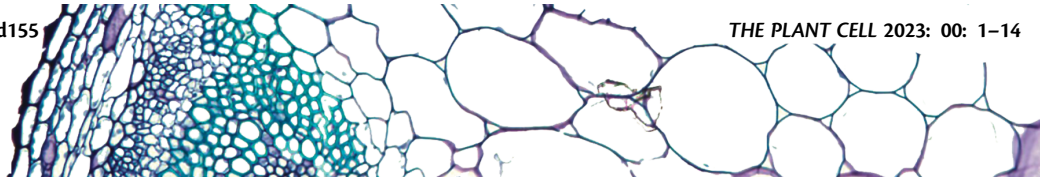
https://doi.org/10.1093/plcell/koad155 THE PLANT CELL 2023: 00: 1–14



Review

Bacterium-enabled transient gene activation by artificial transcription factors for resolving gene regulation in maize

Mingxia Zhao [1,†] Zhao Peng [2,3,†] Yang Qin [4,†] Tej Man Tamang [1] Ling Zhang [5] Bin Tian [1,‡] Yueying Chen [1,§] Yan Liu [4] Junli Zhang [2,**] Guifang Lin [1] Huakun Zheng [1,††] Cheng He [1] Kaiwen Lv [6] Alina Klaus [7] Caroline Marcon [7] Frank Hochholdinger [7] Harold N. Trick [1] Yunjun Liu [4] Myeong-Je Cho [8] Sunghun Park [9] Hairong Wei [5] Jun Zheng [4] Frank F. White [2,*] and Sanzhen Liu [1,*]

1 Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA
2 Department of Plant Pathology, University of Florida, Gainesville, FL 32611, USA
3 College of Plant Protection, Jilin Agricultural University, Changchun, Jilin 130118, China
4 Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
5 College of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA
6 State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Heilongjiang 150040, China
7 INRES, Institute of Crop Science and Resource Conservation, Crop Functional Genomics, University of Bonn, Bonn 53113, Germany
8 Innovative Genomics Institute, University of California, Berkeley, CA 94704, USA
9 Department of Horticulture and Natural Resources, Kansas State University, Manhattan, KS 66506, USA

*Author for correspondence: ffwhite@ufl.edu (F.F.W.), liu3zhen@ksu.edu (S.L.)
†These authors contributed equally to this work.
‡Present address: Seeds Research, Syngenta Crop Protection, LLC, Research Triangle Park, North Carolina 27703, USA.
§Present address: DeltaMed Solutions, Inc., 220 Davidson Avenue, Suite 201, Somerset, NJ 08873, USA.
**Present address: Center for Environmental and Human Toxicology, University of Florida, Gainesville, FL 32611-0680, USA.
††Present address: National Engineering Research Center of JUNCAO Technology, College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China.
The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) are: Sanzhen Liu (liu3zhen@ksu.edu) and Frank F. White (ffwhite@ufl.edu).

Abstract

Understanding gene regulatory networks is essential to elucidate developmental processes and environmental responses. Here, we studied regulation of a maize (*Zea mays*) transcription factor gene using designer transcription activator-like effectors (dTALEs), which are synthetic Type III TALEs of the bacterial genus *Xanthomonas* and serve as inducers of disease susceptibility gene transcription in host cells. The maize pathogen *Xanthomonas vasicola* pv. *vasculorum* was used to introduce 2 independent dTALEs into maize cells to induced expression of the gene *glossy3* (*gl3*), which encodes a MYB transcription factor involved in biosynthesis of cuticular wax. RNA-seq analysis of leaf samples identified, in addition to *gl3*, 146 genes altered in expression by the 2 dTALEs. Nine of the 10 genes known to be involved in cuticular wax biosynthesis were upregulated by at least 1 of the 2 dTALEs. A gene previously unknown to be associated with *gl3*, Zm00001d017418, which encodes aldehyde dehydrogenase, was also expressed in a dTALe-dependent manner. A chemically induced mutant and a CRISPR-Cas9 mutant of Zm00001d017418 both exhibited glossy leaf phenotypes, indicating that Zm00001d017418 is involved in biosynthesis of cuticular waxes. Bacterial protein delivery of dTALEs proved to be a straightforward and practical approach for the analysis and discovery of pathway-specific genes in maize.

Received February 05, 2021. Accepted May 04, 2023. Advance access publication May 26, 2023

Introduction

Precise control of transcriptional regulation is essential for proper cellular differentiation and appropriate responses to environmental signals. Transcription factors (TFs) are key components modulating gene expression, and therefore understanding TF function is critical for elucidating gene regulation networks. Traditional approaches to analyze transcription pathways involve transient or stable ectopic expression, genetic mutation, and analysis of transcription. Additional useful analyses include the identification of TF binding sites using chromatin immunoprecipitation sequencing (ChIP-Seq), DNA footprinting, and in vitro DNA affinity purification sequencing (DAP-Seq) (Bartlett et al. 2017). Each approach has its advantages and limitations (Lai et al. 2019). Although ectopic expression or knockout constructs are used to elucidate the phenotypic effects and transcriptional roles of TFs, the approaches are often not available or require considerable time to construct. Techniques involving transient gene activation offer alternative approaches for inducing transcriptional changes in target cells typically involving *Agrobacterium tumefaciens*-mediated gene transfer, viral infection, or direct bombardment with DNA-coated particles (Gleba et al. 2014). Emerging nanoparticles technologies and techniques, including, for example, nonbiolistic delivery of nano-gold spheres, to enhance plant cell wall penetration also offer options for delivery of nucleotides and proteins (Cunningham et al. 2018; Demirer et al. 2020). However, transient expression methods can be difficult to carry out, depending on the target plant species, and are challenging in maize (*Zea mays*).

Delivery of proteins by natural bacterial processes represents a promising alternative method, although the approach has not been extensively used. Many pathogenic species of bacteria have evolved secretion systems that inject proteins into host cells to induce changes in host metabolism and, thus, facilitate colonization of host tissues (Costa et al. 2015; Deslandes and Rivas 2012; Block et al. 2008). Principal among these systems is the Type III secretion system (T3SS), which is a supramolecular protein complex that delivers bacterial effector proteins into target cells. In nature, the T3SS delivers effectors that suppress host defenses and condition host cells for susceptibility (Green and Mecsas 2016). Most members of the genus *Xanthomonas* require a functional T3SS for virulence and cause a variety of diseases on hundreds of plant species, including most major crop species (White et al. 2009; Büttner and Bonas 2010).

Among the many protein effectors that transit the T3SS, the transcription activator-like effector (TALe) family is a group of Type III effectors that primarily direct expression of specific host disease susceptibility genes. The C-terminal portion of the TALe contains domains specifying eukaryotic nuclear localization and transcription activation (Yang *et al.* 2000; Zhu *et al.* 1998; Van den Ackerveken *et al.* 1996), and the central repetitive sequence consists of a variable number of repeats. Each repeat contains 34–35 nearly identical amino acid residues and 2 highly variable residues at the 12th and 13th positions. These 2 residues are known as the repeat variable diresidue (RVD). Each repeat forms 2 alpha helices, and the RVD in a repeat determines the specific recognition of 1 of the 4 DNA nucleotides at the target site (Boch et al. 2009; Moscou and Bogdanove 2009). The specific recognition between an RVD and nucleotide bases in the target DNA provides a rationale for the construction of artificial designer TALes (dTALes) that can target specific DNA sequences in host plants (Joung and Sander 2013; Li et al. 2013b; Bogdanove 2014; Morbitzer et al. 2010; Streubel et al. 2013).

In maize, as in other monocotyledonous plants, transient expression of genes in intact plants is challenging. To overcome this limitation, we used dTALe-mediated targeting delivered by *X. vasicola* pv. *vasculorum* (Xvv) to activate transcription in intact maize plants. Xvv infects at least 17 members of the Poaceae, either symptomatically or asymptomatically (Hartman et al. 2020). Specifically, the approach was applied to the pathway for cuticular wax development in maize (Perez-Quintero et al. 2020). Cuticular waxes are derivatives of very-long-chain fatty acids that provide an external hydrophobic barrier that protects plant tissues from water loss and other environmental stresses, including, for example, drought (Kunst and Samuels 2003; Lee and Suh 2013; Fehling and Mukherjee 1991). Cuticular waxes are secreted through the plasma membrane of epiderma cells and are deposited on the plant surface. In maize, mutants deficient in cuticular wax production have leaves that can hold water droplets, giving the leaves a distinct glossy phenotype. To date, more than 30 loci responsible for this phenotype have been discovered in these mutants, and at least 11 genes have been shown to be responsible for the glossy leaf phenotype, including *gl1*, *gl2*, *gl3*, *gl4*, *gl6*, *gl8*, *gl13*, *gl14*, *gl15*, *gl26*, and *cer8* (Tacke et al. 1995; Li et al. 2013a, 2019; Liu et al. 2012, 2009; Moose and Sisco 1996; Zheng et al. 2019; Hansen et al. 1997; Xu et al. 1997). Among them, *gl13* encodes an ABC transporter that secretes cuticular waxes through the plasma membrane (Li et al. 2013a). We used dTALes to activate *gl3*, which is the focus of this research and encodes a TF that functions early in the pathway for biosynthesis of cuticular waxes (Liu et al. 2012). The dTALes were also proved useful for identifying additional members of the glossy pathway and, more specifically, were used to discover downstream genes regulated by GL3.

Results

A bacterium-enabled protein delivery system in maize

Xvv strain Xv1601, a maize pathogen that lacks TALe genes (Perez-Quintero et al. 2020), was used for delivery of dTALes into maize plants. Phylogenetic analysis of 10 *Xanthomonas* species (Supplemental Fig. S1) indicated that Xvv is genetically close to *Xanthomonas oryzae* (Xo), a pathogen that causes bacterial blight or rice and employs

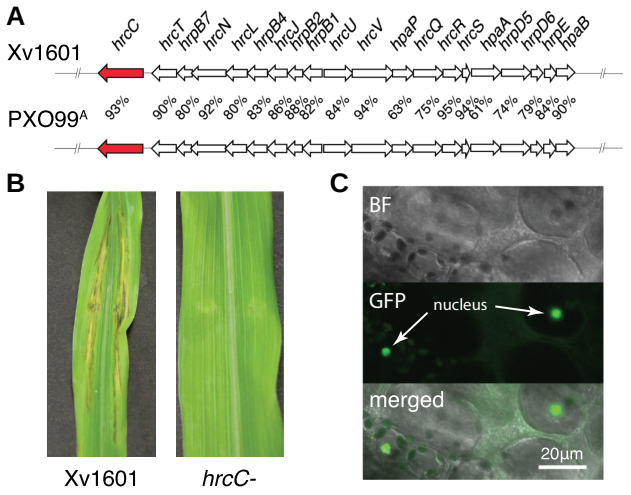


Figure 1. The T3SS of Xvv functions in delivering proteins to maize cells. **A)** Comparison of the Type III secretion gene cluster of Xv1601 and the *Xanthomonas oryzae* strain PX099^{A}. The DNA identity between each orthologous gene pair is listed. The *hrcC* genes are highlighted in red. **B)** Leaf phenotype 5 days after inoculation with the WT strain (Xv1601) and the *hrcC*$^{-}$ knockout mutant strain. **C)** Bright field (BF) and fluorescence (GFP) images of maize cells after 24 h of the infection with bacteria carrying a gene of AvrBs2::T3SS signal peptide-NLS::eGFP-NLS (Supplemental Fig. S2).

critical TALes during the disease process (Mew 1987, Oliva et al. 2019). Xv1601 contains a gene cluster that is syntenic with the T3SS genes from the Xo reference strain PXO99^{A}, which is known to deliver TALes during infection (Fig. 1A) (Yang et al. 2006). Furthermore, Xv1601 *hrcC*$^{-}$ is a knockout mutant of *hrcC*, an essential gene of the T3SS system, and was dramatically reduced in virulence on maize leaves, indicating that the T3SS is functional and critical for bacterial virulence (Fig. 1B). To test the ability of Xv1601 to deliver proteins into intact maize leaf cells, a plasmid bearing a gene encoding an enhanced green fluorescent protein (eGFP) fused to the promoter and N-terminal Type III secretion peptide of the Type III effector *AvrBs2*, which allows non-Type III proteins to transit the T3SS (Mudgett et al. 2000), was introduced into Xv1601 and Xv1601 *hrcC*$^{-}$ (Supplemental Fig. S2). To enhance detection by concentration of the protein into nuclei, a nuclear localization signal was incorporated into the protein (Khang et al. 2010). Following inoculation of leaf tissue, GFP fluorescence was detected in the nuclei of host cells (Fig. 1C and Supplemental Fig. S3), while no signal was detected in leaves infected with the *hrcC*$^{-}$ mutant (Supplemental Fig. S3).

TALe-induced expression of host genes

Two dTALes, referred to as dT1 and dT2, were constructed to target 2 nonoverlapping 16-bp effector binding elements (EBEs) of the *gl3* promoter at positions 5 bp (dT1) and 48 bp (dT2) upstream of the transcription start site (Fig. 2, A and B). The EBEs targeted by dT1 and dT2, respectively, overlap with 2 predicted TATA boxes in the *gl3* promoter region. Expression of *gl3* in untreated seedling leaves of inbred line A188 was observed at germination and dropped to undetectable levels after 14 days (Supplemental Fig. S4). Therefore, 14-day-old seedlings were used to test for dTALe-mediated induction of *gl3*. Bacteria carrying either dT1 or dT2 activated *gl3* expression within 24 h after inoculation (Fig. 2C). Compared with dT2, dT1 promoted stronger induction of *gl3*, as measured by reverse transcription quantitative PCR (RT-qPCR) (Supplemental Fig. S5). A time-course analysis of *gl3* expression induced by dT1 showed that the relative expression levels (compared to EV, the empty vector control) were 22 times and 82 times higher at 24 and 48 h after inoculation, respectively, (Fig. 2D).

Two additional maize genes were targeted by dTALe activation, including a homolog of an AP2/ERF TF gene *bbm*, Zm00001eb144510, and a homolog of a WUSCHEL-related homeobox TF gene *wus2*, Zm00001eb433010. A single dTALe was designed for each gene (Supplemental Fig. S6, A and C), and Xv1601 strains harboring each dTALe were used to inoculate leaves of young seedlings of maize inbred line B73. Quantification of expression of the targeted genes showed that each gene was specifically upregulated by at least 16-fold by the respective dTALe, indicating that dTALe activation of maize genes was not limited to *gl3* or the A188 line (Supplemental Fig. S6, B and D).

GL3 downstream genes identified through RNA-seq

To determine the genes that are expressed in association with *gl3* TALe-mediated induction, RNA-seq was performed using leaf tissues inoculated 24 h earlier with bacteria carrying dT1, dT2, or the EV control. The basal expression level of *gl3* in young leaves of 14-day-old seedlings was low, while treatments with dT1 or dT2 caused a 191- and 74-fold induction of *gl3*, respectively (Fig. 3A, Table 1). Distribution of RNA-seq reads on the *gl3* locus indicated that transcripts induced by dT1 and dT2 included the intact *gl3* coding sequence, as seen with native endogenous transcripts (Supplemental Fig. S7). The comparison of dT1 with the EV control identified 1,249 differentially expressed genes (DEGs) at the false discovery rate (FDR) of 5%, of which 499 were upregulated. A comparison of dT2 versus EV resulted in 430 DEGs being identified at the FDR of 10%, of which 156 were upregulated (Fig. 3B, Supplemental Data Set 1). Note that a higher FDR value was used in dT2 due to a lower level of induction of *gl3* expression. The 92 common upregulated DEGs of dT1 and dT2, which did not include *gl3*, and 54 common downregulated DEGs were associated with TALe-mediated *gl3* expression. Gene ontology (GO) analysis showed that genes related to fatty acid biosynthesis and the endoplasmic reticulum (ER) were overrepresented in the 92 upregulated genes (Fig. 3C). RNA-seq was also performed on leaf samples collected 18 h post-inoculation (hpi) to compare transcriptional responses between dT1 and EV treatments. From 18-hpi RNA-seq data and a FDR of 5%, only 3 genes upregulated by dT1 were in common with the 92 genes upregulated by both dTALes at

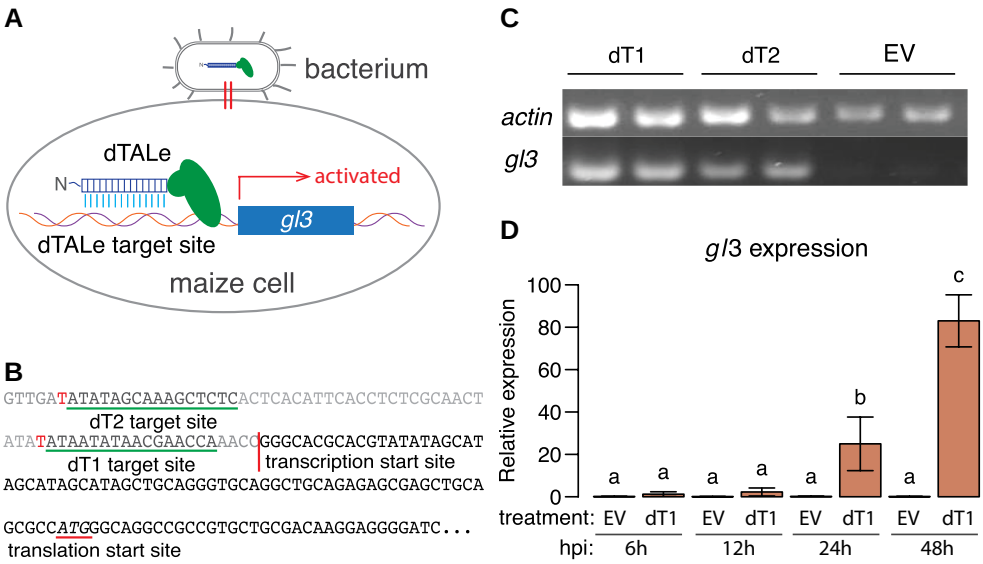


Figure 2. dTALe-dependent *gl3* gene expression. **A)** Schematic of bacterium-mediated delivery of dTALes for the expression activation of maize *gl3*. **B)** Target sequences for dT1 and dT2 (underlined in green). The transcription start site is indicated by a vertical line. The translation start site ATG for GL3 is underlined in red. **C)** Semi-quantitative RT-qPCR of the *gl3* expression in 14-day-old seedling leaves. Treatments with 2 replicates are shown for bacteria carrying either dT1, dT2, or the EV. The constitutively expressed *actin* gene was used for loading controls. **D)** RT-qPCR of the *gl3* expression at 6, 12, 24, and 48 hpi. The bar heights are the average of 3 biological replicates per treatment per time point. Error bars represent standard deviation. Values with the same letter do not differ at the significance level of 0.05 as determined by ANOVA and Tukey's honestly significant difference.

24 hpi, indicating that RNA-seq analysis 18 hpi had a limited ability to discover *gl3* downstream genes (Supplemental Data Set 2).

Of 10 known glossy genes, not including *gl15*, which is excluded due to the indirect role played in the cuticular wax biosynthesis pathway (Moose and Sisco 1996), 7 were among the 93 upregulated DEGs that were upregulated by both dTALes (*gl1*, *gl3*, *gl4*, *gl6*, *gl8*, *gl26*, and *cer8*), and additional 2 genes (*gl2*, *gl14*) were only upregulated by dT1 (Table 1). All 9 genes, regardless of significance, showed the same expression profile by both dTALes (Fig. 3D). The only glossy gene that was unaffected was *gl13*, which is an ABC transporter functioning in the secretion of cuticular waxes across the plasma membrane (Fig. 3D) (Li et al. 2013a). Additional genes upregulated by the dTALes included 6 genes encoding 3-ketoacyl-CoA synthases, of which *gl4* is also a member (Liu et al. 2009); 2 genes encoding HXXXD-type acyltransferase-related proteins (similar to *gl2*) (Tacke et al. 1995); 3 genes encoding GDSL esterase/lipase proteins, which were reported to be involved in wax biosynthesis (Tang et al. 2020); and 2 genes encoding aldehyde dehydrogenases (Supplemental Data Set 1). The 54 genes that were downregulated in association with both dT1 and dT2 in comparison to the EV did not include any known glossy genes.

Most glossy genes were previously reported to be clustered in a module (termed the turquoise module) of a gene co-expression network (GCN295), which was constructed using 295 RNA-seq data sets (Zheng et al. 2019). Of the 92 genes that did not include *gl3* and upregulated by dTALes, 61 were present in GCN295, and 38 of the 61 were assigned to the turquoise module. In contrast, only 3 genes from the 54 genes downregulated by dTALes were included in the turquoise module.

To determine the probability that a gene is regulated by GL3, we used a convolutional neural network (CNN) deep learning approach in conjunction with 739 publicly available RNA-seq data sets of the inbred line B73. To train the prediction model, the gene pairs corresponding to TFs and their targeted genes, which are mapped from *Arabidopsis* gene regulation data, were used as positive control pairs (Yilmaz et al. 2011). Random gene pairs that did not overlap with positive pairs were used as negative control. The deep learning approach predicted that 60% of GL3 associated genes were regulated by GL3, and, of the 594 control gene pairs that were unaffected by both dT1 and dT2, 18% were predicted to be regulated by GL3 (Supplemental Table S1, Supplemental Data Set 3). The CNN prediction supported the notion that most of the *gl3* downstream genes revealed were indeed regulated by *gl3*.

Probability-based identification of genes tightly regulated by GL3

A top-down Gaussian graphical model (GGM) algorithm using RNA-seq data from both 18 and 24 hpi samples was used to find genes whose expression was highly dependent on expression of *gl3*, assuming that they were more likely to be directly regulated by GL3 (Lin et al. 2013, Wei 2019,

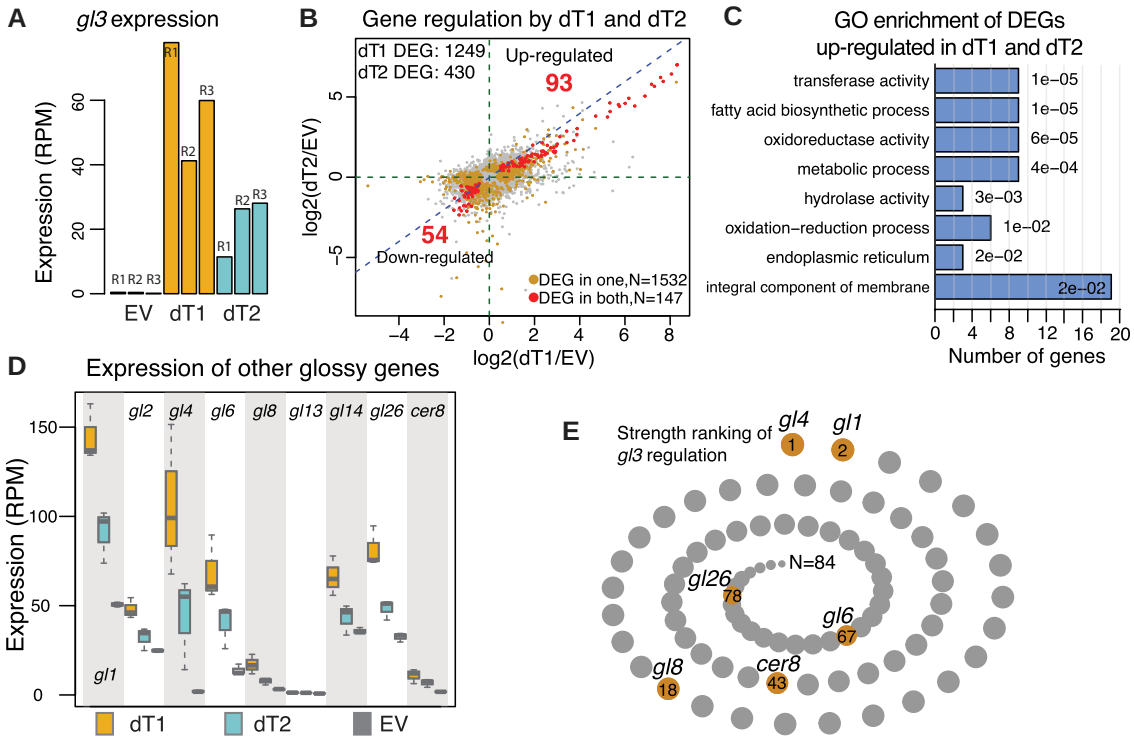


Figure 3. Gene expression associated with TALe-dependent expression of *gl3*. **A)** Expression in RPM (reads per million) of *gl3* from RNA-seq data. R1-R3: biological replicates. EV, dT1, dT2: constructs of EV, dT1, and dT2. **B)** Scatter plot between log2 fold changes of gene expression in the comparison of dT1 versus EV and that in the comparison of dT2 versus EV. The 93 genes upregulated by both dT1 and dT2 include *gl3*. Gray, orange, and red points represent unaffected, DEGs in 1 comparison, and DEGs in both comparisons, respectively. **C)** GO terms enriched in the DEGs in both comparisons. Numbers besides bars are *P*-values of GO enrichment tests. **D)** Expression in RPM of 9 glossy genes that affect cuticular wax accumulation in the 3 treatment groups. **E)** Ranking of regulation strengths by GL3 determined through a top-down GGM analysis. Dot sizes are correlated with regulation strengths. Known glossy genes are highlighted by orange circles with ranking numbers.

Table 1. Differential expression of known glossy genes

Gene	Glossy	dT1 vs. EV (24 hpi)			dT2 vs. EV (24 hpi)		
		Up fc[a]	*q*-value	Significant	Up fc[a]	*q*-value	Significant
Zm00001d020557	*gl1*	2.87	3.74E-32	Yes	1.91	2.57E-07	Yes
Zm00001d002353	*gl2*	1.94	1.26E-10	Yes	1.38	0.208	No
Zm00001d052397	*gl3*	191.2	4.57E-89	Yes	74.31	1.46E-18	Yes
Zm00001d051787	*gl4*	52.53	4.24E-95	Yes	22.99	1.37E-08	Yes
Zm00001d041578	*gl6*	4.96	1.59E-25	Yes	3.08	3.20E-08	Yes
Zm00001d017111	*gl8*	5.51	3.28E-20	Yes	2.63	1.61E-05	Yes
Zm00001d039631	*gl13*	1.57	0.482	No	1.40	0.906	No
Zm00001d004198	*gl14*	1.85	5.62E-07	Yes	1.28	0.481	No
Zm00001d008622	*gl26*	2.52	1.25E-20	Yes	1.58	2.82E-04	Yes
Zm00001d024723	*cer8*	5.93	6.13E-15	Yes	3.96	5.80E-08	Yes

[a]Upregulated fold change.

Wei et al. 2020). We focused on 92 genes presumably upregulated by *gl3*, and any 2 genes from these 92 were combined with *gl3* to form a triple gene block (see Methods) for evaluation. If the adjusted *P*-value with a multivariate delta method (see Methods) for each triple gene block was less than 0.05, *gl3* was scored as interfering with the 2 genes once. All triple gene blocks were evaluated, and the interference count for each gene was calculated. As a result, 84 genes that were strongly interfered with by *gl3* were identified (Fig. 3E, Supplemental Data Set 4), including 6 glossy genes: *gl4*, *gl1*, *gl8*, *cer8*, *gl6*, and *gl26*. The result supports that *gl3* expression had a strong impact on the expression of other genes in the pathway for cuticular wax biosynthesis.

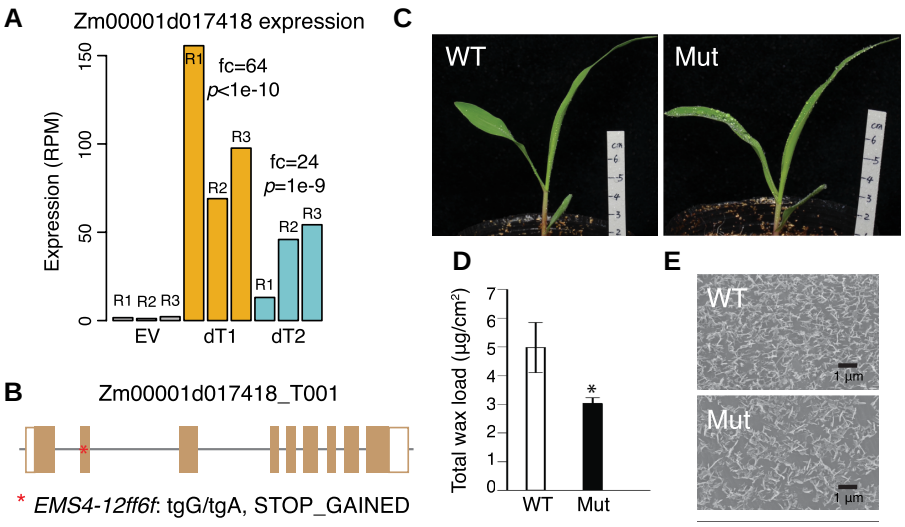


Figure 4. An EMS mutant of Zm00001d017418 has a glossy phenotype. **A)** Expression of the candidate gene Zm00001d017418. R1-R3: biological replicates; EV, dT1, and dT2: constructs of EV, dTALe 1, and dTALe 2; fc: fold change in expression relative to EV; *p*: adjusted *P*-value from RNA-seq analysis. **B)** Gene structure of the isoform of Zm00001d017418_T001. Boxes are exons and blank boxes represent untranslated regions. Asterisk points at the EMS mutation location, which produces a premature stop codon. **C)** Glossy phenotype of the EMS mutant and the WT. Water drops were present on the surfaces of mutant seedling leaves due to reduced epicuticular waxes. **D)** Total cuticular wax loads and wax components of mutants and WTs. Asterisk indicates the statistical significance between the 2 groups. **E)** Epicuticular wax content on the leaf in the WT and the mutant detected via SEM (×10,000 magnification).

A *gl3* downstream gene functions in cuticular wax accumulation

Due to the presence of most known glossy genes among the DEGs upregulated by *gl3*, the genes upregulated by dTALes may contain other, yet unknown, genes involved in biosynthesis of cuticular waxes. Genes that were upregulated by both dTALes and assigned to the turquoise module of GCN295 were selected as candidate glossy genes for subsequent validation. Ethyl methanesulfonate (EMS)-induced mutants of 4 candidate genes were obtained from a maize EMS mutant stock collection and screened for the glossy phenotype (Lu et al. 2018). No glossy phenotype was observed for mutants carrying premature stop codons in 3 of the genes, Zm00001d046642, Zm00001d028241, and Zm00001d032719, which encode a GDSL esterase/lipase, a 3-ketoacyl-CoA synthase, and a long-chain alcohol oxidase FAO4B, respectively (Supplemental Table S2). Zm00001d017418, which encodes an aldehyde dehydrogenase, was upregulated by both dT1 and dT2 (Fig. 4A). The EMS mutant (*ems4-12ff6f*), which had a premature stop codon in the second exon of Zm00001d017418, displayed a glossy phenotype, indicating reduced accumulation of cuticular waxes (Fig. 4, B and C). Total leaf waxes of *ems4-12ff6f* mutant plants were reduced by ∼40% of the amount found in the (WT) plants (Fig. 4D). Microscopic examination of wax components on the leaf surface revealed that fewer wax crystals had accumulated on leaf surfaces of mutant lines as compared to WT leaf surfaces (Fig. 4E). Detailed analysis of wax components revealed a decrease in C30 and longer-chain primary alcohols, alkanes, and fatty acids (Supplemental Fig. S8).

A second mutant allele of Zm00001d017418, designated as *20147-8*, was generated by CRISPR/Cas9 editing. At the 2 CRISPR targeting sites, the *20147-8* allele contained a 1-bp insertion and a 2-bp deletion in exons 1 and 4, respectively (Fig. 5A). Homozygous mutants ($N = 42$) displayed typical glossy phenotypes, with 24 mutants showing a strong glossy phenotype, for which the leaves were largely covered by water droplets, and 18 showing a weak glossy phenotype, meaning the water droplet density was present and visibly reduced, supporting the idea that Zm00001d017418 is a glossy gene involved in biosynthesis of cuticular waxes (Fig. 5B).

Discussion

In this work, the maize pathogen Xvv and the bacterial T3SS system were used for protein delivery into intact maize cells via TALes to characterize and identify genes involved in the cuticular wax pathway. Although typically considered destructive for host plant tissues in compatible interactions, *Xanthomonas* species can be considered as hemi-biotrophic because the pathogens interact with intact cells for some time before destruction of cells is evident. In our initial demonstration to determine the feasibility of the dTALe approach to modulate transcription in maize, the T3SS signal of the effector AvrBs2 was used to direct GFP to intact cells. A nuclear localization signal was added to the effector construct to concentrate the protein in the nucleus and to facilitate detection of the proteins in plant cells (Khang et al. 2010). To demonstrate its utility, this approach was used to study the consequences of ectopic expression of maize

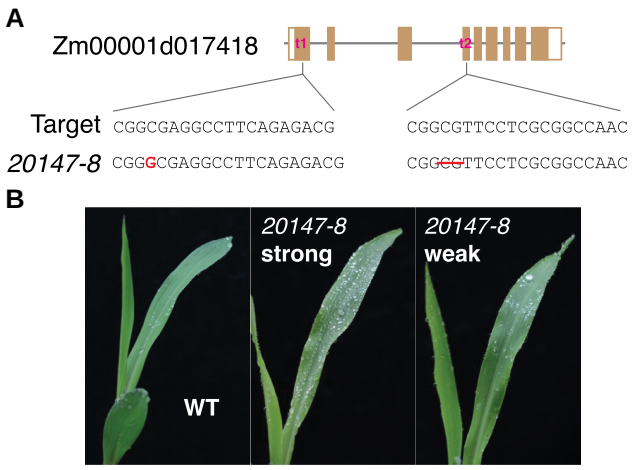


Figure 5. CRISPR mutants of Zm00001d017418 have glossy phenotypes. **A)** Two CRISPR targeting sites (t_1 and t_2) were designed. Details of a CRISPR mutant allele (*20147-8*) are provided. Red bases represent inserted bases. Sequences with red strike-throughs stand for deleted sequences. **B)** WT and mutants of the CRISPR allele showing either strong and weak glossy phenotypes.

GL3 by designing synthetic TFs known as dTALEs. TALEs are particularly useful for the approach as they already have T3SS secretion signal sequences and a nuclear localization signal to direct them into host cell nuclei (Nowack et al. 2022). Proteins that have evolved for transit through the T3SS do not require dimerization or other structural features that might constrain delivery from the bacteria to the host. Although Xvv does not contain endogenous TALEs, the presence of TALEs with biological functions in promoting plant disease in closely related strains, including TALEs that target host TFs, indicated that TALe delivery would be successful. Previous experience with strains of Xo that lack endogenous TALe genes has indicated that TALEs can be delivered by TALe-deficient strains (Tran et al. 2018). In this study, 2 dTALEs were designed to be targeted to 2 distinct DNA sequences in the promoter of *gl3*. Both dTALEs resulted in *gl3* induction, as shown by both RT-qPCR and RNA-seq analyses. In addition to observing modulated expression of *gl3* itself by dTALEs, we also discovered GL3-regulated genes that were identified due to dTALe activation of *gl3*. As clear evidence of this, 1 of the apparent GL3 downstream genes, Zm00001d017418, was upregulated and had the glossy phenotype, which caused reduced wax deposition on leaves when mutated. Note that the failure of mutants in the 3 other genes to cause a glossy phenotype does not indicate that these genes are not involved in GL3-dependent pathways, as the lack of the glossy phenotype upon mutation may be due to functional redundancy. In addition, GL3 downstream genes identified included most known glossy genes. The results indicate the master regulatory role of GL3 in biosynthesis of cuticular waxes and provide strong evidence for the efficacy of the dTALe system for studying gene regulation. Future experiments can be conducted to identify a possible DNA-binding motif of GL3. Also, given the fact that *gl3* expression is low at the adult stages of development, it will be interesting to examine the impacts of constitutive expression of *gl3* on wax biogenesis (Zheng et al. 2019). Using similar dTALe activation and/or knockout constructs, additional genes, particularly TFs downstream of GL3, can be examined to further elucidate the regulation network for cuticular wax biosynthesis.

The dTALe activation system is easy to manipulate, and *Xanthomonas* strains are easy to culture. Some strains, including Xvv, have a relatively wide host range in terms of either virulence or asymptomatic phenotypes (Hartman et al. 2020). Besides the simplicity of the system, it is easy to control the bacterial load by adjusting the concentration and amount of bacterial inoculum, which may fine-tune the efficacy of the approach. The dTALe activation method we presented here can detect transcriptional responses in a short period of time (e.g. 18 h), reducing the time scale of transgenic approaches. At the same time, several limitations should be considered for experimental design. First, multiple independent dTALEs are needed to reduce the potential impact of off-target gene induction. Use of multiple dTALEs can help to identify off-target gene induction, because targeting of independent binding sites are predicted to result in induction of the different off-target genes. Given that conserved domain or motif is required (Moscou and Bogdanove 2009), candidate dTALe-binding sites are relatively abundant. Second, bacterial infection and other T3SS effectors could potentially interfere with host gene expression or host protein function, if related to defense responses. Bacteria carrying an EV as the control, as implemented in this study, control for the impact from the bacteria themselves. Third, transcriptional regulation is a complex process, involving a concerted action of TFs and other cofactors (Spitz and Furlong 2012). Finally, the specific tissue type and the developmental stage of the host plant may be critical to ensure sufficient expression of cofactors necessary for downstream gene activation. In future experiments, the dTALe activation system can be tested at different stages of leaf development and in other tissues. With more tissue options, the tissue in which a TF may function could be chosen.

Genes downstream of a dTALe-targeted gene include direct or indirect targets of the dTALe-binding gene. Expression data of multiple samples with gradient activation could provide clues for direct vs. indirect regulation. The top-down GGM algorithm has been shown to discriminate directly regulated from indirectly regulated genes with more than 90% accuracy (Wei 2019; Lin et al. 2013), and it showed about 80% accuracy for RNA-seq data from stably transgenic lines (Wei et al. 2020). In this study, we found several types of variation of *gl3* expression induction: within a dTALe treatment (probably due to the variation in the bacterial amount injected during inoculation), between 2 dTALEs mimicking multiple levels of *gl3* induction as in time-course experiments, and between 2 time points of bacterial treatment. The data, therefore, enabled the top-down GGM algorithm to identify the genes that closely followed changes in *gl3*

expression, which are more likely to involve the genes directly regulated by *gl3*. Alternatively, the result from dTALe experiments can be combined with the results from DAP-Seq or ChIP-Seq experiments examining protein-binding sites to identify direct targets.

Xanthomonas bacteria can be used as a general tool for protein delivery to a variety of plant host cells because *Xanthomonas* bacterial strains are pathogens of numerous crops and have a well-documented ability to deliver diverse proteins into host plants. Gene activation using dTALes represents a unique and efficient system to study transcriptional regulation, especially in maize. This protein delivery system can also be utilized to study plant–pathogen interactions. For example, any effector gene can be engineered into the *Xanthomonas* bacterium and then delivered to host cells for examining defense responses. Also, to reduce pathogenic effects from *Xanthomonas*, the bacterium can be modified for reduced host cell toxicity by modifying pathogen-associated molecular pattern (PAMP) proteins and removing some Type III effector genes (Li et al. 2023).

Materials and methods

Genetic materials

Xvv strain Xv1601 was isolated from maize (Perez-Quintero et al. 2020). A *hrcC* knockout mutant was generated following the protocol previously described (Peng et al. 2016). The maize inbred lines A188 (PI 693339) and B73 (PI 550473) were obtained from the North Central Regional Plant Introduction Station and maintained at Kansas State University. Plants were grown in a growth chamber at 27 °C during the daytime and 21 °C at night with a 16-h/8-h light/dark photoperiod. Fluorescent lights were used, and the intensity of lights was between 375 and 400 μmol/m^2/s. EMS mutants were ordered from the Maize EMS-induced Mutant Database (MEMD) (Lu et al. 2018).

Design and assembly of protein delivery constructs

The pENTR 11 Dual Selection Vector (Thermo Fisher Scientific, USA) was digested with *Kpn*I and *Xho*I, and the DNA fragments containing the *AvrBs2* promoter and the Type III signal peptide (Mudgett et al. 2000) and eGFP coding sequences were synthesized (GenScript Biotech, USA) and cloned according to the NEBuilder HiFi DNA Assembly protocol (New England Biolabs, USA). The assembled entry construct was then cloned into the broad host-range vector modified from pHM1 (Hopkins et al. 1992) by the Gateway cloning system (Supplemental Fig. S2) and transformed into the Xv1601 strain by electroporation using a Bio-Rad MicroPulser (Peng et al. 2019).

Design and assembly of dTALes

The promoter elements targeted by TAL effectors are typically not far away from and upstream of transcriptional start sites (Moscou and Bogdanove 2009). Based on previous reports, most TAL effectors (e.g. PthA4, AvrBs3, PthXo2, PthXo3, AvrXa7, PthXo6, and PthXo7) bound to TATA-box regions, whereas some (e.g. PthXo1 and Tal8) targeted regions a few base pairs upstream of TATA boxes (Kay et al. 2007; Sugio et al. 2007; Antony et al. 2010; Hu et al. 2014; Zhou et al. 2015; Peng et al. 2019). The 2 dTALes created here, dT1 and dT2, were designed to specifically target a TATA-box region (dT1) and a region upstream of the TATA box (dT2) in the promoter of *gl3*. In addition, for proper design, a T nucleotide preceding each dTALe-binding element was required (Moscou and Bogdanove 2009). The Golden Gate TALEN assembly protocol was followed to construct these 2 dTALes (Cermak et al. 2011). Briefly, the kit (Golden Gate TALEN and TAL Effector Kit 2.0) consisting of 86 library vectors was ordered from Addgene (www.addgene.org). To assemble the dTALe harboring 16 repeats, a 10-repeat TAL array and a 5-repeat TAL array were constructed in the destination vectors pFUS_A and pFUS_B5, respectively. The 16 repeats targeting a 16-bp DNA sequence were designed to provide sufficient binding strength and specificity. The resulting vectors, the last-repeat plasmid, and the destination vector pTAL1 were digested with *Esp*3I (Thermo Fisher Scientific, USA) and ligated with T4 ligase (New England Biolabs, USA) to insert all TAL repeat arrays into the pTAL1 destination vector. The dTALes were then cloned into the broad host-range vector pHM1 and transformed into Xv1601 by electroporation using a Bio-Rad MicroPulser (Peng et al. 2019).

Bacterial culture and inoculation

Xv1601 bacteria were grown on tryptone sucrose agar medium at 28 °C (Peng et al. 2016). For plant inoculation, the inoculum was prepared using bacteria grown to OD_{600} of 0.2 to 0.3 and then suspended in water. The second leaf of 14-day-old seedlings of the inbred lines A188 or B73 was inoculated by infiltration with a needleless syringe. The region filled with bacterial solution was ~6 cm of the second leaf at a distance 2 cm away from the tip to the leaf base.

Generation and validation of the Xv1601 *hrcC*$^-$ mutant

The Xv1601*hrcC*$^-$ mutant was generated by PCR amplification, gene transfer, and homologous recombination. PCR primers XvhrcC-F and XvhrcC-R (Supplemental Table S3) were used to amplify partial sequence of the *hrcC* gene in Xv1601, which was then cloned into suicide vector pKnock-Gm (Addgene plasmid #46260). Bacterial conjugation was employed to transfer the suicide vector into Xv1601 for *hrcC*$^-$ mutants. The mutants were validated by PCR using primer XvhrcC-Out located upstream of the partial sequence and primer Forall-Val in the vector (Supplemental Table S3). The validation PCR fragment was sequenced to further confirm the insertion mutation in the *hrcC* gene.

GFP delivery via bacteria

The second leaf of 3-leaf maize seedlings was inoculated with *Xanthomonas* by infiltration with a needleless syringe.

After 24 or 48 h, leaf peels were mounted on a slide and stained with 20 μL of VECTASHIELD Antifade mounting medium (Vector Laboratories, USA) containing DAPI (4′,6-diamidino-2-phenylindole). DAPI staining was utilized to identify nuclei and monitor the nuclear localization of GFP, which was secreted by the bacterium. Confocal imaging was performed by visualizing leaf peels with a Zeiss LSM780 confocal microscope system using water immersion objectives, C-Apochromat 63×/1.2WCorr. Excitation/emission wavelengths were 488/550 nm for eGFP and 405/500 nm for DAPI. The detection wavelength was 493–598 nm for eGFP and 410–590 nm for DAPI. Image acquisition and processing and fluorescence intensity line scans were generated using the Zeiss ZEN 2010 software. The detector gain was set at 800 for eGFP and 400 for DAPI.

RT-qPCR to determine *gl3* expression at multiple seedling stages

Shoot or second-leaf samples from A188 seedlings were collected at 3, 4, 5, 8, and 14 days following seed germination. RNA was extracted from sample tissues using the Qiagen RNeasy Plant Mini Kit (Qiagen, Germany) and treated with DNase (Qiagen, Germany) to remove DNA contamination. First-strand cDNA was synthesized using the Thermo Scientific Verso cDNA Kit (Thermo Fisher Scientific, USA) with anchored oligo dT primers. RT-qPCR was performed with *gl3*-specific primers (Supplemental Table S3) and iQ SYBR Green Supermix (Bio-Rad, USA). Reactions were conducted on a Bio-Rad CFX with 96-well reaction blocks under the following PCR conditions: 95 °C for 3 min, followed by 40 cycles of 15 s at 95 °C and 30 s at 55 °C. An *actin* gene (GenBank accession AY273142) with the *actin* primers (Supplemental Table S3) was used as a reference gene to normalize *gl3* expression levels (Liu et al. 2009). The mean cycle threshold (Ct) values from technical replicates were used to calculate relative gene expression. Relative *gl3* expression was determined using the formula $100 \times 2\text{^}(Ct_{actin}\_Ct_{gl3})$, where Ct_{actin} and Ct_{gl3} represent the Ct values of *actin* and *gl3*, respectively.

RT-qPCR to determine *gl3* expression upon dTALe treatments

To examine induction of expression due to the 2 dTALes, the second leaf of 14-day-old seedlings was inoculated with bacteria containing dT1, dT2, or the EV negative control. Inoculated leaf tissues, except for the inoculation position, were collected at 24 hpi. Three plants undergoing the same treatment were pooled in a tissue sample. Bacteria treated with dT1 were used to examine the induction of expression at multiple time points after inoculation. Three biological replicates were conducted with 3 plants in each group. For time-series experiment, the inoculated leaf tissues were collected at 6, 12, 24, and 48 hpi. RT-qPCR as described above was performed for the quantification of *gl3* expression.

RNA-seq and data analysis

An RNA-seq experiment was performed to analyze *gl3* regulation of downstream genes. Bacteria grown to the 0.2–0.3 OD_{600} range and suspended in water were used as the inoculum. The second leaf of 14-day-old seedlings was inoculated with a needleless syringe. Three biological replicates (R1, R2, and R3) were conducted for each of the 3 treatment groups, in which bacteria carried either dT1, dT2, or the EV. Inoculated leaf tissues for all 3 treatments were collected at 24 hpi. In addition, inoculated leaf tissues of dT1 and EV treatments were collected at 18 hpi. Three plants with the same treatment were pooled in a tissue sample. As a result, 15 tissue samples were collected in total. RNA was extracted from sampled tissues with the Qiagen RNeasy Plant Mini Kit. Sequencing libraries were prepared and sequenced on a NovaSeq 6000 Illumina platform at Novogene Inc. (Novogene, USA). Adaptor sequences and low-quality bases of raw reads were trimmed by Trimmomatic (version 0.38) (Bolger et al. 2014). Trimmed reads were aligned to the B73 reference genome (B73Ref4) (Jiao et al. 2017, Schnable et al. 2009) using STAR (2.7.3a) (Dobin et al. 2013). Uniquely mapped reads were used for counting reads per gene. DESeq2 (version 1.26.0) was used to identify DEGs between each of the 2 dTALe groups (dT1 and dT2) and the EV group. Differential analysis was performed separately for samples from 18 and 24 hpi. Multiple tests were performed to determine the FDR with the FDR cutoffs of 5% for dT1 and 10% for dT2 (Benjamini and Hochberg 1995).

Glossy phenotyping

The glossy phenotype was identified by spraying water on seedlings at the 2 or 3 leaf stage. Seedlings whose leaves were covered with small water droplets were identified as glossy mutants. Leaves were considered to show a strong glossy phenotype, if a leaf was fully covered by water droplets.

SEM

The middle region of the second leaves from *ems4-12ff6f* mutant and WT plants was cut into 1 cm × 1 cm squares and used for scanning electron microscopy (SEM) analysis. Collected leaves were fixed in 2.5% glutaraldehyde for 24 h at room temperature and then washed with phosphate buffer (pH 7.2) thoroughly. Samples were post-fixed in 1% osmium tetroxide for 2 h, dehydrated with gradient alcohol solutions and then treated with isoamyl acetate for 12 h. Samples were then dried in a critical point dryer (LEICA EM CPD 030, LEICA, Germany) and treated with an ion sputtering instrument (Eiko IB5 Ion Coater, Eiko, Japan) and observed with scanning electron microscope (SU8010, HITACHI, Japan) (Aharoni et al. 2004).

Analysis of wax composition

Wax extraction and gas chromatography-mass spectrometry (GC-MS) analyses were performed according to previously described methods with some modifications (Chen et al.

2017). The *ems4-12ff6f* mutant and WT plants were grown in the substrate of roseite (Beijing Feng Rui Jia Ye Science and Technology, China) and sand (1:1) at a growth chamber (25 °C) until the 3-leaf stage. The second leaves (about 300 mg) were collected and immersed for 1 min in 3 mL of chloroform that contained 15 μg of dissolved nonadecanoic acid (C19) as an internal standard. Extracted solutions were transferred into new vials and evaporated under a gentle stream of nitrogen gas. The residue was derivatized with 100 μL of *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide and incubated for 1 h at 50 °C. These derivatized samples were then analyzed by GC-MS (Agilent gas chromatograph coupled to an Agilent 5975C quadrupole mass selective detector). The area of leaves was calculated by IMAGEJ software (http://imagej.nih.gov/ij/), and the amount of leaf wax was presented per unit of surface area.

Prediction of *gl3*-regulated genes through deep learning

In total, 739 B73 paired-end RNA-seq data sets from diverse tissues and treatments were collected from the NCBI Sequence Read Archive (SRA) database (Supplemental Data Set 5). Software Trimmomatic (version 0.38) (Bolger et al. 2014) was used to trim the adaptor sequence and remove low-quality bases of raw reads. The remaining paired-end reads were aligned to the B73Ref4 (Jiao et al. 2017) using STAR (version 2.6.0), requiring concordant mapping positions of paired-end reads (Dobin et al. 2013). Raw read counts per gene were calculated by STAR and then normalized using the library sizes of RNA-seq samples to represent gene expression.

The 2,140 pairs of TFs and their putative targeted gene in maize obtained by homologous mapping of Arabidopsis experimentally verified regulatory gene pairs from the Arabidopsis Regulatory Network database (Yilmaz et al. 2011) were used as the positive training data set for deep learning. To generate a negative data set, we randomly generated 2,140 gene pairs that do not contain above positive relationships. The maize transcriptomic data of these 4,280 gene pairs was used for training the CNN model for predictions. The input data sets of these 4,280 pairs of genes were from 739 B73 RNA-seq data. Therefore, the data set contained 4,280 gene pairs and each gene had 739 features. Four-hundred and twenty-eight gene pairs and their expression data, which account for 10% of the whole data set, were randomly chosen and used as the validation data set. The test data set, which contained *gl3* versus 146 *gl3* associated genes, and *gl3* versus 594 *gl3*-unaffected genes, was extracted from the 739 RNA-seq data set.

Besides expression data, 2 additional dimensions, the product and the absolute difference of each pair of 2 genes, *gl3* and a putative target gene, were calculated and added as additional features. We employed Keras and TensorFlow libraries to develop the CNNs using R libraries (Abadi et al. 2015, Chollet 2018). The architecture of the CNN includes 3 components: the input, the feature extractor, and the classifier. The feature extractor contains several building blocks, each containing a convolution layer and a pooling layer. A convolution layer consists of multiple filters that help identify features and activation functions that convert linear input to nonlinear output. A pooling layer provided the down-sampling operation to reduce the dimensions of the feature map. The classifier is made up of a flatten layer and several fully connected layers, and each FC layer is followed by an activation function. The flatten layer takes the results from feature extractor process and flattens them into a single long vector that can be an input for the next fully connected layer, which applies the weights of input to predict the true regulatory relationships and delivers the final output of the network as represented by probability for each pair of genes for prediction. To identify a model with a high performance for the prediction, we tried multiple loss functions. The mean squared logarithmic error (MSLE) loss was selected as the loss function.

Inference of GL3-regulated target genes using the top-down GGM algorithm

The top-down GGM algorithm developed earlier (Wei 2019; Lin et al. 2013) was employed to construct a multilayered gene regulatory network (ML-hGRN) mediated by GL3 in 2 steps, using the dT1, dT2, and EV RNA-seq data as the input data. Briefly, in the first step, the GL3 downstream genes that responded to the *gl3* activation were identified using Fisher's exact test (Fisher 1934) and the probability-based method as we described in our publications mentioned above, and these genes were termed *gl3* responsive genes. In the second step, we further identified those that were interfered with frequently by GL3 from the *gl3* responsive genes by evaluating all triple gene blocks, each consisting of *gl3*, defined as *z*, and 2 *gl3* responsive genes, defined as *x* and *y*. If GL3 significantly interfered with the 2 responsive genes in a triple gene block, the difference (d) between the correlation coefficient, r_{xy}, of 2 responsive genes in expression and the partial correlation coefficient, $r_{xy/z}$, representing the correlation of 2 *gl3* responsive genes conditioning on *gl3* (z) should be significant. The null hypothesis H_0: $d = 0$ was tested with the multivariate delta method (MacKinnon et al. 2002). If d was significantly different from 0, *gl3* was defined as interfering with the 2 responsive genes, and their regulatory relationships were recorded once. After all combined triple gene blocks were evaluated, the interference frequency between *gl3* and each responsive gene was calculated. In this study, the candidate target genes with at least 1 interference frequency were considered to be a gene directly regulated by *gl3*.

Plasmid construction for CRISPR-Cas9 editing of Zm00001d017418

Two 20-bp gRNA sequences (target site 1: CGTCTCTGAAGG CCTCGCCG in the first exon, target site 2: GTTGGCCGC GAGGAACGCCG in the fourth exon of Zm00001d017418) were designed using the online design tool, CRISPR-P 2.0

(http://crispr.hzau.edu.cn/CRISPR2/). Then, the 2 gRNA sequences were cloned into the *Bsa*I site of pBUE411 (Xing et al. 2014). The resulting plasmid was transferred into *Agrobacterium tumefaciens* strain LBA4401 and then transformed into maize immature embryos of inbred line Cal as described previously (Liu et al. 2015).

Identification of CRISPR-edited mutations

Genomic DNA was extracted from T_0 and T_1 transgenic plants. The CRISPR-targeted sites were amplified with primers T1F/T1R and T2F/T2R for Sanger sequencing.

Phylogenetic analysis

DNA sequences of 4 housekeeping genes (*DnaK*, *gyrB*, *GroEL*, and *RecA*) from 10 *Xanthomonas* genomes were used to build the phylogenetic tree (Supplemental Files 1 and 2). Ten genomes include *Xanthomonas albilineans* GPE PC73 (GenBank accession: FP565176.1), *Xanthomonas axonopodis* pv. *citri* str. 306 (GenBank accession: AE008923.1), *Xanthomonas campestris* pv. *campestris* str. 8004 (GenBank accession: CP000050.1), *X. campestris* pv. *raphani* 756C (GenBank accession: CP002789.1), *X. campestris* pv. *vesicatoria* str. 85-10 (GenBank accession: CP017190.1), *X. oryzae* pv. *oryzae* PXO99^A (GenBank accession: CP000967.2), *X. oryzae* pv. *oryzicola* BLS256 (GenBank accession: CP003057.2), *Xanthomonas translucens* pv. *undulosa* strain P3 (GenBank accession: CP043500.1), Xvv strain Xv1601 (GenBank accession: CP025272.1), and *X. axonopodis* pv. *manihotis* XAM668 (GenBank accession: GCA_000266665.1). DNA sequences of these 4 genes from each genome were concatenated, followed by multiple alignments with MAFFT (v7.505) (Katoh et al. 2002). The multiple alignments were input to IQ-TREE2 (v2.2.0) to construct a maximum likelihood phylogenetic tree with 1,000 bootstraps (Nguyen et al. 2015). The optimal model ("GTR + F + I + G4") identified by using IQ-TREE2 was used in generating the final tree.

Statistical analysis

Statistical analyses were performed as described in each figure legend. Statistical data are provided in Supplemental Table S4.

Acknowledgments

We thank Fengxia Zhang from the Metabolomics Facility of the Institute of Genetics and Developmental Biology at Chinese Academy of Sciences for analysis of wax composition, Wei Wang from the Eduard Akhunov laboratory, and Melinda Dalby from the Barbara Valent laboratory at Kansas State University for assistance with confocal microscope experiments.

Author contributions

J.Z., M.C., S.P., H.W., F.F.W, and S.L. conceived and designed experiments. M.Z., Z.P., Y.Q., T.M.T., B.T., Y.C., J.Z., G.L., H.Z., K.L., A.K., C.M., F.H., H.T., Y.L., and J.Z. performed experiments and collected data. M.Z., Z.P., Y.Q., L.Z., C.H., H.W., and S. L. analyzed data. M.Z., Z.P., Y.Q., L.Z., Y.L., M.C., S.P., J.Z., H.W., F.F.W., and S.L. wrote the manuscript with comments from other authors.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Phylogenetic tree of *Xanthomonas* strains.

Supplemental Figure S2. Schematic of the eGFP construct.

Supplemental Figure S3. GFP delivery to the nuclei by Xvv.

Supplemental Figure S4. Time course of *gl3* expression in maize seedlings.

Supplemental Figure S5. TALe-dependent *gl3* expression.

Supplemental Figure S6. Two additional genes specifically activated by dTALe.

Supplemental Figure S7. Read distributions on the *gl3* locus.

Supplemental Figure S8. Epidermal wax composition in WT and the mutant.

Supplemental Table S1. Result summary of CNN deep learning.

Supplemental Table S2. Four stop-gained EMS mutants of candidate genes.

Supplemental Table S3. List of the primers used.

Supplemental Table S4. Statistical data.

Supplemental Data Set 1. Detailed information of DEGs at 24 hpi.

Supplemental Data Set 2. Detailed information of DEGs upregulated by dTALe treatment at 18 hpi.

Supplemental Data Set 3. Deep learning classification for *gl3* downstream genes and dTALe unaffected genes.

Supplemental Data Set 4. Highly *gl3*-responsive genes identified in the top-down GGM analysis.

Supplemental Data Set 5. List of 739 B73 RNA-seq data accessions used for deep learning.

Supplemental File 1. Alignment used for phylogenetic analysis.

Supplemental File 2. Newick tree file for phylogenetic analysis.

Funding

We thank funding support from the National Science Foundation (awards no. 1741090 and 2011500), the National Institute of Food and Agriculture (award no. 2018-67013-28511 and 2021-67013-35724), the Department of Energy (award no. DE-SC0023138), and the Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences. This is contribution number 21-196-J from the Kansas Agricultural Experiment Station.

Conflict of interest statement. None declared.

Data availability

Raw dTALe RNA-seq data are available at NCBI SRA under the project of PRJNA692729.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z. TensorFlow: Large-scale machine learning on heterogeneous systems. 2015. tesorflow.org.

Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A. The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis[W]. Plant Cell 2004:**16**(9): 2463–2480. https://doi.org/10.1105/tpc.104.022897

Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene os-*11N3*. Plant Cell 2010:**22**(11): 3864–3876. https://doi.org/10.1105/tpc.110.078964

Bartlett A, O'Malley RC, Huang S-SC, Galli M, Nery JR, Gallavotti A, Ecker JR. Mapping genome-wide transcription-factor binding sites using DAP-seq. Nat Protoc. 2017:**12**(8):1659–1672. https://doi.org/10.1038/nprot.2017.055

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995:**57**(1): 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Block A, Li G, Fu ZQ, Alfano JR. Phytopathogen type III effector weaponry and their plant targets. Curr Opin Plant Biol. 2008:**11**(4): 396–403. https://doi.org/10.1016/j.pbi.2008.06.007

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 2009:**326**(5959):1509–1512. https://doi.org/10.1126/science.1178811

Bogdanove AJ. Principles and applications of TAL effectors for plant physiology and metabolism. Curr Opin Plant Biol. 2014:**19**:99–104. https://doi.org/10.1016/j.pbi.2014.05.007

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014:**30**(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Büttner D, Bonas U. Regulation and secretion of *Xanthomonas* virulence factors. FEMS Microbiol Rev. 2010:**34**(2):107–133. https://doi.org/10.1111/j.1574-6976.2009.00192.x

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011:**39**(12):e82. https://doi.org/10.1093/nar/gkr218

Chen X, Zhang H, Sun H, Luo H, Zhao L, Dong Z, Yan S, Zhao C, Liu R, Xu C, et al. IRREGULAR POLLEN EXINE1 is a novel factor in anther cuticle and pollen exine formation. Plant Physiol. 2017:**173**(1): 307–325. https://doi.org/10.1104/pp.16.00629

Chollet F. Keras: the python deep learning library. Astrophysics Source Code Library: ascl. 2018:**1806**:022. https://ui.adsabs.harvard.edu/abs/2018ascl.soft06022C/abstract.

Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. Secretion systems in gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 2015:**13**(6):343–359. https://doi.org/10.1038/nrmicro3456

Cunningham FJ, Goh NS, Demirer GS, Matos JL, Landry MP. Nanoparticle-mediated delivery towards advancing plant genetic engineering. Trends Biotechnol. 2018:**36**(9):882–897. https://doi.org/10.1016/j.tibtech.2018.03.009

Demirer GS, Zhang H, Goh NS, Pinals RL, Chang R, Landry MP. Carbon nanocarriers deliver siRNA to intact plant cells for efficient gene knockdown. Sci Adv. 2020:**6**(26):eaaz0495. https://doi.org/10.1126/sciadv.aaz0495

Deslandes L, Rivas S. Catch me if you can: bacterial effectors and plant targets. Trends Plant Sci. 2012:**17**(11):644–655. https://doi.org/10.1016/j.tplants.2012.06.011

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013:**29**(1):15–21. https://doi.org/10.1093/bioinformatics/bts635

Fehling E, Mukherjee KD. Acyl-CoA elongase from a higher plant (*Lunaria annua*): metabolic intermediates of very-long-chain acyl-CoA products and substrate specificity. Biochim Biophys Acta. 1991:**1082**(3):239–246. https://doi.org/10.1016/0005-2760(91)90198-Q

Fisher RA. Statistical methods for research workers chap. 5. Statistical methods for research workers. Edinburgh: Oliver & Boyd; 1934.

Gleba YY, Tusé D, Giritch A. Plant viral vectors for delivery by *Agrobacterium*. Curr Top Microbiol Immunol. 2014:**375**:155–192. https://doi.org/10.1007/82_2013_352

Green ER, Mecsas J. Bacterial secretion systems: an overview. Microbiol Spectr. 2016:**4**(1):1–32. https://doi.org/10.1128/microbiolspec.VMBF-0012-2015

Hansen JD, Pyee J, Xia Y, Wen TJ, Robertson DS, Kolattukudy PE, Nikolau BJ, Schnable PS. The glossy1 locus of maize and an epidermis-specific cDNA from *Kleinia odora* define a class of receptor-like proteins required for the normal accumulation of cuticular waxes. Plant Physiol. 1997:**113**(4):1091–1100. https://doi.org/10.1104/pp.113.4.1091

Hartman T, Tharnish B, Harbour J, Yuen GY, Jackson-Ziems TA. Alternative hosts in the families Poaceae and Cyperaceae for *Xanthomonas vasicola* pv. *vasculorum*, causal agent of bacterial leaf streak of corn. Phytopathology 2020:**110**(6):1147–1152. https://doi.org/10.1094/PHYTO-04-19-0132-R

Hopkins CM, White FF, Choi SH, Guo A, Leach JE. Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact. 1992:**5**(6):451–459. https://doi.org/10.1094/MPMI-5-451

Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, Yang B, White FF, Wang N, Jones JB. *Lateral organ boundaries 1* is a disease susceptibility gene for citrus bacterial canker disease. Proc Nat Acad Sci U S A. 2014:**111**(4):E521–E529. https://doi.org/10.1073/pnas.1313271111

Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin CS, et al. Improved maize reference genome with single-molecule technologies. Nature 2017:**546**(7659):524–527. https://doi.org/10.1038/nature22971

Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol. 2013:**14**(1):49–55. https://doi.org/10.1038/nrm3486

Katoh K, Misawa K, Kuma K-I, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002:**30**(14):3059–3066. https://doi.org/10.1093/nar/gkf436

Kay S, Hahn S, Marois E, Hause G, Bonas U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 2007:**318**(5850):648–651. https://doi.org/10.1126/science.1144956

Khang CH, Berruyer R, Giraldo MC, Kankanala P, Park S-Y, Czymmek K, Kang S, Valent B. Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell 2010:**22**(4):1388–1403. https://doi.org/10.1105/tpc.109.069666

Kunst L, Samuels AL. Biosynthesis and secretion of plant cuticular wax. Prog Lipid Res. 2003:**42**(1):51–80. https://doi.org/10.1016/S0163-7827(02)00045-0

Lai X, Stigliani A, Vachon G, Carles C, Smaczniak C, Zubieta C, Kaufmann K, Parcy F. Building transcription factor binding site models to understand gene regulation in plants. Mol Plant. 2019:**12**(6):743–763. https://doi.org/10.1016/j.molp.2018.10.010

Lee SB, Suh MC. Recent advances in cuticular wax biosynthesis and its regulation in Arabidopsis. Mol Plant. 2013:**6**(2):246–249. https://doi.org/10.1093/mp/sss159

Li L, Du Y, He C, Dietrich CR, Li J, Ma X, Wang R, Liu Q, Liu S, Wang G, et al. Maize glossy6 is involved in cuticular wax deposition and drought tolerance. J Exp Bot. 2019:**70**(12):3089–3099. https://doi.org/10.1093/jxb/erz131

Li T, Huang S, Zhou J, Yang B. Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. oryzae in rice. Mol Plant. 2013b:**6**(3):781–789. https://doi.org/10.1093/mp/sst034

Li L, Li D, Liu S, Ma X, Dietrich CR, Hu H-C, Zhang G, Liu Z, Zheng J, Wang G, et al. The maize glossy13 gene, cloned via BSR-seq and seq-walking encodes a putative ABC transporter required for the normal accumulation of epicuticular waxes. PLoS One 2013a:**8**(12):e82333. https://doi.org/10.1371/journal.pone.0082333

Li C, Wang L, Cseke LJ, Vasconcelos F, Huguet-Tapia JC, Gassmann W, Pauwels L, White FF, Dong H, Yang B. Efficient CRISPR-cas9 based cytosine base editors for phytopathogenic bacteria. Commun Biol. 2023:**6**(1):56. https://doi.org/10.1038/s42003-023-04451-8

Lin YC, Li W, Sun YH, Kumari S, Wei H, Li Q. SND1 Transcription factor–directed quantitative functional hierarchical genetic regulatory network in wood formation in Populus trichocarpa. The Plant Cell 2013:**25**(11): 4324–4341. https://doi.org/10.1105/tpc.113.117697

Liu S, Dietrich CR, Schnable PS. DLA-based strategies for cloning insertion mutants: cloning the *gl4* locus of maize using *Mu* transposon tagged alleles. Genetics 2009:**183**(4):1215–1225. https://doi.org/10.1534/genetics.109.108936

Liu S, Yeh C-T, Tang HM, Nettleton D, Schnable PS. Gene mapping via bulked segregant RNA-seq (BSR-seq). PLoS One 2012:**7**(5): e36406. https://doi.org/10.1371/journal.pone.0036406

Liu Y, Zhang Y, Liu Y, Lu W, Wang G. Metabolic effects of glyphosate on transgenic maize expressing a G2-EPSPS gene from *Pseudomonas fluorescens*. J Plant Biochem Biotechnol. 2015:**24**(2):233–241. https://doi.org/10.1007/s13562-014-0263-9

Lu X, Liu J, Ren W, Yang Q, Chai Z, Chen R, Wang L, Zhao J, Lang Z, Wang H, et al. Gene-indexed mutations in maize. Mol Plant. 2018:**11**(3):496–504. https://doi.org/10.1016/j.molp.2017.11.013

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods. 2002:**7**(1):83–104. https://doi.org/10.1037/1082-989X.7.1.83

Mew TW. Current status and future prospects of research on bacterial blight of rice. Annu Rev Phytopathol. 1987:**25**(1):359–382. https://doi.org/10.1146/annurev.py.25.090187.002043

Moose SP, Sisco PH. Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. Genes Dev. 1996:**10**(23): 3018–3027. https://doi.org/10.1101/gad.10.23.3018

Morbitzer R, Römer P, Boch J, Lahaye T. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci U S A. 2010:**107**(50):21617–21622. https://doi.org/10.1073/pnas.1013133107

Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009:**326**(5959):1501. https://doi.org/10.1126/science.1178817

Mudgett MB, Chesnokova O, Dahlbeck D, Clark ET, Rossier O, Bonas U, Staskawicz BJ. Molecular signals required for type III secretion and translocation of the *Xanthomonas campestris* AvrBs2 protein to pepper plants. Proc Natl Acad Sci U S A. 2000:**97**(24): 13324–13329. https://doi.org/10.1073/pnas.230450797

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015:**32**(1):268–274. https://doi.org/10.1093/molbev/msu300

Nowack MK, Holmes DR, Lahaye T. TALE-induced cell death executors: an origin outside immunity? Trends Plant Sci. 2022:**27**(6): 536–548. https://doi.org/10.1016/j.tplants.2021.11.003

Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom JS, Li C, Nguyen H, Liu B, et al. Broad-spectrum resistance to bacterial blight in rice using genome editing. Nat Biotechnol. 2019:**37**(11):1344–1350. https://doi.org/10.1038/s41587-019-0267-z

Peng Z, Hu Y, Xie J, Potnis N, Akhunova A, Jones J, Liu Z, White FF, Liu S. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. BMC Genomics 2016:**17**(1):21. https://doi.org/10.1186/s12864-015-2348-9

Peng Z, Hu Y, Zhang J, Huguet-Tapia JC, Block AK, Park S, Sapkota S, Liu Z, Liu S, White FF. *Xanthomonas translucens* commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. Proc Natl Acad Sci U S A. 2019:**116**(42):20938–20946. https://doi.org/10.1073/pnas.1911660116

Perez-Quintero AL, Ortiz-Castro M, Lang JM, Rieux A, Wu G, Liu S, Chapman TA, Chang C, Ziegle J, Peng Z, et al. Genomic acquisitions in emerging populations of *Xanthomonas vasicola* pv. *vasculorum* infecting corn in the United States and Argentina. Phytopathology 2020:**110**(6):1161–1173. https://doi.org/10.1094/PHYTO-03-19-0077-R

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. The B73 maize genome: complexity, diversity, and dynamics. Science 2009:**326**(5956):1112–1115. https://doi.org/10.1126/science.1178534

Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012:**13**(9):613–626. https://doi.org/10.1038/nrg3207

Streubel J, Pesce C, Hutin M, Koebnik R, Boch J, Szurek B. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. New Phytol. 2013:**200**(3):808–819. https://doi.org/10.1111/nph.12411

Sugio A, Yang B, Zhu T, White FF. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIA* γ1 and *OsTFX1* during bacterial blight of rice. Proc Natl Acad Sci USA. 2007:**104**(25):10720–10725. https://doi.org/10.1073/pnas.0701742104

Tacke E, Korfhage C, Michel D, Maddaloni M, Motto M, Lanzini S, Salamini F, Döring HP. Transposon tagging of the maize glossy2 locus with the transposable element en/spm. Plant J. 1995:**8**(6): 907–917. https://doi.org/10.1046/j.1365-313X.1995.8060907.x

Tang J, Yang X, Xiao C, Li J, Chen Y, Li R, Li S, Lü S, Hu H. GDSL Lipase occluded stomatal pore 1 is required for wax biosynthesis and stomatal cuticular ledge formation. New Phytol. 2020:**228**(6):1880–1896. https://doi.org/10.1111/nph.16741

Tran TT, Pérez-Quintero AL, Wonni I, Carpenter SC, Yu Y, Wang L, Leach JE, Verdier V, Cunnac S, Bogdanove AJ, et al. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. PLoS Pathog. 2018:**14**(6):e1007092. https://doi.org/10.1371/journal.ppat.1007092

Van den Ackerveken G, Marois E, Bonas U. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. Cell 1996:**87**(7):1307–1316. https://doi.org/10.1016/S0092-8674(00)81825-5

Wei H. Construction of a hierarchical gene regulatory network centered around a transcription factor. Brief Bioinform. 2019:**20**(3): 1021–1031. https://doi.org/10.1093/bib/bbx152

Wei M, Liu Q, Wang Z, Yang J, Li W, Chen Y, Lu H, Nie J, Liu B, Lv K, et al. PuHox52-mediated hierarchical multilayered gene regulatory network promotes adventitious root formation in *Populus ussuriensis*. New Phytol. 2020:**228**(4):1369–1385. https://doi.org/10.1111/nph.16778

White FF, Potnis N, Jones JB, Koebnik R. The type III effectors of *Xanthomonas*. Mol Plant Pathol. 2009:**10**(6):749–766. https://doi.org/10.1111/j.1364-3703.2009.00590.x

Xing H-L, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, Wang X-C, Chen Q-J. A CRISPR/cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol. 2014:**14**(1):327. https://doi.org/10.1186/s12870-014-0327-y

Xu X, Dietrich CR, Delledonne M, Xia Y, Wen TJ, Robertson DS, Nikolau BJ, Schnable PS. Sequence analysis of the cloned glossy8 gene of maize suggests that it may code for a [beta]-ketoacyl

reductase required for the biosynthesis of cuticular waxes. Plant Physiol. 1997:**115**(2):501–510. https://doi.org/10.1104/pp.115.2.501

Yang B, Sugio A, White FF. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci U S A. 2006:**103**(27): 10503–10508. https://doi.org/10.1073/pnas.0604088103

Yang B, Zhu W, Johnson LB, White FF. The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. Proc Natl Acad Sci U S A. 2000:**97**(17):9807–9812. https://doi.org/10.1073/pnas.170286897

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E. AGRIS: the *Arabidopsis* gene regulatory information server, an update. Nucleic Acids Res. 2011:**39**(Database):D1118–D1122. https://doi.org/10.1093/nar/gkq1120

Zheng J, He C, Qin Y, Lin G, Park WD, Sun M, Li J, Lu X, Zhang C, Yeh CT, et al. Co-expression analysis aids in the identification of genes in the cuticular wax pathway in maize. Plant J. 2019:**97**(3): 530–542. https://doi.org/10.1111/tpj.14140

Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom J-S, Huang S, Liu S, Vera Cruz C, Frommer WB, et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. Plant J. 2015:**82**(4):632–643. https://doi.org/10.1111/tpj.12838

Zhu W, Yang B, Chittoor JM, Johnson LB, White FF. Avrxa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. Mol Plant Microbe Interact. 1998:**11**(8): 824–832. https://doi.org/10.1094/MPMI.1998.11.8.824