

TALEs als modulare Aktivatoren pflanzlicher Gene

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Frau Jana Streubel

geb. am 29.01.1987 in Oschatz

Gutachter:

1. PD Dr. Jens Boch
2. Prof Dr. Klaus Humbeck
3. Prof. Dr. Uwe Sonnwald

Tag der Verteidigung: 08.07.2015

Publikationsliste

Publikationen die in diese Arbeit eingegangen sind

Streubel, J., Blücher, C., Landgraf, A., and Boch, J. (2012) TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**: 593-595.

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

Richter A., **Streubel J.**, Blücher C., Szurek B., Reschke M., Grau J. and Boch J. (2014) A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* **5**: 3447.

Zusätzliche Publikationen

Geißler, R., Scholze, H., Hahn, S., **Streubel, J.**, Bonas, U., Behrens, S.-E., and Boch, J. (2011) Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**: e19509.

Scholze H., **Streubel J.** and Boch J. (2011) TAL effectors from *Xanthomonas*: design of a programmable DNA-binding specificity. *B. Insectol.* **64**: 279-280.

Yu Y., **Streubel J.**, Balzergue S., Champion A., Boch J., Koebnik R., Feng J., Verdier V., and Szurek B. (2011) colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 *Os11N3* Gene. *Mol. Plant-Microbe Interact.* **24**: 1102-1113.

Streubel, J., Richter, A., Reschke, M., and Boch, J. (2013) TALEs - Proteine mit programmierbarer DNA-Bindesepezifität. *BIOspektrum* **19**: 370-372.

Zusammenfassung

TALEs aus *Xanthomonas* sind bakterielle Proteine, die im pflanzlichen Zellkern als Transkriptionsaktivatoren wirken. Die spezifische Bindung an Ziel-DNA-Sequenzen wird durch die RVD-Abfolge in der hochkonservierten zentralen *repeat*-Domäne vermittelt. Im Rahmen dieser Arbeit wurden die DNA-Spezifitäten von RVDs bestimmt die auch im *R. solanacearum* TALE-Homolog Brg11 vorkommen. Dies ermöglichte erstmals die Vorhersage einer Ziel-DNA-Sequenz für Brg11. Zusätzlich wurden RVDs in stark und schwach eingeteilt, eine Eigenschaft die vorher unbekannt war und die die Gesamtaktivität von TALEs maßgeblich beeinflusst. TALE-abhängig aktivierte Gene, welche die Virulenz des Pathogens fördern, werden als Suszeptibilitätsgene bezeichnet. Besonders in der *Xoo*-Reis Interaktion ist bekannt, dass die TALE-abhängige Aktivierung von drei funktional redundanten *OsSWEETs* in Reis unerlässlich für die *Xoo*-Virulenz ist. Im Rahmen dieser Arbeit wurden zwei weitere nah verwandte *OsSWEETs* identifiziert, die ebenfalls die Virulenz von *Xoo* fördern. Interessanterweise bilden diese insgesamt fünf *OsSWEETs* eine funktional distinkte Gruppe, die Saccharose-Transporter kodiert. Die TALE-abhängige Aktivierung dieser Saccharose-Transporter ermöglicht vermutlich die Ernährung des Pathogens und fördert damit die Virulenz von *Xoo*. Eine Resistenz gegen die TALE-abhängige Aktivierung von Suszeptibilitätsgenen kann bereits durch Einzel-Nukleotid Deletionen oder Insertionen in der Bindestelle vermittelt werden. Diese verschieben den Leserahmen der Ziel-DNA-Sequenz und resultieren in der Anhäufung von RVD-Nukleotid-Fehlpaarungen, die letztendlich dazu führen, dass der TALE nicht mehr bindet. Interessanterweise konnte in dieser Arbeit gezeigt werden, dass die Insertion eines aberranten *repeats* der im Vergleich zum 34-Aminosäure Standard-*repeat* Deletionen bzw. Insertionen in den *repeat- α* -Helices aufweist, die starre ein-*repeat*-ein-Nukleotid Bindung des TALEs durchbricht. Dies erlaubt die flexible Bindung des TALEs an optimale Sequenzen ebenso wie an -1 *frameshift* Sequenzen, die in resistenten Allelen vorkommen. Aberrante *repeats* könnten somit eine evolutionäre Anpassung an solche Resistenzen darstellen. TALEs werden aufgrund ihrer modularen und programmierbaren *repeat*-Region als biotechnologische Werkzeuge zur gezielten Regulation von Genen verwendet. In dieser Arbeit konnte jedoch beobachtet werden, dass die Effizienz der TALE-abhängigen Zielgenaktivierung positionsabhängig ist. Diese Daten deuten darauf hin, dass TALEs vermutlich nicht wie klassische *Enhancer*-bindende Proteine wirken, sondern nur in bestimmten Bereichen des Promotors positioniert werden können.

Summary

Xanthomonas TALEs are bacterial proteins which act as transcriptional activators in the nucleus of plant cells. The specific binding of TALEs to a consecutive DNA sequence is mediated by the RVD sequence in the highly conserved central repeat region. In this study the nucleotide specificities of certain rare RVDs were determined that are also common in the *R. solanacearum* TALE homolog Brg11. These new RVD specificities were used to predict a potential Brg11 target sequence for the first time. Additionally a new feature of TALE RVDs was observed that has a major impact on TALE activity and allows the classification of RVDs as strong and weak. TALE activated genes that contribute to virulence are also called susceptibility genes. Especially in the *Xoo*-rice interaction the TALE dependent activation of three functional redundant rice *OsSWEETs* is known to be essential for the virulence of *Xoo*. In this study two additional closely related *OsSWEETs* were identified that also support the virulence of *Xoo*. Interestingly these five out of twenty rice *OsSWEETs* form a distinct functional protein group whose special feature might be the transport of sucrose. The TALE dependent activation of these sucrose transporters probably facilitates the nutrition of the pathogen, promotes multiplication and thus virulence of *Xoo*. One possible mechanism to mediate resistance against the TALE dependent activation of susceptibility genes is the insertion or deletion of a single nucleotide in the TALE target sequence. This results in a frameshift in the TALE target sequence which increases the RVD-nucleotide mismatches and thereby prevents TALE binding. This study showed that the insertion of a so called aberrant repeat which contains a deletion or duplication in one of the repeat- α -helices in comparison to a standard 34-amino acid repeat breaks the strict one-to-one binding mode of TALEs to tolerate such frameshift sequences. Thereby aberrant repeats allow the flexible binding of either an optimal target sequence or an otherwise resistant -1 frameshift allele by looping the aberrant repeat out of the repeat array. Hence aberrant repeats might resemble an evolutionary adaptation to such resistances. The modular and programmable DNA-binding region of TALEs promotes their use as biotechnology tool for the targeted regulation of genes. This study showed that the efficiency of the TALE mediated gene activation is dependent on their position in a promoter. Additionally TALEs influence the transcriptional start site. These findings suggest that TALEs do not function comparable to classical enhancer binding proteins and should therefore only be positioned in certain promoter regions.

Danksagung

Hiermit bedanke ich mich bei allen Menschen, die mich auf dem Weg zur Erstellung dieser Arbeit unterstützt haben.

Besonderer Dank gilt Frau Prof. Dr. Ulla Bonas, für die Möglichkeit meine Doktorarbeit in ihrer Arbeitsgruppe anzufertigen. Außerdem bedanke ich mich bei Herrn PD Dr. Jens Boch für die Möglichkeit dieses spannende Thema zu bearbeiten, sowie für seine Betreuung und ständige Diskussions- und Hilfsbereitschaft.

Bei der gesamten Arbeitsgruppe bedanke ich mich für die gute Zusammenarbeit und ein angenehmes Arbeitsklima. Weiterhin bedanke ich mich bei Carola, Marina, Hannelore und Bianca ohne die ein zügiges und effektives Arbeiten nicht möglich gewesen wäre.

Annekatriin, Carola, Maik und Sebastian danke ich besonders für die immer wieder anregenden und lustigen Diskussionen über wissenschaftliche und alltägliche Dinge.

Tom und Anika danke ich für Alles.

Außerdem danke ich meiner Familie, die mich in jeglicher Hinsicht auf meinem Weg unterstützt hat.

Inhaltsverzeichnis

Publikationsliste	I
Zusammenfassung	II
Summary	III
Danksagung	IV
Inhaltsverzeichnis	V
Abbildungsverzeichnis	VIII
Abkürzungsverzeichnis	IX
1. Einleitung	1
1.1 Pflanzenpathogene Bakterien und ihr Leben mit dem Wirt.....	1
1.2 Die Familie der <i>Transcription activator-like effectors</i> (TALEs).....	4
1.2.1 Der Aufbau von TALEs.....	5
1.2.2 Die Entschlüsselung des TALE RVD-DNA-Codes.....	6
1.2.2.1 Der TALE RVD-DNA-Code eröffnet neue Möglichkeiten.....	8
1.2.3 Die Struktur der TALE <i>repeat</i> -Region.....	10
1.2.4 Besondere <i>repeat</i> -Varianten.....	11
1.2.5 TALEs als Virulenzfaktoren.....	13
1.2.5.1 Saccharose-Transporter, Methyltransferasen und Transkriptionsfaktoren sind wichtige Suszeptibilitätsgene für <i>Xoo</i> -TALEs.....	14
1.2.5.2 Weitere <i>Xanthomonas spp.</i> aktivieren ebenfalls Transkriptionsfaktoren.....	16
1.2.6 TALEs als Avirulenzfaktoren.....	17
1.2.6.1 Dominante Resistenzgene.....	17
1.2.6.2 Rezessive Resistenzgene.....	19
1.2.7 TALEs als Transkriptionsaktivatoren.....	21
1.3 Zielstellung.....	23
2. Ergebnisse	25
2.1 Die Erweiterung des TALE-RVD-Codes.....	25
2.1.1 Publikation 1.....	25
2.1.1.1 Anlagen zur Publikation 1.....	28
2.1.1.2 Zusammenfassung Publikation 1.....	42
2.2 Gruppe III <i>OsSWEETs</i> sind wichtige Suszeptibilitätsgene für <i>Xoo</i> -TALEs.....	43
2.2.1 Publikation 2.....	43
2.2.1.1 Anlagen zu Publikation 2.....	55
2.2.1.2 Zusammenfassung Publikation 2.....	62

2.3 Aberrante <i>repeats</i> erlauben die Bindung allelischer Sequenzen mit -1 <i>frameshift</i>	63
2.3.1 Publikation 3.....	63
2.3.1.1 Anlagen zu Publikation 3	73
2.3.1.2 Zusammenfassung Publikation 3.....	97
2.4 Eigenanteil an den Publikationen.....	98
2.5 Zusätzliche Ergebnisse.....	99
2.5.1 Funktionsanalyse von TALEs am <i>OsSWEET14</i> -Promotor.....	99
2.5.2 Funktionsanalyse der <i>OsSWEET14</i> -TALEs im transienten System	103
2.5.3 Funktionsanalyse der <i>OsSWEET14</i> -TALEs im natürlichen Reis-System	107
2.5.4 Analyse des Einflusses von TALEs auf die Transkriptionsinitiation	111
3 Diskussion	116
3.1 Der erste <i>Ralstonia</i> -Code: Vorhersage einer Ziel-DNA-Sequenz für Brg11	116
3.2 NH - ein neuer Guanin-spezifischer RVD	118
3.3 Der RVD N* ermöglicht TALEs methylierte Ziel-DNA-Sequenzen zu binden ..	119
3.4 Effizienz - eine neue Eigenschaft für TALE RVDs.....	119
3.5 Die Entschlüsselung aller 400 möglichen RVD-DNA-Spezifitäten	124
3.6 Vom <i>MtN3</i> zum <i>SWEET</i> - die süße Verwandlung.....	126
3.6.1 Die Gruppe III SWEETs sind Saccharose-Transporter	127
3.6.2 <i>Xoo</i> manipuliert vermutlich die pflanzlichen Saccharose Verteilungswege ...	128
3.6.3 <i>Xoo</i> manipuliert vermutlich Zellen des Phloemparenchyms	131
3.6.4 SWEETs und deren Bedeutung für andere Pathogene.....	132
3.6.5 <i>Xoo</i> ist auf die TALE-abhängige Aktivierung von <i>OsSWEET14</i> fokussiert ...	133
3.7 Aberrante <i>repeats</i> - eine evolutionäre Anpassung an resistente Allele?	135
3.8.1 TALEs können die Transkription von verschiedenen Positionen aktivieren ..	137
3.8.2 TALEs verschieben den natürlichen Transkriptionsstart im <i>OsSWEET14</i> Promotor.....	143
3.8.3 Aktivatorproteine können mit der Transkriptionsmaschinerie interagieren	144
3.8.4 TALEs sind keine klassischen <i>Enhancer</i> -bindenden Proteine.....	148
4. Literaturverzeichnis	151
5. Anhang	169
5.1 Übersicht über die mittels PlantCARE identifizierten <i>cis</i> -regulatorischen DNA- Elemente im 1kb-Fragment des <i>OsSWEET14</i> -Promotors.....	169

5.2 Übersicht der vorhergesagten Bindestellen für DNA-bindende Proteine aus <i>Arabidopsis</i>	171
5.3 Übersicht über alle in dieser Arbeit genannten oder verwendeten Bindestellen natürlicher bzw. künstlicher TALEs im 1kb-Fragment des <i>OsSWEET14</i> -Promotors.....	173
5.4 Übersicht über die RVD-Komposition und Ziel-DNA-Sequenz der in dieser Studie verwendeten künstlichen TALEs.....	174
5.5 Phänotypen der BAI3 Δ <i>talC</i> -Stämme mit oder ohne künstlichen TALE nach Entfärben.....	175
5.6 Alignment der Hax3 <i>repeats</i>	176
Erklärung	177
Lebenslauf	178

Abbildungsverzeichnis

Abbildung 1.1	Aufbau und Funktion von TALEs aus <i>Xanthomonas spp.</i>	8
Abbildung 1.2	Verschiedene <i>repeat</i> -Varianten in <i>Xanthomonas</i> -TALEs.....	12
Abbildung 1.3	TALEs und deren Zielgene.....	17
Abbildung 1.4	TALE vermittelte Resistenzmechanismen.....	21
Abbildung 2.1	TALEs am <i>OsSWEET14</i> -Promotor und deren Aktivität im transienten System.....	103
Abbildung 2.2	Nachweis der Expression der künstlichen TALEs <i>in planta</i>	104
Abbildung 2.3	Analyse der TAL18- und TAL3-Aktivität am 74bp <i>Bs4</i> -Promotor im transienten System.....	105
Abbildung 2.4	Analyse der TAL2- und TAL-2-Aktivität am 74bp <i>Bs4</i> -Promotor im transienten System.....	107
Abbildung 2.5	Funktionsanalyse ausgewählter <i>OsSWEET14</i> -TALEs im natürlichen Reis- <i>Xoo</i> -System.....	110
Abbildung 2.6	Transkriptionsstartpunkte in Anwesenheit verschiedener TALEs in <i>N.benthamiana</i>	114
Abbildung 3.1	Vergleich der in meiner Diplomarbeit und de Lange <i>et al.</i> , 2013 verwendeten Brg11 Reporterkonstrukte.....	117
Abbildung 3.2	Die natürlichen Saccharose Verteilungswege in Reis werden von <i>Xoo</i> manipuliert.....	131
Abbildung 3.3	Modell des Einflusses der TALE-Orientierung auf den PIC.....	142
Abbildung 5.1	<i>Cis</i> -regulatorische Motive im 1kb-Fragment des <i>OsSWEET14</i> -Promotors.....	169
Abbildung 5.2	Vorhersage möglicher Bindestellen für DNA-bindende Proteine aus <i>Arabidopsis</i>	171
Abbildung 5.3	Übersicht über alle genannten oder verwendeten TALEs im 1 kb Fragment des <i>OsSWEET14</i> -Promotors.....	173
Abbildung 5.4	Phänotypen der BAI3 Δ <i>talC</i> -Stämme mit oder ohne künstlichen TALE nach Entfärben.....	175
Abbildung 5.5	Alignment der Hax3 <i>repeats</i>	176

Abkürzungsverzeichnis

A	Adenin (Nukleotid)
A	Alanin (Aminosäure)
Abb.	Abbildung
AD	Aktivierungsdomäne
Avr	Avirulenz
bp	Basenpaare
bzw.	beziehungsweise
cDNA	<i>complementary DNA</i> , komplementäre DNA
C	Cytosin
D	Aspartat, Asparaginsäure (Aminosäure)
DNA	<i>deoxyribonucleic acid</i> , Desoxyribonukleinsäure
ECW	<i>Early Californian Wonder</i> , Kultivar von <i>Capsicum annuum</i>
ev	<i>empty vector</i> , Leervektor
FMO	Flavin-abhängige Monooxygenase
G	Glycin (Aminosäure)
G	Guanin (Nukleotid)
GFP	<i>green fluorescent protein</i> , grün-fluoreszierendes Protein
GUS	β -Glucuronidase
H	Histidin (Aminosäure)
HR	hypersensitive Reaktion
I	Isoleucin (Aminosäure)
K	Lysin (Aminosäure)
kb	Kilobasen
kDa	kilo Dalton
LRR	<i>leucine rich repeat</i>
mRNA	<i>messenger RNA</i>
MU	Methylumbelliferon
MUG	Methylumbelliferyl- β -D-glucuronid
N	Asparagin (Aminosäure)
N	beliebiges Nukleotid
NLS	<i>nuclear localisation signal</i> , Kernlokalisierungssignal
OD	optische Dichte
P	Prolin (Aminosäure)
pv.	Pathovar
PAMP	<i>pathogen associated molecular pattern</i>
PCR	<i>polymerase chain reaction</i> , Polymerase-Kettenreaktion
qRT-PCR	<i>quantitative reverse transcription-PCR</i>
R-Gen	Resistenzgen
RVD	<i>repeat variable diresidue</i>
S	Serin (Aminosäure)
spp.	<i>species pluralis</i> , Spezies
T	Threonin (Aminosäure)

T	Thymin (Nukleotid)
T3S	Typ-III-Sekretionssystem
TSS	<i>transcription start site</i> , Transkriptionsstartposition
TALE	<i>transcription activator-like effector</i>
UPA	<i>upregulated by Avrbs3</i>
UTR	untranslatierte Region
<i>Xc</i>	<i>Xanthomonas citri</i>
<i>Xac</i>	<i>Xanthomonas axonopodis</i> pv. <i>citri</i>
<i>Xam</i>	<i>Xanthomonas axonopodis</i> pv. <i>manihotis</i>
<i>Xcc</i>	<i>Xanthomonas campestris</i> pv. <i>campestris</i>
<i>Xcv</i>	<i>Xanthomonas campestris</i> pv. <i>vesicatoria</i>
<i>Xoc</i>	<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>
<i>Xoo</i>	<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>
z.B.	zum Beispiel

1. Einleitung

Seit der ersten Beschreibung von Bakterien in Teichwasser, Regenwasser und menschlichem Speichel vor mehr als 300 Jahren durch Antoni van Leeuwenhoek ist mittlerweile gewiss: Bakterien sind allgegenwärtig und besiedeln auch extremste Lebensräume (van Leeuwenhoek 1677). Beim Zusammenleben mit Pflanzen und Tieren bieten bakterielle Symbionten dem Wirt viele Vorteile beispielsweise die Versorgung von Pflanzen mit Stickstoff oder den effektiven Aufschluss der Nahrung im tierischen Magen. Seit der Begründung der Bakteriologie ist anhand humaner Infektionskrankheiten jedoch deutlich geworden, dass nicht jede Interaktion mit Bakterien für den Wirt von Vorteil ist. Bakterien, die einen Wirt befallen und ihn zum eigenen Vorteil schädigen, lösen Krankheitssymptome aus und werden als Pathogene bezeichnet. Pathogene Bakterien befallen nicht nur Tiere, sondern auch Pflanzen. Eine der ersten wissenschaftlichen Beschreibungen pflanzenpathogener Bakterien wurde von T. J. Burrill 1878 anhand des Bakteriums *Erwynia amylovora* als Auslöser des Feuerbrandes auf Kernobstgewächsen veröffentlicht (Burrill 1878, 1881). Seitdem wurden verschiedenste bakterielle Pflanzenkrankheiten beschrieben, deren Schadbilder von Flecken über Tumore, Gallen und Pusteln bis hin zur Fäule und Welke von Blättern, Stämmen und Früchten reichen und im Falle einer Infektion von Kulturpflanzen meist zu Ernteaussfällen führen. Die Analyse pflanzenpathogener Bakterien und ihrer Interaktion mit dem Wirt ist daher notwendig, um Infektionsmechanismen besser zu verstehen und daraus Ansätze zur Eindämmung von Infektionen abzuleiten oder Resistenzen zu entwickeln.

1.1 Pflanzenpathogene Bakterien und ihr Leben mit dem Wirt

Pflanzenpathogene Bakterien befallen Wirtspflanzen und manipulieren sie zu ihren eigenen Gunsten. Ziel dieser Manipulation ist es, einen möglichst geeigneten Lebensraum zu schaffen, die Versorgung mit Nährstoffen zu garantieren und damit die Vermehrung und letztendlich die Verbreitung der Bakterien zu ermöglichen. Als Anpassung an die Gefahr bakterieller Infektionen haben Pflanzen Abwehrmechanismen entwickelt, die auf verschiedenste Art und Weise die Ausbreitung der Bakterien in der Pflanze verhindern. Nicht jede Pflanze ist *per se* anfällig für eine Infektion mit bestimmten Pflanzenpathogenen. Grund dafür ist die Nicht-Wirts-Resistenz, welche die Resistenz einer Pflanzenspezies gegenüber allen Isolaten einer Pathogenart beschreibt (Heath 2000). Sie wird durch präformierte Hindernisse (z.B. Zellwand, Kutikula oder

antimikrobielle Substanzen) sowie durch die aktive Einleitung von Abwehrreaktionen vermittelt (Heath 2000; Hückelhoven 2007; Senthil-Kumar & Mysore 2013). Sind Pathogene in der Lage die Nicht-Wirts-Resistenz zu überwinden, wird die besiedelte Pflanze zu einem Wirt und somit suszeptibel. Pflanzen haben spezielle Mechanismen entwickelt, um Pathogene zu erkennen und Abwehrreaktionen einzuleiten, die eine weitere Besiedlung verhindern. Die Erkennung von Nicht-Wirts- und Wirtspathogenen verläuft ähnlich, allerdings können sich die aktivierten Abwehrsignalwege und die Intensität der Abwehrreaktionen unterscheiden (Senthil-Kumar & Mysore 2013). Die Erkennung von Pflanzenpathogenen und die nachfolgende Einleitung von Abwehrmechanismen erfolgt nicht über spezialisierte Immunzellen wie im tierischen Immunsystem, sondern an jeder einzelnen Pflanzenzelle über zwei mögliche Mechanismen, die PTI und die ETI (Nürnberger, *et al.*, 2004; Jones & Dangl 2006). In der PTI (**PAMP triggered immunity**) werden bestimmte bakterielle Strukturen, sogenannte PAMPs (**pathogen associated molecular patterns**), durch Oberflächenrezeptoren, die PRRs (**pattern recognition receptors**), erkannt (Nürnberger, *et al.*, 2004). PAMPs sind meist hochkonservierte bakterielle Komponenten, die essentiell für das Pathogen sind, wie z.B. Bestandteile des Flagellums (Nürnberger, *et al.*, 2004). PRRs sind häufig Rezeptor-ähnliche Kinasen mit einer extrazellulären LRR- (**leucine rich repeat**), einer Transmembran- und einer intrazellulären Kinase-Domäne, die nach Erkennung der PAMPs Signalkaskaden zur Einleitung von Abwehrmechanismen aktiviert (Nürnberger, *et al.*, 2004; Senthil-Kumar & Mysore 2013). Die Abwehrreaktionen die nach Erkennung von Pathogenen im Rahmen der PTI ausgelöst werden, resultieren beispielsweise in der Verdickung der Zellwand durch Callose-Einlagerung oder im Schluss der Spaltöffnungen, wodurch das weitere Eindringen der Pathogene verhindert wird (Jones & Dangl 2006; Boller & He 2009; Dodds & Rathjen 2010; Senthil-Kumar & Mysore 2013). Pathogene, die an die Besiedlung einer Wirtspflanze angepasst sind, können die PTI überwinden, indem sie bakterielle Virulenzfaktoren in die Wirtszelle einschleusen. Gram-negative Bakterien besitzen dafür das sogenannte Typ-III-Sekretionssystem (Büttner & Bonas 2002). Dieses ist in der Lage bakterielle Typ-III-Effektorproteine direkt in das Cytosol der pflanzlichen Zelle zu translozieren (Büttner & Bonas 2002). Typ-III-Effektorproteine greifen in diverse zelluläre Abläufe ein und unterdrücken unter anderem die PTI (Alfano & Collmer 2004). Die Pflanze wird damit erneut suszeptibel und kann weiter durch Bakterien besiedelt werden. Pflanzen besitzen jedoch einen zweiten

Erkennungsmechanismus, die sogenannte ETI (*effector triggered immunity*; Jones & Dangl 2006). Die ETI basiert auf der Erkennung der bakteriellen Typ-III-Effektorproteine die auch Avirulenz-Proteine (Avr-Proteine) genannt werden, durch pflanzliche Resistenz-Gene (*R*-Gene) und deren Genprodukte (R-Proteine), anhand der Effektorstruktur oder durch die Bewachung eines Effektor-Interaktionspartners (*Guard Hypothese*; Van der Biezen & Jones 1998; Jones & Dangl 2006; Block, *et al.*, 2008). Die Abwehrreaktionen, die nach Erkennung von Pathogenen im Rahmen der ETI ausgelöst werden, resultieren in einem lokalen, auf die Infektionsstelle begrenzten Zelltod der sogenannten hypersensitiven Reaktion (HR), der die weitere Ausbreitung der Pathogene verhindert (Jones & Dangl 2006; Boller & He 2009; Dodds & Rathjen 2010; Senthil-Kumar & Mysore 2013). Im ständigen Wettrüsten mit der Wirtspflanze können sich Bakterien anpassen und Typ-III-Effektoren verändern oder verlieren, um so der Erkennung durch R-Proteine zu entgehen oder andere Typ-III-Effektoren entwickeln, um Mechanismen der ETI zu blockieren und so die Pflanze erneut zu besiedeln. Diese gegenseitige Anpassung zwischen Pflanze und Pathogen wird auch als Zick-Zack-Modell beschrieben (Jones & Dangl 2006).

Typ-III-Effektoren sind wichtige Virulenzfaktoren, die zu Gunsten des Pathogens wirken. Demzufolge ist die Charakterisierung von Typ-III-Effektoren wichtig, um deren Wirkungsweise in der Pflanze besser zu verstehen und daraus Möglichkeiten für Resistenzen abzuleiten. Zu den Modellorganismen, in denen die Funktion von Typ-III-Effektoren untersucht wird, gehören beispielsweise diverse *Xanthomonas*-Arten. Vertreter der Gattung *Xanthomonas* sind Gram-negative, stäbchenförmige γ -Proteobakterien, deren Hülle aus dem Exopolysaccharid Xanthan besteht (Swings, *et al.*, 1993; Becker, *et al.*, 1998). Verschiedene *Xanthomonas*-Arten wurden von internationalen Wissenschaftlern in die Top10 der pflanzenpathogenen Bakterien gewählt, da sie wichtige Kulturpflanzen wie Baumwolle, Zitruspflanzen, Maniok, Reis und andere Getreidearten befallen (Mansfield, *et al.*, 2012). Aufgrund ihres breiten Wirtsspektrums werden sie in Pathovaren (pv.) eingeteilt. Beispiele hierfür sind die reispfpathogenen Arten *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), der Auslöser der Weißblättrigkeit und *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) der Auslöser der bakteriellen Streifenkrankheit (Nino-Liu, *et al.*, 2006). *Xanthomonas axonopodis* pv. *citri* (*Xac*)-Stämme befallen Zitruspflanzen und lösen dort den sogenannten Zitruskrebs aus, während *Xanthomonas campestris* pv. *malvacearum* (*Xcm*) als Erreger der eckigen

Blattfleckenkrankheit von Baumwolle bekannt ist (Yang, *et al.*, 1996; Brunings & Gabriel 2003).

1.2 Die Familie der *Transcription activator-like effectors* (TALEs)

Pflanzenpathogene der Gattung *Xanthomonas* kodieren ca. 20 bis 40 verschiedene Typ-III-Effektorproteine (<http://www.xanthomonas.org/t3e.html>). Eine Gruppe solcher Typ-III-Effektorproteine, deren Mitglieder pflanzliche Gene aktivieren, umfasst die *transcription activator-like effectors* (TALEs; Yang, *et al.*, 2006). Diverse *Xanthomonas*-Stämme, die an verschiedenste Wirtspflanzen angepasst sind, besitzen *TALE*-Gene, wobei die Anzahl von 0 bis zu mehr als 20 verschiedener *TALE*-Gene pro Stamm variiert (Scholze & Boch 2011; Schornack, *et al.*, 2013). Der Archetyp der *TALE*-Familie ist AvrBs3 aus *Xanthomonas campestris* pv. *euvesicatoria* (*Xcv*), der aufgrund seiner Erkennung durch ein *R*-Gen in resistenten Paprikapflanzen isoliert wurde (Bonas, *et al.*, 1989). Basierend auf dieser ersten Beschreibung wurden weitere Mitglieder der *TALE*-Familie in verschiedenen Stämmen z.B. *Xcm*, *Xac* und vor allem *Xoo* identifiziert (Swarup, *et al.*, 1991; Yang, *et al.*, 1996; Bai, *et al.*, 2000; Yang, *et al.*, 2006; Antony, *et al.*, 2010). Anhand dieser *TALE*s, zu denen AvrB6 (*Xcm*), PthA (*Xac*), sowie AvrXa7 (*Xoo*) und PthXo1 (*Xoo*) zählen, konnte gezeigt werden, dass *TALE*s entscheidend zur Virulenz des jeweiligen Stammes beitragen (Swarup, *et al.*, 1991; Yang, *et al.*, 1996; Bai, *et al.*, 2000; Yang, *et al.*, 2006; Antony, *et al.*, 2010). Mittlerweile sind mehr als 100 Mitglieder der *TALE*-Familie aus verschiedensten *Xanthomonas*-Stämmen bekannt, wobei bisher jedoch nur für wenige Beispiele zugehörige Ziel-Gene beschrieben wurden (Boch & Bonas 2010). *TALE*-ähnliche Proteine wurden ebenfalls in anderen Bakteriengattungen identifiziert, zu denen pflanzenpathogene *Ralstonia solanacearum*-Stämme, sowie der obligate Pilz-endosymbiont *Burkholderia rhizoxinica* gehören (Cunnac, *et al.*, 2004; Heuer, *et al.*, 2007; Lackner, *et al.*, 2011). *TALE*-Homologe aus diesen Spezies sind analog zu *Xanthomonas*-*TALE*s aufgebaut, unterscheiden sich jedoch stärker untereinander (Heuer, *et al.*, 2007). Zudem war zu Beginn dieser Arbeit noch nicht bekannt, ob diese *TALE*-Homologen ebenfalls DNA binden und Gene aktivieren. Die nähere Charakterisierung der *TALE*-Familie in *Xanthomonas*, sowie auch deren Homologe aus anderen Bakteriengattungen könnte zukünftig Aufschluss über aktivierte Ziel-Gene und deren Funktion in der Pathogen-Wirtspflanzen-Interaktion geben.

1.2.1 Der Aufbau von TALEs

Vergleiche der Proteinsequenz verschiedener TALEs aus *Xanthomonas* zeigen eine hohe Konservierung zwischen 80 und 97% (Schornack, *et al.*, 2006). Anhand spezieller Charakteristika in ihrer Aminosäuresequenz lassen sich TALE-Proteine in drei Bereiche untergliedern, die zusammen für die Aktivität des TALEs verantwortlich sind. Im N-terminalen Bereich, der ca. 283 bis 290 Aminosäuren umfasst, befindet sich das Typ-III-Sekretions- und Translokationssignal, welches für die Translokation der TALEs in die pflanzliche Wirtszelle notwendig ist (Szurek, *et al.*, 2002; Boch & Bonas 2010; Abb. 1.1 A/B). Im C-terminalen Bereich, der ca. 274 bis 297 Aminosäuren beinhaltet, befinden sich drei Kernlokalisierungssignale (NLS), von denen zwei als funktional bestätigt wurden (Van den Ackerveken, *et al.*, 1996; Yang, *et al.*, 2000; Boch & Bonas 2010; Abb. 1.1 A/B). Mit Hilfe der NLS interagieren TALEs mit Importin- α , das zusammen mit Importin- β den Transport in den pflanzlichen Zellkern vermittelt (Szurek, *et al.*, 2001; Abb. 1.1 A/B). Zudem besitzen TALEs im C-terminalen Bereich eine saure Aktivierungsdomäne (AD) (Van den Ackerveken, *et al.*, 1996; Zhu, *et al.*, 1998; Yang, *et al.*, 2000; Abb. 1.1 A/B). Die NLS und die AD sind Charakteristika eukaryotischer Transkriptionsfaktoren. Die Deletion dieser Motive führt zum Verlust der Kernlokalisierung und daraus resultierend der Genaktivierungsfunktion von TALEs (Zhu, *et al.*, 1998; Yang, *et al.*, 2000; Szurek, *et al.*, 2001). Diese Erkenntnisse ließen vermuten, dass TALEs bakterielle Proteine sind, die als Transkriptionsfaktoren im pflanzlichen Zellkern wirken, woraus sich die Namensgebung als *transcription activator-like effector* ergab (Yang, *et al.*, 2006). Der zentrale Bereich von TALEs, die sogenannte *repeat*-Region, vermittelt die spezifische Bindung an DNA-Sequenzen im Promotor von Ziel-Genen und ist daher ebenfalls notwendig für die Funktion von TALEs als Transkriptionsaktivatoren. (Boch & Bonas 2010; Abb. 1.1 A/B). Die *repeat*-Region ist aus Wiederholungen eines meist 34 Aminosäuren umfassenden Motivs (*repeat*) zusammengesetzt. Die Anzahl der Motiv-Wiederholungen variiert dabei von 1,5- bis 33,5-mal, wobei die meisten TALEs 17,5 *repeats* enthalten (Boch & Bonas 2010). Der letzte *repeat* der Region umfasst meist nur 19 bzw. 20 konservierte Aminosäuren und wird daher als halber *repeat* gekennzeichnet (Bonas, *et al.*, 1989; Boch & Bonas 2010; Abb. 1.2). Vergleiche der Aminosäuresequenzen der *repeats* zeigen, dass diese hochkonserviert sind, wobei einige Aminosäuren (Position G14 und L29) sogar absolut invariant sind (Bonas, *et al.*, 1993; Boch & Bonas 2010; Abb. 5.5). Erste Analysen ergaben, dass neben Position 4 und 32 vor allem die

Aminosäurepositionen 12 und 13 sehr variabel sind, weshalb diese auch als hypervariable Aminosäuren oder *repeat variable diresidue* (RVD) bezeichnet werden (Bonas, *et al.*, 1989; Schornack, *et al.*, 2006; Boch & Bonas 2010; Abb. 1.1 B & 5.5). Bereits kurz nach der ersten Beschreibung konnte der *repeat*-Region eine zentrale Rolle für die TALE-abhängige Aktivierung von Ziel-Genen zugeordnet werden. Beispielsweise führten Deletionen in der AvrBs3 *repeat*-Region zum Verlust der Erkennung durch das R-Gen *Bs3* in resistenten Paprikapflanzen (Herbers, *et al.*, 1992). Vergleichbare Studien, in denen die *repeat*-Region von AvrXa7 und AvrXa10 getauscht wurden, zeigten ebenfalls eine vertauschte Aktivierung der Ziel-Gene in anfälligen und resistenten Reispflanzen (Zhu, *et al.*, 1998; Yang, *et al.*, 2000). Zusätzlich zeigten *in vitro* durchgeführte DNA-Bindungsstudien, dass AvrXa7 an doppelsträngige DNA bindet und AvrBs3 mit Hilfe der *repeat*-Region spezifisch die Promotorsequenz der beiden aktivierten Gene *UPA20* und *Bs3* erkennt (Yang, *et al.*, 2000; Kay, *et al.*, 2007; Römer, *et al.*, 2007). Damit konnte bewiesen werden, dass TALEs ihre Ziel-DNA-Sequenzen ohne die Hilfe pflanzlicher Wirtsproteine binden. Diese Experimente brachten den letztendlichen Hinweis, dass TALEs spezifisch an DNA-Sequenzen im Promotor eines Ziel-Gens binden und nachfolgend dessen Transkription aktivieren, wobei die Spezifität für bestimmte Ziel-DNA-Sequenzen der Wirtspflanze von der *repeat*-Region kodiert wird (Yang, *et al.*, 2000; Kay, *et al.*, 2007; Römer, *et al.*, 2007; Abb1.1 A). Wie genau die *repeat*-Region die Bindung an die spezifische Ziel-DNA-Sequenz im Promotor pflanzlicher Gene diktiert war jedoch lange Zeit unklar.

1.2.2 Die Entschlüsselung des TALE RVD-DNA-Codes

In vitro DNA-Bindungsstudien zeigten die spezifische Bindung von AvrBs3 an ein 36bp umfassendes DNA-Fragment des *Upa20*-Promotors (*upa*: *upregulated by AvrBs3*; Kay, *et al.*, 2007). Der Vergleich dieses Promotorfragments mit anderen Promotoren AvrBs3-aktivierter Gene z.B. *Upa10* ergab eine besonders konservierte Region im Bereich der TATA-Box, die sogenannte *upa*-Box (Kay, *et al.*, 2007). Weitere Studien konnten die Notwendigkeit der *upa*-Box für die AvrBs3-abhängige Aktivierung des Ziel-Gens *Bs3* bestätigen und in dieser zusätzlich ein hoch-konserviertes Kernelement mit ca. 14 bis 15bp identifizieren (Römer, *et al.*, 2007). Aufgrund dieser Ergebnisse entstand erstmals die Idee, dass die Anzahl und Abfolge der *repeats* die gebundene Nukleotidsequenz bestimmt (Boch *et al.*, 2009). Die anschließende Projektion der RVD-Abfolge der AvrBs3-*repeats* auf die *upa*-Box ließ vermuten, dass ein *repeat* ein

Nukleotid bindet, wobei die Spezifität für ein bestimmtes Nukleotid vermutlich durch den RVD im *repeat* bestimmt wird (Boch, *et al.*, 2009). Der TALE AvrBs3 mit 17,5 *repeats* bindet damit eine Kernsequenz von 18bp, wobei z.B. der RVD NI (Asparagin, Isoleucin) häufig mit Adenin und HD (Histidin, Aspartat) häufig mit Cytosin in der Ziel-DNA-Sequenz korrelierte. Zur experimentellen Bestätigung dieser Hypothese wurden anhand der RVD-Abfolge verschiedener TALEs Ziel-DNA-Sequenzen für diese abgeleitet und vor einen basal nicht aktiven Promotor eingefügt. Die TALE-abhängige und spezifische Aktivierung der vorhergesagten Ziel-DNA-Sequenz wurde anschließend mit Hilfe eines Reportergens detektiert (Boch, *et al.*, 2009). Gleichzeitig anhand bekannter TALEs und deren Ziel-Gen-Promotoren durchgeführte bioinformatische Studien entdeckten ebenfalls die Korrelation der RVD-Abfolge mit bestimmten Nukleotiden in der Ziel-DNA-Sequenz (Boch, *et al.*, 2009; Moscou & Bogdanove 2009). Diese beiden Studien entschlüsselten den RVD-DNA-Code und zeigten erstmals die Nukleotid-Spezifität bestimmter RVDs in der *repeat*-Region. Interessanterweise sind einige RVDs spezifisch für ein bestimmtes Ziel-Nukleotid. Der RVD HD ist spezifisch für Cytosin, NI für Adenin und NG (Asparagin, Glycin) ist spezifisch für Thymin (Abb. 1.1 C). Andere RVDs erlauben hingegen mehrere Nukleotide in der Ziel-DNA Sequenz, so erkennt der RVD NN (Asparagin, Asparagin) die beiden Purine Guanin und Adenin und der RVD N* (welches im Vergleich zu den Aminosäuresequenzen anderer *repeats* an 12. Position Asparagin trägt, dem in der Konsensussequenz jedoch die Aminosäure an Position 13 fehlt; Abb. 1.2), erkennt laut Vorhersage die beiden Pyrimidine Cytosin und Thymin (Abb. 1.1 C). Der RVD NS (Asparagin, Serin) ist ein Ausnahmefall, der unspezifisch alle vier Nukleotide in der Ziel-DNA-Sequenz akzeptiert (Boch, *et al.*, 2009; Moscou & Bogdanove 2009; Abb. 1.1 C). Zusätzlich konnte in diesen sowie weiteren Studien festgestellt werden, dass der eigentlichen Ziel-DNA-Sequenz ein Thymin vorangestellt ist, welches für die TALE-abhängige Aktivierung notwendig ist (Boch, *et al.*, 2009; Moscou & Bogdanove 2009; Römer, *et al.*, 2009; Abb. 1.1 D). Welche TALE-Region jedoch für die Erkennung dieses initialen Thymins verantwortlich ist, war zu diesem Zeitpunkt nicht bekannt.

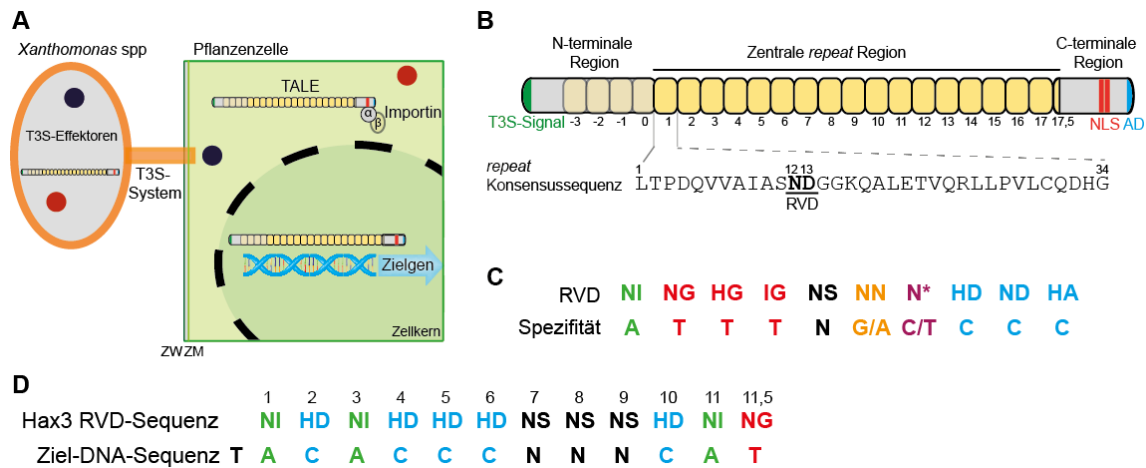


Abbildung 1.1 Aufbau und Funktion von TALEs aus *Xanthomonas spp.*

A Pflanzenpathogene *Xanthomonas spp.* translozieren TALEs mit Hilfe des Typ-III-Sekretionssystems (T3S-System) in die pflanzliche Wirtszelle. Durch die Interaktion mit Importin- α werden TALEs in den Zellkern transportiert. Dort binden sie spezifisch an Promotorsequenzen pflanzlicher Ziel-Gene und aktivieren deren Transkription. **B** TALEs enthalten im N-terminalen Bereich das Typ-III-Sekretions (T3S)-Signal sowie eine erweiterte *repeat*-Region, die zur generellen DNA-Bindung (-3 bis 0) und der Bindung des initialen Thymins (-1) beiträgt. Im C-terminalen Bereich befinden sich zwei funktionale Kernlokalisierungssignale (NLS), sowie eine saure Aktivierungsdomäne (AD). Der zentrale Bereich von TALEs ist aus Wiederholungen eines meist 34 Aminosäuren umfassenden Motivs aufgebaut und wird als *repeat*-Region bezeichnet. Die Aminosäuresequenz der *repeats* ist hochkonserviert und variiert hauptsächlich an Position 12 und 13 dem *repeat variable di-residue* (RVD). **C** Übersicht über die 2009 bekannten RVD-Spezifitäten (N=A,C,G,T). Die RVDs bestimmen welche Nukleotide in der Ziel-DNA-Sequenz gebunden werden. **D** Anhand der RVD-Abfolge im TALE (hier am Beispiel von Hax3, Abb. 5.5) können mögliche Ziel-DNA-Sequenzen abgeleitet werden. Üblicherweise ist der eigentlichen Ziel-DNA-Sequenz ein Thymin vorangestellt.

1.2.2.1 Der TALE RVD-DNA-Code eröffnet neue Möglichkeiten

Insgesamt wurden in *Xanthomonas*-TALEs 23 verschiedene RVDs identifiziert, von denen einige häufig vorkommen (z.B. HD), währenddessen andere eher selten sind (z.B. SN) (Boch & Bonas 2010). Der einfache ein-RVD-bindet-ein-Nukleotid-Code ermöglicht nun die Vorhersage von Ziel-DNA-Sequenzen verschiedener TALEs ausschließlich basierend auf ihrer RVD-Abfolge und erlaubt damit die weitere Identifizierung pflanzlicher Ziel-Gene (Boch *et al.*, 2009; Grau *et al.*, 2013; Perez-Quintero *et al.*, 2013; Abb. 1.1 **D**). Die meisten noch fehlenden RVD-Spezifitäten können anhand von TALEs mit bereits bekannten Ziel-Genen bioinformatisch abgeleitet werden (Moscou & Bogdanove 2009). Dies gilt jedoch nicht für RVDs z.B. HN, NH oder SN die bisher nur sehr selten in TALEs zu finden sind (Boch & Bonas 2010). Die experimentelle Analyse der Spezifität solcher RVDs kann also nützlich sein, um den TALE-RVD-Code zu vervollständigen. Interessanterweise wurden TALE-homologe Proteine auch in *R. solanacearum* beschrieben (Cunnac, *et al.*, 2004; Heuer, *et al.*, 2007). *R. solanacearum* TALE-Homologe z.B. Brg11 besitzen ebenfalls eine zentrale *repeat*-Region, deren Aminosäuresequenz zwar weniger stark konserviert ist, aber

dennoch eine hohe Ähnlichkeit zur *repeat*-Region von *Xanthomonas*-TALEs aufweist (Streubel 2011; de Lange, *et al.*, 2013). In dieser finden sich RVDs beispielsweise HD oder NG deren Spezifität in *Xanthomonas*-TALEs bereits analysiert wurde. Zusätzlich gibt es jedoch auch *Ralstonia*-spezifische RVDs z.B. SH, NP, IS deren Spezifität zu Beginn dieser Arbeit noch nicht bekannt war (Cunnac, *et al.*, 2004; Heuer, *et al.*, 2007). Das Vorhandensein einer zu *Xanthomonas*-TALEs ähnlichen *repeat*-Region und die bereits vor Beginn dieser Arbeit nachgewiesene Kernlokalisierung ließ vermuten, dass *Ralstonia* TALE-homologe ebenfalls spezifisch DNA binden (de Lange, *et al.*, 2013). Allerdings konnte zu diesem Zeitpunkt aufgrund der fehlenden RVD-Spezifitäten noch keine mögliche Ziel-DNA-Sequenz für einen *Ralstonia*-TALE abgeleitet werden, um dessen spezifische DNA-Bindung bzw. Genaktivierungsfunktion zu analysieren. In vorrangegangenen Studien konnte außerdem festgestellt werden, dass vereinzelt TALEs mit bestimmter RVD-Komposition trotz perfekter Ziel-DNA-Sequenz im Promotor eines Reportergens, dieses nicht aktivieren können. Dies ließ einen Einfluss der RVD-Komposition auf die Gesamtaktivität des TALEs vermuten, der bisher noch nicht beschrieben wurde (Boch & Landgraf, unveröffentlicht).

Außerdem eröffnet die Entschlüsselung des TALE-RVD-Codes neue Möglichkeiten zur Anwendung von TALEs als programmierbare DNA-bindende Proteine mit nutzerdefinierter Ziel-Sequenz. Die gezielte Regulation, Mutation oder Reparatur eines Gens ist Gegenstand der Forschung, um beispielsweise Genfunktionen zu analysieren oder aber mittels Gentherapie Krankheiten zu heilen (Carroll 2011). Um solche Modifikationsenzyme jedoch zur gewünschten DNA-Sequenz zu dirigieren, müssen sie an programmierbare DNA-bindende Proteine fusioniert werden. Lange Zeit waren Zink-Finger-Proteine (ZFPs), fusioniert an funktionale Domänen z.B. die Nukleasedomäne der *FokI*-Endonuklease die Grundlage zur Erstellung sequenzspezifischer Gen-Modifikationsenzyme (Klug 2010; Carroll 2011). Ein Zink-Finger bindet mit Hilfe der Aminosäuren an Position -1, 3 und 6 drei Basen in der Ziel-DNA-Sequenz (Klug 2010). Je mehr Zink-Finger kombiniert werden, desto schwieriger ist es deren Ziel-DNA-Sequenz zu bestimmen, weshalb die Variabilität der gebundenen Ziel-DNA-Sequenz häufig auf 9bp, gebunden durch drei Zink-Finger eingeschränkt ist (Klug 2010; Carroll 2011). Außerdem stellte sich heraus, dass die DNA-Spezifität eines Zink-Fingers nicht nur anhand der Aminosäuren an den obengenannten Positionen vorhersagbar ist, sondern durch zusätzliche Kontakte zu umliegenden DNA-Sequenzen sowie durch benachbarte Finger-Module beeinflusst wird (Klug 2010). Daher müssen

ZFPs häufig zeitaufwändige Selektionsschritte durchlaufen, um die spezifische Erkennung der gewünschten Ziel-DNA-Sequenz zu überprüfen (Klug 2010; Carroll 2011). Der modulare Aufbau von TALEs und speziell der einfache und vorhersagbare ein-RVD-ein-Nukleotid-Code hat die Erstellung maßgeschneiderter DNA-Modifikationsenzyme jedoch revolutioniert. Vergleichbar mit einem Baukasten können nun in wenigen Arbeitsschritten einzelne TALE-Module z.B. *repeats* mit gewünschter RVD-Abfolge zusammengesetzt werden, um eine neue DNA-Bindedomäne mit nutzerdefinierter Ziel-DNA-Sequenz zu erstellen (Boch, *et al.*, 2009; Geißler, *et al.*, 2011; Miller, *et al.*, 2011). Durch die Fusion dieser künstlichen DNA-Bindedomäne an andere funktionale Domänen, beispielsweise alternative ADs oder Nukleasen, können TALE-Derivate in pflanzlichen, sowie anderen Modellsystemen als künstliche Genregulatoren oder Gen-Modifikationsenzyme verwendet werden (Bogdanove & Voytas 2011). Daher könnten weiterführende Studien zu RVD-Spezifitäten und RVD-Eigenschaften auch die effiziente Anwendung von TALEs als biotechnologisches Werkzeug erleichtern.

1.2.3 Die Struktur der TALE *repeat*-Region

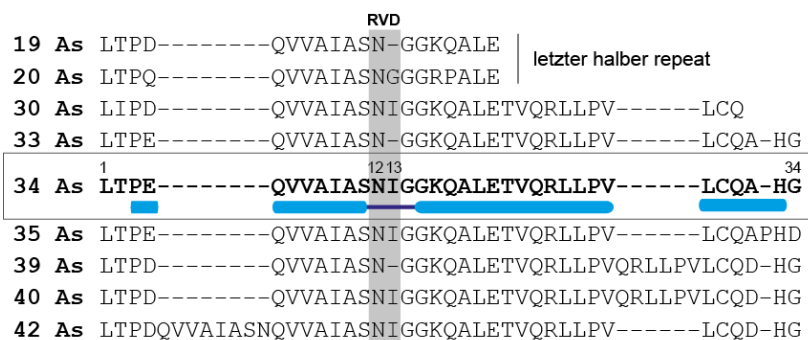
Zu Beginn dieser Arbeit war lediglich die Struktur eines 1,5 *repeats* umfassenden TALE-Fragments bekannt. (Murakami, *et al.*, 2010). Später konnte jedoch die vollständige Struktur der *repeat*-Region anhand eines künstlichen Hax3-basierten TALEs im Komplex mit seiner Ziel-DNA-Sequenz sowie des TALEs PthXo1 gezeigt werden (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Dabei konnte gezeigt werden, dass die *repeat*-Region als rechtswindende Superhelix den *sense*-Strang der DNA bindet, wobei die RVDs zur großen Furche der DNA-Doppelhelix ausgerichtet sind (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Außerdem konnte eine Kompression der TALE-Struktur im DNA-gebundenen-Komplex im Vergleich zum DNA-freien-Komplex beobachtet werden (Deng, *et al.*, 2012a). Die hohe Konservierung der *repeat*-Sequenz resultiert letztendlich darin, dass jeder *repeat* die gleiche Konformation einnimmt, wodurch die gesamte *repeat*-Region einen hohen Grad an struktureller Regelmäßigkeit aufweist (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Im Detail bildet ein *repeat* eine kürzere (Aminosäure 3 bis 11) und eine längere (Aminosäure 15 bis 33) α -Helix (Deng, *et al.*, 2012a). Die zwei α -Helices sind durch eine Schleife verbunden, die den RVD sowie das invariante Glycin an Position 14 beinhaltet, der sogenannte RVD-*loop* (Deng, *et al.*, 2012a; Abb. 1.2). Außerdem konnten anhand dieser Strukturen erstmals Unterschiede

für die Funktion der zwei Aminosäuren im RVD gezeigt werden. Die Aminosäure an Position 12 (die erste Aminosäure im RVD) trägt durch *intra-repeat*-Interaktionen zur Stabilität des RVD-loops bei und ist nicht an der Bindung des Ziel-Nukleotides beteiligt (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Ausschließlich die Aminosäure an Position 13 (die zweite Aminosäure im RVD) ist für die Bindung des Ziel-Nukleotides verantwortlich, wobei die spezifischen Interaktionen zwischen Aminosäure 13 und Nukleotid durch Wasserstoffbrücken oder van-der-Waals-Interaktionen vermittelt werden können (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Ausgehend davon konnte nun der RVD-Code verfeinert werden, sodass unabhängig von der Identität der Aminosäure an Position 12 beispielsweise ein D an Position 13, generell Cytosin spezifiziert oder ein N an Position 13 spezifisch für Adenin und Guanin ist. Zusätzlich konnten im angrenzenden N- und C-terminalen Bereich weitere degenerierte *repeats* identifiziert werden, die zwar anhand ihrer Aminosäuresequenz weniger konserviert sind, jedoch ebenfalls die typische *helix-loop-helix* Struktur aufweisen (Deng, *et al.*, 2012a; Gao, *et al.*, 2012; Mak, *et al.*, 2012; Abb. 1.1 **B**). Dabei konnte gezeigt werden, dass das Tryptophan (W232) im N-terminalen *repeat* -1 für die Koordination des initialen Thymins verantwortlich ist, welches den meisten TALE Ziel-DNA-Sequenzen vorangeht (Römer, *et al.*, 2009; Deng, *et al.*, 2012a; Gao, *et al.*, 2012; Mak, *et al.*, 2012; Abb. 1.1 **B, D**).

1.2.4 Besondere *repeat*-Varianten

Generell sind *repeats* aus *Xanthomonas*-TALEs hochkonserviert und umfassen meist 34 Aminosäuren. Der letzte halbe *repeat* gilt als reguläre Ausnahme zur Standard-*repeat*-Länge (Abb. 1.2). In diesem sind lediglich die ersten 20 Aminosäuren konserviert, wobei jedoch die folgenden nicht-konservierten Aminosäuren die typische *helix-loop-helix* Struktur dieses *repeats* vervollständigen, sodass keine Unregelmäßigkeiten in der Gesamtstruktur der *repeat*-Region entstehen (Boch & Bonas 2010; Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Eine weitere Abweichung ist in *repeats* zu finden, in denen die zweite Aminosäure im RVD fehlt (N*, H* und S*). Diese umfassen 33 bzw. 19 Aminosäuren, wodurch jedoch nicht die Funktion des *repeats* sondern lediglich die Spezifität des RVDs beeinflusst wird (Boch & Bonas 2010; Abb. 1.2). Eine weitere abweichende *repeat*-Länge ist in den TALEs Hax2 und AvrHah1 zu finden. Hax2 besteht ausschließlich aus 35-Aminosäure *repeats*, die ein zusätzliches Prolin an Position 33 enthalten (Kay, *et al.*, 2005; Abb. 1.2). AvrHah1 enthält eine Mischung aus solchen 35-, sowie 34-Aminosäure *repeats* (Schornack, *et al.*, 2008). Allerdings haben

auch diese Abweichungen in der *repeat*-Länge keinen Einfluss auf die DNA-Bindung oder die Aktivierungsfunktion dieser TALEs. (Schornack, *et al.*, 2008; Boch, *et al.*, 2009). Überraschenderweise finden sich im TALE-Repertoire von reispathogenen *Xoo*- und *Xoo*-Stämmen seltene TALEs, die in der Mitte der *repeat*-Region einen einzelnen *repeat* mit deutlich abweichender Länge enthalten (Yang, *et al.*, 2000; Boch & Bonas 2010). Diese *repeats* umfassen verursacht durch Deletionen bzw. Duplikationen in den *repeat- α* -Helices entweder 30, 39/40 oder sogar 42 Aminosäuren (Abb. 1.2). Aufgrund der hohen Konservierung der einzelnen *repeats* windet sich die gesamte *repeat*-Region eines TALEs hochgradig regelmäßig um die DNA (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Ausgehend davon ist es denkbar, dass solche einzelnen *repeats* mit abweichender Länge die hohe strukturelle Regelmäßigkeit der *repeat*-Region unterbrechen und so die DNA-Bindung des TALEs maßgeblich beeinflussen. Interessanterweise enthalten die *Xoo*-TALEs AvrXa7 und PthXo3 einen solchen *repeat* mit abweichender Länge, beide TALEs sind jedoch durch die Aktivierung des Ziel-Gens *OsSWEET14* unerlässlich für die Virulenz des zugehörigen *Xoo*-Stammes (Yang, *et al.*, 2000; Antony, *et al.*, 2010; s. Abschnitt 1.2.5.1). Dies weist darauf hin, dass sie trotz dieses *repeats* effizient Ziel-DNA-Sequenzen in Wirtsgenen binden können. Auffällig ist jedoch, dass für beide TALEs zwar eine Ziel-DNA-Sequenz im *OsSWEET14*-Promotor bekannt ist, diese jedoch in Vorhersageprogrammen nur einen sehr schlechten Rang erhält (Grau, *et al.*, 2013; Perez-Quintero, *et al.*, 2013). Grund dafür ist eine hohe Anzahl von RVD-Nukleotid-Fehlpaarungen, die sich vor allem in der Region nach dem *repeat* mit abweichender Länge befinden (Antony, *et al.*, 2010). Dies deutet darauf hin, dass ein solcher *repeat* die DNA-Bindung des TALEs beeinflusst, sodass dieser in der Lage ist die RVD-Nukleotid-Fehlpaarungen zu tolerieren und dennoch effizient das Ziel-Gen zu aktivieren. Wie genau *repeats* mit abweichender Länge die DNA-Bindung beeinflussen war jedoch zu Beginn dieser Arbeit unklar.



Beschriftung folgt auf der nächsten Seite

Abbildung 1.2 Verschiedene *repeat*-Varianten in *Xanthomonas*-TALEs

Die Aminosäuresequenz der verschiedenen *repeat*-Varianten ist im Ein-Buchstaben-Code angegeben. Die Konsensus-Sequenz der am häufigsten vorkommenden *repeat*-Länge mit 34 Aminosäuren ist fett und umrahmt hervorgehoben. Der RVD an Position 12 und 13 ist grau hinterlegt. Die helicalen Bereiche des 34-Aminosäure *repeats* sind in hellblau, der RVD-*loop* in dunkelblau markiert. Ausgehend von dieser Standard-Sequenz sind weitere *repeat*-Varianten mit Deletionen bzw. Duplikationen dargestellt.

1.2.5 TALEs als Virulenzfaktoren

Xanthomonas Stämme benutzen TALEs zur gezielten Manipulation des pflanzlichen Transkriptoms. Ziel-Gene die TALE-abhängig aktiviert werden und die Virulenz des zugehörigen *Xanthomonas*-Stammes fördern, werden als Suszeptibilitätsgene bezeichnet. Eines der am besten untersuchten Systeme in dem die TALE-abhängige Aktivierung von Ziel-Genen unerlässlich für die Virulenz des jeweiligen Stammes sein kann, ist das *Xoo*-Reis-System (White & Yang 2009). Reis ist eines der ökonomisch wichtigsten Nahrungsmittel, das seit mehreren Jahrtausenden kultiviert wird (Izawa & Shimamoto 1996; Nino-Liu, *et al.*, 2006). Aufgrund der Verfügbarkeit der Genomsequenz von zwei Vertretern der Arten *Oryza sativa* ssp. *japonica* und ssp. *indica* sowie Methoden zur Transformation und Erstellung transgener Pflanzen, ist Reis ein wichtiges molekularbiologisches Modellsystem für monokotyledone Pflanzen, um beispielsweise verschiedene Pathogen-Reis-Interaktionen zu analysieren (Izawa & Shimamoto 1996; Goff, *et al.*, 2002; Kennedy 2002; Yu, *et al.*, 2002; Nino-Liu, *et al.*, 2006). Reis-Pathogene haben sich vermutlich schon seit Jahrtausenden an die Besiedlung dieser Pflanze angepasst. Typische Symptome der bakteriellen Weißblättrigkeit auf Reis wurden erstmals 1884 beschrieben (Mew, *et al.*, 1993; Nino-Liu, *et al.*, 2006). Die Ernteverluste durch eine Infektion mit *Xoo* können bis zu 100% betragen, weshalb die *Xoo*-Reis Interaktion und besonders die Funktion von TALEs für die Virulenz von *Xoo* ein intensiv untersuchtes System darstellt (Mew, *et al.*, 1993; Nino-Liu, *et al.*, 2006). *Xoo* enthält bis zu 19 verschiedene *TALE*-Gene, deren Beitrag zur Virulenz des *Xoo*-Stammes unterschiedlich ist (Schornack, *et al.*, 2013). Es konnte bereits gezeigt werden, dass einzelne TALEs von so großer Bedeutung sind, dass deren Deletion zum kompletten Verlust der Virulenz führt (Bai, *et al.*, 2000; Yang, *et al.*, 2006; Antony, *et al.*, 2010; Yu, *et al.*, 2011). Im Gegensatz dazu haben andere einen moderaten Einfluss oder scheinen nur in bestimmten *Xoo*-Reis-Interaktionen eine Rolle zu spielen, währenddessen scheinen wieder andere TALEs gar keine Rolle zu spielen (Bai, *et al.*, 2000; Yang & White 2004; Yang, *et al.*, 2006; Sugio, *et al.*, 2007; Antony, *et al.*, 2010; Yu, *et al.*, 2011). Besonders für TALEs, welche die Virulenz des Pathogens

beeinflussen ist die Analyse der Ziel-Gene wichtig, um deren Funktion in der Pathogen-Wirt-Interaktion zu verstehen.

1.2.5.1 Saccharose-Transporter, Methyltransferasen und Transkriptionsfaktoren sind wichtige Suszeptibilitätsgene für *Xoo*-TALEs

Eine bedeutende Gruppe von TALE-Ziel-Genen in Reis sind die sogenannten *OsSWEETs*. Diese Suszeptibilitätsgene kodieren Saccharose-Transporter, deren TALE-abhängige Aktivierung zur Ernährung des Pathogens an der Infektionsstelle beitragen könnte (Sonnewald 2011; Baker, *et al.*, 2012; Braun 2012). *Xoo*-Stämme aus unterschiedlichsten Regionen (Afrika, Asien) haben sich auf die TALE-abhängige Aktivierung von Mitgliedern der *OsSWEET*-Familie spezialisiert, wodurch die Funktion dieser als Schlüssel-Suszeptibilitätsgene in der *Xoo*-Reis-Interaktion deutlich wird (Yang, *et al.*, 2006; Antony, *et al.*, 2010; Liu, *et al.*, 2011; Yu, *et al.*, 2011). *OsSWEET11* wird vom TALE PthXo1 aktiviert (Yang, *et al.*, 2006; Abb. 1.3). *OsSWEET13* wird vermutlich durch den TALE PthXo2 aktiviert und *OsSWEET14* wird sogar von drei verschiedenen TALEs (TalC, AvrXa7 und PthXo3) aktiviert (Antony, *et al.*, 2010; Liu, *et al.*, 2011; Yu, *et al.*, 2011; Abb. 1.3). Die Deletion dieser TALEs führt zum kompletten Verlust der Virulenz des jeweiligen *Xoo*-Stammes in anfälligen Reispflanzen (Bai, *et al.*, 2000; Yang & White 2004; Yang, *et al.*, 2006; Antony, *et al.*, 2010; Yu, *et al.*, 2011). Außerdem konnte gezeigt werden, dass allein der Transfer von TalC, AvrXa7 oder PthXo1 in einen TALE-defizienten und somit kaum virulenten *Xoo*-Stamm diesem ermöglicht typische Weißblättrigkeitssymptome in suszeptiblen Reispflanzen auszulösen (Verdier, *et al.*, 2012). Dies verdeutlicht die Bedeutung dieser TALEs und deren Ziel-Gene für die Virulenz der *Xoo*-Stämme.

Die *OsSWEET*-Familie in Reis besteht aus rund 20 Mitgliedern, wobei bereits drei dieser Mitglieder als Suszeptibilitätsgene für *Xoo* bekannt sind. *OsSWEETs* können aufgrund ihrer Aminosäuresequenz in verschiedene Gruppen unterteilt werden (Chen, *et al.*, 2010). Die drei bekannten Suszeptibilitätsgene *OsSWEET11*, *13* und *14* gehören zu Gruppe III, wobei für *OsSWEET11* und *14* bereits die Funktion als Saccharose-Transporter nachgewiesen wurde (Chen, *et al.*, 2010). Zu Beginn dieser Arbeit war jedoch unklar ob auch andere Mitglieder der *OsSWEET*-Familie, deren Proteine möglicherweise andere Zucker transportieren, Suszeptibilitätsgene für *Xoo*-TALEs sein können (Chen, *et al.*, 2010). Die Analyse aller *OsSWEETs* und ihrer Rolle für die

Virulenz von *Xoo* könnte somit Aufschluss über die Transportfunktion bestimmter Mitglieder und deren Bedeutung für das Pathogen geben.

Neben Saccharose-Transportern sind für *Xoo*-TALEs weitere Suszeptibilitätsgene bekannt. Eines dieser Gene ist *OsHen1*, das vermutlich durch PthXo8 aktiviert wird (Moscou & Bogdanove 2009; Abb. 1.3). *OsHen1* ist eine Methyltransferase, die die Stabilität kleiner RNAs beeinflusst. Die Auswirkung der *OsHen1* Aktivierung für die Virulenz von *Xoo* ist jedoch noch nicht bekannt. Interessanterweise ist *OsHen1* das bisher einzige Suszeptibilitätsgen, das nicht nur für einen *Xoo*-TALE sondern auch für den *Xoc*-TALE Tal1c beschrieben wurde (Abb. 1.3). *Xoc* ist ebenfalls ein reispathogenes Bakterium, das eine hohe Anzahl von bis zu 26 TALE-Genen aufweist (Schornack, *et al.*, 2013). Zu Beginn dieser Arbeit war jedoch über den Beitrag verschiedener TALEs zur Virulenz von *Xoc* nur wenig bekannt. Eine weitere Gruppe von Suszeptibilitätsgenen deren TALE-abhängige Aktivierung einen Einfluss auf die Virulenz der *Xoo*-Stämme in Reis hat, kodiert Transkriptionsfaktoren. Beispiele hierfür sind *OsTFXI* und *OsTFIIA γ 1* (Sugio, *et al.*, 2007). *OsTFXI* wird von PthXo6 aktiviert, der neben PXO99 auch in anderen *Xoo*-Stämmen verbreitet ist (Sugio, *et al.*, 2007; Abb. 1.3). Die Deletion von PthXo6 reduziert die Virulenz von PXO99, führt jedoch nicht zum kompletten Verlust dieser (Sugio, *et al.*, 2007). Dies weist darauf hin, dass die TALE-abhängige Aktivierung von *OsTFXI* positiv zur Virulenz von *Xoo* beiträgt, jedoch weniger bedeutend als die Aktivierung der *OsSWEETs* ist. Bisher ist allerdings über mögliche Sekundärziele und damit die Auswirkung von *OsTFXI* für die Virulenz von *Xoo* nichts bekannt. *OsTFIIA γ 1* ist ein Allel, welches eine γ -Untereinheit des Transkriptionsfaktors IIA (TFIIA) kodiert (Sugio, *et al.*, 2007; Iyer-Pascuzzi, *et al.*, 2008). Dieses Gen wird vom PXO99-TALE PthXo7 aktiviert und ist ein Beispiel für ein Suszeptibilitätsgen, welches nur für die Interaktion von *Xoo* mit bestimmten Reisarten von Bedeutung ist (Sugio, *et al.*, 2007; s. Abschnitt 1.2.6.2; Abb. 1.3 & 1.4). Im Gegensatz zu den bereits beschriebenen TALEs hat die Deletion anderer TALEs in *Xoo* lediglich geringe oder gar keine Auswirkung auf die Virulenz dieser Stämme (Bai, *et al.*, 2000; Yu, *et al.*, 2011). Dies könnte einerseits bedeuten, dass diese TALEs nur in bestimmten noch unbekanntem Interaktionen von Bedeutung sind oder generell eine untergeordnete Rolle für die Virulenz des Pathogens spielen. Interessanterweise kann die TALE-abhängige Aktivierung von Transkriptionsfaktoren nicht nur für *Xoo*-Stämme, sondern auch für andere *Xanthomonas*-Arten von Bedeutung sein.

1.2.5.2 Weitere *Xanthomonas spp.* aktivieren ebenfalls Transkriptionsfaktoren

Kürzlich konnte gezeigt werden, dass verschiedene TALEs der PthA-Familie und deren Homologe aus *Xanthomonas citri* (*Xc*)-Stämmen den Transkriptionsfaktor *CsLOB1* aktivieren (Hu, *et al.*, 2014; Pereira, *et al.*, 2014; Abb. 1.3). Die Identifizierung von PthA (*pathogenicity A*) als Pathogenitätsfaktor erfolgte durch die Übertragung von PthA in einen kaum virulenten *Xanthomonas citri* pv. *citrumelo*-Stamm, wodurch dieser in einen hoch-virulenten Stamm transformiert wurde, der typische Zitrus-Krebs-Symptome auslöst (Swarup, *et al.*, 1991). Die Deletion von *PthA* im *Xc*-Stamm 3213 führte zum kompletten Verlust der Virulenz dieses Stammes (Swarup, *et al.*, 1991). Weiterhin war allein die transiente Expression von PthA ausreichend, um die pflanzliche Zellteilung, Zellvergrößerung und letztendlich den Zelltod und somit typische Zitrus-Krebs-Symptome auszulösen (Duan, *et al.*, 1999). Dies zeigt, dass die Aktivierung von *CsLOB1* durch Mitglieder der PthA-Familie unerlässlich für die Virulenz des Pathogens ist und dieser Transkriptionsfaktor über bisher noch unbekannte Sekundärziele zur Entstehung typischer Krankheitssymptome auf Zitruspflanzen führt. Zusätzlich wurde für Mitglieder der PthA-Familie auch die Aktivierung einer Gruppe I *CsSWEETs* beobachtet. Allerdings konnte die Aktivierung dieses Gens durch einen künstlichen TALE die Virulenz des *Xac* Stammes 306, in dem der TALE *PthA4* mutiert wurde, nicht wiederherstellen (Hu, *et al.*, 2014). Dies weist darauf hin, dass im Vergleich zu *Xoo* die Aktivierung von Zuckertransportern in der *Xac*-Zitruspflanzen-Interaktion vermutlich keine Rolle für die Virulenz spielt. Ein weiterer TALE, der einen Transkriptionsfaktor aktiviert, ist AvrBs3 aus *Xcv* (Bonas, *et al.*, 1989). Für AvrBs3 sind mehr als 10 verschiedene Ziel-Gene in Paprika bekannt, zu denen z.B. *UPA20* gehört (Kay, *et al.*, 2007; Kay, *et al.*, 2009). *UPA20* kodiert einen *basic helix-loop-helix* Transkriptionsfaktor, der durch die Aktivierung weiterer Gene die pflanzliche Zellgröße reguliert und dessen TALE-abhängige Induktion ein Anschwellen des Blattgewebes zur Folge hat (Abb. 1.3). Diese Funktion könnte von *Xcv* genutzt werden, um ein Austreten großer Bakterienmengen aus dem Blattgewebe und somit die Verbreitung des Bakteriums zu ermöglichen (Marois, *et al.*, 2002; Kay, *et al.*, 2007). Jedoch ist der Einfluss von AvrBs3 auf die generelle Virulenz von *Xcv* in Feld- wie auch in Laborversuchen nicht detektierbar (Wichmann & Bergelson 2004).

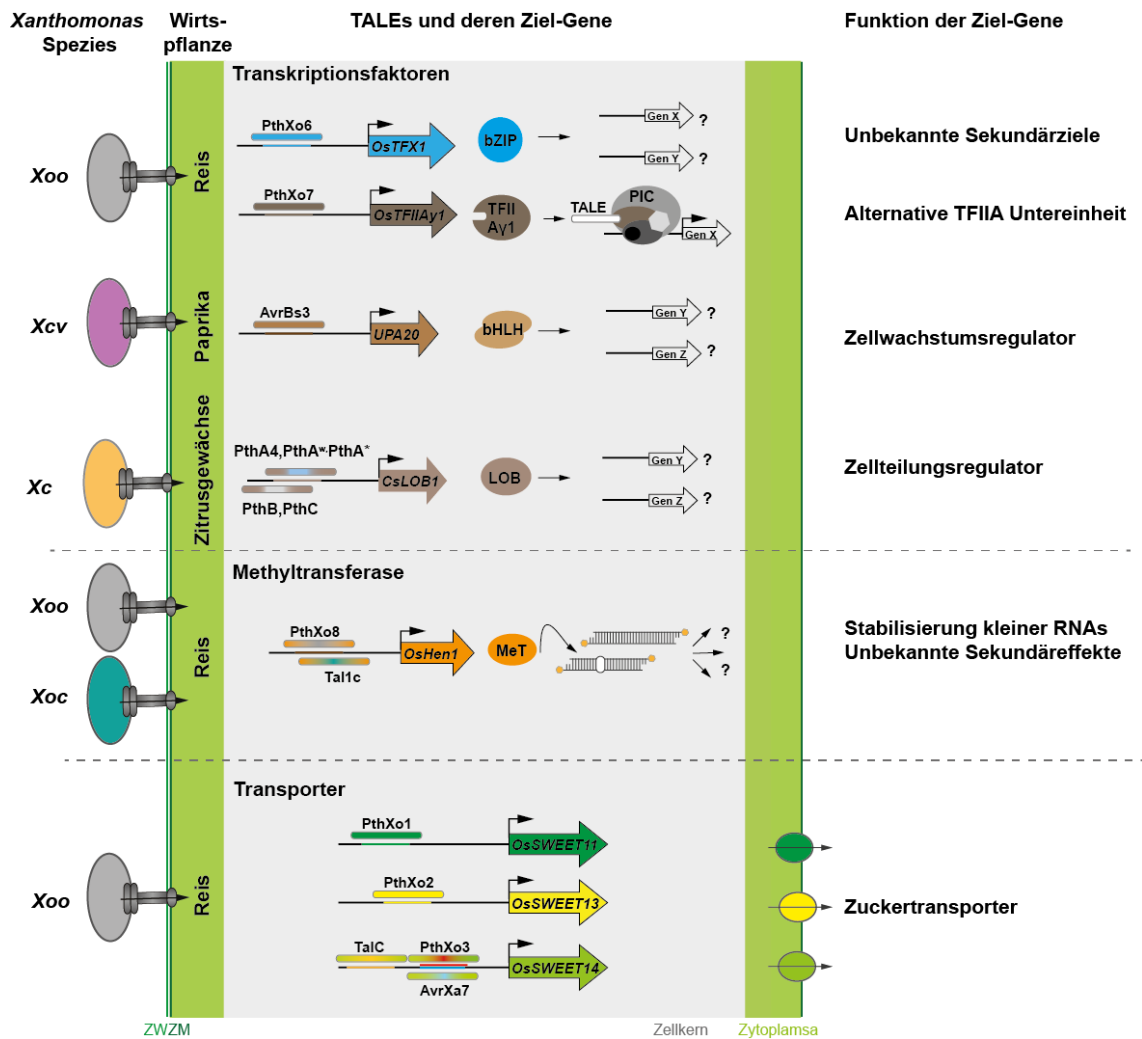


Abbildung 1.3 TALEs und deren Ziel-Gene

Dargestellt sind bekannte Ziel-Gene für TALEs aus *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*), *Xanthomonas citri* (*Xc*) bzw. *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) die für Transkriptionsfaktoren (TF), Methyltransferasen (MeT) oder Transporter kodieren. Die TALE-abhängige Aktivierung dieser Suszeptibilitätsgene trägt wahrscheinlich zur Versorgung des Bakteriums mit Nährstoffen, sowie zur Verbreitung der Bakterien bei. ZW: Zellwand; ZM: Zellmembran

1.2.6 TALEs als Avirulenzfaktoren

1.2.6.1 Dominante Resistenzgene

Im Rahmen der ETI haben Pflanzen unterschiedliche Resistenzmechanismen entwickelt, welche die Aktivität von TALEs detektieren, oder deren Bedeutung für das Pathogen ausnutzen. Eine Möglichkeit, mit der Pflanzen die Aktivität von TALEs erkennen und Resistenzreaktionen einleiten, erweitert die vorher beschriebene *Guard*-Hypothese und zeigt, dass es neben der Resistenz, ausgelöst durch die Interaktion eines Avr-Proteins mit einem *Guard*-Protein (Proteinfalle) auch Promotorfallen gibt. Dabei handelt es sich um TALE-Bindestellen im Promotor von Resistenzgenen (Schornack *et al.*, 2013). Die Resistenzgene werden dadurch TALE-abhängig angeschaltet und lösen Abwehrreaktionen aus, die meist zu einer HR führen (Rouxel & Balesdent 2010;

Schornack *et al.*, 2013). Für einige TALEs wurden die zugehörigen Resistenzgene bereits identifiziert. Allerdings zeigen die kodierten Resistenzproteine häufig keine Ähnlichkeit zueinander. Zudem ist weitgehend unklar über welche Signalwege sie Abwehrreaktionen einleiten. Das populärste Beispiel ist das Resistenzgen *Bs3*, welches in resistenten Paprikapflanzen (ECW-30R) durch *AvrBs3* aus *Xcv* aktiviert wird. *Bs3* wird außerdem vom *Xanthomonas gardneri*-TALE *AvrHah1* aktiviert (Schornack, *et al.*, 2008). Die *Bs3*-Proteinsequenz zeigt Homologien zur Familie der Flavin-Monooxygenasen. Der Mechanismus, über den *Bs3* Abwehrreaktionen aktiviert, ist bisher jedoch nicht bekannt (Römer, *et al.*, 2007; Abb. 1.4). Weitere TALE-abhängig aktivierte Resistenzgene sind vor allem aus den *Xoo*-Reis Interaktionen bekannt (Xia, *et al.*, 2012; Schornack, *et al.*, 2013). Dazu zählen beispielsweise die Resistenzgene *Xa7*, *Xa10*, *Xa23* und *Xa27* die jeweils durch die TALEs *AvrXa7*, *AvrXa10*, *AvrXa23* und *AvrXa27* aktiviert werden (Rouxel & Balesdent 2010; Wang, *et al.*, 2014a; Wang *et al.*, 2014b; Abb. 1.4). *Xa27* wird *AvrXa27*-abhängig aktiviert und vermittelt durch die weite Verbreitung dieses TALEs Resistenz gegen diverse *Xoo*-Stämme (Gu, *et al.*, 2005). Außerdem ist *Xa27* eines der am häufigsten beschriebenen Resistenzgene und wird neben anderen zur Erstellung transgener, resistenter Reispflanzen verwendet (Hummel, *et al.*, 2012; Liu, *et al.*, 2014). In diese wird ein synthetischer Promotor integriert, der Bindestellen für verschiedene TALEs aufweist und an die kodierende Sequenz von *Xa27* fusioniert ist (Hummel, *et al.*, 2012). Dadurch kann die Resistenz dieser Pflanzen auf möglichst viele verschiedene *Xoo*-Stämme erweitert werden (Hummel, *et al.*, 2012). Die Wirkung von *Xa27* ist Dosis-abhängig, wobei bisher über die Funktion von *Xa27* als Resistenzprotein noch wenig bekannt ist (Gu, *et al.*, 2005; Gu, *et al.*, 2009; Liu, *et al.*, 2014). Im Gegensatz dazu ist *Xa10* eines der wenigen klonierten Resistenzgene, für das bereits eine Funktion beschrieben wurde. Es kodiert ein 126 Aminosäuren umfassendes Protein, welches in der Membran des ER (Endoplasmatisches Retikulum) lokalisiert und dort wahrscheinlich in den Ca^{2+} -Ausstoß involviert ist (Tian, *et al.*, 2014). Ca^{2+} wirkt als Signalmolekül, das beispielsweise einen *Respiratorischen Burst* und damit weitere Abwehrreaktionen einleitet (Grant, *et al.*, 2000). Ein ebenfalls kloniertes *Xa10*-Allel ist *Xa23*, dessen Proteinsequenz eine 50%ige Identität zu *Xa10* aufweist (Wang, *et al.*, 2014c; Abb. 1.4). Dies deutet daraufhin, dass *Xa23* Resistenzreaktionen vergleichbar zu *Xa10* auslösen könnte. Bisher konnte dies jedoch noch nicht experimentell bestätigt werden (Wang, *et al.*, 2014c). *Xa23* wird durch den TALE *AvrXa23* aktiviert und löst in monokotyledonen, wie auch dikotyledonen

Pflanzen eine HR aus (Wang, *et al.*, 2014a; Abb. 1.4). AvrXa23 ist ein weit verbreiteter TALE, sodass *Xa23*-Reispflanzen resistent gegen ein breites Spektrum von *Xoo*-Stämmen sind (Wang, *et al.*, 2014a; Wang, *et al.*, 2014b; Wang, *et al.*, 2014c). Das Resistenzgen *Xa7* wird durch den TALE AvrXa7 aktiviert (Antony, *et al.*, 2010; Abb. 1.4). Der Mechanismus der *Xa7*-vermittelten Resistenz ist jedoch noch nicht bekannt. Allerdings ist die Ausnutzung von AvrXa7 als Resistenzgen-Aktivator in Pflanzen, für eine beständige *Xa7*-vermittelte Resistenz vorteilhaft (Vera Cruz, *et al.*, 2000). Dies ist darin begründet, dass die zugehörigen *Xoo*-Stämme auf die AvrXa7-abhängige Aktivierung des Suszeptibilitätsgens *OsSWEET14* angewiesen sind (Antony *et al.*, 2010). Somit ist eine Anpassung an die *Xa7*-Resistenz durch die Mutation oder sogar den Verlust von AvrXa7 kompliziert, weil dadurch die Virulenz des Stammes verloren gehen könnte. Für all diese Promotorfallen ist die Kernlokalisierung des TALEs und die TALE-abhängige Aktivierung des Resistenzgens von Bedeutung.

Das bisher einzige Beispiel Protein-vermittelter TALE-Erkennung ist die Erkennung von AvrBs4 durch das R-Protein Bs4 aus Tomate (Schornack, *et al.*, 2004; Abb. 1.4). Bs4 ist auch an der Erkennung weiterer TALEs z.B. Hax3 und Hax4 beteiligt (Schornack, *et al.*, 2004; Kay, *et al.*, 2005). Worauf genau die Erkennung bestimmter TALEs durch Bs4 beruht, ist bisher jedoch noch nicht vollständig geklärt.

1.2.6.2 Rezessive Resistenzgene

Zusätzlich zur Aktivierung von *R*-Genen haben Pflanzen weitere Mechanismen entwickelt, um die TALE-abhängige Aktivierung von Suszeptibilitätsgenen zu verhindern. Eine Möglichkeit ist die Mutation der TALE-Bindestellen im Promotor von Suszeptibilitätsgenen. Beispiele hierfür sind die rezessiven Reis-Resistenzgene *xa13* (im *OsSWEET11*-Gen) sowie *xa25* (im *OsSWEET13*-Gen). Sie enthalten Insertionen bzw. Deletionen im Promotor dieser Suszeptibilitätsgene, wodurch PthXo1 bzw. vermutlich PthXo2 nicht mehr binden können (Chu, *et al.*, 2006a; Yang, *et al.*, 2006; Liu, *et al.*, 2011; Abb. 1.4). Damit sind diese TALEs nicht mehr in der Lage das jeweilige Suszeptibilitätsgen zu aktivieren und die Pflanze ist resistent. Wie bereits beschrieben, wird auch ein weiteres Suszeptibilitätsgen der *OsSWEET*-Familie (*OsSWEET14*) sogar von drei verschiedenen TALEs aktiviert. Für dieses Gen sind bisher noch keine Resistenzallele bekannt. Die Übertragung eines TALEs der *OsSWEET14* aktiviert, kann somit die Resistenz der *xa13*- bzw. *xa25*-Reispflanzen umgehen (Antony, *et al.*, 2010). Bisher ist jedoch unklar, ob die Fokussierung auf *OsSWEET14* im stetigen Wettrüsten

mit der Pflanze als Anpassung an die *xa13*- bzw. *xa25*- Resistenzen erschlossen wurde oder ob sich die zugehörigen TALEs unabhängig davon entwickelt haben. Eine weitere Möglichkeit für Pflanzen die TALE-abhängige Aktivierung von Suszeptibilitätsgenen zu verhindern, ist die Mutation möglicher TALE-Interaktionspartner wie z.B. TFIIA, der Bestandteil des Präinitiationskomplexes (*preinitiation complex*, PIC) der RNA-Polymerase-II ist (Orphanides, *et al.*, 1996; Thomas & Chiang 2006). TFIIA stabilisiert die Bindung des TBP/TFIID-Komplexes (TATA-Bindeprotein, Transkriptionsfaktor IID) an die DNA und fördert dadurch die weitere Bildung des PIC (Thomas & Chiang 2006). Im Reisgenom sind zwei Allele der γ -Untereinheit des TFIIA kodiert, ein Allel auf Chromosom 1 (*TFIIA γ 1*) und eines auf Chromosom 5 (*TFIIA γ 5*) (Iyer & McCouch 2004). Das rezessive Reis R-Gen *xa5* kodiert eine mutierte TFIIA γ 5-Untereinheit mit einem einzelnen Aminosäureaustausch von Valin (V) zu Glutaminsäure (E) an Position 39 (Iyer & McCouch 2004; Abb. 1.4). Dieser Austausch hat keinen Einfluss auf die TFIIA-Funktion im PIC, kann aber die TALE-abhängige Aktivierung von Ziel-Genen verringern (Iyer & McCouch 2004; Sugio, *et al.*, 2007; Gu, *et al.*, 2009; Abb. 1.4). So ist beispielsweise der *Xoo*-Stamm PXO86 auf homozygoten *xa5* Reispflanzen nicht virulent, vermutlich weil AvrXa7 *OsSWEET14* nicht mehr effizient aktivieren kann (Sugio, *et al.*, 2007). Auch die TALE-abhängige Aktivierung von Resistenzgenen ist in homozygoten *xa5*-Reispflanzen reduziert (Gu, *et al.*, 2009). Eventuell verringert die TFIIA γ 5 Mutation eine Interaktion mit TALEs und somit die Aktivierung des Ziel-Gens (Iyer & McCouch 2004; Sugio, *et al.*, 2007; Gu, *et al.*, 2009; Abb. 1.4). Interessanterweise kann PXO99 als bisher einziger *Xoo*-Stamm die *xa5*-Resistenz umgehen. PXO99 kodiert den TALE PthXo7, der das Allel *TFIIA γ 1* aktiviert (Abb. 1.3). Dadurch ist dieser Stamm in der Lage Krankheitssymptome auf homozygoten *xa5*-Reispflanzen auszulösen (Sugio, *et al.*, 2007). Die Übertragung von PthXo7 in PXO86 ermöglicht auch diesem Stamm wieder Krankheitssymptome auf homozygoten *xa5*-Reispflanzen auszulösen (Sugio, *et al.*, 2007). Durch die Aktivierung des alternativen Allels eines möglichen TALE-Interaktionspartners hilft PthXo7 also vermutlich die *xa*-Resistenz zu überwinden. Damit stellt PthXo7 ein exzellentes Beispiel dar, das zeigt wie sich Pathogene durch die Entwicklung alternativer Effektoren im ständigen Wettrüsten mit der Pflanze an Resistenzen anpassen können, um diese zu überwinden.

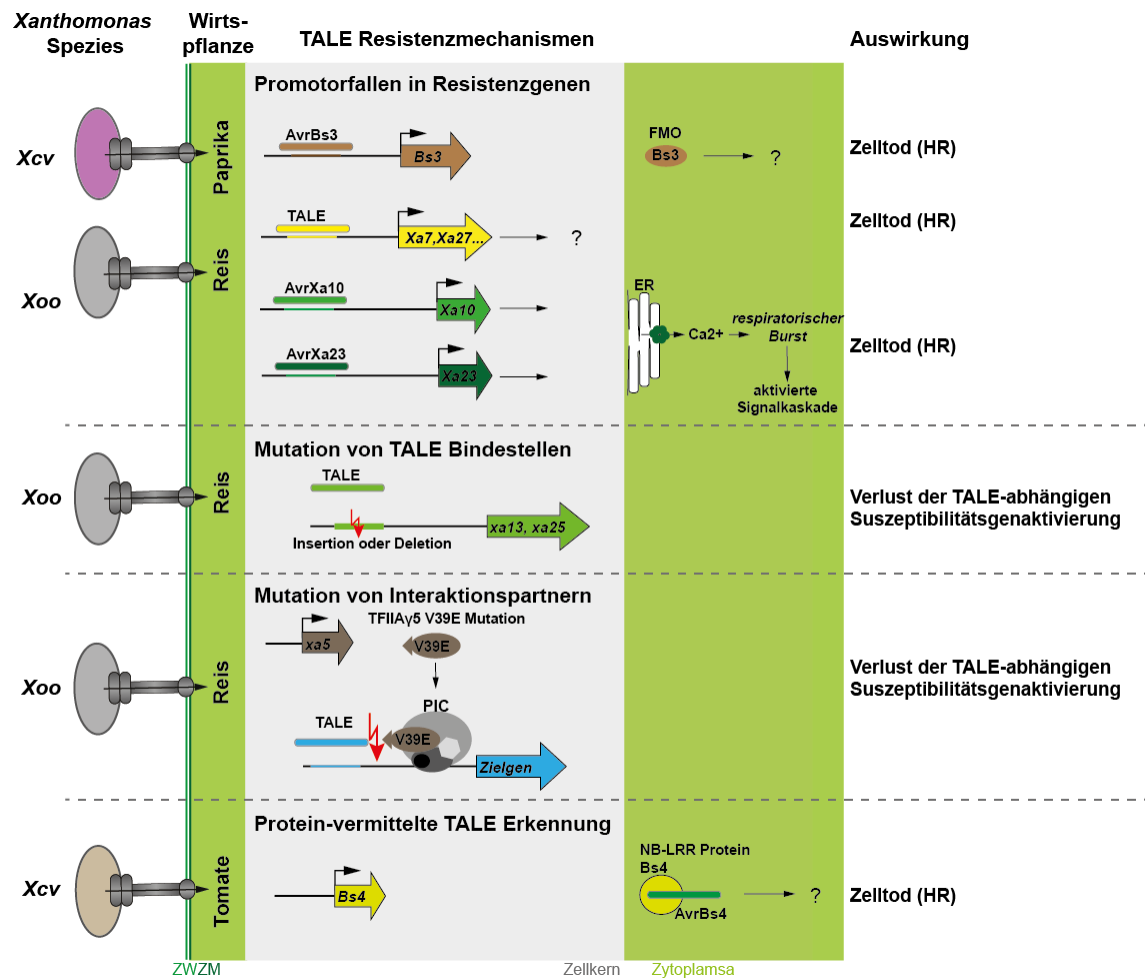


Abbildung 1.4 TALE-vermittelte Resistenzmechanismen

Dargestellt sind bekannte Resistenzmechanismen die zur Erkennung von TALEs aus *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) und *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) führen bzw. deren DNA-Bindung und Aktivität verhindern. Die TALE-abhängige Aktivierung von Resistenzgenen führt über bisher häufig unbekannte Signalwege zur Auslösung der hypersensitiven Reaktion (HR). *Xa10* ist eines der wenigen *R*-Gene dessen Wirkungsweise bereits bekannt ist. *Xa23* ist vermutlich ein Allel von *Xa10* und könnte eine vergleichbare Funktion erfüllen. Weitere Resistenzmechanismen werden durch die Mutation von TALE-Bindestellen im Promotor von Suszeptibilitätsgenen bzw. durch die Mutation von TALE-Interaktionspartnern vermittelt. Dadurch kann die TALE-abhängige Aktivierung dieser Gene verhindert werden. Die einzige bisher bekannte Protein-basierte Erkennung von TALEs wird durch das Tomaten R-Protein Bs4 vermittelt, welches über bisher noch weitgehend unbekannte Signalwege zur HR führt. ER: Endoplasmatisches Retikulum ZW: Zellwand; ZM: Zellmembran; FMO: Flavin-abhängige-Monooxygenase

1.2.7 TALEs als Transkriptionsaktivatoren

TALEs aktivieren Gene. Die Notwendigkeit einer AD für die TALE-Aktivität lässt vermuten, dass TALEs in der Lage sind Teile der Polymerase-II-abhängigen Transkriptionsmaschinerie zu rekrutieren. Wie genau dies passiert bzw. mit welchen Wirtspoteinen TALEs interagieren ist bisher jedoch weitgehend unklar. Interessanterweise verschieben TALEs abhängig von der Position ihrer Ziel-DNA-Sequenz den Transkriptionsstart am Promotor des Ziel-Gens, sodass dieser ca. 40-60bp *downstream* der TALE-Bindestelle liegt und sich damit auch vom natürlichen

Transkriptionsstart unterscheiden kann (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Antony, *et al.*, 2010; Hummel, *et al.*, 2012). Dies könnte darauf hinweisen, dass TALEs eigenständig den PIC an den Promotor rekrutieren können. Diese Rekrutierung könnte über die Interaktion mit Hilfsproteinen, z.B. dem Mediator, oder sogar über die direkte Interaktion mit Bestandteilen des PIC ablaufen. Hinweise auf letzteres gibt das bereits beschriebene Resistenzallel *xa5* welches die γ -Untereinheit des generellen TFIIA betrifft. Ob TALEs jedoch direkt mit TFIIA interagieren oder weitere Proteine benötigen bzw. weitere Interaktionen mit anderen Komponenten der Transkriptionsmaschinerie stattfinden, ist aus diesen Studien bisher nicht hervorgegangen.

Des Weiteren konnte beobachtet werden, dass manchmal eine perfekte TALE-Bindestelle im Promotor allein nicht ausreicht um ein Reportergen effizient zu aktivieren (Scholze 2011; Bultmann, *et al.*, 2012). Dies deutet daraufhin, dass entweder weitere Promotorelemente vorhanden sein müssen, um die TALE-abhängige Aktivierung zu ermöglichen oder die Zugänglichkeit des Promotors nicht gegeben ist, wodurch die Bindung des TALEs verhindert wird. Speziell die Zugänglichkeit des Promotors kann durch verschiedene Möglichkeiten eingeschränkt sein. Beispielsweise könnten DNA-bindende Wirtsproteine TALE-Bindestellen blockieren. Bisher ist jedoch noch nicht bekannt ob und wie TALEs mit diesen um mögliche Bindestellen konkurrieren. Außerdem konnte bereits gezeigt werden, dass DNA-Methylierung die Bindung des TALEs verhindern kann (Bultmann, *et al.*, 2012). Dies ist darin begründet, dass das Cytosin-spezifische RVD HD keine methylierten Cytosine bindet, wodurch die Bindung des gesamten TALEs verhindert werden kann (Bultmann, *et al.*, 2012; Valton, *et al.*, 2012; Deng, *et al.*, 2012b). Möglichkeiten dies zu umgehen sind entweder vor allem bei bekanntem Methylierungsstatus die Verwendung der RVDs N*, H* oder NG zur Bindung methylierter Cytosine oder andernfalls die gleichzeitige chemische Hemmung epigenetischer Modifikationsenzyme (Bultmann, *et al.*, 2012; Valton, *et al.*, 2012; Deng, *et al.*, 2012b). Weiterhin können Nukleosomen die DNA-Bindestelle eines TALEs blockieren (Bultmann, *et al.*, 2012). Daher wird auch hier entweder die chemische Hemmung epigenetischer Modifikationsenzyme oder aber die Positionierung der TALE-Bindestelle in DNaseI-hypersensitiven Regionen als vorteilhaft betrachtet (Bultmann, *et al.*, 2012; Maeder, *et al.*, 2013). Natürlicherweise binden viele TALEs in der Nähe oder direkt an der TATA-Box eines Ziel-Gens und damit in Bereichen von ca. 50 bis zu 300bp vor dem Translationsstart des Ziel-Gens (Grau, *et al.*, 2013).

Interessanterweise zeigten Studien im tierischen System, dass künstliche TALE-Aktivatoren ausgehend von verschiedensten Bindestellen im Promotor beispielsweise in der kodierenden Sequenz eines Gens, oder sogar *reverse* am Gegenstrang bindend ein Ziel-Gen aktivieren können (Maeder, *et al.*, 2013; Perez-Pinera, *et al.*, 2013; Uhde-Stone, *et al.*, 2014). Diese Experimente wurden jedoch mit TALE-Derivaten durchgeführt, die eine alternative AD (VP16 bzw. VP64) anstelle der natürlichen AD enthielten. Die VP16-AD vermittelt vermutlich die Transkriptionsaktivierung unabhängig von Position oder Orientierung der Bindestelle vergleichbar mit einem *Enhancer*-bindenden Protein (Seipel, *et al.*, 1992; Bulger & Groudine 2011; Levine, *et al.*, 2014). Bisher ist jedoch unklar ob TALEs mit natürlicher AD ebenfalls ausgehend von verschiedensten Positionen im Promotor ein Ziel-Gen effektiv aktivieren können.

Die Analyse, wo TALEs in Promotoren positioniert werden können, welche Promotorelemente gegeben sein müssen, wie TALEs mit Wirtsproteinen um Bindestellen konkurrieren oder die Transkriptionsmaschinerie rekrutieren, kann weiter zum Verständnis der TALE-abhängigen Transkriptionsaktivierung beitragen. Außerdem könnten diese Erkenntnisse auch auf die effiziente Anwendung von TALEs als Genaktivatoren, Gen-Modifikationsenzyme oder sogar auf die Anwendung anderer DNA-bindender Proteine übertragen werden.

1.3 Zielstellung

Seit der Entschlüsselung des RVD-DNA-Codes sind TALEs Gegenstand intensiver Forschungen, die sich mit der Anwendung von TALEs als künstliche DNA-Bindedomäne oder der Analyse pflanzlicher Ziel-Gene beschäftigen. Ziel dieser Arbeit war es daher weitere TALE-Charakteristika zu analysieren. Die Vorhersage pflanzlicher Ziel-Gene setzt die möglichst vollständige Entschlüsselung fehlender RVDs voraus. Speziell die Spezifitäten selten vorkommender *Xanthomonas*-TALE-RVDs bzw. *Ralstonia*-TALE-RVDs sind nicht anhand bioinformatischer Analysen schon bekannter TALEs und deren Ziel-Gene vorherzusagen und sollten daher experimentell ermittelt werden. Da TALE-Homologe aus *R. solanacearum* vermutlich ebenfalls mit Hilfe ihrer *repeat*-Region an DNA binden, sollte die Analyse der in *Ralstonia* vorkommenden RVD-Spezifitäten die Vorhersage einer möglichen Ziel-DNA-Sequenz ermöglichen. Interessanterweise finden sich im TALE-Repertoire reispathogener *Xoo*- und *Xoc*-Stämme TALEs die *repeats* mit von der Norm abweichender Länge enthalten. Diese *repeats* könnten die strukturelle Regelmäßigkeit der *repeat*-Region und damit

möglicherweise auch den DNA-Bindungsmechanismus der betroffenen TALEs beeinflussen. Daher sollte die Funktionalität von TALEs mit solchen *repeats* analysiert werden. Die TALE-abhängige Aktivierung von Suszeptibilitätsgenen fördert die Virulenz des zugehörigen *Xanthomonas*-Stammes. Interessanterweise scheinen sich besonders *Xoo*-Stämme auf die Aktivierung von Mitgliedern der *OsSWEET*-Familie zu fokussieren. Im Rahmen dieser Arbeit sollte analysiert werden ob auch weitere Mitglieder dieser Familie als Suszeptibilitätsgene für *Xoo*-TALEs dienen können. Bisher ist noch weitgehend unbekannt ob TALEs beliebig an verschiedene Positionen im Promotor positioniert werden können oder ob es Positionen gibt, die eine besonders effiziente Aktivierung des Ziel-Gens ermöglichen. Anhand solcher Effekte der Positionierung auf die Ziel-Genaktivierung könnten eventuell Rückschlüsse auf die Kooperation von TALEs mit weiteren Promotorelementen gezogen werden. Um Einblicke in solche Positionseffekte zu erhalten, sollte die Aktivität eines Repertoires künstlicher TALEs mit verschiedenen Bindestellen in einem natürlichen Promotor analysiert werden. Die Ergebnisse solcher Analysen könnten nicht nur dazu dienen die Positionierung natürlicher TALEs zu verstehen, sondern auch Hinweise für die geeignete Auswahl von Bindestellen künstlicher TALEs geben, wenn diese als Werkzeug zur gezielten Aktivierung von Genen angewendet werden.

2. Ergebnisse

2.1 Die Erweiterung des TALE-RVD-Codes

2.1.1 Publikation 1

be controlled by cognate promoters to allow physiological regulation of activity. Indeed, for many disorders the profile of gene expression required to achieve significant therapeutic effect is achievable with relatively simple expression systems. For example, the frequency of spontaneous hemorrhage is greatly reduced in hemophilia B with a small fraction of normal circulating levels of FIX. Similarly, a fraction of functional circulating phagocytes is sufficient to offer substantial protection against infection in some inherited immunodeficiencies. Diseases most amenable to effective treatment with current gene therapy approaches may therefore be those where the therapeutic window for functional protein expression is relatively broad, and where a large number of target cells are accessible to transduction.

Previous expectations that gene therapy would produce a 'cure-all' solution for intractable diseases were unrealistic. The agents in question are a diverse spectrum of nucleic-acid based medicines that are formulated in very different ways, yet can be used to prevent, alleviate and provide long-lasting treatments for a wide variety of diseases both inherited and acquired. In some cases they can now realistically provide physicians and patients with new therapeutic options where more conventional approaches have failed, a testament to the huge amount of scientific research in the field over the last 10–20 years. By decreasing the requirement for frequent repeated interventions, many gene therapy approaches can lead to substantial savings in the costs of lifetime medication. However, it is also noteworthy that without clinical trials, the field would not have progressed nearly so rapidly. The fact that some gene therapy strategies are finally beginning to deliver on their potential may well herald a raft of new and imaginative interventional approaches designed to exploit recent insights into cell biology and disease processes.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available at <http://www.nature.com/doifinder/10.1038/nbt.2290>.

Leonard W Seymour¹ & Adrian J Thrasher²

¹President, ESGCT-BSGT Collaborative Congress; General Secretary, ESGCT; Department of Oncology, University of Oxford, Oxford, UK. ²President, BSGCT; Centre for Immunodeficiency, University College London Institute of Child Health, London, UK.

1. Abstracts Collaborative Meeting of the European Society of Gene and Cell Therapy and the British Society for Gene Therapy. *Hum. Gene Ther.* **22**, A1–A138 (2011).
2. Bainbridge, J.W. *et al.* *N. Engl. J. Med.* **358**, 2231–2239 (2008).
3. Maguire, A.M. *et al.* *N. Engl. J. Med.* **358**, 2240–2248 (2008).
4. Maguire, A.M. *et al.* *Lancet* **374**, 1597–1605 (2009).
5. Nathwani, A.C. *et al.* *N. Engl. J. Med.* **365**, 2357–2365 (2011).
6. Raj, D., Davidoff, A.M. & Nathwani, A.C. *Expert Rev. Hematol.* **4**, 539–549 (2011).
7. Booth, C., Gaspar, H.B. & Thrasher, A.J. *Curr. Opin. Pediatr.* **23**, 659–666 (2011).
8. Cartier, N. *et al.* *Science* **326**, 818–823 (2009).
9. Cirak, S. *et al.* *Lancet* **378**, 595–605 (2011).
10. Aartsma-Rus, A. *et al.* *Eur. J. Hum. Genet.* **18**, 889–894 (2010).
11. Jessup, M. *et al.* *Circulation* **124**, 304–313 (2011).
12. Sheehy, S.H. *et al.* *Mol. Ther.* **19**, 2269–2276 (2011).
13. Schaller, T. *et al.* *PLoS Pathog.* **7**, e1002439 (2011).
14. Li, H. *et al.* *Nature* **475**, 217–221 (2011).
15. Perez, E.E. *et al.* *Nat. Biotechnol.* **26**, 808–816 (2008).
16. Stroncek, D.F. *et al.* *J. Transl. Med.* **10**, 48 (2012).
17. Gollamudi, R. *et al.* *Invest. New Drugs* **28**, 641–649 (2010).

CORRESPONDENCE

18. Comins, C. *et al.* *Clin. Cancer Res.* **16**, 5564–5572 (2010).
19. Breitbach, C.J. *et al.* *Hum. Gene Ther.* **22**, A12 (2011).
20. Kaufman, H.L. & Bines, S.D. *Future Oncol.* **6**, 941–949 (2010).
21. Heo, J. *et al.* *Mol. Ther.* **19**, 1170–1179 (2011).
22. Breitbach, C.J. *et al.* *Nature* **477**, 99–102 (2011).
23. Gaspar, H.B. *et al.* *Sci. Transl. Med.* **3**, 97ra79 (2011).
24. Hacein-Bey-Abina, S. *et al.* *N. Engl. J. Med.* **363**, 355–364 (2010).
25. Gaspar, H.B. *et al.* *Sci. Transl. Med.* **3**, 97ra80 (2011).
26. Aiuti, A. *et al.* *N. Engl. J. Med.* **360**, 447–458 (2009).
27. Grez, M. *et al.* *Mol. Ther.* **19**, 28–35 (2011).
28. Boztug, K. *et al.* *N. Engl. J. Med.* **363**, 1918–1927 (2010).
29. Galy, A. & Thrasher, A.J. *Curr. Opin. Allergy Clin. Immunol.* **11**, 545–550 (2011).
30. Cavazzana-Calvo, M. *et al.* *Nature* **467**, 318–322 (2010).
31. Porter, D.L., Levine, B.L., Kalos, M., Bagg, A. & June, C.H. *N. Engl. J. Med.* **365**, 725–733 (2011).
32. Kochenderfer, J.N. *et al.* *Blood* **119**, 2709–2720 (2012).
33. Di Stasi, A. *et al.* *N. Engl. J. Med.* **365**, 1673–1683 (2011).
34. Harrington, K.J. *et al.* *Clin. Cancer Res.* **16**, 4005–4015 (2010).
35. Senzer, N.N. *et al.* *J. Clin. Oncol.* **27**, 5763–5771 (2009).

TAL effector RVD specificities and efficiencies

To the Editor:

The DNA-binding domain of transcription activator–like effectors (TALEs) has become an important tool for the programmable and specific targeting of genome-editing nucleases^{1–4}. We determined specificities of TALE DNA-binding modules and discovered varying efficiencies to promote TALE activity. This prompted us to study the basis of TALE efficiency. Natural TALEs from plant-pathogenic *Xanthomonas* sp. bacteria function as eukaryotic transcription factors. TALEs bind to DNA via tandem near-identical 34-amino-acid repeats, each of which recognizes one base pair via two adjacent amino acids (12 and 13) termed repeat-variable diresidues (RVDs)^{5,6}. Three-dimensional structures of TALE–DNA complexes revealed that only amino acid 13 mediates specific recognition of the sense-strand DNA base, whereas amino acid 12 stabilizes the repeat structure^{7,8}. *In vitro* assembly of repeat units yielded TALEs with new and predictable DNA specificities^{2,9}. Different DNA-recognition specificities of individual RVDs have been shown^{4,5,10}. Specific recognition of guanine is difficult because the common RVD NN (Asn–Asn) recognizes guanine and adenine, whereas the guanine-specific RVD NK (Asn–Lys) apparently functions less well than NN does¹¹.

Here we describe new RVD specificities including NH (Asn–His), which is highly specific for guanine. We show that efficiency of RVDs varies and that strong repeats (HD (His–Asn) or NN) are key for overall activity of TALEs.

First, we constructed artificial TALEs (ArtTALs) composed of an unusual arrangement of repeats with predominantly one type of RVD to decipher RVD specificities (Supplementary Methods). With this highly artificial setup, we first aimed to study properties of each RVD type separately. These ArtTALs contained the RVDs NI (Asn–Ile) and NG (Asn–Gly) in the first and last repeat, respectively, to properly position the TALE on the target DNA. We transfected ArtTAL expression constructs together with β -glucuronidase (*GUS*) reporter constructs via *Agrobacterium* into leaf cells of *Nicotiana benthamiana*. As expected, TALEs with poly(HD) (ArtTAL_{polyHD}) and poly(NN) (ArtTAL_{polyNN}) RVDs efficiently induced the polycytosine and polyguanine reporters, respectively (Fig. 1a and Supplementary Fig. 1). To our surprise, ArtTAL_{polyNN} did not induce expression of the polyadenine reporter construct, which is in contrast to the known specificity of the RVD NN for guanine and adenine. Similarly, ArtTALs with poly(NI) (specificity of NI is for adenine),



CORRESPONDENCE

poly(NG) (specificity of NG is for thymine) and other RVDs did not activate any of the reporter constructs. We modified the naturally most variable non-RVD amino acids (positions 4 and 24) of NI, NG and NN repeats, but this did not restore activity of weak RVDs (Supplementary Fig. 2). This demonstrates that TALEs with certain RVD compositions are not functional, a finding that is contradictory to the simple TALE code reported so far^{2,5,6}.

Possibly RVDs have different efficiencies and having many repeats with weak RVDs impedes TALE function. Therefore, we constructed ArtTALs, with an 'anchor' consisting of the first six repeats of the natural TALE Hax3, which contain several RVDs HD, which are efficient according to our ArtTAL_{polyHD} data (Fig. 1a). These six repeats were followed by six repeats with identical RVDs to yield a total of 11.5 repeats (Supplementary Sequences). Reporter constructs for *in planta* GUS assays contained the first half of the Hax3 box and six A, C, G or T, respectively. We used this artificial setup to test known and new RVDs. In total, we analyzed eight new RVD specificities and found that RVDs not present in natural *Xanthomonas* TALEs can be functional (Fig. 1b, Supplementary Figs. 3, 4 and Supplementary Table 1). Notably, the RVD NH showed a specificity for guanine only, the RVD NP showed a novel specificity for three different bases (A, C or T), and the RVD NS, which previously was designated as nonspecific, showed a clear preference for guanine and adenine. Notably, the new guanine-specific RVD NH was much more efficient than the guanine-specific RVD NK. These RVD specificities will allow a highly specific targeting of guanine (NH) and a flexible targeting of closely related or allelic sequences, respectively. Some RVDs (NI, N*; the * indicates a missing residue in the repeat) again did not result in any measurable TALE activity, although the proteins were produced to similar levels as functional ArtTALs (Fig. 1b and Supplementary Fig. 4).

To test whether activity and specificity of weak RVDs can be measured if fewer than six are combined in tandem, we used the Hax3 repeat framework and replaced four, three or two RVDs with RVDs to be tested. GUS reporter assays *in planta* showed that ArtTALs with test RVDs HD and NN activated the cytosine and the guanine reporters, respectively, irrespective of the number of test RVDs (Supplementary Fig. 5). In contrast, ArtTALs with test RVDs NG, NI, N* or NN

(specificity for adenine) led to a substantially increased reporter activation if fewer of these RVDs were present in tandem. The ArtTAL_{HD} derivatives were less active than the ArtTAL_{NN} derivatives in these experiments, whereas the ArtTAL_{polyHD} was more active than the ArtTAL_{polyNN} (Supplementary Fig. 1). This apparent contradiction might be explained by the hypothesis that efficiencies of RVDs HD and NN or the overall functionality of TALEs can be influenced by the nature of amino acid 4 in the repeats. Here the original poly(HD) and poly(NN) ArtTALs contain the polar glutamine (Fig. 1a; Supplementary

Fig. 1), whereas the other ArtTAL constructs contain the negatively charged aspartic acid or glutamic acid (Fig. 1b–e, Supplementary Figs. 2, 4–7 and Supplementary Table 2). For example, a direct comparison shows a stronger activity of ArtTAL_{polyNN} with aspartic acid than with glutamine at position 4, respectively (Supplementary Fig. 2).

To test whether certain RVDs are only weak if identical RVDs are present in tandem, we constructed ArtTALs with mixed RVDs. The repeat domain consisted of 10.5 repeats, which can be sufficient for maximal transcriptional activation⁵. We combined three alternating

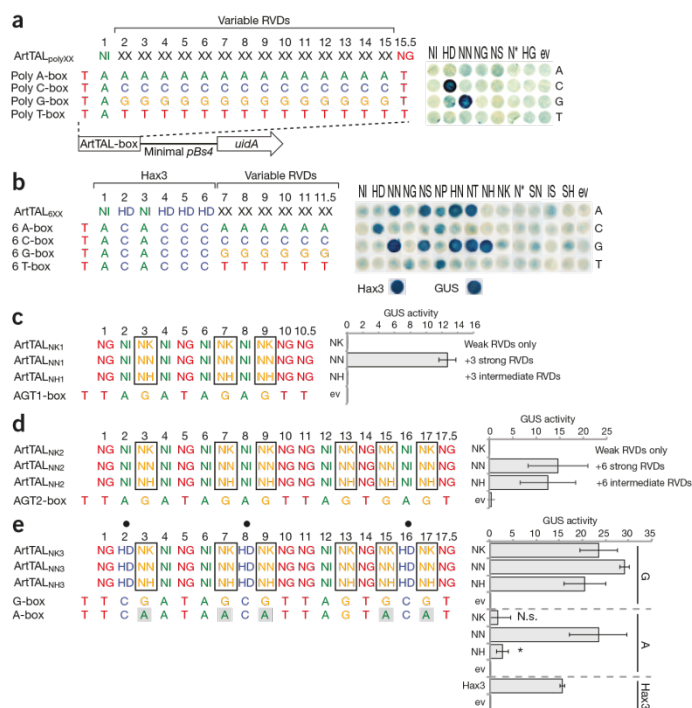


Figure 1 TALE RVD specificities and efficiencies. (a–e) ArtTALs with varying RVD compositions were constructed. Possible target DNA boxes were cloned in front of the minimal *Bs4* promoter into a *GUS* reporter vector. The target boxes are preceded by a T, which is specified by the TALE N-terminal region^{5,6,8}. Constitutive *35S*-driven *artTALS* or empty T-DNA (ev) were codelivered via *A. tumefaciens* with *GUS* reporter constructs into *N. benthamiana*. The *35S::uidA* (*GUS*) and Hax3 with Hax3 box (Hax3) served as positive controls and empty vector (ev) as negative control. Leaf disks were stained with X-Gluc. *GUS* activity is the mean of three samples (expressed in pmol 4-methyl-umbelliferone min⁻¹ per µg protein; error bars indicate s.d.). ArtTAL_{polyXX} with 14 identical RVDs (a). ArtTAL_{6XX} with the first six repeats of Hax3 and six repeats with RVDs to be tested (b). Short ArtTALs with mixed RVDs and same DNA specificity: ArtTAL_{NK1} (only weak RVDs), ArtTAL_{NH1} and ArtTAL_{NH2} (weak and strong repeats) (c). Long ArtTALs with mixed RVDs and same DNA specificity: ArtTAL_{NK2} (only weak RVDs), ArtTAL_{NH2} and ArtTAL_{NH2} (weak and strong repeats) (d). ArtTALs with mixed RVDs including three RVDs HD marked by • (e). Significance (*) or non-significance (N.s.) above empty vector control ($P < 0.05$) is indicated.

Table 1 Recommendations for the design of efficient TALEs

Recommendation	Description
1	Incorporate at least some (–3–4) strong RVDs (e.g., HD or NN)
2	Position strong RVDs to avoid stretches (≥6) of weak RVDs, especially at ends
3	For high guanine specificity use NH (to promote TALE activity) or NK (if enough strong RVDs are present)
4	Use NN for guanine if only a few other strong RVDs are present

repeats with weak RVDs (NI, NG and NK) of different specificities. As an alternative for a guanine-specific RVD, we replaced the weak RVD NK with the more efficient RVDs NN or NH. All resulting ArtTALs should recognize the same DNA sequence. This avoided sequence-specific effects, such as competition with host proteins that might otherwise differently affect target sequences. GUS reporter assays *in planta* showed that only the ArtTAL containing the strong RVD NN, but not the ArtTALs containing the RVDs NK or NH, led to transcriptional activation of the reporter gene (Fig. 1c). This example suggests that also a combination of different weak RVDs compromises TALE activity and that function of TALEs relies on the presence of strong RVDs in the repeat domain. To our surprise, the RVD NN was efficient in this assay, but the RVD NH was not, whereas in the ArtTAL_{6XX} constructs both exhibited similar efficiency. To test whether the particularly short length of the repeat domain was responsible for this, we constructed ArtTALs with 17.5 repeats and mixed weak RVDs (NI, NG and NK) or NN and NH substitutions. GUS reporter assays *in planta* showed that the ArtTAL_{NK2} with mixed 17.5 weak RVDs was not functional, whereas the ArtTALs with RVDs NH and NN produced high GUS reporter activity (Fig. 1d). Protein levels of the ArtTALs with mixed RVDs were comparable (Supplementary Fig. 6). We conclude from these examples that efficiency of weak RVDs cannot be compensated by simply increasing the number of repeats. In addition, the RVD NH can promote activity of TALEs but with less efficiency than NN (thus, NH was classified as 'intermediate', Fig. 1c,d). Here the total number of repeats apparently had an effect, but additional experiments are needed to generalize this.

Our classification of RVDs into weak, intermediate and strong is an approximation of the variable RVD efficiencies. It is likely that a more quantitative assessment will eventually reveal subtle differences in different

contexts. Apparent efficiencies of certain RVDs might also depend on the nature of neighboring RVDs in the TALE repeat domain. With the exception of tandem repeats with RVDs HD that form inter-repeat hydrogen bonds⁷, no influence between neighboring RVDs has so far been reported to our knowledge^{6–8}. Nevertheless, flanking strong RVDs might stabilize TALE protein–DNA interaction sufficiently to also allow non-optimal RVD-base combinations without decreasing overall TALE activity.

To corroborate our hypothesis that insertion of strong RVDs is essential for TALE function, we incorporated three or four strong RVDs HD into the long mixed ArtTALs and placed a corresponding target box with three or four cytosines into the reporter construct. *In planta* GUS reporter assays demonstrated that the RVDs HD efficiently transformed the nonfunctional ArtTAL_{NK} to full activity but had no profound impact on the already active ArtTAL_{NH} and ArtTAL_{NN} (Fig. 1e and Supplementary Fig. 7). To compare specificity of guanine-specific RVDs, we replaced guanine in the target box at five of the six positions with adenine. GUS reporter assays showed that the RVDs NH and NK were more specific for guanine than the RVD NN was (Fig. 1e). Our experiments show that TALE activity depends on RVD composition and that tandem weak RVDs can compromise TALE function. The data set we present here comprises 55 artificial TALEs, but many RVD combinations remain to be tested. Future analyses will define quantitative efficiencies, combinatorial effects and possible positional influence of RVDs in the TALE repeat domain.

In conclusion, our findings have a direct impact on the design of artificial TALEs. We suggest that TALEs should contain at least some (for example, 3–4) properly spaced strong RVDs to enable full activity. If high specificity for guanine is required, the RVDs NH or NK are preferred, but RVD NN should be chosen if weak overall efficiency

of the TALE needs to be increased. In particular, TALEs with predominantly weak RVDs (NI, NG or NK) should be designed carefully to also include strong RVDs (HD or NN). We compiled recommendations to summarize this in Table 1 and believe that these will help the community construct reliably functional TALEs. We believe these recommendations will be important for TALE nuclease design and in other eukaryotic systems, but this awaits experimental proof. Understanding the molecular basis of RVD efficiency will eventually help to define rules for TALE design and RVD choices. Our data are supported by TALE–DNA complex structures^{7,8}. Strong RVDs (NN or HD) form hydrogen bonds with DNA bases, whereas some weak RVDs (for example, NI or NG) only make van der Waals contacts^{7,8}. RVD efficiency is an important general extension to the simple DNA-specificity code of TALEs.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank R. Szczesny and T. Schreiber for helpful discussions on the manuscript, S. Thieme (Department of Genetics, Martin Luther University Halle-Wittenberg) for providing the binary vector, C. Kretschmer for expert technical assistance and U. Bonas for continued support. J.B. was supported by a grant from the Deutsche Forschungsgemeinschaft (SPP 1212).

AUTHOR CONTRIBUTIONS

J.B. and J.S. designed the experiments. J.S., C.B. and A.L. performed the experiments. J.B. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Jana Streubel, Christina Blücher, Angelika Landgraf & Jens Boch

Department of Genetics, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany.
e-mail: jens.boch@genetik.uni-halle.de

1. Baker, M. *Nat. Methods* **9**, 23–26 (2012).
2. Bogdanove, A.J. & Voytas, D.F. *Science* **333**, 1843–1846 (2011).
3. DeFrancesco, L. *Nat. Biotechnol.* **29**, 681–684 (2011).
4. Miller, J.C. *et al. Nat. Biotechnol.* **29**, 143–148 (2011).
5. Boch, J. *et al. Science* **326**, 1509–1512 (2009).
6. Moscou, M.J. & Bogdanove, A.J. *Science* **326**, 1501 (2009).
7. Deng, D. *et al. Science* **335**, 720–723 (2012).
8. Mak, A.N., Bradley, P., Cernadas, R.A., Bogdanove, A.J. & Stoddard, B.L. *Science* **335**, 716–719 (2012).
9. Zhang, F. *et al. Nat. Biotechnol.* **29**, 149–153 (2011).
10. Morbitzer, R., Römer, P., Boch, J. & Lahaye, T. *Proc. Natl. Acad. Sci. USA* **107**, 21617–21622 (2010).
11. Huang, P. *et al. Nat. Biotechnol.* **29**, 699–700 (2011).

PRRRAAQPSDASPAAQVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALG
 TVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAV
 EAVHAWRNALTGAPLNLTPAQVVAIASNI GGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASNI GGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQA
 LESIVAQLSRPDPALAAALNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHA
 QVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHLLQLFRRVGVTELEARSGLPPASQRWDRILQASGM
 KRAKPSPTSTQTPDQASLHAFADSLERDLADAPSPMHEGDQTRASSRKRSDRAVTGPSAQQSFEVVRV
 PEQRDALHLLPLLSWGVKRPRTTRIGLLDPGTPMDADLVASSTVVVEQDADPFAGTADDFPAFNEEELA
 WLMELLPQ*

^aRVDs are shown in green. XX indicates different RVDs according to Figure 1b.

SUPPLEMENTARY METHODS

Plant growth and inoculations

Nicotiana benthamiana plants were cultivated in the greenhouse with 16h light, 40-60% humidity and day/night temperatures of 23°C/19°C. Leaves of five- to seven-week old plants were inoculated with *Agrobacterium tumefaciens* GV3101 containing binary vector derivatives as described¹. Inoculated plants were transferred to a Percival growth chamber (Percival Scientific) with 22°C/18°C day/night temperatures and 16 h light.

Construction of artificial TAL effectors (ArtTALs)

ArtTAL_{polyXX} were constructed as described¹ using random ligation of single repeat *Esp3I* fragments into an ENTRY vector containing a Hax3 backbone. The number of repeats was analyzed by *StuI HincII* digest and TAL effectors with a total of 15.5 repeats chosen. The ArtTAL_{polyXX} were cloned by GATEWAY LR-recombination into pGWB6² for expression of N-terminal GFP-TAL effector fusions under control of the constitutive 35S cauliflower mosaic viral promoter. The repeat region of all constructs was sequenced.

ArtTAL_{6XX}, ArtTAL_{4XX}, ArtTAL_{3XX}, ArtTAL_{2XX}, were constructed using a modified Golden TAL Technology³. Individual repeats of all tested RVDs were subcloned for repeat positions one to six in a hexa-repeat array. The repeat backbone apart from the RVDs is identical (Supplementary Table 2). The first six repeats of Hax3 (NI HD NI HD HD HD) and the last 5.5 repeats containing the RVDs to be tested were assembled into two independent modules and inserted together with the Hax3 N- and C-terminal regions as N-terminal GFP fusion into a binary vector allowing expression of the constructs under control of the constitutive 35S promoter. ArtTAL_{NK}, ArtTAL_{NN}, ArtTAL_{NH}, were constructed using the original Golden TAL Technology in the same binary vector³.

GUS reporter constructs and reporter assay

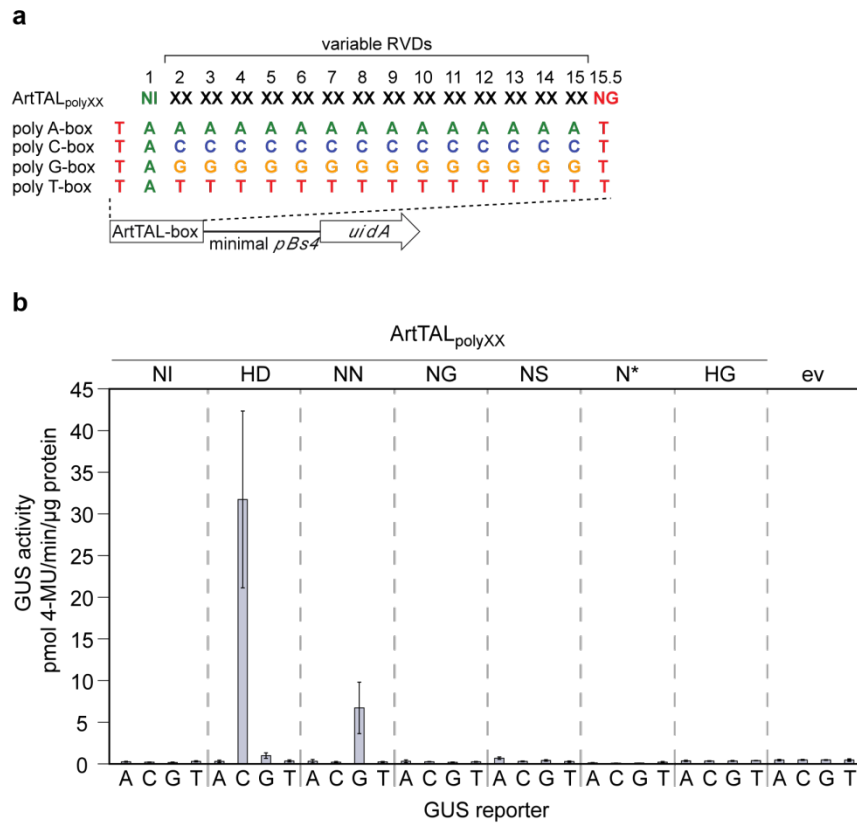
The ArtTAL target DNA boxes were amplified together with the minimal *Bs4* promoter and inserted into pENTR/D-TOPO (Invitrogen). The promoter derivatives were recombined into pGWB3² via LR recombination. Transient GUS reporter assays were performed as described¹.

Briefly, *Agrobacterium* strains delivering ArtTAL constructs and GUS reporter constructs were mixed 1:1 and inoculated into *N. benthamiana* leaves with a total OD₆₀₀ of 0.8. The constructs integrate into the plant chromosomes. Leaf disks (0.9 cm diameter) were sampled two days later and GUS activity determined. For qualitative GUS assays, leaf disks were stained in X-Gluc (5-bromo-4-chloro-3-indolyl- β -D-glucuronide) solution⁴, destained in ethanol and dried between acetate foil. For quantitative GUS assays two leaf disks were pooled, the plant tissue homogenized, diluted, and incubated with 4-methyl-umbelliferyl- β -D-glucuronide (MUG) as described⁵. Proteins were quantified using Bradford assay (BioRad). Triplicate samples from different plants were combined into one data point. All experiments were at least done twice with similar results.

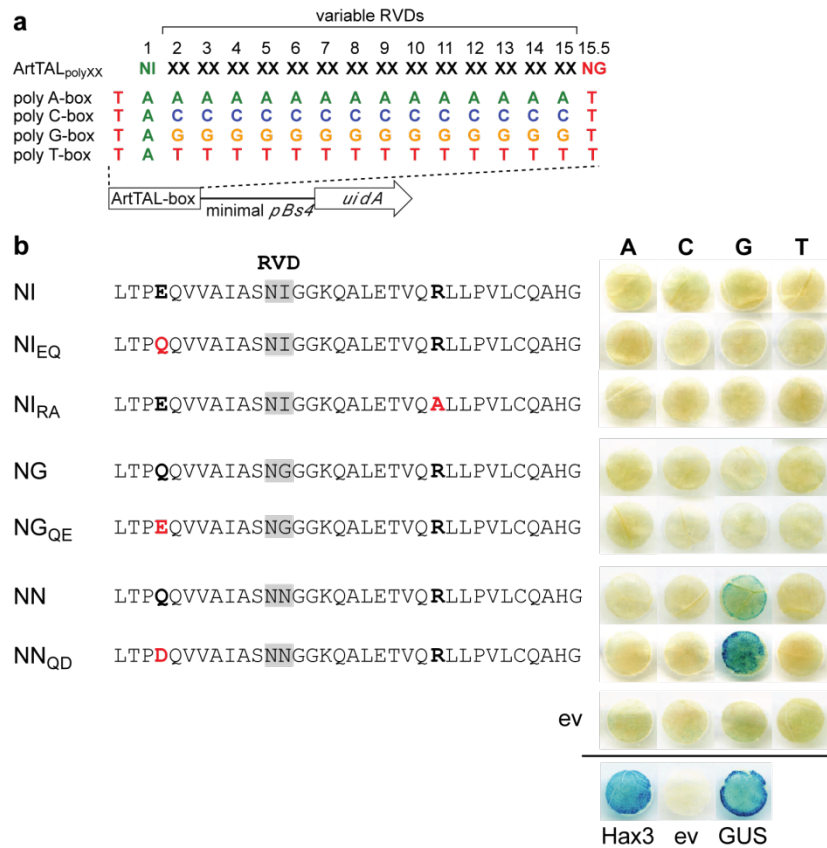
SUPPLEMENTARY REFERENCES

1. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512 (2009).
2. Nakagawa, T. *et al.* Development of series of Gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J Biosci Bioeng* **104**, 34-41 (2007).
3. Geißler, R. *et al.* Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**, e19509 (2011).
4. Schornack, S., Peter, K., Bonas, U. & Lahaye, T. Expression levels of *avrBs3*-like genes affect recognition specificity in tomato *Bs4*- but not in pepper *Bs3*-mediated perception. *Mol Plant-Microbe Interact* **18**, 1215-1225 (2005).
5. 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).

SUPPLEMENTARY FIGURES

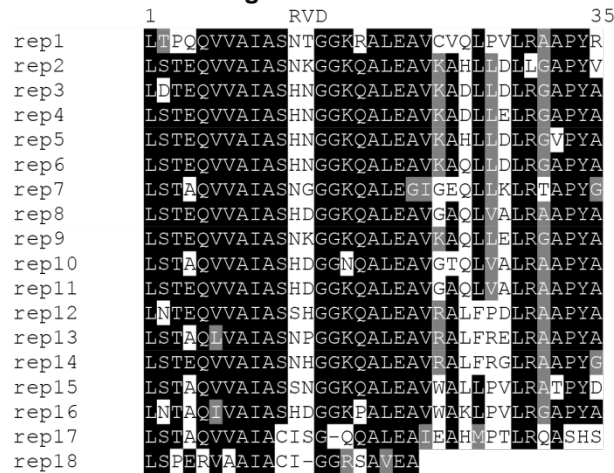


Supplementary Fig. 1 Weak and strong TAL effector RVDs. Quantitative data to Fig. 1a. (a) Artificial TAL effectors (ArtTAL_{polyXX}) composed of predominantly one repeat type were constructed to test specificities and efficiencies of RVDs. DNA boxes with polyA, -C, -G, or -T as target for the poly-RVD repeats were cloned in front of the minimal *Bs4* promoter into a GUS reporter vector. (b) *35S*-driven *artTALs* or empty T-DNA (ev) were codelivered via *A. tumefaciens* with GUS reporter constructs into *N. benthamiana* and quantitative GUS activity was measured (error bars indicate SD; n = 3 samples). 4-MU, 4-methyl-umbelliferone. Note that only two RVD-DNA base combinations (HD=C and NN=G) yield reporter activity although specificity of other RVDs (NI=A, NN=A, NG=T, NS=A/C/G/T, N*=T, HG=T) is known. This demonstrates the effect of strong and weak RVDs, respectively.



Supplementary Fig. 2 Variation of non-RVDs does not restore efficiency of weak RVDs. **(a)** Artificial TAL effectors (ArtTAL_{polyXX}) composed of predominantly one repeat type were constructed. DNA boxes with polyA, -C, -G, or -T as target bp for the poly-repeats were cloned in front of the minimal *Bs4* promoter into a GUS reporter vector. **(b)** 35S-driven *artTALs* or empty T-DNA (ev) were codelivered via *A. tumefaciens* with GUS reporter constructs into *N. benthamiana*. Leaf disks were collected 2 dpi and stained for qualitative GUS activity using X-Gluc. Amino acid substitutions (E4Q, R24A, Q4E, Q4D; bold and red letters) that are frequently found in nature were introduced into the poly-RVD repeats, respectively. Constitutive 35S-driven *uidA* (GUS), *hax3* (Hax3) or empty vector (ev) codelivered with Hax3 box in the reporter vector were used as controls. Note that none of the poly-NI or poly-NG ArtTALs produced reporter activity indicating that the weak efficiency of these RVDs can not be rescued by non-RVD exchanges. Activity of the strong RVD NN is not abolished by non-RVD exchange.

a *Ralstonia solanacearum* Brg11



R. solanacearum Brg11 RVDs:

NT, NK, HN, NG, HD, SH, NP, NH, SN, IS

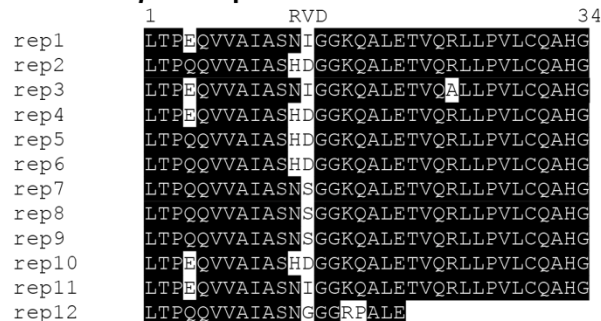
Xanthomonas repeat backbone:

LTPEQVVAIASXXGGKQALETVQRLLPVLCQAHG

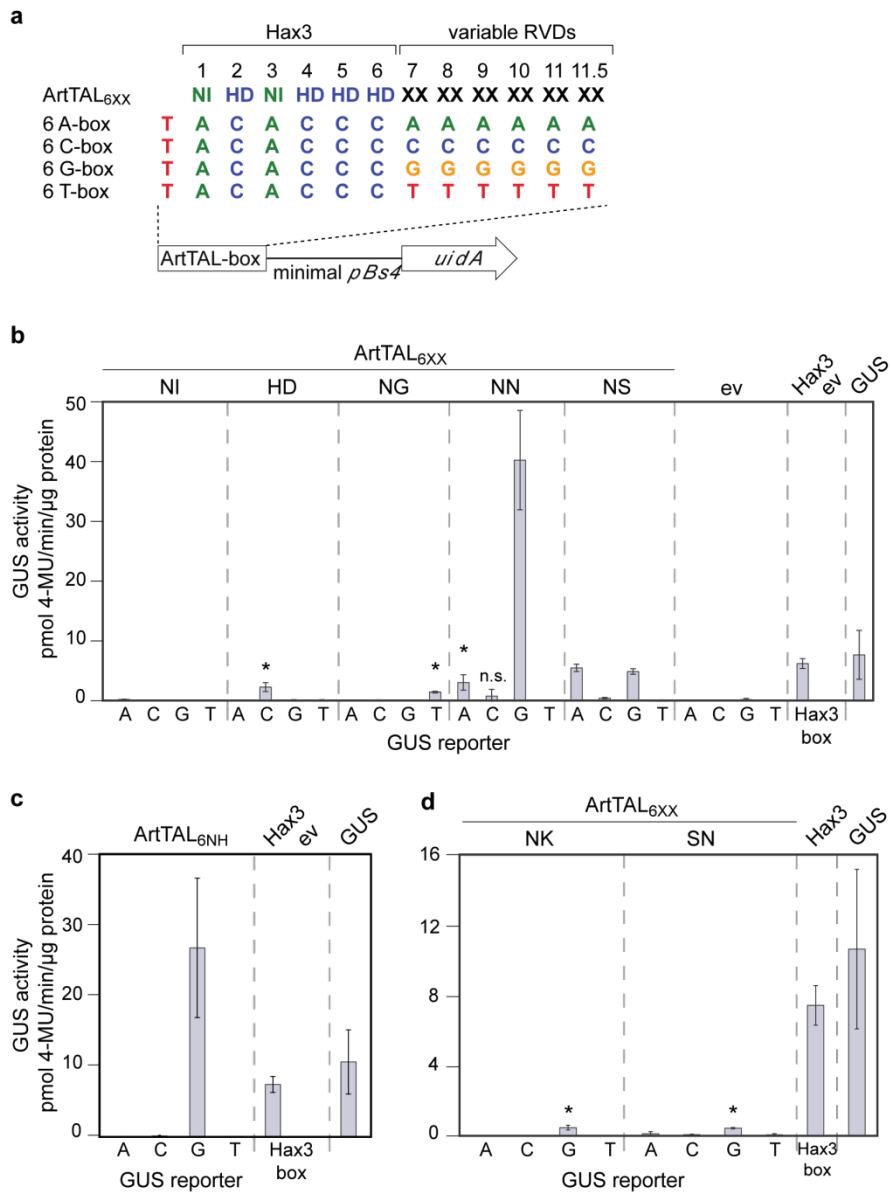
Predicted Brg11 target sequences:

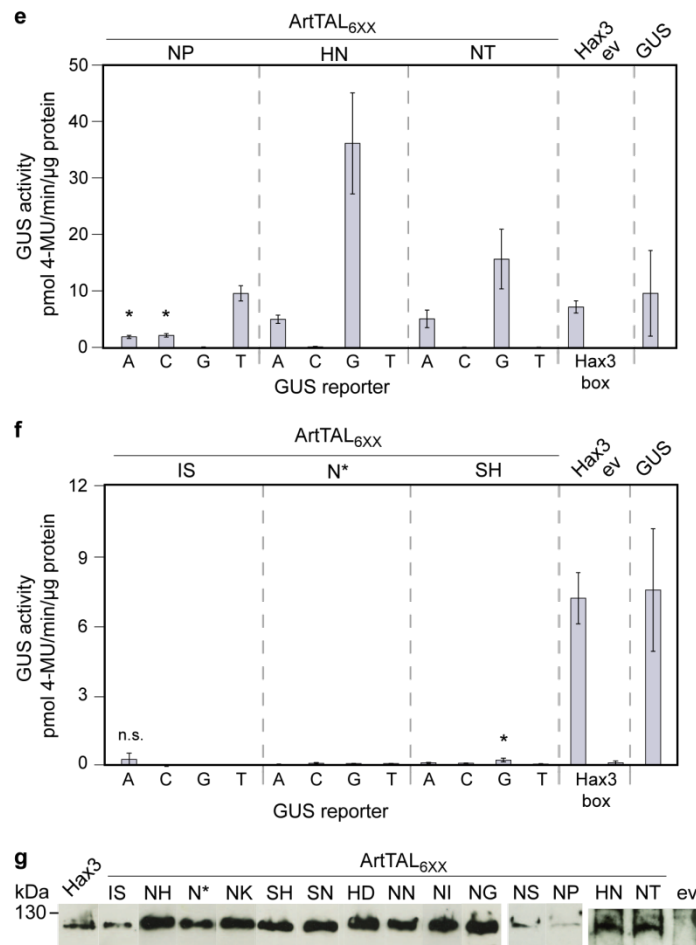
NT-NK-HN-HN-HN-HN-NG-HD-NK-HD-HD-SH-NP-NH-SN-HD-IS-IG
 G G G G G G T C G C C G T G G C N T
 A A A A A A A C

b *Xanthomonas campestris* pv. *armoraciae* Hax3

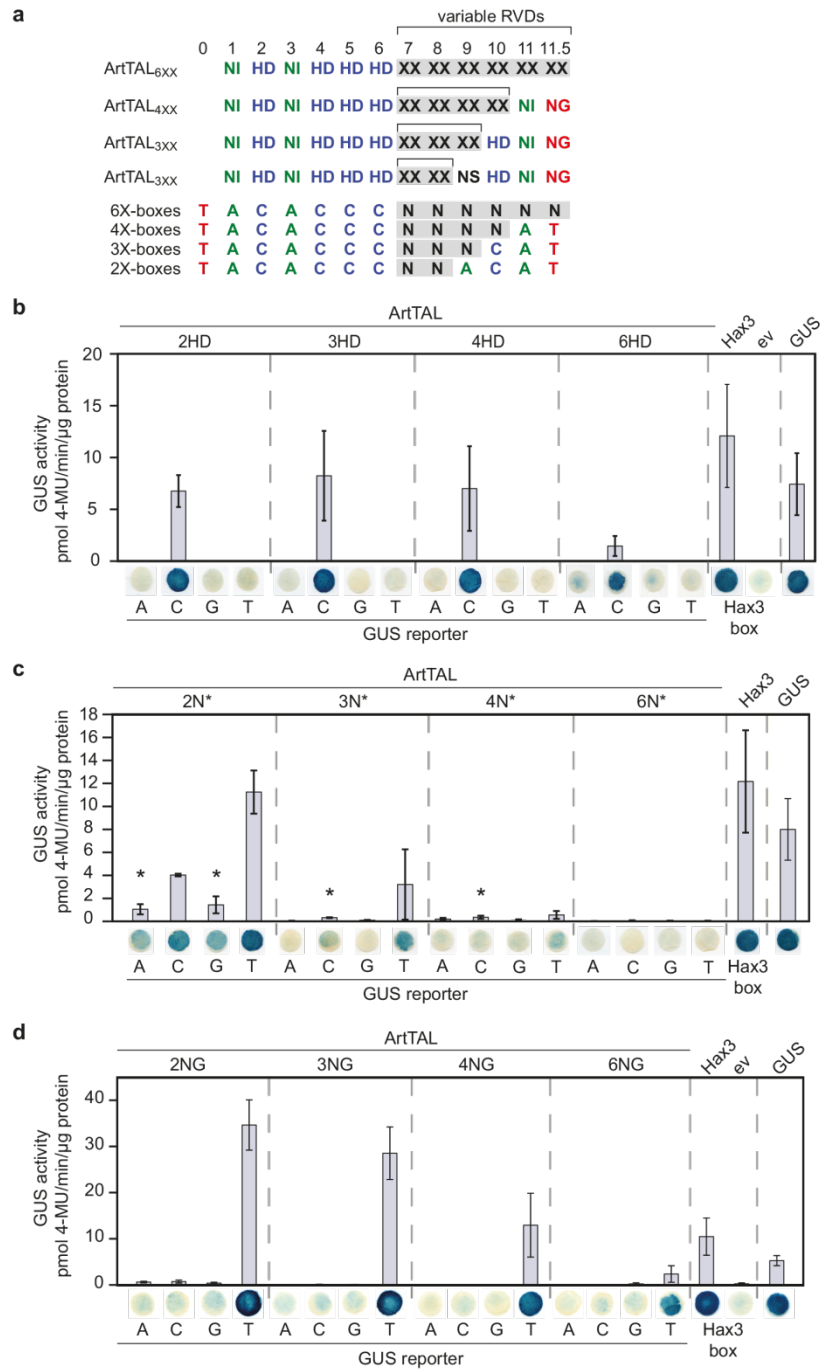


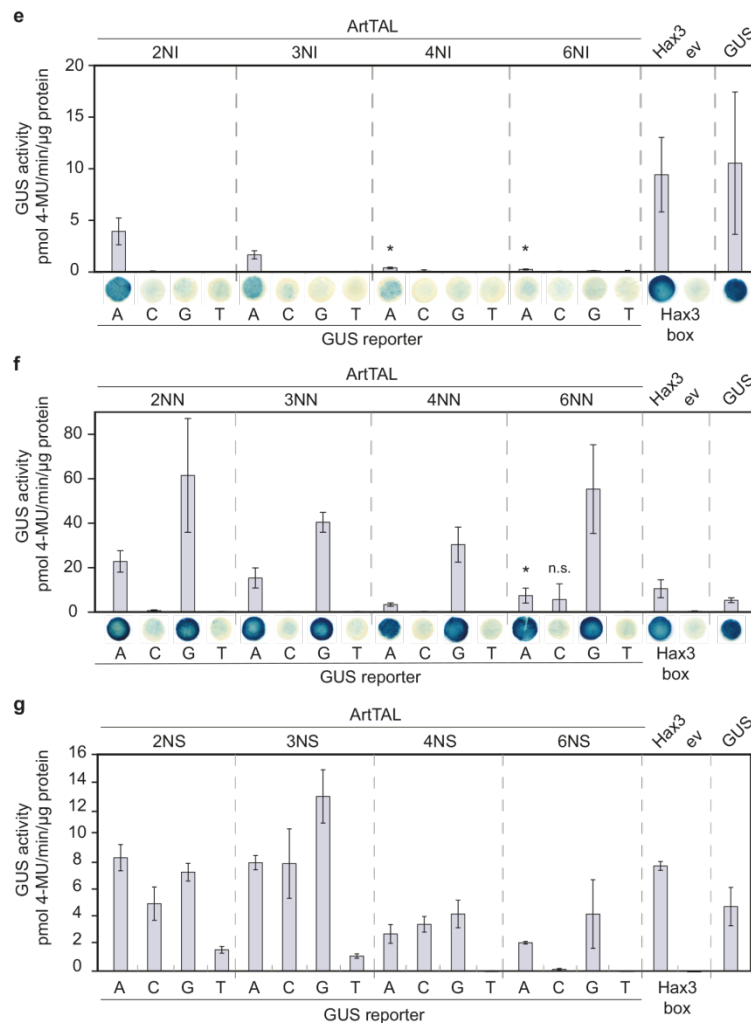
Supplementary Fig. 3 Comparison of *R. solanacearum* GMI1000 Brg11 (accession no. CAD15517.1) repeats and *X. campestris* pv. *armoraciae* Hax3 repeats. Repeat amino acids were aligned, and conserved amino acids shaded using BOXSHADE (ch.embnet.org). The first and last amino acids are numbered, and the RVDs are indicated above the alignment. **a)** *R. solanacearum* Brg11 repeats. Apparent RVDs from Brg11 were identified and inserted into the backbone of *Xanthomonas* TAL effector repeats (RVD positions indicated by red "XX") to determine their specificity. Experimentally derived specificities (Supplementary Fig. 4) were used to predict possible Brg11 target sequences. **b)** *X. campestris* pv. *armoraciae* Hax3 repeats.



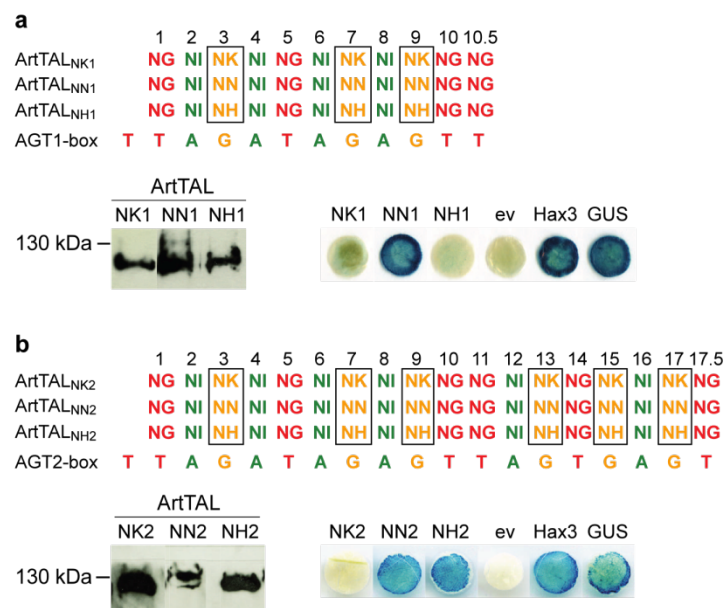


Supplementary Fig. 4 Specificities and efficiencies of RVDs. Quantitative data to Fig. 1b. **(a)** Artificial TAL effectors (ArtTAL_{6XX}) with the first six repeats of Hax3 and six RVDs to be tested were constructed. Target DNA boxes with A, C, G, or T at the last six positions were cloned in front of the minimal *Bs4* promoter into a GUS reporter vector. **(b-f)** Constitutive *35S*-driven *artTALs* or empty T-DNA (ev) were codelivered via *A. tumefaciens* with GUS reporter constructs into *N. benthamiana* and quantitative GUS activity measured (error bars indicate SD; n = 3 samples). 4-MU, 4-methyl-umbelliferone. *35S*-driven *uidA* (GUS), *hax3* (Hax3) or empty vector (ev) codelivered with Hax3 box in the reporter vector were used as controls. Low GUS activities were subjected to *t*-test analysis. Significance (*) or non-significance (n.s.) above empty vector control ($p < 0.05$) is indicated. Please note that the scale of the y-axis differs to display specificities of weak and strong RVDs. **(g)** ArtTAL_{6XX} protein integrity was determined by Western-blot using an anti-GFP antibody directed against the N-terminally fused GFP in all ArtTAL_{6XX} proteins. Please note that although overall ArtTAL_{6XX} protein amount differs, this does not correlate with ArtTAL_{6XX} activity, i.e. some ArtTAL_{6XX} with weak RVDs cause low GUS reporter activity, but have high protein levels.

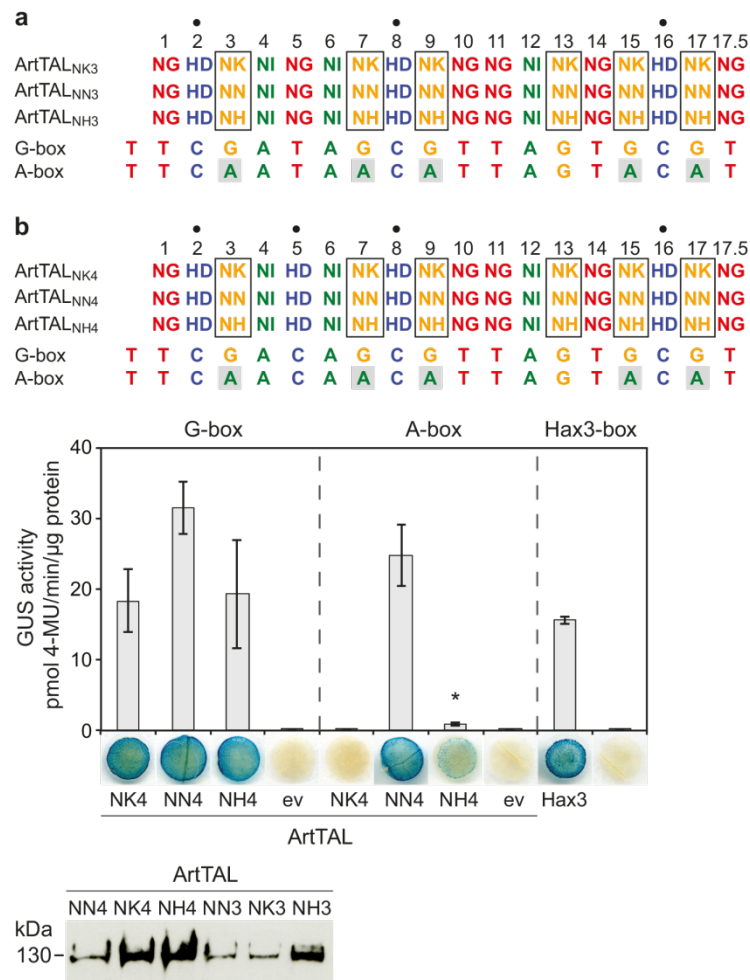




Supplementary Fig. 5 Efficiency of TAL effectors depends on the number of tandem repeats with weak RVDs. **(a)** Artificial TAL effectors ($\text{ArtTAL}_{6\text{XX}}$, $\text{ArtTAL}_{4\text{XX}}$, $\text{ArtTAL}_{3\text{XX}}$, $\text{ArtTAL}_{2\text{XX}}$) composed of Hax3 repeats with substituted six, four, three, or two repeats, respectively, with RVDs to be tested were constructed. Target DNA boxes with A, C, G, or T were cloned in front of the minimal *Bs4* promoter into a GUS reporter vector. **(b-g)** Constitutive *35S*-driven *artTALs* or empty T-DNA (ev) were codelivered with GUS reporter constructs via *A. tumefaciens* into *N. benthamiana*. Qualitative and quantitative GUS activity was determined (error bars indicate SD; $n = 3$ samples). 4-MU, 4-methyl-umbelliferone. *35S::uidA* (GUS) and Hax3 with the Hax3-box served as controls. *t*-test show significance (*) or non-significance (n.s.) above empty vector control ($p < 0.05$). Please note that the activity of ArtTALs depends on the RVD specificity, but also the number of tandem repeats with weak RVDs. Specificity of some RVDs (NN=A, N*, NS) appears to vary with the number of tandem identical RVDs which likely reflects different RVD efficiencies to the DNA bases.



Supplementary Fig. 6 Efficiency of different G-specific RVDs in mixed-RVD ArtTALs. ArtTALs with 10.5 (a) and 17.5 repeats (b), respectively, and mixed weak RVDs (NI, NG, NK) or substitutions (boxed) with more efficient G-specific RVDs (NN, NH) were assembled. The short and long ArtTALs, respectively, recognize the same DNA sequence. Constructs producing ArtTALs and GUS reporter constructs were codelivered into *N. benthamiana* leaf cells using *Agrobacterium*. Leaf disks were collected 2 dpi and stained for qualitative GUS activity using X-Gluc. Please note that mixed weak RVDs resulted in non-functional ArtTALs (ArtTAL_{NK1}, ArtTAL_{NK2}) whereas 3 (ArtTAL_{NN1}) or 6 (ArtTAL_{NN2}) strong NN RVDs, respectively, ensured ArtTAL activity. The intermediately strong RVD NH resulted in a functional ArtTAL (ArtTAL_{NH2}) if 17.5 repeats were used, but a non-functional one (ArtTAL_{NH1}) if only 10.5 repeats were used. Western blots using anti-GFP antibody directed against the N-terminally GFP fused to all ArtTALs indicate stable protein production.



Supplementary Fig. 7 Efficiency and specificity of ArtTALs with mixed and HD RVDs. ArtTALs with 17.5 repeats and mixed weak RVDs (NI, NG, NK), substitutions (boxed) with more efficient G-specific RVDs (NN, NH), and 3 (a) or 4 (b) HD RVDs (•) were assembled. Constructs producing ArtTALs and GUS reporter constructs were codelivered into *N. benthamiana* leaf cells using *Agrobacterium*. Leaf disks were stained for qualitative GUS activity and quantitative GUS activity was measured (error bars indicate SD; n = 3 samples). 4-MU, 4-methyl-umbelliferone. 35S-driven *uidA* (GUS), *hax3* (Hax3) or empty vector (ev) codelivered with Hax3-box in the reporter vector were used as controls. *t*-test show significance (*) above empty vector control (p<0.05). Please note that all ArtTALs were active on the G-box indicating that 3 and 4 HD RVDs efficiently rescue ArtTAL function (compare to Supplementary Fig. 6). Multiple exchanges of G to A in the target DNA box revealed that NK and NH RVDs are more specific for G than NN. Western blots using anti-GFP antibody directed against the N-terminally GFP fused to all ArtTALs indicate stable protein production.

SUPPLEMENTARY TABLES

Supplementary Table 1. RVD specificities and efficiencies^a

RVD ^a	Specificity ^b	Efficiency ^c
<i>Known RVD specificities</i>		
NI	A	weak
HD	C	strong
NN	G > A	strong (G), intermediate (A)
NK	G	weak
NS	G, A > C > T	intermediate
NG	T	weak
<i>Novel RVD specificities</i>		
NH	G	intermediate
N*	T > C >> G, A	weak
NP	T > A, C	intermediate ^d
HN	G > A	intermediate ^d
NT	G > A	intermediate ^d
SN	G	weak
SH	G	weak
IS	-	weak

^a Amino acids in one letter code. A star (*) indicates a missing amino acid in the alignment.

^b DNA base of leading strand. > and >>: decreasing efficiencies. -: no efficiency was detected

^c Strong RVDs were functional in all ArtTALs; intermediate RVDs supported ArtTAL_{6XX} function, but depended on the overall RVD composition of the TAL effector; weak RVDs were not functional in ArtTAL_{6XX}.

^d It was not determined whether RVDs are strong.

Supplementary Table 2. Analyzed repeats^a

RVD	Repeat sequence
ArtTAL_{polyXX}	
NI	LTPEQVVAIAS NI GGKQALETVQRLLPVLCQAHG
HD	LTPQVVAIAS HD GGKQALETVQRLLPVLCQAHG
NS	LTPQVVAIAS NS GGKQALETVQRLLPVLCQAHG
NG	LTPQVVAIAS NG GGKQALETVQRLLPVLCQAHG
NN	LTPQVVAIAS NN GGKQALETVQRLLPVLCQAHG
N*	LTPDQVVAIAS N* GGKQALETVQRLLPVLCQDHG
HG	LTPDQVVAIAS HG GGKQALETVQRLLPVLCQDHG
Repeats with non-RVD exchanges ^b	
NI _{EQ}	LTP Q VVAIAS NI GGKQALETVQRLLPVLCQAHG
NI _{RA}	LTPEQVVAIAS NI GGKQALETVQ A LLPVLCQAHG
NG _{QE}	LTP E QVVAIAS NG GGKQALETVQRLLPVLCQAHG
NN _{QD}	LTP D QVVAIAS NN GGKQALETVQRLLPVLCQAHG
ArtTAL_{6,4,3,2XX}	
NI	LTPEQVVAIAS NI GGKQALETVQRLLPVLCQAHG
HD	LTPEQVVAIAS HD GGKQALETVQRLLPVLCQAHG
NN	LTPEQVVAIAS NN GGKQALETVQRLLPVLCQAHG
NG	LTPEQVVAIAS NG GGKQALETVQRLLPVLCQAHG
NS	LTPEQVVAIAS NS GGKQALETVQRLLPVLCQAHG
NP	LTPEQVVAIAS NP GGKQALETVQRLLPVLCQAHG
HN	LTPEQVVAIAS HN GGKQALETVQRLLPVLCQAHG
NT	LTPEQVVAIAS NT GGKQALETVQRLLPVLCQAHG
NH	LTPEQVVAIAS NH GGKQALETVQRLLPVLCQAHG
NK	LTPEQVVAIAS NK GGKQALETVQRLLPVLCQAHG
SN	LTPEQVVAIAS SN GGKQALETVQRLLPVLCQAHG
IS	LTPEQVVAIAS IS GGKQALETVQRLLPVLCQAHG
N*	LTPEQVVAIAS N* GGKQALETVQRLLPVLCQAHG
SH	LTPEQVVAIAS SH GGKQALETVQRLLPVLCQAHG

^a amino acids in one letter code. A star (*) indicates a missing amino acid in the alignment. RVDs are in green.

^b non-RVD exchanges are in red.

2.1.1.2 Zusammenfassung Publikation 1

Die Entschlüsselung des TALE-RVD-Codes war ein bedeutender Schritt um die Funktion von TALEs als Transkriptionsaktivatoren, sowie deren Ziel-Gene besser zu verstehen und TALEs als nützliches Werkzeug für biotechnologische Anwendungen verfügbar zu machen. Im Rahmen dieser Arbeit wurden die Spezifitäten selten vorkommender *Xanthomonas*- bzw. *Ralstonia*-TALE-RVDs analysiert. Da TALE-Homologe aus *R. solanacearum* vermutlich ebenfalls mit Hilfe ihrer *repeat*-Region an DNA binden, sollte die Analyse der in *Ralstonia* vorkommenden RVD-Spezifitäten erstmals die Vorhersage einer Ziel-DNA-Sequenz für einen *Ralstonia*-TALE ermöglichen. Der modulare Aufbau von TALEs ermöglicht die Erstellung künstlicher TALEs mit nutzerdefinierter RVD-Abfolge. Zur Analyse der RVD-Spezifitäten wurden daher künstliche TALEs erstellt, die eine gewisse Anzahl des zu untersuchenden RVDs in Folge enthielten. Die Aktivität dieser TALEs wurde anhand verschiedener Reporterkonstrukte analysiert, welche an den entsprechenden Positionen des zu untersuchenden RVDs die Nukleotide Adenin, Cytosin, Guanin oder Thymin besaßen. Anhand der künstlichen TALEs konnte die Spezifität für 8 der 9 analysierten RVDs bestimmt werden und so z.B. ausschließlich Guanin-spezifische RVDs (z.B. NH, NK) identifiziert werden. Diese könnten zukünftig das Purin-spezifische RVD NN ersetzen, um z.B. die *off-target* Rate von TALE-Werkzeugen im Ziel-Genom zu verringern. Außerdem konnte nun erstmals anhand der RVD-Abfolge eine mögliche Ziel-DNA-Sequenz des *R. solanacearum* TALEs Brg11 abgeleitet werden. Weiterhin wurden erstmalig einzelne TALEs beschrieben, die einen Reporter trotz perfekter Ziel-DNA-Sequenz nicht aktivieren konnten. Interessanterweise konnte dies auf die RVD-Komposition des jeweiligen TALEs zurückgeführt werden. TALEs die nur eine Auswahl bestimmter RVDs (z.B. NI, NG, NK) enthielten konnten das Reporter-konstrukt nicht aktivieren. Die Aktivität dieser TALEs konnte jedoch durch die Insertion von HD- oder NN-RVDs wiederhergestellt werden. Eine mögliche Erklärung dafür könnte sein, dass die Eigenschaften bestimmter RVDs verschieden sind und damit die Gesamt-TALE-Aktivität unterschiedlich beeinflussen. Einige RVDs z.B. HD oder NN sind starke RVDs mit einem positiven Einfluss auf die TALE-Aktivität, währenddessen sind andere RVDs z.B. NI, NG oder NK schwach und haben lediglich einen geringen oder keinen Einfluss auf die TALE-Gesamt-Aktivität. Diese Erkenntnisse erweiterten den ursprünglichen TALE-RVD-Code und ermöglichten die

Ableitung von Richtlinien zur Erstellung effizienter TALEs mit ausgewogener RVD-Komposition.

2.2 Gruppe III *OsSWEETs* sind wichtige Suszeptibilitätsgene für *Xoo*-TALEs

2.2.1 Publikation 2



Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*

Jana Streubel^{1*}, Céline Pesce^{2,3*}, Mathilde Hutin², Ralf Koebnik², Jens Boch¹ and Boris Szurek²

¹Department of Genetics, Martin Luther University Halle-Wittenberg, Weinbergweg 10, D-06120, Halle (Saale), Germany; ²UMR 186 IRD-UM2-Cirad 'Résistance des Plantes aux Bioagresseurs', BP 64501, 34394, Montpellier Cedex 5, France; ³Earth and Life Institute, Applied Microbiology Phytopathology, Université Catholique de Louvain, 1348, Louvain-la-Neuve, Belgium

Authors for correspondence:
Boris Szurek
Tel: +33 467 416212
Email: boris.szurek@ird.fr

Jens Boch
Tel: +49 345 5526292
Email: jens.boch@genetik.uni-halle.de

Received: 17 April 2013
Accepted: 3 June 2013

New Phytologist (2013) **200**: 808–819
doi: 10.1111/nph.12411

Key words: designer TALE, plant pathogen, plant resistance, rice, TALE, type-III effector, virulence, *OsSWEET*.

Summary

- Bacterial plant-pathogenic *Xanthomonas* strains translocate transcription activator-like (TAL) effectors into plant cells to function as specific transcription factors. Only a few plant target genes of TAL effectors have been identified, so far. Three plant *SWEET* genes encoding putative sugar transporters are known to be induced by TAL effectors from rice-pathogenic *Xanthomonas oryzae* pv. *oryzae* (*Xoo*).
- We predict and validate that expression of *OsSWEET14* is induced by a novel TAL effector, Tal5, from an African *Xoo* strain. Artificial TAL effectors (ArtTALs) were constructed to individually target 20 *SWEET* orthologs in rice. They were used as designer virulence factors to study which rice *SWEET* genes can support *Xoo* virulence.
- The Tal5 target box differs from those of the already known TAL effectors TalC, AvrXa7 and PthXo3, which also induce expression of *OsSWEET14*, suggesting evolutionary convergence on key targets. ArtTALs efficiently complemented an *Xoo talC* mutant, demonstrating that specific induction of *OsSWEET14* is the key target of TalC. ArtTALs that specifically target individual members of the rice *SWEET* family revealed three known and two novel *SWEET* genes to support bacterial virulence.
- Our results demonstrate that five phylogenetically close *SWEET* proteins, which presumably act as sucrose transporters, can support *Xoo* virulence.

Introduction

Plant pathogens threaten the production of food crops worldwide. The bacterial pathogen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) causes bacterial leaf blight (BLB), one of the most devastating diseases of rice (Nino-Liu *et al.*, 2006), which is a primary food plant for half of the world's population. Virulence of *Xoo* relies on injection of a cocktail of effector proteins into plant cells via a specialized type III secretion system (White & Yang, 2009). TAL (transcription activator-like) effectors constitute a major *Xoo* effector family as many as 19 genes per strain (Salzberg *et al.*, 2008). TAL effectors employ tandem, nearly identical 34-amino-acid repeats to directly bind specific DNA sequences in promoter regions. Guided by a repeat variable diresidue (RVD) specific for one of the four DNA bases, or combinations thereof, each repeat binds one base pair in a contiguous sequence (Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Boch & Bonas, 2010). The DNA-binding domain, together with a C-terminal activation domain,

enables TAL effectors to function as specific transcription factors that induce expression of target plant genes to the benefit of the pathogen (Boch & Bonas, 2010). The modular TAL effector architecture also allows one to artificially rearrange repeats to generate artificial TAL effectors with novel, predictable DNA specificities (Boch *et al.*, 2009). Artificial TAL effectors (also called designer TAL effectors) have been used to analyze the specificities and efficiencies of RVDs (Boch *et al.*, 2009; Christian *et al.*, 2012; Cong *et al.*, 2012; Streubel *et al.*, 2012) and induce target gene expression in different organisms (Morbitzer *et al.*, 2010; Geißler *et al.*, 2011; Miller *et al.*, 2011; Zhang *et al.*, 2011; Bultmann *et al.*, 2012; Li *et al.*, 2013; Maeder *et al.*, 2013; Perez-Pinera *et al.*, 2013). Consequently, TAL effectors have become widely used in biotechnology as specific activators, repressors, and nucleases in many different organisms (Bogdanove & Voytas, 2011; Mussolino & Cathomen, 2012).

Several TAL effectors contribute significantly to bacterial virulence (Boch & Bonas, 2010), but in general little is known about their biological virulence targets. Because virulence of most *Xoo* strains mainly relied on the presence of individual TAL effectors and thus the ability to induce specific host genes, these targets

*These authors contributed equally to this work.

have been termed susceptibility genes (White & Yang, 2009). The rice *SWEET*/*nodulin-3* gene family holds the best-studied examples of TAL effector virulence targets so far. Two *SWEET* genes, *Os8N3* and *Os11N3* (also called *OsSWEET11* and *OsSWEET14*, respectively), have been identified as targets of four different TAL effectors from various *Xoo* strains (Chu *et al.*, 2006a; Yang *et al.*, 2006; Antony *et al.*, 2010; Chen *et al.*, 2012). A third *SWEET* gene, *Xa25*, was also shown to be induced after infection with *Xoo* strain PXO339 (Liu *et al.*, 2011), and it is possible that *Xa25* is induced by a hitherto not identified TAL effector. The *SWEET* proteins form a heterogeneous family, which is divided into at least four clades in plants (Chen *et al.*, 2010). *SWEET*s are membrane proteins that are involved in diverse functions, such as pollen development (Guan *et al.*, 2008), senescence (Quirino *et al.*, 1999), and microbe–plant interactions (Gamas *et al.*, 1996; Chu *et al.*, 2006a,b; Yang *et al.*, 2006; Antony *et al.*, 2010; Liu *et al.*, 2011). *Arabidopsis* and rice genomes include 17 and 22 paralogs, respectively. By contrast, chordates and arthropods contain only one and *Caenorhabditis elegans* seven *SWEET* genes, respectively. This suggests that *SWEET*s may have undergone higher diversification in plants to potentially accompany the evolution of the vascular system (Baker *et al.*, 2012). Recent data demonstrated that some members of the *SWEET* family facilitate sucrose and glucose efflux from the phloem parenchyma cells to the apoplast (Chen *et al.*, 2010, 2012). This gave rise to the hypothesis that the pathogen-mediated induction of certain *SWEET* genes might increase sugar availability in the apoplast, and therefore plays a role in *Xoo* nutrition. Members of the *SWEET* gene family in *Arabidopsis* are also induced after infection with *Pseudomonas syringae* strain DC3000 (Chen *et al.*, 2010), which indicates that this plant gene family has a general function for pathogen proliferation in the host.

Here we describe that a novel TAL effector, *Tal5* from *Xoo*, also targets the rice *OsSWEET14* gene. This observation provides compelling evidence that evolution of TAL effectors can converge on the same or highly related virulence targets. The importance of *SWEET*s for susceptibility towards *Xoo* prompted us to systematically study the role of this gene family for *Xoo* virulence. We constructed a series of artificial TAL effectors, which were designed to induce expression of individual *SWEET* genes in rice, and analyzed the ability of these TAL effectors to function as artificial virulence factors for *Xoo*. Our findings demonstrated that only the particular subgroup of rice *SWEET*s that are predicted to be sucrose transporters can be exploited as *Xoo* virulence targets. They also indicated that diverse *SWEET* family members probably fulfill different physiological functions. Finally, our approach demonstrates that artificial TAL effectors are suitable tools to study if candidate host genes are bona fide virulence targets for *Xanthomonas*.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used in this study were *Escherichia coli* DH5 α (Stratagene, La Jolla, CA, USA), *Agrobacterium tumefaciens*

GV3101 (Van Larebeke *et al.*, 1974) and *Xoo* strains BAI3 (Gonzalez *et al.*, 2007), BAI3 Δ *talC* (Yu *et al.*, 2011) and PXO99^A (Hopkins *et al.*, 1992). *E. coli* cells were cultivated at 37°C in lysogenic broth (LB) medium, *Xoo* strains at 28°C in PSA medium (10 g peptone, 10 g sucrose, 1 g glutamic acid, 16 g agar, 1⁻¹ H₂O) and *A. tumefaciens* GV3101 at 28°C in YEB medium. Plasmids were introduced into *E. coli* and *A. tumefaciens* strains by electroporation and into *Xoo* by conjugation using pRK2013 as a helper plasmid in triparental mating (Figurski & Helinski, 1979). Rifampicin- and gentamicin-resistant clones were selected upon plating on PSA medium and one isolate was chosen for further experiments. Antibiotics were added to the medium at the following final concentrations: rifampicin, 100 μ g ml⁻¹; gentamicin, 20 μ g ml⁻¹.

Plant material and plant inoculations

Experiments were performed under glasshouse conditions under cycles of 12 h of light at 28°C and 80% relative humidity (RH) and 12 h of dark at 25°C and 70% RH. *Oryza sativa* ssp. *japonica* cv Nipponbare and the near-isogenic *O. sativa* ssp. *indica* lines IR24 and IRBB13 were used for virulence assays, growth curves and quantitative reverse transcription polymerase chain reaction (qRT-PCR). Leaves of 3-wk-old plants were infiltrated with a bacterial suspension with an optical density at 600 nm (OD₆₀₀) of 0.5 using a needleless syringe, as previously described (Reimers & Leach, 1991), and symptoms (water-soaked lesions) were scored 6 d postinoculation (dpi). Leaf-clip inoculation was performed on 4- to 5-wk-old rice plants using a bacterial suspension with an OD₆₀₀ of 0.2 (Kauffman *et al.*, 1973), and sizes of lesions were measured 15 dpi. Statistical significance of the results was assessed using the Tukey honest significant difference test for post-ANOVA pairwise comparisons, set at 5% ($P < 0.05$).

Nicotiana benthamiana plants were grown under 16 h of light, 40–60% RH, at 23 : 19°C, day : night in the growth chamber. Leaves of 4- to 6-wk-old plants were inoculated with *A. tumefaciens* strains using a needleless syringe.

β -Glucuronidase (GUS) reporter constructs and GUS assays

β -Glucuronidase assays from plant samples were essentially performed as described by Boch *et al.* (2009). Briefly, coding regions of *talC* and *tal5* were inserted into pGWB2 under control of a 35S promoter (Nakagawa *et al.*, 2007). A PCR-amplified 341 bp fragment containing the *OsSWEET14* promoter region was cloned into pENTR/D-TOPO (Life Technologies GmbH, Darmstadt, Germany) and then fused to the *uidA* reporter gene by LR recombination into pGWB3 (Nakagawa *et al.*, 2007; Yu *et al.*, 2011).

To assay reporter activity, *A. tumefaciens* strains delivering TAL effector constructs and GUS reporter constructs were resuspended in infiltration medium (0.1 M 2-(N-morpholino) ethanesulfonic acid (MES), 1 M MgCl₂, 0.1 M acetosyringone), mixed in equal amounts and inoculated into *N. benthamiana* leaves at a total OD₆₀₀ of 0.8. Two days postinfiltration, leaf discs were

sampled and GUS activities were quantified using 4-methylumbelliferyl- β -D-glucuronide (MUG). Protein concentrations were determined using Bradford assays. Data were compiled from triplicate samples originating from different plants.

RNA isolation and qRT-PCR

Leaves of 3-wk-old rice plants of the varieties Nipponbare, IR24 or IRBB13 were infiltrated with water or with the different *Xoo* strains using an OD₆₀₀ of 0.5. At 1 dpi, leaf segments from three plants were ground into a fine powder using the Qiagen TissueLyser system (30 rps for 30 s). Rice total RNA was isolated using the Qiagen RNeasy kit following the manufacturer's recommendations (Qiagen). cDNA was generated from 2 μ g RNA using the Fermentas first-strand cDNA synthesis kit following the manufacturer's recommendations (Thermo Fisher Scientific Inc., Waltham, MA, USA). Real-time PCR was performed using the iCycler (Bio-Rad, München, Germany) and 10 μ l (50%) ABso-lute qPCR SYBR Green Fluorescein Mix (ABgene Limited, Hamburg, Germany), 1 pmol forward, 1 pmol reverse oligonucleotide primer and 12.5 ng template cDNA. For each gene, a minimum of two independent biological replicates were analyzed, each with two technical qPCR replicates. The specificity of the primer pairs was checked by a melting curve and the PCR product was documented by agarose gel electrophoresis. The amplification efficiency for each primer pair was analyzed using a standard curve plot of a dilution series. cDNA amounts were normalized using actin as a reference gene. All RT primer sequences and corresponding amplification efficiencies are provided in Supporting Information, Table S1.

In planta growth assays

Leaves of 3-wk-old IR24 or IRBB13 plants were inoculated with a bacterial suspension of *Xoo* strains using an OD₆₀₀ of 0.2. A 1 cm² infected leaf segment was collected 6 d after infiltration, and ground into a fine powder using the Qiagen

TissueLyser system (30 rps for 30 s). Ground material was resuspended in 1 ml sterile water, and 5 μ l drops of a dilution series were spotted as triplicates onto selective PSA plates containing rifampicin and gentamicin. This experiment was performed three times.

Expression analysis by Western blotting

The *X. oryzae* pv. *oryzae* BAI3 Δ *talC* strains carrying a plasmid encoding an artificial or native TAL effector, or empty vector, were grown in liquid PSA medium supplemented with 20 μ g ml⁻¹ gentamicin at 30°C. Cells of 1 ml of a bacterial suspension at an OD₆₀₀ of 0.2 were harvested and the TAL effector expression was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting using an anti-Flag antibody (Sigma-Aldrich).

Results

Tal5 from *Xoo* strain MAI1 is a novel TAL effector that targets *OsSWEET14*

We analyzed the TAL effector repertoire of an African *Xoo* isolate from Mali, MAI1, which harbors a relatively small set of only eight *TAL* genes (Gonzalez *et al.*, 2007). We screened a MAI1 genomic DNA cosmid library for clones containing *TALs* and subcloned eight *Bam*HI fragments of different sizes containing putative *TAL* genes. Sequencing revealed that one of them, named *tal5*, encodes a new TAL effector with a unique order of 17.5 DNA-binding repeats (Fig. 1a).

Using a program to scan rice promoters for potential TAL effector binding sites (Pérez-Quintero *et al.*, 2013), we predicted which candidate rice genes are targeted by Tal5 (Table S2). Interestingly, the top score was obtained for a binding site located within the promoter region of the gene *Os11g31190* (hereafter *OsSWEET14*), which is a member of the *SWEET1/nodulin-3* gene family (Baker *et al.*, 2012). Remarkably, *OsSWEET14* is a host

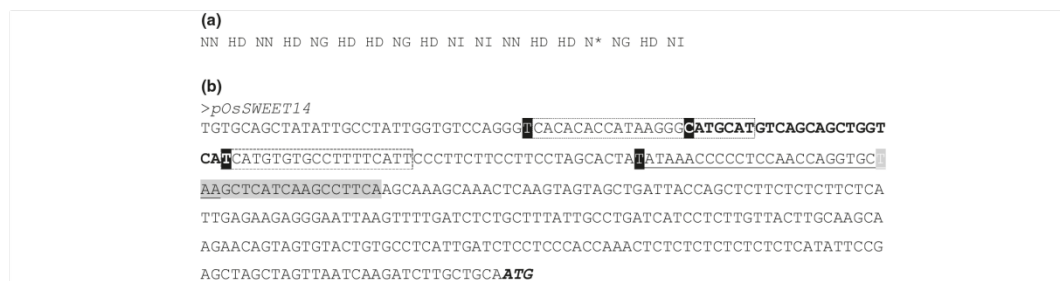


Fig. 1 Tal5 is a novel transcription activator-like (TAL) effector predicted to target *OsSWEET14*. (a) Hypervariable amino acids at positions 12 and 13 (repeat variable diresidue) of the 17.5 repeats of Tal5. A single amino acid code was used for each amino residue (* represents a missing 13th residue). (b) The 370 bp sequence of the *OsSWEET14* promoter region is depicted by the TalC box in bold (Yu *et al.*, 2011), the AvrXa7 box underlined (Römer *et al.*, 2010) and the candidate Tal5 box highlighted in gray. The artificial TAL effector (ArtTAL) boxes for ArtTAL14-1 and ArtTAL14-2 are indicated by dotted and dashed frames, respectively. The 5' terminal nucleotide of each target box is indicated in white font. The predicted start codon of *OsSWEET14* is indicated in bold italic font.

susceptibility gene that was previously reported as a target of the major virulence TAL effectors AvrXa7, PthXo3 and TalC from *Xoo* strains PXO86, JXO1^A and BAI3, respectively (Chu *et al.*, 2006; Antony *et al.*, 2010; Yu *et al.*, 2011). The target boxes for TalC and AvrXa7/PthXo3 are upstream of and overlapping the predicted *OsSWEET14* TATA box, respectively. By contrast, the predicted Tal5 box is located downstream of the TATA box (Fig. 1b).

To test if *OsSWEET14* is a direct virulence target of Tal5, we performed two complementary assays. First, we used a *talC* deletion mutant of *Xoo* BAI3, BAI3 Δ *talC*, which is impaired in *OsSWEET14* induction, resulting in fewer symptoms and reduced proliferation of the bacteria *in planta* compared with the parental strain (Yu *et al.*, 2011). Leaves of the susceptible rice variety IR24 were infiltrated with derivatives of BAI3 Δ *talC* containing the empty cloning vector or plasmids that encode *talC* or *tal5* (Methods S1). Water and wildtype bacteria of strain BAI3 served as negative and positive controls, respectively. qRT-PCR analyses showed that *OsSWEET14* was induced by both TalC and Tal5, but not upon inoculation of rice leaves with BAI3 Δ *talC* carrying the empty vector (Fig. 2b). Second, we used *Xoo* PXO99^A, which is not able to cause disease on *xa13* rice lines (e.g. IRBB13) as a result of a deletion in the PthXo1 target box which prevents PthXo1-mediated induction of the susceptibility gene *OsSWEET11* (Yang *et al.*, 2006). Similar to the previous experiment, only *Xoo* PXO99^A containing *tal5* or *talC* induced *OsSWEET14* expression in *xa13* rice lines (Fig. 2b).

To analyze the functionality of the putative target box for Tal5, a 341 bp fragment containing the *OsSWEET14* promoter region was cloned in front of a promoterless *uidA* reporter gene (Yu *et al.*, 2011). The reporter construct and *tal5* under control of a constitutive promoter were cotransformed into *N. benthamiana* using *Agrobacterium*. GUS assays showed that the *OsSWEET14* promoter fragment was recognized by Tal5, thus leading to reporter gene activity (Fig. 2c).

Finally, to analyze whether or not Tal5-mediated *OsSWEET14* induction affects infection, IR24 and IRBB13 rice lines were challenged with BAI3 Δ *talC* and PXO99^A derivatives, respectively, and disease symptoms were scored upon leaf infiltration and leaf clipping. As shown in Fig. 3(a), *tal5* partially complemented the BAI3 Δ *talC* mutant strain. Similarly, introduction of *tal5* into PXO99^A resulted in increased aggressiveness in IRBB13 lines (Fig. 3b). Altogether, our data indicate that Tal5 can function as a virulence factor in *Xoo* by up-regulating *OsSWEET14*, and potentially facilitates propagation of *Xoo* strain MAI1 in rice leaves.

Artificial TAL effectors targeting *OsSWEET14* effectively substitute for the virulence contributions of TalC and PthXo1

Although TAL effectors have a high target site specificity, they nevertheless induce the expression of a significant number of genes in host cells (Marois *et al.*, 2002; Kay *et al.*, 2009). While some of them are directly induced, others are up-regulated indirectly. Collectively, most of the (directly and indirectly) induced genes probably represent collateral targets that are not related to

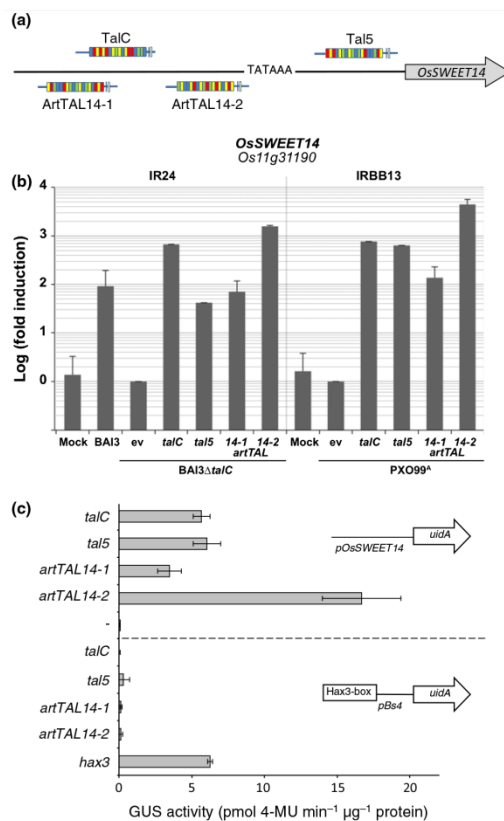


Fig. 2 *OsSWEET14* is induced by Tal5 and artificial transcription activator-like (ArtTAL) effectors. (a) Schematic representation of the *OsSWEET14* promoter region and associated natural TAL and ArtTAL effectors next to their target boxes. The coding sequence is represented by an open arrow and the TATA box is indicated. (b) *OsSWEET14* expression levels. Leaves of *Oryza sativa* IR24 variety were infiltrated with water (Mock), *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strain BAI3, and *Xoo* BAI3 Δ *talC* derivatives carrying the empty cloning vector (ev) or plasmids containing the TAL effector genes *tal5*, *talC*, *artTAL14-1* or *artTAL14-2*. Rice IRBB13 leaves were infiltrated with water or with *Xoo* strain PXO99^A derivatives carrying the same constructs as used for BAI3 Δ *talC*. *OsSWEET14* transcript abundances were determined 1 d postinoculation (dpi) by quantitative reverse transcription polymerase chain reaction (qRT-PCR). The bars represent a minimum of two biological samples taken from a collection of three leaves. The fold change was calculated in comparison to leaves treated with BAI3 Δ *talC* (ev) or PXO99^A (ev). Error bars represent + SD. Actin was used as a reference gene to normalize cDNA amounts. (c) The *OsSWEET14* promoter is recognized by Tal5 and ArtTALS. A 341 bp fragment of the *OsSWEET14* promoter region was cloned into a GUS reporter vector and codelivered via *Agrobacterium tumefaciens* into *Nicotiana benthamiana* leaves, along with 35S-driven *talC*, *tal5*, *artTAL14-1*, *artTAL14-2* or *hax3* (error bars indicate \pm SD; n = 3). MU, 4-methyl-umbelliferone. The Hax3-box cloned in front of the minimal *pBs4* promoter served as the specificity control.

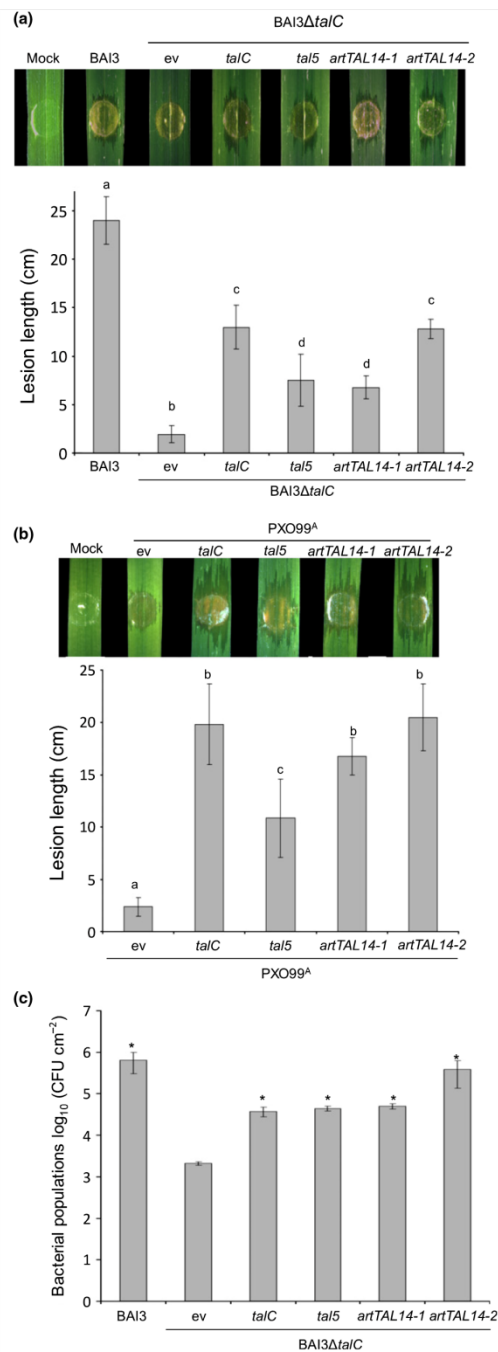


Fig. 3 Tal5 and artificial transcription activator-like (TAL) effectors act as virulence factors. Phenotypes of *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strains on *Oryza sativa* leaves. (a) Leaves of IR24 plants were inoculated with strains BAI3 or BAI3ΔtalC carrying empty vector (ev) or TAL effectors talC, tal5, artTAL14-1, or artTAL14-2. (b) Leaves of IRBB13 plants were inoculated with strain PXO99^A carrying the same constructs as used for BAI3ΔtalC. (a, b) Leaves were photographed 6 d postinoculation (dpi). Lesion length was measured at 15 d after leaf-clip inoculation. Data are the mean of eight measurements. Error bars represent \pm SD. Bars with same letters are not statistically different based on a Tukey's honest significant difference test ($\alpha = 0.05$). (c) Bacterial growth in IR24 rice leaves. Bacterial counts in leaf segments infiltrated with *Xoo* strains BAI3 or BAI3ΔtalC carrying empty vector (ev) or TAL effectors talC, tal5, artTAL14-1, or artTAL14-2 were analyzed 6 dpi. Error bars represent SD ($n = 3$). An asterisk indicates a significant difference in Student's *t*-test at $P < 0.01$ when comparing samples with BAI3ΔtalC (ev).

the infection process. Therefore, it is difficult to know which of these genes is responsible for the virulence effect of a given TAL effector and thus represents a bona fide host susceptibility gene.

We aimed to develop a functional test to determine if *OsSWEET14* is indeed a valid virulence target for *Xoo*. Artificial TAL effectors can be constructed with a designed repeat array and a tailored DNA-binding specificity to induce expression of specified target genes (Morbitzer *et al.*, 2010; Bogdanove & Voytas, 2011). We constructed two artificial TAL effectors (ArtTAL14-1 and ArtTAL14-2) that target sequences in the *OsSWEET14* promoter and that are different from those that are recognized by the four natural TAL effectors (Methods S2; Figs 1b, 2a), for more information on the design of the artificial TAL effectors, see Methods S1 and S2. The rationale was that both sets, ArtTALS and natural TAL effectors, will induce expression of *OsSWEET14*, but it is unlikely that they have common collateral targets. Indeed, the prediction of target sites in rice for the ArtTALS, Tal5, and TalC supported the fact that they probably share no target genes besides *OsSWEET14* (Table S3). We numbered the artificial TAL effectors according to their target *OsSWEET* paralog (i.e. ArtTAL14 for *OsSWEET14*).

As envisaged, both ArtTALS induced expression of *OsSWEET14* when delivered by *Xoo* BAI3ΔtalC and PXO99^A, respectively (Fig. 2b). In addition, both ArtTALS also recognized the *OsSWEET14* promoter in transient GUS reporter studies in *N. benthamiana* (Fig. 2c). We then tested whether the ArtTALS contribute to the virulence of *Xoo*, similar to the natural TAL effectors targeting *OsSWEET14*. To this end, *Xoo* BAI3ΔtalC derivatives, complemented with plasmid-borne copies of artTAL14-1 or artTAL14-2, were infiltrated into leaves of the rice line IR24. Strikingly, both strains caused typical disease symptoms (water-soaked lesions), basically identical to *Xoo* strains BAI3 and BAI3ΔtalC containing talC or tal5 (Fig. 3a). Quantitative leaf-clipping assays using these strains showed that both ArtTALS partially restored virulence of the *Xoo* BAI3ΔtalC strain, reaching levels similar to those observed upon complementation with talC. Similarly, both artTAL genes conferred virulence to the *Xoo* strain PXO99^A when assayed by leaf infiltration or leaf clipping of otherwise resistant *xa13* rice plants (Fig. 3b).

Next, we quantified bacterial populations upon leaf infection with *Xoo* strains BAI3 and BAI3ΔtalC derivatives (Fig. 3c). At

Table 1 Artificial transcription activator-like (ArtTAL) effector repeat variable diresidue (RVDs), target sequences and *in planta* activity

Gene ID ^a	Gene name ^b	ArtTAL ^c	RVD Sequence ^d	Target box ^e	Position from ATC ^f	Water soaking ^g	qRT-PCR ^h
Os01_g65880	O5SWEET1a	ArtTAL1a-1	NG-NG-HD-NG-HD-NG-NI-HD-NI-HD-NI-HD-NI-NK-NG-NK-NG	TTTCTACACACCAGTGT	-221	-	+
Os05_g35140	O5SWEET1b	ArtTAL1a-2	HD-HD-HD-NG-NG-NI-NG-NI-NG-NI-NK-NI-NK-HD-NI-HD-NI-NI	TCCCTTATAAAGCACACA	-184	-	+
Os01_g36070	O5SWEET2a	ArtTAL1b-1	NI-NG-NI-NG-NI-NG-NI-HD-HD-NK-NK-NK-NI-NK-NG-NG	TATATATAACCCGGCAGTT	-260	-	+
Os01_g50460	O5SWEET2b	ArtTAL1b-2	HD-HD-NI-HD-NI-HD-NG-HD-NG-NG-HD-NG-HD-HD-NG-HD	TCACACCCTCTCCCTC	-291	-	+
Os01_g40960	O5SWEET2c	ArtTAL2a-1	NI-NI-HD-NK-NK-HD-NI-NG-NI-NG-NI-NK-NI-HD-HD-HD	TCACGGACATATCCCCC	-215	-	+
Os05_g12320	O5SWEET3a	ArtTAL2a-2	NI-NG-NI-NG-NG-NK-NK-NK-NI-HD-NN-NK-NI-NI-HD	TATCTTGGACGGGAC	-181	-	+
Os02_g19820	O5SWEET4	ArtTAL2b-1	HD-NG-HD-NG-HD-NI-HD-NG-NK-NI-HD-NK-NI-HD-NG-NK	TCCTCACTGACCCACTCG	-110	-	+
Os05_g51090	O5SWEET5	ArtTAL2b-2	NI-NG-NI-NG-NI-NK-HD-NI-HD-NK-NI-NI-NG-NG-NG	TATAATGCCACGGAAATT	-81	-	+
Os07_g42110	O5SWEET6a	ArtTAL2c	HD-NI-NG-NG-HD-NI-HD-NK-NG-NK-NK-NI-NG-NG-NI	TCATTCCAGTGGCTTGA	-91	-	+
Os01_g71730	O5SWEET3b	ArtTAL3a-1	HD-HD-HD-NI-HD-NI-NG-NG-NI-NI-HD-NI-NK-HD-NI-NG	TCCCACATTAACAGCCAT	-117	-	+
Os02_g19820	O5SWEET7a	ArtTAL3a-2	NI-NI-NG-NI-NK-HD-NG-NK-HD-HD-HD-NI-NG-NI-NG-NG	TAATAGCTGCCCATTT	-89	-	+
Os05_g51090	O5SWEET7b	ArtTAL3b-1	NK-HD-NG-NG-HD-NI-NG-HD-NI-HD-NI-HD-NI-NG-NG-HD	TGCTTCCACCACTTCT	-143	-	+
Os07_g42110	O5SWEET7c	ArtTAL3b-2	NI-NG-NI-NG-NI-NG-NI-NK-NI-HD-NI-NK-NI-HD-NI-NI	TATAATAGACAGACCCAC	-92	-	+
Os08_g42350	O5SWEET7d	ArtTAL4-1	NI-NG-NI-HD-NG-NI-HD-NG-NI-HD-NI-HD-NG-NK-NI-NK	TATACTACTACACTTCGAG	-155	-	+
Os09_g08440	O5SWEET7e	ArtTAL4-2	HD-NK-HD-NK-NG-NG-HD-NK-NI-HD-NG-HD-NI-HD-NG-NI	TCGGTCTCGACTCCACTA	-185	-	+
Os12_g07860	O5SWEET11	ArtTAL5	NN-NK-NI-NN-NN-HD-NI-NN-NI-NN-NN-HD-NK-HD-HD-NI	TCGAGGCGAGAGCCGCCA	-283	-	+
Os12_g29220	O5SWEET12	ArtTAL6a-1	HD-NI-HD-NK-HD-NG-NI-NG-NI-NG-NI-NI-NI-HD-NI-NI-HD- HD-NI-HD	TCACGCTATAAACAAACCAC	-126	-	na
Os11_g31190	O5SWEET14	ArtTAL6a-2	NI-HD-NG-HD-NK-NI-NG-HD-NK-NI-NG-HD-NK-HD-NG-HD-HD	TACTCGATCGATCCCTCC	-163	-	na
Os12_g07860	O5SWEET15	ArtTAL6b-1	NG-NG-HD-NI-HD-NG-NN-NN-NG-HD-NG-NI-NG-NN-NG-NG	TITCACTGTTCGTTGTT	-123	-	na
Os08_g42350	O5SWEET16	ArtTAL6b-2	NG-HD-NN-HD-NN-HD-NK-NG-HD-NG-NI-NG-NI-NG-NI-NK-HD	TCCGCGCTATAAAGC	-106	-	nd
Os03_g22590	O5SWEET17	ArtTAL7a-1	NI-HD-NN-HD-NN-NG-NI-NG-NI-NG-NI-NG-HD-HD-NI-NN-NI	TACCGGTATATATCCAGA	-175	-	-
Os03_g22590	O5SWEET18	ArtTAL7a-2	NI-NG-NI-NG-NI-NG-HD-NI-NK-NI-NK-NG-HD-NK-HD	TATATATCCAGAGCTCCGC	-168	-	-
Os03_g22590	O5SWEET19	ArtTAL7b	HD-NK-NK-NI-HD-NG-HD-NK-NI-NI-NG-NK-NK-NG-NK-NK-HD	TCGGACTCGGAATGTGGC	-239	-	na
Os03_g22590	O5SWEET20	ArtTAL7c	HD-NI-HD-NK-NK-NI-HD-NG-HD-NG-HD-NK-NI-HD-NI-HD	TCACCGGACTCCGACCC	-222	-	-
Os03_g22590	O5SWEET21	ArtTAL11-1	HD-HD-NI-HD-NG-NI-NG-NI-NG-NI-NI-HD-NI-HD-NG-NN	TCCAACATATAAACACTG	-211	+/-	+
Os03_g22590	O5SWEET22	ArtTAL11-2	NN-HD-NI-NN-NG-NK-NG-NI-NK-NG-NI-NK-NI-HD-NI-NG-NG	TCGAGTTGTAGCACTT	-166	+	+
Os03_g22590	O5SWEET23	ArtTAL12-1	NI-NI-NG-NI-HD-HD-NI-NK-NK-NI-HD-NK-NI-HD-NG-NK-HD	TAATACACAGGACCATGCC	-188	-	-
Os03_g22590	O5SWEET24	ArtTAL12-2	NI-NG-NI-NG-NN-NG-NN-NN-NN-NN-NN-NN-NN-NN-NN-NG	TATAAATGTGGCTCTCC	-220	+	+
Os03_g22590	O5SWEET25	ArtTAL13-1	NI-NK-NI-NK-NK-HD-NI-HD-NI-HD-NI-NI-NK-NG-NK-NI-NG	TAGAGGCAACCAAGTAT	-269	-	-
Os03_g22590	O5SWEET26	ArtTAL13-2	HD-HD-NG-NI-NG-NI-NG-NI-NI-NK-HD-NI-HD-NI-HD-NI-NI- HD-NG	TCCTATAAAGCACCAACT	-230	+	+
Os03_g22590	O5SWEET27	ArtTAL14-1	HD-NI-HD-NI-HD-NI-HD-NG-NI-NG-NI-NN-NN-NN-NN-NN-NG	TCACACCATAAGGGCAT	-335	+	+
Os03_g22590	O5SWEET28	ArtTAL14-2	HD-NI-NG-NN-NG-NN-NG-NN-HD-HD-NG-NG-NG-HD-NI-NG-NG	TCATGTGCTCTTTCATT	-300	+	+
Os03_g22590	O5SWEET29	ArtTAL15-1	HD-NI-HD-HD-NG-HD-NG-HD-NG-HD-NG-HD-HD-NG-NI-NG-NI	TCACCTCTCTCCCTATA	-242	+	+
Os03_g22590	O5SWEET30	ArtTAL15-2	NI-NG-NI-NG-NI-NG-NI-NN-NN-NN-NN-NN-NN-NN-NN-NG	TATATAGGAGCCCTGCT	-227	+/-	+
Os03_g22590	O5SWEET31	ArtTAL16-1	NI-NG-NG-NK-NG-NG-NI-NI-NG-NK-NG-NI-HD-NG-NK-HD-NG	TATTGTTAATGCACTGCT	-147	+	+
Os03_g22590	O5SWEET32	ArtTAL16-2	NG-NN-NG-NN-HD-NG-NI-NG-NI-NG-NI-NG-NG-HD-NK-NG-HD-NK	TTGCTATATATTCGTCG	-112	-	+

^aMSU Version 7.
^bBased on the nomenclature and clade classification from Chen *et al.* (2010).
^cFor each gene two ArtTAL effectors binding to distinct target sequences were designed.
^dAmino acids in one letter code.
^eThe 5' terminal nucleotide of each target box is indicated in bold.
^fPosition of ArtTAL binding box including the initial T counted from ATG (not included) in bp.
^gLeaves of *Oryza sativa* cultivar IR24 were infiltrated using a blunt syringe and scored at 6 d postinoculation (dpi). No, partial, and water-soaked lesion are indicated as (-), (+/-) and (+), respectively.
^hO5SWEET expression analysis using quantitative reverse transcription polymerase chain reaction (qRT-PCR). Induction: +, no induction; -, no amplification product; na, not determined, nd.

6 dpi, the sizes of bacterial populations of all BAI3 Δ *talC* strains harboring natural or artificial TALs were significantly larger than those of BAI3 Δ *talC* carrying the empty vector.

Together, these data show that the artificial TAL effectors were functionally equivalent to natural ones and further demonstrate that induction of *OsSWEET14* is sufficient to confer virulence to the *Xoo* strains tested. These data also emphasize that induction of *OsSWEET14* can fully compensate the lack of *OsSWEET11* expression in the PXO99^A-IRBB13 combination. Together with previous studies (Antony *et al.*, 2010), these findings strongly suggest that these two SWEETs have a similar function in respect of *Xoo* virulence.

A series of artificial TAL effectors induce expression of individual *SWEET* genes and unmask new potential virulence targets of *Xoo*

To investigate whether additional *SWEET* genes may act as virulence targets of *Xoo*, we adopted the same strategy as used for

OsSWEET14 (Figs 2, 3) for the remaining 21 *OsSWEET* rice paralogs (Table 1). Generally, two ArtTALs (named according to the *OsSWEET* paralog) were designed for each rice *SWEET* gene, targeting specific sequences in the corresponding promoter regions, with the exception of *OsSWEET7d* and *OsSWEET7e*, for which no specific ArtTALs could be generated (Table 1). We predicted the top 20 possible target sequences for each ArtTAL and found that they target almost exclusively nonoverlapping sets of genes. Next, we evaluated the capacity of each of these ArtTALs to confer virulence to the *Xoo* strain BAI3 Δ *talC*. Only ArtTALs targeting *OsSWEET11* to *OsSWEET15* led to the production of typical water-soaked lesions (Fig. 4a). Strikingly, no disease symptoms were observed upon infiltration of plants with *Xoo* strains producing ArtTALs targeting any of the other *SWEET* paralogs (Table 1), as exemplified for *OsSWEET1a*, *OsSWEET6b* and *OsSWEET7a* in Fig. 4(a).

To exclude the possibility that lack of symptom formation was the result of a failure of *SWEET* gene induction, we checked whether or not the ArtTAL proteins are stably produced in *Xoo*

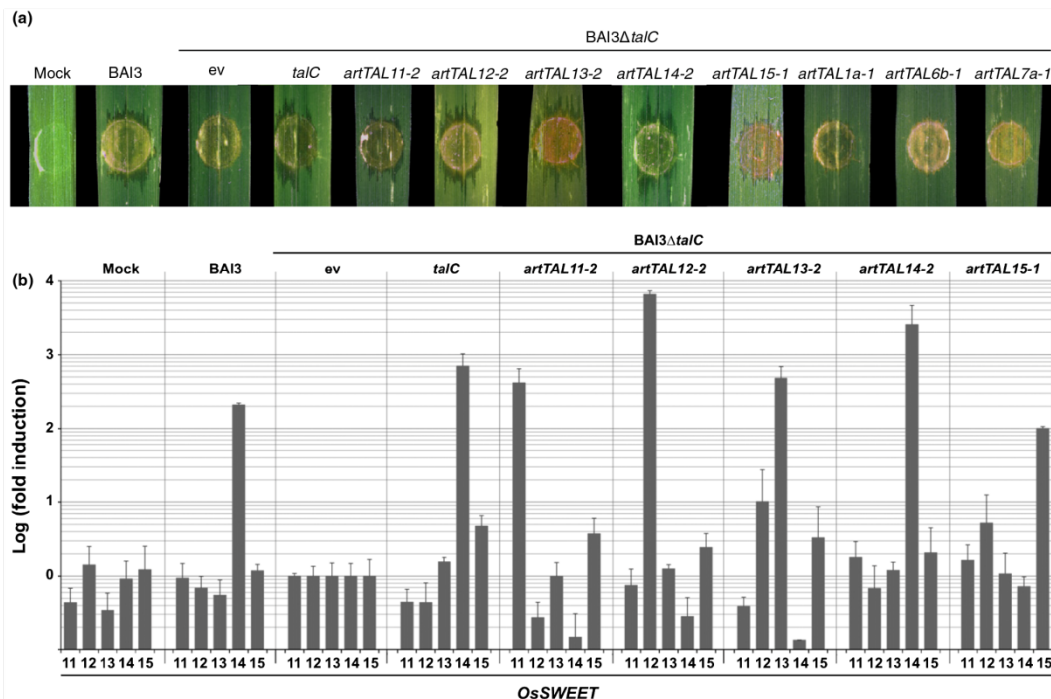


Fig. 4 Artificial transcription activator-like (TAL) effectors that specifically induce clade-III *OsSWEET* genes confer virulence to *Xanthomonas oryzae* pv. *oryzae* (*Xoo*). (a) Phenotypes of *Xoo* strains on rice leaves. *Oryza sativa* cultivars IR24 were inoculated with water (Mock), *Xoo* BAI3, or *Xoo* BAI3 Δ *talC* carrying empty vector (ev), *talC* or artificial transcription activator-like (*artTAL*) effectors targeting the five clade-III *OsSWEET11* to *OsSWEET15* genes, as well as *OsSWEET1a*, *OsSWEET6b* and *OsSWEET7a*. (b) Transcript abundances of clade-III *OsSWEET* genes *SWEET11* to *SWEET15* were determined 1 d postinoculation by quantitative reverse transcription polymerase chain reaction (qRT-PCR). The bars represent a minimum of two biological samples taken from three leaves. The fold change was calculated in comparison to BAI3 Δ *talC* (ev). Error bars represent + SD. Actin was used as a reference gene to normalize cDNA amounts.

and the target genes in rice are induced. Indeed, Western blot analysis indicated that all ArtTALs accumulated at significant levels and showed protein integrity in *Xoo* (Fig. S1). Quantitative RT-PCR experiments showed that most *SWEET* genes were induced in an ArtTAL-dependent manner, except for *OsSWEET2c*, *OsSWEET5* and *OsSWEET7c* (Fig. S2). *OsSWEET6a* and *OsSWEET7b* were not analyzed, because no specific amplification products could be detected, although several pairs of primers have been tested (Table 1).

Interestingly, all five rice *SWEET* paralogs whose activation was associated with the development of typical disease symptoms belong to clade III of the *SWEET* gene family, which comprises *OsSWEET11* (*Os8N3/Xa13*), *OsSWEET12*, *OsSWEET13* (*Xa25*), *OsSWEET14*, and *OsSWEET15*. *OsSWEET11* and *OsSWEET14* are known major *Xoo* susceptibility genes of rice (Yang *et al.*, 2006; Antony *et al.*, 2010), which confirms the suitability of our strategy. Furthermore, we evaluated the basal and induced *OsSWEET* gene expression levels (Table S4), which suggest that clade III *SWEETs* are generally not induced at higher levels than nonclade III ones. However, we cannot rule out the possibility that a few of them act as susceptibility genes when induced at higher levels and/or under different conditions.

To demonstrate that the observed gain of virulence phenotypes can be solely attributed to the induction of the targeted *SWEET* genes, we evaluated the induction profile of each of the five clade-III paralogs by qRT-PCR during infection of plants with *Xoo* strains delivering different ArtTALs. Among the five candidate susceptibility genes we identified, only *OsSWEET14* was activated upon infiltration of rice leaves with *Xoo* wildtype strain BAI3. None of them was induced upon infiltration with water or *Xoo* strain BAI3 Δ *talC* (Fig. 4b). In addition, only the expected target *SWEET* gene was induced upon infection of *Xoo* strains delivering the cognate ArtTAL, indicating that the observed phenotypes are not the result of induction of collateral susceptibility genes.

Finally, we quantified bacterial populations *in planta* in order to further demonstrate the function of the clade-III *SWEET* genes as virulence targets of *Xoo*. As shown in Fig. 5, BAI3 Δ *talC* derivatives carrying ArtTALs inducing any of the five clade-III *SWEET* genes grew to significantly larger population sizes than BAI3 Δ *talC* or BAI3 Δ *talC* carrying ArtTALs that target the other *SWEET* clade members.

Discussion

Modern agriculture relies on effective control of pathogens. Better knowledge about host molecular determinants underlying plant diseases is thus essential to develop new strategies leading to plant resistance. However, only a few host genes have been characterized until now that are critical to promote bacterial diseases. These include members of the *SWEET/modulin-3* family, which were recently identified as essential alternative *Xoo* rice virulence targets, also called susceptibility genes (White & Yang, 2009). Current models suggest that bacteria hijack individual *SWEETs* that function as sugar transporters, leading to the accumulation of nutrients in the apoplast where *Xoo* multiplies. Previous studies

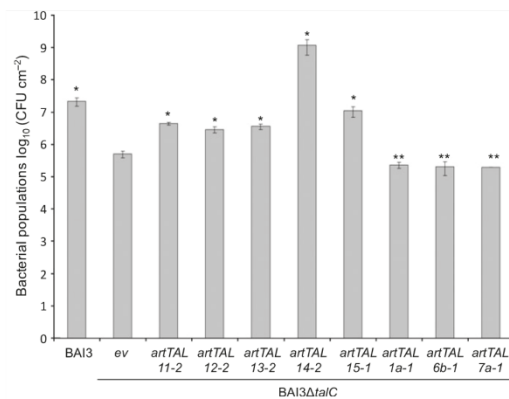


Fig. 5 Artificial transcription activator-like (TAL) effectors that induce clade-III *OsSWEET* genes restore bacterial growth in rice leaves. *Oryza sativa* IR24 leaves were infiltrated with *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) BAI3, *Xoo* BAI3 Δ *talC* carrying empty vector (ev), or artTAL11-2, artTAL12-2, artTAL13-2, artTAL14-2, artTAL15-1, artTAL1a-1, artTAL6b-1, and artTAL7a-1, which target *OsSWEET11*, *OsSWEET12*, *OsSWEET13*, *OsSWEET14*, *OsSWEET15*, *OsSWEET1a*, *OsSWEET6b* and *OsSWEET7a*, respectively. Bacterial populations were scored 6 d postinfiltration. Data are the arithmetic means from three plants and three independent experiments. Error bars represent \pm SD. *Significant difference in Student's *t*-test at $P < 0.01$ when comparing samples with BAI3 Δ *talC* (ev); **, no differences were detected at $P < 0.01$.

identified *OsSWEET11* and *OsSWEET14* as susceptibility genes, as their induction by cognate native TAL effectors provides virulence to strains carrying them (Talbot, 2010).

Here, we further corroborate the key role of *OsSWEET14* in bacterial virulence as a direct target of natural *Xanthomonas* TAL effectors. Tal5, a novel major virulence TAL effector from a Malian strain of *Xoo*, recognizes a distinct sequence within the *OsSWEET14* promoter sequence, which differs from those that are recognized by strains from the Philippines and Burkina Faso. This functional convergence of virulence factors illustrates the evolutionary pressure that is imposed on *Xoo* strains of different geographic origins and genetic lineages to evolve TAL genes that induce this particular *SWEET* gene.

In a recent report, Li *et al.* (2013) identified *OsSWEET12* as a third potential susceptibility gene. Upon infection of rice with *Xoo* mutant strains defective for *OsSWEET11* activation, the authors showed that complementation with two independent ArtTALs activating *OsSWEET12* leads to a gain of virulence. We significantly extend this study by providing additional important experiments. We analyzed bioinformatically that the ArtTALs target no other common genes besides *OsSWEET14* (Table S2). We verified the ArtTALs' stability (Fig. S1) and assessed their binding specificity to their target promoter using GUS reporter assays (Fig. 2c). Importantly, we demonstrate that ArtTAL-mediated *OsSWEET* activation is not the result of collateral expression of other clade III paralogs (Fig. 4). Finally, we extended this strategy to the full *SWEET* gene family, systematically addressing the function of 20 of the 22 *OsSWEET* paralogs.

These studies highlight the functionality of ArtTALs as useful tools to investigate the function of host genes during the plant–pathogen interaction. In total, we generated 36 ArtTALs, 26 of which (72%) were functional, and induced their target plant genes upon delivery by *Xanthomonas*. Few ArtTALs were not functional in our study, which might be because of nonoptimal choice of the target region or a noninducible status of the target gene. Overall, our findings emphasize that ArtTALs are suitable tools to induce host genes of choice and that bacterial delivery is a highly efficient means to do so. In summary, we confirm that *OsSWEET11*, *OsSWEET12*, and *OsSWEET14* are essential bona fide host virulence targets of *Xoo*; we formally validate the function of *OsSWEET13/Xa25* as a fourth susceptibility gene, as up to this point only indirect evidence has been available (Liu *et al.*, 2011); we identify *OsSWEET15* as a novel potential susceptibility gene which can support the development of disease symptoms triggered by *Xoo*; and, in total, we analyzed 20 *OsSWEET* paralogs, only five of which could be demonstrated to function as susceptibility genes.

Our data suggest that, in addition to the three already identified *SWEET* genes (*OsSWEET11*, *OsSWEET12* and *OsSWEET14*), two more (*OsSWEET13* and *OsSWEET15*) may be exploited by *Xoo*, presumably to satisfy the pathogen's nutritional needs. *OsSWEET13* is activated by *Xoo* strain PXO339, but the corresponding TAL effector(s) remain(s) to be identified

(Liu *et al.*, 2011). By contrast, no native TAL effectors have been reported so far for *OsSWEET12* and *OsSWEET15*. Interestingly, screening GenBank for TAL effectors predicted to recognize DNA boxes within their promoter indicated that the TAL protein AAW76267 of *Xoo* strain KACC10331 (Lee *et al.*, 2005) and Tal7b/Tal8b of *Xoo* strain PXO99^A (Salzberg *et al.*, 2008) may target *OsSWEET12* and *OsSWEET15*, respectively (Grau *et al.*, 2013). Although transcriptomic data confirmed an increase in *OsSWEET15* expression in plants challenged with PXO99^A as compared with water (Grau *et al.*, 2013), *OsSWEET15* induction might not be sufficient for growth on IRBB13 (Fig. 3b), indicating that a threshold expression level might be required to support infection. Indeed, the target box for Tal7b/Tal8b appears to be suboptimal, as a G is found at the 5' end of the predicted DNA box, which is known to be less efficient for gene activation (Boch *et al.*, 2009).

Our study highlighted the fact that five *OsSWEET* genes may be manipulated by *Xoo* to favor disease. Interestingly, these five *SWEET* genes, *OsSWEET11* to *OsSWEET15*, all belong to phylogenetic clade III (Fig. 6). Several *A. thaliana* *SWEET*s, as well as rice *OsSWEET11* and *OsSWEET14*, are both low-affinity glucose and sucrose transporters (Chen *et al.*, 2010, 2012). *OsSWEET11* and *OsSWEET14* are bidirectional sugar transporters and differ in their transport direction from the SUT/SUC high-affinity H⁺ sugar symporters of plants (Kühn & Grof, 2010;

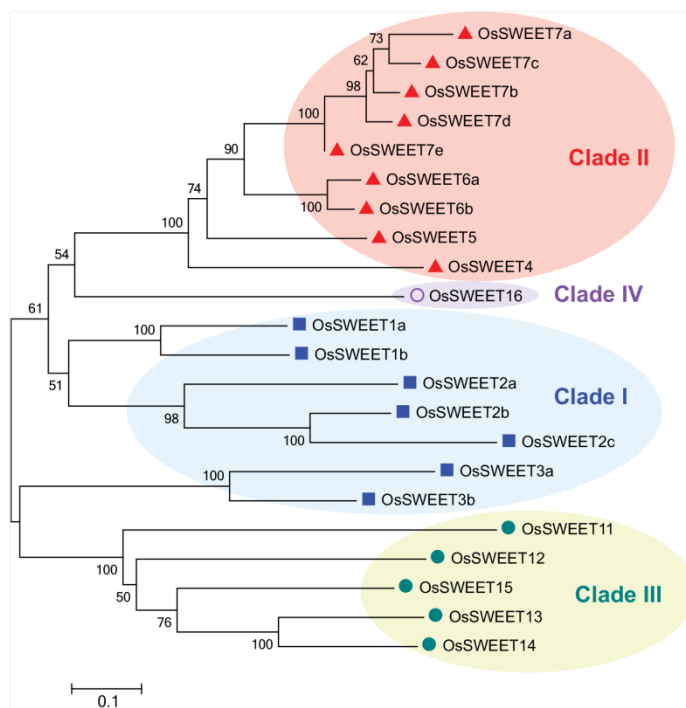


Fig. 6 Phylogenetic tree of *OsSWEET* proteins. Amino acid sequences were obtained from UniProt (<http://www.uniprot.org/>; Bairoch *et al.*, 2005) and aligned using MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>; Edgar, 2004) with default parameters. The sequence of *OsSWEET7a* was corrected, as UniProt reports an erroneous tandem duplication of 54 amino acids at the N terminus. Alignments were analyzed by the neighbor-joining method using MEGA5.10 (<http://www.megasoftware.net>; Tamura *et al.*, 2011), assuming a Poisson substitution model, uniform rates among sites, and pairwise deletions for missing data and gaps. The numbers at the nodes represent bootstrap percentage values based on 1000 replications. Proteins of clade I, blue squares; clade II, red triangles; clade III, green dots; clade IV, purple circle.

Chen *et al.*, 2012). Therefore, increased expression of *SWEET*s alone might lead to a net efflux of sugars from the plant cells. Under physiological conditions, *SWEET* sugar transporters facilitate flux of sucrose from the phloem parenchyma cells into the apoplast, whereas SUT/SUC transporters load sucrose from the apoplast into the phloem companion cells and sieve elements (Kühn & Grof, 2010; Braun, 2012). Our data imply that all clade-III *OsSWEET*s share a function that can support multiplication of *Xoo* and that this function is absent from the other *SWEET* paralogs in rice. It is tempting to speculate that this common function is the export of a particular sugar, possibly sucrose.

At present, it is unclear whether *Xoo* relies on *SWEET*s only at certain infection stages. In a study involving *Xoo* strain BAI3, only vascular spreading, but not local growth of *Xoo*, is prevented in the absence of *SWEET* induction (Yu *et al.*, 2011). Curiously, the locally restricted rice pathogen *X. oryzae* pv. *oryzicola* is also able to benefit from TAL effectors that mediate *SWEET* gene induction, although no native *SWEET*-targeting TAL effectors have been reported in *Xoc* (Verdier *et al.*, 2012). Other pathogens, such as *P. syringae* pv. *tomato* DC3000 and the fungal pathogens *Golovinomyces cichoracearum* and *Botrytis cinerea*, also induce the expression of *SWEET* genes in *Arabidopsis*, including those that facilitate sugar transport (Chen *et al.*, 2010, 2012). This suggests that different pathogens exploit plant sugar transporters by different means to gain access to carbohydrate nutrients.

The hallmark of our approach relies on its great potential to discover as yet unknown plant susceptibility genes. ArtTALs perfectly mimic the action of natural TAL effectors and will therefore enable the major plant genes controlling bacterial proliferation in any TAL-mediated plant disease to be pinpointed, such as Citrus canker and cassava bacterial blight (Duan *et al.*, 1999; Rao *et al.*, 2002). In addition, searches for unresponsive alleles will offer new resources to achieve sustainable plant resistance by marker-assisted breeding, thus circumventing transgenic approaches. Strategies aimed at producing rice lines with improved and broad resistance against BLB already incorporate the concept of gain of 'resistance by defective susceptibility', as successfully done by pyramiding the unresponsive allele of *OsSWEET11*, *xa13* (Rao *et al.*, 2002).

In addition to *OsSWEET11*, hijacking of *OsSWEET14* by *Xoo* appears to be a widespread strategy to provoke disease. As reported by Li *et al.* (2012), 32 out of 40 *Xoo* strains collected worldwide were impaired for virulence when inoculated on TALEN-edited *OsSWEET14* rice lines mutated in the *AvrXa7*/*PthXo3* box. Our results on *Tal5* and *ArtTAL14s* increase the number of experimentally proven native and artificial TAL effectors activating *OsSWEET14* to six, which recognize five different target boxes (Fig. 1). This finding illustrates the impressive potential of TAL effectors to induce genes by binding to differently located target sequences. Therefore, it will be essential to take the natural diversity of TAL effector repertoires in *Xoo* populations into account before exploiting or generating unresponsive susceptibility target gene variants.

Acknowledgements

We thank C. Kretschmer for technical assistance, and U. Bonas for continued support. We are grateful to Y. Yu, M. Soto-Suárez and V. Verdier for providing the MAII DNA cosmid library. This project was supported by an EMBO junior grant to J.S. (ASTF no: 355-2011), a grant from the Deutsche Forschungsgemeinschaft (SPP1212) to J.B., a grant from the Agence Nationale de la Recherche (ANR-2010-GENM-013) to R.K., and a grant from the Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture (093604) to C.P. We are grateful to Alvaro Pérez-Quintero and Sébastien Cunnac for bioinformatic prediction of TAL targets and advice on statistical analyses, respectively.

References

- Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. 2010. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* 22: 3864–3876.
- Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M *et al.* 2005. The Universal Protein Resource (UniProt). *Nucleic Acids Research* 33: D154–D159.
- Baker RF, Leach KA, Braun DM. 2012. *SWEET* as sugar: new sucrose effluxers in plants. *Molecular Plant* 5: 766–768.
- Boch J, Bonas U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* 48: 419–436.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326: 1509–1512.
- Bogdanove AJ, Voytas DF. 2011. TAL effectors: customizable proteins for DNA targeting. *Science* 333: 1843–1846.
- Braun DM. 2012. *SWEET!* The pathway is complete. *Science* 335: 173–174.
- Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, Lahaye T, Leonhardt H. 2012. Targeted transcriptional activation of silent *oct4* pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Research* 40: 5368–5377.
- Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B *et al.* 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468: 527–532.
- Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB. 2012. Sucrose efflux mediated by *SWEET* proteins as a key step for phloem transport. *Science* 335: 207–211.
- Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. 2012. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS ONE* 7: e45383.
- Chu Z, Fu B, Yang H, Xu C, Li Z, Sanchez A, Park YJ, Bennetzen JL, Zhang Q, Wang S. 2006a. Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. *Theoretical and Applied Genetics* 112: 455–461.
- Chu Z, Yuan M, Yao J, Ge X, Yuan B, Xu C, Li X, Fu B, Li Z, Bennetzen JL *et al.* 2006b. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes and Development* 20: 1250–1255.
- Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F. 2012. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature Communications* 3: 968.
- Duan YP, Castañeda A, Zhao G, Erdos G, Gabriel DW. 1999. Expression of a single, host-specific, bacterial pathogenicity gene in plant cells elicits division, enlargement, and cell death. *Molecular Plant-Microbe Interactions* 12: 556–560.

- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- Figurski D, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 is dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences, USA* 76: 1648–1652.
- Gamas P, Niebel Fde C, Lescure N, Cullimore J. 1996. Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Molecular Plant–Microbe Interactions* 9: 233–242.
- Geißler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, Boch J. 2011. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* 6: e19509.
- Gonzalez C, Szurek B, Manceau C, Mathieu T, Sere Y, Verdier V. 2007. Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. *Molecular Plant–Microbe Interactions* 20: 534–546.
- Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J. 2013. Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Computational Biology* 9: e1002962.
- Guan YF, Huang XY, Zhu J, Gao JF, Zhang HX, Yang ZN. 2008. RUPTURED POLLEN GRAIN1, a member of the MtN3/saliva gene family, is crucial for exine pattern formation and cell integrity of microspores in Arabidopsis. *Plant Physiology* 147: 852–863.
- Hopkins CM, White FF, Choi SH, Guo A, Leach JE. 1992. Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant–Microbe Interactions* 5: 451–459.
- Kauffman HE, Reddy AS, Hsieh SP, Merca SD. 1973. An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. *Plant Disease Management Reports* 57: 537–541.
- Kay S, Hahn S, Marois E, Wieduwild R, Bonas U. 2009. Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3Δrep16. *Plant Journal* 59: 859–871.
- Kühn C, Grof CP. 2010. Sucrose transporters of higher plants. *Current Opinion in Plant Biology* 13: 288–298.
- Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS *et al.* 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Research* 33: 577–586.
- Li T, Huang S, Zhou J, Yang B. 2013. Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. *Molecular Plant* 6: 781–789.
- Li T, Liu B, Spalding MH, Weeks DP, Yang B. 2012. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature Biotechnology* 30: 390–392.
- Liu Q, Yuan M, Zhou Y, Li X, Xiao J, Wang S. 2011. A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant, Cell & Environment* 34: 1958–1969.
- Maeder ML, Linder SJ, Reynon D, Angstman JF, Fu Y, Sander JD, Joung JK. 2013. Robust, synergistic regulation of human gene expression using TALE activators. *Nature Methods* 10: 243–245.
- Marois E, Van den Ackerveken G, Bonas U. 2002. The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Molecular Plant–Microbe Interactions* 15: 637–646.
- Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ *et al.* 2011. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* 29: 143–148.
- Morbiter R, Römer P, Boch J, Lahaye T. 2010. Regulation of selected genome loci using *de novo*-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences, USA* 107: 21617–21622.
- Moscou MJ, Bogdanov AJ. 2009. A simple cipher governs DNA recognition by TAL effectors. *Science* 326: 1501.
- Mussolino C, Cathomen T. 2012. TALE nucleases: tailored genome engineering made easy. *Current Opinion in Biotechnology* 23: 644–650.
- Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T. 2007. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *Journal of Bioscience and Bioengineering* 104: 34–41.
- Nino-Liu DO, Ronald PC, Bogdanov AJ. 2006. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Molecular Plant Pathology* 7: 303–324.
- Perez-Pinera P, Ousterout DG, Brunger JM, Farin AM, Glass KA, Guilak F, Crawford GE, Hartemink AJ, Gersbach CA. 2013. Synergistic and tunable gene activation in human cells by combinations of synthetic transcription factors. *Nature Methods* 10: 239–242.
- Pérez-Quintero A, Rodríguez RL, Dereeper A, Lopez C, Koebnik R, Szurek B, Cunnac S. 2013. An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLoS ONE* 8: e68464.
- Quirino BF, Normanly J, Amasino RM. 1999. Diverse range of gene activity during *Arabidopsis thaliana* leaf senescence includes pathogen-independent induction of defense-related genes. *Plant Molecular Biology* 40: 267–278.
- Rao KK, Lakshminarasu M, Jena KK. 2002. DNA markers and marker-assisted breeding for durable resistance to bacterial blight disease in rice. *Biotechnology Advances* 20: 33–47.
- Reimers PJ, Leach JE. 1991. Race-specific resistance to *Xanthomonas oryzae* pv. *oryzae* conferred by bacterial blight resistance gene *Xa-10* in rice (*Oryza sativa*) involves accumulation of a lignin-like substance in host tissues. *Physiological and Molecular Plant Pathology* 38: 39–55.
- Römer P, Recht S, Strauß T, Elsaesser J, Schornack S, Boch J, Wang S, Lahaye T. 2010. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist* 187: 1048–1057.
- Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, Furutani A, Ochiai H, Delcher AL, Kelley D *et al.* 2008. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99^A. *BMC Genomics* 9: 204.
- Streubel J, Blücher C, Landgraf A, Boch J. 2012. TAL effector RVD specificities and efficiencies. *Nature Biotechnology* 30: 593–595.
- Talbot NJ. 2010. Cell biology: raiding the sweet shop. *Nature* 468: 510–511.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Van Larebeke N, Engler G, Holsters M, Van den Elsacker S, Zaenen I, Schilperoot RA, Schell J. 1974. Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* 252: 169–170.
- Verdier V, Triplett LR, Hummel AW, Corral R, Cernadas RA, Schmidt CL, Bogdanov AJ, Leach JE. 2012. Transcription activator-like (TAL) effectors targeting *OxSWEET* genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. *New Phytologist* 196: 1197–1207.
- White FF, Yang B. 2009. Host and pathogen factors controlling the rice–*Xanthomonas oryzae* interaction. *Plant Physiology* 150: 1677–1686.
- Yang B, Sugio A, White FF. 2006. *OxN3* is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences, USA* 103: 10503–10508.
- Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B. 2011. Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice Nodulin-3 *Ox11N3* Gene. *Molecular Plant–Microbe Interactions* 24: 1102–1113.
- Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature Biotechnology* 29: 149–153.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 ArtTAL protein expression analysis in *Xanthomonas*.

Fig. S2 Quantification of *OsSWEET* transcript abundance.

Table S1 Sequences of oligonucleotides used for qRT-PCR

Table S2 Talvez ranking of Tal5 target candidates upon whole promoterome search

Table S3 Target candidates of artificial TAL effectors ArtTAL14-1 and ArtTAL 14-2 upon whole promoterome search

Table S4 Basal and induced *OsSWEET* gene expression levels

Methods S1 Generation of *talC* and *tal5* expression constructs.

Methods S2 Choice of target sites and generation of artificial TAL effectors.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <25 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit www.newphytologist.com to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@ornl.gov)
- For submission instructions, subscription and all the latest information visit www.newphytologist.com

2.2.1.1 Anlagen zu Publikation 2

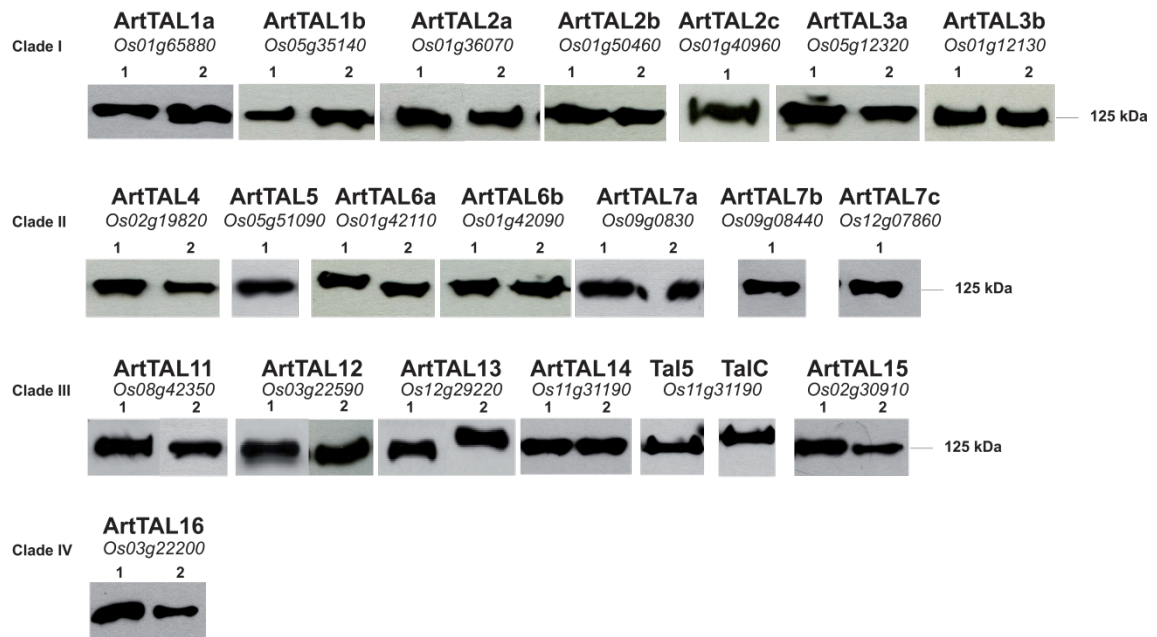


Fig. S1. ArtTAL protein expression analysis in *Xanthomonas oryzae* pv. *oryzae*.

For each OsSWEET gene (except for OsSWEET7d and OsSWEET7e, see text), one or two artificial TAL effectors were generated. Artificial TAL effector constructs as well as *tal5* and *talC* were inserted into the broad host vector pSKX1 allowing a *lac* promoter-driven constitutive expression and the fusion to a C-terminal FLAG epitope. *Xanthomonas oryzae* pv. *oryzae* strains BAI3Δ*talC* carrying the corresponding TAL effector construct were grown in PSA medium. Cells of 1 ml of a bacterial suspension at an OD600 of 0,2 were harvested and analysed by SDS-PAGE followed by immunoblotting. The TAL effectors are ordered according to the different OsSWEET gene clades. The expected size for TAL effectors with 17.5 repeats is 125 kDa and for TAL effectors with 21.5 or 22.5 repeats is 135 or 140 kDa, respectively.

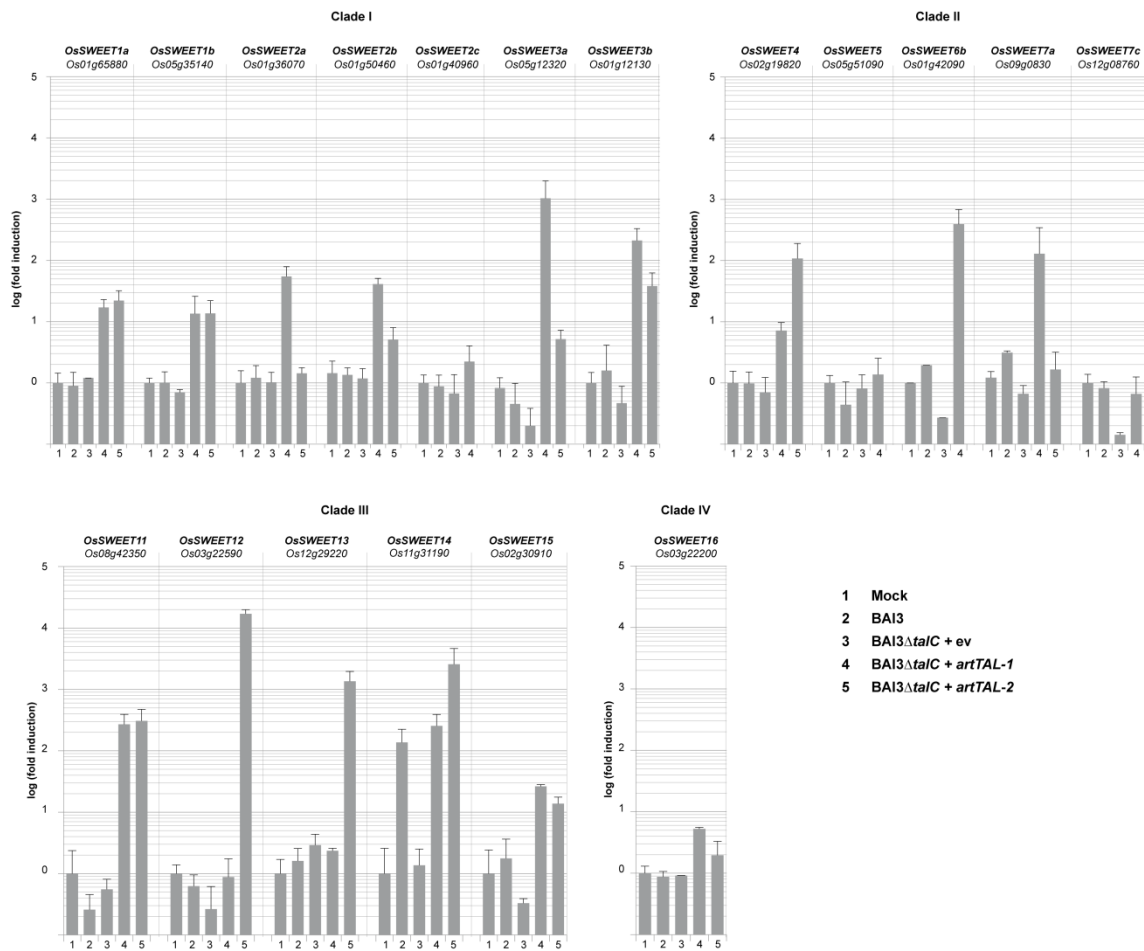


Fig. S2. Quantification of *OsSWEET* transcript levels.

Three-weeks old *Oryza sativa* plants from the cultivar Nipponbare or IR24 (*OsSWEET14*) were inoculated with water (Mock, 1) or with *Xanthomonas oryzae* pv. *oryzae* strains BAI3 (2), BAI3 Δ talC carrying an empty vector (3), or BAI3 Δ talC carrying the corresponding artificial TAL effector (4: *artTAL-1*; 5: *artTAL-2*). Please notice that for genes *Os01g40960*, *Os05g51090*, *Os01g42090*, and *Os12g07860*, only one artificial TAL effector was analysed. Genes *Os01g42110* (*OsSWEET6a*) and *Os09g08440* (*OsSWEET7b*) were not analyzed, because no specific amplification product was detected. The expression level of the gene *Os01g42090* in the sample treated with ArtTAL6b-2 was not determined. Transcript levels of the respective *OsSWEET* genes were determined one day post inoculation by quantitative RT-PCR. The bars represent the mean of two to three biological samples. The fold change was calculated in comparison to water and plotted on a logarithmic scale. Error bars represent the standard deviation. Actin was used as a reference gene to normalize cDNA amounts. The *OsSWEET* genes are ordered in the different clades according to the phylogenetic tree shown in Fig.6.

Supporting Information Tables S1–S4

Table S1. Sequences of Oligonucleotides used for qRT-PCR.

Gene ID	Gene name	Forward primer 5'→3	Reverse primer 5'→3	E ¹	E in %
Os01g65880	OsSWEET1a	TCTCCATCTGCATGTACGCC	TGCATTCTCCGTGCGTGTC	1.817	90.83%
Os05g35140	OsSWEET1b	ACTGCCTCCTCTCTGCATG	GTAGATGAACCACGACGTGC	1.590	79.48%
Os01g36070	OsSWEET2a	TGGCGCTGGGATTGCAG	GACGCAACGCTGAGGTAAC	1.931	96.57%
Os01g50460	OsSWEET2b	ATGATATCTCCTGCTTTGCTG	CCGAATCACAACTCCCATGAC	1.894	94.69%
Os01g40960	OsSWEET2c	GGCAACAAGAGCGAGCAGC	TAGAAAGATGAAGAGACAAATGTAGG	2.026	101.28%
Os05g12320	OsSWEET3a	CCTTCCTGATATACGTTTTATTGTG	TTACGTATATGATGGTTGTGTATTG	1.907	95.37%
Os01g12130	OsSWEET3b	CGTGTAGCAGTAGGAATTCTAG	TGCGTAAACCGTGCATG	1.895	94.77%
Os02g19820	OsSWEET4	GGGGTCGGTGGAGCAGTAC	CCGTTGGGGATCGTGATG	1.840	92.12%
Os05g50190	OsSWEET5	CGTCCGCAATGTCGTTGG	GCTCTTTGTCGCGATCAC	1.756	87.82%
Os01g42090	OsSWEET6b	CCAACAAGAAGCGCCTGAG	ACGGAGACGACACTGCC	1.767	88.33%
Os09g08030	OsSWEET7a	TGGCGTACTCTTAGGCGTG	CTGGATAGTGATATCGATGGTG	1.952	97.61%
Os12g07860	OsSWEET7c	TAGCGGTCGGTGTGCTTTTG	CTGACAGGGGCGACGATG	1.893	94.67%
Os08g42350	OsSWEET11	AGTCGACGGGAGGGTACAG	TTCGGGTACATGACGTAGGG	1.897	94.85%
Os03g22590	OsSWEET12	CATTGGTCTTTGCTGTGGGG	CGGTCGGTGGCCTTGAC	1.957	97.86%
Os12g29220	OsSWEET13	GGCCTGTCCTGCAGCATC	CCCGAACACCCCCACGTTC	1.909	95.47%
Os11g31190	OsSWEET14	CTACCTGGCCCCACTGC	GTGCGCACCACCAGCC	1.963	98.15%
Os02g30910	OsSWEET15	TGTCCATGGAACGCAGCAC	CAGCCGAGGACGTGGAC	1.878	93.91%
Os03g22200	OsSWEET16	GATGGGGCATCAGGTGGAG	GGTCCATCAGCTTCATCAAAATTC	1.781	89.04%
	<i>Actin</i>	TGTGTGTGACAATGGAACCTGGC	GAGTCCAACACGATACCAGTTG	1.884	94.22%

¹Efficiency was calculated with $E=10^{-1/m}$; m was calculated by using a standard curve plot of a dilution series.

Table S2. Talvez ranking of Tal5 target candidates upon whole promoterome search.

Prediction rank ¹	Gene ID ²	Talvez prediction score	Position from ATG ³	Predicted target box ⁴	Annotation
1	<i>Os11g31190</i>	15.043	-235	TAAGCTCATCAAGCCTTCA	nodulin MtN3 family protein, putative, expressed
2	<i>Os02g29150</i>	13.856	-69	TCCACTCCCCAAACCCTCG	OsFBO11 - F-box and other domain containing protein, expressed
3	<i>Os06g48680</i>	13.634	-97	TCCACTCCTCGCGCCTTCA	expressed protein
4	<i>Os01g11720</i>	13.551	-470	TGCACTCATAAAACCCTAA	putative protein containing Pfam profile: PF03004, Transposase_24
5	<i>Os07g46560</i>	13.517	-227	TGAGCTCATAAAGCCTTCA	seven in absentia protein family domain containing protein, expressed
6	<i>Os01g21542</i>	13.467	-81	CGCGCTCCTTAAGCCCTTA	retrotransposon protein, putative, unclassified, expressed
7	<i>Os03g05250</i>	13.373	-459	TACCATCCTCAAATCCTCA	expressed protein
8	<i>Os08g11450</i>	13.263	-219	TCCACTCCTCCACCACTCA	expressed protein
9	<i>Os03g54750</i>	13.244	-215	CACACTCCTCACACCACCA	COBRA-like 3 protein precursor, putative, expressed
10	<i>Os01g04340</i>	13.223	-219	CCCCTCCTCGACCCTTCA	hsp20/alpha crystallin family protein, putative, expressed

¹Talvez prediction for the transcription activator-like (TAL)-DNA box pair in a whole promoterome screening

²MSU version 7

³Position of TAL binding box including the 5' terminal nucleotide counted from ATG (not included)

⁴The 5' terminal nucleotide of each target box is indicated in bold

Table S3. Target candidates of artificial TAL effectors ArtTAL14-1 and ArtTAL 14-2 upon whole promoterome search.

ArtTAL	Prediction rank ¹	Gene ID ²	Talvez prediction score	Position from ATG ³	Predicted target box ⁴	Annotation
ArtTAL14-1	1	<i>Os11g31190</i>	19.309	-335	TCACACACCATAAGGGCA	nodulin MtN3 family protein, putative, expressed
ArtTAL14-1	2	<i>Os04g17890</i>	15.363	-459	TCAAAAACCATAAAAAACAT	retrotransposon protein, putative, unclassified, expressed
ArtTAL14-1	3	<i>Os04g17850</i>	15.363	-459	TCAAAAACCATAAAAAACAT	retrotransposon protein, putative, unclassified, expressed
ArtTAL14-1	4	<i>Os08g07940</i>	15.210	-387	TCATACACCATACGAGCAT	disease resistance protein RPM1, putative, expressed
ArtTAL14-1	5	<i>Os11g09130</i>	14.944	-240	TCACACGCCATAAAAAAAT	receptor-like protein kinase 5 precursor, putative, expressed
ArtTAL14-1	6	<i>Os04g03210</i>	14.097	-444	CTACACACCATAAGAAAAT	receptor kinase, putative, expressed
ArtTAL14-1	7	<i>Os07g07860</i>	14.028	-106	TCACACACCATCAAGCTAT	LTP1.76 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
ArtTAL14-1	8	<i>Os05g03934</i>	14.008	-127	CCACACACCACACACACAT	expressed protein
ArtTAL14-1	9	<i>Os04g17810</i>	13.858	-212	TCAAAAACCACAAAAACAT	retrotransposon protein, putative, unclassified
ArtTAL14-1	10	<i>Os03g02514</i>	13.412	-346	TCAAAACCCACAAAGGCCAC	hydrolase, alpha/beta fold family protein, putative, expressed
ArtTAL14-1	11	<i>Os05g35274</i>	13.306	-406	CCAAACCCACAAAAACAT	RNA recognition motif containing protein, putative, expressed
ArtTAL14-1	12	<i>Os02g17000</i>	13.285	-36	CACAAAACCATAGAACAT	OsSub14 - Putative Subtilisin homologue, expressed
ArtTAL14-1	13	<i>Os04g38290</i>	13.225	-234	TCACACACAACACGCACAT	expressed protein
ArtTAL14-1	14	<i>Os03g10310</i>	13.216	-357	TCACAAAACCTACGCACAT	expressed protein
ArtTAL14-1	15	<i>Os12g30000</i>	13.135	-476	TCATACAACATAAGAAAAT	expressed protein
ArtTAL14-1	16	<i>Os06g12760</i>	13.107	-124	CCACACACCATACACATAT	hypothetical protein
ArtTAL14-1	17	<i>Os04g47960</i>	13.038	-315	TCCCCAACCATAAAAACAT	expressed protein
ArtTAL14-1	18	<i>Os02g03940</i>	12.887	-160	CCGCACACCATAAAAAACC	expressed protein
ArtTAL14-1	19	<i>Os06g08740</i>	12.711	-443	TCACACACCATCAACGCAG	expressed protein
ArtTAL14-1	20	<i>Os03g61590</i>	12.659	-456	TCACACACCACAAAAGAAA	expressed protein
ArtTAL14-2	1	<i>Os11g31190</i>	18.951	-300	TCATGTGTGCTTTTCATT	nodulin MtN3 family protein, putative, expressed

Ergebnisse

ArtTAL14-2	2	<i>Os05g36220</i>	14.795	-154	T CACATGAACCTTTTCATT	transposon protein, putative, CACTA, En/Spm sub-class, expressed
ArtTAL14-2	3	<i>Os11g41780</i>	14.795	-176	T CACATGAACCTTTTCATT	transposon protein, putative, CACTA, En/Spm sub-class, expressed
ArtTAL14-2	4	<i>Os07g05850</i>	14.795	-154	T CACATGAACCTTTTCATT	transposon protein, putative, CACTA, En/Spm sub-class, expressed
ArtTAL14-2	5	<i>Os10g38314</i>	14.704	-156	T CATATCTCACTTTTCATT	glutathione S-transferase, N-terminal domain containing protein, expressed
ArtTAL14-2	6	<i>Os12g02830</i>	14.331	-189	C CATCTCTGCCTTTTCCTT	SNF7 domain containing protein, putative, expressed
ArtTAL14-2	7	<i>Os05g44530</i>	13.739	-217	C CATATATGATTTTCATT	BTB3 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain, expressed
ArtTAL14-2	8	<i>Os02g09820</i>	13.305	-396	C CCCATGTGCCTTTTAATT	zinc finger, C3HC4 type domain containing protein, expressed
ArtTAL14-2	9	<i>Os03g22050</i>	13.268	-383	T AATGTGTACCTTACCATT	CAMK_KIN1/SNF1/Nim1_like.16 - CAMK includes calcium/calmodulin dependent protein kinases, expressed
ArtTAL14-2	10	<i>Os05g12140</i>	13.081	-211	T CCATAAATGCATTTTCATT	Leucine Rich Repeat family protein, expressed
ArtTAL14-2	11	<i>Os01g29850</i>	13.075	-433	T CATATGCCCTTTTCATT	expressed protein
ArtTAL14-2	12	<i>Os05g25160</i>	13.075	-194	T CATATGCCCTTTTCATT	transposon protein, putative, CACTA, En/Spm sub-class, expressed
ArtTAL14-2	13	<i>Os11g10880</i>	13.067	-86	T CATACCTGCCTTTTCCTT	expressed protein
ArtTAL14-2	14	<i>Os09g32190</i>	12.956	-482	C CATACACACCTTTTAATT	tRNA-splicing endonuclease positive effector-related, putative, expressed
ArtTAL14-2	15	<i>Os04g03890</i>	12.955	-367	T TATGTATGACTTTTAATT	cytochrome P450, putative, expressed
ArtTAL14-2	16	<i>Os12g30580</i>	12.780	-127	T CATCTGTCCATTTCCTT	expressed protein
ArtTAL14-2	17	<i>Os08g43670</i>	12.743	-263	T CATGCGTACCTTTCCCTT	RING-H2 finger protein ATL2B, putative, expressed
ArtTAL14-2	18	<i>Os03g26260</i>	12.704	-451	C AATGTGGCCATTTCATT	NB-ARC domain containing protein, expressed
ArtTAL14-2	19	<i>Os09g19330</i>	12.697	-343	T AACATCTACCTTTTCATT	stripe rust resistance protein Yr10, putative, expressed
ArtTAL14-2	20	<i>Os07g06830</i>	12.632	-374	T CACGTGTGTTTTCATT	gibberellin receptor GID1L2, putative, expressed

¹Talvez prediction for the transcription activator-like (TAL)-DNA box pair in a whole proteome screening

²MSU version 7

³Position of TAL binding box including the 5' terminal nucleotide counted from ATG (not included)

⁴The 5' terminal nucleotide of each target box is indicated in bold

Table S4. Basal and induced *OsSWEET* gene expression levels.

Clade	Gene ID ¹	Gene name	Sample ²	2 Δ Ct ³
I	<i>Os01g65880</i>	<i>OsSWEET1a</i>	Water	0,043941
			ArtTAL1a-1	1,285167
			ArtTAL1a-2	1,520956
I	<i>Os05g35140</i>	<i>OsSWEET1b</i>	Water	0,084838
			ArtTAL1b-1	1,562180
			ArtTAL1b-2	5,531725
I	<i>Os01g36070</i>	<i>OsSWEET2a</i>	Water	0,046305
			ArtTAL2a-1	3,675345
			ArtTAL2a-2	0,075831
I	<i>Os01g50460</i>	<i>OsSWEET2b</i>	Water	0,489184
			ArtTAL2b-1	17,895277
			ArtTAL2b-2	1,849142
I	<i>Os01g40960</i>	<i>OsSWEET2c</i>	Water	0,000115
			ArtTAL2c-1	0,000307
I	<i>Os05g12320</i>	<i>OsSWEET3a</i>	Water	0,000143
			ArtTAL3a-1	0,619059
			ArtTAL3a-2	0,001398
I	<i>Os01g12130</i>	<i>OsSWEET3b</i>	Water	0,232770
			ArtTAL3b-1	1,360545
			ArtTAL3b-2	8,662907
II	<i>Os02g19820</i>	<i>OsSWEET4</i>	Water	0,078664
			ArtTAL4-1	1,456052
			ArtTAL4-2	37,445649
II	<i>Os05g50190</i>	<i>OsSWEET5</i>	Water	0,000033
			ArtTAL5-1	0,000049
II	<i>Os01g42090</i>	<i>OsSWEET6b</i>	Water	0,003852
			ArtTAL6b-1	4,999260
II	<i>Os09g08030</i>	<i>OsSWEET7a</i>	Water	0,017625
			ArtTAL7a-1	0,059922
			ArtTAL7a-2	0,000550
II	<i>Os12g07860</i>	<i>OsSWEET7c</i>	Water	0,000804

			ArtTAL7c-1	0,000544
III	<i>Os08g42350</i>	<i>OsSWEET11</i>	Water	0,001342
			ArtTAL11-1	0,744088
			ArtTAL11-2	0,364775
III	<i>Os03g22590</i>	<i>OsSWEET12</i>	Water	0,000795
			ArtTAL12-1	0,000703
			ArtTAL12-2	12,130544
III	<i>Os12g29220</i>	<i>OsSWEET13</i>	Water	0,011003685
			ArtTAL13-1	0,029532199
			ArtTAL13-2	14,0792389
III	<i>Os11g31190</i>	<i>OsSWEET14</i>	Water	0,000946
			ArtTAL14-1	0,159974
			ArtTAL14-2	2,450459
III	<i>Os02g30910</i>	<i>OsSWEET15</i>	Water	0,049225
			ArtTAL15-1	1,519168
			ArtTAL15-2	1,190815
IV	<i>Os03g22200</i>	<i>OsSWEET16</i>	Water	0,000183
			ArtTAL16-1	0,003405
			ArtTAL16-2	0,000287

¹MSU version 7

²The expression levels of the respective genes were analyzed in RNA samples from *Oryza sativa* cultivar Nipponbare or IR24 plants that were inoculated with water or with *Xanthomonas oryzae* pv. *oryzae* BA13 Δ talC carrying the respective ArtTAL

³ $2\Delta Ct = Ct(\text{gene of interest}) - Ct(\text{Actin})$ with the amplification efficiencies set to 2; values represent the mean of at least two biological replicates.

2.2.1.2 Zusammenfassung Publikation 2

Xanthomonas spp. translozieren TALEs in pflanzliche Wirtszellen. Die TALE-abhängige Aktivierung pflanzlicher Gene, sogenannter Suszeptibilitätsgene resultiert in einem veränderten Wirtsmetabolismus, der diverse Vorteile für das Pathogen bietet und essentiell für dessen Virulenz ist. Besonders in der *Xoo*-Reis Interaktion wichtige Suszeptibilitätsgene sind Mitglieder der *OsSWEET*-Familie, die Saccharose-Transporter kodieren. Der TALE PthXo1 aktiviert *OsSWEET11*, PthXo2 aktiviert vermutlich *OsSWEET13* und *OsSWEET14* wird von drei verschiedenen TALEs (TalC, Avrxa7, PthXo3) aktiviert. Im Rahmen dieser Arbeit wurde zusätzlich ein weiterer TALE (Tal5) aus dem afrikanischen *Xoo*-Stamm MAI1 identifiziert der ebenfalls *OsSWEET14* aktiviert. Die Fokussierung diverser *Xoo*-Stämme auf die TALE-abhängige Aktivierung bestimmter *OsSWEETs* lässt vermuten, dass diese Pflanzengene wichtige Suszeptibilitätsgene für *Xoo* darstellen und dass möglicherweise weitere *OsSWEETs* eine vergleichbare Funktion erfüllen. Im Reisgenom finden sich 20 *OsSWEETs*, deren Proteine phylogenetisch in vier Gruppen unterteilt werden können. Um systematisch deren mögliche Funktion als Suszeptibilitätsgene zu analysieren wurden künstliche TALEs erstellt, die jeweils einzelne *OsSWEETs* aktivieren. Es ist bereits bekannt, dass der *Xoo*-Stamm BAI3 durch die Deletion von *talC* nicht mehr virulent ist, da *OsSWEET14* nicht mehr aktiviert wird. Durch die Übertragung eines TALEs der das gleiche oder ein anderes *OsSWEET* aktiviert, kann die Virulenz von BAI3 Δ *talC* jedoch wiederhergestellt werden. Daher wurde unter Verwendung von BAI3 Δ *talC* untersucht ob künstliche TALEs, die alternative *OsSWEETs* aktivieren zur Wiederherstellung der Virulenz dieses Stammes führen. Interessanterweise konnten ausschließlich die fünf Gruppe III *OsSWEETs*, zu denen neben *OsSWEET12* und *15* auch *OsSWEET11*, *13* und *14* zählen, die Virulenz von BAI3 Δ *talC* wiederherstellen. Analysen mittels qRT-PCR zeigten, dass die künstlichen TALEs die Mehrheit der *OsSWEETs* spezifisch aktivieren konnten. Damit konnte gezeigt werden, dass Designer TALEs als künstliche Virulenzfaktoren zur Untersuchung möglicher Suszeptibilitätsgene verwendet werden können. Außerdem zeigten die Ergebnisse, dass Gruppe III SWEETs in Reis eine bestimmte Funktion erfüllen, die sich von der anderer *OsSWEETs* unterscheidet. Für *OsSWEET11* und *14* konnte bereits die Funktion als Saccharose-Transporter gezeigt werden. Möglicherweise transportieren auch die restlichen Gruppe III *OsSWEETs* Saccharose. Die TALE-abhängige Aktivierung solcher Saccharose-Transporter könnte somit zur Nährstoffversorgung der Bakterien an der Infektionsstelle beitragen.

2.3 Aberrante *repeats* erlauben die Bindung allelischer Sequenzen mit -1 *frameshift*

2.3.1 Publikation 3



ARTICLE

Received 7 Nov 2013 | Accepted 13 Feb 2014 | Published 11 Mar 2014

DOI: 10.1038/ncomms4447

A TAL effector repeat architecture for frameshift binding

Annekatrin Richter^{1,*}, Jana Streubel^{1,*}, Christina Blücher¹, Boris Szurek², Maik Reschke¹, Jan Grau³ & Jens Boch¹

Transcription activator-like effectors (TALEs) are important *Xanthomonas* virulence factors that bind DNA via a unique tandem 34-amino-acid repeat domain to induce expression of plant genes. So far, TALE repeats are described to bind as a consecutive array to a consecutive DNA sequence, in which each repeat independently recognizes a single DNA base. This modular protein architecture enables the design of any desired DNA-binding specificity for biotechnology applications. Here we report that natural TALE repeats of unusual amino-acid sequence length break the strict one repeat-to-one base pair binding mode and introduce a local flexibility to TALE-DNA binding. This flexibility allows TALEs and TALE nucleases to recognize target sequence variants with single nucleotide deletions. The flexibility also allows TALEs to activate transcription at allelic promoters that otherwise confer resistance to the host plant.

¹Department of Genetics, Martin Luther University Halle-Wittenberg, Weinbergweg 10, D-06120 Halle (Saale), Germany. ²UMR 186 IRD-UM2-Cirad 'Résistance des Plantes aux Bioagresseurs', BP 64501, 34394 Montpellier cedex 5, France. ³Institute of Computer Science, Martin Luther University Halle-Wittenberg, von-Seckendorff-Platz 1, D-06120 Halle (Saale), Germany. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.B. (email: jens.boch@genetik.uni-halle.de).

Transcription activator-like effectors (TALEs) employ a programmable DNA-binding domain and have revolutionized biotechnology approaches that use sequence-specific DNA-binding proteins^{1–3}. Natural TALEs are injected by plant pathogenic *Xanthomonas* spp. bacteria via a type-III secretion system into plant cells⁴. Inside the plant cell they function as transcription factors that activate target gene expression and support bacterial colonization^{2,5}. TALEs contain a unique DNA-binding domain of tandem near-identical 34-amino-acid repeats. The repeats are highly conserved and mainly differ in two adjacent amino acids (positions 12 and 13) termed repeat-variable diresidue (RVD)^{5,6}. In the array each repeat recognizes one nucleotide in the target DNA sequence, and the RVD specifies which base is bound^{6,7}. In addition to the repeat region, the N-terminal part of TALEs contains four degenerated repeats (termed repeat 0, –1, –2 and –3) that contribute to DNA-binding⁸. Repeat –1 specifies thymine that typically precedes TALE target sequences^{9,10}. The C-terminal part of TALEs contains two functional nuclear localization signals as well as an acidic activation domain that is important for gene activation^{4,11}.

Xanthomonas oryzae pv. *oryzae* (*Xoo*) and *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*) cause bacterial leaf blight and bacterial leaf streak, respectively, which are two of the most devastating diseases of the staple crop rice^{12,13}. Both pathogens contain a particularly large number (7–28) of TALE genes per strain¹⁴ and the TALE-dependent modulation of host gene expression profiles is an important feature of *Xoo* and *Xoc* diseases. In some cases, loss of a single TALE gene severely compromised bacterial virulence⁵. The best-studied group of TALE virulence targets is the *SWEET* gene family in rice. Three members of this gene family are targets of several TALEs from different *Xoo* strains. *OsSWEET11* and *OsSWEET13* are targeted by PthXo1 and probably PthXo2, respectively, whereas *OsSWEET14* is targeted by four different TALEs, TalC, AvrXa7, PthXo3 and TAL5, originating from different *Xoo* strains^{15–21}. Furthermore, it was shown that *OsSWEET12*, *OsSWEET13* and *OsSWEET15* support bacterial colonization if induced by designer TALEs^{16,19}. Considering that >100 *Xanthomonas* TALEs are known the overall number of identified plant target genes is low. One approach to overcome this is by using computational algorithms to predict possible TALE targets^{22–24}. Using the TALE RVD–DNA code the promoter sequences of host plants are scanned for potential TALE target sites. Promising target candidates contain a TALE target site and are TALE-dependently activated^{23,24}.

Plants have evolved resistances that are based on TALE-dependent activation of genes that trigger a resistance response towards *Xanthomonas* strains delivering the matching TALE^{25–27}. A different resistance mechanism is based on mutation of potential TALE target DNA sequences^{17,20,28,29}. These mutations can render virulence targets non-responsive to TALEs, thereby efficiently preventing the contribution of a given TALE to bacterial virulence. If target gene upregulation is important for virulence of the pathogen, the plants become resistant. Accordingly, rice varieties carrying insertions, deletions or substitutions in the promoter sequences of *OsSWEET11* and *OsSWEET12* are not susceptible to *Xoo* strains carrying the TALEs PthXo1 and PthXo2, respectively^{15,17,28–30}.

The modular TALE architecture allows a free combination of repeats to generate any desired DNA-binding specificity^{7,31,32}. This flexibility resulted in adoption of the TALE repeats as specific DNA-binding domain that can be fused to executor domains to generate different biotechnological tools². The most widespread use is based on fusions of TALEs with a nuclease domain (termed TALEN) to edit eukaryotic genomes at specific sites^{3,31,33,34}. Other executor domains were applied for gene activation, gene repression, chromatin modification,

fluorescent tagging of chromosomal loci and chromatin affinity purification^{35–40}.

The three-dimensional (3D) structures of TALE–DNA complexes have been solved^{8–10,41}. TALE repeats form a right-handed superhelix that wraps around the DNA double strand. The helix–loop–helix structure of a single repeat exposes the RVD amino acid at position 13 to interact with a leading strand base from the DNA major groove. In contrast, the RVD amino acid at position 12 stabilizes the RVD loop by interaction with the carbonyl of position 8 of the same repeat^{9,10,41}. This highly regular repeat architecture is likely necessary to position consecutive repeats correctly to a continuous string of DNA bases. The 3D structures show that in a TALE repeat non-RVD amino acids mediate inter- and intra-repeat interactions as well as connections to the DNA phosphate backbone. This probably supports the correct positioning of the repeat array to the DNA^{9,10,41}. The TALE 3D structures also suggest that significant changes in the repeat length may have a profound impact on DNA binding. Two fairly common exceptions to the typical 34-amino-acid length of TALE repeats exist. These are either 33-amino-acid repeats that have a deletion of the RVD amino acid 13 or 35-amino-acid repeats containing a proline following amino acid 32. In both cases the DNA-binding specificity of TALEs is not altered^{7,36,42}. Thirty-five-amino-acid repeats are also typical in TALE homologues from *Ralstonia solanacearum* for which a RVD-guided DNA-binding activity analogous to *Xanthomonas* TALEs was shown^{43,44}. Interestingly, some naturally occurring *Xanthomonas* TALEs possess a single repeat of aberrant length besides the canonical 33–35-amino-acid repeats^{5,45}. These aberrant repeats consist of 30 amino acids (with a 4-amino-acid deletion in the second helix of the repeat), 39 or 40 amino acids (duplication in the second helix of the repeat), or 42 amino acids (duplication in the first helix of the repeat)⁵. It is not known how repeats that differ in length from the normal 33–35 amino acids influence the DNA-binding behaviour of TALEs.

Here we analyse how the insertion of a repeat of aberrant length in the canonical repeat array influences TALE function. We find that such TALEs exhibit two possible binding conformations, the normal one and a novel one that tolerates single nucleotide deletions in the target sequence. We show that this binding behaviour expands the recognition specificity of TALEs and TALENs, respectively. Furthermore, we find that the flexible TALE-binding behaviour provides a potential evolutionary solution for *Xanthomonas* to overcome plant resistance.

Results

Natural TALEs with aberrant repeats. AvrXa7 and PthXo3 are natural TALEs from *Xoo* with 25.5 and 28.5 repeats, respectively, that both contain one aberrant repeat of 39 amino acids approximately in the middle of the repeat array (Fig. 1; Supplementary Figs 1 and 2 and Supplementary Table 1). Both TALEs are important virulence factors that support growth of *Xoo* strains on rice¹⁵, which demonstrates that they are functional. Indeed, both TALEs induce expression of the rice sugar exporter *OsSWEET14* by binding to overlapping target boxes (Fig. 1b and Supplementary Figs 1 and 2)^{15,18,19}. In contrast, computational predictions of AvrXa7 and PthXo3 target sequences rank the PthXo3 site in the *OsSWEET14* promoter at position 119 (ref. 23). We noticed a high number of non-matching RVD base combinations in the 3' part of the PthXo3 target box following the aberrant repeat (Fig. 1c,d and Supplementary Fig. 2). If the C-terminal half of the repeat array is shifted one nucleotide upstream by looping out the long repeat, the PthXo3 target box in *OsSWEET14* will rank at position 1 in target predictions (Fig. 1c and Supplementary Table 2; ref. 23).

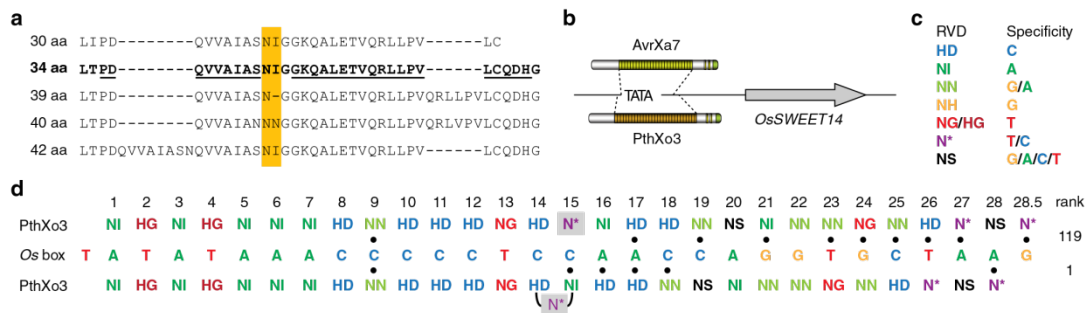


Figure 1 | Aberrant repeat length variants from *Xanthomonas* spp. TALEs. (a) Alignment of amino acid (aa) sequences of natural TALE repeats with different length. The repeat-variable diresidues (RVDs) are boxed in yellow. A typical 34-aa repeat is in bold face and the residues forming the two α -helices are underlined. (b) Cartoon of AvrXa7 and PthXo3 binding to their overlapping target sequences at the TATA-box in the *OsSWEET14* (*Os11g31190*) promoter. (c) TALE RVDs and their DNA base specificities. (d) PthXo3 RVDs with the aberrant repeat in normal or looped-out conformation aligned to the *OsSWEET14* promoter target sequence (Os box). A dot indicates a non-matching RVD-base combination. The rank of both alignments in a target site prediction is indicated.

To test the recognition specificities of AvrXa7 and PthXo3, we build reporter constructs with an optimal target box according to the TALE specificity code, and derivatives with nucleotide deletions (-1 , -2) or insertions ($+1$, $+2$) at the nucleotide position behind the one corresponding to the aberrant repeat, respectively (Supplementary Figs 1a and 2a). The boxes are inserted upstream of a minimal promoter that has no basal activity⁷. In addition, reporter constructs with either the natural box of AvrXa7 in front of the minimal promoter or the *OsSWEET14* promoter fragment containing the natural AvrXa7/PthXo3 box are used (Supplementary Figs 1a and 2a, ref. 21). The reporter constructs are then co-transformed with expression constructs of AvrXa7 or PthXo3 using *Agrobacterium* into leaves of *Nicotiana benthamiana* plants. β -glucuronidase (GUS) assays reveal that both TALEs induce expression of the reporter containing either one of the target boxes although they exhibit several non-matching RVD-base combinations (Supplementary Figs 1b–g and 2b–h). Possible explanations for this surprising result are that AvrXa7 and PthXo3 contain a large number of repeats, several RVDs with broad specificity like NS, NN and N*, and one repeat of aberrant length. We aim to test the role of the aberrant repeat in a more controlled TALE design.

Aberrant repeats permit a novel TALE-DNA-recognition mode. We construct artificial TALEs⁴⁶ with a 17.5-repeat array that is designed to result in a maximum of non-matching RVD-base combinations upon a possible frameshift in the target sequence. TALEs are assembled with and without single aberrant repeats of 30-, 40- and 42-amino-acid length at position 8 (Figs 1a and 2a and Supplementary Table 1), and analysed *in planta* for the activation of GUS reporters with optimal and frameshift (-1 , -2 , $+1$ and $+2$) target boxes (Fig. 2a,b). The aberrant repeat sequences including their RVDs correspond to the natural *X. oryzae* TALE sequences (Supplementary Table 1).

TALEs with or without an aberrant repeat trigger GUS activity with the reporter construct containing the optimal box. This indicates that all TALEs tested recognize the optimal box and that aberrant repeats function similarly to normal ones in a repeat array. Significantly, only TALEs with an aberrant repeat result also in strong GUS activity with reporter constructs containing the -1 box (Fig. 2b). The same effect is apparent in artificial TALEs with a different repeat composition (Supplementary

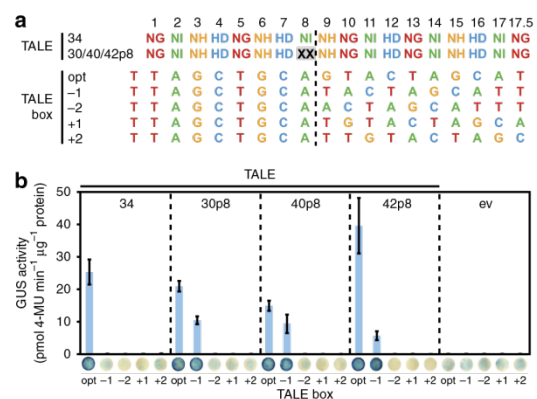


Figure 2 | Aberrant repeats allow a flexible recognition of target DNA sequences with a -1 nucleotide frameshift. (a) RVDs of artificial TALEs and target boxes. A TALE with all 34-aa repeats or TALEs with an aberrant 30-aa NI-repeat, 40-aa NN-repeat and 42-aa NI-repeat, respectively, at position 8 are constructed. An optimal target box (opt) and derivatives with deletions or insertions of one or two nucleotides ($-$ or $+1$ or $+2$) after position 8 (dashed line) are fused to a minimal promoter and a promoterless GUS reporter gene. (b) GUS assays of TALEs and reporter constructs ($n=3$). 35S-driven GFP expression serves as empty vector (ev) control in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown.

Fig. 3). Thus, a TALE containing exclusively canonical 34-amino-acid repeats is not able to recognize a -1 box, because of many (in this case nine) non-matching RVD-base combinations in either the front or rear part of the box, depending on how the TALE RVDs are aligned to the box. In contrast, a TALE with an aberrant repeat efficiently activates the reporter. These data suggest that an aberrant repeat confers a flexible binding to TALEs. Interestingly, this flexible binding mode is possible with different aberrant repeats that are shorter (30 amino acids) or longer (39, 42 amino acids). The -2 , $+1$ and $+2$ boxes are not recognized by any of the TALEs, indicating that other binding modes are not supported.

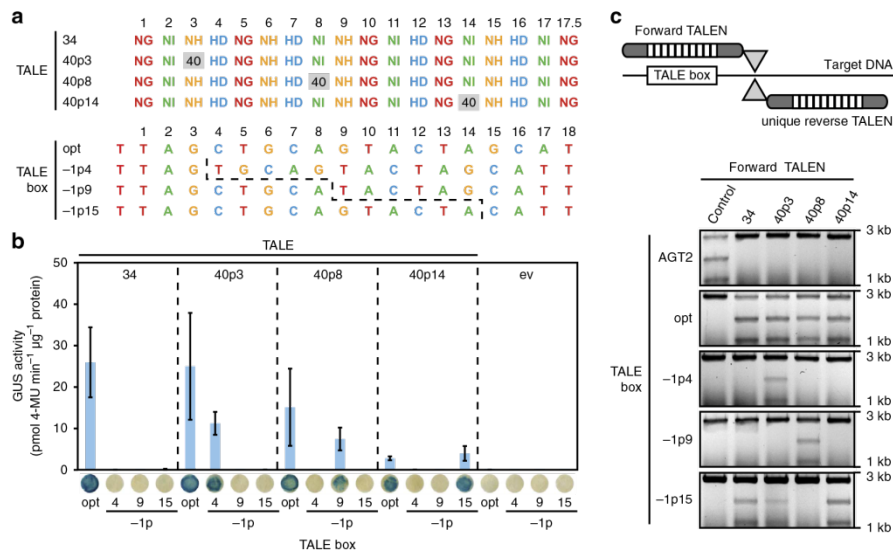


Figure 3 | Aberrant repeats function at different positions in the repeat array of TALEs and TALENs. (a) TALE RVDs and target boxes. Artificial TALEs or TALENs are constructed with all 34-aa repeats or an aberrant 40-aa NN-repeat inserted at position 3 (40p3), position 8 (40p8) or position 14 (40p14). The NN-repeat recognizes both G and A DNA bases. The boxes are either perfectly matching the specificity of normal repeats (opt) or have one base pair deleted to the right of the dashed line at position 4 (-1p4), position 9 (-1p9) or position 15 (-1p15). (b) GUS assays of TALEs and reporter constructs ($n=3$). 35S-driven GFP expression serves as empty vector (ev) control in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown. (c) Cartoon of TALENs bound to DNA and *in vitro* TALEN restriction assay. The TALEN pairs are placed such that the FokI domain (grey triangle) can dimerize and cut the DNA. The unique reverse TALEN is constant in all assays. The control TALEN recognizes a different target box (AGT2). Target DNA is incubated with *in vitro*-transcribed and translated TALEN pairs. Restriction fragments are documented on agarose gels.

Aberrant repeats function at different repeat array positions.

As in nature single aberrant repeats are positioned more or less in the central region of the repeat array (Supplementary Table 1), we investigate whether TALEs tolerate only this arrangement. We insert the aberrant 40-amino-acid repeat at different positions (positions 3, 8 or 14) in the previously used 17.5 TALE repeat array that is highly susceptible to frameshift in its target sequence (Fig. 3a). As target boxes we use either the optimal one or sequences with frameshift at positions 4, 9 and 15, respectively (Fig. 3a). GUS reporter assays reveal that TALEs that carry the aberrant repeat at any of the three positions are functional. All TALEs with aberrant repeats recognize the optimal box and the one with a frameshift following the position corresponding to the aberrant repeat in the TALE (Fig. 3b). However, the TALE with an aberrant repeat at position 14 displays only weak activity on both the optimal and the corresponding -1 frameshift box (Fig. 3b). This suggests that it is not favourable for the overall activity of TALEs if only a small number of repeats (in this case four) follow an aberrant repeat in the array.

Flexible TALENs with aberrant repeats. TALENs have become state-of-the-art tools for genome editing³. TALENs are TALE-DNA-binding domain fusions to the *FokI* endonucleolytic domain, which act in pairs to enable *FokI* dimerization and DNA cleavage³. We test whether aberrant repeats can change the DNA recognition behaviour of TALENs. TALENs with and without an aberrant repeat are assembled. The aberrant 40-amino acid repeat is placed at positions 3, 8 and 14, respectively, in a TALEN of 17.5 repeats (Fig. 3a). These 'forward' placed TALENs

are combined with a common 'reverse' placed TALEN with 34-amino-acid repeats (Fig. 3c and Supplementary Fig. 4b) and linear DNA fragments containing the target boxes in *in vitro* restriction assays. The optimal box is cleaved using TALENs with or without an aberrant repeat emphasizing that the normal binding mode is also supported in TALEN assays. In contrast, the boxes with -1 frameshift at position 4 and 9 are only cleaved in the presence of the TALENs with an aberrant repeat at position 3 and 8, respectively (Fig. 3c and Supplementary Fig. 5). This is in accordance with our observation for TALEs before (Fig. 3b) that aberrant repeats infer a local flexibility to the repeat array.

In contrast, the box with the -1 frameshift at position 15 is recognized not only by the TALEN with an aberrant repeat at position 14 but also by the normal all 34-amino-acid repeat TALEN, and to a lesser degree by the TALEN with an aberrant repeat at position 3. Apparently, these TALENs tolerate three non-matching RVD-base combinations at the end of the repeat array (Fig. 3c). This mismatch tolerance of rear repeats in TALENs was not observed in the previous experiments using TALEs in reporter gene activation assays (Fig. 3b). It suggests that binding of the rear TALE repeats to DNA is less important for dimerization of the nuclease domains than for function of the natural activation domain.

Furthermore, we test whether the TALEN with the aberrant 40-amino-acid repeat at position 8 can perform cleavage of -2, +1 or +2 frameshift boxes (Supplementary Fig. 4). We observe that the TALENs with or without the long repeat do not exhibit activity on these boxes (Supplementary Fig. 4c), suggesting that other binding modes are not supported by TALENs similarly to our observation for TALEs (Fig. 2). In summary, our experiments

demonstrate that aberrant repeats initiate a flexible DNA-binding behaviour *in vivo* and *in vitro*, and that they can be used to expand the recognition specificities of artificial TALEs as well as TALENs.

The aberrant repeat is excluded from the interaction. So far, it is not clear which repeat of the repeat array is excluded from the interaction, for example by looping out. Our results allow several possible explanations: either the aberrant repeat itself is excluded or the repeat up- or downstream of the aberrant repeat. Therefore, we test TALEs with aberrant repeats at position 8 combined with target boxes (–1) deleted in either the nucleotide at position 7, 8 or 9 in GUS reporter assays (Fig. 4a). We expect that the TALE will yield the highest reporter activity in combination with the –1 box that has the deletion exactly in opposite to the repeat that is excluded, because the other boxes will produce at least one non-matching RVD–base combination. Indeed, the target sequence with the nucleotide deleted at position 8 shows the highest activity of the three –1 boxes for the TALEs with 40- and 42-amino-acid aberrant repeats (Fig. 4b). The TALE with a 30-amino-acid aberrant repeat has a similarly high activity with the –1 boxes deleted at position 8 and 9, respectively (Fig. 4b). This suggests that it is the aberrant repeat itself that is excluded in frameshift binding.

To corroborate this finding, we use a TALE with a 40-amino-acid aberrant repeat placed at position 3 in either a 17.5 repeat array or an 11.5 repeat array (Fig. 4c). It has been described that the initial repeats have a stronger impact on the overall TALE binding than later repeats^{37,47}. Therefore, we reason that these TALEs will have a clear preference for the –1 box in which the nucleotide is missing that corresponds to the repeat that is excluded. The TALEs were combined with an optimal target box or boxes deleted in the nucleotide at position 3 or 4 (Fig. 4c). Indeed, the

TALEs induce a higher reporter activity with the box carrying a –1 deletion at position 3 that corresponds to the position of the aberrant repeat than with the box with the deletion at position 4 (Fig. 4d). The shorter TALEs have a significantly weaker activity suggesting that the overall DNA recognition was partly compromised (Fig. 4d). In summary, we postulate that the aberrant repeat itself loops out of the repeat array when the TALE is bound to a target sequence with a single nucleotide deletion, but is inserted into the array when bound to an optimal box (Fig. 7).

Tandem aberrant repeats are not flexible. We analyse whether tandem aberrant repeats are accepted in a TALE repeat array. For this, we generate a TALE with aberrant 40-amino-acid repeats at position 8 and 9 in the frameshift-sensitive 17.5 repeat array used before (Fig. 5a). The TALE with the two aberrant repeats recognizes exclusively the box with a –1 frameshift at position 9 and neither the optimal one nor a –2 or +1 frameshift box (Fig. 5b). Apparently, the tandem aberrant repeats can only be compensated by one repeat being excluded from the interaction. This suggests that two neighbouring 40-amino-acid repeats can neither loop out simultaneously nor be arranged in the consecutive array of normal repeats. It cannot be excluded that aberrant repeats destabilize the TALE protein structure in certain repeat arrangements although we do not detect strong differences in protein amounts of the artificial TALEs used here and in the previous experiments with or without aberrant repeats, respectively (Supplementary Fig. 6).

Aberrant repeats can participate in TALE–DNA recognition. We want to clarify whether the RVD of the aberrant repeat participates in DNA base recognition when the TALE binds in the regular fashion to an optimal box. For this, we compare how

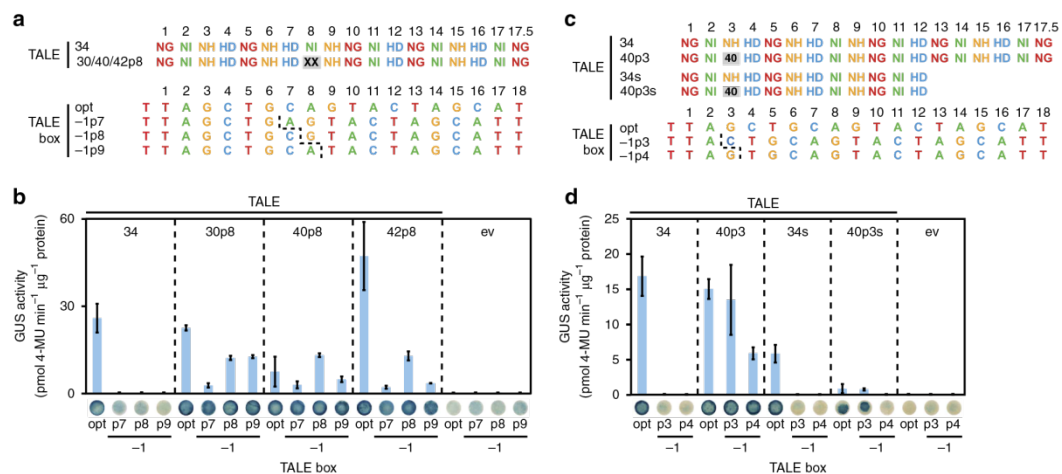


Figure 4 | Aberrant repeats allow recognition of target box frameshift close to their position. (a) RVDs of artificial TALEs and target boxes. A TALE with all 34-aa repeats or TALEs with an aberrant 30-aa NI-repeat, 40-aa NN-repeat and 42-aa NI-repeat, respectively, at position 8 are constructed. TALE boxes are either perfectly matching the specificity of normal repeats (opt) or have one base pair deleted to the right of the dashed line at position 7 (–1p7), position 8 (–1p8) or position 9 (–1p9). (b) GUS assays ($n=3$) of TALEs and reporter constructs described in (a). (b,d) 35S-driven GFP expression serves as empty vector (ev) control in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown. (c) TALE RVDs and target boxes. Artificial TALEs with either 17.5 or 11.5 repeats are constructed with a normal 34-aa NH-repeat or an aberrant 40-aa NN-repeat (grey box) inserted at position 3. The target boxes are either perfectly matching the specificity of normal repeats (opt) or have one base pair deleted to the right of the dashed line at position 3 (–1p3) or position 4 (–1p4). (d) GUS assays ($n=3$) of TALEs and reporter constructs described in (c).

TALEs with 34-amino-acid repeats and TALEs with an aberrant repeat deal with non-matching bases around the position of the aberrant repeat. We expect that a TALE with an aberrant repeat will display a reduced activity at the mismatch boxes if the RVD of the aberrant repeat participates in base recognition. TALEs

with only 34-amino-acid repeats or an aberrant repeat of 30-, 40- or 42-amino-acid length, respectively, at position 8 are used. The aberrant repeat RVDs are NI or NN, which are compatible with the base adenine, but not thymine. We combine the TALEs with target boxes that contain 1–3 non-matching bases starting at position 8 or 9 (Fig. 6a). The TALE with normal 34-amino-acid repeats trigger significantly decreasing reporter activity with increasing number of mismatches (Fig. 6b). The TALEs with aberrant repeats show a very similar pattern. In particular, at the box with one mismatch at the position of the aberrant repeat (box 8.1; Fig. 6b) reporter activity is significantly less than at the optimal one. This indicates that all tested aberrant repeats contribute to DNA–base recognition when the TALE binds an optimal box (Fig. 7).

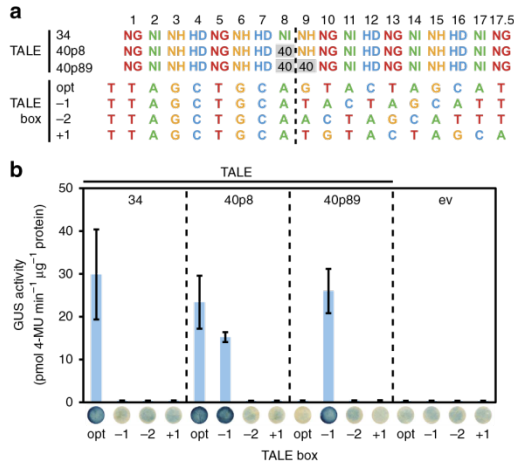


Figure 5 | Two aberrant repeats can be combined in tandem. (a) RVDs of artificial TALEs and target boxes. A TALE with all 34-aa repeats or TALEs with one or two 40-aa NN-repeats at position 8 (40p8) or 8 and 9 (40p89) are constructed. An optimal target box (opt) and derivatives with deletions of one or two nucleotides (–1 or –2) or an insertion of one nucleotide (+1) after position 8 (dashed line) are fused to a minimal promoter and a promoter-less GUS reporter gene. (b) GUS assays of TALEs and reporter constructs (n = 3). 35S-driven GFP expression serves as empty vector (ev) control in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown.

Aberrant repeats influence AvrXa7 and PthXo3 target range. The natural aberrant repeats provide an extended flexibility to the repeat array that is not present in TALEs with standard repeats, and it is tempting to speculate that this has evolved to benefit *Xanthomonas* virulence. One possibility is that aberrant repeats enable *Xanthomonas* to recognize promoter variants in different plant cultivars or species. We investigate whether there are natural insertion/deletion (indel) mutations in the *O_sSWEET14* promoter region targeted by PthXo3. We compare this region in full genomic alignments of *Oryza sativa japonica* cv. Nipponbare, *O. sativa indica*, *O. brachyantha* and *O. glaberrima*. While we do not observe a difference between *O. sativa japonica* and *O. sativa indica*, we indeed find a single base-pair insertion at position 18/19 of the PthXo3 box in *O. brachyantha*, which may be compensated by the aberrant repeat of PthXo3 (Supplementary Fig. 7). Interestingly, while binding of PthXo3 to its target box in *O. sativa* is preferred in the looped-out conformation (6 instead of 10 mismatches), binding in the normal conformation should be preferred in *O. brachyantha* (4 instead of 9 mismatches; Supplementary Fig. 7).

To study the contribution of the aberrant repeat in the natural TALE AvrXa7 on target recognition, we assemble an artificial TALE with the same RVD composition as AvrXa7 (termed ArtXa70) and a derivative (ArtXa71) with a normal 34-amino-acid

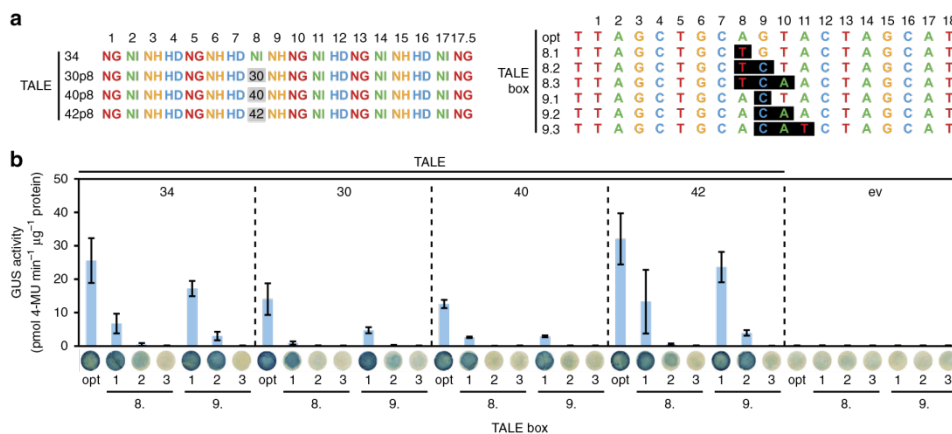


Figure 6 | Aberrant repeats participate in TALE–DNA base pair recognition. (a) TALE RVDs and target boxes. Artificial TALEs are constructed with all 34-aa repeats or aberrant repeats. An aberrant 30-aa NI-repeat, 40-aa NN-repeat and 42-aa NI-repeat is inserted at position 8, respectively. Target DNA boxes are either perfectly matching the specificity of normal repeats (opt) or have one to three non-matching bases (transversions) between position 8 and 11 (black boxes). (b) GUS assays of TALEs and reporter constructs (n = 3). 35S-driven GFP expression serves as empty vector (ev) control in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown.

repeat instead of the aberrant 39-amino-acid-repeat at position 13 (Supplementary Fig. 8a,b). Both artificial TALEs recognize the optimal box in GUS assays similarly well as AvrXa7, but the TALE without the aberrant repeat (ArtXa71) is significantly compromised on the -1 target box (Supplementary Fig. 8c). This indicates that the aberrant repeat also contributes to the flexible target box recognition of the natural TALE AvrXa7 (Supplementary Fig. 1).

TALEs with aberrant repeats can break plant resistance. To test whether TALEs with aberrant repeats contribute to bacterial virulence, we design TALEs based on a naturally occurring

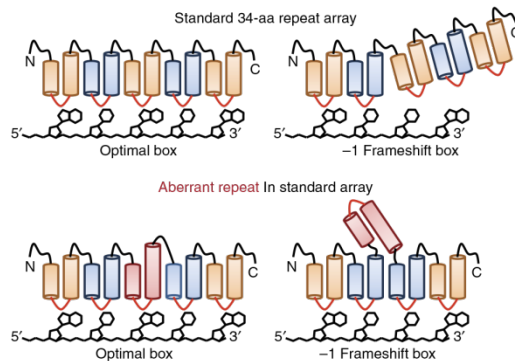


Figure 7 | Cartoons of repeat arrays in the normal and the looped-out conformation. Model of TALE repeats consisting of standard 34-aa repeats and repeat arrays containing one repeat of aberrant length, respectively, aligned to an optimal box or a -1 frameshift box. The aberrant repeat is shown in dark red.

promoter allele that confers plant resistance. Rice *Xa25* (*OsSWEET13*) encodes a SWEET protein that supports *Xanthomonas* virulence probably facilitated by the TALE PthXo2 (refs 16,19). A natural single base-pair insertion mutation in the recessive *xa25* allele results in plants resistant to *Xoo* likely by prohibiting binding of PthXo2, which does not contain any aberrant repeats¹⁷ (Supplementary Fig. 9a). To test whether TALEs with aberrant repeats can overcome this resistance, we construct TALEs targeting the PthXo2-binding site in the recessive *xa25* allele either with 17.5 exclusively regular 34-amino-acid repeats (ArtXa251, Fig. 8 and Supplementary Fig. 9b) or one aberrant repeat at position 4 (ArtXa252, Fig. 8 and Supplementary Fig. 9c). The *xa25* and *Xa25* promoter regions are cloned from *O. sativa* cv. Nipponbare and *O. sativa* cv. Zhenshan 97, respectively¹⁷. Agrobacterium-mediated production of ArtXa251 triggers reporter gene activation of the *xa25* (Nipponbare; that is, optimal box), but not the *Xa25* (Zhenshan 97; that is, -1 frame-shift box) promoter (Supplementary Fig. 9d), supporting that normal TALEs are highly compromised by indel mutations. In contrast, the TALE with an aberrant repeat, ArtXa252, induces expression of both reporter constructs (Supplementary Fig. 9d).

We further analyse the contribution of TALEs with aberrant repeats to *Xanthomonas* spp. infections. The *Xoo* strain BAI3 causes disease on rice cultivars Azucena and Zhenshan 97 in part because it contains the TALE TalC that directs expression of the gene encoding the sugar exporter OsSWEET14 (ref. 21) (Fig. 8). Deletion of *talC* renders the bacteria unable to cause disease²¹ (Fig. 8), but this can be compensated by artificial TALEs that trigger expression of *OsSWEET13* (ref. 19). We introduce ArtXa251 and ArtXa252 into *Xoo* BAI3 Δ talC and infect rice cultivars Azucena (*xa25*) and Zhenshan 97 (*Xa25*), which differ in their *OsSWEET13* promoter region by the 1-bp indel mutation (Fig. 8). The *Xoo* strain containing ArtXa251 with exclusively normal repeats causes disease symptoms on *O. sativa* cv. Azucena, but not *O. sativa* cv. Zhenshan 97 (Fig. 8), because of

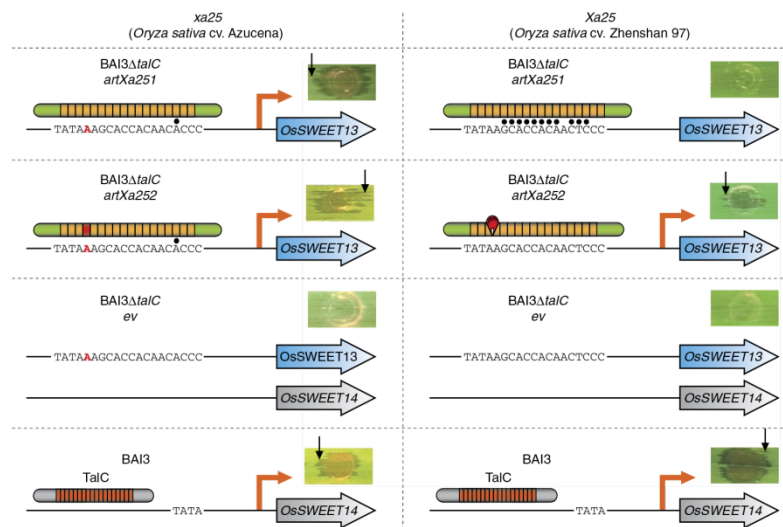


Figure 8 | TALEs with aberrant repeat can overcome plant resistance. Disease phenotypes of *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strains on *Oryza sativa* leaves. Leaves of 3–4 week-old Azucena or Zhenshan 97 plants are inoculated with strain BAI3 or BAI3 Δ talC carrying empty vector (ev) or TALE *artXa251* or *artXa252*. Pictures are taken 4 days post inoculation. Azucena and Zhenshan 97 contain alleles of *OsSWEET13*, termed *xa25* and *Xa25*, respectively, which differ in their promoter sequences by an additional nucleotide (red). The strain harbouring ArtXa252 with a 40-aa repeat (red) causes water soaking disease symptoms (black arrows) on both alleles. Black dots indicate RVD-base-mismatches. The experiments are performed twice with similar results.

the indel mutation in the target box. In contrast, *Xoo* containing ArtXa252 with the aberrant repeat causes disease symptoms on both rice cultivars (Fig. 8). This demonstrates that aberrant repeats can enable *Xanthomonas* to compensate for indels in target promoters and break natural plant resistances. Apparently, aberrant repeats are an evolutionary solution for *Xanthomonas* spp. to overcome small indel mutations that otherwise efficiently block TALE binding.

Discussion

We have described an exceptional and surprising recognition pattern of TALEs that has implications for the general concept of TALE–DNA interaction as well as the evolutionary adaptation of these important virulence factors. The near-identical amino-acid sequence of *Xanthomonas* TALE repeats implies that this regularity is required for binding to the highly symmetric structure of the DNA double helix^{9,10,14}. Nevertheless, in nature the glycine and the leucine at positions 14 and 29, respectively, are the only amino acids conserved in all TALE repeat sequences⁵. Even more striking are natural repeat variants that deviate from the typical 34- or 35-amino-acid length of TALE repeats, because they are expected to impose a structural problem to the overall repeat array. Here, we show that either shorter (30 amino acids) or longer repeat variants (39, 40 or 42 amino acids) change the DNA-binding behaviour of TALEs. They can either insert into the repeat array like normal repeats or be excluded from the interaction to facilitate a shift of the following repeats forward by one position depending on the best fit of all RVDs to a given target sequence (Figs 2 and 7). This behaviour is not possible for TALEs with normal 34-amino-acid repeats. In the absence of structural data, we favour a model that the aberrant repeat loops out of the repeat array, because this seems to be sterically the easiest solution; alternative scenarios are possible though. Surprisingly, all aberrant repeats tested functioned in a comparable way although we have not compared them in all assays in our study. In part this might be explained by the fact that the aberrant repeat sequences are all related to a normal 34-amino-acid TALE repeat. They contain either a short deletion in the second α -helix (30-amino-acid repeat) or a duplicated first (42-amino-acid repeat) or second α -helix (39/40-amino-acid repeats) flanking the RVD (Fig. 1). Possibly, these duplications of structural elements still allow the hydrophobic repeat-to-repeat interactions that stabilize the normal repeat array^{9,10}.

At the same time, the aberrant repeats probably do not fit perfectly into the array, thereby weakening inter-repeat interactions and causing a local flexibility and structural tension that allows the aberrant repeat to loop out. In general, TALEs exhibit a highly flexible structure. Molecular dynamics simulations show that the TALE repeat region has a high conformational plasticity⁴⁸. In addition, structural data and computational simulations indicated that the TALE repeat superstructure condenses upon interaction with cognate DNA resulting in more densely packed repeats than in the DNA-free form^{9,41,48}. Likely, the flexible nature of the repeat region allows the unique binding mode conferred by aberrant repeats.

How does an aberrant repeat influence the dynamics of TALE protein–DNA interaction? Insertion of aberrant repeats in the rear part of the repeat array compromised TALE-mediated gene induction in our experiments (Fig. 3). We postulate that the aberrant repeat weakens the protein–DNA interaction and that a minimum number of repeats following the aberrant one is required to allow subsequent condensation of the rear repeats onto the DNA double helix. The N-terminal domain of TALEs mediates unspecific interaction to DNA as well as recognition of the initial thymine^{7,8,10}. In addition, the initial repeats after the

N-terminal domain are more important for binding than downstream repeats^{37,47}. Together, these observations support a model where the N-terminal TALE region scans the DNA and the repeats subsequently condense at target sequences onto the DNA in a consecutive fashion starting from the N-terminal region. Possibly, an aberrant repeat that is excluded from binding in a frameshift scenario only allows condensation of further repeats, if their number is sufficiently high to condense onto the DNA in a separate event. Accordingly, the aberrant repeat might separate the repeat domain into two binding domains. In general, the tolerance of aberrant repeats in a repeat array likely depends on the overall binding energies of preceding and following repeats. Surprisingly, two tandem aberrant repeats only recognize a -1 frameshift box and not a normal one nor a -2 frameshift box (Fig. 5). This implies that the loop out conformation has specific requirements and does not allow looping out of more than one repeat. Furthermore, the additional structural perturbation of two near aberrant repeats does not allow a normal repeat arrangement. It is presently difficult to estimate how many aberrant repeats may be accepted in a TALE repeat domain.

The novel binding mode comprises interesting potential for biotechnology applications that are otherwise difficult to achieve. TALEs and TALENs with aberrant repeats can be used to simultaneously recognize allelic variants that differ by single nucleotide frameshifts. For synthetic TALE biology⁴⁹, master regulators with aberrant repeats might be used to control two subsets of target sequences that differ by indel mutations. This is difficult to achieve with CRISPR/Cas9 recognition of DNA sequences, because its RNA-guided recognition mechanism cannot compensate for indel mutations⁵⁰. In addition, it is tempting to speculate that the aberrant repeats can function as an insertion point for more complex peptide tags or proteins within the TALE repeat domain to generate unique fusion proteins.

A significant aspect of our study is that these aberrant repeats are naturally evolved variants. Their presence in important virulence factors implies that they confer a selective advantage. The battle between pathogen and host is characterized by a constant struggle for innovation and counteracting activities to prevent losing the evolutionary race. Inherently, TALEs are especially sensitive to frameshift mutations in their target sequence and this has been exploited by natural plant resistances^{16,20,25}. An aberrant repeat contributes a new degree of flexibility to the DNA-binding activity of a TALE without strong penalty on overall activity. We showed that TALEs with aberrant repeats can enable *Xoo* to overcome a natural resistance that is caused by a 1-bp deletion. As another example, the aberrant repeat in PthXo3 and AvrXa7 might have evolved to recognize unknown target box variants in response to a specific mutation that occurred in the *OsSWEET14* promoter in some rice cultivars. Alternatively, our computational search of available *OsSWEET14* promoter sequences raises the possibility that this potent virulence factor might favour colonization of *Xoo* on the different rice species *O. sativa* and *O. brachyantha*. So far, the possibility that TALEs of one bacterial strain support virulence in different host species has not been explored, but it is typical for type III effectors that they function in different plants and even non-plants^{51,52}. Further analysis of natural TALEs with aberrant repeats and their target promoters will clarify the role of these unique virulence factors for the pathogen. Our data reveal that designer resistances based on mutations in TALE target boxes¹⁶ have to be considered carefully to be effective.

Methods

Bacterial strains and growth conditions. *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) BAI3 and BAI3 Δ talC are used in this study^{21,53}. Plasmids are introduced into *Xoo* by conjugation using pRK2013 as a helper plasmid in triparental matings⁵⁴.

Rifampicin (100 $\mu\text{g ml}^{-1}$) and gentamicin (20 $\mu\text{g ml}^{-1}$)-resistant clones are selected upon plating on PSA medium and one isolate is chosen for further experiments. *Escherichia coli* is cultivated at 37 °C in lysogeny broth and *Agrobacterium tumefaciens* GV3101 at 30 °C in yeast extract broth supplemented with appropriate antibiotics.

Plant growth and inoculations. *Nicotiana benthamiana* plants are cultivated in the greenhouse with 16 h light, 40–60% humidity and day/night temperatures of 23 °C/19 °C. Plants inoculated with *Agrobacterium* strains are transferred to a Percival growth chamber (Percival Scientific) with 22 °C/18 °C day/night temperatures and 16 h light. Rice experiments are performed under greenhouse conditions with cycles of 12 h of light at 26 °C, 70% relative humidity and 12 h of dark at 25 °C, 70% relative humidity. *Oryza sativa* subsp. *japonica* cv. Azucena and *O. sativa* subsp. *indica* cv. Zhenshan 97 are used for virulence assays. Leaves of 3–4-week-old plants are infiltrated with a bacterial suspension at an optical density of 0.5 at 600 nm (OD_{600}) using a needle-less syringe⁵⁵. Pictures from disease symptoms (water-soaked lesions) are taken 4 days post inoculation.

Construction of repeat modules with aberrant repeats. Primer pairs encoding natural repeats of aberrant length are designed such that forward and reverse primer overlap in their 3' part (Supplementary Tables 1 and 3). Phusion polymerase is used to extend the primers and the resulting DNA fragments are subcloned into pUC57 via *SmaI* cut-ligation. The resulting plasmids are used as template to amplify the aberrant repeats with primers that add *Bpil* restriction sites matching the existing Golden TAL Technology kit to insert the aberrant repeats at position 2 or 3 into a hexa-repeat array⁴⁶.

Construction of artificial TALEs. TALEs are constructed using the Golden TAL Technology⁴⁶. Up to six individual repeats with selected RVDs are subcloned in an assembly vector. To construct AvrXa7-derivatives with 25.5 repeats, we extend the original Golden TAL kit⁴⁶ with two novel assembly vectors (Supplementary Note 1). The repeat backbone apart from the RVDs is identical (Supplementary Table 4). The repeats are pre-assembled in 2–5 assembly vectors and inserted together with the Hax3 N- and C-terminal regions⁵⁶. For expression *in planta*, N-terminal GFP-TALE fusion are assembled in a Golden Gate-compatible binary vector allowing expression of the constructs under control of the constitutive 35S promoter. For examples of DNA and amino-acid sequences see Supplementary Note 1. For expression in *Xanthomonas* spp., TALE-FLAG fusions are assembled in a Golden Gate-compatible broad host range vector¹⁹.

GUS reporter constructs and reporter assay. OsSWEET13 (*Xa25/xa25*, Os12N3 and Os12g29220) 1 kb promoter fragments (Supplementary Note 1), an Os-SWEET14 (*Os11N3* and *Os11g31190*) 341-bp promoter fragment, and artificial and natural TALE boxes (Supplementary Table 3) together with the minimal *Bs4* promoter, respectively, are inserted into pENTR/D-TOPO (Invitrogen). The promoter derivatives are recombined into pGWB3 (ref. 57) via LR recombination. Transient GUS reporter assays are performed⁷. Briefly, *Agrobacterium* strains delivering TALE constructs and GUS reporter constructs are mixed 1:1 and inoculated into leaves of 5–7-week-old *N. benthamiana* plants with a total OD_{600} of 0.8. The T-DNAs integrate into the plant chromosomes. Leaf disks (0.9 cm diameter) are sampled 2 days later and GUS activity is determined. For qualitative GUS assays, leaf disks are stained in X-Gluc (5-bromo-4-chloro-3-indolyl- β -D-glucuronide) solution, destained in ethanol and dried between acetate foil. For quantitative GUS assays two leaf disks are pooled, the plant tissue homogenized, diluted and incubated with 4-methyl-umbelliferyl- β -D-glucuronide (MUG). Proteins are quantified using Bradford assay (Roth). Values from three plants are combined into one data point. All experiments are done at least twice with similar results.

Construction of TALENs. TALENs are constructed from modules matching the Golden TAL Technology⁴⁶. The repeats are assembled in hexa-repeat modules and inserted together with modified short N- and C-terminal modules into a compatible ENTRY vector (pEGG). The N-terminal TALEN module contains amino acids 153–288 of Hax3, a SV40 nuclear localization sequence, and a tag (c-myc-tag for forward TALEN and FLAG-tag for reverse TALEN) sequence (see Supplementary Note 1). The C-terminal TALEN module contains amino acids 1–63 of the C-terminal region of Hax3 and a heterodimeric (DS for forward TALEN and RR for reverse TALEN) 'sharkey' FokI endonuclease domain⁵⁸. The TALEN are transferred via GATEWAY (Invitrogen) LR recombination into pDEST17 under control of a T7 promoter.

TALEN *in vitro* cleavage assay. TALENs are expressed using the TnT T7 Quick Coupled Transcription/Translation System (Promega) following the manufacturer's instructions. Five hundred nanograms of DNA from each TALEN construct are used. The target DNA fragment is generated by linearization of pENTR containing the respective target box upstream of the minimal *Bs4* promoter using the restriction enzyme *Alw44I* (Thermo Scientific) following the

manufacturer's instructions and subsequent purification with the GeneJET PCR Purification Kit (Thermo Scientific). For the *in vitro* cleavage assay 4 μl of TnT reaction containing the TALEN pair proteins is mixed with 200 ng of target DNA in 1 \times NEBuffer 3 (New England Biolabs) supplied with 2.5 $\mu\text{g ml}^{-1}$ BSA to a total volume of 20 μl . After incubation for 60 min at 37 °C the reaction is inactivated at 65 °C for 20 min and centrifuged at 16,000g for 3 min. The supernatant (16 μl) is analysed on a 1% agarose gel. All experiments are at least done twice with similar results.

Immunoblotting. For plant immunoblotting TALEs are transiently expressed in *N. benthamiana* for 2 days. *Agrobacterium* strains are inoculated with an OD_{600} of 0.4. Two leaf disks are pooled, the plant tissue homogenized, resuspended in 90 μl Lämmli buffer and incubated at 95 °C for 10 min. Debris are pelleted and SDS-PAGE is performed with 15 μl of each sample. The proteins are transferred to a PROTRAN nitrocellulose membrane (Whatman). To detect GFP-tagged proteins membranes are incubated with anti-GFP rabbit serum (Life Technologies; dilution 1:2,000). For detection by enhanced chemiluminescence (ECL) ECL anti-rabbit IgG (GE Healthcare; dilution 1:10,000) is used.

References

- DeFrancesco, L. Move over ZFNs. *Nat. Biotechnol.* **29**, 681–684 (2011).
- Doyle, E. L., Stoddard, B. L., Voytas, D. F. & Bogdanove, A. J. TALE effectors: highly adaptable phyto-bacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol.* **23**, 390–398 (2013).
- Gaj, T., Gersbach, C. A. & Barbas, 3rd C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
- Van den Ackerveken, G., Marois, E. & Bonas, U. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell* **87**, 1307–1316 (1996).
- Boch, J. & Bonas, U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–436 (2010).
- Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TALE effectors. *Science* **326**, 1501 (2009).
- Boch, J. *et al.* Breaking the code of DNA binding specificity of TALE-type III effectors. *Science* **326**, 1509–1512 (2009).
- 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).
- Deng, D. *et al.* Structural basis for sequence-specific recognition of DNA by TALE effectors. *Science* **335**, 720–723 (2012).
- Mak, A. N., Bradley, P., Cernadas, R. A., Bogdanove, A. J. & Stoddard, B. L. The crystal structure of TALE effector PthXo1 bound to its DNA target. *Science* **335**, 716–719 (2012).
- Zhu, W., Yang, B., Chittoor, J. M., Johnson, L. B. & White, F. F. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Mol. Plant Microbe Interact.* **11**, 824–832 (1998).
- Nino-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).
- Verdier, V., Vera Cruz, C. & Leach, J. E. Controlling rice bacterial blight in Africa: needs and prospects. *J. Biotechnol.* **159**, 320–328 (2012).
- Scholze, H. & Boch, J. TALE effectors are remote controls for gene activation. *Curr. Opin. Microbiol.* **14**, 47–53 (2011).
- 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).
- Li, T., Huang, S., Zhou, J. & Yang, B. Designer TALE effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. *Mol. Plant* **6**, 781–789 (2013).
- 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).
- Römer, P. *et al.* Promoter elements of rice susceptibility genes are bound and activated by specific TALE effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **187**, 1048–1057 (2010).
- Streubel, J. *et al.* Five phylogenetically close rice SWEET genes confer TALE effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **200**, 808–819 (2013).
- 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).
- Yu, Y. *et al.* Colonization of rice leaf blades by an african strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TALE effector that induces the rice Nodulin-3 *Os11N3* Gene. *Mol. Plant Microbe Interact.* **24**, 1102–1113 (2011).
- Doyle, E. L. *et al.* TALE Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TALE effector design and target prediction. *Nucleic Acids Res.* **40**, W117–W122 (2012).
- Grau, J. *et al.* Computational predictions provide insights into the biology of TALE effector target sites. *PLoS Comput. Biol.* **9**, e1002962 (2013).

24. 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).
25. Gu, K. *et al.* R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122 (2005).
26. Römer, P. *et al.* Plant-pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**, 645–648 (2007).
27. Strauß, T. *et al.* RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proc. Natl Acad. Sci. USA* **109**, 19480–19485 (2012).
28. Chu, Z. *et al.* Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. *Theor. Appl. Genet.* **112**, 455–461 (2006).
29. 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).
30. Yuan, T., Li, X., Xiao, J. & Wang, S. Characterization of *Xanthomonas oryzae*-responsive cis-acting element in the promoter of rice race-specific susceptibility gene *Xa13*. *Mol. Plant* **4**, 300–309 (2011).
31. Christian, M. *et al.* TAL effector nucleases create targeted DNA double-strand breaks. *Genetics* **186**, 757–761 (2010).
32. Morbitzer, R., Römer, P., Boch, J. & Lahaye, T. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc. Natl Acad. Sci. USA* **107**, 21617–21622 (2010).
33. Li, T. *et al.* TAL nucleases (TALENs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2010).
34. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
35. Byrum, S. D., Taverna, S. D. & Tackett, A. J. Purification of a specific native genomic locus for proteomic analysis. *Nucleic Acids Res.* **41**, e195 (2013).
36. Cong, L., Zhou, R., Kuo, Y. C., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3**, 968 (2012).
37. Garg, A., Lohmueller, J. J., Silver, P. A. & Armel, T. Z. Engineering synthetic TAL effectors with orthogonal target sites. *Nucleic Acids Res.* **40**, 7584–7595 (2012).
38. Mendenhall, E. M. *et al.* Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.* **31**, 1133–1136 (2013).
39. Miyazaki, Y., Ziegler-Birling, C. & Torres-Padilla, M. E. Live visualization of chromatin dynamics with fluorescent TALEs. *Nat. Struct. Mol. Biol.* **20**, 1321–1324 (2013).
40. Politz, M. C., Copeland, M. F. & Pfleger, B. F. Artificial repressors for controlling gene expression in bacteria. *Chem. Commun. (Camb)* **49**, 4325–4327 (2013).
41. Stella, S. *et al.* Structure of the AvrBs3-DNA complex provides new insights into the initial thymine-recognition mechanism. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1707–1716 (2013).
42. Streubel, J., Blücher, C., Landgraf, A. & Boch, J. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**, 593–595 (2012).
43. de Lange, O. *et al.* Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol.* **199**, 773–786 (2013).
44. Li, L. *et al.* Characterization and DNA-binding specificities of *Ralstonia* TAL-like effectors. *Mol. Plant* **6**, 1318–1330 (2013).
45. Yang, B., Zhu, W., Johnson, L. B. & White, F. F. The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proc. Natl Acad. Sci. USA* **97**, 9807–9812 (2000).
46. Geißler, R. *et al.* Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One* **6**, e19509 (2011).
47. Meckler, J. F. *et al.* Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res.* **41**, 4118–4128 (2013).
48. Wan, H., Hu, J. P., Li, K. S., Tian, X. H. & Chang, S. Molecular dynamics simulations of DNA-free and DNA-bound TAL effectors. *PLoS One* **8**, e76045 (2013).
49. Li, Y., Moore, R., Guinn, M. & Bleris, L. Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci. Rep.* **2**, 897 (2012).
50. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
51. Salomon, D., Dar, D., Sreeramulu, S. & Sessa, G. Expression of *Xanthomonas campestris* pv. *vesicatoria* type III effectors in yeast affects cell growth and viability. *Mol. Plant Microbe Interact.* **24**, 305–314 (2011).
52. Deslandes, L. & Rivas, S. Catch me if you can: bacterial effectors and plant targets. *Trends Plant Sci.* **17**, 644–655 (2012).
53. Gonzalez, C. *et al.* Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. *Mol. Plant Microbe Interact.* **20**, 534–546 (2007).
54. Figurski, D. & Helinski, D. R. Replication of an origin-containing derivative of plasmid RK2 is dependent on a plasmid function provided in trans. *Proc. Natl Acad. Sci. USA* **76**, 1648–1652 (1979).
55. Reimers, P. J. & Leach, J. E. Race-specific resistance to *Xanthomonas oryzae* pv. *oryzae* conferred by bacterial blight resistance gene *Xa-10* in rice (*Oryza sativa*) involves accumulation of a lignin-like substance in host tissues. *Physiol. Mol. Plant Pathol.* **38**, 39–55 (1991).
56. Kay, S., Boch, J. & Bonas, U. Characterization of AvrBs3-like effectors from a *Brassicaceae* pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Mol. Plant Microbe Interact.* **18**, 838–848 (2005).
57. Nakagawa, T. *et al.* Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* **104**, 34–41 (2007).
58. Guo, J., Gaj, T. & Barbas, 3rd C. F. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J. Mol. Biol.* **400**, 96–107 (2010).

Acknowledgements

We thank T. Schreiber, S. Thieme and U. Bonas for providing the GoldenGate-compatible vectors; K. Schlien for extending the Golden TAL kit, ArtXa70 and ArtXa71 constructs; T. Schreiber and U. Bonas for suggestions on the manuscript; R. Szczesny for providing FokI constructs; J. Leach for providing *avrXa7*; C. Kretschmer for technical assistance and U. Bonas for support. This work was supported by grants from the Deutsche Forschungsgemeinschaft (SPP 1212, BO 1496/6-1, BO 1496/7-1), and from the European Regional Development Fund of the European Commission. B.S. is supported by a grant from GRiSP for New Frontiers Research projects (MENERGEP).

Author contributions

A.R., J.S. and J.B. conceived and designed the experiments. A.R., J.S. and C.B. performed the experiments. A.R., J.S., C.B. and J.B. analysed the data. M.R. established the TALEN assays. J.G. performed the bioinformatic analyses. B.S. provided material and performed additional experiments. A.R., J.S. and J.B. wrote the manuscript with support from all authors.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: J.B. is a part owner of the patent US 8420782 B2 'Modular DNA-Binding Domains And Methods Of Use' concerning the use of TALEs. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Richter, A. *et al.* A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* 5:3447 doi: 10.1038/ncomms4447 (2014).

2.3.1.1 Anlagen zu Publikation 3

Supplementary Information

A TAL effector repeat architecture for frame shift binding

Annekatriin Richter¹†, Jana Streubel¹†, Christina Blücher¹, Boris Szurek², Maik Reschke¹, Jan
Grau³ and Jens Boch^{1*}

¹Department of Genetics, Martin Luther University Halle-Wittenberg, Weinbergweg 10, D-
06120 Halle (Saale), Germany.

²UMR 186 IRD-UM2-Cirad "Résistance des Plantes aux Bioagresseurs", BP 64501, 34394
Montpellier cedex 5, France.

³Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale),
Germany.

*To whom correspondence should be addressed.

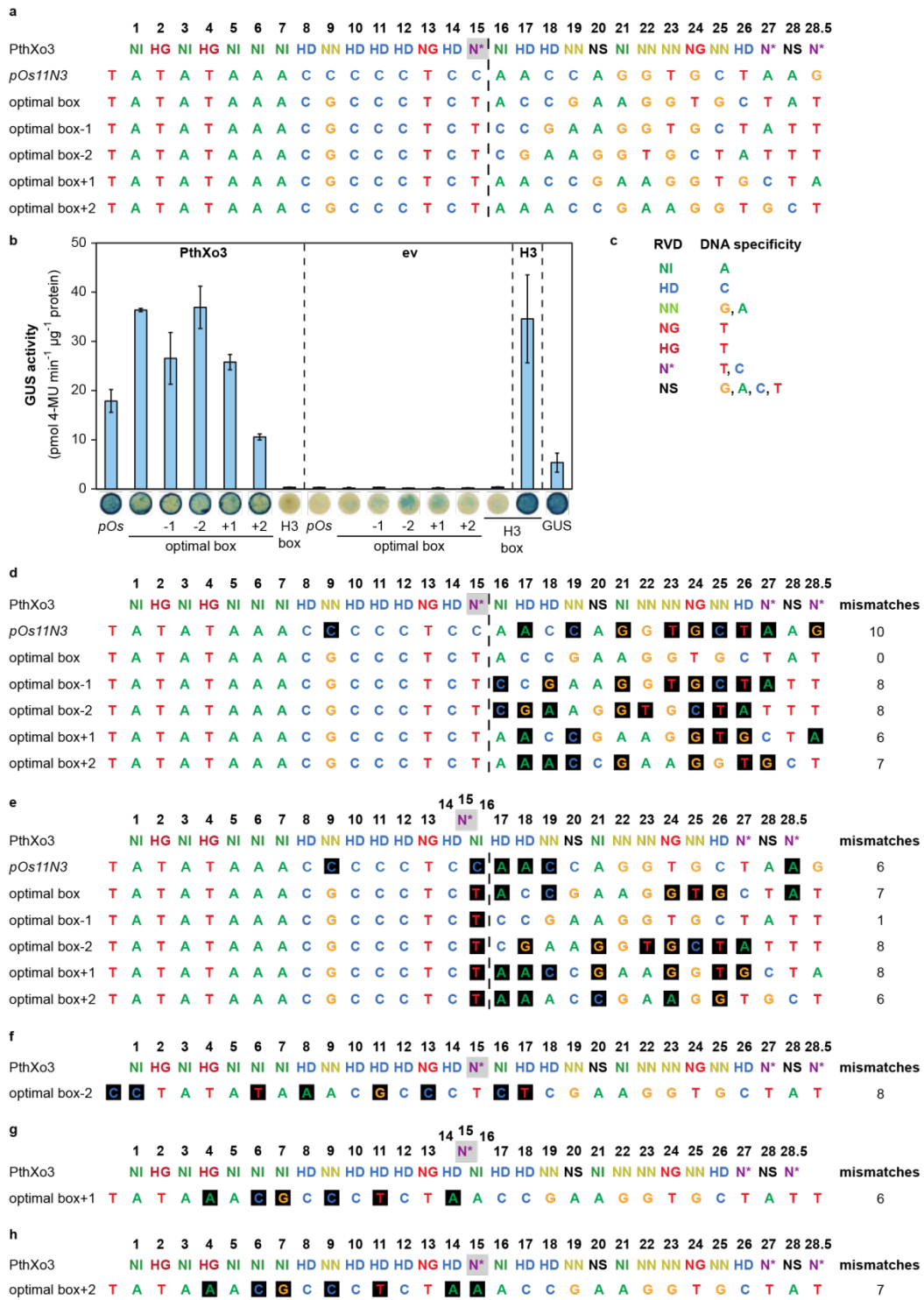
E-mail: jens.boch@genetik.uni-halle.de.

†These authors contributed equally to the manuscript.

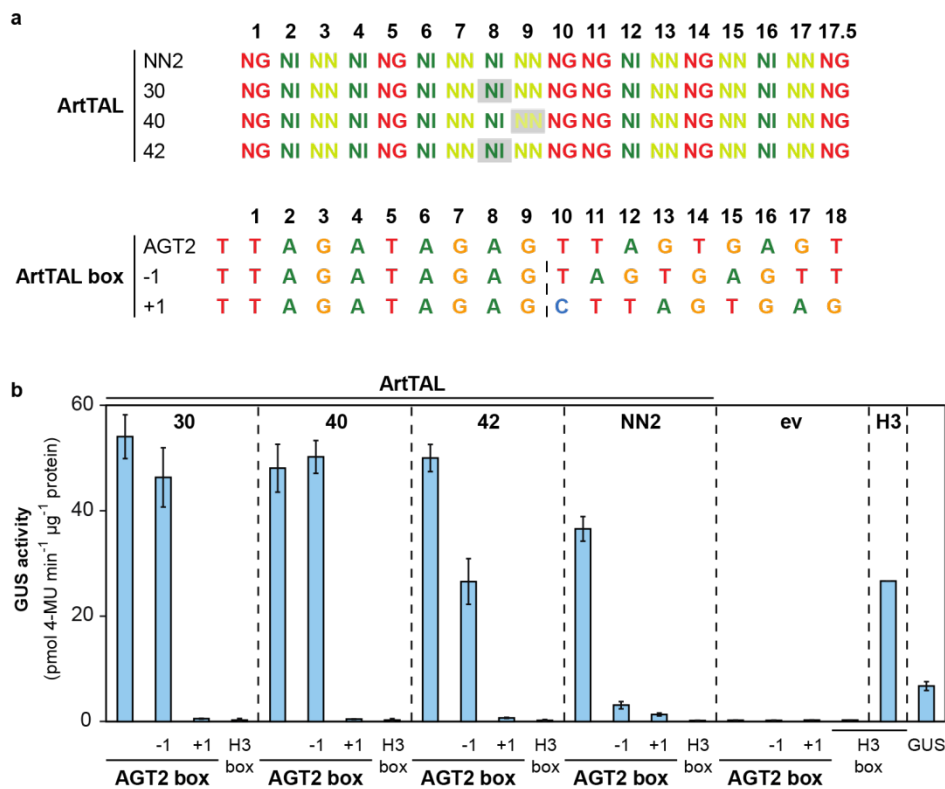
This file includes:

Supplementary Figures 1-9
Supplementary Tables 1-4
Supplementary Note 1
Supplementary References

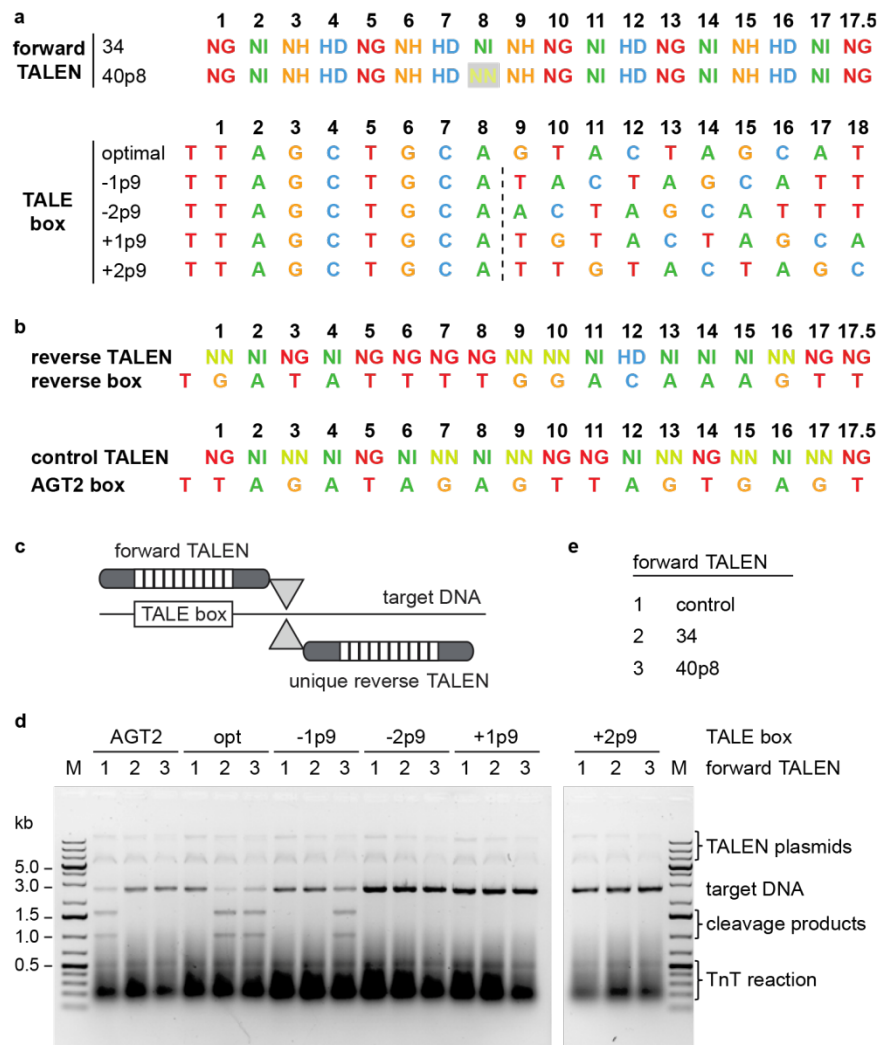
boxes containing one or two base pair deletions (optimal box-1, optimal box-2) or insertions (optimal box+1, optimal box+2) at position 14 (to the right of the dashed line) are fused to the minimal *Bs4* promoter and a promoterless GUS reporter gene. **(b)** GUS assay of AvrXa7 with target boxes. GUS reporter constructs are codelivered by *A. tumefaciens* into *N. benthamiana* leaf cells together with constitutive 35S-driven *avrXa7*, and *GFP* (*ev*) constructs, respectively ($n=3$, error bars indicate s.d.). 35S::*uidA* (GUS) and the natural TALE Hax3 (H3) with its perfect Hax3 box (H3 box) serve as controls. For qualitative assays, leaf disks are stained with X-Gluc, and a blue color indicates GUS activity. **(c)** RVD specificities. **(d-g)** Mismatch analysis of AvrXa7 on different target boxes. Non-matching RVD-DNA base combinations are boxed in black. **(d, e)** Aberrant repeat of AvrXa7 in regular base pair alignment. **(f, g)** Aberrant repeat of AvrXa7 in looped-out alignment. **(d, f)** AvrXa7 with natural box from the rice *Os11N3* promoter aligned either to the leftmost possible sequence described by Römer *et al.*¹ (bottom sequence) or to the second possible position described by Antony *et al.*² (top sequence). **(e, g)** AvrXa7 aligned to optimal boxes.



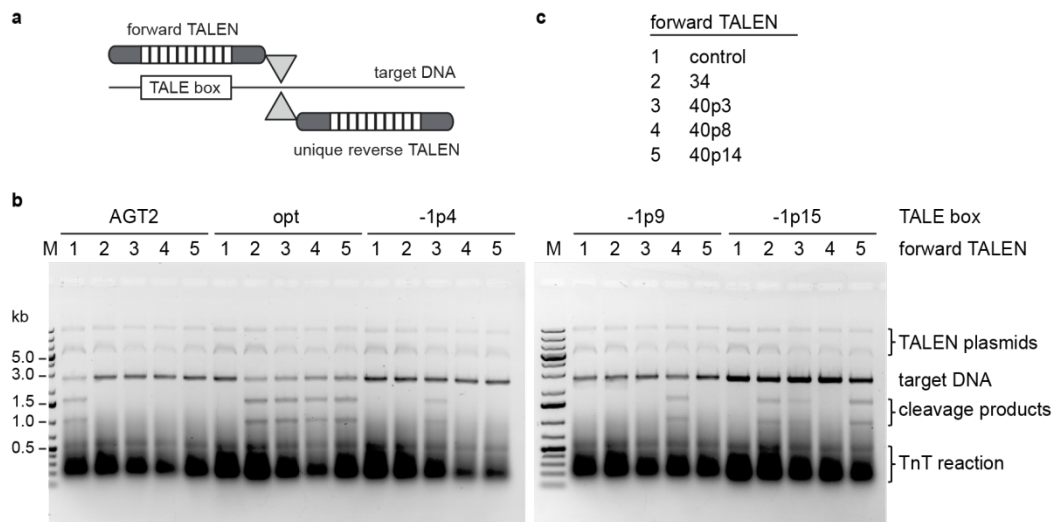
Supplementary Figure 2 | PthXo3 recognizes a variety of target DNA boxes. (a) PthXo3 RVDs and target boxes. The natural TALE PthXo3 carries an aberrant N*-repeat of 39 amino acids at position 15 (grey square). The natural PthXo3 target box in the rice *Os11N3* (*Os11g31190*, *OsSWEET14*) promoter² (*pOs11N3*, *pOs*) or an optimal box, and optimal boxes containing one or two base pair deletions (optimal box-1, optimal box-2) or insertions (optimal box+1, optimal box+2) at position 16 (to the right of the dashed line), upstream of the minimal *Bs4* promoter are fused to a promoterless GUS reporter gene. (b) GUS assay of PthXo3 with target boxes. GUS reporter constructs are codelivered by *A. tumefaciens* into *N. benthamiana* leaf cells together with constitutive *35S*-driven *pthXo3*, and *GFP* (ev) constructs, respectively ($n=3$, error bars indicate s.d.). *35S::uidA* (GUS) and the natural TALE Hax3 (H3) with its perfect Hax3 box (H3 box) serve as controls. For qualitative assays, leaf disks are stained with X-Gluc, and a blue color indicates GUS activity. (c) RVD specificities. (d-h) Mismatch analyses of PthXo3 on different target boxes. Non-matching RVD-DNA base combinations are boxed in black. (d) Aberrant repeat of PthXo3 in regular base pair alignment. (e) Aberrant repeat of PthXo3 in looped-out alignment. (f-h) Alternative alignments of PthXo3 to (f) the optimal box-2, (g) the optimal box+1 and (h) the optimal box+2.



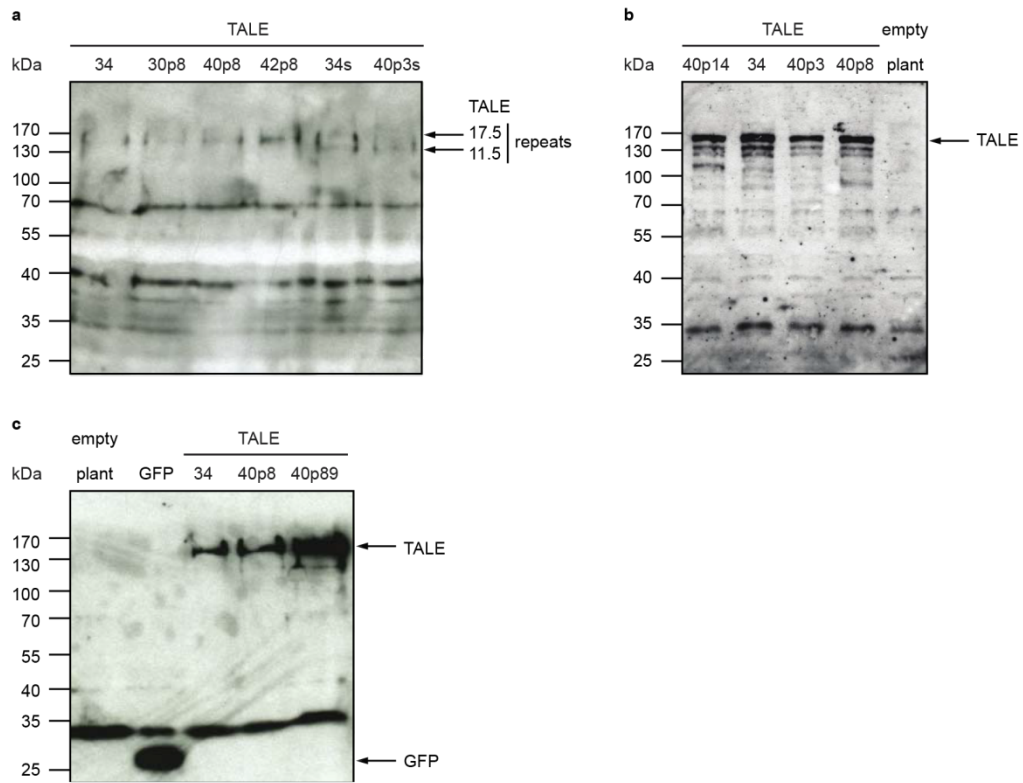
Supplementary Figure 3 | Aberrant repeats enable frame shift binding with an alternative TALE repeat array. (a) ArtTAL RVDs and target boxes. ArtTALs with all 34-amino acid (aa) repeats (NN2) according to Streubel *et al.*³ or ArtTALs with an aberrant 30-aa NI-repeat, 40-aa NN-repeat, and 42-aa NI-repeat, respectively, at position 8 or 9 are constructed. The boxes are either perfectly matching the specificity of normal repeats (AGT2) or have one base pair deleted (AGT2-1) or inserted (AGT2+1) at position 10 (to the right of the dashed line). Target ArtTAL-boxes are fused to a minimal promoter and a promoterless *uidA* (GUS) reporter gene. (b) GUS assays of ArtTALs and reporter constructs. GUS reporter constructs are codelivered by *A. tumefaciens* into *N. benthamiana* leaf cells together with constitutive *35S*-driven *artTAL* genes, and *GFP* (ev), respectively ($n=3$, error bars indicate s.d.). *35S::uidA* (GUS) and the natural TALE Hax3 (H3) with its perfect Hax3-box (H3-box) serve as controls.



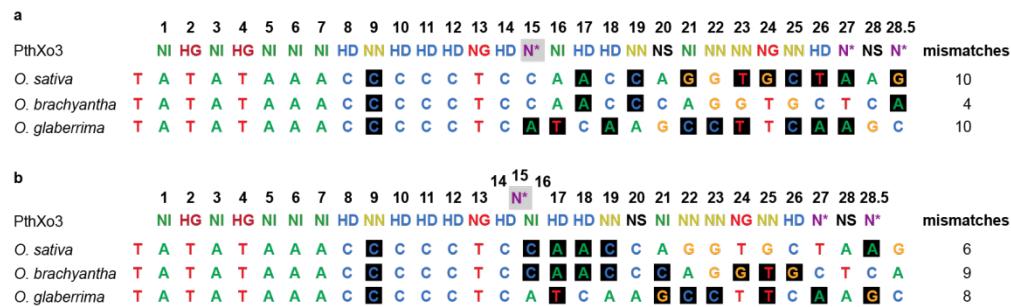
Supplementary Figure 4 | Aberrant repeats allow a -1 nucleotide frame shift tolerance in TALENs. (a) RVDs of forward TALENs and target boxes. TALENs are constructed with all 34-amino acid repeats or an aberrant 40-amino acid NN-repeat inserted at position 8 (40p8). The NN-repeat recognizes both, G and A DNA bases. The boxes are either perfectly matching the specificity of normal repeats (optimal) or have one or two base pairs either deleted (-1p9, -2p9) or inserted (+1p9, +2p9) at position 9 (to the right of the dashed line). The TALEN repeat array is designed such that for normal repeats mainly non-matching repeat-DNA base combinations occur following the position of the indels in the DNA sequence. (b) RVDs of reverse and control TALEN. The control TALEN recognizes a different target box (AGT2). (c) Cartoon of TALENs bound to DNA. The TALEN pairs are placed such that the FokI domain (grey triangle) dimerizes and cuts the DNA. The reverse TALEN is constant in all assays. (d) *In vitro* TALEN restriction assay. Target DNA is incubated with *in vitro* transcribed and translated TALEN pairs. The cleavage products are documented on agarose gels. Additional signals of low molecular weight and TALEN plasmids originate from the *in vitro* transcription and translation (TnT) reaction. GeneRuler 1 kb DNA Ladder (Thermo Scientific) is used as molecular weight marker (M). (e) Legend of forward TALENs.



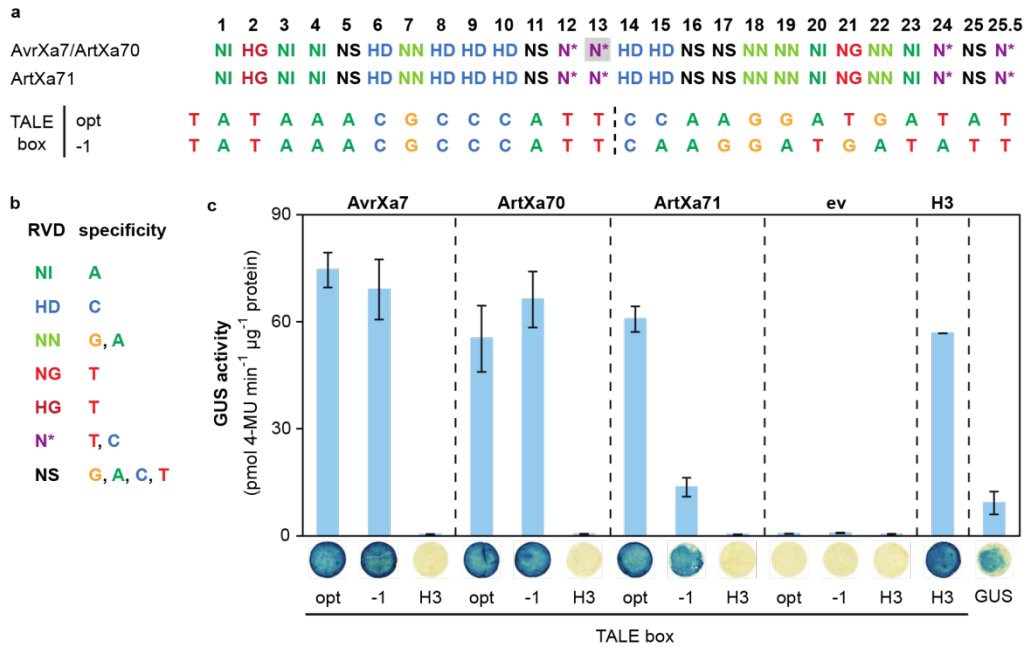
Supplementary Figure 5 | Full gel images of the *in vitro* TALEN restriction assay in Figure 3c. (a) Cartoon of TALENs bound to DNA. The TALEN pairs are placed such that the FokI domain (grey triangle) dimerizes and cuts the DNA. The reverse TALEN is constant in all assays. **(b)** *In vitro* TALEN restriction assay. Target DNA is incubated with *in vitro* transcribed and translated TALEN pairs. The cleavage products are documented on agarose gels. Additional low molecular weight signals and TALEN plasmids originate from the *in vitro* transcription and translation (TnT) reaction. GeneRuler 1 kb DNA Ladder (Thermo Scientific) is used as molecular weight marker (M). **(c)** Legend of forward TALENs. The control TALEN recognizes a different target box (AGT2).



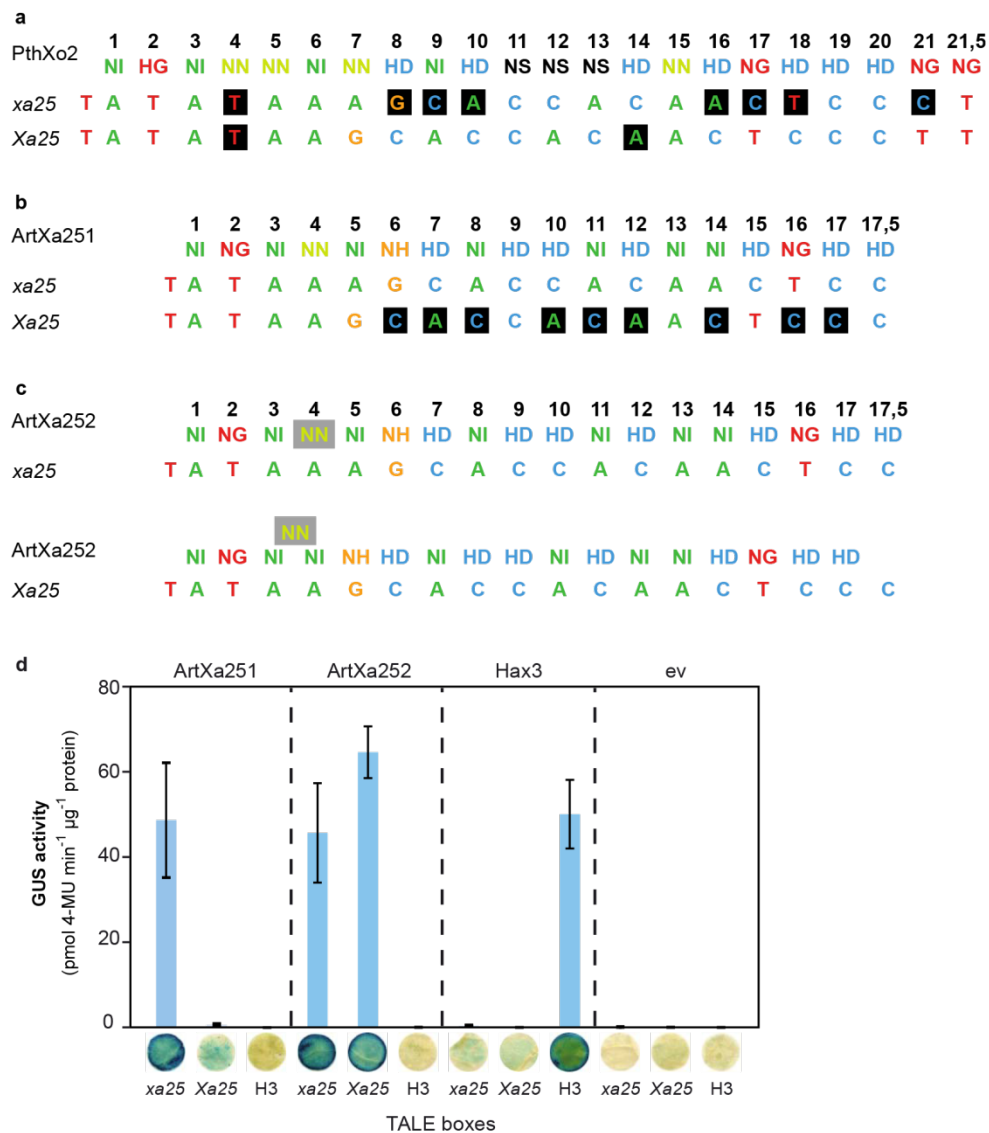
Supplementary Figure 6 | Immunoblotting of GFP-TALEs produced *in planta*. Detection of the TALE proteins used in GUS-assays of (a) Figure 2, 4 and 6, (b) Figure 3 and 4c-d, (c) Figure 5, respectively, using an anti-GFP antiserum.



Supplementary Figure 7 | PthXo3 RVDs aligned to predicted target boxes in the *OsSWEET14* promoters of *O. sativa*, *O. brachyantha* and *O. glaberrima*. (a) Aberrant repeat of PthXo3 in regular base pair alignment. (b) Aberrant repeat of PthXo3 in looped-out alignment.



Supplementary Figure 8 | The aberrant repeat contributes to frame shift box recognition of artificial AvrXa7-derivatives. (a) AvrXa7, ArtXa70 and ArtXa71 RVDs and target boxes. AvrXa7 derivatives are assembled with repeats resembling either the natural AvrXa7 repeats including the aberrant 39-aa N*-repeat at position 13 (grey square, ArtXa70) or all 34 aa repeats (ArtXa71). The optimal box (opt) and one containing an one base pair deletion (-1) at position 14 (to the right of the dashed line) are fused to the minimal *Bs4* promoter and a promoterless GUS reporter gene. (b) RVD specificities. (c) GUS assay. GUS reporter constructs are codelivered by *A. tumefaciens* into *N. benthamiana* leaf cells together with constitutive *35S*-driven *avrXa7*, *artXa70*, *artXa71*, and *GFP* (ev), respectively ($n=3$, error bars indicate s.d.). *35S::uidA* (GUS) and the natural TALE Hax3 (H3) with its Hax3 box (H3 box) serve as controls. For qualitative assays, leaf disks are stained with X-Gluc. A blue color indicates GUS activity.



Supplementary Figure 9 | An artificial TALE with aberrant repeat that targets the PthXo2 box recognizes different allelic *OsSWEET13* (*Xa25*) promoters. (a) PthXo2 RVDs aligned to *xa25* and *Xa25* promoter boxes of *O. sativa* cv. Nipponbare and *O. sativa* cv. Zhenshan 97⁴. (b) RVDs of the artificial TALE ArtXa251 aligned to *xa25* and *Xa25*. (c) RVDs of the artificial TALE ArtXa252 containing an aberrant 40-aa NN-repeat at position 4 (grey square) aligned to *xa25* and *Xa25* in normal and looped out conformation. (d) GUS assays ($n=3$) of artificial TALEs with *OsSWEET13* promoter sequences derived from *Oryza sativa* cv. Nipponbare (*xa25*) and Zhenshan 97 (*Xa25*) or the Hax3 box (H3) fused to a minimal promoter and a promoterless GUS-reporter gene. 35S-driven expression of the natural TALE Hax3 or GFP as empty vector (ev) serve as controls in quantitative and qualitative assays. Error bars indicate the s.d. in the quantitative assay. One representative leaf disk of the qualitative assay is shown.

Supplementary Tables

Supplementary Table 1 | Aberrant repeat sequences.

TALE	Strain ^a	Feature ^b	Sequence ^c	Accession
30 amino acid-TALE repeats				
XOCORF_4248	<i>Xoo</i> BLS256	RVDs	NI-NN-HN-NN-NI-HG-HD-NN-HD-HD-HG-HG-HD-HD-NG (15.5 repeats)	cmr.jcvi.org
		AA	LTPDQVVAIASNIGGKQALETVQRLLPVLC	
		DNA	CTGATCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTCTGTGC	
39 amino acid-TALE repeats				
PthXo3	<i>Xoo</i> PXO61	RVDs	NI-HG-NI-HG-NI-NI-NI-HD-NN-HD-HD-HD-NG-HD- N* -NI-HD-HD-NN-NS-NI-NN-NN-MG-NN-HD-N*-NS-N* (29.5 repeats)	AAS46027
		AA	LTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG	
AvrXa7	<i>Xoo</i> PXO86	RVDs	NI-HG-NI-NI-NS-HD-NN-HD-HD-NG-NS- N* -HD-HD-NS-NS-NN-NN-NI-NG-NN-NI-N*-NS-N* (25.5 repeats)	AAG02079
		AA	LTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG	
		DNA	CTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTACAGCGGCTGTTGCCGGTGTCTGTGC	
40 amino acid-TALE repeats				
XOO2866_MAFF	<i>Xoo</i> MAFF 311018	RVDs	NI-HG-NI-NI-HG-HD-NN-HD-HD-HD-NG-NI-NI- NI -NI-HD-HD-HD-HG-NN-NN-HD-NS-NN-HD-NG-NS-N* (26.5 repeats)	YP_451895
		AA	LTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG	
XOO3014_KACC	<i>Xoo</i> KACC 10331	RVDs	NI-HG-NI-NI-NS-HD-NN-HD-HG-NI-NI-HG-HD-NN-HD-HD-ND-ND- NI -NI-HD-HD-HD-HG-NN-NN-HD-NN- HD-N*-NS-N* (33.5 repeats)	YP_201653
		AA	LTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG	
		DNA	CTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATACCGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTACAGCGGCTGGTGGTGTCTGTGC	
42 amino acid-TALE repeats				
Tal9b	<i>Xoo</i> PXO99 ^A	RVDs	HD-HD-NN-NN-NG-NG-HD-NS-HG-HD-NG-N* -NI -HD-HD- NI -NN-HD-HI-ND-HD-HG-NN-HG-N* (26.5 repeats)	YP_00191509
		AA	LTPDQVVAIASNQVVAIASNIGGKQALETVQRLLPVLCQDHG	
XOO1136_MAFF	<i>Xoo</i> MAFF 311018	RVDs	HD-HD-NN-NN-NS-NG-HD-S* H -HG-HD-NG-N* -NI -HD-HD- NI -NN-HD-HI-ND-HD-HG-NN-HG-N* (26.5 repeats)	YP_450165
		AA	LTPDQVVAIASNQVVAIASNIGGKQALETVQRLLPVLCQDHG	
		DNA	CTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATACCGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTACAGCGGCTGGTGGTGTCTGTGC	

^a *Xoo*, *Xanthomonas oryzae* pv. *oryzicola*; *Xoo*, *Xanthomonas oryzae* pv. *oryzae*.
^b RVDs, repeat variable diresidues; AAs, amino acids; DNA, DNA sequence.
^c RVDs and AAs are in one letter amino acid code; RVDs of aberrant repeats and their coding bases are in red.

13

Supplementary Table 2 | Prediction of PthXo3 target sequences in rice promoters^a.

TALE ^b	Rank ^c	Annotation	Start	ATC ^d	Sequence	Matches ^e	Score	p-value ^f
PthXo3	1	Os01g63580.1: glycerol-3-phosphate acyltransferase, putative, expressed	268	90	TATATTAACCTCCCTCCCGCTGTCCACT	M x x x :xx x x x:	-27,37985934	4,42548E-08
PthXo3	2	Os01g17170.1: magnesium-protoporphyrin IX monomethyl ester cyclase, chloroplast precursor, putative, expressed	323	44	TATATACCGCCCAACCTCCCTCCCC	M xx : x xxxx x x:	-27,38871869	4,42548E-08
PthXo3	3	Os01g17170.2: magnesium-protoporphyrin IX monomethyl ester cyclase, chloroplast precursor, putative, expressed	323	44	TATATACCGCCCAACCTCCCTCCCC	M xx : x xxxx x x:	-27,38871869	4,42548E-08
PthXo3	4	Os01g69160.1: regulatory protein, putative, expressed	197	73	CATAAAAAAGCCCTATAAAAAAATCCAG	m x : x :xxx xx x :x	-27,64408828	4,42548E-08
PthXo3	5	Os03g08880.1: purine permease, putative, expressed	328	52	CATCAACACACCTAATCCAAAGGTGC	m xx x x xxx x : : : : : : : :	-27,85167952	8,85096E-08
PthXo3	6	Os02g49350.1: plastocyanin-like domain containing protein, putative, expressed	269	82	TATATAACTCCCGCGGACAGACACCGC	M x x xx xx :xxx x:	-27,95934861	8,85096E-08
PthXo3	7	Os02g57120.1: HEAT repeat family protein, putative, expressed	268	96	TATATAAACTCCCTCCCAAAAGCCCACT	M xx x :xxx :xxx xx:	-27,98858864	8,85096E-08
PthXo3	8	Os02g03280.2: transmembrane BAX inhibitor motif-containing protein, putative, expressed	269	150	TATATAAAACCCCCCCCCCGCTCCCTCT	M x x x :xxxx x x:	-28,06685038	1,32764E-07
PthXo3	9	Os01g19330.1: MYB family transcription factor, putative, expressed	265	98	TATATAAATCCCGCCCAAGAGCTACCC	M xx x :xxx x x x x x:	-28,23286173	1,32764E-07
PthXo3	10	Os05g45430.1: TOO MANY MOUTHS precursor, putative, expressed	264	65	CACATAAACTCCCCCAACCCATCTCTCT	m x x x xx xx x x:	-28,24252784	1,32764E-07
PthXo3	118	Os03g57240.1: ZOS3-19 - C2H2 zinc finger protein, expressed	273	140	TATAGAAACAACTCCTGATTTCCCC	M x x x xx x x x x x:	-31,90820162	3,58464E-06
PthXo3	119	Os11g31190.1: nodulin MeN3 family protein, putative, expressed	267	203	TATATAAAACCCCTCCAACCCAGGTCTAAG	M x x x x:xxxx x	-31,91182322	3,58464E-06
PthXo3	120	Os01g42820.1: RNA recognition motif containing protein, putative, expressed	264	121	CCGATAAAACCCCTCGACCGATCCCACCC	mxxx xx x :xxxx x x:	-31,9561958	3,76166E-06
PthXo3_short	1	Os11g31190.1: nodulin MeN3 family protein, putative, expressed	267	204	TATATAAAACCCCTCCAACCCAGGTCTA	M x x xxxx : : : : : :	-21,3989497	0
PthXo3_short	2	Os02g57120.1: HEAT repeat family protein, putative, expressed	268	97	TATATAAACTCCCTCCCAAAAGCCCACT	M xx x x x:xx : :	-21,69820182	0
PthXo3_short	3	Os01g46570.1: CTP synthase,	269	189	TATTTAAAGCCCTCACTCGAGTGTCTCC	M x x xxxx : : : : : :	-23,44192339	0

14

		putative, expressed						
PthXo3_short	4	Os01g46570.2: CTP synthase, putative, expressed	269	189	TATTTAAAAGCCCTCACCTCGAGTGCTCC	M x x: xxxx : x:	-23,44192339	0
PthXo3_short	5	Os03g09150.1: pumilio-family RNA binding repeat domain containing protein, expressed	266	205	TATATAAACCCCCCGCGCTAGTGCGGC	M x x p :xxx : xx:	-23,92888744	0
PthXo3_short	6	Os02g40010.1: phosphoribosyl transferase, putative, expressed	267	85	TATATATACCTCCTCACCCACAGGCCAC	M x xx xx xxx : :	-24,84380499	0
PthXo3_short	7	Os02g40010.2: phosphoribosyl transferase, putative, expressed	267	85	TATATATACCTCCTCACCCACAGGCCAC	M x xx xx xxx : :	-24,84380499	0
PthXo3_short	8	Os02g40010.3: phosphoribosyl transferase, putative, expressed	267	85	TATATATACCTCCTCACCCACAGGCCAC	M x xx xx xxx : :	-24,84380499	0
PthXo3_short	9	Os02g40010.4: phosphoribosyl transferase, putative, expressed	267	85	TATATATACCTCCTCACCCACAGGCCAC	M x xx xx xxx : :	-24,84380499	0
PthXo3_short	10	Os09g13460.1: expressed protein	51	220	TATATAATCCCTCCCTCCCGAGGTGATTC	M x x x x : : xx x:	-24,93731372	0

^a Predictions were done by applying TALgetter⁵ to *Oryza sativa* Nipponbare promoters -300 to +200 of the transcriptional start site.

^b PthXo3, bound in normal mode; PthXo3_short, bound with aberrant repeat looped out.

^c Ranks 1 to 10 for both binding modes and ranks 118-120 for the normal mode are shown.

^d Distance of TALE box to ATG.

^e Categories of matches per position: M/m, first position match/mismatch; |, match; :, weak match; x, mismatch

^f Empirical p-values were computed by generating a random data base of at least the total length of all scanned input sequences from a homogeneous Markov model of order 2 trained on the input sequences. For these random data the likelihoods of sub-sequences were computed in the same manner as for the actual input sequences. The p-value of a putative target site was calculated by determining the percentile corresponding to the observed likelihood value on the distribution of likelihoods obtained for the random data base⁵.

Supplementary Table 3 | Selected primers used in this study.

Name	Nucleotide sequence ^a
Cloning of aberrant repeats	
30-F	CTGATCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
30-R	GCACAGCACCGGCAACAGCCGCTGCACCGTCTCCAGCGCCTGCTTGCCGCG
39-F	CTGACCCCGGACCAGGTGCTGGCCATCGCCAGCAATGGCGGCAAGCAGGCG
39-R	GCCATGGTCTTGGCACAGCACCGGCAACAGCCGCTGTACCGGCAACAGCCGCTGCACCGTCTCCAGCGCCTGCTTGCCGCG
40-F	CTGACCCCGGACCAGGTGGTGGCCATCGCCAAACAATAACGGCGGCAAGCAGGCG
40-R	GCCATGGTCTTGGCACAGCACCGGCAACAGCCGCTGTACCGGCAACAGCCGCTGCACCGTCTCCAGCGCCTGCTTGCCGCG
42-F	CTGACCCCGGACCAGGTGGTGGCCATCGCCAGTAATCAAGTCGTCGCGATTGCCAGTAATATTGGCGGCAAGCAGGCG
42-R	GCCATGGTCTTGGCACAGCACCGGCAACAGCCGCTGCACCGTCTCCAGCGCCTGCTTGCCGCG
Cloning of aberrant repeats as Golden TAL Technology modules	
Position 2	
2-F_30	TTTGAAGACTTCTGATCCCGGACCAGGTGG
2-R_30	TTTGAAGACTTCGGTATCAGGCACAGCACCGGCAACAGCCGC
2-F_39-42	TTTGAAGACTTCTGACCCCGGACCAGGTGGTG
2-R_39-42	TTTGAAGACTTCGGTGTACAGGCCATGGTCTTGGCACAGCAC
Position 3	
3-F_30-42	TTTGAAGACTTACCGGACCAGGTGGTGGCCATCGCC
3-R_30	TTTGAAGACTTGGTGAGGCACAGCACCGGCAACAGCCGC
3-R_39-42	TTTGAAGACTTGGTGAGGCCATGGTCTTGGCACAGCAC
Cloning of TALE boxes in reporter vector	
Forward primer^b	
AvrXa7 nat_Box	CACCTATATAAACCCCTCCAACCAGGTGCTAATTCCTTTCTGTATATAACTTTGTCC
AvrXa7 opt_Box	CACCTATAAACGCCCATTTCCAAGGATGATATTTCTTTCTGTATATAACTTTGTCC
AvrXa7 oBox-1	CACCTATAAACGCCCATTTCAAGGATGATATTTCTTTCTGTATATAACTTTGTCC
AvrXa7 oBox-2	CACCTATAAACGCCCATTAAGGATGATATTTCTTTCTGTATATAACTTTGTCC
AvrXa7 oBox+1	CACCTATAAACGCCCATTTACCAAGGATGATATTTCTTTCTGTATATAACTTTGTCC
AvrXa7 oBox+2	CACCTATAAACGCCCATTAACCAAGGATGATATTTCTTTCTGTATATAACTTTGTCC
AGT2-box-1_40	CACCTTAGATAGAGTAGTGAGTTTCTTTCTGTATATAACTTTGTCC
AGT2-box+1_40	CACCTTAGATAGAGCTTAGTGAGTTTCTTTCTGTATATAACTTTGTCC
sTAL-box	CACCTTAGCTGCAGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box-1	CACCTTAGCTGCATACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box-2	CACCTTAGCTGCACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box+1	CACCTTAGCTGCATGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box40pos3	CACCTTAGCTGCAGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box40pos14	CACCTTAGCTGCAGTACTACATTTCTTTCTGTATATAACTTTGTCC
sTAL-box-1p6F	CACCTTAGCTGAGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
sTAL-box-1p7F	CACCTTAGCTGCGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
-1p2/C/F	CACCTTA CTGCAGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
1mm p8 F	CACCTTAGCTGCTGTACTAGCATTTCTTTCTGTATATAACTTTGTCC
2mm p8 F	CACCTTAGCTGCTCTACTAGCATTTCTTTCTGTATATAACTTTGTCC
3mm p8 F	CACCTTAGCTGCTCACTAGCATTTCTTTCTGTATATAACTTTGTCC
1mm F	CACCTTAGCTGCAGTACTAGCATTTCTTTCTGTATATAACTTTGTCC

2mm F	CACCTTAGCTGCACAAC TAGCAT TTCTTTCTTGTATATAACTTTGTCC
sTAL- box3mm	CACCTTAGCTGCACATCTAGCATTTCTTTCTTGTATATAACTTTGTCC
PthXo3box	CACCTATATAAACGCCCTCTACCGAAGGTGCTATTTCTTTCTTGTATATAACTTTGTCC
PthXo3box- 1	CACCTATATAAACGCCCTCTCCGAAGGTGCTATTTCTTTCTTGTATATAACTTTGTCC
PthXo3box- 2	CACCTATATAAACGCCCTCTCGAAGGTGCTATTTCTTTCTTGTATATAACTTTGTCC
PthXo3box+ 1	CACCTATATAAACGCCCTCTAACCGAAGGTGCTATTTCTTTCTTGTATATAACTTTGTCC
PthXo3box+ 2	CACCTATATAAACGCCCTCTAACCGAAGGTGCTATTTCTTTCTTGTATATAACTTTGTCC
Reverse primer	
Bs4PR/B01_ R	AGATTCGATTAATAATAAATTGTATGGATGAGATC
Cloning of <i>OsSWEET13</i> (<i>OsI2N3</i>, <i>OsI2g29220</i>) promoters in reporter vector	
Forward primer	
Xa25Zh/Nip T65 1kb /F	CACCAGGGATGTCTACTGCAGGTGAAAAC
Reverse primer	
Xa25Zh/Nip T65 1kb /R	TTTTGTGTGCTAAAAGGGGGTAATTG

^a *Bpi*I restriction sites are shown in bold.

^b Sequences bound by TALEs are in red. Forward primer start with CACC to enable cloning into pENTR/D-TOPO (Invitrogen). 3' black letters indicate sequences complementary to the *Bs4* promoter

Supplementary Table 4 | TALE repeat amino acid sequences^a

RVD	Repeat sequence
NI	LTPEQVVAIASNI ^g GGKQALETVQRLLPVLCQAHG
HD	LTPEQVVAIASHD ^g GGKQALETVQRLLPVLCQAHG
NN	LTPEQVVAIASNN ^g GGKQALETVQRLLPVLCQAHG
NG	LTPEQVVAIASNG ^g GGKQALETVQRLLPVLCQAHG
HG	LTPEQVVAIASHG ^g GGKQALETVQRLLPVLCQAHG
NS	LTPEQVVAIASNS ^g GGKQALETVQRLLPVLCQAHG
NH	LTPEQVVAIASNH ^g GGKQALETVQRLLPVLCQAHG
N*	LTPEQVVAIASN* ^g GGKQALETVQRLLPVLCQAHG

^a amino acids in one letter code. A star (*) indicates a missing amino acid in the alignment. RVDs are in green.

Supplementary Note 1 | DNA and amino acid sequences

GFP - TALE34

DNA sequence^a

ATGGTGAGCAAGGGCGAGGAGCTGTTCAACGGGGTGGTGCCATCCTGGTTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG
TTCATCTGCACCACCGGAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTG
CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGC
TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTC
GAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC
ATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCATCGGGCAGCGCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGGATC
ACTCTCGGCATGGACGAGCTGTACAAGCATATGGATCCCATTCGTTTCGCGCACACCAAGTCTGCCC
GAGCTTCTGTCCGGACCCCAACCCGATGGGGTTCAGCCGACTGCAGATCGTGGGGTGTCTCCGCCTGCC
GGCGGCCCCCTGGATGGCTTGCCCGCTCGGCGGACGATGTCCCGACCCGGCTGCCATCTCCCCCTGCC
CCCTCACCTGCGTTCGCGCGACAGCTTCAGTGACCTGTTACGTTCAGTTCGATCCGTCACTTTTTAAT
ACATCGTTTTTTGATTATTGCCTCCCTTCGGCGCTCACCATACAGAGGCTGCCACAGGCGAGTGGGAT
GAGGTGCAATCGGGTCTGCGGGCAGCCGACGCCCCCACCACCATGCGCGTGGCTGTCACTGCCGCG
CGGCGCGCGCGCAAGCCGGCGCGCGACGACGTGCTGCGCAACCTCCGACGCTTCGCGGCGCGG
CAGGTGGATCTACGCACGCTCGGCTACAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTCTGTCG
ACAGTGGCGCAGCACACGAGGCACTGGTTCGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGC
CAACACCCGGCAGCGTTAGGGACCGTTCGCTGTCAAGTATCAGGACATGATCGCAGCGTTGCCAGAGGCG
ACACACGAAGCGATCGTTGGCGTCGGCAACAGTGGTCCGGCGCACGCGCTCTGGAGGCTTGTCCAG
GTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAGCGT
GGCGGCGTGACCGCAGTGGAGGCACTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCTGAACTT
ACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGG
CTGTTGCCGGTGTGTCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATT
GGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACA
CCGGAGCAGGTGGTGGCCATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTG
TTGCCGGTGTGTCAGGCCCATGGCCTACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGC
GGCAAGCAGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACTCCG
GAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTG
CCGGTGTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATCATGGCGGC
AAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCGCATGGCCTTACCCCGGAG
CAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAG
CAGGTGGTGGCCATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGGTGGCCATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGT
CTGTGCCAGGCCCATGGCCTACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAG
GCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACTCCGGAGCAGGTG
GTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTG
TGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCGCATGGCCTTACCCCGGAGCAGGTGGT
GCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGC
CAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTG
GAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACACCGGAGCAGGTGGTGGCC
ATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAG
GCCCATGGCCTACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATC
GCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCC
CATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGAGC
ATTGTTGCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCTTGACCAACGACCACCTCGTTCGCTTG
GCCTGCCTCGGCGGACGCTCTGCGCTGGATGCAGTGAAGGAGGATTGCCGCACGCGCCGGCCTTGATC
AAAAGAACAATCGCCGATTTCCGAACGCACATCCCATCGGTTGCCGACCACGCGCAAGTGGTTTCG
GTGCTGGGTTTTTTCCAGTGCCACTCCACCCAGCGCAAGCATTTGATGACCCATGACGCAGTTTCGGG
ATGAGCAGGACGGGTTGTTACAGCTCTTTCGAGAGTGGGCGTACCAGAACTCGAAGCCCGCAGTGGAA
ACGCTCCCCCAGCCTCGCAGCGTTGGGACCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCAAACCG
TCCCCTACTTCAACTCAAACGCCGGATCAGGCGTCTTTCATGCATTTCGCCGATTTCGCTGGAGCGTGAC
CTTGATGCGCCTAGCCCAATGCACGAGGGAGATCAGACGCGGGCAAGCAGCCGTAACCGTCCCAGTTCG
GATCGTGTGTACCCGGTCCCTCCGCACAGCAATCGTTCGAGGTGCGCGTTCCCGAACAGCGCGATGCG

TTGCATTTGCCCTCCTCAGCTGGGGTGTAAAACGCCCGGTACCAGGATCGGCGGCCTCCTGGATCCT
GGTACGCCCATGGATGCCGACCTGGTAGCGTCCAGCACCGTGGTTTGGGAACAAGATGCGGACCCCTTC
GCAGGGACAGCGGATGATTTCCCGGCATCAACGAAGAGGAGCTCGCATGGTTGATGGAGCTATTGCCT
CAGTGA

^aThe sequence encoding the N-terminal GFP-tag is underlined.

GFP-TALE34

Amino acid sequence^b

MVSKGEELFTGVVPIILVELDGDVNGHKFSVSGEGEGDATYGKLTLLKFICTTGKLPVWPPTLVTTTLTYGV
QCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNIL
GHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNI EDGSVOLADHYQONTPIGDGPVLLPDNHYLSTQSA
LSKDPNEKRDMVLLLEFVTAAGITLGMDELYKHMDPIRSRTPSPARELLSGPQPDGVQPTADRGVSPPA
GGPLDGLPARRTMSRTRLPSPPAPSPAFSADSFSDLRLQFDPFLFNTSLFDSLPPFGAHHTEAATGEWD
EVQSGLRAADAPPPTMRVAVTAARPPRAKPAARRRAAQPSDASPAQVLDLRTLGYSSQQQEKIKPKVRS
TVAQHHEALVGHGFTHAHI VALSQHPAALGTAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLT
VAGELRGPPLQLDTGQLLKI AKRGGVTAVEAVHAWRNALTGAPLNLTP EQVVAIAS **NGGGKQALETVQR**
LLPVLCQAHGLTPEQVVAIAS **NI**GGKQALETVQRLLPVLCQAHGLTPEQVVAIAS **NHGGKQALETVQR**
LPVLCQAHGLTPEQVVAIAS **HDGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS** **NGGGKQALETVQRLL**
PVLCQAHGLTPEQVVAIAS **NHGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS** **HDGGKQALETVQRLLP**
VLCQAHGLTPEQVVAIAS **NI**GGKQALETVQRLLPVLCQAHGLTPEQVVAIAS **NHGGKQALETVQRLLP**
LVCQAHGLTPEQVVAIAS **NGGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS** **NI**GGKQALETVQRLLPVLC
CQAHGLTPEQVVAIAS **HDGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS** **NGGGKQALETVQRLLPVLC**
QAHGLTPEQVVAIAS **NI**GGKQALETVQRLLPVLCQAHGLTPEQVVAIAS **NHGGKQALETVQRLLPVLCQ**
AHGLTPEQVVAIAS **HDGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS** **NI**GGKQALETVQRLLPVLCQA
HGLTPEQVVAIAS **NGGGKQALE** SIVAQLSRPDPALAALTDNHLVALACLGRPALDAVKKGLPHAPALI
KRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARS
TLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDL DAPS PMHEGDQTRASSRKR SRS
DRAVTGPSAQQSFEVVRVPEQRDALHLPLLSWGVKRPTRIRIGLLDPGTPMDADLVASSTVVWEQDADPF
AGTADDFPAFNEEELAWLMELLPQ*

^bRVDs are shown in red. The N-terminal GFP-tag is underlined.

Forward TALEN34

DNA sequence^c

ATGGAACAAAAATTAATCTCAGAAGAAGACTTGGCC **AAGAAGAAGAGGAAGGTG** CAGGTGGATCTACGC
ACGCTCGGCTACAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTCTGTTTCGACAGTGGCGCAGCAC
CACGAGGCACCTGGTCGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCG
TTAGGGACCGTCTGCTCAAGTATCAGGACATGATCGCAGCGTTGCCAGAGGCGACACACGAAGCGATC
GTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTG
AGAGGTCCACCGTTACAGTTGGACACAGGCCAAC'TTCTCAAGATTGCAAAGCGTGGCGGGCGTGACCGCA
GTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCTGAACCTTACCCGGGAGCAGGTG
GTGCCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTC
TGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTGACACCGGAGCAGGTGGT
GCCATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGC
CAGGCCATGGCCTCACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTG
GAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTGACTCCGGAGCAGGTGGTGGCC
ATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAG
GCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATCATGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCGCATGGCCTTACCCCGGAGCAGGTGGTGGCCATC
GCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCC
CATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACG
GTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTGACACCGGAGCAGGTGGTGGCCATCGCC
AGCAATCATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCAT
GGCCTCACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTG
CAGCGGCTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATCGCCAGC
AATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCATGGC
CTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCAGTGGCGGCAAGCAGGCGCTGGAGACGGTGCAG
CGGCTGTTGCCGGTGTGTTGCCAGGCGCATGGCCTTACCCCGGAGCAGGTGGTGGCCATCGCCAGCAAT
GGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCATGGCCTG
ACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGG
CTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTGACACCGGAGCAGGTGGTGGCCATCGCCAGCAATCAT
GGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTTGCCAGGCCCATGGCCTCAC

CCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTG
 TTGCCGGTGTGTGCCAGGCCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGT
 GGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCCATGGCCTGACCCCG
 GAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGAGCATTGTTGCCAGTTA
 TCTCGCCCTGATCCGTGTTGGCCGCGTTAACCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGA
 CGTCTGCGCTGGATGCAGTGAAAAAGGGATTGCCGACGCGCCGGCCTTGATCAAAAGAACCAATCGC
 CGTATTCCCGAACGACATCCCATCGCGTTGCCGGATCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAG
AAGTCCGAGCTGCGGCACAAGCTGAAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGG
AACCCACCCAGGACCGCATCCTGGAGATGAAGGTGATGGAGTTCCTCATGAAGGTGTACGGCTACAGG
GGAGAGCACCTGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGATTAC
GCGTGTGATCGTGGACACAAAGGCCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAGGCCGACGCGATG
CAGAGCTACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCCAACGAGTGGTGAAGGTGTAC
CCTAGCAGCGTGACCGAGTTCAGTTCCTGTTCTGCTGAGCGGCCACTTCAAGGGCAACTACAAGGCCAG
CTGACCAGGCTGAACCACATCACCACCTGCAATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGC
GGCGAGATGATCAAAGCCGGCACCCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAAACAACGGCGAGATC
AACCTCAGATCTTGA

^cThe sequences encoding the N-terminal myc-tag and nuclear localisation sequence (NLS) are underlined and in blue, respectively. The sequence encoding the C-terminal Sharkey-DS *FokI* is in bold.

Forward TALEN34

Amino acid sequence^d

MEQKLISEEDLAKKKRKVQVDLRLTLGYSQQQEQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAA
 LGTVAVKYQDMI AALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKI AKRGGVTA
 VEAVHAWRNALTGAPLNLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQA
 LETVQRLLPVLCQAHGLTPEQVVAIASNHGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL
 ETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNHGGKQALE
 TVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALET
 VQRLLPVLCQAHGLTPEQVVAIASNHGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETV
 QRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQ
 RLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQ
 LLPVLCQAHGLTPEQVVAIASNHGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRL
 LPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALESIVAQL
 SRPDPSLAALTNHDLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSQLVKSELEEK
KSELRHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDY
GVIVDTKAYSGGYNLPIGQADAMQSYVEENQTRNKHINPNEWKVPSSVTEFKFLFVSGHFKGNYKAQ
LTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINFRS*

^dRVDs are shown in red. The N-terminal myc-tag and nuclear localisation sequence (NLS) are underlined and in blue, respectively. The C-terminal Sharkey-DS *FokI* domain is in bold.

Reverse TALEN_{B84}

DNA sequence^e

ATGGATTACAAGGATGACGATGACAAGGCCAGAAGAAGAGGAAGGTGCAGGTGGATCTACGCACGCTC
 GGCTACAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTCTGTTCCGACAGTGGCGCAGCACCCAGG
 GCACATGGTTCGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGG
 ACCGTGCGTGTCAAGTATCAGGACATGATCGCAGCGTTGCCAGAGGCGACACAGGAGCGATCGTTGGC
 GTCGGCAAACAGTGGTCCGGCGCACGCGCTCTGGAGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGT
 CCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAGCGTGGCGGCGTGACCGCAGTGGAG
 GCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCTGAACCTTACCCCGGAGCAGGTGGTGGCC
 ATCGCCAGCAATAACGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAG
 GCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAG
 ACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACACCCGGAGCAGGTGGTGGCCATC
 GCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCC
 CATGGCAATCACCCTGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACG
 GTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATCGCC
 AGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCAT
 GGCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTG
 CAGCGGCTGTTGCCGGTGTGTGCCAGGCGCATGGCCTTACCCCGGAGCAGGTGGTGGCCATCGCCAGC
 AATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGC
 CTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAG
 CGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACACCCGGAGCAGGTGGTGGCCATCGCCAGCAAT
 AACGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTC
 ACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAACGGTGGCAAGCAGGCGCTGGAGACGGTGGCAAG
 CTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATCGCCAGCAATATT
 GGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACC
 CCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTG
 TTGCCGGTGTGTGCCAGGCGCATGGCCTTACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGT
 GGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACCCCG
 GAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTG
 CCGGTGTGTGCCAGGCCATGGCCTGACACCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGC
 AAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTCACCCCGGAG
 CAGGTGGTGGCCATCGCCAGCAATAACGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCC
 GTGCTGTGCCAGGCCATGGCCTGACTCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAG
 CAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGTGTGCCAGGCCATGGCCTGACCCCGGAGCAG
 GTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGAGCATTGTTGCCAGTTATCTCGC
 CCTGATCCGTCGTTGGCCGCGTTAACCAACGACCACCTCGTCGCCTTGGCCTGCCCTGGCGGACGTCCT
 GCGCTGGATGCAGTGAAAAAGGATTGCCGCACGCGCCGGCCTTGATCAAAGAACAATCGCCGTATT
 CCCGAACGCACATCCCATCGCGTTGCCGGATCC**CAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCC**
GAGCTGCGGCACAAGCTGAAGTACGTGCCACAGTACATCGAGCTGATCGAGATCGCCAGGAACCCC
ACCCAGGACCGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAGAG
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCATCGATTACGGCGTG
ATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGCCAGGCCAGAGATGCAGAGA
TACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCAAACGAGTGGTGGAAAGGTGTACCCTAGC
AGCGTGACCGAGTTCAAGTTCCTGTTCTGAGCGGCCACTTCAAGGGCAACTACAAGGCCAGCTGACC
AGGCTGAACCACATCACCAACTGCAATGGCGCGTGTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAG
ATGATCAAAGCCGGCACCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTC
AGATCTTGA

^eThe sequences encoding the N-terminal FLAG-tag and nuclear localisation sequence (NLS) are underlined and in blue, respectively. The sequence encoding the C-terminal Sharkey-RR *FokI* is in bold.

Reverse TALEN_{B84}

Amino acid sequence^f

MDYKDDDDKAKKKRKVQVDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSQHPAALG
 TVAVKYQDMI AALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKI AKRGGVTAVE
 AVHAWRNALTGAPLNLTP EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI GGKQALE
 TVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI GGKQALET
 VQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
 QRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQ
 RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQR
 LLPVLCQAHGLTPEQVVAIASNI GGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLL
 LPVLCQAHGLTPEQVVAIASNI GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI GGKQALETVQRLL
 PVLCQAHGLTPEQVVAIASNI GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLP
 VLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLP
 PDPSLAALNDHLVALACLGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGS**QLVKSELEEKKS**
ELRHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGSRKPDGAIYTVGSPIDYGV
IVDTKAYSGGYNLP I GQAREMQRYVEENQTRNKHINPNEWKVPSSVTEFKFLFVSGHFKGNYKAQLT
RLNHI TNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINFRS*

^fRVDs are shown in red. The N-terminal FLAG-tag and nuclear localisation sequence (NLS) are underlined and in blue, respectively. The C-terminal Sharkey-RR *FokI* domain is in bold.

Os12g29220 allelic promoters

CLUSTAL 2.1 multiple DNA sequence alignment

> *Oryza sativa* cv Nipponbare
 > *Oryza sativa* cv Zhenshan 97

```

AGGGATGTCTACTGCAGGTGAAAAACAATCCTTCGACAAAAATAAGTTACTTTTGGTAAA 60
AGGGATGTCTACTGCAGGTGAAAAACAATCCTTCGACAAAAATAAGTTACTTTTGGTAAA 60

GACAGTTAAATAATAAGCAGCTATATCACGCGCATGGGAGAATTGCATATTCAATTACAA 120
GACAGTTAAATAATAAGCAGCTATATCACGCGCATGGGAGAATTGCATATTCAATTACAA 120

TCATTATTTTTTTTTTCAGAACACTGTGCGCGACATTGAGAATTAATCTACCCGTGCAAAC 180
TCATTATTTTTTTTTTCAGAACACTGTGCGCGACATTGAGAATTAATCTACCCGTGCAAAC 180

AAAGAACAGAGAAACTATAGTATACCTACATGTATCTATCACCCAATAATTGCAAGATCA 240
AAAGAACAGAGAAACTATAGTATACCTACATGTATCTATCACCCAATAATTGCAAGATCA 240

TGTTACAAAACGGTTCTAATTAATATATAGAAACAAGGCAGAGAATTCTACCTTTCTTTT 300
TGTTACAAAACGGTTCTAATTAATATATAGAAACAAGGCAGAGAATTCTACCTTTCTTTT 300

GTCTAAGTACAATTATCTTTTTCTCCGCGATTAATATTTTTTCGAGTAGTAAAATTTAAGT 360
GTCTAAGTACAATTATCTTTTTCTCCGCGATTAATATTTTTTCGAGTAGTAAAATTTAAGT 360

CAAAGCCGTATCAGGATTCAGGAATAATCCTTCACCTGGGAGAGATCTCATGTGATTTGC 420
CAAAGCCGTATCAGGATTCAGGAATAATCCTTCACCTGGGAGAGATCTCATGTGATTTGC 420

TGTTGCACTCGCGGCTATCTTTTACC GTTCCCAGCAGGAAGCTGCAGACGTTGGAGAGA 480
TGTTGCACTCGCGGCTATCTTTTACC GTTCCCAGCAGGAAGCTGCAGACGTTGGAGAGA 480

TCGATCTCTACTGACAATGCACAAAGCAATTACTCACTAAATTTGGCTATGGCTAGTGAGA 540
TCGATCTCTACTGACAATGCACAAAGCAATTACTCACTAAATTTGGCTATGGCTAGTGAGA 540

GGTGCGCTGCGCACAAAGCCAATGCAACTTTTTTTTAAAAATTAGCCAGGATTATCTCCAA 600
GGTGCGCTGCGCACAAAGCCAATGCAACTTTTTTTTAAAAATTAGCCAGGATTATCTCCAA 600

CAGTAGCTCATTTTTGTAAAAGCCTAATTATTGTGCGTGTCCAAAAGACTTTCCTCAAAA 660
CAGTAGCTCATTTTTGTAAAAGCCTAATTATTGTGCGTGTCCAAAAGACTTTCCTCAAAA 660

GCAAATAAAGAAAAAAAATCTTTGCATAATTATTCTATGATTACTTTGATGCGTACGTGA 720
GCAAATAAAGAAAAAAAATCTTTGCATAATTATTCTATGATTACTTTGATGCGTACGTGA 720

```

ATGGCCATGGGTAGGAGGCAACCAAGTGATTC**C**CACCTAGCTAGCTTT**C**CTCCTATATAA 780
 ATGGCCATGGGTAGGAGGCAACCAAGTGATTC**C**CACCTAGCTAGCTTT**C**CTCCT**TATATAA** 780

AGCACACAACCTCCCTTCATTCTCTCCAAGAGTTTTCAGCCAACACATTGAACTCTTCT 840
-GCACACAACCTCCCTTCATTCTCTCCAAGAGTTTTCAGCCAACACATTGAACTCTTCT 839

TCAGAGCTCTCCCTTCCTCCACAAAGGGGGTCTAGGGTTAGAGTGTGTGTGTCTGTGAC 900
 TCAGAGCTCTCCCTTCCTCCACAAAGGGG-TCTAGGGTTAGAGTGTGTGTGTCTGTGAC 898

AAGTTCCAAGCTAGCAACAACAAGCTCAATTCCTTGCTTGTGTTGCTTCCATATTACACTA 960
 AAGTTCCAAGCTAGCAACAACAAGCTCAATTCCTTGCTTGTGTTGCTTCCATATTACACTA 958

CATCTCTTCCCTTCAATTACCCCCCTTTTAGCACACAAAA 1000
 CATCTCTTCCCTTCAATTACCCCCCTTTTAGCACACAAAA 998

ArtXa251/ArtXa252 target sequences are underlined. The PthXo2 target sequence is in bold. Single nucleotide exchanges are shown in grey.

New assembly vectors for the "Golden TAL Technology" toolbox⁶

One to six repeats can be assembled using Golden Gate cloning into *BpiI* sites (underlined) of an assembly vector. The repeats in the assembly vectors are flanked by *BsaI* sites (black, bold) to facilitate assembly into the final TALE. Two new assembly vectors were designed to expand the maximum number of repeats per TALE from 23.5⁶ to 29.5. Part of the assembly vector sequences are shown. *BsaI* overhangs are in colors (C: blue, D: magenta, R: dark grey; according to⁶). *BpiI* overhangs matching to the first and last repeat of the repeat assembly are shaded in light and dark grey, respectively. Relevant repeat codons are translated.

Assembly vector C-D

GTC**GGTCTC**AGTGCCAGGCGCATGGCC**T**TACAAGTCTTCCTT•••*lacZ*•••
 C Q A H G L T

•••GGCGAAACTT**TGG**A**G**AGCGGTGCAGCGGCTGTG**CC**GGT**G****CTGT****GAG**ACCAC
 L E T V Q R L L P V L

Assembly vector D-R

GTC**GGTCTC**ACTCA**CTGT**GCCAGGCGCATGGCC**T**TACAAGTCTTCCTT•••*lacZ*•••
 L C Q A H G L T

•••GGCGAAACTT**TGG**A**G**AGCAT**GAG**ACCAC
 L E S

Supplementary References

1. Römer, P. *et al.* Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **187**, 1048-1057 (2010).
2. 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).
3. Streubel, J., Blücher, C., Landgraf, A. & Boch, J. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**, 593-595 (2012).
4. 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).
5. Grau, J. *et al.* Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Comput. Biol.* **9**, e1002962 (2013).
6. Geißler, R. *et al.* Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**, e19509 (2011).

2.3.1.2 Zusammenfassung Publikation 3

TALEs binden ihre Ziel-DNA-Sequenz in einer hochgradig regelmäßigen ein-*repeat*-ein-Nukleotid Abfolge. Ein Nachteil dieser Bindung ist die geringe Flexibilität, besonders im Bezug auf Suszeptibilitätsgene in denen TALE-Bindestellen bereits durch Deletionen oder Insertionen (Indel) einzelner Nukleotide zerstört werden können. Dadurch wird die Bindung des TALEs und damit die Aktivierung dieses Suszeptibilitätsgens blockiert und die Pflanze ist resistent. Interessanterweise findet man in reispathogenen *Xoo*- und *Xoc*-Stämmen TALEs die einen einzelnen *repeat* mit abweichender Länge, einen sogenannten aberranten *repeat* in der Mitte ihrer *repeat*-Region besitzen. Aberrante *repeats* weisen in ihren α -Helices Deletionen bzw. Insertionen auf und sind somit im Gegensatz zur Standardlänge die 34 Aminosäuren beträgt mit 30 Aminosäuren wesentlich kürzer bzw. mit 39/40/42 Aminosäuren deutlich länger. Diese Veränderungen in der Länge der *repeat*- α -Helices könnten die regelmäßige Struktur der *repeat*-Region stören und somit die Gesamt-Bindung des TALEs verringern. Einen Hinweis, dass dies nicht zutrifft, geben die TALEs PthXo3 und AvrXa7, die einen solchen aberranten *repeat* tragen und trotzdem das Suszeptibilitätsgen *OsSWEET14* aktivieren. Allerdings stellte sich heraus, dass besonders nach der Position des aberranten *repeats* gehäuft RVD-Nukleotid-Fehlpaarungen auftreten. In einem hypothetischen Experiment wurde der aberrante *repeat* aus der *repeat*-Region herausgewunden, sodass alle folgenden RVDs eine Position nach vorn verschoben wurden. Damit konnte die Anzahl der RVD-Nukleotid-Fehlpaarungen signifikant verringert werden, wodurch die Hypothese entstand, dass aberrante *repeats* aus der *repeat*-Region herausgewunden werden können um einen *frameshift* in der Ziel-DNA-Sequenz zu tolerieren. Um dies experimentell zu bestätigen wurden künstliche TALEs erstellt, die einen aberranten *repeat* an verschiedenen Positionen ihrer *repeat*-Region tragen. Anhand verschiedener Reporterkonstrukte konnte dabei festgestellt werden, dass TALEs mit aberrantem *repeat* neben der optimalen Ziel-DNA-Sequenz auch eine Ziel-DNA-Sequenz in der ein Nukleotid *downstream* der Position des aberranten *repeats* deletiert wurde (-1 *frameshift*), binden und das dahinterliegende Gen effizient aktivieren können. TALEs mit ausschließlich 34-Aminosäure *repeats* waren dazu nicht in der Lage. Diese Flexibilität bezüglich der Ziel-DNA-Sequenz scheint besonders nützlich für die Anpassung an mögliche Resistenzgene zu sein, in denen die TALE-Bindestelle durch die Deletion eines Nukleotides zerstört wurde. Tatsächlich konnte anhand einer natürlichen *Xoo*-Reis

Infektion gezeigt werden, dass ein künstlicher TALE mit aberrantem *repeat* in der Lage ist das resistente Allel des *OsSWEET13*-Promotors, mit einem Indel in der PthXo2-Bindestelle, ebenso zu aktivieren wie das Allel mit optimaler Bindestelle. Im Gegensatz dazu konnte der TALE mit ausschließlich 34-Aminosäure *repeats* lediglich das Allel mit optimaler Bindestelle, nicht jedoch das resistente Allel aktivieren. Daher könnten aberrante *repeats* eine nützliche Anpassung an allelische Suszeptibilitätsgene sein, in denen eine TALE-Bindestelle durch ein einzelnes Indel zerstört wurde, um diese Form der Resistenz zu umgehen.

2.4 Eigenanteil an den Publikationen

Publikation 1 Kapitel 2.1:

Streubel, J., Blücher, C., Landgraf, A., and Boch, J. (2012) TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**: 593-595.

Eigenanteil: Die Planung der Experimente erfolgte in Zusammenarbeit mit J. Boch. Mit Ausnahme der in Abbildung 1A dargestellten Konstrukte und Ergebnisse, wurden alle künstlichen TALEs, sowie zugehörige Reporterstrukturen von mir erstellt und die nachfolgenden Analysen der Reporteraktivität bzw. der Nachweis mittels Western Blot selbstständig von mir durchgeführt und ausgewertet. Das Manuskript wurde in Zusammenarbeit mit J. Boch erstellt.

Publikation 2 Kapitel 2.2:

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

*geteilte Erstautorenschaft

Eigenanteil: Die Klonierung und Aktivitätsanalyse von Tal5 mittels β -Glucuronidase Assay erfolgte durch C. Pesce unter meiner Anleitung während ihres 1-monatigen Aufenthalts in Halle. Alle in dieser Publikation verwendeten künstlichen TALE-Konstrukte für *OsSWEET*-Promotoren wurden selbstständig von mir geplant und erstellt. Die Transformation der in dieser Publikation verwendeten Stämme mit den künstlichen TALE-Konstrukten, sowie die ersten Virulenzanalysen in Reis wurden von mir während meines ca. 5 wöchigen Aufenthalts im IRD (Montpellier) begonnen und

von C. Pesce beendet. Außerdem führte C. Pesce die *in planta* Wachstumskurven durch. Die RNA-Proben zur Bestimmung der *OsSWEET*-Transkriptlevel nach der Infektion mit den in dieser Publikation verwendeten *Xoo*-Stämme wurde von C. Pesce vorbereitet. Die folgende anschließende Analyse der Expression der künstlichen TALEs in den verwendeten *Xoo*-Stämmen, sowie die Analyse der Transkriptmenge der ca. 20 *OsSWEETs* mittels qRT wurden eigenständig von mir geplant, durchgeführt und ausgewertet: die. Das Manuskript wurde in Zusammenarbeit mit J. Boch, B. Szurek, R. Koebnik und C. Pesce erstellt.

Publikation 3 Kapitel 2.3:

Richter A.*, Streubel J.*, Blücher C., Szurek B., Reschke M., Grau J. and Boch J. (2014) A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* **5**: 3447

*geteilte Erstautorenschaft

Eigenanteil: Die Planung der Experimente erfolgte in Zusammenarbeit mit A. Richter und J. Boch. Folgende Arbeiten wurden von mir in Zusammenarbeit mit A. Richter geplant, durchgeführt und analysiert: Erstellung der verwendeten TALE- und Reporterstrukture, Nachweis der TALE-Expression *in planta*, Analyse der TALE-Aktivität anhand des β -Glucuronidase Reportergens. Folgende Arbeiten wurden von A. Richter geplant, durchgeführt und ausgewertet: Erstellung, Aktivitätsanalyse und Nachweis der TALEN Konstrukte. Die unter Abbildung 8 gezeigten Konstrukte und Experimente wurden eigenständig von mir geplant, durchgeführt und ausgewertet. Das Manuskript wurde in Zusammenarbeit mit A. Richter, J. Grau und J. Boch erstellt.

2.5 Zusätzliche Ergebnisse

2.5.1 Funktionsanalyse von TALEs am *OsSWEET14* Promotor

Aktivatorproteine können die Transkription nicht nur ausgehend von Kernpromotorsequenzen, sondern auch ausgehend von sogenannten *Enhancer*-Elementen initiieren (Maston, *et al.*, 2006; Bulger & Groudine 2011; Levine, *et al.*, 2014). *Enhancer*-Elemente sind DNA-Sequenzen, welche die Transkription eines Gens über Entfernungen von bis zu 1Mb kontrollieren können (Jeong, *et al.*, 2006; Amano, *et al.*, 2009). Die Bindung an *Enhancer*-Elemente ermöglicht Aktivatoren die Transkription eines Ziel-Gens unabhängig von ihrer Position oder der Orientierung ihrer Bindestelle relativ zum Kernpromotor zu aktivieren (Maston, *et al.*, 2006; Bulger &

Groudine 2011; Levine, *et al.*, 2014). Für TALEs ist bisher unklar, ob sie ähnlich wie *Enhancer*-bindende Proteine die Transkription eines Ziel-Gens ausgehend von verschiedensten Positionen bzw. unabhängig von der Orientierung ihrer Ziel-DNA-Sequenz aktivieren können. Möglicherweise sind für TALEs zusätzlich andere Faktoren wie beispielsweise umliegende Promotorelemente oder Wirtsproteine, welche die Zugänglichkeit des Promotors beeinflussen, von Bedeutung. Studien die sich diesem zusätzlichen Level an Komplexität widmen, wurden bisher allerdings kaum durchgeführt. Die zugrundeliegenden Zusammenhänge und Mechanismen sind deshalb bisher nur unzureichend verstanden. Ihre Analyse könnte neue Einblicke in das Zusammenwirken zwischen TALEs mit spezifischen Promotorelementen, sowie der Transkriptionsmaschinerie im Allgemeinen geben.

Studien mit TALE-Aktivatoren im humanen System zeigen, dass selbst TALE-Derivate deren Ziel-DNA-Sequenzen in der kodierenden Sequenz eines Gens, mehr als 500bp vom ATG entfernt oder sogar in *reverse* (umgekehrter) Orientierung liegen, ein Ziel-Gen effektiv aktivieren können (Maeder, *et al.*, 2013; Perez-Pinera, *et al.*, 2013; Uhde-Stone, *et al.*, 2014; Zhang, *et al.*, 2014). Dabei zeigte sich, dass zusätzlich die Positionierung der Bindestelle in DNaseI-sensitiven Promotorregionen vorteilhaft war (Maeder, *et al.*, 2013; Zhang, *et al.*, 2014). Mit natürlichen TALEs im pflanzlichen System wurden solche systematischen Studien dagegen noch nicht durchgeführt. Hier ist lediglich bekannt, dass TALEs den Transkriptionsstart eines Ziel-Gens beeinflussen können (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Antony, *et al.*, 2010; Hummel, *et al.*, 2012). Wobei jedoch häufig künstliche Promotoren analysiert wurden. In diesen wurde entweder eine vorhandene TALE-Bindestelle verschoben oder zusätzlich zu den bereits vorhandenen TALE-Bindestellen bis zu fünf weitere Bindestellen verschiedener TALEs inseriert (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Hummel, *et al.*, 2012). Die dafür verwendeten TALE-Bindestellen enthielten jedoch häufig eine TATA-Box, sodass diese ebenfalls entweder verschoben oder mehrfach inseriert wurde (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Hummel, *et al.*, 2012). Die Möglichkeit künstliche TALEs für beliebige Ziel-DNA-Sequenzen zu erstellen, erlaubt es nun TALEs an verschiedene Stellen eines natürlichen Promotors zu positionieren und deren Aktivität zu vergleichen, ohne dabei den Promotor selbst zu verändern. Im Rahmen dieser Arbeit sollte die Aktivität verschiedener künstlicher TALEs an einem Beispieldpromotor im transienten, sowie im natürlichen System analysiert werden. Publikation 2 beschreibt, dass künstliche TALEs, die für Promotoren bestimmter *OsSWEET*-Gene aus Reis erstellt wurden, zu erhöhter

Virulenz von *Xoo*-Stämmen führen können (Streubel, *et al.*, 2013). Ausgehend davon bot sich vor allem der *OsSWEET14* Promotor für weiterführende Studien an. *OsSWEET14* kodiert einen Saccharose-Transporter, dessen TALE-abhängige Induktion, vermutlich durch die Versorgung mit Nährstoffen, zu erhöhter Virulenz von *Xoo* führt (Chen, *et al.*, 2012; Braun 2012). Für diesen Promotor sind Bindestellen für vier verschiedene natürliche *Xoo*-TALEs bekannt, die bereits als funktional beschrieben wurden. (Yang, *et al.*, 2006; Antony, *et al.*, 2010; Yu, *et al.*, 2011; Streubel, *et al.*, 2013). Darüber hinaus wurden in Publikation 2 bereits zwei künstliche TALEs für diesen Promotor konstruiert und im *Xoo*-Reis-System analysiert (Streubel, *et al.*, 2013). Zusätzlich war ein Reporterkonstrukt verfügbar, um die Funktion der TALEs am *OsSWEET14*-Promotor auch im transienten System (*Agrobacterium*-vermittelter Transfer in *N. benthamiana*) zu untersuchen. Im Rahmen dieser Arbeit wurden weitere künstliche TALEs erstellt, die Aufschluss über Positionierungsmöglichkeiten von TALE Ziel-DNA-Sequenzen im *OsSWEET14*-Promotor geben sollten. Eine Übersicht über deren RVD-Komposition, sowie Ziel-DNA-Sequenzen kann dem Anhang entnommen werden (s. Anhang 5.4). Zur besseren Übersicht wurden die erstellten TALEs bezüglich ihrer Bindestellen im *OsSWEET14*-Promotor in 3 Gruppen aufgeteilt. Die erste Gruppe umfasst künstliche TALEs deren Ziel-DNA-Sequenzen im Bereich der Bindestellen natürlicher TALEs am *OsSWEET14*-Promotor (TalC, AvrXa7/PthXo3 und Tal5), ca. 220 bis 335bp *upstream* des ATG liegen. (Antony, *et al.*, 2010; Yu, *et al.*, 2011; Streubel, *et al.*, 2013; Abb. 2.1A). Zu dieser Kategorie zählen die künstlichen TALEs TAL1, 2, 4 und 5. Ihr zugehöriger Promotorabschnitt ist in Abbildung 3.1A durch ein **b** gekennzeichnet. Die zweite Gruppe umfasst künstliche TALEs, deren Ziel-DNA-Sequenzen entweder in einem Bereich von 336 bis 1000bp *upstream* vom ATG liegen (der zugehörige Promotorbereich ist in Abbildung 2.1A mit **a** markiert) oder ausgehend vom ATG in einem Bereich von 1 bis zu 220bp *upstream* im Promotor bzw. *downstream* in der Protein-kodierenden Sequenz liegen (der zugehörige Promotorbereich ist in Abbildung 2.1A mit **c** markiert). Mit Hilfe dieser künstlichen TALEs zu denen im Bereich **a** TAL11, 14 und 17 und im Bereich **c** TAL3, 6, 7, 8, 9, 10, 18, sowie der TALGUS (der nur für das Reporterkonstrukt im transienten System in *N. benthamiana* analysiert wird) gehören, sollte überprüft werden ob und von welchen Positionen TALEs wirken können. Als dritte Gruppe wurden künstliche TALEs erstellt, die in umgekehrter Orientierung den Gegenstrang (*antisense* Strang) schon vorhandener TALE Ziel-DNA-Sequenzen binden, sogenannte *reverse*-TALEs. Dazu gehören TAL-1,

-2, -AvrXa7, -3, -7, -81, -82 und -9. Durch die Bindung an *Enhancer*-Elemente können Aktivatorproteine die Transkription nicht nur unabhängig von ihrer Position sondern auch unabhängig von der Orientierung ihrer Bindestelle relativ zum Kernpromotor aktivieren (Bulger & Groudine 2011; Levine, *et al.*, 2014). Bisher ist jedoch unklar, ob dies für TALEs ebenfalls zutrifft. Anhand des TALEs AvrBs3 wurde gezeigt, dass dieser das Ziel-Gen *Bs3* nicht mehr aktiviert, sobald dessen Ziel-DNA-Sequenz im *Bs3* Promotor invertiert wird (Römer, *et al.*, 2009). Dadurch wurde allerdings auch die in der AvrBs3-Bindestelle enthaltene TATA-Box invertiert, weshalb nicht auszuschließen ist, dass damit die generelle Aktivierbarkeit des Promotors beeinflusst wird (Römer, *et al.*, 2009). Außerdem sind für natürliche TALEs bisher nur *forward*-orientierte Bindestellen in Ziel-Promotoren bekannt (Boch, *et al.*, 2009, Moscou & Bogdanove 2009). Daher entstand die weitverbreitete Meinung, dass TALEs die Transkription nur über *forward*-orientierte Bindestellen aktivieren können. Ausgehend davon wird auch in verfügbaren Vorhersage-Programmen für TALE-Ziel-Gene lediglich nach *forward*-orientierten Bindestellen in Promotoren gesucht (Grau *et al.*, 2013; Perez-Quintero *et al.*, 2013). Da bisher jedoch noch nicht experimentell analysiert wurde, ob TALEs mit natürlicher Aktivierungsdomäne auch über *reverse* auf dem Gegenstrang positionierte Bindestellen effizient die Transkription aktivieren können, sollte dies mit Hilfe der obengenannten *reverse*-TALEs untersucht werden. Weiterhin kann mit dem Repertoire verschieden positionierter künstlicher TALEs systematisch analysiert werden inwiefern der Transkriptionsstart durch diese am Promotor beeinflusst wird. Um auszuschließen, dass die erstellten TALEs eine zusätzliche Bindestelle im 1kb-Fragment des *OsSWEET14* Promotors oder in einem anderen schon bekannten Suszeptibilitätsgen des Reisgenoms besitzen, wurde deren Spezifität mit Hilfe eines Programms zur Vorhersage potenzieller TALE-DNA-Bindestellen überprüft (Storyteller; Perez-Quintero, *et al.*, 2013). Als Referenz wurde das annotierte *Nipponbare*-Genom Version 7.0 verwendet (Rice MSU Annotation Project). Mögliche Ziel-DNA-Sequenzen der TALEs werden dabei mit einem Rang bewertet, der sich aus der Position der Bindestelle im Promotor, sowie der Bindespezifität zusammensetzt. Die beste Wertung erhielt dabei immer die vorher definierte Ziel-DNA-Sequenz im *OsSWEET14* Promotor. Unabhängig von Position oder Orientierung wurde für keinen der künstlichen TALEs eine weitere mögliche Ziel-DNA-Sequenz in einem anderen bekannten Suszeptibilitätsgen oder eine zweite Bindestelle im *OsSWEET14* Promotor vorhergesagt.

2.5.2 Funktionsanalyse der *OsSWEET14*-TALEs im transienten System

Zuerst wurde die Funktion der künstlichen TALEs im transienten System analysiert. Hierzu wurden die TALEs zusammen mit einem Reporterkonstrukt *Agrobacterium*-vermittelt in *N. benthamiana* transferiert. Das verwendete Reporterkonstrukt enthielt, beginnend vom ATG, die ersten 1000bp des *OsSWEET14*-Promotors, fusioniert an die kodierende Sequenz des Reportergens *uidA* (Abb. 2.1A & Abb. 5.3). Die TALE-abhängige Aktivierung des Reportergens wurde mittels quantitativer β -Glucuronidase Messung analysiert. Es zeigte sich, dass die verschieden positionierten TALEs das Reporterogen unterschiedlich effizient aktivieren (Abb. 2.1B).

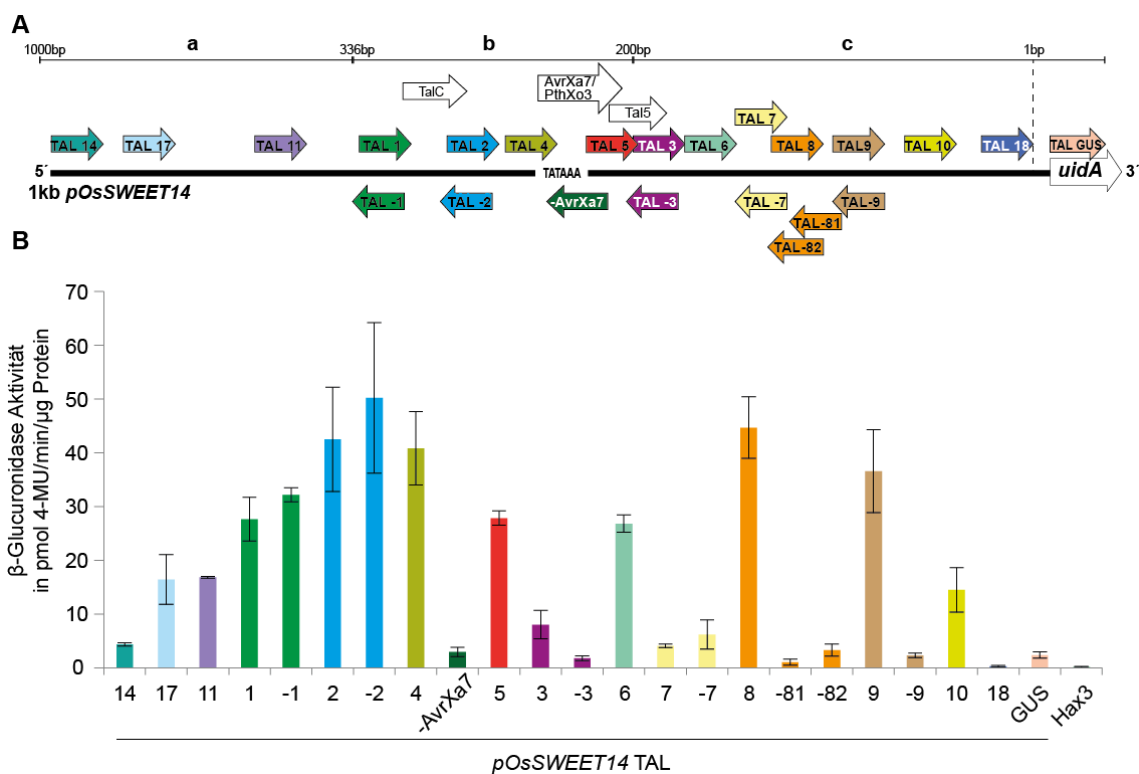


Abbildung 2.1 TALEs am *OsSWEET14*-Promotor und deren Aktivität im transienten System

A Dargestellt ist das 1kb Fragment des *OsSWEET14*-Promotors fusioniert an das Reporterogen *uidA*. Die Bindestellen der künstlichen TALEs, sowie deren Orientierung sind mit farbigen Pfeilen gekennzeichnet. Die Pfeilspitze kennzeichnet jeweils die Lage der AD. Die Bindestellen der natürlichen TALEs TalC, AvrXa7 und Tal5 sind jeweils durch einen weißen Pfeil gekennzeichnet. Der Promotor wurde zur besseren Übersicht in die drei Bereiche **a**, **b** und **c** unterteilt. Die relative Position des ATG der *OsSWEET14*-kodierenden Sequenz in dem verwendeten Reporterkonstrukt ist durch eine gestrichelte Linie gekennzeichnet. Die Darstellung ist rein schematisch und nicht maßstabsgetreu. Eine genauere Übersicht findet sich im Anhang unter Abb. 5.3. **B** Aktivitäten der künstlichen TALEs im transienten System. Die Messung der TALE-abhängigen Aktivierung des Reportergens erfolgte nach *Agrobacterium*-vermittelter transienter Expression der TALEs und ihrer Reporterkonstrukte in *N. benthamiana*. Dazu wurden die TALE-Konstrukte, deren Expression durch den konstitutiven Promotor 35S kontrolliert wird, sowie das Reporterkonstrukt mit einer finalen OD₆₀₀ von jeweils 0,4 ko-inokuliert und 2dpi die Aktivität des Reportergens bestimmt. Als Negativkontrolle diente der TALE Hax3 der keine Bindestelle im *OsSWEET14*-Promotor besitzt. Die Fehlerbalken kennzeichnen die Standardabweichung (n=3). 4-MU=4-Methylumbelliferon. Das Experiment wurde dreimal mit vergleichbaren Ergebnissen durchgeführt.

Die Unterschiede in der Aktivierung des Reportergens können nicht auf die unterschiedliche Expression oder Stabilität der TALEs *in planta* zurückgeführt werden. Der Nachweis der TALE-Expression zeigte, dass die Mehrheit der TALEs in vergleichbaren Mengen *in planta* synthetisiert wird (Abb. 2.2). Lediglich für TAL2, TAL8 und TAL-81 konnte ein schwächeres Signal detektiert werden (Abb. 2.2). Dies scheint jedoch keine Auswirkung auf die Aktivität dieser TALEs zu haben, da besonders TAL2 und 8 sehr effizient sind (Abb. 2.1).

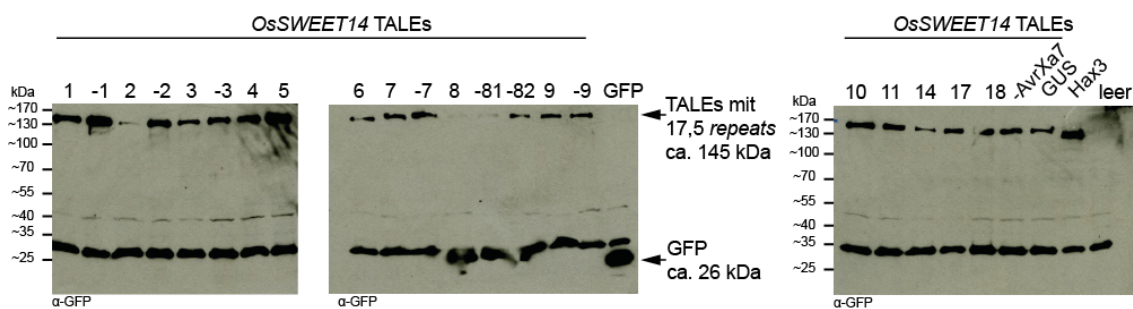


Abbildung 2.2 Nachweis der künstlichen TALE-Proteine *in planta*

Parallel zur Messung der β -Glucuronidase Aktivität wurden die TALE-Proteine mittels Immunoblot nachgewiesen. Die TALE-Konstrukte sind N-Terminal an GFP fusioniert und wurden wie unter Abbildung 2.1 beschrieben in *N. benthamiana* exprimiert. 2dpi wurden je zwei Blattscheiben aus 3 biologischen Replikaten geerntet und die Proteine nach Aufschluss und SDS-PAGE mit Hilfe eines GFP spezifischen Antikörpers nachgewiesen. Als negativ-Kontrolle diente einerseits der Vektor ohne TALE (GFP), sowie eine nicht-inokulierte Pflanzenprobe (leer). Als Positivkontrolle diente Hax3. Die erwarteten Größen für TALEs mit 17,5 repeats bzw. für GFP sind mit Pfeilen markiert. Hax3 enthält lediglich 11,5 repeats und ist daher kleiner.

Daher kann vermutet werden, dass die Position der TALEs im Promotor die Aktivierung des Reportergens beeinflusst. Besonders effiziente *forward*-positionierte TALEs sind TAL1, 2, 4 und 5 des Promotorbereichs **b**, sowie die TALEs TAL11 und 17 aus dem Promotorbereich **a** und TAL6, 8, 9 und 10 des Promotorbereichs **c** (Abb. 2.1 **B**). Die stärkste Aktivierung des Reportergens (TAL-2) wurde als Maximum und damit als 100% gewertet, ausgehend davon erzielten die vorhergenannten TALEs eine Aktivierung des Reportergens im Bereich von ca. 20-85%. Weniger bis gar nicht aktive TALEs sind TAL3, 7, 14, 18 sowie TALGUS deren Aktivität im Bereich zwischen 0 und 20% liegt (Abb. 2.1 **B**). Eine Erklärung für die geringe Aktivität dieser TALEs könnte sein, dass sie trotz ausgeglichener RVD-Komposition nur schwach an ihre Ziel-DNA-Sequenz binden können und somit das Reportergen nicht effizient aktivieren (Streubel, *et al.*, 2012). Um dies auszuschließen wurde exemplarisch jeweils die 19bp umfassende TAL18- und TAL3- Ziel-DNA-Sequenz vor das 74bp-Fragment des *Bs4*-Promotors und ein promotorloses *uidA*-Gen inseriert (Abb. 2.3A). Das Fragment des

Bs4-Promotors besitzt keine Basalaktivität, kann jedoch TALE-abhängig aktiviert werden. Dieses System wurde bereits in vorhergehenden Studien benutzt, um die Aktivität künstlicher TALEs zu untersuchen und sollte einen Einfluss durch *OsSWEET14*-spezifische Promotorelemente ausschließen (Boch, *et al.*, 2009; Yu, *et al.*, 2011). Wie Abbildung 2.3B zeigt, aktivieren TAL18 und Tal3 das Reporterkonstrukt vergleichbar mit der Positivkontrolle Hax3. Die Bindung der jeweiligen Ziel-DNA-Sequenz ist spezifisch, da ein vergleichbares Reporterkonstrukt, welches anstatt der *TAL18*- bzw. der *TAL3*-box die *Hax3*-box enthält, nicht aktiviert wird (Abb. 2.3B). Dies zeigt, dass beide TALEs ihre Ziel-DNA-Sequenz effizient binden und ein dahinterliegendes Gen aktivieren können.

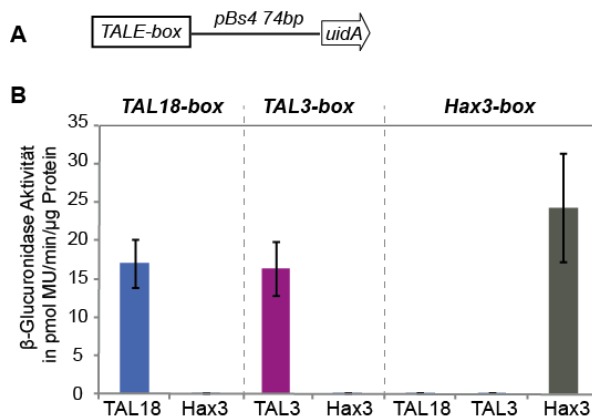


Abbildung 2.3 Analyse der TAL18- und TAL3- Aktivität am 74bp-Fragment des *Bs4*-Promotors im transienten System

A Zur Aktivitätsanalyse wurde die jeweilige *TALE*-box vor das 74bp-Fragment des *Bs4*-Promotors inseriert, dadurch sollten *OsSWEET14* spezifische Promotoreigenschaften ausgeschlossen werden. **B** 2 Tage nach *Agrobacterium*-vermittelter Expression der Konstrukte wurde die TAL18- bzw. TAL3-abhängige Aktivierung des β-Glucuronidase Reportergens gemessen. Als Negativkontrolle diente ein vergleichbares Konstrukt, welches anstelle der *TAL18*- bzw. *TAL3*- die *Hax3*-box enthält. Zusätzlich wurden die *TAL18*- und *TAL3*-box mit Hax3 ko-exprimiert um Basalaktivität dieser Reporterkonstrukte auszuschließen. Als Positivkontrolle diente Hax3 in Kombination mit der *Hax3*-box. Die Fehlerbalken kennzeichnen die Standardabweichung (n=3).

Der Grund, dass TAL18 und TAL3 im *OsSWEET14*-Promotorkontext nicht bzw. nur gering aktiv sind, könnte somit entweder auf die Position der Bindestelle relativ zu anderen Promotorelementen bzw. speziell für TAL18 die Nähe zum ATG oder aber auf die eingeschränkte Zugänglichkeit des Promotors durch Methylierung oder eine Blockierung der Bindestelle durch Wirtsproteine zurückzuführen sein. Interessanterweise sind auch einige *reverse*-bindenden TALEs in der Lage das Reportergen sehr effizient zu aktivieren. TAL-1 und -2 deren Ziel-DNA-Sequenzen *upstream* der TATA-Box in Promotorbereich **a** liegen aktivieren das Reportergen mit ca. 60 bis 100% sehr effizient und vergleichbar mit ihren zugehörigen *forward*-

bindenden TALEs (Abb. 2.1B). Zu Beginn wurde bereits mittels Vorhersageprogrammen überprüft, dass weder TAL-1, noch TAL-2 eine zusätzliche Bindestelle, speziell in *forward*-Orientierung, im *OsSWEET14* Promotor besitzen. Da beide TALEs jedoch zu einer sehr effizienten Aktivierung des Reportergens führten, sollte dies zusätzlich experimentell am Beispiel von TAL-2 untersucht werden. Wie vorher bereits beschrieben wurde dazu die 21bp umfassende Ziel-DNA-Sequenz, in der die Bindestellen von TAL2 und TAL-2 enthalten sind, vor das 74bp-Fragment des *Bs4*-Promotors und ein promotorloses *uidA*-Gen inseriert (Abb. 2.4 A,B). Der *Bs4*-Promotor sollte keinerlei Sequenzähnlichkeiten mit dem *OsSWEET14*-Promotor besitzen und somit eine Analyse der TAL2 und besonders TAL-2 Bindestelle unabhängig vom *OsSWEET14*-Promotorkontext erlauben. Die Analyse der beiden TALEs ergab, dass beide in der Lage waren das *Bs4*-Reporterkonstrukt zu aktivieren (Abb. 2.4B,D). Diese Aktivierung konnte als spezifisch bestimmt werden, da beide TALEs ein vergleichbares Reporterkonstrukt, welches anstelle der TAL2/-2-Ziel-DNA-Sequenz die Hax3-Ziel-DNA-Sequenz enthielt nicht aktivierten (Abb. 2.4B,D). Um jedoch zu überprüfen, dass TAL-2 tatsächlich die *reverse*-orientierte Ziel-DNA-Sequenz im *antisense* Strang bindet wurde zusätzlich ein weiteres *Bs4*-Reporterkonstrukt erstellt (Abb. 2.4C,D). In diesem wurden die Guanin-Nukleotide in der TAL2 Bindestelle des *sense* DNA-Strangs zu Adenin mutiert (Abb. 2.4C). Dadurch sollte TAL2 mit seinen Purin-spezifischen NN RVDs immer noch in der Lage sein diese Ziel-DNA-Sequenz zu binden. Im Gegensatz dazu erzielten diese Mutationen einen Cytosin zu Thymin Austausch in der TAL-2-Ziel-DNA-Sequenz im *antisense* Strang (Abb. 2.4C). Da TAL-2 an diesen Positionen das hochgradig Cytosin-spezifische RVD HD enthält, sollte TAL-2 nicht mehr in der Lage sein diese mutierte Ziel-DNA-Sequenz zu binden und somit das Reporter gen nicht aktivieren. Die Analyse dieses mutierten Reporterkonstruktes zeigte, dass TAL2 wie erwartet das mutierte Reporterkonstrukt trotz der G- zu A-Mutationen aktiviert (Abb. 2.4C,D). Im Gegensatz dazu konnte TAL-2 dieses Reporterkonstrukt nicht mehr aktivieren, wodurch gezeigt wird, dass die Aktivität dieses TALEs tatsächlich auf die Bindung der *reverse*-orientierten Ziel-DNA-Sequenz im *antisense*-Strang zurückzuführen ist (Abb. 2.4C,D). Auffällig bei der Analyse der *reverse*-bindender TALEs war, dass speziell die *downstream* der TATA-Box gelegenen *reverse* TALEs aus Promotorbereich **b** und **c** das Reporter gen nur gering oder gar nicht aktivieren (TAL-AvrXa7, -3, -7, -81/-82, -9; Abb. 2.1B). Dies überrascht insofern, da für einige dieser TALEs der zugehörige *forward*-bindende TALE zu einer sehr effizienten Aktivierung

des Reportergens führt (Abb. 2.1B). Dies weist darauf hin, dass die Bindestellen in diesen Bereichen grundlegend zugänglich und nicht durch Wirtsproteine blockiert sein sollten und trifft vor allem für die Paarungen TAL8/-81/-82 und TAL9/-9 zu (Abb. 2.1B). Zusammenfassend zeigen diese Ergebnisse erstmalig, dass TALEs mit natürlicher Aktivierungsdomäne im transienten System in *N. benthamiana* ein Reporter gen ausgehend von verschiedensten Positionen im Promotor, sowie in *reverse*-Orientierung aktivieren können.

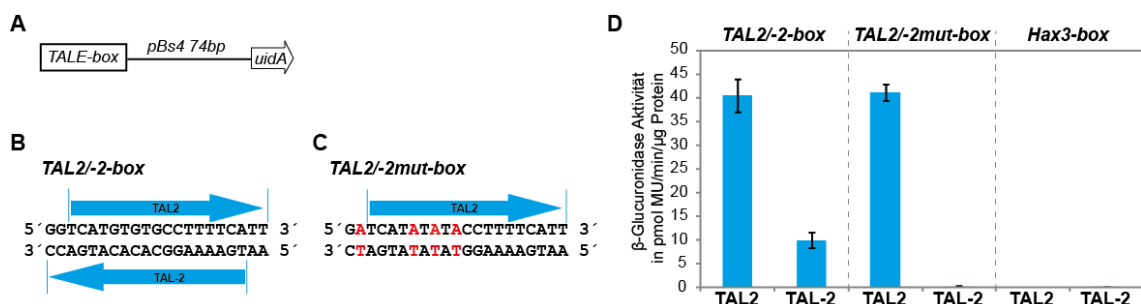


Abbildung 2.4 Analyse der TAL2- und TAL-2-Aktivität am 74bp-Fragment des *Bs4*-Promotors im transienten System

A Zur Aktivitätsanalyse wurde die jeweilige *TALE-box* vor das 74bp-Fragment des *Bs4*-Promotors inseriert, dadurch sollten *OsSWEET14* spezifische Promotoreigenschaften ausgeschlossen werden. **B** und **C** Zur Analyse wurden zwei Reporterkonstrukte erstellt. Das Reporterkonstrukt *TAL2/-2-box* enthielt ein 21bp-Fragment mit den Bindestellen für beide TALEs. Das Reporterkonstrukt *TAL2/-2mut-box* enthielt das 21bp-Fragment in dem im *sense*-Strang Guanin zu Adenin mutiert wurde (rot markiert). TAL2 sollte mit Hilfe seiner Purin-spezifischen NN-RVDs in der Lage sein diese Mutation zu tolerieren. Dagegen führt diese Mutation zu einem Cytosin zu Thymin Austausch im *antisense*-Strang, wodurch TAL-2 nicht mehr binden kann. **D** 2 Tage nach *Agrobacterium*-vermittelter Expression der Konstrukte wurde die TAL2- bzw. TAL-2-abhängige Aktivierung des β-Glucuronidase Reportergens gemessen. Als Negativkontrolle diente ein vergleichbares Konstrukt, welches anstelle der *TAL2/-2-box* die *Hax3-box* enthält. Die Fehlerbalken kennzeichnen die Standardabweichung (n=3).

2.5.3 Funktionsanalyse der *OsSWEET14*-TALEs im natürlichen Reis-System

Es ist nicht auszuschließen, dass nach transienter Expression des *OsSWEET14*-Promotors in *N. benthamiana* spezifische Charakteristika z.B. die Belegung mit Pflanzenproteinen nicht mit denen im natürlichen Reis-System vergleichbar sind. Daher sollten einige *OsSWEET14*-TALEs auch im natürlichen *Xoo*-Reis-System analysiert werden. Hierzu wurden die *OsSWEET14*-TALEs TAL14, 17, 11, 1/-1, 2/-2, 3/-3, 8/-82, 10 sowie 18 ausgewählt (Abb. 2.5A) und jeweils in ein Plasmid inseriert, das eine konstitutive Expression des entsprechenden TALEs in *Xanthomonas* erlaubt. Diese Plasmide wurden in den nicht-virulenten *Xoo*-Stamm BAI3Δ*talC* übertragen. Durch die Deletion des *TALE*-Gens *talC* fehlt diesem Stamm die Fähigkeit das Suszeptibilitätsgen *OsSWEET14* zu aktivieren. Werden allerdings alternative *OsSWEET14*-TALEs in diesem Stamm exprimiert, können diese die Aktivierung von *OsSWEET14* und damit

die Virulenz des Stammes wiederherstellen (Streubel, *et al.*, 2013). Dies kann sowohl anhand von Virulenzphänotypen, als auch durch die Untersuchung des *OsSWEET14* Expressionslevels nachgewiesen werden. Um die Virulenz der jeweiligen *Xoo*-Stämme mit oder ohne *OsSWEET14*-TALEs zu analysieren, wurden diese mittels Spritze in die suszeptible Reislinie *Nipponbare* (*Oryza sativa* ssp. *japonica*) inokuliert. Fünf Tage nach Inokulation wurden die Phänotypen ausgewertet und fotografiert. Als Positivkontrolle diente der Wildtypstamm BAI3. Er enthält *talC*, der das Suszeptibilitätsgen *OsSWEET14* aktiviert und somit typische Virulenzsymptome auslöst, die als dunkelgrüne wässrige Läsionen rund um die Inokulationsstelle sichtbar sind (Abb. 2.5B). Der Stamm BAI3 Δ *talC* kann *OsSWEET14* nicht aktivieren und ist somit nicht virulent. Wird *talC* auf einem Plasmid wieder in diesen Stamm transferiert, kann der Virulenzphänotyp in Form von wässrigen Läsionen wiederhergestellt werden (Abb. 2.5B). Die Übertragung der künstlichen TALEs TAL1/-1, 2/-2, sowie TAL8 führte ebenfalls zur Wiederherstellung der Virulenzsymptome (Abb. 2.5B). Währenddessen konnten TAL14, 17, 11, 3/-3, 7/-7, -82, 10 oder 18 den Virulenzphänotyp von BAI3 Δ *talC* nicht wiederherstellen (Abb. 2.5B). Um zu untersuchen, ob der beobachtete Phänotyp der jeweiligen *Xoo*-Stämme tatsächlich mit der Expression von *OsSWEET14* zusammenhängt, wurde diese mittels quantitativer RT-PCR analysiert. Dazu wurden die jeweiligen *Xoo*-Stämme mittels Spritze in die suszeptible Reislinie *Nipponbare* inokuliert. 24h nach Inokulation wurde das Blattmaterial geerntet und zur Extraktion von Gesamt-RNA verwendet. Als Referenz wurde der Stamm BAI3 Δ *talC* mit einem Plasmid ohne *TALE* (*ev*) verwendet und die *OsSWEET14*-Transkriptmenge in den zu untersuchenden Proben als n-fach im Vergleich zu diesem Stamm berechnet. Übereinstimmend mit dem Virulenzphänotyp konnte für TAL2/-2 und 8 eine signifikante Erhöhung des *OsSWEET14*-Transkripts gemessen werden (Abb. 2.5C). Interessanterweise konnte auch für TAL10, der die Virulenz von BAI3 Δ *talC* nicht wiederherstellen konnte, eine signifikant erhöhte *OsSWEET14*-Expression, vergleichbar mit TAL8, gemessen werden (Abb. 2.5C). In Anbetracht dessen sollte dieser Stamm ebenfalls virulent sein. Auffällig bei der Auswertung der Virulenz-Phänotypen war jedoch, dass BAI3 Δ *talC* mit TAL10 fünf Tage nach Inokulation statt wässriger Läsionen, braune Läsionen hervorrief (Abb. 5.4). Dies könnte ein Hinweis auf eine Abwehrreaktion sein, die einen lokalen Zelltod zur Folge hat (Verdier, *et al.*, 2012). Möglicherweise induziert TAL10 eine Abwehrreaktion z.B. durch die zufällige Aktivierung eines Resistenzgens. In diesem Fall wird

OsSWEET14 zwar aktiviert, die weitere Besiedlung der Reispflanze und damit die Bildung wässriger Läsionen wird jedoch durch die Abwehrreaktion verhindert. Für die anderen analysierten TALEs (TAL14, 3/-3, 7/-7, -82 und 18) konnte keine signifikant erhöhte *OsSWEET14*-Expression gemessen werden (Abb. 2.5C). Für TAL1/-1, 11 und 17 muss die *OsSWEET14*-Transkriptmenge in zukünftigen Experimenten noch bestimmt werden. Im direkten Vergleich stimmen die Ergebnisse aus dem transienten β -Glucuronidase Assay mit den Ergebnissen aus dem natürlichen *Xoo*-Reis-System, überein. Alle dafür ausgewählten TALEs, die schon im transienten System in *N. benthamiana* eine effiziente Reporter-gen Aktivierung zeigten (TAL1/-1, 2/-2, 8 und 10), konnten das Suszeptibilitäts-gen auch im natürlichen Kontext aktivieren und so den Virulenzphänotyp eines nicht-virulenten *Xoo*-Stammes wiederherstellen. Im Gegensatz dazu waren TALEs, die im transienten System eine niedrige Aktivierung des Reporter-gens (im oben gezeigten Experiment zwischen 0 und 20%; Abb. 2.1B) hervorriefen, im natürlichen *Xoo*-Reis-System nicht in der Lage das Ziel-Gen zu aktivieren (TAL14, 3/-3, 7/-7 und -82; Abb. 2.5B,C). Diese TALEs sind im transienten System in *N. benthamiana* eventuell nur deshalb gering aktiv, weil die Gesamtkopienzahl der TALE- und vor allem Reporter-DNA pro Pflanzenzelle deutlich höher ist, als im natürlichen Kontext und so selbst bei ineffizienter TALE-Wirkung eine Reporteraktivierung gemessen werden kann. Für TAL11 und 17 kann noch keine eindeutige Aussage getroffen werden. Beide TALEs riefen im transienten System eine effiziente Aktivierung des Reporter-gens hervor, waren im natürlichen System jedoch nicht in der Lage den Virulenzphänotyp von BAI3 Δ *talC* wiederherzustellen. Dies könnte einerseits darin begründet sein, dass im natürlichen Reis-System die Bindestellen dieser TALEs durch Wirtsproteine blockiert sind, die im transienten System nicht vorhanden sind. Andererseits wäre es auch denkbar, dass beide TALEs zwar die Expression von *OsSWEET14* aktivieren, aber durch eine Verschiebung des Transkriptionsstarts die 5'UTR der mRNA verlängert wird. Diese verlängerte 5'UTR beinhaltet speziell im Fall von TAL17 zusätzliche ATG Motive, sodass bei der Translation Proteine entstehen können, die durch Verschiebung des Leserahmens nicht funktional sind. Selbst bei Beibehaltung des Leserahmens kann ein verfrühter Translationsstart zur Fusion zusätzlicher Aminosäuren am N-Terminus führen und das Signal zur Membranintegration von *OsSWEET14* blockieren. Dadurch würde das Protein seine Funktion als Transporter verlieren. Diese Fragestellung könnte durch die

Analyse der *OsSWEET14*-Transkriptmenge, sowie der Transkriptionsstarts in Anwesenheit von TAL11 oder TAL17 weiter aufgeklärt werden.

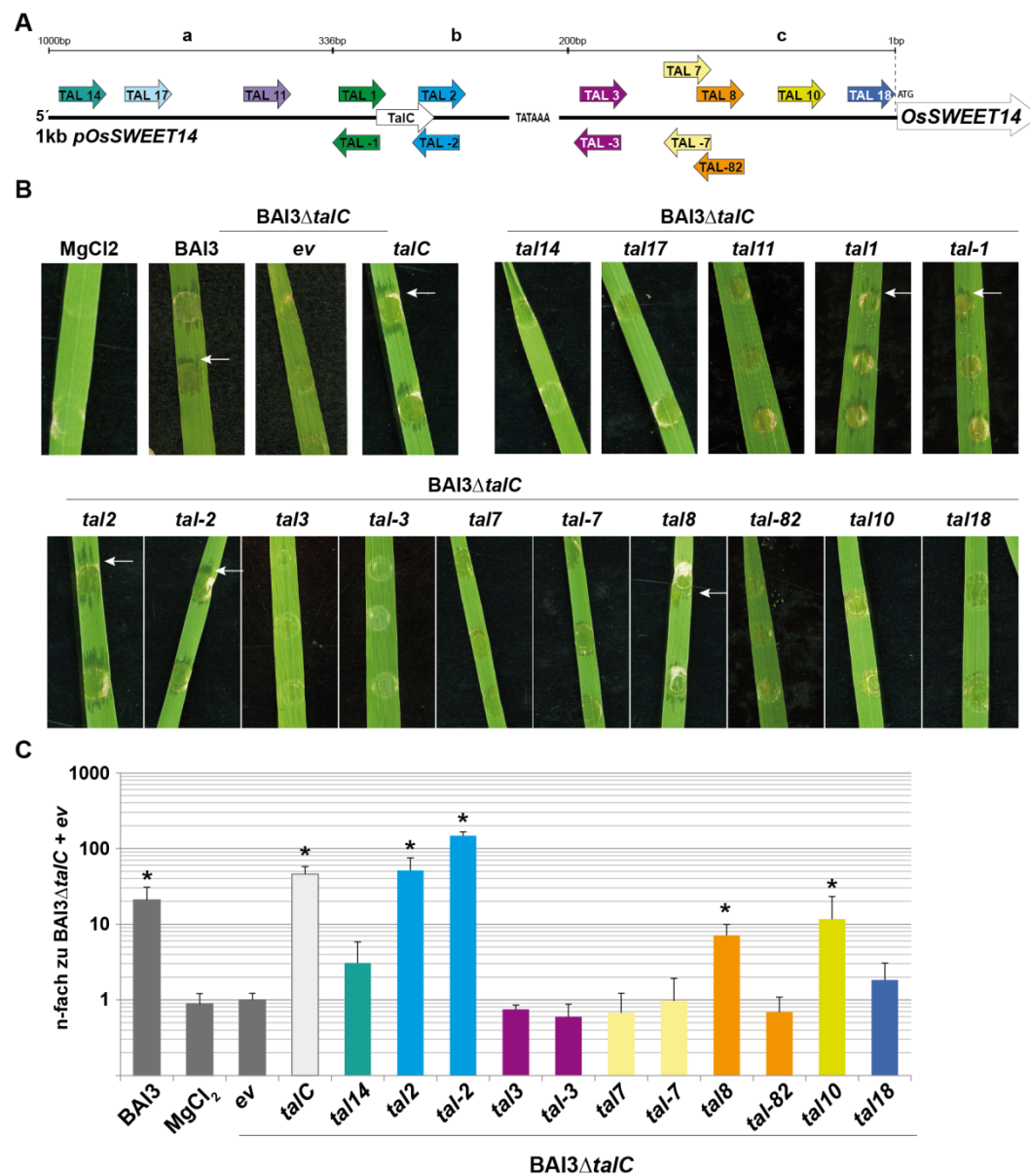


Abbildung 2.5 Funktionsanalyse ausgewählter *OsSWEET14*-TALEs im natürlichen *Xoo*-Reis System

A Schematische Übersicht über die Positionierung der verwendeten TALEs in den Bereichen **a,b,c** des *OsSWEET14*-Promotors. Die Position und Orientierung der künstlichen TALEs ist durch farbige Pfeile dargestellt. Die Position des ATG ist durch eine gestrichelte Linie gekennzeichnet. Die Abbildung ist nicht maßstabsgetreu. **B** Analyse der Virulenzphänotypen nach Inokulation verschiedener *Xoo*-Stämme. Die *Xoo*-Stämme BAI3, BAI3ΔtalC+ev, sowie BAI3ΔtalC mit natürlichen oder künstlichen TALEs wurden mittels Spritze und einer OD₆₀₀ von 0,5 in die susceptible Reislinie *Nipponbare* (*Oryza sativa* ssp *japonica*) inokuliert. Pro Stamm wurden je zwei Blätter mit je 4-6 Inokulationspunkten versehen. 5 Tage nach Inokulation wurden die Virulenzphänotypen, sichtbar als dunkelgrüne wässrige Läsionen (mit weißem Pfeil gekennzeichnet), um den Inokulationspunkt ausgewertet und fotografiert. Dargestellt sind jeweils repräsentative Ausschnitte der inokulierten Blätter. **C** Analyse des *OsSWEET14* Expressionslevels mittels quantitativer RT-PCR. Die *Xoo*-Stämme wurden wie unter **B** beschrieben in die susceptible Reislinie *Nipponbare* inokuliert. Pro Stamm wurde ein 5cm Bereich von je zwei Blättern inokuliert. 24h nach Inokulation wurde das Blattmaterial geerntet und zur Extraktion von Gesamt-RNA verwendet. Die Menge des *OsSWEET14* Transkripts wurde nach cDNA Synthese mittels quantitativer PCR bestimmt

(iQ5 Cycler von Biorad). Als Referenzgen wurde Aktin verwendet. Die Aktivierung des *OsSWEET14* Gens wurde als n-fach im Vergleich zu *BAI3 Δ talC+ev* berechnet und ist anhand einer logarithmischen Skala aufgetragen. Fehlerbalken repräsentieren die Standardabweichung von 3 unabhängigen biologischen Replikaten. Die signifikante Erhöhung der *OsSWEET14* Transkriptmenge im Vergleich zu *BAI3 Δ talC+ev* wurde mit Hilfe des t-Tests ($p \leq 0,05$) berechnet und mit einem * markiert.

2.5.4 Analyse des Einflusses von TALEs auf die Transkriptionsinitiation

Die vorhergehenden Ergebnisse zeigen, dass TALEs effizient die Transkription eines Gens aktivieren können, selbst wenn deren Bindestellen an verschiedensten Positionen im Promotor und sogar *reverse* im Gegenstrang liegen. Vorhergehende Studien zeigten, dass der Transkriptionsstart eines Gens TALE-abhängig verändert wird und in den betrachteten Fällen ca. 40-60 bp *downstream* der TALE-Bindestelle liegt (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Antony, *et al.*, 2010). Allerdings enthielten die dabei analysierten TALE-Bindestellen die TATA-Box. Die TATA-Box dirigiert den Transkriptionsstart, sodass dieser ca. 25 bis 30 Nukleotide *downstream* liegt (Gannon, *et al.*, 1979; Thomas & Chiang 2006). Daher ist nicht auszuschließen, dass die Verschiebung des Transkriptionsstarts im Vergleich zum natürlichen Transkriptionsstart lediglich mit der Blockierung der TATA-Box durch den TALE zusammenhängt. Weiterhin zeigte eine Studie anhand eines künstlichen Promotors in dem mehrere TALE-Bindestellen kombiniert wurden, dass die Mehrheit der analysierten TALEs unabhängig von ihrer Position den natürlichen Transkriptionsstart ca. 27bp *downstream* der TATA-Box zur Folge hat (Hummel, *et al.*, 2012). Allerdings enthielt dieser Promotor durch die Insertion mehrerer TALE-Bindestellen auch mehrere TATA-Elemente. Daher ist es möglich, dass zur Transkriptionsinitiation an diesem Promotor möglicherweise immer die dem ATG am nächsten gelegene TATA-Box verwendet wird. Mit Hilfe der im Rahmen dieser Arbeit erstellten *OsSWEET14*-TALEs können nun systematisch die Transkriptionsstarts in Anwesenheit verschieden positionierter TALEs analysiert werden, ohne dabei wichtige Promotorelemente zu verschieben oder zu inserieren. Interessant sind in diesem Zusammenhang vor allem die *reverse*-bindenden TALEs deren AD im Vergleich zu den *forward*-bindenden TALEs entgegengesetzt positioniert ist. Die Transkriptionsstartpunkte am *OsSWEET14*-Promotor wurden anhand des vorher bereits beschriebenen Reporterkonstruktes (1kb *OsSWEET14*-Promotor fusioniert an *uidA*; Abb. 2.1A) im *N. benthamiana* System mittels 5'RACE analysiert (Abb. 2.6A-D). Dazu wurden die TALEs TAL11, 1/-1, 2/-2, 3/-3, 9 und 10 ausgewählt und zusammen mit dem Reporterkonstrukt über *Agrobacterium* in *N. benthamiana* ko-transferiert (Abb. 2.6A). 24h nach Inokulation wurde das Blattmaterial geerntet und zur Extraktion von Gesamt-RNA verwendet. Nach

cDNA Synthese, wurden die zu analysierenden Transkripte mit Hilfe eines Reporter-gen-spezifischen Primers amplifiziert, zur weiteren Analyse in einen Vektor inseriert und anschließend bis zu 20 der erhaltenen Klone sequenziert.

Um zu überprüfen, ob die in *N. benthamiana* erhaltenen Transkriptionsstarts, mit denen im Reis-System vergleichbar sind, wurde zu Beginn der Transkriptionsstart am *OsSWEET14*-Promotor in Anwesenheit von AvrXa7 bestimmt, da dieser aus vorhergehenden Studien in Reis bereits bekannt war (Antony, *et al.*, 2010). Wie Abbildung 2.6B zeigt, stimmt der für AvrXa7 im transienten System bestimmte Transkriptionsstart, mit dem aus dem natürlichen Reis-System bekannten Start überein. Das transiente System in *N. benthamiana* kann somit zur weiteren Analyse der Transkriptionsstartpunkte verwendet werden. Der natürliche, TALE-unabhängige Startpunkt am *OsSWEET14*-Promotor konnte im transienten System nicht überprüft werden, da das verwendete Reporterkonstrukt keine Basalaktivität besitzt. Die ermittelten Transkriptionsstarts in Anwesenheit ausgewählter TALEs sind in Abbildung 2.6B-D dargestellt. Der natürliche in Reis bestimmte Transkriptionsstart am *OsSWEET14*-Promotor ist mit einem roten A gekennzeichnet (Antony, *et al.*, 2010; Abb. 2.6B). Das erste Nukleotid der ermittelten Transkriptionsstartpunkte ist in schwarz bzw. wenn es mit dem natürlichen Transkriptionsstartpunkt übereinstimmt, in rot angegeben. Zu beachten ist dabei, dass die Abstände der ermittelten Transkriptionsstarts zum 3'-Ende der Bindestellen für *forward*-TALEs bzw. zum 5'-Ende der Bindestellen für *reverse*-TALEs angegeben sind. Zuerst wurden die Transkriptionsstartpunkte für die künstlichen TALEs TAL3/-3, 9 und 10 aus Promotorfragment c bestimmt (Abb. 2.6C). Es konnte eine Verschiebung des Transkriptionsstarts am *OsSWEET14* Promotor in Anwesenheit der verschieden positionierten TALEs beobachtet werden. Die Mehrheit der Transkriptionsstarts lag bei allen vier analysierten TALEs zwischen 40 und 60bp *downstream* der Bindestelle. Dies entspricht einer Verschiebung des Transkriptionsstarts, die mit Daten früherer Analysen übereinstimmt (Kay, *et al.*, 2007; Römer, *et al.*, 2009; Antony, *et al.*, 2010). Interessanterweise stimmt auch die Mehrheit der Transkriptionsstarts in Anwesenheit von TAL3 bzw. TAL-3 überein, obwohl im Fall von TAL-3 die AD entgegengesetzt und somit an einer anderen Stelle im Promotor positioniert ist. Zusätzlich wurden für TAL-3 allerdings auch Transkriptionsstarts ermittelt, die mit 117 bzw. 138bp wesentlich weiter von der TALE-Bindestelle entfernt liegen. Anschließend wurden die Transkriptionsstartpunkte für die TALEs TAL11, 1/-1 und 2/-2 aus Promotorfragment a und b bestimmt (Abb. 2.6D). Hier ergab sich ein

weniger eindeutiges Bild. Für TAL2 und -2 wurden jeweils Startpunkte ermittelt, die 40-60bp *downstream* der TALE Ziel-DNA-Sequenz und damit überwiegend 3bp vor oder am natürlichen Transkriptionsstart liegen. Obwohl deren Bindestellen weiter *upstream* liegen wurden für TAL1 und -1 vergleichbare Startpunkte ermittelt, die ebenfalls entweder 3bp vor oder direkt am natürlichen Start liegen. Möglicherweise werden die Transkriptionsstarts in Anwesenheit von TALEs, die in einem bestimmten Bereich *upstream* der TATA-Box binden durch die TATA-Box dirigiert und sind somit unabhängig von der Position der TALE-Bindestelle. Die Mehrheit der für TAL11 bestimmten Startpunkte wiederum, liegt ca. 40-60bp *downstream* der TALE Ziel-DNA-Sequenz und wird somit vermutlich durch die Position des TALEs dirigiert (Abb. 2.6D). Zusätzlich dazu wurden für TAL11 Transkriptionsstarts ermittelt, die erneut 3bp vor dem natürlichen Startpunkt liegen. Dies deutet daraufhin, dass hier möglicherweise beide Effekte (die Position des TALEs und die Position der TATA-Box) den Transkriptionsstart beeinflussen. Um eine umfassendere Übersicht zu erhalten ist es notwendig weitere Transkriptionsstarts für TAL11, sowie Startpunkte für andere TALEs aus Promotorfragment **a** z.B. TAL17 oder 14 zu bestimmen. Zusammenfassend lässt sich jedoch vermerken, dass für die meisten TALEs individuelle Transkriptionsinitiationspunkte ermittelt werden konnten, die mit Daten früherer Publikationen übereinstimmen. Interessanterweise sind die ermittelten Startpunkte für *forward*- und *reverse*-bindende TALE-Paare (TAL1/-1, 2/-2 und 3/-3) oft vergleichbar, obwohl die AD, die als Plattform für die Bildung des Initiationskomplex dienen sollte, entgegengesetzt positioniert ist. Es ist jedoch möglich, dass es sich bei der Mehrheit der für TAL2/-2, sowie TAL1/-1 ermittelten Startpunkte um den TALE-unabhängigen, TATA-Box-dirigierten Startpunkt handelt, der aus Studien in Reis nur 3bp *downstream* liegt. Weiterführende Studien sollten daher in Reis nach natürlicher Infektion mit *Xanthomonas* durchgeführt werden, sodass TALE-abhängige Startpunkte besser mit dem natürlichen Startpunkt für *OsSWEET14* verglichen werden können.

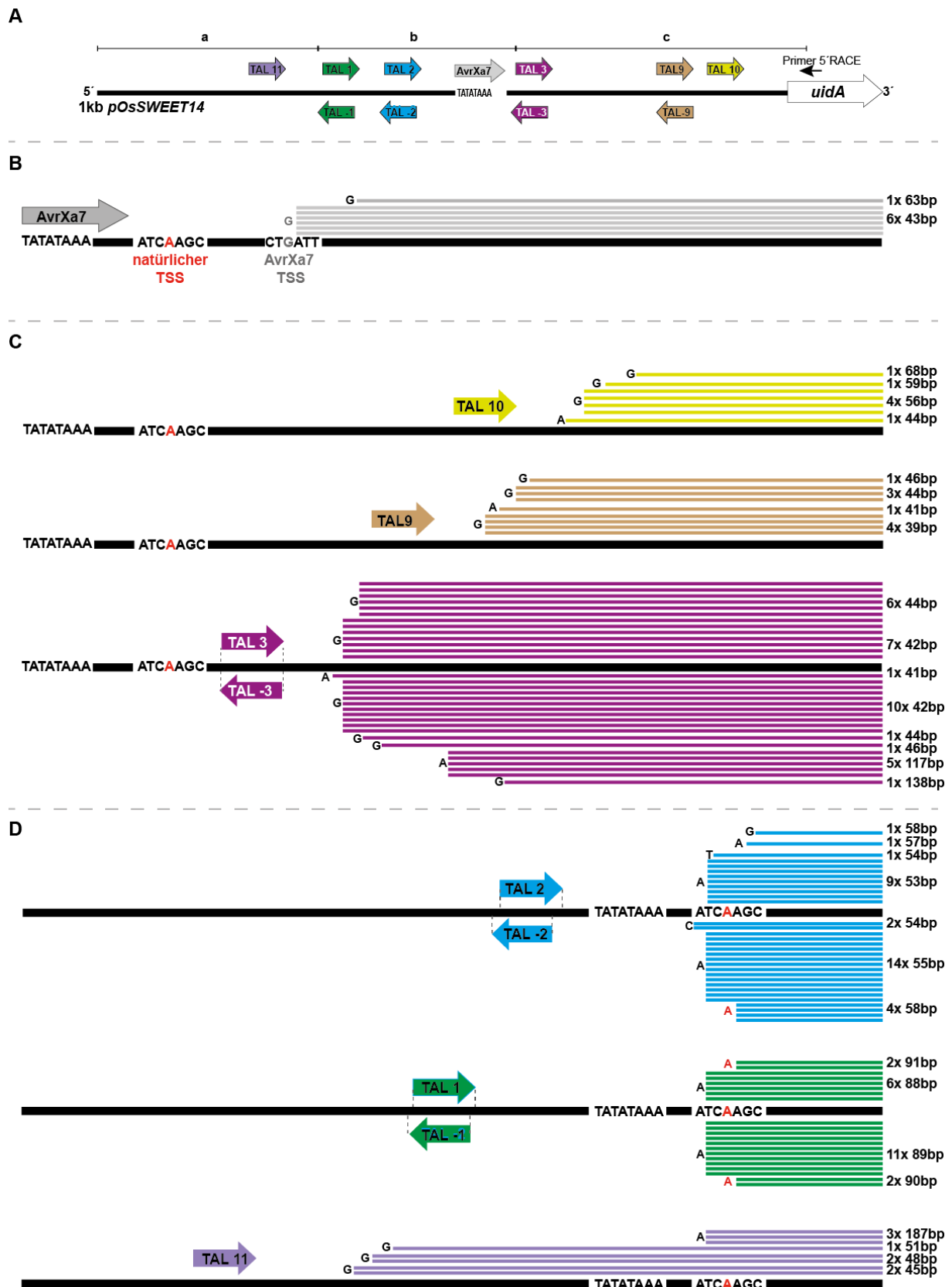


Abbildung 2.6 Transkriptionsstartpunkte in Anwesenheit verschiedener TALEs in *N. benthamiana*
 A Übersicht über die verwendeten TALEs aus den verschiedenen Promotorabschnitten des Reporterkonstruktes. Die Lage und Orientierung der TALEs ist mit Pfeilen angegeben. Die Lage der verwendeten 5'RACE Primers ist ebenfalls gekennzeichnet. Die Abbildung ist nicht maßstabsgetreu
 B Die Abstände der ermittelten Startpunkte zur TALE Bindestelle sind in bp angegeben. Der natürliche aus Reis bekannte Startpunkt am *OsSWEET14*-Promotor ist mit einem roten A gekennzeichnet, der

Ergebnisse

Startpunkt in Anwesenheit von AvrXa7 ist mit einem grauem G gekennzeichnet. Die im Rahmen dieser Arbeit ermittelten Startpunkte in Anwesenheit von AvrXa7 sind durch graue Linien markiert. Das jeweils erste Nukleotid der ermittelten Startpunkte ist ebenfalls angegeben. **C** Ermittelte Startpunkte in Anwesenheit der TALEs *downstream* der TATA-Box (im Promotorbereich) **c)** **D** Ermittelte Startpunkte in Anwesenheit der TALEs *upstream* der TATA-Box (Promotorbereich **a,b**).

3 Diskussion

3.1 Der erste *Ralstonia*-Code: Vorhersage einer Ziel-DNA-Sequenz für Brg11

Im Rahmen der Publikation 1 wurden erstmals die Spezifitäten für RVDs bestimmt, die auch in TALE-Homologen z.B. Brg11 aus *R. solanacearum* zu finden sind (Cunnac, *et al.*, 2004; Heuer, *et al.*, 2007). Mit Hilfe dieser neuen RVD-Spezifitäten konnten nun zum ersten Mal anhand der Brg11-RVD-Abfolge mögliche Ziel-DNA-Sequenzen vorhergesagt werden (Streubel 2011; Streubel *et al.*, 2012). Im Rahmen meiner Diplomarbeit wurde analysiert ob die vorhergesagten Ziel-DNA-Sequenzen durch die Brg11-*repeat*-Region gebunden werden. Dazu wurden verschiedene Reporterkonstrukte erstellt, die sich hauptsächlich in dem für den RVD HN abgeleiteten Ziel-Nukleotid (ausschließlich Adenin oder ausschließlich Guanin bzw. beide im Wechsel) sowie in der Identität des initialen Nukleotids, welches der Ziel-DNA-Sequenz vorangestellt ist, unterschieden (A, C, G oder T; Boch, *et al.*, 2009; Römer, *et al.*, 2009; Streubel 2011). Diese Ziel-DNA-Sequenzen wurden an das 74bp-Fragment des *Bs4*-Promotors fusioniert und vor das promotorlose *uidA*-Reporter gen inseriert. Die Analyse ergab jedoch, dass Brg11 nicht in der Lage war eines der Reporterkonstrukte zu aktivieren (Streubel 2011). In nachfolgenden Studien, die ebenfalls anhand des *Ralstonia TALE-like* (RTL) Proteins Brg11 durchgeführt wurden, konnten die in Publikation 1 ermittelten Spezifitäten jedoch bestätigt und damit die Aktivität von Brg11 anhand eines Reporterkonstruktes gezeigt werden (de Lange, *et al.*, 2013). Zusätzlich konnte für Brg11 die Notwendigkeit eines initialen Guanins anstatt des für *Xanthomonas*-TALEs üblichen Thymins vor der eigentlichen Ziel-DNA-Sequenz beschreiben werden (de Lange, *et al.*, 2013). Aus diesen Ergebnissen können mögliche Gründe abgeleitet werden, die erklären warum Brg11 die im Rahmen meiner Diplomarbeit verwendeten Reporterkonstrukte nicht aktivierte. Durch die Notwendigkeit des initialen Guanins kam nur eines der verwendeten Reporterkonstrukte (Box1G, mit initialem Guanin) als Ziel-DNA-Sequenz für Brg11 in Frage (Abb. 3.1). Trotzdem war Brg11 in diesen Studien nicht in der Lage dieses Reporterkonstrukt zu aktivieren. Möglicherweise ist die fehlende Brg11-Aktivität darin begründet, dass im Box1G-Reporterkonstrukt für den RVD HN ausschließlich Adenin als Ziel-Nukleotid gewählt wurde (Abb. 3.1). Wie für den RVD NN bereits bekannt, ist Asparagin an Position 13 in Kombination mit dem Ziel-Nukleotid Adenin ein schwaches RVD (s. Abschnitt 3.4). Daher wäre es denkbar, dass die Häufung der schwachen HN-Adenin-Kombinationen zusammen mit weiteren schwachen Interaktionen (NK-G; NG-T) besonders im N-terminalen Bereich der Brg11

repeat-Region die Gesamtaktivität von Brg11 reduziert (Abb. 3.1). Ein weiterer Unterschied der in beiden Studien verwendeten Reporterkonstrukte ist an Position 1 (G vs. A) der Ziel-DNA-Sequenz zu finden, die durch den RVD NT gebunden wird (Abb. 3.1). Da jedoch in beiden Studien die Spezifität von NT für A und G gezeigt wurde, sollte dieser Unterschied keinen Einfluss auf die Bindung des TALEs haben (Streubel 2011; Streubel *et al.*, 2012; de Lange, *et al.*, 2013). Weiterhin unterscheiden sich die beiden Reporterkonstrukte im verwendeten Promotor (*Bs4* vs. *Bs3*; Abb. 3.1). Allerdings konnte bereits für beide Promotorfragmente gezeigt werden, dass sie TALE-abhängig aktivierbar sind, daher sollte auch dieser Unterschied keinen Einfluss auf die Brg11 Aktivität ausüben (Boch, *et al.*, 2009; Römer, *et al.*, 2007, Römer, *et al.*, 2009). Dennoch kann zusammenfassend gesagt werden, dass anhand der in Publikation 1 bestimmten RVD-Spezifitäten bereits erstmals eine Ziel-DNA-Sequenz für den *Ralstonia*-TALE Brg11 vorhergesagt wurde, die in nachfolgenden Studien bestätigt werden konnte. Dies bildet den Ausgangspunkt um in zukünftigen Studien die Rolle der RTLs und deren Ziel-Gene für die Virulenz von *R. solanacearum* genauer zu untersuchen.

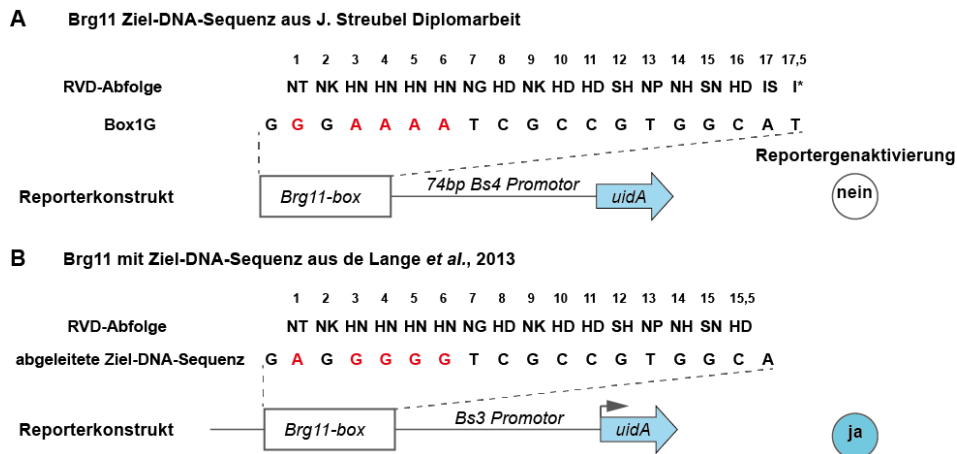


Abbildung 3.1 Vergleich der in meiner Diplomarbeit und de Lange *et al.*, 2013 verwendeten Brg11 Reporterkonstrukte

A Dargestellt ist die RVD-Abfolge der in meiner Diplomarbeit angenommenen *repeat*-Region von Brg11 mit 17,5 *repeats*. Die letzten zwei *repeats* IS und I* sind weniger hoch konserviert (Streubel 2011; Streubel *et al.*, 2013; de Lange *et al.*, 2013). Da bisher unklar ist ob diese *repeats* an der Bindung teilnehmen wurden mögliche Zielnukleotide im Reporterkonstrukt angepasst. Darunter ist das Reporterkonstrukt Box1G dargestellt. Unterschiede zur Ziel-DNA-Sequenz des in **B** verwendeten Reporterkonstrukts sind rot markiert. **B** Dargestellt ist die RVD-Abfolge der in de Lange 2013 angenommenen *repeat*-Region von Brg11, die zwei *repeats* weniger enthält als in **A**. Da bisher unklar ist ob der weniger konservierte *repeat* IS an der DNA-Bindung teilnimmt wurde zusätzlich das IS-Ziel-Nukleotid an Position 17 in die Ziel-DNA-Sequenz inseriert. Das in dieser Studie verwendete Reporterkonstrukt konnte Brg11-abhängig aktiviert werden.

3.2 NH - ein neues Guanin-spezifisches RVD

Interessanterweise konnten unter den in Publikation 1 analysierten RVDs auch vier Guanin-spezifische RVDs identifiziert werden (NK, NH, SN, SH). Für die Erstellung künstlicher TALEs wurde zuvor der RVD NN zur Bindung an Guanin verwendet. NN bindet jedoch neben Guanin auch Adenin (Boch, *et al.*, 2009; Moscou & Bogdanove 2009, Deng *et al.*, 2014). Deshalb könnte durch die Verwendung des NN RVD die *off-target* Rate für TALE-basierte Modifikationsenzyme im entsprechenden Ziel-Genom erhöht werden. Die Entdeckung eines neuen RVDs, das ausschließlich Guanin-spezifisch ist, könnte somit die Spezifität solcher TALE-Derivate im Ziel-Genom signifikant verbessern. NK wurde bereits 2010 von Morbitzer *et al.*, als Guanin-spezifisch beschrieben (Morbitzer, *et al.*, 2010). Allerdings konnte in Publikation 1 gezeigt werden, dass NK ein schwacher RVD ist, welcher besonders in Kombination mit anderen schwachen RVDs (z.B. NI, NG) die Gesamtaktivität eines TALEs signifikant verringern kann (s. Abschnitt 3.4). Da aber für die biotechnologische Anwendung hochspezifische und hocheffiziente TALE-Derivate benötigt werden, sollte NK nur in TALEs verwendet werden, die zusätzlich eine gewisse Anzahl gleichmäßig positionierter starker RVDs in der *repeat*-Region aufweisen (Streubel, *et al.*, 2012). SN und SH sind ebenfalls schwache RVDs und damit nicht zur Erstellung effizienter TALE-Derivate geeignet (Streubel *et al.*, 2013). Zudem ist mittlerweile bekannt, dass lediglich die zweite Aminosäure im RVD die Nukleotid Spezifität bestimmt, sodass für SN davon ausgegangen werden kann, dass er vergleichbar zu NN ebenfalls Adenin und Guanin binden kann (Deng, *et al.*, 2012a; Mak, *et al.*, 2012). Im Gegensatz dazu erwies sich der erstmals in Publikation 1 beschriebene RVD-Typ NH als vielversprechende Alternative zu NN. Anhand der in Publikation 1 beschriebenen Ergebnisse ist NH ein intermediär-starker RVD, der wesentlich Guanin-spezifischer als NN ist und die Gesamtaktivität des TALEs positiv beeinflusst (Streubel, *et al.*, 2012). Anhand von TALE-Aktivatoren und TALE-Nukleasen bestätigten nachfolgende Studien die in Publikation 1 beschriebene höhere Guanin-Spezifität von NH im Vergleich zu NN und die höhere Effizienz von NH im Vergleich zu NK (Cong *et al.*, 2012). Daher wurde der erstmals in Publikation 1 beschriebene RVD NH als neues hochspezifisches RVD zur Bindung von Guanin in das Basis-Repertoire zur Erstellung künstlicher TALE-Derivate aufgenommen (Booher & Bogdanove 2014; Richter, *et al.*, 2014; Yang, *et al.*, 2014).

3.3 Der RVD N* ermöglicht TALEs methylierte Ziel-DNA-Sequenzen zu binden

In Publikation 1 konnte die für den RVD N* vorhergesagte Spezifität (für Cytosin und Thymin) nun erstmals experimentell bestätigt werden (Moscou & Bogdanove 2009; Streubel, *et al.*, 2012). Die ermittelte Spezifität von N* für Pyrimidine wies auf die Möglichkeit hin, dass N* zusätzlich methylierte Cytosine erkennen könnte. Diese erlauben, ähnlich zu Thymin, die Ausbildung von van-der-Waals-Interaktionen (mit Glycin an Position 14 im *repeat*) in der TALE-DNA Interaktionsfläche (Deng, *et al.*, 2012b; Valton, *et al.*, 2012). Eine solche Erkennung methylierter Cytosine konnte für N* sowie auch H* und NG anschließend gezeigt werden (Deng, *et al.*, 2012b; Valton, *et al.*, 2012). Diese Eigenschaft ist besonders daher interessant, da TALEs mit einem HD-RVD an der betreffenden Position, methylierte Cytosine in der Ziel-DNA-Sequenz nicht binden können (Deng, *et al.*, 2012b; Valton, *et al.*, 2012). Besonders bei unbekanntem Methylierungsstatus ist die Verwendung von N*-RVDs statt HD-RVDs (spezifisch für Cytosin) oder NG-RVDs (spezifisch für Thymin und methyliertes Cytosin) vorteilhaft, da diese neben methylierten Cytosinen auch unmethylierte Cytosine erkennen.

RVDs die methylierte Cytosine erkennen, können eingesetzt werden, um einerseits die Sensitivität von TALEs gegenüber potentiell methylierten Sequenzen zu umgehen und andererseits den Methylierungsstatus einzelner Cytosine zu analysieren. (Bultmann, *et al.*, 2012; Deng, *et al.*, 2012b; Valton, *et al.*, 2012; Kubik, *et al.*, 2014, Kubik & Summerer 2015).

3.4 Effizienz - eine neue Eigenschaft für TALE-RVDs

Der TALE "ein-RVD-bindet-ein-Nukleotid"-Code ist einfach und geradlinig. Im Rahmen der Publikation 1 wurde jedoch erstmals ein Phänomen beschrieben, dass dem TALE-RVD-Code zusätzliche Komplexität verleiht. Dabei konnte ein Einfluss der RVD-Komposition auf die Gesamtaktivität des TALEs gezeigt werden, der aus vorhergehenden Studien nicht erkennbar war. Besonders TALEs die eine bestimmte Anzahl von beispielsweise NI, NG oder NK RVDs in Folge enthielten oder ausschließlich aus einer Mischung dieser RVDs zusammengesetzt waren, erzielten lediglich eine sehr geringe oder sogar gar keine Aktivierung des zugehörigen Reporters (Streubel, *et al.*, 2012). Erst die Integration von HD oder NN und in den meisten Fällen auch NH RVDs, konnte die Aktivität dieser TALEs wiederherstellen. Dies ließ erstmals vermuten, dass RVDs unterschiedlich effizient sind und die Gesamtaktivität eines TALEs somit auf der ausgewogenen Kombination starker (HD und NN), intermediär

starker (NH) oder schwacher (NI, NG, NK) RVDs beruht. Diese neue RVD-Eigenschaft (Effizienz) war unerwartet und wurde erstmals in Publikation 1 beschrieben. Nachfolgende Studien konnten den Effekt der RVD-Effizienz beispielsweise im Vergleich verschiedener Guanin-spezifischer RVDs (NN, NK, HN und NH) anhand von TALE-Nukleasen oder TALE-Aktivatoren im humanen System bzw. in Zebrafisch ebenfalls beobachten (Huang, *et al.*, 2011; Christian, *et al.*, 2012; Cong, *et al.*, 2012; Meckler, *et al.*, 2013; Yang, *et al.*, 2014). In Übereinstimmung mit den Daten aus Publikation 1, beschrieben diese Studien ebenfalls, dass NK im Vergleich zu HN und NN zwar spezifischer für Guanin ist, jedoch zu einer geringeren Gesamtaktivität des TALEs führt (Christian, *et al.*, 2012; Cong, *et al.*, 2012; Meckler, *et al.*, 2013; Yang, *et al.*, 2014). Im Gegensatz dazu erzielen die RVDs HN, NN und NH eine hohe TALE-Gesamtaktivität, wobei NH die höchste Guanin-Spezifität aufweist (Christian, *et al.*, 2012; Cong, *et al.*, 2012; Meckler, *et al.*, 2013; Yang, *et al.*, 2014). Anhand verfügbarer Strukturdaten, die die Interaktion von TALEs mit der Ziel-DNA zeigen, können nicht nur Einblicke in die Grundlagen der RVD-vermittelten DNA-Spezifität gewonnen werden, sondern auch mögliche Erklärungen für die in Publikation 1 beschriebene neue RVD-Eigenschaft, die Effizienz, abgeleitet werden. Dabei zeigt sich, dass die unterschiedliche Effizienz von RVDs vermutlich in der Interaktion der 13. Aminosäure mit dem Ziel-Nukleotid begründet ist (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Deng, *et al.*, 2014). Die hohe Effizienz der starken RVDs geht meist mit der Ausbildung von stabilisierenden Wasserstoffbrücken zwischen RVD (z.B. NN, NH, HD) und Ziel-Nukleotid einher, wogegen diese in der Interaktion zwischen schwachen RVDs und dem Ziel-Nukleotid häufig fehlen (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Deng, *et al.*, 2014). Beispielsweise bildet die Aminosäure Aspartat im starken RVD HD ausschließlich mit der Aminogruppe des Cytosin eine stabilisierende Wasserstoffbrücke, wodurch dieses Nukleotid in der Ziel-DNA-Sequenz gegenüber allen anderen Nukleotiden favorisiert wird (Deng, *et al.*, 2012a; Wicky, *et al.*, 2013; Deng, *et al.*, 2014). Asparagin an Position 13 (z.B. im RVD NN) bildet stabilisierende Wasserstoffbrücken zum O6 oder N7 von Guanin bzw. ausschließlich zum N7 von Adenin (Deng, *et al.*, 2014). Das O6 des Guanin weist eine negativere Ladung auf weshalb diese Wasserstoffbrückenbindung zu Asparagin vermutlich stärker ist als die Wasserstoffbrückenbindung die von Asparagin zum N7 von Adenin ausgebildet wird (Deng, *et al.*, 2014). Dies könnte auch die Beobachtung erklären, dass NN nur in Kombination mit Guanin und nicht mit Adenin in der Ziel-DNA-Sequenz ein starker

RVD ist (Streubel, *et al.*, 2012). Die hohe Spezifität von NH für Guanin ist vermutlich darin begründet, dass der Imidazol-Ring von Histidin ausschließlich zum N7 von Guanin eine Wasserstoffbrücke ausbildet, wodurch außerdem die hohe Effizienz dieser Interaktion erklärt werden kann (Deng, *et al.*, 2014). Im Gegensatz zu NN bildet NH keine Wasserstoffbrücke zu Adenin aus (Deng, *et al.*, 2014). Diese Unterschiede in der Interaktion von NH mit A und G im Vergleich zu NN mit A und G erklären die hohe Guanin-Spezifität von NH. Der RVD NK ist ein Ausnahmefall. NK ist zwar ebenfalls in der Lage eine Wasserstoffbrücke zum O6 von Guanin auszubilden, wodurch dessen Spezifität erklärt wird (Deng, *et al.*, 2014). Allerdings fügt sich die lange Seitenkette des Lysin vermutlich nicht perfekt in die RVD-DNA Schnittstelle ein und wirkt dort eher störend (Deng, *et al.*, 2012a; Wicky, *et al.*, 2013; Deng, *et al.*, 2014). Daraus resultierend können zu viele solcher störenden RVDs die Gesamtaktivität des TALEs verringern, wodurch die geringe Effizienz dieses RVD erklärt werden könnte. Die geringe Effizienz der RVDs NI und NG dagegen, könnte darin begründet sein, dass beide RVDs keine stabilisierenden Wasserstoffbrücken zum Ziel-Nukleotid ausbilden (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Wicky, *et al.*, 2013; Deng, *et al.*, 2014). Beispielsweise scheint die Spezifität von Isoleucin an Position 13 eher darauf zu beruhen, dass im Vergleich zu Cytosin, Thymin und Guanin nur in Kombination mit Adenin keine sterischen Kollisionen auftreten, wodurch dieses Nukleotid toleriert wird (Wicky, *et al.*, 2013; Deng, *et al.*, 2014). Die Spezifität von NG für Thymin ist darin begründet, dass an Position 13 nur Aminosäuren mit einer kleinen Seitenkette (z.B. Glycin in NG) oder RVDs mit fehlender 13. Aminosäure (z.B. N*) die Methylgruppe des Thymins tolerieren können (Deng, *et al.*, 2012a, Deng, *et al.*, 2014). Da allerdings auch in diesen Interaktionen keine Wasserstoffbrücken sondern nur van-der-Waals-Interaktionen zu Thymin ausgebildet werden ist der Beitrag dieser RVDs zur Gesamtaktivität eher schwach (Mak, *et al.*, 2012; Deng, *et al.*, 2014).

NG ist der bisher einzige RVD für den in einer nachfolgenden Studie widersprüchliche Ergebnisse beschrieben wurden. Darin wurde im Vergleich zur in Publikation 1 beschriebenen schwachen Aktivität, NG als der am stärksten bindende RVD definiert (Meckler, *et al.*, 2013). In dieser Studie wurde die DNA-Bindung verschiedener künstlicher TALEs mit 17,5 *repeats* analysiert. Das Grundgerüst dieser TALEs enthielt an jeder zweiten RVD-Position (P2, P4, P6 usw. bis P16) entweder NI, NG, NK oder HD. Insgesamt kam somit jedes Grundgerüst-RVD zweimal in der *repeat*-Region vor. Ausgehend von diesem Grundgerüst wurden drei Sets erstellt, in denen die Abfolge der

Grundgerüst RVDs verändert wurde (z.B. der Grundgerüst RVD HD: in Set1 auf P4 und P16, in Set2 auf P6 und P12). Anschließend wurden alle verbliebenen "Gast"-Positionen (P1, P3, P5 usw. bis P17,5) durch ausschließlich HD, NG, NI, NK oder NN RVDs besetzt. Zu allen verwendeten künstlichen TALEs wurden perfekt passende Ziel-DNA-Sequenzen abgeleitet und die Bindung der TALEs an die jeweils passende Ziel-Sequenz mittels *in vitro* DNA-Bindungsstudien (*electrophoretic mobility shift assay*) analysiert. Anhand der erhaltenen Signal-Stärke des TALE-DNA-Komplexes auf einem Röntgenfilm, wurden anschließend Dissoziationskonstanten für jeden TALE berechnet. Aus den erhaltenen Dissoziationskonstanten wurde der "relative" Beitrag verschiedener RVDs zur Bindung an die perfekte DNA-Sequenz bewertet und NG dabei als bester DNA-bindender RVD definiert ($NG (1) > NN (0,18) \sim HD (0,16) \gg NI (0,0016) > NK (0,00016)$; Meckler, *et al.*, 2013). Da jedoch bekannt ist, dass NG keine stabilisierenden Wasserstoffbrücken zu Thymin in der Ziel-DNA-Sequenz ausbildet ist unklar worauf diese hohe Bindung beruht (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Deng, *et al.*, 2014). Eine mögliche Hypothese wäre, dass NG die Gesamt-DNA-Bindung des TALEs am wenigsten stört. Geht man davon aus, dass sich die Bindung eines TALEs an der perfekten Ziel-DNA-Sequenz aus der Bindung des N-Terminus (Gao, *et al.*, 2012), des *repeat*-Rückgrats (Aminosäuren G14, G15, K16, Q17 im *repeat*; Wicky, *et al.*, 2013; Deng, *et al.*, 2014) und den stabilisierenden positiven Interaktionen des RVDs mit dem Zielnukleotid z.B. durch Wasserstoffbrücken ergibt (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Deng, *et al.*, 2014), dann sollte der generelle Beitrag von NG zur Bindung sehr gering sein, da es keine Wasserstoffbrückenbindungen zu Thymin ausbildet. Allerdings fügt sich das im Verhältnis zu anderen Aminosäuren relativ kleine Glycin an Position 13 gut in die RVD-DNA-Schnittstelle ein. Dies könnte Spannungen vermeiden, sodass die Gesamt-Bindung nicht negativ beeinflusst wird. Zusätzlich wird die DNA durch die fehlenden stabilisierenden Bindungen eventuell nicht zu fest bzw. zu starr zum TALE herangezogen, sondern lokal an den Positionen der NG-RVDs "locker" gehalten. Dadurch könnten möglicherweise benachbarten störenden Einflüssen z.B. die von Lysin im RVD NK mehr Raum geboten werden und somit ausgeglichen werden. Dadurch verhilft NG dem TALE Spannungen zu verhindern und so seine hohe Bindungsaffinität beizubehalten. Die in diesem Fall gemessene hohe DNA-Gesamtaffinität ist also nicht auf die starke Bindung von NG an die DNA, sondern vielmehr auf den großen Beitrag der anderen Faktoren zurückzuführen (z.B. den Wasserstoffbrücken zwischen HD und Cytosin im Grundgerüst) der nicht durch NG gestört wird. Die TALEs mit den "Gast"-

RVDs NN und HD zeigen eine ähnliche im Vergleich zu NG ca. um Faktor 10 verringerte Gesamt-Affinität zur perfekt passenden Ziel-DNA-Sequenz (Meckler, *et al.*, 2013). Dies könnte einerseits darauf beruhen, dass HD und NN zwar durch die Ausbildung von Wasserstoffbrücken zur generellen DNA-Bindung beitragen, allerdings könnten vor allem im Fall des negativ geladenen Aspartats zu viele HD-RVDs in Folge zu Spannungen in der TALE-Bindung an die ebenfalls negativ geladene DNA führen. Dadurch könnte die Gesamtbindung reduziert werden. Andererseits könnte die starke Bindung von HD auch zu einer zu starren TALE-Konformation führen, die sich dann negativ auf die Gesamtbindung und vermutlich daraus resultierend auch auf die Gesamtaktivität auswirken könnte. Ein ähnlicher Effekt konnte auch für die TALE-abhängige Genaktivierung in Publikation 1 beobachtet werden. Dabei reduzierte die steigende Anzahl von HD-RVDs in TALEs mit bestimmter RVD-Komposition auch die Gesamtaktivität des TALEs (Streubel *et al.*, 2012). Die "Gast"-RVDs NI bzw. NK zeigen die geringste Gesamt-Affinität (Faktor 100 bzw. 1000; Meckler, *et al.*, 2013). NI kann ebenfalls keine Wasserstoffbrücken zu Adenin ausbilden. Im Gegensatz zu NG könnte das Isoleucin an Position 13 jedoch erneut Spannungen erzeugen, die möglicherweise die Gesamt-Bindung des TALEs reduzieren (Deng, *et al.*, 2014). NK wirkt wie bereits beschrieben vermutlich störend, wodurch die Gesamt-Bindung signifikant verringert wird. Interessanterweise spiegelt sich die in Meckler *et al.*, 2013 beschriebene hohe Bindungsaffinität der NG-TALEs nicht in der Aktivierung des Reportergens wieder. Besonders beim Vergleich der drei verwendeten TALE Sets wird deutlich, dass die Aktivität der NG-"Gast"-RVD-TALEs lediglich mit der Position der starken HD RVDs im Grundgerüst zusammenhängt. In Set1 und Set2 sind die zwei Grundgerüst-HD-RVDs relativ gleichmäßig in der *repeat*-Region verteilt, sodass weder am Ende noch am Beginn der *repeat*-Region zu viele schwache RVDs in Folge auftreten (Streubel, *et al.*, 2013). Daher weisen beide Sets eine hohe, sogar mit HD-"Gast"-RVD-TALEs vergleichbare Aktivität auf (Meckler, *et al.*, 2013). In Set3 wird die Position der zwei Grundgerüst-HD-RVDs jedoch auf P2 und P8 verschoben, sodass bis zum Ende der *repeat*-Region eine Abfolge von 10 schwachen RVDs entsteht (NI, NG, NK). Durch diese lange Abfolge schwacher RVDs, die wie oben beschrieben durch die Anhäufung der NG-RVDs möglicherweise zu einer etwas "lockeren" DNA-Bindung führen, fehlt ein Anker im C-terminalen Bereich der *repeat*-Region, sodass die Aktivität dieses TALEs im Vergleich zu Set1 und 2 signifikant reduziert wird (Streubel, *et al.*, 2012; Meckler, *et al.*, 2013). Dies zeigt, dass NG allein keinen positiven Einfluss auf

die Gesamtaktivität des TALEs hat, sondern diese nur auf das Vorhandensein gleichmäßig positionierter starker RVDs z.B. HD zurückzuführen ist. Diese Auswertung der in Meckler *et al.*, 2013 gezeigten Daten bringt die Rolle von NG auch mit den in Publikation 1 gezeigten Daten wieder in Einklang, die ebenfalls ergaben dass NG solange er von starken RVDs umgeben ist durchaus zu einer sehr hohen Genaktivierung führen kann (Streubel, *et al.*, 2012).

Grundlegend ist klar, zur TALE-abhängigen Genaktivierung ist die effiziente Bindung des TALEs an die passende Ziel-DNA-Sequenz notwendig. Trotzdem sind die genauen Zusammenhänge zwischen der TALE-DNA-Bindung und der biologischen Genaktivierungsfunktion von TALEs noch nicht vollständig aufgeklärt. Jedoch konnte aus den in Publikation 1 gewonnenen Erkenntnissen in Kombination mit anderen Studien zumindest die Hypothese entwickelt werden, dass für die starke DNA-Bindung von TALEs an die perfekte Ziel-DNA-Sequenz wenig störende RVDs z.B. NG durchaus vorteilhaft sein können, solange auch eine gewisse Anzahl starker RVDs vorhanden ist. Dabei ist bisher allerdings noch nicht bekannt, wo genau der Schwellenwert der DNA-Bindung liegen muss, um eine effiziente Genaktivierung zu erzeugen. Im Vergleich dazu scheint für die effiziente Genaktivierung hauptsächlich der positive Beitrag der starken RVDs und vor allem die gleichmäßige Positionierung dieser in der *repeat*-Region ausschlaggebend zu sein. Diese fördern möglicherweise das "einrasten" der gesamten *repeat*-Region mit der passenden Ziel-DNA-Sequenz und erlauben so die Genaktivierung. Die hier beschriebene Hypothese bezieht sich ausschließlich auf die Kombination eines TALEs mit der perfekt passenden Ziel-DNA-Sequenz. Weitere durch z.B. RVD-DNA Fehlpaarungen ausgelöste Auswirkungen auf die DNA-Bindung oder Genaktivierung wurden dabei nicht mit einbezogen.

3.5 Die Entschlüsselung aller 400 möglichen RVD-DNA-Spezifitäten

Seit bekannt ist, dass hauptsächlich die zweite Aminosäure im RVD für die Nukleotid-Spezifität verantwortlich ist, können die Spezifitäten für alle RVDs die eine bereits charakterisierte Aminosäure enthalten vorhergesagt werden (Deng, *et al.*, 2012a, Mak, *et al.*, 2012). Dennoch entschlüsselten zwei unabhängige Studien systemtisch die Spezifitäten aller 400 möglichen RVD-Aminosäurekombinationen (Yang, *et al.*, 2014; Juillerat, *et al.*, 2015). Yang *et al.*, 2014 nutzte dafür einen TALE-Aktivator mit 14,5 *repeats*, der in der Mitte der *repeat*-Region dreimal (Position 7-9) das zu untersuchende RVD in Folge enthielt (Yang, *et al.*, 2014). Aus den Ergebnissen dieser ersten Analysen

wurden einige vielversprechende RVD-Kombinationen mit möglichst hoher Spezifität für eine genauere Untersuchung ausgewählt. Darunter waren als Adenin-spezifisch neben NI weitere Kombinationen z.B. KI und RI. Als Cytosin-spezifisch wurden die beiden schon bekannten RVDs HD und ND ausgewählt. Unter den Guanin-spezifischen RVDs waren die bereits bekannten NH, NK, NN und HN, sowie einige neue Kombinationen z.B. RN, QN, RQ, FR. Und als Thymin-spezifisch wurde das bekannte NG sowie z.B. RG, KA, CA und YA ausgewählt. Zur genaueren Analyse der Effizienz dieser ausgewählten RVDs wurde vergleichbar mit dem in Publikation 1 verwendeten System zusätzlich zu den anfänglich verwendeten TALEs mit dreimal demselben RVD, TALEs mit einer Abfolge von sechsmal und zwölfmal dem gleichen RVD verwendet (Streubel, *et al.*, 2012; Yang, *et al.*, 2014). Interessanterweise konnte dabei festgestellt werden, dass keines der ausgewählten bisher noch nicht beschriebenen RVDs spezifischer oder effizienter als die bereits bekannten Standard-*repeats* HD, NN, NH, NG und NI war (Boch, *et al.*, 2009; Streubel, *et al.*, 2012). Ähnliche Ergebnisse wurden auch anhand der systematischen Analyse aller RVD-Kombinationen in TALE-Nukleasen (TALEN) gezeigt (Juillerat, *et al.*, 2015). Hier wurde lediglich die Alternative MI für die spezifische Bindung von Adenin vorgeschlagen. In Yang *et al.*, 2014 zeigte MI jedoch keine Aktivierung des Reporters (Yang, *et al.*, 2014). Dieser Unterschied könnte darin begründet sein, dass TALEN als Paar funktionieren und somit vermutlich mehr Toleranz gegenüber schwachen RVD-Kombinationen zeigen, solange zumindest einer der beiden TALENs effizient DNA bindet. Generell zeigten beide Studien außerdem, dass für beide Positionen im RVD nur bestimmte Aminosäuren erlaubt sind. Beispielsweise sind an Position 2 im RVD die Aminosäuren W, Y, E, F, K, M selten bzw. gar nicht funktional (Yang, *et al.*, 2014; Juillerat, *et al.*, 2015). Vermutlich weil sie große Seitenketten enthalten und somit in der RVD-DNA Interaktion eher stören (Mak *et al.*, 2012; Wicky *et al.*, 2013; Deng *et al.*, 2014). An Position 1 im RVD werden hauptsächlich die auch natürlicherweise häufigen Aminosäuren H und N favorisiert und zusätzlich noch weitere bisher nicht beschriebenen Aminosäuren z.B. K und R akzeptiert (Yang, *et al.*, 2014; Juillerat, *et al.*, 2015). Möglicherweise sind diese Aminosäuren durch die Vermittlung von *intra-repeat* Interaktionen besonders zur Stabilisierung des RVD-*loops* geeignet (Deng, *et al.*, 2012a; Mak, *et al.*, 2012; Deng, *et al.*, 2014). Im Gegensatz dazu sind auch an Position 1 im RVD Aminosäuren mit großen Seitenketten z.B. W und P, die vermutlich stören selten aktiv. (Yang *et al.*, 2014; Juillerat *et al.*, 2015). Zusammenfassend wurde in

beiden systematischen Studien deutlich, dass die schon bekannten natürlicherweise vorkommenden RVDs NI, NG, HD, NN und NH bereits das größte Potential bezüglich Spezifität und Effizienz ausschöpfen und daher auch in Zukunft nicht durch alternative RVDs ersetzt werden (Juillerat, *et al.*, 2015).

3.6 Vom *MtN3* zum *SWEET* - die süße Verwandlung

OsSWEETs der Gruppe III sind die Schlüssel-Suszeptibilitätsgene für die Virulenz von *Xoo* in Reis (Chu, *et al.*, 2006a; Yang, *et al.*, 2006; Antony, *et al.*, 2010; Yu, *et al.*, 2011; Streubel, *et al.*, 2013). Lange Zeit war jedoch unklar welche Funktion sie in der Pflanze erfüllen und welche Rolle sie für die Virulenz des Pathogens spielen. Eines der ersten identifizierten Suszeptibilitätsgene für TALEs aus *Xanthomonas* war *Os8N3* (*Xa13*; *OsSWEET11*), dessen PthXo1-abhängige Aktivierung unerlässlich für die Virulenz des *Xoo*-Stammes PXO99 ist (Chu, *et al.*, 2006a; Chu, *et al.*, 2006b; Yang, *et al.*, 2006; s. Abschnitt 1.2.5.1). Weitere Analysen ergaben, dass Reis sowie weitere Pflanzen, Insekten und Tiere verschiedene *Os8N3*-Homologe kodieren (Yang, *et al.*, 2006; Chong, *et al.*, 2014; Yuan, *et al.*, 2014). Aufgrund ihrer großen Ähnlichkeit (68%) zum *Medicago truncatula* Nodulin3 (*MtN3*), welches während der Etablierung der Wurzelknöllchen in der Symbiose mit *Rhizobium* aktiv ist, wurden *Os8N3* und Homologe als *MtN3*-Proteine bezeichnet (Gamas, *et al.*, 1996; Chu, *et al.*, 2006a; Yang, *et al.*, 2006). Interessanterweise konnte gezeigt werden, dass die TALE-abhängige Aktivierung des nah verwandten *Os11N3* (*OsSWEET14*) die Virulenz des *Xoo*-Stammes PXO99 wiederherstellen kann, wenn diesem der TALE PthXo1 fehlt oder PthXo1 in resistenten *xa13*-Pflanzen *Os8N3* nicht mehr aktivieren kann (Yang, *et al.*, 2006; Antony, *et al.*, 2010; s. Abschnitt 1.2.6.2). Dies zeigte erstmals das zwei Gene der *MtN3*-Familie funktional austauschbar sind. Zusätzlich entdeckte man, dass *MtN3*-Mitglieder nicht nur für Pathogene, sondern auch für die Pflanze von Bedeutung sind. *Os8N3* ist notwendig für die Pollenentwicklung, während *Os11N3* vermutlich in verschiedene Abläufe der pflanzlichen Entwicklung involviert ist (Chu, *et al.*, 2006a; Yang, *et al.*, 2006; Antony, *et al.*, 2010). Zusätzlich wurden für weitere *MtN3*-Homologe beispielsweise aus Reis oder *A. thaliana* Funktionen für die Nektarproduktion (Lin, *et al.*, 2014) oder Seneszenz (Quirino, *et al.*, 1999; Zhou, *et al.*, 2014) beschrieben. Erst 2010 konnte in der Studie von Chen *et al.*, gezeigt werden, dass die für die Pflanze und für das Pathogen wichtige Funktion der *MtN3*-Proteine vermutlich in ihrer Aktivität als Zuckertransporter zu finden ist (Chen, *et al.*, 2010). In

dieser Studie konnte gezeigt werden, dass Mitglieder der MtN3-Familie aus *Arabidopsis* (AtSWEET1, 4, 5, 7, 8, 13), *Caenorhabditis elegans* (CeSWEET1), *Homo sapiens* (HsSWEET1) und Reis (OsSWEET11, 14) Glucose transportieren, weshalb die MtN3-Familie seitdem als SWEET-Familie bezeichnet wird (Chen, *et al.*, 2010).

3.6.1 Die Gruppe III SWEETs sind Saccharose-Transporter

Als Zuckertransporter sind SWEETs vermutlich an der Verteilung dieser Energiequelle in der Pflanze beteiligt, wodurch deren Rolle in den verschiedenen oben genannten pflanzliche Prozessen erklärt werden könnte (Chen, *et al.*, 2010; Chen, *et al.*, 2012; Guo, *et al.*, 2014; Lin, *et al.*, 2014; Zhou, *et al.*, 2014). In einer weiterführenden Studie wurde für alle SWEET-Mitglieder der Gruppe III aus *Arabidopsis* (AtSWEET10 bis 15) sowie für die Gruppe III Reisproteine OsSWEET11 und OsSWEET14 eine Funktion als Saccharose-Transporter nachgewiesen (Chen, *et al.*, 2012). Zusätzlich konnte gezeigt werden, dass die Gruppe III Mitglieder AtSWEET11 und 12 im Phloemparenchym lokalisiert sind. Daraus wurde abgeleitet, dass diese SWEETs vermutlich an der Beladung des Phloems mit Saccharose beteiligt sind (Sonnewald 2011; Baker, *et al.*, 2012, Braun 2012; Chen *et al.*, 2012). Durch diese Analyse der SWEETs, die ursprünglich mit ihrer Funktion als Suszeptibilitätsgene für *Xoo* begonnen hatte, konnte nun auch endlich die Lücke im Saccharose Verteilungsweg der Pflanze geschlossen werden (Sonnewald 2011; Baker, *et al.*, 2012, Braun 2012; Chen *et al.*, 2012).

Die in Chen *et al.*, 2012 beschriebene Saccharose-Transportfunktion der *Arabidopsis* Gruppe III SWEET-Mitglieder zusammen mit den in Publikation 2 gewonnenen Ergebnissen, dass nur OsSWEETs der Gruppe III wichtig für die Virulenz von *Xoo* sind und dass die Gruppe III OsSWEETs funktional austauschbar sind, weisen darauf hin, dass vermutlich auch in Reis ausschließlich Gruppe III OsSWEETs Saccharose transportieren. Diese Funktion könnte von *Xoo* ausgenutzt werden um die pflanzlichen Verteilungswege der Saccharose zu verändern, die Versorgung von *Xoo* mit Nährstoffen und damit bakterielles Wachstum zu gewährleisten (Sonnewald 2011; Baker, *et al.*, 2012, Braun 2012; Chen *et al.*, 2012).

Zusätzlich wurde von Yuan *et al.*, 2010 eine andere Funktion für OsSWEETs als Kupferexporter vorgeschlagen (Yuan, *et al.*, 2011). Diese Studie zeigte, dass OsSWEET11 mit den Kupfertransportern COPT1 und COPT5 aus Reis interagiert und so vermutlich an der Umverteilung des pflanzlichen Kupfers beteiligt ist. Der *Xoo*-Stamm PXO99 ist sensitiv gegenüber Kupfer im pflanzlichen Xylem (Yuan, *et al.*,

2011). Die Aktivierung von *OsSWEET11* durch den PXO99-TALE PthXo1 könnte daher den Export des Kupfers aus dem Xylem fördern und somit das Wachstum von PXO99 ermöglichen. Allerdings ist diese Funktion nicht generell für *Xoo*-Stämme relevant. Der Stamm PXO86 der AvrXa7-abhängig das Homolog *OsSWEET14* aktiviert, zeigt keine Sensitivität gegenüber Kupfer (Yuan, *et al.*, 2011). In Publikation 2 wurde jedoch gezeigt, dass die alternative Aktivierung von *OsSWEET14* nach Transfer der TALEs Tal5 oder AvrXa7 in PXO99 zur Wiederherstellung der Virulenz führt, wenn dieser in PthXo1-resistente Reislinien (*xal3*) inokuliert wird bzw. *PthXo1* deletiert ist (Antony, *et al.*, 2010; Streubel, *et al.*, 2013). Diese Austauschbarkeit der verschiedenen Gruppe III *OsSWEETs* deutet daraufhin, dass deren Funktion für die Virulenz von *Xoo* redundant ist. Der Kupfer-Export ist möglicherweise eine zusätzliche Funktion von *OsSWEET11*, die für PXO99 von Bedeutung ist, jedoch nicht generell auf andere *Xoo*-Stämme zutrifft.

3.6.2 *Xoo* manipuliert vermutlich die pflanzlichen Saccharose Verteilungswege

Saccharose ist die Hauptform in der Zucker über lange Strecken von photosynthetisch aktiven Zellen (dem *source*) zu den energieverbrauchenden Zellen (dem *sink*) in der Pflanze transportiert wird (Kühn & Grof 2010; Baker, *et al.*, 2012; Braun 2012). Dazu wird die Saccharose zuerst symplastisch aus den *source*-Zellen zu den Parenchymzellen des Phloems transportiert (Baker, *et al.*, 2012, Braun 2012). Aus diesen Zellen wird die Saccharose von SWEET-Proteinen in den pflanzlichen Apoplast exportiert (Baker, *et al.*, 2012; Chen *et al.*, 2012). Vom Apoplast wird die Saccharose durch eine zweite Klasse von Saccharose-Transportern den SUC/SUT-H⁺-Symportern in die Geleitzellen des Phloems importiert (Kühn & Grof 2010). Die Geleitzellen sind durch Plasmodesmata mit den Siebelementen des Phloems verbunden, sodass die Saccharose symplastisch eingeleitet und zum *sink* transportiert wird (Braun 2012). Anschließend kann die Saccharose vom Phloem über Plasmodesmata in die *sink*-Zellen importiert werden und dort zur Weiterverwendung durch cytoplasmatische bzw. vakuoläre Invertasen oder Synthasen in Glucose und Fructose gespalten oder zu Stärke umgewandelt und gespeichert werden (Ruan 2014). Eine weitere Möglichkeit des Saccharose-Exportes aus dem Phloem ist die direkte Entlassung in den Apoplast, der die *sink*-Zellen umgibt. Von dort wird die Saccharose entweder durch SUT/SUC-H⁺ Symporter in die *sink*-Zelle importiert und wie oben beschrieben weiterverarbeitet oder direkt durch Zellwand-assoziierte Invertasen (*Cell Wall associated INVertases*,

CWINV) in Glucose und Fructose gespalten (Ruan 2014). Die Spaltprodukte können anschließend durch entsprechende Hexose-Transporter in die *sink*-Zelle aufgenommen werden (Ruan 2014).

Interessanterweise konnte gezeigt werden, dass die Expression der homologen Saccharose-Transporter *AtSWEET11* und *12* in *Arabidopsis*, mit der Expression weiterer Gene für die Saccharose-Synthese, sowie den Transport gekoppelt ist (Chen, *et al.*, 2012). Die Aktivierung der *OsSWEETs* in Reis und der Verbrauch der Saccharose durch *Xoo* könnte daher ebenfalls rückkoppelnd zum generellen Anstieg der Saccharose Produktion und des Transportes an die Infektionsstelle führen. Dadurch wird an der Infektionsstelle eine Art künstliches *sink* geschaffen in das aktiv Saccharose transportiert. Eine immer noch offene Frage ist jedoch in welcher Form die Saccharose von *Xoo* genutzt wird.

Eine Möglichkeit ist die direkte Aufnahme durch bakterielle Saccharose-Transporter. In anderen Pflanzen-Pathogen Interaktionen konnten Saccharose-Transporter bereits als wichtige Virulenzfaktoren beschrieben werden. So exprimiert der Pilz *U. maydis* einen Saccharose-Transporter (*Srt1*) der zur effizienten Saccharose-Aufnahme an der Pflanzen-Pilz-Interaktionsfläche dient und damit entscheidend zur Virulenz des Pathogens beiträgt. (Wahl, *et al.*, 2010). Außerdem konnte im Genom von *Xcc* der sogenannte *sux-Locus* identifiziert werden, der Proteine zum Import bzw. zur Verwertung der Saccharose kodiert (Blanvillain, *et al.*, 2007). Mutationen im *sux-Locus* reduzierten das Wachstum auf Saccharose-haltigem Medium, sowie die Virulenz von *Xcc* auf *Arabidopsis* (Blanvillain, *et al.*, 2007). Im Genom des *Xoo*-Stammes KACC wurde ein ähnlicher *Locus* beschrieben, jedoch gibt es bisher keine Untersuchungen ob diese Gene die Virulenz des Pathogens beeinflussen (Blanvillain, *et al.*, 2007). Im Gegensatz dazu hat die Mutation des Saccharose-Importers *suxC* im Baumwollpathogen *Xam* keinen Einfluss auf die Virulenz (Cohn, *et al.*, 2014). Da die Virulenz von *Xam* ebenfalls auf der TALE-abhängigen Aktivierung des Gruppe III *MeSWEET10a* beruht, könnte dies auf eine alternative Möglichkeit der Saccharose-Verwertung hinweisen (Cohn, *et al.*, 2014). Dabei wird die Saccharose zur Infektionsstelle transportiert, dort aber extrazellulär gespalten und in Form von Hexosen vom Pathogen aufgenommen (Cohn, *et al.*, 2014). Der Stamm *XooT7174* kodiert einen H⁺/Glucose/Galaktose Symporter der auch in anderen *Xanthomonas*-Arten konserviert ist, sodass dieser Weg generell von *Xoo*-Stämmen genutzt werden könnte (Tsuge, *et al.*, 2001). Die extrazelluläre Spaltung von Saccharose kann durch Invertasen katalysiert werden.

Bisher ist jedoch unklar ob *Xoo* bakterielle Invertasen sekretiert oder die pflanzlichen CWINVs zur Saccharose-Spaltung nutzt (Ruan 2014). Interessanterweise konnte die erhöhte Expression der CWINVs *OsCINI*, 4 und 7 in Reispflanzen ca. 48h nach der Infektion mit dem *Xoo*-Stamm PXO99 beobachtet werden (Sun, *et al.*, 2014). Auch in Paprika bzw. Tomate konnte eine erhöhte CWINV-Aktivität ca. 3 bis 4 Tage (je nach verwendeter bakterieller Dichte) nach der Infektion mit *Xcv* gemessen werden (Biemelt & Sonnewald 2006; Kocal, *et al.*, 2008; Sonnewald, *et al.*, 2012). Außerdem erzielte die Blockierung der CWINV Expression (*silencing*) in transgenen Tomatenpflanzen im Vergleich zu Wildtyppflanzen eine signifikante Verzögerung der Symptomentwicklung nach Infektion mit *Xcv* (Kocal, *et al.*, 2008). Ausgehend von diesen Ergebnissen ist die Ausnutzung pflanzlicher CWINVs zur Spaltung der Saccharose für *Xoo* durchaus denkbar, wobei bisher unklar ist, ob CWINVs direkt durch das Pathogen oder aber durch die Pflanze selbst aktiviert werden. Letzteres könnte besonders dahingehend eine Rolle spielen, da bereits bekannt ist, dass pflanzliche CWINVs als Reaktion auf biotischen Stress aktiviert werden (Roitsch, *et al.*, 2003; Rolland, *et al.*, 2006; Sonnewald, *et al.*, 2012; Sun, *et al.*, 2014; Ruan 2014). Dadurch steigt die Produktion von Hexosen, die als Signalmoleküle wirken und Abwehrreaktionen einleiten können (Roitsch, *et al.*, 2003; Rolland, *et al.*, 2006; Sonnewald, *et al.*, 2012; Sun, *et al.*, 2014; Ruan 2014). Durch diese mögliche bivalente Funktion von CWINVs zum Einen vorteilhaft für das Pathogen zum Anderen vorteilhaft für die Pflanze, ist es für das Pathogen vermutlich wichtig die Balance zu halten (Sonnewald *et al.*, 2012; Ruan 2014). Ob und welche bakteriellen Faktoren *Xoo* nutzt um die CWINV Expression bzw. Aktivität zu beeinflussen ist bisher noch weitgehend unklar.

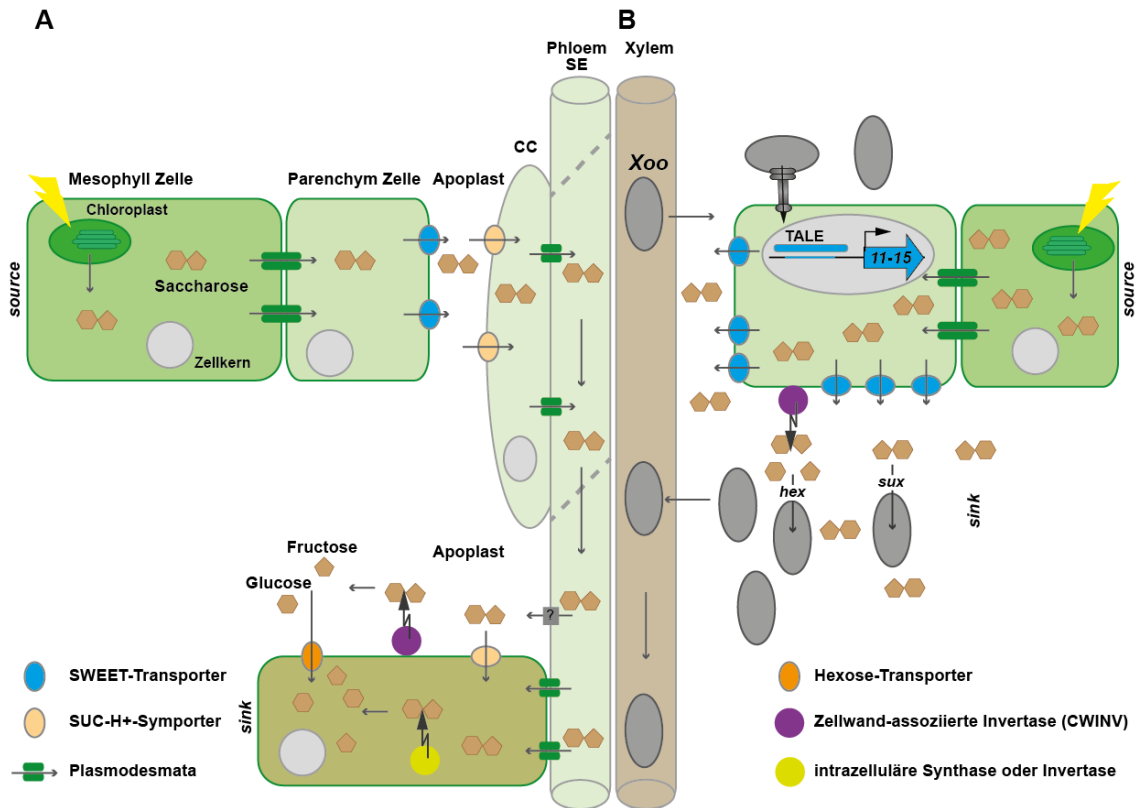


Abbildung 3.2 Die natürlichen Saccharose Verteilungswege in Reis werden von *Xoo* manipuliert
 Schematische Darstellung der Saccharose Verteilungswege in Reis und ihre Manipulation durch *Xoo* **A** Saccharose wird in photosynthetisch aktiven Zellen (*source*) produziert und über Plasmodesmata in umliegende Parenchymzellen transportiert. Anschließend exportieren SWEET Proteine die Saccharose in den pflanzlichen Apoplast. Von dort wird sie über Saccharose-H⁺-Symporter (SUC-H⁺) in die Companion Zellen (CC) importiert. Zum Langstreckentransport wird die Saccharose von den CC über Plasmodesmata in die Siebelemente (SE) des Phloems eingeleitet. An den Verbrauchenden Zellen (*sink*) wird die Saccharose entweder symplastisch importiert und durch intrazelluläre Synthesen bzw. Invertasen in Glucose und Fructose gespalten oder in den Apoplast entlassen. Vom Apoplast kann die Saccharose entweder von Saccharose-H⁺Symporter importiert werden oder durch Zellwand assoziierte Invertasen (CWINV) in Glucose und Fructose gespalten werden, die anschließend durch Hexose-Transporter in die *sink* Zelle importiert werden. **B** *Xoo* dringt über Hydathoden oder Wunden in das pflanzliche Xylem ein. Mit Hilfe sekretierter Enzyme z.B. Xylanasen kann es das Xylem verlassen und umliegende Parenchymzellen manipulieren. Dazu transloziert es TALEs in die Zelle, um die Expression der Saccharose-Transporter (*OsSWEET11-15*) zu aktivieren. Die gesteigerte Expression der SWEET-Transporter erhöht den Saccharose Export in den Apoplast und erzeugt so ein *sink*. Rückkoppelnd könnten durch den erhöhten Export auch Saccharose-Synthese Enzyme in den *source*-Zellen aktiviert werden. Die Saccharose kann vom Apoplast entweder direkt durch bakterielle Saccharose-Transporter (*sux*) aufgenommen werden oder mit Hilfe pflanzlicher CWINVs in Glucose und Fructose vorgespalten und so mittels Hexose-Transportern (*hex*) aufgenommen werden. Die verbesserte Nährstoffversorgung ermöglicht die Vermehrung des Pathogens und dessen weitere Ausbreitung über das Xylem.

3.6.3 *Xoo* manipuliert vermutlich Zellen des Phloemparenchym

Xoo ist ein systemisches Pathogen, das durch Hydathoden oder Wunden ins Reis-Xylem eintritt und sich dort ausbreitet (Swings *et al.*, 1993). Da das Xylem jedoch aus abgestorbenen, verholzten Zellen besteht, die keine Proteine synthetisieren, muss *Xoo* Zellen außerhalb des Xylems manipulieren (Sperry 2003). Welche Zellen dies betrifft ist bisher allerdings unklar. Für *Xcc* ein ebenfalls systemisches Pathogen konnte bereits gezeigt werden, dass diese Bakterien nicht nur im Xylem sondern auch in umliegenden

Epidermiszellen bzw. dem Apoplast zu finden sind (Akimoto-Tomiyama, *et al.*, 2014). Dies könnte ebenfalls für *Xoo* zutreffen. Es ist bereits bekannt, dass Typ-II-sekretierte Zellwand-abbauende Enzyme beispielsweise Xylanasen wichtige Virulenzfaktoren für *Xoo* sind (Sun, *et al.*, 2005; Rajeshwari, *et al.*, 2005). Xylanasen degradieren Xylan, welches besonders in Gräsern Bestandteil von Zellwänden ist (Vogel 2008). Die Sekretion von Xylanasen könnte somit den systemischen *Xoo*-Bakterien erlauben Xylem-Bestandteile abzubauen, um diese Barriere zu durchbrechen und Effektoren z.B. TALEs in umliegende lebende Zellen zu injizieren (Abb. 3.2). Die manipulierten Zellen liefern dann Nährstoffe, die im Apoplast von *Xoo* aufgenommen und zur Vermehrung verwendet werden können. Anschließend könnten sich die Bakterien weiter systemisch über das Xylem in noch nicht infizierte Blattbereiche ausbreiten (Abb. 3.2).

3.6.4 SWEETs und deren Bedeutung für andere Pathogene

Die Verfügbarkeit von Nährstoffen an der Infektionsstelle ist eine wichtige Grundlage für die Besiedlung der Wirtspflanze durch Pathogene. Die Möglichkeiten für Pathogene dies zu erreichen können unterschiedlich sein (Howlett 2006; Divon & Fluhr 2007; Chen, *et al.*, 2010; Chen, *et al.*, 2012). Die Versorgung des Reispathogens *Xoo* ist abhängig von der TALE-vermittelten Aktivierung der *OsSWEET* Saccharose-Transporter (Chen, *et al.*, 2012; Streubel, *et al.*, 2013). Interessanterweise beruht die Virulenz des ebenfalls reispathogenen *Xoc* nicht auf der TALE-abhängigen Aktivierung von *OsSWEETs* (Cernadas, *et al.*, 2014). Die beiden Reispathogene *Xoo* und *Xoc* unterscheiden sich hauptsächlich in der Gewebespezifität während der Infektion (Nino-Liu, *et al.*, 2006). Im Gegensatz zum systemischen *Xoo* ist *Xoc* ein apoplastisches Pathogen. Da für *Xoo* bekannt ist, dass die Deletion eines *OsSWEET*-aktivierenden TALEs hauptsächlich die Ausbreitung im Xylem, nicht jedoch die lokale Vermehrung von *Xoo* an der Eintrittsstelle verhindert, wurde vermutet, dass die TALE-abhängige Aktivierung von *OsSWEETs* eventuell mit der Gewebespezifität von *Xoo* zusammenhängt (Yu, *et al.*, 2011; Verdier, *et al.*, 2012). Der Transfer von AvrXa7 (aktiviert *OsSWEET14*) bzw. PthXo1 (aktiviert *OsSWEET11*) in den *Xoc* Stamm BLS256 konnte zwar dessen Virulenzsymptome verstärken, ermöglichte ihm jedoch nicht das Xylem zu besiedeln (Verdier, *et al.*, 2012). Daher ist die Aktivierung der *OsSWEETs* entweder nicht mit der Gewebespezifität verknüpft oder es gibt neben diesen noch weitere Faktoren die *Xoo* erlauben das Xylem zu besiedeln. Warum *Xoo*-Stämme im Gegensatz zu *Xoc* so auf die Aktivierung der Saccharose-Transporter

angewiesen sind, ist bisher nicht bekannt. Möglicherweise ist *Xoc* im Gegensatz zu *Xoo* in der Lage Nährstoffe über andere Wege zu gewinnen, die möglicherweise nicht mit der Genaktivierung durch TALEs zusammenhängen oder *Xoc* erschließt generell andere Nahrungsgrundlagen, sodass die Verfügbarkeit von Saccharose nicht grundsätzlich notwendig ist (Verdier, *et al.*, 2012; Cernadas, *et al.*, 2014).

Im Gegensatz zu *Xoc* beruht die Virulenz des apoplastischen Maniok-pathogenen Stammes *Xam668* ebenfalls auf der TALE-abhängigen Aktivierung des Gruppe III *MeSWEET10a* (Cohn, *et al.*, 2014). Ob *Xam668* und *Xoo* jedoch Gemeinsamkeiten im Zuckerverarbeitenden Metabolismus haben, welche die Notwendigkeit der TALE-abhängigen *SWEET*-Aktivierung erklären, ist bisher unklar.

Ein weiteres bakterielles Pflanzenpathogen *Pseudomonas syringae* DC3000 aktiviert ebenfalls die Expression verschiedener *AtSWEETs* (Chen, *et al.*, 2010). Diese gehören jedoch keiner distinkten Gruppe an (z.B. *AtSWEET2* (Gruppe I); *AtSWEET4*, 7, 8 (Gruppe II) *AtSWEET10*, 11, 15 (Gruppe III), sodass in dieser Interaktion nicht nur der Saccharose-Transport, sondern vermutlich auch der Transport anderer Metabolite beeinflusst wird. Allerdings ist in dieser Interaktion bisher unklar ob die *SWEET*-Gene gezielt durch *Pseudomonas* oder als Antwort der Wirtspflanze auf die Infektion aktiviert werden. Interessanterweise scheinen *SWEETs* auch in anderen nicht-bakteriellen Pflanzen-Pathogen-Interaktionen eine Rolle zu spielen. In der Infektion von *Arabidopsis* mit *Golovinomyces cichoracearum* werden ebenfalls die *AtSWEETs* 1, 2, 4, 11 und 12 aktiviert (Chen, *et al.*, 2010). Die Rolle dieser für die Virulenz des Pathogens ist jedoch nicht bekannt. Weiterhin wird der Glucose-Transporter *VvSWEET4* spezifisch durch den necrotrophen Pilz *Botrytis cinerea* nach Befall der Wirtspflanze *Vitis vinifera* aktiviert (Chong, *et al.*, 2014). *Arabidopsis*-Pflanzen in denen das Homolog *AtSWEET4* deletiert wurde, sind weniger anfällig für eine Infektion durch *B. cinerea*. Eventuell dient die Aktivierung dieses Glucose-Transporters somit ebenfalls der Nährstoffversorgung des Pilzes (Chong, *et al.*, 2014). Allerdings ist auch für diese beiden Pathogene bisher nicht bekannt, wodurch die *SWEET*-Gene aktiviert werden.

3.6.5 *Xoo* ist auf die TALE-abhängige Aktivierung von *OsSWEET14* fokussiert

In Publikation 2 konnte der mittlerweile vierte TALE identifiziert werden, der *OsSWEET14* aktiviert (Antony, *et al.*, 2010; Yu, *et al.*, 2011; Streubel, *et al.*, 2013). Im Gegensatz dazu sind bisher noch keine natürlichen TALEs für *OsSWEET12* und 15

bekannt (Streubel, *et al.*, 2013). Die Frequenz von TALEs aus *Xoo*-Stämmen unterschiedlichster Herkunft (Asien, Afrika) am *OsSWEET14* Promotor kann entweder Zufall sein oder auf eine spezielle Eigenschaft dieses Gens im Vergleich zu den anderen Gruppe III Genen hinweisen. Eine Möglichkeit wäre, dass der *OsSWEET14*-Promotor und damit die TALE-Bindestellen im Vergleich zu anderen Gruppe III *OsSWEETs* hochkonserviert ist. Dadurch wäre die Virulenz von *Xoo* in vielen verschiedenen Reislinien gewährleistet (Verdier, *et al.*, 2012; J. Streubel, unveröffentlicht). Im Gegensatz dazu sind für andere Gruppe III *OsSWEETs* bereits resistente Promotorvarianten bekannt (*OsSWEET11* R-Gen: *xa13* und *OsSWEET13* R-Gen: *xa25*) in denen TALE Bindestellen durch Indels oder Substitutionen zerstört wurden (Chu, *et al.*, 2006a; Liu, *et al.*, 2011). Außerdem könnte es sein, dass der *OsSWEET14* Promotor auch in Nicht-Wirtspflanzen z.B. anderen Gräsern konserviert ist. *Xoo* ist auf solchen Pflanzen zwar nicht virulent, durch die Aktivierung des Zuckertransporters könnte allerdings zumindest übergangsweise das Überleben von *Xoo* gesichert werden. Allerdings ist fraglich ob *Xoo* in Nicht-Wirtspflanzen in der Lage ist TALEs in pflanzliche Zellen zu translozieren. Eine weitere Möglichkeit wäre, dass *OsSWEET14* absolut unerlässlich für die Pflanze ist. Eventuell hat es im Gegensatz zu den anderen *OsSWEETs* eine höhere Affinität für bestimmte Zucker oder kann zusätzlich andere Metabolite transportieren. Dadurch könnte dieses Gen essentiell für die Pflanze sein und somit im Wettrüsten mit dem Pathogen nicht so einfach verändert werden oder sogar verloren gehen. Dadurch wäre *OsSWEET14* ein stabiles Ziel-Gen für *Xoo*. Hinweise für dessen Bedeutung in Reis geben transgene Pflanzen in denen *OsSWEET14* durch *Gene silencing* nicht mehr exprimiert wird (Antony, *et al.*, 2010). Diese Pflanzen sind bezüglich verschiedener Entwicklungsprozesse eingeschränkt, wodurch u.a. das generelle Wachstum verringert wird (Antony, *et al.*, 2010). Ähnliches konnte jedoch auch für *OsSWEET11* gezeigt werden (Chu, *et al.*, 2006a; Yang, *et al.*, 2006). Transgene Pflanzen, die *OsSWEET11* nicht mehr exprimieren, sind steril, da dieses Gen in der Pollenentwicklung involviert ist (Chu, *et al.*, 2006a; Yang, *et al.*, 2006). Daher scheint auch *OsSWEET11* unerlässlich für die Reis-pflanze zu sein. Die Notwendigkeit der anderen Gruppe III *OsSWEETs* für die Entwicklungsprozesse der Pflanze sind bisher nicht bekannt, sodass über die mögliche besondere Funktion von *OsSWEET14* keine eindeutige Aussage getroffen werden kann. Eine letzte Möglichkeit könnte sein, dass *OsSWEET14* eine höhere Transportrate für Saccharose aufweist, sodass *Xoo* effizienter mit Nährstoffen versorgt werden könnte. Tatsächlich konnte im Vergleich

der Transportfunktion von OsSWEET11 und 14 festgestellt werden, dass OsSWEET14 Saccharose effizienter transportiert (Chen, *et al.*, 2012). Allerdings konnte in Publikation 2 gezeigt werden, dass die alternative Aktivierung von OsSWEET11 bzw. anderen Gruppe III OsSWEETs die Funktion von OsSWEET14 ersetzen und so die Virulenz von BAI3 Δ *talC* wiederherstellen kann, wodurch diese Möglichkeit vermutlich nicht zutrifft.

3.7 Aberrante *repeats* - eine evolutionäre Anpassung an resistente Allele?

In Publikation 3 konnte erstmals gezeigt werden, dass die *repeat*-Region eines TALEs weniger starr ist als vermutet und die strikt konsekutive ein-*repeat*-ein-Nukleotid Abfolge durch die Insertion von aberranten *repeats* durchbrochen werden kann (Boch, *et al.*, 2009; Moscou & Bogdanove 2009; Richter, *et al.*, 2014). Dadurch wird eine Flexibilität erzeugt, die dem TALE erlaubt je nach Ziel-DNA-Sequenz den aberranten *repeat* an der Bindung teilnehmen zu lassen (optimale Ziel-DNA-Sequenz) oder aber diesen aus der *repeat*-Region herauszuwinden, wodurch alle nachfolgenden *repeats* eine Position nach vorn verschoben werden können, um eine -1 *frameshift* Sequenz zu binden. Die in Publikation 3 gewonnenen Erkenntnisse zeigen erstmals eine Möglichkeit wie sich TALE *repeat*-Regionen an bestimmte Resistenzen anpassen könnten. TALE-abhängig aktivierte Ziel-Gene wie z.B. die *OsSWEETs* sind wie bereits beschrieben oft auch für die Wirtspflanze selbst von Bedeutung (Chu, *et al.*, 2006a; Yang, *et al.*, 2006; Antony, *et al.*, 2010). Daher ist es der Pflanze im Wettrüsten mit dem Pathogen nicht möglich das betreffende Ziel-Gen zu verändern oder sogar zu verlieren ohne sich selbst zu schaden. Eine Möglichkeit die TALE-abhängige Aktivierung des Ziel-Gens zu unterbinden, ist daher die Veränderung der TALE-Bindestelle im Promotor. TALEs sind aufgrund ihrer konsekutiven ein-*repeat*-ein-Nukleotid Bindung äußerst anfällig für Veränderungen in ihrer Bindestelle, die bereits durch die Insertion oder Deletion (Indel) eines einzelnen Nukleotids hervorgerufen werden können (Chu, *et al.*, 2006a; Yang, *et al.*, 2006; Liu, *et al.*, 2011). Solche Indels verursachen eine Verschiebung des Leserahmens in der Ziel-DNA-Sequenz, sodass es in den meisten Fällen zu einer Anhäufung von RVD-DNA-Fehlpaarungen kommt. Diese können letztendlich dazu führen, dass TALEs nicht mehr binden und so auch das Ziel-Gen nicht mehr aktivieren. Pflanzen mit solchen Promotorveränderungen sind damit resistent. Publikation 3 zeigt, dass aberrante *repeats* TALEs erlauben könnten andernfalls resistente Allele die eine einzelne Nukleotid-Deletion enthalten zu

umgehen. Damit könnten aberrante *repeats* eine evolutionäre Anpassung von *Xanthomonas* TALEs an solche Allele darstellen. Bisher ist jedoch für keinen der natürlichen TALEs mit aberrantem *repeat* ein solcher Fall bekannt. PthXo3 und AvrXa7 sind bisher die einzigen TALEs mit aberrantem *repeat* für die bekannt ist, dass sie notwendig für die Virulenz sind (Antony, *et al.*, 2010; Richter, *et al.*, 2014). Die AvrXa7 bzw. PthXo3 Bindestelle im *OsSWEET14* Promotor scheint jedoch bereits ein *frameshift* Allel zu repräsentieren da es besonders in der Region nach dem aberranten *repeat* eine Anhäufung von RVD-DNA-Fehlpaarungen aufweist die signifikant verringert werden können, wenn der aberrante *repeat* aus der *repeat*-Region herausgewunden wird (Richter, *et al.*, 2014). Zusätzlich könnten jedoch auch die unspezifischen RVDs (z.B. NS, N*) die besonders im C-terminalen Bereich der Avrxa7 bzw. PthXo3 *repeat*-Region zu finden sind die Flexibilität dieser TALEs ermöglichen (Richter *et al.*, 2014). Daher ist bisher nicht klar ob die *OsSWEET14* Aktivierung nur durch den aberranten *repeat* oder zusätzlich durch das Vorhandensein unspezifischer RVDs vermittelt wird. Aufgrund des Fehlens natürlicher Beispiele in denen TALEs mit aberranten *repeats* einen evolutionären Vorteil bieten ist, bisher nicht klar ob aberrante *repeats* nur übrig gebliebene Fehler der Replikation sind, die bisher nicht gestört haben oder ob diese *repeats* tatsächlich in bestimmten *Xoo* bzw. *Xoc*-Reis-Interaktionen durch positive Selektion beibehalten wurden.

Zusätzlich könnten aberrante *repeats* auch für die Biotechnologie interessant sein, da sie als flexible Verknüpfung genutzt werden könnten, um funktionale Domänen (z.B. Nukleasen) in die *repeat*-Region zu inserieren. Dadurch könnte z.B. die biotechnologische Verwendung von TALE-Nukleasen verbessert werden. Derzeit werden TALENs als Paar verwendet, die je ein *FokI* Monomer enthalten (Miller, *et al.*, 2011; Bogdanove 2014). Durch die Verwendung zweier unabhängiger DNA-bindender Proteine kann außerdem die Spezifität für die gewählte Ziel-DNA-Sequenz erhöht werden (Bogdanove 2014; Tsai, *et al.*, 2014). Der Nachteil dabei ist jedoch, dass zwei Konstrukte gleichzeitig in die verwendeten Ziel-Zellen transferiert werden müssen. Dies reduziert möglicherweise die Erfolgsrate. Durch die Integration der zwei Nuklease-Domänen in die *repeat*-Region könnte ein TALE das TALEN-Paar ersetzen, und trotzdem seine hohe Spezifität durch das Vorhandensein von zwei unabhängigen DNA-bindenden-Domänen (getrennt durch die Nuklease-Domäne) erhalten.

3.8.1 TALEs können die Transkription von verschiedenen Positionen aktivieren

Die Analyse eines TALE-Repertoires mit Bindestellen im *OsSWEET14* Promotor im transienten System (*N. benthamiana*) ergab, dass die Mehrheit der *forward*-bindenden TALEs das Ziel-Gen aktivieren konnte, wobei die Position der Bindestelle im Promotor die Effizienz der Aktivierung beeinflusste (Abb. 2.1B). Für ausgewählte TALEs konnten diese Ergebnisse auch im natürlichen *Xoo*-Reis System bestätigt werden. Einige TALEs die das Reporter-gen in *N. benthamiana* nur wenig effizient (TAL3, 7, 14) oder gar nicht aktivierten (TAL18), konnten auch das Endogen in Reis nicht aktivieren (Abb. 2.4B,C). Die geringe Effizienz dieser TALEs könnte verschiedene Gründe haben. Entweder die TALEs sind trotz ausgewogener RVD-Komposition generell schwach und können ihre Ziel-DNA-Sequenz nicht effektiv binden, oder die Zugänglichkeit der Bindestelle ist eingeschränkt (Streubel, *et al.*, 2012). Letzteres könnte entweder auf die Blockierung der Bindestelle durch Wirtsproteine oder auf den Methylierungsstatus der Bindestelle zurückzuführen sein (Bultmann, *et al.*, 2012; Deng, *et al.*, 2012b; Valton, *et al.*, 2012). Zusätzlich könnten in der TALE-Bindestelle zufällig DNA-Elemente vorkommen, die häufig im entsprechenden Genom zu finden sind, sodass der TALE austitriert wird und seine eigentliche Bindestelle nicht effizient findet. Für TAL18 und TAL3 konnte anhand des *Bs4*-Reporterkonstruktes bereits gezeigt werden, dass beide TALEs außerhalb des *OsSWEET14* Promotor-Kontext das Reporter-gen effizient aktivieren (Abb. 2.3). Für diese TALEs kann somit eine "schwache" RVD-Komposition ausgeschlossen werden, für TAL14 und TAL7 sollte dies allerdings noch überprüft werden. Daher weisen diese Ergebnisse darauf hin, dass möglicherweise die Bindestelle dieser TALEs durch Wirtsproteine oder Methylierung blockiert ist. Um Wirtsproteine zu identifizieren, die im *OsSWEET14* Promotor binden und so möglicherweise mit den TALEs um deren Bindestellen konkurrieren, wurden zwei verschiedene Vorhersagen durchgeführt. In der ersten Vorhersage wurde der *OsSWEET14*-Promotor auf mögliche Bindestellen bekannter DNA-bindender Proteine aus *Arabidopsis* überprüft (Abb. 5.2, Liste DNA-bindender Proteine und Analyse zur Verfügung gestellt von Jan Grau). Die zweite Vorhersage wurde mit Hilfe der Software PlantCARE (CARE: *Cis-acting-regulatory-element*) durchgeführt (Rombauts, *et al.*, 1999, Lescot, *et al.*, 2002; Abb. 5.1). In dieser werden alle in der PlantCARE Datenbank verfügbaren regulatorischen DNA-Elemente die bereits in verschiedenen Pflanzenspezies (z.B. Monokotyledone und Dikotyledone) beschrieben wurden, einbezogen. Anhand der beiden Analysen konnten für einige TALE-Bindestellen

Wirtsproteine identifiziert werden, die möglicherweise für die Blockierung dieser verantwortlich sind. Mittels PlantCARE-Analyse wurden für die TAL7-Bindestelle zwei mögliche regulatorische DNA-Elemente identifiziert, die von Wirtsproteinen gebunden werden könnten (Abb. 5.1). Das erste ist das TCA-Element, ein *cis*-wirkendes Element, welches an der Signalweiterleitung ausgelöst durch Salicylsäure beteiligt ist (Rombauts, *et al.*, 1999, Lescot, *et al.*, 2002). Dieses Element überlappt jedoch nicht nur mit der TAL7 sondern auch mit der TAL8 Bindestelle (Abb. 5.1 & 5.3). Da TAL8 in den durchgeführten Studien jedoch effektiv das entsprechende Ziel-Gen aktivieren konnte spielt dieses Element vermutlich keine Rolle in der Blockierung der TAL7 Bindestelle (Abb. 5.3). Das zweite DNA-Element ist das F-box-DNA-Element (Rose, *et al.*, 1999). Das F-box-DNA-Element wurde als mögliches regulatorisches Element im *RBCS3A*-Gen in Tomate beschrieben (Rose, *et al.*, 1999). Im *RBCS3A*-Gen überlappt es mit der I-box die vom Transkriptionsfaktor LeMYB1 gebunden wird (Rose, *et al.*, 1999). Die I-box kommt jedoch nicht im *OsSWEET14* Promotor vor, weshalb unklar ist ob das F-Box-DNA-Element in diesem eine Rolle spielt. Weiterhin auffällig ist die Pyrimidin-reiche Sequenz am Beginn der TAL7 Bindestelle (Abb. 5.3). Aus Studien anhand pflanzlicher Promotorregionen ist bekannt, dass Pyrimidin-reiche Sequenzen, der sogenannte Y-Patch, pflanzenspezifisch sind und vermutlich regulatorisch auf die Transkription wirken (Yamamoto, *et al.*, 2007a; Yamamoto, *et al.*, 2007b; Civan & Svec 2009; Kumari & Ware 2013). Solche zusätzlichen regulatorischen Sequenzen sind für die komplexe Kontrolle der Transkription notwendig (Kumari & Ware 2013). Sie können einerseits mit der TATA-Box zusammenwirken und andererseits besonders in Promotoren ohne TATA-Box für die korrekte Positionierung des PIC sorgen (Thomas & Chiang 2006; Yamamoto, *et al.*, 2007a; Yamamoto, *et al.*, 2007b; Civan & Svec 2009; Kumari & Ware 2013). TATA-lose Promotoren sind in Reis mit einem Anteil von 81% häufig zu finden (Kumari & Ware 2013). Bisher ist jedoch noch nicht bekannt welche Aufgabe der Y-Patch bei der Kontrolle der Transkription hat oder ob und welche Proteine an diese Sequenzen binden. Daher ist nicht auszuschließen, dass TAL7 mit Y-Patch bindenden Proteinen konkurriert. Andererseits könnte es auch sein, dass TAL7 durch die Häufigkeit dieser Sequenz im Genom austitriert wird. Der Y-Patch konnte in ca. 67% der analysierten Promotoren in Reis identifiziert werden (Kumari & Ware 2013). Die Häufigkeit dieser Sequenz könnte dazu führen, dass TAL7 mit seinen N-terminalen Y-patch-bindenden *repeats* (betrifft mindestens *repeat* 1-7) länger an diesen Sequenzen verweilt, selbst wenn die restliche *repeat*-Region nicht binden kann.

Dadurch ist TAL7 möglicherweise nicht in der Lage die ausgewählte Bindestelle im *OsSWEET14*-Promotor effizient genug zu finden. Da die TAL7-Bindestelle den Y-Patch nur im vorderen Teil der *repeat*-Region bindet, die folgenden *repeats* jedoch spezifisch vermutlich seltenere Ziel-DNA-Sequenzen erkennen, wurden solche Y-Patch Sequenzen vermutlich auch nicht als mögliche *off-targets* in der vorher durchgeführten Analyse identifiziert (s. Abschnitt 2.1). Außerdem ist der Pflanzen-spezifische Y-Patch mit CpG Inseln im tierischen System vergleichbar. Daher könnte ein weiteres Problem der TAL7-Bindestelle Methylierung sein. Methylierung dient hauptsächlich dem Stilllegen von Genen, kann aber auch andere Funktionen erfüllen (Jones 2012). In Pflanzen kann Methylierung nicht nur an CpG, sondern auch an CHG und CHH Positionen erfolgen (H=A,C,T; Cokus, *et al.*, 2008; Saze, *et al.*, 2012). Das bedeutet, dass fast jede Cytosin-haltige TALE-Bindestelle im *OsSWEET14*-Promotor methyliert sein könnte. Der verwendete Cytosin-spezifische RVD HD bindet nicht an methylierte Cytosine, sodass die Gesamtbindung des TALEs und somit auch die Aktivierung des Ziel-Gens verhindert werden könnte. (Bultmann, *et al.*, 2012; Deng, *et al.*, 2012b; Valton, *et al.*, 2012). Der Methylierungsstatus des *OsSWEET14*-Promotors in Reis ist derzeit nicht bekannt, sodass diese Ursache durchaus denkbar wäre.

TAL18 ist der einzige TALE der im transienten System in *N. benthamiana* keine messbare Aktivierung des Reportergens hervorrief, obwohl er seine Ziel-DNA-Sequenz grundsätzlich binden kann (Abb. 2.1 & 2.3). Für diesen TALE ergab die Analyse DNA-bindender Proteine aus *Arabidopsis*, dass dessen Ziel-DNA-Sequenz die Bindestelle des Proteins ARR10 enthält (Jan Grau; Abb. 5.2). ARR10 ist auch in Reis zu finden und gehört zur Familie der GARP-Proteine (*glycoprotein A repetitions predominant*), die vermutlich an der Weiterleitung von Cytokinin Signalen beteiligt sind (Ishida, *et al.*, 2008). Allerdings ergab die Suche nach ARR10-Homologen mittels BLAST, dass dieses Protein bisher noch nicht in *N. benthamiana* identifiziert wurde (Johnson, *et al.*, 2008). Außerdem liegt die Bindestelle dieses Proteins zentral in der TAL18 Bindestelle, sodass ARR10 auch im *Bs4*-Reporterkonstrukt binden und somit die Bindestelle blockieren könnte. Da TAL18 an diesem Reporterkonstrukt jedoch aktiv ist, könnte dies darauf hinweisen, dass entweder ein möglicherweise noch nicht identifiziertes ARR10-Homolog in *N. benthamiana* noch weitere umliegende Sequenzen benötigt bzw. generell eine andere Sequenz bevorzugt oder aber ein anderes, bisher noch nicht identifiziertes Protein die TAL18 Bindestelle blockiert.

Für TAL14 ergab die Vorhersage mittels PlantCARE, dass an dieser Position eine Bindestelle für das Protein SORLREP3 vorhanden ist, die sich häufig in microRNA-Promotoren in Reis und *Arabidopsis* befindet (Zhao & Li 2013; Abb.5.1). Daher könnte es sein, dass TAL14 mit diesem Protein um die Bindestelle konkurriert oder aber wie bereits für TAL7 beschrieben durch die Häufigkeit ähnlicher Bindestellen in anderen Promotoren austitriert wird.

Zusätzlich ist für die fehlende Wiederherstellung der Virulenz in Reis durch die TALEs TAL18 und TAL14 auch noch eine andere Erklärung im Zusammenhang mit der TALE-abhängigen Verschiebung des Transkriptionsstarts denkbar. Im Rahmen dieser Arbeit konnte anhand anderer TALEs am *OsSWEET14* Promotor bereits gezeigt werden, dass sie den Transkriptionsstart verschieben, sodass dieser meist ca. 40-60bp *downstream* der Bindestelle liegt. Die TAL18 Bindestelle liegt im Reis *OsSWEET14* Promotor direkt vor dem ATG (Abb. 5.3). Wenn nach der TAL18-abhängigen Genaktivierung in Reis der Transkriptionsstart ebenfalls 40-60bp *downstream* der Bindestelle liegen würde, könnte durch die Nutzung eines weiter *downstream* liegenden alternativen Startkodons (nächstes Startkodon *in frame* ca. 69bp oder 174bp *downstream* vom *OsSWEET14* ATG) ein N-Terminal verkürztes *OsSWEET14* Protein produziert werden. Diesem fehlt dann vermutlich das Signal zur Membranverankerung, welches häufig im N-terminalen Bereich von Proteinen lokalisiert ist (Keenan, *et al.*, 2001). Da SWEETs Transporter sind, könnte ein solches Protein seine Funktion nicht mehr erfüllen. Allerdings sollte in diesem Fall trotzdem *OsSWEET14*-Transkript nachweisbar sein. Analysen der *OsSWEET14*-Transkriptmenge in Anwesenheit von TAL18 zeigten jedoch im Vergleich zur Negativkontrolle keine erhöhte Transkriptmenge. Außerdem müsste TAL18 in *N. benthamiana* trotz Transkriptionsstartverschiebung ein funktionales β -Glucuronidase Protein produzieren, da sich durch die Insertion des Promotorfragmentes in den verwendeten Vektor mittels LR-Reaktion zusätzliche Sequenzen zwischen der TAL18 Bindestelle und dem ATG der β -Glucuronidase befinden. Dies ist jedoch nicht der Fall. All dies lässt vermuten, dass die TAL18-Bindestelle wie oben bereits beschrieben durch Wirtsproteine blockiert sein könnte. Die Transkriptionsstartverschiebung könnte auch die fehlende Wiederherstellung der Virulenz durch TAL14 in Reis erklären. Verschiebt TAL14 den Transkriptionsstart, wird ein Transkript mit einer ca. um 640bp verlängerten 5'UTR im Vergleich zum natürlichen Transkript synthetisiert (Antony *et al.*, 2010). Diese verlängerte 5'UTR enthält zusätzliche ATGs, die als alternative Startpunkte für die

Translation zur Synthese fehlerhafter Proteine führen könnten. Die minimale TAL14 Restaktivität im transienten System, könnte dadurch erklärt werden, dass wie am Beispiel von TAL11 beobachtet einige Transkripte am natürlichen Startpunkt beginnen (Abb. 2.1 & 2.6). Dieser Anteil reicht für die Aktivierung des Reportergens im transienten System (*N. benthamiana*) vielleicht aus, nicht jedoch für die Wiederherstellung des Virulenzphänotyps im natürlichen Reis-System. Allerdings müsste auch in diesem Fall *OsSWEET14* Transkript messbar sein. Dies war jedoch in Reis nicht der Fall (Abb. 2.5). Möglicherweise sind die verlängerten Transkripte instabil oder werden frühzeitig abgebrochen und sind somit nicht mit den verwendeten Oligonukleotiden, die in der kodierenden Sequenz binden, nachweisbar.

Interessanterweise waren die zwei *reverse* und *upstream* der TATA-Box bindenden TALEs TAL-1 und -2 in der Lage das Reportergen vergleichbar mit TAL1 und 2 zu aktivieren und auch den Virulenzphänotyp im natürlichen System wiederherzustellen (Abb. 2.1 & 2.5). Da zufällige alternative Bindestellen im *OsSWEET14*-Promotor bzw. im Promotor anderer *OsSWEET*-Gene ausgeschlossen wurden, ist es denkbar, dass die TALEs TAL1/-1 und TAL2/-2 hauptsächlich eine Rekrutierungsfläche für den PIC bieten, in der die Orientierung der AD keinen Einfluss auf die Aktivierung des Ziel-Gens hat. Dies ist im Einklang mit Studien in denen gezeigt werden konnte das TALE-Derivate, die statt der natürlichen sauren AD die saure AD des *Herpes simplex virus* Proteins VP16 enthielten, im Human-System ein Ziel-Gen auch *reverse* orientiert effizient aktivieren, solange sie in zugänglichen Bereichen binden (Maeder, *et al.*, 2013; Uhde-Stone, *et al.*, 2014). Im Gegensatz dazu führten *reverse* TALEs die *downstream* der TATA-Box binden, zu keiner effizienten Aktivierung des Ziel-Gens (Abb. 2.1 & 2.5). Für TAL-3 und -7 kann dies dadurch erklärt werden, dass bereits die *forward*-bindenden TALEs kaum aktiv sind (Abb. 2.1 & 2.5). Allerdings überlappen die Bindestellen von TAL-81, -82, -9 und -AvrXa7 mit den Bindestellen aktiver TALEs (Abb. 2.1 & 2.5). Dies deutet daraufhin, dass deren Bindestellen grundsätzlich zugänglich sein sollten. Möglicherweise, ist dies ein Hinweis, dass die beiden sauren ADs (TALE vs. VP16) trotz ähnlicher Eigenschaften unterschiedlich mit der Transkriptionsmaschinerie interagieren, weshalb die Lage der natürlichen AD für TALEs an bestimmten Positionen im Promotor doch eine Rolle spielt. Ein mögliches Problem bestimmter *reverse*-orientierter TALEs wäre, dass in *forward* gerichteten TALEs die AD am C-Terminus und damit in Transkriptionsrichtung positioniert ist (Abb. 3.3). Dadurch kann ein PIC an den C-terminalen Teil des TALEs rekrutiert

werden, liegt damit hinter dem TALE und kann ohne Behinderung durch den TALE die Transkription initiieren (Abb. 3.3). Im Fall der *reverse*-gerichteten TALEs ist die C-terminale AD andersherum positioniert. Rekrutiert sie nun den PIC bindet dieser vor dem TALE (Abb. 3.3). Dadurch ist der PIC entweder falsch herum orientiert oder wird durch den TALE blockiert (Abb. 3.3). Um dieses Problem zu umgehen bzw. den PIC wieder in die richtige Orientierung zum Ziel-Gen zu positionieren, könnte im Fall der *reverse*-gerichteten TALEs möglicherweise die DNA gewunden werden (Levine, *et al.*, 2014; Abb. 3.3). Ähnlich wie bei *Enhancer*-Element bindenden Aktivatoren könnte dies z.B. durch den Mediator vermittelt werden (Milbradt, *et al.*, 2011; Vojnic, *et al.*, 2011; Levine, *et al.*, 2014). Durch diesen *DNA-loop* benötigen *reverse*-orientierte TALEs möglicherweise mehr Platz an der DNA. In diesem Fall könnte es problematisch sein, wenn *reverse*-bindende TALEs z.B. TAL-81/-81/-9 zu nah an der kodierenden Sequenz liegen. Durch den *DNA-loop* könnte der Transkriptionsstart dadurch in die kodierende Sequenz verschoben werden, sodass kein funktionales Protein entsteht (Abb. 3.3).

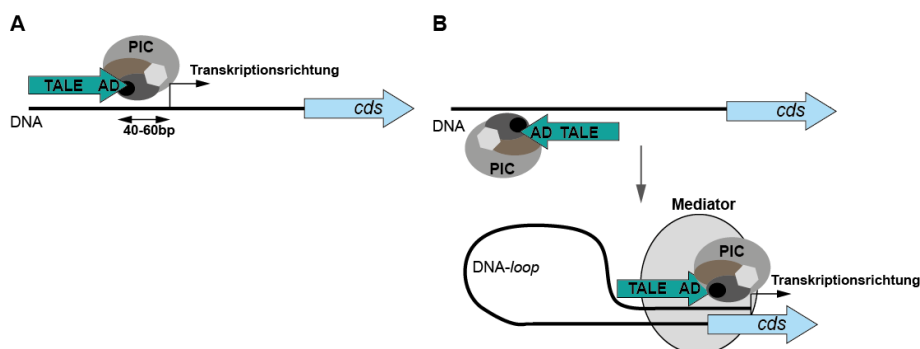


Abbildung 3.3 Modell des Einflusses der TALE-Orientierung auf den PIC

A TALEs rekrutieren den PIC vermutlich über die C-terminale Aktivierungsdomäne (AD). Dadurch liegt der PIC (Präinitiationskomplex) im Fall von *forward*-orientierten TALEs hinter dem TALE und kann in Richtung der kodierenden Sequenz (*cds*) transkribieren. **B** In *reverse*-gerichteten TALEs ist die AD in die entgegengesetzte Richtung orientiert. Dadurch ist der PIC (relativ zur *cds*) vor dem TALE positioniert. Dies könnte einerseits die Transkriptionsrichtung beeinflussen, andererseits könnte der TALE den PIC blockieren, sodass dieser sich nicht in Richtung der *cds* bewegen kann. Um dies zu umgehen könnte die DNA durch Adapterproteine z.B. den Mediator gewunden werden, sodass TALE und PIC wieder korrekt zur *cds* ausgerichtet sind. Bei *reverse* TALEs die zu nah an der *cds* binden, könnte dadurch der Transkriptionsstart in die *cds* verschoben werden sodass kein funktionales Protein entsteht.

Allerdings konnte anhand der Transkriptionsstarts für TAL1/-1, TAL2/-2 und TAL3/-3 gezeigt werden, dass diese unabhängig von der Orientierung der TALE-Bindestelle übereinstimmen, wodurch diese Erklärung vermutlich ausgeschlossen werden kann (Abb. 2.6). Eine andere Erklärung könnte sein, dass zufällig in unmittelbarer Nähe der AD *reverse*-bindender TALEs ein Wirtsprotein bindet, welches die Rekrutierung des PICs blockiert. Dies wäre z.B. im Fall von TAL-81 und -82 denkbar, da die

davorliegende TAL7 DNA-Sequenz wie bereits beschrieben möglicherweise blockiert wird. Dies könnte auch für TAL-9 und -AvrXa7 zutreffen, muss jedoch noch analysiert werden. Ob außerdem die Position der *reverse*-TALEs relativ zu anderen regulatorischen Promotorelementen z.B. *upstream* oder *downstream* der TATA-Box, einen Einfluss auf das Aktivierungspotential hat könnte anhand weiterführender Analysen im *OsSWEET14* Promotor oder an anderen Promotoren analysiert werden.

3.8.2 TALEs verschieben den natürlichen Transkriptionsstart im *OsSWEET14* Promotor

Interessanterweise konnte im Rahmen dieser Arbeit gezeigt werden, dass die Mehrheit der in *N. benthamiana* analysierten TALEs einen individuellen Transkriptionsstart am *OsSWEET14* Promotor zur Folge hatte, der sich vom natürlichen Transkriptionsstart unterscheidet. Diese Ergebnisse stimmen mit denen vorheriger Studien überein und konnten bestätigen, dass TALEs den Transkriptionsstart nicht durch die räumliche Blockierung der TATA-Box, sondern vermutlich durch die eigenständige Rekrutierung des Polymerase-II-Initiationskomplexes (PIC) beeinflussen. Wie genau TALEs den PIC rekrutieren ist bisher jedoch weitgehend unklar. Interessanterweise gleichen sich die Transkriptionsstarts der *forward* und *reverse* orientierten TALEs (TAL1/-1 bzw. TAL2/-2 bzw. TAL3/-3) obwohl die AD unterschiedlich positioniert ist. Dies könnte darauf hinweisen, dass die AD nur die Rekrutierungsoberfläche für den PIC darstellt. Durch die Bindung an zusätzliche regulatorische Sequenzen, die für einige Komponenten des PIC z.B. TFIID und TFIIB bereits beschrieben wurden, ist dieser anschließend möglicherweise in der Lage die korrekte Position einzunehmen und den nächstmöglichen Transkriptionsstart auszuwählen (Kutach & Kadonaga 2000; Thomas & Chiang 2006; Kostrewa, *et al.*, 2009, Liu, *et al.*, 2010; Kumari & Ware 2013). Transkriptionsstarts enthalten häufig die sogenannte Initiator-Sequenz, allerdings sind solche Initiator-Sequenzen in pflanzlichen Promotoren lediglich durch die YR-Regel (C/T A/G⁺¹; +1 entspricht dem Startpunkt) beschrieben, sodass mögliche Startpunkte im *OsSWEET14* Promotor nicht vorhergesagt werden können (Thomas & Chiang 2006; Kumari & Ware 2013). Das Vorhandensein einer passenden Initiator-Sequenz könnte auch erklären, warum die Transkriptionsstarts der Paare 1/-1 und 2/-2 übereinstimmen. Möglicherweise ist der für beide TALEs erhaltene Transkriptionsstart die einzige nächstliegende Initiator-Sequenz.

3.8.3 Aktivatorproteine können mit der Transkriptionsmaschinerie interagieren

Die Übersetzung der Erbsubstanz in RNA ist die Grundlage für zelluläre Abläufe in jedem Organismus. Dazu besitzen Eukaryoten verschiedene RNA-Polymerasen, wobei lediglich die Polymerase-II (Pol-II) an der Produktion von Proteinkodierenden mRNAs beteiligt ist (Thomas & Chiang 2006). Die Pol-II benötigt sogenannte generelle Transkriptionsfaktoren (GTFs), die z.B. für die Bindung an Kernpromotorelemente, die Rekrutierung der Pol-II, das Öffnen der DNA oder für den Wechsel von Initiation zu Elongation benötigt werden (Thomas & Chiang 2006; He, *et al.*, 2013; Grünberg & Hahn 2013). Die Assemblierung des PIC kann entweder schrittweise in der Abfolge TFIID->TFIIA->TFIIB->TFIIF+PolII->TFIIE->TFIIH oder durch die Rekrutierung eines bereits vorhandenen Holoenzymen erfolgen (Thomas & Chiang 2006). Aktivatoren können durch Interaktionen mit verschiedenen Komponenten des PIC oder mit Co-Aktivatoren bzw. Adapterproteinen verschiedene Vorgänge der Transkriptionsinitiation fördern und verstärken (Triezenberg 1995). Sie können durch die Interaktion mit Chromatin-remodeling Komplexen die Zugänglichkeit des Promotors erhöhen, mögliche repressorische Komponenten entfernen, oder den PIC rekrutieren und stabilisieren bzw. dessen Affinität oder Häufigkeit am Promotor erhöhen und vermutlich auch am Übergang der Initiation zur Elongation beteiligt sein (Triezenberg 1995; Thomas & Chiang 2006). Wichtig für die Interaktion von Aktivatorproteinen mit Komponenten der Transkriptionsmaschinerie ist das Vorhandensein einer AD (Triezenberg 1995). Trotz intensiver Forschung anhand diverser Aktivatorproteine ist bisher noch wenig über die strukturellen Voraussetzungen von ADs bekannt. Dies ist vermutlich darin begründet, dass sie erst nach Kontakt mit dem Interaktor eine Struktur annehmen (Triezenberg 1995; Shen, *et al.*, 1996; Uesugi, *et al.*, 1997). Lediglich die Einteilung anhand ihrer dominanten Aminosäurekomposition in glutaminreiche (z.B. SP1), prolinreiche (z.B. AP2, c-jun) oder saure ADs (z.B. Gal4, Gcn4, VP16, p53) wurde bisher vorgenommen (Seipel, *et al.*, 1992; Triezenberg 1995). Dabei konnte jedoch festgestellt werden, dass nicht zwangsläufig die namensgebenden Aminosäuren funktionsrelevant sind, sondern vielmehr die Abfolge bestimmter Aminosäuren z.B. der Wechsel der sauren Aminosäuren mit Hydrophoben Aminosäuren mit großen Seitenketten (Lin, *et al.*, 1994; Triezenberg 1995; Seipel, *et al.*, 1992). Diese bilden dann vermutlich die geeignete Oberfläche für die Interaktion mit anderen Proteinen (Seipel, *et al.*, 1992; Lin, *et al.*, 1994; Triezenberg 1995). Die meisten Studien an ADs wurden im Human- oder Hefe-System durchgeführt, sodass der Vergleich mit

pflanzlichen ADs oft fehlt. Trotzdem können einige Übereinstimmungen gefunden werden. Anhand der Aminosäuresequenz der TALE-AD wurde diese der Gruppe der sauren ADs zugeordnet (Zhu, *et al.*, 1998; Szurek, *et al.*, 2001). Dies könnte darauf hinweisen, dass sie ähnlich wie andere Aktivatoren mit saurer AD wirken. Der Prototyp der sauren AD ist die VP16-AD aus dem *Herpes simplex virus* (Sadowski, *et al.*, 1988, Triezenberg, *et al.*, 1988). Für diese sowie weitere saure ADs (z.B. des Tumorsuppressor Proteins p53 oder des Hefeproteins Gcn4) konnten bereits verschiedene Interaktionen beschrieben werden. Beispielsweise interagieren diese sauren ADs mit Histonacetyltransferasen bzw. Chromatin-remodeling-Komplexen und ermöglichen so die Zugänglichkeit des Promotors für den PIC (Utley, *et al.*, 1998; Ikeda, *et al.*, 1999; Neely, *et al.*, 1999; Barlev, *et al.*, 2001; Hall & Struhl 2002). Außerdem können sie direkt mit GTFs interagieren und somit diverse Schritte des PIC Zusammenbaus beeinflussen. Die VP16-AD und die p53-AD interagieren mit Komponenten des TFIID-Komplexes, zu dem das TATA-Bindeprotein (TBP) sowie TBP-assoziierte Faktoren (TAFs) gehören (Goodrich, *et al.*, 1993; Thut, *et al.*, 1995; Uesugi, *et al.*, 1997; Goodrich & Tjian 2010; Bieniossek, *et al.*, 2013). Durch diese Interaktionen werden die ersten Komponenten des PIC an den Promotor rekrutiert und vermutlich die Bindung von TFIID an die DNA stabilisiert (Thomas & Chiang 2006). Durch die Rekrutierung weiterer Komponenten des PIC z.B. TFIIA durch die VP16- oder p53-AD kann der TFIID-DNA Komplex stabilisiert und der Zusammenbau des PIC weiter gefördert werden. (Stringer, *et al.*, 1990; Ozer, *et al.*, 1994; Kobayashi, *et al.*, 1995; Coleman, *et al.*, 1999; Kraemer, *et al.*, 2001; Xing, *et al.*, 2001; Thomas & Chiang 2006). Zusätzliche Interaktionen mit TFIIB (VP16) und TFIIF (VP16 und p53) können weiter zur Stabilisierung des PICs beitragen und an der Auswahl des Transkriptionsstartpunktes oder der Aktivierung der RNA-Polymerisation beteiligt sein (Lin, *et al.*, 1991; Lin & Green 1991; Xiao, *et al.*, 1994; Hayashi, *et al.*, 1998).

Für die saure AD der TALEs ist bisher weitgehend unklar ob und mit welchen Teilen der Transkriptionsmaschinerie sie interagiert. Hinweise, dass TALEs vergleichbar zu den beschriebenen sauren ADs, ebenfalls mit GTFs interagieren können, sind anhand der *xa5*-Mutation der γ -Untereinheit des TFIIA Alleles auf Chromosom 5 in Reis zu finden (s. Abschnitt 1.2.6.2). Die TFIIA γ Untereinheit konnte bereits in Studien im Human-System als wichtiger Interaktionspartner für die saure VP16-AD beschrieben werden, wodurch sie als Interaktionspartner für TALEs ebenfalls in Frage kommen könnte (Ozer, *et al.*, 1994; Kobayashi, *et al.*, 1995). Der humane, sowie auch

pflanzliche TFIIA setzt sich aus den drei Untereinheiten α , β und γ zusammen und (Li, *et al.*, 1999). Wie auch andere einkeimblättrige Pflanzen kodiert Reis zwei Allele der γ -Untereinheit (Chromosom 1 (TFIIA γ 1) und Chromosom 5 (TFIIA γ 5)) (Iyer-Pascuzzi & McCouch 2007). Das rezessive *xa5* Allel beschreibt einen einzelnen Aminosäureaustausch (V39E) in der γ -Untereinheit des TFIIA auf Chromosom 5. Interessanterweise sind einige TALEs in homozygoten *xa5* Reispflanzen nicht mehr in der Lage Suszeptibilitäts- oder Resistenzgene zu aktivieren (Sugio, *et al.*, 2007; Gu, *et al.*, 2009). Dies lässt vermuten, dass TALEs direkt oder indirekt mit der TFIIA γ Untereinheit interagieren können und diese Interaktion durch die *xa5*-Mutation gestört wird. Ein Aminosäureaustausch (V39 in *Xa5* zu E39 in *xa5*) in der möglichen TALE-Interaktionsfläche von TFIIA γ 5 ist vermutlich eine der wenigen Möglichkeiten eine Resistenz des TFIIA gegenüber TALEs zu vermitteln, da andere Mutationen bzw. die Deletion des TFIIA den generellen Zusammenbau des PIC stören und somit vermutlich letal sind (Ozer, *et al.*, 1994; Kang, *et al.*, 1995; Tan, *et al.*, 1996). Durch die Übertragung des PXO99-TALEs PthXo7 der das alternative TFIIA γ 1 Allel aktiviert konnte die Virulenz von PXO86 wiederhergestellt werden, was darauf hindeutet das TFIIA γ 1 die Funktion von TFIIA γ 5 ersetzen kann (Sugio, *et al.*, 2007; s. Abschnitt 1.2.6.2). Ein vergleichbares Experiment zur Wiederherstellung der Aktivierung von Resistenzgenen wurde jedoch noch nicht durchgeführt. Allerdings ist anzumerken, dass die PthXo7-abhängige TFIIA γ 1 Induktion nur möglich ist, wenn entweder zumindest eine gewisse Grundaktivität von TALEs trotz der *xa5*-Mutation in *TFIIA γ 5* oder eine geringe Basalexpression von *TFIIA γ 1* gegeben ist. Dadurch sollte zumindest anfangs eine gewisse TALE-Aktivität ermöglicht werden, welche anschließend durch die PthXo7-aktivierte Expression von *TFIIA γ 1* verstärkt werden kann. Hinweise dafür sind einerseits darin zu finden, dass beide Allele basal in verschiedenen Organen exprimiert sind und andererseits darin, dass *Xa27* trotz der *xa5*-Mutation *AvrXa27*-abhängig in geringen Mengen in homozygoten Pflanzen aktiviert wird (Jiang, *et al.*, 2006; Gu, *et al.*, 2009). Solche geringen Mengen reichen zwar nicht aus, um eine Resistenzreaktion auszulösen jedoch weist diese Beobachtung daraufhin, dass TALEs trotz *xa5*-Mutation Gene zumindest minimal aktivieren können (Gu, *et al.*, 2005, Iyer-Pascuzzi & McCouch 2007, Gu, *et al.*, 2009; Jiang, *et al.*, 2006). Bisher konnte eine mögliche Interaktion von TALEs mit TFIIA jedoch noch nicht experimentell nachgewiesen werden. Daher ist weiter unklar ob TALEs überhaupt mit TFIIA interagieren und wenn ja, ob die Aminosäure V39 in der TFIIA γ 5 Untereinheit (*Xa5*) tatsächlich an dieser

Interaktion beteiligt ist. Zudem ist ebenfalls unbekannt ob die Interaktion mit TFIIA direkt ist oder über ein bisher unbekanntes Adapterprotein vermittelt wird.

In verschiedenen Studien konnten zudem weitere möglicherweise funktionsrelevante Interaktoren für TALEs mittels Hefe-Di-Hybrid Untersuchungen identifiziert werden. Beispielsweise konnte die Interaktion zwischen AvrBs3 und zwei Homologen *Capsicum annuum* Importin- α Proteinen nachgewiesen werden (Szurek, *et al.*, 2001). Diese Interaktion wird über die NLS in AvrBs3 vermittelt und ist notwendig für den Import von TALEs in den pflanzlichen Zellkern (Szurek, *et al.*, 2001). Weitere Interaktionspartner, die möglicherweise an der Transkription beteiligt sind, konnten in dieser Studie jedoch nicht identifiziert werden. Erst eine weitere Studie identifizierte anhand der vier homologen TALEs der PthA-Familie mögliche Interaktionspartner, die an Prozessen der Chromatin-Umstrukturierung, der Transkriptionsregulation oder an der Stabilisation und Modifikation von RNAs beteiligt sind (Domingues, *et al.*, 2010; Domingues, *et al.*, 2012; de Souza, *et al.*, 2012; Soprano, *et al.*, 2013). Allerdings interagierte die Mehrheit dieser Kandidaten lediglich mit einem oder zwei der vier homologen TALEs (Domingues, *et al.*, 2010; Domingues, *et al.*, 2012; de Souza, *et al.*, 2012; Soprano, *et al.*, 2013). Dies ist in sofern kritisch da die PthA-Homologe bis auf einzelne Aminosäureaustausche hochkonserviert sind und daher unterschiedliche Interaktoren nicht mit Sequenzunterschieden erklärt werden können (Domingues, *et al.*, 2010). Außerdem ist bekannt, dass TALEs generell hochkonserviert und innerhalb eines bzw. zwischen pflanzlichen Systemen austauschbar sind, ohne ihre Aktivierungsfähigkeit zu verlieren (Morbiter, *et al.*, 2010; Boch, *et al.*, 2009; Yu, *et al.*, 2011, Streubel, *et al.*, 2013). Dies lässt vermuten, dass mögliche Interaktionspartner aufgrund der hohen Konservierung möglicher Interaktionsoberflächen z.B. der TALE-AD, mit allen TALEs interagieren sollten. Lediglich das Citrus *high mobility group* (HMG) Protein interagierte mit allen vier PthA-Homologen (de Souza, *et al.*, 2012). Die Funktion der TALE-CsHMG Interaktion wurde durch zwei Möglichkeiten beschrieben. Einerseits war bekannt, dass HMG-Proteine im Human-System mit dem TBP interagieren und so die Transkription hemmen (de Souza, *et al.*, 2012). Diese Interaktion kann durch TFIIA gelöst werden, wodurch die Initiation der Transkription ermöglicht werden könnte. In diesem Zusammenhang könnten TALEs eine ähnliche Rolle wie TFIIA einnehmen und zur effizienten Transkriptionsinitiation hemmende Proteine z.B. HMGs entfernen (de Souza, *et al.*, 2012). Andererseits konnte für CsHMG, sowie für PthA4 eine Bindung an poly(U)RNA nachgewiesen werden (de

Souza, *et al.*, 2012). Daher wurde vermutet, dass der Interaktor CsHMG zusammen mit weiteren identifizierten Interaktoren einen großen Proteinkomplex bildet, der durch TALEs rekrutiert wird und zur Stabilisierung und Prozessierung der entstandenen RNA dient. Dies würde eine bisher noch nicht für TALEs beschriebene Funktion darstellen die zeigt, dass TALEs nicht nur an der Transkriptionsaktivierung sondern auch an posttranskriptionellen Abläufen z.B. der mRNA-Stabilisierung und Reifung beteiligt sein könnten (de Souza, *et al.*, 2012).

3.8.4 TALEs sind keine klassischen *Enhancer*-bindenden Proteine

Die saure VP16-AD aktiviert Gene von proximalen-Kernpromotorelementen ebenso wie von distalen-DNA-Elementen (*Enhancer*) und ist daher vermutlich ein *Enhancer*-bindendes Protein (Seipel, *et al.*, 1992; Maston, *et al.*, 2006). Trotz der sauren AD die TALEs und das VP16-Protein gemeinsam haben, gibt es jedoch Hinweise darauf, dass beide Proteine unterschiedlich mit der Transkriptionsmaschinerie interagieren. Beispielsweise konnte gezeigt werden, dass TALEs mit natürlicher AD im Human-System weniger effizient Gene aktivieren als ein vergleichbares TALE-Derivat mit VP16-AD (Geißler, *et al.*, 2011). Ebenso sind TALE-Derivate mit VP16-AD im pflanzlichen System weniger effizient als TALEs mit natürlicher AD. Dies deutet darauf hin, dass beide ADs jeweils an spezifische humane (VP16) bzw. pflanzliche (TALE) Interaktionspartner der Transkriptionsmaschinerie angepasst sein könnten. Weitere Unterschiede konnten im Human-System anhand von TALE-Derivaten gezeigt werden in denen die natürliche AD durch die VP16-AD ausgetauscht wurde. Diese konnten ein Ziel-Gen von unterschiedlichsten Positionen aktivieren (bis zu 1,5kb *upstream* vom Transkriptionsstart entfernt oder bis zu 500bp in der Protein-kodierenden Sequenz; Perez-Pinera, *et al.*, 2013; Maeder, *et al.*, 2013; Uhde-Stone, *et al.*, 2014). Allerdings wurden in diesen Studien keine Transkriptionsstartpunkte bestimmt. Die Fähigkeit dieser TALE-Derivate trotz weit entfernter Bindestelle oder Bindestelle in der kodierenden Sequenz das Ziel-Gen zu aktivieren deutet jedoch daraufhin, dass sie den Transkriptionsstart nicht beeinflussen, sondern vergleichbar mit *Enhancer*-bindenden Proteinen wirken, die vermutlich durch Interaktion mit anderen Proteinen z.B. dem Mediator Kontakt zum PIC am Kernpromotor herstellen. Eine solche Interaktion z.B. der VP16-AD mit dem Mediator konnte bereits beschrieben werden (Milbradt, *et al.*, 2011; Vojnic, *et al.*, 2011). Im Vergleich dazu konnte im Rahmen dieser Arbeit gezeigt werden, dass natürliche TALEs zwar ebenfalls von verschiedensten Positionen und

Orientierungen im Promotor wirken können, dabei allerdings den Transkriptionsstart verschieben (Abb. 2.1, 2.5 & 2.6). Dies deutet daraufhin, dass sie im Gegensatz zur VP16 nicht wie klassische *Enhancer*-bindende Proteine wirken können, sondern ihre Funktion als Virulenzfaktor nur dann erfüllen, wenn sie so positioniert sind, dass die entstehende mRNA auch zu einem funktionalen Protein führt. Weitere Unterstützung dafür kann darin gefunden werden, dass anhand einer Computer-basierten Vorhersage natürlicher TALEs in Reis gezeigt werden konnte, dass deren Bindestellen in Ziel-Genen häufig in einem Bereich von ca. 300bp *upstream* bzw. bis zu 200bp *downstream* um den Transkriptionsstartpunkt liegen (Grau, *et al.*, 2013). Diese Positionierung von TALEs im Kernpromotor sichert möglicherweise die Entstehung einer mRNA, die zur Synthese eines funktionalen Proteins führt. Auch die damit verbundene Nähe zur TATA-Box und möglicherweise weiteren regulatorischen Sequenzen des Kernpromotors könnte vorteilhaft sein, um die effiziente Rekrutierung des PICs und somit die effiziente Genaktivierung zu ermöglichen (Scholze 2011; Kumari & Ware 2013). Welche regulatorischen Kernpromotorelemente neben der TATA-Box jedoch für die TALE-vermittelte Genaktivierung wichtig sind ist bisher weitgehend unklar (Scholze 2011).

Zusammenfassend konnte gezeigt werden, dass TALEs von verschiedensten Positionen im Promotor und sogar in *reverse* Orientierung funktionieren, wobei deren Positionierung in Kernpromotorregionen bzw. in der Nähe des natürlichen Transkriptionsstarts durchaus vorteilhaft ist, um die Synthese eines funktionalen Proteins zu gewährleisten. Besonders die Erkenntnis, dass TALEs mit natürlicher AD *reverse*-positioniert Gene aktivieren können, ist für alle bisher verwendeten Programme zur Vorhersage möglicher TALE Ziel-Gene bzw. möglicher *off-targets* für künstliche TALEs von Interesse, da in diese Analyse bisher lediglich *forward*-positionierte Bindestellen einbezogen werden. Außerdem geben die in dieser Arbeit gewonnenen Kenntnisse Hinweise für die Auswahl von Bindestellen künstlicher TALEs zur gezielten Aktivierung pflanzlicher Gene. Die in dieser Arbeit gewonnenen Erkenntnisse sowie zukünftige Analysen regulatorischer Promotorelemente oder Interaktoren die relevant für die TALE-abhängige Ziel-Genaktivierung sind, könnten in Zukunft Einblicke in die TALE-vermittelte Transkriptionsaktivierung geben und möglicherweise auch Rückschlüsse auf die generelle pflanzliche Transkriptionsaktivierung zulassen. Offensichtlich kann auch die Zugänglichkeit des Promotors die Bindung des TALEs beeinflussen. Genauere Analysen ob und wie TALEs mit Wirtsproteinen am Promotor

konkurrieren, könnten zukünftig nicht nur die Anwendung von TALE-Aktivatoren, sondern auch die Anwendung anderer TALE-basierter biotechnologischer Werkzeuge verbessern und möglicherweise sogar auf andere DNA-bindende biotechnologische Werkzeuge übertragen werden.

Literaturverzeichnis

- Akimoto-Tomiya, C., Furutani, A. and Ochiai, H.** 2014. Real time live imaging of phytopathogenic bacteria *Xanthomonas campestris* pv. *campestris* MAFF106712 in 'plant sweet home'. *PLoS ONE* **9(4)**: e94386
- Alfano, J. R. and Collmer, A.** 2004. Type III secretion effector proteins: double agents in bacterial disease and plant defense. *Annual Review of Phytopathology* **42**: 385-414
- Amano, T., Sagai, T., Tanabe, H., Mizushima, Y., Nakazawa, H. and Shiroishi, T.** 2009. Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Developmental Cell* **16**: 47-57
- Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F. and Yang, B.** 2010. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* **22**: 3864-3876
- Bai, J., Choi, S.-H., Ponciano, G., Leung, H. and Leach, J. E.** 2000. *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. *Molecular Plant-Microbe Interactions* **13**: 1322-1329
- Baker, R. F., Leach, K. A. and Braun, D. M.** 2012. SWEET as sugar: new sucrose effluxers in plants. *Molecular Plant* **5**: 766-8
- Barlev, N. A., Liu, L., Chehab, N. H., Mansfield, K., Harris, K. G., Halazonetis, T. D. and Berger, S. L.** 2001. Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Molecular Cell* **8**: 1243-1254
- Becker, A., Katzen, F., Puhler, A. and Ielpi, L.** 1998. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Applied Microbiology and Biotechnology* **50**: 145-152
- Biemelt, S. and Sonnewald, U.** 2006. Plant-microbe interactions to probe regulation of plant carbon metabolism. *Journal of Plant Physiology* **163**: 307-318
- Bieniossek, C., et al.,** 2013. The architecture of human general transcription factor TFIID core complex. *Nature* **493**: 699-702
- Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., Vasse, J., Lauber, E. and Arlat, M.** 2007. Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**: e224
- Block, A., Li, G., Fu, Z. Q. and Alfano, J. R.** 2008. Phytopathogen type III effector weaponry and their plant targets. *Current Opinion in Plant Biology* **11**: 396-403

- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U.** 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509-1512
- Boch, J. and Bonas, U.** 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* **48**: 419-436
- Bogdanove, A. J. and Voytas, D. F.** 2011. TAL effectors: customizable proteins for DNA targeting. *Science* **333**: 1843-1846
- Bogdanove, A. J.** 2014. Principles and applications of TAL effectors for plant physiology and metabolism. *Current Opinion in Plant Biology* **19**: 99-104
- Boller, T. and He, S. Y.** 2009. Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* **324**: 742-744
- Bonas, U., Stall, R. E. and Staskawicz, B. J.** 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet* **218**: 127-136
- Bonas, U., Conrads-Strauch, J. and Balbo, I.** 1993. Resistance in tomato to *Xanthomonas campestris* pv. *vesicatoria* is determined by alleles of the pepper-specific avirulence gene *avrBs3*. *Molecular and General Genetics* **238**: 261-269
- Booher, N. J. and Bogdanove, A. J.** 2014. Tools for TAL effector design and target prediction. *Methods* **69**: 121-127
- Braun, D. M.** 2012. SWEET! The pathway is complete. *Science* **335**: 173-174
- Brunings, A. M. and Gabriel, D. W.** 2003. *Xanthomonas citri*: breaking the surface. *Molecular plant pathology* **4**: 141-157
- Bulger, M. and Groudine, M.** 2011. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**: 327-339
- Bultmann, S., Morbitzer, R., Schmidt, C. S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T. and Leonhardt, H.** 2012. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Research* **40**: 5368-5377
- Burrill, T. J.** 1878. Report on botany and vegetable physiology. *Transactions of the Illinois State Horticultural Society* **11**: 114-116
- Burrill, T. J.** 1881. Bacteria as a cause of disease in plants. *American Naturalist* **15**: 527-531
- Büttner, D. and Bonas, U.** 2002. Port of entry - the type III secretion translocon. *Trends in Microbiology* **10**: 186-192

- Carroll, D.** 2011. Genome engineering with zinc-finger nucleases. *Genetics* **188**: 773-782
- Cernadas, R. A., et al.,** 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathogens* **10**: e1003972
- Chen, L. Q., et al.,** 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**: 527-532
- Chen, L. Q., Qu, X. Q., Hou, B. H., Sosso, D., Osorio, S., Fernie, A. R. and Frommer, W. B.** 2012. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**: 207-211
- Chong, J., Piron, M. C., Meyer, S., Merdinoglu, D., Bertsch, C. and Mestre, P.** 2014. The SWEET family of sugar transporters in grapevine: VvSWEET4 is involved in the interaction with *Botrytis cinerea*. *Journal of Experimental Botany* **65**: 6589-6601
- Christian, M. L., et al.,** 2012. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS ONE* **7**: e45383
- Chu, Z., et al.,** 2006a. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes and Development* **20**: 1250-1255
- Chu, Z., et al.,** 2006b. Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. *Theoretical and Applied Genetics* **112**: 455-461
- Civan, P. and Svec, M.** 2009. Genome-wide analysis of rice (*Oryza sativa* L. subsp. *japonica*) TATA box and Y Patch promoter elements. *Genome* **52**: 294-297
- Cohn, M., et al.,** 2014. *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in Cassava. *Molecular Plant Microbe Interactions* **27**: 1186-1198
- Cokus, S. J., et al.,** 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**: 215-219
- Coleman, R. A., Taggart, A. K., Burma, S., Chicca, J. J., 2nd and Pugh, B. F.** 1999. TFIIA regulates TBP and TFIID dimers. *Mol Cell* **4**: 451-457
- Cong, L., Zhou, R., Kuo, Y. C., Cunniff, M. and Zhang, F.** 2012. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature Communications* **3**: 968
- Cunnac, S., Occhialini, A., Barberis, P., Boucher, C. and Genin, S.** 2004. Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum*.

identification of novel effector proteins translocated to plant host cells through the type III secretion system. *Molecular Microbiology* **53**: 115-128

de Lange, O., Schreiber, T., Schandry, N., Radeck, J., Braun, K. H., Koszinowski, J., Heuer, H., Strauss, A. and Lahaye, T. 2013. Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytologist* **199**: 773-786

de Souza, T. A., Soprano, A. S., de Lira, N. P., Quaresma, A. J., Pauletti, B. A., Paes Leme, A. F. and Benedetti, C. E. 2012. The TAL effector PthA4 interacts with nuclear factors involved in RNA-dependent processes including a HMG protein that selectively binds poly(U) RNA. *PLoS ONE* **7**: e32305

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J. K., Shi, Y. and Yan, N. 2012a. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**: 720-723

Deng, D., et al., 2012b. Recognition of methylated DNA by TAL effectors. *Cell Research* **22**: 1502-1504

Deng, D., Yan, C. Y., Wu, J. P., Pan, X. J. and Yan, N. 2014. Revisiting the TALE repeat. *Protein & Cell* **5**: 297-306

Divon, H. H. and Fluhr, R. 2007. Nutrition acquisition strategies during fungal infection of plants. *FEMS Microbiology Letters* **266**: 65-74

Dodds, P. N. and Rathjen, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics* **11**: 539-548

Domingues, M. N., De Souza, T. A., Cernadas, R. A., de Oliveira, M. L., Docena, C., Farah, C. S. and Benedetti, C. E. 2010. The *Xanthomonas citri* effector protein PthA interacts with citrus proteins involved in nuclear transport, protein folding and ubiquitination associated with DNA repair. *Molecular Plant Pathology* **11**: 663-675

Domingues, M. N., Campos, B. M., de Oliveira, M. L., de Mello, U. Q. and Benedetti, C. E. 2012. 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

Duan, Y. P., Castañeda, A., Zhao, G., Erdos, G. and Gabriel, D. W. 1999. Expression of a single, host-specific, bacterial pathogenicity gene in plant cells elicits division, enlargement, and cell death. *Molecular Plant-Microbe Interactions* **12**: 556-560

Gamas, P., Niebel Fde, C., Lescure, N. and Cullimore, J. 1996. Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Molecular Plant-Microbe Interactions* **9**: 233-242

- Gannon, F., et al.,** 1979. Organisation and sequences at the 5' end of a cloned complete ovalbumin gene. *Nature* **278**: 428-434
- Gao, H., Wu, X., Chai, J. and Han, Z.** 2012. Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Research* **22**: 1716-1720
- Geißler, R., Scholze, H., Hahn, S., Streubel, J., Bonas, U., Behrens, S.-E. and Boch, J.** 2011. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**: e19509
- Goff, S. A., et al.,** 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**: 92-100
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A. and Tjian, R.** 1993. Drosophila TAFII40 interacts with both a Vp16 activation domain and the basal transcription factor-TFIIB. *Cell* **75**: 519-530
- Goodrich, J. A. and Tjian, R.** 2010. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nature Reviews Genetics* **11**: 549-558
- Grant, M., Brown, I., Adams, S., Knight, M., Ainslie, A. and Mansfield, J.** 2000. The *RPM1* plant disease resistance gene facilitates a rapid and sustained increase in cytosolic calcium that is necessary for the oxidative burst and hypersensitive cell death. *The Plant Journal* **23**: 441-450
- Grau, J., Wolf, A., Reschke, M., Bonas, U., Posch, S. and Boch, J.** 2013. Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Computational Biology* **9**: e1002962
- Grünberg, S. and Hahn, S.** 2013. Structural insights into transcription initiation by RNA polymerase II. *Trends in Biochemical Sciences* **38**: 603-611
- Gu, K., et al.,** 2005. *R* gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122
- Gu, K., Tian, D., Qiu, C. and Yin, Z.** 2009. Transcription activator-like type III effector AvrXa27 depends on OsTFIIA γ 5 for the activation of *Xa27* transcription in rice that triggers disease resistance to *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology* **10**: 829-835
- Guo, W. J., Nagy, R., Chen, H. Y., Pfrunder, S., Yu, Y. C., Santelia, D., Frommer, W. B. and Martinoia, E.** 2014. SWEET17, a facilitative transporter, mediates fructose transport across the tonoplast of Arabidopsis roots and leaves. *Plant Physiology* **164**: 777-789

- Hall, D. B. and Struhl, K.** 2002. The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking *in vivo*. *Journal of Biological Chemistry* **277**: 46043-46050
- Hayashi, F., Ishima, R., Liu, D., Tong, K. I., Kim, S., Reinberg, D., Bagby, S. and Ikura, M.** 1998. Human general transcription factor TFIIB: conformational variability and interaction with VP16 activation domain. *Biochemistry* **37**: 7941-7951
- He, Y., Fang, J., Taatjes, D. J. and Nogales, E.** 2013. Structural visualization of key steps in human transcription initiation. *Nature* **495**: 481-486
- Heath, M. C.** 2000. Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology* **3**: 315-319
- Herbers, K., Conrads-Strauch, J. and Bonas, U.** 1992. Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature* **356**: 172-174
- Heuer, H., Yin, Y.-N., Xue, Q.-Y., Smalla, K. and Guo, J.-H.** 2007. Repeat domain diversity of *avrBs3*-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Applied and Environmental Microbiology* **73**: 4379-4384
- Howlett, B. J.** 2006. Secondary metabolite toxins and nutrition of plant pathogenic fungi. *Current Opinion in Plant Biology* **9**: 371-375
- Hu, Y., et al.,.** 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 521-529
- Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S. and Zhang, B.** 2011. Heritable gene targeting in zebrafish using customized TALENs. *Nature Biotechnology* **29**: 699-700
- Hückelhoven, R.** 2007. Cell wall-associated mechanisms of disease resistance and susceptibility. *Annual Review of Phytopathology* **45**: 101-127
- Hummel, A. W., Doyle, E. L. and Bogdanove, A. J.** 2012. Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**: 883-893
- Ikeda, K., Steger, D. J., Eberharter, A. and Workman, J. L.** 1999. Activation domain-specific and general transcription stimulation by native histone acetyltransferase complexes. *Molecular and Cellular Biology* **19**: 855-863
- Ishida, K., Yamashino, T., Yokoyama, A. and Mizuno, T.** 2008. Three type-B response regulators, ARR1, ARR10 and ARR12, play essential but redundant roles in

cytokinin signal transduction throughout the life cycle of *Arabidopsis thaliana*. *Plant and Cell Physiology* **49**: 47-57

Iyer, A. S. and McCouch, S. R. 2004. The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. *Molecular Plant-Microbe Interactions* **17**: 1348-1354

Iyer-Pascuzzi, A. S. and McCouch, S. R. 2007. Recessive resistance genes and the *Oryza sativa-Xanthomonas oryzae* pv. *oryzae* pathosystem. *Molecular Plant-Microbe Interactions* **20**: 731-739

Iyer-Pascuzzi, A. S., Jiang, H., Huang, L. and McCouch, S. R. 2008. Genetic and functional characterization of the rice bacterial blight disease resistance gene *xa5*. *Phytopathology* **98**: 289-295

Izawa, T. and Shimamoto, K. 1996. Becoming a model plant: The importance of rice to plant science. *Trends In Plant Science* **1**: 95-99

Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M. and Epstein, D. J. 2006. A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* **133**: 761-772

Jiang, G. H., Xia, Z. H., Zhou, Y. L., Wan, J., Li, D. Y., Chen, R. S., Zhai, W. X. and Zhu, L. H. 2006. Testifying the rice bacterial blight resistance gene *xa5* by genetic complementation and further analyzing *xa5* (*Xa5*) in comparison with its homolog *TFIIAyl*. *Molecular Genetics and Genomics* **275**:354-366

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T. L. 2008. NCBI BLAST: a better web interface. *Nucleic Acids Research* **36**: W5-9

Jones, J. D. and Dangl, J. L. 2006. The plant immune system. *Nature* **444**: 323-329

Jones, P. A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* **13**: 484-492

Juillerat, A., et al.,. 2015. Optimized tuning of TALEN specificity using non-conventional RVDs. *Scientific Reports* **5**: 8150

Kang, J. J., Auble, D. T., Ranish, J. A. and Hahn, S. 1995. Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Molecular and Cellular Biology* **15**: 1234-1243

Kay, S., Boch, J. and Bonas, U. 2005. Characterization of AvrBs3-like effectors from a *Brassicaceae* pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Molecular Plant-Microbe Interactions* **18**: 838-848

- Kay, S., Hahn, S., Marois, E., Hause, G. and Bonas, U.** 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**: 648-651
- Kay, S., Hahn, S., Marois, E., Wieduwild, R. and Bonas, U.** 2009. Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3 Δ rep16. *Plant J* **59**: 859-871
- Keenan, R. J., Freymann, D. M., Stroud, R. M. and Walter, P.** 2001. The signal recognition particle. *Annual Review of Biochemistry* **70**: 755-775
- Kennedy, D.** 2002. The importance of rice. *Science* **296**: 13
- Klug, A.** 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry* **79**: 213-231
- Kobayashi, N., Boyer, T. G. and Berk, A. J.** 1995. A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. *Molecular and Cellular Biology* **15**: 6465-6473
- Kocal, N., Sonnewald, U. and Sonnewald, S.** 2008. Cell wall-bound invertase limits sucrose export and is involved in symptom development and inhibition of photosynthesis during compatible interaction between tomato and *Xanthomonas campestris* pv. *vesicatoria*. *Plant Physiol.* **148**: 1523-1536
- Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M. and Cramer, P.** 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**: 323-330
- Kraemer, S. M., Ranallo, R. T., Ogg, R. C. and Stargell, L. A.** 2001. TFIIA interacts with TFIID via association with TATA-binding protein and TAF40. *Molecular and Cellular Biology* **21**: 1737-46
- Kubik, G., Schmidt, M. J., Penner, J. E. and Summerer, D.** 2014. Programmable and highly resolved *in vitro* detection of 5-methylcytosine by TALEs. *Angewandte Chemie International Edition* **53**: 6002-6006
- Kubik, G. and Summerer, D.** 2015. Achieving Single-Nucleotide Resolution of 5-Methylcytosine Detection with TALEs. *ChemBioChem* **16**: 228-231
- Kühn, C. and Grof, C. P.** 2010. Sucrose transporters of higher plants. *Current Opinion in Plant Biology* **13**: 288-298
- Kumari, S. and Ware, D.** 2013. Genome-wide computational prediction and analysis of core promoter elements across plant monocots and dicots. *PLoS ONE* **8**: e79011

- Kutach, A. K. and Kadonaga, J. T.** 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Molecular and Cellular Biology* **20**: 4754-4764
- Lackner, G., Moebius, N., Partida-Martinez, L. and Hertweck, C.** 2011. Complete genome sequence of *Burkholderia rhizoxinica*, an Endosymbiont of *Rhizopus microsporus*. *Journal of Bacteriology* **193**: 783-784
- Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. and Rombauts, S.** 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* **30**: 325-327
- Levine, M., Cattoglio, C. and Tjian, R.** 2014. Looping back to leap forward: transcription enters a new era. *Cell* **157**: 13-25
- Li, Y. F., Le Gourierrec, J., Torki, M., Kim, Y. J., Guerineau, F. and Zhou, D. X.** 1999. Characterization and functional analysis of Arabidopsis TFIIA reveal that the evolutionarily unconserved region of the large subunit has a transcription activation domain. *Plant Molecular Biology* **39**: 515-525
- Lin, I. W., et al.,** 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546-549
- Lin, J. Y., Chen, J. D., Elenbaas, B. and Levine, A. J.** 1994. Several hydrophobic amino-acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus-5 E1b 55-kd protein. *Genes & Development* **8**: 1235-1246
- Lin, Y. S. and Green, M. R.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. *Cell* **64**: 971-981
- Lin, Y. S., Ha, I., Maldonado, E., Reinberg, D. and Green, M. R.** 1991. Binding of general transcription factor TFIIB to an acidic activating region. *Nature* **353**: 569-571
- Liu, Q., Yuan, M., Zhou, Y., Li, X., Xiao, J. and Wang, S.** 2011. A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant Cell Environ* **34**: 1958-69
- Liu, W., Liu, J., Triplett, L., Leach, J. E. and Wang, G.-L.** 2014. Novel insights into rice innate immunity against bacterial and fungal pathogens. *Annual Review of Phytopathology* **52**: 213-241
- Liu, X., Bushnell, D. A., Wang, D., Calero, G. and Kornberg, R. D.** 2010. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327**: 206-209

- Maeder, M. L., Linder, S. J., Reyon, D., Angstman, J. F., Fu, Y., Sander, J. D. and Joung, J. K.** 2013. Robust, synergistic regulation of human gene expression using TALE activators. *Nature Methods* **10**: 243-245
- Mak, A. N., Bradley, P., Cernadas, R. A., Bogdanove, A. J. and Stoddard, B. L.** 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**: 716-719
- Mansfield, J., et al.,.** 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology* **13**: 614-629
- Marois, E., Van den Ackerveken, G. and Bonas, U.** 2002. The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Molecular Plant-Microbe Interactions* **15**: 637-646
- Maston, G. A., Evans, S. K. and Green, M. R.** 2006. Transcriptional regulatory elements in the human genome. *Annual Review of Genomics and Human Genetics* **7**: 29-59
- Meckler, J. F., et al.,.** 2013. Quantitative analysis of TALE–DNA interactions suggests polarity effects. *Nucleic Acids Research* **41**: 4118-4128
- Mew, T. W., Alvarez, A. M., Leach, J. E. and Swings, J.** 1993. Focus on bacterial-blight of rice. *Plant Disease* **77**: 5-12
- Milbradt, A. G., Kulkarni, M., Yi, T., Takeuchi, K., Sun, Z. Y., Luna, R. E., Selenko, P., Naar, A. M. and Wagner, G.** 2011. Structure of the VP16 transactivator target in the Mediator. *Nature Structural and Molecular Biology* **18**: 410-415
- Miller, J. C., et al.,.** 2011. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* **29**: 143-148
- Morbitzer, R., Römer, P., Boch, J. and Lahaye, T.** 2010. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 21617-21622
- Moscou, M. J. and Bogdanove, A. J.** 2009. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**: 1501
- Murakami, M. T., Sforça, M. L., Neves, J. L., Paiva, J. H., Domingues, M. N., Pereira, A. L., de Mattos Zeri, A. C. and Benedetti, C. E.** 2010. The repeat domain of the type III effector protein PthA shows a TPR-like structure and undergoes conformational changes upon DNA interaction. *Proteins* **78**: 3386-3395
- Neely, K. E., Hassan, A. H., Wallberg, A. E., Steger, D. J., Cairns, B. R., Wright, A. P. and Workman, J. L.** 1999. Activation domain-mediated targeting of the SWI/SNF

complex to promoters stimulates transcription from nucleosome arrays. *Molecular Cell* **4**: 649-655

Nino-Liu, D. O., Ronald, P. C. and Bogdanove, A. J. 2006. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Molecular Plant Pathology* **7**: 303-324

Nürnberg, T., Brunner, F., Kemmerling, B. and Piater, L. 2004. Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews* **198**: 249-266

Orphanides, G., Lagrange, T. and Reinberg, D. 1996. The general transcription factors of RNA polymerase II. *Genes and Development* **10**: 2657-2683

Ozer, J., Moore, P. A., Bolden, A. H., Lee, A., Rosen, C. A. and Lieberman, P. M. 1994. Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription. *Genes and Development* **8**: 2324-35

Pereira, A. L., Carazzolle, M. F., Abe, V. Y., de Oliveira, M. L., Domingues, M. N., Silva, J. C., Cernadas, R. A. and Benedetti, C. E. 2014. Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. *BMC Genomics* **15**: 157

Perez-Pinera, P., Ousterout, D. G., Brunger, J. M., Farin, A. M., Glass, K. A., Guilak, F., Crawford, G. E., Hartemink, A. J. and Gersbach, C. A. 2013. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nature Methods* **10**: 239-42

Perez-Quintero, A. L., Rodriguez, R. L., Dereeper, A., Lopez, C., Koebnik, R., Szurek, B. and Cunnac, S. 2013. An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLoS ONE* **8**: e68464

Quirino, B. F., Normanly, J. and Amasino, R. M. 1999. Diverse range of gene activity during *Arabidopsis thaliana* leaf senescence includes pathogen-independent induction of defense-related genes. *Plant Molecular Biology* **40**: 267-278

Rajeshwari, R., Jha, G. and Sonti, R. V. 2005. Role of an *in planta*-expressed xylanase of *Xanthomonas oryzae* pv. *oryzae* in promoting virulence on rice. *Molecular Plant-Microbe Interactions* **18**: 830-837

Richter, A., Streubel, J., Blucher, C., Szurek, B., Reschke, M., Grau, J. and Boch, J. 2014. A TAL effector repeat architecture for frameshift binding. *Nature Communications* **5**: 3447

Roitsch, T., Balibrea, M. E., Hofmann, M., Proels, R. and Sinha, A. K. 2003. Extracellular invertase: key metabolic enzyme and PR protein. *Journal of Experimental Botany* **54**: 513-24

- Rolland, F., Baena-Gonzalez, E. and Sheen, J.** 2006. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annual Review of Plant Biology* **57**: 675-709
- Rombauts, S., Dehais, P., Van Montagu, M. and Rouze, P.** 1999. PlantCARE, a plant *cis*-acting regulatory element database. *Nucleic Acids Research* **27**: 295-296
- Römer, P., Hahn, S., Jordan, T., Strauß, T., Bonas, U. and Lahaye, T.** 2007. Plant-pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science* **318**: 645-648
- Römer, P., Strauß, T., Hahn, S., Scholze, H., Morbitzer, R., Grau, J., Bonas, U. and Lahaye, T.** 2009. Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. *Plant Physiology* **150**: 1697-1712
- Rose, A., Meier, I. and Wienand, U.** 1999. The tomato I-box binding factor LeMYBI is a member of a novel class of Myb-like proteins. *Plant Journal* **20**: 641-652
- Rouxel, T. and Balesdent, M.-H.** 2010. Avirulence genes. *John Wiley & Sons, Ltd*
- Ruan, Y. L.** 2014. Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annual Review of Plant Biology* **65**: 33-67
- Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M.** 1988. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**: 563-564
- Saze, H., Tsugane, K., Kanno, T. and Nishimura, T.** 2012. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant and Cell Physiology* **53**: 766-784
- Scholze, H. and Boch, J.** 2011. TAL effectors are remote controls for gene activation. *Current Opinion in Microbiology* **14**: 47-53
- Scholze, H.** 2011. DNA-Bindesepezifität von TAL-Effektoren aus Xanthomonaden. *Dissertation vorgelegt an der Naturwissenschaftlichen Fakultät I Biowissenschaften der Martin-Luther-Universität Halle-Wittenberg*
- Schornack, S., Ballvora, A., Gürlebeck, D., Peart, J., Baulcombe, D., Ganai, M., Baker, B., Bonas, U. and Lahaye, T.** 2004. The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *Plant Journal* **37**: 46-60
- Schornack, S., Meyer, A., Römer, P., Jordan, T. and Lahaye, T.** 2006. Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. *Journal of Plant Physiology* **163**: 256-72
- Schornack, S., Minsavage, G. V., Stall, R. E., Jones, J. B. and Lahaye, T.** 2008. Characterization of AvrHah1, a novel AvrBs3-like effector from *Xanthomonas gardneri* with virulence and avirulence activity. *New Phytologist* **179**: 546-556

- Schornack, S., Moscou, M. J., Ward, E. R. and Horvath, D. M.** 2013. Engineering plant disease resistance based on TAL effectors. *Annual Review of Phytopathology* **51**: 383-406
- Seipel, K., Georgiev, O. and Schaffner, W.** 1992. Different activation domains stimulate transcription from remote (enhancer) and proximal (promoter) positions. *Embo Journal* **11**: 4961-4968
- Senthil-Kumar, M. and Mysore, K. S.** 2013. Nonhost resistance against bacterial pathogens: retrospectives and prospects. *Annual Review of Phytopathology* **51**: 407-427
- Shen, F., Triezenberg, S. J., Hensley, P., Porter, D. and Knutson, J. R.** 1996. Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. *Journal of Biological Chemistry* **271**: 4827-4837
- Sonnewald, S., Priller, J. P., Schuster, J., Glickmann, E., Hajirezaei, M. R., Siebig, S., Mudgett, M. B. and Sonnewald, U.** 2012. Regulation of cell wall-bound invertase in pepper leaves by *Xanthomonas campestris* pv. *vesicatoria* type three effectors. *PLoS ONE* **7**: e51763
- Sonnewald, U.** 2011. SWEETS-The missing sugar efflux carriers. *Frontiers in Plant Science* **2**: 7
- Soprano, A. S., Abe, V. Y., Smetana, J. H. and Benedetti, C. E.** 2013. Citrus MAF1, a repressor of RNA polymerase III, binds the *Xanthomonas citri* canker elicitor PthA4 and suppresses citrus canker development. *Plant Physiology* **163**: 232-242
- Sperry, J. S.** 2003. Evolution of water transport and xylem structure. *International Journal of Plant Sciences* **164**: 115-127
- Streubel, J.** 2011. DNA Spezifität von TAL-Effektoren. *Diplomarbeit, Martin-Luther-Universität Halle-Wittenberg*
- Streubel, J., Blücher, C., Landgraf, A. and Boch, J.** 2012. TAL effector RVD specificities and efficiencies. *Nature Biotechnology* **30**: 593-595
- Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J. and Szurek, B.** 2013. Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist* **200**: 808-819
- Stringer, K. F., Ingles, C. J. and Greenblatt, J.** 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* **345**: 783-786
- Sugio, A., Yang, B., Zhu, T. and White, F. F.** 2007. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OstFIIAγ1* and

OsTFXI during bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 10720-10725

Sun, L., Yang, D. L., Kong, Y., Chen, Y., Li, X. Z., Zeng, L. J., Li, Q., Wang, E. T. and He, Z. H. 2014. Sugar homeostasis mediated by cell wall invertase GRAIN INCOMPLETE FILLING 1 (GIF1) plays a role in pre-existing and induced defence in rice. *Molecular Plant Pathology* **15**: 161-73

Sun, Q. H., Hu, J., Huang, G. X., Ge, C., Fang, R. X. and He, C. Z. 2005. Type-II secretion pathway structural gene *xpsE*, xylanase- and cellulase secretion and virulence in *Xanthomonas oryzae* pv. *oryzae*. *Plant Pathology* **54**: 15-21

Swarup, S., de Feyter, R., Brlansky, R. H. and Gabriel, D. W. 1991. A pathogenicity locus from *Xanthomonas citri* enables strains from several pathovars of *X. campestris* to elicit cankerlike lesions on citrus. *Phytopathology* **81**: 802-809

Szurek, B., Marois, E., Bonas, U. and Van den Ackerveken, G. 2001. Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. *Plant Journal* **26**: 523-534

Szurek, B., Rossier, O., Hause, G. and Bonas, U. 2002. Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Molecular Microbiology* **46**: 13-23

Tan, S., Hunziker, Y., Sargent, D. F. and Richmond, T. J. 1996. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**: 127-151

Thomas, M. C. and Chiang, C. M. 2006. The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology* **41**: 105-178

Thut, C. J., Chen, J. L., Klemm, R. and Tjian, R. 1995. P53 Transcriptional Activation Mediated by Coactivators TAF(II)40 and TAF(II)60. *Science* **267**: 100-104

Tian, D. S., et al., 2014. The rice TAL effector-dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. *Plant Cell* **26**: 497-515

Triezenberg, S. J., Kingsbury, R. C. and McKnight, S. L. 1988. Functional dissection of VP16, the trans-activator of *herpes simplex virus* immediate early gene expression. *Genes and Development* **2**: 718-729

Triezenberg, S. J. 1995. Structure and function of transcriptional activation domains. *Current Opinion in Genetics and Development* **5**: 190-196

- Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., Goodwin, M. J., Aryee, M. J. and Joung, J. K.** 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnology* **32**: 569-576
- Tsuge, S., Furutani, A., Kubo, Y. and Horino, O.** 2001. Identification of a H⁺/glucose and galactose symporter gene *glt* from *Xanthomonas oryzae* pv. *oryzae*. *Microbiology Immunology* **45**: 543-547
- Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. and Verdine, G. L.** 1997. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277**: 1310-1313
- Uhde-Stone, C., Cheung, E. and Lu, B.** 2014. TALE activators regulate gene expression in a position- and strand-dependent manner in mammalian cells. *Biochemical and Biophysical Research Communications* **443**: 1189-1194
- Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S. and Workman, J. L.** 1998. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**: 498-502
- Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Marechal, A., Macmaster, R., Melliand, K., Juillerat, A. and Duchateau, P.** 2012. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *Journal of Biological Chemistry* **287**: 38427-38432
- Van den Ackerveken, G., Marois, E. and Bonas, U.** 1996. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell* **87**: 1307-1316
- Van der Biezen, E. A. and Jones, J. D. G.** 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences* **23**: 454-456
- van Leeuwenhoek, A.** 1677. Concerning little animals by him observed in rain-well sea and snow water; as also in water wherein pepper had lain infused. . *Philosophical Transactions of the Royal Society* **12**: 821-831
- Vera Cruz, C. M., Bai, J., Ona, I., Leung, H., Nelson, R. J., Mew, T.-W. and Leach, J. E.** 2000. Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation *Proceedings of the National Academy of Sciences of the United States of America* **97**: 13500-13505
- Verdier, V., Triplett, L. R., Hummel, A. W., Corral, R., Cernadas, R. A., Schmidt, C. L., Bogdanove, A. J. and Leach, J. E.** 2012. 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 Phytologist* **196**: 1197-1207
- Vogel, J.** 2008. Unique aspects of the grass cell wall. *Current Opinion in Plant Biology* **11**: 301-307
- Vojnic, E., et al.,** 2011. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat Structural and Molecular Biology* **18**: 404-409
- Wahl, R., Wippel, K., Goos, S., Kamper, J. and Sauer, N.** 2010. A novel high-affinity sucrose transporter is required for virulence of the plant pathogen *Ustilago maydis*. *PLoS Biology* **8**: e1000303
- Wang, C. , Qin, T. F., Yu, H. M., Zhang, X. P., Che, J. Y., Gao, Y., Zheng, C. K., Yang, B. and Zhao, K. J.** 2014a. The broad bacterial blight resistance of rice line CBB23 is triggered by a novel transcription activator-like (TAL) effector of *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology* **15**: 333-341
- Wang, C., Fan, Y., Zheng, C., Qin, T., Zhang, X. and Zhao, K.** 2014b. High-resolution genetic mapping of rice bacterial blight resistance gene *Xa23*. *Molecular Genetics and Genomics* **289**: 745-753
- Wang, C., et al.,** 2014c. XA23 is an executor R protein and confers broad-spectrum disease resistance in rice. *Molecular Plant*
- White, F. F. and Yang, B.** 2009. Host and pathogen factors controlling the rice-*Xanthomonas oryzae* interaction. *Plant Physiology* **150**: 1677-1686
- Wichmann, G. and Bergelson, J.** 2004. Effector genes of *Xanthomonas axonopodis* pv. *vesicatoria* promote transmission and enhance other fitness traits in the field. *Genetics* **166**: 693-706
- Wicky, B. I., Stenta, M. and Dal Peraro, M.** 2013. TAL effectors specificity stems from negative discrimination. *PLoS ONE* **8**: e80261
- Xia, C., Hongqi, C. and Zhu, X.** 2012. Identification, mapping and Isolation of the genes resisting to bacterial blight and application in rice. *Molecular Plant Breeding* **3**: 120-130
- Xiao, H., et al.,** 1994. Binding of basal transcription factor TFIID to the acidic activation domains of VP16 and p53. *Molecular and Cellular Biology* **14**: 7013-7024
- Xing, J., Sheppard, H. M., Corneillie, S. I. and Liu, X.** 2001. p53 Stimulates TFIID-TFIIA-promoter complex assembly, and p53-T antigen complex inhibits TATA binding protein-TATA interaction. *Molecular and Cellular Biology* **21**: 3652-3661

- Yamamoto, Y. Y., Ichida, H., Abe, T., Suzuki, Y., Sugano, S. and Obokata, J.** 2007a. Differentiation of core promoter architecture between plants and mammals revealed by LDSS analysis. *Nucleic Acids Research* **35**: 6219-6226
- Yamamoto, Y. Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K. and Abe, T.** 2007b. Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* **8**:
- Yang, B., Zhu, W., Johnson, L. B. and White, F. F.** 2000. The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 9807-9812
- Yang, B. and White, F. F.** 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular Plant-Microbe Interactions* **17**: 1192-1200
- Yang, B., Sugio, A. and White, F. F.** 2006. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 10503-10508
- Yang, J., et al.,.** 2014. Complete decoding of TAL effectors for DNA recognition. *Cell Research* **24**: 628-631
- Yang, Y., Yuan, Q. and Gabriel, D. W.** 1996. Watersoaking function(s) of XcmH1005 are redundantly encoded by members of the *Xanthomonas avr/pth* gene family. *Molecular Plant-Microbe Interactions* **9**: 105-113
- Yu, J., et al.,.** 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**: 79-92
- Yu, Y., Streubel, J., Balzergue, S., Champion, A., Boch, J., Koebnik, R., Feng, J., Verdier, V. and Szurek, B.** 2011. Colonization of rice leaf blades by an african strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice Nodulin-3 *Os11N3* Gene. *Molecular Plant Microbe Interactions* **24**: 1102-1113
- Yuan, M., Zhao, J., Huang, R., Li, X., Xiao, J. and Wang, S.** 2014. Rice *MtN3/saliva/SWEET* gene family: Evolution, expression profiling, and sugar transport. *Journal of Integrative Plant Biology* **56**: 559-570
- Yuan, T., Li, X., Xiao, J. and Wang, S.** 2011. Characterization of *Xanthomonas oryzae*-responsive *cis*-acting element in the promoter of rice race-specific susceptibility gene *Xa13*. *Molecular Plant* **4**: 300-309
- Zhang, H., et al.,.** 2014. Engineered TAL Effector modulators for the large-scale gain-of-function screening. *Nucleic Acids Research* **42**: e114

Zhao, X. and Li, L. 2013. Comparative analysis of microRNA promoters in Arabidopsis and rice. *Genomics Proteomics Bioinformatics* **11**: 56-60

Zhou, Y., Liu, L., Huang, W., Yuan, M., Zhou, F., Li, X. and Lin, Y. 2014. Overexpression of *OsSWEET5* in rice causes growth retardation and precocious senescence. *PLoS ONE* **9**: e94210

Zhu, W., Yang, B., Chittoor, J. M., Johnson, L. B. and White, F. F. 1998. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Molecular Plant-Microbe Interactions* **11**: 824-832

5. Anhang

5.1 Übersicht über die mittels PlantCARE identifizierten *cis*-regulatorischen DNA-Elemente im 1kb Fragment des *OsSWEET14* Promotors

```

+GTGTGTGCCACTCCAACTGGATAACTCACCATATAATAACTCTACAGTTATTGATAGATTATATACATT
-CACACACGGTGAGGTTGACTATTGAGTGGTATATTATTGAGATGTCAATAACTTATCTATATATGTAA

+ACGTCCCTTTTTTTAAATAATATCATCCAAATATTAGTTAAAAAGGGGGCGTAGTGTATATAGCCAA
-TGCAGGGAAAAAATTTATTATAGTAGGTTTATAATCAATTTTTCCCCCGCATCACATATATCGGTT

+CAGCTGTCAAAAAATTGAAAATTACAACGAATATACGTACAATTGCAAAGTGCAAATCACTCTTGTA
-GTCGACAGTTTTTTAACTTTTAAATGTTGCTTATATGCATGTTAACGTTTCACGTTTAGTGAGAACAT

+CTATCTTAAAAAGAAATGCAGAGTTTGATCGATTTTGTCTCTAATTTTAGGCAACTGGCATTGTGAT
-GATAGAATTTTTCTTTACGTCTCAAACTAGCTAAAACAGAGATTAAAATCCGTTGACCGTAAGACTA

+TTTTTACAAGTGGGTTAATAGAGCCAGCTGGTCTCAGATCGTCCGTTGGGGAATTAAAGGAGGTGT
-AAAAATGTTCACCCAATTAATCTCGGTCGACCAGAGTCTAGCAGGCAACCCCTTAATTTCTCCACA

+GCTCCTTTGCGGCTCATCAGTTTCTCTAAGCTCTCACCATTCATTCCACTATACAAGCCTAAGGCAG
-CGAGGAAACGCCGAGTAGTCAAAGAGATTCGAGAGTGGTAAGTAAGGTGATATGTTCGGATTCCGTC

+CTAGCTTAGTTAATTACCTAATAACTATAGCTTGCCCAACTCTAGATCCCTTAACTAGGACAACTTG
-GATCGAATCAATTAATGGATTATTGATATCGAACGGGTTGAGATCTAGGGAATTGATCCTGTTGAAC

+GAGTACACAACAATGTTACTAATCCCATGCATTGAGGACAGAGTTATGAAGGAAACAAAAAAGCT
-CTCATGTGTTGTTACAATGATTAGGGTACGTAACTCCTGTCTCAATACTTCCCTTTGTTTTTTTTCGA

+AGCAGATTGGCACTTCTGTGCATGCATGGGTGCTGATGATTATCTTGTATCTAATTTAATCAATCCC
-TCGTCTAACCGTGAAAGACAGTACGTACCCACGACTACTAATAGAACATAGATTAAAATTAGTTAGGG

+ATGGCTGTGATTGATCAGGAATAGTTTGTGTGTGCAGCTATATTGCCTATTGGTGTCCAGGGTCACA
-TACCGACACTAACTAGTCCTTATCAAACACACACGTCGATATAACGGATAACCACAGGTCCAGTGT

+CACCATAAGGGCATGCATGTCAGCAGCTGGTCATGTGTGCCTTTTCATTCCCTTCTTCCTTCCTAGC
-GTGGTATTTCCCGTlalCACGTACAGTCGTCGACCAGTACACACGGAAAAGTAAAGGGAAGAAGGAGGATCG

+ACTATATAAACCCCCTCCAACCAGGTGCTAAGCTCATCAAGCCTTCAAGCAAAGCAAACTCAAGTAG
PthXo3 _____ Tal5
AvrXa7 _____
-TGATATATTTGGGGAGGTTGGTCCACGATTCGAGTAGTTCGGAAGTTCGTTTCGTTTGAGTTCATC

+TAGCTGATTACCAGCTCTTCTCTTCTCATTGAGAAGAGGGAATTAAGTTTTGATCTCTGCTTTAT
-ATCGACTAATGGTCGAGAAGAGAGAAGATAACTCTTCTCCCTTAATTCAAAACTAGAGACGAAATA

+TGCCTGATCATCCTCTTGTTACTTGCAAGCAAGAACAGTAGTGTACTGTGCCTCATTGATCTCCTCC
-ACGGACTAGTAGGAGAACAATGAACGTTCGTTCTTGTCATCACATGACACGGAGTAACTAGAGGAGG

+CACCAAACTCTCTCTCTCTCTCATATTCCGAGCTAGCTAGTTAATCAAGATCTTGCTGCAATG
-GTGGTTTGAGAGAGAGAGAGAGATAAAGGCTCGATCGATCAATTAGTTCTAGAACGACGTTAC

```

Abbildung 5.1 Cis-regulatorische Motive im 1kb Fragment des *OsSWEET14* Promotors Die Analyse des *OsSWEET14* Promotors (1kb) wurde mittels der frei verfügbaren PlantCARE Datenbank durchgeführt. In der PlantCARE Datenbank sind *cis*-regulatorische DNA-Elemente hinterlegt, die aus verschiedenen Pflanzen bekannt sind (149 von Monokotyledonen, 281 von Dikotyledonen und 5 von anderen Pflanzen). Nachfolgend sind die mittels PlantCARE erhaltenen *cis*-regulatorischen Motife und deren Funktion, soweit auf PlantCARE verfügbar, sowie zusätzliche Markierungen aufgeführt.

PlantCARE-Motiv: Funktion:

5UTR Py-rich stretch	cis-acting element conferring high transcription levels
CAAT Box	common cis-acting element in promoter and enhancer regions
CAT-Box	cis-acting regulatory element related to meristem expression
CATT-Motif	part of a light responsive element
F-Box	unknown
GA Motif	part of a light responsive element
GAG Motif	part of a light responsive element
GT1 Motif	light responsive element
SPI	light responsive element
TCA Motif	cis-acting element involved in salicylic acid responsiveness
circadian	cis-acting element involved in salicylic acid responsiveness

Zusätzliche Markierungen im *OsSWEET14* Promotor:

TATA-Box

Fett markiert sind die Bindestellen der künstlichen TALEs,

Fett und *kursiv* markiert sind überlappende TALE-Bindestellen

Unterstrichen ist die TalC, PthXo3 (nach Antony 2010), AvrXa7 (nach Antony 2010) und Tal5 Bindestelle

Referenz:

Lescot, M., Déhais, P., Moreau, Y., De Moor, B., Rouzé, P., and Rombauts, S. 2002 PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research Database* issue 30:325-327.

5.3 Übersicht über alle in dieser Arbeit genannten oder verwendeten Bindestellen natürlicher bzw. künstlicher TALEs im 1kb Fragment des *OsSWEET14* Promotors

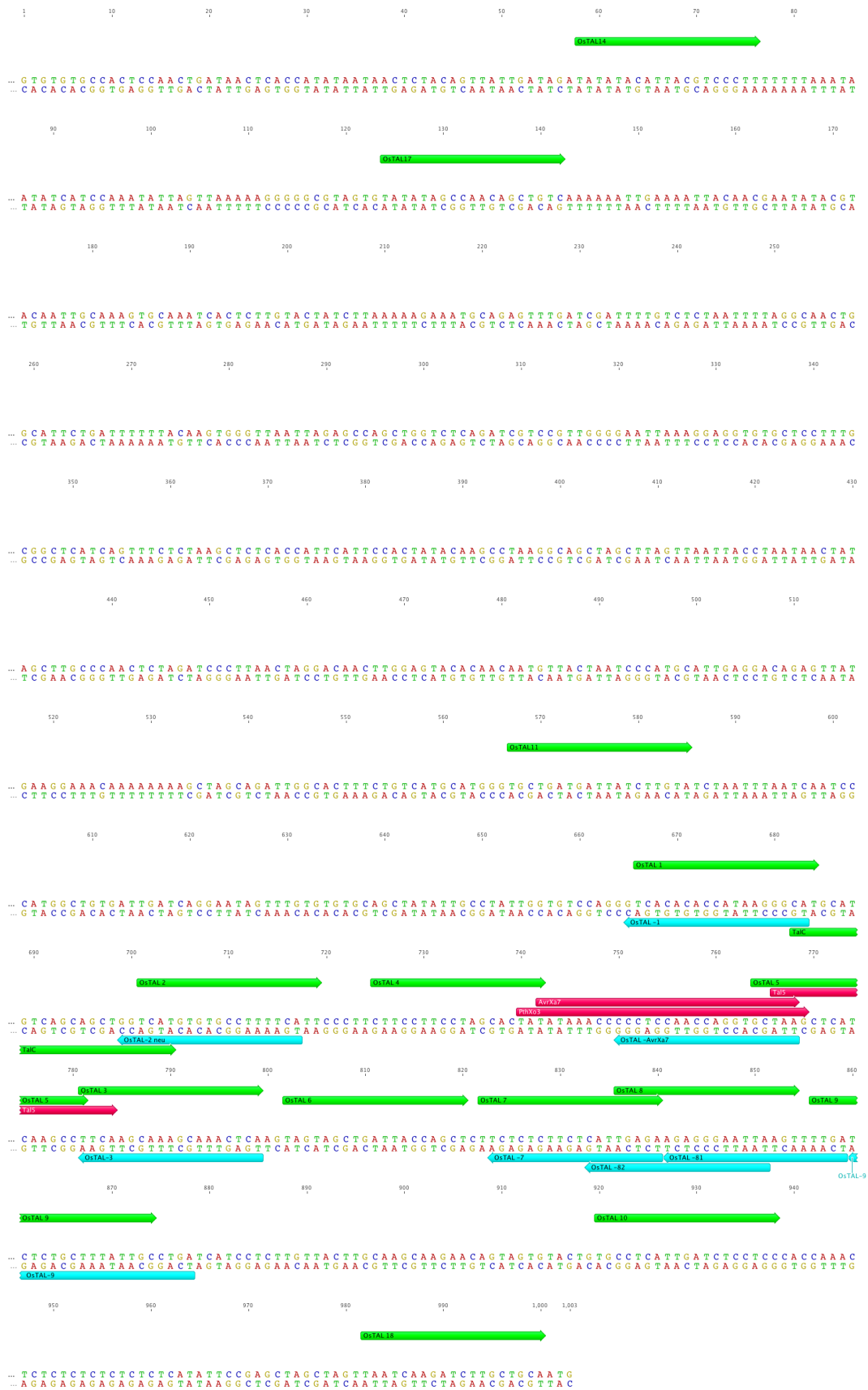


Abbildung 5.3 Übersicht über alle genannten oder verwendeten TALEs im 1kb-Fragment des *OsSWEET14*-Promotors Dargestellt ist das 1kb-Fragment des *OsSWEET14*-Promotors. Das *OsSWEET14* Startkodon ist gelb markiert. Die Bindestellen aller in dieser Arbeit verwendeten künstlichen *forward*-bindenden TALEs sind mit grünen Pfeilen am oberen Strang die *reverse*-bindenden TALEs mit blauen Pfeilen am unteren Strang markiert. Die natürlichen am *OsSWEET14*-Promotor bindenden TALEs TalC, PthXo3, AvrXa7 und Tal5 sind mit pinken Pfeilen markiert.

5.4 Übersicht über die RVD-Komposition und Ziel-DNA-Sequenz der in dieser Studie verwendeten künstlichen TALEs

OstAL 1

HD-NI-HD-NI-HD-NI-HD-HD-NI-NG-NI-NI-NN-NN-NN-HD-NI-NG
TC A C A C A C C A C A A G G G C A T

OstAL -1

NH-HD-HD-HD-NG-NG-NI-NG-NH-NH-NG-NH-NG-NH-NG-NH-NI-HD
TG C C C T T A T G G T G T G T G A C

OstAL 2

HD-NI-NG-NN-NG-NN-NG-NN-HD-HD-NG-NG-NG-NG-HD-NI-NG-NG
TC A T G T G T G C C T T T T C A T T

OstAL -2

NN-NI-NI-NI-NI-NN-NN-HD-NI-HD-NI-HD-NI-NG-NN-NI-HD-HD
TG A A A A G G C A C A C A T G A C C

OstAL 3

NG-HD-NI-NI-NN-HD-NI-NI-NI-NN-HD-NI-NI-NI-HD-NG-HD-NI
TT C A A G C A A A G C A A A C T C A

OstAL -3

NN-NI-NN-NG-NG-NG-NN-HD-NG-NG-NG-NN-HD-NG-NG-NN-NI-NI
TG A G T T T G C T T T G C T T G A A

OstAL 4

HD-NG-NG-HD-HD-NG-NG-HD-HD-NG-NI-NN-HD-NI-HD-NG-NI-NG
TC T T C C T T C C T A G C A C T A T

OstAL 5

NN-HD-NG-NI-NI-NN-HD-NG-HD-NI-NG-HD-NI-NI-NN-HD-HD-NG
TG C T A A G C T C A T C A A G C C T

OstAL 6

NI-NN-NG-NI-NN-HD-NG-NN-NI-NG-NG-NI-HD-HD-NI-NN-HD-NG
TA G T A G C T G A T T A C C A G C T

OstAL 7

NG-HD-NG-HD-NG-HD-NG-NG-HD-NG-HD-NI-NG-NG-NN-NI-NN-NI
TT C T C T C T T C T C A T T G A G A

OstAL -7

HD-NG-HD-NI-NI-NG-NN-NI-NN-NI-NI-NN-NI-NN-NI-NN-NI-NI
TC T C A A T G A G A A G A G A G A A

OstAL 8

NN-NI-NN-NI-NI-NN-NI-NN-NN-NN-NI-NI-NG-NG-NI-NI-NN-NG
TG A G A A G A G G G A A G G A A G T

OstAL -8 (EBE1)

HD-NI-NI-NI-NI-HD-NG-NG-NI-NI-NG-NG-HD-HD-HD-NG-HD-NG
TC A A A A C T T A A T T C C C T C T

OstAL -8 (EBE2)

NI-NI-NG-NG-HD-HD-HD-NG-HD-NG-NG-HD-NG-HD-NI-NI-NG-NH
 TA A T T C C C T C T T C T C A A T G

OstAL 9

NG-NN-NI-NG-HD-NG-HD-NG-NN-HD-NG-NG-NG-NI-NG-NG-NN-HD
 TT G A T C T C T G C T T T A T T G C

OstAL 10

NN-HD-HD-NG-HD-NI-NG-NG-NN-NI-NG-HD-NG-HD-HD-NG-HD-HD
 TG C C T C A T T G A T C T C C T C C

OstAL 11

NN-HD-NG-NN-NI-NG-NN-NI-NG-NG-NI-NG-HD-NG-NG-NN-NG-NI
 TG C T G A T G A T T A T C T T G T A

OstAL 14

NI-NG-NI-NG-NI-HD-NI-NG-NG-NI-HD-NN-NG-HD-HD-HD-NG-NG
 TA T A T A C A T T A C G T C C C T T

OstAL 17

NI-NG-NI-NG-NI-NN-HD-HD-NI-NI-HD-NI-NN-HD-NG-NN-NG-HD
 TA T A T A G C C A A C A G C T G T C

OstAL 18

NI-NI-NG-HD-NI-NI-NH-NI-NG-HD-NG-NG-NH-HD-NG-NH-HD-NI
 TA A T C A A G A T C T T G C T G C A

OstAL -AvrXa7

NG-NI-NN-HD-NI-HD-HD-NG-NN-NN-NG-NG-NN-NN-NI-NN-NN-NN
 TT A G C A C C T G G T T G G A G G G

OstAL GUS

NI-NN-NI-NI-NI-NI-HD-HD-HD-HD-NI-NI-HD-HD-HD-NN-NG-NN-NI
 TA G A A A C C C C A A C C C G T G A

5.5 Phänotypen der BAI3 Δ talC-Stämme mit oder ohne künstlichen TALE nach Entfärben

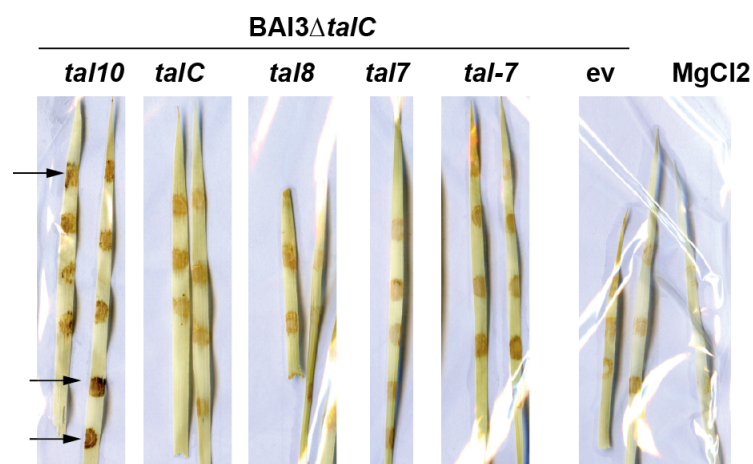


Abbildung 5.4 Phänotypen der BAI3 Δ talC-Stämme mit oder ohne künstlichen TALE nach Entfärben Der Stamm BAI3 Δ talC mit TAL10 konnte den Virulenzphänotyp in Reis 5dpi nicht wiederherstellen. Allerdings konnte 1dpi *OsSWEET14* Transkript gemessen werden (Abb2.5). Nach Entfärben der in Abb. 2.5 dargestellten Reisblätter mit 95% Ethanol und 5% Glycerin konnte für TAL10 dunkelbraune Läsionen (Pfeil) nachgewiesen werden. Im Vergleich dazu sind die hellbraunen Läsionen beispielhaft für die Stämme mit *talC*, *tal8*, *tal7*, *tal-7* sowie das *ev* Konstrukt und nach Inokulation mit MgCl₂ dargestellt. Die dunkelbraunen Läsionen lassen eine HR ausgelöst durch *tal10* vermuten.

5.6 Alignment der Hax3 repeats

Hax3 TALE Protein



Alignment Hax3 repeats

	1	RVD	34
rep1	LTP	EQVVAIASN I	GGKQALETVQRLLPVLCQAHG
rep2	LTP	QQVVAIAS HD	GGKQALETVQRLLPVLCQAHG
rep3	LTP	EQVVAIAS NI	GGKQALETVQ ALLPVLCQAHG
rep4	LTP	EQVVAIAS HD	GGKQALETVQRLLPVLCQAHG
rep5	LTP	QQVVAIA SHD	GGKQALETVQRLLPVLCQAHG
rep6	LTP	QQVVAIAS HD	GGKQALETVQRLLPVLCQAHG
rep7	LTP	QQVVAIASN S	GGKQALETVQRLLPVLCQAHG
rep8	LTP	QQVVAIASN S	GGKQALETVQRLLPVLCQAHG
rep9	LTP	QQVVAIASN S	GGKQALETVQRLLPVLCQAHG
rep10	LTP	EQVVAIAS HD	GGKQALETVQRLLPVLCQAHG
rep11	LTP	EQVVAIASN I	GGKQALETVQRLLPVLCQAHG
rep11,5	LTP	QQVVAIASN	GGRP ALE

Abbildung 5.5 Alignment der Hax3 repeats

Dargestellt ist das Hax3 Protein mit N- und C-terminaler Region sowie der zentralen *repeat*-Domäne. Das Alignment der Aminosäuresequenz (mittels Boxshade) der 11,5 repeats von Hax3 zeigt eine hohe Konservierung (schwarz). Die größte Variabilität ist an Position 12 und 13 dem RVD zu finden (grün umrahmt)

Erklärung

Hiermit erkläre ich, dass ich die vorliegende wissenschaftliche Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Ich erkläre weiterhin, dass keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und diesen Werken wörtlich oder inhaltlich entnommene Stellen als solche kenntlich gemacht wurden.

Mit dieser Arbeit bewerbe ich mich erstmals um die Erlangung des Doktorgrades.

Halle (Saale), 26. Februar 2015

Jana Streubel

Lebenslauf

Persönliche Daten

Name: Jana Streubel
Anschrift: Schieloer Str. 9
06124 Halle
Geburtsdatum: 29.01.1987
Geburtsort: Oschatz (Sachsen)
Familienstand: ledig

Ausbildung

1997-2005 Thomas - Mann - Gymnasium Oschatz
Abschluss: Allgemeine Hochschulreife (Abitur)
2005-2011 Studium der Biologie an der Martin-Luther-Universität Halle-
Wittenberg (MLU)
2010-2011 Diplomarbeit am Institut für Genetik, Abteilung Pflanzengenetik
der MLU
Thema: „DNA Spezifität von TAL- Effektoren“
Seit August 2011 Promotion an der MLU, Abteilung Pflanzengenetik, Betreuung
durch Dr. Jens Boch

Teilnahme an wissenschaftlichen Seminaren

Vorträge:

Effectome Meeting vom 28.-30. September 2011 in Lauret (Frankreich)
Vortrag: "Golden-TAL-Technology"

Concluding Colloquium SPP1212 (Plant-Micro) vom 08.-10. Juni 2013 in Freising

Vortrag: "Molecular Manipulation of Arabidopsis by type-III-effectors from bacterial pathogens *Pseudomonas syringae* and *Xanthomonas*"

Effectome Meeting vom 25.-27. September 2013 in Lauret (Frankreich)

Vortrag: "TAL effector repeats and their specificities"

COST FA1208 Meeting vom 08.-10. Oktober 2013 in Birnam (Schottland)

Vortrag: "Rice SWEET genes as virulence targets for *Xanthomonas* TALEs"

Poster:

XGC Congress vom 09.-12. Juli 2012 in Anger (Frankreich)

Titel: TAL-effector RVD specificities and efficiencies

XVI IS-MPMI Meeting vom 06.-10. Juli 2014 in Rhodos (Griechenland)

Titel: A flexible TALE-DNA-binding domain

Förderung

EMBO short-term fellowship:

Förderung des Aufenthalts am Institute de Recherche pour le developpement (IRD) in Montpellier (Frankreich)

ASTF Nummer: 355-2011

Beginn: 29.08.2011

Dauer: 38 Tage

Titel des Antrags: Rice MtN3 genes as virulence targets of TAL effectors

Halle (Saale), 04.02.2015