Characterization of the Transcription Activator-Like Effectors of *Xanthomonas axonopodis* pv. *manihotis* and identification of susceptibility targets in the host cassava

by

Megan Michelle Cohn

The dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Brian J. Staskawicz, Chair
Professor Shauna C. Somerville
Professor Russell E. Vance

Summer 2015

Abstract

Characterization of the Transcription Activator-Like Effectors of *Xanthomonas axonopodis* pv. *manihotis* and identification of susceptibility targets in the host cassava

by

Megan Michelle Cohn

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Brian Staskawicz, Chair

This work provides a comprehensive study of the molecular events mediated by the transcription activator-like (TAL) effectors of *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) during infection of its host plant cassava (*Manihot esculenta*). TAL effectors are secreted via the bacterial type III secretion system into plant cells where they localize to the nucleus, bind specific sequences of nucleotides in plant promoters, and activate the expression of downstream genes. The DNA binding activity of *Xam* TAL effectors is carried out by a central domain containing a variable number of 34 amino acid repeats, where the 12th and 13th amino acid residues, the repeat variable diresidues (RVDs), of each repeat dictate which nucleotide will be bound at that site. Genetic analysis of the TAL effectors of highly virulent *Xam* strain Xam668 showed a virulence role for TAL20$_{Xam668}$, which contributes to *in planta* bacterial growth and watersoaking symptom development, and TAL14$_{Xam668}$, which contributes to *in planta* bacterial growth. RNA-Sequencing (RNA-Seq) revealed a single host gene target of TAL20$_{Xam668}$, *MeSWEET10a*, which is a member of the SWEET family of sugar transporters. Designer TAL effectors that activate *MeSWEET10a* complemented the Xam668ΔTAL20 mutant phenotype, revealing the role of *MeSWEET10a* in cassava's susceptibility to *Xam*. SWEETs are known *X. oryzae* pv. *oryzae* TAL effector-targeted susceptibility genes in rice and thus this result establishes the activation of SWEET sugar transporters as a common mechanism of disease promotion for diverse species of *Xanthomonas*. In contrast to TAL20$_{Xam668}$ which has a single target gene, RNA-Seq showed that TAL14$_{Xam668}$ activates over 50 genes during the infection process. A subset of the TAL14$_{Xam668}$-targeted genes were tested for activation by TAL14$_{CIO151}$ from *Xam* strain CIO151 and, although TAL14$_{CIO151}$ and TAL14$_{Xam668}$ differ by only a single RVD, they display differential activation of host gene targets. TAL14$_{CIO151}$ complements the TAL14$_{Xam668}$ mutant defect, implying that shared target genes are important for TAL14$_{Xam668}$-mediated disease susceptibility. This type of complementation with closely related TAL effectors is a novel approach to narrowing down biologically relevant susceptibility genes of TAL effectors with multiple targets. The comparison of the activities of TAL14$_{Xam668}$ and TAL14$_{CIO151}$ provides an example of how TAL effector

target activation by two strains within a single species of *Xanthomonas* can be dramatically affected by a small change in RVD-nucleotide affinity at a single site and reflects the parameters of RVD-nucleotide interaction determined using artificial TAL effectors in transient systems. Finally, we present proof-of-concept experiments showing TAL effector-induced resistance to *Xam* in cassava.

Table of Contents

.

List of Figures

List of Tables

List of Abbreviations

668	Xam668
A	adenine
AAD	acidic activation domain
bp	base pairs
BSR	base-specifying residue
C	cytosine
CBB	cassava bacterial blight
CFU	colony forming unit
CHX	cycloheximide
DNA	deoxyribonucleic acid
dpi	days post inoculation
dTALE, dT	designer transcription activator-like effector
E gene	executor gene
EBE	effector binding element
EMSA	electromobility shift assay
ETI	effector triggered immunity
EV	empty vector
FEC	friable embryonic calli
FPKM	fragments per kilobase transcript per million mapped reads
FRET	Förster resonance energy transfer
G	guanine
HD	histidine-aspartic acid
hpi	hours post inoculation
HR	hypersensitive response
kb	kilobase pairs
kDa	kilodaltons
m	mock
NBS-LRR	nucleotide binding site leucine rich repeat
NG	asparagine-glycine
NI	asparagine-isoleucine
NLS	nuclear localization signal
NN	asparagine-asparagine
NS	asparagine-serine
NT	not tested
OD	optical density
PCR	polymerase chain reaction
ps	plasmid segment
pv	pathovar
R gene	resistance gene
RLU	relative light units
RNA	ribonucleic acid
RNA-Seq	RNA sequencing

RT-PCR	reverse transcriptase polymerase chain reaction
RVD	repeat variable diresidue
s	seconds
S gene	susceptibility gene
sbsp	subspecies
SD	standard deviation
SE	standard error
SWEET	sugars will eventually be exported transporters
T	thymine
T13	TAL13
T14	TAL14
T15	TAL15
T20	TAL20
T22	TAL22
TAL	transcription activator-like
TIR	toll-interleukin 1 receptor
Xam	*Xanthomonas axonopodis* pv. *manihotis*
Xcc	*Xanthomonas citri sbsp. citri, X. campestris pv. campestris*
Xe	*Xanthomonas euvesicatoria*
Xoc	*Xanthomonas oryzae pv. oryzicola*
Xoo	*Xanthomonas oryzae pv. oryzae*

Acknowledgments

I would like to acknowledge all of the people who influenced, guided, and contributed to this body of work. First, I'd like to thank Brian Staskawicz for being an optimistic and enthusiastic mentor. He has been supportive of me both professionally and personally, and I am grateful for all of his guidance. I am deeply grateful to Doug Dahlbeck for his vast knowledge of molecular biology techniques and the patience he showed when guiding me through new experiences in the lab. His contributions to this work and to my practical skills are immeasurable. I would like to thank my committee members Shauna Somerville and Russell Vance for their guidance and encouragement over the years. I'd like to thank my parents, John and Jody, and my sister Alyssa for their unending support and love. Finally, I'd like to thank my husband David for his unwavering optimism and encouragement. He has been my equal partner in life throughout my time in graduate school and I could not have finished my PhD if it were not for his love and support. Most importantly, I'd like to thank my son Roland who is 22 months old at the time of this writing. He has brought more joy to my life than I can express. This dissertation is dedicated to him. I hope that by having worked hard to accomplish my goal of attaining my PhD, he is inspired to work hard toward something great in his own life.

I'd like to acknowledge all the people who carried out experiments that are included in this dissertation. Rebecca Bart did the colony lift hybridization, Southern blot (Figure 2-1), and the TAL20$_{Xam668}$ RNA-Seq analysis (Figure 3-3A, 3-10). Mikel Shybut generated the TAL effector knock-out strains with suicide vector pLVC18 and the western blots in figures 2-5 and 2-8A. Doug Dahlbeck did the growth assays in figure 2-6 and the EMSA (Figure 3-4B). Michael Gomez generated the Xam668ΔsuxC growth curve (Figure 3-7A). Bi-Huei Hou and Wolf Frommer determined the sugar transport capabilities of MeSWEET10a (Figure 3-8). Davide Sosso named *MeSWEET10a* based on its phylogenetic relationship to other SWEETs. Thomas Lahaye and Robert Morbitzer made the designer TAL effectors featured in figures 3-5 and 4-10. My graduate work was funded by an NSF Graduate Research Fellowship and an NIH Genetics Training Grant.

1. Introduction

Xanthomonas axonopodis pv. *manihotis* (*Xam*), the causal agent of bacterial blight of cassava (CBB), delivers type III effector proteins into the plant cell to suppress or modulate host innate immunity and promote pathogenesis. Most *Xam* strains studied to date contain between one and five transcription activator-like (TAL) effectors (Bart *et al.*, 2012; Castiblanco *et al.*, 2012), which in other xanthomonads have been shown to elicit disease symptoms by inducing expression of host genes that allow the pathogen to achieve full virulence (Yang *et al.*, 2006; Antony *et al.*, 2010; Verdier *et al.*, 2012; Cernadas *et al.*, 2014; Hu *et al.*, 2014; Li *et al.*, 2014). Following secretion, TAL effectors are translocated to the plant nucleus where they physically interact with DNA in a sequence-specific manner to activate host gene expression (Kay *et al.*, 2007; Römer *et al.*, 2007; Boch & Bonas, 2010; Bogdanove *et al.*, 2010; Mak *et al.*, 2012; Deng *et al.*, 2012). Consistent with their function, the modular structure of TAL effectors consists of an N-terminal type III secretion signal, C-terminal nuclear-localization sequences, a eukaryotic acidic activation domain (AAD), and a central repeat domain that is responsible for sequence-specific protein-DNA interactions (Figure 1-1). This central repeat domain consists of a variable number of nearly identical 33-35 amino acid repeats. Each repeat mediates binding to a single consecutive nucleotide and the sequence of nucleotides bound by a TAL effector is termed the effector binding element (EBE). The majority of polymorphism between repeats within a TAL effector DNA binding domain is at the 12^{th} and 13^{th} amino acids, termed the repeat variable diresidues (RVDs). By observing the correspondence of RVDs with specific nucleotides in known TAL effector/EBE pairs, the TAL effector-DNA binding "code" was elucidated (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). Based on these association frequencies, the RVDs NG (Asn-Gly), NI (Asn-Ile), and HD (His-Asp) were shown to be highly specific for thymine (T), adenine (A), and cytosine (C), respectively, while NS (Asn-Ser) and NN (Asn-Asn) were less specific in their nucleotide preference (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). Crystal structures of TAL effectors bound to their DNA targets confirmed the one-to-one nature of RVD-nucleotide interaction and showed that the 13^{th} amino acid of the repeat directly interacts with the DNA base, and is now referred to as the base-specifying residue (BSR) (Mak *et al.*, 2012; Deng *et al.*, 2012; de Lange *et al.*, 2014).The majority of natural TAL effector EBEs are directly preceded by a thymine (T_0) which is required for efficient TAL effector binding and gene activation (Römer *et al.*, 2009b; 2010; Doyle *et al.*, 2013a).

The knowledge of the TAL effector-DNA binding code made it possible to predict potential EBEs in a genome, albeit with a large number of false positive predictions, and enabled the development of TAL effector-based biotechnological applications such as TAL effector nucleases (Bogdanove & Voytas, 2011; Joung & Sander, 2012; Sun & Zhao, 2013; Carroll, 2014). Driven by the need for more specific and efficient DNA targeting by TAL effector-based biotechnologies, much has been clarified in recent years regarding parameters that influence TAL effector-DNA binding such as RVD efficiency, RVD-nucleotide affinity, polarity effects, and the vicinity of EBEs to core

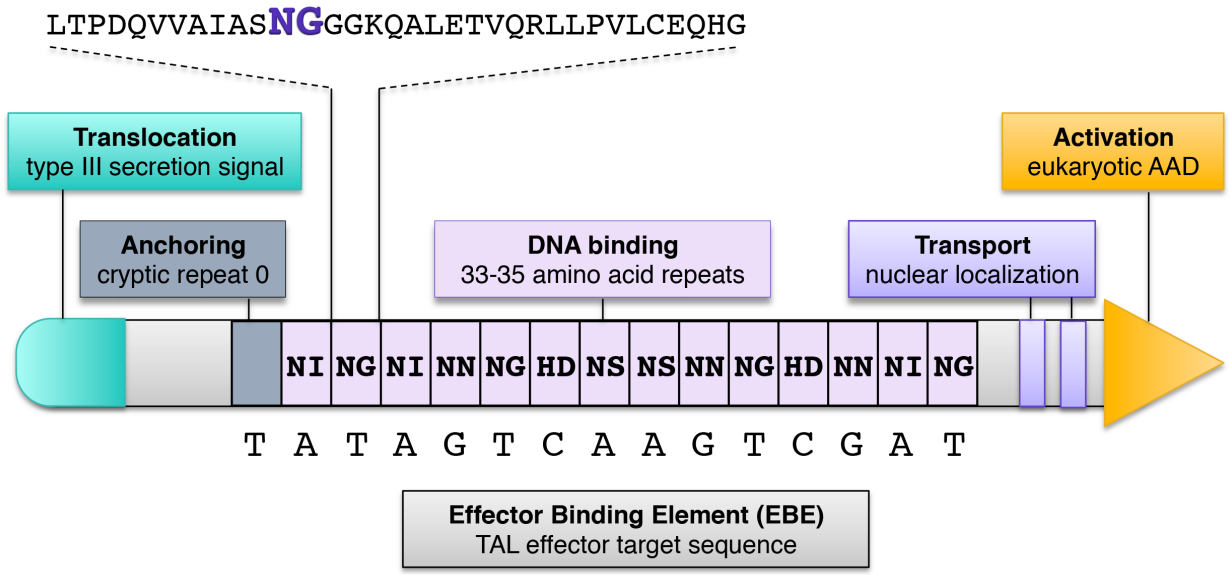


Figure 1-1. Transcription activator-like (TAL) effector structure. TAL effectors are modular proteins with conserved structural domains. The N-terminal region contains a type III secretion signal, and the C-terminal region contains nuclear localization signals and a eukaryotic acidic activation domain (AAD). The central DNA binding domain consists of a variable number of nearly perfect 33-35 amino acid repeats termed the repeat variable diresidues (RVDs). The sequence of a single repeat is shown with the RVD highlighted in purple. The RVDs dictate the specificity of DNA binding and the target nucleotide sequence is called the effector binding element (EBE).

promoter elements (Streubel *et al.*, 2012; Cong *et al.*, 2012; Grau *et al.*, 2013; Meckler *et al.*, 2013; Cernadas *et al.*, 2014; Moore *et al.*, 2014). These studies have primarily been done using artificial TAL effector-EBE pairs assayed transiently for activation strength. The way in which RVD substitutions affecting TAL effector binding efficiency may effect the function of natural TAL effectors in the context of their plant hosts is an open question.

TAL effectors can function in effector-triggered immunity (ETI) if recognized by resistance (*R*) genes or as virulence determinants activating plant genes whose increased or ectopic expression facilitates bacterial growth, disease symptom formation, or both [reviewed in (Schornack *et al.*, 2013)]. Several TAL effector targeted genes have been identified as susceptibility (*S*) genes, providing insight into strategies used by *Xanthomonas* during infection (Hutin *et al.*, 2015). *UPA20,* activated by the TAL effector AvrBs3 from *X. euvesicatoria* (*Xe*), is a transcription factor that regulates cell size in pepper (*Capsicum annuum*), leading to the cell hypertrophy that is a symptom of infection by *Xe* (Kay *et al.*, 2007). *CsLOB1*, a transcription factor in sweet orange (*Citrus sinensis*) associated with expression of genes involved in cell expansion, is activated by multiple TAL effectors of *X. citri* subspecies *citri* (*Xcc*). Expression of *CsLOB1* promotes

citrus bacterial canker pustule formation and *in planta* bacterial growth (Hu *et al.*, 2014; Li *et al.*, 2014). TAL2g of the rice (*Oryza sativa*) mesophyll colonizer *X. oryzae* pv. *oryzicola* (*Xoc*) activates the predicted sulfate transporter *OsSULTR3;6* to promote lesion expansion and bacterial exudation from the leaf surface (Cernadas *et al.*, 2014). *OsSWEET11, OsSWEET14,* and *OsSWEET13* from rice*,* each activated by at least one *X. oryzae* pv. *oryzae* (*Xoo*) TAL effector, are members of the SWEET family of sugar transporters, and were shown to transport both glucose and sucrose (Yang *et al.*, 2006; Chen *et al.*, 2010; Antony *et al.*, 2010; Chen *et al.*, 2012; Zhou *et al.*, 2015). Expression of host SWEETs at the site of bacterial infection is thought to result in an excess of sugars in the apoplasm, providing bacteria with a source of carbon (Chen, 2014).

The recessive rice resistance gene *xa13* was found to be an allele of *OsSWEET11* that contains a modified promoter sequence that can neither be bound nor activated by the TAL effector PthXo1, rendering strains that utilize PthXo1 as their primary source of virulence unable to cause disease (Chu, 2006; Yang *et al.*, 2006) Similarly, the recessive resistance gene *xa25* encodes *OsSWEET13* and avoids activation by PthXo2 through promoter polymorphism (Liu *et al.*, 2011; Yuan & Wang, 2013; Zhou *et al.*, 2015). Inspired by naturally occurring plant alleles, researchers have shown that plant disease resistance can be engineered by employing genome editing technologies to manipulate EBEs in the promoters of *S* genes such that they can no longer be activated in a TAL effector-dependent manner (Li *et al.*, 2012). Engineering recessive resistance by modifying host *S* gene promoters is a strategy that could be combined (stacked) with dominant *R* gene-mediated strategies employed by breeders and researchers (Schornack *et al.*, 2013). However, in order to be a widely applicable strategy, susceptibility targets must first be identified in the pathosystem of interest.

CBB is the one of the most serious and widespread diseases of cassava (*Manihot esculenta*). CBB is found in cassava growing regions in Central and South America, Africa, and Asia (CABI, current year. Crop Protection Compendium. Wallingford , UK: CAB International. www.cabi.org/cpc), and outbreaks can result in extensive crop damage [reviewed in (Lozano, 1986; López & Bernal, 2012)]. *Xam* colonizes both mesophyll and vascular tissues leading to a variety of symptoms that include watersoaked angular lesions on leaves, shoot dieback and necrosis (Figure 1-2). CBB is spread primarily by rain splash, propagation of infected cuttings, and through use of infected cutting tools by farmers (Lozano & Sequeira, 1973; Boher & Verdier, 1994). Cassava is the third most important source of calories for people living in the tropics behind rice and corn and its production is increasing in part because of the low input costs and drought tolerance of this staple food source. However, cassava remains a relatively understudied crop (FAO, 2008; Howeler *et al.*, 2013). The development of resistant cassava varieties could protect farmers from the crop losses associated with CBB (Wydra *et al.*, 2004). A robust cassava transformation protocol has been developed which will allow genetic engineering approaches to address this problem (Bull *et al.*, 2009). However, few reports of resistance to CBB exist and the genes underlying disease tolerance and resistance have not been extensively studied resulting

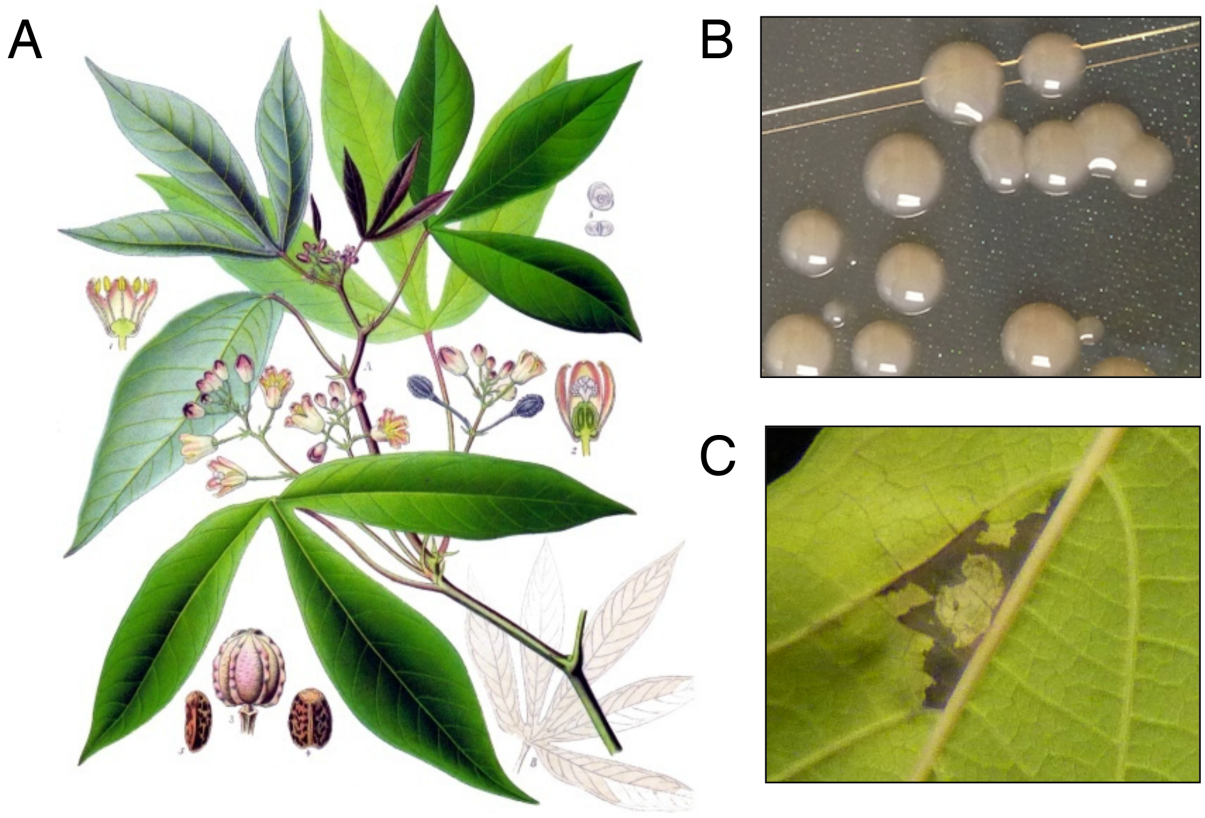


Figure 1-2. *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) is the causal agent of cassava bacterial blight (CBB). A. Anatomy of cassava (*Manihot esculenta*). Illustration by Franz Eugen Köhler, Köhlers Medizinal-Pflanzen (List of Koehler Images) [Public domain], via Wikimedia Commons. B. Picture of *Xam* mucoid white colonies. C. Characteristic watersoaking symptoms of CBB.

in a lack of resources for genetic engineering of resistant cultivars [reviewed in (López & Bernal, 2012)].

All sequenced *Xam* strains contain TAL effector-like sequences, indicating that TAL effectors are a conserved component of *Xam*'s virulence arsenal (Bart *et al.*, 2012; Arrieta-Ortiz *et al.*, 2013). In addition, a virulence role for TALE1$_{Xam}$ from *Xam* strain CFBP1851 has been demonstrated (Castiblanco *et al.*, 2012). Therefore, the development of resistance based on modification of TAL effector susceptibility targets in the plant is a promising approach to developing cassava that is resistant to CBB. However, little is known about the TAL effector-mediated molecular interactions that occur between *Xam* and cassava during disease progression.

The objective of this work is to shed light on the molecular mechanisms of *Xam* infection of cassava by elucidating the virulence roles and host *S* gene targets of *Xam* TAL effectors. The findings of this study are summarized on the following page.

- Highly virulent *Xam* strain Xam668 has five TAL effectors: $TAL13_{Xam668}$, $TAL14_{Xam668}$, $TAL15_{Xam668}$, $TAL20_{Xam668}$, and $TAL22_{Xam668}$, which have 13, 14, 15, 20, and 22 RVDs, respectively.
- $TAL20_{Xam668}$ contributes to *in planta* bacterial growth and watersoaking symptom development.
- $TAL14_{Xam668}$ contributes to *in planta* bacterial growth.
- RNA-Sequencing (RNA-Seq) revealed that $TAL20_{Xam668}$ has a single cassava gene target, *MeSWEET10a*, a member of the SWEET family of sugar transporters.
- The *MeSWEET10a* promoter is directly bound and activated by $TAL20_{Xam668}$.
- MeSWEET10a can transport both sucrose and glucose.
- Designer TAL effectors (dTALEs) showed that *MeSWEET10a* activation is an important factor in the susceptibility of cassava to CBB.
- RNA-Seq revealed that $TAL14_{Xam668}$ has 52 cassava gene targets, including pectate lyases and EamA-like/MtN21 transporters.
- $TAL14_{ClO151}$ of *Xam* strain ClO151 activates a subset of the $TAL14_{Xam668}$ targets despite differing from $TAL14_{Xam668}$ by only a single repeat.
- $TAL14_{ClO151}$ has an RVD at position 5 with a stronger preference for its corresponding nucleotide than does $TAL14_{Xam668}$, making it less likely to bind mismatched nucleotides and dramatically affecting its target gene repertoire.
- $TAL14_{ClO151}$ complements a Xam668 $TAL14_{Xam668}$ mutant strain, narrowing down candidate *S* genes to targets activated by both TAL14 proteins.
- Proof-of-concept experiments show that a TAL effector-activated resistance-triggering construct can inhibit *Xam* growth *in planta.*

2. Characterization of the transcription activator-like (TAL) effectors of virulent *Xanthomonas axonopodis* pv. *manihotis* strain Xam668

Background

The causal agent of cassava bacterial blight (CBB), *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) uses TAL effectors as part of its arsenal of type III effector proteins, but little is known about the specific roles of *Xam* TAL effectors in virulence (Bart *et al.*, 2012; Castiblanco *et al.*, 2012). TAL effectors are a unique class of type III effector in that they are defined as transcriptional activators, but their specific function in the host plant is dependent on their repeat variable diresidue (RVD) sequence, which dictates the nucleotide sequences that are bound within the host (Boch & Bonas, 2010). Therefore, the first step toward understanding the molecular events mediated by the TAL effectors of *Xam* is the elucidation of their DNA sequences. From this we can know the RVD sequence and predict host targets based on the TAL effector-DNA binding code, and also generate mutant strains to test the roles of the individual TAL effectors in *Xam* virulence (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). For this body of work we chose to study the TAL effectors of *Xam* strain Xam668 because of its high level of virulence (Bart *et al.*, 2012). In this chapter we present RVD sequences, host target predictions, and virulence roles of the TAL effectors of Xam668.

Results

Sequences and features of the TAL effectors of Xam668

Southern and western blot analyses indicate that Xam668 has five differently sized TAL effectors (Figure 2-1, 2-5) ranging from 105 to 137 kilodaltons (kDa). A Sanger sequencing approach was taken to elucidate the Xam668 TAL effector sequences. A Xam668 genomic DNA cosmid library was screened for TAL effector-containing clones by colony lift hybridization and twenty-one clones were selected for further analysis. Cosmids were isolated, digested with BamHI which cuts conserved sites just 3' of the TAL effector ATG and 150 bp 5' of the stop codon, and probed by Southern blot analysis. Five sizes of TAL effector were identified and sequenced. Nucleotide and amino acid sequences are provided (Supplement A). $TAL13_{Xam668}$, $TAL14_{Xam668}$, $TAL15_{Xam668}$, $TAL20_{Xam668}$, and $TAL22_{Xam668}$, have 13, 14, 15, 20, and 22 RVDs in their DNA binding domains, respectively (Table 1). Amino acid alignments of the five TAL effector N- and C-terminal regions reveals a high level of sequence conservation, with amino acid polymorphism at 3% (17/570) of residues. The program cNLS Mapper identified two conserved nuclear localization signals (NLS) in the TAL effector C-terminal region, and these NLS overlap with those previously identified and

Table 1. Xam668 transcription activator-like (TAL) effector repeat variable diresidues (RVDs)

TALE	RVD Sequence
TAL13$_{Xam668}$	NI NS NN HD NG HD NI NG HD NN NI NI NG
TAL14$_{Xam668}$	NI NG NI NN NG HD NS NS NN NG HD NN NI NG
TAL15$_{Xam668}$	NI NG NI NN HD HD NS NS NS HD HD NS HD NG NG
TAL20$_{Xam668}$	NI NG NI NN NI HD NS NS NN NG HD NS HD NN HD HD HD NI NG NG
TAL22$_{Xam668}$	NI NG HD NG NG NG HD HD NG NG HD NG HD HD NG NG HD NG NG HD NG NG

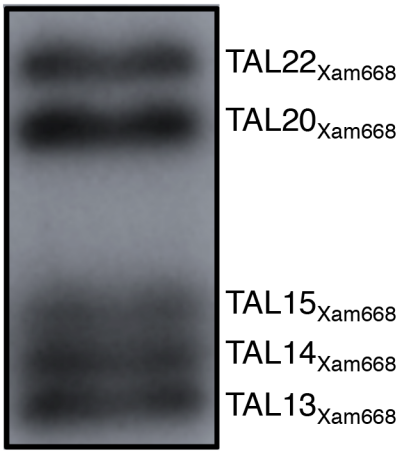


Figure 2-1. The Xam668 genome encodes 5 differently sized transcription activator-like (TAL) effectors. Southern blot of BamHI-digested Xam668 genomic DNA probed for TAL effector sequences.

validated in the TAL effectors AvrBs3 and TALE1$_{Xam}$ (Szurek *et al.*, 2001; Kosugi *et al.*, 2009a,b; Castiblanco *et al.*, 2012). Amino acid polymorphisms, NLS, and acidic activation domains [consensus T–V–M–x–E–Q–D–[EA]–[DA]–P–F–A–G–A–A–D–D– F–P–A–F–N–E(3)] of the Xam668 TAL effectors are shown in figure 2-2.

Xam668 TAL effector target prediction

The TAL effector-DNA binding code allows for the prediction of TAL effector binding elements (EBEs) in a promoterome based on observed RVD-nucleotide specificity (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). The RVD-nucleotide association frequencies elucidated by Moscou and Bogdanove (2009) were used to visualize possible EBEs for the five Xam668 TAL effectors (Figure 2-3). The 5' nucleotides of the predicted EBEs together with the observed preference for a thymine directly upstream of the EBE (T_0) reveal a likelihood of TATA-box or TATA-box variant

A

```
TAL13(Xam668)      MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPP
TAL14(Xam668)      MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPP
TAL20(Xam668)      MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPP
TAL15(Xam668)      MDPIRPRTSSPADELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPP
TAL22(Xam668)      MDPIRPRTSSPADELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPP
                   ******** ***.********************************************

TAL13(Xam668)      APLPAFSAGSFSDLLRQFDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQ
TAL14(Xam668)      APLPAFSAGSFSDLLRQFDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQ
TAL20(Xam668)      APLPAFSAGSFSDLLRQFDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQ
TAL15(Xam668)      APLPAFSAGSFSDLLRQFDPSLLDTSLFDSMSAFGAPHTEAASGEGDEVQSGLRAADDPH
TAL22(Xam668)      APLPAFSAGSFSDLLRQFDPSLLDTSLFDSMSAFGAPHTEAASGEGDEVQSGLRAADDPH
                   ***************************:******************************:

TAL13(Xam668)      ATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQ
TAL14(Xam668)      ATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQ
TAL20(Xam668)      ATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQ
TAL15(Xam668)      PTVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQ
TAL22(Xam668)      PTVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQVDLCTLGYSQQQQEKIKLKARSTVAQ
                    ********************************************************

TAL13(Xam668)      HHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGGKQWSGARAL
TAL14(Xam668)      HHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGGKQWSGARAL
TAL20(Xam668)      HHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGGKQWSGARAL
TAL15(Xam668)      HHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGGKQWSGARAL
TAL22(Xam668)      HHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGVGKQWSGARAL
                   ********************************************** **********

TAL13(Xam668)      EALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
TAL14(Xam668)      EALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
TAL20(Xam668)      EALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
TAL15(Xam668)      EALLTVSGELRGPPLQLDPGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
TAL22(Xam668)      EALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
                   ****************** *****************************
```

B

```
TAL13(Xam668)      STFAQLSRPDQALAALTNDHLVALACLGGRPALEAVRKGLPHAPTLIKRTNRRLPERTSH
TAL15(Xam668)      SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRRLPERTSH
TAL20(Xam668)      SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRRLPERTSH
TAL14(Xam668)      STFAQLSRPDQALAALTNDHLVALACLGGRPALEAVRKGLPHAPTLIKRTNRRLPERTSH
TAL22(Xam668)      SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRHLPERTSH
                   * ********************************:***************:*******

TAL13(Xam668)      RVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEAHSGTLPP
TAL15(Xam668)      RVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEAHSGTLPP
TAL20(Xam668)      RVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEAHSGTLPP
TAL14(Xam668)      RVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTGLEARSGTLPP
TAL22(Xam668)      RVADHAQVARVLGFFQCHSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEARSGTLPP
                   ******************************************.*** ***:*******

TAL13(Xam668)      ASQRWHRILQASGMKRAEPSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSS
TAL15(Xam668)      ASQRWHRILQASGMKRAEPSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSS
TAL20(Xam668)      ASQRWHRILQASGMKRAEPSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSS
TAL14(Xam668)      APQRWHRILQASGMKRAEPSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSS
TAL22(Xam668)      APQRWHRILQASGMKRAEPSGASAQTPDQASLHAFADALERELDAPSPIDRAGQALASSS
                   * ********************************************************

TAL13(Xam668)      RKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLPDPGTPMDAD
TAL15(Xam668)      RKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLPDPGTPMDAD
TAL20(Xam668)      RKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLPDPGTPMDAD
TAL14(Xam668)      RKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLPDPGTPMDAD
TAL22(Xam668)      RKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLPDPGTPMDAD
                   ******************************: ***************************

TAL13(Xam668)      LAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ
TAL15(Xam668)      LAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ
TAL20(Xam668)      LAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ
TAL14(Xam668)      LAASSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ
TAL22(Xam668)      LAASSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ
                   ** ***********************************
```

Figure 2-2 (previous page). Amino acid alignments of the Xam668 transcription activator-like (TAL) effector N- and C-terminal domains. A. The N terminal domains (start codon to beginning of DNA binding domain) of the Xam668 TAL effectors are highly similar with polymorphisms at 7 sites (highlighted in turquoise). B. The C terminal domains (end of DNA binding domain to stop codon) are highly similar with polymorphisms at 10 sites (highlighted in turquoise). Predicted nuclear localization signals (NLS) are underlined and predicted acidic activation domains (AAD) are highlighted in gray. Sequence alignments were generated with Clustal Omega 1.2.1 (Goujon *et al.*, 2010; Sievers *et al.*, 2011).

Figure 2-3. Visualization of predicted Xam668 transcription activator-like (TAL) effector target effector binding elements (EBEs). Repeat variable diresidue (RVD)-nucleotide association frequencies established in Moscou (2009) were used to display the probability of nucleotide binding at each RVD in the Xam668 TAL effector DNA binding domains.

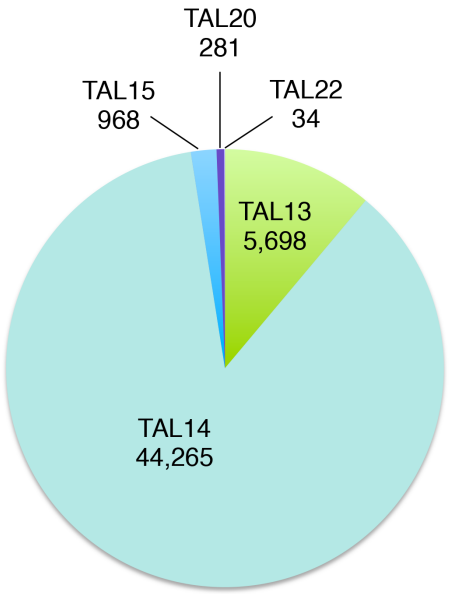


Figure 2-4. Xam668 transcription activator-like (TAL) effector target prediction. A pie chart shows the number of computationally predicted targets for each Xam668 TAL effector within the 1 kb cassava promoterome (promoters defined as 1 kb upstream of an annotated start codon). Predictions were made using TALE-NT 2.0 Target Finder scanning the promoterome of cassava AM560-2 (Doyle *et al.*, 2012).

binding for $TAL14_{Xam668}$, $TAL15_{Xam668}$, and $TAL20_{Xam668}$ (Römer *et al.*, 2009b; Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Bernard *et al.*, 2010; Römer *et al.*, 2010; Doyle *et al.*, 2013a). $TAL22_{Xam668}$ may bind a TATA-box variant or a TC-promoter motif (Bernard *et al.*, 2010).

TAL Effector-Nucleotide Targeter (TALE-NT) 2.0 Target Finder was used to predict potential Xam668 TAL effector EBEs in the cassava promoterome, defined as 1 kb upstream of annotated transcriptional start sites (*M. esculenta* AM560-2 genome version 4.1). Target Finder utilizes a scoring function based on RVD-nucleotide association frequencies found in a set of known TAL effector-DNA target pairs and assigns scores to predicted EBEs based on how likely it is that the TAL effector will bind to that nucleotide sequence (Moscou & Bogdanove, 2009; Cermak *et al.*, 2011; Doyle *et al.*, 2012). All five Xam668 TAL effectors are predicted to bind target sites in the cassava promoterome, with $TAL14_{Xam668}$ having the highest number of predicted EBEs (Figure 2-4). Since the Target Finder EBE prediction algorithm is solely based on RVD-nucleotide association frequencies and does not take into account positional effects of the RVD within the DNA binding region, positional effects of the EBE within the target promoter, or RVD binding efficiencies, the number of truly functional EBEs is much smaller than the number predicted, making transcriptomic analysis essential to identify true target genes.

The TAL effectors of Xam668 contribute differentially to virulence

To assess the role of TAL effectors in *Xam* virulence, we generated knockout strains by integration of a suicide vector into the TAL effector coding region. Thirty four independent insertions were screened by western blot analysis and in each case at least one TAL effector was disrupted. Single knockouts for each of the five Xam668 TAL effectors were selected and virulence levels were assessed through observation of watersoaking symptom development and measurement of bacterial growth *in planta*. Disruption of $TAL13_{Xam668}$, $TAL15_{Xam668}$ and $TAL22_{Xam668}$ did not cause a visible alteration in the development of watersoaked lesions or measurable differences in bacterial growth (Figure 2-5, 2-6A). Disruption of either $TAL14_{Xam668}$ or $TAL20_{Xam668}$ resulted in a reduction in bacterial growth in both the leaf midvein (Figure 2-6A), and the leaf apoplast (Figure 2-6B). An independent $TAL14_{Xam668}$ mutant strain missing the $TAL14_{Xam668}$ plasmid segment (Xam668ΔTAL14ps) also showed a growth defect that was fully complemented by wild type $TAL14_{Xam668}$ (Figure 2-7A, 4-9A).

In addition to its growth defect and in contrast to $TAL14_{Xam668}$, disruption of $TAL20_{Xam668}$ resulted in a reduction in watersoaking symptom development (Figure 2-5). A $TAL20_{Xam668}$ plasmid-cured strain (Xam668ΔTAL20p) mimicked the symptom defect observed for Xam668ΔTAL20 and reintroducing $TAL20_{Xam668}$ into the plasmid cured strain [Xam668ΔTAL20p($TAL20_{Xam668}$)] fully complemented the symptom defect phenotype (Figure 2-5, 2-7B). Because the $TAL20_{Xam668}$ plasmid cured strain was marker free, we were able to generate a Xam668ΔTAL14/ΔTAL20p double knockout by integration of a suicide vector into the $TAL14_{Xam668}$ coding region. The double mutant showed a reduction in watersoaking when compared to wild-type Xam668 and less

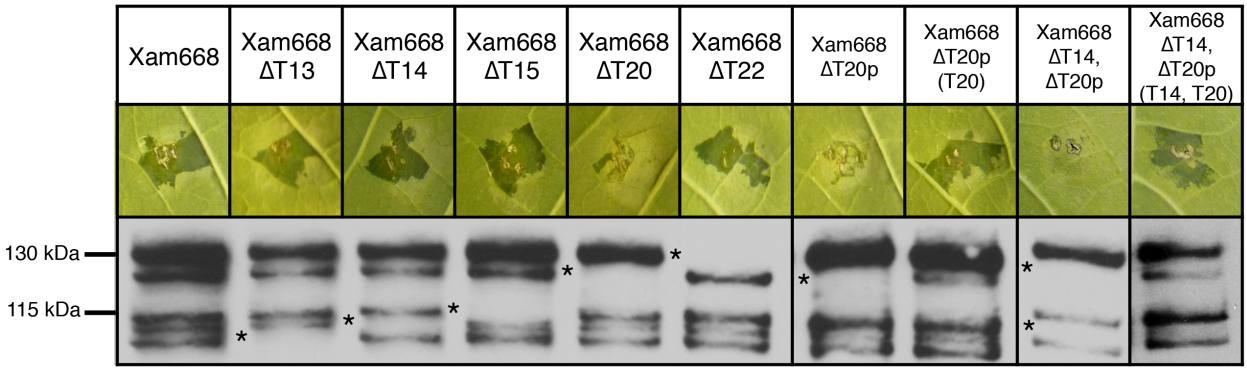


Figure 2-5. *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strain Xam668 contains five distinct transcription activator-like (TAL) effectors that contribute unequally to symptom development. Mutant strains were generated for each TAL effector and confirmed by western blot analysis. Missing TAL effectors are indicated with asterisks. TAL effector knockouts, $TAL20_{Xam668}$ plasmid-loss (ΔT20p), and double mutant lines were inoculated into cassava leaves (OD_{600} = 0.01) and symptom development was recorded after 5 days.

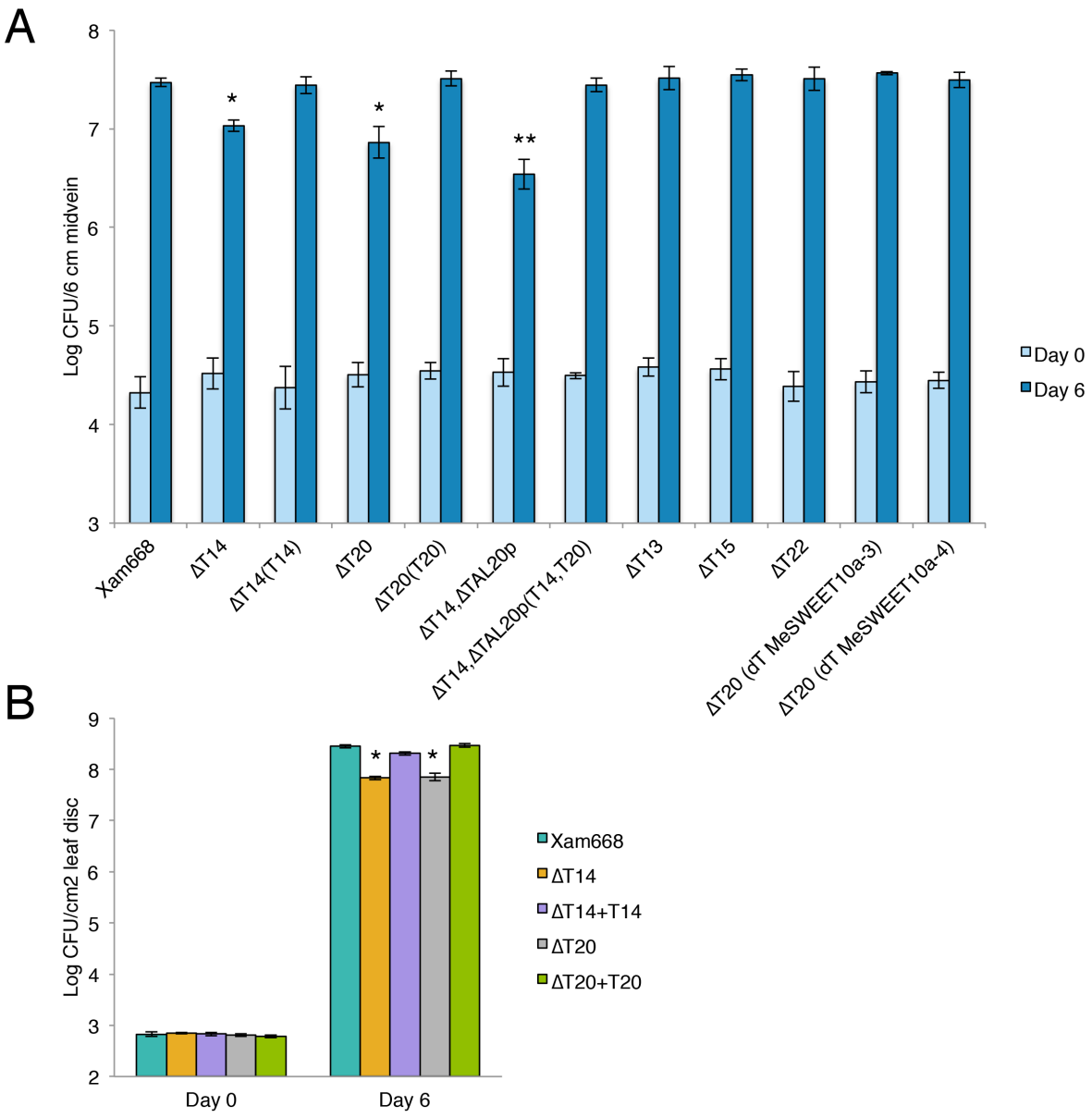


Figure 2-6. Xam668 transcription activator-like (TAL) effectors contribute differentially to *in planta* bacterial growth. A. Bacterial populations in leaf midveins were measured at days 0 and 6. Data are represented as mean colony forming units (CFU) per 6 cm of leaf midvein (+/- SD). * denotes significantly lower growth than Xam668, 2-tailed t-test, p value < 0.005. ** denotes lower growth than Xam668ΔTAL14 and Xam668ΔTAL20, 2-tailed t-test, p value = 0.006 and 0.06, respectively. B. Contributions of $TAL14_{Xam668}$ and $TAL20_{Xam668}$ to apoplastic growth were determined by measuring bacterial populations after infiltration of bacterial suspensions into the leaf apoplast. Data are represented as mean CFU/cm^2 (+/- SD). * denotes significantly lower growth than Xam668, 2-tailed t-test, p value < 0.005. Growth assays were repeated at least 3 times with similar results.

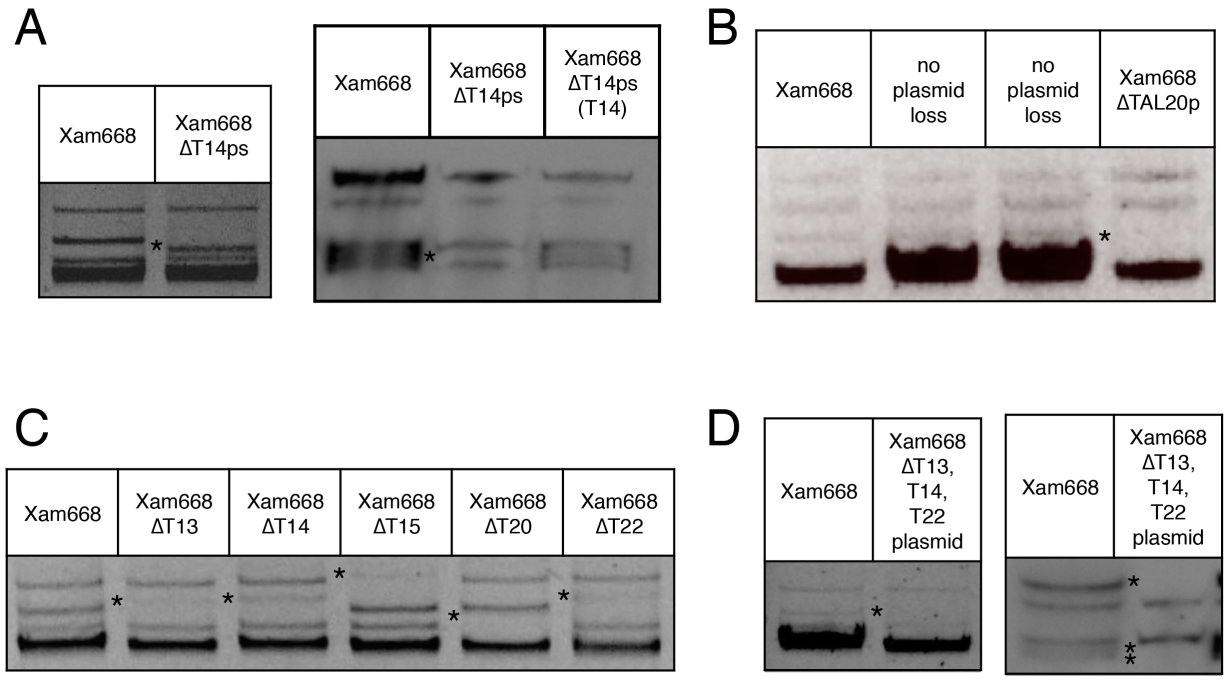


Figure 2-7. The transcription activator-like (TAL) effectors of Xam668 are on plasmids. Megaplasmid preparations were visualized by agarose gel electrophoresis. Preparations of wild type Xam668 have 3 upper bands which are plasmids and a lower band which is chromosomal DNA. A. Plasmid preparation showing downward shift of the middle Xam668 plasmid (asterisk) due to loss of the *TAL14*$_{Xam668}$ plasmid segment (ps) (left). Western blot shows loss of TAL14$_{Xam668}$ protein expression (asterisk) in Xam668ΔTAL14ps (right). B. Plasmid preparation showing loss of the smallest Xam668 plasmid (asterisk) which results in a loss of *TAL20*$_{Xam668}$. C. Plasmid preparations showing plasmids shifted up due to integration of the 18 kb pLVC18 suicide vector into the TAL effector coding region (asterisks). D. Plasmid preparation showing loss of the middle Xam668 plasmid (asterisk) (left). Western blot shows loss of TAL13$_{Xam668}$, TAL14$_{Xam668}$, and TAL22$_{Xam668}$ protein expression (asterisks) resulting from loss of the middle Xam668 plasmid (right).

growth than Xam668 and either single mutant (Figure 2-5, 2-6A). These results suggest that TAL14$_{Xam668}$ and TAL20$_{Xam668}$ make distinct contributions to *Xam* virulence.

Xam668 TAL effectors are on plasmids

Xanthomonas strains can have large plasmids that range in size from 2 kb to 183 kb as part of their genomes, and many of these plasmids carry genes associated with virulence (Sundin, 2007; Ryan *et al.*, 2011; Vivian *et al.*, 2015). The Xam668 genome contains 3 differently sized plasmids which can be visualized by agarose gel

electrophoresis (Figure 2-7). TAL effector knockout strains have an 18 kb suicide vector integrated into the TAL effector region causing the plasmid containing the TAL effector to be shifted up when visualized by agarose gel electrophoresis. Megaplasmid preparations of the TAL effector knockout strains show that *TAL20*$_{Xam668}$ is on the smallest plasmid, *TAL13*$_{Xam668}$, *TAL14*$_{Xam668}$, and *TAL22*$_{Xam668}$ are on the middle plasmid, and *TAL15*$_{Xam668}$ is on the largest plasmid. We found that plasmids likely containing the suicide vector were also reduced in copy number (Figure 2-7C). Plasmid-loss strains also give clues as to where in the genome the Xam668 TAL effectors are encoded. Xam668ΔTAL20p is missing the smallest plasmid (Figure 2-7B), Xam668ΔTAL14ps is missing a segment of the middle plasmid (Figure 2-7A), and a strain cured for the middle plasmid is missing *TAL13*$_{Xam668}$, *TAL14*$_{Xam668}$, and *TAL22*$_{Xam668}$ (Figure 2-7D). Lastly, two Xam668 genomic library cosmids (10-F12 and 10-H10) contained both *TAL14*$_{Xam668}$ and *TAL22*$_{Xam668}$ on a single DNA insert, indicating that these TAL effectors are within 40 kb of one another.

Western blot analysis was carried out in order to identify the TAL effector size repertoire of a number of *Xam* strains collected from South America and Africa (Table S1). All strains tested contain TAL effectors of similar sizes to those in Xam668 (Figure 2-8A). Strain plasmid profiles were examined by megaplasmid preparation and agarose gel electrophoresis. We did not see any correlation between TAL effector size repertoire and plasmid profile, indicating that the plasmid structure and genomic context of the TAL effectors in various *Xam* strains may be diverse (Figure 2-8B).

Discussion

TAL effectors constitute a major class of virulence factors used by xanthomonads during infection of their host plants. Sequenced *Xanthomonas* strains typically encode zero to six TAL effectors, with some encoding more than thirty. RVD number ranges from two to thirty-four, with an average of eighteen (Doyle *et al.*, 2013b). Experimental evidence indicates that ten repeats is the minimum for TAL effector function and smaller TAL effectors are likely remnants of recombination events (Boch *et al.*, 2009; Doyle *et al.*, 2013b). Xam668 is typical in its TAL effector content with a total of five differently sized TAL effectors that range in their RVD number from 13-22. All five Xam668 TAL effectors have predicted target sites in the cassava promoterome, with TAL14$_{Xam668}$ predicted to target the highest number of genes. The predicted EBEs of TAL14$_{Xam668}$, TAL15$_{Xam668}$, and TAL20$_{Xam668}$ together with the observed preference for a thymine directly upstream of the EBE (T_0) reveal a likelihood of TATA-box or TATA-box variant binding for these TAL effectors (Römer *et al.*, 2009b; Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Bernard *et al.*, 2010; Römer *et al.*, 2010; Doyle *et al.*, 2013a). TAL22$_{Xam668}$ may bind a TATA-box variant or a TC-promoter motif (Bernard *et al.*, 2010). Implications for TAL effector binding to core plant promoter motifs are discussed in section 4 of this work.

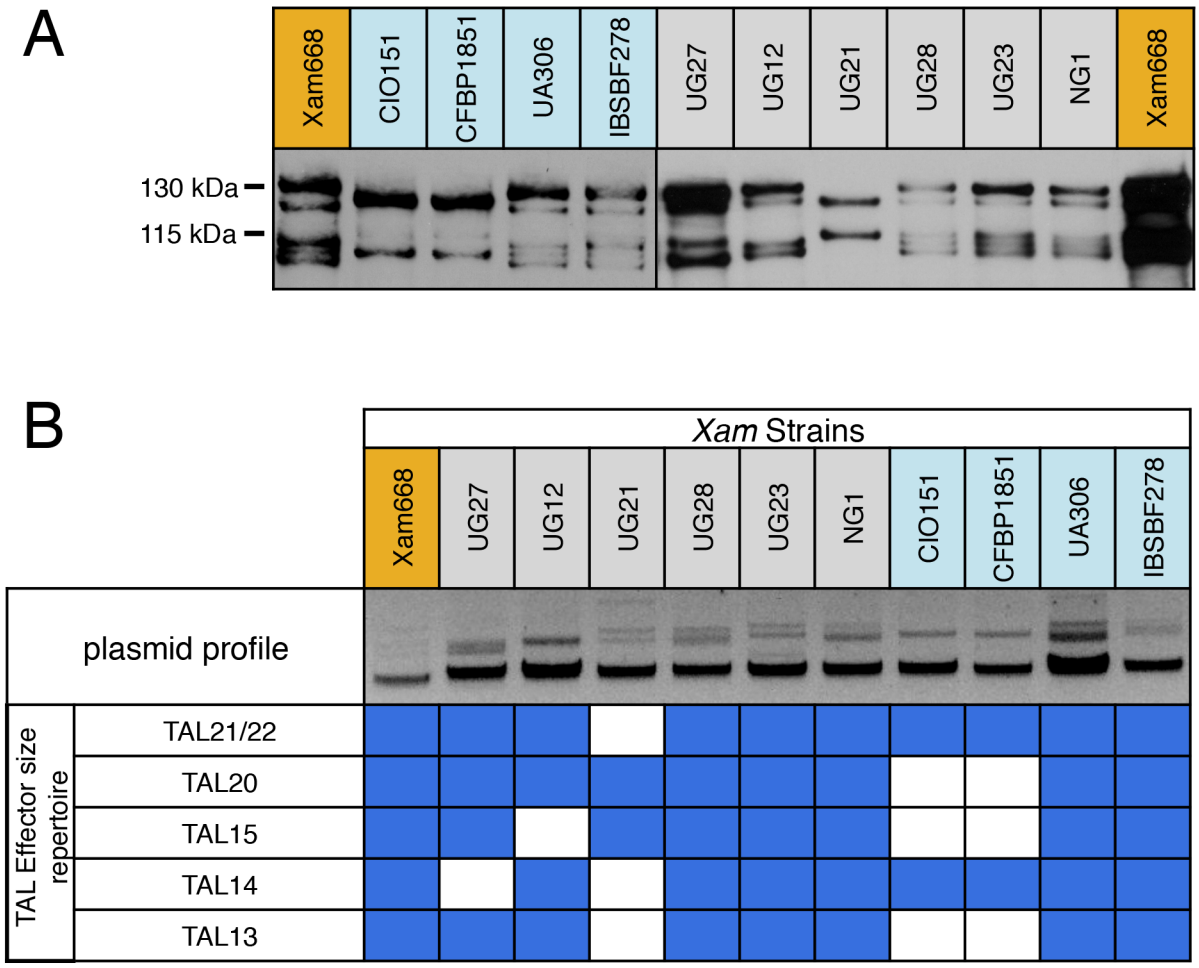


Figure 2-8. Diverse *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strain plasmid profiles do not correlate with transcription activator-like (TAL) effector repertoires. A. *Xam* strains are from Asia, Africa, and South America (highlighted in orange, grey, and light blue, respectively).TAL effector sizes for each strain were estimated based on western blot. B. Plasmid profiles of diverse *Xam* strains were visualized by agarose gel electrophoresis. Results of the western blot are displayed figuratively with blue and white squares indicating the presence or absence of TAL effectors with the specified number of repeat variable diresidues (RVDs), respectively.

A differential contribution to virulence was found for the five TAL effectors of Xam668 which is consistent with what is reported in the literature for other TAL effector family members. $TAL20_{Xam668}$ influences bacterial growth and symptom formation, $TAL14_{Xam668}$ contributes to growth, and $TAL13_{Xam668}$, $TAL15_{Xam668}$, and $TAL22_{Xam668}$ show no detectable effect on virulence. Like $TAL20_{Xam668}$, TAL effectors avrXa7 and PthXo1 of *Xoo,* PthA of *Xcc* and $TALE1_{Xam}$ of *Xam* strain CFBP1851 have been reported to contribute to both growth and symptom formation (Swarup *et al.*, 1991; Bai

et al., 2000; Yang & White, 2004; Castiblanco *et al.*, 2012). The phenotype of Xam668ΔTAL14 shows that an influence on bacterial growth can be uncoupled from symptom development. Similarly, Tal2g of *Xoc* contributes to lesion length and watersoaking in rice, but not to bacterial growth (Cernadas *et al.*, 2014), and AvrB6 from *X. campestris* pv. *malvacearum* (*Xcm*) contributes to watersoaking and necrosis, but not to growth, in cotton (*Gossypium hirsutum*) (Yang *et al.*, 1994). TAL effectors that have no effect when mutated have also been reported numerous times in the literature (Yang *et al.*, 1996; Bai *et al.*, 2000; Cernadas *et al.*, 2014).

The five TAL effectors of Xam668 appear to be encoded on plasmids. Other *Xanthomonas* TAL effectors reside on plasmids including *avrBs3* of *Xe*, several TAL effectors from *X. campestris* pv. *malvacearum*, and $TALE1_{Xam}$ from *Xam* strain CFBP1851 (Bonas *et al.*, 1989; Yang *et al.*, 1996; Castiblanco *et al.*, 2012). All of the *Xam* strains analyzed in this work contain TAL effectors which are likely encoded on plasmids similarly to Xam668. Interestingly, strains that have identical TAL effector size profiles do not necessarily have identical plasmid profiles. Repetitive DNA regions are active sites for homologous recombination and it has been shown that intergenic recombination can occur between TAL effectors (Yang & Gabriel, 1995). These recombination events may lead to rearrangements of the structure of TAL effector-containing plasmids, resulting in variations in plasmid profiles among strains.

The cassava gene targets of $TAL14_{Xam668}$ and $TAL20_{Xam668}$ are of interest because of the virulence defects seen in Xam668ΔTAL14 and Xam668ΔTAL20. Knowledge of the RVD sequences of $TAL14_{Xam668}$ and $TAL20_{Xam668}$ allowed us to predict target sites for these TAL effectors within the cassava promoterome using the TAL effector-DNA binding code. We discovered a large number of predicted targets, especially for $TAL14_{Xam668}$, making a transcriptomic approach necessary to identify *bona fide* targets. In the next sections of this work we use RNA-Sequencing (RNA-Seq) to discover the cassava genes targeted by $TAL14_{Xam668}$ and $TAL20_{Xam668}$.

3: *Xanthomonas axonopodis* pv. *manihotis* virulence is promoted by a $TAL20_{Xam668}$-mediated induction of the SWEET sugar transporter *MeSWEET10a* in cassava

Background

Xanthomonas transcription activator-like (TAL) effectors function by directly binding and activating host genes whose over- or ectopic expression aids in disease progression (Boch & Bonas, 2010; Bogdanove *et al.*, 2010, Boch *et al.*, 2014). A growing number of *Xanthomonas* TAL effector targeted genes have been shown to contribute to host susceptibility, including citrus *CsLOB1*, pepper *UPA20*, rice *OsSULTR3;6*, and members of the SWEET family of sugar transporters in rice (Yang *et al.*, 2006; Kay *et al.*, 2007; Chen *et al.*, 2010; Antony *et al.*, 2010; Chen *et al.*, 2012; Cernadas *et al.*, 2014; Hu *et al.*, 2014; Li *et al.*, 2014; Hutin *et al.*, 2015). In chapter 2 of this work, we showed that $TAL20_{Xam668}$ of the cassava bacterial blight pathogen *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) contributes to *in planta* bacterial growth and watersoaking symptom development. In this chapter we show that *MeSWEET10a*, a member of the SWEET family of sugar transporters, is directly activated by $TAL20_{Xam668}$ and contributes to cassava's susceptibility to *Xam*.

Results

RNA-Seq identifies genes upregulated by $TAL20_{Xam668}$

RNA-Sequencing (RNA-Seq) was utilized to identify cassava genes induced by $TAL20_{Xam668}$. A clone from our Xam668 genomic DNA cosmid library encoding *$TAL20_{Xam668}$* (clone 3-H3) was conjugated into *X. euvesicatoria* (*Xe)* 85-10. *Xe* has no endogenous TAL effectors and does not cause disease on cassava, so served to deliver $TAL20_{Xam668}$ independently of the other TAL effectors of Xam668 (Figure 3-1). The RNA-Seq experimental design encompassed three time points (8, 24, 50 hours post inoculation) and four experimental conditions [mock, Xam668, *Xe* and *Xe*($TAL20_{Xam668}$)]. Genes upregulated in Xam668 compared to mock infiltration were considered potential targets of the five Xam668 TAL effectors, while genes upregulated in *Xe*($TAL20_{Xam668}$) compared to *Xe* were considered to be $TAL20_{Xam668}$-specific (Figure 3-2). In total, three biological replicates for each condition were compared.

$TAL20_{Xam668}$ directly targets *MeSWEET10a*

The RNA-Seq data analysis focused on identifying genes induced similarly in timing and intensity between mock versus Xam668 and *Xe* versus *Xe*($TAL20_{Xam668}$). These

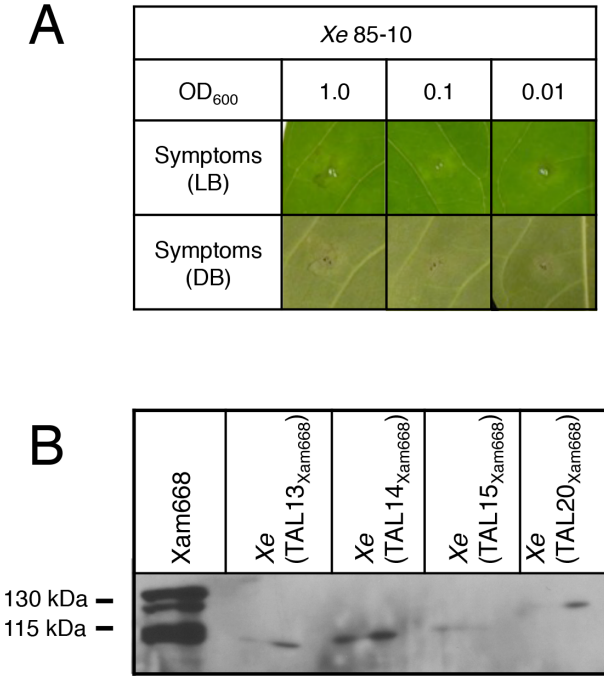


Figure 3-1. *Xanthomonas euvesicatoria* (*Xe*) does not invoke a hypersensitive response (HR) or disease symptoms when infiltrated into cassava leaves. A. *Xe* 85-10 cells were infiltrated into cassava leaves as OD_{600} = 1.0, 0.1, 0.01. Two replicates were inoculated for each OD_{600} with similar results. Leaves were photographed 3 days post inoculation on a light box (LB, better for visualizing HR) and on a dark background (DB, better for visualizing watersoaking disease symptoms). B. Xam668 cosmids or pVSP61 plasmids expressing single transcription activator-like (TAL) effectors were conjugated into *Xe* 85-10 and western blot analysis was used to confirm TAL effector expression. $TAL15_{Xam668}$, $TAL20_{Xam668}$, and $TAL22_{Xam668}$ are expressed from cosmid clones, while $TAL13_{Xam668}$ and $TAL14_{Xam668}$ driven by the $TAL20_{Xam668}$ promoter are expressed from the pVSP61 plasmid.

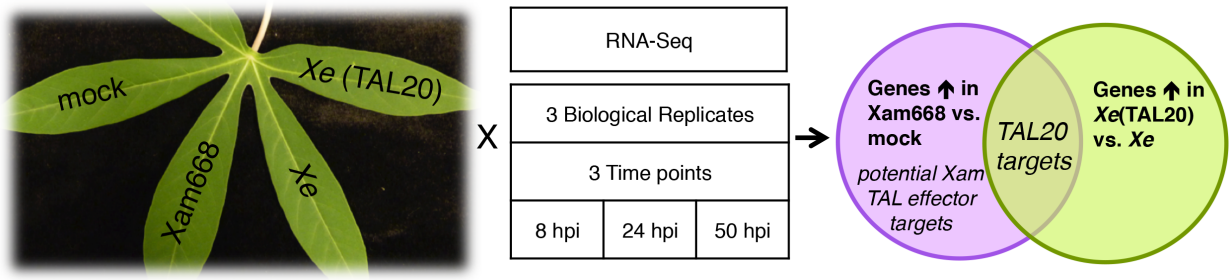


Figure 3-2. Cartoon of strategy for experimental identification of $TAL20_{Xam668}$ targets by RNA sequencing (RNA-seq). hpi, hours post inoculation.

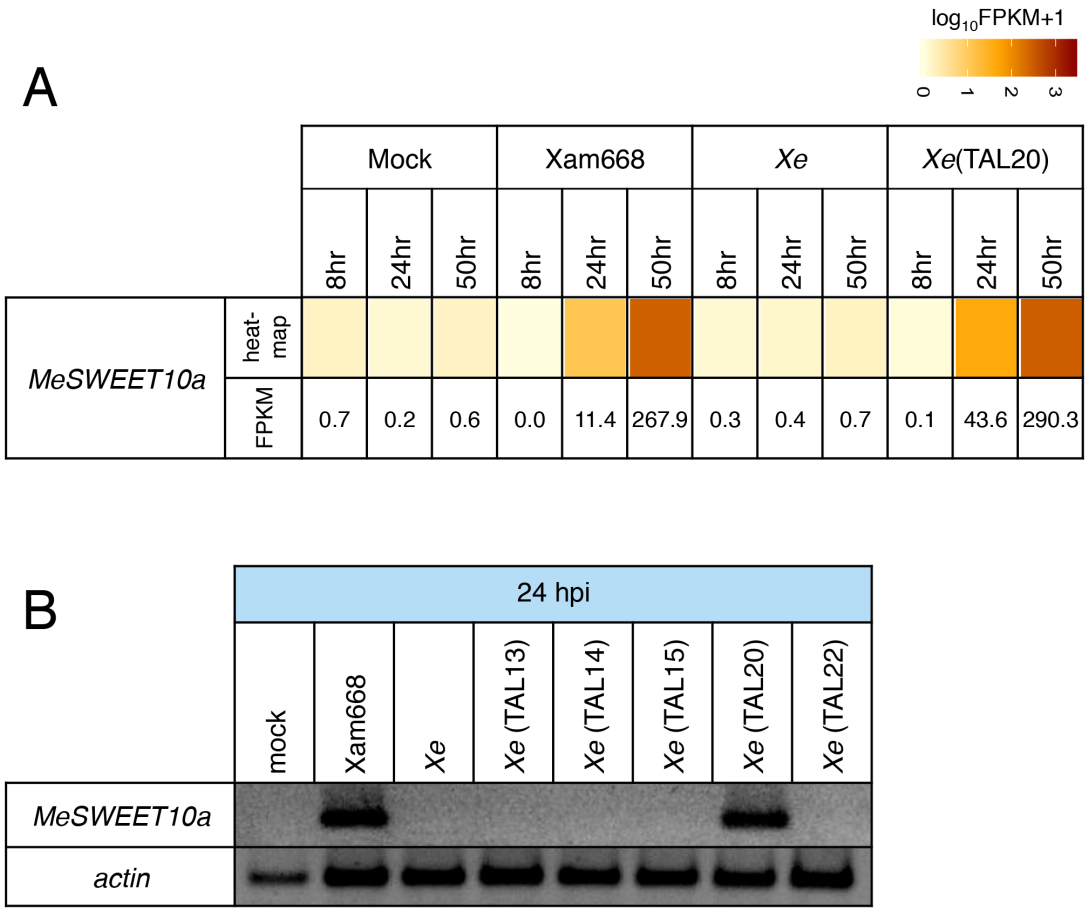


Figure 3-3. TAL20$_{Xam668}$ activates expression of *MeSWEET10a.* A. RNA-Sequencing data showing expression of *MeSWEET10a* in mock (wounding), Xam668 [5 distinct transcription activator-like (TAL) effectors], *Xanthomonas euvesicatoria (Xe)* 85-10 (no TAL effectors) and *Xe* 85-10 (TAL20$_{Xam668}$)-inoculated cassava leaves 8, 24, and 50 hours post inoculation (hpi). Heatmap of log_{10}FPKM values represents the mean of 3 biological replicates. B. Semiquantitative RT-PCR of *MeSWEET10a* expression 24 hpi in leaf tissue inoculated at an OD_{600} of 0.5 with *Xe* 85-10 delivering each of the 5 Xam668 TAL effectors. Experiment was repeated 3 times with similar results.

analyses yielded a single putative target for TAL20$_{Xam668}$, *cassava4.1_013474*, henceforth referred to as *MeSWEET10a*, a member of the SWEET sugar transporter family of TAL effector-targeted susceptibility genes (Chen *et al.*, 2010; 2012) (Figure 3-3A). TAL effector dependence for activation was tested using *Xe* delivering individual Xam668 TAL effectors. Only TAL20$_{Xam668}$ was able to activate *MeSWEET10a* (Figure 3-3B). In addition, *MeSWEET10a* was not activated by Xam668ΔTAL20, suggesting that TAL20$_{Xam668}$ is solely responsible for induction of this gene (Figure 3-5C).

To test if *MeSWEET10a* activation is directly induced by TAL20$_{Xam668}$ we pursued a transient approach in which the *MeSWEET10a* promoter was used to drive expression

of the reporter gene luciferase in an *Agrobacterium* binary vector. The resulting construct was co-expressed with 35S-driven *TAL20*$_{Xam668}$ delivered via *Agrobacterium* in *Nicotiana benthamiana* leaf cells. Luciferase activity was measured 24 hours post inoculation. *Agrobacterium* delivering *TAL20*$_{Xam668}$ specifically induced expression of the *MeSWEET10a* promoter. Previous research has demonstrated that a *X. oryzae* pv. *oryzae* TAL effector binding element (EBE) placed in the *Bs3* promoter from pepper (normally targeted by *Xe* TAL effector AvrBs3) retains its function (Römer *et al.*, 2009a), so we tested if this was true for the predicted TAL20$_{Xam668}$-targeted EBE from the *MeSWEET10a* promoter. While TAL20$_{Xam668}$ did not induce expression of the wild type *Bs3* promoter, when the predicted TAL20$_{Xam668}$ EBE was inserted into the *Bs3* promoter, activation was observed (Figure 3-4A). We note that the tested promoters do not display identical background levels and speculate that this is a result of their differential interaction with endogenous transcription factors.

Electromobility Shift Assays (EMSAs) confirmed direct interaction between TAL20$_{Xam668}$ and an oligonucleotide encompassing the predicted TAL20$_{Xam668}$ EBE from the *MeSWEET10a* promoter (Figure 3-4B). Additionally, *MeSWEET10a* gene activation was seen in the presence of 50 μM cycloheximide (CHX), an inhibitor of eukaryotic protein synthesis (Figure 3-4C). Taken together, our data show that the *MeSWEET10a* promoter is directly bound and activated at the predicted EBE by TAL20$_{Xam668}$.

Designer TAL effectors restore Xam668ΔTAL20 virulence defects

To test the role of *MeSWEET10a* as a host susceptibility gene, designer TAL effectors (dTALEs) targeting distinct places within the *MeSWEET10a* promoter (dT$_{MeSWEET10a-3}$ and dT$_{MeSWEET10a-4}$) were assembled with code optimized repeat variable diresidue (RVD) sequences (Figure 3-5A, B) (Boch *et al.*, 2009; Moscou & Bogdanove, 2009; Morbitzer *et al.*, 2010). dT$_{MeSWEET10a-3}$ and dT$_{MeSWEET10a-4}$ were conjugated into Xam668ΔTAL20 and the resulting strains were assessed for restoration of virulence phenotypes. Both dT$_{MeSWEET10a-3}$ and dT$_{MeSWEET10a-4}$ were able to activate expression of *MeSWEET10a*, and both complemented symptom development (Figure 3-5C) and bacterial growth (Figure 2-6A) in Xam668ΔTAL20. Complementation of the Xam668ΔTAL20 mutant phenotype by dTALEs that activate *MeSWEET10a* shows that TAL20$_{Xam668}$ contributes to *Xam* virulence through the activation of *MeSWEET10a*, and that MeSWEET10a carries out a function in the plant that, when co-opted by *Xam*, promotes disease susceptibility.

MeSWEET10a is a sugar transporter

Phylogenetic analysis indicates that *MeSWEET10a* belongs to clade III of the SWEET sugar transporter family and is most closely related to the *Arabidopsis* sucrose

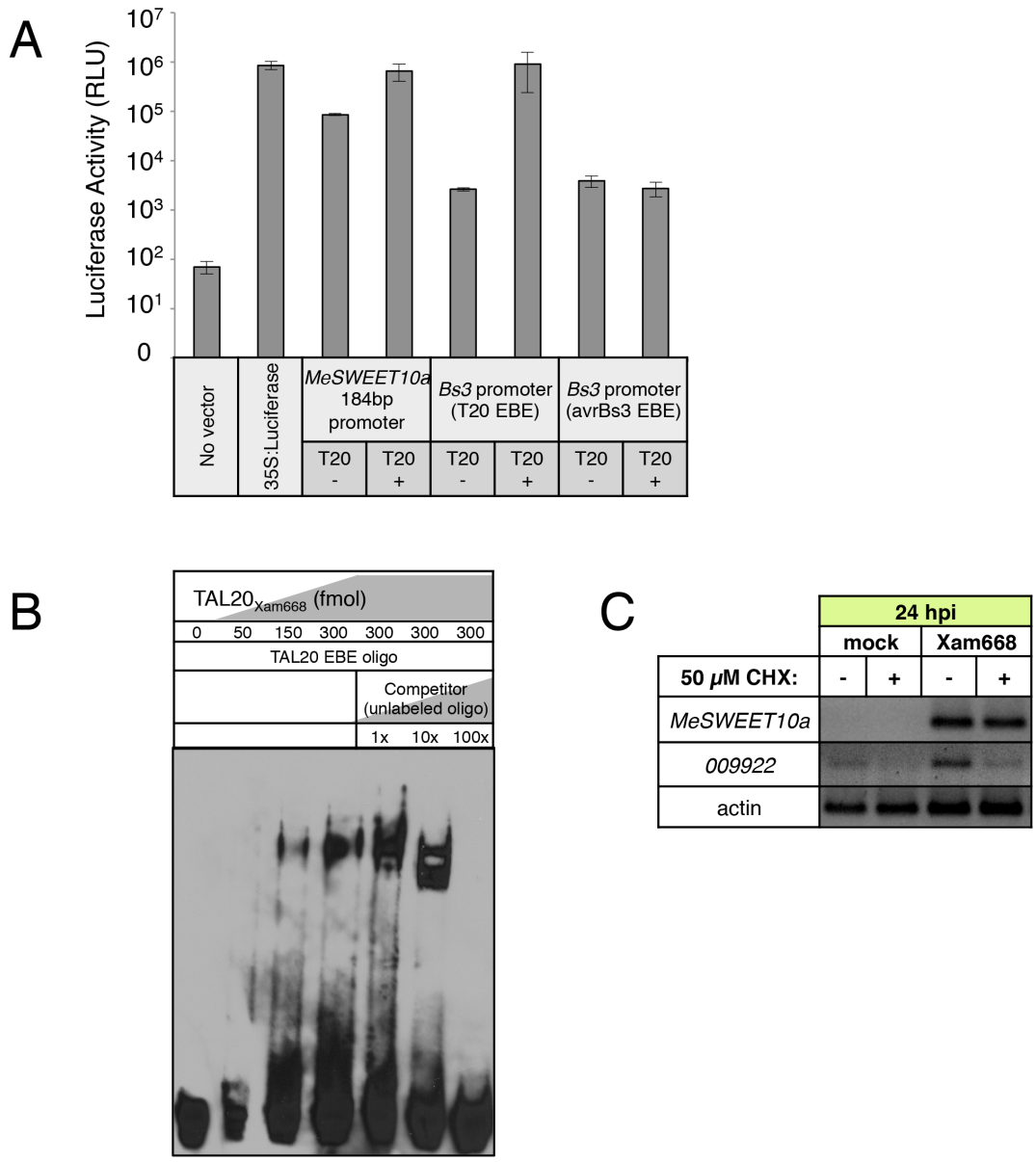


Figure 3-4. TAL20$_{Xam668}$ directly activates expression of *MeSWEET10a.* A. *Agrobacterium* was used to deliver promoter-luciferase fusion constructs and 35S-driven *TAL20$_{Xam668}$* into *Nicotiana benthamiana* cells. Luciferase activity (relative light units, RLU) was quantified 24 hours post inoculation (hpi). Data are displayed as mean technical replicates (+/- SD). Experiments were repeated 3 times with similar results; EBE, effector binding element. B. Electromobility shift assay (EMSAs) demonstrates binding of TAL20$_{Xam668}$ to its predicted EBE from the *MeSWEET10a* promoter. C. Co-infiltration of 50 μM cycloheximide (CHX) was used to block expression of secondary transcripts prior to semiquantitative RT-PCR. Leaf tissue was collected 24 hpi. Actin expression is shown as a loading control, and expression of *009922* is shown as a control for the effect of CHX on secondary transcript accumulation. Displayed gels are representative of 3 independent experiments.

A *MeSWEET10a* promoter

TGTTCGGATCATTTTAACGTAAAAAGCACTTAGTAGGACCAACTTAAAAGTAACCATAAAATTAATATTGTTTTC
TCTTTCTTCAAAAAAAATAATAAAAGAAACAAGGCCACTGTTACATTGACATATTTTATTCACTTTAATCATGCA
TGCAACTTGACTTCATTCCGTTCCCTGGATTCCTCCCCTATATAAACGCTTCTCGCCCATCCATCATTGCACAAC
ATAGCTAGAGTTTCCTCTTGAGAAAGAGAGTTTCCTCTGCACAAGGGAAAGAGAGTCTCTACTCTCGCCGGAGAA

B

dTALE	RVD sequence	Target EBE
dTALE$_{MeSWEET10a-3}$	NG NN NI HD NG NG HD NI NG NG HD HD NN NG NG HD HD HD	TGACTTCATTCCGTTCCC
dTALE$_{MeSWEET10a-4}$	NG HD HD HD NG NN NN NI NG NG HD HD NG HD HD HD HD NG	TCCCTGGATTCCTCCCCT

C

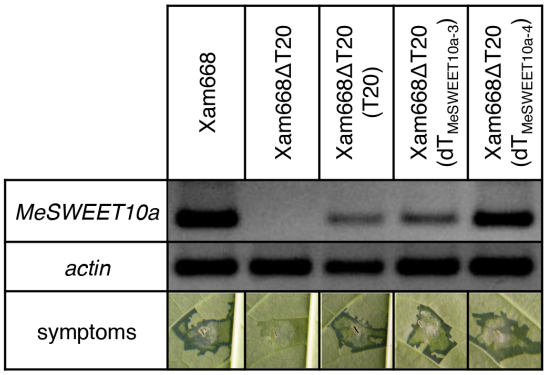


Figure 3-5. *MeSWEET10a*-targeted designer transcription activator-like (TAL) effectors (dTALEs) restore virulence defects of Xam668ΔTAL20. A. 300 bp of the *MeSWEET10a* promoter directly upstream of the ATG are shown with the TAL20$_{Xam668}$ effector binding element (EBE) boxed, the dTALE$_{MeSWEET10a-3}$ EBE highlighted in grey, and the dTALE$_{MeSWEET10a-4}$ EBE underlined. B. Table of dTALE repeat variable diresidues (RVDs) and sequences of corresponding target EBEs. C. dTALEs designed to target locations within the *MeSWEET10a* promoter were conjugated into Xam668ΔTAL20. Semiquantitative RT-PCR was used to confirm *MeSWEET10a* transcript induction 24 hours post inoculation. Watersoaking symptom development was assessed at 5 days post inoculation for Xam668ΔTAL20 and dTALE-complemented strains inoculated at OD_{600} 0.1. Experiments were repeated 3 times with similar results.

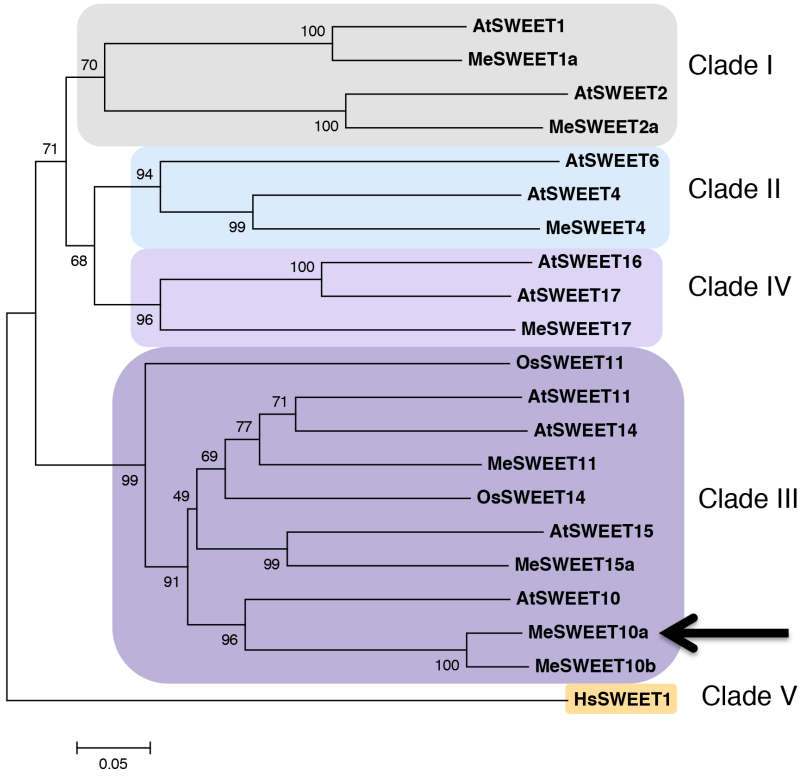
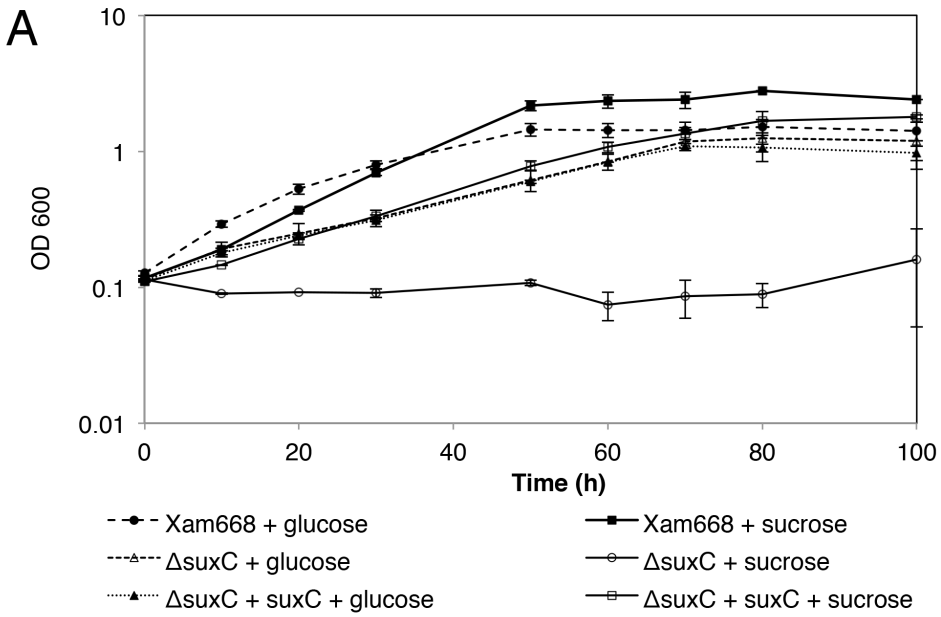


Figure 3-6. MeSWEET10a is a member of the clade III family of SWEET sugar transporters. SWEET proteins from *Arabidopsis thaliana* (At), *Oryzae sativa* (Os), and *Homo sapiens* (Hs) were chosen to represent clades of the SWEET superfamily (Chen *et al.*, 2010). The phylogenetic tree was made using the neighbor-joining method in MEGA version 5.2.2 (Tamura *et al.*, 2011). Bootstrap values (percentage of 1000) are displayed.

transporter *AtSWEET10* (Chen *et al.*, 2012) (Figure 3-6). Several recent studies indicate that pathogens induce sucrose export from cells into the apoplasm to be used as a carbon source (Chen *et al.*, 2010; Sonnewald, 2011; Chen *et al.*, 2012; Chen, 2014). If indeed the *Xam* virulence strategy follows this model, we hypothesized that disrupting the ability of *Xam* to import sucrose should result in a similar phenotype as that observed from blocking induction of *MeSWEET10a* (Xam668ΔTAL20). *SuxC* was identified as part of an operon from *X. campestris* pv. *campestris* (*Xcc*) that is specific for sucrose import and essential to *Xcc* fitness (Blanvillain *et al.*, 2007). To test the above hypothesis, we generated a deletion mutant in the orthologous *SuxC* gene of *Xam* and confirmed that Xam668ΔsuxC was unable to grow when sucrose was provided as the only carbon source (Figure 3-7A). Notably, in minimal media, Xam668ΔsuxC could still utilize glucose. We inoculated this mutant into cassava leaves and observed no significant reduction in growth (Figure 3-7B). While the watersoaking caused by Xam668ΔTAL20 was reduced, watersoaking by Xam668ΔsuxC was similar to wild type Xam668 (Figure 3-7C). These results indicate that the induction of *MeSWEET10a*



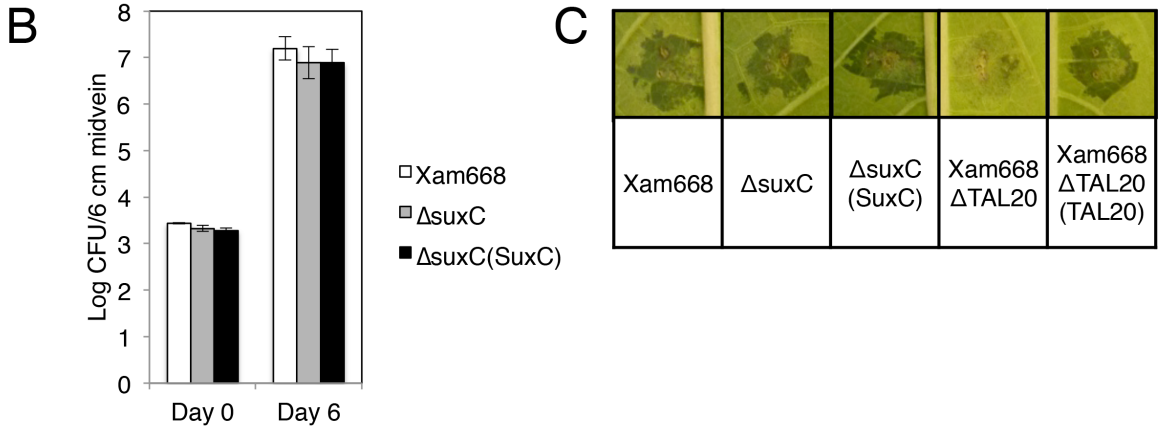


Figure 3-7. Xam668ΔsuxC does not have the same virulence defects as Xam668ΔTAL20. A. A full gene knockout was constructed for the *SuxC* gene in Xam668. Bacterial growth of XamΔsuxC was compared to wild type Xam668 and Xam668ΔsuxC(SuxC) in minimal media supplemented with either sucrose or glucose. B. *In planta g*rowth of Xam668ΔsuxC was compared to Xam668 and Xam668ΔsuxC(SuxC). Bacterial populations in leaf midveins were measured at days 0 and 6. Data are represented as mean colony forming units (CFU)/6 cm leaf midvein (+/- SD). Growth assays were repeated at least 3 times with similar results. C. Symptom development was compared for Xam668, Xam668ΔsuxC, Xam668ΔTAL20, and complemented mutant lines. Pictures representative of 7 independent replicates were taken 5 days post inoculation (OD_{600} = 0.1).

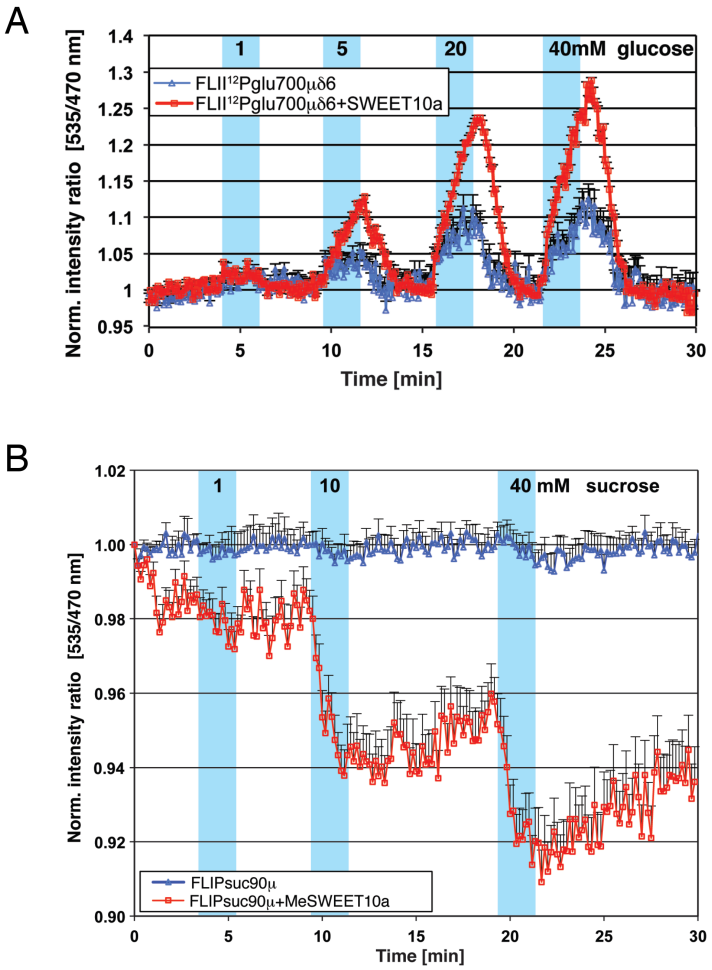


Figure 3-8. MeSWEET10a is a functional sugar transporter. A. MeSWEET10a mediates glucose transport activity as detected by co-expression with the cytosolic Förster resonance energy transfer (FRET) glucose sensor FLII^{12}Pglu700$\mu\partial$6 in HEK293T cells (Wilkins *et al.*, 2015). Individual cells were analysed by quantitative ratio imaging of eCFP and Citrine emission (acquisition interval 5 s). HEK293T cells were perfused with culture medium, followed by square pulses of increasing glucose concentrations. Blue line indicates cells expressing sensor alone; red line indicates cells co-expressing sensor and MeSWEET10a; accumulation of glucose is indicated by a positive FRET ratio change (mean +SE; $n > 15$). Experiments were repeated with comparable results at least four times. B. MeSWEET10a mediates sucrose transport activity as detected by co-expression with the cytosolic FRET sucrose sensor FLIPsuc90$\mu\Delta$1a in HEK293T cells. Individual cells were analysed by quantitative ratio imaging of eCFP and Aphrodite emission (acquisition interval 10 s). HEK293T cells were perfused with culture medium, followed by square pulses of increasing sucrose concentrations. Blue line indicates cells expressing sensor alone; red line indicates cells co-expressing sensor and MeSWEET10a; accumulation of sucrose is indicated by a negative FRET ratio change (mean +SE; $n > 19$). Experiments were repeated with comparable results at least four times.

contributes to bacterial growth and watersoaking independently of the bacteria's ability to directly utilize sucrose in the apoplasm.

While clade III transporters typically function as sucrose transporters, some members of this clade have been found to transport both glucose and sucrose. To test the sugar transport capabilities of MeSWEET10a, the protein was expressed in mammalian cells together with sucrose and glucose Förster resonance energy transfer (FRET) sensors (Chen *et al.*, 2010). Sucrose and glucose-induced FRET responses indicate that MeSWEET10a can transport both sucrose and glucose (Figure 3-8).

TAL20$_{Xam668}$ increases watersoaking conferred by TAL20-deficient strain *Xam* CIO151

Xam strain CIO151 lacks TAL20 and causes less watersoaking than Xam668 (Bart *et al.*, 2012). To test if TAL20$_{Xam668}$ would increase watersoaking of *Xam* CIO151, native promoter-driven *TAL20$_{Xam668}$* was transferred to *Xam* CIO151 by conjugation. Watersoaking caused by *Xam* CIO151(TAL20$_{Xam668}$) was greater than that caused by *Xam* CIO151 (Figure 3-9A). Semiquantitative RT-PCR verified *MeSWEET10a* activation by *Xam* CIO151(TAL20$_{Xam668}$) indicating that MeSWEET10a plays a role in the increased watersoaking seen in this strain (Fig. 3-9B). However, *Xam* CIO151(TAL20$_{Xam668}$) did not grow to significantly higher levels *in planta* than *Xam* CIO151 (Figure 3-9C).

Correlating induced genes with specific TAL effectors

To identify targets of TAL13$_{Xam668}$, TAL14$_{Xam668}$, TAL15$_{Xam668}$, and TAL22$_{Xam668}$, we examined our RNA-Seq data for genes that were highly induced by Xam668 as compared to mock in at least two of the three time points (8 and 24 hpi; 24 and 50 hpi, or in all three time points). We cross-referenced the set of induced genes with the list of putative EBEs predicted by Target Finder and found that many induced genes contained EBEs with what are considered good scores (within a 3-fold cutoff of the best possible score for the given TAL effector) within their promoters (Doyle *et al.*, 2012). A heatmap for a manually selected subset of the most highly induced candidate TAL effector targets is presented (Figure 3-10).

We assigned TAL effector specificity to the subset of putative target genes by carrying out semiquantitative RT-PCR on cassava leaves inoculated with *Xe* strains delivering individual Xam668 TAL effectors. Of the 17 genes tested, we found that 1 was upregulated by TAL13$_{Xam668}$, 9 by TAL14$_{Xam668}$, 2 by TAL15$_{Xam668}$, and 4 by TAL22$_{Xam668}$. No additional genes were found to be upregulated by TAL20$_{Xam668}$ (Figure 3-10, right). To test whether the TAL effectors were directly activating these target genes, we carried out inoculations in the presence of 50 μM CHX. All genes tested were

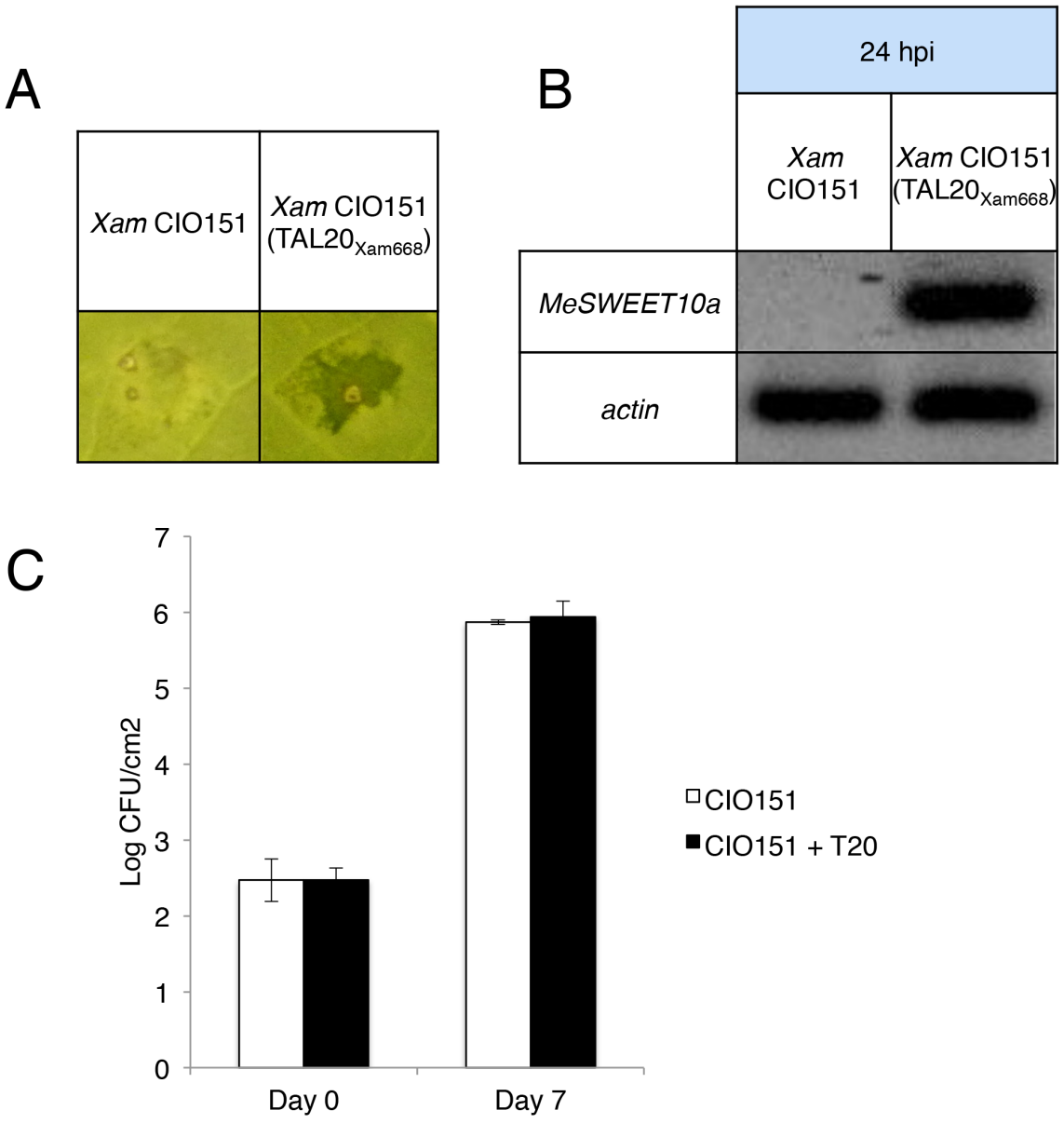


Figure 3-9. TAL20$_{Xam668}$ increases watersoaking symptoms of *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strain CIO151. A. Native promoter-driven TAL20$_{Xam668}$ was introduced into *Xam* CIO151 by bacterial conjugation. Strains were inoculated into cassava leaves (OD_{600} = 0.01) and symptom development was recorded after 6 days. B. Semiquantitative RT-PCR was used to confirm *MeSWEET10a* transcript induction by *Xam* CIO151(TAL20$_{Xam668}$) 24 hours post inoculation (hpi). C. Contribution of TAL20$_{Xam668}$ to growth of *Xam* strain CIO151 was determined by measuring bacterial populations after infiltration of bacterial suspensions into the leaf apoplast. Data are represented as mean colony forming units (CFU)/cm^2 (+/- SD). Experiment was repeated 2 times with similar results.

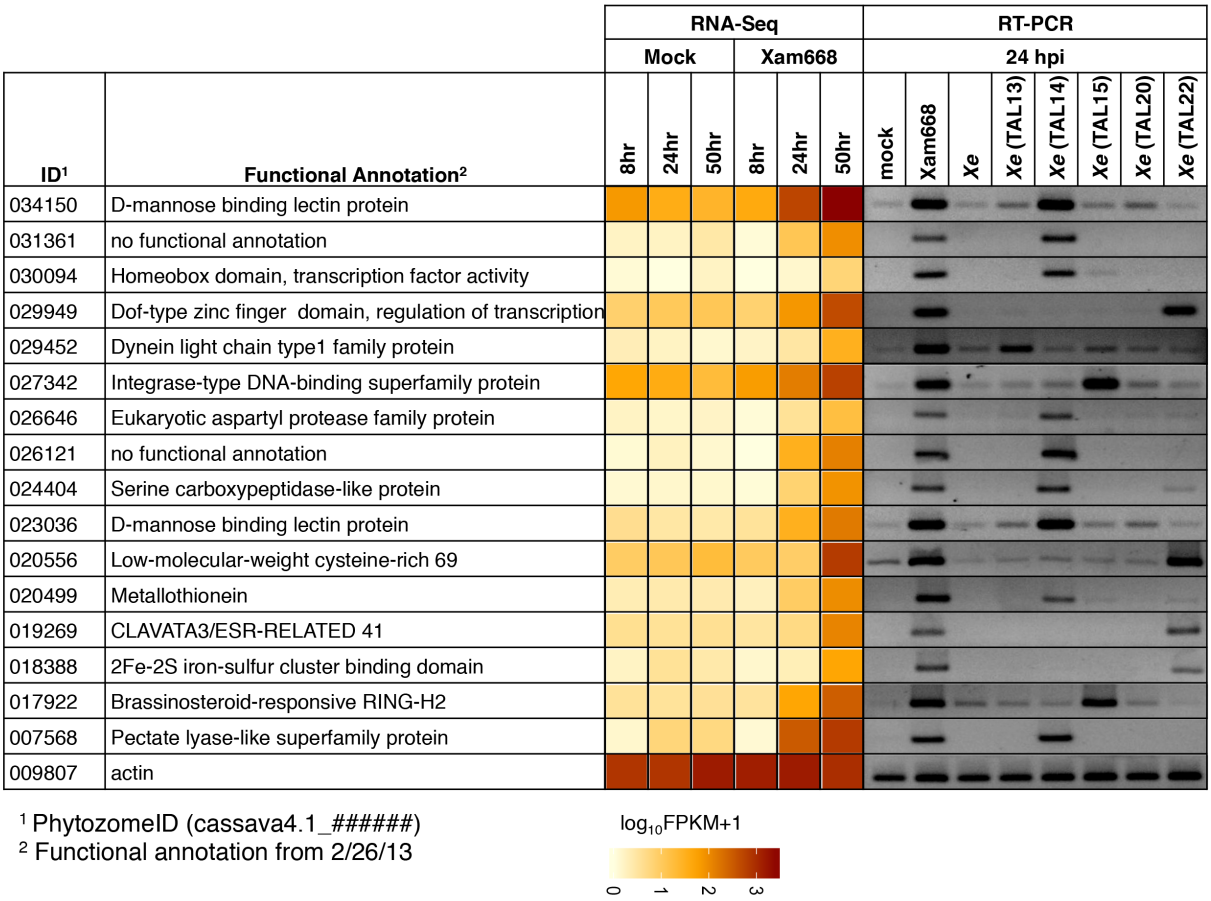


Figure 3-10. Specific gene upregulation during *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) infection can be attributed to individual transcription activator-like (TAL) effectors. Subset of RNA-Sequencing (RNA-Seq) data showing genes induced by Xam668 (5 distinct TAL effectors) but not in mock (wounding) inoculated cassava leaves shows potential TAL effector targets. Heatmap of log_{10}FPKM values represents the mean of 3 biological replicates. Semiquantitative RT-PCR of leaf tissue inoculated (OD_{600} = 0.5) with *X. euvesicatoria (Xe)* 85-10 delivering each of the 5 Xam668 TAL effectors shows TAL effector-dependent gene activation. Experiment was repeated 3 times with consistent results. The PCR cycle number for each gene primer set is listed in Supplementary Table 2.

activated in a TAL effector dependent manner both in the presence and absence of CHX, indicating that activation is direct (Figure 3-11). A slight decrease in expression was observed for *cassava4.1_024404* and *cassava4.1_026646* suggesting that these two genes may be indirect targets. However, the expression was not reduced to background levels, so the decrease is likely due to biological variation between samples. Most of the validated targets have one or more EBEs specific to the TAL effectors by which they are activated in their promoters. However, two direct target

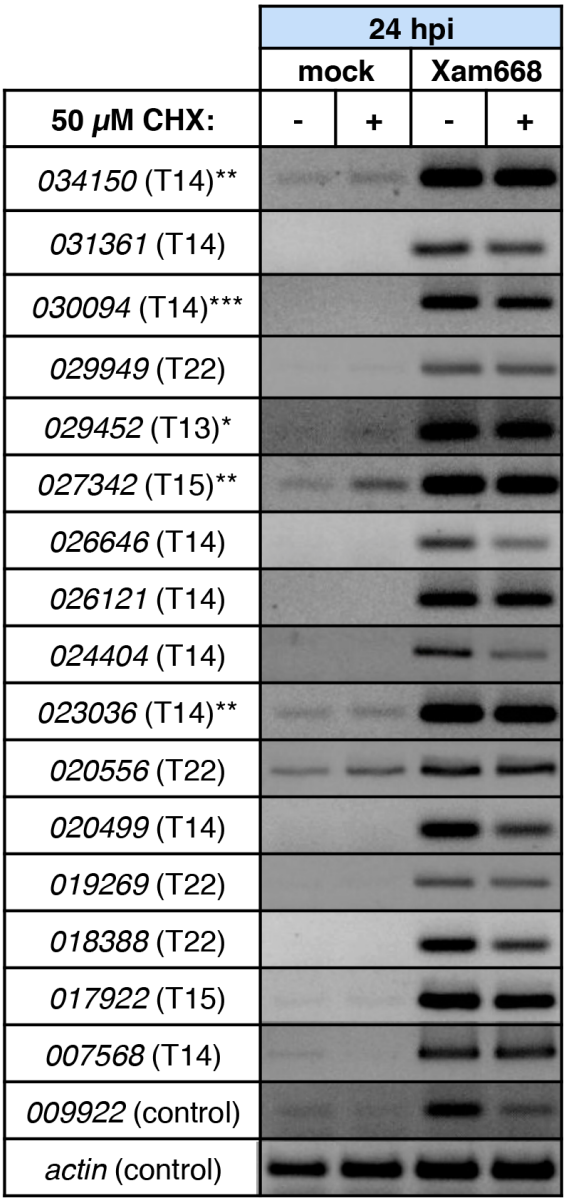


Figure 3-11. Xam668 transcription activator-like (TAL) effectors directly activate cassava genes. For each target gene, activation by mock infiltration and Xam668 was measured by semiquantitative RT-PCR after inoculation in the presence and absence of 50 μM cycloheximide (CHX), which was used to block activation of secondary transcripts. The TAL effector specific to each target gene is indicated in parentheses. Leaf tissue was collected 24 hours post inoculation (hpi). The PCR cycle number for each gene primer set is listed in Supplementary Table 2. Displayed gels are representative of 2 independent experiments. Actin expression is shown for all samples as a loading control, and expression of *009922* is shown as a control for the effect of CHX on secondary transcript accumulation.

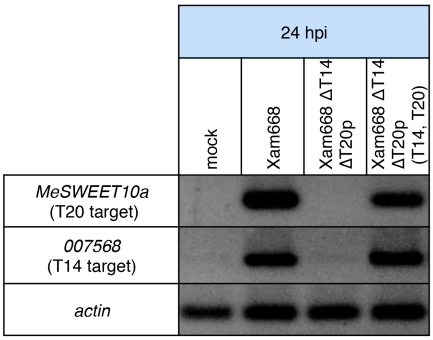


Figure 3-12. Confirmation of target gene induction by TAL14$_{Xam668}$ and TAL20$_{Xam668}$. Semiquantitative RT-PCR was used to confirm *MeSWEET10a* and *007568* transcript induction by Xam668ΔT14/ΔT20p(T14,T20) 24 hours post inoculation (hpi).

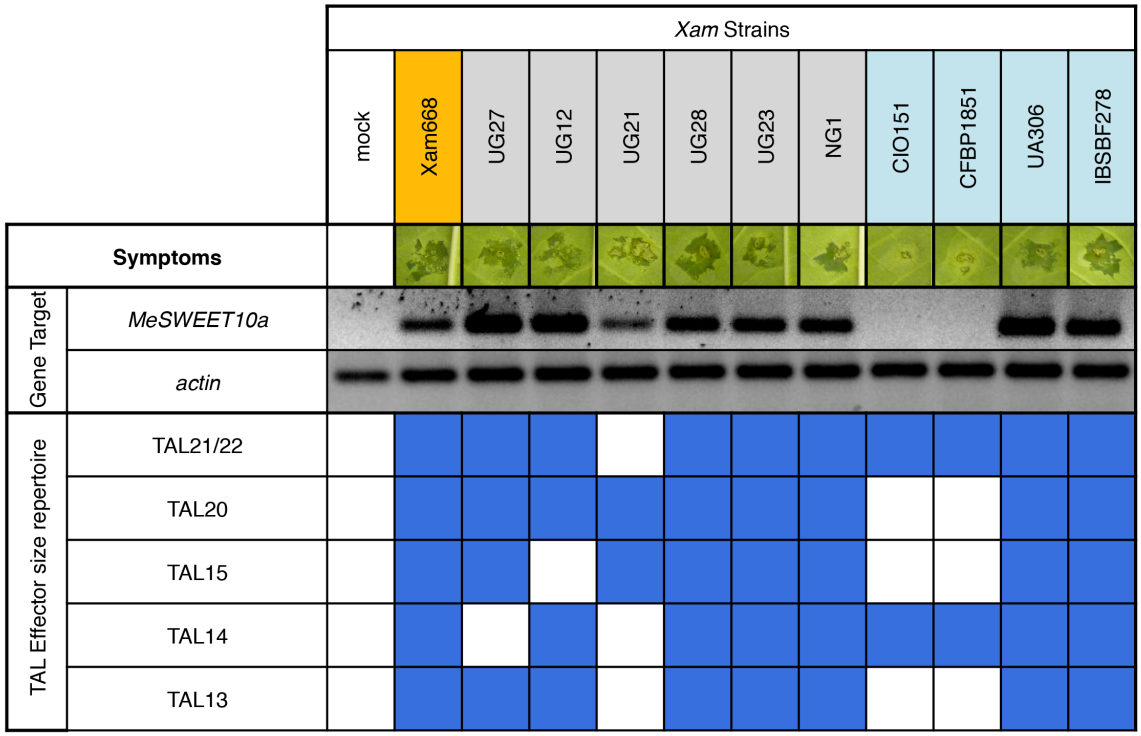


Figure 3-13. Multiple *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strains from Asia, Africa, and South America (highlighted in orange, grey, and light blue, respectively) activate *MeSWEET10a*. Semiquantitative RT-PCR shows activation of *MeSWEET10a* by various strains. Results of the western blot in figure 2-8A are displayed figuratively with blue and white squares indicating the presence or absence of transcription activator-like (TAL) effectors with the specified number of repeat variable diresidues (RVDs), respectively. Symptom development for each strain is shown. Leaves were imaged 7 days post inoculation (OD_{600} = 0.01).

promoters (*cassava4.1_020556* and *cassava4.1_018388*) do not contain predicted EBEs specific to the TAL effectors that activate them. The most plausible explanation for this is the presence of promoter polymorphisms between experimental cultivar TMS 60444 and the reference genome AM560-2. The *cassava4.1_020499* promoter contains multiple $TAL14_{Xam668}$ EBEs when the stringency for EBE score is relaxed to be within 3.5-fold of the best possible score. Similarly, the *cassava4.1_029452* promoter contains 2 $TAL13_{Xam668}$ EBEs when the score stringency is reduced to be within 4-fold the best possible score.

Lastly, because the TAL effector western blot in figure 2-5 did not convincingly show $TAL14_{Xam668}$ complementation of our $TAL14_{Xam668}$/$TAL20_{Xam668}$ double mutant (Xam668ΔTAL14/TAL20p(TAL14,TAL20)), we used semiquantitative RT-PCR to detect activation of the TAL14 target *007568*, confirming TAL14 complementation (Figure 3-12).

Presence of *TAL20* in diverse *Xam* strains correlates with *MeSWEET10a* induction and watersoaking

To determine the level of conservation among *Xam* isolates in the ability to induce *MeSWEET10a* during infection, we selected a variety of *Xam* field isolates and tested *MeSWEET10a* induction by semi-quantitative RT-PCR (Bart *et al.*, 2012). The number and approximate sizes of TAL effectors in these strains is shown in figure 2-8A. We found that *MeSWEET10a* induction is conserved in 9 of 11 strains and correlates with relatively higher levels of watersoaking (Figure 3-13). All strains capable of inducing *MeSWEET10a* expressed a TAL effector similar in size to $TAL20_{Xam668}$.

Discussion

We chose to carry out RNA-Seq to identify targets of $TAL20_{Xam668}$ because of its role in both *in planta* bacterial growth and watersoaking symptom formation. Our RNA-Seq experimental design allowed us to identify transcriptional changes induced by all five TAL effectors encoded by the Xam668 genome (within the mock vs. Xam668 data set) and also to distinguish the transcriptional changes caused specifically by the TAL effector $TAL20_{Xam668}$ (*Xe* vs. *Xe* delivering $TAL20_{Xam668}$). Despite the fact that 281 annotated cassava promoters contain potential $TAL20_{Xam668}$ EBEs, only a single transcript, *MeSWEET10a*, was induced to similar levels by both Xam668 and *Xe*($TAL20_{Xam668}$). Notably, there are predicted EBEs for $TAL13_{Xam668}$ and $TAL14_{Xam668}$ in the *MeSWEET10a* promoter in addition to the $TAL20_{Xam668}$ EBE, yet activation is dependent only on $TAL20_{Xam668}$. Mechanisms of TAL effector target specificity not mediated by RVD-base pair interaction are being elucidated. It has been shown that TAL effector dependent activation is limited by epigenetic modifications, so chromatin state is a likely factor contributing to specificity (Bultmann *et al.*, 2012). Recently, it was shown that adding constraints such as a lower relative EBE score (ratio of actual score

to the score of a hypothetical perfect match) requirement and increased stringency for distance from the transcriptional start site can reduce the incidence of falsely predicted targets (Cernadas *et al.*, 2014). Other factors contributing to specificity may include precision gained by members of the transcriptional complex or binding constraints built into TAL effectors themselves (Streubel *et al.*, 2012; Grau *et al.*, 2013).

To identify targets of $TAL13_{Xam668}$, $TAL14_{Xam668}$, $TAL15_{Xam668}$, and $TAL22_{Xam668}$, we turned to our mock vs. Xam668 RNA-Seq data set. Seventeen likely targets were selected based on expression profile and were subsequently shown to be activated by TAL effectors. Some targets lend themselves to hypotheses of function: TAL14-targeted *cassava4.1_007568*, a pectate lyase, may aid bacterial colonization of the leaf and vascular tissues. Pectate lyases have been reported to promote disease in plants: bacteria in the genus *Erwinia* secrete pectate lyase enzymes via the Type II secretion pathway to break down plant cell walls and cause symptoms of soft rot (reviewed in (Brencic & Winans, 2005)), *PMR6,* an *Arabidopsis* pectate lyase-like gene, is required for susceptibility to powdery mildew infection (Vogel *et al.*, 2002), and cell wall pectic structures were previously shown to be altered during *Xam* infection of cassava (Boher *et al.*, 1995). Interestingly, *Xe* AvrBs3 was found to induce pepper pectate lyase *upa8*, and the *Xcc* PthA-targeted *CsLOB1* transcription factor was associated with expression of pectate lyase *Cit.39387* (Marois *et al.*, 2002; Hu *et al.*, 2014). $TAL14_{Xam668}$-targeted pectate lyases are discussed further in section 4 of this work. *Cassava4.1_030094,* cassava4.1*_029949,* and *cassava4.1_027342* are annotated as transcription factors which is consistent with reports of TAL effectors affecting multiple changes in the host transcriptome by directly activating a host transcription factor (Kay *et al.*, 2007; Hu *et al.*, 2014; Li *et al.*, 2014). While $TAL20_{Xam668}$ has a single host target, $TAL14_{Xam668}$ has many which may contribute collectively to virulence. The absence of a mutant phenotype for $TAL13_{Xam668}$, $TAL15_{Xam668}$, and $TAL22_{Xam668}$ despite having *bona fide* host targets may be attributed to effector functional redundancy or to the challenge of measuring bacterial fitness in a laboratory setting. Laboratory experiments are hindered by an inability to include all potential pathogen life stages or environmental stresses, measure infection over long periods of time or include large numbers of replicates. Additionally, $TAL13_{Xam668}$, $TAL15_{Xam668}$, and $TAL22_{Xam668}$ may have susceptibility targets in host cultivars other than our experimental cultivar TMS 60444.

MeSWEET10a is a sugar transporter related to the susceptibility genes *OsSWEET11* and *OsSWEET14* in rice. *OsSWEET11* and *OsSWEET14* are activated by multiple TAL effectors from *Xoo,* and their activation promotes disease progression (Yang *et al.*, 2006; Antony *et al.*, 2010; Yu *et al.*, 2011; Streubel *et al.*, 2013). *MeSWEET10a* resides in clade III of the SWEET transporters along with *OsSWEET11* and *OsSWEET14.* Interestingly, another clade III SWEET, *cassava4.1_014124* or *MeSWEET15b,* was upregulated in our mock vs. Xam668 data set and was found to be dependent on $TAL22_{Xam668}$ (Figure 3-14). Since the $TAL22_{Xam668}$ mutant is unaffected in its ability to cause watersoaking, the function of this SWEET in disease progression remains unclear. Despite this, clade III SWEETs appear to be particularly relevant for disease promotion, as a clade I SWEET from citrus, *CsSWEET1*, was identified as an

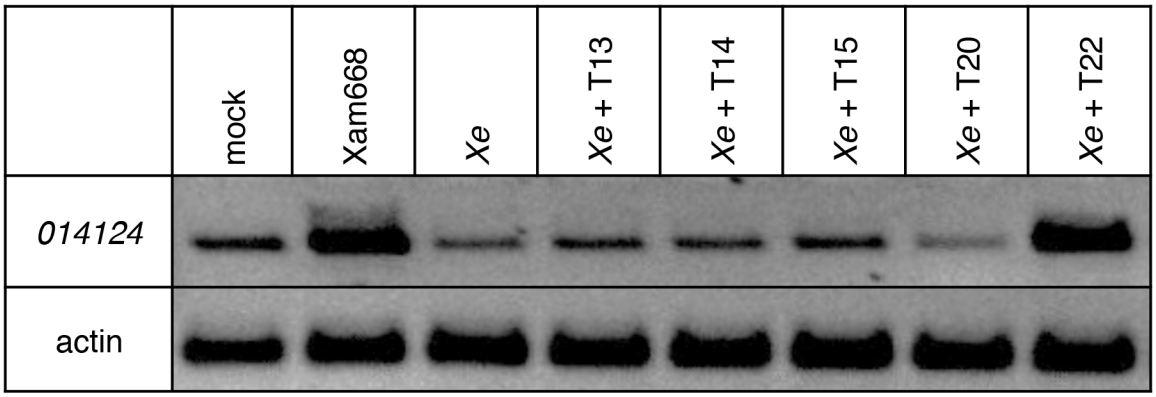


Figure 3-14. *Cassava4.1_014124 (MeSWEET15b),* a clade III SWEET transporter, is activated by TAL22$_{Xam668}$. Semiquantitative RT-PCR of leaf tissue inoculated (OD_{600} = 0.5) with *Xanthomonas euvesicatoria (Xe)* 85-10 delivering each of the 5 Xam668 transcription activator-like (TAL) effectors shows TAL22$_{Xam668}$ -dependent activation of *cassava4.1_014124*. Experiment was repeated 2 times with similar results.

Xcc TAL effector target but no role in susceptibility was found (Hu *et al.*, 2014; Hutin *et al.*, 2015). Clade III SWEETs in rice not yet shown to be TAL effector targets increased virulence of *Xoo* when specifically activated by dTALEs, while SWEETs from other clades did not (Streubel *et al.*, 2013). The cassava genome encodes 23 putative SWEET members, 13 in clade III, based on annotations available on Phytozome (Goodstein *et al.*, 2011; Prochnik *et al.*, 2012). The reason for the specific role of *MeSWEET10a* in susceptibility is unknown.

SWEETs facilitate glucose and sucrose efflux from mesophyll cells into the apoplasm, and are likely involved in phloem loading [reviewed in (Chen, 2014)]. Since clade III SWEETs were found to preferentially transport sucrose, it was hypothesized that pathogens induce expression of *SWEET* genes to increase sucrose export from cells into the apoplasm to be used as a pathogen carbon source (Chen *et al.*, 2010; 2012). We explored the proposed link between SWEET transporter induction and pathogen sucrose utilization by comparing disease phenotypes of Xam668ΔTAL20, which does not activate *MeSWEET10a*, and Xam668ΔsuxC, which cannot import sucrose. Xam668ΔsuxC did not show a significant defect in growth *in planta*, which is likely explained by its ability to utilize carbon sources in the apoplasm other than sucrose. Glucose exported by MeSWEET10a may be used directly, while exported sucrose may be broken down into hexose sugars by a host cell wall invertase enzyme (Sturm, 1999; Kocal *et al.*, 2008). This hypothesis is supported by our RNA-Seq data that show that one cassava gene annotated as an invertase is expressed and another is specifically induced during Xam668 infection (1.9 fold induction at 50 hpi). Both contain a signal peptide predicted by SignalP 4.1 (Petersen *et al.*, 2011). In contrast to the phenotype of Xam668ΔTAL20, watersoaking was largely unaffected by the mutation of

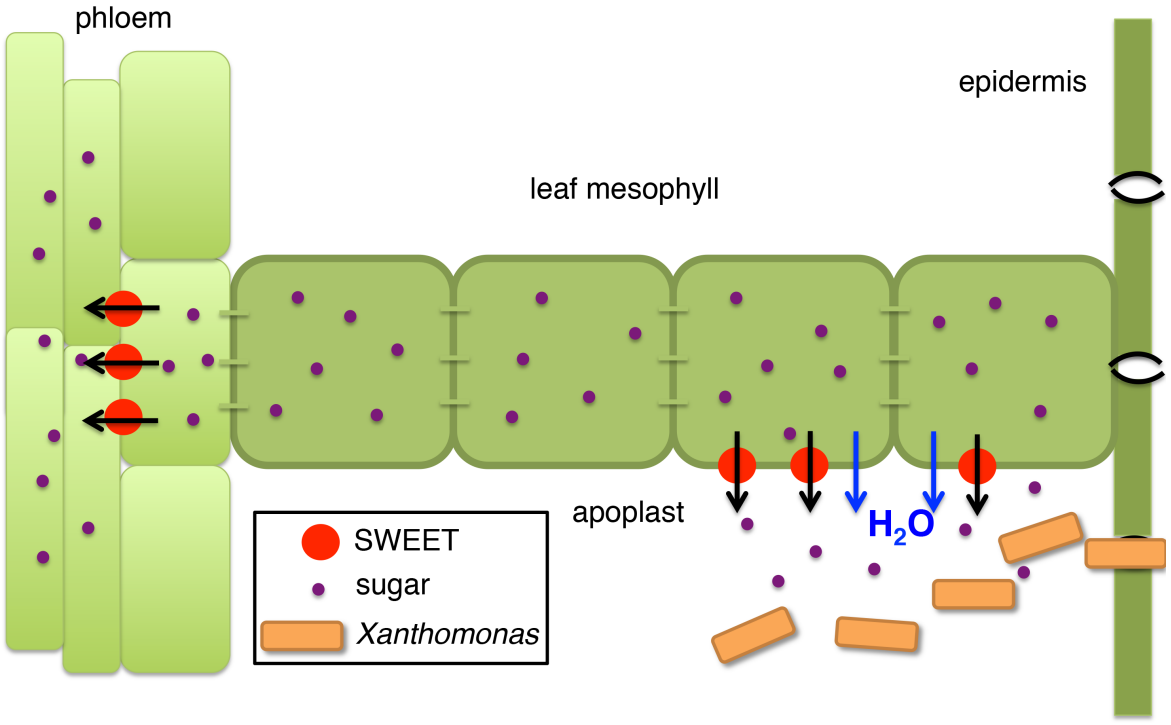


Figure 3-15. Model of SWEET sugar transporter function during pathogen infection [adapted from (Chen, 2014)]. SWEET transporters are normally expressed at the phloem parenchyma and are involved in phloem loading of sugars. Upon *Xanthomonas* infection, SWEETs are ectopically overexpressed in mesophyll cells in a transcription activator-like (TAL) effector-dependent manner. SWEET overexpression causes sugars to be exported to the apoplasm where they may be used as a pathogen carbon source. An excess of sugar in the apoplasm may also alter the osmotic gradient of the tissue such that water is pulled into the apoplasm, leading to the watersoaking that is attributed to $TAL20_{Xam668}$-dependent activation of *MeSWEET10a*.

SuxC. Taken together, our results show that *MeSWEET10a*'s contribution to virulence can be independent of *Xam*'s ability to directly utilize apoplastic sucrose.

While both the absence of $TAL14_{Xam668}$ and $TAL20_{Xam668}$ lead to a reduction in bacterial growth, only when *MeSWEET10a* is not activated (Xam668ΔTAL20) do we observe a significant reduction in watersoaking. In addition, $TAL20_{Xam668}$ increases watersoaking by TAL20-deficient strain *Xam* CIO151, but does not cause a significant increase in growth. The fact that growth can be uncoupled from watersoaking indicates that the role of symptom formation is not solely to aid in, nor is it necessarily a result of, bacterial growth. Watersoaking is caused by most xanthomonads and it is thought to contribute to bacterial dissemination (Brunings & Gabriel, 2003). Mutation of *Tal2g* in *Xoc* and *AvrB6* in *Xcm* resulted in decreased bacterial exudation from the leaf surface, a phenotype that was attributed to the decreased watersoaking conferred by the mutant strains. Notably, neither Tal2g nor AvrB6 directly promote bacterial growth (Yang *et al.*,

1994; Cernadas *et al.*, 2014). AtSWEET9, a clade III SWEET sucrose uniporter from *Arabidopsis thaliana,* was recently shown to be involved in nectar secretion. AtSWEET9 exports sucrose into the apoplasm where it is likely broken down into hexose sugars by a plant invertase, thereby maintaining a sufficient osmotic gradient for water secretion (Lin *et al.*, 2014). An intriguing possibility is that MeSWEET10a aids bacterial growth by providing an apoplastic carbon source and independently promotes water accumulation by altering the osmotic gradient of the mesophyll, leading to *Xam*'s characteristic watersoaking symptoms (Figure 3-15).

4. Comparison of gene activation by TAL14 proteins from *Xanthomonas axonopodis* pv. *manihotis* reveals parameters of virulence in cassava

Background

The transcription activator-like (TAL) effectors of *Xanthomonas* directly bind and activate host gene promoters via a central repeat domain (Boch & Bonas, 2010; Bogdanove *et al.*, 2010). The repeat variable diresidues (RVDs) of each DNA binding domain repeat mediate binding to a single consecutive nucleotide through direct interaction (Mak *et al.*, 2012; Deng *et al.*, 2012). By observing natural TAL effector/effector binding element (EBE) pairs, RVDs were found to be specific for one or a few nucleotides, resulting in the elucidation of the TAL effector-DNA binding code (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). EBE predictions based on observed RVD specificity alone result in many false positive predictions, indicating that additional nuanced parameters influence *bona fide* TAL effector-DNA interaction. Studies in recent years have found that RVD efficiency, RVD-nucleotide affinity, polarity, and the vicinity of EBEs to core promoter elements are parameters that influence TAL effector binding strength (Streubel *et al.*, 2012; Cong *et al.*, 2012; Grau *et al.*, 2013; Meckler *et al.*, 2013; Moore *et al.*, 2014). In this chapter we show that a single repeat difference affecting both RVD specificity and nucleotide affinity in two highly similar TAL14 proteins from different *Xam* strains has a dramatic effect on host target repertoires. Despite differentially activating host targets, the two TAL14 proteins complement the opposite parent strain in *in planta* bacterial growth assays and are therefore functionally equivalent. Insights into potential *Xam* virulence mechanisms based on shared gene targets are discussed. The work in this section reveals the dramatic impact of a single RVD difference on host targets in two naturally occurring TAL effectors within a single species of *Xanthomonas* and demonstrates the use of highly similar TAL effectors as tools to rapidly eliminate false-positive candidate susceptibility genes.

Results

$TAL14_{Xam668}$ is predicted to target many cassava promoters

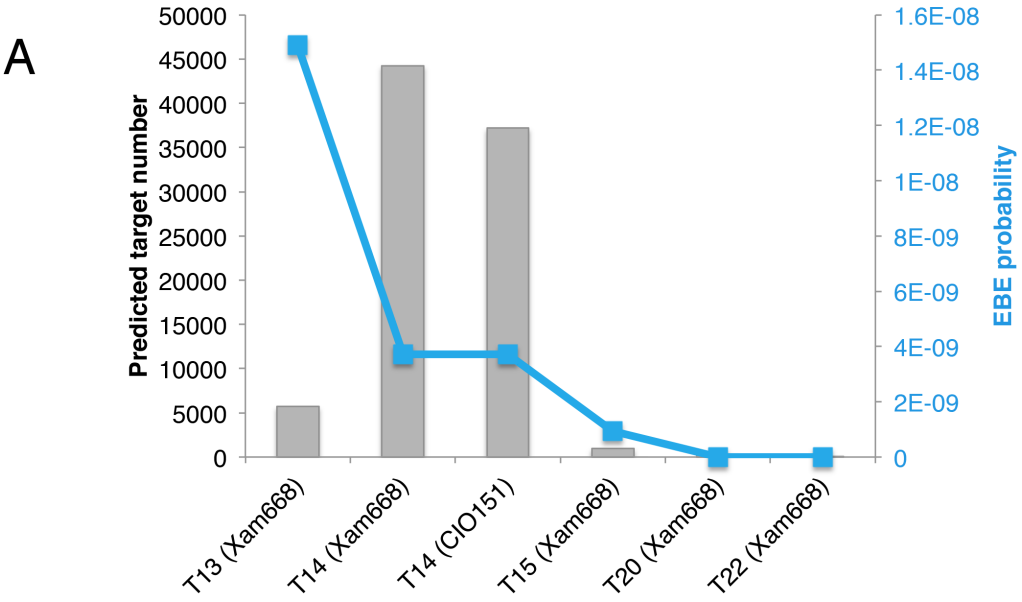
$TAL14_{Xam668}$ has a large number of predicted targets relative to the other TAL effectors of Xam668 (Cohn *et al.*, 2014). Target predictions were generated by TALE-NT (2.0) Target Finder, which identifies potential targets based on observed RVD-nucleotide association frequencies (*i.e.* RVD specificity) but does not take into account RVD efficiency or location (Doyle *et al.*, 2012). The observation that $TAL14_{Xam668}$ has a relatively high number of predicted targets held when we ran Target Finder on the 1 kb cassava promoterome with more stringent parameters than previously reported (Figure

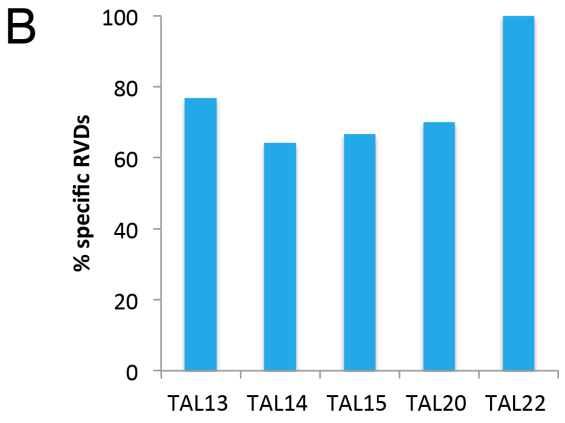
4-1A). We first hypothesized that the number of TAL14$_{Xam668}$ targets was large because its small size (low repeat number) makes it more likely to match a suitable nucleotide sequence at random than a TAL effector with more repeats. We compared the number of predicted targets for the Xam668 TAL effectors which have RVD numbers ranging from 13 to 22 and found that TAL14$_{Xam668}$ has more targets than one would expect given that it differs in length from TAL13$_{Xam668}$ and TAL15$_{Xam668}$ by only 1 RVD (Figure 4-1A).

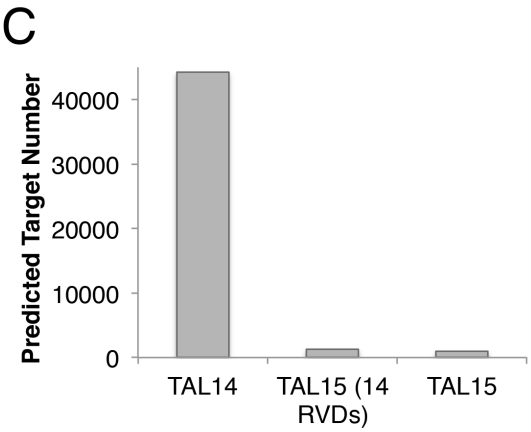
We next speculated that TAL14$_{Xam668}$ has a large number of predicted targets because it contains a higher proportion of RVDs with relaxed specificity. Of the RVDs present in Xam668 TAL effectors, HD, NG, and NI are specific with a preference for binding C, T, and A, respectively, while the RVDs NN and NS are less stringent in their nucleotide specificity (Moscou & Bogdanove, 2009). We calculated the percent of specific RVDs (NI, NG, HD) present in the RVD sequence for each TAL effector to see if TAL14$_{Xam668}$ was generally less specific. TAL14$_{Xam668}$ has the lowest percent of specific RVDs (64.3%, 9/14), however its percentage does not differ greatly from that of TAL15$_{Xam668}$ which has 66.7% (10/15) specific RVDs (Figure 4-1B). Furthermore, the first 14 RVDs of TAL15$_{Xam668}$ consist of the same percent of specific RVDs as TAL14$_{Xam668}$, yet maintain a much smaller number of predicted targets in the 1 kb cassava promoterome (Figure 4-1C). Both TAL14$_{Xam668}$ and TAL15$_{Xam668}$ begin with the RVDs NI-NG-NI which, together with a 5' T (T_0), prefer to bind the sequence TATA. Therefore, both TAL14$_{Xam668}$ and TAL15$_{Xam668}$ are likely to have predicted target EBEs that are anchored in a TATA-box sequence (TATAWA) showing that predicted binding to a core promoter motif may contribute to TAL14$_{Xam668}$'s apparent target promiscuity, but is not the only explanation (Butler & Kadonaga, 2002). We conclude that TAL14$_{Xam668}$ is predicted to bind a group of nucleotide sequences that are relatively common in the cassava promoterome compared to the groups of sequences predicted to be bound by the other TAL effectors of Xam668. The consensus sequence for the top 5,000 predicted targets of TAL14$_{Xam668}$ is displayed (Figure 4-1D).

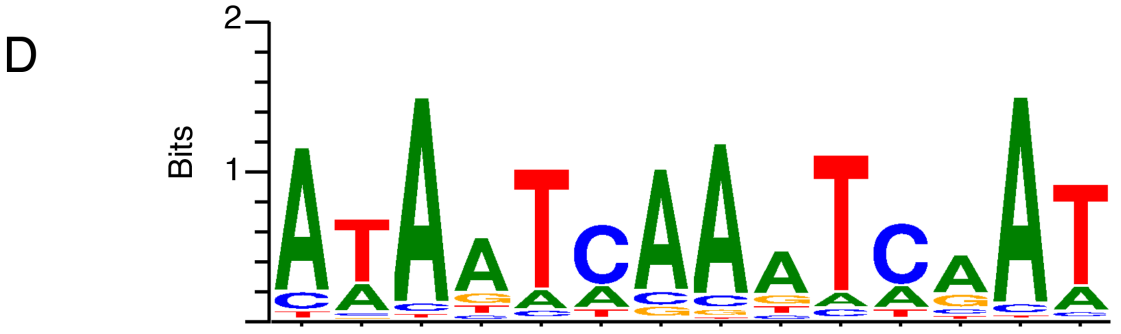
RNA-Seq identifies targets of TAL14$_{Xam668}$

Predicting TAL effector target EBEs using only RVD-nucleotide association frequencies results in many false positive predictions, making transcriptomic approaches necessary to identify true TAL effector targets (Grau *et al.*, 2013; Cohn *et al.*, 2014). We conducted an RNA-Seq experiment to identify the full repertoire of genes activated by TAL14$_{Xam668}$ because of its large number of predicted targets and also because of the virulence defect seen in the TAL14$_{Xam668}$ mutant strain (Cohn *et al.*, 2014). *Xanthomonas euvesicatoria* (*Xe*) strain 85-10, which does not cause disease on cassava and has no TAL effectors of its own, was used to deliver TAL14$_{Xam668}$ to cassava leaf tissue independently of the other TAL effectors of Xam668 (Cohn *et al.*, 2014). We compared genes that were upregulated in leaf tissue infiltrated with mock inoculation (10 mM $MgCl_2$) versus Xam668, and those that were upregulated between *Xe* and *Xe*(TAL14) at a single time point (48 hours post inoculation, hpi). The genes present in both comparisons were designated as potential targets of TAL14$_{Xam668}$









RVD position	1	2	3	4	5	6	7	8	9	10	11	12	13	14
T14 (Xam668)	NI	NG	NI	NN	NG	HD	NS	NS	NN	NG	HD	NN	NI	NG

Figure 4-1 (previous page). $TAL14_{Xam668}$ is predicted to target many cassava promoters. A. Target effector binding elements (EBEs) for the transcription activator-like (TAL) effectors of Xam668 and $TAL14_{ClO151}$ were predicted using Target Finder (TALE-NT 2.0) in the 1 kb cassava promoterome. Predicted EBEs were required to be within 3-fold of the best possible binding score, on the forward DNA strand, and preceded by a thymine (T) at position 0. Predicted target number was plotted (left y-axis, bar graph) for each TAL effector. To visualize the effect of repeat number on EBE prediction, we determined the rate at which a correct EBE sequence would occur in the genome at random, where each RVD has a ¼ chance of being aligned to its preferred nucleotide ($0.25^{RVD\#}$) (right y-axis, line graph). B. The percentage of specific repeat variable diresidues (RVDs) (NI, NG, HD) for each Xam668 TAL effector is plotted for comparison. C. EBEs for the first 14 RVDs of $TAL15_{Xam668}$ were predicted as in A and plotted alongside the number of predicted targets of $TAL14_{Xam668}$ and $TAL15_{Xam668}$ for comparison. D. The consensus predicted EBE for the top 5,000 predicted $TAL14_{Xam668}$ EBEs in the 1 kb cassava promoterome is shown along with the RVD sequence of $TAL14_{Xam668}$.

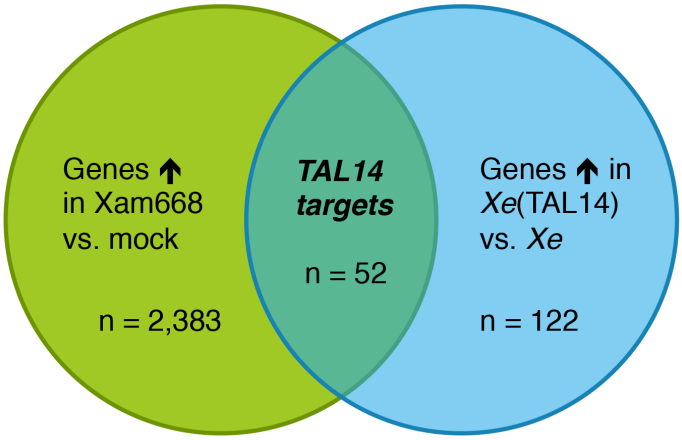


Figure 4-2. Venn diagram illustrating the RNA-Sequencing experimental approach to identify $TAL14_{Xam668}$ targeted genes. The total number of genes identified in each group is displayed [$\log_2$ (fold change) > 1; FPKM difference > 5].

(Figure 4-2). We found a total of 52 genes upregulated in a $TAL14_{Xam668}$ dependent manner ($\log_2$ fold change >1, FPKM difference > 5) (Figure 4-3). These target genes encode a range of proteins, including pectate lyases and other cell wall modifying enzymes, EamA-like/MtN21 transporters, several proteases, glyceraldehyde-3-phosphate dehydrogenases, LRR kinases, and proteins of unknown function.

gene name*	FPKM (48 hpi)				log$_2$ fold change		Gene Annotations (Phytozome)	group
	mock	Xam668	*Xe*	*Xe*(T14)	m v. 668	*X* v *X*(14)		
022805	0.6	85.6	0.3	83.9	8.2	9.9	Arabidopsis protein of unknown function	2
007568	3.3	818.0	0.7	632.1	7.9	9.9	pectate lyase	1
026121	0.4	290.3	0.4	382.0	9.5	9.8	no annotated domains	2
024404	0.4	113.2	0.4	328.7	8.3	9.8	serine protease	1
020743	0.2	185.5	0.5	369.4	9.5	9.5	no annotated domains	2
007516	96.7	1835.7	2.3	912.4	4.2	8.6	pectate lyase	1
033289	9.4	67.4	1.2	422.6	2.8	8.5	subtilase serine protease	1
031361	2.0	212.5	4.3	852.1	6.8	7.6	no annotated domains	2
023665	0.0	14.6	0.2	41.6	inf	7.6	acyl transferase	2
001042	0.5	31.7	0.2	35.3	6.1	7.4	glycosyl transferase, starch synthase	2
034150	18.3	5031.6	29.4	4483.9	8.1	7.3	mannose-binding lectin	1
022871	0.2	85.5	1.2	104.3	8.5	6.5	oxidoreductase	2
011345	12.3	469.2	5.4	442.0	5.3	6.3	glyceraldehyde 3-phosphate dehydrogenase	1
012090	2.6	27.6	0.9	59.9	3.4	6.1	oxidoreductase	2
022534	1.5	11.9	0.2	10.2	2.9	5.8	EamA-like transporter	2
016646	0.8	29.0	2.3	116.9	5.1	5.6	no annotated domains	2
020499	1.2	128.4	0.6	23.8	6.7	5.2	metallothionein	2
026299	1.3	27.3	1.1	38.9	4.4	5.1	no annotated domains	1
019005	29.9	558.2	14.4	463.4	4.2	5.0	no annotated domains	1
025591	0.3	120.4	19.0	571.4	8.9	4.9	LEA group 1	2
015102	7.7	101.2	2.2	49.4	3.7	4.5	vesicle associated membrane associated protein	2
009347	17.4	1789.2	74.5	1350.7	6.7	4.2	EamA-like transporter	1
026646	0.1	15.2	1.7	26.1	7.1	3.9	aspartyl protease	2
024542	4.4	59.4	5.3	64.0	3.7	3.6	putative serine esterase	2
011524	17.6	197.7	25.5	291.0	3.5	3.5	clathrin light chain, vesicle mediated transport	1
023036	0.3	55.5	14.2	159.3	7.6	3.5	mannose-binding lectin	2
021136	1.8	12.6	2.5	26.7	2.8	3.4	no annotated domains	NT
028527	2.6	23.2	1.6	15.9	3.2	3.3	cytochrome P450	NT
025868	8.8	93.6	13.0	112.7	3.4	3.1	no annotated domains	NT
007819	7.9	79.9	14.6	110.1	3.3	2.9	vitamin B6 photo-protection and homeostasis	NT
026095	7.8	84.0	13.1	94.9	3.4	2.9	no annotated domains	NT
011062	10.4	122.0	20.3	139.1	3.5	2.8	WRKY TF	NT
008375	94.4	386.7	146.8	997.3	2.0	2.8	thioredoxin	NT
031238	0.3	17.9	12.5	83.7	5.8	2.7	xylose isomerase	NT
020441	22.8	908.4	179.9	901.2	5.3	2.3	no annotated domains	NT
003505	61.2	331.8	61.2	285.5	2.4	2.2	alcohol dehydrogenase	NT
027076	5.6	37.3	2.2	8.7	2.7	1.9	EamA-like transporter	NT
007307	17.7	45.8	5.2	19.9	1.4	1.9	vacuolar iron transporter	NT
033504	4.0	10.6	4.2	14.1	1.4	1.7	PBS lyase HEAT-like repeat	NT
030745	13.8	511.8	229.1	675.2	5.2	1.6	LRR kinase	NT
025744	0.0	14.2	7.2	19.9	9.3	1.5	LRR kinase	NT
022442	1.3	21.0	6.2	17.2	4.0	1.5	auxin-responsive protein	NT
000922	7.8	22.5	16.4	45.2	1.5	1.5	PB1 domain, RWP-RK domain (NLP7)	NT
019559	219.1	465.1	119.8	326.6	1.1	1.4	ATLS1, macrophage inhibitory factor related	NT
022935	2.4	8.7	6.9	18.2	1.9	1.4	Rare lipoprotein A (RlpA)-like double-psi beta-barrel	NT
029833	0.3	17.7	5.5	14.3	5.9	1.4	BCS1 AAA-type ATPase	NT
027926	84.7	273.1	340.1	809.3	1.7	1.3	amino acid transporter	NT
005942	2.1	43.9	8.1	18.7	4.4	1.2	serine protease	NT
003660	14.4	229.0	30.2	68.9	4.0	1.2	LRR kinase	NT
014976	46.1	235.3	60.9	138.5	2.4	1.2	heat shock transcription factor	NT
011175	87.9	182.5	81.1	169.7	1.1	1.1	Glyceraldehyde 3-phosphate dehydrogenase	NT
002412	72.5	149.0	98.4	200.5	1.0	1.0	X-BOX transcription factor-related, cellulose synthase	NT

Figure 4-3 (previous page). RNA-Sequencing reveals 52 genes activated in a $TAL14_{Xam668}$-dependent manner. Genes activated by $TAL14_{Xam668}$ are defined as showing log_2(fold change) > 1 and FPKM difference > 5 in both the mock versus Xam668 and *Xe* versus *Xe*(T14) comparisons. For each cassava gene, FPKM values for mock, Xam668, *Xe*, and *Xe*(T14) are listed and color coded by value, along with the log_2(fold change) between mock and Xam668 (m v 668), and between *Xe* and *Xe*(T14) [*X* v *X*(14)]. Genes are listed by decreasing values for the latter comparison. Phytozome gene annotations are listed, along with the assigned promoter group (NT, not tested). * Only gene numbers are shown, full gene names are *cassava4.1_xxxxxx.*

We validated the $TAL14_{Xam668}$-dependent activation of the 26 most highly activated genes through semiquantitative RT-PCR of cassava leaf tissue inoculated with Xam668 and a strain that is missing the plasmid segment (ps) encoding $TAL14_{Xam668}$ (Xam668ΔTAL14ps) (Figure 2-7A, 4-4A). Xam668ΔTAL14ps is a markerless strain that shows the same bacterial growth defect as the previously studied $TAL14_{Xam668}$ knockout strain (Xam668ΔTAL14), and is fully complemented by the introduction of wild type *$TAL14_{Xam668}$* on a plasmid (Figure 4-5A) (Cohn *et al.*, 2014). Twenty-five of the 26 tested genes are upregulated by $TAL14_{Xam668}$ alone. Since the activation of *cassava4.1_009347* was induced by Xam668 but also by Xam668ΔTAL14ps, we tested for activation by the other Xam668 TAL effectors and found that both $TAL14_{Xam668}$ and $TAL22_{Xam668}$ activate this gene (Figure 4-4B).

$TAL14_{ClO151}$ from *Xam* strain ClO151 activates a subset of $TAL14_{Xam668}$ targets

Xam strain ClO151 has two TAL effectors, $TAL14_{ClO151}$ and $TAL21_{ClO151}$ (Bart *et al.*, 2012; Cohn *et al.*, 2014). Growth assays of ClO151 and a $TAL14_{ClO151}$ knockout strain (ClO151ΔTAL14) show that $TAL14_{ClO151}$ promotes *in planta* bacterial growth (Figure 4-5B, 4-6). $TAL14_{Xam668}$ and $TAL14_{ClO151}$ have identical amino acid sequences except for 3 amino acid differences in the fifth DNA binding domain repeat (Figure 4-7A) (Bart *et al.*, 2012). The fifth RVD of $TAL14_{Xam668}$ is NG which preferentially interacts with T, while the fifth RVD of $TAL14_{ClO151}$ is NI which preferentially interacts with A. Similar to what is seen for $TAL14_{Xam668}$, $TAL14_{ClO151}$ has a unexpectedly high number of predicted EBEs in the 1 kb cassava promoterome (Figure 4-1A). Using semiquantitative RT-PCR we tested the ability of $TAL14_{ClO151}$ to activate the top targets of $TAL14_{Xam668}$. Given the similarity between the two TAL14 proteins, we were surprised to find that $TAL14_{ClO151}$ is only able to activate 10 of the 26 top $TAL14_{Xam668}$ targets (38%) (Figure 4-7B). We will subsequently refer to the group of genes activated by both $TAL14_{Xam668}$ and $TAL14_{ClO151}$ as group 1, and to the group of genes activated by only $TAL14_{Xam668}$ as group 2. Gene targets were tested for activation in the presence of 50 μM cycloheximide (CHX), an inhibitor of eukaryotic protein synthesis, to determine if they were direct targets of the TAL14 proteins (Figure 4-8). All targets are directly activated except the group 2 gene

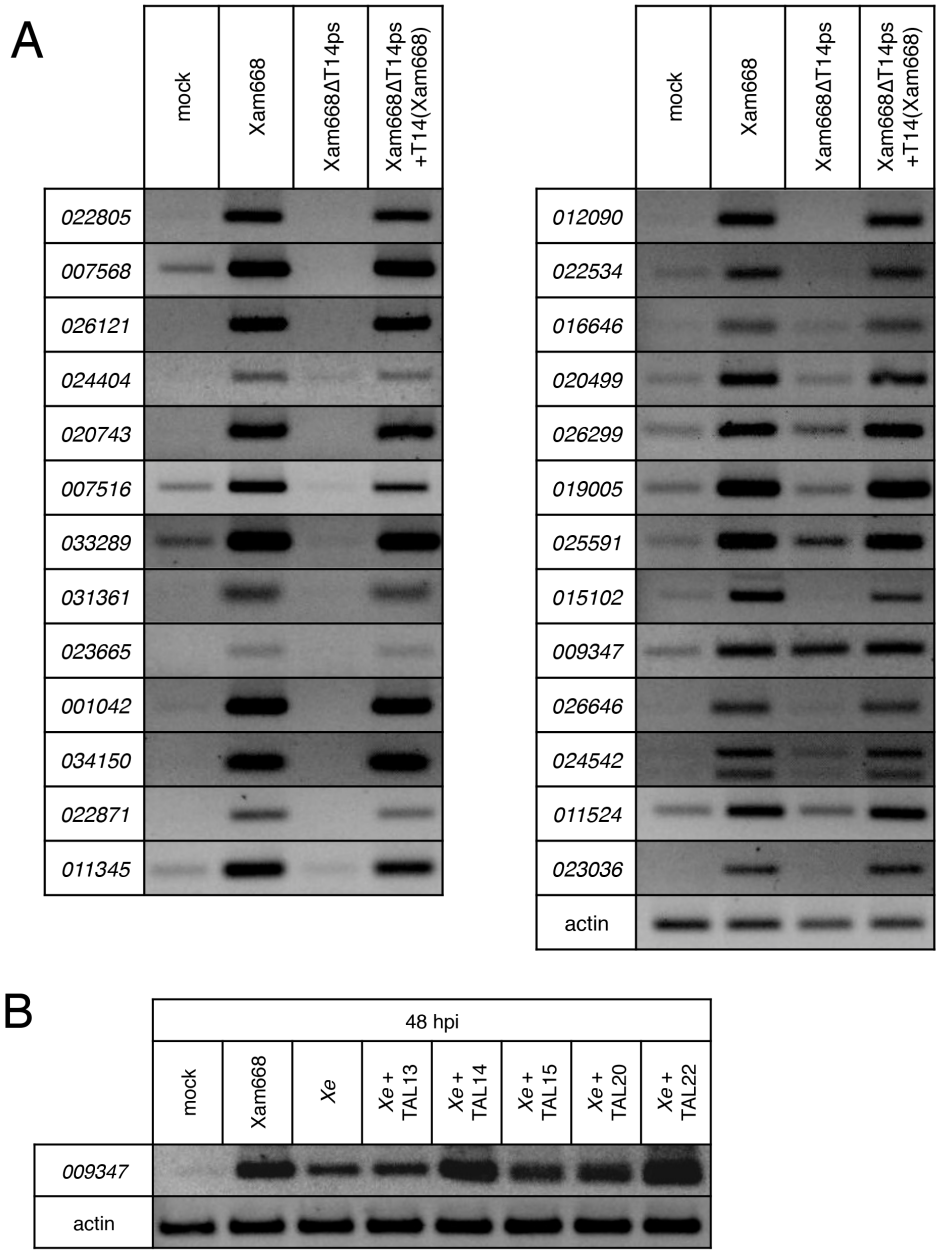


Figure 4-4. Confirmation of TAL14$_{Xam668}$-dependent target activation by semiquantitative RT-PCR. A. The 26 most highly activated TAL14 targets discovered by RNA-Sequencing were confirmed by semiquantitative RT-PCR. Gene expression was measured 48 hours post inoculation (hpi) in leaf tissue inoculated with mock infiltration (10mM $MgCl_2$), Xam668, Xam668ΔT14ps, and Xam668ΔT14ps(T14) strains at OD_{600} 0.5. B. Semiquantitative RT-PCR of *cassava4.1_009347* expression 48 hpi in leaf tissue inoculated at an OD_{600} of 0.5 with *Xe* 85-10 delivering each of the 5 Xam668 TAL effectors. The PCR cycle number for each gene primer set is listed in Supplementary Table 2.

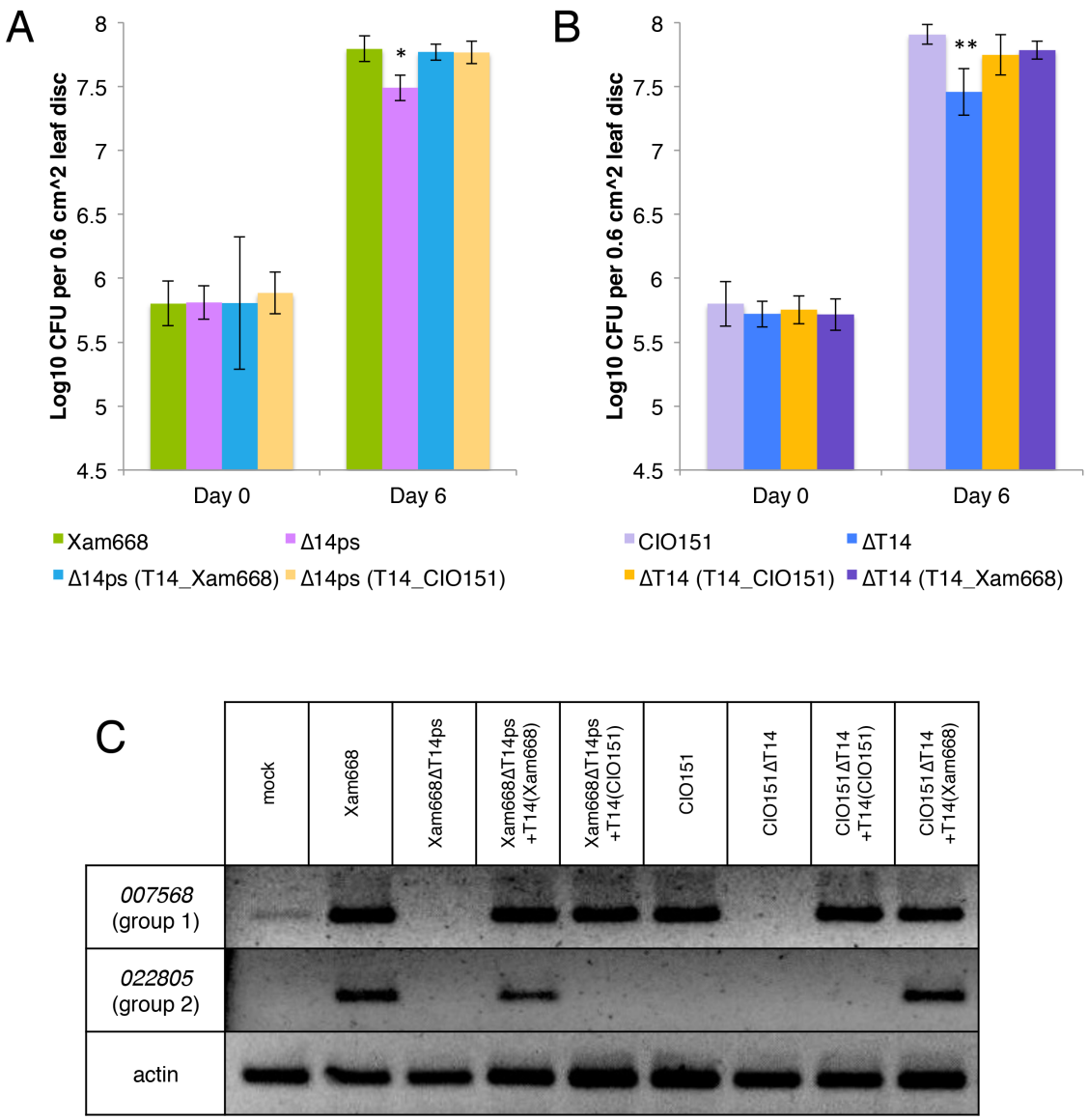


Figure 4-5. TAL14$_{ClO151}$ complements Xam668ΔT14ps and TAL14$_{Xam668}$ complements ClO151ΔT14 in bacterial growth assays. A, B. Bacterial populations at leaf midvein inoculation points were measured at days 0 and 6. Data are represented as mean colony forming units (CFU) per 0.6 cm^2 leaf disc encompassing the inoculation point (+/- SD). * denotes significantly lower growth than Xam668, 2-tailed t-test, p value < 0.001. ** denotes significantly lower growth than ClO151, 2-tailed t-test, p value < 0.005. ClO151ΔTAL14(TAL14$_{Xam668}$) and ClO151ΔTAL14(TAL14$_{ClO151}$) grow to higher levels than ClO151ΔTAL14 (p-values = 0.054 and 0.016, respectively). Growth assays were repeated 4 times with similar results. C. Semiquantitative RT-PCRs of a group 1 gene (*cassava4.1_007568*) and a group 2 gene (*cassava4.1_022805*) are displayed to show that the TAL14 proteins retain their gene activation specificities in a non-native context.

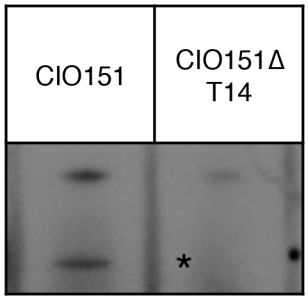


Figure 4-6. ClO151ΔTAL14 does not express TAL14$_{ClO151}$. Western blot analysis shows the absence of TAL14$_{ClO151}$ protein (asterisk) in *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strain ClO151ΔTAL14.

cassava4.1_022534. Gene groups 1 and 2 were not distinguished by common function: Group 1 contains two pectate lyases, two proteases, and a mannose-binding lectin among others. Group 2 contains an acyl transferase, a glycosyl transferase, 2 EamA-like transporters, and multiple proteins with no annotated domains among others.

To gain more information as to why TAL14$_{ClO151}$ is able to activate some of the targets of TAL14$_{Xam668}$ and not others, we decided to compare the consensus of the predicted EBEs for TAL14$_{Xam668}$ and TAL14$_{ClO151}$ in the group 1 and group 2 directly-targeted promoters using two EBE prediction programs: Target Finder (TALE-NT 2.0) and TALgetter (Doyle *et al.*, 2012; Grau *et al.*, 2013). Target Finder takes into account only RVD-nucleotide association frequencies (*i.e.* binding specificity) while TALgetter takes into account RVD binding specificity, the positive contribution of binding efficiency (*i.e.* the strength of a matching RVD-nucleotide interaction), as well as the negative penalty for non-matching RVD-nucleotide pairs. Group 1 and group 2 gene promoters were defined as 300 bp upstream of the start codon, a region which has been shown to be enriched in TAL effector target sites (Grau *et al.*, 2013).

Both the Target Finder and TALgetter analyses showed a greater occurrence of an A at position 5 in the group 1 consensus TAL14$_{ClO151}$ EBE than in the group 2 consensus TAL14$_{ClO151}$ EBE. This was in contrast to position 5 of the consensus TAL14$_{Xam668}$ EBE for the group 1 and group 2 promoters, which showed a roughly equal occurrence of an A or T at this position in both promoter groups (Figure 4-9, S1). There was also a better consensus for a canonical TATA-box (TATAWA) in the group 1 consensus TAL14$_{ClO151}$ EBE versus in group 2 (Butler & Kadonaga, 2002) (Figure 4-9, S1). We compared average normalized EBE scores for the TAL14$_{Xam668}$ and TAL14$_{ClO151}$ Target Finder-predicted EBEs in the two promoter groups. While the average normalized EBE scores were the same for TAL14$_{Xam668}$ and TAL14$_{ClO151}$ in the group 2 promoters (3.1), the average normalized score was lower (*i.e.* better) for the TAL14$_{ClO151}$ EBEs in the group 1 promoters (2.8) versus the group 1 promoters (3.0) indicating a stricter requirement for ideal alignment in promoters which are activated by

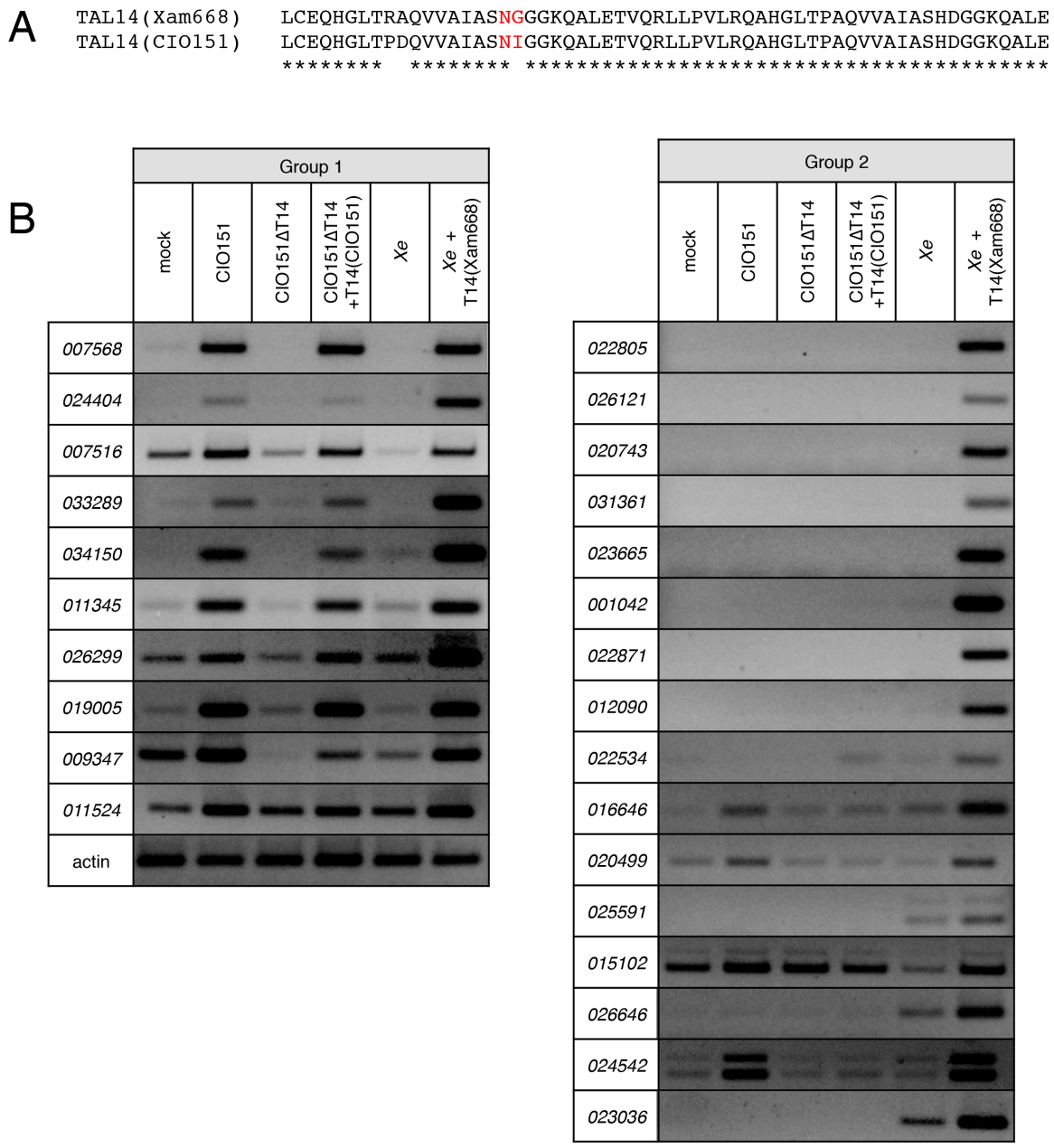


Figure 4-7. TAL14$_{Xam668}$ and TAL14$_{CIO151}$ differ by one repeat variable diresidue (RVD) and differentially activate host target genes. A. Amino acid alignment of the 5th DNA binding domain repeat of TAL14$_{Xam668}$ and TAL14$_{CIO151}$ showing three amino acid differences. Repeat variable diresidues (RVDs) are colored red. B. Semiquantitative RT-PCR on leaf tissue inoculated with mock inoculation (10mM $MgCl_2$), *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strain CIO151, CIO151ΔTAL14, CIO151ΔTAL14(TAL14$_{CIO151}$), *X. euvesicatoria* (*Xe*), and *Xe*(TAL14$_{Xam668}$) at OD_{600} 0.5 and collected 48 hours post inoculation (hpi) shows activation of genes by both TAL14$_{Xam668}$ and TAL14$_{CIO151}$ (Group 1 genes), or by TAL14$_{Xam668}$ only (Group 2 genes). The PCR cycle number for each gene primer set is listed in Supplementary Table 2.

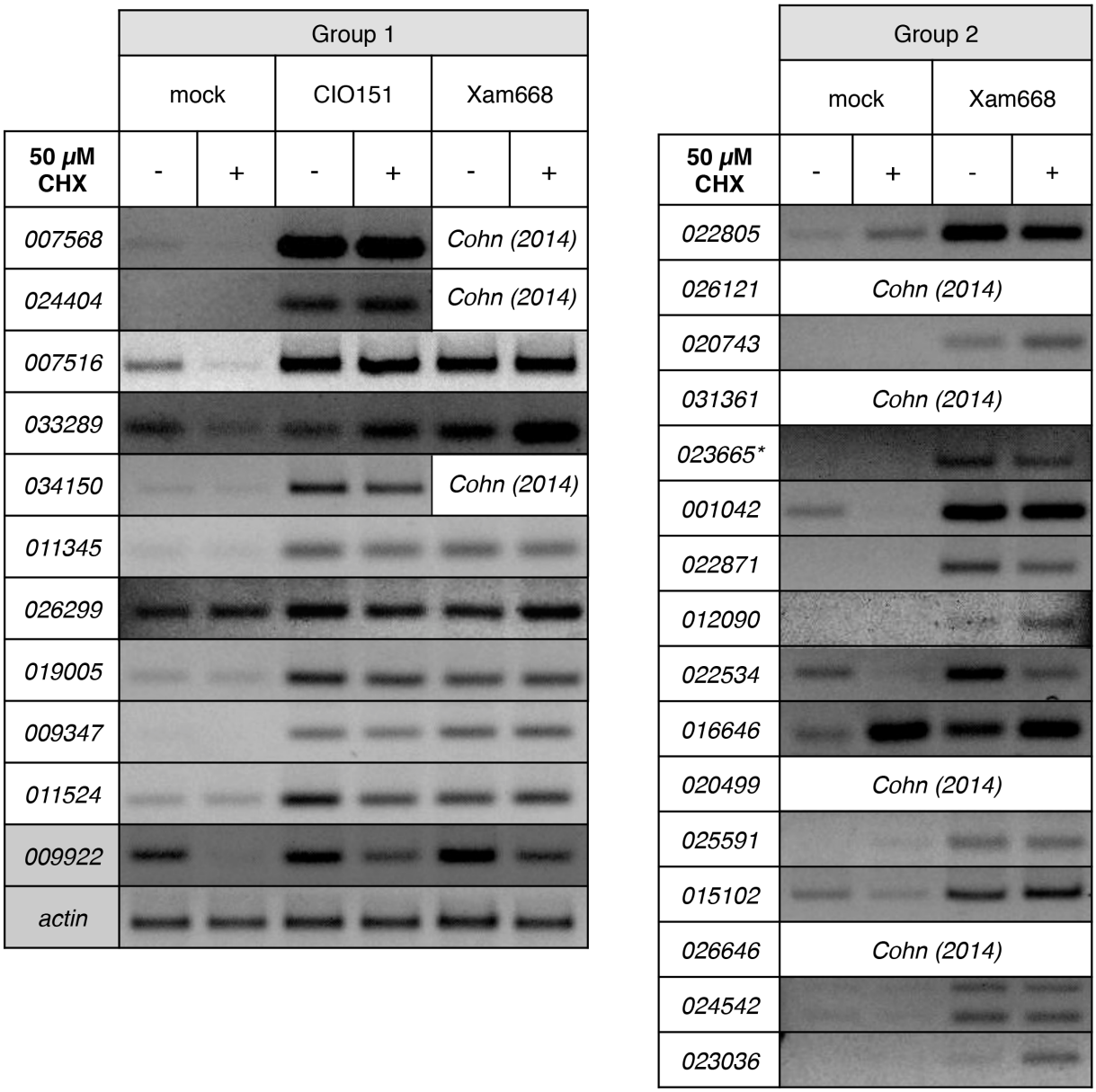
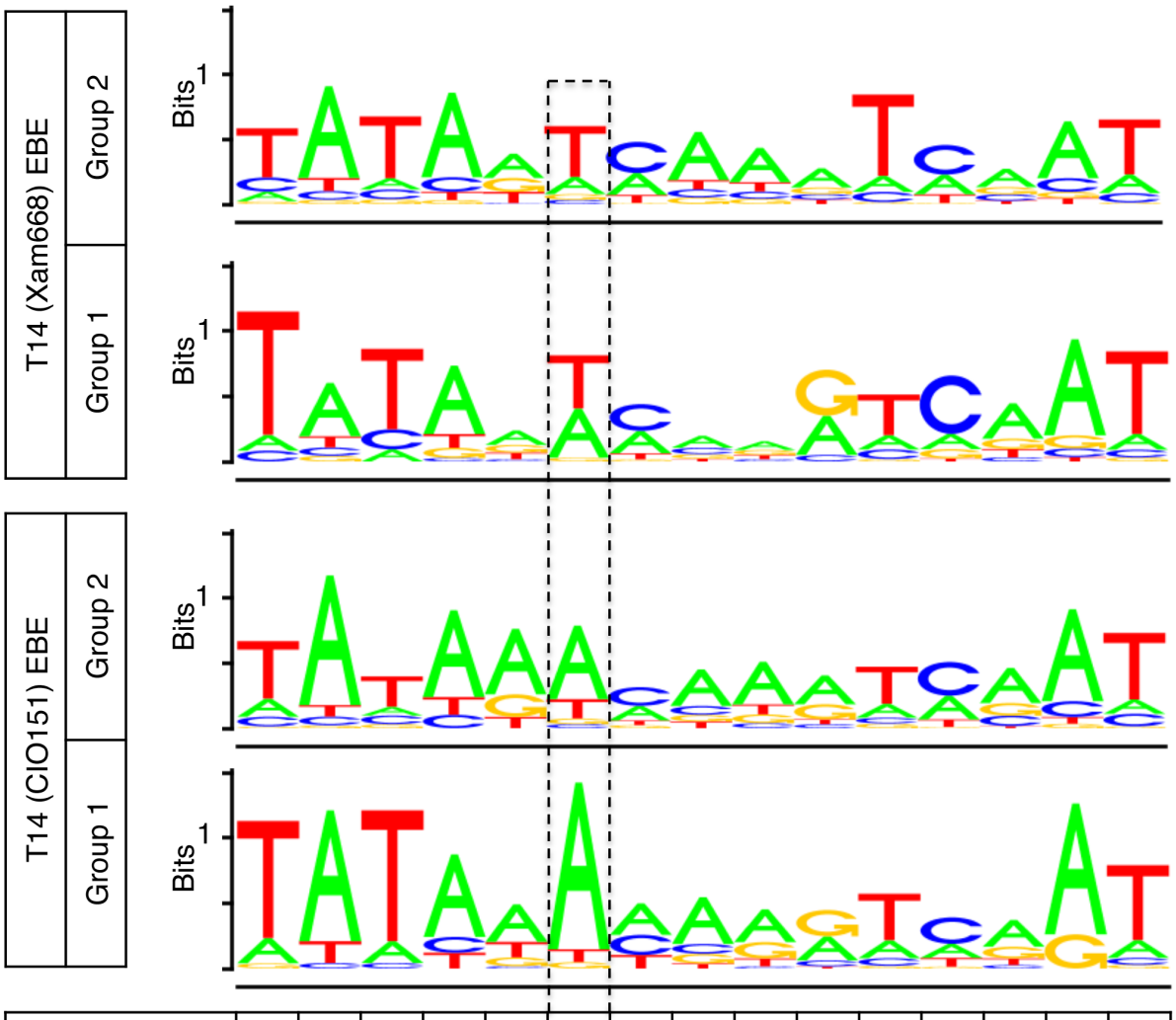


Figure 4-8. Group 1 and Group 2 genes are directly activated by *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) TAL14 proteins. *Xam* strains Xam668 and ClO151 and mock (10 mM $MgCl_2$) were inoculated into cassava leaves in the presence of 50 μM cycloheximide (CHX) and harvested 24 hours post inoculation (hpi). Genes activated above background levels in the presence of CHX are direct targets of their corresponding TAL14 proteins. The PCR cycle number for each gene primer set is listed in Supplementary Table 1. * 30 PCR cycles were required for visualization of *cassava4.1_023665* transcripts at 24 hpi. Actin expression is shown for all samples as a loading control, and expression of *cassava4.1_009922* is shown as a control for the effect of CHX on secondary transcript accumulation. Genes that were previously shown to be direct targets of $TAL14_{Xam668}$ are referenced (Cohn *et al.*, 2014).



RVD position	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
T14 (Xam668)	-	NI	NG	NI	NN	NG	HD	NS	NS	NN	NG	HD	NN	NI	NG
T14 (ClO151)	-	NI	NG	NI	NN	NI	HD	NS	NS	NN	NG	HD	NN	NI	NG

Figure 4-9. Consensus predicted TAL14$_{ClO151}$ effector binding elements (EBEs) in group 1 promoters show a strict requirement for an A at position 5. TALgetter (1.0) was used to predict EBEs for TAL14$_{Xam668}$ and TAL14$_{ClO151}$ in group 1 and 2 promoters (Grau *et al.*, 2013). The consensus EBE for each group is displayed with position 5 outlined by dashed lines. The height of the consensus indicates the nucleotide conservation at that site (measured in bits) while the height of the nucleotides within the consensus reflect their relative frequency. TAL14$_{Xam668}$ and TAL14$_{ClO151}$ repeat variable diresidue (RVD) sequences are displayed.

$TAL14_{ClO151}$. The lower average score for $TAL14_{ClO151}$ EBEs in group 2 promoters is consistent with the observation that lower scores are indicative of true targets versus false positive target predictions (Cernadas *et al.*, 2014).

$TAL14_{ClO151}$ complements Xam668ΔTAL14ps in growth assays

Designer TAL effectors (dTALEs) that directly activate *Xanthomonas* TAL effector targets have been used to complement TAL effector mutant phenotypes and attribute the promotion of disease susceptibility to activation of specific genes (Morbitzer *et al.*, 2010; Cernadas *et al.*, 2014; Hu *et al.*, 2014; Li *et al.*, 2014; Cohn *et al.*, 2014). dTALEs were made for 8 of the highly upregulated previously published $TAL14_{Xam668}$ targets and conjugated into Xam668ΔTAL14ps (Cohn *et al.*, 2014). Of the dTALE targets, 4 are in group 1 (*cassava4.1_007568, cassava4.1_007516, cassava4.1_034150*, and *cassava4.1_024404*), and 4 are in group 2 (*cassava4.1_031361, cassava4.1_026121, cassava4.1_026646*, and *cassava4.1_020499*). Xam668ΔTAL14ps(dTALE) strains were then tested for target gene activation and complementation of the $TAL14_{Xam668}$ mutant growth defect (Figure 4-10). No single tested target or target class promoted growth to the same level as complementation with $TAL14_{Xam668}$. Given the level of variability inherent in growth assays, we were not able to detect any statistically significant partial complementation by the tested dTALEs. When RNA-Seq revealed the large number of genes activated by $TAL14_{Xam668}$ the dTALE approach became unfeasible in the timeframe of our study and we turned to $TAL14_{ClO151}$ as a tool to narrow down candidate *S* genes that might be important for susceptibility to *Xam*.

$TAL14_{ClO151}$ was conjugated into Xam668ΔTAL14ps on a plasmid and the resulting strain was tested for complementation of the mutant bacterial growth defect. $TAL14_{ClO151}$ promoted growth to wild type levels in Xam668ΔTAL14ps while maintaining its target specificity (Figure 4-5A, C) indicating that while $TAL14_{ClO151}$ activates only a subset of $TAL14_{Xam668}$-activated genes, the two TAL effectors are functionally interchangeable. We found that $TAL14_{Xam668}$ also complemented the growth defect seen in ClO151ΔTAL14 in a manner similar to $TAL14_{ClO151}$ (Figure 4-5B). This result indicates that one or more of the group1 genes are the biologically relevant $TAL14_{Xam668}$-targeted susceptibility genes.

Discussion

$TAL14_{Xam668}$ of highly virulent *Xam* strain Xam668 promotes bacterial growth in the host plant cassava (Cohn *et al.*, 2014). In this study we used RNA-Seq to identify the full repertoire of genes activated by $TAL14_{Xam668}$ in the cassava genome. We tested the 26 most highly upregulated $TAL14_{Xam668}$ targets for activation by $TAL14_{ClO151}$ from *Xam* strain ClO151, whose RVD sequence differs from $TAL14_{Xam668}$ at the 5^{th} repeat in the DNA binding domain. The 5^{th} RVD of $TAL14_{Xam668}$ is NG, while the 5^{th} RVD of

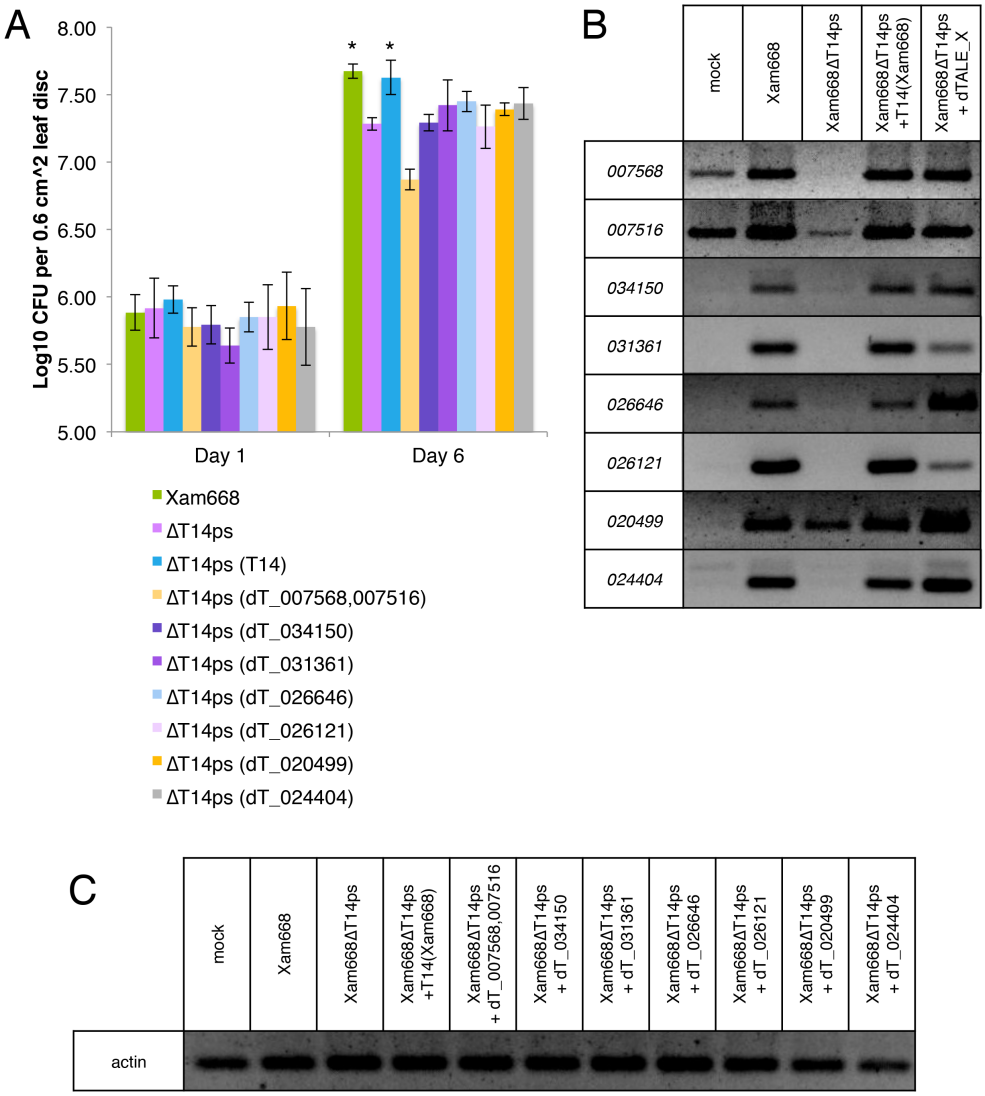


Figure 4-10. Designer TAL effectors (dTALEs) do not complement the TAL14$_{Xam668}$ mutant growth defect. A. dTALEs activating TAL14$_{Xam668}$ target genes *cassava4.1_007568/007516, cassava4.1_034150, cassava4.1_031361, cassava4.1_026646, cassava4.1_026121, cassava4.1_020499,* and *cassava4.1_024404* were conjugated into Xam668ΔTAL14ps and tested for their ability to complement the growth defect seen in the TAL14$_{Xam668}$ mutant strain. Bacterial populations at leaf midvein inoculation points were measured at days 1 and 6. Data are represented as mean colony forming units (CFU) per 0.6 cm^2 leaf disc encompassing the inoculation point (+/- SD). * denotes significantly higher growth than Xam668ΔTAL14ps, 2-tailed t-test, p value < 0.005. Growth assays were repeated 2 times with similar results. B. Semiquantitative RT-PCR shows that the dTALEs activate their intended target genes. For each gene listed in the left column, Xam668ΔT14ps + dTALE_X is the strain carrying the dTALE that activates that gene. C. Actin expression is shown for each dTALE strain.

$TAL14_{ClO151}$ is NI. We found that $TAL14_{ClO151}$ only activates a subset of the $TAL14_{Xam668}$ targets due to the stricter binding requirement of NI to A, yet is able to complement the *in planta* growth defect of a Xam668 strain lacking $TAL14_{Xam668}$, implying that shared targets of the two TAL14 variants are responsible for $TAL14_{Xam668}$'s contribution to virulence.

$TAL14_{Xam668}$ and $TAL14_{ClO151}$ are predicted to bind a large number of EBEs in the cassava promoterome when those EBEs are predicted based on the original TAL effector DNA binding code which assigns binding scores based on RVD specificity (Moscou & Bogdanove, 2009; Doyle *et al.*, 2012). The disproportionately large number of predicted targets for $TAL14_{Xam668}$ and $TAL14_{ClO151}$ cannot be explained by DNA binding domain repeat number or non-specific RVD content alone. By comparing EBE predictions among the Xam668 TAL effectors, we believe that the best explanation for the large number of predicted targets is that the TAL14 proteins are predicted to bind a group of nucleotide sequences that are relatively common in the cassava promoterome compared to the groups of sequences predicted to be bound by the other TAL effectors of Xam668. The consensus of the top 5,000 predicted $TAL14_{Xam668}$ targets reveals that $TAL14_{Xam668}$ is predicted to target AT-rich promoter sequences, and that the motif TATA-T-AA-T- - AT may be a relatively common sequence in cassava promoters, where dashes indicate variable sites.

The original TAL effector DNA binding code revealed a level of degeneracy in RVD-nucleotide specificity (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). Given this ability for RVDs to occasionally bind imperfectly to mismatched nucleotides, it was surprising that $TAL14_{ClO151}$ only activates 10 of the 26 tested targets of $TAL14_{Xam668}$ despite the two proteins differing by only a single repeat. However, this result is supported by what studies in recent years have elucidated regarding the additional parameters that influence TAL effector binding which include RVD efficiency and nucleotide affinity, polarity effects, the distance of the EBE to the transcriptional start site, and the vicinity of EBEs to core promoter elements such as the TATA-box (Cong *et al.*, 2012; Grau *et al.*, 2013; Meckler *et al.*, 2013; Streubel *et al.*, 2013; Cernadas *et al.*, 2014; Moore *et al.*, 2014; Pereira *et al.*, 2014). We compared the consensus EBEs for $TAL14_{Xam668}$ and $TAL14_{ClO151}$ in the promoters of group 1 genes (which are activated by both $TAL14_{Xam668}$ and $TAL14_{ClO15}$) and group 2 genes (which are only activated by $TAL14_{Xam668}$). While the 5^{th} RVD of $TAL14_{Xam668}$ (NG) appears to tolerate binding to both T and A, the 5^{th} RVD of $TAL14_{ClO151}$ (NI) has a stricter requirement for binding A in order to activate gene expression. The RVDs NG and NI form relatively weak van der Waals interactions with T and A, respectively, while HD and NN form strong hydrogen bonds with C and A/G, respectively (Mak *et al.*, 2012; Deng *et al.*, 2012). The nature of these RVD-nucleotide interactions is consistent with the observation that RVD efficiency varies, with NI and NG classified as weak RVDs, and HD and NN classified as strong RVDs (Streubel *et al.*, 2012). While both NG and NI are considered weak RVDs in terms of their binding efficiency, they are still highly specific with their preference for T and A, respectively. However, their relative affinity for their corresponding nucleotides is different, with NI having an affinity for A that is about 3x the strength of the affinity of NG for T. Furthermore, the affinity of NI for A is 22x its affinity for T, while the affinity of NG

for T is only about 2.5x its affinity for A (Cong *et al.*, 2012; Moore *et al.*, 2014). $TAL14_{ClO151}$, therefore, has an RVD at position 5 with a stronger preference for its corresponding nucleotide than does $TAL14_{Xam668}$, and, consistent with our findings, is less likely to functionally bind mismatched nucleotides.

The RVD difference between $TAL14_{Xam668}$ and $TAL14_{ClO151}$ occurs toward the N terminus of the DNA binding domain (RVD 5 of 14), where mismatches are less likely to be tolerated (Römer *et al.*, 2010; Meckler *et al.*, 2013). In addition, $TAL14_{Xam668}$ and $TAL14_{ClO151}$ both appear to prefer binding to a TATA-rich sequence, with $TAL14_{ClO151}$ showing a strong requirement for binding the canonical TATA-box (TATAWA) if gene activation is to be achieved. Because a preference for binding a TATA-box has been seen for other TAL effectors as well, it is widely speculated that EBE overlap with this core promoter motif has functional significance (Boch *et al.*, 2009; Römer *et al.*, 2010; Antony *et al.*, 2010; Grau *et al.*, 2013; Pereira *et al.*, 2014). Several non-mutually exclusive reasons why EBEs may overlap with the TATA-box motif have been proposed. These include a requirement to bind regions of open chromatin, to cooperate with host factors for transcription initiation, to create a functional gene product by having the correct translational start site, and to bind in the "safe haven" of the TATA-box where sequence changes through mutation are unlikely due to the resulting negative effects on endogenous transcription by the host (Grau *et al.*, 2013).

While the TAL effector-DNA binding code is a powerful tool, many false positive predictions are made when using association frequencies alone to predict TAL effector targets (Grau *et al.*, 2013; Cohn *et al.*, 2014). We anticipated that most of our predicted $TAL14_{Xam668}$ targets were likely false positives, prompting us to undertake a transcriptomic approach to identify true targets of $TAL14_{Xam668}$. RNA-Seq revealed that $TAL14_{Xam668}$ activates 52 cassava genes. In contrast, $TAL20_{Xam668}$ was shown to target only a single gene, *MeSWEET10a*, a sugar transporter that functions in promoting both bacterial growth and watersoaking symptom development during *Xam* infection (Grau *et al.*, 2013; Cohn *et al.*, 2014). It is unclear whether $TAL14_{Xam668}$ targets a conserved core promoter motif that activates many off-target genes but only one is truly required to promote disease, or if $TAL14_{Xam668}$ activates multiple genes whose collective over- or ectopic expression promotes disease. What we do know is that $TAL14_{ClO151}$ complements the growth defect seen in Xam668ΔTAL14ps in a manner identical to $TAL14_{Xam668}$ indicating that one or multiple of the target genes in group 1 are important for disease susceptibility. We point out the caveat that we only tested the 26 most highly upregulated genes, so other group 1 genes may exist. For example, *cassava4.1_014976* was predicted to be a virulence target of $TALE_{Xam1}$ from *Xam* strain CFBP1851 (which has an RVD sequence identical to $TAL14_{ClO151}$) based on gene expression analysis and EBE prediction, although its role in susceptibility was not experimentally validated (Muñoz-Bodnar *et al.*, 2014). This gene was also activated by $TAL14_{Xam668}$ with a mock versus Xam668 $\log_2$(fold change) of 2.4, and was thus outside of the group of targets that we tested for $TAL14_{ClO151}$ activation.

Designer TAL effectors (dTALEs) have been a useful tool for implicating single TAL effector targets as susceptibility (*S*) genes (Cernadas *et al.*, 2014; Hu *et al.*, 2014; Li *et al.*, 2014; Cohn *et al.*, 2014). Previously, it has been observed that when a TAL

effector targets more than one gene, the activation of only one of those genes is sufficient to complement the TAL effector mutant phenotype. Of the two gene targets of *Xoc* Tal2g in rice, only dTALEs targeting *OsSULTR3;6* complement the mutant phenotype, and of the two gene targets of the TAL effector PthA4 in citrus, only dTALEs targeting *CsLOB1* complement the mutant phenotype (Cernadas *et al.*, 2014; Hu *et al.*, 2014; Li *et al.*, 2014). We tested a 8 of the highly upregulated targets that were identified in our previously published *Xam* TAL effector study for their role in promoting growth in the $TAL14_{Xam668}$ mutant background, but did not identify a single *S* gene (Cohn *et al.*, 2014).

Given the deduced role of group 1 genes in promoting virulence, we utilized the program Mapman to assign these genes to functional categories and found that half of them are predicted to be involved in responses to biotic stress (Thimm *et al.*, 2004). These include the proteases *cassava4.1_033289* and *cassava4.1_024404*, the mannose-binding lectin *cassava4.1_034150,* and the pectate lyases *cassava4.1_007516* and *cassava4.1_007568.* The pectate lyases are particularly interesting given that plant pectin modification is a hallmark of *Xam* infection of cassava (Boher *et al.*, 1995). *Xam*-infected xylem vessels were found to be occluded by pectinaceous material associated with tyloses, which are outgrowths of the xylem parenchyma that form blockages in xylem vessels as a defense against the spread of vascular pathogens (Yadeta & Thomma, 2013). Secretion of pectin during tylose formation is thought to plug spaces not occluded by the tylose itself (Rioux *et al.*, 1998). The assumption that TAL effectors activate genes important for susceptibility would lead one to hypothesize that *Xam* is activating plant pectate lyases as a way to degrade these occlusive materials, overcome defense responses, and promote bacterial growth and movement through the vasculature. Interestingly, a dTALE activating the TAL effector-targeted pectate lyases (dT*_007568*, *007516*) expressed by Xam668ΔTAL14ps resulted in fewer colony forming units (CFU) at the site of midvein inoculation than Xam668ΔTAL14ps. One explanation for this is that the bacteria expressing dT*_007568*, *007516* were better able to move out of the inoculated area due to the breakdown of pectic barriers in the absence of possible movement-inhibiting “off-target” effects of the other $TAL14_{Xam668}$ targets. However, we were unable to test this as we could not reproducibly measure Xam668 movement through the leaf. Another possibility is that pectin degradation must be combined with the effects of other $TAL14_{Xam668}$ targets in order to promote bacterial growth to wild type levels. A third and contrastive possibility is that the activities of the pectate lyases actually have negative effects on *Xam* growth by aiding in the secretion of pectinaceous vessel-occluding material. In this scenario, the pectate lyases would be similar to executor (*E*) genes which are TAL effector-activated genes that promote disease resistance (Gu *et al.*, 2005; Römer *et al.*, 2007; Strauß *et al.*, 2012; Schornack *et al.*, 2013; Tian *et al.*, 2014). It would then follow that the negative effects of pectate lyase activation would be overcome by other targets of $TAL14_{Xam668}$ which promote disease susceptibility. Another group 1 target, the MtN21/EamA-like transporter *cassava4.1_009347,* is of interest since it is targeted by $TAL14_{ClO151}$, $TAL14_{Xam668}$ and $TAL22_{Xam668}$, and may therefore be a virulence hub (Hutin *et al.*, 2015). MtN21/EamA-like transporters have been associated with both amino acid

and auxin transport (Denance *et al.*, 2014). The *Arabidopsis* MtN21/EamA-like transporter AtUMAMIT18/SlAR1 is thought to be involved in amino acid loading of the apoplasm and the xylem and is primarily expressed in vascular tissues of source leaves (Ladwig *et al.*, 2012). Similar to $TAL20_{Xam668}$ which likely activates *MeSWEET10a* to promote the accumulation of a carbon source in the apoplasm at the site of bacterial infection, $TAL14_{Xam668}$ may be activating *cassava4.1_009347* to export amino acids into the apoplasm to be used as a bacterial nitrogen and carbon source. The remaining group 1 genes are a glyceraldehyde 3-phosphate dehydrogenase (*cassava4.1_011345*), and 3 genes of unknown function (*cassava4.1_019005, cassava4.1_011524*, and *cassava4.1_026299*).

Studies of RVD efficiency have been primarily done using highly expressed artificial TAL effector constructs and transient reporter assays. In this study, we show that a single repeat difference in two TAL effectors within one species of *Xanthomonas* can have a dramatic effect on target activation that is consistent with the RVD-nucleotide binding parameters that have been determined through studies of artificial TAL effectors. TAL14 proteins from *Xam* strains Xam668 and CIO151 promote bacterial growth in the host plant cassava making strategies of resistance in response to these proteins likely to be effective (Dangl *et al.*, 2013). One resistance strategy against TAL effectors that activate known *S* genes is targeted editing of the TAL effector EBE in the *S* gene promoter so that it is no longer activated upon infection (Li *et al.*, 2012). The identification of biologically relevant TAL effector-targeted S genes is made difficult when TAL effectors have multiple targets, as is the case with $TAL14_{Xam668}$ (Wilkins *et al.*, 2015). While we demonstrated the use of $TAL14_{CIO151}$ as a tool to narrow down the biologically relevant *S* gene targets of *Xam*, identification of a single TAL14-targeted *S* gene awaits future study. In the absence of known *S* genes, a feasible resistance strategy is an executor (*E*) gene approach where a resistance-triggering gene is engineered downstream from a TAL14-targeted promoter so that *Xam* is “tricked” into activating resistance upon infection (Römer *et al.*, 2009a; Hummel *et al.*, 2012; Schornack *et al.*, 2013; Boch *et al.*, 2014). The ideal *E* gene promoter would be activated by as many strains as possible. The results of this study show that in order to know whether an *E* gene construct will be effective against a group of TAL14-containing *Xam* strains requires knowledge of not only whether they contain a TAL effector of the proper size, but also the TAL effector RVD sequence and binding capacity for the *E* gene promoter.

5. A proof-of-concept executor gene approach to engineering resistance to cassava bacterial blight

Background

All *Xam* strains sequenced to date contain transcription activator-like (TAL) effector sequences, and virulence roles have been attributed to $TALE1_{Xam}$ from strain CFBP1851, $TAL14_{ClO151}$, $TAL14_{Xam668}$, and $TAL20_{Xam668}$ (Bart et al. 2012; Arrieta-Ortiz et al. 2013; Cohn et al. 2014, Castiblanco *et al.*, 2012). Effective resistance strategies against plant pathogens should respond to type III effectors that are important for virulence. When this is the case, the pathogen is less likely to evade recognition by effector loss, and, if the effector is dispensed, the resulting pathogen is impaired in virulence (Dangl *et al.*, 2013). Since TAL effectors are conserved in *Xam* and have been shown to be important virulence components, resistance strategies based on TAL effectors is a promising approach to engineering cassava plants that are resistant to bacterial blight of cassava (CBB).

Three natural plant mechanisms of resistance in response to TAL effectors are known [reviewed in (Schornack *et al.*, 2013; Boch *et al.*, 2014)]. The rarest is recognition of the TAL effector structure by a classic nucleotide binding site-leucine rich repeat (NBS-LRR) protein. The only known NBS-LRR protein that recognizes TAL effectors is the toll-interleukin 1 receptor (TIR) NBS-LRR, Bs4 (Ballvora *et al.*, 2001; Schornack *et al.*, 2004). A second mechanism of resistance against TAL effectors is the suppression of virulence function through promoter polymorphisms that abolish susceptibility (*S*) gene target activation. For example, the rice recessive resistance gene *xa13* is an allele of the TAL effector PthXo1-targeted *S* gene *OsSWEET11* with promoter polymorphisms that prevent PthXo1 binding. This renders strains of the rice pathogen *X. oryzae* pv. *oryze* (*Xoo*) that utilize PthXo1 as their primary determinant of virulence unable to cause disease on *xa13* plants (Chu, 2006; Yang *et al.*, 2006). Genome editing technologies have been employed to manipulate effector binding elements (EBEs) in the promoters of *S* genes such that they can no longer be activated in a TAL effector-dependent manner, generating resistant plants (Li *et al.*, 2012). With the discovery of a *Xam*-induced *S* gene in cassava this resistance strategy became a possibility and CRISPR-Cas9 mediated promoter editing is currently underway to eliminate the $TAL20_{Xam668}$ binding site in the *MeSWEET10a* promoter (Sander & Joung, 2014; Cohn *et al.*, 2014).

The third mechanism of plant resistance against TAL effectors is a form of the decoy model of resistance (*R*) gene-effector recognition whereby the invading effector is tricked into interacting with a host component that is similar to the natural virulence target, but instead triggers a robust resistance response (van der Hoorn & Kamoun, 2008). In this scenario, the plant genome encodes a TAL effector-targeted promoter driving an *R* gene, which in this context is termed an executor (*E*) gene. *Bs3* and *Bs4C* in pepper and *Xa10* and *Xa27* in rice are examples of natural TAL effector-induced *E* genes which are activated by AvrBs3 and AvrBs4 from *X. euvesicatoria*, and AvrXa10

and AvrXa27 from *Xoo*, respectively (Gu *et al.*, 2005; Römer *et al.*, 2007; Strauß *et al.*, 2012; Tian *et al.*, 2014). While TAL14$_{Xam668}$ has a virulence role, its *S* gene target(s) remain unknown. However, a TAL14$_{Xam668}$-induced *E* gene strategy remains a possibility. In this final chapter we provide proof-of-concept experiments that show that TAL14$_{Xam668}$ and TAL20$_{Xam668}$ targeted promoters can be engineered to drive expression of a resistance-triggering gene, *avrGf2*, whose expression can limit *Xam* growth and symptom formation in cassava.

Results

AvrGf2 elicits a hypersensitive response in cassava

Knowledge of cassava *R* genes is limited, so in order to find a protein capable of eliciting a resistance response in cassava, we screened through a number of bacterial effectors for elicitation of a strong hypersensitive response (HR). We found that AvrGf2 from *X. fuscans* sbsp. *aurantifolii* triggers a robust HR in cassava leaves when delivered by *Xam* strain Xam668 (Figure 5-1).

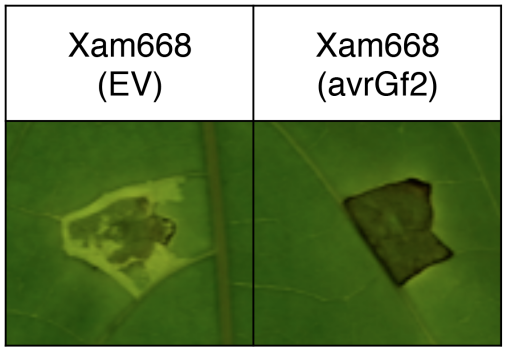


Figure 5-1. AvrGf2 from *Xanthomonas fuscans* sbsp. *arantifolii* elicits a hypersensitive response (HR) on cassava. A pLAFR3 cosmid library clone containing *avrGf2* was conjugated into Xam668. Xam668(avrGf2) was inoculated into cassava leaves (OD_{600} 0.01). Representative pictures were taken 7 days post inoculation (dpi) on a light box. Watersoaking symptoms which appear translucent on the light box are the result of inoculation with Xam668 (empty vector, EV), while Xam668-delivered avrGf2 induces HR.

Cassava4.1_007568 and *cassava4.1_020743* are direct targets of TAL14$_{Xam668}$

In order to employ an *E* gene strategy for resistance, a promoter that is directly bound and activated by a TAL effector is required. The promoter of *MeSWEET10a* is directly bound and activated by TAL20$_{Xam668}$ (Figure 3-4). We tested the promoters of

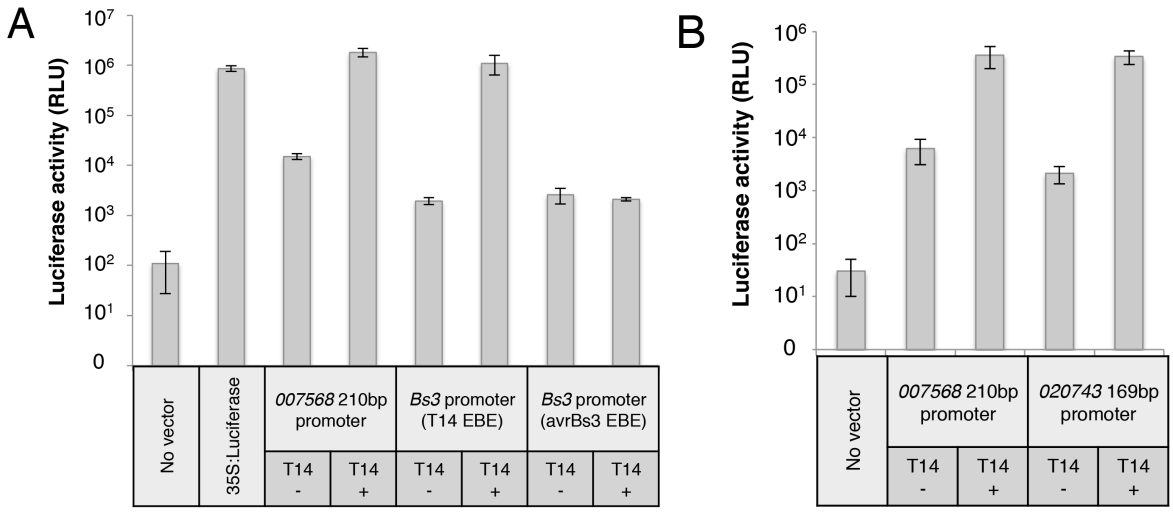


Figure 5-2. *Cassava4.1_007568* and *cassava4.1_020743* promoters are directly activated by TAL14$_{Xam668}$. *Agrobacterium* was used to deliver promoter-luciferase fusion constructs and 35S-driven TAL14$_{Xam668}$ into *Nicotiana benthamiana* cells. Luciferase activity (relative light units, RLU) was quantified 24 hours post inoculation (hpi). Data are displayed as mean technical replicates (+/- SD). Experiments were repeated 3 times with similar results.

cassava4.1_007568 and *cassava4.1_020743,* which are upregulated in a TAL14$_{Xam668}$-dependent manner, for direct activation using a luciferase reporter assay. The *cassava4.1_ 007568* promoter was selected because it is robustly activated by TAL14 proteins from both *Xam* strain CIO151 and strain Xam668. The *cassava4.1_020743* promoter was chosen because it is almost entirely shut off in the absence of *Xam* infection, a desirable trait when attempting to reduce leaky activation of a resistance-triggering gene. To test for direct activation the *cassava4.1_007568* and *cassava4.1_020743* promoters were used to drive expression of the reporter gene luciferase in an *Agrobacterium* binary vector. The constructs were individually co-expressed with 35S-driven TAL14$_{Xam668}$ delivered via *Agrobacterium* in *Nicotiana benthamiana* leaf cells. Luciferase activity was measured 24 hours post inoculation. *Agrobacterium* delivering TAL14$_{Xam668}$ specifically induced expression of both promoters (Figure 5-2A, B). Previous research has demonstrated that an *Xoo* TAL effector targeted EBE placed in the *Bs3* promoter from pepper (normally targeted by *Xe* TAL effector AvrBs3) retains its function (Römer *et al.*, 2009a), so we tested if this was true for the predicted TAL14$_{Xam668}$-targeted EBE from the *cassava4.1_007568* promoter (TATAAAAAGGTCGGT). While TAL14$_{Xam668}$ did not induce expression of the wild type *Bs3* promoter, when the predicted TAL14$_{Xam668}$ EBE was inserted into the *Bs3* promoter, activation was observed (Figure 5-2A). After validating direct TAL effector dependent activation, we moved forward with the *MeSWEET10a*, *cassava4.1_007568*, and *cassava4.1_020743* promoters driving *avrGf2* as our executor constructs in cassava (Figure 5-3).

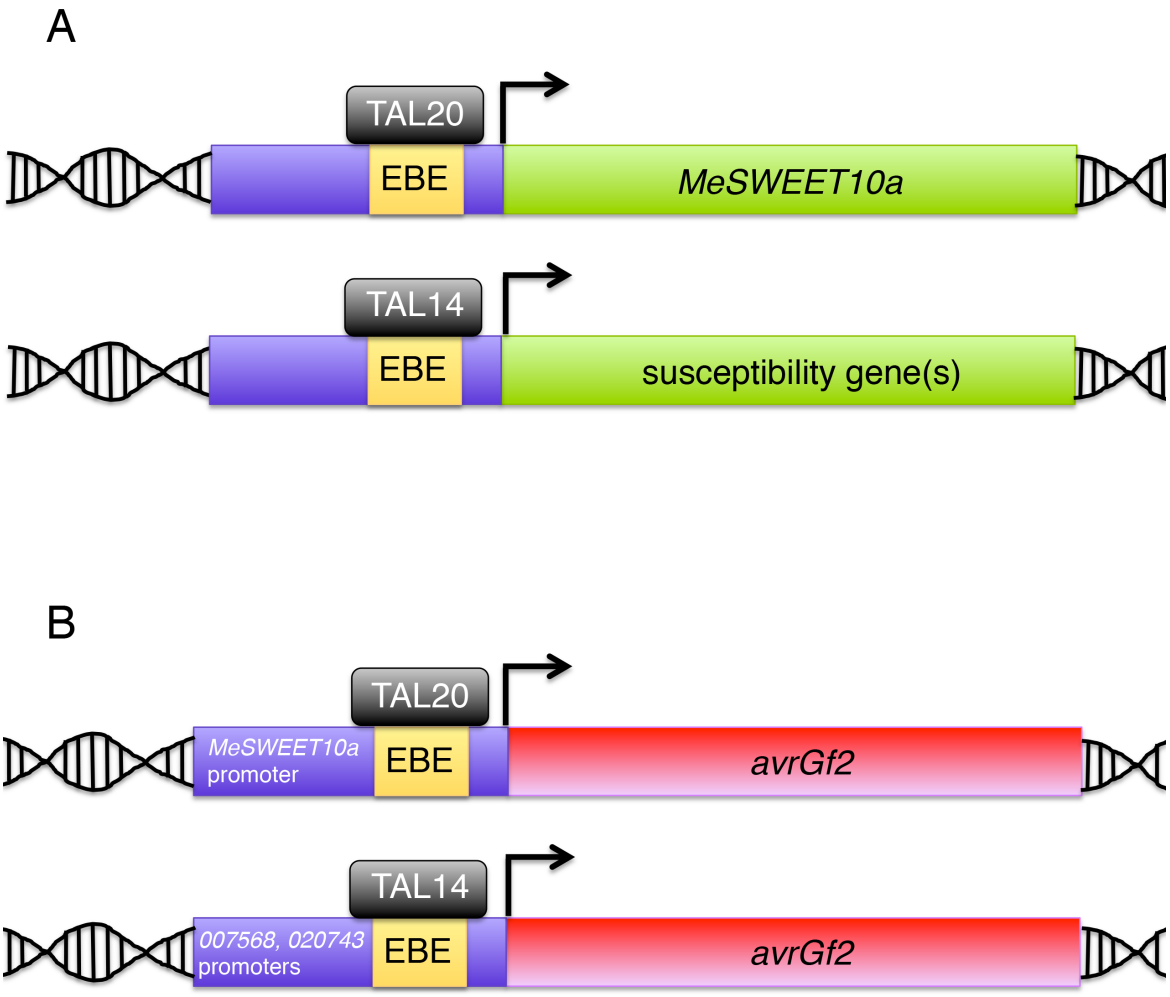


Figure 5-3. Illustration of the executor (*E*) gene resistance strategy in cassava. A. Typical *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) infection where TAL20$_{Xam668}$ activates expression of *MeSWEET10a* and TAL14$_{Xam668}$ activates unknown susceptibility gene(s) leading to disease. B. Cassava engineered with an *E* gene construct consisting of transcription activator-like (TAL) targeted promoters driving expression of *avrGf2* from *X. fuscans* sbsp. *aurantifolii*. TAL20$_{Xam668}$ and/or TAL14$_{Xam668}$ are "tricked" into activating a resistance response.

Leakiness of executor gene constructs is minimal

In order for an *E* gene strategy to be successful, the executor constructs cannot have leaky expression in the absence of the pathogen due to the detrimental effects of HR-eliciting genes on normal growth and development. We tested the elicitation of HR by *Agrobacterium*-delivered executor constructs in the absence of *Xam*

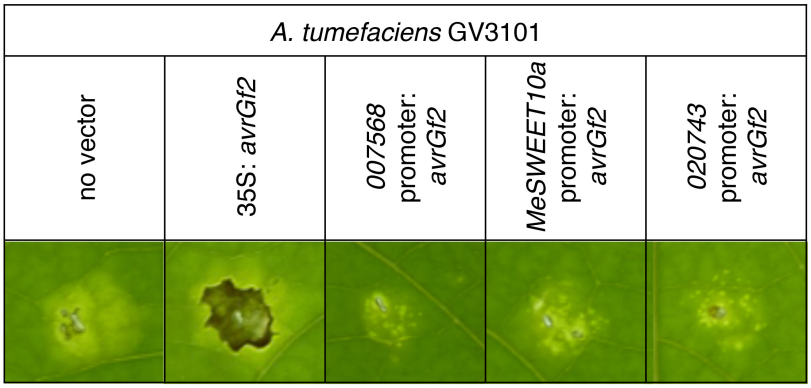


Figure 5-4. Executor constructs are not leaky when transiently expressed in cassava leaves. Constructs containing the hypersensitive response (HR)-eliciting gene *avrGf2* driven by the promoters of *MeSWEET10a, cassava4.1_007568* and *cassava4.1_020743* were delivered by *Agrobacterium tumefaciens* strain GV3101 into cassava leaves (OD_{600} = 0.1). HR was pictured 10 days post inoculation (dpi).

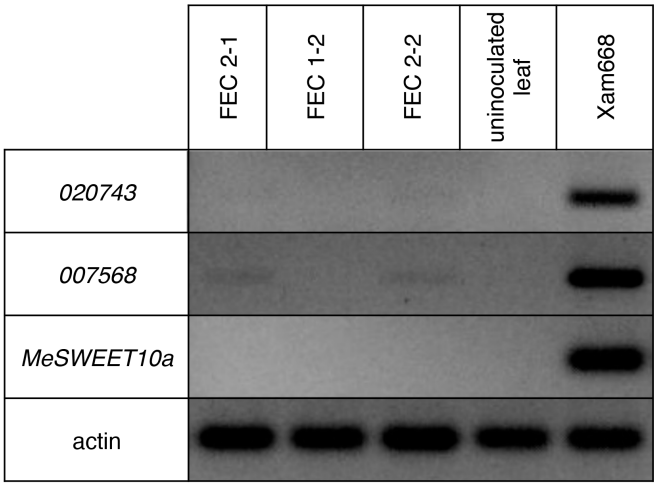


Figure 5-5. Executor gene promoter activity in friable embryonic calli (FEC). Semiquantitative RT-PCR was used to check for executor construct promoter activity in FEC of cassava cultivar TMS 60444. The results of RT-PCR on 3 replicates of FEC tissue are shown in addition to uninoculated leaf tissue and leaf tissue inoculated with Xam668. RT-PCR primers and corresponding cycle numbers are listed in table S2.

in cassava leaves. While our 35S-driven *avrGf2* construct triggered an HR response, we did not see any cell death elicited by our executor constructs alone (Figure 5-4).

Similarly, it is important that the executor constructs are not activated during the process of *Agrobacterium*-mediated transformation to produce stable transgenic plants, which in cassava is carried out on friable embryonic calli (FEC) (Bull *et al.*, 2009). We looked for endogenous activity of the cassava promoters used in our *E* gene constructs

in 3 FEC replicates as well as in uninoculated leaf tissue by semiquantitative RT-PCR and saw no gene expression for the *MeSWEET10a* and *cassava4.1_020743* promoters, and a small level of expression for the *cassava4.1_007568* promoter in 2/3 FEC (Figure 5-5).

Executor constructs decrease *Xam* growth and symptom development in transient assays

Before investing in the production of transgenic cassava plants, we wanted to test if our *E* gene constructs have the capacity to limit *Xam* growth and symptom formation in transient assays. We co-inoculated *Agrobacterium* strains delivering our *E* gene constructs with a low inoculum of Xam668 carrying a pLAFR-*lux* plasmid [Xam668(pLAFR-*lux*)] into cassava leaves. Xam668(pLAFR-*lux*) constitutively expresses the components necessary for luciferase reporter activity, allowing for visualization by X-ray film exposure (Cohn *et al.*, 2015). At 6 and 10 days post inoculation (dpi) we observed robust elicitation of HR by *007568*_promoter:*avrGf2* and *MeSWEET10a*_promoter:*avrGf2* in response to Xam668(pLAFR-*lux*). The *020743*_promoter:*avrGf2* construct showed a weaker HR and allowed the formation of some watersoaking symptoms. HR strength was negatively correlated with the amount of Xam668(pLAFR-*lux*) in the inoculated spot as visualized by X-ray film exposure (Figure 5-6A). We measured colony forming units of Xam668(pLAFR-*lux*) in the presence of our *E* gene constructs delivered by *Agrobacterium* and saw at 10 dpi a reduction in *Xam* growth in the presence of *007568*_promoter:*avrGf2* and *MeSWEET10a*_promoter:*avrGf2* that was similar to what was seen in the presence of *35S:avrGf2*. While the *020743*_promoter:*avrGf2* construct also limited Xam668(pLAFR-*lux*) growth, the effect was not as great (Figure 5-6B).

Discussion

In this study we present proof-of-concept experiments showing that a TAL effector-targeted promoter activating a resistance-triggering *E* gene can reduce *Xam* growth and CBB symptom formation in cassava leaves. Since there are no known naturally occurring *E* genes in cassava, we screened a variety of bacterial effectors for their ability to elicit HR in cassava. AvrGf2 of *X. fuscans* subspecies *aurantifolii* was found to elicit a strong HR when expressed from a plasmid by *Xam* and also when transiently delivered by *Agrobacterium*. We created *E* gene constructs by cloning avrGf2 downstream of the *MeSWEET10a* promoter which is activated by $TAL20_{Xam668}$, and the promoters of *cassava4.1_007568* and *cassava4.1_020743* which are activated by $TAL14_{Xam668}$. We then expressed these *E* gene constructs transiently in cassava leaves and measured their ability to inhibit growth of Xam668. The three constructs were

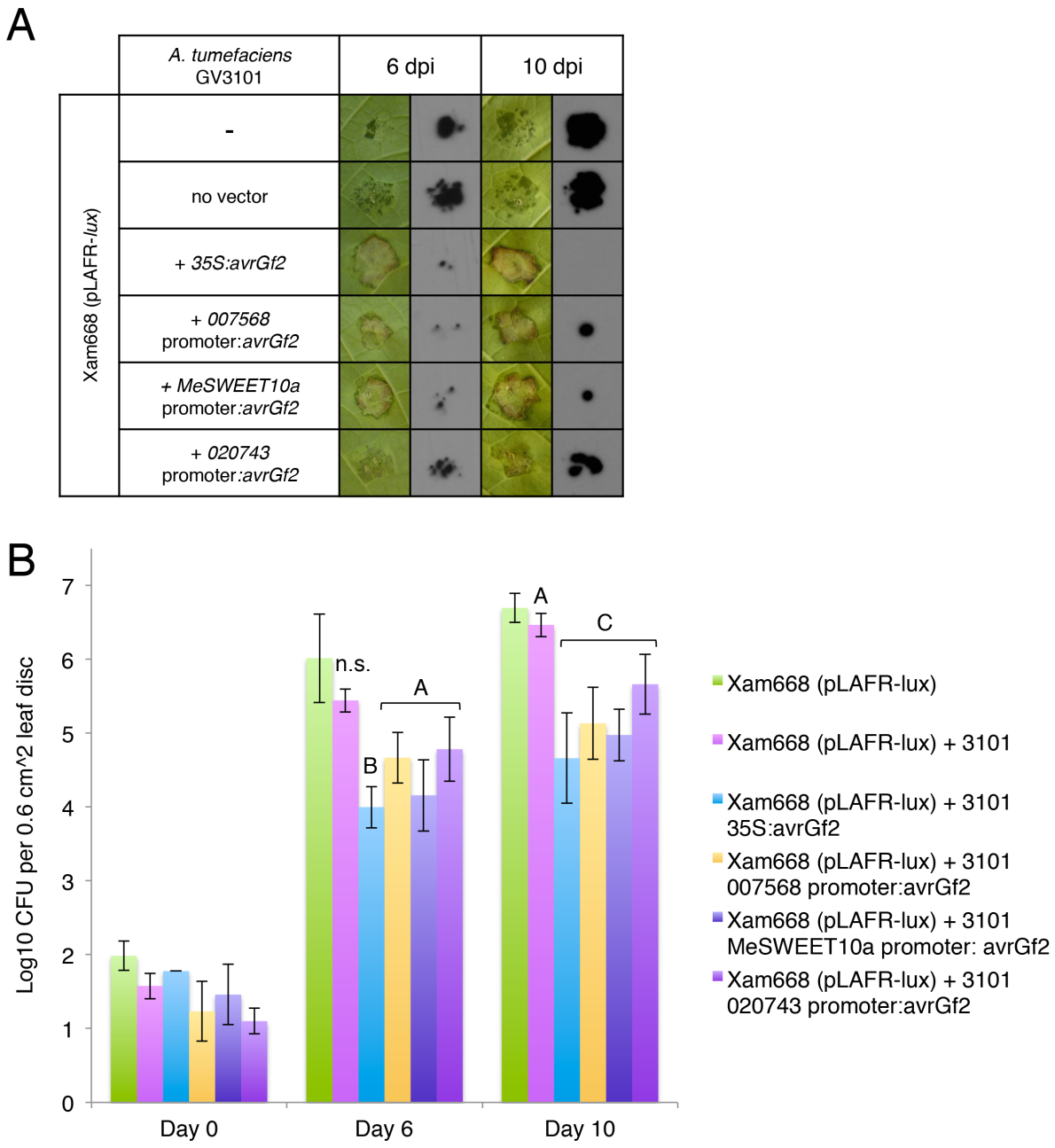


Figure 5-6. Executor (*E*) gene constructs limit *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) growth and symptom development in transient assays. *Agrobacterium* strains delivering the *E* gene constructs *007568*_promoter:*avrGf2, MeSWEET10a*_promoter:*avrGf2* and *020743*_promoter:*avrGf2* were co-inoculated with Xam668(pLAFR-*lux*) in cassava leaves (*Agrobacterium* OD_{600} = 0.4, *Xam* OD_{600} = 0.0001). A. Hypersensitive responses (HR) and luciferase activity were recorded 6 and 10 days post inoculation (dpi). B. Colony forming units (CFU) of Xam668(pLAFR-*lux*) were measured at 6 and 10 dpi (t-test; A, $p < 0.05$; B, $p < 0.01$; C, $p < 0.001$; n.s., not significant).

successful at inhibiting *Xam* growth and symptom formation, with the *MeSWEET10a*_promoter*:avrGf2* and the *cassava4.1_007568*_promoter*:avrGf2* constructs showing the highest levels of inhibitory capability.

A prerequisite to a successful *E* gene strategy is to have an *E* gene that is silent in the absence of the pathogen. As HR results in cell death, an *E* gene construct with leaky expression could result in either an inability to be transformed, or if transformation is successful, a sickly plant. Naturally occurring *E* genes are silent in the absence of pathogen attack (Gu *et al.*, 2005; Römer *et al.*, 2007; Strauß *et al.*, 2012; Tian *et al.*, 2014). The cassava promoters we chose for our *E* gene constructs appear to be silent when transiently expressed in cassava leaves in the absence of *Xam*, and the *MeSWEET10a* and *cassava4.1_020743* promoters are silent in FECs. The *cassava4.1_007568* showed a low level of expression in 2 out of 3 FEC replicates, indicating that the rate of transformation may be lower for this construct than the others due to detrimental effects of leaky expression in some transformants. However, RNA-Sequencing indicates a very low level of activity for our *E* gene promoters in mock inoculated tissue (Figure 3-3, 4-3). This small level of expression may prevent successful transformation events or may cause growth to be stunted in transformed plants. Nevertheless, we hope to generate stable transformants with the *E* gene constructs presented in this work.

E gene promoters can be engineered to contain multiple functioning TAL effector binding sites (Römer *et al.*, 2009a; Hummel *et al.*, 2012). While the cassava gene promoters used in our *E* gene constructs are targeted by single *Xam* TAL effectors, additional EBEs can be inserted in order to ensure recognition of multiple TAL effectors. Since $TAL14_{Xam668}$ and $TAL20_{Xam668}$ promote *Xam* virulence, placing code-optimized EBEs for these TAL effectors in *E* gene promoters could enable durable resistance. We suggest stacking promoter mutations that abolish $TAL20_{Xam668}$-mediated *MeSWEET10a* activation with transformation of a $TAL14_{Xam668}$/$TAL20_{Xam668}$ targeted *E* gene construct in order to generate cassava plants with increased resistance to CBB.

6. Materials and Methods

6-1. Xam668 genomic DNA cosmid library construction

DNA extraction:
A 200 ml overnight culture of Xam668 was centrifuged for 5 minutes at 3,619 x *g*, washed with H_2O, resuspended in 1 ml H_2O, and mixed with 15 ml DNA extraction buffer (0.3 M NaCl, 50 mM Tris pH 7.5, 20 mM EDTA, 2% sarkosyl, 0.5% SDS, 5 M urea, 5% phenol in H_2O). 15 ml phenol:choloroform:isoamyl alcohol (24:24:1) was added and the mixture was shaken and poured into 50 ml Phase Lock Gel tubes (5 PRIME) and centrifuged for 1 minute at 5,211 x *g*. 15 ml of the upper phase was transferred to a new tube and mixed with 20 ml isopropanol by inverting slowly 50 times, then centrifuged for 5 minutes at 5,211 x *g*. The pellet was washed with 70% ethanol, then resuspended in 20 ml T5E (50 mM Tris pH 8.0, 10 mM EDTA pH 8.0).

Cesium chloride (CsCl) equilibrium gradient:
20 g CsCl was dissolved in 20 ml T5E/DNA. The T5E/DNA/CsCl mixture was loaded into a 40 ml Quick-Seal tube (25 x 89 mm, Beckman Coulter) along with 3 ml ethidium bromide. The tube was heat-sealed after filling with 1 g CsCl per 1 ml TE stock solution until only a 0.5 ml bubble remained at the top. The mixture was centrifuged in a Vti 50 rotor (Beckman Coulter) at 196,000 x *g* (15° C, deceleration 9) for 24 hours. DNA bands were illuminated with long wave UV light and drawn out of the tube side using a syringe fitted with an 18 gauge needle. The addition, mixture, and removal of 1 volume of water saturated butanol was carried out 4 times to extract the ethidium bromide from the sample. Sample was dialyzed against TE overnight at 4° C to remove cesium chloride. DNA integrity was evaluated by agarose gel electrophoresis.

DNA digestion and size fractionation:
100 μg of Xam668 genomic DNA was digested with Tsp509I (NEB) at 65° C in a total volume of 1 ml. 200 μl digestion reaction aliquots were stopped at 3, 4.25, 5.5, 6.75, and 8 minutes by the addition of 20 μl 0.5 M EDTA pH 8.0. Digestion was confirmed by agarose gel electrophoresis. Two 12 ml 6-step (10-40% sucrose in TE) gradients were prepared in 13.2 ml centrifuge tubes (14 x 89 mm, Beckman Coulter), loaded with 500 μl partially digested DNA, and centrifuged in a SW-41 rotor (Beckman Coulter) at 150,000 x *g* (20° C, deceleration 9) for 18 hours. The gradient was divided into 0.5 ml fractions and DNA size was determined by running 5 μl aliquots on a 0.75% agarose gel. Fractions with approximately cosmid sized inserts (20-40 kb) were pooled for precipitation. 1 μg CsCl gradient-purified pLAFR3 vector was digested with EcoRI (NEB) for 1 hour at 37° C.

Ligation and library packaging:
1 μg of size-fractionated Xam668 genomic DNA was combined with 0.1 μg digested pLAFR3, ethanol precipitated, resuspended in 9 μl water, heated for 5 minutes at 65° C, and cooled at room temperature for 1 hour. Ligation with T4 DNA ligase (high

concentration, NEB) was carried out overnight at 12° C. 12 μl of Gigapack III lambda packaging extract (Agilent Technologies) was added to 5 μl of ligation product and kept at room temperature for 2 hours. 125 μl lambda dilution buffer (10 mM Tris pH 7.5, 10 mM $MgCl_2$) was added and mixture stored at 4° C.

Library transduction:
E. Coli DH5α OD_{600} 0.25 was grown at 37° C, shaking, for 4 hours in LB + 20% maltose. 100 μl pre-grown *E. Coli* DH5α was mixed with 25 μl packaging product, 75 μl lambda dilution buffer, and 100 μl salts (10 mM $MgCl_2$, 10 mM $CaCl_2$). The mixture was placed at 37° C (not shaking) for 30 minutes, and then allowed to shake at 37° C for 1 hour before plating on LA with X-gal (40 μg/ml), tetracycline (10 μg/ml), and cycloheximide (50 μg/ml). Approximately 1,000 white colonies were selected for the library.

6-2. Transcription activator-like (TAL) effector sequencing

Sequencing of $TAL13_{Xam668}$, $TAL14_{Xam668}$, and $TAL15_{Xam668}$:
TAL effectors were sequenced from cosmid clones using oligonucleotide primers ATG F (ATGGATCCCATTCGTCCGCGCA), stop R (TCACTGAGGAAATAGCTCCAT), 744 F (AGATTGCAAAACGTGGCGGCG), 2409 R (TGCGGCAATCCCTTTCTCACT), 5' rpt jxn (CTGCTCCGTCAGTTCGATCCG), and 2409 F (AGTGAGAAAGGGATTGCCGCA).

Sequencing of $TAL20_{Xam668}$:
Sequences from primers 744 F and 2409 R were not sufficient to span the repeat region of TAL20Xam668. A unique SmaI site within the TAL20Xam668 repeat region allowed the generation of a pENTR/D-TOPO TAL20Xam668 (NotI-SmaI) deletion construct. An M13 forward vector primer was then used to sequence the core repeat sequence of TAL20Xam668.

Sequencing of $TAL22_{Xam668}$:
Sequences from primers 744 F and 2409 R were not sufficient to span the repeat region of $TAL22_{Xam668}$. Nested deletions of the $TAL22_{Xam668}$ BamHI fragment in pUC19 were generated using the Deletion Kit for Kilosequencing (Takara). 12 deletion constructs were sequenced from the vector-located M13 reverse primer site and assembled as 22 repeats. The sequence was confirmed by verifying the predicted banding pattern of BspmI-digested $TAL22_{Xam668}$.

6-3. TAL effector mutant and complementation constructs

Transcription activator-like (TAL) effector mutant constructs:
Knockout strains were generated through integration of a suicide vector into the coding region of the TAL effectors. A 665 bp region 5' of the repeats of $TAL13_{Xam668}$ was

modified to remove the ATG start site using primers 5'- CACCCGAACCCGTCTCCCG TCTC -3' and 5'- ACCGCGGTCACGCCGCCACGT -3' and cloned into a pENTR/D-TOPO vector (Life Technologies). The gateway compatible destination cassette ccdB-B (Life Technologies) was cloned into the EcoRI site (filled in by T4 DNA polymerase) of the suicide vector pLVC18 (DNA Plant Technology, Oakland, CA). The modified TAL effector construct was mobilized into this destination suicide vector pLVC18Tc^r Des by LR recombination and then conjugated into Xam668 or CIO151 via tri-parental mating with *E. coli* helper strain pRK600. Strains carrying endogenous plasmids with single homologous integrations of the suicide construct into TAL effectors were selected on NYGA plates with rifampicin (100 μg/mL) and tetracycline (10 μg/mL). To generate single Xam668ΔTAL mutants these strains were further cycled on NYGA plates with rifampicin (100 μg/mL) and tetracycline (10 μg/mL) to allow the elimination of the corresponding non-tagged endogenous TAL plasmid copies and screened via western blot for loss of the corresponding TAL effector expression.

Xam668ΔT20p and Xam668ΔTAL14ps were generated by cycling Xam668 on rifampicin and screening for TAL effector loss by Southern blot. Megaplasmid profiling revealed loss of the smallest plasmid in the case of Xam668ΔT20p, and loss of the $TAL14_{Xam668}$ plasmid segment in the case of Xam668ΔTAL14ps.

***$TAL14_{Xam668}$*, *$TAL20_{Xam668}$*, and *$TAL14_{CIO151}$* complementation constructs:**
The Xam668ΔTAL20 knock-out strain was complemented with *$TAL20_{Xam668}$* driven by its 1 kb native promoter. The *$TAL20_{Xam668}$* promoter was PCR amplified from a cosmid clone template using forward and reverse oligonucleotide primers 5'-CACCCCCTCATG GCCGAGCTGC-3' and 5'-GAATTCCCCCCCGGTACCCAGGCATACCTGTTTTATA-3' which added a KpnI site and an EcoRI site at the 3' end of the promoter. *$TAL13_{Xam668}$* was PCR-amplified from a cosmid clone template using forward and reverse oligonucleotide primers 5'-CCCCCGGTACCATGGATCCCATTCGTCCGC GC-3' and 5'-CCCCCGAATTCTCACTGAGGAAATAGCTCCAT-3' which added a KpnI site 5' of the start codon and an EcoRI site 3' of the stop codon. The amplified fragments were ligated into pENTR/D-TOPO (Life Technologies). A pENTR/D-TOPO vector containing *$TAL20_{Xam668}$* was constructed by modifying the above vector containing *$TAL13_{Xam668}$*. The *$TAL13_{Xam668}$* internal BamHI fragment was replaced with a *$TAL20_{Xam668}$* internal BamHI fragment that had been subcloned from a cosmid clone. *$TAL13_{Xam668}$* and *$TAL20_{Xam668}$* have identical sequences outside of the BamHI cut-sites, so this strategy resulted in a pENTR/D-TOPO vector containing a wild-type copy of *$TAL20_{Xam668}$*. *$TAL20_{Xam668}$* was directionally cloned into the pENTR/D-TOPO *$TAL20_{Xam668}$* promoter vector by digestion with KpnI-HF and EcoRI-HF (New England Biolabs) and ligation with T4 DNA ligase (Roche Rapid DNA Ligation Kit).

The Xam668ΔTAL14 knock-out line and Xam668ΔTAL14ps were complemented with *$TAL14_{Xam668}$* driven by the 1 kb *$TAL20_{Xam668}$* promoter, necessitated by the lack of sequence for the *$TAL14_{Xam668}$* promoter region. *$TAL14_{Xam668}$* was cloned as described above for *$TAL13_{Xam668}$*. *$TAL14_{Xam668}$* was directionally cloned into the pENTR/D-TOPO

TAL20$_{Xam668}$ promoter vector by digestion with KpnI-HF and EcoRI-HF (New England Biolabs) and ligation with T4 DNA ligase (Roche Rapid DNA Ligation Kit).

Complementation construct TAL effector sequences 5' and 3' of the repeat region were verified using forward and reverse oligonucleotide primers 5'-ATGGATCCCATTCGTCC GCGC-3' and 5'-TCACTGAGGAAATAGCTCCAT-3'. The repeat regions were sequenced with forward and reverse oligonucleotide primers 5'-AGATTGCAAAACGTG GCGGCG-3' and 5'- TGCGGCAATCCCTTTCTCACT-3' located just 5' (plus strand) and just 3' (minus strand) of the repeats. Sequences from these primers verified the entire repeat region for *TAL14*$_{Xam668}$ and most of the repeat region for *TAL20*$_{Xam668}$. A unique SmaI site within the TAL20$_{Xam668}$ repeat region allowed the generation of a pENTR/D-TOPO TAL20$_{Xam668}$ (NotI-SmaI) deletion construct. An M13 forward vector primer was then used to sequence the core repeat sequence of *TAL20*$_{Xam668}$.

The gateway compatible destination cassette ccdB-B (Life Technologies) was cloned into the T4 DNA polymerase filled in HindIII site of the stable broad host-range vector pVSP61. *TAL20*$_{Xam668}$ promoter-driven *TAL20*$_{Xam668}$ and *TAL14*$_{Xam668}$ were transferred into pVSP61Des by LR recombination (Life Technologies). The resulting constructs were conjugated into Xam668ΔTAL20 or Xam668ΔTAL14/Xam668ΔTAL14ps), respectively, with *E. coli* helper strain pRK600. Expression of TAL14$_{Xam668}$ and TAL20$_{Xam668}$ was verified by western blot.

A TAL14$_{ClO151}$ complementation construct was generated as described above for TAL14$_{Xam668}$.

6-4. TAL effector target prediction

TALE-NT (2.0) Target Finder:

For prediction of targets of Xam668 TAL effectors, Target Finder searched the cassava promoterome (1 kb upstream of annotated transcriptional start sites, cassava genome version 4.1) (Prochnik *et al.*, 2012; Doyle *et al.*, 2012). Only the forward DNA strand was searched and a 5' T (T_0) was required. Each TAL effector was assigned a best binding score given its RVD sequence, and sites within a 3-fold cut-off of this score were considered potential effector binding elements (EBEs).

For prediction of targets of TAL14$_{Xam668}$ and TAL14$_{ClO151}$ in group 1 and group 2 target promoters, Target Finder searched 300 bp upstream of the gene start codons (cassava genome version 4.1). Only the forward DNA strand was searched and a 5' T (T_0) was required. Each TAL effector was assigned a best binding score given its RVD sequence, and sites within a 3.5-fold cut-off of this score were considered potential EBEs.

TALgetter:
TALgetter (version 1.0) predicted targets of $TAL14_{Xam668}$ and $TAL14_{ClO151}$ in group 1 and group 2 target promoters 300 bp upstream of the gene start codons (cassava genome version 4.1) (Prochnik *et al.*, 2012; Grau *et al.*, 2013). Computation of p-values was fine-grained, and the maximum p-value was set at 1e-2. The TALgetter standard model was used for model training.

6-5. Plant inoculations and virulence assays

Assays of bacterial growth in leaf midvein:
For the midvein growth assays presented in section 3 of this work, 2 mm holes were made with a glass pasteur pipette dipped in 10 mM $MgCl_2$ in the midvein approximately ¼ of the way in from the leaf tip. A 5 μl drop of bacterial suspension at OD_{600} = 0.2 was immediately placed in the hole and allowed to dry undisturbed for 15 minutes. At each time point, 3 cm midvein sections starting 0.5 cm in from the inoculation point were cut out and ground in 10 mM $MgCl_2$ with a 3 mm glass bead in a beadbeater. Two midvein sections were combined per replicate, for a total of 3 replicates. Serial dilutions were plated on the appropriate selection plus cycloheximide to inhibit fungal growth.

A detailed protocol for the midvein growth assays presented in section 4 of this work is published (Cohn *et al.*, 2015).

Assays of bacterial growth in leaf apoplast:
Xam strains were washed and resuspended in 10 mM $MgCl_2$ at OD_{600} = 0.00005. Abaxial nicks were made on leaves of cassava cultivar 60444 with a razor blade and culture was injected into the leaf via a 1 mL needleless syringe. At each time point, leaf punches were taken from the inoculation sites in triplicate and ground in 10 mM $MgCl_2$ with a 3-mm glass bead in a beadbeater. Serial dilutions were plated on the appropriate selection plus cycloheximide to inhibit fungal growth.

6-6. Plasmid preparation for visualization by agarose gel electrophoresis

1 ml of OD_{600} = 0.25 cell suspension was pelleted and resuspended in 175 μl fresh lysis buffer (125 mM NaOH, 30 mM NaCl, 50 mM Tris, 5 mM EDTA, 3% SDS). Suspensions were vortexed for 1 second, then incubated at 65°C for 5 minutes. 400 μl Phenol Chloroform (1:1) was added to each sample and the mixtures were vortexed for 30 seconds then centrifuged at maximum speed for 5 minutes. 20 μl of the top phase was loaded onto 0.75% agarose gels for visualization.

6-7. Western Analysis

Minimal Media Expression and Protein Extraction:
Overnight liquid cultures of *Xam* were re-suspended in 125μL Minimal Media A pH 5.14 to OD_{600} = 0.2 along with appropriate antibiotics: *Xam* strains rifampicin (100 μg/mL), ΔTALEs rifampicin (100 μg/mL) and tetracycline (10 μg/mL), Xam668ΔTALEs(pVSP61 TALE) rifampicin (100 μg/mL), tetracycline (10 μg/mL), and kanamycin (25 μg/mL), and shaken at 28ºC for 4 hrs. An equal volume of 3x Laemmli buffer was added and samples were boiled 6 minutes, vortexed, spun down and stored as aliquots at -80ºC.

Protein extraction from inoculated leaf tissue:
Xam strains were inoculated via needleless syringe as described above at OD_{600} = 0.5. 30-50 mg inoculated leaf tissue was frozen and protein was extracted with 300 μl 3x Laemmli buffer. Samples were boiled for 5 minutes, vortexed, spun down and stored as aliquots at -80ºC.

Western blot:
Protein lysates were separated via SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using either NuPAGE Novex 4-12% Bis-Tris precast gels in NuPAGE MOPS SDS running buffer, or 3-8% Tris-Acetate precast gels in NuPAGE Tris-Acetate SDS running buffer (Life Technologies). TAL effector content was analyzed using standard immunoblotting techniques with a polyclonal antibody capable of recognizing diverse TAL effectors.

6-8. $TAL20_{Xam668}$ RNA sequencing (RNA-Seq)

Strain Construction for RNA-Seq:
Xe($TAL20_{Xam668}$) was made by conjugating a cosmid library clone (pLAFR3) containing *$TAL20_{Xam668}$* into *Xe* 85-10 by triparental mating using *E. coli* helper strain pRK600. Expression of $TAL20_{Xam668}$ was verified by western blot.

Plant Inoculations for RNA-Seq:
Xam and *Xe* grown on NYGA plates supplemented with rifampicin (100 μg/mL) or rifampicin and tetracycline (10 μg/mL) for *Xe*($TAL20_{Xam668}$) were re-suspended in 10 mM $MgCl_2$ at OD_{600} = 0.5. Abaxial nicks were made on leaves of cassava cultivar 60444 with a razor blade and culture was injected into the leaf via a 1mL needleless syringe. Mock infiltrations were done with 10 mM $MgCl_2$.

Library preparation:
Total RNA from inoculated leaf tissue (30-50 mg) was extracted with the Spectrum Plant Total RNA Kit (Sigma-Aldrich), with an on-column DNaseI digestion step included. RNA quality was checked by Bioanalyzer (RNA pico chip, Functional Genomics Laboratory (FGL), UC Berkeley). RNA-Seq libraries were made using the TruSeq RNA sample

preparation kit, v2 (Illumina), starting with 3 μg total RNA. Library quality was assessed by Bioanalyzer (DNA 1000 chip, FGL, UC Berkeley). Quantification by qPCR and pooling of samples was carried out at UC Berkeley's Genomics Sequencing Laboratory (GSL). Libraries were sequenced at the GSL on an Illumina HiSeq 2000, generating 100 bp paired-end reads.

Data Analysis:
Raw read files were trimmed for quality and adaptor removal using Trimmomatic (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) (Lohse *et al.*, 2012). After trimming, gene expression changes were identified using the Tuxedo suite of tools with the default setting (Trapnell *et al.*, 2012). In brief, bowtie-build was used to make a bowtie index from the only publicly available and annotated cassava genome, AM560-2 v 4.1 available from Phytozome (http://www.phytozome.net/) (Prochnik *et al.*, 2012). TOPHAT (-I 1000 -i20) was used to align reads to identify gene models, map reads to gene models and estimate expression levels. Analyses were first conducted individually on three biological replicates and global gene expression patterns were examined using the R-implementation CummeRbund (Trapnell *et al.*, 2012) to identify outlying biological replicates. As all three replicates had similar expression profiles, final analysis was conducted on the union of all three biological replicates for increased statistical power. Cufflinks was run with default settings on the "accepted hits" bam files from the TOPHAT analysis. Cuffmerge was used to join all gene models from all Cufflinks analyses and then Cuffdiff was used to estimate differential gene expression values. By default, Cuffdiff analyzes all possible combinations of input conditions. In-house python scripts were used to pare data down to comparisons of mock vs. Xam668 and *Xe* 85-10 vs. *Xe*(TAL20) for each time point. Comparisons that showed at least 3-fold induction at one time point were chosen for further consideration. In addition, comparisons that yielded non-statistically significant data were excluded. A heat map of the resulting list was generated to compare expression patterns across all samples. Rows that showed highly variable or inconsistent trends in expression were excluded.

6-9. Semiquantitative RT-PCR analysis
Xam strains were inoculated into cassava leaves by needleless syringe as described at OD_{600} = 0.5. Total RNA was extracted with the Spectrum Plant Total RNA Kit (Sigma-Aldrich), with an on-column DNaseI digestion step included. Starting with 1 μg total RNA, cDNA was made using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) and diluted 1:10. 25 μl RT-PCR reactions contained 10 μl of the diluted cDNA sample, 0.2 μM forward primer, 0.2 μM reverse primer, and 1X OneTaq Master Mix with Standard Buffer (New England Biolabs). Reactions cycled 25 times (unless otherwise stated in figure legends) and were run on high % TAE agarose gels. For cycloheximide treatments, bacterial infiltrations were done with suspensions containing 50 μM cycloheximide. Actin (gene ID *cassava4.1_009807*) was used as a control for all RT-PCRs.

6-10. Construction of plasmids for Luciferase transient assays

Promoter-luciferase reporter constructs:
184 bp of the *cassava4.1_013474* promoter (up to the ATG) was PCR amplified from cassava cultivar 60444 genomic DNA using forward and reverse oligonucleotide primers 5'- CCCGCGGCCGCGTTACATTGACATATTTTATTCAC-3' and 5'- CCCGG CGCGCCTTCTCCGGCTATAGTAGAGACTCTC-3'. No polymorphisms were seen in the $TAL20_{Xam668}$ EBE compared to the AM560-2 reference genome sequence. The promoter was directionally cloned into pENTR/D-TOPO vectors (Life Technologies) by digestion with Not-HF and AscI (New England Biolabs) and ligation with T4 DNA ligase (Roche Rapid DNA Ligation Kit).

The *Bs3* promoter sequence containing the $TAL20_{Xam668}$ EBE was constructed with the QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene). The AvrBs3 EBE was changed to the $TAL20_{Xam668}$ EBE using the forward and reverse mutagenic oligonucleotide primers 5'- GCCTGACCAATTTTATAAACGCTTCTCGCCCATCCTCA CAACTTCAAG-3' and 5'- CTTGAAGTTGTGAGGATGGGCGAGAAGCGTTTATAAAA TTGGTCAGGC-3'. The promoter sequence was transferred to binary vector pGWB35 by LR recombination (Life Technologies), making a fusion with the reporter gene luciferase.

Construction of TAL effector constructs for transient expression:
The vectors for transient *Agrobacterium*-mediated delivery of $TAL20_{Xam668}$ was constructed by transferring *$TAL20_{Xam668}$* from the pENTR/D-TOPO vector to the binary vector pEG100 by LR recombination (Life Technologies), resulting in 35S-driven $TAL20_{Xam668}$. The construct was mobilized into *Agrobacterium* strain GV3101 by conjugation with *E. coli* helper strain pRK600.

6-11. *Nicotiana benthamiana* transient expression analyses
Leaves of *N. benthamiana* were injected with *Agrobacterium* GV3101 suspensions (OD_{600} = 0.4) using a 1mL needleless syringe. 24 hours post infection, sites of *Agrobacterium* inoculation were infiltrated with 1 mM luciferin. For each site, four 0.28 cm^2 leaf punches were taken and placed in separate wells of a black microtiter plate, suspended on 100 µl of water. Luciferase activity was read using a Wallace Envision machine.

6-12. Electromobility shift assay (EMSA)
EMSA analysis was performed as described (Römer *et al.*, 2007). Biotin labeled 35 bp oligonucleotides were mixed with 0, 50, 150 and 300 fmol of purified $TAL20_{Xam668}$ effector protein. For competition assays, unlabeled oligonucleotides were pre-incubated with the protein at 1x, 10x and 100x the concentration of the biotin labeled

oligonucleotides. *MeSWEET10a* EMSA oligo sequence: TCCCCTATATAAACGCTTCTCGCCCATCCATCATT.

6-13. Construction of Xam668ΔsuxC and complemented strain

Construction of pLVC18tcSacBRdes:
Vector pLVC18 (DNA Plant Technologies, Oakland, CA) Tet^r Amp^r, pBR322 derivative containing pRSF1010 mob was modified by cloning the 2.8 kb PstI *sacBR* region from pSD800 (Gassmann *et al.*, 2000) into the unique PstI site. This pLVC18tcSacBR vector was converted into a Gateway-compatible destination vector by cloning the EcoRV ccdB cassette (Life Technologies) into the unique SmaI site.

Construction of Xam668ΔsuxC:
The 1.2 kb region upstream of the Xam668 *SuxC* gene was PCR amplified using forward and reverse oligonucleotide primers 5'- CACCGGTCGTTGAGGTATTTGCCG-3' and 5'- AAGCTTACTGTAACTTGCCACCGTTG-3', and cloned into pENTR/D-TOPO (Life Technologies) with a unique HindIII site on the 3' end. The 1.1 kb region downstream of the Xam668 *SuxC* gene was PCR amplified using forward and reverse oligonucleotide primers 5'- CACCAAGCTTCAATACCTAGCTGACATCGC-3' and 5'- AAGCTTACTGTAACTTGCCACCGTTG-3', and cloned into pENTR/D-TOPO (Life Technologies) with a unique HindIII site on the 5' end. A pENTR/D-TOPO encoding the concatenated 1.2 kb upstream + 1.1 kb downstream segments was constructed by cloning the HindIII-AscI downstream insert onto the HindIII-AscI C-terminal sites of the upstream pENTR/D-TOPO clone. The upstream-downstream entry clone was then introduced into the Gateway-compatible (Life Technologies) suicide destination vector pLVC18tcSacBRdes. The construct pLVC18tcSacBR(SuxC upstream + downstream) was then conjugated into Xam668 and single homologous recombinants were selected on tetracycline (10 μg/mL). After growing single homologous recombinants without tetracycline overnight, double homologous recombinant resolution events were selected on media supplemented with 5% sucrose, as the *sacBR* genes confer sensitivity to sucrose in gram negative bacteria (Logue *et al.*, 2009).

Construct for complementation of SuxC deletion mutant:
The Xam668 *SuxC* gene and promoter region was amplified by PCR using forward and reverse oligonucleotide primers 5'-CACCACGCTGGACTTGAGTTT-3' and 5'-GGTATT GCTAAGAATCTG-3' and ligated into pENTR/D-TOPO vector (Life Technologies). The *SuxC* fragment was transferred to pVSP61 broad host-range vector by LR recombination (Life Technologies). The pVSP61-*SuxC* construct was introduced into XamΔsuxC by bacterial conjugation alongside the *E. coli* pRK600 helper strain.

Characterization of SuxC mutant
Xam strains were grown in a 28°C shaker in 20 ml MME minimal medium (Arlat *et al.*, 1991) containing 20 mM sucrose or glucose and rifampicin (100 μg/mL). Following

inoculation at an OD_{600} = 0.1 from a washed starter culture, optical densities at 600 nm were measured every 10-20 hours over a period of 100 hours.

6-14. Cloning of *MeSWEET10a*

The *MeSWEET10a* full length cDNA was synthesized by GeneArt with unique restriction sites (NotI & AscI) flanking the coding sequence. The coding sequence was subsequently cloned into the pENTR/D-TOPO (Life Technologies) backbone by digestion with Not-HF and Asc-I (New England Biolabs) and ligation with T4 DNA ligase (Roche Rapid DNA Ligation Kit). The *MeSWEET10a* cDNA was mobilized into the mammalian expression vector pcDNA3.2/V5-DEST by LR recombination (Life Technologies).

6-15. Förster resonance energy transfer (FRET) analysis

Mammalian cell culture, transfection, image acquisition and FRET analysis were performed essentially as described (Takanaga & Frommer, 2010; Hou *et al.*, 2011). In short, HEK293T cells were grown in DMEM (Invitrogen) with 10% FBS, 50 U/mL penicillin and 50 μg/mL streptomycin in a 37°C incubator at 5% CO_2. For FRET imaging, cells were cultured on collagen-coated glass cover slips. Cells were transfected when the confluence of cells reached 35-40%. FLIPsuc90$\mu\Delta$1a (a FRET sucrose sensor modified from FLIPsuc90μ (Lager *et al.*, 2006) carrying full eCFP at the N-terminus and Aphrodite (Deuschle, 2006) at its C-terminus) in pcDNA3.1(-) and FLII^{12}Pglu700$\mu\partial$6 in pcDNA3.1(-) (Takanaga *et al.*, 2008) were used to monitor sucrose and glucose uptake activities respectively.

6-16. $TAL14_{Xam668}$ RNA-Seq

Strain Construction for RNA-Seq:

Xe($TAL14_{Xam668}$) was made by conjugating the $TAL14_{Xam668}$ complementation construct described above into *Xe* 85-10 by triparental mating using *E. coli* helper strain pRK600. Expression of $TAL14_{Xam668}$ was verified by western blot.

Plant Inoculations for RNA-Seq:

Xam and *Xe* grown on NYGA plates supplemented with rifampicin (100 μg/mL) or rifampicin and kanamycin (25 μg/mL) for *Xe*($TAL14_{Xam668}$) were re-suspended in 10 mM $MgCl_2$ at OD_{600} = 0.5. Abaxial nicks were made on leaves of cassava cultivar 60444 with a razor blade and culture was injected into the leaf via a 1mL needleless syringe. Mock infiltrations were done with 10 mM $MgCl_2$. Tissue was harvested and frozen 48 hours post inoculation.

Library preparation:
Total RNA from inoculated leaf tissue (30-50 mg) was extracted with the Spectrum Plant Total RNA Kit (Sigma-Aldrich), with an on-column DNaseI digestion step included. RNA quality was checked by Bioanalyzer (RNA pico chip, Functional Genomics Laboratory (FGL), UC Berkeley). RNA-Seq libraries were made using the TruSeq RNA sample preparation kit, v2, adapter set B (Illumina), starting with 1 μg total RNA. Library quality was assessed by Bioanalyzer (DNA 1000 chip, FGL, UC Berkeley). Quantification by qPCR and pooling of samples was carried out at UC Berkeley's Genomics Sequencing Laboratory (GSL). Libraries were sequenced at the GSL in a single lane of an Illumina HiSeq 2000, generating 100 bp paired-end reads.

Data Analysis:
$TAL14_{Xam668}$ RNA-Seq data analysis was carried out using the Galaxy platform (Blankenberg *et al.*, 2001; Giardine *et al.*, 2005; Goecks *et al.*, 2010). Reads were trimmed using Trim Galore with default settings (Cutadapt version 1.2.1) (Martin, 2011). Data quality was assessed using FastQC:Read QC (version 0.51) (Andrews, 2015). Trimmed reads were aligned to cassava reference genome AM560-2 version 4.1 by Tophat2 (version 0.5, mean inner distance 150, N mode, min intron 45 bp, max intron 5 kb, all other parameters set at default) (Trapnell *et al.*, 2009; Prochnik *et al.*, 2012). We allowed 2 mismatches per read to accommodate SNPs between our experimental cultivar TMS 60444 and the sequenced cultivar AM560-2. Transcript assembly was carried out by Cufflinks (version 0.0.5) using the reference annotation as a guide (Trapnell *et al.*, 2010). Quartile normalization, bias correction, and multi read correction were enabled. Cuffmerge (version 0.0.5) merged the cufflinks output from the various data sets (treatments) and differential gene expression analysis was carried out by Cuffdiff (version 0.0.5, false discovery rate 0.05, enabled quartile normalization, bias correction, and multi read correction) (Trapnell *et al.*, 2012). Results were filtered for $\log_2$(fold change) > 1.0 and fragments per kilobase of transcript per million mapped reads (FPKM) difference > 5 for genes activated by Xam668 versus mock infiltration, and by *Xe*(TAL14) versus *Xe* alone. Genes that were statistically significantly upregulated ($p<0.001$) in at least one of the two comparisons were considered.

6-17. Construction of designer TALEs (dTALEs) and complemented strains
The DNA binding domain of dTALEs were assembled as described in (Morbitzer *et al.*, 2010) and used to replace the RVDs of AvrBs3. For the *MeSWEET10a* promoter, we designed two distinct dTALEs with code optimized RVD sequences ($dTALE_{MeSWEET10a-3}$ and $dTALE_{MeSWEET10a-4}$) (Boch *et al.*, 2009; Moscou & Bogdanove, 2009). These dTALEs were designed to target sites distinct from the validated $TAL20_{Xam668}$ EBE in the *MeSWEET10a* promoter. For $TAL14_{Xam668}$-targeted promoters, we designed single dTALEs with code optimized RVD sequences. These dTALEs were designed to target sites distinct from the predicted $TAL14_{Xam668}$ EBEs in the target promoters. For all dTALEs, the ratio of the EBE score to the hypothetical best possible score is 1. dTALE

constructs driven by the *TAL20*$_{Xam668}$ promoter in pVSP61 were conjugated into Xam668ΔTAL20 or Xam668ΔTAL14ps using *E. coli* helper strain pRK600.

7. Supporting Information

Supplement A. *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) transcription activator-like (TAL) effector nucleotide and amino acid sequences

1. TAL effector nucleotide sequences described in this study

The DNA binding domains are highlighted in gray.

TAL13$_{Xam668}$

ATGGATCCCATTCGTCCGCGCACGCCAAGTCCTGCCCACGAACTTCTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCCGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCACCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTAATTCGATGTCTGCCTTCGGCGCTCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCAAGCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCTCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGGCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACACAGGTCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCG:GTGGAGGCAGTGCATGCATGGCGCAAT
GCACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGC
GCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCTGGACCAGGTGGTGGCCATCG
CCAGCAATAGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTG
ACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCC
GGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCCGGCGC
TGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCC
AGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGAC
CCCGGCGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGG
TGCTGTGCGAGCAACATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTG
GAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGCGCAGGTGGTGGCCATCGCCAG
CAATGGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCC
CGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTG
CTGCGCCAGGCCCATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCCGGCGCTGGA
GACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCTATCGCCAGCA
ATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCG
GACCAGGTGGTGGCTATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCT
GTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCCGGCGCTGGAGA
GCACTTTTGCCCAGTTATCTCGCCCTGATCAGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCCTGC
CTCGGCGGGCGTCCTGCGCTGGAGGCAGTGAGAAAGGGATTGCCGCACGCGCCGACCTTGATCAAAAGAACCAATCG
CCGTCTTCCCGAACGCACGTCCCATCGCGTTGCCGACCACGCGCAAGTGGCTCGCGTGCTGGGTTTTTTCCAGTGCC
ACTCCCACCCAGCGCAAGCATTTGATGAAGCCATGACGCAGTTCGGGATGAGCAGGCACGGGTTGTTACAGCTATTT
CGCAGAGTGGGCGTCACCGAACTCGAGGCCCACAGTGGAACGCTCCCCCCAGCCTCGCAGCGTTGGCACCGTATCCT
CCAGGCATCAGGGATGAAAAGGGCCGAACCGTCCGGTGCTTCCGCTCAAACGCCGGACCAGGCGTCTTTGCATGCAT
TCGCCGATGCGCTGGAGCGTGAGCTGGATGCGCCCAGCCCAATAGACCGGGCGGGCCAGGCGCTGGCAAGCAGCAGC
CGTAAACGGTCCCGATCGGAGAGTTCTGTCACCGGCTCCTTCGCACAGCAAGCTGTCGAGGTGCGCGTTCCCGAACA
GCGCGATGCGCTGCATTTCCTCCCCCTCAGCTGGGGTGTAAAACGCCCGCGTACCAGGATCGGGGGCGGCCTCCCGG
ATCCTGGTACGCCCATGGACGCCGACCTGGCACCGTCCAGCACCGTGATGTGGGAACAAGATGCTGACCCCTTCGCA
GGGGCAGCGGATGATTTTCCGGCATTCAACGAAGAGGAGATGGCATGGTTGATGGAGCTATTTCCTCAGTGA

TAL14$_{Xam668}$

ATGGATCCCATTCGTCCGCGCACGCCAAGTCCTGCCCACGAACTTCTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCCGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCACCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTAATTCGATGTCTGCCTTCGGCGCTCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCAAGCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCTCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGGCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACACAGGTCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCGGTGGAGGCAGTGCATGCATGGCGCAATG
CACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCTGGACCAGGTGGTGGCCATCGC
CAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGA
CCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCCA
GCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACC
CCGGCGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGGT
GCTGTGCGAGCAACATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGC
AATAGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCC
GGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC
TGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGG
CGCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCTATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATG
GCGGCGGCAAGCCGGCGCTGGAGAGCACTTTTGCCCAGTTATCTCGCCCTGATCAGGCGTTGGCCGCGTTGACCAAC
GACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGAGGCAGTGAGAAAGGGATTGCCGCACGCGCC
GACCTTGATCAAAAGAACCAATCGCCGTCTTCCCGAACGCACGTCCCATCGCGTTGCCGACCACGCGCAAGTGGCTC
GCGTGCTGGGTTTTTTCCAGTGCCACTCCCACCCAGCGCAAGCATTTGATGAAGCCATGACGCAGTTCGGGATGAGC
AGGCACGGGTTGTTACAGCTATTTCGCAGAGTGGGCGTCACCGGACTCGAGGCCCGCAGTGGAACGCTCCCCCCAGC
CCCGCAGCGTTGGCACCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCGAACCGTCCGGTGCTTCGGCTCAAACGC
CGGACCAGGCGTCTTTGCATGCATTCGCCGATGCGCTGGAGCGTGAGCTGGATGCGCCCAGCCCAATAGACCGGGCG
GGCCAGGCGCTGGCAAGCAGTAGCCGTAAACGGTCCCGATCGGAGAGTTCTGTCACCGGCTCCTTCGCACAGCAAGC
TGTCGAGGTGCGCGTTCCCGAACAGCGCGATGCGCTGCATTTACCCCCCCTCAGCTGGGGTGTAAAACGCCCGCGTA
CCAGGATCGGGGGCGGCCTCCCGGATCCTGGTACACCCATGGACGCCGACCTGGCAGCGTCCAGCACCGTGATGTGG
GAACAAGATGCTGACCCCTTCGCAGGGGCAGCGGATGATTTTCCGGCATTCAACGAAGAGGAGATGGCATGGTTGAT
GGAGCTATTTCCTCAGTGA

TAL15$_{Xam668}$

ATGGATCCCATTCGTCCGCGCACGTCAAGTCCTGCCGACGAACTTTTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCTGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCCCCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTGATTCGATGTCTGCCTTCGGCGCGCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCACCCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCTCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGGCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACCCAGGTCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCGGTGGAGGCAGTGCATGCATGGCGCAATG
CACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCTGGACCAGGTGGTGGCCATCGC
CAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGTCTGA
CCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCCA
GCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACC
CCGGCGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGGT
GCTGTGCGAGCAACATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGC
AATAGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCC
GGCGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC
TGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGG
ACCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATG
GCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCCGGCG
CAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGAGCATTTTTGCCCAGTTATCTCGCCCTGA
TCAGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGAGGCAG
TGAAAAAGGGATTGCCGCACGCGCCGACCTTGATCAAAAGAACCAATCGCCGTCTTCCCGAACGCACGTCCCATCGC
GTTGCCGACCACGCGCAAGTGGCTCGCGTGTTGGGTTTTTTCCAGTGCCACTCCCACCCAGCGCAAGCATTTGATGA
AGCCATGACGCAGTTCGGGATGAGCAGGCACGGGTTGTTACAGCTATTTCGCAGAGCCGGCGTCACCGAACTCGAGG
CCCACAGTGGAACGCTCCCCCCAGCCTCGCAGCGTTGGCACCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCGAA
CCGTCCGGTGCTTCCGCTCAAACGCCGGACCAGGCGTCTTTGCATGCATTCGCCGATGCGCTGGAGCGTGAGCTGGA
TGCGCCCAGCCCAATAGACCGGGCGGGCCAGGCGCTGGCAAGCAGCAGCCGTAAACGGTCCCGATCGGAGAGTTCTG
TCACCGGCTCCTTCGCACAGCAAGCTGTCGAGGTGCGCGTTCCCGAACAGCGCGATGCGCTGCATTTCCTCCCCCTC
AGCTGGGGTGTAAAACGCCCGCGTACCAGGATCGGGGGCGGCCTCCCGGATCCTGGTACGCCCATGGACGCCGACCT
GGCACCGTCCAGCACCGTGATGTGGGAACAAGATGCTGACCCCTTCGCAGGGGCAGCGGATGATTTTCCGGCATTCA
ACGAAGAGGAGATGGCATGGTTGATGGAGCTATTTCCTCAGTGA

TAL20$_{Xam668}$

ATGGATCCCATTCGTCCGCGCACGCCAAGTCCTGCCCACGAACTTCTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCCGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCACCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTAATTCGATGTCTGCCTTCGGCGCTCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCAAGCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCTCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGGCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACACAGGTCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCGGTGGAGGCAGTGCATGCATGGCGCAATG
CACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCTGGACCAGGTGGTGGCCATCGC
CAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGA
CCCCGGACCAGGTGGTGGCTATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCTATCGCCA
GCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACC
CCGGCGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGGT
GCTGTGCGAGCAACATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGC
AATAGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCC
GGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC
TGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGG
AGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
YGCCAGSCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATA
ACGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGAC
CAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTG
CGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGG
TGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCCACGAT
GGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCA
GGTGGTGGCTATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCG
AGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTG
CAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATGGCGG
CGGCAGGCCGGCGCTGGAGAGCATTTTTGCCCAGTTATCTCGCCCTGATCAGGCGTTGGCCGCGTTGACCAACGACC
ACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGAGGCAGTGAAAAAGGGATTGCCGCACGCGCCGACC
TTGATCAAAAGAACCAATCGCCGTCTTCCCGAACGCACGTCCCATCGCGTTGCCGACCACGCGCAAGTGGCTCGCGT
GTTGGGTTTTTTCCAGTGCCACTCCCACCCAGCGCAAGCATTTGATGAAGCCATGACGCAGTTCGGGATGAGCAGGC
ACGGGTTGTTACAGCTATTTCGCAGAGCCGGCGTCACCGAACTCGAGGCCCACAGTGGAACGCTCCCCCCAGCCTCG
CAGCGTTGGCACCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCGAACCGTCCGGTGCTTCCGCTCAAACGCCGGA
CCAGGCGTCTTTGCATGCATTCGCCGATGCGCTGGAGCGTGAGCTGGATGCGCCCAGCCCAATAGACCGGGCGGGCC
AGGCGCTGGCAAGCAGCAGCCGTAAACGGTCCCGATCGGAGAGTTCTGTCACCGGCTCCTTCGCACAGCAAGCTGTC
GAGGTGCGCGTTCCCGAACAGCGCGATGCGCTGCATTTCCTCCCCCTCAGCTGGGGTGTAAAACGCCCGCGTACCAG
GATCGGGGGCGGCCTCCCGGATCCTGGTACGCCCATGGACGCCGACCTGGCACCGTCCAGCACCGTGATGTGGGAAC
AAGATGCTGACCCCTTCGCAGGGGCAGCGGATGATTTTCCGGCATTCAACGAAGAGGAGATGGCATGGTTGATGGAG
CTATTTCCTCAGTGA

TAL22$_{Xam668}$

ATGGATCCCATTCGTCCGCGCACGTCAAGTCCTGCCGACGAACTTTTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCTGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCCCCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTGATTCGATGTCTGCCTTCGGCGCGCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCACCCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCGCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGTCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCGGTAGAGGCAGTGCATGCATGGCGCAATG
CACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGC
CAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGA
CCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCA
GCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACC
CCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGT
GCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGC
CACGATGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCC
GGCGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC
TGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGACCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGG
ACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACG
ATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGCG
CAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTG
CGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGG
TGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGACCAGGTGGTGGCCATCGCCAGCCACGAT
GGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGCGCA
GGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCG
AGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTG
CAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGATCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGG
CGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGC
CTGATCCCGGCGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTT
GCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGG
CGCTGGAGAGCATTTTTGCCCAGTTATCTCGCCCTGATCAGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCC
TTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGAGGCAGTGAAAAAGGGATTGCCGCACGCGCCGACCTTGATCAAAAG
AACCAATCGCCATCTTCCCGAACGCACGTCCCATCGCGTTGCCGACCACGCGCAAGTGGCTCGCGTGTTGGGTTTTT
TCCAGTGCCACTCCCACCCAGCGCAAGCATTTGATGAAGCCATGACGCAGTTCGGGATGAGCAGGCACGGGTTGTTA
CAGCTATTTCGCAGAGCCGGCGTCACCGAACTCGAGGCCCGCAGTGGAACGCTCCCCCCAGCCCCGCAGCGTTGGCA
CCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCGAACCGTCCGGTGCTTCGGCTCAAACGCCGGACCAGGCGTCTT
TGCATGCATTCGCCGATGCGCTGGAGCGTGAGCTGGATGCGCCCAGCCCAATAGACCGGGCGGGCCAGGCGCTGGCA
AGCAGCAGCCGTAAACGGTCCCGATCGGAGAGTTCTGTCACCGGCTCCTTCGCACAGCAAGCTGTCGAGGTGCGCGT
TCCCGAACAGCGCGATGCGCTGCATTTACCCCCCCTCAGCTGGGGTGTAAAACGCCCGCGTACCAGGATCGGGGGCG
GCCTCCCGGATCCTGGTACACCCATGGACGCCGACCTGGCAGCGTCCAGCACCGTGATGTGGGAACAAGATGCGGAC
CCCTTCGCAGGGGCAGCGGATGATTTTCCGGCATTCAACGAAGAGGAGATGGCATGGTTGATGGAGCTATTTCCTCA
GTGA

TAL14$_{ClO151}$

ATGGATCCCATTCGTCCGCGCACGCCAAGTCCTGCCCACGAACTTCTGGCCGGACCCCAGCCGGATAGGGTTCAGCC
GCAGCCGACTGCAGATCGTGGGGGGGCTCCGCCTGCCGGCAGCCCCCTGGATGGCTTGCCCGCTCGACGGACGATGT
CCCGAACCCGTCTCCCGTCTCCCCCTGCACCCTTGCCTGCGTTCTCAGCGGGCAGTTTCAGCGATCTGCTCCGTCAG
TTCGATCCGTCGCTTCTTGATACATCGCTTTTTAATTCGATGTCTGCCTTCGGCGCTCCTCATACAGAGGCTGCCTC
AGGAGAGGGGGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCAAGCCACCGTGCAGGTCGCTGTGACGG
CCGCGCGACCGCCGCGCGCCAAGCCGGCGCCGCGACGGCGTGCTGCGCACACCTCTGACGCTTCGCCGGCCGGGCAG
GTCGATCTATGCACGCTCGGCTACAGCCAGCAGCAGCAAGAGAAGATCAAACTGAAGGCGCGTTCGACAGTAGCACA
GCACCACGAGGCACTGATCGGCCATGGGTTTACACGTGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCCTTAG
GGACCGTCGCTGTCAAGTACCAGGCCATGATCGCGGCGTTGCCGGAGGCGACACACGAAGACATCGTTGGCGGCGGC
AAACAGTGGTCCGGCGCACGCGCCCTGGAAGCATTGCTCACGGTGTCGGGAGAGTTGAGAGGTCCACCGTTACAGTT
GGACACAGGTCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCGGTGGAGGCAGTGCATGCATGGCGCAATG
CACTGACGGGCGCTCCCCTGAACCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCTGGACCAGGTGGTGGCCATCGC
CAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGA
CCCCGGACCAGGTGGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCCGGACCAGGTGGTGGCCATCGCCA
GCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACC
CCGGCGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCAGCTGTTGCCGGT
GCTGTGCGAGCAACATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGCCAGCAATAGCGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGC
AATAGCGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGTCTGACCCC
GGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC
TGTGCGAGCAACATGGCCTGACCCGGGCGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGCGCCAGGCCCATGGCCTGACCCCGG
CGCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCTATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCGAGCAACATGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATG
GCGGCGGCAAGCCGGCGCTGGAGAGCACTTTTGCCCAGTTATCTCGCCCTGATCAGGCGTTGGCCGCGTTGACCAAC
GACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGAGGCAGTGAGAAAGGGATTGCCGCACGCGCC
GACCTTGATCAAAAGAACCAATCGCCGTCTTCCCGAACGCACGTCCCATCGCGTTGCCGACCACGCGCAAGTGGCTC
GCGTGCTGGGTTTTTTCCAGTGCCACTCCCACCCAGCGCAAGCATTTGATGAAGCCATGACGCAGTTCGGGATGAGC
AGGCACGGGTTGTTACAGCTATTTCGCAGAGTGGGCGTCACCGGACTCGAGGCCCGCAGTGGAACGCTCCCCCCAGC
CCCGCAGCGTTGGCACCGTATCCTCCAGGCATCAGGGATGAAAAGGGCCGAACCGTCCGGTGCTTCGGCTCAAACGC
CGGACCAGGCGTCTTTGCATGCATTCGCCGATGCGCTGGAGCGTGAGCTGGATGCGCCCAGCCCAATAGACCGGGCG
GGCCAGGCGCTGGCAAGCAGCAGCCGTAAACGGTCCCGATCGGAGAGTTCTGTCACCGGCTCCTTCGCACAGCAAGC
TGTCGAGGTGCGCGTTCCCGAACAGCGCGATGCGCTGCATTTACCCCCCCTCAGCTGGGGTGTAAAACGCCCGCGTA
CCAGGATCGGGGGCGGCCTCCCGGATCCTGGTACACCCATGGACGCCGACCTGGCAGCGTCCAGCACCGTGATGTGG
GAACAAGATGCTGACCCCTTCGCAGGGGCAGCGGATGATTTTCCGGCATTCAACGAAGAGGAGATGGCATGGTTGAT
GGAGCTATTTCCTCAGTGA

2. TAL effector amino acid sequences described in this study

The DNA binding domain repeats are shown with Repeat Variable Diresidues (RVDs) highlighted in gray.

TAL13$_{Xam668}$

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGG
KQWSGARALEALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTLDQVVAIASNSGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKPALETVQRLLPVLCEQHG
LTRAQVVAIASNGGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNIGGKQALETVQRLLPVLCEQHG
LIPAQVVAIASNGGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASNNGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKPALE

STFAQLSRPDQALAALTNDHLVALACLGGRPALEAVRKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTELEAHSGTLPPASQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLP
DPGTPMDADLAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

TAL14$_{Xam668}$

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGG
KQWSGARALEALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTLDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTRAQVVAIASNGGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNSGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTRAQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASNNGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKPALE

STFAQLSRPDQALAALTNDHLVALACLGGRPALEAVRKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTGLEARSGTLPPAPQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLP
DPGTPMDADLAASSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

TAL15$_{Xam668}$

MDPIRPRTSSPADELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFDSMSAFGAPHTEAASGEGDEVQSGLRAADDPHPTVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGG
KQWSGARALEALLTVSGELRGPPLQLDPGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTLDQVVAIASNGGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTRAQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNSGGKQALETVQRLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNSGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPAQVVAIASNGGGRPALE

SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEAHSGTLPPASQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLP
DPGTPMDADLAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

TAL20$_{Xam668}$

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGG
KQWSGARALEALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTLDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNSGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKPALETVQRLLPVLCEQHG
LTRAQVVAIASNGGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPAQVVAIASNGGGRPALE

SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEAHSGTLPPASQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHFLPLSWGVKRPRTRIGGGLP
DPGTPMDADLAPSSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

TAL22$_{Xam668}$

MDPIRPRTSSPADELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFDSMSAFGAPHTEAASGEGDEVQSGLRAADDPHPTVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGVG
KQWSGARALEALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKPALETVQRLLPVLCEQHG
LIPAQVVAIASNGGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LIPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LIPAQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LIPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LIPAQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LIPDQVVAIASHDGGKQALETVQRLLPVLCEQHG
LIPAQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGRPALE

SIFAQLSRPDQALAALTNDHLVALACLGGRPALEAVKKGLPHAPTLIKRTNRHLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRAGVTELEARSGTLPPAPQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLP
DPGTPMDADLAASSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

TAL14$_{CIO151}$

MDPIRPRTPSPAHELLAGPQPDRVQPQPTADRGGAPPAGSPLDGLPARRTMSRTRLPSPPAPLPAFSAGSFSDLLRQ
FDPSLLDTSLFNSMSAFGAPHTEAASGEGDEVQSGLRAADDPQATVQVAVTAARPPRAKPAPRRRAAHTSDASPAGQ
VDLCTLGYSQQQQEKIKLKARSTVAQHHEALIGHGFTRAHIVALSQHPAALGTVAVKYQAMIAALPEATHEDIVGGG
KQWSGARALEALLTVSGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTLDQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASHDGGKQALETVQQLLPVLCEQHG
LTPAQVVAIASNSGGKQALETVQRLLPVLRQAHG
LTPDQVVAIASNSGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCEQHG
LTRAQVVAIASNGGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASHDGGKQALETVQRLLPVLRQAHG
LTPAQVVAIASNNGGKPALETVQRLLPVLCEQHG
LTPDQVVAIASNIGGKQALETVQRLLPVLCEQHG
LTPDQVVAIASNGGGKPALE

STFAQLSRPDQALAALTNDHLVALACLGGRPALEAVRKGLPHAPTLIKRTNRRLPERTSHRVADHAQVARVLGFFQC
HSHPAQAFDEAMTQFGMSRHGLLQLFRRVGVTGLEARSGTLPPAPQRWHRILQASGMKRAEPSGASAQTPDQASLHA
FADALERELDAPSPIDRAGQALASSSRKRSRSESSVTGSFAQQAVEVRVPEQRDALHLPPLSWGVKRPRTRIGGGLP
DPGTPMDADLAASSTVMWEQDADPFAGAADDFPAFNEEEMAWLMELFPQ-

Supplement B. Supplemental Tables

Table S1. *Xanthomonas axonopodis* pv. *manihotis* (*Xam*) strains utilized in this work.

Strain	Country of collection	Year of collection
Xam668	Indonesia	1978
CIO151	Colombia	1995
CFBP1851	Colombia	1974
UA306	Colombia	2008
IBSBF278	Brazil	1965
UG27	Uganda	2011
UG12	Uganda	2011
UG21	Uganda	2011
UG28	Uganda	2011
UG23	Uganda	2011
NG1	Nigeria	2011

Table S2. RT-PCR primers used in this study with corresponding PCR cycle number.

Primer	Sequence (5'-3')	Gene	PCR cycle number
022805 F	GGCGAATGATCCCAAAGCTG	*cassava4.1_022805*	28
022805 R	GGATGTTGAGTAGAGACACCCT		
007568 F2	GGTGGAAACAAGTACAGGCAT	*cassava4.1_007568*	25-26
007586 R2	TTGGCCAAATCTAATTCATCAG		
026121 F	AAATCCAATGCCAGCAGCAC	*cassava4.1_026121*	26
026121 R	GCAGCTTCTCCTTCTTGGGT		
024404 F	GGGGATTCCAGCTCTCCTTG	*cassava4.1_024404*	26
024404 R	TGGGGCTGCTCCAAATTCTT		
020743 F	CCAGAATCTTTGGGCGTGGA	*cassava4.1_020743*	28
020743 R	ACCACGTTGAATACGCCACA		
007516 F	AGACAGCTGCAAGCCAATG	*cassava4.1_007516*	26
007516 R	GGCCAAATTTAATTGATTAGTAAG		
033289 F	TGTCTTCGGGACTTTTCTGCT	*cassava4.1_033289*	26
033289 R	GCTCGAGGGGAGATCGTCAT		

Table S2. (continued)

031361 F	AGCAAGACGAAGATGGGGAC	*cassava4.1_031361*	26
031361 R	GGCAGCTTCGAAAAGTTGGT		
023665 F	AACAAGACTTGGGCGGATCA	*cassava4.1_023665*	26
023665 R	ATCAGTAGTCGCCTGCTTCG		
001042 F	ATGTAGGGAGCAATCTGAGAGA	*cassava4.1_001042*	30
001042 R	CAACTGGAGCAATATTTCCTGG		
034150 F3	GTTGTGGAGTATGGTCCCG	*cassava4.1_034150*	23
034150 R3	CAACGCATAAGTGACCTAGAA		
022871 F	CGAGTTGGGTGTTGCAGGC	*cassava4.1_022871*	26
022871 R	GATGATTAGGGATTCTCAAGACCTCG		
011345 F	CTGCCACATAGCCTCTGCC	*cassava4.1_011345*	26
011345 R	GTAAACTGCATCCTAAGCCGG		
012090 F	TTGGCCTGAAGGAAACGAGG	*cassava4.1_012090*	26
012090 R	GGGACTCATGCGTCTGCTTTG		
022534 F	AGAGGCTTCATTTGGGGAGC	*cassava4.1_022534*	30
022534 R	TCAGCAGCAATGGTTTTGCC		
016646 F	CTACCTCCTCCCTGTCAACG	*cassava4.1_016646*	28
016646 R	TACAACTTGTCCCGCTGTTAAT		
020499 F	GTGGAGGAAATTGCGGTTGC	*cassava4.1_020499*	28
020499 R	AGAGCCAGAGTCTCCTTGGT		
026299 F	AACCGCTTGCAGCATAACCA	*cassava4.1_026299*	28
026299 R	GTGAGGGATGGAGCTTGCTG		
019005 F	CTCTTCATCGTCTCGCTGTG	*cassava4.1_019005*	26
019005 R	CCCAGTATTTCTCCGACTGGG		
025591 F	ACCAAAGCCACCGTTCAAGA	*cassava4.1_025591*	26
025591 R	CCGCCTTCCTCTCTTCCTTC		

Table S2. (continued)

015102 F	CTGGCGTCATTTTGCCTGG	*cassava4.1_015102*	26
015102 R	GCATATCATGGGGTGCCTCC		
009347 F	CTGCAGCTCTGGGAACCATT	*cassava4.1_009347*	23
009347 R	CAATGCTGGTGGTTGTGAGC		
026646 F	GTTGTGGAGGATCTGGGGTG	*cassava4.1_026646*	28
026646 R	TTGTCCTCCAGTTGATGCCC		
024542 F	TGCATACTGCCCCATCCATT	*cassava4.1_024542*	28
024542 R	CGTCTTCAATGTGGCAGGGA		
011524 F	CCGATCCACGTTGAGAATGGC	*cassava4.1_011524*	26
011524 R	GAATTGCATTTTGACGCCGC		
023036 F2	ATTGTGGAGTATGGTGCCA	*cassava4.1_023036*	26
023036 R2	CGCATAAGTGACTCAGAC		
013474 F	TCCTCACCTTGACTGCGGTG	*cassava4.1_013474*	25-26
013474 R	AGCACCATCTGGACAATCCCA		
017922 F	ATGGGCTGCTTCTGGGATTC	*cassava4.1_017922*	25-26
017922 R	TATACACATCACGGCCACGC		
018388 F	TTGACGCTTTGGAGAAGGCT	*cassava4.1_018388*	25-26
018388 R	AACCCAACCAGCTGCAATCT		
019269 F	ATCTCTGCTGCTACCACGAC	*cassava4.1_019269*	25-26
019269 R	GTTTGGTCCGCTGGGAACTT		
020556 F	GCGTCTCTTCCCAGCTGTTT	*cassava4.1_020556*	25-26
020556 R	GCACCTTTGCAATCTCCACC		
027342 F	GAGCCATTTCTTTCGCGGAC	*cassava4.1_027342*	23
027342 R	AGTACGGACTCCCAAACTGC		
029452 F	AGAGCTGTCCTTGACGCAAA	*cassava4.1_029452*	28
029452 R	GGCCTATAACGCAGTGCCAT		

Table S2. (continued)

030094 F	TGATGGAGGTGGTGGCTCTA	*cassava4.1_030094*	30
030094 R	TCAAGCCCAGGTTGATCCTC		
009922 F	TCCCTGGTGCTTTGGTCATC	*cassava4.1_009922*	30
009922 R	ATTGGGTGCTTGTTGTCCGA		
009807 F (actin)	ACAGTGTCTGGATCGGAGGATC	*cassava4.1_009807*	25-26
009807 R (actin)	GAAGCACTTCCTGTGGACGATG		
014124 F	AGATGCTGCTGTATGCAGTGT	*cassava4.1_014124*	28
014124 R	GGAACCTTTATGGCTCGAGGA		

Table S3. TAL14$_{Xam668}$ TALE-NT (2.0) predicted effector binding elements (EBEs).

gene name	predicted EBE	score	start position	group
cassava4.1_007516	T ATAAAAAGGTCGGT	14.88	165	1
cassava4.1_007568	T ATAAAAAGGTCGGT	14.88	163	1
cassava4.1_009347	T ATAAAAAGGCCAGT	17.08	213	1
cassava4.1_011345	T ATAAACCCAACAAT	12.21	218	1
cassava4.1_011345	T AAATTCAAATCATA	15.33	38	1
cassava4.1_011345	T ATATATAAACCCAA	17.93	214	1
cassava4.1_011345	T GTGATCCAACCACT	18.16	101	1
cassava4.1_011345	T ACAATACTATCTAA	18.48	11	1
cassava4.1_011345	T ATGGTCCCCAAACT	18.74	142	1
cassava4.1_011524	T ATAAAAAGGTCAAA	13.56	148	1
cassava4.1_011524	T CAAATAAAAGCGAT	15.14	158	1
cassava4.1_011524	T AAAATCCAAAATTT	18.77	68	1
cassava4.1_019005	T ATAGAAGGGGCGAT	15.62	165	1
cassava4.1_024404	T ATAAACCAGTCAAG	11.89	231	1
cassava4.1_024404	T TTAATTAAAACAAT	14.74	161	1
cassava4.1_024404	T TTATTTAAATAAAT	16.9	86	1
cassava4.1_024404	T ATAGAAGGCTTAAT	17.12	118	1
cassava4.1_024404	T GCAATGGAGTCTAT	18.13	217	1
cassava4.1_024404	T CAAAACCTGTCAAA	18.16	265	1
cassava4.1_026299	T ATAAAAATGTCAAT	13.32	155	1
cassava4.1_033289	T ATAAAAGGGACAAC	17.03	27	1
cassava4.1_034150	T AAATTAAAATAAAT	14.47	36	1
cassava4.1_034150	T ATATACCAACCAAA	14.94	262	1
cassava4.1_034150	T ATAAAAGGGTACAT	16.2	233	1
cassava4.1_034150	T TTAGACAAAAGAAT	16.83	83	1
cassava4.1_034150	T AAAATTAAATGAAG	17.75	147	1
cassava4.1_001042	T ATAAGAAAGTTAAC	16.34	218	2
cassava4.1_001042	T CCCAACAAGAAAAT	18.19	151	2
cassava4.1_001042	T ATAGACGAAGAACA	18.6	184	2
cassava4.1_001042	T AGAATCTCGTCAGA	18.62	129	2

Table S3. (continued)

cassava4.1_012090	T TCAGTCAAATCCAC	14.15	125	2
cassava4.1_012090	T ACCAAGAAAACAAT	17.56	246	2
cassava4.1_012090	T ATAGTAAGATATCA	17.77	274	2
cassava4.1_012090	T TTATTCATATTAAT	17.79	29	2
cassava4.1_012090	T GTAATAGTATCACT	18.53	192	2
cassava4.1_015102	T TTCTTCAAATCCTT	17.74	71	2
cassava4.1_015102	T ATATAGATTTCAAT	18.41	250	2
cassava4.1_016646	T ATAATCTTCTCGCT	15.1	230	2
cassava4.1_020743	T ATAATAAGACCAAT	11.76	197	2
cassava4.1_020743	T ATATTTAAATAAAT	13.38	43	2
cassava4.1_020743	T ATTATCCACTCAAA	13.46	156	2
cassava4.1_020743	T AAATTAAAATATAT	16.31	33	2
cassava4.1_020743	T ATTTAAAAATAAAT	17.98	23	2
cassava4.1_020743	T AAAAATAAATTAAA	18.47	27	2
cassava4.1_022805	T ATGTTCCAATATAT	16.61	220	2
cassava4.1_022805	T AAAATTAATCAAAT	17.39	51	2
cassava4.1_022805	T ATAACAGAAGCACA	18.8	79	2
cassava4.1_022871	T GTAAAAAAATAAAT	16.14	148	2
cassava4.1_022871	T CAAATCATAACAAA	17.24	89	2
cassava4.1_022871	T AAAATAAAATTTAA	17.39	110	2
cassava4.1_023036	T ATAAAAGAGTACAT	14.74	233	2
cassava4.1_023036	T ATATACCAACCAAA	14.94	262	2
cassava4.1_023036	T ATAAACAAAATCAA	16.09	121	2
cassava4.1_023036	T ACAATAACTTCTCT	17.14	39	2
cassava4.1_023665	T ATAATAATATTCAT	15.62	128	2
cassava4.1_023665	T AAAAATAAAAAAAT	17.55	230	2
cassava4.1_023665	T TTAATAACCAAAAT	18.11	88	2
cassava4.1_023665	T ATAAAAGGGTAGGT	18.66	251	2
cassava4.1_024542	T TCATTAAAGGCAAT	18.45	139	2
cassava4.1_024542	T CTTATCTAACCAAA	18.75	178	2
cassava4.1_025591	T ATAATCAAAACACC	12.02	231	2
cassava4.1_025591	T GAAGTCGGACCAAA	18.22	5	2
cassava4.1_026121	T ATAATCATATCGAT	8.99	208	2
cassava4.1_026121	T ACCTCTCAATCAAT	18.22	225	2
cassava4.1_026646	T ATAATAACAAAGAA	15.54	168	2
cassava4.1_026646	T ACAGACAAAACATC	17.3	133	2
cassava4.1_026646	T ATATTTACTTCCTT	18.38	281	2
cassava4.1_031361	T ATAATAAGATCAGT	13.27	122	2
cassava4.1_031361	T ATAGACATAACTTT	18.35	156	2
cassava4.1_020499	T CTGGTCAACTCAGC	17.7	23	2
cassava4.1_020499	T CTCCTCAACTGCCT	18.2	200	2
cassava4.1_020499	T ATAAAAGGAGCAAC	18.22	214	2

Table S4. TAL14$_{Xam668}$ TALgetter predicted effector binding elements (EBEs).

gene name	predicted EBE	score	p-value	position	group
cassava4.1_007516	TATAAAAAGGTCGGT	-14.789	2.13E-04	163	1
cassava4.1_007568	TATAAAAAGGTCGGT	-14.789	2.13E-04	161	1
cassava4.1_009347	TTTTTTCTCGTCTAT	-17.804	2.34E-03	199	1
cassava4.1_009347	TATAAAAAGGCCAGT	-18.486	3.76E-03	211	1
cassava4.1_011345	TATAAACCCAACAAT	-15.095	2.80E-04	216	1
cassava4.1_011345	TAAATTCAAATCATA	-17.536	1.94E-03	36	1
cassava4.1_011345	CATACTTTTACCGAT	-18.769	4.53E-03	47	1
cassava4.1_011345	TATGGTCCCCAAACT	-19.359	6.63E-03	140	1
cassava4.1_011345	TGTGATCCAACCACT	-19.572	7.57E-03	99	1
cassava4.1_011345	TACAATACTATCTAA	-19.828	8.87E-03	9	1
cassava4.1_019005	TATAGAAGGGGCGAT	-15.317	3.38E-04	163	1
cassava4.1_019005	CACCGTCTTGTATAT	-19.687	8.13E-03	151	1
cassava4.1_024404	TATAAACCAGTCAAG	-13.42	5.95E-05	229	1
cassava4.1_024404	TTTAATTAAAACAAT	-16.567	9.33E-04	159	1
cassava4.1_024404	TTTATTTAAATAAAT	-17.69	2.16E-03	84	1
cassava4.1_024404	TATAGAAGGCTTAAT	-18.533	3.88E-03	116	1
cassava4.1_024404	ACCTGTCAAATCCAC	-19.238	6.15E-03	268	1
cassava4.1_024404	TGCAATGGAGTCTAT	-19.406	6.83E-03	215	1
cassava4.1_024404	TCAAAACCTGTCAAA	-19.528	7.37E-03	263	1
cassava4.1_026299	TATAAAAATGTCAAT	-13.67	7.72E-05	153	1
cassava4.1_026299	AATTCTCATGTGAAT	-19.236	6.14E-03	165	1
cassava4.1_026299	TCTGGTAAAAAGGAT	-19.886	9.18E-03	282	1
cassava4.1_033289	TATAAAAGGGACAAC	-19.654	7.97E-03	25	1
cassava4.1_034150	TAAATTAAAATAAAT	-16.463	8.58E-04	34	1
cassava4.1_034150	TATAAAAGGGTACAT	-17.374	1.72E-03	231	1
cassava4.1_034150	TATATACCAACCAAA	-17.76	2.27E-03	260	1
cassava4.1_034150	TTTAGACAAAAGAAT	-18.994	5.26E-03	81	1
cassava4.1_034150	AACTTGCAAGTCAAT	-19.187	5.96E-03	183	1
cassava4.1_026121	TATAATCATATCGAT	-9.029	4.80E-07	206	2
cassava4.1_016646	TATAATCTTCTCGCT	-13.321	7.66E-05	228	2
cassava4.1_015102	AATAATCTCTTCGAT	-14.016	1.48E-04	176	2
cassava4.1_020743	TATAATAAGACCAAT	-14.295	1.88E-04	195	2
cassava4.1_012090	TTCAGTCAAATCCAC	-14.312	1.90E-04	123	2
cassava4.1_025591	TATAATCAAAACACC	-14.617	2.46E-04	229	2
cassava4.1_020743	TATTATCCACTCAAA	-14.618	2.47E-04	154	2
cassava4.1_020743	TATATTTAAATAAAT	-14.848	3.01E-04	41	2
cassava4.1_025591	CATATCCAAATCGAA	-14.996	3.43E-04	254	2
cassava4.1_031361	TATAATAAGATCAGT	-15.322	4.51E-04	120	2
cassava4.1_023036	TATAAAAGAGTACAT	-15.753	6.36E-04	231	2
cassava4.1_023665	TATAATAATATTCAT	-16.212	9.17E-04	126	2
cassava4.1_022805	TATGTTCCAATATAT	-17.171	1.88E-03	218	2
cassava4.1_020499	CATAGTTGTTTCAAT	-17.259	2.00E-03	243	2
cassava4.1_015102	TTTCTTCAAATCCTT	-17.523	2.41E-03	69	2
cassava4.1_026646	TATAATAACAAAGAA	-17.598	2.54E-03	166	2
cassava4.1_012090	TTTATTCATATTAAT	-17.696	2.72E-03	27	2
cassava4.1_020743	TAAATTAAAATATAT	-17.747	2.81E-03	31	2
cassava4.1_023036	TATATACCAACCAAA	-17.76	2.84E-03	260	2
cassava4.1_022805	ATTAATCAAATTGCT	-17.779	2.88E-03	53	2
cassava4.1_001042	TATAAGAAAGTTAAC	-17.802	2.93E-03	216	2

Table S4. (continued)

cassava4.1_020499	TCTGGTCAACTCAGC	-17.898	3.12E-03	21	2
cassava4.1_022871	TGTAAAAAAATAAAT	-18.091	3.56E-03	146	2
cassava4.1_023036	TACAATAACTTCTCT	-18.278	4.04E-03	37	2
cassava4.1_031361	TATAGACATAACTTT	-18.286	4.06E-03	154	2
cassava4.1_026646	TATATTTACTTCCTT	-18.33	4.18E-03	279	2
cassava4.1_024542	TTCATTAAAGGCAAT	-18.475	4.60E-03	137	2
cassava4.1_016646	GATCGGCTCATCAAT	-18.581	4.92E-03	214	2
cassava4.1_020499	CAGCGTCCAGAAAAT	-18.603	4.99E-03	32	2
cassava4.1_020743	AATTGGCCAATCAAA	-18.717	5.37E-03	129	2
cassava4.1_001042	TAGAATCTCGTCAGA	-18.818	5.73E-03	127	2
cassava4.1_031361	AATAAATAAATAAAT	-18.852	5.85E-03	106	2
cassava4.1_023665	TATAAAAGGGTAGGT	-18.883	5.97E-03	249	2
cassava4.1_015102	CACAGCCCATTCAAG	-18.896	6.01E-03	34	2
cassava4.1_024542	AAAAAAAAAGTAAAT	-18.985	6.37E-03	114	2
cassava4.1_020743	AAAAGACGAATACAT	-18.989	6.38E-03	110	2
cassava4.1_023036	TATAAACAAAATCAA	-19.012	6.47E-03	119	2
cassava4.1_015102	CCCATTCAAGTTGAA	-19.04	6.59E-03	39	2
cassava4.1_026646	TACAGACAAAACATC	-19.043	6.60E-03	131	2
cassava4.1_016646	CAGAGTAAACCCCCT	-19.101	6.85E-03	59	2
cassava4.1_012090	TATAGTAAGATATCA	-19.107	6.87E-03	272	2
cassava4.1_012090	CATATTAATGTTGAA	-19.323	7.86E-03	33	2
cassava4.1_001042	CATCATCAATTAGCA	-19.438	8.43E-03	253	2
cassava4.1_020743	TATTTAAAAATAAAT	-19.454	8.51E-03	21	2
cassava4.1_020499	TCTCCTCAACTGCCT	-19.488	8.68E-03	198	2
cassava4.1_023665	TTTAATAACCAAAAT	-19.496	8.72E-03	86	2
cassava4.1_020499	AATCGGTAGATCAAT	-19.51	8.80E-03	264	2
cassava4.1_012090	TGTAATAGTATCACT	-19.514	8.82E-03	190	2
cassava4.1_020743	CAAAGTTTTGTAAAT	-19.604	9.31E-03	140	2
cassava4.1_023665	CAAAAACAAAAAAAT	-19.605	9.31E-03	201	2
cassava4.1_015102	TATCACATACTCCAC	-19.621	9.40E-03	17	2
cassava4.1_024542	GAAAGTGAACTTGAT	-19.624	9.42E-03	67	2
cassava4.1_022805	TAAAATTAATCAAAT	-19.68	9.74E-03	49	2
cassava4.1_024542	TCTTATCTAACCAAA	-19.697	9.84E-03	176	2

Table S5. TAL14$_{ClO151}$ TALE-NT (2.0) predicted effector binding elements (EBEs).

gene name	predicted EBE	score	start position	group
cassava4.1_007516	T ATAAAAAGGTCGGT	12.77	165	1
cassava4.1_007568	T ATAAAAAGGTCGGT	12.77	163	1
cassava4.1_009347	T ATAAAAAGGCCAGT	14.97	213	1
cassava4.1_011345	T ATAAACCCAACAAT	10.1	218	1
cassava4.1_011345	T ATATATAAACCCAA	15.82	214	1
cassava4.1_011345	T CTAAAAAAAGAAAA	17.56	21	1
cassava4.1_011345	T ATCTAAAAAAAGAA	18.21	19	1
cassava4.1_011524	T ATAAAAAGGTCAAA	11.46	148	1
cassava4.1_019005	T ATAGAAGGGGCGAT	13.51	165	1
cassava4.1_024404	T ATAAACCAGTCAAG	9.78	231	1
cassava4.1_024404	T ATAGAAGGCTTAAT	15.02	118	1
cassava4.1_024404	T CAAAACCTGTCAAA	16.05	265	1
cassava4.1_024404	T ATTAATAAATTTAT	17.71	102	1
cassava4.1_024404	T TTAAATAAATTTAT	17.71	90	1
cassava4.1_024404	T TTAATTAAAACAAT	18.15	161	1
cassava4.1_026299	T ATAAAAATGTCAAT	11.21	155	1
cassava4.1_033289	T ATAAAAGGGACAAC	14.92	27	1
cassava4.1_034150	T ATATACCAACCAAA	12.83	262	1
cassava4.1_034150	T ATAAAAGGGTACAT	14.1	233	1
cassava4.1_034150	T TTAGACAAAAGAAT	14.72	83	1
cassava4.1_034150	T AAATTAAAATAAAT	17.88	36	1
cassava4.1_034150	T ACCAACCAAACCCA	17.93	266	1
cassava4.1_034150	T TAAAATAAATGAAT	18.17	40	1
cassava4.1_001042	T CCCAACAAGAAAAT	16.08	151	2
cassava4.1_001042	T ATAGACGAAGAACA	16.49	184	2
cassava4.1_001042	T ATAAGAAAGTTAAC	16.76	218	2
cassava4.1_001042	T AAAGAGGTTTCAAT	17.73	109	2
cassava4.1_012090	T ACCAAGAAAACAAT	15.45	246	2
cassava4.1_012090	T TCAGTCAAATCCAC	17.56	125	2
cassava4.1_015102	T ATATAGATTTCAAT	16.3	250	2
cassava4.1_016646	T AGAAACTACAGGAT	18.28	104	2
cassava4.1_020743	T ATAATAAGACCAAT	15.17	197	2
cassava4.1_020743	T ATTTAAAAATAAAT	15.87	23	2
cassava4.1_020743	T AAAAATAAATTAAA	16.37	27	2
cassava4.1_020743	T ATATTTAAATAAAT	16.79	43	2
cassava4.1_020743	T ATTATCCACTCAAA	16.87	156	2
cassava4.1_020743	T ACCAAACACTCCAA	17.07	212	2
cassava4.1_022871	T GTAAAAAAATAAAT	14.03	148	2
cassava4.1_022871	T CCTAAAAATTCAAT	17.59	59	2

Table S5. (continued)

cassava4.1_023036	T ATAAAAGAGTACAT	12.63	233	2
cassava4.1_023036	T ATATACCAACCAAA	12.83	262	2
cassava4.1_023036	T ATAAACAAAATCAA	13.98	121	2
cassava4.1_023036	T TTAAAAAAATTAAA	16.96	153	2
cassava4.1_023036	T AGAGATAAGATCAT	17.82	96	2
cassava4.1_023036	T TAAAAAAATTAAAT	17.88	154	2
cassava4.1_023036	T ACCAACCAAACCCA	17.93	266	2
cassava4.1_023665	T AAAAATAAAAAAAT	15.45	230	2
cassava4.1_023665	T ATAAAAGGGTAGGT	16.55	251	2
cassava4.1_023665	T ATACATCAACAAAA	17.42	10	2
cassava4.1_025591	T ATAATCAAAACACC	15.43	231	2
cassava4.1_025591	T AAGAAGAAATCGAA	16.73	286	2
cassava4.1_025591	T AATCAAAACACCAT	18.29	233	2
cassava4.1_026121	T ATAATCATATCGAT	12.4	208	2
cassava4.1_026121	T ACCTCTCAATCAAT	18.15	225	2
cassava4.1_026646	T ACAGACAAAACATC	15.19	133	2
cassava4.1_026646	T AATAACAAAGAAAT	16.71	170	2
cassava4.1_031361	T ATAGACATAACTTT	16.25	156	2
cassava4.1_031361	T ATAATAAGATCAGT	16.68	122	2
cassava4.1_031361	T ATATAGACATAACT	16.72	154	2
cassava4.1_031361	T ATTAAAAAACTAAT	16.96	96	2
cassava4.1_020499	T ATAAAAGGAGCAAC	16.11	214	2

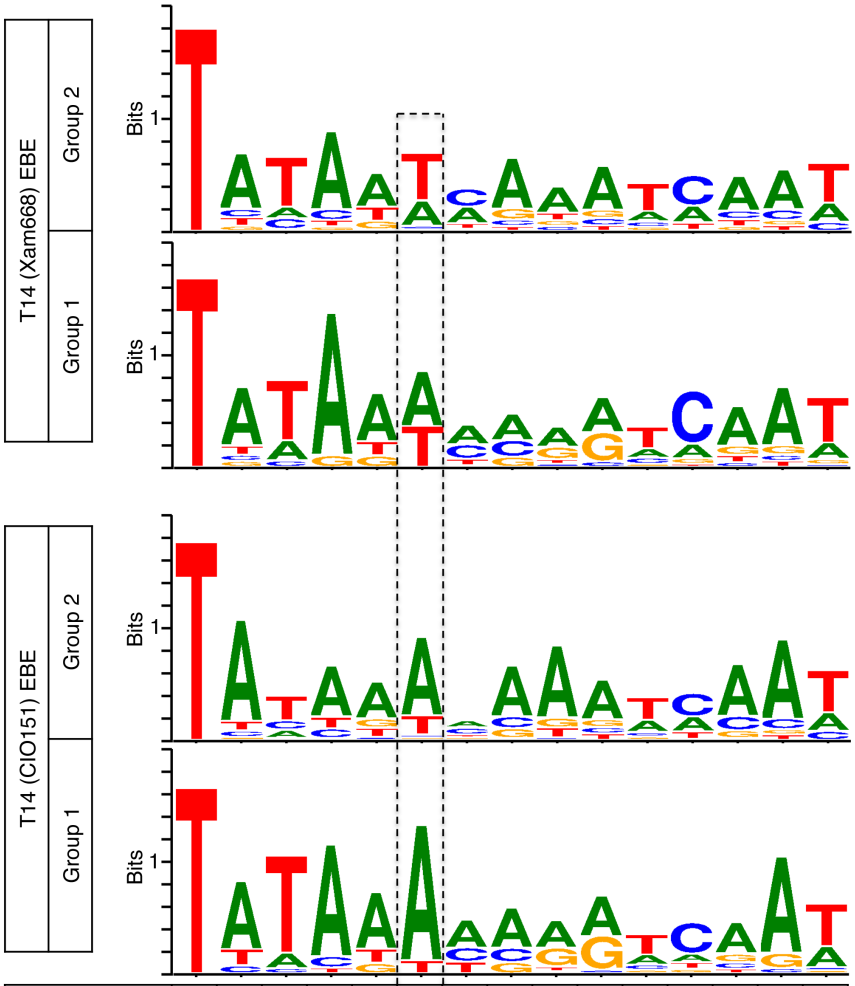
Table S6. TAL14$_{ClO151}$ TALgetter predicted effector binding elements (EBEs).

gene name	predicted EBE	score	p-value	position	group
cassava4.1_007516	TATAAAAAGGTCGGT	-12.386	2.67E-05	163	1
cassava4.1_007568	TATAAAAAGGTCGGT	-12.386	2.67E-05	161	1
cassava4.1_009347	TATAAAAAGGCCAGT	-16.083	8.04E-04	211	1
cassava4.1_009347	AAAAAAAAAGTTTAT	-18.768	5.38E-03	164	1
cassava4.1_011345	TATAAACCCAACAAT	-12.693	3.86E-05	216	1
cassava4.1_011345	TATATATAAACCCAA	-17.746	2.73E-03	212	1
cassava4.1_011345	TATCTAAAAAAAGAA	-19.6	9.00E-03	17	1
cassava4.1_019005	TATAGAAGGGGCGAT	-12.762	4.09E-05	163	1
cassava4.1_019005	TATCTACATCTCTGT	-18.446	4.38E-03	136	1
cassava4.1_019005	TAACGAAAAGTAGGC	-18.663	5.03E-03	30	1
cassava4.1_024404	TATAAACCAGTCAAG	-11.018	6.17E-06	229	1
cassava4.1_024404	TATAGAAGGCTTAAT	-15.978	7.39E-04	116	1
cassava4.1_024404	TCAAAACCTGTCAAA	-17.125	1.76E-03	263	1
cassava4.1_024404	AATCCACAGCTCAAA	-17.645	2.55E-03	276	1
cassava4.1_024404	TATTAATAAATTTAT	-18.632	4.93E-03	100	1
cassava4.1_024404	TTTAAATAAATTTAT	-18.67	5.06E-03	88	1
cassava4.1_024404	TATAAATTTATTAAT	-18.727	5.24E-03	65	1
cassava4.1_024404	TTTAATTAAAACAAT	-19.516	8.56E-03	159	1
cassava4.1_026299	TATAAAAATGTCAAT	-11.268	7.93E-06	153	1
cassava4.1_026299	TATTAAAAAAAAGGT	-19.548	8.72E-03	73	1
cassava4.1_033289	TATAAAAGGGACAAC	-17.252	1.93E-03	25	1
cassava4.1_033289	AATTAATTAGTAGAT	-18.868	5.74E-03	66	1
cassava4.1_034150	TATAAAAGGGTACAT	-14.972	3.24E-04	231	1
cassava4.1_034150	TATATACCAACCAAA	-15.055	3.47E-04	260	1
cassava4.1_034150	TTTAGACAAAAGAAT	-16.438	1.05E-03	81	1
cassava4.1_034150	ATTAAACAAAAAAAT	-18.288	3.95E-03	60	1
cassava4.1_034150	GATATAAAATTAAAT	-18.878	5.78E-03	141	1
cassava4.1_034150	TAAATTAAAATAAAT	-19.228	7.19E-03	34	1
cassava4.1_034150	AACTTGCAAGTCAAT	-19.295	7.49E-03	183	1
cassava4.1_001042	TCCCAACAAGAAAAT	-17.75	3.82E-03	149	2
cassava4.1_001042	TATAAGAAAGTTAAC	-17.901	4.22E-03	216	2
cassava4.1_001042	TATAGACGAAGAACA	-18.574	6.49E-03	182	2
cassava4.1_001042	TAAAGAGGTTTCAAT	-18.762	7.28E-03	107	2
cassava4.1_001042	TAGCAAAAAGGCGTT	-18.959	8.21E-03	263	2
cassava4.1_001042	CAAAGAAAGGACAAC	-18.99	8.37E-03	230	2
cassava4.1_012090	TTCAGTCAAATCCAC	-17.313	2.85E-03	123	2
cassava4.1_012090	AAAAAAACAGTATAT	-18.997	8.40E-03	13	2
cassava4.1_012090	TACCAAGAAAACAAT	-19.13	9.11E-03	244	2
cassava4.1_015102	AATAATCTCTTCGAT	-16.965	2.23E-03	176	2
cassava4.1_015102	TATATAGATTTCAAT	-17.267	2.76E-03	248	2
cassava4.1_015102	CACAGCCCATTCAAG	-18.75	7.23E-03	34	2
cassava4.1_016646	TATAATCTTCTCGCT	-16.269	1.35E-03	228	2
cassava4.1_016646	GATCGGCTCATCAAT	-18.702	7.02E-03	214	2
cassava4.1_020499	AAGTGACAAGTGGAT	-17.485	3.19E-03	162	2
cassava4.1_020499	TATAAAAGGAGCAAC	-18.73	7.14E-03	212	2
cassava4.1_020499	TAGCAATCGGTAGAT	-19.003	8.43E-03	260	2
cassava4.1_020743	AAAAGACGAATACAT	-16.434	1.52E-03	110	2
cassava4.1_020743	TATTTAAAAATAAAT	-16.538	1.64E-03	21	2
cassava4.1_020743	TATAATAAGACCAAT	-17.244	2.71E-03	195	2

Table S6. (continued)

cassava4.1_020743	TATTATCCACTCAAA	-17.367	2.95E-03	154	2
cassava4.1_020743	TATATTTAAATAAAT	-17.613	3.48E-03	41	2
cassava4.1_020743	CTCAAACACGACGCT	-18.436	5.95E-03	163	2
cassava4.1_020743	AATAAATTGGCCAAT	-18.742	7.20E-03	125	2
cassava4.1_020743	AATTGGCCAATCAAA	-18.833	7.61E-03	129	2
cassava4.1_020743	TACCAAACACTCCAA	-18.879	7.82E-03	210	2
cassava4.1_022871	TGTAAAAAAATAAAT	-15.688	8.64E-04	146	2
cassava4.1_022871	TCCTAAAAATTCAAT	-18.835	7.62E-03	57	2
cassava4.1_023036	TATAAAAGAGTACAT	-13.351	1.18E-04	231	2
cassava4.1_023036	TATATACCAACCAAA	-15.055	5.26E-04	260	2
cassava4.1_023036	TATAAACAAAATCAA	-16.61	1.73E-03	119	2
cassava4.1_023036	TAGAGATAAGATCAT	-18.095	4.79E-03	94	2
cassava4.1_023036	TCTTAACCTGAAGAT	-18.664	6.86E-03	49	2
cassava4.1_023036	TTTAAAAAAATTAAA	-18.772	7.33E-03	151	2
cassava4.1_023665	TATAAAAGGGTAGGT	-16.481	1.58E-03	249	2
cassava4.1_023665	CAAAAACAAAAAAAT	-17.203	2.64E-03	201	2
cassava4.1_023665	AATAAAAAAATACCT	-17.729	3.77E-03	232	2
cassava4.1_023665	TAAAAATAAAAAAAT	-18.408	5.85E-03	228	2
cassava4.1_023665	TATAATAATATTCAT	-19.16	9.27E-03	126	2
cassava4.1_024542	AAAAAAAAAGTAAAT	-16.583	1.69E-03	114	2
cassava4.1_025591	CATATCCAAATCGAA	-14.843	4.41E-04	254	2
cassava4.1_025591	TATAATCAAAACACC	-17.566	3.37E-03	229	2
cassava4.1_025591	AATAAACTTGCCCCT	-18.457	6.03E-03	113	2
cassava4.1_025591	CTCAAAAAACTCGAC	-18.769	7.32E-03	94	2
cassava4.1_026121	TATAATCATATCGAT	-11.978	3.12E-05	206	2
cassava4.1_026646	TACAGACAAAACATC	-16.487	1.58E-03	131	2
cassava4.1_026646	TAATAACAAAGAAAT	-18.941	8.12E-03	168	2
cassava4.1_031361	TATAGACATAACTTT	-15.731	8.93E-04	154	2
cassava4.1_031361	AATAAATAAATAAAT	-16.449	1.54E-03	106	2
cassava4.1_031361	AAAAAACTAATAAAT	-17.863	4.12E-03	98	2
cassava4.1_031361	TATAATAAGATCAGT	-18.27	5.36E-03	120	2
cassava4.1_031361	TATTAAAAAACTAAT	-18.661	6.84E-03	94	2
cassava4.1_031361	TATATAGACATAACT	-19.254	9.80E-03	152	2

Supplement C. Supplemental Figure



RVD position	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
T14 (Xam668)	-	NI	NG	NI	NN	NG	HD	NS	NS	NN	NG	HD	NN	NI	NG
T14 (ClO151)	-	NI	NG	NI	NN	NI	HD	NS	NS	NN	NG	HD	NN	NI	NG

Figure S1. Consensus TAL14$_{Xam668}$ and TAL14$_{ClO151}$ effector binding elements (EBEs) predicted by TALE-NT (2.0) Target Finder in group 1 and group 2 promoters. EBEs were predicted by TALE-NT (2.0) Target Finder in the 300 bp promoters of group 1 and group 2 genes by searching only the forward DNA strand and requiring a thymine (T) at position 0 (Doyle *et al.*, 2012). EBEs with scores within 3.5-fold of the best possible TAL14 binding score were considered. The consensus EBE for each group is displayed with position 5 outlined by dashed lines. Consensus sequences were generated with Weblogo 3.4 (Schneider & Stephens, 1990; Crooks *et al.*, 2004). The height of the consensus indicates the nucleotide conservation at that site (measured in bits) while the height of the nucleotides within the consensus reflect their relative frequency. TAL14$_{Xam668}$ and TAL14$_{ClO151}$ repeat variable diresidue (RVD) sequences are displayed.

6. References

Andrews S. **2015**. Babraham Bioinformatics - FastQC: a quality control tool for high throughput sequence data. Available at: *http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.*

Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. **2010**. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell Online* **22**: 3864–3876.

Arlat M, Gough CL, Barber CE, Boucher C, Daniels MJ. **1991**. *Xanthomonas campestris* contains a cluster of *hrp* genes related to the larger *hrp* cluster of *Pseudomonas solanacearum*. *Molecular Plant Microbe Interactions* **4**: 593–601.

Arrieta-Ortiz ML, Rodríguez-R LM, Pérez-Quintero ÁL, Poulin L, Díaz AC, Rojas NA, Trujillo C, Benavides MR, Bart R, Boch J, *et al.* 2013. Genomic survey of pathogenicity determinants and VNTR markers in the cassava bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis* strain CIO151. *PloS one* **8**: e79704.

Bai J, H CS, Ponciano G, Leung H, Leach JE. **2000**. *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. *Molecular Plant Microbe Interactions* **13**: 1322–1329.

Ballvora A, Pierre M, Van den Ackerveken G, Schornack S, Rossier O, Ganal M, Lahaye T, Bonas U. **2001**. Genetic mapping and functional analysisof the tomato Bs4 locus governing recognition of the *Xanthomonas campestris* pv. *vesicatoria* AvrBs4 protein. *Molecular Plant Microbe Interactions* **14**: 629–638.

Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, *et al.* 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences* **109**: E1972–E1979.

Bernard V, Brunaud V, Lecharny A. **2010**. TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics* **11**: 166–15.

Blankenberg D, Kuster GV, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. **2001**. Galaxy: a web-based genome analysis tool for experimentalists. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M. **2007**. Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria (R Redfield, Ed.). *PloS one* **2**: e224–21.

Boch J, Bonas U. **2010**. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* **48**: 419–436.

Boch J, Bonas U, Lahaye T. **2014**. TAL effectors - pathogen strategies and plant resistance engineering. *New Phytologist* **204**: 823–832.

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

Bogdanove AJ, Voytas DF. **2011**. TAL effectors: customizable proteins for DNA targeting. *Science* **333**: 1843–1846.

Bogdanove AJ, Schornack S, Lahaye T. **2010**. TAL effectors: finding plant genes for disease and defense. *Current opinion in plant biology* **13**: 394–401.

Boher B, Verdier V. **1994**. Cassava bacterial blight in Africa: the state of knowledge and implications for designing control strategies. *African Crop Science Journal* **2**: 505–509.

Boher B, Kpemoua K, Nicole M, Luisetti J, Geiger JP. **1995**. Ultrastructure of interactions between cassava and *Xanthomonas campestris* pv. *manihotis*: cytochemistry of cellulose and pectin degradation in a susceptible cultivar. *Phytopathology* **85**: 777–788.

Bonas U, Stall RE, Staskawicz BJ. **1989**. Genetic and structural characterization of the avirulence gene avrBs3 from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular Genetics and Genomics* **218**: 127–136.

Brencic A, Winans SC. **2005**. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiology and Molecular Biology Reviews* **69**: 155–194.

Brunings AM, Gabriel DW. **2003**. *Xanthomonas citri*: breaking the surface. *Molecular Plant Pathology* **4**: 141–157.

Bull SE, Owiti JA, Niklaus M, Beeching JR, Gruissem W, Vanderschuren H. **2009**. *Agrobacterium*-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nature Protocols* **4**: 1845–1854.

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.

Butler JEF, Kadonaga JT. **2002**. The RNA polymerase II core promoter: a key

component in the regulation of gene expression. *Genes and Development* **16**: 2583–2592.

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

Castiblanco LF, Gil J, Rojas A, Osorio D, Gutiérrez S, Muñoz-Bodnar A, Pérez-Quintero ÁL, Koebnik R, Szurek B, López C, *et al.* 2012. TALE1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. *Molecular Plant Pathology* **14**: 84–95.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. **2011**. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**: gkr218–e82.

Cernadas RA, Doyle EL, Niño-Liu DO, Wilkins KE, Bancroft T, Wang L, Schmidt CL, Caldo R, Yang B, White FF, *et al.* 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene (JM Zhou, Ed.). *PLoS Pathogens* **10**: e1003972.

Chen L-Q. **2014**. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytologist* **201**: 1150–1155.

Chen L-Q, Hou B-H, Lalonde S, Takanaga H, Hartung ML, Qu X-Q, Guo W-J, Kim J-G, Underwood W, Chaudhuri B, *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 AR, Frommer WB. **2012**. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**: 207–211.

Chu Z. **2006**. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes and Development* **20**: 1250–1255.

Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T, Staskawicz BJ. **2014**. *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector–mediated induction of a SWEET sugar transporter in cassava. *dx.doi.org* **27**: 1186–1198.

Cohn M, Shybut M, Dahlbeck D, Staskawicz BJ. **2015**. Assays to assess virulence of *Xanthomonas axonopodis* pv. *manihotis* on cassava. *Bio-protocol* **5**: e1522.

Cong L, Zhou R, Kuo Y-C, Cunniff M, Zhang F. **2012**. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature Communications* **3**: 968.

Crooks GE, Hon G, Chandonia JM, Brenner SE. **2004**. WebLogo: a sequence logo generator. *Genome Research* **14**: 1188–1190.

Dangl JL, Horvath DM, Staskawicz BJ. **2013**. Pivoting the plant immune system from dissection to deployment. *Science* **341**: 746–751.

de Lange O, Binder A, Lahaye T. **2014**. From dead leaf, to new life: TAL effectors as tools for synthetic biology. *The Plant Journal* **78**: 753–771.

Denance N, Szurek B, Noel LD. **2014**. Emerging functions of nodulin-like proteins in non-nodulating plant species. *Plant and Cell Physiology* **55**: 469–474.

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

Deuschle K. **2006**. Rapid metabolism of glucose detected with FRET dlucose nanosensors in epidermal cells and intact roots of *Arabidopsis* RNA-silencing mutants. *The Plant Cell Online* **18**: 2314–2325.

Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, VanDyk JK, Bogdanove AJ. **2012**. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Research* **40**: W117–W122.

Doyle EL, Hummel AW, Demorest ZL, Starker CG, Voytas DF, Bradley P, Bogdanove AJ. **2013a**. TAL effector ppecificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat –1 (A Jeltsch, Ed.). *PloS one* **8**: e82120–17.

Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ. **2013b**. TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. *Trends in Cell Biology* **23**: 390–398.

Food and Agriculture Organization of the United Nations: Cassava. **2008**. Available at: http://www.fao.org/ag/AGP/agpc/gcds/index_en.html.

Gassmann W, Dahlbeck D, Chesnokova O, Minsavage GV, Jones JB, Staskawicz BJ. **2000**. Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. *Journal of Bacteriology* **182**: 7053–7059.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, *et al.* 2005. Galaxy: A platform for interactive large-scale genome analysis. *Genome Research* **15**: 1451–1455.

Goecks J, Nekrutenko A, Taylor J, Galaxy Team T. **2010**. Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology* **11**: R86–13.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, *et al.* 2011. Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research* **40**: D1178–D1186.

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* **38**: W695–W699.

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 (A Tresch, Ed.). *PLOS Comput Biol* **9**: e1002962.

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

Hou B-H, Takanaga H, Grossmann G, Chen L-Q, Qu X-Q, Jones AM, Lalonde S, Schweissgut O, Wiechert W, Frommer WB. 2011. Optical sensors for monitoring dynamic changes of intracellular metabolite levels in mammalian cells. *Nature Protocols* **6**: 1818–1833.

Howeler R, Lutaladio N, Thomas G. 2013. Save and grow cassava: a guide to sustainable production intensification. : 1–24.

Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, Yang B, White FF, Wang N, Jones JB. 2014. Lateral Organ Boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences* **111**: E521–E529.

Hummel AW, Doyle EL, Bogdanove AJ. 2012a. 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.

Hutin M, Pérez-Quintero ÁL, López C, Szurek B. 2015. MorTAL Kombat: The story of defense against TAL effectors through loss-of-susceptibility. *Frontiers in Plant Science* **6**: 1–12.

Joung JK, Sander JD. 2012. TALENs: A widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology* **14**: 49–55.

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

Kocal N, Sonnewald U, 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 Physiology* **148**: 1523–1536.

Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, Yanagawa H. **2009a**. Six classes of nuclear localization signals specific to different binding grooves of importin α. *Journal of Biological Chemistry* **284**: 478–485.

Kosugi S, Hasebe M, Tomita M, Yanagawa H. **2009b**. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences* **106**: 10171–10176.

Ladwig F, Stahl M, Ludewig U, Hirner AA, Hammes UZ, Stadler R, Harter K, Koch W. **2012**. Siliques Are Red1 from *Arabidopsis* acts as a bidirectional amino acid transporter that is crucial for the amino acid homeostasis of siliques. *Plant Physiology* **158**: 1643–1655.

Lager I, Looger LL, Hilpert M, Lalonde S, Frommer WB. **2006**. Conversion of a putative *Agrobacterium* sugar-binding protein into a FRET sensor with high selectivity for sucrose. *The Journal of Biological Chemistry* **281**: 30875–30883.

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.

Li Z, Zou L, Ye G, Xiong L, Ji Z, Zakria M, Hong N, Wang G, Chen G. **2014**. A potential disease susceptibility gene CsLOB of citrus is targeted by a major virulence effector PthA of *Xanthomonas citri* subsp. *citri*. *Molecular Plant* **7**: 912–915.

Lin IW, Sosso D, Chen L-Q, Gase K, Kim S-G, Kessler D, Klinkenberg PM, Gorder MK, Hou B-H, Qu X-Q, *et al.* 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546–549.

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.

Logue C-A, Peak IRA, Beacham IR. **2009**. Facile construction of unmarked deletion mutants in *Burkholderia pseudomallei* using sacB counter-selection in sucrose-resistant and sucrose-sensitive isolates. *Journal of Microbiological Methods* **76**: 320–323.

Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B. **2012**. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research* **40**: W622–W627.

Lozano JC. **1986**. Cassava bacterial blight: a manageable disease. *Plant Disease* **70**: 1089–1093.

Lozano JC, Sequeira L. **1973**. Bacterial blight of cassava in Colombia: epidemiology and control. *Phytopathology* **64**: 83–88.

López CE, Bernal AJ. **2012b**. Cassava bacterial blight: using genomics for the elucidation and management of an old problem. *Tropical Plant Biology* **5**: 117–126.

Mak ANS, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. **2012**. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**: 716–719.

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.

Martin M. **2011**. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10–12.

Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, Yu A, Lockwood SH, Morbitzer R, Elsaesser J, *et al.* 2013. Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Research* **41**: 4118–4128.

Moore R, Chandrahas A, Bleris L. **2014**. Transcription activator-like effectors: a toolkit for synthetic biology. *ACS Synthetic Biology* **3**: 708–716.

Morbitzer R, Römer P, Boch J, Lahaye T. **2010**. Regulation of selected genome loci using denovo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences* **107**: 21617–21622.

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

Muñoz-Bodnar A, Pérez-Quintero ÁL, Gomez-Cano F, Gil J, Michelmore R, Bernal A, Szurek B, López C. **2014**. RNAseq analysis of cassava reveals similar plant responses upon infection with pathogenic and non-pathogenic strains of *Xanthomonas axonopodis* pv. *manihotis*. *Plant Cell Reports* **33**: 1901–1912.

Pereira AL, Carazzolle MF, Abe VY, de Oliveira ML, Domingues MN, Silva JC, Cernadas RA, Benedetti CE. **2014**. Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. *BMC Genomics* **15**: 157.

Petersen TN, Brunak S, Heijne von G, Nielsen H. **2011**. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**: 785–786.

Prochnik S, Marri PR, Desany B, Rabinowicz PD, Kodira C, Mohiuddin M, Rodriguez F, Fauquet C, Tohme J, Harkins T, *et al.* 2012. The cassava genome: current progress, future directions. *Tropical Plant Biology* **5**: 88–94.

Rioux D, Nicole M, Simard M, Ouellette GB. **1998**. Immunocytochemical evidence that secretion of pectin occurs during gel (gum) and tylosis formation in trees. *Phytopathology* **88**: 494–505.

Römer P, Hahn S, Jordan T, Strauß T, Bonas U, Lahaye T. **2007**. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**: 645–648.

Römer P, Recht S, Lahaye T. **2009a**. A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proceedings of the National Academy of Sciences* **106**: 20526–20531.

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.

Römer P, Strauß T, Hahn S, Scholze H, Morbitzer R, Grau J, Bonas U, Lahaye T. **2009b**. Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper Bs3 alleles. *Plant Physiology* **150**: 1697–1712.

Ryan RP, Vorhölter F-J, Potnis N, Jones JB, Van Sluys M-A, Bogdanove AJ, Dow JM. **2011**. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. *Nature Reviews Microbiology* **9**: 344–355.

Sander JD, Joung JK. **2014**. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology* **32**: 347–355.

Schneider TD, Stephens RM. **1990**. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research* **18**: 6097–6100.

Schornack S, Ballvora A, Gürlebeck D, Peart J, Ganal M, Baker B, Bonas U, Lahaye T. **2004**. The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *The Plant Journal* **37**: 46–60.

Schornack S, Moscou MJ, Ward ER, Horvath DM. **2013**. Engineering plant disease resistance based on TAL effectors. *Annual Review of Phytopathology* **51**: 383–406.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, ding JSO, *et al.* 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**: 1–6.

Sonnewald U. **2011**. SWEETS – The missing sugar efflux carriers. *Frontiers in Plant Science* **2**: 1–2.

Strauß T, van Poecke RMP, Strauß A, Römer P, Minsavage GV, Singh S, Wolf C, Strauß A, Kim S, Lee H, *et al.* 2012. RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proceedings of the National Academy of Sciences* **109**: 19480–19485.

Streubel J, Blücher C, Landgraf A, Boch J. **2012**. TAL effector RVD specificities and efficiencies. *Nature biotechnology* **30**: 593–595.

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

Sturm A. **1999**. Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiology* **121**: 1–7.

Sun N, Zhao H. **2013**. Transcription Activator-Like Effector Nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnology and Bioengineering* **110**: 1811–1821.

Sundin GW. **2007**. Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. *Annual Review of Phytopathology* **45**: 129–151.

Swarup S, De Feyter R, Brlansky RH, Gabriel D. **1991**. A pathogenicity locus from *Xanthomonas citri* enables strains from several pathovars of *X. campestris* to elicit cankerlike lesions on citrus. *Molecular Plant Pathology* **81**: 802–809.

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

Takanaga H, Frommer WB. **2010**. Facilitative plasma membrane transporters function during ER transit. *The FASEB Journal* **24**: 2849–2858.

Takanaga H, Chaudhuri B, Frommer WB. **2008**. GLUT1 and GLUT9 as major contributors to glucose influx in HepG2 cells identified by a high sensitivity intramolecular FRET glucose sensor. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778**: 1091–1099.

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.

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA,

Rhee SY, Stitt M. **2004**. Mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* **37**: 914–939.

Tian D, Wang J, Zeng X, Gu K, Qiu C, Yang X, Zhou Z, Goh M, Luo Y, Murata-Hori M, *et al.* 2014. The rice TAL effector–dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. *The Plant Cell Online* **26**: 497–515.

Trapnell C, Pachter L, Salzberg SL. **2009**. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. **2012**. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**: 562–578.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. **2010**. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**: 511–515.

van der Hoorn RAL, Kamoun S. **2008**. From guard to decoy: a new model for perception of plant pathogen effectors. *The Plant Cell Online* **20**: 2009–2017.

Verdier V, Triplett LR, Hummel AW, Corral R, Cernadas RA, Schmidt CL, Bogdanove AJ, Leach JE. **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.

Vivian A, Murillo J, Jackson RW. **2015**. The roles of plasmids in phytopathogenic bacteria: mobile arsenals? *Microbiology* **147**: 763–780.

Vogel JP, Raab TK, Schiff C, Somerville SC, Somerville. **2002**. PMR6, a pectate lyase-like gene required for powdery mildew susceptibility in *Arabidopsis*. *The Plant Cell Online* **14**: 2095–2106.

Wilkins KE, Booher NJ, Wang L, Bogdanove AJ. **2015**. TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. *Frontiers in Plant Science* **6**: 1–15.

Wydra K, Zinsou V, Jorge V, Verdier V. **2004**. Identification of pathotypes of *Xanthomonas axonopodis* pv. *manihotis* in Africa and detection of quantitative trait loci and markers for resistance to Bacterial Blight of Cassava. *Phytopathology* **94**: 1084–1093.

Yadeta KA, Thomma BPHJ. **2013**. The xylem as battleground for plant hosts and vascular wilt pathogens. *Frontiers in Plant Science* **4**: 1–12.

Yang B, White FF. **2004**. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular Plant Microbe Interactions* **17**: 1192–1200.

Yang B, Sugio A, White FF. **2006**. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* **103**: 10503–10508.

Yang Y, Gabriel DW. **1995**. Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. *Journal of Bacteriology* **177**: 4963–4968.

Yang Y, De Feyter R, Gabriel DW. **1994**. Host-specific symptoms and increased release of *Xanthomonas citri* and *X. campestris* pv. *malvacearum* from leaves are determined by the 102-bp tandem repeats of pthA and avrb6, respectively. *Molecular Plant Microbe Interactions* **7**: 345–355.

Yang Y, Gabriel DW, Yuan Q. **1996**. Watersoaking functlon(s) of XcmH1005 are redundantly encoded by members of the *Xanthomonas avr/pth* gene family. *Molecular Plant Microbe Interactions* **9**: 105–113.

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 *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, Wang S. **2013**. Rice MtN3/Saliva/SWEET family genes and their homologs in cellular organisms. *Molecular Plant* **6**: 665–674.

Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom J-S, Huang S, Liu S, Vera Cruz C, Frommer WB, *et al.* 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant Journal* **82**: 632–643.