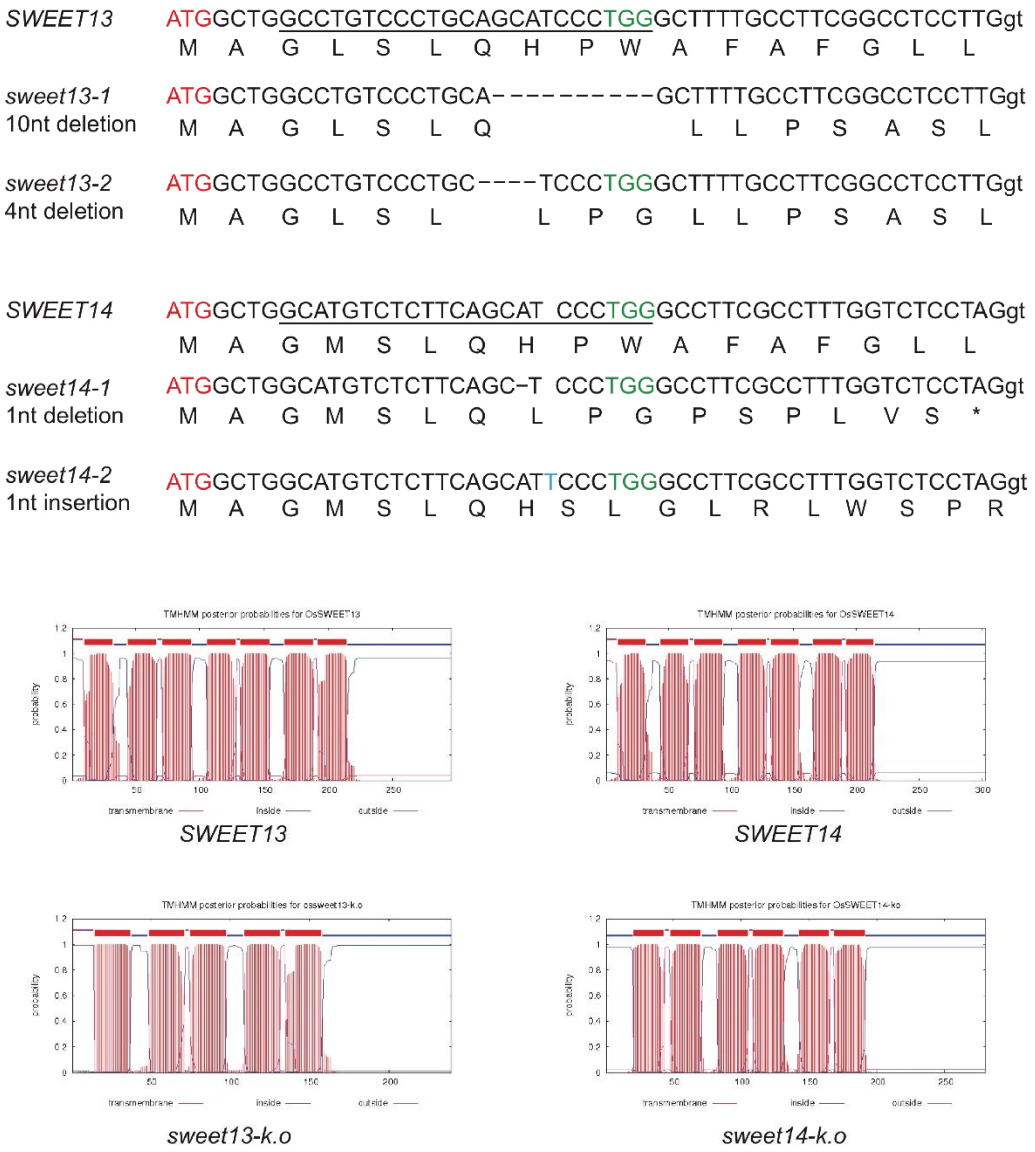


Figure 33 : Mutagenèse de *SWEET13* et *SWEET14* par édition du génome par CRISPR-Cas9. A. Les sites ciblés pas les ARNs guides sont soulignés et les motif PAM sont représentés en vert. Les tirets indiquent les nucléotides supprimés. Une insertion de 1nt dans l'un des mutants *sweet14* (*sweet14*-2) est mis en surbrillance en bleu. B. Prédictions de structures des sous-unités SWEET traduites depuis les gènes mutés. Dans les deux mutants, les mutations conduisent à la perte des deux premiers domaines transmembranaires, conduisant très probablement à des transporteurs non fonctionnels. Figure « Supplementary Figure 10 » de Eom *et al.*, 2019 [288].

2 Recherche des cibles requises pour le développement des *Xoo* africaines chez le riz.

2.1 Identification des *SWEETs* du clade III nécessaires à l'infection

2.1.1 Choix des lignées mutantes

Dans le cadre du projet « Transformative strategies for controlling rices diseases », phase II, des lignées de riz mutantes pour les différents gènes *OsSWEET* du clade III utilisables par *Xoo* ont été produites dans la variété Kitaake (*japonica*). Des mutants ponctuels pour *OsSWEET11*, *OsSWEET13* et *OsSWEET14* ont été obtenus par édition dans la partie codante du gène via transformation d'un plasmide portant le système CRISPR-Cas9 et un ARN guide adapté à chaque gène. Par la suite, des mutants doubles, ou triples, ont été générés par croisement entre les différents mutants simples. De nombreuses données produites par qPCR, analyses transcriptomiques et analyses de TALomes ont montré que les *Xoo* africaines n'étaient pas capables d'utiliser *OsSWEET11* (Yu *et al.*, 2011 [292] ; Streubel *et al.*, 2013 [289] ; Blanvillain-Baufumé *et al.*, 2017 [294]). Par ailleurs, des données obtenues au début du projet par inoculation de 3 souches africaines (Malienne, Burkinabé et Camerounaise) sur le mutant « knock-out » *sweet11* ont montré que ces souches pouvaient tout à fait se développer en l'absence du gène (Eom *et al.*, 2019 [288]). Ainsi, nos études se sont focalisées sur les gènes *OsSWEET13* et *OsSWEET14*. Les mutations des différents mutants obtenus pour ces deux gènes ainsi que les prédictions des séquences et conformation des sous-unités traduites sont représentées en figure 33.

2.1.2 Recherche d'un éventuel gain de résistance acquis chez des plantes mutantes pour *OsSWEET13* et *OsSWEET14*

Afin de déterminer quelle était la dépendance des *Xoo* africaines à l'utilisation des SWEET du clade III du riz, nous avons testé un panel de 21 souches de *Xoo* représentant huit pays d'Afrique de l'Ouest sur des mutants simples *sweet13*, *sweet14* et un double mutant *sweet13* ; *sweet14* (figure 34). Les résultats obtenus montrent clairement que contrairement à ce qui est observé pour la souche contrôle PXO86 issue des Philippines (TALE AvrXa7 ciblant *OsSWEET14*), l'absence de *sweet14* n'empêche pas la colonisation du mutant ponctuel par les *Xoo* africaines. En revanche, un double mutant *sweet13* ; *sweet14* présente un phénotype drastique de résistance. En effet, la moyenne des longueurs de lésions obtenues sur le sauvage est de 12,3 cm, contre 8 cm sur le mutant *sweet14* (ce qui correspond à une diminution des longueurs de lésions d'environ 35%), tandis que l'on passe à une moyenne de 3,4 cm sur le double mutant *sweet13* ; *sweet14*, ce qui équivaut à une diminution de 72,4 %. Ces résultats corroborent les premières données accumulées sur ce double mutant par Eom *et al.* en 2019 [288], et laissent envisager une importance élevée de *OsSWEET13* dans la virulence des *Xoo* africaines. Cependant, lorsque l'on observe les lésions obtenues sur un mutant simple *sweet13*, on constate qu'il n'existe pas ou peu de différences avec le sauvage. En effet, la moyenne des longueurs de lésions reste la même, mais des profils de virulence différenciés sont observés pour certaines souches du panel. Les *Xoo* situées sur la partie supérieure de la figure (BAI4, Nati Park, Karfiguela 13, Toula 13, BAI3 et Toula 20) produisent des lésions plus longues sur le KO *sweet13* que sur le sauvage, tandis que la tendance inverse est visible pour les souches AXO1947, MAI1, CFBP1951, CFBP1949 et NAI7. Ces souches proviennent toutes de différents pays et à ce jour, nous n'avons pas pu mettre en évidence une origine à ce phénotype étonnant. Ces résultats, en plus des précédentes données montrant le phénotype de virulence Africain en l'absence de l'induction de *SWEET14* remettent en cause l'aspect essentiel de ce gène dans le développement de la maladie pour les *Xoo* africaines (Blanvillain-Beaufumé *et al.*, 2017 [294]; Oliva *et al.*, 2019 [287] ; Eom *et al.*, 2019 [288]).

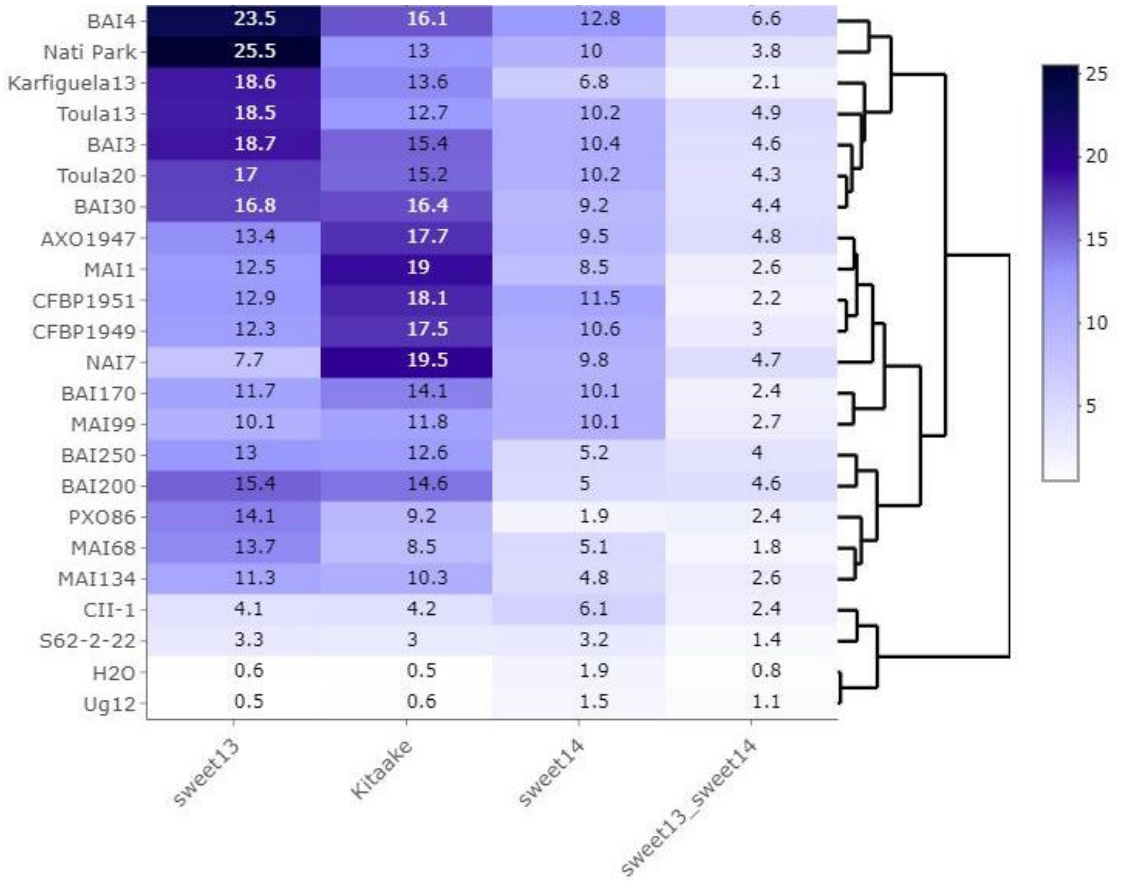
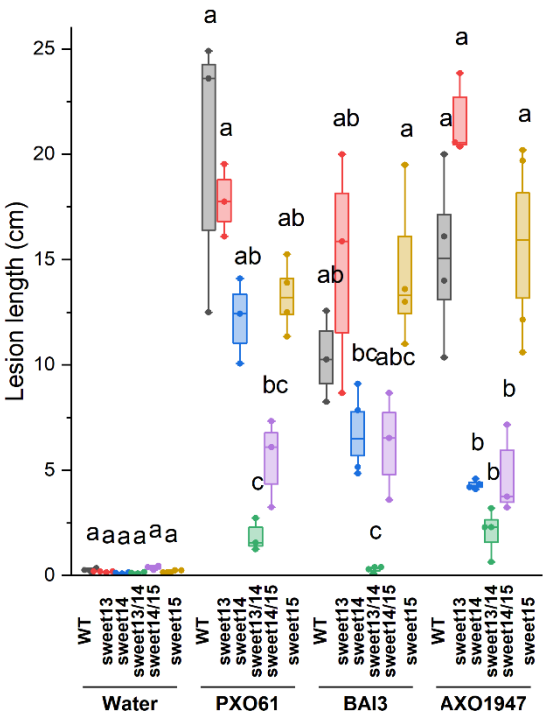


Figure 34 : Gain de résistance à large spectre contre les Xoo africaines par perte de fonction *sweet13 ; sweet14.* Lésions (en cm) obtenues par l'inoculation d'une vingtaine de *Xoo* représentant 8 pays Africains présents en collection à l'IRD sur Kitaake sauvage et mutants *sweet13*, *sweet14* et *sweet13* ; *sweet14*. PXO86 est une souche contrôle Asiatique nécessitant l'induction de *SWEET14* pour pouvoir proliférer dans la plante. Deux événements de mutations différents ont été testés pour chaque KO.

Figure 35 : Un double mutant *sweet14 ; sweet15* ne confère pas de résistance supplémentaire aux *Xoo* africaines. Longueurs de lésions (en cm) obtenues sur différents mutants dans les gènes *SWEET* du clade III par l'inoculation de deux souches africaines (BAI3 et AXO1947) et une souche contrôle Asiatique (PXO61). PXO61 est capable d'induire *SWEET13* et *SWEET14* via les TALEs PthXo2 et PthXo3 respectivement. Les données d'une expérience sont représentées. Les données présentées et leur représentation graphique ont été produites par nos partenaires à l'HHU en Allemagne.



2.1.3 Investigation d'un potentiel effet cumulatif de l'absence de *SWEETs* du clade III

Dans l'hypothèse où le phénotype de résistance du double mutant *sweet13* ; *sweet14* est dû à un effet cumulatif de la mutation de *SWEET14* et de n'importe quel autre SWEET du clade III, d'autres combinaison de mutants ont été testées. Les résultats obtenus par Eom *et al.* en 2019 [288] n'indiquent aucune différence entre le mutant simple pour *SWEET14* et un double mutant *sweet11* ; *sweet14* pour 3 souches africaines : BAI3 (Burkina Faso), MAI1 (Mali), et AXO1947 (Cameroun). D'autre part, des données obtenues par nos partenaires de l'Institut de Physiologie Moléculaire de l'Heinrich Heine University (HHU) montrent la même tendance sur un double mutant *sweet14* ; sweet15 (figure 35). Il semblerait donc que la résistance conférée au double mutant soit spécifiquement liée à la mutation combinée de *OsSWEET13* et *OsSWEET14*.

2.1.4 Confirmation du phénotype résistant par criblage bactérien à grande échelle.

Dans le but d'évaluer le spectre de résistance apporté par la mutation de *OsSWEET13* et *OsSWEET14* chez le double mutant en fond Kitaake, nous avons inoculé par leaf-clipping 324 *Xoo* originaires de 8 pays différents d'Afrique de l'Ouest : Burkina Faso (201), Mali (63), Sénégal (41), Niger (11), Bénin (3), Cameroun (2), Côte d'Ivoire (2) et Ouganda (1). Au final, aucune des souches testées n'a pu contourner la résistance apportée par la double mutation, aucune des lésions enregistrées ne dépassant 8 cm. La résistance à large spectre contre les *Xoo* d'Afrique de l'Ouest que prodiguent les mutations combinées de *OsSWEET13* et *OsSWEET14* suggère l'utilisation de SWEET13 par les souches africaines.

2.2 Recherche de voies de régulation bactériennes de *OsSWEET13*

2.2.1 Quantification relative de transcrits par PCR quantitative.

La voie classique d'interaction des gènes SWEET du riz avec *Xoo* est leur induction transcriptionelle via la fixation d'un TALE de la bactérie au niveau de l'EBE en amont du gène. Hors, l'induction de *SWEET13* par les *Xoo* africaines n'a encore jamais été mise en évidence (Streubel et a., 2013 [289] ; Blanvillain-Baufumé *et al.*, 2017 [294]). Nous avons donc cherché à confirmer les schémas d'induction des différents SWEETs du clade III par les *Xoo* africaines. Des quantifications relatives des transcrits par qPCR sur des échantillons de plantes sauvages infectées ou non par des *Xoo* africaines et asiatiques ont été réalisées (figure 36). Tout d'abord, une forte induction de *SWEET14* est bien observée lors de l'inoculation de souches contenant des TALEs majeurs ayant des EBEs dans le promoteur du gène. C'est le cas des souches africaines MAI1 (TalC, TalF), BAI3 et AXO1947 (TalC), ainsi que des souches asiatiques PXO86 (AvrXa7) et PXO61 (PthXo3). L'induction de *SWEET11* n'est détectable qu'en présence de la seule souche pouvant l'induire dans cet essai, PXO99, qui contient PthXo1. Dans le cas de *SWEET13*, l'augmentation des transcrits attendue par inoculation de la souche Asiatique contrôle PXO61 (PthXo2) est assez variable. De même, une induction de *SWEET13* par les souches BAI3, AXO1947 et MAI1 est observée ; plus ou moins robuste selon les essais (Figure 36). Dans l'optique de lever le doute sur ces inductions, des expérimentations équivalentes ont été menées dans les laboratoires de nos partenaires de HHU. Comme précédemment, les inductions de *SWEET13* sont peu reproductibles. D'autres amorces dessinées par nos partenaires pour la qPCR de suivi de *SWEET13* ont été testées (tableau S1). Ces amorces ont elles aussi donné des résultats peu robustes, même dans l'induction par la souche contrôle PXO61.

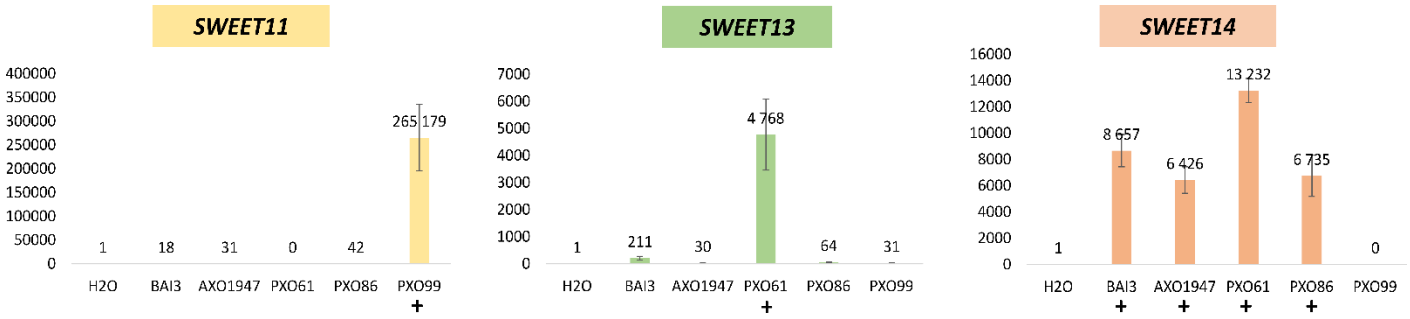


Figure 36 : Potentielle induction de *SWEET13* par les *Xoo* africaines. Niveaux relatifs de transcrits obtenus pour *SWEET11*, *SWEET13* et *SWEET14* sur des plantes infiltrées par des souches africaines et asiatiques de *Xoo*. Les souches avec un signe « + » en indice sont détentrices d'au moins un TALE pouvant induire le gène étudié. Les valeurs de Ct de chaque échantillon ont été normalisées par le niveau d'expression du gène EF1α, puis chaque niveau d'expression a été rapporté à l'eau (ΔΔCt). Inflitrations à DO_{600} = 0.5, récolte à 48 hpi.

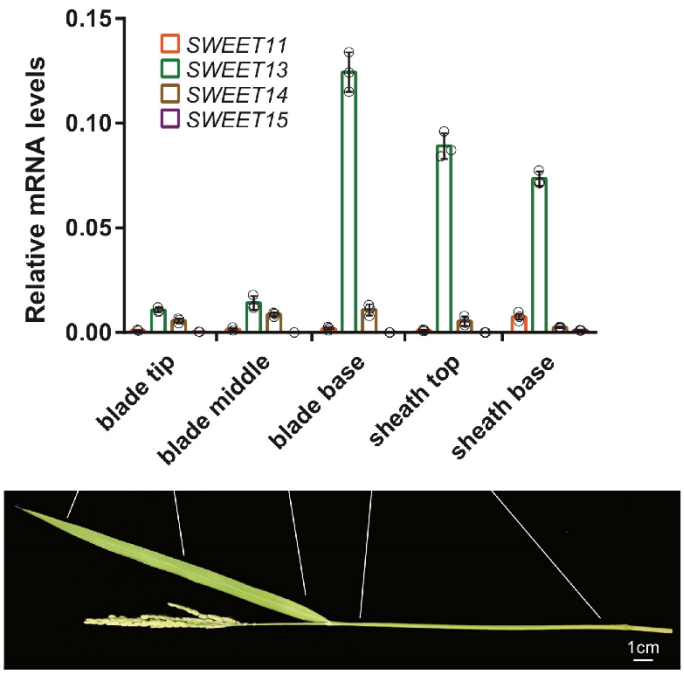


Figure 37 : Niveaux de transcrits des gènes *SWEET* du clade III dans les feuilles de riz non infectées. Niveaux d'ARNm relatifs (RT-PCR quantitative) de *SWEET11*, *SWEET13*, *SWEET14* et *SWEET15* dans différentes régions des parties aériennes de la plante. 3 réplicas biologiques issus de 3 plantes différentes ont été traités. L'expression des différents SWEET est normalisée par rapport au niveau de transcrits de l'Ubiquitine1 du riz. 3 expériences ont été réalisées avec des résultats similaires. Figure « Supplementary Figure 4 » de Eom *et al.*,2019 [288].

Les résultats obtenus n'excluent donc pas l'induction de *SWEET13* par les souches de *Xoo* africaines lors de l'infection. En temps normal, les gènes *SWEET11* et *SWEET14* sont très faiblement exprimés dans les tissus végétaux. En comparaison, le niveau basal d'expression de *SWEET13* est bien plus élevé, notamment dans les feuilles et la tige de la plante (figure 37). La sensibilité nécessaire à la détection de l'induction de ce gène doit donc être bien plus forte que lors du suivi de SWEET11 ou *SWEET14*, où l'on se retrouve dans une situation présence de bactéries/induction, absence de bactéries/pas d'induction. Ceci soulève un problème de sensibilité de la méthode de qPCR employée, qui nécessite un approfondissement de la mise au point pour la détection spécifique de *SWEET13*. Une étude des profils d'induction obtenus en présence de *Xoo* africaines par d'autres techniques plus sensibles, comme la PCR numérique, est souhaitable.

2.2.2 Investigation d'une potentielle régulation post-transcriptionnelle des SWEETs par les *Xoo* africaines.

Dans le cas où les *Xoo* africaines seraient capables de moduler la quantité de protéines *SWEET13* autrement que par l'induction de sa transcription et afin d'éventuellement de se soustraire aux biais de la qPCR, nous nous sommes tournés vers des méthodes de quantification des protéines SWEET produites en présence et en absence de bactéries. Les *Xoo* africaines pourraient par exemple, favoriser la stabilité et la traduction des transcrits *SWEET13*. Les mécanismes de régulation post-transcriptionelle de ce type sont variés, il pourrait s'agir d'une répression de l'export de micro ARNs dont le rôle est de réprimer la traduction des *SWEETs*, ou bien d'une amélioration de la protection des transcrits *SWEET* contre la dégradation.

La première méthode retenue constitue une voie de quantification indirecte des SWEET par l'utilisation de plantes transgéniques. Ces lignées rapportrices, construites par Eom *et al.* en 2019 [288] dans le cadre du projet « Transformative Strategies for Controlling Rice Diseases » Phase II, contiennent un plasmide p*SWEET*:g*SWEET-GUSplus* codant pour des protéines de fusion traductionnelle entre les sous unités SWEET et la β-glucuronidase (GUS) sous le contrôle du promoteur natif du *SWEET* étudié. La mesure de l'activité GUS de ces protéines fusionnées permet ainsi la quantification relative d'une accumulation potentielle des protéines SWEET en présence de bactéries. L'activité mesurée ici correspond à la production d'un composé fluorescent par la GUS via l'hydrolyse d'un substrat particulier, le 4-Methylumbelliferylβ-D-glucuronide (MUG). Cette quantification est effectuée après l'infiltration de *Xoo* africaines sauvages ou mutantes en comparaison avec des souches asiatiques (figure 38). La méthode offre une détection claire de l'induction de la production de sous unités SWEET11 par PXO99 (PthXo1), de même, l'accumulation de sous-unités SWEET14 est facilement observable lors de l'inoculation des différentes souches susceptibles de l'induire (*talC* africain ou PthXo3 asiatique). En revanche, il est encore une fois difficile de conclure sur les résultats obtenus avec les lignées transgéniques p*SWEET13*:g*SWEET13*-GUSplus. En effet, même les activités mesurées en présence de PXO61 (souche contrôle contenant PthXo2, inducteur de *SWEET13*), ne sont pas robustes. La variabilité inter réplicas biologiques est très élevée. Il s'est avéré que cette variabilité est due à une hétérozygotie du matériel biologique utilisé. Il n'est donc pas possible à ce jour de conclure sur les taux de protéines *SWEET13* relatifs engendrés au cours de l'infection par nos différentes souches de *Xoo*. Afin de poursuivre ces tests, l'obtention de graines homozygotes pour cette construction est primordiale.

Un protocole d'appréciation directe de la quantité de protéines SWEET accumulées au cours de l'infection de plantes mutantes ou sauvage est en cours de mise au point. Ce protocole repose sur l'utilisation d'anticorps anti-SWEET produit par nos partenaires de HHU. Bien que ces anticorps permettent une détection sur des extraits d'ovocytes (Eom *et al.*, 2019 [288]), nos premiers essais sur extraits de protéines totales de feuilles ne permettent pas de visualiser nos transporteurs.

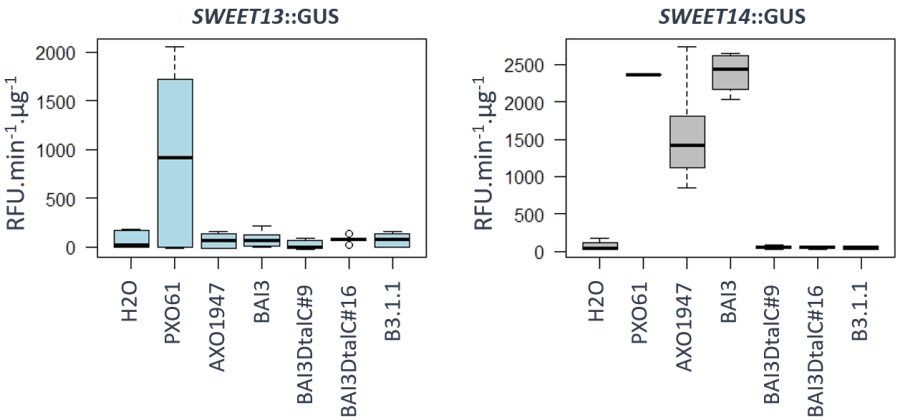


Figure 38 : Investigation d'un potentielle accumulation de protéines *SWEET13* en présence de *Xoo* africaines. Activité en RFU par minute par microgrammes de protéines totales d'échantillons de plantes *pSWEET*:g*SWEET*-GUSplus infiltrés par des souches mutantes et sauvages de *Xoo*. Inflitrations à DO_{600} = 0.5, récolte à 48 hpi.

Tableau 11 : Liste des gènes présents dans la région *RegtalC* et leurs annotations respectives. Les gènes en gras sont des gènes candidats intéressants dans la recherche de déterminants de virulence potentiellement impliqués dans SWIIVA présents dans cette région.

Locus	Position	Type de locus	ID protéine	Produit
gene-C0L90_02850	CP025610.1:581093-581428:+	CDS	NA	hypothetical protein
gene-C0L90_02855	CP025610.1:581547-582923:+	CDS	AVU01659.1	IS5/IS1182 family transposase
gene-C0L90_02860	CP025610.1:582943-583695:+	CDS	NA	hypothetical protein
gene-C0L90_02865	CP025610.1:583828-584790:-	CDS	AVU01660.1	IS1595 family transposase
gene-C0L90_02870	CP025610.1:584883-585035:+	CDS	NA	PAAR domain-containing protein
gene-C0L90_02875	CP025610.1:585183-585832:-	CDS	NA	IS5/IS1182 family transposase
gene-C0L90_02880	**CP025610.1:585939-589763:+**	**CDS**	**AVU01661.1**	**TalC (avirulence protein)**
gene-C0L90_02885	CP025610.1:589856-590666:+	CDS	NA	IS3 family transposase
gene-C0L90_02890	CP025610.1:590833-592416:-	CDS	AVU01662.1	oxidoreductase
gene-C0L90_02895	CP025610.1:592656-592731:-	ncRNA	NA	sX9 sRNA
id-C0L90_02895	CP025610.1:592656-592731:-	ncRNA	NA	sX9 sRNA
gene-C0L90_02900	CP025610.1:592807-592998:-	CDS	AVU01663.1	hypothetical protein
gene-C0L90_02905	CP025610.1:593010-593183:-	CDS	NA	IS5/IS1182 family transposase
gene-C0L90_02910	CP025610.1:593356-595302:+	CDS	NA	peptidase S9
gene-C0L90_02915	CP025610.1:595362-595428:+	ncRNA	NA	sX9 sRNA
id-C0L90_02915	CP025610.1:595362-595428:+	ncRNA	NA	sX9 sRNA
gene-C0L90_02920	CP025610.1:595499-597823:-	CDS	AVU01664.1	S9 family peptidase
gene-C0L90_02925	CP025610.1:598244-599667:+	CDS	NA	hypothetical protein
gene-C0L90_02930	CP025610.1:599727-600818:-	CDS	AVU01665.1	DUF1615 domain-containing protein
gene-C0L90_02935	CP025610.1:601202-601278:-	ncRNA	NA	sX9 sRNA
id-C0L90_02935	CP025610.1:601202-601278:-	ncRNA	NA	sX9 sRNA
gene-C0L90_02940	CP025610.1:601783-602661:-	CDS	AVU04587.1	SAM-dependent methyltransferase
gene-C0L90_02945	CP025610.1:602850-603734:+	CDS	AVU01666.1	EamA/RhaT family transporter
gene-C0L90_02950	**CP025610.1:604012-605784:-**	**CDS**	**AVU01667.1**	**cellulase**
gene-C0L90_02955	**CP025610.1:606182-607882:-**	**CDS**	**AVU01668.1**	**CbsA (1,4-beta-cellobiosidase)**
gene-C0L90_02960	CP025610.1:608030-608361:+	CDS	NA	hypothetical protein
gene-C0L90_02965	CP025610.1:609037-610413:+	CDS	AVU01669.1	D-serine/D-alanine/glycine transporter
gene-C0L90_02970	**CP025610.1:610500-611495:-**	**CDS**	**AVU04588.1**	**XopV (type III secretion system effector protein)**
gene-C0L90_02975	CP025610.1:611741-612652:+	CDS	AVU01670.1	magnesium transporter
gene-C0L90_02980	CP025610.1:612856-613067:-	CDS	NA	hypothetical protein
gene-C0L90_02985	CP025610.1:613187-613471:-	CDS	AVU01671.1	hypothetical protein
gene-C0L90_02990	CP025610.1:613801-614247:+	CDS	AVU01672.1	autotransporter

3 Identification des déterminants bactériens à l'origine de la capacité des *Xoo* africaines à maintenir de la virulence sans l'induction de *SWEET14*

3.1 Etude d'une région génomique contenant *talC*

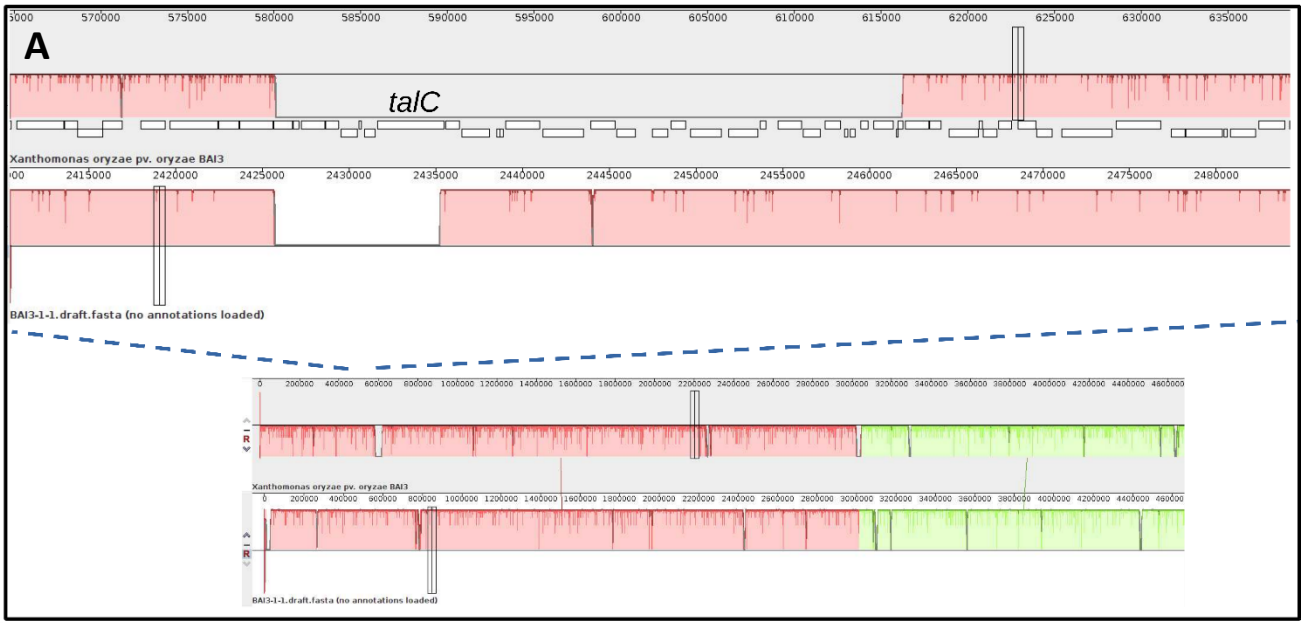
3.1.1 Découverte et implications

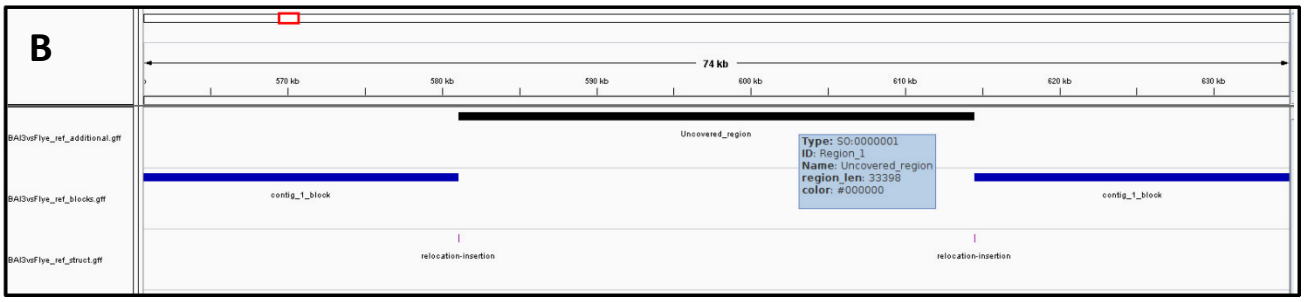
L'étude d'un mutant d'insertion du plasmide pSM7 dans le gène *talC* de la souche Burkinabée BAI3 par Yu *et al.* [292] a révélé qu'elle avait perdu toute capacité d'infection, tous génotypes confondus. Cependant, le mutant corollaire de l'EBE *talC* chez l'hôte ne présentait pas de résistance à la bactérie (Blanvillain-Baufumé *et al.*, 2017[294]). Ces résultats étaient inattendus, car si la contribution de *talC* au pouvoir pathogène des souches de *Xoo* africaines est entièrement *OsSWEET14*-dépendant, tel que le suggère son effet sur l'induction de ce gène, l'édition de l'EBE TalC sur le promoteur de *OsSWEET14* aurait dû conférer de la résistance à la plante. C'est en tout cas ce que l'on observe dans l'ingénierie de la résistance aux souches asiatiques (Li *et al.*, 2014 ; Oliva *et al.*, 2019 [287]).

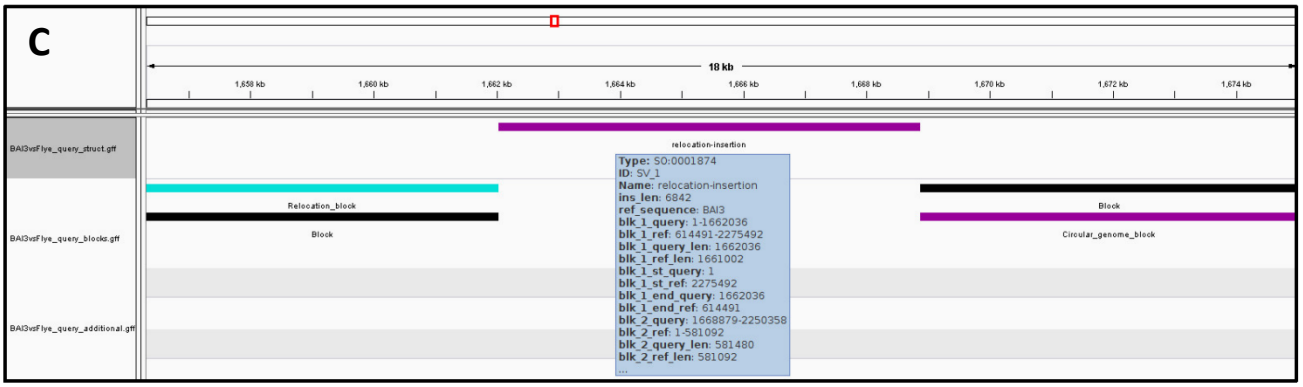
Pour tenter d'expliquer ces résultats, une étude à laquelle j'ai participé a exploré l'hypothèse selon laquelle *talC* aurait au moins une autre cible transcriptionnelle génétiquement redondante avec *OsSWEET14* dans le génome du riz (voir annexe 4).

Au cours de ce travail de thèse, nous nous sommes intéressés à une autre hypothèse, concernant la souche BAI3.1.1, décrite précédemment dans l'équipe comme un mutant d'insertion dans le gène *talC* obtenu à l'aide du plasmide suicide psM7 (Yu *et al.*, 2011 [292]). Afin de vérifier la nature exacte de cet évènement, nous avons séquencé le génome de BAI3.1.1 à l'aide de la technologie Oxford Nanopore (ONT). Les *reads* obtenus ont été assemblés à l'aide du logiciel Flye pour aboutir à un contig unique s'alignant correctement sur le génome de BAI3 (Figure 38.A). Cependant, la comparaison des séquences du génome de BAI3 (GenBank: CP025610.1) et du génome de BAI3.1.1 au niveau de la région où se trouve le gène *talC* n'est pas cohérente avec l'insertion de pSM7 à l'intérieur de la séquence du gène. L'alignement entre ces deux séquences fait ressortir que plusieurs dizaines de gènes annoté dans la région sur le génome BAI3 (y compris *talC*) ne s'alignent pas avec la séquence BAI3.1.1 (Figure 39.A). Une caractérisation plus fine des différences structurales entre ces deux génomes, réalisée à l'aide du logiciel NucDiff fait ressortir, qu'une région de ~33kb contenant *talC* et que nous avons nommé Reg*talC* est manquante sur le génome de BAI3.1.1 (Figure 39.B). A sa place, se trouve une séquence de ~6,3kb (Figure 39.C) dont la nature est probablement plasmidique (Figure 39.D). Nous n'avons pour l'instant pas cherché à reconstituer par quelle série d'évènements la configuration observée dans le génome de BAI3.1.1 est apparue. Il n'est cependant pas exclu que les nombreux éléments génétiques mobiles de la région (voir tableau 11) interviennent dans ce processus.

Cette découverte indique cependant que le phénotype de la souche BAI3.1.1 initialement attribué à l'inactivation de *talC* seul, est potentiellement la conséquence de la délétion involontaire de plusieurs gènes dans la région. Ceci étant d'autant plus crédible que la complémentation de la virulence de BAI3.1.1 à l'aide d'un plasmide exprimant *talC* en *trans* n'est que partielle (Yu *et al.*, 2011 [292]). Cette découverte soulève donc la question du phénotype d'un véritable mutant ponctuel pour le gène *talC*, et l'éventuel rôle des autres séquences de la région délétée dans le mutant BAI3.1.1.


A
talC
Xanthomonas oryzae pv. oryzae BAI3
BAI3-1-1.draft.fasta (no annotations loaded)


B
74 kb
BAI3vsFlye_ref_additional.gff
BAI3vsFlye_ref_blocks.gff
BAI3vsFlye_ref_struct.gff
Uncovered_region
Type: SO:0000001
ID: Region_1
Name: Uncovered_region
region_len: 33398
color: #000000
contig_1_block
relocation-insertion


C
18 kb
BAI3vsFlye_query_struct.gff
BAI3vsFlye_query_blocks.gff
BAI3vsFlye_query_additional.gff
relocation-insertion
Type: SO:0001874
ID: SV_1
Name: relocation-insertion
ins_len: 6842
ref_sequence: BAI3
blk_1_query: 1-1662036
blk_1_ref: 614491-2275492
blk_1_query_len: 1662036
blk_1_ref_len: 1661002
blk_1_st_query: 1
blk_1_st_ref: 2275492
blk_1_end_query: 1662036
blk_1_end_ref: 614491
blk_2_query: 1668879-2250358
blk_2_ref: 1-581092
blk_2_query_len: 581480
blk_2_ref_len: 581092
Relocation_block
Block
Circular_genome_block

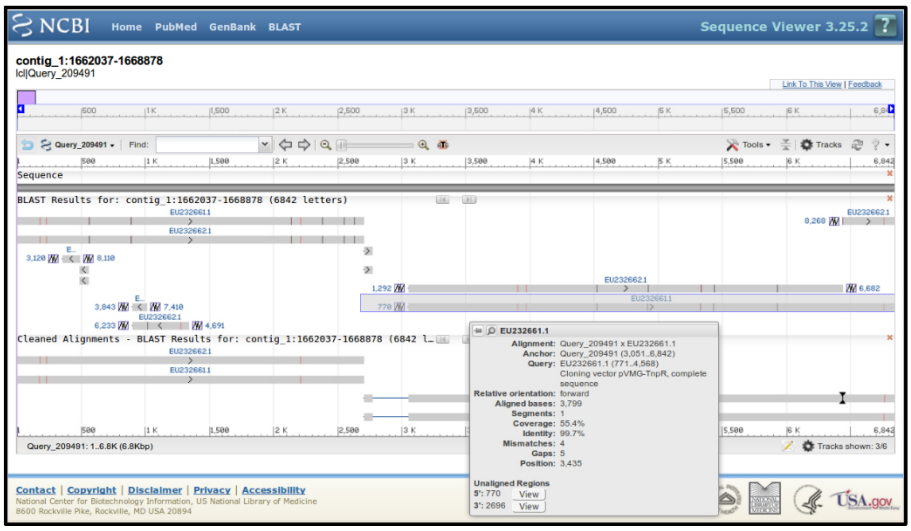

D
NCBI
Home
PubMed
GenBank
BLAST
Sequence Viewer 3.25.2
contig_1:1662037-1668878
lcl|Query_209491
Sequence
BLAST Results for: contig_1:1662037-1668878 (6842 letters)
Cleaned Alignments - BLAST Results for: contig_1:1662037-1668878 (6842 l...
EU232661.1
Alignment: Query_209491 x EU232661.1
Anchor: Query_209491 (3,051..6,842)
Query: EU232661.1 (771..4,568)
Cloning vector pVMG-TnpR, complete sequence
Relative orientation: forward
Aligned bases: 3,799
Segments: 1
Coverage: 55.4%
Identity: 99.7%
Mismatches: 4
Gaps: 5
Position: 3,435
Unaligned Regions
5': 770 View
3': 2696 View
Contact | Copyright | Disclaimer | Privacy | Accessibility
National Center for Biotechnology Information, US National Library of Medicine
8600 Rockville Pike, Rockville, MD USA 20894
USA.gov

Figure 39 : Comparaison des séquences du génome de référence BAI3 et du génome du mutant BAI3.1.1 au niveau de la région *talC*. A. Alignement des séquences BAI3 et BAI3.1.1 obtenu avec le logiciel Mauve (https://darlinglab.org/mauve/mauve.html) au niveau de *RegtalC*. Les régions conservées entre les deux génomes sont repérées par le profil de conservation ombré en rouge. Les coordonnées en nucléotides sont indiquées pour chacun des génomes. B et C. Analyse des variations structurales entre les génomes BAI3 et BAI3.1.1 au niveau de *RegtalC* à l'aide de l'outil NucDiff. B. identification grâce à NucDiff sur la référence du génome BAI3, d'une région de 33.398pb qui n'est pas couverte par des séquences dans le génome de BAI3.1.1. Corollairement, en C, NucDiff identifie une région de 6.842pb du génome BAI3.1.1 qui n'est pas couverte par des séquences dans le génome de BAI3. D. Alignement des hits BlastN sur la région de 6.842pb du génome BAI3.1.1 dont la séquence a été utilisée comme requête dans la base nr du NCBI.

3.1.2 Un mutant ponctuel dans le gène *talC* est moins affecté que le mutant BAI3.1.1 dans sa virulence

Dans le but de mettre au clair l'influence de *talC* seul dans la virulence des *Xoo* africaines, nous avons cherché à comparer les profils de virulence obtenus par des mutants ponctuels pour ce gène versus le mutant délété pour la région *talC* BAI3.1.1. Un criblage par southern blot d'une banque de mutants TALEs de BAI3 créée par Tran *et al.*, en 2018 [291] a permis de détecter deux premiers mutants dépourvus de *talC* : BAI3Δ*talC*#9 et BAI3Δ*talC*#16. Le séquençage ONT des mutants BAI3Δ*talC*#9 et BAI3Δ*talC*#16 générés confirme l'insertion de pSM7 dans *talC*. La caractérisation des différents mutants BAI3 sur riz japonica (Kitaake et Nipponbare) montre que la mutation ponctuelle de *talC* ou de Reg*talC* abolit l'induction de *OsSWEET14* chez l'hôte (Figure 40.A). Il existe un différentiel net dans les longueurs de lésions entre les mutants ponctuels, et le mutant Reg*talC* (figure 40.B). Les mutants ponctuels produisent des lésions bien plus courtes que le sauvage, aux alentours de 10 cm, mais plus longues que le mutant Reg*talC* (maximum 5 cm). Parmi d'autres hypothèses, ce différentiel laisse envisager que la région absente chez ce mutant contribue à la virulence des souches africaines de *Xoo*.

3.1.3 Dans la région *talC*, des facteurs autres que *talC* contribuent à la virulence de BAI3.

Dans le but de confirmer que la perte sévère de virulence du mutant BAI3.1.1 par rapport à un mutant *talC* ponctuel provient bien de l'absence de la région contenant *talC* (Reg*talC*) et non pas d'une autre différence ailleurs dans le génome (voir raisonnement en figure S1), nous avons reproduit la délétion de ~33kb par transformation d'un plasmide suicide basé sur le vecteur pK18mobsacB, permettant de sélectionner une recombinaison sur chacune des régions encadrant Reg*talC*. Les mutants nouvellement créés présentent le même phénotype non-virulent que le premier mutant BAI3.1.1 (figure 41). Ces données indiquent qu'il s'agit bien de la délétion de Reg*talC* dans BAI3.1.1 et non pas un autre polymorphisme dans le génome qui est responsable de la perte de virulence observée. Cette région semble donc avoir un rôle prépondérant dans la virulence de BAI3, allant au-delà du seul rôle de *talC*.

Nous avons par conséquent examiné les annotations en gènes de cette région de ~33kb qui contient 32 loci annotés par le NCBI, dont 6 loci classés comme ARN non-codants et similaires à des petits ARN précédemment décrits chez *X. campestris* pv. *vesicatoria* (Schmidtke *et al.*, 2012 [295]) et 26 loci prédits comme codant pour des protéines (tableau 11). Afin d'identifier les facteurs déterminants de Reg*talC* dans la virulence résiduelle de BAI3, nous souhaitons procéder à la ré-introduction de fragments de cette région dans le mutant ΔReg*talC*, également en combinaison avec *talC*. Notre première approche, avec a priori, nous a amené à rechercher des candidats prioritaires au sein de Reg*talC*. Parmi les gènes

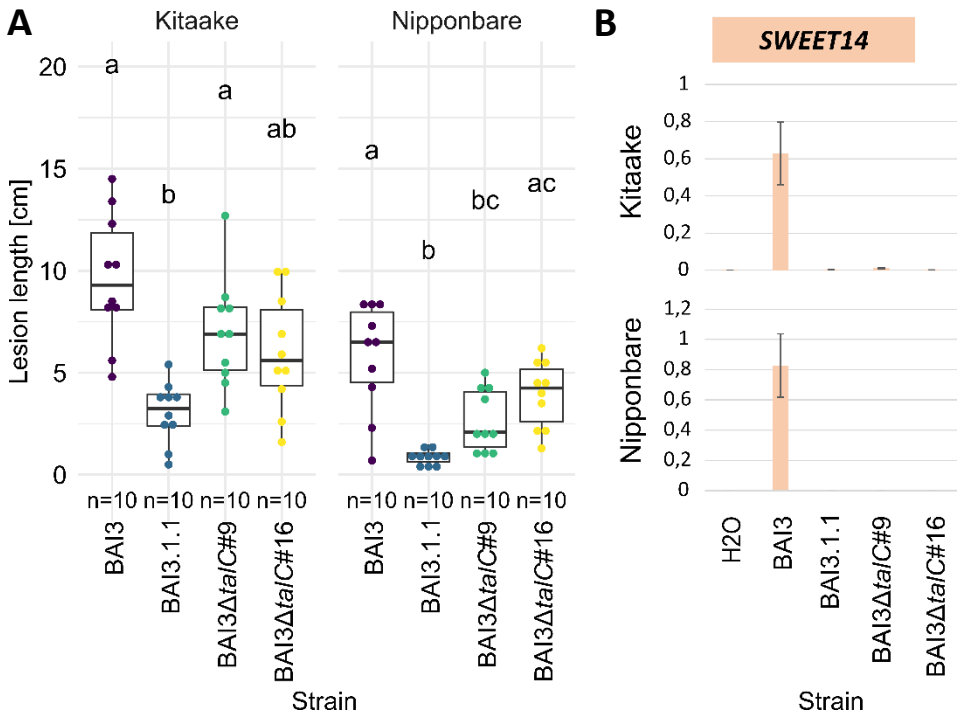


Figure 40 : La région délétée dans le mutant BAI3.1.1 a un impact dans la virulence de BAI3. A. Longueurs de lésions (en cm) obtenues sur deux variétés de riz *japonica* sauvages, Kitaake et Nipponbare, suite à l'inoculation par leaf-clipping de BAI3, le mutant pour la région *talC* (BAI3.1.1), et deux mutants ponctuels pour *talC* (BAI3Δ*talC* # 9 et #16). Pour l'explication complète du boxplot, voir la partie « Matériel et méthodes ». B. Niveaux relatifs de transcrits *SWEET14* après l'infiltration des mêmes génotypes par les souches citées précédemment. Les valeurs de Ct de chaque échantillon ont été normalisées par le niveau d'expression du gène EF1α (ΔCt). Les Inflitrations ont été réaliseés à une DO_{600} = 0.5 et la récolte des échantillons à 48 hpi. En A et en B, les données d'une expérimentation sont représentées. Ces expériences ont été réalisées une fois sur Nipponbare et reproduites de nombreuses fois (N>5) sur Kitaake, avec des résultats similaires.

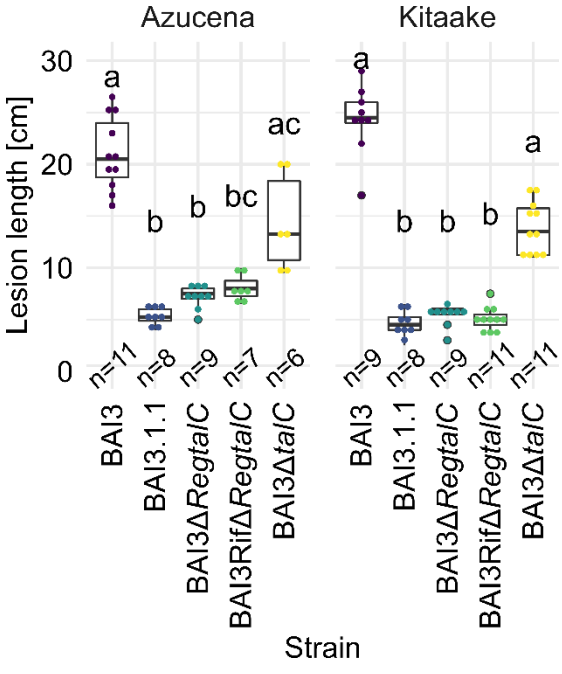


Figure 41 : Validation du phénotype produit par la délétion de *RegtalC*. La délétion de *RegtalC* a été reproduite dans deux fonds génétiques différents, BAI3 sauvage et un BAI3 spontanément résistant à la rifampicine (fond génétique originel du mutant BAI3.1.1). Les deux nouveaux mutants *RegtalC* ont été phénotypés par leaf-clipping en comparaison à la souche sauvage BAI3, au mutant BAI3.1.1, et à un mutant ponctuel pour *talC* BAI3Δ*talC*. Les leaf-clippings ont étés réalisés à DO600=0.2, et les longueurs de lésions ont étés enregistrés 2 semaines après inoculation. Les résultats d'une expérience sont représentés. Cette expérience a été reproduite 5 fois sur Kitaake, avec des résultats similaires. Pour l'explication complète du boxplot, voir la partie « Matériel et méthodes ».

annotés, on trouve celui d'un effecteur de type III (C0L90_02970) homologue à *xopV*. Bien qu'il soit conservé chez *Xoo* et retrouvé chez différentes espèces de Xanthomonas, il n'a pas été montré à ce jour que XopV ait un rôle propre dans la virulence. Un mutant individuel pour *xopV* dans la souche de *Xoo* Asiatique PXO99A n'est pas affecté dans sa capacité à induire de la maladie sur des variétés de riz sauvages. Une réduction significative de la longueur des lésions sur Kitaake est néanmoins observée avec un triple mutant xopV, xopZ (deux copies) et xopN. Dans l'ensemble, il semblerait que XopV soit un contributeur redondant à la virulence de *Xoo* (Furutani *et al.*, 2009 [1] ; Song et Yang, 2010 [1] ; Long *et al.*, 2018 [1]). Un autre locus d'intérêt (C0L90_02955), dont l'importance dans la colonisation des tissus de l'hôte a été démontrée chez *Xoo*, est *cbsA* qui code pour une cellobiohydrolase potentiellement impliquées dans la dégradation de la paroi végétale et l'élicitation de l'immunité de l'hôte (Tayi *et al.*, 2017 [173]) et qui a été identifié comme un facteur de spécificité tissulaire déterminant la colonisation des vaisseaux du xylème (Gluck-Thaler *et al.*, 2020 [296]). A ses côtés, le locus C0L90_02950 code pour une cellulase putative, avec donc, là aussi, une fonction potentielle dans la dégradation des polysaccharides de la paroi végétale libérant sucres simples dans l'apoplaste, fournissant un apport énergétique à la multiplication bactérienne.

Comme nous disposions de la construction miniTn*7*::*cbsA*$_{Xoo}$ utilisée par Gluck-Thaler *et al.* [296] pour la complémentation génétique de cbsA chez *Xoo*, ce transposon a été introduit (en parallèle avec le transposon vide miniTn*7*) dans les mutants *RegtalC.* Les tests de virulence sur Kitaake des souches dérivées correspondantes ne font pas apparaître de gain en termes de longueurs de lésion par rapport aux contrôles (figure 42). Ceci laisse donc supposer que la contribution de *cbsA* est potentiellement dépendante d'un autre facteur génétique présent sur la région, notamment *talC*.

3.1.4 Analyse génétique de l'implication de *talC* et Reg*talC* dans la virulence spécifique des *Xoo* africaines et leur dépendance aux gènes *OsSWEET* du clade III.

Nous avons voulu évaluer le rôle de *talC* seul et celui de la région lui étant associée dans le phénotype de virulence indépendant de l'activité de *OsSWEET14* propre aux *Xoo* africaines (SWIIVA). Dans ce but, nous avons inoculé les lignées Kitaake sauvage, *sweet13*, *sweet14* et *sweet13* ; *sweet14* avec les mutants BAI3.1.1 ou Δ*talC*, transformés ou non avec un plasmide exprimant *talC* (ou le vecteur vide correspondant). Les souches sauvages de référence PXO99 (Philippines) et BAI3 ont été inoculées en parallèle comme contrôles (Figure 43).

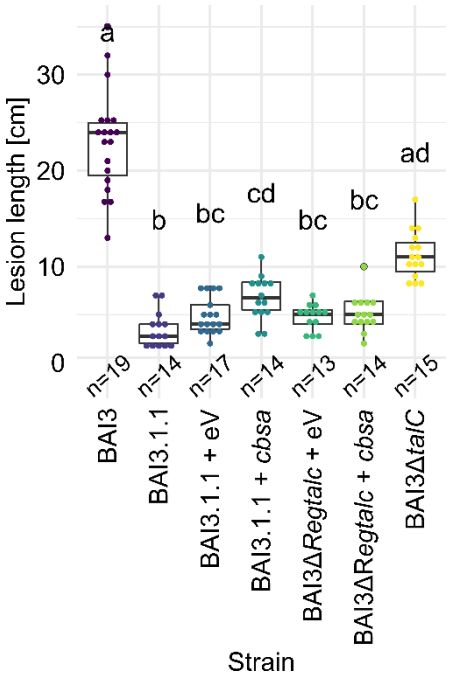
Comme observé précédemment, toutes les souches sont contrôlées par un double mutant *sweet13* ; *sweet14*.

Sur les autres génotypes, les lésions obtenues par l'inoculation des mutants ponctuels *talC* confirment l'importance de la région bordant ce gène dans la virulence des *Xoo* africaines. En effet, ces mutants ponctuels produisent des symptômes deux à trois fois plus longs que le mutant ΔReg*talC* (figure 43).

On remarque que *talC* semble être d'une certaine manière impliqué dans le phénotype de virulence résiduelle des *Xoo* africaines, car une diminution significative de la longueur des lésions est visible pour les mutants BAI3Δ*talC* par rapport à la souche sauvage sur le KO *sweet14*. Dans le même ordre d'idée, la complémentation du mutant BAI3Δ*talC* par p*talC* ramène les longueurs de lésions à des valeurs équivalentes ou supérieures à BAI3 sauvage sur tous les génotypes, *sweet14* compris (figure 43). Cette observation met en lumière l'implication de *talC* dans le phénotype SWIIVA.

En fond ΔReg*talC* (mutant BAI3.1.1), la complémentation par *talC* est très partielle, comme déjà observé précédemment par Blanvillain-Baufumé *et al.* en 2017[1]. Cette complémentation partielle indique que des déterminants génétiques de la région, autres que *talC*, sont requis pour la virulence.

Figure 42 : La complémentation par *cbsA* seul des mutants Reg*talC* ne procure pas de gain de virulence. Les mutants Reg*talC* (BAI3.1.1 et le nouveaux mutant en fond sauvage) ont été complémentés par *cbsA*. Les différentes souches complémentées par l'enzyme, ou un vecteur vide, ont été phénotypées par leaf-clipping en comparaison à la souche sauvage et au mutant ponctuel pour *talC* BAI3Δ*talC*. Les leaf-clippings ont étés réalisés à DO600=0.2, et les longueurs de lésions ont étés mesurées 2 semaines après inoculation. Les résultats de deux expériences sont représentés. Cette expérience a été reproduite 4 fois avec des résultats similaires. Pour l'explication complète du boxplot, voir la partie « Matériel et méthodes ».



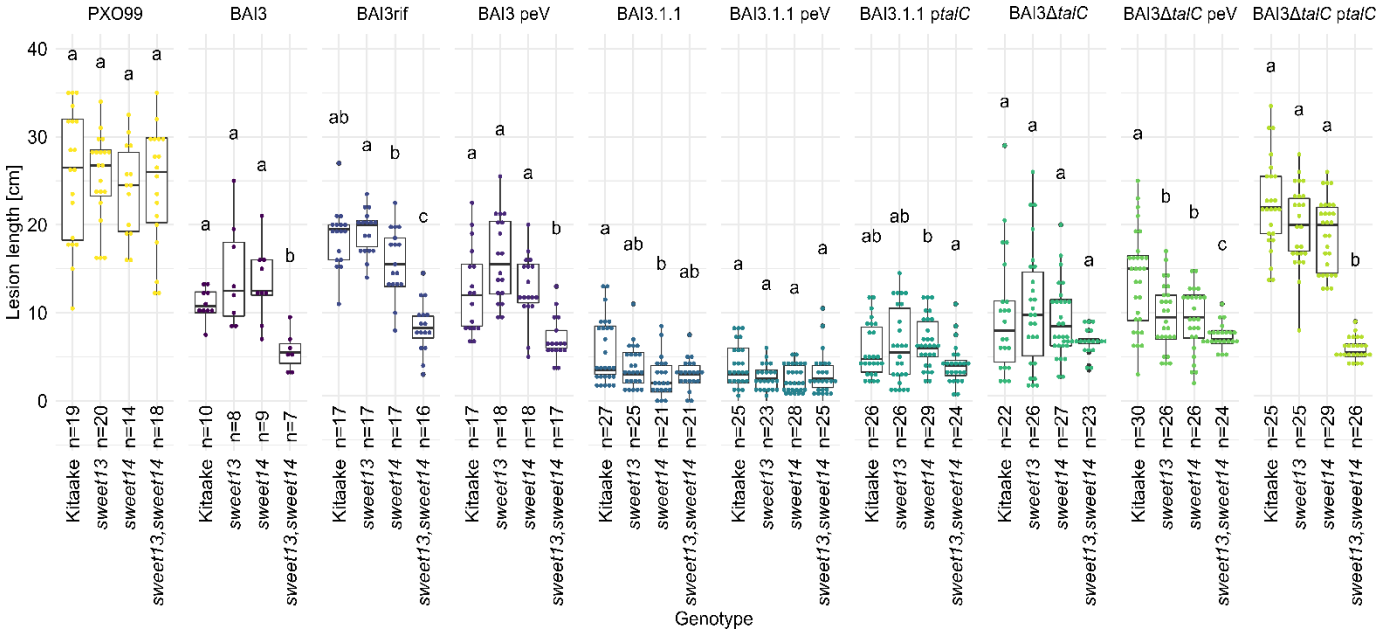


Figure 43 : Mise en évidence d'un rôle potentiel de la région *talC* dans le phénotype de virulence Africain. Longueur des lésions (en cm) obtenues après l'inoculation des mutants BAI3ΔReg*talC* et BAI3Δ*talC* complémentés par pSKX1 (eV) ou pSKX1-*talC* sur Kitaake sauvage et mutants *sweet13*, *sweet14* et *sweet13* ; *sweet14.* Inoculations par technique leaf-clipping à DO_{600} = 0.2, lecture des symptômes 14 dpi, 3 expériences sont représentées, environ 25 lectures par traitement.

Se pose alors la question de savoir si les fonctions associées à ces déterminants contribuent exclusivement à l'activité *talC* (synergie) ou ont un effet global sur le potentiel infectieux (additivité).

Enfin, au vu de la résistance obtenue pour un double mutant *sweet13* ; *sweet14*, nous aurions pu nous attendre à un phénotype similaire lors de l'inoculation des souches Δ*talC* (incapables d'induire *OsSWEET14*) sur un KO *sweet13*, hors ce n'est pas le cas (figure 43). Il semblerait donc que l'activité du gène *OsSWEET13* seule ne permette pas le phénotype SWIIVA.

3.2 Evaluation des performances du métabolisme du sucre chez les *Xoo* africaines

Etant donné qu'aucune induction d'un autre *SWEET* du clade III ne soit détectée de manière robuste en présence de *Xoo* africaines, une autre hypothèse serait que ces souches auraient la capacité de métaboliser certains substrats inutilisables par les *Xoo* asiatiques (figure S1). Nous avons donc entrepris d'étudier les capacités de dégradations d'un panel de 9 souches, comprenant des *Xoo* et *Xoc* asiatiques ou africaines, ainsi que le mutant BAI3.1.1. En partenariat avec l'équipe Emersys de l'Institut de Recherche en Horticulture et Semences (INRAE de Beaucouzé, France), nous avons testé ce panel sur une sélection de substrats en plaques Phenotype Microarray (PM) de Biolog®. Nous nous sommes focalisés sur le suivi de la réduction au cours du temps des 192 sucres des plaques PM1 et PM2 . Après 72h de culture, les différences dans les profils de réduction des souches testées nous permettent d'établir un groupement net selon le pathovar et l'origine de la souche considérée (figure 44). Les *Xoo* africaines forment un cluster aux profils de dégradation plus proches des *Xoc* que des *Xoo* asiatiques. Cette parenté a déjà été mise en évidence par de précédentes données de typage moléculaire MLVA (Poulin *et al.* en 2015 [285]), d'ANI (Average Nucleotide Identity), d'une phylogénie SNP ainsi que d'une phylogénie basée sur des variations dans 33 gènes de ménage (Lang *et al.* en 2019 [286]).

Les profils redox enregistrés nous apprennent que les *Xoc* ont la plus grande diversité de dégradation de substrats, suivies des *Xoo* asiatiques et enfin des *Xoo* africaines (figure 44). D'autre part, on apprend que la souche mutante non virulente BAI3.1.1 ne semble pas altérée dans son métabolisme des sucres. A première vue, nous n'avons pas détecté de carbohydrate spécifiquement dégradé par les *Xoo* africaines. Il semblerait donc que les *Xoo* africaines et asiatiques disposent d'un arsenal d'enzymes commun, ou du moins avec des fonctions conservées. Néanmoins, une analyse plus approfondie de ces données métaboliques permettra peut-être de mettre en évidence pour certains substrats des efficacités de réduction contrastées selon l'origine des souches.

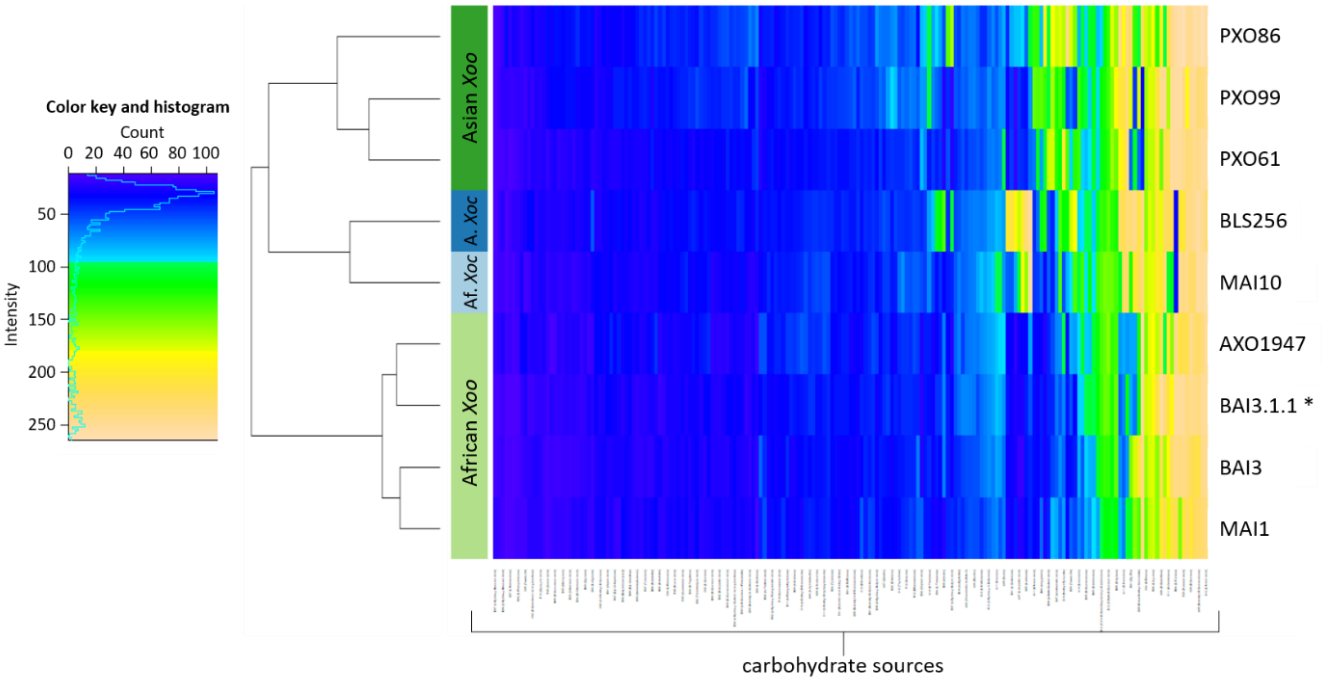


Figure 44 : Métabolisme différentiel des sucres chez *Xanthomonas oryzae.* Heatmap montrant l'intensité de la couleur obtenue par réaction au cristal violet au cours de l'incubation des différentes souches test dans un milieu minimal en plaques contenant 2x96 carbohydrates différents. La production de couleur violette correspond à une dégradation du substrat présent dans le puit. L'intensité de la production de pigment est représentée par l'échelle de couleur à gauche : bleu = substrats qui n'ont pas été dégradés, jaune= substrats totalement oxydés. Les suivis colorimétriques ont été effectués sur 48h, deux réplicas techniques sont représentés.

4 Matériel et méthodes

4.1 Conduite des expérimentations de phytopathologie

4.1.1 Origine et conditions de croissance du matériel végétal utilisé

Les graines sont directement semées dans un terreau complémenté par 3 g.L-1 d'engrais à diffusion lente . Les semis sont placés en serres, suivants des cycles de 12 h de lumière à 28 °C et 80% d'humidité relative (HR) et 12 h d'obscurité à 25 °C et 60% HR. Les plantes sont utilisées à 3 semaines après semis pour les expérimentations par infiltrations et à 4 semaines pour les tests par leaf-clipping (voir 4.1.3, « Tests de pathogénie »).

4.1.2 Souches bactériennes utilisées et conditions de cultures

La liste des différentes souches de *Xoo* utilisées au cours des expérimentations présentées dans le chapitre 3 est disponible dans le tableau S2. Les bactéries sont cultivées sur un milieu PSA (Peptone Sucrose Agar, 1% peptone, 1% sucrose, 0,1% acide glutamique, 1,6% agar) pendant quatre jours à 28°C. Une colonie isolée par souche est prélevée puis mise en culture sur milieu PSA pendant 24h avant d'être re-suspendue dans de l'eau stérile. La Densité Optique à 600 nm (DO_{600}) utilisée pour les essais par leaf-clipping et pour les infiltrations est de 0,2 et 0,5 respectivement. Ces DO_{600} correspondent approximativement à 43 et 108 Unités Formant des Colonies (UFC) par millilitre.

4.1.3 Déroulement et suivi des tests de pathogénie

L'évaluation de la virulence des différentes souches de *Xoo* a été réalisée par la méthode leaf-clipping. La plus jeune feuille entièrement déployée est coupée à 2-3 cm de l'extrémité de la feuille avec des ciseaux préalablement trempés dans une suspension bactérienne à une DO_{600} de 0.2 ou de l'eau stérile pour le contrôle. La longueur des lésions produite est mesurée 14 jours après l'inoculation. La production des échantillons pour les expériences de qPCR et de qGUS a étés réalisé par infiltrations. Les plus jeunes feuilles de plants de riz âgés de 3 semaines ont été infiltrées avec des suspensions bactériennes à une DO_{600} de 0.5 ou de l'eau pour le contrôle à l'aide d'une seringue sans aiguille, appliquée sur la face inférieure de la feuille.

4.2 Techniques de biologie moléculaire

4.2.1 Suivi d'inductions de gènes par PCR quantitative

Trois plantes différentes par traitement et par génotype sont infiltrées. Une seule feuille par plante est utilisée. Deux à trois répétitions biologiques sont pratiquées. Les échantillons sont collectés 48 h après l'inoculation. L'ARN total est extrait à l'aide du réactif TRI-Reagent (EUROMEDEX)). Après l'extraction, un traitement DNAse est effectué à l'aide du kit TURBO DNA-free (Thermo Fisher Scientific,). La synthèse d'ADN complémentaire à partir d'oligo-dT est réalisée via la SuperScriptIII (Thermo Fisher Scientific). Les réactions de PCR quantitative sont préparées avec le Mesa Blue qPCR Mastermix (EUROGENTEC). Concernant le suivi de l'expression de SWEET11, *SWEET13* et *SWEET14*, les amorces utilisées dans cette étude ainsi que le gène témoin choisi (Elongation factor-1 alpha, EF1α, GenBank : GQ848072.1) ont été décrits par Blanvillain-Baufumé et al en 2017, et répertoriées dans le tableau S1. Trois réplicas techniques sont préparés pour chaque échantillon. Les valeurs d'expression sont normalisées en soustrayant les valeurs obtenues pour le gène EF1α au gène étudié (méthode 2-ΔCt). La moyenne des réplicas techniques pour chaque échantillon est calculée.

4.2.2 Reproduction de la délétion RegtalC

La reproduction de la délétion de 35 kb présente chez le mutant originel BAI3.1.1 a été réalisée par double recombinaison homologue d'un vecteur suicide. Ce vecteur de délétion, pK18_Reg*talC*_SpecR a été construit via la technique d'assemblage Gibson à partir du vecteur de contre sélection au sucrose pK18mobsacB. pK18_Reg*talC*_SpecR contient une cassette de résistance à la spectinomycine bordée par des régions flanquantes de Reg*talC*, d'environ 2 kb chacune (voir carte, figure S2) permettant la double recombinaison homologue. Le gène sacB porté par le plasmide code une enzyme qui convertit le sucrose en un métabolite toxique chez les bactéries gram-négatives. Après éléctroporation de pK18_Reg*talC*_SpecR dans les souches BAI3 sauvage et BAI3 résistante à la rifampicine (fond génétique originel du mutant BAI3.1.1), une contre sélection en milieu PSA à 10 % de sucrose et 100 µg.mL-1 de spectinomycine a permis d'éliminer les clones contenant le plasmide pour ne garder que les évènements de recombinaison. Un séquençage sur l'insert a confirmé la délétion.

4.2.3 Complémentations talC

Les mutants ponctuels *talC* (BAI3Δ*talC*#9 et BAI3Δ*talC*#16) ainsi que le mutant Reg*talC* (BAI3.1.1) ont été complémentées par introduction par électroporation d'un vecteur d'expression construit dans le plasmide pSKX1. Ce vecteur, pSKX1.1-*talC* contient la séquence codante de *talC* sous le contrôle de son promoteur et porte un gène de résistance à la gentamycine (voir carte en figure S3). Des clones positifs (séquençage sur l'insert) ont été obtenus et une confirmation par Western-blot de la présence de TalC a été réalisée.

4.2.4 Séquençage de souches mutantes par la technologie Nanopore

Les ADNs pour le séquençage long read par technologie Oxford Nanopore (ONT) ont été extraits depuis des cultures pures de bactéries dont la dernière phase de croissance s'est déroulée sur milieu « PSA » sans sucrose. Les extractions ont été réalisées avec le kit Genomic-tip 100/G, selon le protocole du fabricant (QIAGEN). Le kit « Rapid library preparation » (SQK-RBK110-96, Oxford Nanopore) a été utilisé pour la préparation des banques. Le séquençage a été effectué sur un séquenceur MinION Mk1C (Oxford Nanopore)avec des flow cells R10.3 (FLO-MIN111, Oxford Nanopore). Le base calling et le dé-multiplexage des signaux ONT ont été réalisés avec le logiciel ONT Guppy (v6.0.1+652ffd179). L'assemblage des reads a été réalisé avec le pipeline CulebrONT (v2.0.1) (Orjuela *et al.*, 2022 [297]). Les assemblages ont ensuite été raffinés avec polypolish (v0.5.0) (Wick et Holt, 2022 [298]).

4.3 Techniques Biochimiques

4.3.1 Quantification cinétique de l'activité GUS

Les essais de quantification de l'activité β-glucuronidase (GUS) ont été réalisés par fluorométrie sur des extraits de plantes transgéniques contenant les constructions pSWEET:gSWEET-GUSplus, selon le protocole de Cervera 2005 1. Pour rappeler brièvement le principe de cet essai, l'hydrolyse du substrat (Methylumbelliferylβ-D-glucuronide, MUG) par la GUS libère un composé fluorescent, le 4-methylumbelliferone (MU), émettant à 465 nm en milieu basique. Les échantillons de feuilles, collectées à 48h après inoculation sont broyés puis resuspendus dans un Tampon d'Extraction GUS (GEB, 50mM $NaPO_4$ pH 7, 10mM Na_2EDTA pH 8, 0,1% SDS, 0,1% Triton X-100). Les débris de cellules sont éliminés par centrifugation. La quantité de protéines totales dans le surnageant est déterminée par la méthode de Bradford (Roti-Quant 5X reagent, Carl Roth). Pour les mesures d'activité GUS, 2 µL de lysat sont ajoutés à 98 µL d'une solution de 4-MUG dilué dans du GEB. Les mesures en fluorométrie

sont pratiquées avec un lecteur de plaque (Tecan SPARK multimode microplate reader). Le suivi de production de MU par la GUS est réalisé sur 30 min, avec une mesure toutes les minutes. L'enceinte du lecteur est thermosthatée à 31°C. Le MU est excité à 360 nm et émet à 465 nm. L'activité relative de la GUS en RFU. min-1. µg-1 est calculée en divisant la valeur de la pente de la droite de régression de production de MU par la quantité totale de protéines dans le lysat.

4.3.2 Suivis d'activités métaboliques en plaques

Le suivi des performances du métabolisme du sucre a été réalisé sur les plaques de phénotypage spécifiques des carbohydrates PM01 et PM02 de BIOLOG (Phenotype Microarrays). Les bactéries ont été inoculées sur les plaques en suspension à 85% de transmittance T dans un fluide d'inoculation minimal (IF-0a, BIOLOG) en présence d'un indicateur redox à base de tétrazolium (Redox Dye Mix H, BIOLOG). Les plaques ont été incubées à 25°C pendant 96h. La réduction du tétrazolium dans le milieu donne un composé soluble violet, le formazan, dont l'apparition est mesurée au cours du temps par spectrophotométrie à 590 nm directement dans les microplaques. Cette réduction est possible en présence de cellules avec une chaîne de transport d'électrons fonctionnelle et active. Cette réduction n'est donc pas un indicateur de la dégradation du substrat, mais est révélatrice de l'activité métabolique des cellules. Cette réaction est proportionnelle au nombre de cellules présentes (Braissant *et al.*, 2020 [299]).

7 Données supplémentaires :

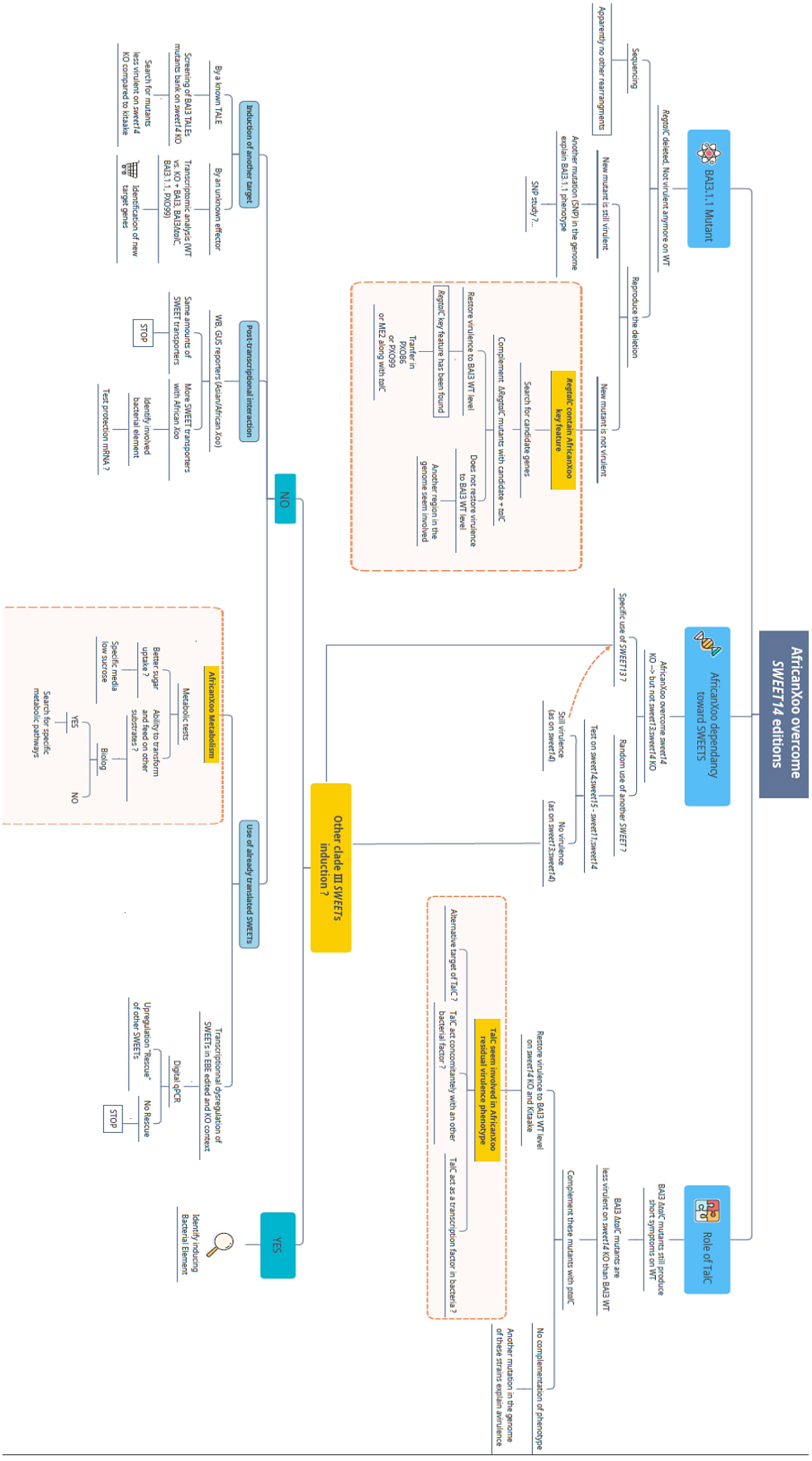


Figure S1 : Organigramme décisionnel pour la résolution de l'énigme Africaine. Les différentes hypothèses ayant été exploitée au cours de ces travaux pour mettre à jour les mécanismes permettant aux *Xoo* africaines la mise en place du phénotype SWIIVA.

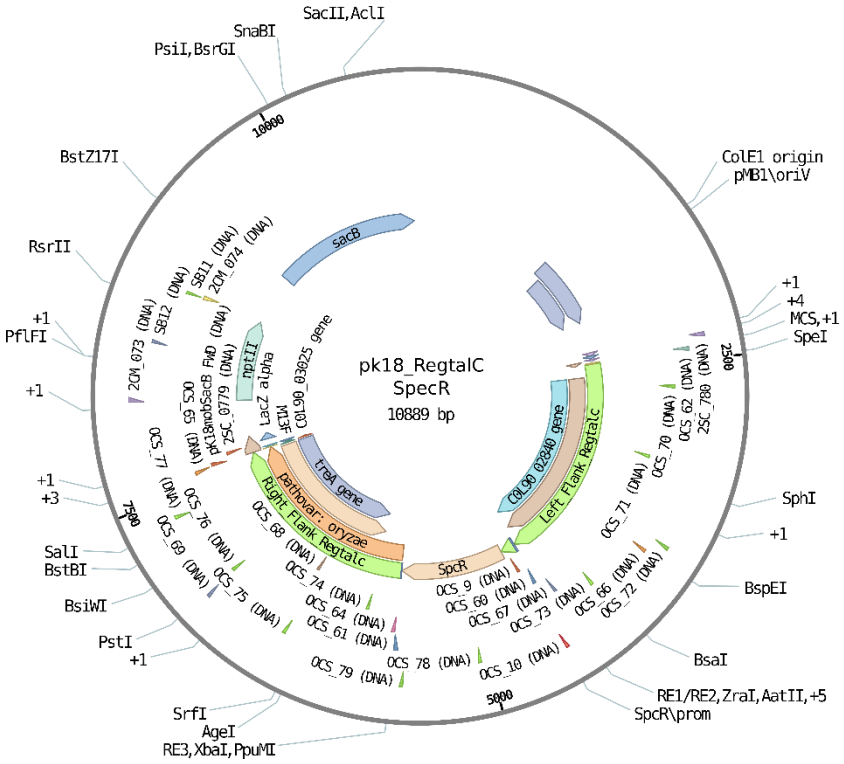


Figure S2 : Vecteur de délétion pK18_Reg*talC*_SpecR. Carte génétique du vecteur de délétion construit pour la reproduction de la délétion de 35 kb ayant donné le mutant non-virulent BAI3Δ*RegtalC* (BAI3.1.1).

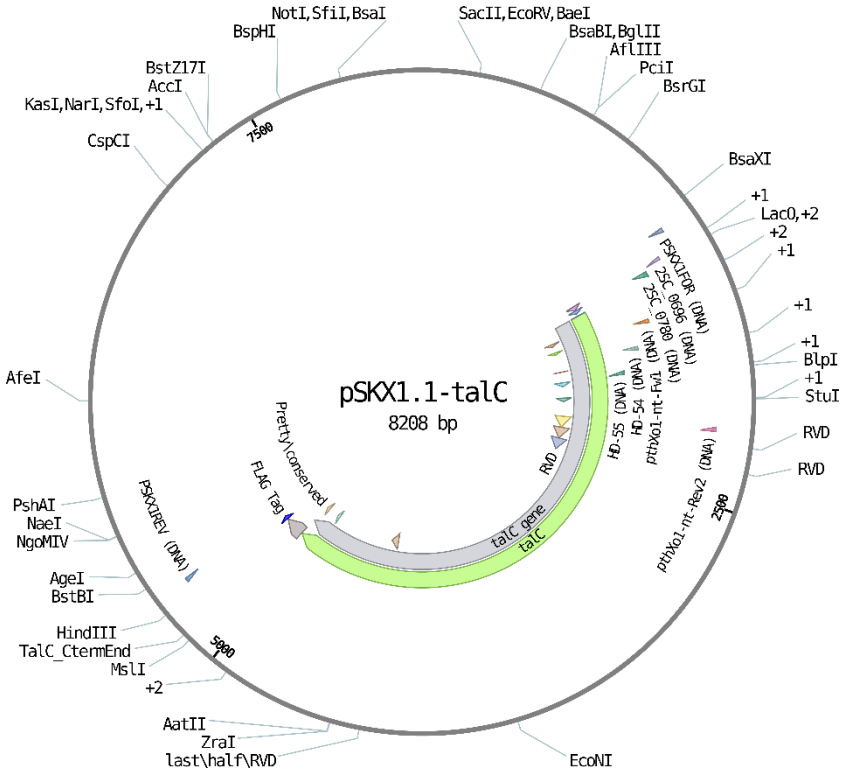


Figure S3 : Vecteur de complémentation *talC*. Carte génétique du vecteur utilisé pour la complémentation plasmidique par *talC*. p*talC* a été construit à partir du vecteur d'expression pSKX1, portant une résistance à la gentamycine.

Tableau S1 : Amorces utilisées pour la qPCR sur les *SWEET* du clade III. Deux couples d'amorces différents pour chaque *SWEET* ont été testés, un couple a été dessiné par l'IRD, et l'autre par l'HHU.

Primer	Origin	Gene cible	Séquence
SWT11 forw	HHU	*OsSWEET11*	GGGATTTCTGGCTAGTTTCT
SWT11 rev	HHU	*OsSWEET11*	CGAGGTAGAGGACGATGTAG
SWT13 forw	HHU	*OsSWEET13*	AGAGTTTTCAGCCAACACAT
SWT13 rev	HHU	*OsSWEET13*	TGTAGATCCGGTAGAACGTC
SWT14 forw	HHU	*OsSWEET14*	TATTGCCTGATCATCCTCTT
SWT14 rev	HHU	*OsSWEET14*	GTGAACATCTTGGCCTTCT
SWT15 forw	HHU	*OsSWEET15*	CTTCACCTTTGGCATCTTAG
SWT15 rev	HHU	*OsSWEET15*	AACGCGTAGTACATCCACA
OsEF-1alpha_qPCR_F	IRD	*osEF1α*	GAAGTCTCATCCTACCTGAAGAAG
OsEF-1alpha_qPCR_R	IRD	*osEF1α*	GTCAAGAGCCTCAAGCAAGG
*sweet11*_FW	IRD	*OsSWEET11*	AGTCGACGGGAGGGTACAG
*sweet11*_RV	IRD	*OsSWEET11*	TTCGGGTACATGACGTAGGG
*sweet13*_FW	IRD	*OsSWEET13*	GGCCTGTCCCTGCAGCATC
*sweet13*_RV	IRD	*OsSWEET13*	CCCGAACACCCCCACGTTC
*sweet14*_FW	IRD	*OsSWEET14*	ACTTGCAAGCAAGAACAGTAGT
*sweet14*_RV	IRD	*OsSWEET14*	ATGTTGCCTAGGAGACCAAAGG
sweet15_FW	IRD	*OsSWEET15*	TGTCCATGGAACGCAGCAC
sweet15_RV	IRD	*OsSWEET15*	CAGCCGAGGACGTGGAC

Tableau S2 : Souches sauvages et mutantes de *Xoo* et *Xoc* utilisées au cours de des études présentées dans le Chapitre III et la discussion.

Name	Pathovar	Origine	Résistance	Construction
MAI68	*Xoo*	Mali	/	WT
MAI99	*Xoo*	Mali	/	WT
AXO1947 -1	*Xoo*	Cameroun	/	WT
B3.1.1	*Xoo*	Burkina Faso	Rif - Kan	Souche BAI3 KO TalC (BAI3ΔTalC), la mutation, produite via insertion d'un plasmide suicide, a mené à une large délétion de 35 kb des séquences bordantes à l'insertion.
B3.1.1 pSKX1 #1	*Xoo*	Burkina Faso	Rif Kn Gm	conjugaison de pSKX1 empty dans la souche mutante B3.1.1
B3.1.1 p*talC* #2	*Xoo*	Burkina Faso	Rif Kn Gm	conjugaison de pSKX1 avec insert *talC* dans la souche mutante B3.1.1
BAI170	*Xoo*	Burkina Faso	/	WT
BAI200	*Xoo*	Burkina Faso	/	WT
BAI250	*Xoo*	Burkina Faso	/	WT
BAI3	*Xoo*	Burkina Faso	/	WT
BAI3 Δ16 pSM7	*Xoo*	Burkina Faso	Rif - Kan	souche BAI3 *talC* mutante par insertion de pSM7 (Tu Tran Tuan).

BAI3 Δ9 pSM7	*Xoo*	Burkina Faso	Rif Kn	souche BAI3 *talC* mutante par insertion de pSM7 (Tu Tran Tuan).
BAI3 Rif	*Xoo*	Burkina Faso	Rif	Rifampicine résistant spontané
BAI3ΔReg*talC*#13	*Xoo*	Burkina Faso	spec	électroporation de pK18_Reg*talC*_Spec#02 (pSCIAL_01) dans BAI3 WT (CIX2679, CSX_57)
BAI3Δ*talC*#16 pSKX1 #2	*Xoo*	Burkina Faso	Rif Kn Gm	conjugaison de pSKX1 empty vector dans la souche mutante BAI3 Δ16 pSM7 (CIX1790)
BAI3Δ*talC*#16 p*talC* #1	*Xoo*	Burkina Faso	Rif Kn Gm	conjugaison de pSKX1 insert *talC* dans la souche mutante BAI3 Δ16 pSM7 (CIX1790)
BAI30	*Xoo*	Burkina Faso	/	WT
BAI3-1-1	*Xoo*	Burkina Faso	Rif Gm	conjugaison de pSKX1 empty dans le BAI3 rif (CIX531)
BAI3RifΔReg*talC*#1	*Xoo*	Burkina Faso	rif, spec	électroporation de pK18_Reg*talC*_Spec#02 (pSCIAL_01) dans BAI3 rifR (CIX151, CSX_23)
BAI4 *Xoo*	*Xoo*	Burkina Faso	/	WT
CFBP 1949	*Xoo*	Mali	/	WT
CFBP 1951	*Xoo*	Mali	/	WT
Karfiguela 13	*Xoo*	Burkina Faso	/	WT
MAI134	*Xoo*	Mali	/	WT
NAI7	*Xoo*	Niger	/	WT
Nati Park	*Xoo*	Bénin	/	WT
PXO61	*Xoo*	Philippines	/	WT
PXO86	*Xoo*	Philippines	/	WT
PXO99	*Xoo*	Philippines	/	WT
S62-2-22	*Xoo*	Sénégal	/	WT
Toula 13	*Xoo*	Niger	/	WT
Toula 20	*Xoo*	Niger	/	WT
Ug 12	*Xoo*	Ouganda	/	WT
XOO CII-1	*Xoo*	Côte-d'Ivoire	/	WT
VXO32	*Xoc*	Vietnam	/	WT
BLS279	*Xoc*	Philippines	/	WT
MAI10	*Xoc*	Mali	/	WT
Bur1	*Xoc*	Burundi	/	WT
BLS256	*Xoc*	Philippines	/	WT
BLS256 hrcC	*Xoc*	Philippines	Kn	Mutant T3SS, insertion du plasmide pSM7 dans le gène hrcC
BAI11	*Xoc*	Burkina Faso	/	WT

DISCUSSION

1 LE CONUNDRUM AFRICAIN

L'origine du conundrum Africain réside en l'observation de la capacité des souches de *Xoo* africaines à infecter des plantes éditées pour les EBEs de TalC et TalF qui se situent au sein du promoteur du gène *OsSWEET14*. Ce phénotype, que nous avons appelé SWIIVA pour « *SWEET* Induction Independant Virulence of African *Xoo* », traduit la capacité des *Xoo* africaines à maintenir de la virulence en l'absence de l'induction d'un gène *SWEET* du clade III. Ce phénotype était pour le moins surprenant, car en l'état des connaissances sur le pathosystème *Xoo* / riz, l'induction des *SWEETs* du clade III semblait être un prérequis à l'infection généralisable à l'ensemble des souches de *Xoo*. Notre questionnement principal au cours de ces travaux de thèse avait donc pour but de mettre à jour l'origine de cette virulence particulière et a donné lieu à la formulation de plusieurs hypothèses, illustrées dans la figure S1 du Chapitre III.

1.1 Concernant les déterminants de pathogénie de la région *talC*

1.1.1 Recherche de cibles alternatives de TalC

Pour tenter d'expliquer SWIIVA, une étude à laquelle j'ai participé à poursuivi l'hypothèse selon laquelle TalC aurait au moins une autre cible transcriptionnelle génétiquement redondante avec *OsSWEET14* dans le génome du riz (Doucouré *et al.*, 2022 [300], voir annexe 3). Ces travaux ont permis d'identifier un locus dont la transcription est induite de manière bidirectionnelle par TalC, que nous avons nommé Alternative Target of TalC (ATAC). Ce locus se compose du gène INCREASED LEAF INCLINATION 2 (*ILI2*), impliqué dans l'élongation des cellules du parenchyme, et d'un long ARN non codant putatif (lncRNA). Bien qu'ATAC soit une cible réelle de TalC, l'introduction de TALEs artificiels ciblant ce locus dans une souche asiatique dépourvue de TALEs, ne confère pas de gain de virulence. De la même manière, l'édition chez le riz de l'EBE ciblé par TalC dans le locus ATAC prévient bien l'induction des deux gènes du locus, mais n'apporte pas de résistance contre les *Xoo* africaines. D'autre part, il a été montré par des partenaires au cours du projet que la complémentation de la souche ME2 par *talC* seul, ne promeut pas l'apparition de symptômes sur un mutant *sweet14* ou une lignée éditée au niveau de l'EBE de TalC dans le promoteur de *SWEET14* (phénotype de virulence des *Xoo* africaines) (Eom *et al.*, 2019 [288]; Oliva *et al.*, 2019 [287]). Ce résultat remet en question l'hypothèse selon laquelle une cible alternative de TalC promeut la virulence des *Xoo* africaines et nous montre que *talC* seul n'est pas suffisant à SWIIVA. En conclusion, il semblerait que l'induction d'ATAC par TalC est fortuite et ne contribue pas à la pathogénèse de la BLB.

1.1.2 Recherche de l'élément conditionnant la virulence apportée par Reg*talC*

Comme exposé dans le chapitre III, la région *talC* contient différents éléments présentant un intérêt évident pour une étude plus approfondie.

Le principal candidat est *cbsA*, un gène codant pour une cellobiohydrolase, dont l'activité consiste à dégrader la cellulose de la paroi cellulaire végétale en hydrolysant une par une des unités de cellobiose. Dans la suite du processus de dégradation de la cellulose, le cellobiose est converti en glucose par les β-glucosidases (Jensen *et al.*, 2018 [301]), ce qui laisse penser qu'en présence de bactéries, la concentration en sucres libres augmente. La présence de CbsA induit des réactions de défense non

spécifiques du riz, comme l'accumulation de callose, pouvant correspondre à la détection du signal généré par une concentration de sucres anormale chez l'hôte. La mutation de ce gène chez *Xoo* a un fort impact sur la virulence (Tayi *et al.*, 2017 [173]). De récentes observations ont d'ailleurs révélé que le mode de vie vasculaire ou extravasculaire de *Xoo* et d'autres espèces de *Xanthomonas* était conditionné par la présence de *cbsA* (Gluck-Thaler *et al.*, 2020 [174]).

Étonnamment, la complémentation par *cbsA* des mutant Δ*RegtalC* n'apporte pas de gain de virulence (voir figure 42, Chapitre III). Au vu des connaissances sur ce gène dans la pathogénie de *Xoo*, on peut supposer qu'il nécessite l'interaction d'autres facteurs pour moduler la virulence.

Plusieurs études ont mis en évidence que *cbsA* était induit par OryR, une protéine homologue à LuxR, qui est un facteur de transcription dont l'activité est soumise à la perception de molécules du Quorum Sensing (AHLs) (Ferluga *et al.*, 2007 [302]). Sadoine *et al.* ont montré en 2022 [303] que la transcription des récepteurs d'AHLs était induite par l'efflux de sucrose déclenché dans la plante lors de l'infection par *Xoo*. On pourrait alors supposer que l'induction de *cbsA* nécessite l'efflux de sucrose médié par TalC lors de l'infection. La complémentation des mutants Δ*RegtalC* + *cbsA* par *talC* permettrait alors l'activité de la cellobiohydrolase et pourrait potentiellement restaurer la virulence du mutant.

Comme vu dans le Chapitre III, il existe deux autres gènes candidats tout particulièrement intéressants dans *RegtalC* : l'effecteur de type III *xopV* et une cellulase. Il est prévu d'utiliser un système de base editing CRISPR-Cas9 afin de générer des mutants ponctuels pour chacun de ces trois gènes et déterminer la potentielle perte de virulence induite.

Enfin, nous avons vu au Chapitre III que l'expression de *talC* à partir d'un plasmide est capable de complémenter la virulence de la souche BAI3Δ*talC* sur un mutant KO *sweet14*. Il est prévu de mieux cerner la spécificité de TalC dans le phénotype SWIIVA, en complémentant la souche BAI3Δ*talC* avec d'autres TALEs ciblant *SWEET14* ou un autre *SWEET* du clade III comme par exemple AvrXa7, TalF ou PthXo1. Cette souche hybride sera ensuite inoculée sur plantes sauvages et la lignée mutante KO *sweet* correspondante, afin d'observer ou non un phénotype SWIIVA. Si la complémentation permet la virulence, *talC* ne sera donc pas considéré comme partie prenante du phénotype. Dans le cas contraire, cela mettra en évidence un statut particulier pour TalC.

1.1.3 Validation des déterminants par transfert dans un fond génétique de *Xoo* asiatique

A l'issue de cette recherche de facteurs contribuant à la pathogénie de BAI3 dans la région *talC* (*RegtalC*), il serait intéressant de transférer les potentiels candidats dans un fond génétique de souche asiatique de *Xoo*. Cette expérience de complémentation permettrait de tenter de reproduire un phénotype de virulence Africain par l'ajout du ou des fragments choisis de *RegtalC*, avec ou sans l'apport de *talC* lui-même.

Parmi les complémentations envisagées, nous souhaiterions utiliser la souche des Philippines ME2 qui est dépourvue de son TALE de virulence majeur *pthXo1*, dans laquelle il s'agirait de transférer *talC* accompagné de la ou des portions choisies de *RegtalC*. Nous souhaiterions également utiliser PXO86, une autre souche des Philippines qui possède le TALE AvrXa7, inducteur de *SWEET14*, mais ne présentant pas de virulence sur le mutant KO *sweet14*. Dans cette configuration, nous souhaiterions ne rajouter que les éléments déterminants de *RegtalC*. Enfin, il serait intéressant de voir si l'induction d'un autre *SWEET* que *SWEET14* permettrait d'arriver au phénotype de virulence africain. Ainsi, nous voudrions ajouter les déterminants de *RegtalC* dans la souche modèle PXO99 qui contient *pthXo1*, le TALE inducteur de *SWEET11*. Le phénotypage de ces souches asiatiques complémentées sur la variété Kitaake, les lignées dérivatives mutantes KO pour *sweet14* ou *sweet11,* et sur les lignées éditées dans

les promoteurs de *SWEET14* et *SWEET11* nous permettra alors d'observer ou non un transfert du phénotype SWIIVA chez les souches de *Xoo* asiatiques.

Dans le cas de l'acquisition effective de SWIIVA à l'issue de ces complémentations, nous pourrons conclure que la région *talC* seule permet de conférer ce phénotype. Dans le cas contraire, nous pourrons supposer que *RegtalC* n'est pas suffisante pour aboutir à SWIIVA, et que le fond génétique Africain renferme d'autres éléments nécessaires à son exécution.

1.1.4 Perspectives : Criblage sans a priori d'un élément génétique clé pour l'apparition du phénotype SWIIVA

Dans la perspective évoquée ci-dessus de la non complémentation des souches asiatiques de *Xoo* par *RegtalC* seule, une autre approche serait le criblage sans à priori du génome des *Xoo* africaines pour la recherche du ou des déterminants de SWIIVA. Cette approche consiste à créer une banque de cosmides d'ADN génomique de BAI3, notre souche de *Xoo* Burkinabé modèle dont le génome a été séquencé, et d'introduire ces fragments d'ADNg dans la souche asiatique de *Xoo* PXO86 via l'utilisation du plasmide réplicatif pLAFR3. Le phénotypage sur la lignée éditée dans le promoteur de *OsSWEET14* dans et le mutant KO pour *sweet14* de la souche PXO86 complémentée par la banque de cosmides de BAI3 nous permettra de recenser les clones offrant un gain de virulence. Afin de diminuer le nombre d'inoculations en parallèle, ces inoculations prendront la forme de pools d'inoculation d'une dizaine de souches complémentées.

1.2 Au sujet de la résistance des KOs *sweets*

1.2.1 Précisions sur l'étude du phénotype SWIIVA sur les lignées mutantes *sweet*

Le phénotype SWIIVA correspond à la capacité des *Xoo* africaines à générer de la maladie chez le riz, sans induction d'un *SWEET* du clade III. Cette faculté est avant tout observable sur des lignées de riz dont les EBEs dans les promoteurs *SWEETs* ont été rendus inutilisables pour la bactérie. Au-delà d'une absence d'induction du gène, ce que l'on observe sur un KO *sweet14* relève d'une perte de fonction du gène chez la plante. Les conclusions biologiques lors des expériences d'inoculation sur ces différents fonds génétiques sont donc assimilées, mais pas identiques. Le KO *sweet14* sert en quelques sortes de proxy à l'étude du phénotype SWIIVA et nous permet surtout d'évaluer l'importance même des *SWEETs* du clade III pour les *Xoo* africaines. Au cours de cette évaluation, nous avons vu que les *Xoo* africaines étaient capables de coloniser un mutant *sweet14* simple, mais étaient contrôlées par un double mutant KO *sweet13 ;sweet14*. Pourtant, dans un fond sauvage ou KO *sweet14*, ces souches n'ont pas montré d'induction robuste de *SWEET13*. Au-delà des considérations de génétique bactérienne et d'un potentiel contrôle post-transcriptionnel exercé par les *Xoo* africaines sur *SWEET13*, on peut aussi s'interroger sur l'origine de la résistance du double mutant *sweet13* ;*sweet14* coté plante.

1.2.2 Potentiel priming de l'immunité du double mutant *sweet13 ;sweet14*

Comme abordé dans l'introduction générale de ce manuscrit de thèse, le rôle des sucres dans l'immunité est complexe. Plusieurs hypothèses coexistent au sujet du rôle des transporteurs de sucres dans les interactions plantes/microorganismes (voir introduction, 4.4.2.3 « Exploitation par les microorganismes »). Il est envisageable que l'altération du transport du sucrose chez le double mutant *sweet13*;*sweet14* ait un impact dans la signalisation de l'immunité chez la plante. D'autre part, il n'est pas à exclure que le mutant en lui-même ait subi un remaniement ailleurs dans son génome, pouvant être à l'origine d'une modification de son immunité. Afin de vérifier l'existence d'un potentiel

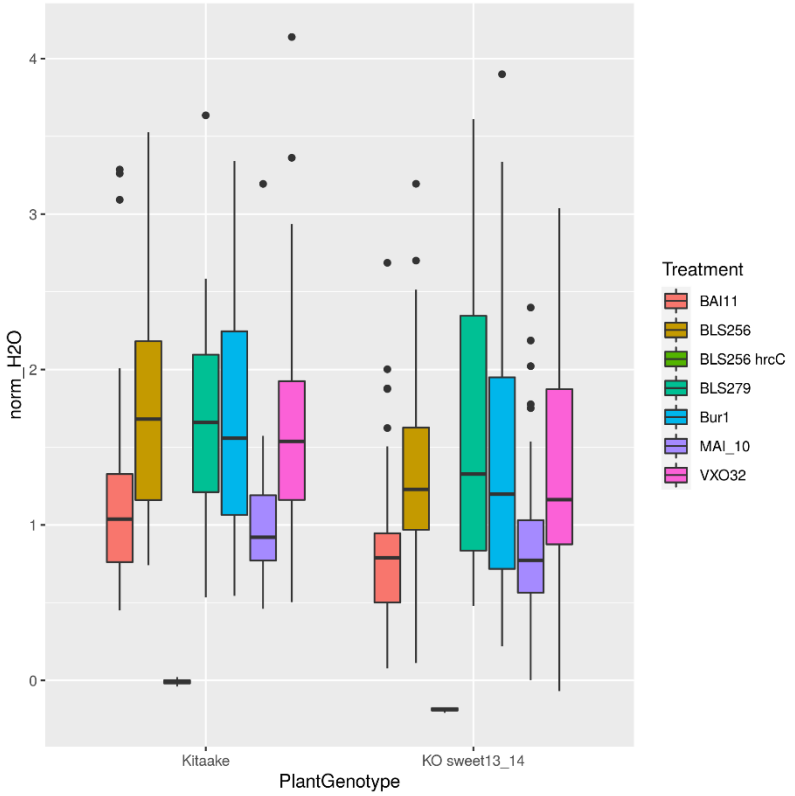


Figure 45 : La double mutation de *SWEET13* et *SWEET14* chez Kitaake ne fournit pas de résistance à *Xoc*. Longueurs de lésions obtenues pour différentes souches de *Xoc* asiatiques (BLS256, BLS279, VXO32) et africaines (BAI11, Bur1, MAI_10) et d'un mutant non-virulent (BLS256hrcC). Les inoculations ont été pratiquées par infiltration à une DO_{600}=0.5, et les lésions ont été observées à 7 jours post infiltration.

	AtSWEET1	AtSWEET2	AtSWEET3	AtSWEET4	AtSWEET5	AtSWEET6	AtSWEET7	AtSWEET8	AtSWEET9	AtSWEET10	AtSWEET11	AtSWEET12	AtSWEET13	AtSWEET14	AtSWEET15	AtSWEET16	AtSWEET17
AtSWEET1																	
AtSWEET2																	
AtSWEET3																	
AtSWEET4																	
AtSWEET5																	
AtSWEET6																	
AtSWEET7																	
AtSWEET8																	
AtSWEET9																	
AtSWEET10																	
AtSWEET11																	
AtSWEET12																	
AtSWEET13																	
AtSWEET14																	
AtSWEET15																	
AtSWEET16																	
AtSWEET17																	

Figure 46 : Assemblages de sous-unités SWEET en homo et hétéropolymères décrits chez *A. thaliana*. Cette matrice montre les différents types d'associations rencontrées lors de l'assemblage des transporteurs SWEETs de *A. thaliana*. Les cases colorées représentent l'interaction entre deux SWEETs. Les cases oranges indiquent les homopolymères, et les cases grises les hétéropolymères. Reproduit de Ji et al., 2020 [304].

amorçage de l'immunité chez le double mutant *sweet13*;*sweet14*, nous avons évalué sa résistance à d'autres pathogènes : le proche pathovar *Xoc*, et le champignon phytopathogène *Magnaporthe oryzae*. Différentes souches de *Xoc*, africaines et asiatiques, ainsi qu'un mutant non virulent déficient pour le T3SS (BLS256Δ*hrcC*) furent infiltrés sur les variétés Kitaake sauvage et KO *sweet13*;*sweet14* (Figure 45). Sur nos deux expériences, la mesure de la dispersion du pathogène sept jours après inoculation ne s'est pas montrée significativement différente selon le génotype, toutes bactéries confondues. De la même manière, la comparaison du nombre de points d'infections dus à *M. oryzae* entre la variété sauvage et le mutant n'a pas montré de résistance accrue du double KO au champignon.

1.2.3 Formation d'hétéromères de SWEET

Les protéines issues de la traduction des gènes KO *sweet* créés au cours du projet forment des sous-unités SWEET tronquées (un à deux domaines transmembranaires manquants, Eom *et al.*, 2019 [288]). Par ailleurs, comme évoqué dans l'introduction de ce manuscrit, nous savons que les transporteurs SWEETs forment des structures quaternaires formées d'homomères ou d'hétéromères, comme cela a été montré chez *A. thaliana* (figure 46). Dans le cas où un transporteur fonctionnel SWEET14 ou SWEET13 ne nécessitait que des homomères, on imagine que nous obtiendrions des transporteurs totalement non fonctionnels chez les plantes KO. Hors, s'il s'agit d'hétéromères SWEET13/SWEET14, on pourrait imaginer que la mutation simple de SWEET13, ou SWEET14, conduirait à la formation d'un transporteur toujours fonctionnel, mais avec une efficacité réduite. Cette observation va dans le sens des résultats obtenus sur les KO simples *sweet13* et *sweet14*, sur lesquels nous observons toujours des lésions en présence de *Xoo* induisant *SWEET14*. Dans le même ordre d'idée, un double mutant *sweet13*;*sweet14* verrait la totalité des sous-unités de son transporteur tronquées, et ne pourrait pas (ou quasiment pas) assurer sa fonction, ce qui expliquerait pourquoi les *Xoo* n'arrivent pas à s'octroyer une quantité suffisante de sucre à l'intérieur de l'apoplaste d'un double KO pour ces gènes. Poursuivant cette idée, nous pouvons également inférer une stœchiométrie dans la proportion des sous-unités d'un transporteur. Pour rappel, on observe une diminution importante de la longueur des lésions sur un KO *sweet14* en présence de *Xoo* induisant *SWEET14*. Au vu de ces résultats, on peut imaginer que la proportion de sous-unités SWEET14 est plus importante dans le transporteur (deux sous-unités SWEET14 pour une sous unité SWEET13 par exemple).

Aussi, dans le cas de l'assemblage d'hétéromères de SWEET13 natives et SWEET14 tronquées, donnant un transporteur à l'efficacité amoindrie mais fonctionnel, une induction de la production de sous-unités SWEET14, même tronquées, aurait toujours un effet. C'est ce que l'on observe lorsque l'on compare les lésions produites sur un mutant KO *sweet14* par des mutants de la souche de *Xoo* BAI3Δ*talC* par rapport à la souche BAI3 sauvage, ou de manière plus flagrante encore avec les souches BAI3Δ*talC* complémentée avec *talC*. Des résultats préliminaires de qPCR sur les lignées KO nous ont d'ailleurs montré qu'il existait toujours une induction du gène *SWEET14* chez les lignées de riz mutantes KO *sweet14* et *sweet13*;*sweet14*.

1.2.4 Concernant les profils de longueurs de lésions contrastés sur les KO *sweet*

Bien que l'on s'attendrait à observer des symptômes de lésions plus courtes sur une lignée de riz KO pour un gène *SWEET* du clade III, quel qu'il soit, ce n'est pas ce que l'on observe pour une partie de nos souches africaines. Lors de l'inoculation de notre panel de 21 souches africaines originaires de 8 pays différents, nous avons pu remarquer que certaines souches (environ la moitié) produisaient des

symptômes de lésions plus longues sur la lignée de riz KO pour *sweet13* que sur la variété Kitaake sauvage. Ces résultats suggèrent que les populations de *Xoo* africaines ne sont pas uniformes dans leur pathogénie. Ces deux profils différents sont engendrés par des isolats d'origines variées, et des souches d'un même pays (comme le Burkina Faso) peuvent faire partie des deux catégories.

Nous savons déjà que les populations de *Xoo* retrouvées au sein de certains pays d'Afrique de l'Ouest ont une diversité forte de TALomes. Neuf types de TALEs ont été décrits dans les populations d'Afrique de l'Ouest, de TalA à TalI, et plusieurs études ont montré que leur présence varie selon les souches, ainsi que leur taille et parfois leur séquence (Doucouré *et al.*, 2018 [293]). Une étude réalisée en 2021 ayant pour but de caractériser la diversité génétique neutre et celle des TALomes des souches de *Xoo* du Burkina Faso a révélé que différents TALomes coexistent au Burkina Faso, parfois même au niveau de souches d'un même champ ou appartenant à un même haplotype MLVA (Diallo *et al*., in prep).

La totalité des génomes des souches de ce panel sont disponibles (soit séquencés en amont, soit par nos soins au cours de l'étude). Il serait intéressant de rechercher s'il existe des différences génétiques notables entre ces deux groupes de souches, notamment au niveaux de la composition de leurs TALomes.

1.2.5 Effet du fond génétique de la variété de riz sur le phénotype des mutants

Contrairement à ce que l'on a pu observer sur nos mutants générés dans la variété de riz Kitaake, une étude menée par Zeng *et al.*, en 2020 [305] a démontré qu'un mutant KO *sweet14* dans le fond génétique Zhonghua 11 confère une résistance totale à la souche Camerounaise modèle AXO1947. Nos données (Figure 34, Chapitre 3) montrent qu'une mutation équivalente dans le fond génétique Kitaake s'accompagne d'une réduction de moitié seulement des longueurs de lésions provoquées par cette même souche. Selon des données non publiées du laboratoire de notre collaborateur Bing Yang, il en serait de même dans un fond de type *indica* pour la variété IR24. Ces différentes réponses suggèrent que le fond génétique pourrait avoir un effet sur la sensibilité à la maladie dans un contexte *sweet14* muté.

Il est connu que le génotype d'une variété influe sur l'efficacité ou la durabilité des résistances génétiques. Par exemple, dans le cas des QTL (Quantitative Trait Loci) de résistance, on observe souvent une additivité de l'effet des différents QTL présents dans le génome sur le niveau de résistance global. Mais il existe aussi des effets que l'on dit épistatiques, où l'influence d'un gène en particulier sur le niveau de résistance est différente selon les autres allèles auxquels il est associé dans un génotype donné (Gallois *et al.*, 2018 [306]). Chez le riz, il a été montré que la résistance apportée par la protéine NB-LRR Xa3/Xa26 était plus efficace dans le fond génétique *japonica* que dans les variétés *indica*, ceci étant dû à la présence d'un QTL associé à une expression accrue du gène chez *japonica* (Deng *et al.*, 2018 [307]). De la même manière, on pourrait imaginer qu'un QTL dans le génome Zhonghua ou IR24 apporte une ligne de défense supplémentaire ou potentialise l'effet de la mutation, en influant par exemple sur la quantité de sucres libres disponibles.

2 CRÉATION DE RÉSISTANCE CONTRE LA BACTÉRIOSE VASCULAIRE DU RIZ.

2.1 Méthodes d'amélioration du riz

2.1.1 Généralités

En diminuant ou supprimant totalement les intrants chimiques, l'amélioration végétale est un moyen écologique d'optimiser les rendements et de lutter contre les phytopathogènes.

Les méthodes classiques d'amélioration du riz consistent en l'introduction de caractéristiques choisies depuis une variété donneuse vers une variété élite receveuse. Cette introduction se fait par le biais de croisement puis de rétro-croisements et de sélection de la descendance. Le processus classique de sélection peut être très long et il peut s'écouler 8 à 9 ans avant la mise sur le marché d'une variété élite améliorée. Avec la venue de la Sélection Assistée par des Marqueurs (MAS), ce temps de sélection a été réduit à 5 à 6 années (Dossa *et al.*, 2015 [308]). Cette technique se base sur la sélection d'individus dont un trait facilement détectable est lié génétiquement à un QTL ou un gène d'intérêt. La sélection se fait généralement par des méthodes moléculaires et peut donc être effectuée à n'importe quel stade de développement ce qui augmente la rapidité du processus (Collard *et al.*, 2008 [309]). La MAS a permis le pyramidage de nombreux gènes de résistance chez le riz. La majorité de ces gènes proviennent de donneurs *O. sativa*, mais quelques-uns sont issus de variétés de riz sauvages, comme *O. longistaminata*, *O. rufipogon*, *O. minuta* et *O. officinalis* (Kumar *et al.*, 2020 [310]).

Plus récemment, l'application de la MAS aux rétro-croisements a encore augmenté la rapidité de production de variétés élites. Cette technique, nommée MAS Backcrossing (ou MABC) utilise les marqueurs afin d'effectuer des rétro-croisements avec la lignée parentale receveuse, pour se débarrasser plus rapidement et avec une précision plus fine des portions de génome non-souhaitées, provenant de la lignée donneuse (Collard *et al.*, 2008 [309]). Le MABC permet de conserver les caractéristiques souhaitées de la lignée parentale élite receveuse (haut rendement, meilleure qualité des grains…), tout en apportant les traits d'intérêt. Cette méthode a réduit le temps de sélection à 4 voir 2 ans (Dossa *et al.*, 2015 [308]).

La création de résistance contre les phytopahtogènes peut aussi être obtenue par transformation de gènes de résistance préalablement clonés. Cette technique est bien plus rapide que la MAS ou MABC, et s'affranchit des problèmes de « bruits de fond » générés par l'introgression fortuite de matériel génétique non voulu du donneur. Les premières lignées de riz transgéniques produites étaient issues de la transformation du gène de résistance *Xa21* dans un fond *japonica*. Cependant, pour des raisons de réglementation, ces lignées de riz n'ont pas pu être commercialisées. Aujourd'hui encore, les sélectionneurs préfèrent utiliser les techniques MAS ou MABC pour le développement de variétés résistantes à la BLB (Kumar *et al.*,2020 [310]).

2.1.2 L'édition de génome pour le gain de résistance à la BLB

Outre l'introduction classique de gènes de résistance ou de QTL, ou la création de lignées transgéniques, d'autres approches par le biais d'outils moléculaires permettent de modifier des variétés avec précision, en éditant une séquence d'ADN donnée. Ces techniques d'édition de génomes sont aujourd'hui réalisées par l'utilisation de TALEN (Transcription Activator Like Effector Nucleases) ou par le système CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats). Un TALEN est un TALE modifié auquel est fusionnée une nucléase, qui va induire une cassure double brin de la séquence reconnue par le TALE. Cette technique, mise au point par Christian *et al.* en 2010 a permis d'améliorer l'efficacité et la simplicité des éditions. Elle a par exemple été utilisée avec succès chez le

riz pour augmenter la stabilité des semences et la résistance aux herbicides (Siddiq and Lakshminarayana, 2021 [311]).

Le système CRISPR/Cas9 nécessite la co-expression d'une endonucléase Cas9 et d'un ou plusieurs ARN guides (ARNg) spécifiques. L'appariement du complexe Cas9/ARNg avec l'ADN cible déclenche le clivage de l'ADN par Cas9, entraînant une cassure double brin sur l'ADN. Cette cassure sera prise en charge par les systèmes de réparation de la plante, qui généreront une réparation parfaite, ou avec une erreur induite, conduisant généralement à l'insertion ou la délétion d'un nucléotide (Impens *et al.*, 2022 [312]). Cette technique est devenue l'outil d'édition le plus utilisé en génétique et a été employée chez le riz pour diverses améliorations, comme la résistance aux herbicides (Xu *et al.* 2014 [313]), la résistance à la pyriculariose (Wang *et al.*, 2016 [314]), ou l'augmentation de la teneur en amylose (Sun *et al.* 2017 [315]).

Un intérêt majeur de l'édition de génome, que ce soit par l'utilisation de TALENs ou du système CRISPR-Cas9, est que les produits générés sont dépourvus de tout matériel génétique étranger. Les variétés produites en utilisant ces méthodes ne sont donc pas considérées comme transgéniques (Siddiq and Vemireddy, 2021 [311]). Ainsi, les variétés de riz résistantes à la BLB créées dans le cadre du projet « Transformative Strategies For Controlling Rice Diseases Phase II » ont été générées par édition de génome via CRISPR-Cas9. D'autres avantages de cette technique sont que la Cas9 est une nucléase universelle ne nécessitant pas d'ingénierie préalable et que plusieurs cibles peuvent être éditées avec une seule construction par le biais de différents ARNg dont la conception est relativement simple (Tremblay, 2015 [316]).

Dans le cadre du projet, l'outil CRISPR-Cas9 a majoritairement été utilisé pour l'édition des EBEs dans les promoteurs de gènes *SWEET*. Mais plus récemment, à l'aide d'une nouvelle déclinaison de ce système, l'équipe d'un des membres du consortium Healthy Crops, Bing Yang, a généré des insertions d'EBEs reconnues par les TALEs naturels de *Xoo*, par exemple TalC, en amont d'un gène de résistance de type exécuteur, comme *Xa10* ou *Xa23* afin d'en élargir le spectre de reconnaissance (stratégie dite 'knock in').

2.2 Réglementation

Si les méthodes d'amélioration du riz par croisement sont exemptes de considérations sociétales, ce n'est pas le cas des plantes génétiquement modifiées. Les différentes controverses sur la sécurité alimentaire et les contraintes réglementaires propres à chaque pays ont mis un frein à l'utilisation de ces variétés (Kumar *et al.*, 2020 [310]). Cependant, l'arrivée des techniques d'édition de génomes par l'utilisation de Site-Directed Nucleases (SDN) comme CRISPR-Cas9 et les TALENs, a conduit de nombreux pays à réévaluer leur législation (Schmidt *et al.*, 2020 [317]).

Les éditions réalisées par SDN ont été classées en 3 catégories : les éditions classiques, donnant des indels imprévisibles d'un seul ou quelques nucléotides sont dans la catégorie SDN-1; les éditions générant des indels précis de quelques nucléotides à l'aide de la recombinaison homologue d'un petit fragment d'ADN sont dans la catégorie SDN-2; enfin, l'insertion de fragments d'ADN plus larges (typiquement des gènes entiers) entre dans la catégorie SDN-3 (Schmidt *et al.*, 2020 [317]). Il est à noter que des outils plus récents encore, consistant en une Cas9 modifiée et néofonctionnalisée à l'aide de domaines enzymatique hétérologues, peuvent permettre le remplacement d'une base par une autre de manière précise, sans recourir à la recombinaison homologue, une technique que l'on appelle base-editing (Buchholzer et Frommer, 2022 [318]).

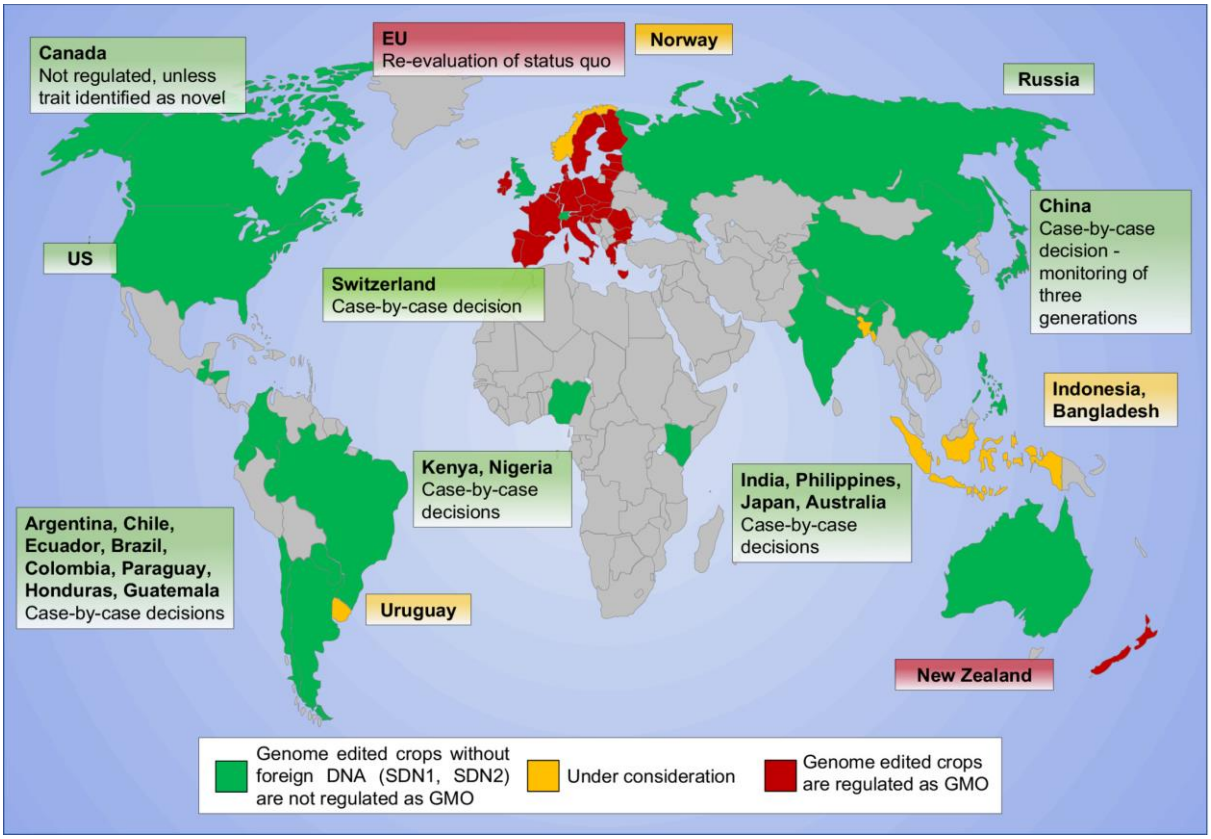


Figure 47 : État de la législation sur l'édition de génome dans le monde en juin 2022. Les pays en gris ne possèdent pas encore de régulation relatives aux organismes générés par biologie moléculaire. GMO, « Genetically Modified Organism ». Issu de Buchholzer et Frommer 2022 [318].

La question est donc de savoir comment réguler ces nouvelles variétés. Selon les pays, le débat est toujours en cours. Certains, comme la quasi-totalité des pays européens et la Nouvelle-Zélande, les placent directement dans la case OGM (figure 47). D'autres prennent en compte la modalité d'édition, et considèrent les lignées éditées sans transgènes (c'est-à-dire SDN-1 et SDN-2) comme non OGM. C'est le cas pour la Chine, l'Inde, l'Australie, la Russie, le Kenya, le Nigéria, et la quasi-totalité du continent Américain (figure 47, Buchholzer et Frommer, 2022 [318]).

Dans toute démarche d'amélioration végétale par édition, il faut donc prendre en compte l'avancée de la réglementation dans les pays ciblés, au risque de ne pas pouvoir répondre aux attentes de la population.

2.3 Méthodes de maximisation de la durabilité des résistances

Au-delà des pratiques de protection des cultures, comme le désherbage des parcelles pour éliminer les hôtes potentiels ou l'élimination des débris végétaux, la pérennité de la résistance passe par la gestion du matériel végétal, tant au niveau des constructions génétiques qu'à leur répartition dans l'espace et le temps.

Les pathogènes s'adaptent rapidement à leur niche écologique et sont capables d'outrepasser les résistances mises en place. Ces adaptations sont d'autant plus fréquentes s'il s'agit de résistances qualitatives, c'est-à-dire contrôlées par des gènes de résistance majeurs, car elles exercent une pression de sélection forte sur les populations. Pour cette raison, les résistances de type quantitatives ont gagné en popularité au fil des années.

Une résistance quantitative peut être définie comme un phénotype de résistance modéré (Pilet-Nayel *et al.*, 2017 [319]). Les propriétés de résistance des QTLs ont été catégorisées en différents groupes fonctionnels : les gènes entrant dans le développement et la morphologie de l'hôte, les gènes de la PTI, de détoxification, de transduction des signaux de défense, ou des allèles affaiblis de gènes de résistance majeurs (Poland *et al.*, 2009 [320]). Bien qu'il s'agisse de résistances partielles, il semblerait que les défenses fournies par les QTLs soient moins facilement érodées que les résistances qualitatives. Plusieurs hypothèses ont été émises quant à l'origine de ce phénomène : (i) une faible pression de sélection sur la population pathogène, (ii) des pressions de sélection contradictoires, (iii) une faible probabilité pour que l'agent pathogène puisse surmonter plusieurs QTLs, et (iv) une succession de mécanismes de résistance agissant à différentes temporalités (Pilet-Nayel *et al.*, 2017 [319])

Le type de déploiement est peut-être le facteur le plus influent pour la durabilité de la résistance (Dossa *et al.*, 2015 [308]). La probabilité d'un contournement de résistance est d'autant plus élevée que la variété portant le gène est utilisée en monoculture et de façon prolongée. La gestion durable des résistances passe donc par le déploiement de génotypes variés, ou de façon limitée dans le temps, afin de ne pas exercer une pression de sélection directionnelle de la population pathogène.

Elle s'exerce aussi directement au sein de la plante, en associant ou « pyramidant » différents gènes de résistance, ou en choisissant des constructions génétiques mêlant gènes de résistance majeurs et QTLs variés issus du fond génétique de la variété (Quenouille *et al.*, 2014 [321]). Le pyramidage de deux ou plus gènes de résistance diminue la probabilité pour le pathogène d'évoluer de manière à outrepasser les défenses de l'hôte. Par exemple, les combinaisons des gènes *Xa4*/*xa5* et *xa5*/*Xa21* offrent un niveau de résistance plus élevé que celui des lignées parentales, ou d'un gène seul (Kumar *et al.*, 2020 [310]). La mise en place de telles construction n'est pas aisée, car si la résistance apportée

par un gène de résistance ou QTL en particulier est soumise à un fort effet du fond génétique, les progénitures issues de croisements n'auront pas forcément une résistance aussi efficace, aussi durable ou un spectre d'action aussi étendu que les lignées parentales (Gallois *et al.*, 2018 [306]). De plus, l'éloignement des lignées parentales peut rendre difficile l'introduction des gènes d'intérêt, voir poser des problèmes de stérilité dans la descendance (Siddiq et Lakshminarayana 2021 [322]).

Enfin, une dernière bonne pratique pour maximiser la durabilité des résistances au champ est de ne pas donner un potentiel avantage sélectif aux variants bactériens en les exposant inutilement à des gènes de résistance inadaptés. Autrement dit, la surcharge de résistances ou le déploiement préventif de lignées résistantes dans des zones de production où elles ne sont pas nécessaires pour le contrôle de la BLB est à proscrire. On préfèrera la mise en place ciblée de résistances spécifiques aux populations de *Xoo* présentes dans chaque zone de production, afin de ne pas laisser l'opportunité au pathogène de renforcer sa virulence.

3 AU SUJET DE LA DURABILITÉ DES RÉSISTANCES CONTRE *XANTHOMONAS*

3.1 Émergence de nouveaux pathogènes

L'émergence de nouveaux pathogènes est d'une part facteur de la stochasticité évolutive et de l'environnement, et d'autre part modelée par l'activité humaine.

Tout d'abord, les aléas de l'évolution. L'évolution des pathogènes, par des mécanismes de mutation et de HGT, est à l'origine de nouveaux variants (voir 3.4.2 « Plasticité du génome et potentiel évolutif »). Ces variants peuvent constituer ou non une menace épidémique en fonction de la modification génétique apparue, selon le potentiel gain de virulence ou l'augmentation de fitness qu'elle apporte au pathogène. Les modifications du génome peuvent aussi conduire des organismes commensaux à devenir pathogènes, ou bien des pathogènes connus à effectuer un saut de niche écologique ou d'hôte, créant ainsi des populations inédites dans un modèle donné.

Les *Xo* étant capables de coloniser d'autres poacées que le riz, ou des riz sauvages, nous pourrions par exemple imaginer que des pathogènes de ces plantes non-cultivées puissent un jour s'attaquer au riz (ou inversement). Il est donc important de se préoccuper d'ores et déjà des populations microbiennes présentes dans les plantes hôtes au voisinage du riz. De même, il serait intéressant de mieux connaître le microbiote associé au riz et à sa rhizosphère.

Outre les potentiels sauts d'hôtes, la cohabitation de diverses bactéries et virus au sein d'une niche écologique est propice aux HGT. L'assimilation potentielle d'ADN étranger par la bactérie via transformation, conjugaison ou transduction va alors apporter de la nouveauté génétique et participer à l'évolution des génomes.

Nous savons que les pathovars *oryzae* et *oryzicola* peuvent cohabiter au sein d'une même lésion (Cunnac *et al.*, communication personnelle). La proximité génétique entre ces pathovars favorise la recombinaison potentielle de matériel génétique. De plus, nous avons montré au cours de cette étude qu'à l'instar des *Xoo* et *Xoc* asiatiques, les *Xoo* d'Afrique de l'Ouest possédaient des plasmides et des systèmes conjugatifs opérationnels (voir Chapitre I). Bien que l'échange d'ADN chromosomique par conjugaison soit peu fréquent et limité à l'espèce, le transfert de plasmides est considéré comme le principal moyen de HGT chez les microorganismes, et ce, même inter-règne (Frost, 2009 [323]). Notre pathobiome d'intérêt possède donc des moyens efficaces pour enrichir sa population.

Évidemment, l'établissement de ces nouvelles populations est soumis à la sélection par l'environnement. Si les modifications n'apportent aucun avantage à l'organisme, il ne prolifèrera pas, et dans le cas d'une modification désavantageuse, il sera éliminé. Si l'organisme est avantagé, sa persistance et / ou sa reproduction seront favorisées, et le nouveau génotype se répandra dans la population. Certains facteurs environnementaux fournissent même une pression de sélection que l'on appelle directionnelle, car orientant drastiquement la fréquence d'un génotype donné. Il peut s'agir d'un changement dans les conditions abiotiques, comme une augmentation des températures ou une baisse de l'humidité, ou bien d'un changement chimique, comme la présence d'un insecticide ou d'un composé antimicrobien, ou enfin d'un changement biologique, comme la présence d'une variété particulière au champ.

Dans le cas des phytopathogènes des cultures nous nous attarderons ici sur le dernier point. En effet, l'agriculture moderne consiste principalement en de larges monocultures, facilitant la récolte et la gestion du matériel végétal. Ces monocultures sont l'occasion pour les pathogènes de proliférer simplement dès lors qu'ils sont capables de se développer sur la variété utilisée. D'autre part la monoculture va engendrer une sélection directionnelle forte vers les génotypes bactériens capables de contourner les gènes de résistance mis en place. La monoculture favorise donc l'émergence de populations dangereuses pour les variétés cultivées.

La persistance d'un nouveau génotype est également soumise à une balance gain de fonction / coût de fitness. C'est par exemple ce qu'ont pu observer Vera Cruz *et al.* en 2000 [324] au cours de l'étude de l'agressivité et de la persistance des populations de *Xoo* présentes sur plusieurs saisons de plantation de la variété IRBB7. Cette lignée possède le gène de résistance *Xa7*, dont l'éliciteur est le TALE AvrXa7, inducteur du gène *OsSWEET14*. Bien que les souches possédant une version mutée d'*avrXa7* se montraient virulentes sur IRBB7, la pénalité de fitness associée à la perte de fonction du TALE était apparemment suffisante pour empêcher les épidémies de BLB sur les lignées de riz avec le gène *Xa7*.

Enfin, la dissémination de matériel végétal via le commerce ou les marchés des sélectionneurs est un facteur important dans l'introduction de populations pathogènes sur de nouveaux territoires. Nous avons d'ailleurs étés en mesure de détecter de tels évènements en Tanzanie et à Madagascar. Nous avons notamment proposé qu'à l'occasion d'une collaboration entre la Chine et la Tanzanie pour l'amélioration des systèmes d'irrigation sur les sites de Dakawa, l'utilisation de variétés chinoises dans des centres de formations a pu être à l'origine de l'introduction de souches de *Xoo* asiatiques dans le pays.

L'émergence de nouvelles populations pathogènes représente un risque évident pour les cultures, en particulier les monocultures. Dans le cas de l'apparition de génotypes avantagés, on peut être en situation de rupture de résistance, conduisant à la destruction partielle ou totale des récoltes et à un rétrécissement du panel de gènes de résistance efficaces utilisables en amélioration variétale.

3.2 Précisions sur l'évolution des TALomes

Comme présenté dans la partie 3.4.2.3 de l'introduction, les TALEs ont une capacité d'évolution importante. Cette capacité est notamment due à la présence de nombreuses régions répétées dans leur structure, et leur fréquente organisation en clusters. Ces caractéristiques sont propices aux recombinaisons, duplications, insertions et délétions, qui sont des mécanismes majeurs de l'évolution des génomes. De plus, il a été montré que comparé aux autres effecteurs de type TALEs retrouvés chez d'autres organismes (RipTALEs, BATs), les TALEs de Xanthomonas possèdent des répétitions aux séquences très conservées (Perez-Quintero et Szurek, 2019 [202]). Cette caractéristique est très favorable aux différents mécanismes de réarrangement précédemment cités, augmentant la

probabilité de la survenue de tels évènements. Le potentiel évolutif des TALomes de Xanthomonas est donc considéré important, permettant l'apparition fréquente de nouveaux variants de TALEs chez la bactérie, lui offrant une rapidité d'adaptation remarquable.

Cette rapidité a d'ailleurs été mise en évidence chez *X. citri* par Teper et Wang en 2021 [201], où 30 cycles d'infections et de ré-inoculations successives de souches non virulentes (*TALEs* mutés) ont suffi à faire émerger des variants TALEs capables d'induire *CsLOB1*, qui est le gène de sensibilité majeur dans la genèse du chancre des agrumes.

Les potentiels transferts de gènes horizontaux entre souches de *Xo* s'ajoutent à la plasticité naturelle des TALEs dans l'évolution des TALomes. En effet, comme expliqué précédemment, *Xoo* et *Xoc* peuvent cohabiter au sein d'une même lésion, et il a été montré par A. Diallo au cours de ses travaux de thèse sur la caractérisation des souches de *Xoo* au Burkina Faso que différents génotypes, présentant différents TALomes, pouvaient être présents sur une même parcelle. Ainsi, le transfert de TALEs entre différentes souches voir différents pathovars et d'éventuelles recombinaisons entre ces effecteurs constitue une source d'armement supplémentaire pour *Xoo*.

L'émergence de souches de *Xoo* possédant des variants de TALEs leur permettant de contourner des résistances mises en place chez le riz a d'ailleurs été d'ores et déjà documentée. Par exemple, des souches isolées en Corée du sud se sont révélées virulentes sur des lignées multi-éditées dans l'ensemble des EBEs des gènes *SWEET* connus alors. Cette capacité d'outrepasser ces allèles résistants par perte de sensibilité s'est avérée due à la présence dans leur TALome de variants du TALE PthXo2 qui leur permettent d'induire *SWEET13* (Oliva *et al.*, 2019 [287]). Des analyses préliminaires d'évolution expérimentale in silico des TALEs menées par notre collègue A. Perez-Quintero ont effectivement révélé que la stochasticité de l'évolution des TALomes pourrait très bien conduire au contournement des éditions réalisées dans le projet, par l'utilisation de nouvelles EBEs. Outre l'apparition de variants capables de cibler de nouvelles EBEs, l'évolution des TALomes a aussi donné lieu à l'apparition de TALEs tronqués, les iTALEs (voir introduction), ayant pour rôle de bloquer tout bonnement la reconnaissance des TALEs par les gènes de résistances *Xa1*/*Xo1* chez le riz.

Les souches de *Xoo* africaines possèdent un TALome plus réduit, et moins diversifié que les *Xoo* asiatiques (Doucouré *et al.*, 2018 [293] ; Lang *et al.*, 2019 [123]). Cependant, ces TALEs sont aussi sujets à recombinaison d'autant plus que la diversité de leurs répétitions est plus faible. De plus, l'introduction de nouveaux génotypes depuis l'Asie en Afrique de l'Est dont nous avons été témoins, renforce la menace d'épidémies de BLB sur le continent Africain, car met en péril les résistances en place.

Ces différents résultats soulèvent la fragilité relative des résistances basées sur les TALEs (ou TALE-ETI), dont l'efficacité peut être rapidement mise à l'épreuve, notamment car elles représentent une pression de sélection forte. Nos approches moléculaires pour le gain de résistance contre *Xoo* doivent donc prendre en compte dès à présent la future rupture des résistances. Cette prise en compte comprends : i) le design d'autres allèles de résistance en préventif, ii) la mise en place raisonnée des résistances au champs, et iii) l'éventuel pyramidage des résistances quantitatives et qualitatives, comme c'est le cas dans la lignée Komboka éditée au cours de ce projet (voir Chapitre II).

4 FUTURS DÉFIS ET PERSPECTIVES

4.1 Changement climatique

La phytopathologie définit les maladies des plantes comme le résultat de l'interaction entre le pathogène, l'hôte et l'environnement, ces deux derniers formant la niche écologique du pathogène. Cette niche détermine la zone géographique et la période de l'année qui permettent aux pathogènes d'attaquer les cultures compatibles. Un changement dans l'un de ces paramètres peut donc modifier le spectre d'hôtes ou l'apparition dans le temps et l'espace de la maladie (Chaloner *et al.*, 2020 [325]).

Plusieurs études ont notamment montré que les conditions abiotiques avaient une influence sur l'efficacité des gènes de résistance à la BLB. Par exemple, la lignée IRBB4 (*Xa4*) est beaucoup plus sensible à des températures de 35/27°C (jour/nuit) que 29/21°C (Kumar *et al.*, 2020 [310]). Le stress hydrique est lui aussi connu pour influer sur l'efficacité des défenses du riz. La lignée IRBB7 (*Xa7*) s'est montrée plus résistante à *Xoo* en conditions de sécheresse qu'en conditions normales (Dossa *et al.*, 2017 [326]).

Ces résultats suggèrent que nous devons nous tenir prêts à trouver d'éventuelles parades à l'affaiblissement potentiel des résistances du riz, soit en modulant les conditions abiotiques, soit en changeant de résistance génétique dans les zones concernées.

Les changements environnementaux vont également fortement influer sur l'apparition de nouvelles populations pathogènes. Différentes hypothèses sur l'émergence de populations liée au changement climatique ont été émises par Morris *et al.* en 2017 [327]: (i) l'apparition de conditions favorables à la résurgence d'anciens pathogènes; (ii) un déplacement des populations pathogènes pour satisfaire leur niche écologique, iii) la dissémination des agents pathogènes via l'augmentation des évènements météorologiques violents, (iv) l'expansion de la gamme d'hôtes des agents pathogènes liée à l'apparition de conditions satisfaisants leur niche, et (v) de nouveaux agents pathogènes résultant de HGT entre des souches devenues voisines.

L'augmentation de la probabilité d'émergence de nouvelles populations pathogènes liée au changement climatique constituera donc un défi pour les sciences épidémiologiques dont le rôle sera crucial dans l'anticipation et la réponse aux maladies.

4.2 Nécessité d'un suivi épidémiologique régulier des populations de *Xoo*

Afin de protéger efficacement les cultures contre les génotypes émergents et de ne pas fragiliser inutilement les résistances du riz en les disposant à l'aveugle, il est impératif d'avoir une vue d'ensemble des populations de *Xoo* présentes au champ. Cette connaissance passe par un suivi épidémiologique rigoureux de la BLB.

Certaines connaissances basiques de l'épidémiologie de la BLB sont toujours manquantes ou floues. C'est le cas par exemple de la transmission et de la dissémination du pathogène. Notamment, la question de la transmission via les graines doit être élucidée. Bien qu'encore non formellement prouvée, il semble plus que probable que *Xoo* soit capable de rester en dormance sur les graines puis d'infecter la plante au moment de sa germination. Cette hypothèse est d'autant plus plausible lorsque l'on est confronté à des mouvements de populations sur de très longues distances, ou l'émergence soudaine du pathogène sur des territoires isolés comme nous avons pu l'observer à Madagascar et en Tanzanie. Effectivement, il paraît très peu crédible que ces diffusions de populations de *Xoo* soient le résultat d'évènements météorologiques ayant transporté des bactéries viables depuis la Chine ou la Thaïlande jusqu'au continent Africain.

Nous manquons également de données statistiques systématiques sur l'incidence de la maladie et les pertes de rendements lui étant associées, toutes localisations confondues. Ces données, collectées au cours des différentes saisons de récoltes seraient de précieux indicateurs de l'émergence de nouvelles populations, et renseigneraient sur l'efficacité des résistances et des pratiques de contrôle de la maladie mises en place au champ (Straub *et al.*, 2021 [328]). Ces renseignements permettraient de diriger de manière optimale les efforts des sélectionneurs pour la mise au point et le déploiement des variétés (Dossa *et al.* 2015 [308]).

La compréhension de la structure génétique des populations de *Xoo* par l'utilisation de marqueurs moléculaires ou de séquençages de génomes complets peut nous permettre de retracer l'origine et les voies d'introduction du pathogène, et d'estimer sa rapidité d'évolution. Ces connaissances pourront servir à la prévention de futures invasions (Straub *et al.*, 2021 [328]).

Le rôle majeur des TALEs dans la virulence et la spécificité d'hôtes de *Xoo* rend centrale l'étude des TALomes au champ. Les avancées techniques du séquençage « long reads » permettent d'envisager une caractérisation systématique des TALomes des échantillons prélevés, ouvrant la voie à un suivi précis du potentiel de virulence de la population (Dossa *et al.*, 2015 [308]).

D'autres caractérisations, au niveau du phénotype de virulence des souches prélevées, permettent de déduire les capacités infectieuses des populations de *Xoo*. Notamment, la détermination de leur profil de virulence sur une sélection de NILs informe des gènes de résistance capables de les contrôler. Le diagnostic des gènes *SWEET* du clade III utilisés pour l'infection par le biais du suivi des inductions en qPCR, ou de tests de pathogénie sur les différents lignées KO *sweet* et rapportrices GUS du projet, signale quant à lui quelles sont les cibles à modifier pour empêcher la prolifération bactérienne chez le riz, du moins pour les *Xoo* de type asiatique.

La centralisation de toutes ces données d'incidence, de races, de phylogénies, de gènes de sensibilité majeurs et de TALomes dans un outil associant la localisation GPS des isolats, fournirait aux sélectionneurs une force de frappe inédite pour la mise en place de réponses variétales adaptées au niveau local.

Un outil de ce type, nommé Pathotracer, a été créé au cours du projet mené par Healthy Crops (Eom *et al.*, 2019 [288]). PathoTracer intègre les données GPS, génotypiques et phénotypiques des *Xoo* collectées et permet de visualiser quels *SWEETs* sont utilisés. Cette plateforme suggère également les variétés éditées efficaces à planter lors de la prochaine saison.

La collection des données nécessaires à la surveillance et la caractérisation du pathogène devra impliquer à la fois les producteurs, les sélectionneurs et les organismes de recherche pour aboutir à une démarche intégrative et efficace dans le contrôle de la BLB au champ.

4.3 Recherche de nouvelles sources de résistance

La surveillance des populations de *Xoo* constitue la première condition pour une protection des cultures optimale. La seconde est bien entendu la mise en service de sources de résistance, permettant le contrôle des nouveaux variants. Une manière de procéder est la création de résistance par voies moléculaires, comme précédemment exposé. Dans notre pathosystème, le design de ces résistances repose principalement sur le fonctionnement des TALEs, par la modification ou l'introduction d'EBEs. Ce type de résistance qualitative, en étant correctement déployée (rotations de génotypes dans le temps, cultures multi-variétés, apport de résistance ciblée en réponse aux populations bactériennes) peut voir sa durée de vie augmentée et son efficacité maximisée dans le temps.

Néanmoins, *Xoo* a un potentiel évolutif fort, notamment au niveau de son Talome qui possède par nature, une forte fluidité génétique. Nous ne devons donc pas négliger la recherche d'autres sources de résistance, notamment de type quantitatives.

Il faudra donc mettre à profit le pool génétique du riz et étudier les variations génétiques disponibles pour l'amélioration des variétés cultivées (Kumar *et al.*, 2020 [310]). Bien qu'il existe une grande diversité de riz sauvages, variétés traditionnelles et poacées proches du riz, la diversité génétique des cultivars agricoles est relativement faible. En effet, une poignée de lignées parentale a fourni le matériel génétique de 80% des variétés cultivées actuellement (Siddiq et Lakshminarayana 2021 [322]). La grande majorité du pool génétique du riz reste donc inutilisée. Pourtant, de nombreux gènes et QTLs de résistance ont été identifiés dans des espèces de riz sauvages (voir Tableau 3 dans l'introduction).

Il est donc inévitable d'avancer dans la connaissance du capital de résistance disponible dans le patrimoine génétique global du riz, pour se prémunir des potentiels contournement de résistance à venir.

REFERENCES

1. Lu, B. Taxonomy of the genus Oryza (Poaceae): historical perspective and current status. *IRRN Int. Rice Res. Notes* **24**, 4–8 (1999).
2. Chang, T.-T. & Bardenas, E. A. THE MORPHOLOGY AND VARIETAL CHARACTERISTICS OF THE RICE PLANT. *IRRI Tech. Bull.* **4**, 40 (1965).
3. Coste, H. & Flahault, C. *Flore descriptive et illustrée de la France, de la Corse et des contrées limitrophes*. (Librairie des sciences et des arts, 1937).
4. Clement, G. & Châtel, M. LA SELECTION DU RIZ AU CIRAD. 14 (2002).
5. THE ANGIOSPERM PHYLOGENY GROUP. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III: APG III. *Bot. J. Linn. Soc.* **161**, 105–121 (2009).
6. Xu, Z. & Zhou, G. *Identification and Control of Common Weeds: Volume 1*. (Springer, 2017).
7. Soreng, R. J. *et al.* A worldwide phylogenetic classification of the Poaceae (Gramineae): Phylogenetic classification of the grasses. *J. Syst. Evol.* **53**, 117–137 (2015).
8. Linnaeus, Carolus. *Species plantarum: exhibentes plantas rite cognitas ad genera relatas, cum diferentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas. Tomus 1*. (Zenodo, 1753). doi:10.5281/ZENODO.3931989.
9. Chen, E., Huang, X., Tian, Z., Wing, R. A. & Han, B. The Genomics of *Oryza* Species Provides Insights into Rice Domestication and Heterosis. *Annu. Rev. Plant Biol.* **70**, 639–665 (2019).
10. Vaughan, D. A., Morishima, H. & Kadowaki, K. Diversity in the Oryza genus. *Curr. Opin. Plant Biol.* **6**, 139–146 (2003).
11. Stein, J. C. *et al.* Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nat. Genet.* **50**, 285–296 (2018).
12. Kumagai, M. *et al.* Rice Varieties in Archaic East Asia: Reduction of Its Diversity from Past to Present Times. *Mol. Biol. Evol.* **33**, 2496–2505 (2016).

13. Chen, L. *et al.* Unveiling a Novel Source of Resistance to Bacterial Blight in Medicinal Wild Rice, Oryza officinalis. *Life Basel Switz.* **12**, 827 (2022).

14. Doebley, J. F., Gaut, B. S. & Smith, B. D. The Molecular Genetics of Crop Domestication. *Cell* **127**, 1309–1321 (2006).

15. Hammer, K. Das Domestikationssyndrom. *Kult.* **32**, 11–34 (1984).

16. Allaby, R. G. Domestication Syndrome in Plants. in *Encyclopedia of Global Archaeology* (ed. Smith, C.) 2182–2184 (Springer New York, 2014). doi:10.1007/978-1-4419-0465-2_2416.

17. Kesh, H. & Kaushik, P. Impact of Marker Assisted Breeding for Bacterial Blight Resistance in Rice: A Review. *Plant Pathol. J.* **19**, 151–165 (2020).

18. Hilbert, L. *et al.* Evidence for mid-Holocene rice domestication in the Americas. *Nat. Ecol. Evol.* **1**, 1693–1698 (2017).

19. Cubry, P. *et al.* The Rise and Fall of African Rice Cultivation Revealed by Analysis of 246 New Genomes. *Curr. Biol.* **28**, 2274-2282.e6 (2018).

20. Sasaki, A. *et al.* Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature* **416**, 701–702 (2002).

21. Muthayya, S., Sugimoto, J. D., Montgomery, S. & Maberly, G. F. An overview of global rice production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* **1324**, 7–14 (2014).

22. Xu, K. *et al.* Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705–708 (2006).

23. Toledo, A. M. U. *et al.* Development of improved Ciherang-Sub1 having tolerance to anaerobic germination conditions. *Plant Breed. Biotechnol.* **3**, 77–87 (2015).

24. Gutaker, R. M. *et al.* Genomic history and ecology of the geographic spread of rice. *Nat. Plants* **6**, 492–502 (2020).

25. Choi, J. Y. & Purugganan, M. D. Multiple Origin but Single Domestication Led to Oryza sativa. *G3 GenesGenomesGenetics* **8**, 797–803 (2018).

26. Alam, O. *et al.* Genome Analysis Traces Regional Dispersal of Rice in Taiwan and Southeast Asia. *Mol. Biol. Evol.* **38**, 4832–4846 (2021).

27. The 3,000 rice genomes project. The 3,000 rice genomes project. *GigaScience* **3**, 7 (2014).

28. Civáň, P. *et al.* Origin of the *Aromatic* Group of Cultivated Rice (*Oryza sativa* L.) Traced to the Indian Subcontinent. *Genome Biol. Evol.* **11**, 832–843 (2019).

29. Chen, Q., Li, W., Tan, L. & Tian, F. Harnessing Knowledge from Maize and Rice Domestication for New Crop Breeding. *Mol. Plant* **14**, 9–26 (2021).

30. Mew, T.-W. & Vera Cruz, C. Part I. General Introduction, Section 2. The "Shifty Enemies" of Rice Production in Asia. in *Rice Diseases: Biology & Selected Management Practices* (Rina Opulencia and G.P. Hettel, 2018).

31. Wang, M. *et al.* The genome sequence of African rice (Oryza glaberrima) and evidence for independent domestication. *Nat. Genet.* **46**, 982–988 (2014).

32. Murray, S. S. Searching for the origins of African rice domestication. 5.

33. Konishi, S. *et al.* An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392–1396 (2006).

34. Win, K. T. *et al.* A single base change explains the independent origin of and selection for the nonshattering gene in African rice domestication. *New Phytol.* **213**, 1925–1935 (2017).

35. Prasad, R., Shivay, Y. S. & Kumar, D. Current Status, Challenges, and Opportunities in Rice Production. in *Rice Production Worldwide* (eds. Chauhan, B. S., Jabran, K. & Mahajan, G.) 1–32 (Springer International Publishing, 2017). doi:10.1007/978-3-319-47516-5_1.

36. Courtois, B. Une brève histoire du riz et de son amélioration génétique. 13 (2007).

37. FAO. *FAO Rice Market Monitor (RMM), April 2018, VOLUME XXI ISSUE No. 1*. (FAO, 2018).

38. USDA. Grain: World Markets and Trade. *Glob. Mark. Anal.* (2022).

39. Elert, E. Rice by the numbers: A good grain. *Nature* **514**, S50–S51 (2014).

40. Timmer, C. P. The Changing Role of Rice in Asia's Food Security. 24.

41. Sié, M. *et al.* Towards a Rational Use of African Rice (Oryza glaberrima Steud.) for Breeding in Sub-Saharan Africa. 7.

42. Zenna, N., Senthilkumar, K. & Sie, M. Rice Production in Africa. in *Rice Production Worldwide* (Springer International Publishing, 2017).

43. Seck, P. A., Touré, A. A., Coulibaly, J. Y., Diagne, A. & Wopereis, M. Africa's rice economy before and after the 2008 rice crisis. *Realizing Africa's rice promise* https://agritrop.cirad.fr/572523/ (2013).

44. Tondel, F., D'Alessandro, C., Hathie, I. & Blancher, C. Rice trade and value chain development in West Africa: an approach for more coherent policies. (2020).

45. AfricaRice. 50 YEARS OF RICE RESEARCH FOR DEVELOPMENT. (2022).

46. Venuprasad, R., Lafitte, H. R. & Atlin, G. N. Response to Direct Selection for Grain Yield under Drought Stress in Rice. *Crop Sci.* **47**, 285–293 (2007).

47. Yang, X., Wang, B., Chen, L., Li, P. & Cao, C. The different influences of drought stress at the flowering stage on rice physiological traits, grain yield, and quality. *Sci. Rep.* **9**, 3742 (2019).

48. Pathak, H., Kumar, M., Molla, K. & Chakraborty, K. Abiotic stresses in rice production: Impacts and management. *ORYZA- Int. J. Rice* **58**, 103–125 (2021).

49. Gomez-Zavaglia, A., Mejuto, J. C. & Simal-Gandara, J. Mitigation of emerging implications of climate change on food production systems. *Food Res. Int. Ott. Ont* **134**, 109256 (2020).

50. Kumar, M., Sandhya, Kumar, P., Gaurav Singh, A. & Kumar Jukanti, A. Abiotic Stress Tolerance in Rice: Insight in Climate Change Scenario. in *Integrative Advances in Rice Research* (ed. Huang, M.) (IntechOpen, 2022). doi:10.5772/intechopen.98909.

51. Colmer, T. D. & Pedersen, O. Underwater photosynthesis and respiration in leaves of submerged wetland plants: gas films improve CO2 and O2 exchange. *New Phytol.* **177**, 918–926 (2008).

52. Flowers, T. J. Improving crop salt tolerance. *J. Exp. Bot.* **55**, 307–319 (2004).

53. Umeh, E. D. N., Joshi, R. C. & Ukwungwu, M. N. Field insect pests of rice in Africa: biology and control. *IITA Res. Guide* **43**, 32 (1995).

54. Pathak, M. D. & Khan, Z. R. *Insect pests of rice*. (International Rice Research Institute, 1994).

55. Laha, G. S. *et al.* Importance and Management of Rice Diseases: A Global Perspective. in *Rice Production Worldwide* (eds. Chauhan, B. S., Jabran, K. & Mahajan, G.) 303–360 (Springer International Publishing, 2017). doi:10.1007/978-3-319-47516-5_13.

56. Uehara-Ichiki, T. *et al.* Detection and diagnosis of rice-infecting viruses. *Front. Microbiol.* **4**, 289 (2013).

57. Séré, Y. *et al.* Managing the major diseases of rice in Africa. in *Realizing Africa's rice promise* (eds. Wopereis, M. C. S., Johnson, D. E., Ahmadi, N., Tollens, E. & Jalloh, A.) 213–228 (CABI, 2013). doi:10.1079/9781845938123.0213.

58. Dean, R. *et al.* The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* **13**, 414–430 (2012).

59. Koiso, Y. *et al.* Isolation and structure of an antimitotic cyclic peptide, ustiloxin F: chemical interrelation with a homologous peptide, ustiloxin B. *J. Antibiot. (Tokyo)* **51**, 418–422 (1998).

60. Khan, M. R. & Ahamad, F. Incidence of Root-Knot Nematode (Meloidogyne graminicola) and Resulting Crop Losses in Paddy Rice in Northern India. *Plant Dis.* **104**, 186–193 (2020).

61. Thio, B., Ouedraogo, L. S., Sanon, E., Sankara, P. & Kiemde, S. Les nématodes parasites associés au riz dans les trois (3) principales écologies rizicoles au Burkina Faso. *Int. J. Biol. Chem. Sci.* **11**, 1178–1189 (2017).

62. Johnson, D. E. *Les adventices en riziculture en Afrique de l'Ouest =: Weeds of rice in West Africa*. (Association pour le développement de la riziculture en Afrique de l'Ouest ; Centre technique de coopération agricole et rurale ; Dép. pour le développement international, 1997).

63. Dongya, W., Sangting, L. & Longjiang, F. De-Domestication: An Extension of Crop Evolution | Elsevier Enhanced Reader. https://reader.elsevier.com/reader/sd/pii/S1360138521000327?token=BA4FDEE05357607BB6

28C18462C313CCB6363E16AEC1F2E648253BA8E34BF2555640763C0ED47A17E71A77493DE6A841&originRegion=eu-west-1&originCreation=20220708131906 (2021) doi:10.1016/j.tplants.2021.02.003.

64. Takeuchi, H. Occurrence of bacterial blight of rice. *Bull. Agric. Exp. Stn. Gov.-Gen. Chosen* (1930).

65. *Bacterial blight of rice: proceedings of the International Workshop on Bacterial Blight of Rice, 14-18 March 1988*. (IRRI, 1989).

66. Naqvi, S. A. H. Bacterial Leaf Blight of Rice: An Overview of Epidemiology and Management with Special Reference to-Indian-Sub-Continent. *Pak. J. Agric. Res.* **32**, (2019).

67. Buddenhagen, I. W., Vuong, H. H. & Ba, D. D. Bacterial blight found in Africa. *Int. Rice Res. Newsl.* **4**, (1979).

68. Reckhaus, P. M. Occurrence of Bacterial Blight of Rice in Niger, West Africa. *Plant Dis.* **67**, 1039 (1983).

69. Aldrick, S. J., Buddenhagen, I. W. & Reddy, A. P. K. The occurrence of bacterial leaf blight in wild and cultivated rice in northern Australia. *Aust. J. Agric. Res.* **24**, 219–227 (1973).

70. Lozano, J. C. Identification of bacterial leaf blight in rice, caused by Xanthomonas oryzae, in America [Central America, South America]. *Plant Dis. Report. USA* (1977).

71. Jeger, M. *et al.* Pest categorisation of Xanthomonas oryzae pathovars oryzae and oryzicola. *EFSA J.* **16**, e05109 (2018).

72. Naidu, V. D., Rao, B. S. & Rao, C. S. Effects of Nitrogen Nutrition and Bacterial Leaf Blight on Rice Leaves. *J. Phytopathol.* **96**, 83–86 (1979).

73. Reddy, A. P. K. Bacterial blight: crop loss assessment and disease management. in (IRRI, 1989).

74. Mew, T. W. Focus on Bacterial Blight of Rice. *Plant Dis.* **77**, 5 (1993).

75. Noh, T.-H. *et al.* Effects of Bacterial Leaf Blight Occurrence on Rice Yield and Grain Quality in Different Rice Growth Stage. *Res. Plant Dis.* **13**, 20–23 (2007).

76. Ou, S. H. *Rice Diseases*. (IRRI, 1985).

77. Exconde, O. R. & Opina, O. S. Yield losses due to bacterial leaf blight of rice. *Philipp. Agric.* (1973).

78. Adhikari, T., Leach, J. & Mew, T. W. Bacterial blight of rice and its control in Nepal. *Integr. Pest Manag. Rev.* **1**, 91–95 (1996).

79. Sere, Y. *et al.* Rice Bacterial Leaf Blight in West Africa: Preliminary Studies on Disease in Farmers` Fields and Screening Released Varieties for Resistance to the Bacteria. *Asian J. Plant Sci.* **4**, 577–579 (2005).

80. Tall, H. *et al.* Confirmation Report of Bacterial Blight Caused by Xanthomonas oryzae pv. oryzae on Rice in Senegal. *Plant Dis.* **104**, 968–968 (2019).

81. Corral, R., Leach, J. E. & Vera Cruz, C. Recovery Plan for Xanthomonas oryzae Causing Bacterial Blight and Bacterial Leaf Streak of Rice. (2013).

82. CABI & EPPO. Xanthomonas oryzae. (2002).

83. PM 3/78 (2) Consignment inspection of seed and grain of cereals. *EPPO Bull.* **51**, 387–396 (2021).

84. Niño-Liu, D. O., Ronald, P. C. & Bogdanove, A. J. Xanthomonas oryzae pathovars: model pathogens of a model crop. *Mol. Plant Pathol.* **7**, 303–324 (2006).

85. Mew, T. W., Cottyn, B. & Vera Cruz, C. Part II. Biology of Rice Diseases Section 2: Bacterial Diseases. in *Rice Diseases: Biology & Selected Management Practices* (Rina Opulencia and G.P. Hettel, 2018).

86. Mizukami, T. & Wakimoto, S. Epidemiology and Control of Bacterial Leaf Blight of Rice. *Annu. Rev. Phytopathol.* **7**, 51–72 (1969).

87. Fang, C., Liu, C.-F. & Chu, C. A PRELIMINARY STUDY ON THE DISEASE CYCLE OF THE BACTERIAL LEAF BLIGHT OF RICE. *Acta Phytopathol. Sin.* (1956).

88. Reitsma, J. & Schure, P. S. J. *'Kresek': a bacterial disease of rice*. (Archipel, 1950).

89. Reddy, P. R. Kresek phase of bacterial blight of rice. *Oryza* **21**, 179–187 (1984).

90. T, N., N, S. & Z, S. Occurrence of kresek symptoms [wilting] of bacterial leaf blight on rice seedlings [Oryza sativa] injured by rice plant weevil, Echinocnemus squameus Billberg. *Ann. Phytopathol. Soc. Jpn.* (1981).

91. Cottyn, B., Cerez, M. T. & Mew, T. W. Chapter 7 : Bacteria. in *A manual of rice seed health testing* (1994).

92. Benedict, A. A. *et al.* Pathovar-Specific Monoclonal Antibodies for Xanthomonas campestris pv. oryzae and for Xanthomonas campestris pv. oryzicola. *Phytopathology* **79**, 322 (1989).

93. Lane, D. J. 16S/23S rRNA sequencing,. in *Nucleic acid techniques in bacterial systematics.* 115–175 (John Wiley and Sons, 1991).

94. Lang, J. M. *et al.* Genomics-Based Diagnostic Marker Development for Xanthomonas oryzae pv. oryzae and X. oryzae pv. oryzicola. *Plant Dis.* **94**, 311–319 (2010).

95. Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.* **28**, e63 (2000).

96. Lang, J. M. *et al.* Sensitive Detection of Xanthomonas oryzae Pathovars oryzae and oryzicola by Loop-Mediated Isothermal Amplification. *Appl. Environ. Microbiol.* **80**, 4519–4530 (2014).

97. Noda, T. & Kaku, H. Growth of Xanthomonas oryzae pv. oryzae In Planta and in Guttation Fluid of Rice. *Jpn. J. Phytopathol.* **65**, 9–14 (1999).

98. Devadath, S. & Dath, A. P. Infected chaff as a source of inoculum of Xanthomonas campestris pv. oryzae to the rice crop / Infizierte Spreu als Infektionsquelle für Xanthomonas campestris pv. oryzae an Reis. *Z. Für Pflanzenkrankh. Pflanzenschutz J. Plant Dis. Prot.* **92**, 485–488 (1985).

99. Sakthivel, N., Mortensen, C. & Mathur, S. Detection of Xanthomonas oryzae pv. oryzae in artificially inoculated and naturally infected rice seeds and plants by molecular techniques. *Appl. Microbiol. Biotechnol.* **56**, 435–441 (2001).

100. Thri Murty, V. S. & Devadath, S. Role of seed in survival and transmission of Xanthomonas campestris pv. oryzae causing bacterial blight of rice. *Phytopathol. Z. Ger. FR* (1984).

101. Singh, D., Vinther, F. & Mathur, S. B. Seed transmission of bacterial leaf blight in rice. *Seed Pathol. News* 11 (1983).

102. Mondal, S. *et al.* Survival and transmission of Xanthomonas oryzae pv. oryzae in rice seeds. *Fundam. Appl. Agric.* **4**, 680 (2019).

103. T.W. Mew, B. Cottyn, & C.M. Vera Cruz. Section 2. Bacterial diseases. in *Rice diseases: their Biology and selected management practices.* 100 (2018).

104. EPPO. Xanthomonas oryzae pv. oryzae (XANTOR), Hosts. *EPPO Global Database* https://gd.eppo.int/taxon/XANTOR/hosts (2022).

105. Assis, R. A. B. *et al.* A comparative genomic analysis of Xanthomonas arboricola pv. juglandis strains reveal hallmarks of mobile genetic elements in the adaptation and accelerated evolution of virulence. *Genomics* **113**, 2513–2525 (2021).

106. Huang, C.-J. *et al.* Comparative Genomic Analysis Uncovered Evolution of Pathogenicity Factors, Horizontal Gene Transfer Events, and Heavy Metal Resistance Traits in Citrus Canker Bacterium Xanthomonas citri subsp. citri. *Front. Microbiol.* **12**, 2605 (2021).

107. An, S.-Q. *et al.* Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen Xanthomonas. *FEMS Microbiol. Rev.* **44**, 1–32 (2019).

108. Cirac, A. D. *et al.* Rational Design of Cyclic Antimicrobial Peptides Based on BPC194 and BPC198. *Mol. Basel Switz.* **22**, E1054 (2017).

109. Inui Kishi, R. N. *et al.* Evaluation of cytotoxicity features of antimicrobial peptides with potential to control bacterial diseases of citrus. *PloS One* **13**, e0203451 (2018).

110. Camó, C. *et al.* Antimicrobial peptide KSL-W and analogues: Promising agents to control plant diseases. *Peptides* **112**, 85–95 (2019).

111. Li, Y. *et al.* Synthesis and biological evaluation of stilbene-based peptoid mimics against the phytopathogenic bacterium Xanthomonas citri pv. citri. *Pest Manag. Sci.* **77**, 343–353 (2021).

112. Kanugala, S. *et al.* Chumacin-1 and Chumacin-2 from Pseudomonas aeruginosa strain CGK-KS-1 as novel quorum sensing signaling inhibitors for biocontrol of bacterial blight of rice. *Microbiol. Res.* **228**, 126301 (2019).

113. Yasmin, S. *et al.* Biocontrol of Bacterial Leaf Blight of Rice and Profiling of Secondary Metabolites Produced by Rhizospheric Pseudomonas aeruginosa BRp3. *Front. Microbiol.* **8**, (2017).

114. Sanya, D. R. A. *et al.* A review of approaches to control bacterial leaf blight in rice. *World J. Microbiol. Biotechnol.* **38**, 113 (2022).

115. Hop, D. V. *et al.* Biological Control of Xanthomonas Oryzae pv. Oryzae Causing Rice Bacterial Blight Disease by Streptomyces toxytricini VN08-A-12, Isolated from Soil and Leaf-litter Samples in Vietnam. *Biocontrol Sci.* **19**, 103–111 (2014).

116. Vidhyasekaran, P. *et al.* Induction of systemic resistance byPseudomonas fluorescens Pf1 againstXanthomonas oryzae pv.Oryzae in rice leaves. *Phytoparasitica* **29**, 155–166 (2001).

117. Chae, J.-C., Hung, N. B., Yu, S.-M. & Lee, H. K. L. and Y. H. Diversity of Bacteriophages Infecting Xanthomonas oryzae pv. oryzae in Paddy Fields and Its Potential to Control Bacterial Leaf Blight of Rice. **24**, 740–747 (2014).

118. Williams, K. P. *et al.* Phylogeny of Gammaproteobacteria. *J. Bacteriol.* **192**, 2305–2314 (2010).

119. Jacques, M.-A. *et al.* Using Ecology, Physiology, and Genomics to Understand Host Specificity in Xanthomonas. *Annu. Rev. Phytopathol.* **54**, 163–187 (2016).

120. Pruvost, O. *et al.* The Dynamic World of the Genus Xanthomonas. in *Virulence Mechanisms of Plant-Pathogenic Bacteria* 39 (2015).

121. Swings, J. *et al.* Reclassification of the Causal Agents of Bacterial Blight (Xanthomonas campestris pv. oryzae) and Bacterial Leaf Streak (Xanthomonas campestris pv. oryzicola) of Rice as Pathovars of Xanthomonas oryzae (ex Ishiyama 1922) sp. nov., nom. rev. *Int. J. Syst. Bacteriol.* **40**, 309–311 (1990).

122. Triplett, L. R. *et al.* Genomic Analysis of Xanthomonas oryzae Isolates from Rice Grown in the United States Reveals Substantial Divergence from Known X. oryzae Pathovars. *Appl. Environ. Microbiol.* **77**, 3930–3937 (2011).

123. Lang, J. M. *et al.* A Pathovar of Xanthomonas oryzae Infecting Wild Grasses Provides Insight Into the Evolution of Pathogenicity in Rice Agroecosystems. *Front. Plant Sci.* **10**, 507 (2019).

124. Diekmann, M. & Bogyo, T. P. Distribution of bacterial leaf blight of rice [Xanthomonas campestris pv. oryzae (Ishiyama) Dye] depending on climatological factors / Verbreitung der Weissblättrigkeit des Reises [Xanthomonas campestris pv. oryzae (Ishiyama) Dye] in Abhängigkeit von Klimafaktoren. *Z. Für Pflanzenkrankh. Pflanzenschutz J. Plant Dis. Prot.* **99**, 127–136 (1992).

125. Reddy, P. R. & Yin, S. Survival of Xanthomonas campestris pv. oryzae, the causal organism of bacterial blight of rice. in *Bacterial of Blight Rice* (1989).

126. Kauffman, H. E., Reddy, A. P. K., Hsieh, S. P. Y. & Merca, S. D. An improved technique for evaluating resistance of rice varieties to Xanthomonas oryzae. *Plant Dis. Report.* **57**, 537–541 (1973).

127. McCann, H. C. & Guttman, D. S. Evolution of the type III secretion system and its effectors in plant–microbe interactions. *New Phytol.* **177**, 33–47 (2008).

128. Alvarez-Martinez, C. E. *et al.* Secrete or perish: The role of secretion systems in Xanthomonas biology. *Comput. Struct. Biotechnol. J.* **19**, 279–302 (2021).

129. Hueck, C. J. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev. MMBR* **62**, 379–433 (1998).

130. Notti, R. Q. & Stebbins, C. E. The Structure and Function of Type III Secretion Systems. *Microbiol. Spectr.* **4**, 10.1128/microbiolspec.VMBF-0004–2015 (2016).

131. Marlovits, T. C. *et al.* Structural insights into the assembly of the type III secretion needle complex. *Science* **306**, 1040–1042 (2004).

132. Ji, H. & Dong, H. Key steps in type III secretion system (T3SS) towards translocon assembly with potential sensor at plant plasma membrane: Type III translocon and plasma membrane sensor. *Mol. Plant Pathol.* **16**, 762–773 (2015).

133. Wagner, S. *et al.* Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiol. Lett.* **365**, (2018).

134. Studholme, D. J. *et al.* Draft Genome Sequences of Xanthomonas sacchari and Two Banana-Associated Xanthomonads Reveal Insights into the Xanthomonas Group 1 Clade. *Genes* **2**, 1050–1065 (2011).

135. Teper, D., Pandey, S. S. & Wang, N. The HrpG/HrpX Regulon of Xanthomonads—An Insight to the Complexity of Regulation of Virulence Traits in Phytopathogenic Bacteria. *Microorganisms* **9**, 187 (2021).

136. Cesbron, S. *et al.* Comparative Genomics of Pathogenic and Nonpathogenic Strains of Xanthomonas arboricola Unveil Molecular and Evolutionary Events Linked to Pathoadaptation. *Front. Plant Sci.* **6**, (2015).

137. Ramnarine, S. D. B. J., Jayaraman, J. & Ramsubhag, A. Comparative genomics of the black rot pathogen Xanthomonas campestris pv. campestris and non-pathogenic co-inhabitant Xanthomonas melonis from Trinidad reveal unique pathogenicity determinants and secretion system profiles. *PeerJ* **9**, e12632 (2022).

138. Yan, Q. & Wang, N. High-throughput screening and analysis of genes of Xanthomonas citri subsp. citri involved in citrus canker symptom development. *Mol. Plant-Microbe Interact. MPMI* **25**, 69–84 (2012).

139. Darsonval, A. *et al.* The Type III Secretion System of Xanthomonas fuscans subsp. fuscans Is Involved in the Phyllosphere Colonization Process and in Transmission to Seeds of Susceptible Beans. *Appl. Environ. Microbiol.* **74**, 2669–2678 (2008).

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

141. Mondal, K. K. *et al.* Rice transcriptome upon infection with Xanthomonas oryzae pv. oryzae relative to its avirulent T3SS-defective strain exposed modulation of many stress responsive genes. *3 Biotech* **12**, 130 (2022).

142. WHITE, F. F., POTNIS, N., JONES, J. B. & KOEBNIK, R. The type III effectors of Xanthomonas. *Mol. Plant Pathol.* **10**, 749–766 (2009).

143. Dangl, J. L. The Enigmatic Avirulence Genes of Phytopathogenic Bacteria. in *Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms* (ed. Dangl, J. L.) 99–118 (Springer, 1994). doi:10.1007/978-3-642-78624-2_5.

144. Sinha, D., Gupta, M. K., Patel, H. K., Ranjan, A. & Sonti, R. V. Cell Wall Degrading Enzyme Induced Rice Innate Immune Responses Are Suppressed by the Type 3 Secretion System Effectors XopN, XopQ, XopX and XopZ of Xanthomonas oryzae pv. oryzae. *PLOS ONE* **8**, e75867 (2013).

145. Popov, G., Fraiture, M., Brunner, F. & Sessa, G. Multiple *Xanthomonas euvesicatoria* Type III Effectors Inhibit flg22-Triggered Immunity. *Mol. Plant-Microbe Interactions®* **29**, 651–660 (2016).

146. Üstün, S., Bartetzko, V. & Börnke, F. The Xanthomonas campestris Type III Effector XopJ Targets the Host Cell Proteasome to Suppress Salicylic-Acid Mediated Plant Defence. *PLoS Pathog.* **9**, e1003427 (2013).

147. Kim, J.-G., Stork, W. & Mudgett, M. B. Xanthomonas Type III Effector XopD Desumoylates Tomato Transcription Factor SlERF4 to Suppress Ethylene Responses and Promote Pathogen Growth. *Cell Host Microbe* **13**, 143–154 (2013).

148. Szczesny, R. *et al.* Suppression of the AvrBs1-specific hypersensitive response by the YopJ effector homolog AvrBsT from Xanthomonas depends on a SNF1-related kinase. *New Phytol.* **187**, 1058–1074 (2010).

149. Teper, D. *et al.* Xanthomonas euvesicatoria type III effector XopQ interacts with tomato and pepper 14–3–3 isoforms to suppress effector-triggered immunity. *Plant J.* **77**, 297–309 (2014).

150. Deb, S., Ghosh, P., Patel, H. K. & Sonti, R. V. Interaction of the Xanthomonas effectors XopQ and XopX results in induction of rice immune responses. *Plant J.* **104**, 332–350 (2020).

151. Carter, M. E. *et al. A TAL effector-like protein of symbiotic* Mycetohabitans *increases stress tolerance and alters the transcriptome of the fungal host* Rhizopus microsporus. http://biorxiv.org/lookup/doi/10.1101/2020.03.04.968529 (2020) doi:10.1101/2020.03.04.968529.

152. Nowack, M. K., Holmes, D. R. & Lahaye, T. TALE-induced cell death executors: an origin outside immunity? *Trends Plant Sci.* **27**, 536–548 (2022).

153. Boch, J. *et al.* Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**, 1509–1512 (2009).

154. 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* **335**, 716–719 (2012).

155. Boch, J. *et al.* Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**, 1509–1512 (2009).

156. Moscou, M. J. & Bogdanove, A. J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **326**, 1501–1501 (2009).

157. Rinaldi, F. C., Doyle, L. A., Stoddard, B. L. & Bogdanove, A. J. The effect of increasing numbers of repeats on TAL effector DNA binding specificity. *Nucleic Acids Res.* **45**, 6960–6970 (2017).

158. Gao, H., Wu, X., Chai, J. & Han, Z. Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Res.* **22**, 1716–1720 (2012).

159. Yu, Y. *et al.* Colonization of rice leaf blades by an African strain of Xanthomonas oryzae pv. oryzae depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. *Mol. Plant-Microbe Interact. MPMI* **24**, 1102–1113 (2011).

160. Scholze, H. & Boch, J. TAL effector-DNA specificity. *Virulence* **1**, 428–432 (2010).

161. Domingues, M. N., de Campos, B. M., de Oliveira, M. L. P., de Mello, U. Q. & Benedetti, C. E. TAL Effectors Target the C-Terminal Domain of RNA Polymerase II (CTD) by Inhibiting the Prolyl-Isomerase Activity of a CTD-Associated Cyclophilin. *PLoS ONE* **7**, e41553 (2012).

162. Doucouré, H. *et al.* The Rice ILI2 Locus Is a Bidirectional Target of the African Xanthomonas oryzae pv. oryzae Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility. *Int. J. Mol. Sci.* **23**, 5559 (2022).

163. Sgro, G. G. *et al.* Bacteria-Killing Type IV Secretion Systems. *Front. Microbiol.* **10**, 1078 (2019).

164. Smith, T. J., Sondermann, H. & O'Toole, G. A. Type 1 Does the Two-Step: Type 1 Secretion Substrates with a Functional Periplasmic Intermediate. *J. Bacteriol.* **200**, e00168-18 (2018).

165. Nivaskumar, M. & Francetic, O. Type II secretion system: A magic beanstalk or a protein escalator. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1843**, 1568–1577 (2014).

166. Leo, J. C., Grin, I. & Linke, D. Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1088–1101 (2012).

167. Choi, Y. *et al.* Characterization of Type VI Secretion System in Xanthomonas oryzae pv. oryzae and Its Role in Virulence to Rice. *Plant Pathol. J.* **36**, 289–296 (2020).

168. Kim, S. *et al.* Time-resolved pathogenic gene expression analysis of the plant pathogen Xanthomonas oryzae pv. oryzae. *BMC Genomics* **17**, 345 (2016).

169. Shen, Y. & Ronald, P. Molecular determinants of disease and resistance in interactions of Xanthomonas oryzae pv. oryzae and rice. *Microbes Infect.* **4**, 1361–1367 (2002).

170. Das, A., Rangaraj, N. & Sonti, R. V. Multiple Adhesin-Like Functions of *Xanthomonas oryzae* pv. *oryzae* Are Involved in Promoting Leaf Attachment, Entry, and Virulence on Rice. *Mol. Plant-Microbe Interactions®* **22**, 73–85 (2009).

171. Girija, A. M. *et al.* Rice Leaf Transcriptional Profiling Suggests a Functional Interplay Between Xanthomonas oryzae pv. oryzae Lipopolysaccharide and Extracellular Polysaccharide in Modulation of Defense Responses During Infection. *Mol. Plant-Microbe Interact. MPMI* **30**, 16–27 (2017).

172. Rajeshwari, R., Jha, G. & Sonti, R. V. Role of an in planta-expressed xylanase of Xanthomonas oryzae pv. oryzae in promoting virulence on rice. *Mol. Plant-Microbe Interact. MPMI* **18**, 830–837 (2005).

173. Tayi, L. *et al.* A mutation in an exoglucanase of Xanthomonas oryzae pv. oryzae, which confers an endo mode of activity, affects bacterial virulence, but not the induction of immune responses, in rice. *Mol. Plant Pathol.* **19**, 1364–1376 (2017).

174. Gluck-Thaler, E. *et al. Repeated gain and loss of a single gene modulates the evolution of vascular pathogen lifestyles*. 2020.04.24.058529 https://www.biorxiv.org/content/10.1101/2020.04.24.058529v1 (2020) doi:10.1101/2020.04.24.058529.

175. Nathawat, R., Maku, R. V., Patel, H. K., Sankaranarayanan, R. & Sonti, R. V. Role of the FnIII domain associated with a cell wall-degrading enzyme cellobiosidase of Xanthomonas oryzae pv. oryzae. *Mol. Plant Pathol.* **23**, 1011–1021 (2022).

176. He, Y.-W., Wu, J., Cha, J.-S. & Zhang, L.-H. Rice bacterial blight pathogen Xanthomonas oryzae pv. oryzae produces multiple DSF-family signals in regulation of virulence factor production. *BMC Microbiol.* **10**, 187 (2010).

177. Jeong, K.-S. *et al.* Virulence Reduction and Differing Regulation of Virulence Genes in rpf Mutants of Xanthomonas oryzae pv. oryzae. *Plant Pathol. J.* **24**, 143–151 (2008).

178. Salzberg, S. L. *et al.* Genome sequence and rapid evolution of the rice pathogen Xanthomonas oryzae pv. oryzae PXO99A. *BMC Genomics* **9**, 204 (2008).

179. Huguet-Tapia, J. C. *et al.* Complete Genome Sequence of the African Strain AXO1947 of Xanthomonas oryzae pv. oryzae. *Genome Announc.* **4**, e01730-15 (2016).

180. Chen, F., Yan, B., Gong, X., Li, H. & He, Z. Genome sequencing of the bacterial blight pathogen DY89031 reveals its diverse virulence and origins of Xanthomonas oryzae pv. oryzae strains. *Sci. China Life Sci.* **64**, 2175–2185 (2021).

181. Amuthan, G. & Mahadevan, A. Plasmid and pathogenicity in Xanthomonas oryzae pathovar oryzae, the bacterial blight pathogen of Oryza sativa. *J. Appl. Bacteriol.* **76**, 529–538 (1994).

182. Niu, X.-N. *et al.* Complete sequence and detailed analysis of the first indigenous plasmid from Xanthomonas oryzae pv. oryzicola. *BMC Microbiol.* **15**, 233 (2015).

183. Carpenter, S. C. D. *et al.* An xa5 Resistance Gene-Breaking Indian Strain of the Rice Bacterial Blight Pathogen Xanthomonas oryzae pv. oryzae Is Nearly Identical to a Thai Strain. *Front. Microbiol.* **11**, (2020).

184. Lee, B.-M. *et al.* The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* **33**, 577–586 (2005).

185. Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A. & Kaku, H. Genome sequence of Xanthomonas oryzae pv. oryzae suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *JARQ Jpn.* (2005).

186. Booher, N. J. *et al.* Single molecule real-time sequencing of Xanthomonas oryzae genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Microb. Genomics* **1**, (2015).

187. Quibod, I. L. *et al.* Effector Diversification Contributes to Xanthomonas oryzae pv. oryzae Phenotypic Adaptation in a Semi-Isolated Environment. *Sci. Rep.* **6**, 34137 (2016).

188. Doucouré, H. *et al.* Functional and Genome Sequence-Driven Characterization of tal Effector Gene Repertoires Reveals Novel Variants With Altered Specificities in Closely Related Malian Xanthomonas oryzae pv. oryzae Strains. *Front. Microbiol.* **9**, 1657 (2018).

189. Oliva, R. *et al.* Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **37**, 1344–1350 (2019).

190. Najafi, M. B. H. & Pezeshki, P. BACTERIAL MUTATION; TYPES, MECHANISMS AND MUTANT DETECTION METHODS: A REVIEW. 11 (2013).

191. Chevallereau, A., Meaden, S., van Houte, S., Westra, E. R. & Rollie, C. The effect of bacterial mutation rate on the evolution of CRISPR-Cas adaptive immunity. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180094 (2019).

192. Metzgar, D., Liu, L., Hansen, C., Dybvig, K. & Wills, C. Domain-Level Differences in Microsatellite Distribution and Content Result from Different Relative Rates of Insertion and Deletion Mutations. *Genome Res.* **12**, 408–413 (2002).

193. Patel, S. Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. *Infect. Genet. Evol.* **45**, 151–164 (2016).

194. McDonald, B. A. & Linde, C. PATHOGEN POPULATION GENETICS, EVOLUTIONARY POTENTIAL, AND DURABLE RESISTANCE. *Annu. Rev. Phytopathol.* **40**, 349–379 (2002).

195. Dorman, C. J. H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid* **75**, 1–11 (2014).

196. Zhao, R., Feng, J., Huang, J., Li, X. & Li, B. Reponses of microbial community and antibiotic resistance genes to the selection pressures of ampicillin, cephalexin and chloramphenicol in activated sludge reactors. *Sci. Total Environ.* **755**, 142632 (2021).

197. Ferreira, R. M. *et al.* A TALE of Transposition: Tn3-Like Transposons Play a Major Role in the Spread of Pathogenicity Determinants of Xanthomonas citri and Other Xanthomonads. *mBio* **6**, e02505-14 (2015).

198. Lima, W. C., Paquola, A. C. M., Varani, A. M., Van Sluys, M.-A. & Menck, C. F. M. Laterally transferred genomic islands in Xanthomonadales related to pathogenicity and primary metabolism. *FEMS Microbiol. Lett.* **281**, 87–97 (2008).

199. Timilsina, S. *et al.* Xanthomonas diversity, virulence and plant–pathogen interactions. *Nat. Rev. Microbiol.* **18**, 415–427 (2020).

200. Tran, T. T. *et al.* Functional analysis of African Xanthomonas oryzae pv. oryzae TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLOS Pathog.* **14**, e1007092 (2018).

201. Teper, D. & Wang, N. Consequences of adaptation of TAL effectors on host susceptibility to Xanthomonas. *PLOS Genet.* **17**, e1009310 (2021).

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

203. de Vries, S., Stukenbrock, E. H. & Rose, L. E. Rapid evolution in plant–microbe interactions – an evolutionary genomics perspective. *New Phytol.* **226**, 1256–1262 (2020).

204. Tynkkynen, S., Satokari, R., Saarela, M., Mattila-Sandholm, T. & Saxelin, M. Comparison of Ribotyping, Randomly Amplified Polymorphic DNA Analysis, and Pulsed-Field Gel Electrophoresis in Typing ofLactobacillus rhamnosus and L. caseiStrains. *Appl. Environ. Microbiol.* **65**, 3908–3914 (1999).

205. Sheeja, T. E. *et al.* Amplified Fragment Length Polymorphism: Applications and Recent Developments. *Methods Mol. Biol. Clifton NJ* **2222**, 187–218 (2021).

206. Vaneechoutte, M. & Heyndrickx, M. 9 - Application and Analysis of ARDRA Patterns in Bacterial Identification, Taxonomy and Phylogeny. in *New Approaches for the Generation and Analysis of Microbial Typing Data* (eds. Dijkshoorn, L., Towner, K. J. & Struelens, M.) 211–247 (Elsevier Science B.V., 2001). doi:10.1016/B978-044450740-2/50010-1.

207. Garaizar, J., Rementeria, A. & Porwollik, S. DNA microarray technology: a new tool for the epidemiological typing of bacterial pathogens? *FEMS Immunol. Med. Microbiol.* **47**, 178–189 (2006).

208. Sharma, A., Lee, S. & Park, Y.-S. Molecular typing tools for identifying and characterizing lactic acid bacteria: a review. *Food Sci. Biotechnol.* **29**, 1301–1318 (2020).

209. Poulin, L. *et al.* New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by Xanthomonas oryzae. *Appl. Environ. Microbiol.* **81**, 688–698 (2015).

210. Tehrani, N. A. *et al.* Overview of typing techniques as molecular epidemiology tools for bacterial characterization. 9 (2021).

211. Ogawa, T., Yamamoto, T., Khush, G. S. & Mew, T.-W. Breeding of Near-Isogenic Lines of Rice with Single Genes for Resrstance to Bacterial Blight Pathogen (Xanthomonas campestris pv. oryzae). *Jpn. J. Breed.* **41**, 523–529 (1991).

212. Li, G., Song, C.-F., Pang, X.-M., Yang, Y. & Wang, J.-S. Analysis of Pathotypic and Genotypic Diversity of Xanthomonas oryzae pv. oryzae in China. *J. Phytopathol.* **157**, 208–218 (2009).

213. Yang, S.-Q. *et al.* Molecular and pathogenic characterization of new Xanthomonas oryzae pv. oryzae strains from the coastline region of Fangchenggang city in China. *World J. Microbiol. Biotechnol.* **29**, 713–720 (2013).

214. Mishra, D. *et al.* Pathotype and Genetic Diversity amongst Indian Isolates of Xanthomonas oryzae pv. oryzae. *PLoS ONE* **8**, e81996 (2013).

215. Suryadi, Y. *et al.* Determination of Pathotypes from Indonesian Xanthomonas oryzae Pv. Oryzae Population causing Bacterial Leaf Blight and their Reactions on Differential Rice. *Makara J. Sci.* **20**, 109–118 (2016).

216. Rashid, Md. M. *et al.* Pathotype profiling, distribution and virulence analysis of Xanthomonas oryzae pv. oryzae causing bacterial blight disease of rice in Bangladesh. *J. Phytopathol.* **169**, 438–446 (2021).

217. Furuya, N. *et al.* Diversity in Virulence of Xanthomonas oryzae pv. oryzae from Northern Vietnam. *Jpn. Agric. Res. Q. JARQ* **46**, 329–338 (2012).

218. Gonzalez, C. *et al.* Molecular and Pathotypic Characterization of New *Xanthomonas oryzae* Strains from West Africa. *Mol. Plant-Microbe Interactions®* **20**, 534–546 (2007).

219. Tekete, C. *et al.* Characterization of New Races of *Xanthomonas oryzae* pv. *oryzae* in Mali Informs Resistance Gene Deployment. *Phytopathology™* **110**, 267–277 (2020).

220. 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.* **104**, 10720–10725 (2007).

221. Freeman. An Overview of Plant Defenses against Pathogens and Herbivores. *Plant Health Instr.* (2008) doi:10.1094/PHI-I-2008-0226-01.

222. Wang, W. *et al.* Rice Secondary Metabolites: Structures, Roles, Biosynthesis, and Metabolic Regulation. *Molecules* **23**, 3098 (2018).

223. Chen, X. & Ronald, P. C. Innate immunity in rice. *Trends Plant Sci.* **16**, 451–459 (2011).

224. Li, P., Lu, Y.-J., Chen, H. & Day, B. The Lifecycle of the Plant Immune System. *Crit. Rev. Plant Sci.* **39**, 72–100 (2020).

225. Ji, Z., Wang, C. & Zhao, K. Rice Routes of Countering Xanthomonas oryzae. *Int. J. Mol. Sci.* **19**, (2018).

226. Song, W. Y. *et al.* A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* **270**, 1804–1806 (1995).

227. Park, C.-J., Han, S.-W., Chen, X. & Ronald, P. C. Elucidation of XA21-mediated Innate Immunity. *Cell. Microbiol.* **12**, 1017–1025 (2010).

228. Pruitt, R. N. *et al.* The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. *Sci. Adv.* **1**, e1500245 (2015).

229. Zhao, J., Fu, J., Li, X., Xu, C. & Wang, S. Dissection of the factors affecting development-controlled and race-specific disease resistance conferred by leucine-rich repeat receptor kinase-type R genes in rice. *Theor. Appl. Genet.* **119**, 231–239 (2009).

230. Xiang, Y., Cao, Y., Xu, C., Li, X. & Wang, S. 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 (2006).

231. Zhao, J., Fu, J., Li, X., Xu, C. & Wang, S. Dissection of the factors affecting development-controlled and race-specific disease resistance conferred by leucine-rich repeat receptor kinase-type R genes in rice. *Theor. Appl. Genet.* **119**, 231–239 (2009).

232. Yang, Y. *et al.* Research Progress on Cloning and Function of Xa Genes Against Rice Bacterial Blight. *Front. Plant Sci.* **13**, 847199 (2022).

233. Li, P., Long, J., Huang, Y., Zhang, Y. & Wang, J. AvrXa3: A novel member of avrBs3 gene family from Xanthomonas oryzae pv. oryzae has a dual function. *Prog. Nat. Sci.* **14**, 774–780 (2004).

234. Hu, K. *et al.* Improvement of multiple agronomic traits by a disease resistance gene via cell wall reinforcement. *Nat. Plants* **3**, 1–9 (2017).

235. Furukawa, T., Inagaki, H., Takai, R., Hirai, H. & Che, F.-S. Two distinct EF-Tu epitopes induce immune responses in rice and Arabidopsis. *Mol. Plant-Microbe Interact. MPMI* **27**, 113–124 (2014).

236. Desaki, Y. *et al.* Bacterial lipopolysaccharides induce defense responses associated with programmed cell death in rice cells. *Plant Cell Physiol.* **47**, 1530–1540 (2006).

237. Yuan, M. *et al.* Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* **592**, 105–109 (2021).

238. Gómez-Gómez, L. & Boller, T. FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. *Mol. Cell* **5**, 1003–1011 (2000).

239. Wang, S. *et al.* Rice OsFLS2-Mediated Perception of Bacterial Flagellins Is Evaded by Xanthomonas oryzae pvs. oryzae and oryzicola. *Mol. Plant* **8**, 1024–1037 (2015).

240. Ramirez-Prado, J. S., Abulfaraj, A. A., Rayapuram, N., Benhamed, M. & Hirt, H. Plant Immunity: From Signaling to Epigenetic Control of Defense. *Trends Plant Sci.* **23**, 833–844 (2018).

241. Lee, D. *et al.* Regulation of reactive oxygen species during plant immunity through phosphorylation and ubiquitination of RBOHD. *Nat. Commun.* **11**, 1838 (2020).

242. Nguyen, Q.-M., Iswanto, A. B. B., Son, G. H. & Kim, S. H. Recent Advances in Effector-Triggered Immunity in Plants: New Pieces in the Puzzle Create a Different Paradigm. *Int. J. Mol. Sci.* **22**, 4709 (2021).

243. Peng, Y., Bartley, L. E., Canlas, P. & Ronald, P. C. OsWRKY IIa Transcription Factors Modulate Rice Innate Immunity. *Rice N. Y. N* **3**, 36–42 (2010).

244. Moon, S.-J. *et al.* OsTGA2 confers disease resistance to rice against leaf blight by regulating expression levels of disease related genes via interaction with NH1. *PLOS ONE* **13**, e0206910 (2018).

245. Li, J., Henty-Ridilla, J. L., Staiger, B. H., Day, B. & Staiger, C. J. Capping protein integrates multiple MAMP signalling pathways to modulate actin dynamics during plant innate immunity. *Nat. Commun.* **6**, 7206 (2015).

246. 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.* **6**, (2015).

247. 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* **318**, 648–651 (2007).

248. 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.* **114**, E897–E903 (2017).

249. Hu, Y. *et al.* Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc. Natl. Acad. Sci.* **111**, E521–E529 (2014).

250. Zhang, M. *et al.* Resistance of Xanthomonas oryzae pv. oryzae to Lytic Phage X2 by Spontaneous Mutation of Lipopolysaccharide Synthesis-Related Glycosyltransferase. *Viruses* **14**, 1088 (2022).

251. Cernadas, R. A. *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.* **10**, e1003972 (2014).

252. Chen, L.-Q. *et al.* Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527–532 (2010).

253. Ji, J. *et al.* Plant SWEET Family of Sugar Transporters: Structure, Evolution and Biological Functions. *Biomolecules* **12**, 205 (2022).

254. Wu, L.-B. *et al.* OsSWEET11b, a sixth leaf blight susceptibility gene involved in sugar transport-dependent male fertility. 2021.08.21.457078 Preprint at https://doi.org/10.1101/2021.08.21.457078 (2021).

255. Chen, L.-Q. *et al.* Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport. *Science* **335**, 207–211 (2012).

256. Tao, Y. *et al.* Structure of a eukaryotic SWEET transporter in a homotrimeric complex. *Nature* **527**, 259–263 (2015).

257. Xuan, Y. H. *et al.* Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. *Proc. Natl. Acad. Sci.* **110**, E3685–E3694 (2013).

258. Riesmeier, J. W., Willmitzer, L. & Frommer, W. B. Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. *EMBO J.* **13**, 1–7 (1994).

259. Bezrutczyk, M. *et al.* Sugar flux and signaling in plant–microbe interactions. *Plant J.* **93**, 675–685 (2018).

260. Fei, H. *et al.* OsSWEET14 cooperates with OsSWEET11 to contribute to grain filling in rice. *Plant Sci.* **306**, 110851 (2021).

261. Zhou, Y. *et al.* Overexpression of OsSWEET5 in Rice Causes Growth Retardation and Precocious Senescence. *PLOS ONE* **9**, e94210 (2014).

262. Aked, J. & Hall, J. L. The uptake of glucose, fructose and sucrose into pea powdery mildew (Erysiphe pisi DC) from the apoplast of pea leaves. *New Phytol.* **123**, 277–282 (1993).

263. Sadoine, M. *et al.* Sucrose-dependence of sugar uptake, quorum sensing and virulence of the rice blight pathogen Xanthomonas oryzae pv. oryzae. 2021.08.22.457195 Preprint at https://doi.org/10.1101/2021.08.22.457195 (2022).

264. Gebauer, P. *et al.* Sugar Accumulation in Leaves of Arabidopsis sweet11/sweet12 Double Mutants Enhances Priming of the Salicylic Acid-Mediated Defense Response. *Front. Plant Sci.* **8**, (2017).

265. Streubel, J. *et al.* Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to Xanthomonas oryzae pv. oryzae. *New Phytol.* **200**, 808–819 (2013).

266. Yang, B., Sugio, A. & White, F. F. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl. Acad. Sci.* **103**, 10503–10508 (2006).

267. Antony, G. *et al.* Rice xa13 Recessive Resistance to Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene Os-11N3. *Plant Cell* **22**, 3864–3876 (2010).

268. Zhou, J. *et al.* Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* **82**, 632–643 (2015).

269. Schornack, S. *et al.* 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. Cell Mol. Biol.* **37**, 46–60 (2004).

270. Jiang, N. *et al.* Resistance Genes and their Interactions with Bacterial Blight/Leaf Streak Pathogens (Xanthomonas oryzae) in Rice (Oryza sativa L.)-an Updated Review. *Rice N. Y. N* **13**, 3 (2020).

271. Ji, C. *et al.* Xa1 Allelic R Genes Activate Rice Blight Resistance Suppressed by Interfering TAL Effectors. *Plant Commun.* **1**, 100087 (2020).

272. Read, A. C. *et al.* Suppression of Xo1-Mediated Disease Resistance in Rice by a Truncated, Non-DNA-Binding TAL Effector of Xanthomonas oryzae. *Front. Plant Sci.* **7**, 1516 (2016).

273. Chen, X. *et al.* Xa7, a new executor R gene that confers durable and broad-spectrum resistance to bacterial blight disease in rice. *Plant Commun.* **2**, 100143 (2021).

274. Wang, C. *et al.* Xa7, a Small Orphan Gene Harboring Promoter Trap for AvrXa7, Leads to the Durable Resistance to Xanthomonas oryzae Pv. oryzae. *Rice* **14**, 48 (2021).

275. Zhang, J., Yin, Z. & White, F. TAL effectors and the executor R genes. *Front. Plant Sci.* **6**, (2015).

276. Gu, K. *et al.* R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122–1125 (2005).

277. Tian, D. *et al.* The rice TAL effector-dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. *Plant Cell* **26**, 497–515 (2014).

278. Wang, C. *et al.* XA23 is an executor R protein and confers broad-spectrum disease resistance in rice. *Mol. Plant* **8**, 290–302 (2015).

279. Chu, Z. *et al.* Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes Dev.* **20**, 1250–1255 (2006).

280. Chen, H., Wang, S. & Zhang, Q. New gene for bacterial blight resistance in rice located on chromosome 12 identified from minghui 63, an elite restorer line. *Phytopathology* **92**, 750–754 (2002).

281. Liu, Q. *et al.* A paralog of the MtN3/saliva family recessively confers race-specific resistance to Xanthomonas oryzae in rice. *Plant Cell Environ.* **34**, 1958–1969 (2011).

282. Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R. & Szurek, B. A knowledge-based molecular screen uncovers a broad-spectrum OsSWEET14 resistance allele to bacterial blight from wild rice. *Plant J. Cell Mol. Biol.* **84**, 694–703 (2015).

283. Jiang, G.-H. *et al.* 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 (2006).

284. Han, J. *et al.* TALEN-based editing of TFIIAy5 changes rice response to Xanthomonas oryzae pv. Oryzae. *Sci. Rep.* **10**, 2036 (2020).

285. Poulin, L. *et al.* New Multilocus Variable-Number Tandem-Repeat Analysis Tool for Surveillance and Local Epidemiology of Bacterial Leaf Blight and Bacterial Leaf Streak of Rice Caused by Xanthomonas oryzae. *Appl. Environ. Microbiol.* **81**, 688–698 (2015).

286. Lang, J. M. *et al.* A Pathovar of Xanthomonas oryzae Infecting Wild Grasses Provides Insight Into the Evolution of Pathogenicity in Rice Agroecosystems. *Front. Plant Sci.* **10**, 507 (2019).

287. Oliva, R. *et al.* Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **37**, 1344–1350 (2019).

288. Eom, J.-S. *et al.* Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* **37**, 1372–1379 (2019).

289. Streubel, J. *et al.* Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to Xanthomonas oryzae pv. oryzae. *New Phytol.* **200**, 808–819 (2013).

290. Chu, Z. *et al.* Targeting xa13, a recessive gene for bacterial blight resistance in rice. *Theor. Appl. Genet.* **112**, 455–461 (2006).

291. Tran, T. T. *et al.* Functional analysis of African Xanthomonas oryzae pv. oryzae TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **14**, e1007092 (2018).

292. Yu, Y. *et al.* Colonization of rice leaf blades by an African strain of Xanthomonas oryzae pv. oryzae depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. *Mol. Plant-Microbe Interact. MPMI* **24**, 1102–1113 (2011).

293. Doucouré, H. *et al.* Functional and Genome Sequence-Driven Characterization of tal Effector Gene Repertoires Reveals Novel Variants With Altered Specificities in Closely Related Malian Xanthomonas oryzae pv. oryzae Strains. *Front. Microbiol.* **9**, 1657 (2018).

294. Blanvillain-Baufumé, S. *et al.* 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 (2017).

295. Schmidtke, C. *et al.* Genome-wide transcriptome analysis of the plant pathogen Xanthomonas identifies sRNAs with putative virulence functions. *Nucleic Acids Res.* **40**, 2020–2031 (2012).

296. Gluck-Thaler, E. *et al.* Repeated gain and loss of a single gene modulates the evolution of vascular plant pathogen lifestyles. *Sci. Adv.* **6**, eabc4516 (2020).

297. Orjuela, J. *et al.* CulebrONT: a streamlined long reads multi-assembler pipeline for prokaryotic and eukaryotic genomes. *Peer Community J.* **2**, (2022).

298. Wick, R. R. & Holt, K. E. Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLOS Comput. Biol.* **18**, e1009802 (2022).

299. Braissant, O., Astasov-Frauenhoffer, M., Waltimo, T. & Bonkat, G. A Review of Methods to Determine Viability, Vitality, and Metabolic Rates in Microbiology. *Front. Microbiol.* **11**, (2020).

300. Doucouré, H. *et al.* The Rice ILI2 Locus Is a Bidirectional Target of the African Xanthomonas oryzae pv. oryzae Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility. *Int. J. Mol. Sci.* **23**, 5559 (2022).

301. Jensen, M. S. *et al.* Discovery and characterization of a thermostable two-domain GH6 endoglucanase from a compost metagenome. *PLOS ONE* **13**, e0197862 (2018).

302. Ferluga, S., Bigirimana, J., Höfte, M. & Venturi, V. A LuxR homologue of Xanthomonas oryzae pv. oryzae is required for optimal rice virulence. *Mol. Plant Pathol.* **8**, 529–538 (2007).

303. Sadoine, M. *et al.* Sucrose-dependence of sugar uptake, quorum sensing and virulence of the rice blight pathogen Xanthomonas oryzae pv. oryzae. 2021.08.22.457195 Preprint at https://doi.org/10.1101/2021.08.22.457195 (2022).

304. Ji, J. *et al.* Plant SWEET Family of Sugar Transporters: Structure, Evolution and Biological Functions. *Biomolecules* **12**, 205 (2022).

305. Zeng, X. *et al.* CRISPR/Cas9-mediated mutation of OsSWEET14 in rice cv. Zhonghua11 confers resistance to Xanthomonas oryzae pv. oryzae without yield penalty. *BMC Plant Biol.* **20**, 313 (2020).

306. Gallois, J.-L., Moury, B. & German-Retana, S. Role of the Genetic Background in Resistance to Plant Viruses. *Int. J. Mol. Sci.* **19**, 2856 (2018).

307. Deng, Y. *et al.* Exploring the mechanism and efficient use of a durable gene-mediated resistance to bacterial blight disease in rice. *Mol. Breed.* **38**, 18 (2018).

308. Dossa, G. S., Sparks, A., Cruz, C. V. & Oliva, R. Decision tools for bacterial blight resistance gene deployment in rice-based agricultural ecosystems. *Front. Plant Sci.* **6**, (2015).

309. Collard, B. C. Y., Vera Cruz, C. M., McNally, K. L., Virk, P. S. & Mackill, D. J. Rice Molecular Breeding Laboratories in the Genomics Era: Current Status and Future Considerations. *Int. J. Plant Genomics* **2008**, e524847 (2008).

310. Kumar, A. *et al.* Deployment of Genetic and Genomic Tools Toward Gaining a Better Understanding of Rice-Xanthomonas oryzae pv. oryzae Interactions for Development of Durable Bacterial Blight Resistant Rice. *Front. Plant Sci.* **11**, 1152 (2020).

311. *Rice Improvement: Physiological, Molecular Breeding and Genetic Perspectives*. (Springer International Publishing, 2021). doi:10.1007/978-3-030-66530-2.

312. Impens, L., Jacobs, T. B., Nelissen, H., Inzé, D. & Pauwels, L. Mini-Review: Transgenerational CRISPR/Cas9 Gene Editing in Plants. *Front. Genome Ed.* **4**, (2022).

313. Xu, R. *et al.* Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice. *Rice* **7**, 5 (2014).

314. Wang, F. *et al.* Enhanced Rice Blast Resistance by CRISPR/Cas9-Targeted Mutagenesis of the ERF Transcription Factor Gene OsERF922. *PLOS ONE* **11**, e0154027 (2016).

315. Sun, Y. *et al.* Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. *Front. Plant Sci.* **8**, (2017).

316. Tremblay, J. P. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *médecine/sciences* **31**, 1014–1022 (2015).

317. Schmidt, S. M., Belisle, M. & Frommer, W. B. The evolving landscape around genome editing in agriculture: Many countries have exempted or move to exempt forms of genome editing from GMO regulation of crop plants. *EMBO Rep.* **21**, (2020).

318. Buchholzer, M. & Frommer, W. B. An increasing number of countries regulate genome editing in crops. *New Phytol.* **n/a**,.

319. Pilet-Nayel, M.-L. *et al.* Quantitative Resistance to Plant Pathogens in Pyramiding Strategies for Durable Crop Protection. *Front. Plant Sci.* **8**, (2017).

320. Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C. & Nelson, R. J. Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* **14**, 21–29 (2009).

321. Quenouille, J., Paulhiac, E., Moury, B. & Palloix, A. Quantitative trait loci from the host genetic background modulate the durability of a resistance gene: a rational basis for sustainable resistance breeding in plants. *Heredity* **112**, 579–587 (2014).

322. Siddiq, E. A. & Lakshminarayana, R. V. Advances in Genetics and Breeding of Rice: An Overview. in *Rice Improvement Physiological, Molecular Breeding and Genetic Perspectives* 1–29 (2021).

323. Frost, L. S. Conjugation, Bacterial. in *Encyclopedia of Microbiology (Third Edition)* (ed. Schaechter, M.) 517–531 (Academic Press, 2009). doi:10.1016/B978-012373944-5.00007-9.

324. Vera Cruz, C. M. *et al.* Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13500–13505 (2000).

325. Chaloner, T. M., Gurr, S. J. & Bebber, D. P. Plant pathogen infection risk tracks global crop yields under climate change. *Nat. Clim. Change* **11**, 710–715 (2021).

326. Dossa, G. S. *et al.* Rice response to simultaneous bacterial blight and drought stress during compatible and incompatible interactions. *Eur. J. Plant Pathol.* **147**, 115–127 (2017).

327. Morris, C. E., Barny, M., Berge, O., Kinkel, L. L. & Lacroix, C. Frontiers for research on the ecology of plant-pathogenic bacteria: fundamentals for sustainability. *Mol. Plant Pathol.* **18**, 308–319 (2017).

328. Straub, C., Colombi, E. & McCann, H. C. Population Genomics of Bacterial Plant Pathogens. *Phytopathology®* **111**, 23–31 (2021).

ANNEXE 1 :

Tableau récapitulant les différents types d'effecteurs de type 3 de Xanthomonas.

Class	Family	Synonyms	Related proteins	Function / Features	Xcv	Xcc	Xcr	Xoo	Xoc	Xvm	Xvv	Aac	Pa	Dd	Pph	Psy	Pta	Pto	Rs
AvrBs1	AvrBs1	.	AvrA	Unknown	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
AvrBs2	AvrBs2	.	.	Glycerophosphoryl diester phosphodiesterase	+	+	-	+	+	+	+	-	-	-	-	-	-	-	-
AvrBs3	AvrBs3	Pth, TAL	.	AvrBs3/PthA-type transcription activator; nuclear localization	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+
XopB	XopB	.	HopD1 (HopPtoD1, AvrPphD1)	Unknown	+	-	-	-	-	+	+	+	-	-	+	-	-	+	+
XopC	XopC1	.	.	Phosphoribosyl transferase domain and haloacid dehalogenase-like hydrolase	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+
XopC	XopC2	.	RSp1239	Haloacid dehalogenase-like hydrolase	+	-	-	+	+	+	+	+	-	-	-	-	-	-	+
XopD	XopD	.	.	C48-family SUMO cysteine protease (Ulp1 protease family) (Clan CE); EAR motif; DNA binding; nuclear localization	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
XopE	XopE1	AvrXacE1	HopX (AvrPphE)	Putative transglutaminase	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
XopE	XopE2	AvrXacE3, AvrXccE1	HopX (AvrPphE)	Putative transglutaminase	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
XopE	XopE3	AvrXacE2	HopX (AvrPphE)	Putative transglutaminase	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+
XopF	XopF1	Hpa4	.	Unknown	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
XopF	XopF2	.	.	Unknown	+	-	-	-	-	+	+	+	-	-	-	-	-	-	-
XopG	XopG1	.	HopH (HopPtoH), HopAP	M27-family peptidase (Clostridium toxin)	+	-	-	+	-	+	+	+	-	-	-	+	-	+	+
XopH	XopH1	AvrBs1.1	HopAO1 (HopPtoD2)	Putative tyrosine phosphatase	+	+	-	-	-	+	+	+	-	-	-	-	-	+	-
XopI	XopI1	.	.	F-box protein	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
XopJ	XopJ1	.	XopJ	C55-family cysteine protease or Ser/Thr acetyltransferase (Clan CE)	+	-	-	-	-	-	-	+	-	-	-	+	-	-	-
XopJ	XopJ2	AvrBsT	XopJ, HopZ2	C55-family cysteine protease or Ser/Thr acetyltransferase (Clan CE)	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
XopJ	XopJ3	AvrRxv	XopJ	C55-family cysteine protease or Ser/Thr acetyltransferase (Clan CE)	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-
XopJ	XopJ4	AvrXv4	XopJ, PopP1	C55-family SUMO cysteine protease or Ser/Thr acetyltransferase (Clan CE)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
XopJ	XopJ5	AvrXccB	XopJ	Putative C55-family cysteine protease or Ser/Thr acetyltransferase (Clan CE)	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-
XopK	XopK	.	.	Unknown	+	+	-	+	+	+	+	+	-	-	-	-	-	-	-
XopL	XopL	.	.	LRR protein	+	+	-	+	+	+	+	-	-	-	-	-	-	-	+
XopN	XopN	.	HopAU1	ARM/HEAT repeat	+	+	-	+	+	+	+	+	-	-	+	-	-	-	-

XopO	XopO	.	HopK1 (HopPtoK, HolPtoAB) N-terminal domain, AvrRps4	Unknown	+	-	-	-	+	-	-	+	-	-	+	-	-	+	-	-	+	-
XopP	XopP	.	Hlk	Unknown	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
XopQ	XopQ	.	HopQ1 (HolPtoQ), RipB	Putative i-sine-uridine nucleoside N-ribohydrolase	+	+	-	+	+	+	+	+	-	-	+	-	-	+	+
XopR	XopR	.	.	Unknown	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
XopT	XopT	.	.	Unknown	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
XopU	XopU	.	.	Unknown	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-
XopV	XopV	.	.	Unknown	+	-	-	+	+	+	+	+	-	-	-	-	-	-	-	+
XopW	XopW	.	.	Unknown	-	-	-	+	+	-	+	-	-	-	-	-	-	-	-	-
XopX	XopX	.	HopAE1 (HolPsyAE)	Unknown	+	+	-	+	+	+	+	+	-	-	-	-	-	-	-	-
XopY	XopY	.	.	Unknown	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-
XopZ	XopZ1	.	HopAS1	Unknown	+	+	-	+	+	+	+	+	-	-	+	-	-	+	+
XopAA	XopAA	Ecf	HopAE1 (HolPsyAE)	Early chlorosis factor; Proteasome/cyclosome repeat	+	-	-	+	+	+	+	-	-	-	+	+	+	-	-
XopAB	XopAB	.	weakly related to XopN	Unknown	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-
XopAC	XopAC	AvrAC	.	LRR-Fic/DOC protein; vascular hypersensitivity in Arabidopsis landrace Col-0	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
XopAD	XopAD	.	RSc3401	SKWP repeat protein	+	-	+	+	+	+	+	-	-	-	+	-	-	-	+
XopAE	XopAE	HpaF/HpaG	PopC	LRR protein	+	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-
XopAF	XopAF1	AvrXv3	HopAF1 (HopPtoJ)	Unknown	-	-	-	-	-	+	+	+	-	-	+	+	-	+	-
XopAF	XopAF2	AvrXv3	HopAF1 (HopPtoJ)	Unknown	-	-	-	-	+	+	+	+	-	-	+	+	-	+	-
XopAG	XopAG	AvrGf1	HopG1 (HopPtoG), HolPtoW	Unknown	-	+	-	-	-	+	+	+	-	-	+	-	-	+	+
XopAH	XopAH	AvrXccC	AvrB/AvrC (AvrPphC)	Unknown	-	+	-	+	+	-	-	-	-	-	+	+	-	-	-
XopAI	XopAI	.	HopO1 (HopPtoO, HopPtoS), HopAI1 (HolPtoAI)	Putative ADP-ribosyltransferase	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-
XopAJ	XopAJ	AvrRxo1	.	Unknown	+	-	-	-	+	-	-	+	-	-	-	-	-	-	-
XopAK	XopAK	.	HopK1 (HopPtoK, HolPtoAB) C-terminal domain	Deamidase	+	-	-	-	+	+	+	-	-	-	-	-	-	+	-
XopAL	XopAL1	.	Eop3	Unknown	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-
XopAL	XopAL2	.	Eop3	Unknown	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
XopAM	XopAM	.	HopR1	Unknown	-	+	-	-	-	+	+	-	-	-	+	-	+	+	+
XopA	XopA	Hpa1/HpaG	.	Harpin,	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
HpaA	HpaA	.	.	Type III secretion control protein,	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
HrpW	HrpW	.	PopW	Pectate lyase,	-	+	+	-	-	-	-	+	+	-	+	+	+	+	+
AvrXccA	AvrXccA1	AvrXca	.	Unknown	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-
AvrXccA	AvrXccA2	.	.	Unknown	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-

Xcv	Xanthomonas campestris pv. vesicatoria
Xcc	Xanthomonas campestris pv. campestris
Xcr	Xanthomonas campestris pv. raphani
Xoo	Xanthomonas oryzae pv.oryzae
Xoc	Xanthomonas oryzae pv.oryzicola
Xvm	Xanthomonas vasicola pv.musacearum
Xvv	Xanthomonas vasicola pv. vasculorum
Aac	Acidovorax citrulli
Pa	Pectobacterium atrosepticum
Dd	Dickeya dadantii
Pph	Pseudomonas syringae pv. phaseolicola
Psy	Pseudomonas syringae pv. syringae
Pta	Pseudomonas amygdali pv. tabaci
Pto	Pseudomonas syringae pv. tomato
Rs	Ralstonia solanacearum

ANNEXE 2 : Article en co-auteure “Broad-spectrum resistance to bacterial blight in rice using genome editing”

ARTICLES
https://doi.org/10.1038/s41587-019-0267-z

nature biotechnology

OPEN

Broad-spectrum resistance to bacterial blight in rice using genome editing

Ricardo Oliva[1,12*], Chonghui Ji[2,12], Genelou Atienza-Grande[1,10,12], José C. Huguet-Tapia[3,12], Alvaro Perez-Quintero[4,11,12], Ting Li[5], Joon-Seob Eom[6], Chenhao Li[2], Hanna Nguyen[1], Bo Liu[2], Florence Auguy[4], Coline Sciallano[4], Van T. Luu[6], Gerbert S. Dossa[7], Sébastien Cunnac[4], Sarah M. Schmidt[6], Inez H. Slamet-Loedin[1], Casiana Vera Cruz[1], Boris Szurek[4], Wolf B. Frommer[6,8*], Frank F. White[3] and Bing Yang[2,9*]

Bacterial blight of rice is an important disease in Asia and Africa. The pathogen, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), secretes one or more of six known transcription-activator-like effectors (TALes) that bind specific promoter sequences and induce, at minimum, one of the three host sucrose transporter genes *SWEET11*, *SWEET13* and *SWEET14*, the expression of which is required for disease susceptibility. We used CRISPR–Cas9-mediated genome editing to introduce mutations in all three *SWEET* gene promoters. Editing was further informed by sequence analyses of TALe genes in 63 *Xoo* strains, which revealed multiple TALe variants for *SWEET13* alleles. Mutations were also created in *SWEET14*, which is also targeted by two TALes from an African *Xoo* lineage. A total of five promoter mutations were simultaneously introduced into the rice line Kitaake and the elite mega varieties IR64 and Ciherang-Sub1. Paddy trials showed that genome-edited *SWEET* promoters endow rice lines with robust, broad-spectrum resistance.

Xanthomonas oryzae pv. *oryzae* (*Xoo*) is the etiological agent of bacterial blight disease in rice. The disease is most severe in southeast Asia but is increasingly damaging in west African countries, and results in substantial yield loss[1]. TALes from *Xoo* are injected by a type III secretion system into plant cells and recognize effector-binding elements (EBEs) in cognate *SWEET* host gene promoters, which results in induction of *SWEET* genes and production of sugars that enable disease susceptibility in rice[2,3]. An array of central repeats, which are 34–35 amino acids long, are present in each TALe and interact with EBEs via two repeat variable di-residues (RVDs) at the 12th and 13th position of each repeat[4,5]. Aberrant repeats, longer than 35 amino acids, are hypothesized to allow looping out of the repeat to accommodate alternate sequence binding for a particular TALe[6].

Bacterial blight depends on TALe-mediated induction of at least one member of a family of sugar-transporter genes. Although rice has more than 20 *SWEET* genes, only those of clade III are reported to be induced by *Xoo*[7–10]. Although all five of the known clade III *SWEET* genes in rice can function as susceptibility genes for bacterial blight, only three are known to be targeted in nature[10]. More specifically, *SWEET11* expression is induced by strains encoding the TALe PthXo1, *SWEET13* by PthXo2 and *SWEET14* by any one of several TALes, namely AvrXa7, PthXo3, TalC and TalF (originally Tal5) [7,9–15] (Table 1). Effectors of *Xoo* that target clade III *SWEET* genes are referred to as major TALes owing to their strong virulence effect.

Naturally occurring resistance has arisen as the result of nucleotide polymorphisms in EBEs of *SWEET* promoters. EBE alleles of *SWEET11* that are not recognized by PthXo1 are collectively referred to as the recessive resistance gene *xa13*. Rice varieties containing *xa13* are resistant to strains that solely depend on PthXo1 for *SWEET* induction. Most *indica* rice varieties carry a *SWEET13* allele that contains four adenines in the EBE for PthXo2, and rice lines carrying this allele are susceptible to PthXo2-dependent strains[12]. A rare exception is the recessive resistance allele *xa25*, which contains three adenines in the EBE for *SWEET13* in the *indica* cultivar Minghui 63, conferring resistance to strains that depend solely on PthXo2[16]. A similar recessive resistance allele in *japonica* rice varieties is equally effective against strains relying on PthXo2 (ref. [12]). Additional naturally occurring recessive EBE polymorphisms that confer resistance to strains carrying PthXo2, and the overlapping EBEs for PthXo3, TalF and AvrXa7 have subsequently been identified in the promoters of *SWEET13* and *SWEET14*, respectively, in germplasm collections[17,18].

Rice susceptibility genes are good targets for genome editing for disease resistance. TALe-mediated susceptibility is particularly modifiable. For instance, transcription-activator-like effector nuclease (TALEN)-directed mutations in *SWEET14* created lines resistant to strains carrying PthXo3/AvrXa7 (refs. [19,20]). Importantly, characterization of TALes and their cognate EBEs indicates that a limited set of EBEs are targeted by extant *Xoo* populations. At the

[1]International Rice Research Institute, Metro Manila, Philippines. [2]Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA. [3]Department of Plant Pathology, University of Florida, Gainesville, FL, USA. [4]IRD, CIRAD, Université Montpellier, IPME, Montpellier, France. [5]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, USA. [6]Institute for Molecular Physiology and Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine Universität Düsseldorf and Max Planck Institute for Plant Breeding Research, Köln, Germany. [7]Erfurt University of Applied Sciences, Erfurt, Germany. [8]Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Aichi, Japan. [9]Donald Danforth Plant Science Center, St. Louis, MO, USA. [10]Present address: College of Agriculture and Food Science, University of the Philippines Los Baños, Los Baños, Philippines. [11]Present address: Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO, USA. [12]These authors contributed equally: Ricardo Oliva, Chonghui Ji, Genelou Atienza-Grande, José C. Huguet-Tapia, Alvaro Perez-Quintero.
*e-mail: r.oliva@irri.org; frommew@hhu.de; yangbi@missouri.edu

Table 1 | Known EBEs in *SWEET* promoters

Host gene	TALe	EBE sequence
SWEET11	PthXo1	GCATCTCCCCCTACTGTACACCAC
SWEET13	PthXo2(A)	ATAAAAGCACCACAACTCCCTT
	PthXo2(A)	ATATAAGCACCACAACTCCCTT
	PthXo2B	ATAAAGCACCACAACTCCCTTC
	PthXo2C	ATAAAGCACCACAACTCCCTTC
SWEET14	TalC	CATGCATGTCAGCAGCTGGTCAT
	PthXo3	ATATAAACCCCCTCCAACCAGGTGCTAAG
	AvrXa7	ATAAACCCCCTCCAACCAGGTGCTAA
	TalF	AAGCTCATCAAGCCTTCA

same time, relatively few strains have been assessed for TALe gene content, and optimized genome editing requires knowledge of the TALe variants present in bacterial populations. Targeting both prevalent and rare TALes should increase the durability of resistance. In addition, the development of better tools for TALe characterization, which is a challenging and laborious procedure owing to their repetitive DNA sequences and presence of multiple copies per genome, would facilitate identification of novel TALes and determination of gene-editing strategies[21–23].

Here we report a detailed characterization of 63 *Xoo* strains from diverse geographic areas which includes genome sequence, the TALes present and virulence assays. Building on the understanding of TALes and their *SWEET* gene target EBEs, we developed a strategy to use CRISPR–Cas9 genome editing to engineer broad and durable resistance to bacterial blight in rice. An accompanying manuscript[24] reports the development of a disease management diagnostic tool kit for customized deployment of the resistant lines we report here.

Diversity of TALes in 63 *Xoo* strains

An assessment of the diversity of the major TALes in extant strains of *Xoo* was obtained by analyzing *Xoo* genome sequence data and virulence assays on rice lines with *SWEET* promoter polymorphisms. TALe gene content was derived from an analysis of complete genome sequences of 33 strains from Asia (Xoo^S) and 30 strains from Africa (Xoo^F), which were either in databases or newly sequenced (Supplementary Table 1). Whole-genome, single nucleotide polymorphism (SNP)-based parsimony trees clearly separated Xoo^S from Xoo^F genomes, revealing two distinct evolutionary lineages (Fig. 1). Trees based on alignments of repeat regions of 856 TALes revealed two major clusters (Supplementary Fig. 1). Expansion and diversification of TALe loci were more frequent in Xoo^S (18–21 TALes per genome) than in Xoo^F genomes (nine TALes per genome). Expansion and diversification of TALe loci is also reflected by the number of *SWEET*-gene-inducing TALes, of which there are 13 variants (unique RVD sequences) in Xoo^S strains and three variants in Xoo^F strains (Supplementary Fig. 2).

TALes predicted to induce *SWEET* genes (major TALes) were present in all *Xoo* strains, regardless of origin (Fig. 1). Indeed, all strains possessed one or more close homologs of the known major TALe groups (PthXo1, PthXo2 and PthXo3/AvrXa7). PthXo3/AvrXa7 is treated as a single class owing to the closely overlapping EBEs in the *SWEET14* promoter. Asian strains had approximately equal numbers of PthXo2 (targeting *SWEET13*) and PthXo3/AvrXa7 *(SWEET14)* (24 versus 23, respectively; Fig. 1). Most Asian strains had multiple major TALes with a combination of PthXo2 with either PthXo3 or AvrXa7 (Fig. 1). None of the strains contained both PthXo3 and AvrXa7 (Fig. 1). In African lineages, the two known major TALes of *Xoo* are TalC and TalF, which each induce *SWEET14* using non-overlapping EBEs[10,13]. Half of the analyzed African strains had both TalC and TalF, while none had TalF alone (Fig. 1). Overall, we found six non-overlapping EBEs in three *SWEET* promoters, if we treat the EBEs of PthXo2 variants that occur in some rice varieties as a single EBE (Table 1). These results support the hypothesis that activation of *SWEET* genes was a critical evolutionary step that occurred independently in different geographic regions during *Xoo* adaptation to rice.

Diversity of TALes from infection trials

To estimate the prevalence and variation of TALes in geographically diverse *Xoo* strains, a collection of 105 *Xoo* strains was screened against the *japonica* rice variety Kitaake and derivative Kitaake lines carrying mutations in the three *SWEET* genes normally targeted by TALes. Kitaake contains a natural variant allele of *SWEET13* that is not recognized by PthXo2[12]. Strains were also screened against lines with alternate alleles of *SWEET11* and *SWEET14* generated by TAL effector nucleases (TALEN) both singly and in combination (Fig. 2 and Supplementary Table 2).

Ten of 105 *Xoo* strains were not virulent on the *japonica* reference line Kitaake. These ten strains were virulent on an *indica* reference line IR24 (Supplementary Fig. 3). IR24 contains few known resistance genes and serves as a recurrent susceptible parent line for breeding diverse near-isogenic resistance gene lines against bacterial blight[25]. These strains may be solely dependent on PthXo2 for induction of the IR24 *indica* allele of *SWEET13*, as the EBEs in the other *SWEET* gene promoters are the same in Kitaake and IR24. Alternatively, Kitaake may have a resistance gene that is effective against the ten strains. These ten strains were excluded from further screening.

Of the remaining 95 Xoo strains, 18 were not virulent on edited rice line 1, which is defective in the EBE targeted by PthXo1 in *SWEET11*, and 69 strains were not virulent on line 5, which is defective in the EBEs of AvrXa7 and PthXo3 within *SWEET14* (Fig. 2 and Supplementary Table 2). Eighty-seven strains (18 + 69) were also not virulent on line 52-1, which harbors mutant alleles of both *SWEET11* and *SWEET14*, while the remaining eight strains (95 – 87) were virulent on line 52-1 (Fig. 2 and Supplementary Table 2). These findings indicate that 87 strains from the screen rely on major TALes that target known EBEs of *SWEET11* and *SWEET14*, whereas eight 'deviant' strains have different virulence abilities and, presumably, different TALes. These eight strains originated from the Philippines (PXO61, PXO364, PXO404, PXO421 and PXO513), the Korean peninsula (JW89011 and KXO85) and the African continent (AXO1947) (Supplementary Table 2). The virulence assay could not distinguish the deviant strains that, individually, carry both *avrXa7*/*pthXo3*-like and *pthXo2*-like or *pthXo1* and *pthXo2*-like TALe genes.

Characterization of deviant *Xoo* strains

Genome sequencing and TALe content analysis of the seven deviant Asian strains identified *pthXo3* in the five Philippine strains (PXO61, PXO364, PXO404, PXO421 and PXO513) and *avrXa7* in the two Korean strains (JW89011 and KXO85) as judged by the similarity of RVDs (Fig. 1, Supplementary Fig. 2 and Supplementary Table 3). *pthXo1* was not found in any of the deviant strains. However, on the basis of similarity of the RVDs, all seven genomes had one of two variant *pthXo2*-like genes (Fig. 1 and Supplementary Table 3). One variant, named PthXo2B, is present in all five strains from the Philippines, while PthXo2C is present in the two Korean strains. All three PthXo2-related effectors have repetitive domains of 22 RVDs. PthXo2B differs in eight RVDs, while PthXo2C contains seven variant RVDs as compared to PthXo2 (Supplementary Fig. 4a). Both PthXo2B and PthXo2C were predicted to bind to the same EBE (TATAAAGCACCACAACTCCCTTC) within the *SWEET13* promoter of Kitaake and other *japonica* varieties.

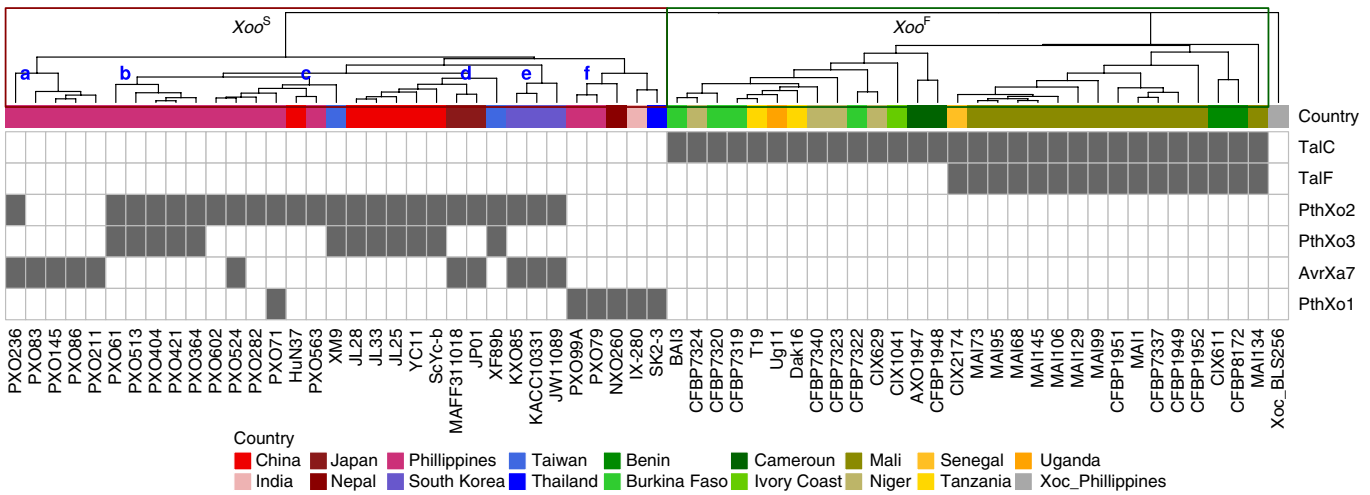


Fig. 1 | Distribution of *SWEET*-inducing TALes in *Xoo* strains. Heat map indicates presence (gray) or absence of TALes shown or predicted to bind to *SWEET* promoters in fully sequenced *Xoo* genomes. Origin is indicated. TALes were grouped by DisTAL (Supplementary Fig. 1). Six groups containing previously characterized *SWEET*-inducing TALes are shown. A parsimony tree (top) that is based on whole-genome SNPs obtained using kSNP is shown, highlighting the two main *Xoo* lineages: *Xoo*^{S} (Asia) and *Xoo*^{F} (Africa). Six sublineages were defined for *Xoo*^{S}, and are indicated in bold blue letters (a–f). Seventy-three percent of branches had support values over 0.9 as determined by FastreeMP implemented in kSNP3.

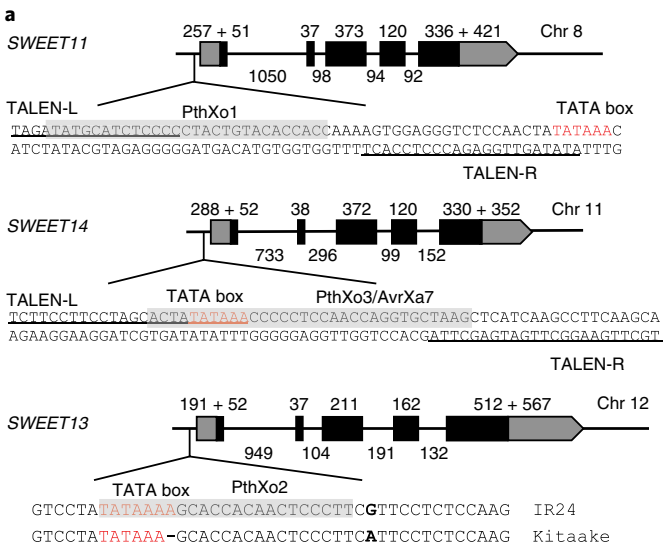


Fig. 2 | TALEN-induced variants and natural *SWEET* variants to assess prevalence of major TALes in 105 *Xoo* strains. a, Schematic gene structures of *SWEET11* and *SWEET14* with promoter sequences targeted by respective TALes shaded and with TALEN-bound sites underlined. Two natural variants of *SWEET13* in two rice varieties are shown. The EBE for the respective TALe is highlighted in gray. **b**, Genotypes of three Kitaake-derived lines for disease assays. All lines have the Kitaake allele for *SWEET13*.

The predicted EBEs of PthXo2B and PthXo2C differ from the EBE for PthXo2 of *SWEET13* in IR24, which is representative of most *indica* alleles (Supplementary Fig. 4b). Realigning RVDs with their respective EBEs and accounting for the shift in sequence owing to the single 'A' variation, the differences in RVD order between PthXo2 and PthXo2B or PthXo2C are reduced to four and five variant RVDs, respectively (Supplementary Fig. 4b). In addition to differences in the RVDs, PthXo2B and PthXo2C have 36-amino-acid repeats at RVD 9 and 12 relative to PthXo2. Atypical repeats have been found to accommodate alternative EBE binding in the major TALe AvrXa7 (ref. [6]). The differences between PthXo2B and PthXo2C and the presence of PthXo3 or AvrXa7 reflect the overall separation of the sublineages from the Philippines and Korea (Fig. 1; sublineages b and e, respectively).

Xoo strains with PthXo2 alone are incompatible on Kitaake-derived lines owing to the inability to induce the *japonica* allele of *SWEET13*[12]. The presence of *pthXo2B* or *pthXo2C* could explain the virulence phenotypes of the seven deviant strains on Kitaake and the double-edited Kitaake line 52-1. Both effectors are predicted to enable induction of the *japonica* allele of *SWEET13*. To test this possibility, the candidate gene for PthXo2B from PXO61 was cloned and introduced into strain PXO99^{A}ME2 (hereafter ME2), which is not pathogenic on any of the tested rice lines owing to a null mutation in the sole major TALe gene *pthXo1* of PXO99^A (ref. [9]). The *japonica* allele of *SWEET13* was induced in Kitaake and Nipponbare only upon inoculation with ME2 carrying *pthXo2B*, while relative expression in the *indica* allele in variety IR24 was low upon inoculation with ME2 or ME2 carrying PthXo2B (Fig. 3a). The *indica* allele of *SWEET13* in IR24 was induced upon inoculation with ME2 harboring PthXo2 (Fig. 3a). Lines in which *SWEET13* was induced scored as susceptible based on lesion length, and lines without *SWEET13* expression scored as resistant (Supplementary Fig. 5). An additional rice line (Zhengshan 97) carrying a different polymorphism in the EBE region for PthXo2-related effectors, namely a two A nucleotide deletion as compared to the *indica* allele, was also tested (Fig. 3b). The *SWEET13* allele in Zhengshan 97 behaved similarly to the IR24 allele and scored as susceptible to ME2 with PthXo2 and resistant to ME2 alone or with PthXo2B (Fig. 3 and Supplementary Fig. 5).

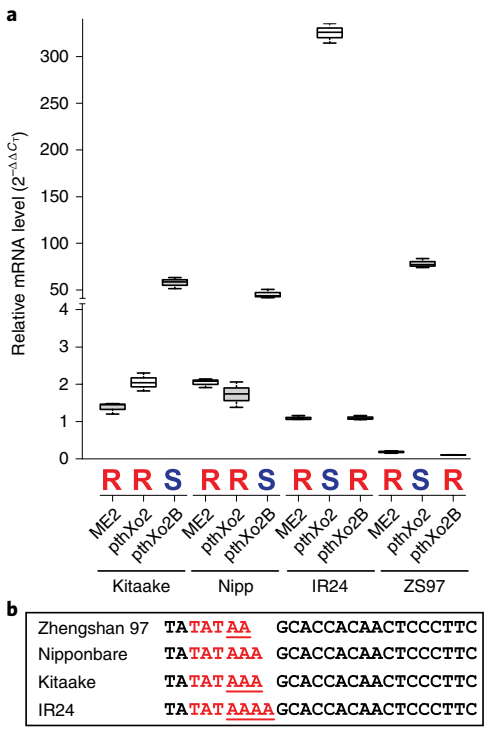


Fig. 3 | Functional analysis of PthXo2 homologs in four rice varieties. The genes *pthXo2* and *pthXo2B* were introduced into ME2 (a non-virulent derivative of PXO99^{A}, lacking *pthXo1*). **a**, *SWEET13* induction in four rice varieties by *Xoo* strains (rice actin 1 as control). Data are plotted using BoxPlotR (see Methods); center lines show medians, box limits indicate the 25th and 75th percentiles. The experiment was repeated three times independently with comparable results. Bold letters above strains indicate susceptibility score (red R, resistance; blue S, susceptibility). Actual data for resistance or susceptibility are shown in Supplementary Fig. 5. ZS, Zhengshan. **b**, Nucleotide variations (underlined) in EBE regions for PthXo2-related TALes in selected rice varieties. The predicted TATA box is highlighted in red.

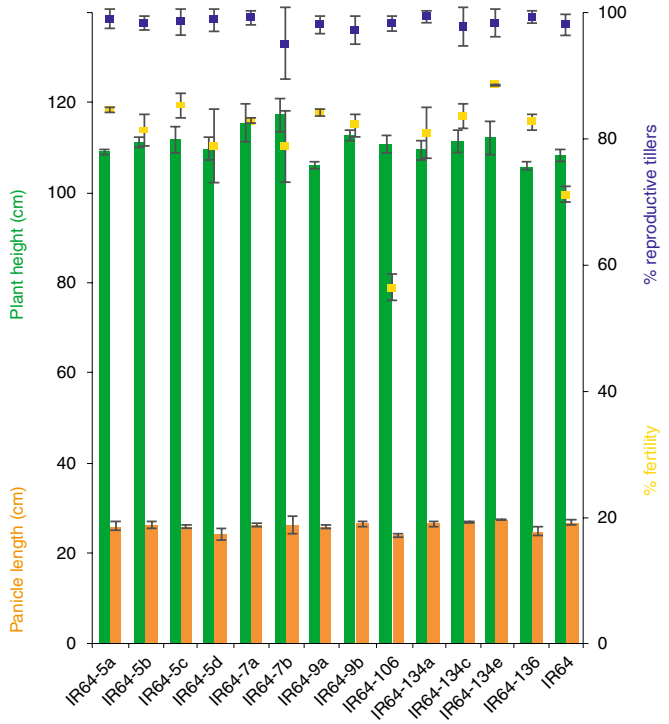


Fig. 4 | Agronomic traits in selected genome-edited mega variety lines as compared to the parental controls. Performance of edited IR64-IRS1132 lines in the T3 generation (left axis: plant height in green, panicle length in orange; right axis: percentage fertility in yellow and percentage of reproductive tillers in blue) relative to parental controls (IR64) ($n=15$, no significant difference between IR64 and edited lines except fertility for line IR64-106, $P<0.01$). Microfield experiments in screenhouses for agronomic trait assessments were conducted using randomized complete block design with three replicates. Trials were done in a single season. Data are presented in Supplementary Fig. 10.

Stacking edited EBEs for broad-spectrum resistance to *Xoo*

We tested whether stacking multiple mutations in the EBEs of three *SWEET* promoters known to be targeted by major TALes would provide resistance to most, if not all, *Xoo* strains in a single rice line. Using a multiplex CRISPR editing approach (Supplementary Figs. 6–8), a series of mutant lines were generated from Kitaake (Supplementary Table 4). The EBE mutant lines were first challenged with ME2 transformants, each carrying a gene for either PthXo1, PthXo2B, AvrXa7, TalC or TalF. The lines with the respective promoter mutations were resistant to ME2 derivatives carrying the corresponding TALes (Supplementary Table 5). When challenged with eleven field strains, including the seven RVD-variant *Xoo* strains that can induce *SWEET13*, one edited Kitaake line (11.1.3) was resistant to all strains (Supplementary Table 5). The 11.1.3 line conferred moderate resistance to the *Xoo*^{F} lineage strain AXO1947, which, of the known TALes, only contains TalC, suggesting the presence of another unknown effector. A similar observation was obtained upon inoculation of *Xoo*^{F} strain MAI1, which has both TalF and TalC (Fig. 1). Parallel tests using *SWEET* knockout mutants showed that AXO1947 likely depends on both *SWEET13* and *SWEET14* (ref. [24]).

Genome-edited *Xoo*-resistant mega varieties

Rice mega varieties, lines cultivated over one million hectares, are semidwarf and high-yielding lines that usually perform uniformly across multiple environments[26]. If the combined *SWEET* editing events tested here do not affect the normal physiological function of the genes, we predict that the same combinations would not impair agronomic traits in rice mega varieties, and would thus be useful for rice breeding programs. Genome editing by CRISPR–Cas9 was used to target *SWEET* promoters in two rice mega varieties, IR64 and Ciherang-Sub1 (Supplementary Fig. 9). Key agronomic traits were measured in confined paddy experiments. Before trials, the CRISPR-edited *SWEET* promoters of all mutant lines were sequenced, and lines were selected on the basis of the diversity of mutations in the respective EBEs (Supplementary Table 6). Overall, 13 IR64-IRS1132 and 18 Ciherang-Sub1-IRS1132 *Cas9*-free lines, spanning 30 combinations of EBE mutations in *SWEET11*, *SWEET13* and *SWEET14* were tested (Supplementary Table 7). Agronomic assessments and multivariant analysis of height, panicle length, number of reproductive tillers and fertility rate indicated that most lines performed similarly to wild-type parents (Fig. 4 and Supplementary Fig. 10). Only line IR64-106 showed phenotypic differences regarding yield, panicle length and fertility (reduced by 10% under paddy conditions). A range of lines with different insertions, deletions and substitutions in *SWEET11* ($n=6$), *SWEET13* ($n=8$) and *SWEET14* ($n=7$) were selected, assuming that larger EBE differences will be less likely to match TALe variation in pathogen populations. Distinct combinations of these 21 *SWEET* variants were distributed across nine rice lines. The nine lines were challenged with tester *Xoo* strains harboring single major TALes. All nine lines were resistant to all *Xoo* strains tested[24] (Figs. 5 and 6). Seven variants of *SWEET13*, ranging from

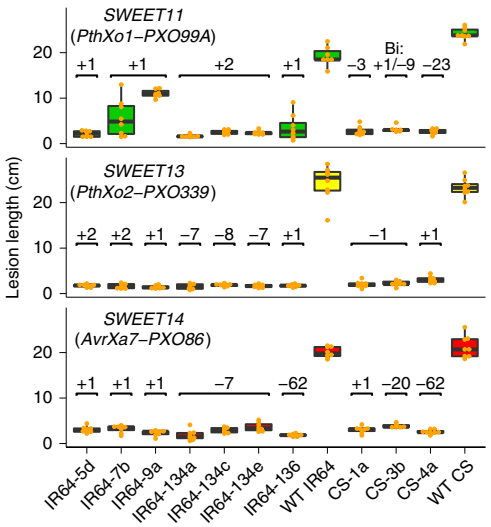


Fig. 5 | Resistance of *SWEET* promoter edited IR64 and Ciherang-Sub1 lines. Reactions of selected mutant lines of T3 IR64-IRS1132 and T2 Ciherang-Sub1-IRS1132 lines as compared to parental IR64 and Ciherang-Sub1 controls to infections with *pthXo1*-, *pthXo2*- and *avrXa7*-dependent *Xoo* strains (PXO99, PXO339 and PXO86, respectively). Different sequence alterations in the target *SWEET* promoters of edited lines resulted in varying levels of resistance to corresponding *Xoo* strains. Lesion lengths (cm) were measured 14 d after infection of specified plants. Types of mutations are indicated. Phenotyping experiments were conducted with four replicates per strain and two plants per replicate ($n=8$), and between three and six inoculated leaf samples were scored per plant. All tested lines carried the same mutations in both alleles, except those marked as 'Bi' (biallelic, that is, two different mutations in EBE). Center lines show medians; box limits indicate 25th and 75th percentiles determined in the R software package; whiskers extend 1.5 times the interquartile range from 25th and 75th percentiles. Means for all edited lines are significantly different from controls (Dunnett's test; $P<2\times10^{-16}$). A two-sided test was used.

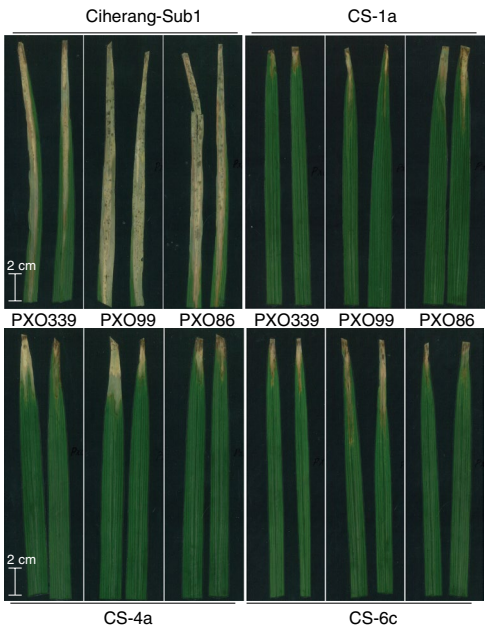


Fig. 6 | Resistance of three genome-edited Ciherang-Sub1 lines to three representative *Xoo* strains. Appearance of lesions resulting from clipping leaf tips using scissors that had been dipped into bacterial cultures of three different *Xoo* strains (PXO399, PXO99 and PXO86, as indicated), for three genome-edited Ciherang-Sub1 mutant lines CS-1a, CS-4a and CS-6c (T2 generation) as compared to the parental control (Ciherang-Sub1). The experiment was repeated twice independently.

two-nucleotide insertions (+2) to a seven-nucleotide deletion (−7), conferred resistance. Line IR64-9a carried only SNPs in *SWEET11*, which were insufficient to abrogate PthXo1 function (Fig. 5).

Discussion

We set out to evaluate our hypothesis that editing known *SWEET* EBEs will endow rice with broad resistance to bacterial blight. The success of our approach depended on extant *Xoo* strains fitting the profile of the relatively few *Xoo* strains that have been genetically characterized. Indeed, analysis of a broad collection of 63 *Xoo* strains found that all strains targeted, or were predicted to target, *SWEET11*, *SWEET13* and *SWEET14*, as previously reported[8,10,27]. Furthermore, most *Xoo* strains, if not all, target some combination of six short EBE sequences present in the promoters of the three *SWEET* genes. We applied multiplexed CRISPR–Cas9 genome editing to systematically interfere with *SWEET* gene induction at all known major TALe EBEs and engineer Kitaake rice resistant to all currently known strains of *Xoo*. Kitaake is an excellent cultivar for proof-of-concept owing to high regeneration and rapid flowering cycle. The lines generated here will serve as diagnostic tools for rapid evaluation of the virulence of novel *Xoo* isolates[24]. Although Kitaake can be used for breeding resistance genes into Japanese and Chinese varieties, the *japonica* cultivar is not optimal as breeding material in large parts of southeast Asia and Africa, which use mainly *indica* varieties. Therefore, we also applied the same editing approach in two mega varieties, IR64[26] and Ciherang-Sub1[28]. The edited mega variety lines grew normally, without yield suppression, and were resistant to three representative strains of bacterial blight carrying the known major TALes.

To devise an efficient genome editing strategy it was important to identify a panoply of TALes and cognate EBEs to guide engineering of broadly applicable bacterial blight resistance. Sequence analysis of the seven deviant *Xoo* strains from Asia and relatedness of TALes with close, but not identical, RVDs led us to propose that variants of PthXo2 target variant EBEs in *SWEET13*. These seven Asian strains retained virulence on the first-generation of genome-edited EBE lines in the Kitaake background, indicating that the variant strains either carry another major TALe targeting an undefined susceptibility gene or a variant major TALe targeting *SWEET11*, *SWEET13* or *SWEE14* at an alternate EBE. The two newly identified *pthXo2* members (*pthXo2B* and *pthXo2C*) were predicted to target the *SWEET13* promoters found in many *japonica* varieties, including Kitaake and Nipponbare. Indeed, PthXo2B conferred virulence to ME2, a strain lacking major TALe genes, on both Kitaake and Nipponbare, concomitantly with *SWEET13* induction. Thus, the presence of PthXo2B or PthXo2C in the deviant strains is sufficient to explain their virulence on the TALEN-edited lines. Ultimately, several lines with CRISPR-mediated edits at the PthXo2B-cognate EBE of Kitaake were resistant to *pthXo2B*-containing strains, indicating that the Kitaake allele for *SWEET13* was responsible for the susceptibility of line 52-1. Recent testing of all TALe genes of the strain PXO61 did not reveal any new major TALe other than PthXo3 and PthXo2B[29]. *Xoo*[8] strains carrying PthXo2B were obtained in the Philippines, where several bacterial blight outbreaks occurred. This sublineage is characterized as local races 1 and 9d, which have not been reported in other countries. At the same time, the Philippine archipelago may offer a safe geographic reservoir for less prevalent strains. Both strains from the Korean peninsula possess PthXo2C. The sampling of Korean strains was limited, and we are unable to state whether genes for PthXo2C are the rule or the exception. What is clear is that different variants of PthXo2 family are present in the

global *Xoo* population, suggesting an ongoing adaptation to nucleotide variations in the *SWEET13* promoter[12,16,17].

African strains were shown to possess less TALe diversity and fall into two sublineages. All strains harbor the major virulence TALe, TalC, which targets *SWEET14*. Members of the second sublineage have TalC and TalF, the latter also targeting *SWEET14*. Another feature of the African strains is that loss of TalC function results in significantly reduced lesion length[13]. *Xoo* strains relying solely on TalC retained full virulence on Kitaake lines mutated in the TalC EBE within *SWEET14* (ref. [19]).

We show here that mutation of the TalC EBE in the promoter of *SWEET14* in Kitaake led to only moderate resistance to the African strain AXO1947. By contrast, double knock out of *SWEET13* and *SWEET14* resulted in complete resistance[24]. Consistent with previous studies[19,30], the dependence of African *Xoo* strains on *SWEET* gene expression to cause symptoms appears more complicated than anticipated. Alternatively, lesion length may not be strictly correlated with ecological fitness or disease severity, and variation in lesion length, particularly at lower value ranges, may be due to variant host defense responses to diverse pathogen lineages.

Broad resistance to bacterial blight at the *SWEET* promoters will not prevent adaptation of the pathogen, and the durability of this approach will depend on the ability of *Xoo* populations to adapt to recessive resistance alleles. As noted, TALe genes have limited diversity, and, when diversity is present, it comprises minor variations. We propose that creating polymorphisms that are larger than a single nucleotide change (ideally modifying the whole EBE) in the EBEs of *SWEET* genes is advisable. These changes will still capture minor TALe variants of the major, for example, for PthXo2 and AvrXa7/PthXo3. The promoter editing will also likely thwart TALe adaptation, assuming that ease of adaptation to new binding sites is inversely related to the number of novel nucleotides in the target sequence. EBEs must meet certain structural requirements for TALe binding within the available sequences of the susceptibility gene promoters. Another consideration is that breeders will likely combine these recessive resistance alleles with other locally effective resistance genes and these combinations will likely severely reduce disease pressure, further delaying strain adaptation.

A diagnostic kit that can both identify the promoter variants best suited to defeat the local pathogen population and monitor variant strain emergence is presented in an accompanying publication[24]. Breeders and pathologists can work together to provide selected lines to farmers, with an overall strategy of deploying multiple lines in a region to disfavor the evolution of novel strains.

Multiplex targeting using CRISPR–Cas9 has created the ability to simultaneously edit all EBEs present in any single rice line. However, this method is limited owing to the constraint of possible edits and problems with the delivery of editing complexes. For example, the most efficient genome editing is via DNA-mediated delivery, which confers regulatory concerns on edited plants. Genome editors such as CRISPR–Cpf1 (also called Cas12a, which allows for larger mutations), potent ribonucleoprotein (RNP) delivery and transgene-free genome editing may mitigate these limitations and improve the durability of disease resistance.

In conclusion, by using a combination of systematic analyses of diverse *Xoo* strains, an understanding of *SWEET* genes and genome editing, we were able to engineer broad-spectrum resistance in Kitaake and two mega varieties IR64 and Ciherang-Sub1. Genome editing can result in off-target effects and the transformation process often leads to mutations and somaclonal variation[31,32], but on- and off-target analysis of rice genomes in NCBI with the Cas9 guide RNAs revealed that all the predicted off-target sites lack PAM sequences (an essential component for Cas9 recognition) and contain mismatches larger than 4 nucleotides to the guide sequences within the single guide RNA (sgRNA) genes (Supplementary Table 8). Furthermore, off-target mutations (if they are present), and other mutations derived from tissue culture and the regeneration process, will be eliminated in crosses during breeding. Full genome sequencing will be needed to identify possible off-target mutations and to ensure that the lines do not contain any T-DNA. Finally, our preliminary analysis of agronomic traits, while promising, is insufficient for direct deployment. Extensive field trials will be necessary to ensure performance of these edited resistant lines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41587-019-0267-z.

Received: 27 February 2019; Accepted: 28 August 2019;
Published online: 28 October 2019

References

1. Niño-Liu, D. O., Ronald, P. C. & Bogdanove, A. J. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Mol. Plant Pathol.* **7**, 303–324 (2006).
2. White, F. F., Potnis, N., Jones, J. B. & Koebnik, R. The type III effectors of *Xanthomonas*. *Mol. Plant Pathol.* **10**, 749–766 (2009).
3. Bezrutczyk, M. et al. Sugar flux and signaling in plant-microbe interactions. *Plant J.* **93**, 675–685 (2018).
4. Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
5. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
6. Richter, A. et al. A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* **5**, 3447 (2014).
7. Antony, G. et al. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* **22**, 3864–3876 (2010).
8. Chen, L. Q. et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527–532 (2010).
9. Yang, B., Sugio, A. & White, F. F. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl Acad. Sci. USA* **103**, 10503–10508 (2006).
10. Streubel, J. et al. Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **200**, 808–819 (2013).
11. 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.* **17**, 1192–1200 (2004).
12. Zhou, J. et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* **82**, 632–643 (2015).
13. Yu, Y. et al. Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 *Os11N3* gene. *Mol. Plant Microbe Interact.* **24**, 1102–1113 (2011).
14. Tran, T. T. et al. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **14**, e1007092 (2018).
15. Doucouré, H. et al. Functional and genome dequence-driven characterization of TAL effector gene repertoires reveals novel variants with altered specificities in closely related malian *Xanthomonas oryzae* pv. *oryzae* strains. *Front Microbiol* **9**, 1657 (2018).
16. Liu, Q. et al. A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant Cell Environ.* **34**, 1958–1969 (2011).
17. Zaka, A. et al. Natural variations in the promoter of *OsSWEET13* and *OsSWEET14* expand the range of resistance against *Xanthomonas oryzae pv. oryzae*. *PLoS ONE* **13**, e0203711 (2018).
18. Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R. & Szurek, B. A knowledge-based molecular screen uncovers a broad-spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. *Plant J.* **84**, 694–703 (2015).
19. Blanvillain-Baufumé, S. et al. 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 (2017).
20. Li, T., Liu, B., Spalding, M. H., Weeks, D. P. & Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **30**, 390–392 (2012).
21. Booher, N. J. et al. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Micro. Genom.* **1**, e000032 (2015).

22. Peng, Z. et al. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics* **17**, 21 (2016).
23. Gochez, A. M. et al. 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 (2018).
24. Eom, J.-S. et al. Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-019-0268-y (2019).
25. Ogawa, T., Yamamoto, T., Khush, G. S. & Mew, T.-W. Breeding of near-isogenic lines of rice with single genes for resistance to bacterial blight pathogen (*Xanthomonas campestris* pv. *oryzae*). *Jpn. J. Breed.* **41**, 523–529 (1991).
26. Mackill, D. J. & Khush, G. S. IR64: a high-quality and high-yielding mega variety. *Rice* **11**, 18 (2018).
27. Chen, L. Q. et al. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**, 207–211 (2012).
28. Toledo, A. M. U. et al. Development of improved Ciherang-Sub1 having tolerance to anaerobic germination conditions. *Plant Breed. Biotech.* **3**, 77–87 (2015).
29. Li, C. et al. An efficient method to clone TAL effector genes from *Xanthomonas oryzae* using Gibson assembly. *Mol. Plant Pathol.* https://doi.org/10.1111/mpp.12820 (2019).
30. Carpenter, S. C. D. et al. A strain of an emerging Indian *Xanthomonas oryzae* pv. *oryzae* pathotype defeats the rice bacterial blight resistance gene *xa13* without inducing a clade III SWEET gene and is nearly identical to a recent Thai isolate. *Front. Microbiol.* **9**, 2703 (2018).
31. Hahn, F. & Nekrasov, V. CRISPR/Cas precision: do we need to worry about off-targeting in plants? *Plant Cell Rep.* **38**, 437–441 (2019).
32. Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* **19**, 84 (2018).

Acknowledgements

We are grateful to A. Crepcia-Pevzner and J. Long for technical assistance; J. Jacobs for help in sequencing *Xoo* genomes; and G. Onaga, K. Wydra, H. Tall and A. von Tiedemann for supporting pathogen collection. We gratefully acknowledge funding from the Bill and Melinda Gates Foundation to Heinrich Heine University, Düsseldorf, with subawards to Iowa State University/University of Missouri, University of Florida, Institut de recherche pour le développement and the International Rice Research Institute. Characterization of *SWEET* promoters was supported by a joint grant from the National Science Foundation to B.Y., F.F.W. and W.B.F. (IOS-1258018 and IOS-1258028). W.B.F. was supported by an Alexander-von-Humboldt Professorship.

Author contributions

R.O., C.V.C., B.S., F.F.W., B.Y. and W.B.F. conceived and designed experiments. C.J., G.A.-G., J.C.H.-T., A.P.-Q., T.L., J.-S.E., C.L., V.T.L., C.S. and H.N. performed experiments. B.L. and I.H.S.-L. performed rice transformation. G.S.D. provided materials. J.S.E., C.J., J.C.H.-T., G.A.-G., A.P.-Q., F.A., S.C., B.Y. and W.B.F. analyzed the data. R.O., F.F.W., S.M.S., J.C.H.-T., B.S., B.Y. and W.B.F. wrote the manuscript, and S.M.S., J.S.E., G.A.-G. and A.P.Q. helped with revisions.

Competing interests

W.B.F., J.-S.E., F.W., B.Y. and R.O. are inventors on a provisional patent application 62832300 that covers Kitaake, IR64 and Ciherang-Sub1 EBE-edited lines.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41587-019-0267-z.

Correspondence and requests for materials should be addressed to R.O., W.B.F. or B.Y.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.
© The Author(s) 2019

Methods

Plant material, bacterial strains, medium and growth conditions. Rice varieties used here were *Oryza sativa* L. ssp. *japonica* Kitaake[12], and *Oryza sativa* L. ssp. *indica* IR24 (ref. [12]), IR64 (ref. [26]) and Ciherang-Sub1 (ref. [28]). Rice plants were grown in growth chambers at 30 °C for a 12-h light period and at 28 °C for a 12-h dark period, with 60–75% relative humidity in the Yang laboratory, or under small-scale field environment in a screenhouse (28 °C ± 7 °C day and 23 °C ± 4 °C night; 80–85% relative humidity) at IRRI. *Escherichia coli* strains were grown in LB medium supplemented with appropriate antibiotics at 37 °C. *Agrobacterium tumefaciens* strains were grown at 30 °C in the dark. All *Xoo* strains were grown at 28 °C in TSA (10 g l^{-1} tryptone, 10 g l^{-1} sucrose and 1 g l^{-1} glutamic acid). Antibiotics were used at the following concentrations if required: 100 µg ml^{-1} ampicillin; 10 µg ml^{-1} cephalexin; 25 µg ml^{-1} chloramphenicol; 25 µg ml^{-1} kanamycin; 100 µg ml^{-1} spectinomycin; and 10 µg ml^{-1} tetracycline.

***Xoo* genome sequencing and assembly.** Total genomic DNA of *Xoo* strains was isolated using MagAttract HMW DNA Mini Kit (QIAGEN). DNA samples were sequenced using a single molecule real-time (SMRT-Pacbio) platform, P4-C6 chemistry. For each strain, at least two SMRT cells were used, generating around 180× coverage per genome. De novo assembly was conducted using the hierarchical assembly pipeline (HGAP) implemented in Canu v.1.5 software[33]. Raw PacBio reads were mapped against the resulting contigs using blasR aligner (https://github.com/PacificBiosciences/blasr) and corrections were conducted with the variant-caller software utilizing the arrow algorithm (https://github.com/PacificBiosciences/GenomicConsensus).

Genome annotation and gene prediction were conducted using the Prokka annotation pipeline[34] and the NCBI Prokaryotic Genome Annotation (PGAP)[35]. Genome sequences were deposited in the GenBank under BioProjects PRJNA497307 and PRJNA497605. Additional reported and publicly available genomes in GenBank were retrieved and used for comparisons (Supplementary Table 1).

Phylogenetic and comparative genomic analyses of *Xoo* strains. Phylogenetic analysis was conducted by identifying pan-genome SNPs in oligonucleotides of length $k = 31$ and generating parsimony trees on the basis of these SNPs, as implemented in the program kSNP (https://doi.org/10.1093/bioinformatics/btv271).

TALe content analysis in *Xoo*. TALe annotation was conducted using the AnnoTALE software[36]. Additional annotation and comparisons were performed using Artemis genome browser and Artemis Comparative Tool[37,38]. A neighbor-joining tree based on the alignments of repeat regions of TALes was obtained using DisTAL[39]. Target prediction for TALes was conducted with Talvez[40]; the target used was sequences in the promoter regions (1 kb upstream from the translation start site) of the rice *japonica* Nipponbare genome version deposited in Phytozome database V11 and of the *indica* variety IR64 (v.CSHL1.0)[41].

CRISPR–Cas9 design and constructs. For gene editing in Kitaake, the polycistronic tRNA–guide RNA (gRNA) system was used to generate multiple sgRNAs with different target sequences by flanking the sgRNAs with a tRNA precursor sequence[42]. Six intermediate vectors were constructed, pTLN-tgRNA-1 to T6, for six individual tRNA–gRNA units. A double-stranded DNA oligonucleotide for each site was produced by annealing two complementary oligonucleotides (24–25 nt) (Supplementary Table 9). Six double-stranded oligonucleotides were individually inserted into the *BsmB*I-digested pTLN-tgRNA-1 to T6 (Supplementary Fig. 7). The DNA sequence of positive clones was confirmed by Sanger sequencing (Supplementary Table 10). All six tRNA–gRNA units were transferred into another intermediate vector named pENTR4-U6.1P-ccdB using the Golden Gate ligation method (Supplementary Fig. 8). The gRNA cassettes were finally mobilized to pBY02-ZmUbiP-OsCas9 (Supplementary Fig. 9) by using Gateway LR Clonase (Thermo Fisher Scientific) as described[43]. For gene editing in IR64 and Ciherang-Sub1, four gRNA genes (IRS1132; Supplementary Fig. 9), targeting PthXo1 EBE in *SWEET11*, PthXo2 EBE in *SWEET13*, or TalC EBE or AvrXa7 EBE in *SWEET14*, were constructed as described[14].

Generation of rice mutant lines. *Agrobacterium* strains containing the respective CRISPR constructs were used for genome editing in Kitaake, IR64 and Ciherang-Sub1. Individual transformants were selected, propagated and regenerated into whole plants (T0) as described[44,45]. Leaf tissues collected from individual samples were homogenized in liquid nitrogen, and genomic DNA was isolated following the CTAB method. Initial screening for transformants (T0) was done via PCR of the phosphotransferase transgene (*hpt*) with primers HptF and HptR. Nuclease surveyor assay using Cel1 enzyme was employed in T0 plants to preselect lines with possible sequence alterations in the target regions. We selected 20 mono- and biallelic events for IR64 and sequenced the EBE regions using site-specific PCR amplicons obtained with primer pair 8NKpI-F5 and 8N3-R for *SWEET11*, primer pair 12N3-F1 and 12N3-R1 for *SWEET13* and primer pair 11N3-F2 and 11N3-R for *SWEET14*. To detect the CRISPR transgenes, paired primers OsCas9-F and OsCas9-R were used for Cas9, primers g8N3-F and g12N3-R for gRNA genes targeting *SWEET11* and *SWEET13*, and primers g11N3-F and g11N3-R for gRNA genes targeting *SWEET14*. These primer pairs were also used for selection of candidate mutant lines in the advanced generations (T1 to T2). Primer information is provided in Supplementary Table 9. Nineteen gRNA-gene-free and Cas9-free IR64 T1 plants were selected and further analyzed. Thirty plants for each of the 19 IR64 T2 lines were phenotyped for resistance to *Xoo* strains PXO339, PXO99 and PXO86, and target promoter regions of the candidate lines were sequenced to confirm mutations. Seeds of T2 plants were bulked (from 30 plants per line), phenotyped for *Xoo* resistance and analyzed for agronomic traits as described below. Seeds from 13 individual T3 plants were advanced to T4 on the basis of three criteria: (1) mutation type, (2) consistent resistance to the three *Xoo* strains and (3) seed count. For Ciherang-Sub1, we obtained seven independent transformation events. Eighteen Cas9-free T1 plants were further analyzed for resistance and amplicon sequencing as described above up to the T3 generation.

Evaluation for agronomic traits of the rice mutant lines. We expected that mutant lines would show potential changes in agronomic traits owing to disruptive expression of the *SWEET* genes, off-target mutations across the genome or somaclonal variation. To assess the performance of the selected mutant lines, four agronomic characters in IR64-IRS1132 and Ciherang-Sub1-IRS1132 lines were assessed under paddy screenhouse conditions in a randomized block design experiment with three replications. At maturity, the plant height, panicle length, percentage of reproductive tillers (number of tillers with panicle/total number of panicles per plant) and percentage fertility (number of filled grains/total number of grains) were measured.

Disease assays. Fully expanded leaves of rice plants (6–8 weeks old) were inoculated using a leaf-tip clipping method. *Xoo* stocks, preserved in a −80 °C freezer, were streaked out on TSA plates supplemented with appropriate antibiotics and incubated at 28 °C for 2–4 d. The cells were harvested from the plates and resuspended in sterilized distilled water (optical density at 600 nm of 0.5, ~10^8 colony-forming units per milliliter). Scissor blades were immersed in *Xoo* suspension and used to clip about 2 cm from the leaf tip. The lesion lengths were measured 14 d after inoculation. Lesions were measured from each test plant. Lesion length measurements ≤ 5 cm were scored as resistant (R), 6–10 cm as moderately resistant (MR), 11–14 cm as moderately susceptible (MS) and ≥ 15 cm as susceptible (S). Three replications with about ten leaves from two to five plants per replicate were inoculated per strain.

Statistical analysis. Data are plotted using BoxPlotR (http://shiny.chemgrid.org/boxplotr/) or are presented as mean ± s.e.m. Numbers of measurements are specified in figure legends or in graphs. One-way analysis of variance (ANOVA) was conducted on all measurements. Tukey's honestly significant difference test was used for post-ANOVA pairwise tests for significance (set at $P < 0.05$), or, alternatively, the two-sided Dunnett's test was used. Exact P values, statistical tests used and sample numbers (n) can be found in figure legends or in graphs.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Links to genome sequencing data are available in Supplementary Table 3 and at BioProjects PRJNA497605 and PRJNA497307. Materials will be made available for non-profit research under a material-transfer agreement (Supplementary Note 1) and a standard material-transfer agreement (SMTA) that complies with the International Treaty on Plant Genetic Resources for Food and Agriculture (IT PGRFA) for IR64 or Ciherang-Sub1 lines (Supplementary Note 2). We aim at obtaining freedom-to-operate for use by low-income farmers and will work with breeders to make the materials available to subsistence farmers. For commercial applications, license agreements will be negotiated that are compliant with national and international law regarding biosafety, import and export, exchange of biological resources, intellectual property and seed systems. Profits will be used to support dissemination to subsistence farmers. Edited IR64- and Ciherang-Sub1-based materials can be obtained from R.O., edited Kitaake lines from B.Y., translational reporter lines from W.B.F. Distribution of *Xoo* strains may be restricted, owing to regulations on *X. oryzae* as a Select Agent by the US government, because of the Nagoya protocol, or because some strains were donated from other groups; these groups should be contacted directly (Supplementary Table 11).

References

33. Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
34. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
35. Haft, D. H. et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.* **46**, D851–D860 (2018).

36. Grau, J. et al. Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Comput. Biol.* **9**, e1002962 (2013).
37. Carver, T. J. et al. ACT: the Artemis comparison tool. *Bioinformatics* **21**, 3422–3423 (2005).
38. Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**, 464–469 (2012).
39. Pérez-Quintero, A. L. et al. QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Front. Plant Sci.* **6**, 545 (2015).
40. Pérez-Quintero, A. L. et al. An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLoS One* **8**, e68464 (2013).
41. Schatz, M. C. et al. 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 (2014).
42. Xie, K., Minkenberg, B. & Yang, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl Acad. Sci. USA* **112**, 3570–3575 (2015).
43. Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H. & Yang, B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucl. Acid Res* **42**, 10903–10914 (2014).
44. Hiei, Y. & Komari, T. *Agrobacterium*-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nat. Protoc.* **3**, 824–834 (2008).
45. Slamet-Loedin, I. H., Chadha-Mohanty, P. & Torrizo, L. *Agrobacterium*-mediated transformation: rice transformation. *Methods Mol. Biol.* **1099**, 261–271 (2014).

nature research

Corresponding author(s): NBT-RA47189

Last updated by author(s): 21 July 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The promoter sequences of rice genes in rice varieties Nipponbare and IR64 were aquired from Phytozome database v11. Genome sequences of published Xoo strains were obtained from NCBI with genome accession numbers indicated in Supplementary Table 1.
Data analysis	Canu V1.5, BlasR aligner, Varient-Caller, Prokka annotation pipeline, and NCBI Prokaryotic Genome Annotation were used for de novo assembly, annotation and gene prediction of Xoo genomes; AnnoTALE, Artemis genome browser and Artemis Comparative Tool were used for annotation and comparison of TALEs; Talvez was used for prediction of TALes. BoxPlotR (http://shiny.chemgird.org/boxplotr/) was use to plotted data; One-way analysis of variance (ANOVA) was conducted on all measurements; the Tukey honestly significant different test was used for post-ANOVA pair-wise tests for significance.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Xoo genome sequences were deposited in GenBank under Bioprojects PRJNA497307 and PRJNA497605.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed.
Data exclusions	No data were excluded in this study.
Replication	Most experiments were repeated at least three times independently. Where explicitly mentioned, two independent experiments are shown.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

ANNEXE 3 : Supplemental Data “Genome editing of an African elite rice variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains”

Supplementary Information

Genome editing of an African elite rice variety confers resistance against endemic and emerging *Xanthomonas oryzae* pv. *oryzae* strains

Van Schepler-Luu[1,7,a], Coline Sciallano[3,a], Melissa Stiebner[1,a], Chonghui Ji[2,a], Gabriel Boulard[3], Amadou Diallo[3], Florence Auguy[3], Si Nian Char[2], Yugander Arra[1], Kyrylo Schenstnyi[1], Marcel Buchholzer[1], Eliza P.I. Loo[1], Atugonza L. Bilaro[4], David Lihepanyama[4], Mohammed Mkuya[5], Rosemary Murori[6], Ricardo Oliva[7,8], Sebastien Cunnac[3], Bing Yang[2,9], Boris Szurek[3,*] & Wolf B. Frommer[1,10,*]

[1] Institute for Molecular Physiology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2] Division of Plant Science and Technology, Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA

[3] Plant Health Institute of Montpellier (PHIM), Université Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France

[4] Tanzania Agricultural Research Institute (TARI)-Uyole Centre, PO Box 400 Mbeya

[5] International Rice Research Institute, Eastern and Southern Africa Region (IRRI ESA), BOX 30709 00100, Nairobi, Kenya

[6] International Rice Research Institute (IRRI), Africa Regional Office, Nairobi, Kenya

[7] International Rice Research Institute, Pili Drive, Los Baños, Laguna 4031, Philippines

[8] new address: World Vegetable Center, Shanhua, Tainan, Taiwan 74151

[9] Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

[10] Institute for Transformative Biomolecules, ITbM, Nagoya University, Nagoya, Japan

7 Extended Data Figures

10 Supplementary Tables

2 Supplementary Data Sets

Supplementary Table 1. List of *Xoo* strains used for phenotyping.

CIX° Number	Strain name	CFBP ID	Country of origin	Region*	Year of isolation	NCBI Accession #
607	Nati Park	8172	Benin	WA	2013	CP033173
609	Tanguieta3	NA	Benin	WA	2013	NA
705	Karfiguela13	NA	Burkina Faso	WA	2013	NA
2679	BAI3	7321	Burkina Faso	WA	2004	GCF_003031385.1
4127	BAI250	NA	Burkina Faso	WA	2016	NA
1917	AXO1947	NA	Cameroon	CA	NA	GCF_001466505.1
2801	CFBP1948	1948	Cameroon	CA	1979	GCF_004355465.1
1042	CII-2	NA	Ivory Coast	WA	NA	NA
4083	CII-1	NA	Ivory Coast	WA	NA	NA
894	MAI145	NA	Mali	WA	2012	GCF_002850095.1
4079	MAI73	NA	Mali	WA	2012	GCF_002850075.1
629	Toula20	NA	Niger	WA	2013	GCF_004321555.1
2787	NAI9	7324	Niger	WA	2004	CP033177
4099	NAI5	7323	Niger	WA	2004	NA
4482	N2-4	NA	Niger	WA	2018	NA
4517	MAI132	NA	Niger	WA	2013	NA
2374	S62-2-22	NA	Senegal	WA	2014	CP036377
2976	S82-4-3	NA	Senegal	WA	2016	NA
4457	Tz1	NA	Tanzania	EA	2019	NA
4458	Tz2	NA	Tanzania	EA	2019	NA
4462	Tz6	NA	Tanzania	EA	2019	CP101704
4505	F1SP1-1d**	NA	Tanzania	EA	2021	NA
4506	F2SP1-1c*	NA	Tanzania	EA	2021	CP101703
4509	F4SP1-1a*	NA	Tanzania	EA	2021	CP101702
2439	PXO86	7202	Philippines	Asia	1977	NA
2831	PXO61	7201	Philippines	Asia	1973	CP021789.1
2840	PXO99	NA	Philippines	Asia	NA	GCA_000019585.2

°CIX, Collection IRD *Xanthomonas*. *WA, West Africa; CA, Central Africa; EA, East Africa. **: has not been sequenced, analysis on-going. *: sequenced but not pathotyped.

Supplementary Table 2. Disease survey in multiple rice growing areas in Tanzania in 2022. Note that the data were obtained from local farmers. Validation and genotyping of bacteria from infected leaf material will be initiated after transfer to our facilities.

FIELD NO.	VILLAGE	DISTRICT	REGION/ PROVINCE	VARIETY	SOWING DATE	STAGE* DURING SAMPLING	DISEASE SCALE**	GEOGRAPHIC LOCATION (GPS)
1	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900

	Lupiro	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.368452, 31.674900
2	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
	Lupiro	Ulanga	Morogoro	Lawama	March	Maturity	3	-8.370881, 36.677017
3	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378

	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
	Idunda	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.482672, 36.692378
4	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281

	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
	Minepa	Ulanga	Morogoro	TXD 306	March	Maturity	3	-8.252905, 36.683281
5	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
	Ifakara	Kilombero	Morogoro	Dunduli	February	Maturity	3	-8.137694, 36.684177
6	Ifakara	Kilombero	Morogoro	EST PN. 326/44	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	EST PN. 326/44	February	Maturity	3	-8.150813, 36.665661

	Ifakara	Kilombero	Morogoro	EST PN. 326/44	February	Maturity	3	-8.150813, 36.665661

	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
	Ifakara	Kilombero	Morogoro	DSR PN. 137/31	February	Maturity	3	-8.150813, 36.665661
8	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780

	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
	Njage	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.253896, 36.181780
9	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054
	Itongowa	Mlimba	Morogoro	TXD 306	March	Maturity	2	-8.352210, 36.080054

10	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
	Mkula	Kilombero	Morogoro	Lawama	March	Dough	2	-7.799772, 36.905489
11	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786

	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.420505, 37.544786
12	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267
	Dakawa	Mvomero	Morogoro	Supa	February	Milk	4	-6.419846, 37.544267

13	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.415932, 37.551897
14	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535

	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
	Dakawa	Mvomero	Morogoro	TXD 306	March	Maturity	3	-6.400923, 37.567535
15	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265

	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	3	-6.395002, 37.567265
16	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610
	Dakawa	Mvomero	Morogoro	TXD 306	March	Flowering	4	-6.392945, 37.60610

17	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
	Dakawa	Mvomero	Morogoro	TXD 306	March	Milk	5	-6.395147, 37.574687
18	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029

	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
	Dakawa	Mvomero	Morogoro	TXD 306	February	Maturity	4	-6.385051, 37.598029
19	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811

	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
	Lukenge	Mvomero	Morogoro	TXD 306	March	Maturity	4	-6.234562, 37.678811
20	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524
	Lukenge	Mvomero	Morogoro	TXD 306	April	Milk	3	-6.238532, 37.679524

21	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
	Kwam-tonga	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.189823, 37.599676
22	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718

	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
	Mkindo	Mvomero	Morogoro	TXD 306	March	Maturity	2	-6.256603, 37.539718
23	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
	Ruvu (Chauru)	Baga-moyo	Coast	TXD 306	March	Maturity	3	-6.742699, 38.668300
24	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098

	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
	Baga-moyo	Baga-moyo	Coast	Local	March	Maturity	2	-6.456316, 38.902098
25	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
	Mombo	Koro-gwe	Tanga	TXD 306	March	Maturity	2	-4.896162, 38.268456
26	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460

	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	3	-4.355418, 33.909460
27	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.350908, 33.914491
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.350908, 33.914491
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.350908, 33.914491
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.350908, 33.914491
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.350908, 33.914491
28	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349097, 33.916132
29	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919

	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.349588, 33.916919
30	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
	Mwama-puli	Igunga	Tabora	TXD 306	February	Reproductive	4	-4.347134, 33.925409
31	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.584835, 33.6997167
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.584835, 33.6997167
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.584835, 33.6997167
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.584835, 33.6997167
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.584835, 33.6997167
32	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116
	Ngana	Kyela	Mbeya	TXD 306	February	Maturity	2	-9.587494, 33.691116

33	Kilasilo	Kyela	Mbeya	SUPA	March	Maturity	2	-9.587998, 33.806142
	Kilasilo	Kyela	Mbeya	SUPA	March	Maturity	2	-9.587998, 33.806142
	Kilasilo	Kyela	Mbeya	SUPA	March	Maturity	2	-9.587998, 33.806142
	Kilasilo	Kyela	Mbeya	SUPA	March	Maturity	2	-9.587998, 33.806142
	Kilasilo	Kyela	Mbeya	SUPA	March	Maturity	2	-9.587998, 33.806142
34	Llilido	Mikin-dani	Mtwara	TARI RIC 1	March	Maturity	2	-10.352557, 39.704874
	Llilido	Mikin-dani	Mtwara	TARI RIC 2	March	Maturity	2	-10.352557, 39.704874
	Llilido	Mikin-dani	Mtwara	TARI RIC 3	March	Maturity	2	-10.352557, 39.704874
	Llilido	Mikin-dani	Mtwara	TARI RIC 4	March	Maturity	2	-10.352557, 39.704874
	Llilido	Mikin-dani	Mtwara	TARI RIC 5	March	Maturity	2	-10.352557, 39.704874
35	Lilido	Mikin-dani	Mtwara	Komboka	March	Maturity	2	-10.351676, 39.705414
	Lilido	Mikin-dani	Mtwara	Komboka	March	Maturity	2	-10.351676, 39.705414
	Lilido	Mikin-dani	Mtwara	Komboka	March	Maturity	2	-10.351676, 39.705414
	Lilido	Mikin-dani	Mtwara	Komboka	March	Maturity	2	-10.351676, 39.705414
	Lilido	Mikin-dani	Mtwara	Komboka	March	Maturity	2	-10.351676, 39.705414
36	Lilido	Mikin-dani	Mtwara	TARI RIC 1	March	Maturity	2	-10.350944, 39.706175
	Lilido	Mikin-dani	Mtwara	TARI RIC 2	March	Maturity	2	-10.350944, 39.706175
	Lilido	Mikin-dani	Mtwara	TARI RIC 3	March	Maturity	2	-10.350944, 39.706175
	Lilido	Mikin-dani	Mtwara	TARI RIC 4	March	Maturity	2	-10.350944, 39.706175
	Lilido	Mikin-dani	Mtwara	TARI RIC 5	March	Maturity	2	-10.350944, 39.706175
37	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278
	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278
	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278
	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278
	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278
	Lilido	Mikin-dani	Mtwara	TXD 306	March	Maturity	2	-10.353980, 39.705278

*According to the IRRI Standard Evaluation System for rice, nine growth stages were considered: 1-Germination, 2-Seedling, 3-Tillering, 4-Stem elongation, 5-Booting, 6-Heading, 7-Milk stage, 8-Dough stage, and 9-Mature grain.

**Bacterial Leaf Blight Disease score scale: 1=0-3%, 2=4-6%, 3=7-12%, 4=13-25%, 5=26-50%, 6=51-75%, 7= 76-87%, 8=88-94%, 9=95-100% (% of the leaf area which has been affected by the disease)

Supplementary Table 3. Accession numbers and main features of genome assemblies used for strain phylogeny inference. See separate Supplementary file.

Supplementary Table 4. List of SNPs and Indels in SWEET promoters after editing of the rice varieties of Komboka and Kitaake

Gene	Position before start codon*	Komboka	Kitaake
SWEET11a	816	+AA	
SWEET11a	857	+T	
SWEET11a	888	+TT	
SWEET13	132	-G	
SWEET13	221	-A (at PthXo2 EBE)	
SWEET13	232	T	G
SWEET13	248	G	C
SWEET13	1013	T	A
SWEET13	1122	+A	
SWEET13	1216	-TA	
SWEET13	1280	C	T
SWEET14	63	T	C
SWEET14	353	A	G
SWEET14	486	G	A
SWEET14	513	A	C

* for details see Supplementary Table 6

Supplementary Table 5. List of T0 lines with biallelic mutations at all targeted EBEs. Same mutation types are highlighted with the same colors. Asterisks indicate mutations outside of the EBE. For example: G/A*, +G, -4bp means one substitution of G to A outside of the EBE, one G insertion and a 4-bp deletion within the EBE.

Trans-forma-tion Experiment	T0 Event	PthXo1		PthXo2		TalC		PthXo3, AvrXa7, TalF	
		Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
1	1.1	-9bp	-12bp	C/T, -17bp*	-11bp	-4bp	-8bp	-11bp	T/A, -4bp
	1.2	-9bp	-12bp	-9bp	-11bp	-4bp	-8bp	-8bp	-7bp
	1.3	-9bp	-12bp	-3bp	-11bp	-4bp	-8bp	-11bp	T/A, -4bp
	1.4	-9bp	-12bp	-14bp*	-11bp	-4bp	-8bp	-8bp	-7bp
	1.5	-9bp	-12bp	-7bp	C/T, -11bp*	-7bp	-8bp	-8bp	G/T, -8bp
	1.6	-9bp	-12bp	A/C, -2bp	-11bp	-7bp	-8bp	-8bp	G/T, -8bp
	1.7	-9bp	-12bp	-10bp	-11bp	-7bp	-8bp	-8bp	G/T, -8bp
2	2	-10bp	-119bp*	C/A, -7bp	-23bp	+A	-8bp, A/C	-24bp	-7bp
	4	-9bp	G/A*, +G, -4bp	-8bp	-5bp	-3bp	+C	-14bp	-4bp, C/G
	7	G/C*, -13bp	-18bp	-15bp	-20bp	-3bp	-6bp	-9bp	-8bp
	9	-9bp	-12bp	A/C, -9bp	-17bp	-3bp	WT	-7bp	-20bp
	10	-13bp	G/A*, -14bp	-37bp	-8bp	-3bp, C/A	-7bp	-2bp	-6bp, T/A
	12	-10bp	-11bp	-5bp	-11bp	-5bp	T/A, -5bp	-6bp	-5bp
	14	-9bp (homo-allelic)		-10bp	-10bp	-7bp	+T	-7bp	-21bp
	16	-12bp	-11bp	C/T, -9bp	-7bp	-3pb	-6pb	-7bp	-12bp
	18	-8bp	-9bp	-5bp	-37bp	-3bp	-8bp, A/T	-13bp	-150bp*
	Total	15 variants		27 variants		16 variants		21 variants	

Supplementary Table 6. EBE sequences in select SWEET promoters of Komboka wild-type and CRISPR-edited lines. DNA sequences shaded in gray are SWEET EBEs for respective TALes. Letters in blue: PAM sequence for LbCpf1; letters in pink: PAM sequence for SpCas9; letters in red: TATA boxes; bold lowercase letters: insertion mutations; dashed lines: deletions; Italic lowercase letters: substitution mutations. Lines in orange and bold font were subjected to the pathogen tests. n.d., not determined.

SWEET11		AGCAAAGGTTAGATATGCATCTCCCCCTACTGTACACCACC AAAAGTGGAGGGTCTCCAACTATATAAACAC (PthXo1 EBE is shaded)	Genotype
	T0 generation		
1.1 to 1.7	Biallelic	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTAGATATGC---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9 -12
2	Biallelic	AGCAAAGGTTAGATATG---------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTAGATA----------------------- //---GCTAGCAT	-10 -119
4	Biallelic	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTA*a*ATATG**g**CAT---- CCCTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9 -4
7	Biallelic	AGCAAAGGTTA*c*ATATG---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTA--------------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-13 -18
9	Biallelic	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTAGATATGC---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9 -12
10	Biallelic	AGCAAAGGTTAGATAT---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTAaATA---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-13 -14
12	Biallelic	AGCAAAGGTTAGATATG---------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-10 -11

		AGCAAAGGTTAGATA---------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	
14	**Homo-zygous**	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
16	Biallelic	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAAC AC AGCAAAGGTTAGATA---------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-12 -11
18	Biallelic	AGCAAAGGTTAGATATGC-------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA ACAC	-8 -9
	T1 generation		
1.1_10	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.1_22	Homo-zygous	AGCAAAGGTTA*G**ATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9;*G or C
1.2_17	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAAC AC	-12
1.2_38	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.2_40	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.2_45	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.2_64	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.3_18	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.3_22	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
1.3_24	Homo-	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA	-9

	zygous	AACAC	
1.3_29	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.3_34	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.3_54	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
1.3_74	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
1.3_78	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.4_03	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.4_12	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.4_15	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
1.4_21	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.5_19	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
1.5_20	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
1.5_24	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
1.7_10	Homo-zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-12
9_32	Biallelic	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC AGCAAAGGTTAGATATGC---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9 /-12
9_33	Biallelic	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA	-9 /-12

		ACAC AGCAAAGGTTAGATATGC---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACA C	
9_34	Homo- zygous	AGCAAAGGTTAGATATGC---------- TGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACA C	-12
12_57	Homo- zygous	AGCAAAGGTTAGATA---------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-11
12_69	Homo- zygous	AGCAAAGGTTAGATA---------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-11
12_81	Homo- zygous	AGCAAAGGTTAGATA---------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-11
14_19	Homo- zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
14_29	Homo- zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
14_32	Homo- zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
14_65	Homo- zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC	-9
16_5	Homo- zygous	AGCAAAGGTTAGATATG---------- CTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAAC AC	-12
18_31	Biallelic	AGCAAAGGTTAGATATGC-------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA ACAC	-8 /-9
18_68	Homo- zygous	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA ACAC	-9
18_118	Biallelic	AGCAAAGGTTAGATATGC-------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATA AACAC AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA ACAC	-8 /-9
18_142	Homo-	AGCAAAGGTTAGATATGC--------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAA	-9

	zygous	ACAC	
18_160	Homo-zygous	AGCAAAGGTTAGATATGC-------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-8
T2 generation			
1.2_45_25	Homo-zygous	AGCAAAGGTTAGATATG---------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-10
1.3_24_1	Homo-zygous	AGCAAAGGTTAGATATG---------- TACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-10
1.3_24_29	Homo-zygous	AGCAAAGGTTAGATATG--------- CTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACAC	-9
SWEET13		CTTTTCTCCTATATAAGCACCACAACTCCCTTCATTCCTCT CCAAGAGTTTTCAGCCAACACATTGAACT (PthXo2A EBE is shaded)	
	T0 generation		
1.1	Biallelic	CTTTTCTCCTATATAAGCAC*t*-------------- CTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCACC---------- CATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-17 -11
1.2	Biallelic	CTTTTCTCCTATATAAGCAC--------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-9 -11
1.3	Biallelic	CTTTTCTCCTATATAAGCACC--- ACTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-3 -11
1.4	Biallelic	CTTTTCTCCTATATAAGCAC---------- TTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14 -11
1.5	Biallelic	CTTTTCTCCTATATAAGCAC------- CCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC*tA*---------- ATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-7 -11
1.6	Biallelic	CTTTTCTCCTATATAAGCACC*c*-- ACTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC----------	-2 -11

		TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	
1.7	Biallelic	CTTTTCTCCTATATAAGCAC---------- TTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCACC---------- CATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10 -11
2	Biallelic	CTTTTCTCCTATATAAGCAC*a*------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T CTTTTCTCCTATATAA---------------- TCTCCAAGAGTTTTCAGCCAACACATTGAACT	-7 -23
4	Biallelic	CTTTTCTCCTATATAAGCACC-------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCACC----- TCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGA ACT	-8 -5
7	Biallelic	CTTTTCTCCTATATAAGCAC----------- TCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTAT---------- CATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-15 -20
9	Biallelic	CTTTTCTCCTATATAA*c*GCACC--------- TTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTA---------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-9 -17
10	Biallelic	CTTTTCTCCTATATAAGCAC----------//--- AACACATTCCTCTC CTTTTCTCCTATATAAGCACC-------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-37 -8
12	Biallelic	CTTTTCTCCTATATAAGCACCA----- CCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAA CT CTTTTCTCCTATATAAGC---------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-5b -11
14	Biallelic	CTTTTCTCCTATATAAGCACC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC---------- TTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10 -10
16	Biallelic	CTTTTCTCCTATATAAGCAC*t*A--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCAC------- CCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAA CT	-9 -7
18	Biallelic	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG	-5

		AACT TTCGCA--//---------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-37
	T1 generation		
1.1_10	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.1_22	Homo-zygous	CTTTTCTCCTATATAAGCA---------- CCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-17
1.2_17	Homo-zygous	CTTTTCTCCTATATAAGCA--------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-9
1.2_38	Homo-zygous	CTTTTCTCCTATATAAGCA--------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-9
1.2_40	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.2_45	Homo-zygous	CTTTTCTCCTATATAAGCA--------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-9
1.2_64	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.3_18	Homo-zygous	CTTTTCTCCTATATAAGCACC--- ACTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATT GAACT	-3
1.3_22	Homo-zygous	CTTTTCTCCTATATAAGCACC--- ACTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATT GAACT	-3
1.3_24	Homo-zygous	CTTTTCTCCTATATAAGCACC--- ACTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATT GAACT	-3
1.3_29	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.3_34	Homo-zygous	CTTTTCTCCTATATAAG---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14
1.3_54	Homo-zygous	CTTTTCTCCTATATAAG---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14
1.3_74	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.3_78	Homo-zygous	CTTTTCTCCTATATAAG---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14
1.4_03	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14
1.4_12	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11

1.4_15	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-14
1.4_21	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.5_19	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.5_20	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.5_24	Homo-zygous	CTTTTCTCCTATATAAGCAC---------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-11
1.7_10	Homo-zygous	CTTTTCTCCTATATAAGCACC--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10
9_32		n.d	
9_33	Homo-zygous	CTTTTCTCCTA---------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-17
9_34	Homo-zygous	n.d.	
12_57	Homo-zygous	CTTTTCTCCTATATAAGCACCA---- CCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAA CT	-5
12_69	Homo-zygous	CTTTTCTCCTATATAAGCACCA---- CCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAA CT	-5
12_81	Homo-zygous	CTTTTCTCCTATATAAG---------- CCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAAC T	-11
14_19	Homo-zygous	CTTTTCTCCTATATAAGCA---------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10
14_29	Biallelic	CTTTTCTCCTATATAAGCA---------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCA---------- C*c*TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10 /-10
14_32	Biallelic	CTTTTCTCCTATATAAGCA---------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCA---------- C*c*TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10/ -10
14_65	Biallelic	CTTTTCTCCTATATAAGCA---------- CTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT CTTTTCTCCTATATAAGCA---------- C*c*TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-10/ -10
16_5	Homo-zygous	CTTTTCTCCTATATAAGCACCA--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-9
18_31	Homo-zygous	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG	-5

		AACT	
18_68	Homo-zygous	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG AACT	-5
18_118	Homo-zygous	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG AACT	-5
18_142	Homo-zygous	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG AACT	-5
18_160	Homo-zygous	CTTTTCTCCTATATAAGCAC----- CTCCCTTCATTCCTCTCCAAGAGTTTTCAGCCAACACATTG AACT	-5
	T2 generation		
1.2_45_25	Homo-zygous	CTTTTCTCCTATATAAGCACCA--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-9
1.3_24_1	Homo-zygous	CTTTTCTCCTATATAAGCACCA--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-9
1.3_24_29	Homo-zygous	CTTTTCTCCTATATAAGCACCA--------- TCATTCCTCTCCAAGAGTTTTCAGCCAACACATTGAACT	-9
SWEET14		CATGCATGTCAGCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGG TGCTAAGCTCATCAAGCCTTCAAGCAAA (TalC, PthXo3, AvrXa7, TalF)	
	T0 generation		
1.1	Biallelic	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ---------AAGCCTTCAA CATGCA-------- GCTGGTCATG...TAATAAACCCCCTCCAACCAGGTGC*a*-- --TCATCAAGCCTTCAA	-4/ -11 -8/ -4
1.2	Biallelic	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA CATGCA-------- GCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC-- -----ATCAAGCCTTCAA	-4/ -8 -8/ -7
1.3	Biallelic	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ---------AAGCCTTCAA CATGCA-------- GCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC*a*- ---TCATCAAGCCTTCAA	-4/ -11 -8/ -4
1.4	Biallelic	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8 -8/ -7

		CATGCA-------- GCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC-- -----ATCAAGCCTTCAA	
1.5, 1.6 and 1.7	Biallelic	CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA CATGCA-------- GCTGGTCATG...TATATAAACCCCCTCCAACCAGGT*t*--- -----ATCAAGCCTTCAA	-7/ -8 -8/ -8
2	Biallelic	CATGCATGTCAGCA**a**GCTGGTCATG...TATATAAACCCCC TCCAACCAGG-------//--CAAAGCAACTA CATGCA-------- GCTGGTC*c*TG...TATATAAACCCCCTCCAACCAGGTG--- ----CATCAAGCCTTCAA	+**a**/ -24 -8/ -7
4	Biallelic	CATGCATGTCAGC--- TGGTCATG...TATATAAACCCCCTCCAACCAGGTGC---- ----------CTTCAA CATGCATGTCAGC**c**AGCTGGTCATG...TATATAAACCCCC TCCAACCAGGTGCTA----CATCAAGC*g*TTCAA	-3/ -14 +**c**/ -4
7	Biallelic	CATGCATGTCAGC--- TGGTCATG...TATATAAACCCCCTCCAACCAGGTG----- ---ATCAAGCCTTCAA CATGCAT------ AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC- --------CAAGCCTTCAA	-3/ -8 -6/ -9
9	Biallelic	CATGCATGT--- CAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG- ------CATCAAGCCTTCAA CATGCATGTCAGCAGCTGGTCATG...TATATAAACCCCCT CCAA-------------------GCCTTCAA	-3/ -7 WT/ -20
10	Biallelic	CATGCATGTCAGC--- TGGT*a*ATG...TATATAAACCCCCTCCAACCAGGTGCTAAG CT--TCAAGCCTTCAA CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGCT AA------TCAAGCC*a*TCAA	-3/ -2 -7/ -6
12	Biallelic	CATGCAT----- CAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG- -----TCATCAAGCCTTCAA CATGCA*a*G----- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGCT AA-----TCAAGCCTTCAA	-5/ -6 -5/ -5
14	Biallelic	CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGT--- ----TCATCAAGCCTTCAA CATGCATGTCAGC**t**AGCTGGTCATG...TATATAAACCCCC TCCAACC-------------------TTCAA	-7/ -7 +**t**/ -21

16	Biallelic	CATGCATGTCAGC--- TGGTCATG...TATATAAACCCCCTCCAACCAGGTG----- --CATCAAGCCTTCAA CATGCAT------ AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC- ----------GCCTTCAA	-3/ -7 -6/ -12
18	Biallelic	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA CATGCA-------- GCTGGTC*t*TG...TATATAAACCCCCTCCAACCAGGTG--- ---------//--TACTG	-3/ -13 -8/ -150
	T1 generation		
1.1_10	Homo- zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.1_22	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.2_17	Homo- zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/-7
1.2_38	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.2_40	Homo- zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.2_45	Homo- zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.2_64	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.3_18	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.3_22	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.3_24	Homo- zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.3_29	Homo- zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.3_34	Homo-	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG--	-4/ -8

	zygous	------ATCAAGCCTTCAA	
1.3_54	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.3_74	Homo-zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.3_78	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.4_03	Homo-zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.4_12	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.4_15	Homo-zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.4_21	Homo-zygous	CATGCATGT---- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-4/ -8
1.5_19	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGT*t* --------ATCAAGCCTTCAA	-8/ -8
1.5_20	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGT*t* --------ATCAAGCCTTCAA	-8/ -8
1.5_24	Homo-zygous	CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-7/ -8
1.7_10	Homo-zygous	CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG-- ------ATCAAGCCTTCAA	-7/ -8
9_32	Homo-zygous	CATGCATGT--- CAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG- ------CATCAAGCCTTCAA	-3/ -7
9_33	Homo-zygous	CATGCATGTCAGCAGCTGGTCATG...TATATAAACCCCCT CCAACCAGGTGCTAAGCTCATCAAGCCTTCAA	WT
9_34	Homo-zygous	CATGCATGT--- CAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG- ------CATCAAGCCTTCAA	-3 /-7
12_57	Homo-zygous	CATGCATG----- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGCT AA-----TCAAGCCTTCAA	-5/ -5
12_69	Homo-zygous	CATGCATG----- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGCT	-5/ -5

		AA-----TCAAGCCTTCAA	
12_81	Homo-zygous	CATGCATG----- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGCT AA-----TCAAGCCTTCAA	-5/ -5
14_19	Homo-zygous	CATGCATGTCAGC**t**GCTGGTCATG...TATATAAACCCCCT CCAACC----------TTCAA	+**t**/ -21
14_29	Homo-zygous	CATGCATGTCAGC**t**GCTGGTCATG...TATATAAACCCCCT CCAACC----------TTCAA	+**t**/ -21
14_32	Homo-zygous	CATGCATGTCAGC**t**GCTGGTCATG...TATATAAACCCCCT CCAACC----------TTCAA	+**t**/ -21
14_65	Homo-zygous	CATGCA------- AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGT--- ----TCATCAAGCCTTCAA	-7/ -7
16_5	Homo-zygous	CATGCAT------ AGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTGC- ----------GCCTTCAA	-6/ -12
18_31	Homo-zygous	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA	-3/ -13
18_68	Homo-zygous	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA	-3/ -13
18_118	Homo-zygous	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA	-3/ -13
18_142	Homo-zygous	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA	-3/ -13
18_160	Homo-zygous	CATGCATGTCAGCAG-- GTCATG...TATATAAACCCCCTCCAACC---------- -ATCAAGCCTTCAA	-3/ -13
	T2 generation		
1.2_45_25	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.3_24_1	Homo-zygous	CAT-------- GCA*t*CTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7
1.3_24_29	Homo-zygous	CAT-------- GCAGCTGGTCATG...TATATAAACCCCCTCCAACCAGGTG C-------ATCAAGCCTTCAA	-8/ -7

Supplementary Table 7. List of primers used in this study.

Primer pair name	Forward Primer sequence 5'-3'	Reverse Primer sequence 5'-3'	Fragment size (bp)	Description	Reference
ZmUbi-Cas9-F1R1	CGTTGTAAAACGACGGCCAGTGCCAAGCTTCTGCAGTGCAGCGTGACC	TTCTTGTACAAAGTGGTGAAGCTTGATCTAGTAACATAGATGACACCG	6412	amplifies Zmubip-Cas9-Tnos expression cassette	This study
cXo1-FR	agatGTGGTGTACAGTAGGGGGAGATG	gcctCATCTCCCCCTACTGTACACCAC		cRNA targeting pthxo1 EBE	This study
cXo2c-FR	agatTCCTATATAAGCACCACAACTCC	gcctGGAGTTGTGGTGCTTATATAGGA		cRNA targeting pthxo2b EBE	This study
cXo2d-FR	agatTTCCTATATAAAGCACCACAACT	gcctAGTTGTGGTGCTTTATATAGGAA		cRNA targeting pthxo2c EBE	This study
cTalF-FR	agatCTTGAAGGCTTGATGAGCTTAGC	gcctGCTAAGCTCATCAAGCCTTCAAG		cRNA targeting pthXo3-TalF EBE	This study
pTLN-seqF	tggcccgtgtctcaaaatctctg			cRNA unit Sanger sequencing	This study
gTalC-FR	tgttGGGCATGCATGTCAGCAGC	aaacGCTGCTGACATGCATGCCC		gRNA targeting TalC EBE	This study
U6P-gRNA-F3R3	tcgcggccgcactcgagatatctagAAGAACGAACTAAGCCGGAC	TTGTACAAGAAAGCTGGGTCTAGA	998	Amplifies gTalC-gRNA expression cassette	This study
pENTR-seqF	GCGTTTCTACAAACTCTTCCTGT			donor plasmid Sanger sequencing	This study
OsSWEET14 F9R9	TGCGGCTCATCAGTTTCTCT	ATGCATGGCCCCTCTAATGT	958	amplifies OsSWEET14 promoter in Komboka	This study
OsSWEET13 F6R6	AAATTACCACGGACGTCTCG	AATGTCGTCGCTAGCTATGC	1557	amplifies OsSWEET13 promoter in Komboka	This study
OsSWEET11 F7R7	CATGCCCGTTTTCAGCAAAC	ACTGCTCAACCGATCTCCAT	2787	amplifies OsSWEET11 promoter in Komboka	This study
OsCas9-FR	GGGTAATGAACTCGCTCTGC	TGGCGTCAAGAACTTCCTTTG	370	checking for Cas9-positive plants	[3]
OsCas9-F1R1	TGACAGGCACTCTATAAAGAAGA	ATAAAAAGCTTGTCAACATCTGA	448	checking for Cas9-positive plants	[3]
OsCas9-F2R2	ATGACTAGAAAGTCCGAAGAAAC	ATGTCCTCCAGTATATCTTCGTT	464	checking for Cas9-positive plants	[3]
Cpf1_F1R1	AAGTGGTCCAATGCCTACGA	TTGTTGGCGATAGGGGAGTT	359	checking for Cpf1-positive plants	This study
HptFR	AGCCTGACCTATTGCATCT	CATATGAAATCACGCCATGT	200	checking for Hygromycin selection marker-positive plants	[3]
GET_SWT11_F1R1	TGAGTGGTCATACGTGTCATATTGC	AGTGCTAGGTCTTTTGAAGTGAGATTC	900	For amplifying and sequencing targeted site on OsSWEET11 promoter	This study

GET_SWT13_F1R1	GCTTAGCTTCATATACGTGGGAGTAG	ATGTCTTAGTTGTGGTAGGTGATGATG	1185	For amplifying and sequencing targeted site on OsSWEET13 promoter	This study
GET_11N3_F1_SWT14	GCTCTCACCATTCATTCCACT	GGG ATG CTG AAG AGA CAT GC	663	For amplifying and sequencing targeted site on OsSWEET14 promoter	This study
GET-SWEET11	GGGATTTCTGGCTAGTTTCT	CGAGGTAGAGGACGATGTAG		qPCR primer	This study
GET-SWEET13 F	AGAGTTTTCAGCCAACACAT	TGTAGATCCGGTAGAACGTC		qPCR primer	This study
GET-SWEET14 F	TATTGCCTGATCATCCTCTT	GTGAACATCTTGGCCTTCT		qPCR primer	This study
GET- EF1α -F	GAAGTCTCATCCTACCTGAAGAAG	GTCAAGAGCCTCAAGCAAGG		qPCR primer	This study
HPT-FAM	CTATCAGAGCTTGGTTGACGG	CTTCTACACAGCCATCGGTC		ddPCR primer/probe	This study
HPT-probe	/56-FAM/ACACAAATC/ZEN/GCCCGCAGAAGC/3IABkFQ/			ddPCR primer/probe	This study
TPI-HEX	CACCCAGAAACAATGCCATC	CGAAGACGGCTATTGAGACTG		ddPCR primer/probe	This study
TPI-probe	/5HEX/CCTGCCTTC/ZEN/ACTTGGTATTACATGTTGC/3IABkFQ/			ddPCR primer/probe	This study
Cas9-FAM	AGACAAACGGAGAAACAGGAG	TTGCTAAACCCACCTGTCTG		ddPCR primer/probe	This study
Cas9-probe	/56-FAM/CCGCCCTTT/ZEN/GTCCCATACGATCT/3IABkFQ/			ddPCR primer/probe	This study
Cpf1-FAM	AACAATCACGGACAGATCAGG	AGGGTTGTGGTTTTCTTGGG		ddPCR primer/probe	This study
Cpf1-probe	/56-FAM/CA GCC AAC T/Zen/C CCC TAT CGC CAA /3IABkFQ/			ddPCR primer/probe	This study
OsEF-1alpha_qPCR	GAAGTCTCATCCTACCTGAAGAAG	GTCAAGAGCCTCAAGCAAGG		monitoring *OsEF1α* housekeeping gene for qPCR	[4]
Sweet11a	AGTCGACGGGAGGGTACAG	TTCGGGTACATGACGTAGGG		monitoring *SWEET11a* mRNA levels by qPCR	[5]
Sweet13	GGCCTGTCCCTGCAGCATC	CCCGAACACCCCCACGTTC		monitoring *SWEET13* mRNA levels by qPCR	[5]
Sweet14	ACTTGCAAGCAAGAACAGTAGT	ATGTTGCCTAGGAGACCAAAGG		monitoring *SWEET14* mRNA levels by qPCR	[5]

XaL-F2 / Xa37d-R1	CCCTCTTGCACACGCCATTGG	ACGGTTCTGAAGGTCGTCAT	6605	Amplification of the whole length *Xa1-like* gene from Komboka	[2]
KS156	GACCATGAGTTGGCTATTAG			Sequencing of *Xa1-like* gene from Komboka	[2]
Xa1upT-F	CGTGTAACAAGTACTGACCA			Sequencing of *Xa1-like* gene from Komboka	This study
Xa14-F4	GAGGCTATGTCGATCCTCA			Sequencing of *Xa1-like* gene from Komboka	[2]
gXa1BE-R4	CATTGTCTTAGATGATGTGT			Sequencing of *Xa1-like* gene from Komboka	[2]
Xa1C-F	TCCAGTGAGAAGTTGGAGCA			Sequencing of *Xa1-like* gene from Komboka	[2]
KS113	CTTAGAAAAGCTACATTTGTCC			Sequencing of *Xa1-like* gene from Komboka	[2]
XaL-F2	CCTTCCGCAATCAGTGGTT			Sequencing of *Xa1-like* gene from Komboka	This study
XaL-R2	AATGTAGCAAGCGACGCACA			Sequencing of *Xa1-like* gene from Komboka	[2]
KS157	TTTACAGCAGCTTGAACGCC			Sequencing of *Xa1-like* gene from Komboka	[2]
Xa4.0FR	GGTTCCAAATTAAGCAACACTTATT	CAGTGTTTTTCGGCCACGAAGG	1163	Amplification of fragment 1 *Xa4* gene from Komboka	This study
Xa4.1FR	CGAGTGATTCTTTGCAGGTA	TGACCGTTAATCTTATT	1616	Amplification of fragment 2 *Xa4* gene from Komboka	This study
Xa4.2FR	TGTTTGACCGTTCGTCTT	CACTCACTATGTTCGCTCTT	1803	Amplification of fragment 3 *Xa4* gene from Komboka	This study
Xa4.3FR	GCGAACATAGTGAGTGCGTC	CTACACGAGTTCGACCGGCG	1702	Amplification of fragment 4 *Xa4* gene from Komboka	This study
M13F	CCCAGTCACGACGTTGTAAAACG			Sequencing of subcloned *Xa4* fragments from Komboka	This study
M13R	CACACAGGAAACAGCTATGACCATG			Sequencing of subcloned *Xa4* fragments from Komboka	This study

Supplementary Table 8: Guide RNA sequences. Green letters indicate the protospacer adjacent motif (PAM) sequence (NGG).

Genome	gRNA/ cRNA	Target locus	Sequence
Oryza sativa japonica (Nipponbare, IRGSP1.0)	cXo1	chr08: 26728846-26728872	TTTG GTGGTGTACAGTAGGGGG AGATG
Oryza sativa indica (Shuhui498)	cXo1	chr08: 28035804-28035830	TTTG GTGGTGTACAGTAGGGGG AGATG
Oryza sativa japonica (Nipponbare, IRGSP1.0)	cXo2c	Komboka specific	TTTC TCCTATATAAGCACCACA ACTCC
Oryza sativa indica (Shuhui498)	cXo2c	Komboka specific	TTTC TCCTATATAAGCACCACA ACTCC
Oryza sativa japonica (Nipponbare, IRGSP1.0)	cXo2d	variety specific	TTTG TTCCTATATAAAG CACCACAACT
Oryza sativa indica (Shuhui498)	cXo2d	variety specific	TTTG TTCCTATATAAAG CACCACAACT
Oryza sativa japonica (Nipponbare, IRGSP1.0)	cTalF	chr11: 18174461-18174487	TTTG CTTGAAGGCTTGATGAGC TTAGC
Oryza sativa indica (Shuhui498)	cTalF	chr11: 20518333-20518359	TTTG CTTGAAGGCTTGATGAGC TTAGC
Oryza sativa japonica (Nipponbare, IRGSP1.0)	gTalC	chr11: 18174551-18174572	GGGCATGCATGTCAGCAG C TGG
Oryza sativa indica (Shuhui498)	gTalC	chr11: 20518423-20518444	GGGCATGCATGTCAGCAG C TGG

Supplementary Table 9. Screening at T1 generation for homozygous mutation of all 6 EBEs.

	T0 event	# of plants screened	Homozygous mutations in all 6 EBEs
Transformation 1	1.1	14	2
	1.2	52	5
	1.3	87	8
	1.4	21	4
	1.5	26	3
	1.6	21	0
	1.7	12	1
	Total	**233**	**23**
Transformation 2	12	26	3
	14	10	4
	16	16	0
	Total	**52**	**7**
	Sum	**285**	**30**

Supplementary Table 10. List of T1 lines with homozygous mutations of all targeted EBEs. n.d. not determined.

EBE / Line ID	SWEET11 PthXo1	SWEET13 PthXo2	SWEET14 TalC	SWEET14 PthXo3	SWEET14 AvrXa7	SWEET14 TalF	SpCas9 gene copy/2n
1.1_10	-9bp	-11bp	-8bp	-5bp	-4bp	-5bp	7
1.1_22	-9bp	-17bp	-4bp	-5bp	-4bp	-6bp	3
1.2_17	-12bp	-9bp	-8bp	-5bp	-4bp	-5bp	5
1.2_38	-9bp	-9bp	-4bp	-5bp	-4bp	-6bp	7
1.2_40	-9bp	-11bp	-8bp	-5bp	-4bp	-5bp	4
1.2_45	-9bp	-9bp	-8bp	-5bp	-4bp	-5bp	4
1.2_64	-9bp	-11bp	-4bp	-5bp	-4bp	-6bp	4
1.3_18	-9bp	-3bp	-4bp	-5bp	-4bp	-6bp	3
1.3_22	-9bp	-3bp	-4bp	-5bp	-4bp	-6bp	1
1.3_24	-9bp	-3bp	-8bp	-5bp	-4bp	-5bp	2
1.3_29	-12bp	-11bp	-4bp	-5bp	-4bp	-6bp	4
1.3_34	-12bp	-14bp	-4bp	-5bp	-4bp	-6bp	4
1.3_54	-9bp	-14bp	-8bp	-5bp	-4bp	-5bp	6
1.3_74	-9bp	-11bp	-4bp	-5bp	-4bp	-6bp	3
1.3_78	-12bp	-14bp	-8bp	-5bp	-4bp	-5bp	7
1.4_03	-12bp	-14bp	-4bp	-5bp	-4bp	-6bp	6
1.4_12	-12bp	-11bp	-8bp	-5bp	-4bp	-5bp	4
1.4_15	-9bp	-14bp	-4bp	-5bp	-4bp	-6bp	4
1.4_21	-12bp	-11bp	-4bp	-5bp	-4bp	-6bp	4
1.5_19	-12bp	-11bp	-8bp	-7bp; G/T	G/T, -4bp	-5bp	4
1.5_20	-9bp	-11bp	-8bp	7bp; G/T	G/T, -4bp	-5bp	6
1.5_24	-9bp	-11bp	-7bp	-5bp	-4bp	-6bp	6
1.7_10	-12bp	-10bp	-7bp	-5bp	-4bp	-6bp	4
12_57	-11bp	-5bp	-5bp	-2bp	-1bp	-4bp	0
12_69	-11bp	-5bp	-5bp	-2bp	-1bp	-4bp	0
12_81	-11bp	-11bp	-5bp	-2bp	-1bp	-4bp	0
14_19	-9bp	-10bp	+T	-12bp	-11bp	-12bp	0
14_29	-9bp	-10bp/ -10bp	+T	-12bp	-11bp	-12bp	0
14_32	-9bp	-10bp/ -10bp	+T	-12bp	-11bp	-12bp	0
14_65	-9bp	-10bp/ -10bp	-7bp	-6bp	-5bp	-4bp	2
16_5	-12bp	-9bp	-6bp	-4bp	-3bp	-11bp	n.d.

Supplementary Data 1. Predicted coding sequences of the TALe genes in the genomes of strains CIX4506, CIX4509 and CIX4462 from the Morogoro region of Tanzania that were examined in this study.

See separate Supplementary file.

Supplementary Data 1. OsSWEET promoter sequences of Komboka and Kitaake

***OsSWEET11a* sequence. PthXo1 EBE is highlighted in yellow. The first exon is highlighted in grey.**

>OsSWEET11a_Komboka_ 1756 bp.

CTCTCTAAGAATAGGCATATTATGGTTTAATTAGCTAGACAAGAAATTAGTTAAGGTTCTTAATGAAGATTCCGT
CACTTTTGCTAGCTTTGAAAACCTGCAGAGTGAATTGTGCAAAACATCTTGGCATGTTGGTGTTAGTGGTACAAC
AGATGCTGATACAAAGAAATTAATGCTGCAATTGTTAGAAGCTCTTTTTTTTTTCTCTCTTTTGAACACCAAAGC
ACTTTTGTTCATGGTGAAAGGGACTACCCTTCTCTTTCATATGTTCCCTTCTTCCTCTGTTCTTTTCATGAGATT
TATTTGATGACTTCCCTCTGCTCTGTTATTTTCATGTTCTTTGATAAATGTGCTGATTGTATTATCAGTGTTATT
GCACTCAATTCATTCTTGAACAGTGTGTGAATTAACTCTTTGTTCGATTTGAGCTTAGTCACTTGATTGCACACG
AACTACTCTGCAATTCTTTTCTGACGATAGAAGTCGATTGATGGATCACCCAACATGTTACTAATTAAGTTGCAT
CATTGTCCATGGTTGTACATCCTTCTACAAATAAAACTACACAAATCAAGAAAATTTTCAATAACATTTCAACTA
TTGTAACAAGTAAAAGAAACCTATATGAGAGCTCCAGCTCTCCAAATGGCAACAGACACACTGAGTGGTCATACG
TGTCATATTGCCCCTCAGTTATGCATTCATATGACCACATATTCAGAGTAGTGGAGAGAGGGACAGATCTAGAGG
TAGAAAAAGAAAATTCATATAAATGATATATCAGAGTGAAAAAGAAATATCAAGCACAAGAAAAAAAAAAGCAAA
GGTTAGATATGCATCTCCCCCTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACACTGAGCCATG
GCCAAGGCCAAACCACACATGCAGTTGTAGTAGCACTTAAGCCTTCCTCTCTAGCTAGCATCTCTTGTGTCAGGA
AGTTGGAAGGGATTTCTGGCTAGTTTCTAGCTGGTGTCTCCTCTCCTCTTCCTAACCTTCTCACTGATTAACACC
TTAGAGTTAGTTAATAACCTTCATCACCAGTAGCAATGGCAGGAGGTTTCTTGTCCATGGCTAACCCGGCGGTCA
CCCTCTCCGGTGTTGCAGGTAAAGCATGCAACCAATGCATAATGCTCAAACTTAATTTCATCATTATCATCATCA
TCATCTTCACAGCCATGATCATCCATGGACAAATGCAACTGAAGATCATTTTAGTTTTCATATGCTAATGATCAA
ATTCAGGTTAATTGCTGTTTAATTTCTCCATACACTAGTTGTCTGCACCATTGCATTGTGCACAGCACACACACG
CTTTTGATGCTTCTAGGAATGCATATCTGTTCAGCAGTTCACACAGTGCAGCAGGGCAATGTTGTTAAAAAATCT
TCTCCTTTTTTTTTATGTCCTTGTATTCTTGAGCTTTCTGTCTCCATTGATCTGCTTTTTTCTTGTTTACAAGTG
ATGGGCACAAGTCACTTCCCTTAGCTTCAGCTCATGCATGGAGCAGGAATCTCACTTCAAAAGACCTAGCACTTT
TTCTCTCTTCACCTTTTTGCCTCAACACATGCCCAGTTTCTGGCCACACAAACATAAACACATATACTATCTAGC
TGCATAATTGCATCAAATTAAGCAGGGTTTGTTTCAGCTAGGAATTCCACACATAGGTCATTAATTAGTATTGCC
AACTTTCTCAACATGCATGCACTCTAGTACT

> OsSWEET11a_Kitaake_OsativaKitaake_499.genome_Chr8 Chr8: 27602383..27604137 (-strand) class=match length=1755 bp

CTCTCTAAGAATAGGCATATTATGGTTTAATTAGCTAGACAAGAAATTAGTTAAGGTTCTTAATGAAGATTCCGT
CACTTTTGCTAGCTTTGAAAACCTGCAGAGTGAATTGTGCAAAACATCTTGGCATGTTGGTGTTAGTGGTACAAC
AGATGCTGATACAAAGAAATTAATGCTGCAATTGTTAGAAGCTCTTTTTTTTTCTCTCTTTTGAACACCAAAGCAC
TTTGTTCATGGTGAAAGGGACTACCCTTCTCTTTCATATGTTCCCTTCTTCCTCTGTTCTTTTCATGAGATTTAT
TTGATGACTTCCCTCTGCTCTGTTATTTTCATGTTCTTTGATAAATGTGCTGATTGTATTATCAGTGTTATTGCA
CTCAATTCATTCTTGAACAGTGTGTGAATTAACTCTTTGTTCGATTTGAGCTTAGTCACTTGATTGCACACGAAC
TACTCTGCAATTCTTTTCTGACGATAGAAGTCGATTGATGGATCACCCAACATGTTACTAATTAAGTTGCATCAT
TGTCCATGGTTGTACATCCTTCTACAAATAAAACTACACAAATCAAGAAAATTTTCAATAACATTTCAACTATTG
TAACAAGTAAAAGAAACCTATATGAGAGCTCCAGCTCTCCAAATGGCAACAGACACACTGAGTGGTCATACGTGT
CATATTGCCCCTCAGTTATGCATTCATATGACCACATATTCAGAGTAGTGGAGAGAGGGACAGATCTAGAGGTAG
AAAAAGAAAATTCATATAAATGATATATCAGAGTGAAAAAGAAATATCAAGCACAAGAAAAAAAAGCAAAGGTTA
GATATGCATCTCCCCCTACTGTACACCACCAAAAGTGGAGGGTCTCCAACTATATAAACACTGAGCCATGGCCAA
GGCCAAACCACACATGCAGTTGTAGTAGCACTTAAGCCTTCCTCTCTAGCTAGCATCTCTTGTGTCAGGAAGTTG
GAAGGGATTTCTGGCTAGTTTCTAGCTGGTGTCTCCTCTCCTCTTCCTAACCTTCTCACTGATTAACACCTTAGA
GTTAGTTAATAACCTTCATCACCAGTAGCAATGGCAGGAGGTTTCTTGTCCATGGCTAACCCGGCGGTCACCCTC
TCCGGTGTTGCAGGTAAAGCATGCAACCAATGCATAATGCTCAAACTTAATTtcatcatcatcatcatcatcatc
atcTTCACAGCCATGATCATCCATGGACAAATGCAACTGAAGATCATTTTAGTTTTCATATGCTAATGATCAAAT
TCAGGTTAATTGCTGTTTAATTTCTCCATACACTAGTTGTTGTCTGCACCATTGCATTGTGCACAGCACACACAC
GCTTTTGATGCTTCTAGGAATGCATATCTGTTCAGCAGTTCACACAGTGCAGCAGGGCAATGTTGTTAAAAAATC
TTCTCCTTTTTTTTTATGTCCTTGTGTTCTTGAGCTTTCTGTCTCCATTGATCTGCTTTTTTCTTGTTTACAAGTG
ATGGGCACAAGTCACTTCCCTAGCTTCAGCTCATGCATGGAGCAGGAATCTCACTTCAAAAGACCTAGCACTTTT
TCTCTCTTCACCTTTTTGCCTCAACACATGCCCAGTTTCTGGCCACACAAACATAAACACATATACTATCTAGCT
GCATAATTGCATCAAATTAAGCAGGGTTTGTTTCAGCTAGGAATTCCACACATAGGTCATTAATTAGTATTGCCA
ACTTTCTCAACATGCATGCACTCTAGTACT

***OsSWEET13* sequence. PthXo2 EBE is highlighted in yellow. The first exon is highlighted in grey.**

>OsSWEET13_Komboka_ 1444 bp

TAAAAGAACAATTAAATACGAGTACCCGTATCAAATTTAGGATTTAAATATATGTATGTATCATATATATTTCGA
AATTAAGTATGCTTATATATATGCAGATTGATTGCGTATATATATTGTGCTATGCTTATAGGTTGATTGGAAAGGCA
AGCAGCCGCGCGCACAGAAAAAAGTAGAGAGAAAAAAAAAAAGCAGATCCACTAGCTTAGCTTCATATACGTGGGA
GTAGAACAAGATCAACGCGCTTCGCAGAAGCAGAAATCGACCTGTCTTCCCAACAAACTGTGTATGATAGCTTAG
TCGACAGGGATGTCTACTGCAGGTGAAAACAATCCTTCGACAAAAAATAAGTTACTTTTGGTAAAGACAGTTAAA
TAATAAGCAGCTATATCACGCGCATGGGAGAATTGCATATTCAATTACAATCATTATTTTTTTTTCAGAACACTG
TCGGCGACATTGAGAATTAATCTACCCGTGCAAACAAAGAACAGAGAAACTATAGTATACCTACATGTATCTATC
ACCCAATAATTGCAAGATCATGTTACAAAACGGTTCTAATTAATATATAGAAACAAGGCAGAGAATTCTACCTTT
CTTTTGTCTAAGTACAATTATCTTTTTCTCCGCGATTAATATTTTTCGAGTAGTAAAATTTAAGTCAAAAGCCGT
ATCAGGATTCAGGAATAATCCTTCACTGGGAGAGATCTCATGTGATTTGCTGTTGCACTCGGCGGCTATCTTTTA
CCGTTCCCAGCAGGAAGCTGCAGACGTTGGAGAGATCGATCTCTACTGACAATGCACAAAGCAATTACTCACTAA
ATTGGCTATGGCTAGTGAGAGGTGCGCTGCGCACAAAGCCAATGCAACTTTTTTTGAAAATTAGCCAGGATTATC
TCCAACAGTAGCTCATTTTTGTAAAAGCCTAATTATTGTGCGTGTCCAAAAGACTTTCCTCAAAAGCAAATAAAG
AAAAAAAATCTTTGCATAATTATTCTATGATTACTTTGATGCGTACGTGAATGGCCATGGGTAGGAGGCAACCAA
GTGATTCGCACCTAGCTAGCTTTTCTCCTATATAAGCACCACAACTCCCTTCATTCCTCTCCAAGAGTTTTCAGC
CAACACATTGAACTCTTCTTCAGAGCTCTCCCTTCCCTCCACAAAGGGGTCTAGGGTTAGAGTGTGTGTGTCTGT
GACAAGTTCCAAGCTAGCAACAACAAGCTCAATTCCTTGCTTGTTTGCTTCCATATTACACTACATCTCTTCCCT
TCAATTACCCCCCTTTTAGCACACAAAAATGGCTGGCCTGTCCCTGCAGCATCCCTGGGCTTTTGCCTTCGGCCT
CCTTGGTATATCATCATCACCTACCACAACTAAGACATTCCCTTCATTGCCAACATTTTACTTCTTTTTATTAGA
AACCATTGAGTTTGTACAT

> *OsSWEET13*_Kitaake_OsativaKitaake_499.genome _Chr12 Chr12:17058815..17060261 (- strand) class=match length=1447 bp

TAAAAGAACAATTAAATACGAGTACCTGTATCAAATTTAGGATTTAAATATATGTATGTATCATATATATTTCGA
AATTAAGTATGCTTATATATATGCAGATTGATTGCGTATATATATTGTGCTATGCTTATAGGTTGATTGGAAAGG
CAAGCAGCCGCGCGCACAGAAAAAAGTAGAGAGAAAAAAAAAAGCAGATCCACTAGCTTAGCTTCATATACGTGGG
AGTAGAACAAGATCAACGCGCTTCGCAGAAGCAGAAATCGACCTGTCTTCCCAACAAACTGTGTATGAAAGCTTA
GTCGACAGGGATGTCTACTGCAGGTGAAAACAATCCTTCGACAAAAAATAAGTTACTTTTGGTAAAGACAGTTAA
ATAATAAGCAGCTATATCACGCGCATGGGAGAATTGCATATTCAATTACAATCATTATTTTTTTTTCAGAACACT
GTCGGCGACATTGAGAATTAATCTACCCGTGCAAACAAAGAACAGAGAAACTATAGTATACCTACATGTATCTAT
CACCCAATAATTGCAAGATCATGTTACAAAACGGTTCTAATTAATATATAGAAACAAGGCAGAGAATTCTACCTT
TCTTTTGTCTAAGTACAATTATCTTTTTCTCCGCGATTAATATTTTTCGAGTAGTAAAATTTAAGTCAAAAGCCG
TATCAGGATTCAGGAATAATCCTTCACTGGGAGAGATCTCATGTGATTTGCTGTTGCACTCGGCGGCTATCTTTT
ACCGTTCCCAGCAGGAAGCTGCAGACGTTGGAGAGATCGATCTCTACTGACAATGCACAAAGCAATTACTCACTA
AATTGGCTATGGCTAGTGAGAGGTGCGCTGCGCACAAAGCCAATGCAACTTTTTTTGAAAATTAGCCAGGATTAT
CTCCAACAGTAGCTCATTTTTGTAAAAGCCTAATTATTGTGCGTGTCCAAAAGACTTTCCTCAAAAGCAAATAAA
GAAAAAAAATCTTTGCATAATTATTCTATGATTACTTTGATGCGTACGTGAATGGCCATGGGTAGGAGGCAACCA
AGTGATTCCCACCTAGCTAGCTTTGCTCCTATATAAAGCACCACAACTCCCTTCATTCCTCTCCAAGAGTTTTCA
GCCAACACATTGAACTCTTCTTCAGAGCTCTCCCTTCCCTCCACAAAGGGGGTCTAGGGTTAGAGTGTGTGTGTC
TGTGACAAGTTCCAAGCTAGCAACAACAAGCTCAATTCCTTGCTTGTTTGCTTCCATATTACACTACATCTCTTC
CCTTCAATTACCCCCCTTTTAGCACACAAAAATGGCTGGCCTGTCCCTGCAGCATCCCTGGGCTTTTGCCTTCGG
CCTCCTTGGTATATCATCATCACCTACCACAACTAAGACATTCCCTTCATTGCCAACATTTTACTTCTTTTTATT
AGAAACCATTGAGTTTGTACAT

***OsSWEET14* sequence. TalC is highlighted in yellow. AvrXa7 EBE is colored in red. PthXo3 is highlighted in light turquoise. TalF EBE is underlined. The first exon is highlighted in grey.**

>OsSWEET14_Komboka_ 888 bp

TGCGGCTCATCAGTTTCTCTAAGCTCTCACCATTCATTCCACTATACAAGCCTAAGGCAGCTAGCTTAGTTAATT
ACCTAATAACTATAGCTTGCCCAACTCTAGATCCCTTAACTAGGACAACTTGGAGTACACAACAATGTTAATAAT
CCCATGCATTGAGGACAGAGTTGTGAAGGAAACAAAAAAAAGCTAGCAGATTGGCACTTTCTGTCATGCATGGGT
GCTGATGATTATCTTGTATCTAATTTAATCAATCCCATGGCTGTGATTGATCAGGAATAGTTTGTGTGTGCAGCT
ATATTACCTATTGGTGTCCAGGGTCACACACCATAAGGGCATGCATGTCAGCAGCTGGTCATGTGTGCCTTTTCA
TTCCCTTCTTCCTTCCTAGCACTATATAAACCCCCTCCAACCAGGTGCTAAGCTCATCAAGCCTTCAAGCAAAGC
AAACTCAAGTAGTAGCTGATTACCAGCTCTTCTCTCTTCTCATTGAGAAGAGGGAATTAAGTTTTGATCTCTGCT
TTATTGCCTGATCATCCTCTTGTTACTTGCAAGCAAGAACAGTAGTGTACTGTGCCTCATTGATCTCCTCTCACC
AAACTCTCTCTCTCTCTCTCATATTCCGAGCTAGCTAGTTAATCAAGATCTTGCTGCAATGGCTGGCATGTCTCT
TCAGCATCCCTGGGCCTTCGCCTTTGGTCTCCTAGGTGTGTTGCCTTTGATCTGATCCAAGGAATTCTCTTGAGA
ATTAATCTTGCATGGTTATTTACTTTTGTTGTTATTATTCTCTACATTTTTAATCATGTACTTTTCCATGTTCCT
CTTTTGTTGCCAAAGCTACTATATTTTTCCTACCAATTCATCCAAAACTACTATATTATAGCA

> OsSWEET14_Kitaake_OsativaKitaake_499.genome Chr11 Chr11:18397214..18398101 (- strand) class=match length=888 bp

TGCGGCTCATCAGTTTCTCTAAGCTCTCACCATTCATTCCACTATACAAGCCTAAGGCAGCTAGCTTAGTTAATT
ACCTAATAACTATAGCTTGCCCAACTCTAGATCCCTTAACTAGGACAACTTGGAGTACACAACAATGTTACTAAT
CCCATGCATTGAGGACAGAGTTATGAAGGAAACAAAAAAAAAGCTAGCAGATTGGCACTTTCTGTCATGCATGGGT
GCTGATGATTATCTTGTATCTAATTTAATCAATCCCATGGCTGTGATTGATCAGGAATAGTTTGTGTGTGCAGCT
ATATTGCCTATTGGTGTCCAGGGTCACACACCATAAGGGCATGCATGTCAGCAGCTGGTCATGTGTGCCTTTTCA
TTCCCTTCTTCCTTCCTAGCACTATATAAACCCCCTCCAACCAGGTGCTAAGCTCATCAAGCCTTCAAGCAAAGC
AAACTCAAGTAGTAGCTGATTACCAGCTCTTCTCTCTTCTCATTGAGAAGAGGGAATTAAGTTTTGATCTCTGCT
TTATTGCCTGATCATCCTCTTGTTACTTGCAAGCAAGAACAGTAGTGTACTGTGCCTCATTGATCTCCTCCCACC
AAACTCTCTCTCTCTCTCATATTCCGAGCTAGCTAGTTAATCAAGATCTTGCTGCAATGGCTGGCATGTCTCT
TCAGCATCCCTGGGCCTTCGCCTTTGGTCTCCTAGGTGTGTTGCCTTTGATCTGATCCAAGGAATTCTCTTGAGA
ATTAATCTTGCATGGTTATTTACTTTTGTTGTTATTATTCTCTACATTTTTAATCATGTACTTTTCCATGTTCCA
CTTTTGTTGCCAAAGCTACTATATTTTTCCTACCAATTCATCCAAAACTACTATATTATAGCA

Supplementary Data 3. Sequences of *Xa1* and *Xa4* genes from *Oryza sativa* cv. Komboka. Capital letters represent exons, while small letters represent introns. Start and stop codons are highlighted in yellow and red, respectfully.

> *Oryza sativa* cv. Komboka *Xa1*

ATGGAGGAGGTGGAAGCCGGTTTGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCCTGGATAACCTGGAC
GCCGACAAGCTGGATGAGTGGATTCGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAG
AAGGTGGATGGGGTGGTCGCTGCCGTGAAGGGGAGGGCGATCGGGAACAGGTCGCTGGCCCGATCTCTCGGCCGT
CTCAGGGGGTTGCTGTACGACGCCGACGATGCGGTCGACGAGCTCGACTACTTCAGGCTCCAGCAGCAGGTCGAG
GGAGGAGgtactgtctttgcatatatccgtgccttttaattaagtttgcaagctgcgttgcctgcaacaatggcg
tattggcgtcagtttccaatccatgcttgtgctacagTTACTACACGGTTTGAGGCTGAAGAGACGGTCGGAGAT
GGAGCAGAGGACGAGGACGATATTCCGATGGACAATACTGATGTACCGGAGGCAGTGGCGGCAGGCAGCAGCAAG
AAACGGTCCAAGGCATGGGAACACTTTACTACCGTAGAGTTCACTGCTGACGGGAAGGATTCTAAAGCACGGTGC
AAGTACTGCCACAAGGACCTATGTTGCACATCTAAGAACGGGACATCAGCTTTGCGCAACCATCTCAATGTTTGC
AAGAGGAAACGTGTAACAAGTACTGACCAACCGGTAAATCCATCAAGgtaatgctaatggagttctgaatttagt
gtaaatccgttgaagtgtaaatttggcccgttacatctgcttaagatctcattctgtctctaatcttctaatagc
caactcatggtcatttttttttcctaatatatagtaccggtgatggtgcaccaaatgtaattagatgcaaggaaac
aaaagtgaacaattgtatatatcaaatataattatatctaaaacatgagtagtgtatcaaatccaattctttcaa
aaatctactatgcaaaattgagtgacaaaatctgctgccttttttttttacagaaagcaaccaattaatataagt
caaatataaaaacgctttgtagtctccaataaaatagctcattgtttcgtttatacttatgtttataaatttaaa
tttaaaacttaattttggagttgattttgtggttttcttttcatcctattttattttacaacatttgattttgaa
tagttaagaatgcgtatataaaaattttacccataagttatttttttaaattgttaataaatcgtaaggataatca
taagtataagtgaaacgattcgctcttcatctacttaagattgcgttatattgctgacctttctaatcgcctaac
cacgatcacatgctcttccagTGCCGGTGAGGGTGCATCAAATGCAACTGGTAATTCAGTTGGCAGAAAAAGGAT
GAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGCACCCTTGGAACAAGGCTGAACTTTCCAACAG
GATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAAATGAGGTTATGAGGCTATGTCGATCCTCAAGTTCAAA
CCAGAGTCGACAGGGTACACCACCGGCCACAAATGCAACAACATCGTCTTATCTTCCGGAGCCCATAGTGTATGG
GAGGGCTGCAGAGATGGAAACCATCAAACAGCTGATCATGAGCAATAGATCTAATGGCATAACCGTCCTGCCAAT
TGTAGGCAATGGAGGGATAGGAAAAACCACTTTGGCGCAACTGGTCTGCAAAGATCTGGTAATTAAAAGTCAGTT
TAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTACAAGGCAGATTTTGGATCATGT
CTCCAACCAGAGCCACGAAGGAATAAGCAACCTTGATACGCTTCAGCAGGATCTTGAGGAACAAATGAAATCTAA
GAAGTTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAAACTACTGGCTCCTTTAAG
ACCTAAAGATCAGGTGAATTCGTCACAGGAAGAGGCAACAGGTAATATGATAATTTTGACAACTCGTATACAGAG
TATTGCCAAAAGTCTTGGAACAGTACAATCAATTAAGTTAGAAGCTCTGAAAGATGACGATATATGGTCACTATT
TAAAGTGCATGCTTTTGGTAATGATAAACATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAGCAAATTGCTAG
CGAGCTAAAAGGCAACCCACTGGCAGCAAAAACTGTGGGTTCACTATTAGGAACGAATCTTACCATCGATCATTG
GGATAGCATTATAAAGAGTGAAGAATGGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGAAGTTGAG
CTATGATCATCTATCCAACCCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCCAAGGGTTATTCTTTCAG
CAAAGCACAACTAATACAAATATGGATTGCTCAAGGATTTGTGGAAGAATCCAGTGAGAAGTTGGAGCAGAAAGG
ATGGAAATATCTAGCTGAGTTGGTAAATTCGGGTTTCCTTCAGCAAGTTGAAAGCACACGGTTTTCATCAGAATA
TTTTGTTATGCACGATCTTATGCATGATTTAGCGCAAAAGGTTTCACAAACAGAATATGCAACTATAGATGGCTC
AGAGTGCACAGAGTTAGCCCCAAGTATACGCCATTTGTCAATAGTAACTGATTCTGCATACCGCAAGGAGAAATA
TAGAAACATATCTCGTAATGAGGTGTTTGAGAAAAGGTTGATGAAAGTTAAGTCAAGGAGTAAGTTGAGGTCACT
GGTATTAATTGGGCAATATGATTCTCATTTTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCG
ACTGCTGCAGATCACTGCAACTTATGCTGATTCTGATTCATTTCTCTCCAGTTTGGTAAATTCTACACATCTCCG
GTATCTGAAAATTGTGACCGAAGAATCCGGCAGAACTTTGCCCCGATCTCTAAGGAAGTATTACCATCTTCAAGT
ACTAGATATTGGCTATAGATTTGGAATTCCCCGTATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCT
TGTTGCATATGATGAAGTGTGTTCTTCCATTGCTAACATTGGTAAAATGACCTCACTTCAGGAACTAGGCAATTT
TATTGTTCAGAATAATTTAAGTGGTTTTGAGGTGACACAATTGAAATCCATGAACAAGCTTGTACAACTTAGTGT
GTCTCAGCTTGAAAATGTTAGAACTCAGGAGGAGGCATGTGGGGCAAAACTGAAAGACAAACAACACTTAGAAAA
GCTACATTTGTCCTGGAAGGATGCATGGAATGGATATGACAGTGACGAAAGCTATGAAGATGAATACGGCAGTGA
TATGAATATAGAAACAGAAGGGGAGGAACTGTCAGTTGGTGATGCCAATGGTGCCCAAAGCTTACAACATCACAG
TAATATAAGCTCTGAACTTGCTTCAAGTGAGGTGCTCGAAGGTCTTGAACCACATCACGGCCTCAAGTATCTACG
GATATCTGGGTATAATGGATCTACCTCCCCAACTTGGCTTCCTTCTTCACTTACCTGTCTGCAAACACTTCATCT
AGAAAAATGTGGAAAATGGCAAATACTTCCTTTAGAAAGGCTAGGGTTACTTGTAAAGCTCGTGTTGATCAAAAT
GAGGAATGCAACAGAACTCTCAATCCCTTCACTGGAGGAGCTTGTGTTAATTGCATTGCCAAGCTTGAACACATG
CTCCTGCACTTCCATCAGGAACTTGAACTCCAGTTTAAAGGTTCTGAAAATTAAGAATTGCCCTGTACTGAAGGT
ATTTCCCTTGTTTGAGATTTCCCAGAAATTTGAAATCGAGCGGACGTCGTCATGGTTGCCCCATCTTAGCAAGCT
TACCATCTATAATTATCCTCTTTCCTGTGTGCACAGTTCTCTGCCACCTTCCGCAATCAGTGGTTATGGAGAATA
TGGAAGGTGTACCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTG
CTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAAACTCAAATTTAGTGTCTCTGCAGCTCCA
TTCCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCT

TGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCA
ATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCT
CCTGAGAAAATTACATGTACTGGGAAACTCAAATTTAGTGTCTCTGCAGCTCCATTCCTGCACAGCACTCGAAGA
GTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCA
GGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACTTTACAT
CCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACT
GGGAAACTCAAATTTAGTGTCTCTGCAGCTCCATTCCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGA
GTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGG
TCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAAAC
TCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAAACTCAAATTTAGTGTC
TCTGCAGCTCCATTCCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGG
CTTGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTG
TATCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTTCAGG
GAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAAACTCAAATTTAGTGTCTCTGCAGCTCCATTCCTGCAC
AGCACTCGAAGTGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAACCT
CAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGA
GGAACTTTACATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAA
ATTACATGTACTGGGAAACTCAAATTTAGTGTCTCTGCAGCTCCATTCCTGCACAGCACTCGAAGTGTTGATAAT
TCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAG
ATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACTTTTCATCAGTGAGTA
TTCTCTAGAAACTCTGCAGCCCTGCTTCCTGACGAATCTCACCTGCTTAAAACAATTAAGGGTATCAGGCACCAC
AAGTTTCAAATCTCTAGAACTGCAATCATGCACTGCACTCGAACATTTGAAGATTCAAGGTTGTGCGTCGCTTGC
TACATTGGAGGGGTTGCAATTCCTCCACGCCCTCAGGCATATGGAAGTATTCAGATGCCCTGGCTTGCCTCCATA
TTTGGGGAGTTCGTCAGAGCAGGGCTATGAGCTATGCCCACGACTGGAAAGGCTCGACATCGATGACCCCTCTAT
CCTTACCACGTCGTTCTGCAAGCACCTCACCTCCCTCCAACGCCTAGAGCTTAACTATTGCGGAAGTGAAGTGGC
AAGACTAACGGATGAGCAAGAGAGAGCGCTTCAGCTCCTCACGTCCCTGCAAGAGCTCCGGTTTAAGTATTGCTA
CAATCTCATAGATCTTCCTGCGGGGCTCCACAGCCTTCCCTCCCTCGAGAGGTTGGAGATCCGGAGTTGCAGGAG
CATCGCGAGGCTGCCGGAGAAGGGCCTCCCACCTTCGTTCGAAGAACTGGATATCATCGCTTGCAGTAATGAGCT
AGCTCAGCAGTGCAGAACTCTAGCAAGCACTCTGAAGGTCAAAATTAATGGGGGATATGTGAACTGA

> *Oryza sativa* cv. Komboka *Xa4*

ATGGCAGCAGCCGCCGCACTTTCTGCGGTGGCGCCTGGGCAGCGGCCCGCGGACTGCCCGAGCGAGTGCGGCGGC
GTGGACATCCCCTACCCATTCGGCGTGGACAACTGCTCCTGGCCGGGTCCTGATGACTTCACCATCATCTGCAAA
TATAGCAGGCCGTACTACAGGGGCGCGGAGATCGTGAACATCTCGGTGGAGGCAGGGGAGATGCGCGTCTACTCT
CCCGTGGTCTCCCAGTGCTACAACTCATCCAACACCACGGACTCTGACGGATTCGAGTTCTTGCGGCTCAACATC
ACCAACACGCCGTTCCTGGTCGCTCCGGAGAGAAACGAGTTCACGGCCATCGGCTGTGCCACGTTGGCGTGGCTA
TGGGGCAGAGACGATGGGAGCTACTTGACCGGCTGCATCTCGACGTGCGCGAGCTTGGCGACGGCTGCCAAGGAC
CGTGACCCATGTACGGGGCTGGGTTGCTGCCAGGTGCCCTCCATCCCGGCCAACCTTAGCGTCTTAAATATTTCC
TTGGGCACAGGCATCGCCAACGTTGCCTGGGAAGAAAGTCCCTGTAGCTACGCCTTCGTGGCCGAAAAACACTGg
tacgtaactaagtacgaacacatacctactacttcatacatctatttttatttttgatagtgatatttctaaatc
agaaaaatttacttttgataggtatatttcaatccaaccatttattcttttaatgactttcttggatttaatgcg
tgaatctccattctttcacacaaaattggctacatgagcatcaagaaatgtaaatattattatatcgcttgttta
cgaggaataactaatactccactagtatgtttaaatggatgataagtagaattatttattcttggtcattgtgcc
aagataaaatatgactatcaaaggactctatctatcgtctagactcacgaatttgaatacatatatatggatgga
tacacggatctatatctctatctctactattttaagttttcgttaatactttagtacgtagtttgtatttgagtc
ggtttttaaatccgttcgcttttcaaaatacataaggagtcatacaaggaattcttttaaaaaacttgcatgcta
actgggatgagcgagaattagactcctaattgcatcagaattcctaccaagacataagaaaacaacgtatatgat
atggttattatccgtcccaacttagatatcaaatactatctcaaagtaaaattcacatttacttttagataaaa
taaatctaataataatatgattaaaatagcattcacatgttgcaacgcacaagtatttttctgtatatatagata
tatattcaaaacccaatacgttaagttcatcttataatatgaaaaagagaatatactccctccgtccctaaatat
ttgacgccgttaacttttttaaatatgtttgaccgttcgtcttattcaaaaaaatttaagtaattattcattctt
tttgattcattgttaaatgtacttttatgtatatatatagttttaaatatttcataaaagttttcaaataagatt
aacggtcaaatatatttaaaaaaatcaacggtgtcaaatatttagggacggagggagtatataagtacacaaaat
ccattagtattataattatatggatctaattattaatgataatgattattcactctatttggtgtccatactagt
ttaatttatttgatgtatatccttatatgtgacggttctttaaaagaaacagcatgtacgtgtatcttcatatc
cgttgtttttactgagaacagcacatatatatcgacatcatatttgcgttggttctcaatgaaacaggcatatga
gtgcagcataatatgcaatattttacacatgcatggtggatatttttatatgtgcatattgtgccgtcacttata
tgtgcatatttttaataacttttaaaattatatgaagttggatgaagataaatttttactatactaagttcataga
gttcaacgagatgtaacaactttgcagataattgatatatagctttttcacttgggattagttatatgtccgaat
atactttagaaactcttggaccaatttggatatataaatagacttttaaaattttaaacgacttcgtacgcgcta
tacaaaagttatacgagaccatcggatggcttattggcttattggggaaataagccaaacggcatatttacaaac
aaaaaataatttgtgaataaaacttttatatacatgttcttaataatctaaaaacaaaggctgaaaaataaactt

caatgaaaaaacctcaaaatcaactccaaatttaaagttgaaaatttaaattttggctgataagtataagcataa
gcgaaaagatgaggccgataaagcttgacgagatctaaaattttgtagtttataacctttcatttgaactcatt
tggggtgcaacatacgcattatgaatttttgacgggagtagtaacaagatacatttgttttctaacgacaaatgc
tcgatagttgagatgataagaaagctgcacatgcagataaaaggtctcggatttgaattccacatttcccatgtg
cacgtatttcgcatgaaaaatctcgtgacttgtgacttgcaaaattagaaatactactctgggttttaaatttgc
acctaggaaattcaaaatgtagcaccactgtatttgccggttttactttagaatcgaaatataatatgcctaaca
ccaaatggtttaagaattaagtgggttaacaaaataaatggcacatggatgactttcatgtgtgatgattcttaa
taaaaattagcacatatataatttttcatttgttacagttcttaataaaaacctgaacatatgaaggtatatatgt
gccggttcttttaactttggctctatatgcatcttggggtgaatgagagattaataccttatttgtgttagtttt
tggataatcagtatatatagtattgtcacttatttaatgtttttcagccgtgaactgattgtgcacctgttgtcc
gtatgtgaatcagGTATAATTTCAACAGGCAAGATTTCGAACGTGACGGTAAAAGCTTTGAACACCGGGATGAGA
AAATGGTAGTTCCGACGGTGCTTGACTGGGCCATCAGGAAAAATGGGTCGTGCCCGTCAACCGGACAGGGTGCTC
CTGCCTGTAAGAGCGAACATAGTGAGTGCGTCAACGCCACCAACGGCAAGGGCTACCTCTGCAATTGCTCCAGGG
GATACGCCGGCAATCCCTATAGGGACGATGGATGCAAAAgttagttacgttcttaattagctcgctagctatcac
ctaagcatagaattcttctctaattttatttttggatgctagattagtatttacagtacgtggtgatgattaatga
cgagcatgcttgcttagtgtttaaatttacttatttatgacttaattgttctaactctttaatttgattgtttat
caccaatacgtctgtagATATTAATGAGTGTAAGGAACCTTCTATTACTTGTTACGGTGGTAGCACATGCCAGGA
CACAGACGGCAGTTACGAGTGCAAATGCCAATTTGGATATAGAGGCGATGGCAGGAAGAATGAGAGTCAGAAGGG
ACGATGCCAACCCATAATTCCCGCTGCTATAGCCAATGCAATAGgtaaaccgatagagtgttaattcttaactac
tccaacacgtagcaatataaatttaatttgctacctattagcaaatgtggatcaacactgtatatctttatctta
tgtttctttgtatgcagCGATTGTTTGCATTGTGATTGTTCTCTTGGGCCTTTTCTGGTTGCCTAAGAGATGGAA
GCGGAGAGTGTTCTTCGATAATAACGGTGGTCGCCTACTGAAGGATATGGACATCATCGTTTTCACTGAGAAGGA
ACTCAACAAGATAACAAACAAGAAGCGCACTAAGATTGGAGAGGGCGCCTTTGGCGAGGTCTACAAGGGGAACCA
CAATAACCAACCGGTTGCTGTCAAGTACTCCATCGCAAAGAACATGACTCAAACACATTACAAAGACGTCGTCGA
GAGCATAAACCAAAACGTATTCCAGACTGTTTTTCGTCAGTCAAAAGTTCCACCATCAACACCGGGCCAGAATGC
AGTCGTCAACGAGATAAAGGTTCAGCTGCAGATCCGGCACCCCAACATTGTCCGCCTCATCGGGTGCTGTATGGA
AACAGAGGTCCCCATGCTTGTCTTTGAGTTCATCCCTAACGGGAGCCTCGAGACAGTTCTCCATGGTATCGATCG
ATGCAGCCTTTCTCTCCAGCAACGCCTGGACATCGCGATCGGCTCTGCAGAGGCTCTTGCCTACATGCATTGGCA
CGGCCATCATCAGATCATACATGGGGATATCAAGCCTGGCAACATTCTCCTTGGTGACAATCTCATGCCTAAGGT
CTCCGACTTTGGATCATCTGAGCTCACGTTGAAAGTCAAGCGTGCAGGGAAGTGGAACGTATATGCTGACATGAA
CTACATCGACCCTGTGTACATCAAGACAGGCGATTTCACGGATAAGAGCGATGTATATAGTTTTGGGGTTGTGCT
CCTAGAGCTCATCACCAGGAAGAAGGCCAAGTATGACGATAGAAGCCTCCCGGTAGAATTTGTCAGCCATTATGA
GGATGAAGACACAAGGAGGAAAATGTATGACCAGGACATGTTGCCTACCGAGGCCTCGCATCCTCACTGCATGGA
GTGTCTTGACAGAATGGCCGATATTGTGCTCCGTTGTCTTGAAAATGAAGTGGGCAAGAGGCCAACCATGGCTGA
GGTGCTAGAGGAGCTTAAGAAGTTGTTACCATTGCTAACCACGACGCCGGTCGAACTCGTGTAG

References

1. Hu, K. *et al.* Improvement of multiple agronomic traits by a disease resistance gene via cell wall reinforcement. *Nat. Plants* **3**, 1–9 (2017).
2. Ji, C. *et al.* Xa1 allelic R genes activate rice blight resistance suppressed by interfering TAL effectors. *Plant Commun.* **1**, 100087 (2020).
3. Eom, J.-S. *et al.* Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* **37**, 1372–1379 (2019).
4. Tran, T. T. *et al.* Functional analysis of African *Xanthomonas oryzae pv. oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **14**, e1007092 (2018).
5. Blanvillain-Baufumé, S. *et al.* 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 (2017).

ANNEXE 4 : Article en co-auteure “ The Rice ILI2 Locus Is a Bidirectional Target of the African *Xanthomonas oryzae* pv. *oryzae* Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility”

International Journal of
Molecular Sciences

MDPI

Article

The Rice *ILI2* Locus Is a Bidirectional Target of the African *Xanthomonas oryzae* pv. *oryzae* Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility

Hinda Doucouré [1,†], Florence Auguy [2,†], Servane Blanvillain-Baufumé [2], Sandrine Fabre [2], Marc Gabriel [2], Emilie Thomas [2], Fleur Dambreville [2], Coline Sciallano [2], Boris Szurek [2], Ousmane Koita [1], Valérie Verdier [2] and Sébastien Cunnac [2,*]

[1] LBMA, Faculté des Sciences et Techniques, University des Sciences Techniques et Technologiques, Bamako E 3206, Mali; doucourehinda@yahoo.fr (H.D.); okoita@icermali.org (O.K.)
[2] PHIM Plant Health Institute, University Montpellier, IRD, CIRAD, INRAE, Institut Agro, 34398 Montpellier, France; florence.auguy@ird.fr (F.A.); servane.baufume@cirad.fr (S.B.-B.); sandrine.fabre@cirad.fr (S.F.); marc.gabriel@curie.fr (M.G.); emilie.thomas@ird.fr (E.T.); fleur.dambreville@gmail.com (F.D.); coline.sciallano@ird.fr (C.S.); boris.szurek@ird.fr (B.S.); valerie.verdier@ird.fr (V.V.)
* Correspondence: sebastien.cunnac@ird.fr
† These authors contributed equally to this work.

check for updates

Citation: Doucouré, H.; Auguy, F.; Blanvillain-Baufumé, S.; Fabre, S.; Gabriel, M.; Thomas, E.; Dambreville, F.; Sciallano, C.; Szurek, B.; Koita, O.; et al. The Rice *ILI2* Locus Is a Bidirectional Target of the African *Xanthomonas oryzae* pv. *oryzae* Major Transcription Activator-like Effector TalC but Does Not Contribute to Disease Susceptibility. *Int. J. Mol. Sci.* **2022**, *23*, 5559. https://doi.org/10.3390/ijms23105559

Academic Editors: Yongmei Bao and Shimin Zuo

Received: 18 March 2022
Accepted: 30 April 2022
Published: 16 May 2022

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

Copyright: © 2022 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: *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strains that cause bacterial leaf blight (BLB) limit rice (*Oryza sativa*) production and require breeding more resistant varieties. Transcription activator-like effectors (TALEs) activate transcription to promote leaf colonization by binding to specific plant host DNA sequences termed effector binding elements (EBEs). *Xoo* major TALEs universally target susceptibility genes of the SWEET transporter family. TALE-unresponsive alleles of clade III *OsSWEET* susceptibility gene promoter created with genome editing confer broad resistance on Asian *Xoo* strains. African *Xoo* strains rely primarily on the major TALE TalC, which targets *OsSWEET14*. Although the virulence of a *talC* mutant strain is severely impaired, abrogating *OsSWEET14* induction with genome editing does not confer equivalent resistance on African *Xoo*. To address this contradiction, we postulated the existence of a TalC target susceptibility gene redundant with *OsSWEET14*. Bioinformatics analysis identified a rice locus named ATAC composed of the *INCREASED LEAF INCLINATION 2* (*ILI2*) gene and a putative lncRNA that are shown to be bidirectionally upregulated in a TalC-dependent fashion. Gain-of-function approaches with designer TALEs inducing ATAC sequences did not complement the virulence of a *Xoo* strain defective for *SWEET* gene activation. While editing the TalC EBE at the ATAC loci compromised TalC-mediated induction, multiplex edited lines with mutations at the *OsSWEET14* and ATAC loci remained essentially susceptible to African *Xoo* strains. Overall, this work indicates that ATAC is a probable TalC off-target locus but nonetheless documents the first example of divergent transcription activation by a native TALE during infection.

Keywords: rice; *Xanthomonas oryzae*; TAL effector; disease

1. Introduction

Bacterial leaf blight (BLB) is a widespread rice disease caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) [1]. The significant economic impact of this disease on rice cultivation [2] has prompted breeding commercial varieties for enhanced resistance traits by conventional introgression of resistance loci from the rice germplasm [3,4]. Recently, pioneering work [5] and the development of straightforward genome editing tools have, however, sparked a wave of reports on the engineering of BLB resistance in elite varieties [6–11]. In this strategy, the modified rice genomic sequences correspond to effector binding elements

(EBEs), the DNA sequence recognized by a translocated bacterial virulence protein of the transcription activator-like effectors (TALEs) family. Editing a relevant EBE thus renders the altered genotype refractory to TALE-mediated BLB susceptibility gene induction and confers disease resistance.

Upon injection into the plant host cell via the molecular syringe of the bacterial type III secretion system (T3SS), TALEs enter the nucleus and promote the transcription of their target locus. TALEs were initially thought to induce transcription in a single direction by activating the synthesis of a protein-coding mRNA molecule on the same strand and 40–60 bp downstream of their target EBE [12–14]. However, this does not seem to be universal because the transcription of sequences on the opposite strand and upstream of the EBE was recently reported to occur in a couple of specific TALE–EBE combinations [15,16]. EBE sequence recognition is mediated by the central domain of the effector that wraps around the major groove of the genomic DNA and makes specific contacts with bases of this sequence. These interactions involve repeat variable diresidues (RVDs) corresponding to positions 12 and 13 of otherwise-conserved ~34 amino acid repeats that make up the central domain. In a colinear manner, amino acid 13 from successive repeats along the TALE central domain sequence recognizes a cognate nucleobase along the EBE strand. Subsequent to its discovery [17,18], the RVD–nucleotide association code has been relatively well characterized and prominent RVDs, such as NI, NG, HD, and NN, were shown to recognize A, T, C, and G/A, respectively [19]. As a general rule and with the exception of RipTALs [20], EBEs start with a T that is not recognized by a canonical repeat but rather by a structurally related helical bundle or a cryptic repeat located before the central domain [21–23]. Because this feature contributes to DNA recognition, it is ordinarily referred to as "repeat 0" and the corresponding location of the recognition site as "position 0."

With a few notable exceptions, including the transcription elongation factor *OsTFIIAγ1* and the b-ZIP transcription factor *OsTFX1*, which were shown to be induced by TALE effectors from Philippine *Xoo* strains [24] or the African-*Xoo*-specific target *OsERF#123* [25], in rice, susceptibility genes exploited by *Xoo* major TALEs, which were thus edited for resistance engineering, belong to clade III of the *OsSWEET* gene family [26]. How their rice membranous sugar transporter products favor parasitism is still an unsettled issue, but they may support bacterial metabolism by releasing sucrose into the apoplast [27]. Virtually all pathogenic *Xoo* strains harbor at least one TALE inducing one of three *OsSWEET* clade III homologs. For instance, PthXo1 from the Philippine strain PXO99A [28] targets *OsSWEET11*, while PthXo2 variants from Asian Xoo target *OsSWEET13* [8,29]. Probably as a result of convergent evolution, overlapping or distinct EBEs have been identified on the *OsSWEET14* promoter. The cognate TALEs AvrXa7 and PthXo3 originate from Asian Xoo strains [30]. Others, namely TalC and TalF, have been described in African *Xoo* strains [26,31].

Both functional characterization and long read sequencing of the African *Xoo* TALE repertoires revealed that TalC is highly conserved and found in all analyzed strains. A large share of African strains also possesses the redundant TalF TALE [8]. Some strains, such as BAI3 [25] and MAI68 [32], contain a TalF variant that is inactive on the *OsSWEET14* promoter and, therefore, rely solely on TalC for *OsSWEET14* induction.

The genetic divide observed in phylogenetic studies between *Xoo* strains of the Asian and African lineages [33,34] also appears in the prospects of engineering BLB resistance with modified, TALE-unresponsive, *OsSWEET* promoter alleles. While this strategy possesses an immense potential for the broad control of Asian *Xoo* strains, as previously demonstrated [8,9], the situation appears less straightforward regarding African *Xoo* bacteria. This is because the dependency of African *Xoo* strains on the SWEET susceptibility component for host colonization differs from that of the Asian strains analyzed so far. Initial clade III *OsSWEET* promoter editing approaches to engineer resistance to African strains were unsuccessful [35]: While *OsSWEET14* induction upon African *Xoo* BAI3 inoculations was abrogated in TalC-EBE-edited rice lines, those lines, in two different Japonica rice

backgrounds, remained essentially susceptible to this Burkinabe strain. A subsequent study obtained equivalent results when challenging Kitaake lines with the Malian MAI68 strain [32]. More recent findings align with this conclusion and provide additional insight. For instance, a single *OsSWEET14* knockout event in Kitaake background exhibits moderately reduced susceptibility (about half the lesion length of wild-type plants) to AXO1947, BAI3, or MAI1 African strains [7]. In contrast, an *OsSWEET14* knockout mutant in rice cultivar Zhonghua 11 appears to completely abolish lesions caused by AXO1947 [11], suggesting that other polymorphic loci may epistatically influence the disease outcome depending on the rice genotype. The susceptibility of quintuple-EBE-edited lines from Oliva et al. [8] is again modestly decreased when challenged with the AXO1947 strain. However, a double *OsSWEET14*/*OsSWEET13* knockout line in Kitaake is completely resistant to the AXO1947 African strain [7], which represents evidence that those clade III genes are redundant. The expression of other *OsSWEET* genes from clade III is not modified upon BAI3 inoculation in either wild-type or edited lines, suggesting that the loss of *OsSWEET14* induction is not compensated by transcriptional upregulation of redundant homologous genes. Thus, in contrast to Asian *Xoo*, African *Xoo* has the capacity to infect rice in the absence of clade III *SWEET* susceptibility gene induction [35]. These experimental data were unexpected because they are inconsistent with the dramatic phenotype of BAI3-1-1, a BAI3 derivative with a defective *talC* gene that has been previously shown to be unable to cause significant BLB symptoms and efficiently colonize susceptible rice leaves, leading to the conclusion that this effector is a major virulence TALE [26,31]. To reconcile these observations, we previously postulated the existence of an unidentified susceptibility gene targeted by TalC and genetically redundant with *OsSWEET14* [35].

This study was primarily conducted to explore the possible existence of such a redundant TalC target. Revisiting available rice transcriptomics data with an original strategy helped identify a particularly interesting candidate TalC target locus. In addition, we provide evidence that *talC* triggers transcription of two divergent transcripts through the recognition of an EBE that is shown to be necessary for induction at this target locus. However, both gain- and loss-of-function approaches could not establish that this rice locus contributes to BLB susceptibility. Overall, this work identifies a second TalC target and provides the first unambiguous results indicating that a native TALE can trigger bidirectional transcription during infection.

2. Results

2.1. Reassessing TalC Target Predictions in the Context of a BAI3-Infected Rice Experimental Transcriptome Yields New Candidate TalC Target Loci

Prior searches for TalC targets in rice identified *OsSWEET14* [31]. A subsequent Nipponbare-genome-wide inquiry of African TALE targets combined both in silico target predictions on upstream sequences of annotated genes and lists of differentially expressed genes upon African Xoo infection based on transcriptomic data [25]. In addition to *OsSWEET14*, the latter approach identified LOC_Os03g58790, a putative ATPase, as a candidate TalC target. However, the induction ratio and prediction rank values are moderately supportive for a TALE target (~4× and 264, respectively). Furthermore, the predicted TalC EBE for this locus starts with a T for position 0 of the RVD array, which contrasts with the functional TalC EBE on the *OsSWEET14* promoter, which uncommonly starts with a C, as previously pointed out [31]. Considering both that TALEs may trigger the synthesis of transcripts overlooked in reference annotations and that previous TalC target searches did not incorporate the possibility of "antisense EBEs," in order to maximize the likelihood of identifying candidate redundant susceptibility genes, we decided to reassess TalC target predictions.

For that, we analyzed a recently documented mRNA-seq dataset [36]. First, an experimental transcriptome (see our dataverse) was inferred with the StringTie software by combining mRNA-seq read alignments to the Nipponbare genome and the reference MSU7 annotation [37]. This experimental transcriptome served to compute read counts

at the gene level, which permitted the assessment of differential expression (annotation and expression data are available on our dataverse). From this, we listed 261 loci deemed as differentially induced in comparisons between African-*Xoo*-strain-BAI3-infected leaves and a BAI3 T3SS defective mutant strain (BAI3H) or mock inoculated Nipponbare leaves. To predict potential TalC EBEs, we ran Talvez [38] using the sequences located 500 bp upstream and 200 bp downstream of the start of the experimental BAI3 T3SS upregulated loci. The resulting list of candidate loci with the 10 highest Talvez prediction scores is shown in Table 1. When available, TalC-dependent induction data (derived from 31) were included in this table. Annotated genome browser views of those candidate TalC target loci with read coverage profiles and predicted EBEs are provided in Figure S1.

As expected, the *OsSWEET14* locus (LOC_Os11g31190, also previously designated as Os11N3) carries the top-scoring predicted EBE and exhibits elevated T3SS- and *talC*-dependent induction ratios (Table 1). Apart from this positive control locus, the *INCREASED LEAF INCLINATION 2* (*ILI2*) locus (LOC_Os11g39000) is the only MSU7 reference annotation locus with a similar profile, being notably induced in a *talC*-dependent manner (adjusted p-value = 6.29×10^{-3}; Table 1). In contrast, however, the predicted TalC EBE is positioned in an antisense orientation relative to the ILI2 coding strand.

Interestingly, our analytical approach also uncovered two novel loci, PIX.7679 and PIX.6359, without MSU7 annotation, which appeared to generate mRNA-seq reads only if the bacteria were able to deliver type III effectors (Figures S1 and 1a,b). They were, therefore, noteworthy TalC target candidates even if they could not be assessed immediately with regard to talC-dependent induction because their expression is not measured in microarray data.

2.2. ATAC, A Candidate Alternative TalC-Target Locus Is Induced Bidirectionally and Comprises a Long ncRNA and the ILI2 Gene

Upon further scrutiny of the genomic positions of the predicted EBEs in Table 1, we unexpectedly realized that the *ILI2* and PIX.6359 loci actually share the same candidate EBE and are arranged in opposite, head-to-head, orientations on Chromosome 11, with the TalC EBE positioned on the same strand as PIX.6359 (Figure 1a). The interval between the PIX.6359 or *ILI2* start site and the proximal end of this EBE is, respectively, 52 bp or 115 bp. The distance from the PIX.6359 transcript start is consistent with distances usually observed between a TALE EBE and its target transcription start (40–60 bp). This configuration is reminiscent of the bidirectional transcription activation by TALEs previously observed in transient plant expression assays [16]. Because of this rather unique feature and because the observed expression profile of the ILI2 and PIX.6359 transcript was at this stage consistent with their being bidirectional targets of TalC (Figure 1b), we tentatively refer to this locus as ATAC (Alternative TalC Target locus), with PIX.6359 designated as *ATAC1* and *ILI2* as *ATAC2*.

In order to examine if the sequence of the predicted EBE in this region (EBE_{TalC}@ATAC) is likely to be targeted by TalC, we aligned this sequence with EBE_{TalC} on the *OsSWEET14* promoter (EBE_{TalC}@OsSWEET14) and with an "optimal" EBE_{TalC} sequence based on the Talvez nucleotide–RVD association matrix (Figure 1c). The nine first 5′ nucleotides of EBE_{TalC}@ATAC and EBE_{TalC}@OsSWEET14 are strictly conserved. This includes the cytosine at position 0, an atypical feature of EBE_{TalC}@OsSWEET14. EBE_{TalC}@SWEET14 displays only one position conflicting with the RVD–nucleotide code. EBE_{TalC}@ATAC contains two such positions but possibly with compensatory better matches at other positions.

The StringTie-inferred ATAC1 transcript is 456 nt long and spans positions 23,222,261 to 23,222,716 on Chromosome 11. It is predicted to be composed of a single exon. The longest predicted ORF is 46aa long (Figure S2), and preliminary searches with blastx on NCBI for proteins similar to putative ATAC1 translation products yielded no meaningful hit outside of *O. sativa* hypothetical proteins. Because this suggested that the ATAC1 transcript does not code for a protein, we used PLncPRO [39] to examine if the ATAC1 transcript is a potential long noncoding RNA (lncRNA). Predicting the ATAC1 transcript status with PlncPRO returned a noncoding probability of 0.965, a value well above 0.8, which is considered the lower bound for high-confidence lncRNAs [39].

Table 1. Candidate TalC targets in the Nipponbare genome.

EBE Position on IRGSP-1.0				EBE on Subject [−500, +200]			Talvez Score	Candidate Target Gene Annotation				mRNA-Seq Diff. Exp. BAI3 vs. BAI3H			Micro-Array Diff. Exp BAI3 vs. BAI3-1-1 *	
Seq name	start	end	strand	start	end	pol ‡		id	ref_gene_id	name_Gramene	description_Gramene	log2FC	pvalue	padj	log2FC	padj
Chr11	18,174,547	18,174,569	-	410	432	S	14.45	PIX.6085	LOC_Os11g31190	11N3	Sugar transporter, TAL effector-mediated susceptibility to bacterial pathogen (Os11t0508600-01)	6.869	2.55×10^{-45}	8.05×10^{-42}	5.97	1.03×10^{-1}
Chr12	17,463,497	17,463,519	+	577	599	S	13.86	PIX.7679	NA §	NA	NA	5.446	1.13×10^{-12}	4.82×10^{-10}	NA	NA
Chr5	29,688,055	29,688,077	-	106	128	AS	12.36	PIX.19089	LOC_Os05g51750	NA	Peptidase A1 domain containing protein (Os05t0596000-01)	7.077	5.47×10^{-37}	9.09×10^{-34}	NA	NA
Chr11	23,222,769	23,222,791	-	426	448	S	11.27	PIX.6359	NA	NA	NA	4.623	2.81×10^{-3}	2.19×10^{-1}	NA	NA
Chr11	23,222,769	23,222,791	-	363	385	AS	11.27	PIX.6360	LOC_Os11g39000	INCREASED LEAF INCLINATION 2	Basic helix–loop–helix dimerization region bHLH domain containing protein (Os11t0603000-01)	6.998	1.09×10^{-13}	5.08×10^{-11}	4.26	6.29×10^{-3}
Chr7	21,798,885	21,798,907	-	496	518	S	10.83	PIX.22524	LOC_Os07g36430	NA	Conserved hypothetical protein (Os07t0549100-01)	4.66	3.82×10^{-3}	2.83×10^{-1}	NA	NA
Chr11	28,760,402	28,760,424	+	84	106	AS	10.74	PIX.6663	LOC_Os11g47600	NA	Similar to Class III chitinase homologue (OsChib3H-h) (Fragment) (Os11t0702100-01)	2.432	7.94×10^{-7}	1.77×10^{-4}	NA	NA
Chr10	18,168,970	18,168,992	+	369	391	S	10.46	LOC_Os10g34040	LOC_Os10g34040	FLOT1	Similar to Flotillin-like protein 1 (Os10t0481500-00)	4.599	1.68×10^{-3}	1.39×10^{-1}	NA	NA
Chr4	25,888,105	25,888,127	+	659	681	AS	10.39	PIX.16030	LOC_Os04g43730	WALL-ASSOCIATED KINASE GENE 51	Similar to OSIGBa0145M07.8 protein (Os04t0517700-01)	7.301	2.80×10^{-3}	2.19×10^{-1}	NA	NA
Chr4	15,500,717	15,500,739	-	360	382	S	10.22	PIX.15197	LOC_Os04g26550	NA	Conserved hypothetical protein (Os04t0332700-01)	3.877	5.94×10^{-13}	2.61×10^{-10}	NA	NA

* Dataset GSE19844; § Not Available; ‡ Polarity; S, sense; AS, antisense.

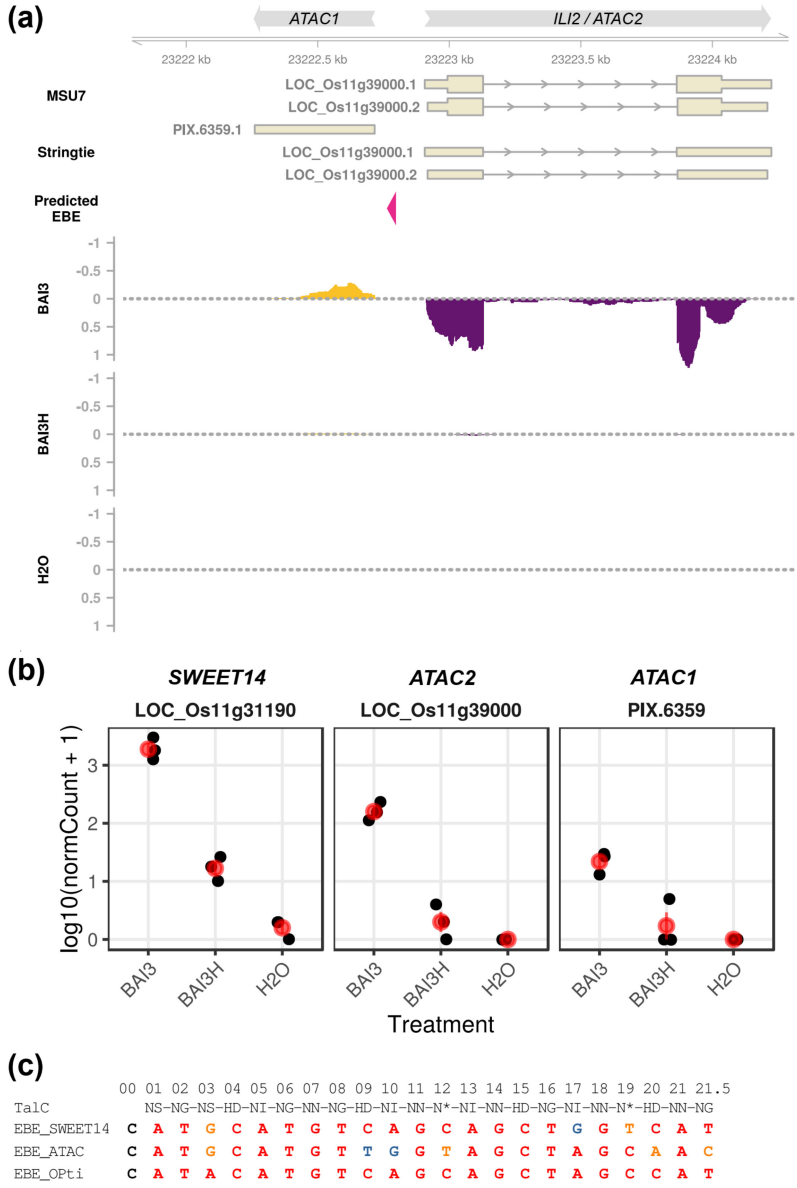


Figure 1. Transcripts at a second candidate target locus are potentially induced bidirectionally by TalC. (**a**) Genome browser snapshot of the ATAC locus with transcript annotation from MSU7, the experimental transcriptome annotation inferred by StringTie, the predicted TalC EBE and read coverage tracks (reads per million) for various experimental treatments in our mRNA-seq data. For the MSU7 annotation, shorter exon boxes correspond to untranslated regions. For coverage tracks, only the reads corresponding to the transcribed strand in paired-end data for a specific treatment were considered. Sequencing libraries were generated with RNA extracted 24 h after infiltration of Nipponbare leaves with the strains BAI3, BAI3H (T3SS-), and H_2O. Coverage of reads mapping to the top genomic strand is represented with positive values and a golden-colored area, whereas coverage of reads mapping to the opposite strand is represented with negative values and a violet-colored area. (**b**) Expression of TalC target genes in our mRNA-seq data. Black dots correspond to individual replicate normalized count values, while the red dots and line ranges denote the mean and the standard error, respectively. (**c**) Alignment of the TalC RVD sequence and putative DNA target sequences. The numbers on top refer to the classical numbering of repeats in the TALE central repeat domain for TalC. The EBE Opti sequence was generated by concatenating the most probable nucleotide for each RVD in the RVD–nucleotide association matrix used by Talvez. In the alignment, nucleotides are colored according to their match to the corresponding TalC RVD in this matrix: red, best score; orange, intermediate; and blue, worst score.

The *ATAC2/ILI2* locus codes for a putative protein belonging to an atypical group of the bHLH transcription factor family that lacks the basic region of the domain and does not bind DNA. In rice, the ILI clade contains seven homologous genes, including *ILI1*, its founding member. These genes are phylogenetically and functionally related to members of the Arabidopsis PRE family [40]. Characterized proteins of this group positively regulate cell elongation in response to environmental and hormonal cues, but this signaling module has also been implicated in the regulation of Arabidopsis PAMP-triggered immunity (PTI) [41].

Under the LOC_Os11g39000 locus identifier, *ATAC2/ILI2* has previously been shown by Yu et al. [31] to be *talC*-regulated based both on microarray and independent RT-PCR data. To ascertain that ATAC1 lncRNA induction is both T3SS and *talC* dependent, we profiled gene expression in response to wild-type BAI3, its *talC*-mutant derivative BAI3-1-1 [31], and complemented strains in leaves of wild-type Kitaake or the EBE_{TalC}@OsSWEET14-edited line *sweet14–32* [35]. The qRT-PCR data in Figure 2 indicate that the *OsSWEET14* control gene is induced on wild-type plants but not on the edited line, in response to the BAI3 strain and the BAI3-1-1 mutant complemented with *talC*. ATAC transcripts expression follow a similar pattern except that they also appear to be *talC* induced on the EBE_{TalC}@OsSWEET14-edited line. This demonstrates that both *ATAC1* and *ATAC2* are *talC* regulated and that this effect occurs independently of *OsSWEET14* induction.

Altogether, these data support the hypothesis that TalC recognizes a second sequence in the rice genome at the ATAC locus and triggers bidirectional transcription of two opposite units.

2.3. Artificial TALEs Targeting the ATAC Locus Do Not Enhance the Virulence of a Xoo Strain Unable to Induce an OsSWEET Gene

To explain why, in contrast to Asian *Xoo*, African *Xoo* virulence is marginally affected on TALE-unresponsive *OsSWEET14* promoter-edited lines, we postulated the existence of another TalC susceptibility gene target, redundant with *OsSWEET14*. Under this model, induction of the redundant factor is expected to compensate for the loss of the susceptibility functions provided by *OsSWEET14* induction. If ATAC acted as such a redundant TalC target, we reasoned that the induction of the ATAC locus should rescue the virulence of an Asian strain that is unable to induce OsSWEET functions.

To address this possibility, we assembled a set of four "designer" or "artificial" TALEs (ArTALEs) as before [26]. Their respective EBEs were selected to maximize specificity toward the ATAC loci and as shown in Figure 3a and Figure S3, are located near EBE_{TalC}@ATAC on the same strand (ArTALE05 and 06) or the opposite, *ATAC2*, strand (ArTALE01 and 04). These constructs, together with the empty vector or plasmids expressing the control $TalF_{MAI1}$ or $TalC_{BAI3}$ native TALEs, were introduced into the ME2 strain, a non-virulent PXO99A mutant derivative deprived of the PthXo1 TALE [42] that is routinely employed for testing virulence complementation by *OsSWEET* targeting TALEs [8,16]. The resulting strains are described in Table 2. To functionally validate these ArTALEs, the transformed strains were infiltrated into Kitaake leaves and their ability to induce gene expression was monitored with qRT-PCR (Figure 3b). The positive control strain expressing TalC induced *OsSWEET14*, *ATAC1*, and *ATAC2* relative to the empty vector strain, whereas the ME2 strain expressing TalF induced *OsSWEET14* only. None of the ArTALEs were capable of dramatically affecting *OsSWEET14* expression relative to water-infiltrated plants. In contrast, all four of them induced *ATAC2* above empty control levels. The situation regarding *ATAC1* is more contrasted. ArTALE05 and 06 induced this transcript at the same level as TalC, whereas ArTALE01 and, to a lesser extent, ArTALE04 were similar to the empty control strain. Overall, it appeared that all our ArTALEs are able to specifically activate ATAC transcription and that those designed to target an EBE with the same orientation as EBE_{TalC}@ATAC, similar to TalC, acted in a bidirectional manner.

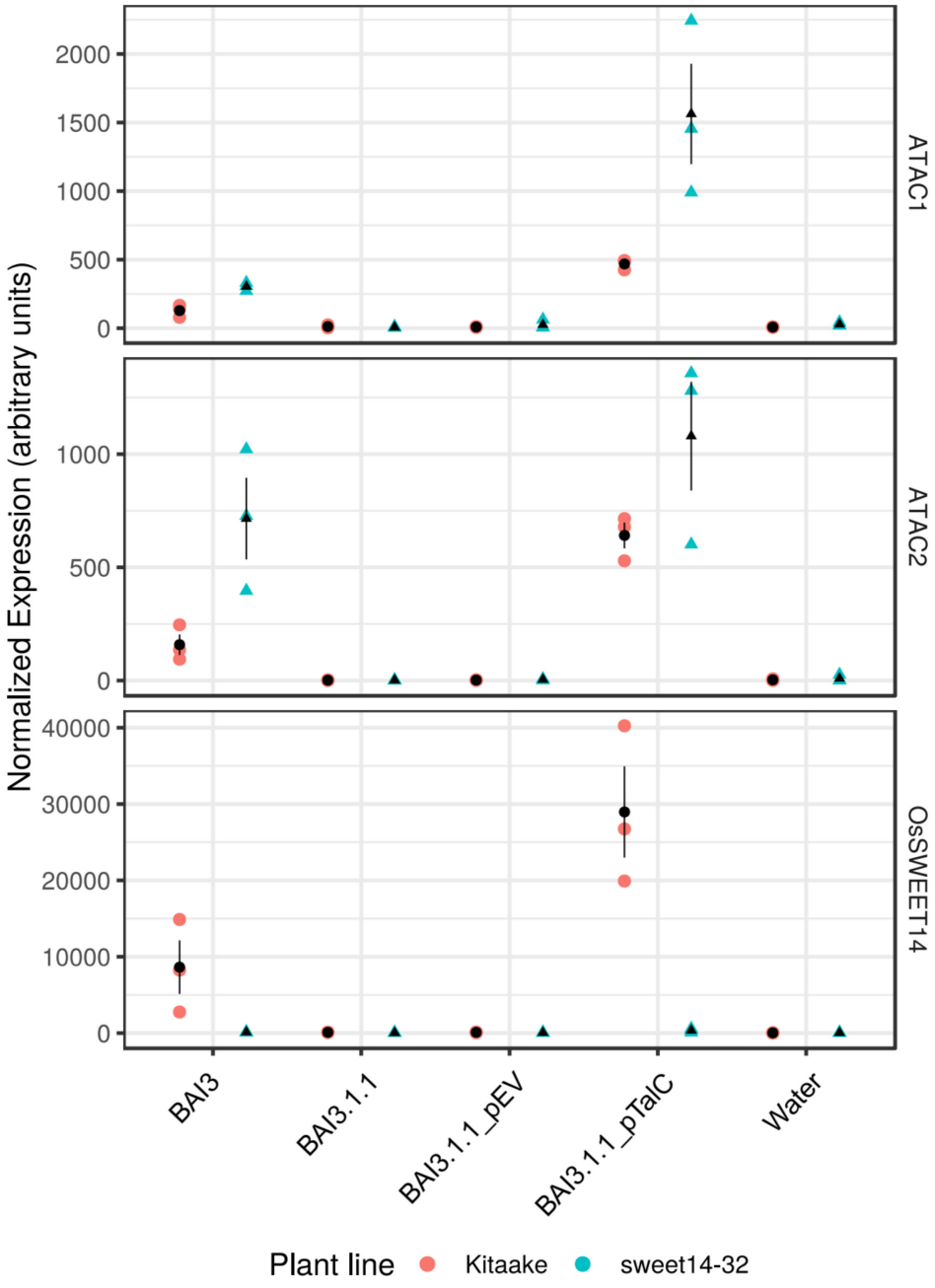


Figure 2. ATAC transcript induction in response to BAI3 is *talC* dependent. Transcript expression of ATAC1, ATAC2, and SWEET14 on the wild-type Kitaake and the edited sweet14–32 lines measured 48 h after inoculation with derivatives of the BAI3 strain or water. The strains are represented on the *x*-axis: BAI3.1.1 is a BAI3 derivative defective for *talC*. "_TalC" or "_EV" suffixes indicate, respectively, that the *talC* strain carried a plasmid expressing TalC or the corresponding empty vector. Expression levels of the different transcripts measured by qRT-PCR are expressed in arbitrary units and calculated by the ΔCt method. Individual biological replicate values ($n = 3$) are denoted with colored symbols as a function of the inoculated plant line. The black dots and line ranges denote the mean and the standard error, respectively.

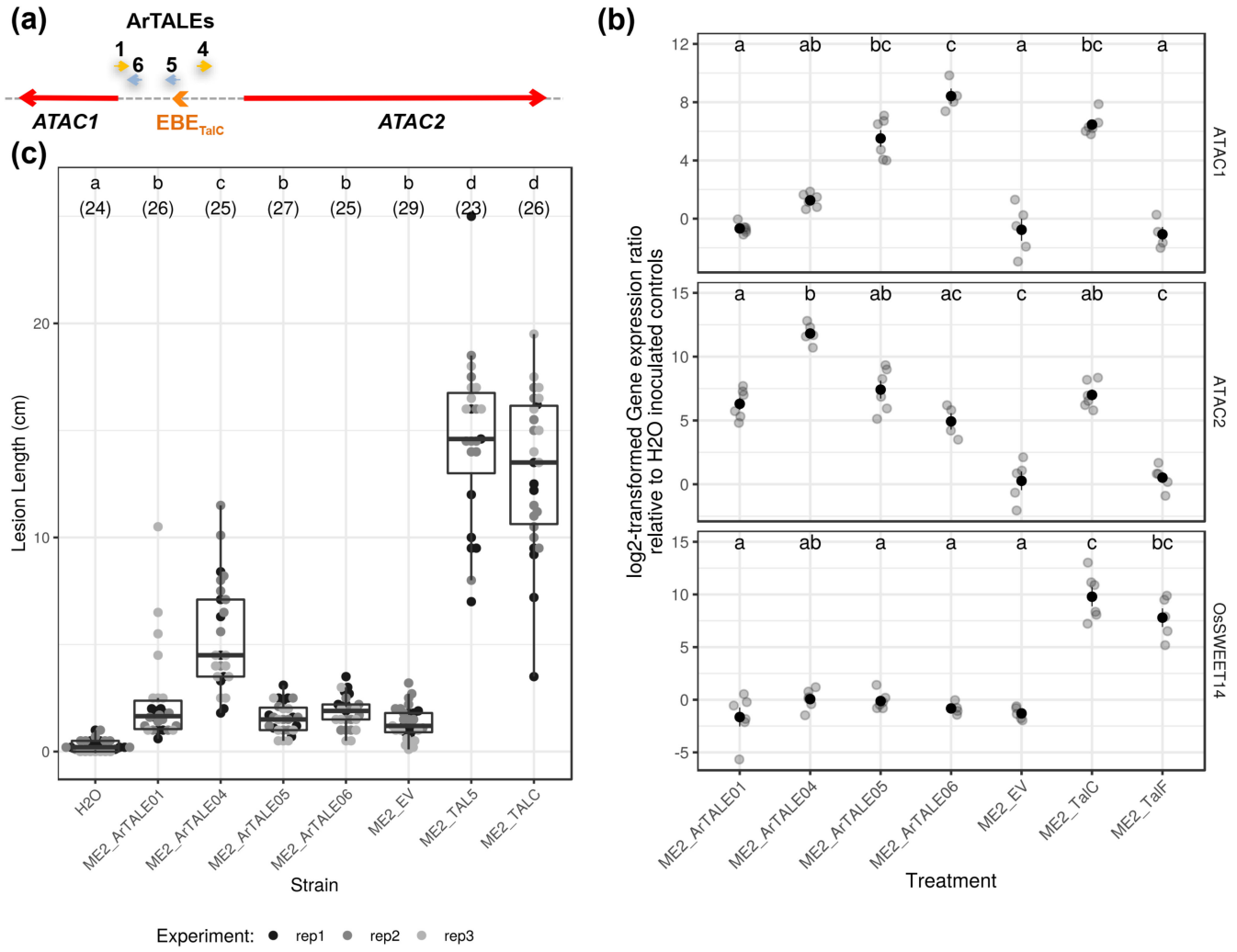


Figure 3. Independent ArTALEs targeting the ATAC loci are functional but do not compensate the loss of clade III OsSWEET induction for the virulence of the Asian PXO99 strain *pthXo1* mutant derivative. (**a**) Diagram depicting the orientations (color and arrowhead) and approximate locations of the different EBEs targeted by ArTALEs relative to the ATAC locus transcripts. Numbers on top of ArTALE EBEs identify individual ArTALEs in panels B and C below. (**b**) TalC target transcript expression ratios in response to ArTALEs designed to target the ATAC locus. qRT-PCR was performed with leaf samples collected 48 h post infiltration of the Kitaake wild-type rice line with derivatives of the ME2 (PXO99 mutated in the major TALE gene *pthXo1*) strain carrying an empty vector (EV) or related plasmids encoding the indicated TALEs. The plotted data (gray dots) derive from two independent experiments (n = 4–6). The black dots and line ranges denote the mean and the standard error, respectively. Treatments with the same letter in an individual transcript panel are not significantly different (Dunn's test for post hoc pairwise multiple comparisons; p-value < 0.05). (**c**) Lesion length measured on wild-type rice line Kitaake 14 days post leaf clipping inoculation with the indicated strains. The plotted data derive from three independent experiments (dots with different gray levels). Additional water control leaves did not show BLB symptoms. The boxplots represent the median and the first and third quartiles, whereas the whisker extends from the hinge to the largest or the lowest value but no further than 1.5 times the distance between the first and third quartiles. Treatments with the same letter in an individual transcript panel are not significantly different (Dunn's test for post hoc pairwise multiple comparisons; p-value < 0.05). Numbers in parentheses represent the count of values for the corresponding treatment.

Table 2. List of *X. oryzae* strains used in this study.

Designation	Strain ID	Description	References
BAI3	CIX151	A rifampicin-resistant and fully virulent derivative of the BAI3 wild-type *Xoo* strain from Burkina Faso	[33]
BAI3H	CIX3878	CIX151 deleted for a portion of the *hrcC* gene	[31]
BAI3.1.1	CIX531	Also referred to as BAI3^{R} Δ*talC*; corresponds to CIX151 with a suicide plasmid insertion in *talC*	[31]
BAI3.1.1_pEV	CIX2533	CIX531 carrying the empty pSKX1 plasmid	[26]
BAI3.1.1_pTalC	CIX2529	CIX531 carrying the pSKX1plasmid containing the *talC* gene from strain BAI3	[26]
ME2	CIX4497	PXO99^{A} derivative with a knockout mutation in *pthXo1*	[42]
ME2_EV	CIX4498	CIX4497 carrying the empty pSKX1 plasmid	This study
ME2_pTALC	CIX4499	CIX4497 carrying the pSKX1plasmid containing the *talC* gene from strain BAI3	This study
ME2_pTALF	CIX4500	CIX4497 carrying the pSKX1plasmid containing the *talF* gene from strain MAI1	This study
ME2_pArTAL01	CIX4501	CIX4497 carrying the pHD62 plasmid expressing ArtTale1	This study
ME2_pArTAL04	CIX4502	CIX4497 carrying the pHD65 plasmid expressing ArtTale4	This study
ME2_pArTAL05	CIX4503	CIX4497 carrying the pHD66 plasmid expressing ArtTale5	This study
ME2_pArTAL06	CIX4504	CIX4497 carrying the pHD67 plasmid expressing ArtTale6	This study

Next, to test if ATAC transcript induction could confer enhanced virulence on the ME2 mutant, we screened those strains in leaf clipping virulence assays on Kitaake plants. As shown in Figure 3c, the TalC- and TalF-expressing strains caused disease lesions longer than 10 cm, well above the empty vector control strain and consistent with previous similar experiments [7,8]. However, with the exception of the ME2 derivative expressing ArTALE04, lesion lengths recorded for other ArTALE-expressing strains were not different from those produced with the empty vector control strain.

We, therefore, conclude that while our ArTALEs could simultaneously induce both *ATAC1* and *ATAC2* transcripts or *ATAC2* alone, they did not consistently compensate for the loss of clade III OsSWEET induction and failed to confer increased virulence on the ME2 mutant strain in our conditions.

2.4. Multiplex Editing of the OsSWEET14 Promoter and the ATAC Locus Using CRISPR/Cas9

Negative results in these gain-of-function experiments could stem from a number of reasons. If ATAC is indeed redundant with *OsSWEET14*, another prediction of our working model is that TalC-unresponsive plants should display enhanced resistance to African Xoo strains. To investigate the significance of ATAC targeting by TalC with a complementary loss-of-function approach, we therefore applied CRISPR/Cas9 genome editing to inactivate TalC target loci in multiplex. To this end, we employed a previously described system [43] to generate a set of binary constructs with *cas9* and an array of 4 gRNAs (Figure S4) for the genetic transformation of the Nipponbare variety. To abrogate TALE-mediated *OsSWEET14* induction, all our three constructs bore a gRNA targeting EBE_{TalF} and another gRNA targeting EBE_{TalC}@OsSWEET14. In order to query different regions of the ATAC loci, they

differed at the level of the two remaining gRNAs: a pair of gRNA targeting EBE_{TalC}@ATAC or a nearby sequence was included in one construct with the goal of abrogating TalC responsiveness. Additional pairs of gRNAs targeting either *ATAC1* or *ATAC2* transcript sequences for the insertion of deleterious mutations in the underlying locus, including the deletion of intervening sequences, were cloned in the second or third constructs (Figure S4 and Table S1). Finally, two complementary control constructs designed for single locus editing and comprising only the *OsSWEET14* promoter gRNAs or the pair of gRNAs targeting *ATAC2* were also included in transformation experiments.

Ultimately, after transformation, genotyping, and propagation of edited plants, we gathered a set of seven homozygous, non-transgenic T3 lines with a range of edition events as described in Figures 4a and S5. The efficiency of EBE_{TalF} editing was too low to select consistently modified lines. However, all sweet14 alleles carry mutations that are likely to prevent TalC recognition. The *atac-1* allele consists in a 100 bp deletion of the *ATAC1* transcript, including a portion of the longest predicted ORF. Furthermore, we obtained several alleles with insertions or deletions upstream of position 8 of EBE_{TalC}@ATAC (*attac2*, *atac4*, and *atac5*) (see Figures 4a and S5). We were, however, unable to recover edited lines combining an unresponsive allele at EBE_{TalC}@OsSWEET14 and a modified allele of *ATAC2*.

2.5. Bidirectional Induction of Transcription at ATAC Is Dependent on the Predicted TalC EBE, but TalC-Unresponsive Plants Remain Essentially Susceptible to an African Xoo Strain

Next, we examined TalC-responsiveness of edited lines using qRT-PCR following leaf infiltration of the BAI3 strain (Figure 4b). As opposed to wild-type Nipponbare or the *atac-6* line, and consistent with our previous results [35], *OsSWEET14* induction was abrogated in all lines carrying a mutation affecting the EBE_{TalC} on its promoter.

Plants containing the *atac-1* allele were not measured for *ATAC1* levels because of the large deletion in the genomic sequence of the transcript. Likewise, the *atac-6* plants were not assayed for *ATAC2* expression because of the deletion of most of the gene sequences, including EBE_{TalC}@ATAC (Figure S5). With these potential caveats in mind, plants with the *atac-1* allele were not compromised for *ATAC2* induction that was similar to the wild-type plants. In contrast, in *atac-6* plants, *ATAC1* was not responsive to BAI3 inoculations, possibly because EBE_{TalC}@ATAC is missing in those individuals. In agreement with this view, induction was essentially abrogated for both ATAC transcripts in other lines with more subtle modifications at EBE_{TalC}@ATAC (*atac2*, *atac4*, and *atac5* alleles). This is indicative that an intact EBE_{TalC}@ATAC is necessary for TalC recognition and subsequent divergent transcription activation.

Having determined that our edition events panel encompasses a range of effects relative to the potential function and the regulation of the ATAC loci combined with TalC-unresponsive *OsSWEET14* promoters, we asked if this resulted in altered susceptibility to the *Xoo* strain BAI3. As shown in Figure 4c, while in our leaf clipping assays with BAI3, and as reported previously [7], the double *OsSWEET14*/*OsSWEET13* knockout plants (KO-S13-S14) displayed strongly reduced lesions relative to the wild-type Kitaake controls, the resistance level of our edited plants in the Nipponbare background was not commensurately affected. Even if the lesion lengths for two lines edited in multiplex at both ATAC and *OsSWEET14* loci (sweet14–38 atac-1 and sweet14–42 atac-5) were significantly different than wild-type and single edited lines (*sweet14–36* and *sweet14–37*), this effect was modest and not consistent across modified lines with similar anticipated genetic defects.

Overall, the production and phenotypical analysis of a series of ATAC-edited Nipponbare lines revealed that the putative TalC EBE is required for bidirectional induction of transcription at this locus by BAI3 but simultaneous abrogation of TalC responsiveness at both *OsSWEET14* and ATAC loci did not confer dramatically enhanced resistance to disease caused by the African *Xoo* BAI3 strain.

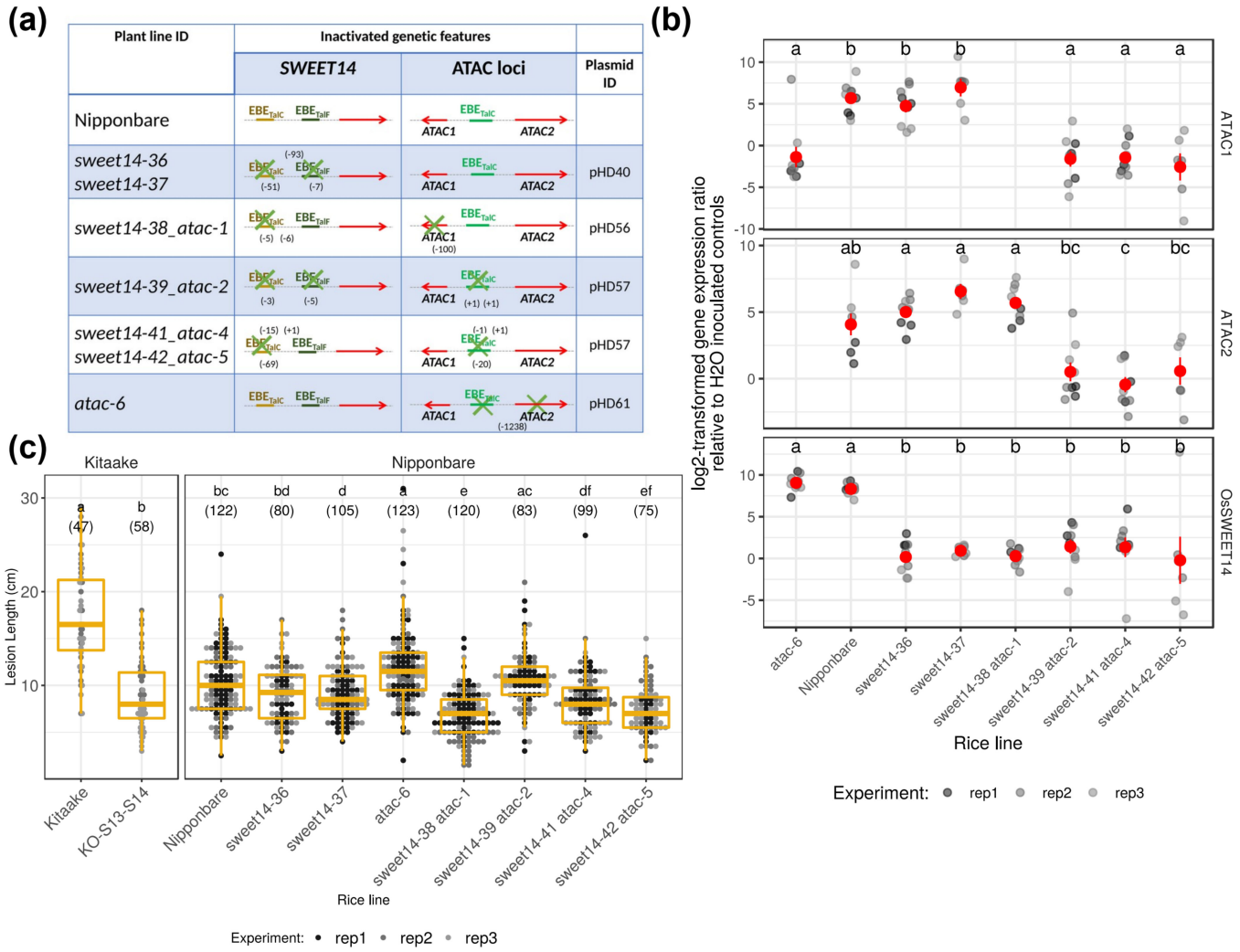


Figure 4. ATAC-locus-edited lines confirm that EBE_{TalC}@ATAC is functional but does not support a role as a BLB disease susceptibility locus. (**a**) A schematic diagram describing the nature of the editing event in the modified Nipponbare homozygous lines characterized below. A cross on top of a functional feature indicates that the sequence of the corresponding allele has been altered. Red arrows represent the *OsSWEET14* or *ATAC* genes. Colored segments depict TALE target EBEs. Values in parentheses aligned with a line identifier inform about the extent and approximate location of DNA editions as in Figure S5. Features are not drawn to scale. (**b**) TalC target transcript expression ratios in multiplex edited lines 48 h post infiltration with BAI3 relative to water-infiltrated control leaves. The plotted data (gray dots) derive from three independent experiments (shades of gray; n = 6–9). The black dots and line ranges denote the mean and the standard error, respectively. Treatments with the same letter in an individual transcript panel are not significantly different (Dunn's test for post hoc pairwise multiple comparisons; p-value < 0.05). For a couple of genotypes, sequence alterations precluded qRT-PCR on a specific target, and no data are reported (no mean difference letter; see the main text for details). (**c**) Leaf clipping susceptibility assays performed with the edited lines described in panel A challenged with the BAI3 strain. Lesion lengths were measured 14 days post inoculation. Additional water control leaves did not show BLB symptoms. The plotted data derive from three independent experiments (dots with different gray levels). The boxplots represent the median and the first and third quartiles, whereas the whisker extends from the hinge to the largest or the lowest value but no further than 1.5 times the distance between the first and third quartiles. Treatment comparison was performed separately depending on the wild-type rice background, and those with the same letter are not significantly different (Dunn's test for post hoc pairwise multiple comparisons; p-value < 0.05). Numbers in parentheses represent the count of values for the corresponding treatment.

3. Discussion

3.1. TalC Functions and the Dependence of Af Xoo on OsSWEET Induction for Infection

In this study, following up on our earlier results, we investigated potential alternative TalC targets in the rice genome and identified a second TalC-responsive locus coined ATAC. Recent data indicate that multi-EBE-edited Kitaake lines with an altered EBE_{TalC}@OsSWEET14 are resistant to the ME2 derivative complemented with *talC*, which is otherwise fully virulent on wild-type Kitaake rice [8]. Similarly, *talC* no longer provides increased lesions to ME2 on the *OsSWEET14* knockout line versus wild-type Kitaake [7]. Thus, at least in the ME2 background, *talC* contribution to virulence is abrogated when either host *OsSWEET14* gene activity or transcriptional induction is compromised. This is consistent with the view that *talC* virulence activity is exclusively channeled via an *OsSWEET14* pathway and contradicts our hypothesis that alternative TalC rice targets, including the ATAC locus, could redundantly contribute to susceptibility.

Considering these observations and our inability to demonstrate a contribution of ATAC transcripts using either ArTALEs (Figure 3c) or edited lines (Figure 4c), we provisionally conclude that transcriptional targeting of ATAC is likely fortuitous and does not contribute to the virulence of African strains causing BLB. In this context, the mild ME2-virulence-enhancing activity of ArTALE04 only (Figure 3c) can only be interpreted as an off-target effect of this particular construct.

A growing list of TALEs has been shown to simultaneously target more than one host genomic locus for transcription activation. Most documented examples, such as AvrXa7 on *OsSWEET14* and the transposon Os04g19960 [5], the *X. oryzae* pv. *orzicola* Tal2g on rice *OsSULTR3;6* and Os06g46500 [44], or the *X. citri* subsp. *citri* TALE PthA4 on *CsLOB1* and *CsSWEET1* [45], involve a genuine susceptibility gene and a second host locus that does not measurably affect disease, as is likely the case with TalC. Nonetheless, dual upregulation of independent bona fide susceptibility genes by single TALEs has been demonstrated. For example, the African Xoo TALE TalB targets *OsTFX1* and *OsERF#123*, both of which increase BLB symptoms upon induction [25]. The tomato (*Solanum lycopersicum*) transcription factor genes *bHLH3* (Solyc03g097820) and *bHLH6* (Solyc06g072520) are both regulated by AvrHah1, a TALE from X. gardneri and can individually promote water soaking, a typical disease readout [46].

Incidentally, regardless of whether this impinges on plant–pathogenic bacteria interactions, TALEs appear to repeatedly regulate genes of the large plant basic/helix–loop–helix (bHLH) transcription factor family [47]. In addition to AvrHah1 (mentioned above) and TalC, with the newly discovered *ILI2/ATAC2* target reported in this study, an archetypal example of this is AvrBs3 from *X. campestris* pv. *vesicatoria* and its bHLH target *UPA20*, which controls pepper (*Capsicum annuum*) cell hypertrophy [13]. Many strains of the *Xoo*-related *oryzicola* pathovar of *X. oryzae* broadly upregulate the rice Os03g07540 locus, which also encodes a bHLH protein [48]. With the BLS256 strain, this effect is mediated by the Tal3c TALE [44]. Intriguingly, Os03g07540, which corresponds to *ILI3* [40], is therefore a close homolog of ILI2/ATAC2, the alternative TalC target identified here.

3.2. Bidirectional Induction of Transcription by TALEs

Arguably, one of the most important outcomes of our work, with relevance to the molecular mechanisms of TALE-mediated transcription activation in plants, is the observation that the TalC bacterial effector triggers divergent bidirectional transcription at ATAC. Both ATAC1 and ATAC2 transcripts are coordinately induced in a TalC-dependent manner (Figure 2), and this requires an intact putative TalC EBE at the ATAC locus (Figure 4).

ATAC1 is a previously undocumented short (456 nt) rice transcript (Figure 1) whose putative translation products are less than 100 amino acids (Figure S2) and that is predicted to be an lncRNA by a classifier trained on a rice-specific dataset. Plant lncRNAs are comparatively less understood than animal lncRNAs, and they comprise a relatively diverse set of molecules having functions in plant development and adaptation to both abiotic and biotic stresses through a variety of mechanisms [49]. Whether *ATAC1* has an intrinsic

biological function or is just an accidental by-product of TalC activity is an interesting question that may be illuminated in future work. In line with another study [50], the discovery of the ATAC1 transcript in a reconstructed transcriptome is, nonetheless, a strong hint that the search for TALE target sequences should not solely rely on incomplete reference plant genome annotations and must incorporate experimental RNA-seq information to enable the identification of novel candidates.

Two previous studies have demonstrated that designer TALEs can activate unidirectional transcription of sequences upstream of the EBE on the opposite strand (antisense transcription). This was shown to occur both in transient assays in *Nicotiana benthamiana* and during leaf infection. ArTALEs designed to target an EBE in an antisense orientation relative to *OsSWEET14* [15,16] or *OsSULTR3;6* [16] susceptibility genes were able to complement the virulence of strains deficient for the cognate natural TALE. Wang et al. [16] provided additional insight and showed that both artificial and natural TALEs can activate bidirectional transcription in Agrobacterium-mediated transient expression assays in *N. benthamiana* with a T-DNA carrying a dual reporter construct. The same authors also described two functional EBEs in opposite orientation for the natural *X. oryzae* pv. *orzicola* TALE Tal3c upstream of the Os09g39810 locus, but they detected no sign of bidirectional transcription in mRNA-seq expression data during infection. Here, in contrast, our results represent the first tangible evidence that a native TALE can activate bidirectional transcription when delivered by the T3SS during infection.

ATAC1 is reminiscent of enhancer-associated lncRNAs (eRNAs) [51,52]. While in animals, enhancers and gene promoters frequently undergo divergent RNA-polymerase-II-mediated transcription, in plants, the evidence for bidirectional transcription from promoters of protein-coding genes is relatively scarce and previously defined bidirectional plant promoters [53–56] may entail a different phenomenon. However, hypomorphic mutations in the Arabidopsis exosome complex or its associated helicase system have recently enabled the detection of about a hundred loci with antisense RNAs upstream of mRNA transcripts and an equivalent number of noncoding regions with signs of bidirectional transcription reminiscent of eRNAs [57]. TALE-driven divergent transcription may be a consequence of an idiosyncratic mode of action of these bacterially encoded transcription factors. This property probably depends on the nature of the additional plant transcription initiation machinery components that are recruited and that are just beginning to be deciphered [58]. The current picture is, however, that TALEs predominantly activate unidirectional transcription downstream of their EBEs; for example, there is no signal of antisense transcription at the *OsSWEET14* (LOC_Os11g31190) promoter (Figure S1). Thus, as suggested before [15,16], other promoter elements (e.g., the TATA-box) but also DNA conformational properties (dictated in part by chromatin modifications) conceivably determine the local mode of transcription activation (i.e., bidirectional, sense or antisense only) by stimulating the release of halted RNA polymerase II complexes or by inhibiting the exosomal processing of nascent transcripts. In this regard, it is probably not fortuitous that our ArTALE05 and 06 with an EBE on the same strand as EBE_{TalC}@ATAC trigger divergent transcription whereas ArTALE01 and 04 with an EBE oriented toward *ATAC2* activate unidirectional sense transcription only (Figure 3).

In conclusion, pursuing the hypothesis that TalC targeted a second redundant susceptibility gene in rice, we did identify a second locus but failed to demonstrate its contribution to disease susceptibility. Hence, the basis of the unique ability of African Xoo to persistently colonize rice in the absence of clade III OsSWEET gene induction remains enigmatic and shall be investigated further. In this endeavor, we nonetheless documented the first example of bidirectional transcription activation by a native TALE during infection, which will be a valuable system for a better understanding of the molecular mechanisms underlying TALE activity.

4. Materials and Methods

4.1. Bacterial Strains, Plant Inoculations, and Growth of Bacteria in Planta

The wild-type rice genotypes Kitaake and Nipponbare varieties used in this study were originally obtained from the International Rice Research Institute and the IRGSP consortium [37], respectively. The *OsSWEET14*/*OsSWEET13* double-knockout edited line corresponds to the *sweet13;14* +1bp/+1bp line from Eom et al. [7]. All rice plants were grown under cycles of 12 h of light at 28 °C and 80% relative humidity (RH) and 12 h of darkness at 25 °C and 70% RH in the greenhouse.

The plasmids and *Xoo* bacterial strains used in this study are described in Tables S1 and S2, respectively. Bacterial strains stored in 15% glycerol at −80 °C were cultivated on PSA (10 g/L peptone, 10 g/L sucrose, 1 g/L glutamic acid, and 16 g/L Bacto Agar). The strains containing plasmids were grown on PSA supplemented as needed with kanamycin, gentamicin, or rifampicin at concentrations of 50 µg/mL, 20 µg/mL, and 50 µg/mL, respectively. The cultures were incubated at 28 °C for 48 h.

Inoculations were conducted as before [31]. Briefly, bacteria grown on plates for 48 h were resuspended in sterile water at an OD600 = 0.2 for leaf clipping assay inoculum preparation or OD600 = 0.5 for leaf infiltration. Leaves of 6-week-old plants were cut with scissors previously dipped in bacterial suspensions, whereas leaves of 4-week-old plants were pressure infiltrated with a 1 mL syringe. Lesion lengths were measured 14 days post inoculation.

4.2. RNA Isolation and qRT-PCR

The leaves of Kitaake, Nipponbare, or edited rice line plants were infiltrated with bacterial suspensions or water, and three independent replicate samples were processed at 48 h after inoculation. Total RNA was extracted using TRI reagent (EUROMEDEX, Souffelweyersheim, France) and further purified using the RNA Clean-Up & Concentration kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. A total of 1 µg of RNA per sample was treated with the TURBO DNA-free kit (Thermo Fisher Scientific, Waltham, MA, USA) and reverse-transcribed into cDNA using SuperScriptIII (Thermo Fisher Scientific, Waltham, MA, USA) with oligo-dT primers. PCR reactions were performed with SYBR-based Mesa Blue qPCR Mastermix (EUROGENTEC, Liège, Belgium).

For the OsSWEET14 gene, qRT-PCR oligos were reported before [35]. For the ATAC locus, a primer pair specific to the ATAC1 transcript (GAGGTGGAGCTGGAGCTATG and TACAACGGGCTACAACCACA) and a primer pair specific to the ATAC2 transcript (CGTCCGATCTCGCCATCAAC and GAGAACATTTTAGGCCGACAGCA) were used. The elongation gene factor-1 alpha (GenBank: GQ848072.1) was used for normalization of the expression values of complementary DNA in each sample as before [35]. Normalized expression values were calculated with the $2^{-\Delta Ct}$ methods, whereas expression ratios were calculated with the $2^{-\Delta\Delta Ct}$ method.

4.3. Plasmid Vector Construction

The design and construction of artificial TAL effector (ArTALEs) expression vectors involved the use of the method of Streubel et al. [26].

For rice genome editing, first, the pHD46 vector was generated by PCR amplifying the dual gRNA cassette from pENTR4:gRNA4 [43] with a pair of primers wherein the forward primer (ACTTAAGCTTGAATTCagaacgaactaagccggaca) comprises the HindIII and EcoRI restriction sites and the reverse primer (ACCTTTCTAGAgcccttcgaagggacaaaaa) comprises an XbaI site. The amplified sgRNA cassette was digested and gel-purified using the DNA Gel Recovery Kit (Zymo Research, Irvine, CA, USA). This new cassette was cloned into the original pENTR4:gRNA4 vector digested at the EcoRI and XbaI restriction sites that are located in between the attL sequences.

The E-CRISPR software [59] was used to design synthetic guide RNAs (sgRNAs). The sgRNA sequences are indicated in Figure S4. When the designed gRNA sequence 5′ nucleotide was not a guanine (G), a starting G was appended to this sequence to promote

U6 promoter transcription initiation at this site. Complementary oligos spanning sgRNA sequences with appropriate 4 bp overhangs were thermally annealed to each other, diluted (1:200) with distilled water, and cloned into pENTR4:gRNA4 or pHD46 vectors using the procedure described by Zhang et al. [43].

For the assembly of an array of four gRNA expression units into a Gateway entry vector, pHD46-derived EcoRI-XbaI cassettes with ATAC locus targeting gRNAs were cloned into pHD23, a pENTR4:gRNA4-based vector bearing the OsSWEET14 promoter targeting gRNAs, digested with the same enzymes. This array was subsequently transferred into a modified pH-Ubi-cas9-7 [60] vector documented by Fayos et al. [61] in a Gateway LR reaction (Thermo Fisher Scientific, Waltham, MA, USA). The gRNA arrays of the resulting binary vectors (see Table S1) were confirmed again by sequencing before Agrobacterium transformation was continued.

4.4. Genetic Transformation of Rice and Modified Allele Genotyping

For rice genetic transformation, *Agrobacterium tumefaciens* EHA105 was transformed with the set of binary vectors by electroporation. The embryo calli derived from Nipponbare were used and co-cultured with bacteria as described [35,62]. T0 plants were screened to characterize multiplex editing events. The MATAB protocol [63] was used to extract DNA from leaf samples. The amplification products of the OsSWEET14 promoter were obtained as described before [35]. The ATAC locus was PCR amplified with specific forward (GACTGTAGCTTTGCCTAGCT) and reverse (CTACAGATGAAGAAGCTAGAG) primers. The resulting amplicons were Sanger sequenced and analyzed with DSDecode software [64] to identify mutations within the *OsSWEET14* promoter. For simultaneous detection of mutation of the *OsSWEET14* promoter and ATAC locus, the PCR products were cloned into the pGMT-T vector according to the instructions of the manufacturer (Promega, Madison, WI, USA). The clones were sequenced and analyzed with DSDecode software to decode the mutations.

4.5. Computational Methods for TalC Candidate Target Identification

The mRNA-seq data used in this study for experimental transcriptome inference and differential expression analysis are available in SRA (NCBI BioProject accession: PRJNA679478) and have been described before [36].

Paired-end reads were trimmed with AdapterRemoval version 2.2.2 [65] using the following relevant parameters: "–combined-output –gzip –trimns –trimqualities –mm 5 –minadapteroverlap 3 –minlength 25". Trimmed reads were subsequently mapped onto the nuclear rice genome. Rice Nipponbare IRGSP-1.0 pseudomolecule sequences and MSU7 gene annotation data [37] were obtained from MSU (http://rice.uga.edu/ (accessed on 10 March 2022)). For mapping, HISAT2 (v.2.1.0) was run with the following parameters: "-q –new-summary –dta". Experimental transcriptome reconstruction essentially followed the recommended procedure [66]. Transcripts were first inferred based on mRNA-seq mappings and the MSU7 reference annotation on an individual sample basis with StringTie (v.1.3.4d) and the following parameters: "-l GXY-1 –fr -m 100 -t -f 0.1 -c 5 -j 2". Individual gtf files were subsequently merged with the reference annotation in a single step with the StringTie –merge mode and default parameter values.

Gene-level and strand-aware mated read pair counts were obtained with R (v.3.5.1) using the summarizeOverlaps function from the GenomicAlignments (v.1.18.1) Bioconductor package [67], discarding reads overlapping multiple features ("Union" mode and "inter.feature" true). StringTie gene loci that accumulated less than 5 total reads across all libraries in the dataset were excluded. Likewise, loci that overlapped rice loci with feature types tRNA, rRNA, or pseudogenic_tRNA in the gff annotation file for NCBI RefSeq assembly accession GCF_001433935.1 were excluded from differential expression (DE) analysis.

Differential expression was estimated with the DESeq2 (v.1.22.2) R Bioconductor package [68]. For normalization, size factors were estimated using the iterate method and

the shorth function to compute a location for a sample. Dispersions were fitted using the parametric method. Two-tailed Wald tests were used for DE tests where the alternative hypothesis was that the absolute value of the log2 transform of the fold change ratio is greater than 0.5. Adjusted *p*-values were computed with the Benjamini & Hochberg method and considered significant when equal to or below 0.1. Outlier filtering based on Cook's distance was implemented.

To generate a list of upstream sequences for EBE prediction, the DESeq2 gene list was arbitrarily further filtered to retain only the loci with a log2(FC) value ≥ 2 and an adjusted *p*-value ≤ 0.01 in both BAI3 vs. BAI3H and BAI3 vs. H_2O comparisons. EBE predictions for the TalC RVD sequence on this set of upstream sequences (−500 bp to +200 bp relative to annotated gene start) used Talvez (v.3.2) with parameters "-t 0 -l 19" and default RVD–nucleotide association matrixes [38].

The microarray expression data used to define TalC-induced genes were published earlier [31] and analyzed as previously described [38]. Only probesets with a log2 transform of the fold change ratio ≥ 1 in the XooBAI3-XooBAI3_Dtalc comparison were considered.

The PlncPRO [39] model for *O. sativa* was built using data and instructions from https://github.com/urmi-21/PlncPRO/blob/master/examples/Build%20and%20test%20monocot%20model.ipynb (accessed on 23 July 2021). The plncpro predict ("-v -r") command was run using the StringTie-predicted ATAC1 transcript sequence.

Data plots were generated in R (v. 3.6.3) with the ggplot2 (v.3.3.2) or Gviz (v.1.30.3) packages, and statistical tests were conducted with functions from the rstatix (v.0.6.0) package.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/ijms23105559/s1.

Author Contributions: Conceptualization, H.D., F.A., S.B.-B. and S.C.; methodology, H.D., F.A., S.B.-B., S.F., C.S., B.S., V.V. and S.C.; software, M.G. and S.C.; validation, H.D., F.A. and S.C.; formal analysis, H.D., F.A., S.B.-B. and S.C; investigation, H.D., F.A., S.B.-B., S.F., E.T., F.D. and C.S.; resources, B.S.; data curation, H.D., F.A., S.B.-B. and S.C.; writing—original draft preparation, H.D., F.A. and S.C.; writing—review and editing, H.D., F.A., S.B.-B., S.F., M.G., E.T., F.D., C.S., B.S., O.K., V.V. and S.C.; visualization, H.D., F.A. and S.C.; supervision, O.K., V.V. and S.C.; project administration, O.K., V.V. and S.C.; funding acquisition, H.D., O.K., V.V. and S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by a Monsanto's Beachell-Borlaug International Scholars Program PhD fellowship awarded to H.D. and a Chercheur d'Avenir grant from the Region Occitanie attributed to S.C.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article and supplementary material on our dataverse at https://doi.org/10.23708/XBTRUV, last accessed on 23 July 2021.

Acknowledgments: The authors acknowledge the ISO 9001-certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing high-performance computing resources.

Conflicts of Interest: The authors declare no conflict of interest. The sponsors had no role in the design, execution, interpretation, or writing of the study.

References

1. Niño-Liu, D.O.; Ronald, P.C.; Bogdanove, A.J. *Xanthomonas oryzae* pathovars: Model pathogens of a model crop. *Mol. Plant Pathol.* **2006**, *7*, 303–324. [CrossRef] [PubMed]
2. Savary, S.; Willocquet, L.; Pethybridge, S.J.; Esker, P.; McRoberts, N.; Nelson, A. The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* **2019**, *3*, 430–439. [CrossRef] [PubMed]
3. Jiang, N.; Yan, J.; Liang, Y.; Shi, Y.; He, Z.; Wu, Y.; Zeng, Q.; Liu, X.; Peng, J. Resistance Genes and their Interactions with Bacterial Blight/Leaf Streak Pathogens (*Xanthomonas oryzae*) in Rice (*Oryza sativa* L.)—An Updated Review. *Rice* **2020**, *13*, 3. [CrossRef] [PubMed]

4. Kumar, A.; Kumar, R.; Sengupta, D.; Das, S.N.; Pandey, M.K.; Bohra, A.; Sharma, N.K.; Sinha, P.; Sk, H.; Ghazi, I.A.; et al. Deployment of Genetic and Genomic Tools Toward Gaining a Better Understanding of Rice-*Xanthomonasoryzae* pv. oryzae Interactions for Development of Durable Bacterial Blight Resistant Rice. *Front. Plant Sci.* **2020**, *11*, 1152. [CrossRef] [PubMed]
5. Li, T.; Liu, B.; Spalding, M.H.; Weeks, D.P.; Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **2012**, *30*, 390–392. [CrossRef] [PubMed]
6. Duy, P.N.; Lan, D.T.; Thu, H.P.; Thu, H.P.T.; Thanh, H.N.; Pham, N.P.; Auguy, F.; Manh, T.B.; Cunnac, S.; Pham, X.H. Improved bacterial leaf blight disease resistance in the major elite Vietnamese rice cultivar TBR225 via editing of the OsSWEET14 promoter. *PLoS ONE* **2021**, *16*, e0255470. [CrossRef]
7. Eom, J.S.; Luo, D.; Atienza-Grande, G.; Yang, J.; Ji, C.; Thi Luu, V.; Huguet-Tapia, J.C.; Char, S.N.; Liu, B.; Nguyen, H.; et al. Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* **2019**, *37*, 1372–1379. [CrossRef]
8. 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]
9. Xu, Z.; Xu, X.; Gong, Q.; Li, Z.; Li, Y.; Wang, S.; Yang, Y.; Ma, W.; Liu, L.; Zhu, B.; et al. Engineering Broad-Spectrum Bacterial Blight Resistance by Simultaneously Disrupting Variable TALE-Binding Elements of Multiple Susceptibility Genes in Rice. *Mol. Plant* **2019**, *12*, 1434–1446. [CrossRef]
10. Zafar, K.; Khan, M.Z.; Amin, I.; Mukhtar, Z.; Yasmin, S.; Arif, M.; Ejaz, K.; Mansoor, S. Precise CRISPR-Cas9 Mediated Genome Editing in Super Basmati Rice for Resistance Against Bacterial Blight by Targeting the Major Susceptibility Gene. *Front. Plant Sci.* **2020**, *11*, 575. [CrossRef]
11. Zeng, X.; Luo, Y.; Vu, N.T.Q.; Shen, S.; Xia, K.; Zhang, M. CRISPR/Cas9-mediated mutation of OsSWEET14 in rice cv. Zhonghua11 confers resistance to *Xanthomonas oryzae* pv. oryzae without yield penalty. *BMC Plant Biol.* **2020**, *20*, 313. [CrossRef] [PubMed]
12. Kay, S.; Hahn, S.; Marois, E.; Wieduwild, R.; Bonas, U. Detailed analysis of the DNA recognition motifs of the Xanthomonas type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant J.* **2009**, *59*, 859–871. [CrossRef] [PubMed]
13. 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]
14. Römer, P.; Strauss, T.; Hahn, S.; Scholze, H.; Morbitzer, R.; Grau, J.; Bonas, U.; Lahaye, T. Recognition of AvrBs3-Like Proteins Is Mediated by Specific Binding to Promoters of Matching Pepper *Bs3* Alleles. *Plant Physiol.* **2009**, *150*, 1697–1712. [CrossRef] [PubMed]
15. Streubel, J.; Baum, H.; Grau, J.; Stuttmann, J.; Boch, J. Dissection of TALE-dependent gene activation reveals that they induce transcription cooperatively and in both orientations. *PLoS ONE* **2017**, *12*, e0173580. [CrossRef]
16. Wang, L.; Rinaldi, F.C.; Singh, P.; Doyle, E.L.; Dubrow, Z.E.; Tran, T.T.; Perez-Quintero, A.L.; Szurek, B.; Bogdanove, A.J. TAL Effectors Drive Transcription Bidirectionally in Plants. *Mol. Plant* **2017**, *10*, 285–296. [CrossRef]
17. 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]
18. Moscou, M.J.; Bogdanove, A.J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **2009**, *326*, 1501. [CrossRef]
19. Streubel, J.; Blücher, C.; Landgraf, A.; Boch, J. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **2012**, *30*, 593–595. [CrossRef]
20. de Lange, O.; Schreiber, T.; Schandry, N.; Radeck, J.; Braun, K.H.; Koszinowski, J.; Heuer, H.; Strauß, A.; Lahaye, T. Breaking the DNA-binding code of Ralstonia solanacearum TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol.* **2013**, *199*, 773–786. [CrossRef]
21. Doyle, E.L.; Hummel, A.W.; Demorest, Z.L.; Starker, C.G.; Voytas, D.F.; Bradley, P.; Bogdanove, A.J. 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* **2013**, *8*, e82120. [CrossRef] [PubMed]
22. Gao, H.; Wu, X.; Chai, J.; Han, Z. Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Res.* **2012**, *22*, 1716–1720. [CrossRef] [PubMed]
23. 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]
24. White, F.F.; Yang, B. Host and Pathogen Factors Controlling the Rice-X*anthomonas oryzae* Interaction. *Plant Physiol.* **2009**, *150*, 1677–1686. [CrossRef] [PubMed]
25. 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] [PubMed]
26. 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]
27. Bezrutczyk, M.; Yang, J.; Eom, J.S.; Prior, M.; Sosso, D.; Hartwig, T.; Szurek, B.; Oliva, R.; Vera-Cruz, C.; White, F.F.; et al. Sugar flux and signaling in plant-microbe interactions. *Plant J.* **2018**, *93*, 675–685. [CrossRef]
28. 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]
29. Zhou, J.; Peng, Z.; Long, J.; Sosso, D.; Liu, B.; Eom, J.-S.; Huang, S.; Liu, S.; Cruz, C.V.; Frommer, W.; 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]

30. 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]
31. 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] [PubMed]
32. Doucouré, H.; Pérez-Quintero, A.L.; Reshetnyak, G.; Tekete, C.; Auguy, F.; Thomas, E.; Koebnik, R.; Szurek, B.; Koita, O.; Verdier, V.; et al. Functional and Genome Sequence-Driven Characterization of *tal* Effector Gene Repertoires Reveals Novel Variants With Altered Specificities in Closely Related Malian *Xanthomonas oryzae* pv. *Oryzae Strains*. *Front. Microbiol.* **2018**, *9*, 1657. [CrossRef] [PubMed]
33. Gonzalez, C.; Szurek, B.; Manceau, C.; Mathieu, T.; Séré, Y.; Verdier, V. Molecular and Pathotypic Characterization of New *Xanthomonas oryzae* Strains from West Africa. *Mol. Plant Microbe Interact.* **2007**, *20*, 534–546. [CrossRef] [PubMed]
34. Poulin, L.; Grygiel, P.; Magne, M.; Gagnevin, L.; Rodriguez-R, L.M.; Serna, N.F.; Zhao, S.; El Rafii, M.; Dao, S.; Tekete, C.; et al. New Multilocus Variable-Number Tandem-Repeat Analysis Tool for Surveillance and Local Epidemiology of Bacterial Leaf Blight and Bacterial Leaf Streak of Rice Caused by Xanthomonas oryzae. *Appl. Environ. Microbiol.* **2015**, *81*, 688–698. [CrossRef]
35. 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]
36. Reshetnyak, G.; Jacobs, J.M.; Auguy, F.; Sciallano, C.; Claude, L.; Medina, C.; Perez-Quintero, A.L.; Comte, A.; Thomas, E.; Bogdanove, A.; et al. An atypical class of non-coding small RNAs is produced in rice leaves upon bacterial infection. *Sci Rep.* **2021**, *11*, 24141. [CrossRef]
37. 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.; et al. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice* **2013**, *6*, 4. [CrossRef]
38. Perez-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]
39. Singh, U.; Khemka, N.; Rajkumar, M.S.; Garg, R.; Jain, M. PLncPRO for prediction of long non-coding RNAs (lncRNAs) in plants and its application for discovery of abiotic stress-responsive lncRNAs in rice and chickpea. *Nucleic Acids Res.* **2017**, *45*, e183. [CrossRef]
40. Zhang, L.Y.; Bai, M.Y.; Wu, J.; Zhu, J.Y.; Wang, H.; Zhang, Z.; Wang, W.; Sun, Y.; Zhao, J.; Sun, X.; et al. Antagonistic HLH/bHLH transcription factors mediate brassinosteroid regulation of cell elongation and plant development in rice and Arabidopsis. *Plant Cell* **2009**, *21*, 3767–3780. [CrossRef]
41. Fan, M.; Bai, M.Y.; Kim, J.G.; Wang, T.; Oh, E.; Chen, L.; Park, C.H.; Son, S.H.; Kim, S.K.; Mudgett, M.B.; et al. The bHLH transcription factor HBI1 mediates the trade-off between growth and pathogen-associated molecular pattern-triggered immunity in Arabidopsis. *Plant Cell* **2014**, *26*, 828–841. [CrossRef] [PubMed]
42. 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] [PubMed]
43. Zhou, H.; Liu, B.; Weeks, D.P.; Spalding, M.H.; Yang, B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* **2014**, *42*, 10903–10914. [CrossRef] [PubMed]
44. 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]
45. 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]
46. 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]
47. Toledo-Ortiz, G.; Huq, E.; Quail, P.H. The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell* **2003**, *15*, 1749–1770. [CrossRef]
48. Wilkins, K.E.; Booher, N.J.; Wang, L.; Bogdanove, A.J. TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen Xanthomonas oryzae pv. oryzicola while strict conservation suggests universal importance of five TAL effectors. *Front. Plant Sci.* **2015**, *6*, 536. [CrossRef]
49. Yu, Y.; Zhang, Y.; Chen, X.; Chen, Y. Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Annu. Rev. Cell Dev. Biol.* **2019**, *35*, 407–431. [CrossRef]
50. Mücke, S.; Reschke, M.; Erkes, A.; Schwietzer, C.-A.; Becker, S.; Streubel, J.; Morgan, R.D.; Wilson, G.G.; Grau, J.; Boch, J. Transcriptional Reprogramming of Rice Cells by *Xanthomonas oryzae* TALEs. *Front. Plant Sci.* **2019**, *10*, 162. [CrossRef]
51. Arnold, P.R.; Wells, A.D.; Li, X.C. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Front. Cell Dev. Biol.* **2020**, *7*, 377. [CrossRef] [PubMed]
52. Chen, L.; Zhu, Q.H.; Kaufmann, K. Long non-coding RNAs in plants: Emerging modulators of gene activity in development and stress responses. *Planta* **2020**, *252*, 92. [CrossRef]

53. Banerjee, J.; Sahoo, D.K.; Dey, N.; Houtz, R.L.; Maiti, I.B. An Intergenic Region Shared by At4g35985, and At4g35987, in Arabidopsis thaliana Is a Tissue Specific and Stress Inducible Bidirectional Promoter Analyzed in Transgenic Arabidopsis and Tobacco Plants. *PLoS ONE* **2013**, *8*, e79622. [CrossRef] [PubMed]
54. Kourmpetli, S.; Lee, K.; Hemsley, R.; Rossignol, P.; Papageorgiou, T.; Drea, S. Bidirectional promoters in seed development and related hormone/stress responses. *BMC Plant Biol.* **2013**, *13*, 187. [CrossRef] [PubMed]
55. Liu, X.; Yang, W.; Li, Y.; Li, S.; Zhou, X.; Zhao, Q.; Fan, Y.; Lin, M.; Chen, R. The intergenic region of the maize defensin-like protein genes Def1 and Def2 functions as an embryo-specific asymmetric bidirectional promoter. *J. Exp. Bot.* **2016**, *67*, 4403–4413. [CrossRef] [PubMed]
56. Wang, Q.; Wan, L.; Li, D.; Zhu, L.; Qian, M.; Deng, M. Searching for bidirectional promoters in Arabidopsis thaliana. *BMC Bioinform.* **2009**, *10*, S29. [CrossRef] [PubMed]
57. Thieffry, A.; Vigh, M.L.; Bornholdt, J.; Ivanov, M.; Brodersen, P.; Sandelin, A. Characterization of *Arabidopsis thaliana* Promoter Bidirectionality and Antisense RNAs by Inactivation of Nuclear RNA Decay Pathways. *Plant Cell* **2020**, *32*, 1845–1867. [CrossRef]
58. 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]
59. Heigwer, F.; Kerr, G.; Boutros, M. E-CRISP: Fast CRISPR target site identification. *Nat. Methods* **2014**, *11*, 122–123. [CrossRef]
60. Miao, J.; Guo, D.; Zhang, J.; Huang, Q.; Qin, G.; Zhang, X.; Wan, J.; Gu, H.; Qu, L.-J. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* **2013**, *23*, 1233–1236. [CrossRef]
61. Fayos, I.; Meunier, A.C.; Vernet, A.; Navarro-Sanz, S.; Portefaix, M.; Lartaud, M.; Bastianelli, G.; Périn, C.; Nicolas, A.; Guiderdoni, E. Assessment of the roles of SPO11-2 and SPO11-4 in meiosis in rice using CRISPR/Cas9 mutagenesis. *J. Exp. Bot.* **2020**, *71*, 7046–7058. [CrossRef] [PubMed]
62. Sallaud, C.; Meynard, D.; van Boxtel, J.; Gay, C.; Bès, M.; Brizard, J.P.; Larmande, P.; Ortega, D.; Raynal, M.; Portefaix, M.; et al. Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor. Appl. Genet.* **2003**, *106*, 1396–1408. [CrossRef] [PubMed]
63. Romero, L.E.; Lozano, I.; Garavito, A.; Carabali, S.J.; Triana, M.; Villareal, N.; Reyes, L.; Duque, M.C.; Martinez, C.P.; Calvert, L.; et al. Major QTLs control resistance to Rice Hoja Blanca Virus and its vector Tagosodes orizicolus. *G3 Genes Genomes Genet.* **2014**, *4*, 133–142.
64. Liu, W.; Xie, X.; Ma, X.; Li, J.; Chen, J.; Liu, Y.G. DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. *Mol. Plant* **2015**, *8*, 1431–1433. [CrossRef]
65. Schubert, M.; Lindgreen, S.; Orlando, L. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **2016**, *9*, 88. [CrossRef]
66. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650–1667. [CrossRef]
67. Lawrence, M.; Huber, W.; Pagès, H.; Aboyoun, P.; Carlson, M.; Gentleman, R.; Morgan, M.; Carey, V.J. Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* **2013**, *9*, e1003118. [CrossRef]
68. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]