7-5-2023

Investigating Cassava Susceptibility to Xanthomonas Induced Bacterial Blight Using Disease Phenotyping and Genome Engineering Strategies

Kiona Rajene Elliott
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Division of Biology and Biomedical Sciences
Plant and Microbial Biosciences

Dissertation Examination Committee:
Rebecca Bart, Chair
Barbara Kunkel
Dmitri Nusinow
Rachel Penczykowski
Nigel Taylor
Joe Vogel

Investigating Cassava Susceptibility to *Xanthomonas* Induced Bacterial Blight Using Disease Phenotyping and Genome Engineering Strategies

by
Kiona R. Elliott

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2023
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CBB = Cassava Bacterial Blight

Me = *Manihot esculenta*

Pv. = Pathovar

Xpm = *Xanthomonas phaseoli* (pv.) *manihotis*

Xam = *Xanthomonas axonopodis* (pv.) *manihotis*

TAL = Transcription Activator-Like

EBE = Effector Binding Element

RVDs = Repeat Variable Di-residues

SWEET = Sugars Will Eventually be Exported Transporter

PLL = Pectate Lyase-Like

CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats

gRNA = Guide RNA

SVM =Support Vector Machine Learning

DPI = Days Post Inoculation

HPI = Hours Post Inoculation

Acknowledgments

First, I would like to thank Washington University and the Donald Danforth Plant Science Center (DDPSC) communities. I am proud to have completed my graduate training at these top-tier institutions. Thank you to the funding sources that have supported me including the NSF Graduate Research Fellowship Program, Gates Millennium Scholarship, Washington University William H. Danforth Plant Science Fellowship, and the DDPSC Danforth Fellowship. Thank you to the student and professional groups that provided me with valuable opportunities and enhanced my skillsets outside of the lab: Initiative to Maximize Student Development (IMSD), Committee for Scientific Training and Mentoring (CSTM), GradCast, and ProSPER.

A dissertation is never completed alone and I would like to thank all the people who have made this work possible. Thank you to my thesis advisor, Rebecca Bart who encouraged me at every step of my dissertation and in my professional development beyond the bench. I have learned a lot about being a good scientist and leader from you. Thank you for being a mentor and advocate, and for cultivating a lab environment that I was happy to walk into every day.

Thank you to the people (both past and present) who made the lab a welcoming and collaborative place. Thank you for always being willing to answer questions, talk through an experiment or scientific concept, and have impromptu lab lunches. Special thanks to Kira, Greg, Dan, Qi, Jeff, and Ben who provided me with much guidance and training over the years. Thank you to Marisa-lab manager extraordinaire, and to my fellow Bart lab grad student comrades: Anne, Taylor, Nastya, Ketra, and Ray.

I would also like to thank my thesis committee members: Barbara Kunkel, Dmitri Nusinow, Rachel Penczykowski, Nigel Taylor, and Joe Vogel. Your thoughtful advice, feedback, and critique of my work over the years have been extremely valuable. I left every committee meeting feeling energized and excited about the direction of my research.

Thank you to my incredible support system of friends who have fortified me. Special thanks to my lifelong friends, Clara and Raiesa, for the long-distance calls, video chats, and messages that I always look forward to. To the PMB girls or "Core Four"-Erin, Natasha, and Kari, you have my sincerest gratitude for the wine nights, girls' trips, venting sessions, and dinner parties that have been a true highlight of graduate school. To Mahliyah, Emilee, Ryan, and M.R., thank you for the Delco parties, the game nights, and for all the ways that we have cared for one another over the years.

Finally, thank you to my mom, grandmother, brother, sister, and extended family for their love and for always believing in me. The biggest thanks to my mother who never doubted that I was capable of anything, worked extremely hard to care for our family, and made sure that I could seize every opportunity.

Kiona R. Elliott

Washington University in St. Louis

August 2023

Dedicated to my mother and grandmother.

ABSTRACT OF THE DISSERTATION

Investigating Cassava Susceptibility to *Xanthomonas* Induced Bacterial Blight Using Disease Phenotyping and Genome Engineering Strategies

by

Kiona R. Elliott

Doctor of Philosophy in Biology and Biomedical Sciences

Plant and Microbial Biosciences

Washington University in St. Louis, 2023

Professor Rebecca Bart, Chair

In nature, plants regularly contend with bacterial, viral, and fungal pathogens, which employ strategies that promote pathogen fitness and subsequently elicit disease in the host. Xanthomonads are a group of bacterial phytopathogens that induce disease in an extensive range of host plants. Mechanisms underlying pathogen disease promotion have been understudied in food staple crops such as cassava. Cassava is a starchy root crop susceptible to cassava bacterial blight (CBB) a disease caused by the pathogen *Xanthomonas phaseoli* pv. *manihotis* (Xpm). Xpm and other bacterial pathogens use effector molecules to manipulate host genes that promote disease susceptibility. Xpm has specialized effector molecules called transcription activator-like (TAL) effectors that enhance pathogen virulence and promote CBB by directly or indirectly inducing the expression of host susceptibility (S) genes. Two TAL effectors, TAL20 and TAL14 respectively target host genes *MeSWEET10a*, a sugar transporter, and two putative cassava pectate lyases

(*MePLLs*). However, the function of these S genes in promoting bacterial blight is not fully characterized.

This thesis describes work done to enhance understanding of the roles that *MeSWEET10*a and the *MePLLs* play in promoting CBB disease susceptibility. First, I developed image analysis-based methods to quantify CBB disease severity in Xpm-infected cassava. Next, I used dual gRNA CRISPR/Cas9-mediated genome editing to develop several *MeSWEEET10a* mutant lines and examined disease susceptibility in mutants. I found that *MeSWEEET10a* mutant lines exhibited reduced disease symptoms and therefore were less susceptible to CBB. Furthermore, editing *MeSWEEET10a* did not appear to detrimentally impact the development or function of cassava flowers where *MeSWEEET10a* is natively expressed. In addition, I investigated the role of putative *MePLLs* in promoting cassava bacterial blight. I compared MePLL protein sequence similarity to validated pectate lyases from the literature and found that both MePLL sequences had the conserved domain pectate lyase C (Pel C). Work identifying the *MePLLs* as TAL14 targeted S genes was previously completed in the cassava cultivar, 60444. I used RT-PCR and bacterial growth assays to confirm if virulence-related phenotypes were observed in a farmer-preferred cassava cultivar, TME419. I generated CRISPR/Cas9 constructs with gRNAs targeting the *MePLLs*, generated several *MePLL* transgenic lines in the TME419 background, and characterized mutant genotypes.

Overall, this thesis work advances the fundamental understanding of host-pathogen interactions, results in *MeSWEEET10a* mutants with reduced CBB susceptibility, and provides

valuable knowledge and resources to further examine the role of *MePLLs* in promoting Xpm virulence.

Chapter 1: Introduction

1.1 Cassava: A Global Food Staple Crop

Cassava (*Manihot esculenta* Crantz, diploid, 2n=36, 700Mb genome) is a perennial shrub from the family *Euphorbiaceae* (Alves, 2002 and Mansfeld et al., 2021). It is a monoecious plant with female flowers that are larger in size and fewer in number compared to male flowers. Additionally, females open 1-2 weeks before males (Alves, 2002). Most cassava cultivars flower 8-16 months after planting, however, it can be longer for some. Cassava can be self or cross-pollinated and natural pollination occurs by bees (Kawano, 1980).

Cassava is mainly grown for its starchy roots which are a significant carbohydrate source in either directly cooked or processed forms (starch, flour, chips, etc.). It is produced in tropical and sub-tropical regions across South America, East Asia, and Sub-Saharan Africa and is considered a food security crop for nearly 800 million people worldwide (Lozano, 1986, Morgan and Choct, 2016, and Strange, 2003). According to the Food and Agricultural Organization (FAO), 34.3, 76.7, and 140.9 million tons of cassava were harvested in Latin America/the Caribbean, Asia, and Sub-Saharan Africa respectively in 2011.

Cassava is a low-input crop often grown without expensive agrochemicals such as fertilizer (El-Sharkawy, 2003). It is especially important for smallholder farmers who grow cassava as a sustenance crop and sell it for revenue when yields allow (Teeken et al., 2018). Farmers cultivate cassava from clonally propagated mature stem cuttings and after about 3 months post-planting,

starch deposition begins in the root (El-Sharkawy 2004). After 6-24 months the storage roots can be harvested or left underground for prolonged periods in case of food shortages. Cassava can be planted from seed but this method is primarily used by breeders who work to enhance cassava nutritional value, increase yields, and improve pest and disease resistance (Ceballos et al., 2004).

1.2 Pathogen Induced Diseases in Cassava

While cassava is tolerant to abiotic stresses such as poor soil quality or drought, there are several biotic stresses that impact production. Cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) result in revenue losses of over one billion dollars per year and both diseases are caused by viral pathogens (Bart and Taylor 2017). The leading bacterial disease impacting cassava production is cassava bacterial blight (CBB) (Howeler et al., 2013). CBB disease symptoms include water-soaked leaf lesions, chlorosis, leaf-wilting, defoliation, and stem browning (**Figure 1.1A-C**). CBB disease severity and crop loss can vary (Wydra and Verdier, 2002 and Fanou et al., 2017). However, up to 100% of the crop can be lost along with the stem cuttings used as planting material (Lozano, 1980). The first reported case of CBB was in 1912 in Brazil, where cassava was originally domesticated (Hillocks et al., 2009). Today, CBB is present in all cassava-growing regions (López and Bernal, 2012).

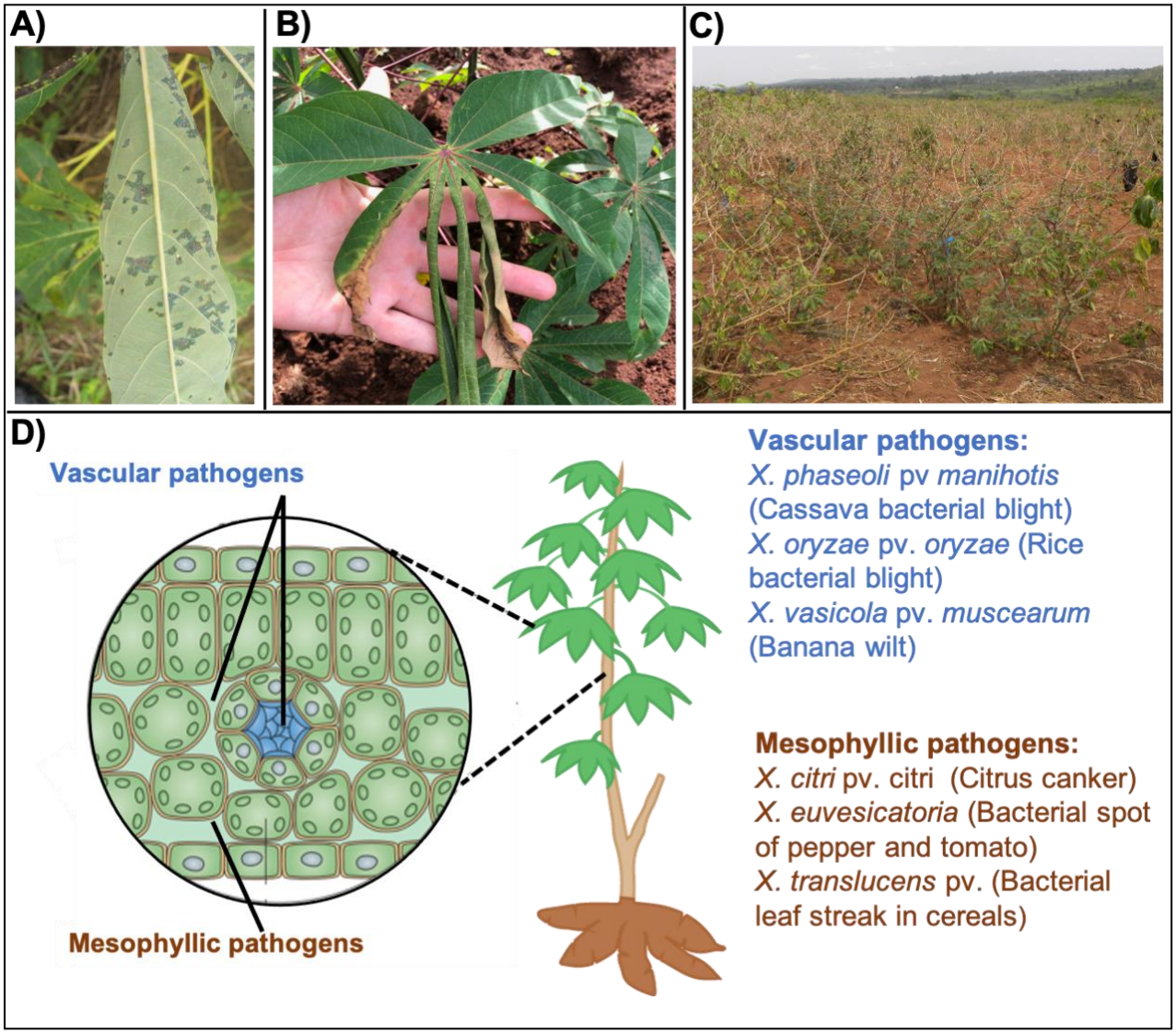


Figure 1.1: Field Plants Exhibit Cassava Bacterial Blight Disease Symptoms
Cassava bacterial blight disease symptoms include **A)** Water-soaked lesions, **B)** leaf-wilting, and **C)** water-soaked lesions (Field images courtesy of Rebecca Bart, Morag Ferguson, and Vincent Kyaligonza). **D)** Xanthomonads are classified as vascular or mesophyllic pathogens. *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) is a vascular pathogen that initially colonizes at the mesophyll cell surfaces and spreads to the vasculature (graphic modified from Ryan et al, 2011).

1.3 The Causal Agent of Cassava Bacterial Blight

CBB is caused by a bacterium in the genus *Xanthomonas*, a group of gram-negative phytopathogens. There are 27 different *Xanthomonas* species that elicit disease in economically important crops including rice, cotton, sorghum, and citrus (Leyns et al.,1984 and Jacques et al., 2016). In total, Xanthomonads are associated with diseases in over 400 plant species (Mhedbi-Hairi et al., 2013). These plant pathogens are obligate aerobes and typically have a single polar flagellum (An et al., 2007). Xanthomonads grow optimally in temperature ranges from 25-30°C. This genus can be further classified by pathovars (pv.) which separate members of the same species into groups based on host specificity and colonization strategies. The Xanthomonad that induces CBB is *X. phaseoli* pv. *manihotis* (Xpm) (Constantin et al., 2015), formerly called *X. axonpodis* pv. *manihotis* (Xam). Throughout this dissertation, I will refer to the CBB pathogen as Xpm, except in chapters from previously published work in which the pathogen was denoted as Xam (**Chapter 2 and Appendix 1**).

Xpm is dispersed from plant to plant through rain, wind, or by propagation of already infected stem cuttings. The Xpm infection cycle consists of epiphytic and endophytic stages. In the epiphytic stage, Xpm persists on the leaf surface and enters the endophytic stage upon invasion of the leaf through open stomata or wounds (Kandel et al., 2017). Xpm is a vascular pathogen that initially infects the cassava leaf and spreads throughout the plant vasculature during the endophytic infection cycle. Bioluminescence data tracking the *in planta* spread of Xpm carrying a luciferase-luciferin plasmid (Xpm:pLUX) validated the Xpm movement in the vasculature (Mutka et al.,

2016). The vasculature is comprised of xylem and phloem vessels vital for plant structural support and transport of water, sugar, and small molecules (Rybel et al., 2016). While some Xanthomonads like Xpm can initially colonize the surface of mesophyll cells and then systemically spread throughout the plant vasculature, others are restricted to colonizing cells at the initial site of infection **(Figure 1.1D**) (Ryan et al., 2011 and An et al., 2019). For example, *Xanthomonas euvesicatoria* (Xe) is a foliar pathogen and the causal agent of bacterial spot disease in pepper and tomato (Jones et al., 2004). Xe is a mesophyllic pathogen that does not enter the vasculature. Presumably, pathogens such as Xe are unable to proliferate throughout the plant vasculature due to a lack of key virulence factors that allow for host manipulation in the vasculature. While the mechanisms that allow some Xanthomonads to enter the vasculature are not fully understood, there has been progress made investigating genetic mechanisms underlying pathogen ability to disperse through the vasculature. For example, it was reported that a hydrolase gene, CbsA, acts as a phenotypic switch for vascular and nonvascular pathogen lifestyles (Gluck-Thaler et al., 2020).

1.4 Plant Response to Pathogen Invasion

Upon pathogen invasion, plants respond using two defense strategies, pattern-triggered immunity (PTI) or effector-triggered immunity (ETI). PTI, the first defense tactic, is triggered when pathogen components such as flagellin or chitin are perceived by pattern recognition receptors located on the plant cell surface. PTI induces a number of molecular, morphological, and physiological changes in the plant for defense against pathogens (Anderson et al., 2010). To overcome PTI, pathogens use effector molecules that prevent plant resistance and manipulate host

responses. Plants use ETI, the second defense tactic, to impede pathogen effector molecules. Effectors are proteins secreted by the pathogen to promote virulence (Toruno et al., 2016). Effectors have a broad range of functions including fostering pathogen entry into the host, evading host defenses, and altering the host environment to preferred pathogen growing conditions. ETI is activated when a plant recognizes a pathogen effector using resistance genes (R genes) (Flor, 1971).

R genes typically encode intracellular proteins that recognize effectors transcribed by pathogen Avr genes either directly or by perceiving effector-induced changes. R gene-mediated perception of effector molecules initiates plant defenses. Currently, no classical R genes that recognize Xpm effectors and protect cassava against CBB have been identified (Bart and Taylor 2017). The relationship between pathogen effectors and R genes is commonly characterized as an evolutionary arms race in which effectors evolve to avoid recognition and R genes evolve to "capture" effectors and prevent disease (Anderson et al., 2010). This phenomenon is also explained by the zig-zag model (Jones and Dangl, 2006). This relationship can be expanded to include susceptibility genes (S genes) which are targeted by effectors to manipulate the host and promote disease (Boch and Bonas, 2010, Hogenhout et al., 2009 and Yang et al., 2006). Plant disease susceptibility caused by effectors is known as effector-triggered susceptibility (ETS) (Van Schie and Takken, 2014 and Eckardt, 2002).

1.5 Transcription Activator-Like Effectors and Host Susceptibility Genes

Pathogen effectors are released into the environment or host cell via bacterial secretion systems. Secreted proteins have numerous functions including promotion of virulence, competition, resource scavenging, and conjugation. Bacterial secretion systems are divided into specialized classes based on the secretion system structure and the types of molecules exuded. A secretion system of relevance to this dissertation is the type three secretion system (T3SS) which is encoded by the hypersensitive response and pathogenicity (*hrp*) gene cluster (Lombardi et al., 2019). The T3SS is present in many bacteria and is a needle-like protrusion that penetrates the host plasma membrane to release effectors into the cytoplasm (**Figure 1.2A**). The T3SS structure is comprised of three main components including a base complex, a needle, and a translocon (Abrusci et al., 2014). There are more than twenty structural proteins that make up the T3SS. However, the exact number of type three effectors (T3Es) secreted by the T3SS is species dependent. Based on Illumina short read sequencing data, Xpm strains were discovered to have nine core T3Es including Hpa2, HpaA, XopAK, XopE1, XopN, HrpF, XopAE, XopV, and XopL (Bart et al., 2012).

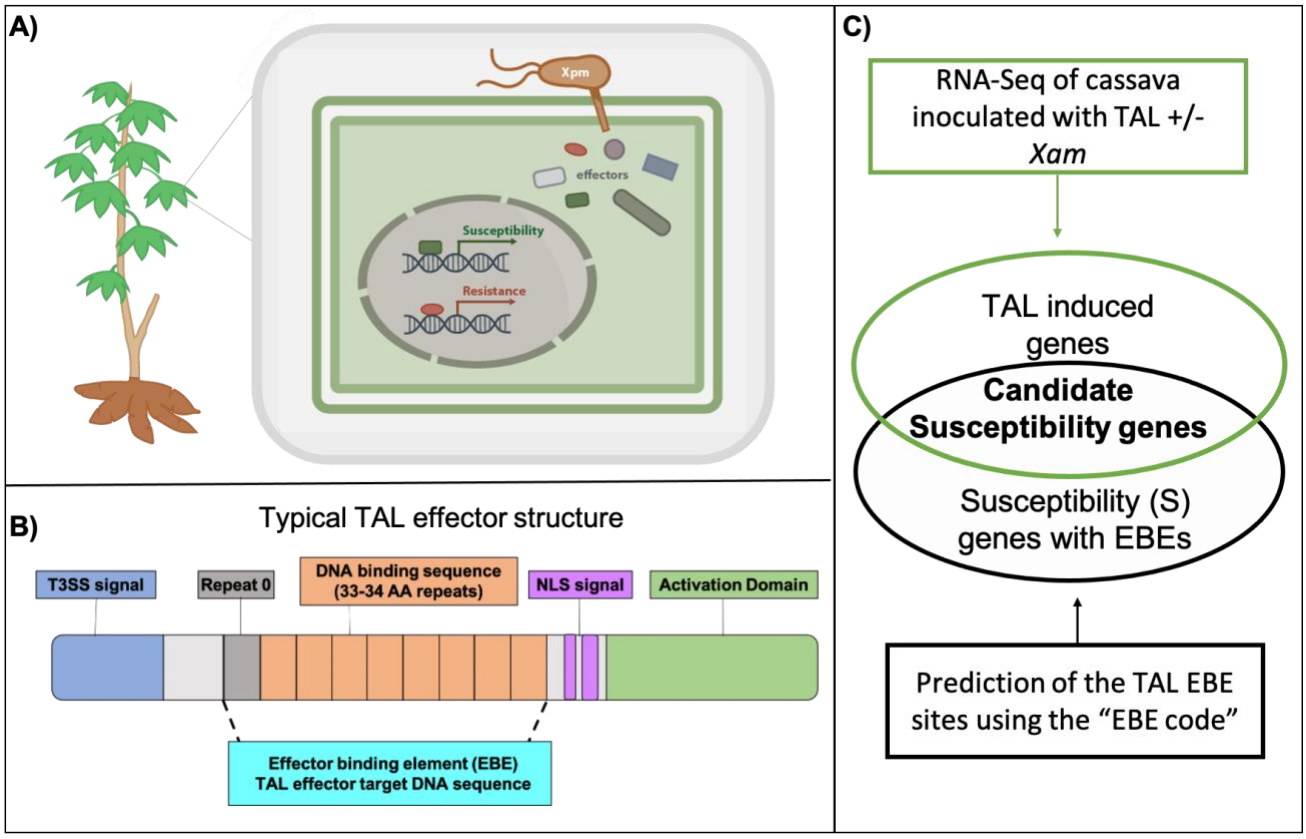


Figure 1.2: *Xanthomonas* Secreted Effectors and Susceptibility Genes
A) Xanthomonads secrete effector proteins into the plant cell via the type three secretion system (T3SS). Effectors can trigger either susceptibility (ETS) or immunity (ETI or resistance) in the plant (graphic modified from Molly Kuhs). **B)** Xanthomonads have a specialized class of effectors called transcription activator-like (TAL) that bind to effector binding element (EBE) sites upstream of susceptibility genes. The typical TAL effector structure includes a T3SS signal, an anchoring repeat 0, a DNA binding sequence with 33-34 amino acid (AA) repeats, nuclear localization signal (NLS), and an activation domain. **C)** Prior work identified candidate S genes by coupling RNA-seq data of Xanthomonads with or without individual TALEs and EBE site prediction software to find overlapping genes

Many bacteria from the *Xanthomonas* and *Ralstonia* genera have specialized T3Es called transcription activator-like (TAL) effectors (Doyle et al., 2013, Liu et al., 2015, and Medina et al., 2017). TAL effector proteins structurally resemble eukaryotic transcription factors and consist of

a transcriptional activation domain, nuclear localization signal, and DNA binding domain that is a series of tandem amino acid repeats (**Figure 1.2B**) (Schornack et al., 2013 and Boher et al., 1995). TAL effector repeat units typically have 33-34 amino acids that are mostly identical except for the 12^{th} and 13^{th} residues called the repeat variable di-residues (RVDs)**.** The 13^{th} residue makes direct contact with DNA at the EBE site while the 12^{th} residue helps with interaction stabilization. Upon secretion into the host cell, TAL effectors localize to the nucleus and bind directly to DNA sites upstream of S genes called effector binding elements (EBEs). Following EBE binding, TAL effectors directly or indirectly induce expression of susceptibility genes to promote disease progression and pathogen virulence. TAL effector binding to EBE sites is sequence-specific and the tandem repeats create a code that can be used to identify candidate binding sites and S genes (Cohn and Bart et al., 2014, Cernadas et al., 2014, Phillips et al., 2017, and Cox et al., 2017). In recent years, the discovery of executor R genes and interfering/truncated TAL (iTAL/truncTAL) effectors have expanded understanding of ETI and ETS. Executor R genes are plant defense genes that recognize pathogen TAL effectors and trigger resistance. iTAL/truncTAL effectors have shortened or altered amino acid sequences producing TAL effectors that serve as decoys to host R genes (Gupta et al., 2021). There are many unanswered questions about the mechanisms of executor R genes and iTAL/truncTAL effectors, but as future work investigates these factors in the plant-pathogen interaction, novel resistance strategies can be developed.

Xpm strains typically carry one to five TAL effectors and the model Xpm strain used in this dissertation, Xpm668, carries five TAL effectors: TAL13, TAL14, TAL15, TAL20, and TAL22 (Cohn and Bart et al., 2014). Prior work found that the loss of effectors TAL20 and TAL14

both cause reduced Xpm virulence *in planta*. Candidate S gene targets were identified by using computational prediction of TAL effector EBE binding sites and transcriptome data of cassava inoculated with either wildtype Xanthomonads or mutant Xanthomonads with single knockouts of individual TAL effectors (**Figure 1.2C**) (Cohn and Bart et al., 2014). However, the role that these candidate S genes play in promoting disease susceptibility and precisely how they contribute to pathogen fitness is yet to be determined. Through my dissertation work, I sought to advance understanding of the molecular mechanism underlying cassava S genes by **Aim I:** Elucidating the role of *MeSWEET10a* and **Aim II**: Elucidating the role of cassava pectate-like lyases (*MePLL*s) in promoting cassava bacterial blight.

Furthermore, I worked to define disease in the context of CBB. Plant disease is characterized in various ways, including the amount of pathogen growth *in planta*, visible plant disease symptoms, reduced crop yield, or plant cellular responses to biotic stress such as the production of reactive oxygen species (Strange 2003 and Sahu et al., 2022). However, as described by Guant et al, a disease model can be "constrained to the specific conditions under which the model was developed" (Guant et al., 1995). In other words, some disease measures may only be robust in certain settings and while one measure is discernable in one condition, another may not be. For example, a recent study reported that hydathode-based entry of bacteria into *Arabidopsis* plant triggered plant immunity and consequently there was low bacterial growth *in planta.* However, bacterial titer was higher when bacteria were syringe inoculated into the plant and hydathode entry was bypassed (Paauw et al., 2023). In this case, the type of inoculation method used to model disease impacted bacterial growth. Therefore, it is important to identify the settings

that increase the robustness of a disease measure. Additionally, different ways to understand or quantify plant disease should be considered.

One marker of CBB disease severity is water-soaked lesions. Prior research demonstrated that loss of the effector TAL20 in Xpm resulted in visibly reduced water-soaked lesions compared to Xpm WT (Cohn and Bart et al., 2014). Image-based methods are valuable tools used to analyze and measure plant health (Gehan et al., 2017, Laflamme et al., 2016, and Lobet et al., 2017). Images captured through different platforms including cell phones, imaging chambers, high-throughput phenotyping facilities, drones, and satellites can be used for analysis (Li et al., 2014 and Zhang and Zhang, 2018). Additionally, open-source image analysis tools such as ImageJ are available (Ferreira and Rasband 2021 [18]). Image-based phenotyping tools have been successfully developed to study a broad range of plant diseases including citrus canker (Bock et al., 2008), grapevine powdery mildew (Bierman et al., 2019), and cereal rust disease (Gallego-Sanchez et al., 2020). The use of machine learning techniques has also been incorporated into image analysis tools for improved trait identification, classification, and faster analysis of plant disease symptoms (Singh et al., 2016 and Tsaftaris et al., 2016). In my dissertation, I developed image analysis tools for efficient segmentation and quantification of water-soaked lesion symptoms as a measure of CBB severity.

1.6 Xpm Susceptibility Gene Targets

The TAL20 effector RVD is predicted to interact with twenty nucleotides in an S gene EBE (**Figure 1.2A**). XpmΔTAL20 (suicide vector knockout) mutants were reported to have lower

bacteria titer of about a one-log difference compared to Xpm WT through midvein-inoculated bacterial growth assays (Cohn and Bart et al., 2014). Additionally, XpmΔTAL20 infection sites had noticeably reduced water-soaked leaf lesions compared to plants infected with Xpm WT (Cohn and Bart et al., 2014). Water-soaked lesions or spots are a common, early symptom of various bacterial diseases (Aung et al., 2018). Water-soaked lesions are the first visible indicators of CBB disease and have been implicated in disease severity (Lozano, 1986). Therefore, I sought to create image analysis tools that allowed for the quantification of water-soaked lesions as a CBB disease severity measure (**Chapter 2**).

Susceptibility genes targeted by TAL20 were discovered using transcriptomic analyses coupled with the EBE site prediction tool, TALE-NT (Cohn and Bart et al., 2014). TAL20 significantly induced expression of the S gene, *MeSWEET10a* (Gene ID: Manes.06G123400). *MeSWEET10a* encodes for a sugar transporter and is a member of the SWEET (Sugar Will Eventually be Exported Transporter) gene family. Ectopic expression of *MeSWEET10a* induced by TAL20 causes the export of sugar out of the plant cell into the apoplast where Xpm proliferates (**Figure 1. 2B**) (Chen et al., 2014, Chen et al., 2010 and Cohn and Bart et al., 2014). Experiments expressing *MeSWEET10a* in HEK293T cells confirmed that it is able to mediate sucrose and glucose transport (Cohn and Bart et al., 2014). We hypothesize that TAL20 induction of *MeSWEET10a* provides sugar as a carbon source or osmolyte for Xpm.

SWEET genes have been identified in plants, animals, and bacteria. A recent review examined SWEETs in over 30 plant species (Gupta et al., 2021). The number of SWEET genes in a plant genome varies across different species. For example, there are 17 SWEETs in *Arabidopsis*

and 108 in hexaploid wheat. The native function of SWEETs in plants also varies. Assorted studies have found SWEET genes with roles in nectar secretion, pollen development, seed filling, and phloem loading in which sucrose is released from the mesophyll cell and then diffuses to the phloem (Feng and Frommer, 2015). In plants, SWEET proteins generally have seven transmembrane domains (TDMs) (Chen et al., 2014). The MeSWEET10a protein sequence was analyzed using the protein domain prediction resource, Interpro, and it was identified as a SWEET sugar transporter with seven transmembrane domains. SWEETs can localize to different cellular compartments such as the plasma membrane or cytoplasmic membrane around vacuoles. Plant SWEET genes are classified into clades I, II, III, and IV based on phylogeny (Gupta et al., 2021). Research in other plant species found SWEET genes in clades I and II are related to the transport of glucose and fructose, clade III SWEETs preferentially transport sucrose but can mediate transport of other sugars, and Clade IV SWEETs are involved in vacuolar transport of fructose (Gupta et al., 2021 and Chen et al., 2012). Phylogenetic analysis of *MeSWEET10a* identified it as a member of clade III (Cohn and Bart et al., 2014). In cassava, there are twenty-three predicted SWEET genes and thirteen are annotated as Clade III SWEETs.

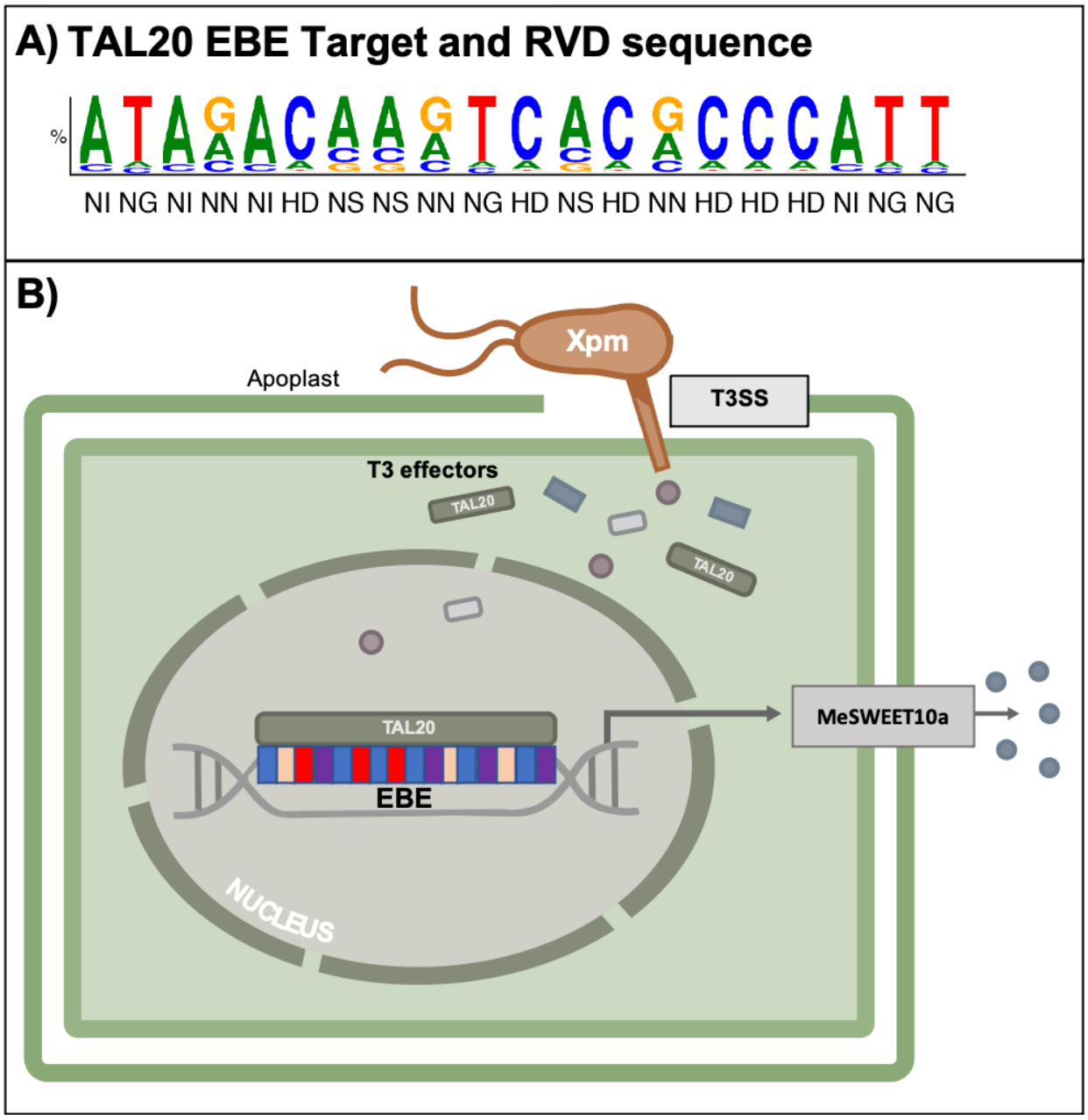


Figure 1.3: TAL20 Binds to the Effector Binding Element (EBE) Site Upstream of *MeSWEET10a*

A) T3SS secreted Xpm TAL20 EBE target sequence and DNA binding domain RVD sequence. EBE sequence displayed with the probability of nucleotide binding at each RVD sequence (adapted from Cohn and Bart et al., 2014) **B)** Graphic depicting *MeSWEET10a* induction by TAL20 binding to the EBE. Following induction, sugars are exported out of the plant cell into the apoplast or extracellular space (graphic modified from Molly Kuhs and Kira Veley).

SWEET genes are established TAL effector susceptibility targets in various plant species including rice, pepper, and cotton (Li et al., 2012, Antony et al., 2010, Cox et al., 2017, and Phillips et al., 2017). SWEET genes have also been implicated in plant disease susceptibility in the absence of TAL effectors. For example, in the plant *Brassica rapa*, expression of SWEET homologs was upregulated in response to infection by the causal agent of clubroot disease, *Plasmodiophora brassicae* (Li et al., 2018). In pathosystems with TAL effector-dependent induction of SWEET genes, numerous studies have demonstrated that preventing TAL effector interaction with the SWEET EBE reduced plant disease susceptibility (Gupta et al., 2021, Veley et al, 2023). For example, research on rice bacterial blight, induced by *X. oryzae* pv. *oryzae* (Xoo), found that CRISPR/Cas9-mediated editing of three rice SWEET genes decreased plant susceptibility to bacterial blight (Olivia et al, 2019). EBE site-specific edits of S genes beyond the SWEET family have also improved plant resistance to disease. For instance, CRISPR/Cas9 generated edits in the EBE of *X. citri* subsp. *citri* (Xcc) S gene target, Cs LATERAL ORGAN BOUNDARIES 1 (LOB1), resulted in canker-resistant sweet orange (Huang et al., 2022). Such studies illustrate that S gene editing strategies provide a useful opportunity to investigate the function of S genes in promoting host susceptibility and pathogen virulence. We hypothesized that editing the *MeSWEET10a* S gene coding sequence and/or EBE site specifically would reduce cassava susceptibility to CBB. Using a dual gRNA CRISPR/Cas9 strategy, I generated several mutant *MeSWEET10a* lines, characterized CBB disease susceptibility, and examined the impact of editing *MeSWEET10a* in cassava flowers where there is native gene expression (**Chapter 3**).

The TAL14 effector RVD is predicted to interact with fourteen nucleotides in an S gene EBE (**Figure 1.4A**). XpmΔTAL14 (suicide vector knockout) is reported to have decreased bacterial growth when inoculated into cassava (Cohn and Bart et al., 2014). TAL14 induces the expression of fifty-two cassava candidate S genes (Cohn and Bart et al., 2014 and Cohn at al., 2016). A table of ten candidate TAL14 S genes targets with potential function in promoting CBB disease susceptibility is provided in **Table 1.1.** Two of the candidate S genes (Gene IDs: Manes.15G048700/Cassava4.1_007516 and Manes.03G152600/Cassava4.1_007568) were identified as putative cassava pectate lyase-like proteins *(MePLLs*). We hypothesize that TAL14-mediated expression of *MePLLs* aids Xpm infection of cassava either by degrading pectate at host cell surfaces or pectinaceous blockages that occur in the vasculature as a plant defense response (**Figure 1.4B)**. In **Chapter 4,** background information on pectate lyases will be presented along with work completed to compare the putative *MePLLs* sequence to pectate lyases validated in the literature, generate dual gRNA CRISPR/Cas9 constructs targeting the *MePLLs*, and recover and genotype *MePLL* transgenic lines**.**

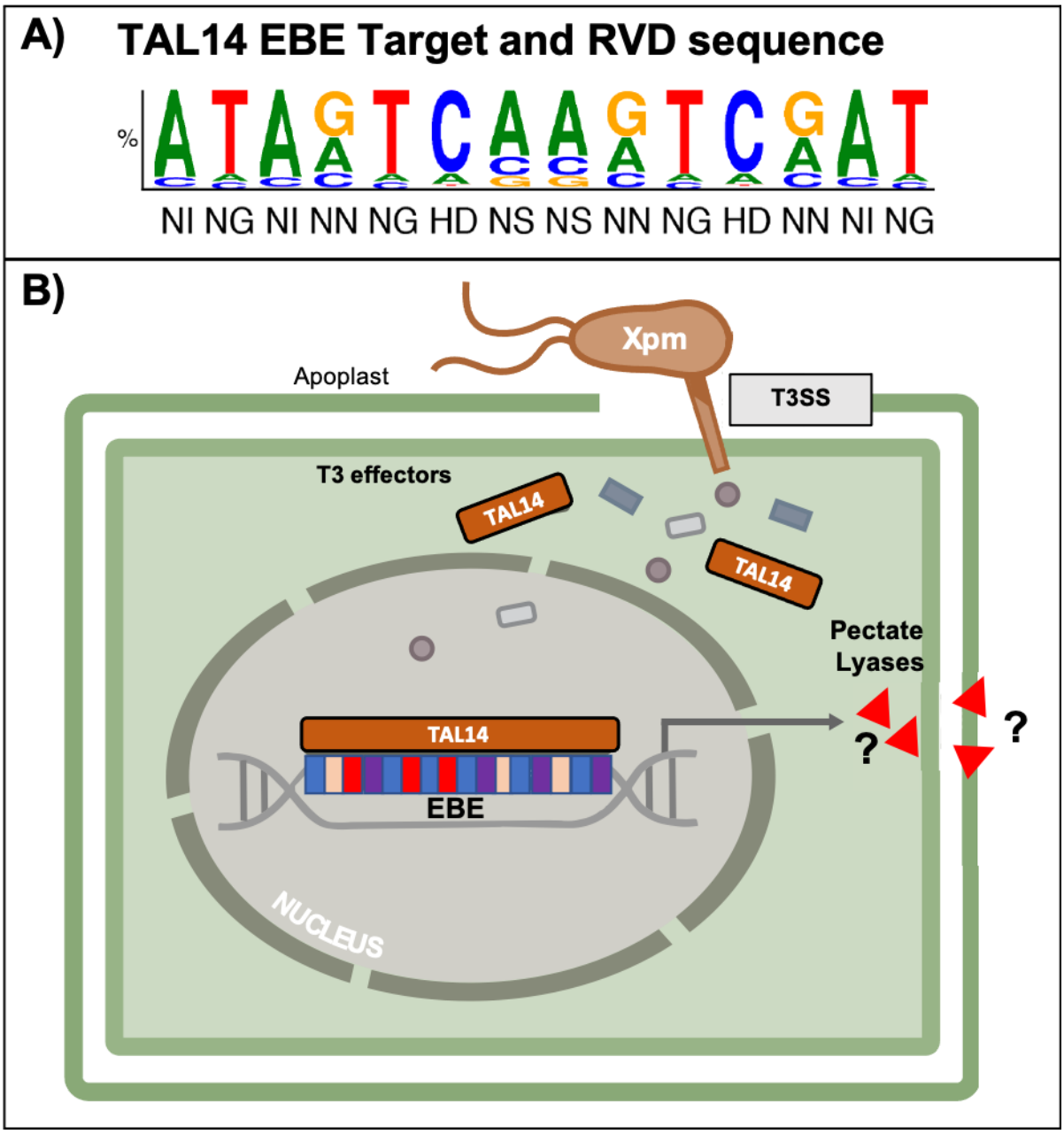


Figure 1.4: TAL14 binds to the Effector Binding Element (EBE) site upstream of *MePLLs*
A) T3SS secreted Xpm TAL14 EBE target sequence and DNA binding domain RVD sequence. EBE sequence displayed with probability of nucleotide binding at each RVD sequence (adapted

from Cohn and Bart et al., 2014) **B)** Graphic depicting MePLLs induction by TAL14 binding to the EBE (graphic modified from Molly Kuhs and Kira Veley).

Gene ID:	Annotated function:
Manes.03G152600	Pectate lyase-like
Manes.15G048700	Pectate lyase-like
Manes.11G151300.1	Serine protease family s10 serine carboxypeptidase
Manes.02G134800.1	Xylose isomerase
Manes.S044400.1	Glycosyltransferase
Manes.04G012000.1	Cyclin-dependent protein kinase inhibitors
Manes.15G026800.1	Membrane associated kinase regulator-1 related
Manes.16G104900.1	Subtilase serine protease
Manes.01G000600.1	EamA-like transporter
Manes.01G268500.2	Vesicle associated protein

Table 1.1: Candidate Susceptibility Genes Targeted by TAL14
Table of ten select candidate TAL14 susceptibility gene targets identified using transcriptomic and EBE site prediction data. Adapted from Cohn and Bart et al., 2014.

1.8 Chapter Summary, Significance, and Scope

According to the Food and Agriculture Organization (FAO), agricultural production will need to increase by sixty percent to meet the food demands of a growing population in 2050 (FAO, Alexandratos and Bruinsma, 2012). Pathogen induced plant diseases are a severe threat to food security. Major crops including wheat, rice, and maize experience global yield losses that range from 21.5, 30.3, and 22.6 percent, respectively (Savary et al., 2019). The burden of plant disease

is expected to magnify as climate change worsens. Higher temperatures and humidity will allow for increased incidences of pathogen invasion on crops and as natural disasters such as hurricanes become more frequent and extreme, pathogens can spread and be introduced to new areas (Ristaino et al., 2021). Smallholder farmers in particular are detrimentally impacted by plant disease because treatment options to combat pathogens can be expensive and in the case of some diseases, no treatment options exist. For example, in Guatemala a severe outbreak of coffee rust caused by a fungal pathogen occurred, the cost of fungicides skyrocketed, and more than 120,000 small-scale farmers lost their livelihoods (Sieff, 2019). Research on fundamental plant-pathogen interactions is vital to enhance our understanding of how pathogens induce plant diseases. Moreover, translating this research into strategies that combat pathogen virulence is necessary to ensure the production of crops needed to sustain future populations.

In my dissertation, I examined cassava susceptibility to *Xanthomonas* induced bacterial blight. I investigated the molecular function of select cassava susceptibility genes targeted by Xpm TAL effectors. To do this, I developed S gene and/or EBE mutant cassava lines using CRISPR/Cas9 technology and applied disease phenotyping strategies to characterize mutant susceptibility to CBB. This research increases basic understanding of how susceptibility gene targets promote CBB and *Xam* virulence and contributes to the development of bacterial blight-resistant cassava. Furthermore, information learned about disease susceptibility in cassava may provide a better understanding of host-pathogen interactions in other Xanthomonad induced plant diseases.

1.9 References

1. Abrusci P, McDowell MA, Lea SM, Johnson S (2014) Building a secreting nanomachine: a structural overview of the T3SS. Current Opinion in Structural Biology 25: 111–117
2. Alves AAC (2002) Cassava botany and physiology. In RJ Hillocks, JM Thresh, eds, Cassava: biology, production and utilization. CABI, Wallingford, pp 67–89
3. An S-Q, Potnis N, Dow M, Vorhölter F-J, He Y-Q, Becker A, Teper D, Li Y, Wang N, Bleris L, et al Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen Xanthomonas. FEMS Microbiol Rev. doi: 10.1093/femsre/fuz024
4. Anderson JP, Gleason CA, Foley RC, Thrall PH, Burdon JB, Singh KB (2010) Plants versus pathogens: an evolutionary arms race. Funct Plant Biol 37: 499–512
5. 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 22: 3864–3876
6. Aung K, Jiang Y, He SY (2018) The role of water in plant–microbe interactions. The Plant Journal 93: 771–780
7. 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. Proc Natl Acad Sci USA 109: E1972-1979
8. Bart RS, Taylor NJ (2017) New opportunities and challenges to engineer disease resistance in cassava, a staple food of African small-holder farmers. PLOS Pathogens 13: e1006287
9. Bierman A, LaPlumm T, Cadle-Davidson L, Gadoury D, Martinez D, Sapkota S, Rea M (2019) A High-Throughput Phenotyping System Using Machine Vision to Quantify Severity of Grapevine Powdery Mildew. Plant Phenomics 2019: 9209727
10. Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48: 419–436
11. Bock CH, Parker PE, Cook AZ, Gottwald TR (2008) Visual Rating and the Use of Image Analysis for Assessing Different Symptoms of Citrus Canker on Grapefruit Leaves. Plant Disease 92: 530–541
12. Boher B, Kpémoua KE, Nicole M, Luisetti J, Geiger J (1995) Ultrastructure of interactions between Cassava and Xanthomonas campestris pv. manihotis : cytochemistry of cellulose and pectin degradation in a susceptible cultivar. doi: 10.1094/phyto-85-777
13. Ceballos H, Iglesias CA, Pérez JC, Dixon AGO (2004) Cassava breeding: opportunities and challenges. Plant Mol Biol 56: 503–516
14. 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. PLOS Pathogens 10: e1003972

15. Chen L-Q (2014) SWEET sugar transporters for phloem transport and pathogen nutrition. New Phytologist 201: 1150–1155
16. 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
17. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T, Staskawicz BJ (2014) Xanthomonas axonopodis virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. Mol Plant Microbe Interact 27: 1186–1198
18. Cohn M, Morbitzer R, Lahaye T, Staskawicz BJ (2016) Comparison of gene activation by two TAL effectors from Xanthomonas axonopodis pv. manihotis reveals candidate host susceptibility genes in cassava. Molecular Plant Pathology 17: 875–889
19. Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B (2016) Genetic characterization of strains named as Xanthomonas axonopodis pv. dieffenbachiae leads to a taxonomic revision of the X. axonopodis species complex. Plant Pathology 65: 792–806
20. Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, Carpenter SCD, Chen L-Q, Zheng H, Gao X, et al (2017) TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. Nature Communications 8: 1–14
21. Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA targeting proteins. Trends Cell Biol 23: 390–398
22. Eckardt NA (2002) Plant Disease Susceptibility Genes? The Plant Cell 14: 1983–1986
23. EL-Sharkawy MA (2003) Cassava biology and physiology. Plant Mol Biol 53: 621–641
24. Fanou AA, Zinsou VA, Wydra K (2017) Cassava Bacterial Blight: A Devastating Disease of Cassava. Cassava. doi: 10.5772/intechopen.71527
25. FAO World agriculture towards 2030/2050: the 2012 revision: ESA Working paper No. 12-03. FAO, Rome, Italy
26. Ferreira T, Rasband W. ImageJ user guide. Madison: University of Wisconsin; 2012.
27. Flor HH (1971) Current Status of the Gene-For-Gene Concept. Annual Review of Phytopathology 9: 275–296
28. Gallego-Sánchez LM, Canales FJ, Montilla-Bascón G, Prats E (2020) RUST: A Robust, User-Friendly Script Tool for Rapid Measurement of Rust Disease on Cereal Leaves. Plants (Basel) 9: 1182
29. Gehan MA, Fahlgren N, Abbasi A, Berry JC, Callen ST, Chavez L, Doust AN, Feldman MJ, Gilbert KB, Hodge JG, et al (2017) PlantCV v2: Image analysis software for high-throughput plant phenotyping. PeerJ 5: e4088
30. Gaunt RE (1995) The Relationship Between Plant Disease Severity and Yield. Annual Review of Phytopathology 33: 119–144
31. Gluck-Thaler E, Cerutti A, Perez-Quintero AL, Butchacas J, Roman-Reyna V, Madhavan VN, Shantharaj D, Merfa MV, Pesce C, Jauneau A, et al (2020) Repeated gain and loss

of a single gene modulates the evolution of vascular plant pathogen lifestyles. Science Advances 6: eabc4516
32. Gupta PK, Balyan HS, Gautam T (2021) SWEET genes and TAL effectors for disease resistance in plants: Present status and future prospects. Mol Plant Pathol 22: 1014–1026
33. Hillocks RJ, Thresh JM, Bellotti A (2002) Cassava: Biology, Production and Utilization. CABI
34. Hogenhout SA, Van der Hoorn RAL, Terauchi R, Kamoun S (2009) Emerging Concepts in Effector Biology of Plant-Associated Organisms. MPMI 22: 115–122
35. Howeler RH, Lutaladio N, Thomas G (2013) Save and grow: cassava: a guide to sustainable production intensification. Food and Agriculture Organization of the United Nations, Rome
36. Huang X, Wang Y, Wang N (2022) Highly Efficient Generation of Canker-Resistant Sweet Orange Enabled by an Improved CRISPR/Cas9 System. Frontiers in Plant Science 12:
37. Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al (2016) Using Ecology, Physiology, and Genomics to Understand Host Specificity in Xanthomonas. Annual Review of Phytopathology 54: 163–187
38. Jones JB, Lacy GH, Bouzar H, Stall RE, Schaad NW (2004) Reclassification of the Xanthomonads Associated with Bacterial Spot Disease of Tomato and Pepper. Systematic and Applied Microbiology 27: 755–762
39. Jones JDG, Dangl JL (2006) The plant immune system. Nature 444: 323–329
40. Kandel SL, Joubert PM, Doty SL (2017) Bacterial Endophyte Colonization and Distribution within Plants. Microorganisms 5: 77
41. Kawano K (1980) Cassava. Hybridization of Crop Plants. John Wiley & Sons, Ltd, pp 225–233
42. Laflamme B, Middleton M, Lo T, Desveaux D, Guttman DS (2016) Image-Based Quantification of Plant Immunity and Disease. MPMI 29: 919–924
43. Leyns F, Cleene MD, Swings J-G, De Ley J (1984) The Host Range of the Genus Xanthomonas. Botanical Review 50: 308–356
44. Li T, Liu B, Spalding M, Weeks D, Yang B (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nature biotechnology 30: 390–2
45. Li X, Si W, Qin Q, Wu H, Jiang H (2018) Deciphering evolutionary dynamics of SWEET genes in diverse plant lineages. Scientific Reports 8: 13440
46. Li L, Zhang Q, Huang D (2014) A Review of Imaging Techniques for Plant Phenotyping. Sensors (Basel) 14: 20078–20111
47. Liu X, Sun Y, Kørner CJ, Du X, Vollmer ME, Pajerowska-Mukhtar KM (2015) Bacterial Leaf Infiltration Assay for Fine Characterization of Plant Defense Responses using the Arabidopsis thaliana-Pseudomonas syringae Pathosystem. J Vis Exp. doi: 10.3791/53364
48. Lobet G (2017) Image Analysis in Plant Sciences: Publish Then Perish. Trends in Plant Science 22: 559–566

49. Lombardi C, Tolchard J, Bouillot S, Signor L, Gebus C, Liebl D, Fenel D, Teulon J-M, Brock J, Habenstein B, et al (2019) Structural and Functional Characterization of the Type Three Secretion System (T3SS) Needle of Pseudomonas aeruginosa. Frontiers in Microbiology 10:
50. López CE, Bernal AJ (2012) Cassava bacterial blight: using genomics for the elucidation and management of an old problem. Tropical Plant Biology 5: 117–126
51. Lozano JC (1986) Cassava Bacterial Blight: A Manageable Disease. Plant Dis 70: 1089
52. Mansfeld BN, Boyher A, Berry JC, Wilson M, Ou S, Polydore S, Michael TP, Fahlgren N, Bart RS (2021) Large structural variations in the haplotype-resolved African cassava genome. The Plant Journal 108: 1830–1848
53. Medina CA, Reyes PA, Trujillo CA, Gonzalez JL, Bejarano DA, Montenegro NA, Jacobs JM, Joe A, Restrepo S, Alfano JR, et al (2017) The role of type III effectors from Xanthomonas axonopodis pv. manihotis in virulence and suppression of plant immunity. Mol Plant Pathol 19: 593–606
54. Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, Saux MF-L, Manceau C, Poussier S, Pruvost O, et al (2013) Evolutionary History of the Plant Pathogenic Bacterium Xanthomonas axonopodis. PLOS ONE 8: e58474
55. Morgan NK, Choct M (2016) Cassava: Nutrient composition and nutritive value in poultry diets. Animal Nutrition 2: 253–261
56. Mutka AM, Fentress SJ, Sher JW, Berry JC, Pretz C, Nusinow DA, Bart R (2016) Quantitative, image-based phenotyping methods provide insight into spatial and temporal dimensions of plant disease. Plant Physiology pp.00984.2016
57. Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom J-S, Li C, Nguyen H, Liu B, et al (2019) Broad-spectrum resistance to bacterial blight in rice using genome editing. Nat Biotechnol 37: 1344–1350
58. Paauw M, van Hulten M, Chatterjee S, Berg JA, Taks NW, Giesbers M, Richard MMS, van den Burg HA (2023) Hydathode immunity protects the Arabidopsis leaf vasculature against colonization by bacterial pathogens. Current Biology 33: 697-710.e6
59. Phillips AZ, Berry JC, Wilson MC, Vijayaraghavan A, Burke J, Bunn JI, Allen TW, Wheeler T, Bart RS (2017) Genomics-enabled analysis of the emergent disease cotton bacterial blight. PLOS Genetics 13: e1007003
60. Ristaino JB, Anderson PK, Bebber DP, Brauman KA, Cunniffe NJ, Fedoroff NV, Finegold C, Garrett KA, Gilligan CA, Jones CM, et al (2021) The persistent threat of emerging plant disease pandemics to global food security. Proceedings of the National Academy of Sciences 118: e2022239118
61. 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
62. Rybel BD, Mähönen AP, Helariutta Y, Weijers D (2016) Plant vascular development: from early specification to differentiation. Nat Rev Mol Cell Biol 17: 30–40

63. Sahu PK, Jayalakshmi K, Tilgam J, Gupta A, Nagaraju Y, Kumar A, Hamid S, Singh HV, Minkina T, Rajput VD, et al (2022) ROS generated from biotic stress: Effects on plants and alleviation by endophytic microbes. Frontiers in Plant Science 13:
64. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A (2019) The global burden of pathogens and pests on major food crops. Nat Ecol Evol 3: 430–439
65. van Schie CCN, Takken FLW (2014) Susceptibility Genes 101: How to Be a Good Host. Annual Review of Phytopathology 52: 551–581
66. Schornack S, Moscou MJ, Ward ER, Horvath DM (2013) Engineering Plant Disease Resistance Based on TAL Effectors. Annual Review of Phytopathology 51: 383–406
67. Sieff K Falling coffee prices drive Guatemalan migration to the United States. Washington Post, https://www.washingtonpost.com/world/2019/06/11/falling-coffee-prices-drive-guatemalan-migration-united-states/
68. Singh A, Ganapathysubramanian B, Singh AK, Sarkar S (2016) Machine Learning for High-Throughput Stress Phenotyping in Plants. Trends in Plant Science 21: 110–124
69. Strange RN (2003) Introduction to plant pathology. Wiley, New York, NY, USA
70. Teeken B, Olaosebikan O, Haleegoah J, Oladejo E, Madu T, Bello A, Parkes E, Egesi C, Kulakow P, Kirscht H, et al (2018) Cassava Trait Preferences of Men and Women Farmers in Nigeria: Implications for Breeding. Econ Bot 72: 263–277
71. Toruño TY, Stergiopoulos I, Coaker G (2016) Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. Annu Rev Phytopathol 54: 419–441
72. Tsaftaris SA, Minervini M, Scharr H (2016) Machine Learning for Plant Phenotyping Needs Image Processing. Trends in Plant Science 21: 989–991
73. Veley KM, Elliott K, Jensen G, Zhong Z, Feng S, Yoder M, Gilbert KB, Berry JC, Lin Z-JD, Ghoshal B, et al (2023) Improving cassava bacterial blight resistance by editing the epigenome. Nat Commun 14: 85
74. Wydra K, Verdier V (2002) Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. Agriculture, Ecosystems & Environment 93: 211–226
75. Yang B, Sugio A, White FF (2006) Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. PNAS 103: 10503–10508
76. Zhang Y, Zhang N (2018) Imaging technologies for plant high-throughput phenotyping: a review. Front Agr Sci Eng 5: 406–419

Chapter 2: Image Analysis Methods to Measure Cassava Bacterial Blight Severity

This chapter was previously published in the journal BMC Plant Methods as:

Elliott, K., Berry, J.C., Kim, H., Bart, R. (2022) A comparison of ImageJ and machine learning based image analysis methods to measure cassava bacterial blight disease severity. *Plant Methods* **18**, 86 https://doi.org/10.1186/s13007-022-00906-x

2.1 Personal Contributions

This manuscript was developed as a collaborative effort of scientists within the Bart Lab. I served as the first author and led project design and progression with advising from Dr. Rebecca Bart. I completed all bacterial inoculations, collected images at each timepoint, and trained/supervised high school summer intern, Hobin Kim to complete ImageJ analysis and interpret results. I tested the performance and accuracy of a PhenotyperCV machine learning tool developed by Data Scientist Jeffrey Berry who also set up the Raspberry Pi camera system. I processed all images analyzed through the machine learning tool and interpreted results. Jeffrey Berry developed the gray correction and image effects color correction methods for image analysis and provided expertise on statistical analyses and developed initial R scripts used to run statistical tests. I wrote the manuscript, completed statistical analysis, and generated all figures.

2.2 Abstract

Methods to accurately quantify disease severity are fundamental to plant pathogen interaction studies. Commonly used methods include visual scoring of disease symptoms, tracking

pathogen growth *in planta* over time, and various assays that detect plant defense responses. Several image-based methods for phenotyping of plant disease symptoms have also been developed. Each of these methods has different advantages and limitations which should be carefully considered when choosing an approach and interpreting the results.

In this chapter, we developed two image analysis methods and tested their ability to quantify different aspects of disease lesions in the cassava-*Xanthomonas* pathosystem. The first method uses ImageJ, an open-source platform widely used in the biological sciences. The second method is a few-shot support vector machine learning tool that uses a classifier file trained with five representative infected leaf images for lesion recognition. Cassava leaves were syringe infiltrated with wildtype *Xanthomonas*, a *Xanthomonas* mutant with decreased virulence, and mock treatments. Digital images of infected leaves were captured overtime using a Raspberry Pi camera. The image analysis methods were analyzed and compared for the ability to segment the lesion from the background and accurately capture and measure differences between the treatment types.

Both image analysis methods presented in this paper allow for accurate segmentation of disease lesions from the non-infected plant. Specifically, at 4-, 6-, and 9- days post inoculation (DPI), both methods provided quantitative differences in disease symptoms between different treatment types. Thus, either method could be applied to extract information about disease severity. Strengths and weaknesses of each approach are discussed.

2.3 Introduction

Annually 20-40% of crops are lost due to plant pests and disease (FAO 2021 [1]). Causal agents of plant disease such as bacteria, viruses, oomycetes, and fungi employ various strategies to promote pathogenesis and elicit disease susceptibility in host plants. Disease susceptibility is commonly measured by the amount of *in planta* pathogen growth, reduction in crop yield/biomass, or by scaled scoring systems that use visible disease symptoms to measure severity (Strange 2003[2] , Liu 2015 [3], Guant 1995 [4], Moore 1943 [5]). Each of these methods have advantages and limitations and no single method can capture the full complexity of plant disease. For instance, it is common to introduce a small number of bacteria into a plant leaf and then quantify pathogen growth overtime (Agrios 5th edition 2004 [6]). This method highly quantitative and can reveal subtle differences in virulence between related pathogen strains or mutants (Bart 2012 [7], Cohn & Bart 2014 [8], Diaz, 2018 [9]). However, this assay probes only one part of the disease cycle and provides limited insight into pathogen spread, plant symptoms or defense responses. Another common method is to visually score disease symptoms on a numerical scale (Jorge & Verdier 2002 [10]). This method can be used in lab to field level experiments, is cost effective, and does not require special techniques or tools. However, accurate identification of pathogen incited symptoms can be difficult, especially in the case of multiple biotic and/or abiotic stresses. Further, disease scores may vary among different scorers and often are not sensitive enough to capture subtle changes in disease severity (Poland and Nelson 2011 [11], Strange and Scott 2005 [12]).

In recent years, there has been an increase in the use of image-based methods to analyze and measure plant health (Gehan 2017 [13], Laflamme 2016 [14], Lobet 2017 [15]). Images can be captured through many different platforms including cell phones, imaging chambers, high-throughput phenotyping facilities, drones, and satellites (Li 2014 [16], Zhang and Zhang 2018 [17]) and many analysis platforms have also been developed, for example, ImageJ (Ferreira and Rasband 2021 [18]). Image-based phenotyping tools have been successfully developed to study a broad range of plant diseases including citrus canker (Bock 2008 [19]), grapevine powdery mildew (Bierman 2019 [20]), and cereal rust disease (Gallego-Sanchez 2020 [21]). At least in some cases, image-based phenotyping can overcome some of the limitations associated with the more traditional methods described above (Mutka and Bart 2015 [22]). For example, a study investigating *Zymoseptoria tritici* infected wheat leaves found that an ImageJ analysis method provided more reliable and reproducible measures of wheat blotch disease compared to a traditional visual scoring system (Stewart 2014 [23], Stewart 2016 [24]). However, manual image analysis based on user selection of disease lesions can also be time-consuming. Some image analysis methods have incorporated machine learning techniques for improved trait identification, classification, and faster analysis of plant disease symptoms (Singh 2016 [25], Tsaftaris 2016 [26]). While machine learning has enhanced the ability to process imaging data, accurate trait classification or quantification often relies on large datasets that can be expensive to acquire. Therefore, more cost-effective, few-shot image analysis tools that allow for efficient segmentation and quantification of disease symptoms are needed.

In this study, we apply image-based phenotyping to cassava (*Manihot esculenta* Crantz), a starchy storage root crop (Morgan 2016 [27]). Cassava is a hardy crop predominantly grown by smallholder farmers in South America, East Asia, and Sub-Saharan Africa (Bart and Taylor 2017 [28], Hillock 2002 [29], El-Sharkawy 2003 [30]). Cassava production is threatened by the disease cassava bacterial blight (CBB). CBB can result in complete crop loss and is present in all cassava growing regions (Howler 2013 [31], Fanuo 2017 [32], Zárate-Chaves 2021 [33]). The causal agent of CBB is *Xanthomonas axonopodis* pv. *manihotis* also referred to as *Xanthomonas phaseoli* pv. *manihoti*s (*Xam* or *Xpm)* (Constantin 2016 [34]). *Xam* infects cassava by entering through open stomata or wounds in the leaf, colonizes the surface of mesophyll cells, and spreads systemically in the plant. The first visible indicators of CBB disease are dark "water-soaked" lesions that appear on the leaf. Water-soaked lesions or spots are a common, early disease symptom of various bacterial diseases. (Aung 2018 [35]). Other CBB disease symptoms include leaf wilt, defoliation, stem browning, and eventual plant death. Like other plant pathogens, *Xam* has a repertoire of effectors that can alter the structure or function of a host cell, create a more ideal environment for pathogen colonization, and overcome plant defense mechanisms (Boch 2010 [36], Hogenhout 2009 [37]). In the *Xanthomonas* and *Ralstonia* bacterial genera, this repertoire includes specialized transcription activator-like (TAL) effectors (Bodnar 2013 [38], Van Schie and Takken 2014 [39], Koseoglou 2021 [40]). TAL effectors are secreted into the plant cell and induce expression of plant susceptibility (S) genes that enhance disease. In many pathosystems, TAL effectors target *SWEET* (*Sugars Will Eventually be Exported Transporters*) genes and preventing this interaction reduces disease symptoms (Li 2012 [41], Phillips 2017 [42], Cox 2017 [43]).The *Xam* strain used in this

study, *Xam668,* carries the effector, TAL20, which induces ectopic expression of *MeSWEET10a* (Cohn and Bart et al., 2014 [8]). *Xam668* mutants with loss of TAL20 (*Xam668ΔTAL20*) exhibit visibly reduced water-soaked lesions compared to wild-type *Xam*. Here, we develop and compare ImageJ and machine learning based image analysis tools that allow for segmentation and quantification of CBB induced water-soaked lesions.

2.4 Results

2.4.1 Xam/Xpm Induction of Water-Soaked Lesions in Cassava

In cassava, water-soaked lesions appear as dark angular spots at the site of infection and spread as the bacteria proliferate **(Figure 2.1A).** To capture the progression of water-soaking in cassava, leaves were syringe-infiltrated with *Xam668*, *Xam668ΔTAL20,* or mock treatments. At 0-, 4-, 6-, and 9-days post inoculation (DPI) infected leaves were detached from the plant and imaged. Images were taken with a Raspberry Pi camera in an enclosed box to increase uniformity of imaging. An X-Rite ColorChecker Passport was included in every image for post-acquisition gray balance color correction (Berry 2018 [44]). At 4DPI, water-soaked spots began to appear in both *Xam668* (Xam WT) and *Xam668ΔTAL20* (XamΔ*TAL20*) infiltration sites **(Figure 2.1B).** Water-soaked lesions spread and increased in visibility at 6 and 9 DPI. However, as previously reported [33] water-soaking appeared reduced in *Xam668ΔTAL20* infection sites as compared to wildtype *Xam668* sites. Additionally, *Xam668ΔTAL20* infection sites appeared lighter in color compared to the darker lesions that develop at wildtype *Xam668* sites. Water-soaked lesions were not observed for any time point in mock infiltrated spots.

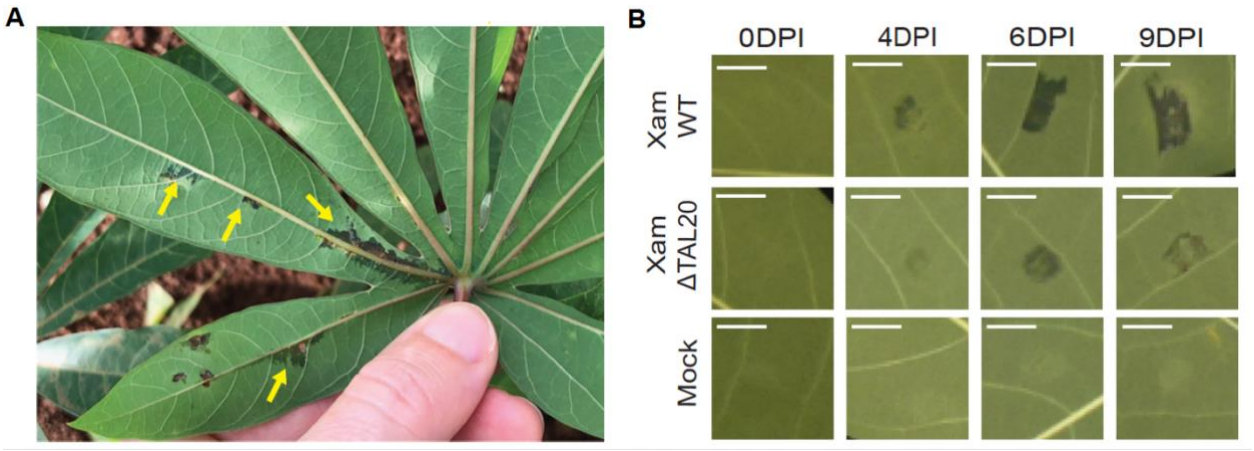


Figure 2.1 *Xanthomonas* Causes Complex Water-soaking Symptoms in Cassava

A) Image of cassava leaf in the field exhibiting water-soaking symptoms characteristic of cassava bacterial blight. Yellow arrows indicate different water-soaked lesions. **B)** Water-soaked symptoms of cassava infiltrated with *Xam668* (Xam WT*)* and a *Xam668* deletion mutant lacking the TAL20 effector (XamΔTAL20) at 0, 4, 6, and 9DPI. Mock inoculations of 10mM $MgCl_2$ at each timepoint were included as controls. Scale bar = 0.5cm.

2.4.2 ImageJ Based Quantification of Water-soaked Symptoms

ImageJ is regularly used for image analysis in biological studies (Ferreira and Rasband 2021 [18]). Here, we applied ImageJ based analysis to extract, quantify, and examine water-soaked lesion traits. Water-soaked lesions induced by *Xam668* and *Xam668ΔTAL20* were segmented using a manual overlay segmentation strategy **(Figure 2.2A).** For segmentation, color corrected images were uploaded and duplicated in ImageJ and the *Xam668* and *Xam668ΔTAL20* lesions were outlined using the pencil tool. Outlined images were converted from RGB to the LAB color space and the "A Channel" was obtained for better separation of the outlined lesions from the leaf background. The A channel images were thresholded and converted to a binary mask. The binary

masks and analyze particle tool in ImageJ were used to define the *Xam668* and *Xam668ΔTAL20* infected sites and an overlay was created for each image. The overlays were applied to the RGB image and measurements for 27 traits were calculated. Mock sites were measured using the rectangle selection tool in the RGB image to capture information about "non-water-soaked" leaf background. ImageJ processing took approximately 6 minutes and 30 seconds per image. A movie example of the image J based analysis method was generated as a tutorial **(Supplemental File 2.1).**

Ten traits were selected and further analyzed using an ANOVA analysis to determine the variance explained (VE) by three terms of interest: (1) inoculation type, (2) DPI and (3) the interaction between inoculation type and DPI **(Figure 2.2B).** Inoculation type and DPI were selected as defining factors because we expected that water-soaking severity is dependent on these terms. Area had the highest amount of VE, with over 60% VE. We selected gray-scale mean as another trait of interest because of the color difference we observed between *Xam668* and *Xam668ΔTAL20* water-soaked lesions. Gray-scale mean, accounted for over 50% VE. Water-soaked area **(Figure 2.2C)** and gray-scale mean **(Figure 2.2D)** were further analyzed as measures of CBB disease severity. The *Xam668* sites had significantly more water-soaked area compared to *Xam668ΔTAL20* at each timepoint. We found there was noise in the gray-scale mean data due to lack of standardization across individual images despite gray balance color correction. To account for this, a linear model was applied to determine the grand mean of all gray values in each image and the *Xam668* and *Xam668ΔTAL20* gray values were centered to mock. In each timepoint,

Xam668 treatment resulted in lesions that had a significantly larger gray-scale mean compared to *Xam668ΔTAL20* treatment. A greater difference in gray-scale mean was observed between *Xam668* and mock treated spots compared to *Xam668ΔTAL20* and mock spots. These results indicate that ImageJ based segmentation allowed for separation of treatment types and for the quantitative analysis of water-soaked lesions over time.

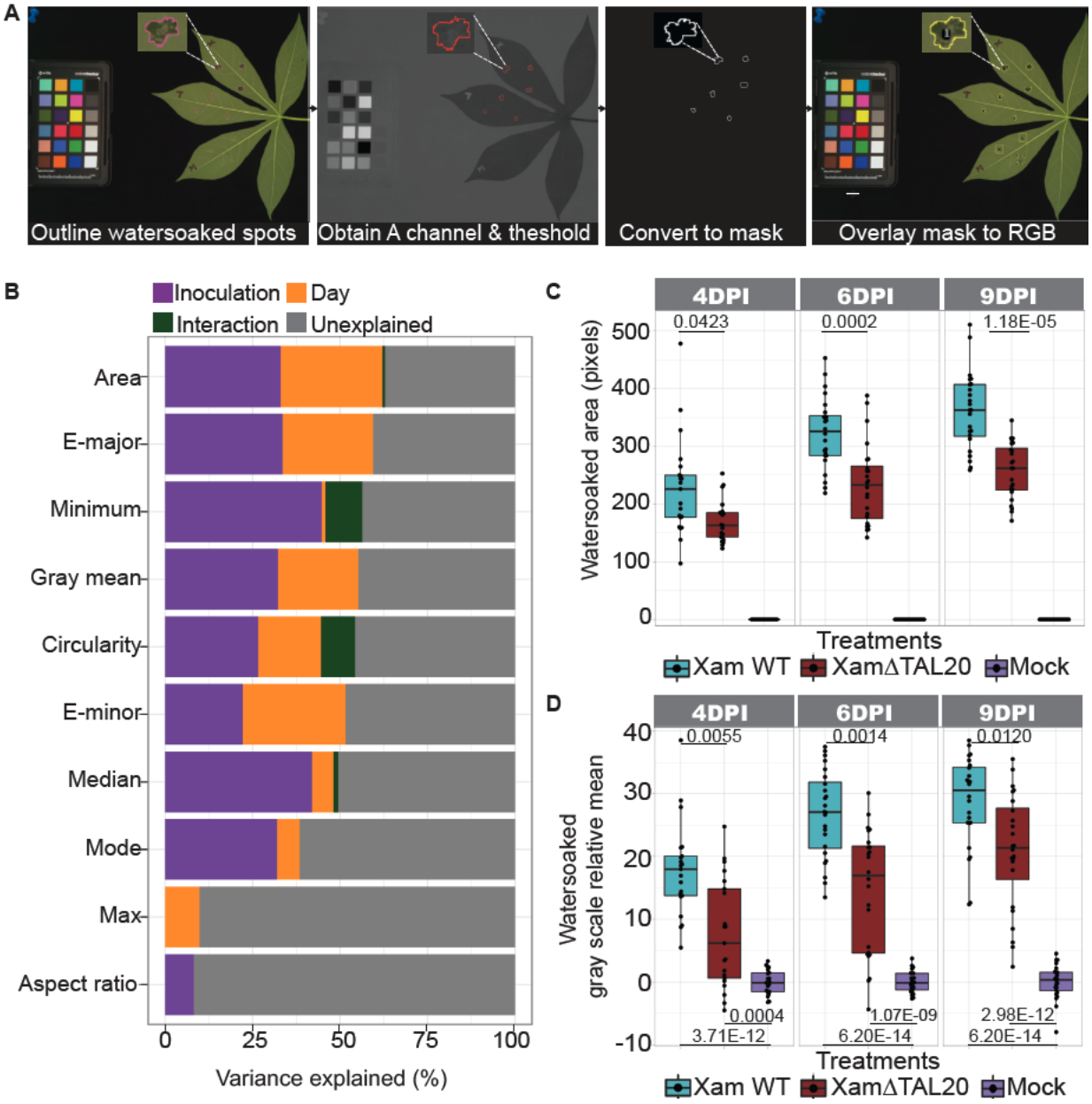


Figure 2.2: Manual ImageJ Analysis of CBB Water-soaking Symptoms

A) Images of cassava leaves infiltrated with Xam WT, XamΔTAL20, and mock treatments were segmented and analyzed using an ImageJ overlay segmentation method. Overlay segmentation analysis depicted by step using a CBB infected cassava leaf image. Images were taken at 0, 4, 6 and 9 DPI. Leaf lobes were labeled by treatment type: X=Xam WT, T= XamΔTAL20, and M=Mock. White lines point to selected regions of a representative water-soaked lesion at each step

of the ImageJ overlay segmentation process. **B)** The variance explained by inoculation type (Xam WT or XamΔTAL20), DPI (4-, 6- and 9-), or the interaction between inoculation type and DPI for ten ImageJ generated measurements. Variances were determined by ANOVA analysis. **C)** Total water-soaked area (pixels, y-axis) for sites infiltrated with each treatment (x-axis). Calculated *p*-values (Kolmogorov-Smirnov test) shown above the line in each plot. **D)** Negative gray-scale mean (y-axis) of water-soaked lesions for Xam WT and XamΔTAL20 relative to mock inoculated spots (x-axis) within the same leaf. Calculated *p*-values (Kolmogorov-Smirnov test) shown above the line in each plot. In ImageJ, the gray-scale mean was measured by averaging the mean of each gray-scale value in the RGB channels.

2.4.3 Machine Learning Based Quantification of Water-soaked Symptoms

While ImageJ provided sufficient segmentation of water-soaked lesions, developing an overlay mask for every individual image is time intensive. Therefore, we sought to develop a machine learning tool that would provide faster segmentation and quantification of diseased leaves. A custom workflow for machine learning disease lesion analysis was developed using the source file from PhenotyperCV, a C++11 library designed for image-based phenotyping (Berry 2018 [44]). The machine learning workflow was run using the Mac terminal. Command syntax specific for each step of the machine learning tool was developed **(Supplemental Table 2.1)**. Five representative images of CBB infected leaves from different DPI were selected and combined into one graphic as a training image for the machine learning tool **(Figure 2.3A).** A binary mask was generated from the combined leaf graphic using ImageJ. The mask was used to generate a support vector machine (SVM) learning classifier (YAML) file. The classifier file was used to process the images and eliminated the need to manually outline each lesion or make individual masks **(Figure 2.3B).** During processing, images were color corrected and manually thresholded using a scale bar built into the program to reduce background noise and enhance segmentation of lesion pixels.

Next, infiltrated spots were manually labelled and color-coded by treatment type. Output images were generated and included color corrected, pseudo-color map, and feature prediction images for every image analyzed **(Figure 2.3C).** Machine learning processing took approximately 2 minutes and 30 seconds per image. Processing speed increased when all images were analyzed using an iteration (for loop) command in terminal allowing the machine learning tool to be executed on several images in succession. A movie example of the machine learning based analysis method was generated as a tutorial **(Supplemental File 2.2).** Additionally, two space separated text (TXT) files were produced with shape and color related measurements of each lesion. A list of the reported measurements is included **(Supplemental Table 2.2)**. Shape data generated by the machine learning tool includes area, hull area, height, width, etc. The color data generated by machine learning is a lightness histogram of 0-255 for each lesion which was used to calculate lesion gray-scale mean.

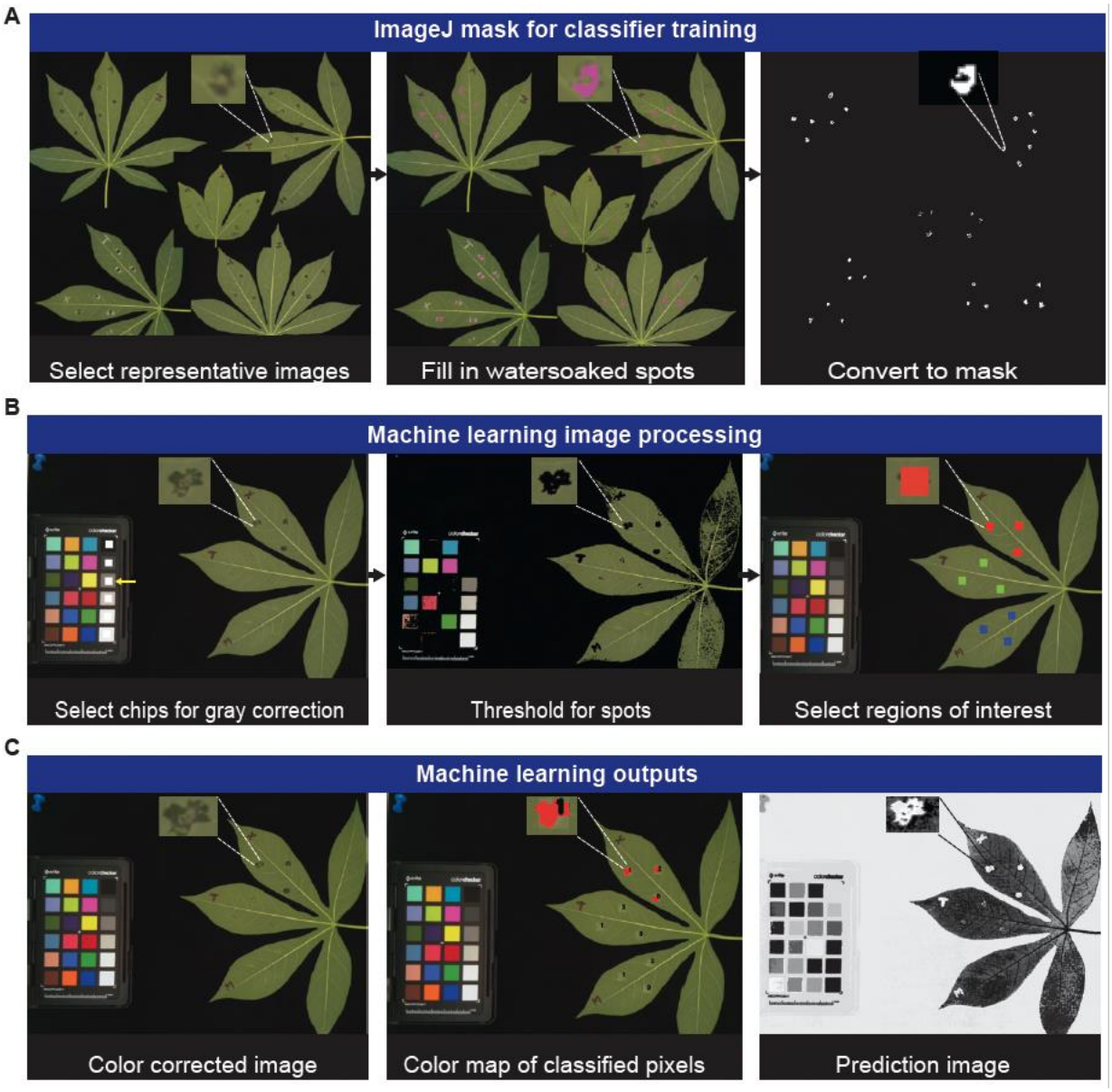


Figure 2.3: Overview of the Support Vector Machine Learning Segmentation and Analysis Method

A) Images of cassava leaves infiltrated with Xam WT, Xam XamΔTAL20, and mock treatments were segmented and analyzed using a support vector machine learning tool. Images depict steps used to generate a classifier training mask for the machine learning tool. A mask was made by combining representative CBB infected images into one graphic and generating a binary mask in ImageJ. White lines showcase a representative water-soaked lesion within the combined leaf graphic and indicate changes at each step. The mask was used to generate a classifier (YAML) file

with PhenotyperCV. **B)** Images depict steps of machine learning processing using a CBB infected cassava leaf image. Images were uploaded into the machine learning tool and processed by gray balance color correction, thresholding, and the inoculated regions of interest were selected and labeled using a color code: Red=Xam WT, Green= XamΔTAL20, and Blue= Mock. White lines showcase a representative water-soaked lesion within the image and indicate changes at each step. **C)** Images exhibit outputs from the machine learning image processing and include the color corrected image (left), a pseudo-colored map of the pixels classified as water-soaked (middle), and a feature prediction image (right). White lines showcase a representative water-soaked lesion within the image and indicate differences in each output image. Text separated files with shapes and color data for each inoculation spot were also generated.

Twelve machine learning derived traits were selected and the ANOVA analysis was used to measure VE by each trait **(Figure 2.4A).** Area measured by the machine learning tool had over 75% VE by the defining factors. As was determined during ImageJ analysis, area also accounts for the highest amount of VE in the machine learning analysis. The gray-scale mean had over 60% VE by the defining factors. Consistent with the ImageJ analysis, the machine learning approach revealed that *Xam668* caused a larger water-soaked area **(Figure 2.4B)** and relative gray-scale mean **(Figure 2.4C)** compared to *Xam668ΔTAL20* infiltrated spots. These data suggest that the machine learning tool adequately distinguished between treatment types and provided quantitative measures of water-soaked lesions using the classifier file created from one training mask.

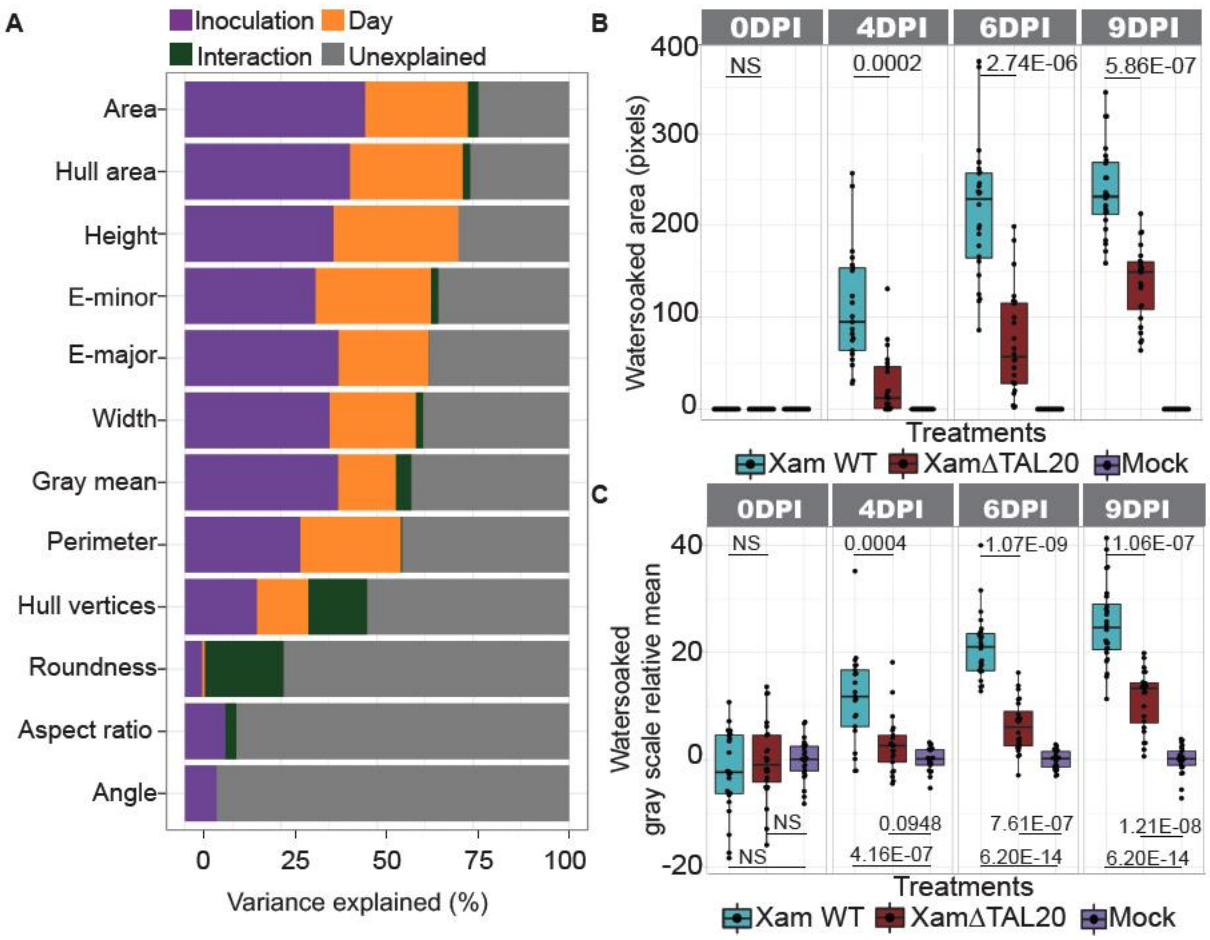


Figure 2.4: Support Vector Machine Learning Analysis of CBB Water-soaked Symptoms
A) The variance explained by inoculation type (Xam WT or XamΔTAL20), DPI (4-, 6- and 9-), or the interaction between inoculation type and DPI for twelve machine learning generated measurements. Variances were determined by an ANOVA. **B)** Total water-soaked area (pixels, y-axis) for sites infiltrated with each treatment (x-axis). Calculated *p*-values (Kolmogorov-Smirnov test) shown above the line in each plot. **C)** Negative gray-scale mean (y-axis) of water-soaked lesions for Xam WT and XamΔTAL20 relative to mock inoculated spots (x-axis) within the same leaf. Calculated *p*-values (Kolmogorov-Smirnov test) shown above the line in each plot. In the machine learning analysis, the gray-scale mean was generated using the average mean of the "L" channel from the LAB color space.

2.4.4 Comparison of the ImageJ and Machine Learning Based Lesion Analysis Methods

The ImageJ and machine learning based methods both successfully distinguished *Xam668* and *Xam668ΔTAL20* and yet the results were not equivalent. To further compare and contrast these methods, representative *Xam668* and *Xam668ΔTAL20* lesions from 4- ,6-, and 9- DPI were selected and visually inspected **(Figure 2.5A).** We observed that machine learning was able to distinguish between water-soaked and "non-water-soaked" pixels within the lesion spot whereas in ImageJ, a boundary was put around the whole spot and could include a mix of both pixel types. This suggests that the machine learning tool is more selective in classification of water-soaked versus non-water-soaked pixels and would explain the trend of overall smaller area measurements generated by machine learning compared to ImageJ. In ImageJ, the lesion boundary is user-selected. However, to completely separate water-soaked from non-water-soaked pixels in lesions where there is a mix, smaller independent boundaries would be required. Having multiple boundaries for one lesion is not ideal as it would impact measures such as gray-scale mean and increase image processing time. The two image analysis methods were statistically compared by pairing the mock, *Xam668* and *Xam668ΔTAL20* area data and performing F-statistic variance tests on each respective treatment type **(Figure 2.5B**). At each timepoint, there was no significant difference in the variance observed between ImageJ and machine learning data suggesting the two methods have equal variation within each treatment type.

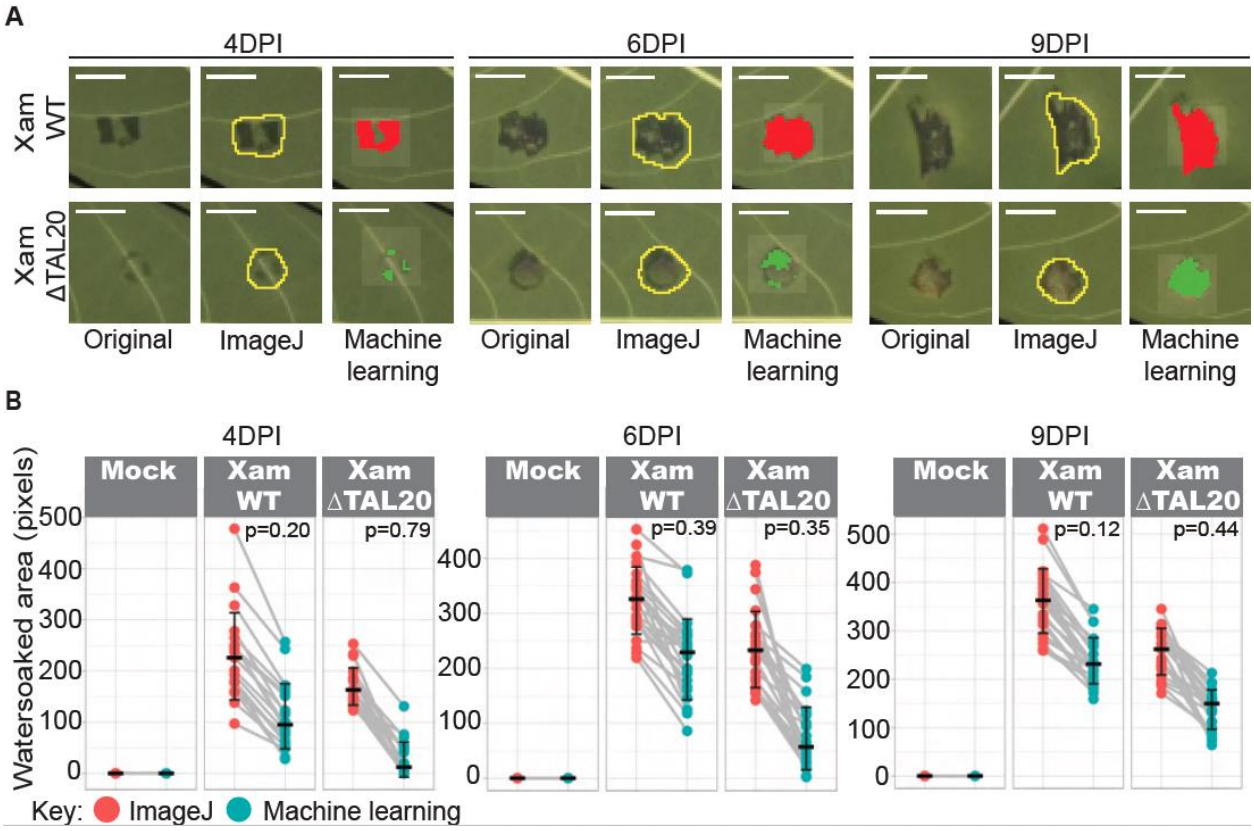


Figure 2.5: Comparison of the ImageJ and Machine Learning Analyses of CBB Infected Leaves

A) Representative images from each timepoint (4-, 6-, and 9- DPI) of a Xam WT (top row) and XamΔTAL20 (bottom row) water-soaked spots were selected, visually inspected, and compared. The original images show the water-soaked spots from the color corrected images without segmentation from the background. The "ImageJ" images show water-soaked spots manually segmented from background and overlaid onto the RGB image. The machine learning images shows water-soaked spots segmented from background and pseudo-colored. Scale bar = 0.5cm **B)** Water-soaked area data generated by ImageJ or machine learning were paired by inoculation location and plotted for 4 DPI (left plot), 6 DPI (middle plot), and 9 DPI (right plot). Calculated *p*-values (F-Variance test) shown in the upper corner of plot. Red=ImageJ Blue=machine learning.

2.5 Discussion

To quantify CBB, we developed and compared ImageJ and machine learning image analysis methods for accurate segmentation and quantification of water-soaked lesion symptoms. We found that an ImageJ overlay segmentation method allowed for adequate separation between cassava infected with mock, *Xam668* and *Xam668ΔTAL20* treatments area and gray-scale mean values of disease lesions. However, the ImageJ analysis was time-consuming because an individual mask had to be made for every image analyzed. Other ImageJ analysis methods tested with this data set such as non-segmentation and color-threshold based segmentation of water-soaked lesions failed to accurately capture the water-soaking phenotype.

Machine learning has previously been applied to detect and measure several cassava diseases including bacterial blight, brown streak and mosaic disease (Sangbamrung 2020 [45], Ramcharan 2019 [46]). However, these tools rely on hundreds to thousands of images for classifier training. Any machine learning tool is heavily reliant on its classifier file for adequate segmentation and measure of an object of interest. If a classifier file does not adequately capture the range of traits for an object of interest, classification of that object will fail. To determine if a classifier file would work accurately for our data set, we tested its predictive capability by spot checking analysis accuracy in a subset of images and visually inspecting classification of pixels defined as water-soaked. We initially developed classifier files based on a single representative CBB infected leaf image and found it could not reliably predict features of interest for all images. However, by combining representative images of cassava infected with three replicates each of

mock, *Xam668,* and *Xam668ΔTAL20* treatments across different timepoints into one training graphic, we developed a classifier that better predicted water-soaked lesions. The accuracy of the combined leaf graphic was tested by again spot-checking a subset of color map images and inspecting classification of pixels defined as “water-soaked”. Similarly, our classifier file was developed using one genotype of cassava, TME419. In future studies, if this approach were to be applied to datasets derived from multiple genotypes or a breeding program, the classifier file would need to be updated with representative images to capture any additional variability in leaf traits.

Another important consideration for classifier file development is the machine learning algorithm used. The machine learning workflow presented here functions with either support vector machine (SVM) or Naïve Bayes learning algorithms. During testing of classifier files, we found that SVM training files predicted water-soaked lesion features in our system more accurately than Naïve Bayes. Similarly, a previous study tested three machine learning methods and reported that SVM had high performance in predicting and classifying cassava diseases (Ramcharan 2017 [47]).

Despite the limitations, we found that the few-shot machine learning based image analysis tool presented here offered a fast and accurate approach to segment water-soaked lesions. Processing for the machine learning tool took less than half the time of ImageJ based analysis for each image. The machine learning tool worked as well as the ImageJ overlay segmentation method for separating lesions by treatment type and extracting quantifiable data. Due to the time needed

to validate a classifier file, we suggest that a machine learning approach for image-based lesion analysis is appropriate when there is a large number of images to be processed. If the data set is small, ImageJ could be a faster approach as the accuracy of the method does not rely on a classifier file. Moreover, manual thresholding is still required for segmentation of the lesions in each image and may be slightly variable within the data set. Thresholding performed within either the machine learning or ImageJ methods requires user decision to determine the threshold cut-off. In the case of the machine learning tool, it is important to inspect the color maps generated for each image analyzed to ensure proper classification of water-soaked lesions. In some cases, we found it necessary to re-process images in the machine learning tool and adjust the threshold for more precise capture of a lesion.

While improvement is still needed in image-based phenotyping, there are several potential uses for the machine learning and ImageJ analyses presented in this study. Image based phenotyping has become increasingly popular for examining the link between disease symptoms and genetics in plant science (Casto 2021 [48]). The tools presented here provide a new resource for experiments investigating CBB disease susceptibility. Additionally, the general framework of the machine learning workflow can be applied to other plant species and disease symptoms using classifier files representative of the disease of interest.

2.7 Conclusions

To quantify CBB, we developed and compared ImageJ and machine learning image analysis methods for accurate segmentation and quantification of water-soaked lesion symptoms.

Both the ImageJ and machine learning image analysis methods are described in detail, along with video tutorials and we hope these resources will help other researchers use these tools and/or design similar tools that can be applied to other pathosystems. We found that both methods accurately distinguished between and quantified different water-soaked lesion types in the cassava-*Xanthomonas* pathosystem. The ImageJ method is best used from smaller datasets as it relies on the user developing a mask for every image. The machine learning based tool is best used for larger datasets as it is more time efficient to develop a single classifier file to process many images. Many machine learning tools rely on thousands of training images for accurate function. However, the machine learning tool presented here is few-shot learning based and functions as well as ImageJ for disease segmentation and measurement.

2.8 Materials and Methods

2.8.1 Plant materials and growing conditions

Cassava plants from the cultivar TME204 were kept in greenhouse conditions set to 28°C; 50% humidity; 16 hrs light / 8 hrs dark and 1000 W light fixtures that supplemented natural light levels below 400 W / m^2. Cuttings were taken from the woody stem of mature plants and propagated to 4-inch pots of Berger45 soil. 4–5-week-old propagated plants that were well established were used for infection experiments. During infection experiments, plants were kept in a post-inoculation room set to 50% humidity, ambient room temperature, 12 hrs light / 12hr dark and 32 W light fixtures.

2.8.2 Bacterial inoculations

Xanthomonas strains were struck from glycerol stocks onto NYG agar plates containing appropriate antibiotics. The strains used for this study were *Xam668* (rifampicin 50 µg/ml) and *Xam668ΔTAL20* (suicide vector knockout (Cohn and Bart et al., 2014 [8]) tetracycline 5 µg/ml, rifampicin 50 µg/ml). *Xanthomonas* strains were grown in a 30°C incubator for 2-3 days. Inoculum for each strain was made by transferring bacteria from plates into 10mM $MgCl_2$ using inoculation loops and brought up to a concentration of $OD_{600} = 0.01$. Leaves from 4–5-week-old cassava plants were inoculated using a 1.0 mL needleless syringe. For each replicate assay, two cassava plants were used for inoculations and four leaves were inoculated on each plant. One bacterial strain was inoculated per leaf lobe with three injection sites. Mock inoculations of 10mM $MgCl_2$ alone were included resulting in nine infiltrated sites per leaf. Four replicate rounds of inoculations were done in total.

2.8.3 Imaging

Cassava leaves were detached and imaged at 0-, 4-, 6-, and 9-days post inoculation (DPI). One leaf from each cassava plant was collected and imaged for a total of two leaves per timepoint. In all, thirty-two leaves were imaged and analyzed across four replicate rounds of inoculations. Leaves were imaged from above using a Raspberry Pi Sony IMX219 camera in an enclosed box with an overhead light. To account for setting inconsistencies between images, images were color-corrected by gray balancing using a X-Rite ColorChecker Passport color card. Images were uploaded to the machine learning workflow and six gray color chips (black-white) were manually selected using a selection tool built into the program. Saturation of each chip was estimated and

the brightness of each image was adjusted accordingly. The gray corrected images were then used for water-soaking analysis. Analytical standardization of the gray values post-image-processing by ImageJ and machine learning was performed separately by estimating the grand mean of all gray values within each image and centering those values to the grand mean across all images. This is achieved by creating a linear model with a single fixed effect term accounting for each image and extracting model residuals.

2.8.4 ImageJ image analysis

Gray corrected images were uploaded to ImageJ version FIJI (Schindelin 2012 [49]) and duplicated. Water-soaked lesions were manually outlined on the duplicate image using the pencil tool (color: #ff00b6 and size 2). The outlined images were converted from RGB to LAB and split to obtain the A color channel. The A channel images were thresholded, converted to a mask and the mask for each spot was added to the ROI manager using the analyze particle tool. The ROI masks were applied to the original RGB gray corrected images. Mock infiltrated spots (no water-soaking, plant background data) were added to the ROI manager using an arbitrarily sized rectangle selection tool consistently set to a W=26 and H=30. Area, gray-scale mean, and eight other measurement data were obtained for each infiltrated spot using the FIJI measure tool. The measurements were saved as a comma separated value (CSV) file. The variance explained by ten image J derived traits were calculated and plotted in the software program R using a custom partial correlations script. Area and gray-scale mean data for all lesions were compared across different treatment types and timepoints using a Kolmogorov-Smirnov (KS) statistical test in R. All plots were generated in R with a dpi=300, width=8.66, and height=6.86.

2.8.5 Machine learning image analysis

Five images of *Xanthomonas* inoculated cassava leaves from different timepoints were selected as representatives to make a classifier file for the machine learning image analysis tool. The images were combined into one graphic, uploaded to ImageJ, and water-soaked spots were outlined and filled in using the pencil tool (color: #ff00b6). The outlined combined leaf image was converted to a binary mask and referred to as the "labeled image". The machine learning image analysis tool is part of PhenotyperCV, a C++11 header-only library designed for image-based plant phenotyping. The machine learning workflow and software download instructions are available on GitHub (https://github.com/jberry47/ddpsc_phenotypercv/wiki/Machine-Learning-Workflow).

All steps of the machine learning workflow were run on the Mac terminal command line. The labeled leaf mask image and original combined leaf graphic were used to create a support vector machine learning classifier or YAML file. Individual images of inoculated cassava leaves were processed in the machine learning tool by uploading the images and gray correcting. The images were thresholded using a scale bar built into the program to set a cut-off for pixels that can be classified as water-soaked. The inoculated sites were manually selected with a color-coded region of interest (ROI) selector (mouse right click-red, left click-green, and middle click-blue). The ROI selector tool size ranges from 0-20. The ROI size was consistently set to 11 for this study. The ROI selector does not restrict the size of the object identified as a water-soaked lesion. If a part of the object defined as a lesion is included in the ROI selection, then the entire object will be labelled and color-coded. For this study, we designated red as *Xam668*, green as *Xam668ΔTAL20,*

and blue as mock inoculation spots. If color-code separation is not required for other studies using the machine learning tool, one click/color type can be used for all lesion selections. Outputs from the workflow include a color corrected image (also used in the ImageJ analysis), a prediction image of what could be captured as pixels of interest, and a pseudo-colored map image showing what was captured as pixels of interest. Additionally, two space separated text files were generated with measurement data about the shape and color of each lesion. The shape file includes nineteen trait measures such as area, height, circularity, etc. The color file includes is a lightness histogram of 0-255 for each lesion. The text files were uploaded into R and processed using a custom script designed to read and format the data and create a comma separated value (CSV) file. For the color file, the histogram data were used to calculate lesion gray-scale mean. The variance explained by twelve machine learning derived traits were calculated and plotted in R using a custom partial correlations script. Area and gray-scale mean data for all lesions were compared across different treatment types using a Kolmogorov-Smirnov (KS) statistical test in R. All plots were generated in R with a dpi=300, width=8.66, and height=6.86.

2.9 Supplemental Information

Supplemental File 2.1: Movie example of ImageJ based analysis method. Available online at https://youtu.be/EtEzRls4Jh4

Supplemental Table 2.1: Machine learning tools commands. A table of the command syntax, function, and description of inputs and outputs for each command.

Tool Commands:	Function:	Inputs Needed:	Outputs:
ML_CREATE	Create .YAML classifier file based on of Naïve Bayes or Support Vector Machine learning algorithms	Representative RGB (input) and mask (labeled) images	.YAML file
ML_PRED	Create prediction image of what pixels will be classified as object of interest based on classifier file	Classifier file, input image, and type of machine learning method used to create classifier file	Prediction .PNG file
ML_PROC	Process all images using appropriate classifier type	Classifier file, input image, and type of machine learning method used to create classifier file, size of ROI selection box	Gray corrected, color-map and prediction .PNG file for each image and .txt files with shape and color data for each object
ML_STAT	Create summary statistics of classifier file accuracy rate	Classifier file, input image, and type of machine learning method used to create classifier file	Summary true positive, false positive, false negative, true negative, precision, recall, accuracy, true positive rate, false positive rate, true negative rate, and false discovery rate statistics for the classifier file

Supplemental File 2.2: Movie example of machine learning based analysis method.
Available online at https://youtu.be/Dw2VebjExZw

Measurement type:	Description:
Area	Area of object
Hull_area	Area of convex hull (shape of the smallest convex set)
Solidity	Ratio: area / convex hull area
Perimeter	Perimeter around object
Width	Width of object
Height	Height of object
Cmx	Center of mass x: position of center of mass
Cmy	Center of mass y: position of center of mass
Hull_verticies	Number of convex hull vertices
Ex	Ellipse center x: position of the center of the minimum bounding ellipse
Ey	Ellipse center y: position of the center of the minimum bounding ellipse
E-major	Ellipse major axis: length of the major axis of the minimum bounding ellipse
E-minor	Ellipse minor axis: length of the minor axis of the minimum bounding ellipse
E-angle	Ellipse angle: angle of rotation of the bounding ellipse major axis
Eccen	Ellipse eccentricity: eccentricity of the bounding ellipse
Circ	Circularity
round	Roundness
Ar	Aspect ratio
Fd	Box count fractal dimension
Gray scale	Lightness channel histogram from 0-255

Supplemental Table 2.2: Machine Learning Measurement Types
A table of measurements generated from the machine learning tool and their descriptions

2.10 Application of ImageJ Analysis Method to Epigenetically Modified Cassava

The work discussed in this section was published in the journal Nature Communications:

Veley KM, **Elliott K**, Jensen G, Zhong Z, Feng S, Yoder M, et al. (2022) Improving cassava bacterial blight resistance by editing the epigenome. *Nat Commun* **14**, 85 (2023). https://doi.org/10.1038/s41467-022-35675-7

In addition to CRISPR/Cas9 editing, our lab took an epigenetic engineering approach to prevent TAL20-mediated induction of *MeSWEET10a.* The epigenetics project is led by research scientist, Dr. Kira Veley who developed the construct DMS3-ZF containing a Zinc-Finger (ZF) targeting the *MeSWEET10a* EBE fused to *Arabidopsis* DMS3, an *RNA-directed* DNA methylation (RdDM) component. DMS3 is part of a POL V recruiting complex and is required for the production of POL V transcribed RNAs during RdDM (Erdmann and Picard 2020 [50]). With this construct, two DMS3-ZF lines with CG methylation at the *MeSWEET10a* EBE (Lines 133 and 204) were made. Additionally, a ZF-only line (#216) without DMS3 was generated as a control. We hypothesized that DMS3-ZF lines would have reduced CBB susceptibility compared to wildtype plants or ZF-only control plants.

I contributed to this project as the second author by completing plant infections, water-soaking analysis, and bacterial growth analysis. For water-soaking analysis, DMS3-ZF lines and control plants were infected with Xam WT, Xam ΔTAL20, and mock (10mM $MgCl_2$) treatments and infected leaf samples were collected at 0 and 4 DPI. ImageJ based analysis of the water-soaked lesions was completed. Results showed that DMS3-ZF plants infected with Xam WT had

significantly reduced water-soaked area and gray scale intensity compared to wildtype and ZF-only infected plants (**Figure A2.1**). For bacterial growth assays, DMS3-ZF and control plants were infected with Xam WT, XamΔTAL20, and mock (10mM $MgCl_2$) treatments, and at 0 and 4 DPI, leaf punches were collected. No difference in bacterial growth was detected between DMS3-ZF plants compared to control plants (**Figure A2.2)**. Results and methodology from the DMS3-ZF plant bacterial growth water-soaking analyses are available in (**Appendix 2**).

2.11 References

1. Access to food in 2020. Results of twenty national surveys using the Food Insecurity Experience Scale (FIES). FAO; 2021. https://doi.org/10.4060/cb5623en.
2. Strange RN. Introduction to plant pathology. New York, NY, USA: Wiley; 2003.
3. Liu X, Sun Y, Kørner CJ, Du X, Vollmer ME, Pajerowska-Mukhtar KM. Bacterial Leaf Infiltration Assay for Fine Characterization of Plant Defense Responses using the Arabidopsis thaliana-Pseudomonas syringae Pathosystem. J Vis Exp 2015. https://doi.org/10.3791/53364.
4. Gaunt RE. The Relationship Between Plant Disease Severity and Yield. Annual Review of Phytopathology 1995;33:119–44. https://doi.org/10.1146/annurev.py.33.090195.001003.
5. Moore WC. The measurement of plant diseases in the field: Preliminary report of a sub-committee of the Society's Plant Pathology Committee n.d.:8.
6. Plant Pathology - 5th Edition n.d. https://www.elsevier.com/books/plant-pathology/agrios/978-0-08-047378-9 (accessed March 30, 2022).
7. Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, et al. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proc Natl Acad Sci USA 2012;109:E1972-1979. https://doi.org/10.1073/pnas.1208003109.
8. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, et al. Xanthomonas axonopodis virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. Mol Plant Microbe Interact 2014;27:1186–98. https://doi.org/10.1094/MPMI-06-14-0161-R.
9. Díaz Tatis PA, Herrera Corzo M, Ochoa Cabezas JC, Medina Cipagauta A, Prías MA, Verdier V, et al. The overexpression of RXam1, a cassava gene coding for an RLK, confers disease resistance to Xanthomonas axonopodis pv. manihotis. Planta 2018;247:1031–42. https://doi.org/10.1007/s00425-018-2863-4.

10. Jorge V, Verdier V. Qualitative and quantitative evaluation of cassava bacterial blight resistance in F1 progeny of a cross between elite cassava clones n.d.:8.
11. Poland JA, Nelson RJ. In the Eye of the Beholder: The Effect of Rater Variability and Different Rating Scales on QTL Mapping. Phytopathology® 2011;101:290–8. https://doi.org/10.1094/PHYTO-03-10-0087.
12. Strange RN, Scott PR. Plant Disease: A Threat to Global Food Security. Annual Review of Phytopathology 2005;43:83–116. https://doi.org/10.1146/annurev.phyto.43.113004.133839.
13. Gehan MA, Fahlgren N, Abbasi A, Berry JC, Callen ST, Chavez L, et al. PlantCV v2: Image analysis software for high-throughput plant phenotyping. PeerJ 2017;5:e4088. https://doi.org/10.7717/peerj.4088.
14. Laflamme B, Middleton M, Lo T, Desveaux D, Guttman DS. Image-Based Quantification of Plant Immunity and Disease. MPMI 2016;29:919–24. https://doi.org/10.1094/MPMI-07-16-0129-TA.
15. Lobet G. Image Analysis in Plant Sciences: Publish Then Perish. Trends in Plant Science 2017;22:559–66. https://doi.org/10.1016/j.tplants.2017.05.002.
16. Li L, Zhang Q, Huang D. A Review of Imaging Techniques for Plant Phenotyping. Sensors (Basel) 2014;14:20078–111. https://doi.org/10.3390/s141120078.
17. Zhang Y, Zhang N. Imaging technologies for plant high-throughput phenotyping: a review. Front Agr Sci Eng 2018;5:406–19. https://doi.org/10.15302/J-FASE-2018242.
18. Ferreira T, Rasband W. ImageJ User Guide n.d.:198.
19. Bock CH, Parker PE, Cook AZ, Gottwald TR. Visual Rating and the Use of Image Analysis for Assessing Different Symptoms of Citrus Canker on Grapefruit Leaves. Plant Disease 2008;92:530–41. https://doi.org/10.1094/PDIS-92-4-0530.
20. Bierman A, LaPlumm T, Cadle-Davidson L, Gadoury D, Martinez D, Sapkota S, et al. A High-Throughput Phenotyping System Using Machine Vision to Quantify Severity of Grapevine Powdery Mildew. Plant Phenomics 2019;2019:9209727. https://doi.org/10.34133/2019/9209727.
21. Gallego-Sánchez LM, Canales FJ, Montilla-Bascón G, Prats E. RUST: A Robust, User-Friendly Script Tool for Rapid Measurement of Rust Disease on Cereal Leaves. Plants (Basel) 2020;9:1182. https://doi.org/10.3390/plants9091182.
22. Mutka AM, Bart RS. Image-based phenotyping of plant disease symptoms. Frontiers in Plant Science 2015;5:734. https://doi.org/10.3389/fpls.2014.00734.
23. Stewart EL, McDonald BA. Measuring Quantitative Virulence in the Wheat Pathogen Zymoseptoria tritici Using High-Throughput Automated Image Analysis. Phytopathology® 2014;104:985–92. https://doi.org/10.1094/PHYTO-11-13-0328-R.
24. Stewart EL, Hagerty CH, Mikaberidze A, Mundt CC, Zhong Z, McDonald BA. An Improved Method for Measuring Quantitative Resistance to the Wheat Pathogen Zymoseptoria tritici Using High-Throughput Automated Image Analysis. Phytopathology 2016;106:782–8. https://doi.org/10.1094/PHYTO-01-16-0018-R.

25. Singh A, Ganapathysubramanian B, Singh AK, Sarkar S. Machine Learning for High-Throughput Stress Phenotyping in Plants. Trends in Plant Science 2016;21:110–24. https://doi.org/10.1016/j.tplants.2015.10.015.
26. Tsaftaris SA, Minervini M, Scharr H. Machine Learning for Plant Phenotyping Needs Image Processing. Trends in Plant Science 2016;21:989–91. https://doi.org/10.1016/j.tplants.2016.10.002.
27. Morgan NK, Choct M. Cassava: Nutrient composition and nutritive value in poultry diets. Animal Nutrition 2016;2:253–61. https://doi.org/10.1016/j.aninu.2016.08.010.
28. Bart RS, Taylor NJ. New opportunities and challenges to engineer disease resistance in cassava, a staple food of African small-holder farmers. PLOS Pathogens 2017;13:e1006287. https://doi.org/10.1371/journal.ppat.1006287.
29. Hillocks RJ, Thresh JM, Bellotti A. Cassava: Biology, Production and Utilization. CABI; 2002.
30. EL-Sharkawy MA. Cassava biology and physiology. Plant Mol Biol 2003;53:621–41. https://doi.org/10.1023/B:PLAN.0000019109.01740.c6.
31. Howeler RH, Lutaladio N, Thomas G. Save and grow: cassava: a guide to sustainable production intensification. Rome: Food and Agriculture Organization of the United Nations; 2013.
32. Fanou AA, Zinsou VA, Wydra K. Cassava Bacterial Blight: A Devastating Disease of Cassava. Cassava 2017. https://doi.org/10.5772/intechopen.71527.
33. Zárate-Chaves CA, Gómez de la Cruz D, Verdier V, López CE, Bernal A, Szurek B. Cassava diseases caused by Xanthomonas phaseoli pv. manihotis and Xanthomonas cassavae. Molecular Plant Pathology 2021;22:1520–37. https://doi.org/10.1111/mpp.13094.
34. Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, et al. Genetic characterization of strains named as Xanthomonas axonopodis pv. dieffenbachiae leads to a taxonomic revision of the X. axonopodis species complex. Plant Pathology 2016;65:792–806. https://doi.org/10.1111/ppa.12461.
35. Aung K, Jiang Y, He SY. The role of water in plant–microbe interactions. The Plant Journal 2018;93:771–80. https://doi.org/10.1111/tpj.13795.
36. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 2010;48:419–36. https://doi.org/10.1146/annurev-phyto-080508-081936.
37. Hogenhout SA, Van der Hoorn RAL, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-associated organisms. Mol Plant Microbe Interact 2009;22:115–22. https://doi.org/10.1094/MPMI-22-2-0115.
38. Muñoz Bodnar A, Bernal A, Szurek B, López CE. Tell Me a Tale of TALEs. Mol Biotechnol 2013;53:228–35. https://doi.org/10.1007/s12033-012-9619-3.
39. Van Schie CCN, Takken FLW. Susceptibility Genes 101: How to Be a Good Host. Annual Review of Phytopathology 2014;52:551–81. https://doi.org/10.1146/annurev-phyto-102313-045854.

40. Koseoglou E, van der Wolf JM, Visser RGF, Bai Y. Susceptibility reversed: modified plant susceptibility genes for resistance to bacteria. Trends in Plant Science 2021:S1360138521002053. https://doi.org/10.1016/j.tplants.2021.07.018.
41. Li T, Liu B, Spalding M, Weeks D, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nature Biotechnology 2012;30:390–2. https://doi.org/10.1038/nbt.2199.
42. Phillips AZ, Berry JC, Wilson MC, Vijayaraghavan A, Burke J, Bunn JI, et al. Genomics-enabled analysis of the emergent disease cotton bacterial blight. PLOS Genetics 2017;13:e1007003. https://doi.org/10.1371/journal.pgen.1007003.
43. Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, et al. TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. Nature Communications 2017;8:1–14. https://doi.org/10.1038/ncomms15588.
44. Berry JC, Fahlgren N, Pokorny AA, Bart RS, Veley KM. An automated, high-throughput method for standardizing image color profiles to improve image-based plant phenotyping. PeerJ 2018;6. https://doi.org/10.7717/peerj.5727.
45. Sangbamrung I, Praneetpholkrang P, Kanjanawattana S. A Novel Automatic Method for Cassava Disease Classification Using Deep Learning. JAIT 2020;11:241–8. https://doi.org/10.12720/jait.11.4.241-248.
46. Ramcharan A, McCloskey P, Baranowski K, Mbilinyi N, Mrisho L, Ndalahwa M, et al. A Mobile-Based Deep Learning Model for Cassava Disease Diagnosis. Frontiers in Plant Science 2019;10:272. https://doi.org/10.3389/fpls.2019.00272.
47. Ramcharan A, Baranowski K, McCloskey P, Ahmed B, Legg J, Hughes DP. Deep Learning for Image-Based Cassava Disease Detection. Frontiers in Plant Science 2017;8. https://doi.org/10.3389/fpls.2017.01852.
48. Picturing the future of food - Casto - 2021 - The Plant Phenome Journal - Wiley Online Library n.d. https://acsess-onlinelibrary-wiley-com.libproxy.wustl.edu/doi/full/10.1002/ppj2.20014?af=R (accessed December 21, 2021).
49. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–82. https://doi.org/10.1038/nmeth.2019.
50. Erdmann RM, Picard CL. RNA-directed DNA Methylation. PLoS Genet 2020;16:e1009034. https://doi.org/10.1371/journal.pgen.1009034.

Chapter 3: Applying CRISPR/Cas9 Genome Editing to Investigate the Role of *MeSWEET10a* in Promoting Cassava Bacterial Blight

This chapter is pending upcoming journal submission.

3.1 Personal Contributions

I served as the first author and led project design and progression with advising from Dr. Rebecca Bart. Dr. Kira Veley designed the CRISPR/Cas9 constructs and developed the first generation of *MeSWEET10a* mutants along with Greg Jensen. I completed additional transformations with all constructs and genotyped *MeSWEET10a* mutant lines recovered from these transformations. I completed bacterial infections, RT-PCR, bacterial growth, and water-soaked lesion analyses. Additionally, I collected, dissected, and imaged all flowers used for mutant versus wild-type morphology comparisons. Joanna Norton and Lukas Kambic completed crosses of mutant and wildtype plants and recovered F1 seed in the field supervised by Dr. Sharon Motomura-Wages. I measured and tested the germination rate of the F1 seed with the aid of Marisa Yoder. I wrote the manuscript, completed statistical analysis, and generated all figures with the exception of the supplemental transgene insertion graphics generated by Kerrigan Gilbert.

3.2 Abstract

Bacteria from the genus *Xanthomonas* are prolific phytopathogens that elicit disease in over 400 plant species. Xanthomonads carry a repertoire of specialized proteins called transcription activator-like (TAL) effectors that promote disease and pathogen virulence by inducing expression of host susceptibility (S) genes. *Xanthomonas phaseoli* pv. *manihotis* (Xpm) causes bacterial blight on the staple food crop, cassava. The Xpm effector, TAL20, induces ectopic expression of the S gene *MeSWEET10a*, a sugar transporter that contributes to cassava bacterial blight susceptibility. We used CRISPR/Cas9 to generate multiple cassava lines with different edits to the *MeSWEET10a* TAL20 effector binding site and/or coding sequence. In several of the transgenic lines, *MeSWEET10a* expression was no longer induced by *Xpm* and in these cases, we observed reduced cassava bacterial blight disease symptoms post-Xpm infection. Prior work showed that *MeSWEET10a* has native expression in cassava flowers. Therefore, we investigated flower development and reproductive function in a *MeSWEET10a* mutant line. We found that the *MeSWEET10a* mutant produced phenotypically wildtype cassava and F1 seed. Thus, blocking *MeSWEET10a* induction is a viable strategy for decreasing cassava susceptibility to CBB.

3.3 Introduction

Cassava (*Manihot esculenta* Crantz) is a starchy root crop that serves as a carbohydrate source and food security crop for nearly 800 million people globally (Alves 2002 and Morgan 2016). Cassava is tolerant to abiotic stressors such as poor soil quality and is often grown without costly inputs like fertilizer (EL-Sharkawy, 2003). Cassava is especially important for smallholder

farmers in Sub-Saharan Africa who grow cassava as a sustenance crop and sell it for revenue when yields allow (Taylor et al., 2003 and Hillocks et al., 2003). A leading biotic factor threatening cassava production is cassava bacterial blight (CBB). CBB disease symptoms include water-soaked leaf lesions, chlorosis, defoliation, and stem browning (Lozano, 1986). CBB is present in all cassava growing regions and can result in total crop loss including the stem used to plant a subsequent crop through clonal propagation (Lozano et al., 1980 and López and Bernal, 2012).

The causal agent of CBB is a gram-negative phytopathogen in the genus *Xanthomonas*. Xanthomonads elicit disease in over 400 plant species including economically important crops such as rice, cotton, sorghum, and citrus (Leyns et al., 1984, Jacques et al., 2017, and Mhedbi-Hajri et al., 2013). The *Xanthomonas* specific to cassava was recently reclassified as *X. phaseoli* pv. *manihotis* (Xpm) (Constantin et al., 2015) and was formerly known as *X. axonpodis* pv. *manihotis* (Xam). Xpm is dispersed from plant to plant through rain, wind, or by propagation of already infected stem cuttings. Xpm persists on the leaf surface and enters the endophytic stage of infection upon invasion of the leaf through open stomata or wounds (Kandel et al, 2017). *In planta*, *Xanthomonas* colonizes the surface of mesophyll cells and some Xanthomonads can systemically spread throughout the plant vasculature (An et al., 2019 and Ryan et al., 2011).

Xpm induces effector-triggered susceptibility (ETS) using an arsenal of effector proteins released into the plant through a needle-like projection that penetrates the host cell wall called the type III secretion system (T3SS) [(Bart et al., 2012) and (Abrusci et al., 2014)]. T3SS effectors manipulate the host to help the pathogen overcome plant defenses and promote disease (Hogenhout et al., 2009). Bacteria in the *Xanthomonas* and *Ralstonia* genera have specialized transcription

activator-like (TAL) effectors that induce expression of host susceptibility (S) genes to enhance pathogenesis (Van Schie and Takken, 2014, and Eckardt, 2002). TAL effectors structurally resemble eukaryotic transcription factors and consist of an activation domain, nuclear localization signal, and a DNA binding domain that is a series of tandem amino acid repeats (Schornack et al., 2013). Upon secretion into the host cell, TAL effectors localize to the nucleus and bind directly to DNA sites, called effector binding elements (EBEs), upstream of S genes. Following EBE binding, TAL effectors induce ectopic or overexpression of susceptibility genes. TAL effector binding to EBE sites is sequence specific and the tandem repeats create a code used to identify candidate binding sites and S genes (Cernadas et al., 2014, Moscou and Bogdanove 2009 and Boch 2009). Xpm strains typically carry between one to five TAL effectors and the model Xpm strain used in this study, Xpm668 (formerly referred to as Xam668), has five TAL effectors: TAL13, TAL14, TAL15, TAL20 and TAL22 (Bart et al., 2012).

Previous work reported that Xpm mutants with loss of the TAL20 effector (Xpm668ΔTal20) exhibit visibly reduced water-soaked lesions and have decreased bacterial growth *in planta*. A member of the SWEET (Sugars Will Eventually be Exported Transporter) gene family, *MeSWEET10a* (Gene ID: Manes.06G123400), was identified as the S gene target for TAL20 (Cohn and Bart et al., 2014). TAL20 binding to the *MeSWEET10a* EBE induces ectopic gene expression in the leaf and results in sugar transport into the apoplast where Xpm proliferates (Chen 2014 and 2010). SWEET genes are established TAL effector targets in several plant species including rice, pepper, and cotton (Antony et al., 2010, Hu et al., 2014, Cox et al., 2017, and Philips et al., 2017). Various studies have demonstrated that preventing TAL effector interaction with a

SWEET gene EBE reduces plant susceptibility to diseases induced by Xanthomonads (Gupta et al., 2021 and Veley et al., 2023). For example, research on rice bacterial blight, induced by *X. oryzae* pv. *oryzae (Xoo*), found that CRISPR/Cas9-mediated editing of three rice SWEET genes decreased plant susceptibility to bacterial blight (Olivia et al., 2019). EBE site-specific edits of S genes beyond the SWEET family have also improved plant resistance to disease. For instance, CRISPR/Cas9 generated edits in the EBE of *X. citri* subsp. *citri* (Xcc) S gene target, *Cs LATERAL ORGAN BOUNDARIES 1 (LOB1)*, resulted in canker-resistant sweet orange (Huang et al., 2022).

In this study, we used a dual gRNA CRISPR/Cas9 strategy to generate *MeSWEET10a* mutant lines with edits to the TAL20 EBE and/or gene coding sequence. We characterized the disease phenotypes of Xpm infected plants and demonstrated that *MeSWEET10a* mutants exhibit reduced cassava bacterial blight symptoms. Additionally, while *MeSWEET10a* is not normally expressed in cassava leaves, prior work showed there is endogenous expression in cyathium or "false flower" tissue (Veley et al., 2021 and Perera et al., 2012). In rice, knocking out the SWEET gene, *OsSWEET15*, led to reduced rice fertility (Hu et al., 2023). Therefore, we investigated the impact of editing *MeSWEET10a* on cassava flower development and reproductive function. We found that *MeSWEET10a* mutant cassava plants developed flowers morphologically similar to wildtype plants based on macro imaging. When these flowers were used for crosses, they produced fruit, and viable F1 seed.

3.4 Results

3.4.1 First Generation of *MeSWEET10a* Mutants

We hypothesized that editing *MeSWEET10a* would reduce cassava susceptibility to *Xpm*. To test this hypothesis, we designed a single CRISPR/Cas9 construct (construct 108) with two guide RNAs (gRNAs), gRNA1 and gRNA2 which target the TAL20 EBE site and the translation start site (ATG), respectively. Additionally, construct 108 contains a repair template with homology arms that flank the EBE to allow for potential CRISPR-mediated homology directed repair (HDR). We predicted that we would recover a mix of mutants with edits at one or both target sites, larger deletions between the two gRNAs, and mutants containing a repair template. The repair template was designed to replace the EBE with a sequence that TAL20 was unlikely to bind while maintaining the annotated TATA box (**Figure 3.1A**). *Agrobacterium*-mediated transformation was carried out in friable embryogenic callus (FEC) from the farmer-preferred cultivar of cassava, TME419, also referred to as WT419 (Chauhan et al., 2015). In total, thirty transgenic lines were recovered. The *MeSWEET10a* region of interest was amplified from each recovered transgenic line. Restriction digest was used to identify lines with potential EBE repair template integration and larger INDELs (**Supplemental Figure 3.1**). If the EBE repair template was integrated, we expected it to abolish a HaeIII restriction enzyme site at the gRNA2 repair template site. Restriction digests with HaeIII were performed on PCR products from each line. Wildtype plants were expected to be cut at the HaeIII site while mutants with repair template integration were not. However, there is low efficiency of HDR in plants (Putcha 2005 and Britt

and May 2003) and transgenic lines with integration of the repair template were not recovered.

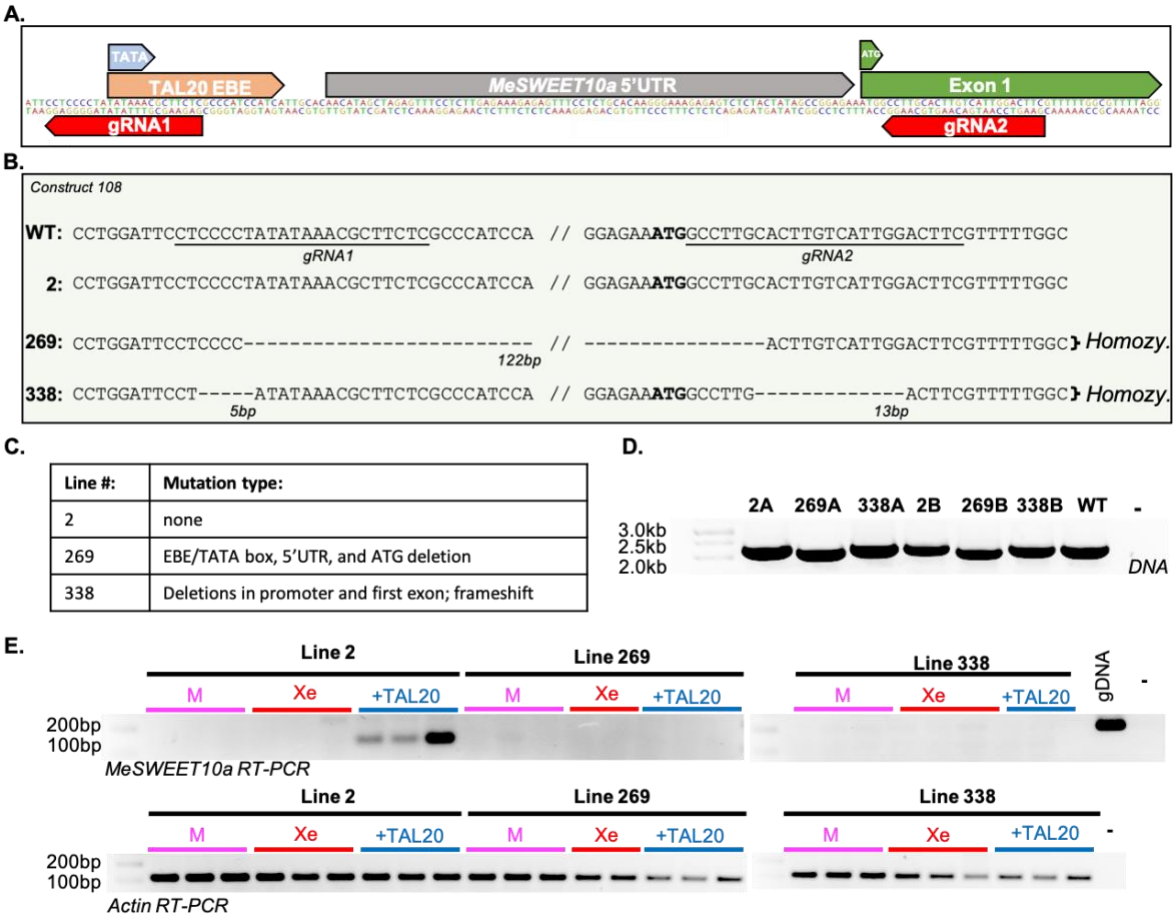


Figure 3.1: First Generation *MeSWEET10a* Mutant Lines Lack TAL20-Mediated Induction **A)** Graphic depicting the *MeSWEET10a* region of interest including the TAL20 EBE (effector binding element), TATA box, translation start site (ATG), and gRNA target sites for construct 108. **B)** Genotyping of *MeSWEET10a* mutant lines recovered from construct 108 based on Sanger-sequencing. Text indicates sequences at the region of interest for wildtype plants and mutant lines # 2, 269, and 338. Underlined text denotes the target site for each gRNA. Deletions are indicated by '-' and the number of deleted base pairs (bp) is indicated below each deletion. Deletions in lines #269 and #338 are homozygous (homozyg.) **C)** Table with description of mutation and location type for each mutant line. **D)** PCR products generated by primers targeting the *MeSWEET10a* region in gDNA from WT and lines 2, 269, and 338. A and B denote different individuals from each line. **E)** RT-PCR of wildtype cassava and *MeSWEET10a* mutant lines infiltrated with mock,

X. euvesicatoria (Xe) alone, and Xe+TAL20 treatments. Top gel shows results of RT-PCR with primers amplifying *MeSWEET10a* with an expected product size of 123 bp. The bottom gel shows results of RT-PCR with primers amplifying the housekeeping gene, *Actin*, as a control for sample loading with an expected product size of 125 bp. DNA from WT419 leaf tissue is included as a positive control and '-' denotes a negative water control. M=Mock (magenta), Xe=(red), and +TAL20=Xe+TAL20 (blue).

Lines that exhibited digest patterns different from WT419 (lines: 243, 269, 323, and 338) were moved forward for further genotyping using Sanger sequencing alone with a potential control line 2 (**Supplemental Figure 3.2A**). Based on genotype results (**Figure 3.1B**) and tissue culture survival, lines 2, 269, and 338 were selected for further examination. Line 2 sequencing results indicate that it is a WT-like transgenic and it was kept as a control. Line 269 has a large 122 bp deletion including the TATA Box, TAL20 EBE, and the *MeSWEET10a* ATG which explains the smaller band size seen in *MeSWEET10a* PCR product from line 269 DNA compared to WT419 DNA (**Figure 3.1D**). Line 338 has a 5 bp deletion upstream of the TATA box and TAL20 EBE and a 13 bp deletion after the ATG causing a frameshift and stop codon in exon 1. Lines 269 and 338 were confirmed as homozygous mutants using genomic DNA (gDNA) clone sequencing (**Supplemental Figure 3.2B**). The mutation and location types for each line were summarized in table form (**Figure 3.1C**).

We previously demonstrated that a TAL effectorless Xanthomonad, *Xanthomonas euvesicatoria* (Xe- a Xanthomonad non-pathogenic to cassava), is able to deliver TAL20 to cassava cells and induce *MeSWEET10a* expression (Cohn and Bart et al., 2014). We used this system to compare Xe, Xe +TAL20, or mock treatments for *MeSWEET10a* induction in the mutant lines 2, 269, and 338. At 48 hours post-infection (HPI), samples were collected for RNA extraction

and reverse transcriptase (RT)-PCR analysis (**Figure 3.1E**). In control line 2 plants infected with Xe +TAL20, RT-PCR results show a 123 bp product indicating TAL20-mediated induction of *MeSWEET10a*. In contrast, no product was present for plants infected with Xe alone or mock treatments. Additionally, lines 269 and 338 infected with Xe +TAL20 have no RT-PCR product indicating the mutations in each line are sufficient to prevent TAL20-mediated induction of *MeSWEET10a.*

3.4.2 Generation of Additional *MeSWEET10a* Mutants

We sought to generate additional mutant lines and ideally create mutants with edits specific to the TAL20 EBE leaving the *MeSWEET10a* gene intact for native plant function. Therefore, two additional CRISPR/Cas9 constructs were made (**Figure 3.2A**). Construct 249 contains gRNA1 and gRNA3 which localize to the TAL20 EBE site and the *MeSWEET10A* 5'UTR (untranslated region) upstream of the ATG. Construct 250 contains gRNA4 which targets upstream of the TATA box and TAL20 EBE and gRNA5 which targets the TAL20 EBE downstream of the TATA box. Four rounds of cassava transformation with all three constructs were performed. In total, twenty-four transgenic lines were recovered with seven mature lines generated from construct 108, eight from construct 249, and nine lines from construct 250. Leaf tissue was sampled from each line at the plantlet stage in tissue culture and plants were genotyped by Sanger sequencing. Twenty-three out of twenty-four lines had edits within the *MeSWEET10A* gene (**Supplemental Figure 3.3A**). One line was recovered with edits within the TAL20 binding site with an intact TATA box, however this line died during the tissue culture process.

The genotypes of five mutant lines of interest (**Figure 3.2B and 3.2C**) were confirmed using gDNA clone-sequencing to determine mutant line zygosity (**Supplemental Figure 3.3B**). Line #27 is homozygous with a 185 base pair (bp) deletion spanning the TATA box, TAL20 binding site, and 5'UTR. Additionally, there is a 1 bp frameshift insertion after the ATG. Line #30 is homozygous with a 1 bp deletion upstream of the TATA box and a 2 bp frameshift deletion downstream of the ATG. Line #41 is a biallelic mutant with one allele that has an 11 bp deletion at the TATA box/ TAL20 EBE site and a 1 bp insertion in the 5'UTR. The second allele has a 5 bp deletion at the TATA box/TAL20 EBE site and a 1 bp insertion in the 5'UTR. Line #54 is biallelic with one allele containing a 5 bp deletion at the TATA box/TAL20 EBE site and a 1bp insertion in the 5'UTR. The second allele contains 76 bp deletion spanning the TATA box/TAL20 EBE site. Line #69A is a biallelic silent mutant with one allele that has a 1 bp deletion upstream of the TATA box and one insertion in the 5'UTR and another allele that has a 1 bp deletion upstream of the TATA box. gDNA from lines 27, 30, 41, 54, and 69A all produced a *MeSWEET10a* PCR product near 2.1kb corresponding with their insertion/deletion types (**Figure 3.2D**). An overview of *MeSWEET10a* mutant types generated from all transformations is provided in **Table 3.1.** Additionally, results from select stages of the transformation pipeline are reported in **Supplemental Table 3.1**. Mutants were moved from tissue culture to soil and characterized for various cassava traits including plant height, node number, internode length, petiole length, central lobe length, central lobe width, and whole lobe length (**Figure 3.2E and Supplemental Table 3.2**). Overall, mutant plant traits were physiologically similar to wildtype cassava.

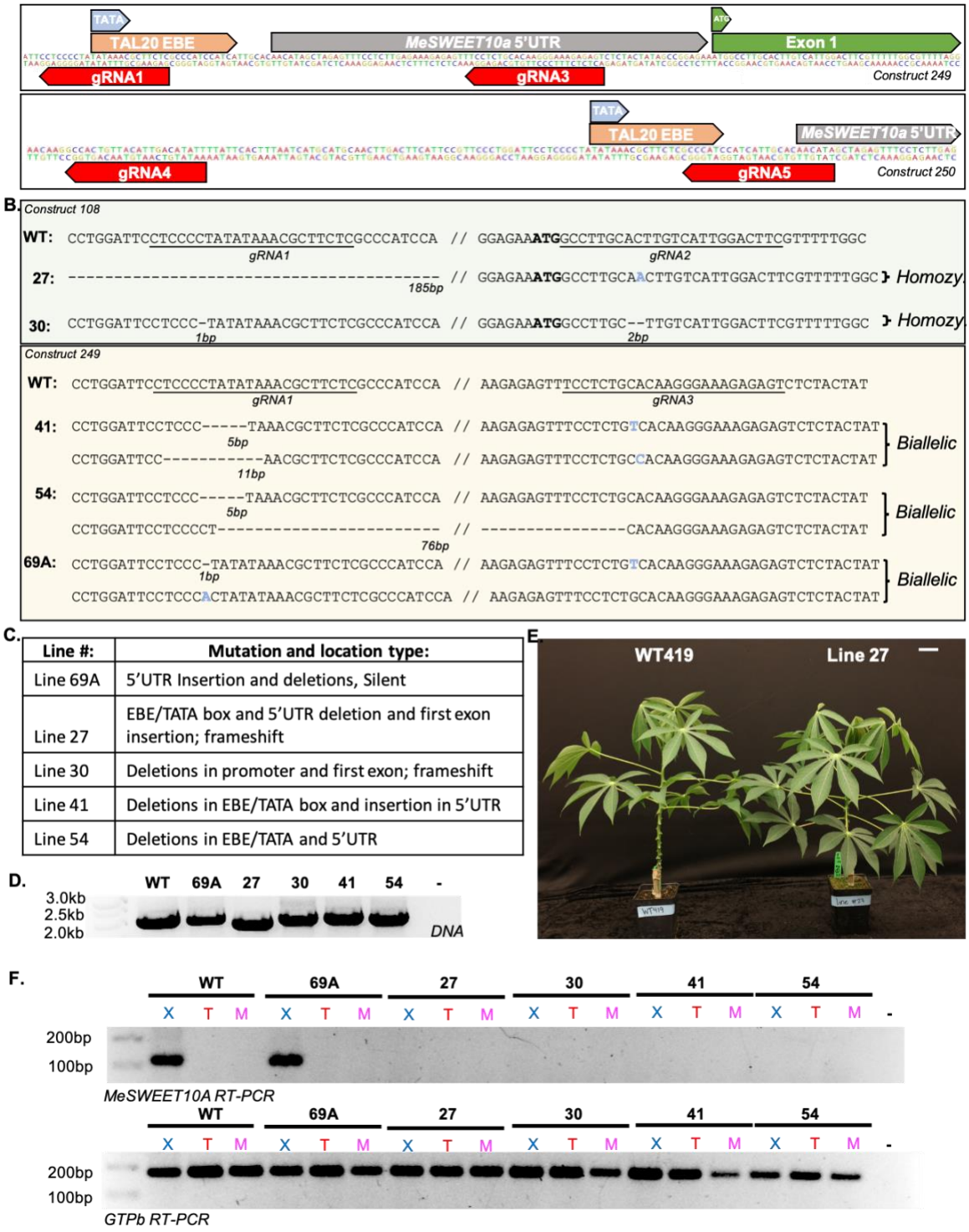


Figure 3.2: Additional *MeSWEET10a* Mutant Lines Lack TAL20-Mediated Induction

A) Graphic depicting the *MeSWEET10a* region of interest, TAL20 EBE site, TATA box, translation start site (ATG), and gRNA target sites for constructs 108, 249, and 250. **B)** Genotyping of *MeSWEET10a* mutant lines recovered from construct 108 and 249 based on Sanger-sequencing. Text indicates sequences at the region of interest for wildtype plants and mutant lines # 27, 30, 41, 54, and 69A. Underlined text denotes the gRNA target site. The italicized text below the underlined text describes gRNA number. The number of base pair (bp) deletions at each target site is depicted at the end of the sequence. Bases in blue indicate insertion events. Mutant line zygosity is indicated by 'Homozy' (homozygous) or 'biallelic' text to the right of the sequence. **C)** Table with description of mutation and location type for each mutant line. **D)** PCR products generated by primers targeting the *MeSWEET10a* region in gDNA from WT and edited lines. **E)** Representative image of wildtype cassava (left) and line 27 (right) plants grown from stake cuttings in greenhouse. Scale bar = 14cm. **F)** RT-PCR of wildtype cassava and *MeSWEET10a* mutant lines infected with Xpm WT, Xpm△TAL20, and mock treatments. The top gel shows results of RT-PCR with primers amplifying *MeSWEET10a* with an expected product size of 123 bp. The bottom gel shows results of RT-PCR with primers amplifying the housekeeping gene GTPB as a control for sample loading with an expected product size of 184 bp. '-' denotes a negative water control. X= Xpm WT (blue), T=Xpm△TAL20 (red), and M=Mock (magenta).

Mutant type:	**General description:**	**# of lines:**
WT-like transgenic	Transgenic lines with no edits at gRNA target sites	1
Silent mutants	Edits that do not impact *MeSWEET10a* coding sequence	13
TAL20 EBE promoter region INDEL mutants	Edits within the TAL20 EBE with intact TATA box	1*
MeSWEET10a promoter region INDEL mutants	Edits in the promoter region impacting the TATA box predicted to impact *MeSWEET10a* expression	11
MeSWEET10a frameshift mutants	Edits after the TSS expected to impact the *MeSWEET10a* coding sequence	3

Table 3.1: Overview of *MeSWEET10a* Mutant Line Genotypes
Mutation type summary of the 29 transgenic lines recovered from all rounds of transformation based on Sanger sequencing results. '*' denotes a line that died during the tissue culture process.

3.4.3 Characterization of *MeSWEET10a* Mutants CBB Disease Severity

To select transgenic lines for disease phenotyping assays, leaves were detached from plantlets in tissue culture, and syringe-infiltrated with Xpm WT, Xpm△TAL20, and mock treatments. Samples were collected at 48HPI for RNA extraction and RT-PCR analysis (**Supplemental Figure 3.4**). In samples from wildtype plants infected with Xpm, RT-PCR results show the expected 123 bp band indicating TAL20-mediated induction of *MeSWEET10a*. However, no product is present for wildtype plants infected with Xpm△TAL20 or mock treatments. Transgenic lines with edits predicted to prevent *MeSWEET10a* induction were screened for RT-PCR results. As predicted based on their mutation types, *MeSWEET10a* was not induced by Xpm WT infection in lines 27, 30, 41, and 54. However, RT-PCR results from the silent mutant line 69A infected with Xpm WT show that *MeSWEET10a* induction by the TAL20 effector occurs in a wildtype-like manner (**Figure 3.2F**).

The *MeSWEET10a* mutant lines were phenotyped for CBB disease severity. Bacterial growth of Xpm WT and Xpm△TAL20 was quantified over time in infected wildtype and *MeSWEET10a* mutant plants (**Figure 3.3A**). In this study, the significant difference in Xpm WT growth between wildtype and *MeSWEET10a* mutant plants varied across experimental replicates (**Supplemental Figure 3.5**). Furthermore, as shown in previously published work (Veley et al., 2023), we did not observe a consistent difference of *in planta* bacterial titer between Xpm WT and Xpm△TAL20. Therefore, water-soaking analysis was used to measure CBB disease severity more consistently. The amount of water-soaking was quantified for Xpm WT-, Xpm△TAL20-, and

mock-treated plants using a machine learning image analysis method (Elliott et al., 2022). Representative images of infected plants show visibly reduced water-soaked lesions in line #27 compared to wildtype cassava and the silent mutant line 69A. Furthermore, the water-soaked lesions for Xpm WT and Xpm△TAL20 spots appear more similar to each other in line 27. No watersoaking was visible in mock-treated spots (**Figure 3.3B**). Images of all infected leaves from each plant background are included in supplemental data. Xpm WT-infected plants from lines 27, 30, 41, and 54 all exhibited significantly reduced water-soaked lesion area compared to Xpm WT-infected wildtype cassava (**Figure 3.3C**). However, Xpm WT-infected plants from the silent mutant line 69A had water-soaked areas not significantly different from wildtype infected plants. There was no significant difference in Xpm△TAL20 lesion area between any of the mutant lines compared to wildtype plants (**Figure 3.3D**). Similar water-soaking assay results were obtained from line 269 and 338 mutant plants infected with Xpm WT and Xpm△TAL20 (S**upplemental Figure 3.6**). Therefore, we conclude that *MeSWEET10a* mutant plants have decreased susceptibility to cassava bacterial blight.

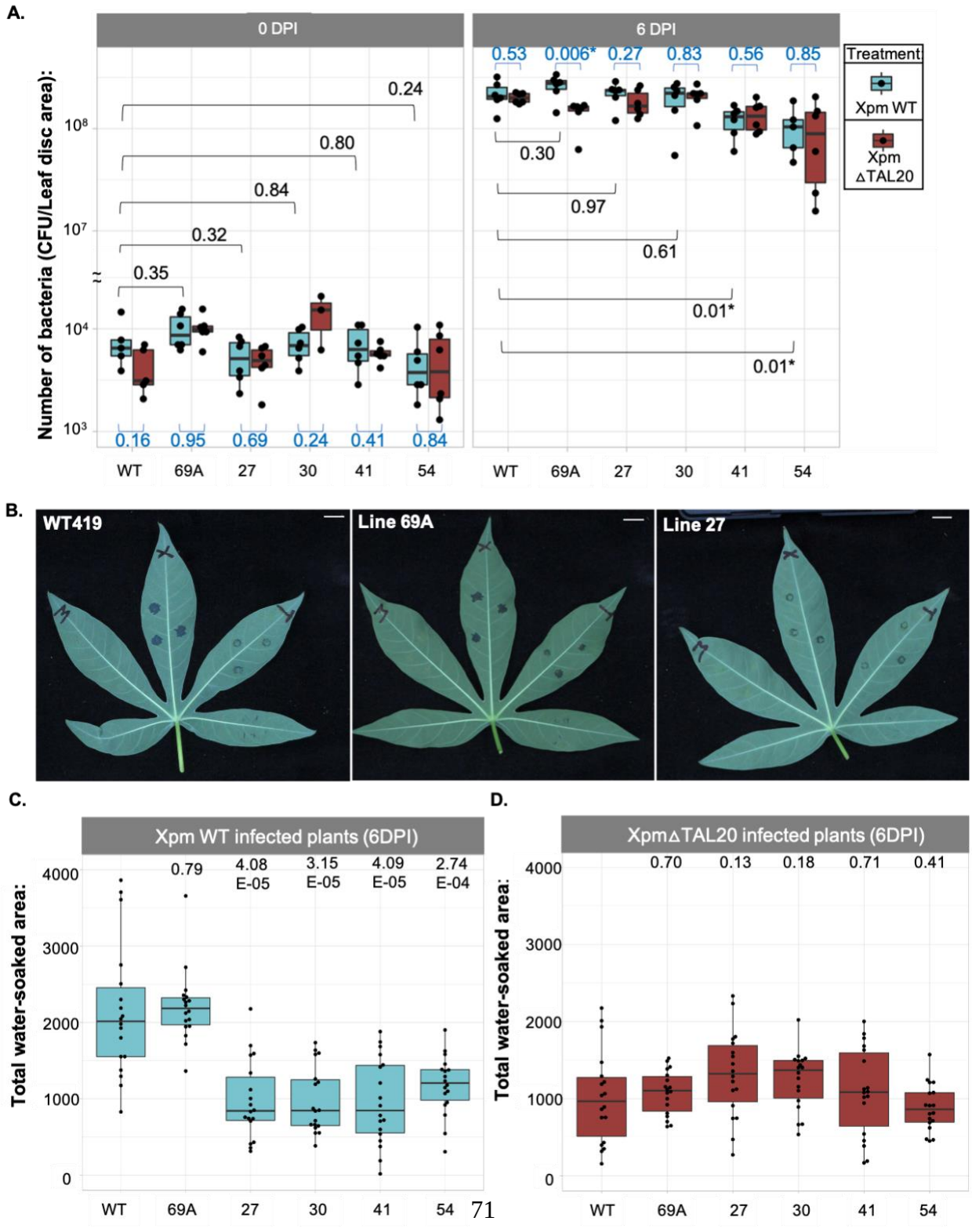

A.
0 DPI
6 DPI
Number of bacteria (CFU/Leaf disc area):
Treatment
Xpm WT
Xpm ΔTAL20
0.24
0.80
0.84
0.32
0.35
0.53
0.006*
0.27
0.83
0.56
0.85
0.30
0.97
0.61
0.01*
0.01*
0.16
0.95
0.69
0.24
0.41
0.84
WT
69A
27
30
41
54
B.
WT419
Line 69A
Line 27
C.
Xpm WT infected plants (6DPI)
Total water-soaked area:
0.79
4.08 E-05
3.15 E-05
4.09 E-05
2.74 E-04
D.
XpmΔTAL20 infected plants (6DPI)
0.70
0.13
0.18
0.71
0.41

Figure 3.3: *MeSWEET10a* Mutant Lines CBB Disease Symptoms Post Xpm Infection
A) Number of bacteria in cassava leaves measured at 0 days post inoculation (DPI) (left) and 6DPI (right) with Xpm (blue) and Xpm△TAL20 (red) treatments. Colony forming units (CFU/leaf disc area, Y-axis) are plotted by plant genotype (X-axis) tested (wildtype or mutant). Black dots represent sample replicates from an independent bacterial growth experiment. **B)** Representative images of infected wildtype (left), silent mutant line 69A (middle), and mutant line 27 (right) cassava leaves detached from the plant and imaged at 4DPI. X= Xpm WT, T=Xpm△TAL20, and M=Mock. Scale bar = 1cm. **C)** Total water-soaked area (pixels, y-axis) of Xpm WT infected plants (genotypes, x-axis) at 6DPI. **D)** Total water-soaked area (pixels, y-axis) of Xpm△TAL20 infected plants (genotypes, x-axis) at 6DPI. Black dots represent individual water-soaked lesions from three independent water-soaking assay experiments combined. In all boxplots, the calculated *p*-values (Unpaired Student's T-test with unequal variance) are shown above or below each box plot. Black text represents Xam WT comparisons between WT and mutant infected plants. Blue text represents comparisons between Xam WT and Xpm△TAL20 within each genotype. Dots outside whiskers represent outliers based on default settings of the R package ggplot2. The horizontal line within the box represents the median sample value. The ends of the boxes represent the 3rd (Q3) and 1st (Q1) quartiles. The whiskers show values that are 1.5 times interquartile range (1.5xIQR) above and below Q1 and Q3.

3.4.4 Characterization of *MeSWEET10a* Mutant Flower Morphology and F1 Progeny

Unlike leaves, cassava flowers have endogenous expression of *MeSWEET10a*. Thus, we wanted to determine if mutating *MeSWEET10a* would impact flower development or reproductive function. Since they were further along in tissue culture, propagated clones from line 338 were planted in a field site along with WT419 plants. WT419 female and male flower buds were collected for RNA extraction and RT-PCR. Expression of *MeSWEET10a* was confirmed for both flower types (**Figure 3.4A**). Eleven months after planting, line 338 plants formed the first inflorescences of male and female flower buds (**Figure 3.4B**). WT419 and line 338 female and male flowers were collected, the petal-like bracts were dissected, and flowers were imaged in the

field (Perera et al., 2012). Macro images of all flowers showed no visible defects in line 338 compared to WT419 (**Figure 3.4C**). Field crosses of WT419 mother plants by line 338 father plants resulted in twelve heterozygous F1 seeds. The seeds were imaged and seed length, width, and weight were measured (**Figure 3.4D and Supplemental Figure 3.7**). WT419x338 F1 seed length ranged from 0.9-1.1 cm while WT419 open pollinated seed length ranged from 0.9-1 cm. There was no significant difference observed between WT419x338 F1 and WT419 open pollinated seed width. One common way of measuring seed viability is through float tests, seeds that sink are expected to regenerate while seeds that float commonly do not (Pegman et al., 2017). Thus, seed float tests were performed. Three out of twelve WT419x338 seeds sank indicating viable seeds for germination. As expected, these three seeds successfully germinated and so did two additional seeds that floated. Representative images of F1 plants are provided (**Figure 3.4E**). The germination rate from all WT419x338 F1 seeds was 41.7 percent. The germination rate for open-pollinated WT419 seeds was 66.7 percent as four out of six seeds germinated (**Supplemental Table 3.3**).

Additionally, whole genome sequencing for lines 338, 269, 27, 30, 41, and 54 was completed and the transgene insertion number and location for each line were identified (**Supplemental Figure 3.8**). This data will be used in the future to strategize for the development of transgene-free *MeSWEET10a* edited plants.

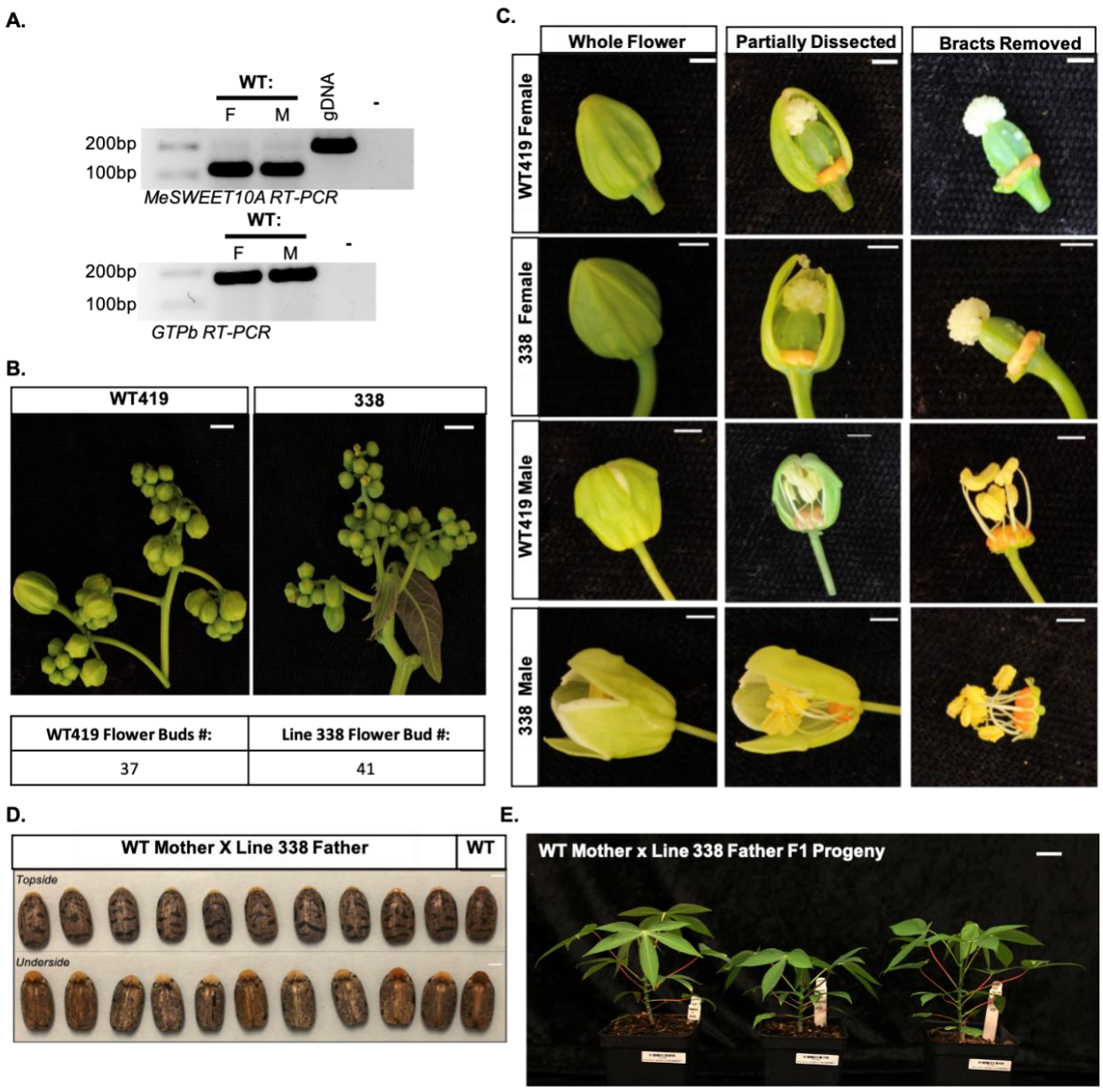


Figure 3.4: *MeSWEET10A* Mutant has Wildtype-like Reproductive Morphology and Progeny

A) RT-PCR of WT419 female (F) and male flowers (M) collected from field grown plants. Top gel shows results of RT-PCR with primers amplifying *MeSWEET10a* with an expected product

size of 123 bp. The bottom gel shows results of RT-PCR with primers amplifying the housekeeping gene GTPB as a control for sample loading with an expected product size of 184 bp. gDNA is from WT419 leaf tissue included as a positive control and '-' denotes a negative water control. **B)** Representative images of WT419 (left) and line 338 (right) inflorescence structures detached from individual field-grown plants for imaging (top). The number of flower buds present on each inflorescence is presented in table format (bottom). Scale bar = 5cm. **C)**. Representative female and male flowers collected from WT419 and line 338 flowering field-grown plants. Images of the same flower were taken as whole flowers (left) partially dissected with one or two petal-like bracts removed (middle, and dissected with all petal-like bracts removed (right). Scale bar = 0.25cm. **D)** Image of ten seeds derived from WT Mother x line 338 Father crosses along with an open-pollinated WT419 seed (top). Scale bar = 0.25cm. **E)** Image of three WT419 Mother x line 338 Father F1 individuals that germinated post-planting in soil.

3.5 Discussion

This study explored how editing the TAL20 targeted S gene, *MeSWEET10a*, could reduce cassava susceptibility to Xpm-induced bacterial blight. Using a dual gRNA CRISPR/Cas9 strategy, we generated several *MeSWEET10a* mutant lines with edits that impacted the TAL20 EBE binding site and/or the *MeSWEET10a* coding sequence. Mutants that lacked TAL20-mediated induction of *MeSWEET10a* were phenotyped for CBB disease severity through bacterial growth assays and water-soaked lesion analysis. While a consistent difference in bacterial titer was not observed, we found that *MeSWEET10a* mutant lines had significantly reduced water-soaked lesions post Xpm WT infection compared to infected wildtype cassava. We also inspected the impact of editing *MeSWEET10a* on cassava flowers as the flowers are known to have endogenous expression of the gene (Veley et al., 2021). Using macro imaging and observations of flowers from the *MeSWEET10a* mutant line 338, we found no obvious defects compared to wildtype flowers. Furthermore, data from field crosses of WT419 mother plants by line 338 father plants suggested

that *MeSWEET10a* mutant plants are fertile and able to produce viable seeds. Overall, these results suggest that editing the *MeSWEET10a* S gene and/or TAL20 EBE site is a viable strategy to improve plant disease resistance and reduce susceptibility to CBB.

The exact mechanics of how *MeSWEET10*a is used by Xpm to promote CBB and pathogen virulence remains unknown. *MeSWEET10*a is a clade III SWEET gene that exports sucrose and glucose from the plant cell into the apoplast where Xpm proliferates. One hypothesis is that Xpm uses these sugars as a carbon source. However, if the *MeSWEET10a* exported sugars were a direct carbon source for Xpm, we would expect that loss of TAL20 would significantly impact bacterial growth. In our bacterial growth assays, we found there was not a consistently significant difference between Xpm WT and Xpm△TAL20 colony-forming units (CFUs). Yet, whenever a bacterial growth difference was observed, Xpm△TAL20 titer trended downward compared to Xpm WT. This suggests limitations in the sensitivity of the bacterial growth assays and that at the timepoints and conditions used in our study, a robust difference between Xpm WT and Xpm△TAL20 was not captured.

An additional hypothesis is that *MeSWEET10a* exported sugars may serve as an osmolyte for Xpm. As sucrose and glucose are exported out of the plant cell, there is also osmotic movement of water. *MeSWEET10*a mutants have consistently reduced water-soaked lesions after infection with Xpm WT compared to infected wildtype cassava. Water-soaked lesions are dark angular spots that occur during pathogenesis as water is moved from the plant cell into the apoplast (Schwartz et al., 2017). Many plant pathogens induce water-soaked leaf lesions during early stages of plant infection (Aung et al., 2018). Other studies have suggested that the role of water-soaking is to

create an aqueous environment to aid in bacterial colonization from the plant surface into the apoplast or to help with bacterial spread once *in planta* (Xin et al., 2016). Perhaps the efflux of sugar and water into the apoplast increases bacterial entry into the plant which would not be captured through syringe infiltration-based infection assays. Additionally, Xpm eventually spreads throughout the plant vasculature after initial colonization at the surface of mesophyll cells. It is possible that induction of *MeSWEET10a* by TAL20 may play a role in bacterial spread. Another study in the *Xanthomonas gardneri*-pepper pathosystem reported that reduced water-soaked lesion symptoms did not correlate to a decrease in bacterial growth (Schornack et al., 2008). In the future, additional work is required to tease apart the role of *MeSWEET10a* exported sugars and water-soaking in Xpm pathogenesis.

In cassava, the native function of *MeSWEET10a* remains unknown. However, *MeSWEET10a* is expressed in the cassava flower. SWEET genes have been implicated in various roles in plants such as nectar secretion, pollen development, seed filling, and phloem loading (Feng and Frommer, 2015). The work done with *MeSWEET10a* mutant flowers in this study did not uncover obvious differences between the mutant and wildtype cassava flowers and serves as initial proof of *MeSWEET10a* mutant flower viability. However, further research is needed to determine if editing *MeSWEET10a* has detrimental impacts on flower development and reproductive function. For example, a full field trial with more crosses of the *MeSWEET10a* mutant lines could further validate the ability of mutant plants to produce viable fruit and seed. Additionally, with a larger field trial more seed could be collected to further compare the germination rate of line *MeSWEET10a* mutant plant progeny to WT419 open pollinated seed. Microscopy comparing the

structure of *MeSWEET10a* mutant and wildtype cassava flowers could also determine if there are differences in flower development, not visible to the naked eye.

The ideal *MeSWEET10a* mutants would contain edits at the TAL20 EBE site while maintaining an intact TATA box and gene coding sequence for native plant function. In this study, one line with EBE only edits and an intact TATA box was recovered, however this line died during the tissue culture process. It is possible that this outcome is selected against for an as of yet unknown reason. More likely, if many additional lines were recovered, we would eventually achieve this outcome. One factor complicating the ability to generate EBE specific edits to the *MeSWEET10a* TAL20 binding site is overlap between the TATA box and EBE. In cassava, other TAL effectors localize to EBE sites that include TATA box motifs (Cohn et al., 2016). Other work shows that EBE overlap within or localization near the host TATA box is common in TAL effector S gene target sites [(Grau et al., 2013), (Pereira et al., 2014) and (Pérez-Quintero et al., 2015)]. Alternative gene editing strategies such as base editing or the use of a single CRISPR/Cas9 gRNA for editing at the TAL20 EBE downstream of the TATA box may increase the chances of recovering *MeSWEET10a* mutants with EBE-only edits (Azameti and Dauda, 2021). It is also possible that other TAL20-carrying Xpm strains may have effectors with redundant functions and therefore may not exhibit reduced virulence after infection in *MeSWEET10a* mutants. Future work screening *MeSWEET10a* mutant susceptibility against other Xpm isolates would help determine the overall viability of a disease resistance strategy based on editing the *MeSWEET10a* S gene alone. Likewise, developing *MeSWEET10a* mutants in additional cassava cultivars and examining CBB susceptibility would be useful. It is important to note that editing one S gene may not be

enough to significantly reduce cassava susceptibility to CBB in a field setting. Thus, investigating the role of additional S genes on Xpm virulence and developing cassava mutants with stacked edits at different S gene targets may be required to develop plants with sustained resistance to CBB.

3.6 Materials and Methods

3.6.1 Construct design and cloning:

The *MeSWEET10a* (Manes.06G123400) FASTA sequence file was downloaded from Phytozome (Manihot esculenta genome v6.1) and uploaded to the software Geneious. Notable promoter regulatory elements were annotated as previously reported by Cohn and Bart et. al, 2014 including the effector binding element (EBE) site where TAL20 binds. The reported EBE sequence was confirmed using the TALEnt target finder tool. The Geneious "find CRISPR sites" function was used to find all potential targets and Cas9 (S. pyogenes)-specific PAM sites (sequence: 5′-NGG-3′). Candidate guide RNA (gRNAs) target sequences were selected based on those whose targets were near the TAL20 EBE and the translational start site of *MeSWEET10a* or within the 5'UTR. Candidate gRNAs were further analyzed by comparing the candidate gRNAs against the cassava genome to identify potential off-targets using NCBI-BLAST. Candidate gRNAs without strong potential off-target hits were moved forward for construct design. All constructs were assembled using a multiple gRNA spacer Csy4 array as previously described (Cermak et al 2017). Three constructs were used for this study. All constructs were designed to carry two gRNAs, were cloned in the pTRANS_220D backbone, and have a kanamycin resistance cassette. Construct 108 carries gRNA1 (GAGAAGCGTTTATATAGGGG) which targets TAL20 EBE site and gRNA2

(GAAGTCCAATGACAAGTGCA) targeting the *MeSWEET10a* translation start site (ATG). Construct 108 also carries an intended EBE repair template (as an attempt to replace the sequence) containing homology arms that flank the EBE (1,079bp 5' homology arm and 727 bp 3' homology arm). Construct 249 was designed to carry the TAL20 EBE site target gRNA1 (GAGAAGCGTTTATATAGGGG) and gRNA3 (ACTCTCTTTCCCTTGTGCAG) which targets the 5'UTR with no repair template. Construct 250 was designed to carry gRNA4 (AAAATATGTCAATGTAACAG) and gRNA5 (TATGTTGTGCAATGATGGAT) which target the 5'UTR and EBE with no repair template. Construct assembly was confirmed through colony PCR, Sanger sequencing, and by Illumina sequencing. Constructs were transformed into LBA4404 Agrobacterium cells for cassava transformations. All construct sequences, maps, and Illumina reads are available in supplementary data.

3.6.2 Plant materials and growing conditions:

Transgenic cassava lines expressing the CRISPR/Cas9 machinery and gRNAs were generated in the cassava cultivar TME419 through *Agrobacterium*-mediated transformation as described (Chauhun at al., 2015). Transgenic FEC cells were selected for resistance using 100mM paramomycin (275uL/L) on spread plates. 100mM paramomycin (450uL/L) was used to further select for resistant transgenic cells at stage 1, 2, and 3 plates. Transgenic FEC cells and eventual transgenic plantlets were maintained in tissue culture in conditions set to 28°C +/- 1°C, 75 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light; 16 hrs light / 8 hrs dark. Plantlets were transferred to soil on a misting bench and covered with domes to maintain high humidity. After establishment to soil, plants were moved from the misting bench and acclimated to greenhouse conditions set to 28°C; 50% humidity; 16 hrs light /

8 hrs dark and 1000 W light fixtures that supplemented natural light levels below 400 W / m^2. Following bacterial infection assays, plants were kept in a post-treatment growth chamber with conditions set to 27°C; 50% humidity and 12 hrs light / 12 hrs dark. F1 seeds generated from WT419 x 338 crosses were planted in soil and kept in a plant growth chamber set to 37°C, 60% humidity, 12 hrs light / 12 hrs dark at 400 $\mu mol \cdot m^{-2} \cdot s^{-1}$. Once seedlings germinated, they were transferred to larger pots and moved to greenhouse conditions listed above.

3.6.3 DNA extraction and transgenic line genotyping:

As an initial pass: transgenic lines recovered from the first transformation with construct 108, a PCR followed by restriction digest and gel electrophoresis strategy was used. Mutant lines were prioritized by screening for those with digest patterns unlike wildtype. Mutants with varying HaeIII digest patterns were suspected to have deletions, which were the primary mutant type of interest at the time, and were moved forward for Sanger sequencing. In subsequent rounds of transformation, mutants with both point mutations and insertions/deletions became of interest.

For later transformations: leaf lobe samples from transgenic lines were collected from 2-3 individual plantlets and pooled into 2mL Eppendorf Safelock tubes with three disposable 3 mm Propper solid glass beads. The sample tubes were flash frozen in liquid nitrogen and ground to a fine powder using a QIAGEN TissueLyser II machine at 30hz for 3 minutes until the sample was fully homogenized. Genomic DNA was extracted using the Sigma GenElute Plant Genomic DNA Miniprep Kit. The *MeSWEET10A* region of interest was amplified using “outer” primers designed to avoid amplification of the EBE repair template present in the construct 108 transgene and a 2.1kb product was generated for each line. All primers used in this study are provided in

Supplemental Table 4. The PCR product was purified using the Qiagen QIAquick PCR Purification kit. The samples were sent for Sanger sequencing using secondary "inner" primers designed to start amplification closer to the gRNA target sites. Transgenic line trace files were compared to wildtype TME419 trace files and edits within and across each gRNA were identified using the Geneious bioinformatics tool. Outer and inner primer sequences used for genotyping are available in the primer list table. Clone-seq was performed on select lines to determine if edits were homozygous for each allele.

3.6.4 Identification of transgene location(s)

For each construct a custom reference genome was created which contained the haplotype-resolved genome assembly for cassava variety TME204 (Qi et al., 2022) along with the vector sequence from the T-DNA Left Border through the T-DNA Right Border of the appropriate construct. The program bwa mem (version 0.7.12-r1039) was used to align the whole genome sequencing data to the custom reference genome (Li H., 2013). Reads where one pair aligned to the T-DNA insertion sequence and the other aligned to the cassava genome were isolated using samtools (version 1.11) (Danecek P, et al., 2021) and used as input for de novo assembly by Trinity (version v2.1.1) (Grabherr MG, et al., 2011). Resultant contigs were then used in a blastn (version 2.12.0+) query against a BLAST database of the custom reference genome initially used for bwa alignments (Sayers et al., 2022). Contigs where a portion matched the cassava genome and another portion matched the T-DNA insertion sequence identified the coordinates of the 5' and 3' ends of an insertion point within the genome. Integrative Genomics Viewer (IGV; version 2.12.3) was

used for manual inspection and visualization of the aligned WGS data to the custom T-DNA insertion plus genome (Thorvaldsdóttir et al., 2013).

3.6.5 Bacterial inoculations:

Xanthomonas strains were struck from glycerol stocks onto NYG agar plates containing appropriate antibiotics. The strains used were Xpm668 (rifampicin 50 µg/ml), Xpm668ΔTAL20 (suicide vector knockout, tetracycline 5 µg/ml, rifampicin 50 µg/ml), Xe85-10 (rifampicin 50 µg/ml) and Xe85-10+TAL20$_{Xpm668}$ (rifampicin 50 µg/ml, kanamycin 50 µg/ml), (Cohn and Bart et al., 2014). Xanthomonas strains were grown in a 30°C incubator for 2-3 days. Inoculum for each strain was made by transferring bacteria from plates into 10mM $MgCl_2$ using inoculation loops and brought up to a concentration of OD600 = 0.01 for bacterial growth and water-soaked lesion assays and OD600 = 1 for RT-PCR. Leaves on cassava plants were inoculated using a 1.0 mL needleless syringe. For each replicate assay, two cassava plants per background (WT or transgenic) were used for inoculations and four leaves were inoculated on each plant. One bacterial strain suspended 10mM $MgCl_2$ in was inoculated per leaf lobe with three injection sites and mock inoculations of 10mM $MgCl_2$ alone were included. In total there were nine infiltrated sites per leaf.

3.6.6 RNA extraction and RT-PCR analysis:

For lines with edits of interest, RNA extraction and RT-PCR was performed at the plantlet stage in tissue culture and on soil established plants in the greenhouse. At the plantlet stage, 9 leaves were detached from every transgenic line and 3 leaves each were syringe infiltrated with either mock (10 mM $MgCl_2$ alone) or *Xanthomonas* (Xpm668 +/- TAL20) on sterile petri dishes.

For each line, a set of 3 infiltrated leaves per treatment were kept on MS2 plates in a post treatment room light shelf. At 48 hours post infection, samples were collected with a 7mm cork borer, and the set of 3 infiltrated leaves per treatment were pooled into Eppendorf safelock tubes with 3mm glass beads. For greenhouse plants, one leaf was selected (3 biological replicates per plant background) and syringe infected with 3 infiltrated sites per treatment (either mock (10 mM $MgCl_2$ alone) or *Xanthomonas* (Xpm668+/- TAL20) or *Xanthomonas euvesicatoria* (Xe85-10+/- TAL20) strains on separate leaf lobes. Samples were flash-frozen in liquid nitrogen and ground using TissueLyser settings described above. Total RNA was extracted from each sample using the Sigma Spectrum Plant Total RNA Kit. 1 µg of RNA was DNase treated using Promega RQ1 DNase enzyme and reverse transcribed into cDNA using Thermo Fisher Scientific SuperScript III Reverse Transcriptase. RT-PCR was performed on each sample using primers specific to *MeSWEET10a* and to cassava GTPb (Manes.09G086600) as a constitutively expressed control. All primers used in this study are provided in **Supplemental Table 3.4**. RT-PCR results were analyzed to identify transgenic lines in which ectopic expression of *MeSWEET10a* was not induced by *Xanthomonas* (+TAL20) infection as is normally seen in wildtype cassava infected with *Xanthomonas* (+TAL20).

3.6.7 Bacterial growth assay:

Cassava leaves were infiltrated with either 10 mM $MgCl_2$ (mock control) or *Xanthomonas* (Xpm668 strains +/- TAL20) suspended in 10 mM $MgCl_2$ as described above. Leaf punch samples were taken at the site of infiltration using a 7 mm core borer (size 4) at 0- , 2-, 4- and 6- days post-inoculation. For day-0 samples, infiltrated spots were allowed to dry down prior to processing.

Individual leaf punches were transferred to 2 mL Eppendorf Safelock tubes with 200 uL of 10 mM $MgCl_2$ and three disposable 3 mm Propper solid glass beads. Samples were ground with a Qiagen Tissuelyzer at 28 hZ for 3 minutes. 200 uL of the ground sample was transferred to the first column of a labeled 96-well plate. Serial dilutions were performed by transferring 20 uL of the non-dilute sample (10^1) to the next well containing 180 ul of 10 mM $MgCl_2$. Samples were serially diluted to 10^4 for day 0, 10^6 for day 2 and 4, and for 10^8 days 6. 10 ul of each serial dilution was pipetted and spread onto labeled quadrants of an NYG plate with cycloheximide and the appropriate antibiotics for the infiltrated bacterial strain. Plates were incubated at 30°C for 2-3 days, and the number of colonies were counted. Colony Forming Units (CFU) reported in this manuscript were transformed by sample area (CFU/leaf disc area where disc area = 0.38cm).

3.6.7 Water-soaked lesion imaging and quantification:

Cassava leaves were detached and imaged at 0-, 6-, and 9-days post-inoculation (DPI). One leaf for every plant was collected for a total of two leaves per plant background at each time point. Line 338 and 269 leaves were imaged from above using a Raspberry Pi Sony IMX219 camera in an enclosed box with an overhead light. To increase image resolution, all subsequent infected plant leaves were imaged from above using a Canon EOS Rebel T5i camera with a 15-85mm lens in an enclosed box with an overhead light. Images were processed and analyzed for water-soaked lesion area and gray-scale color using a previously described custom machine learning image analysis tool developed for CBB disease quantification.

3.6.8 Flower inflorescences and flower bud imaging and dissection:

Flower inflorescences and individual buds were detached from cassava plants (WT419 or line 338) growing in a Hawaii field site. All flower inflorescences and individual buds were imaged in the field from above using a Canon EOS Rebel T5i camera with a 15-85mm lens in a portable, partially enclosed pop-up light box with built in LED lights controlled by a USB power pack. Images were post processed using photoshop for color correction and a scale bar was added using ImageJ version FIJI.

3.7 References

1. Abrusci P, McDowell MA, Lea SM, Johnson S (2014) Building a secreting nanomachine: a structural overview of the T3SS. Current Opinion in Structural Biology 25: 111–117
2. An S-Q, Potnis N, Dow M, Vorhölter F-J, He Y-Q, Becker A, Teper D, Li Y, Wang N, Bleris L, et al Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen Xanthomonas. FEMS Microbiol Rev. doi: 10.1093/femsre/fuz024
3. 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 22: 3864–3876
4. Aung K, Jiang Y, He SY (2018) The role of water in plant–microbe interactions. The Plant Journal 93: 771–780
5. Azameti MK, Dauda WP (2021) Base Editing in Plants: Applications, Challenges, and Future Prospects. Frontiers in Plant Science 12:
6. 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. Proc Natl Acad Sci USA 109: E1972-1979
7. Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48: 419–436
8. Britt AB, May GD (2003) Re-engineering plant gene targeting. Trends Plant Sci 8: 90–95
9. 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. PLoS Pathog 10: e1003972

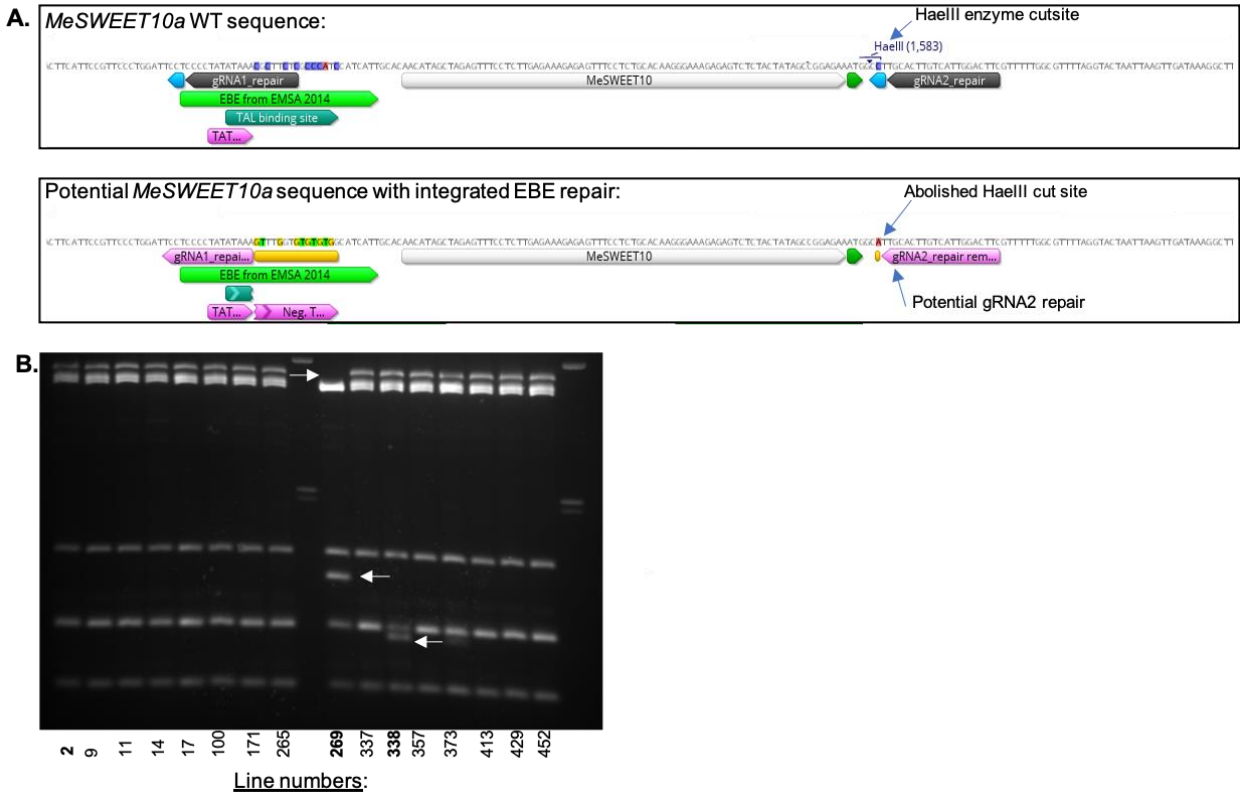
10. Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, et al (2017) A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. Plant Cell 29: 1196–1217
11. Chauhan RD, Beyene G, Kalyaeva M, Fauquet CM, Taylor N (2015) Improvements in Agrobacterium-mediated transformation of cassava (Manihot esculenta Crantz) for large-scale production of transgenic plants. Plant Cell, Tissue and Organ Culture (PCTOC) 121: 591–603
12. Chen L-Q (2014) SWEET sugar transporters for phloem transport and pathogen nutrition. New Phytologist 201: 1150–1155
13. 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
14. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T, Staskawicz BJ (2014) Xanthomonas axonopodis virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. Mol Plant Microbe Interact 27: 1186–1198
15. Cohn M, Morbitzer R, Lahaye T, Staskawicz BJ (2016) Comparison of gene activation by two TAL effectors from Xanthomonas axonopodis pv. manihotis reveals candidate host susceptibility genes in cassava. Molecular Plant Pathology 17: 875–889
16. Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, Cottyn B (2016) Genetic characterization of strains named as Xanthomonas axonopodis pv. dieffenbachiae leads to a taxonomic revision of the X. axonopodis species complex. Plant Pathology 65: 792–806
17. Cox KL, Meng F, Wilkins KE, Li F, Wang P, Booher NJ, Carpenter SCD, Chen L-Q, Zheng H, Gao X, et al (2017) TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. Nature Communications 8: 1–14
18. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al (2021) Twelve years of SAMtools and BCFtools. GigaScience 10: giab008
19. Eckardt NA (2002) Plant Disease Susceptibility Genes? The Plant Cell 14: 1983–1986
20. Elliott K, Berry JC, Kim H, Bart RS (2022) A comparison of ImageJ and machine learning based image analysis methods to measure cassava bacterial blight disease severity. Plant Methods 18: 86
21. EL-Sharkawy MA (2003) Cassava biology and physiology. Plant Mol Biol 53: 621–641
22. Feng L, Frommer WB (2015) Structure and function of SemiSWEET and SWEET sugar transporters. Trends Biochem Sci 40: 480–486
23. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644–652

24. Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J (2013) Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. PLOS Computational Biology 9: e1002962
25. Gupta PK, Balyan HS, Gautam T (2021) SWEET genes and TAL effectors for disease resistance in plants: Present status and future prospects. Mol Plant Pathol 22: 1014–1026
26. Hillocks RJ, Thresh JM, Bellotti A (2002) Cassava: Biology, Production and Utilization. CABI
27. Hogenhout SA, Van der Hoorn RAL, Terauchi R, Kamoun S (2009) Emerging Concepts in Effector Biology of Plant-Associated Organisms. MPMI 22: 115–122
28. Hu Y, Zhang J, Jia H, Sosso D, Li T, Frommer WB, Yang B, White FF, Wang N, Jones JB (2014) Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. Proc Natl Acad Sci USA 111: E521-529
29. Hu Z, Tang Z, Yang J, Bao S, Zhang Y, Ma L, Zheng Q, Yang F, Zhang D, Sun S, et al (2023) Knockout of OsSWEET15 Impairs Rice Embryo Formation and Seed-Setting. Plant and Cell Physiology 64: 258–268
30. Huang X, Wang Y, Wang N (2022) Highly Efficient Generation of Canker-Resistant Sweet Orange Enabled by an Improved CRISPR/Cas9 System. Frontiers in Plant Science 12:
31. Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, et al (2016) Using Ecology, Physiology, and Genomics to Understand Host Specificity in Xanthomonas. Annual Review of Phytopathology 54: 163–187
32. Kandel SL, Joubert PM, Doty SL (2017) Bacterial Endophyte Colonization and Distribution within Plants. Microorganisms 5: 77
33. Leyns F, Cleene MD, Swings J-G, De Ley J (1984) The Host Range of the Genus Xanthomonas. Botanical Review 50: 308–356
34. Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. doi: 10.48550/arXiv.1303.3997
35. López CE, Bernal AJ (2012) Cassava bacterial blight: using genomics for the elucidation and management of an old problem. Tropical Plant Biology 5: 117–126
36. Lozano JC (1986) Cassava Bacterial Blight: A Manageable Disease. Plant Dis 70: 1089
37. Lozano JC, Byrne D, Bellotti A (1980) Cassava/Ecosystem Relationships and their Influence on Breeding Strategy. Tropical Pest Management 26: 180–187
38. Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, Saux MF-L, Manceau C, Poussier S, Pruvost O, et al (2013) Evolutionary History of the Plant Pathogenic Bacterium Xanthomonas axonopodis. PLOS ONE 8: e58474
39. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326: 1501
40. Oliva R, Ji C, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, Eom J-S, Li C, Nguyen H, Liu B, et al (2019) Broad-spectrum resistance to bacterial blight in rice using genome editing. Nat Biotechnol 37: 1344–1350

41. Pegman APM, Perry GLW, Clout MN (2017) Size-based fruit selection by a keystone avian frugivore and effects on seed viability. New Zealand Journal of Botany 55: 118–133
42. 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
43. Perera PIP, Quintero M, Dedicova B, Kularatne JDJS, Ceballos H (2012) Comparative morphology, biology and histology of reproductive development in three lines of Manihot esculenta Crantz (Euphorbiaceae: Crotonoideae). AoB Plants 5: pls046
44. Pérez-Quintero A, Lamy L, Gordon J, Escalon A, Cunnac S, Szurek B, Gagnevin L (2015) QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. Frontiers in Plant Science 6:
45. Phillips AZ, Berry JC, Wilson MC, Vijayaraghavan A, Burke J, Bunn JI, Allen TW, Wheeler T, Bart RS (2017) Genomics-enabled analysis of the emergent disease cotton bacterial blight. PLOS Genetics 13: e1007003
46. Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. Journal of Experimental Botany 56: 1–14
47. Qi W, Lim Y-W, Patrignani A, Schläpfer P, Bratus-Neuenschwander A, Grüter S, Chanez C, Rodde N, Prat E, Vautrin S, et al (2022) The haplotype-resolved chromosome pairs of a heterozygous diploid African cassava cultivar reveal novel pan-genome and allele-specific transcriptome features. Gigascience 11: giac028
48. 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
49. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, et al (2022) Database resources of the national center for biotechnology information. Nucleic Acids Research 50: D20–D26
50. van Schie CCN, Takken FLW (2014) Susceptibility Genes 101: How to Be a Good Host. Annual Review of Phytopathology 52: 551–581
51. Schornack S, Minsavage GV, Stall RE, Jones JB, Lahaye T (2008) Characterization of AvrHah1, a novel AvrBs3-like effector from Xanthomonas gardneri with virulence and avirulence activity. New Phytologist 179: 546–556
52. Schornack S, Moscou MJ, Ward ER, Horvath DM (2013) Engineering Plant Disease Resistance Based on TAL Effectors. Annual Review of Phytopathology 51: 383–406
53. Schwartz AR, Morbitzer R, Lahaye T, Staskawicz BJ (2017) TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. Proceedings of the National Academy of Sciences 114: E897–E903
54. Taylor N, Chavarriaga P, Raemakers K, Siritunga D (2003) Development and application of transgenic technologies in cassava.

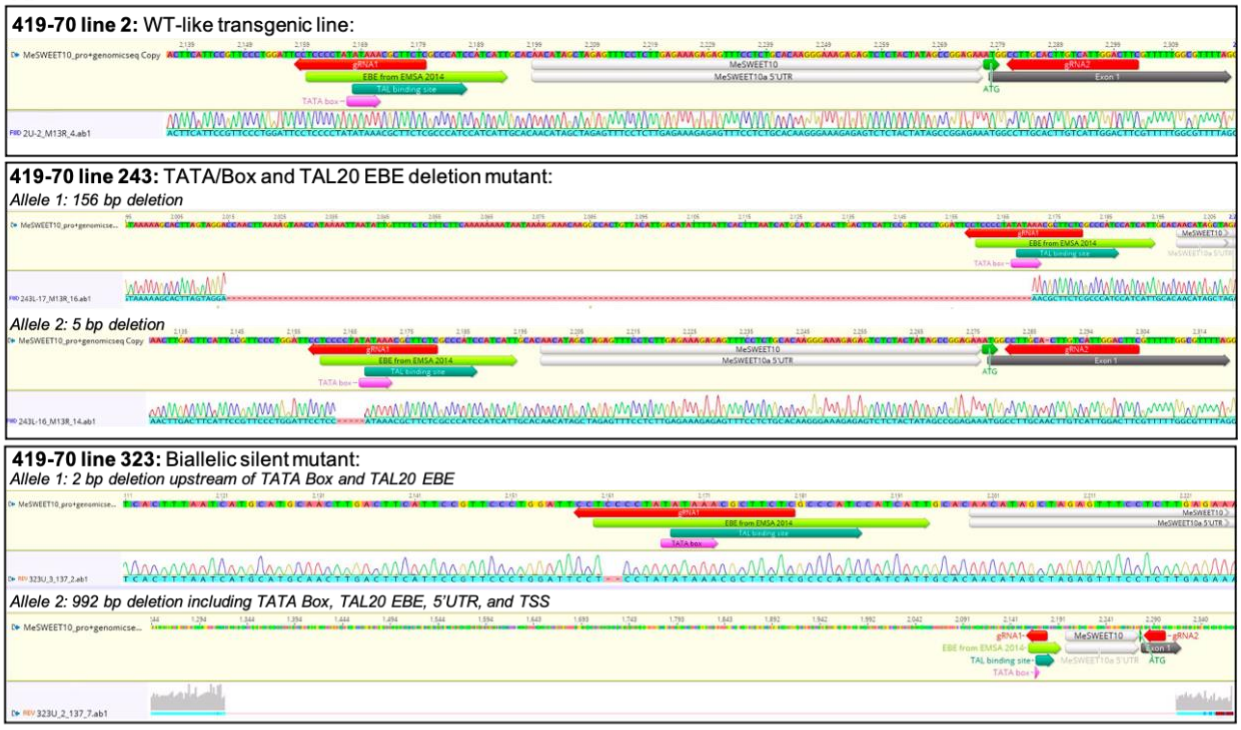
55. Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14: 178–192
56. Veley KM, Elliott K, Jensen G, Zhong Z, Feng S, Yoder M, Gilbert KB, Berry JC, Lin Z-JD, Ghoshal B, et al (2023) Improving cassava bacterial blight resistance by editing the epigenome. Nat Commun 14: 85
57. Veley KM, Okwuonu I, Jensen G, Yoder M, Taylor NJ, Meyers BC, Bart RS (2021) Gene tagging via CRISPR-mediated homology-directed repair in cassava. G3 (Bethesda) 11: jkab028
58. Xin X-F, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, Chang JH, Zipfel C, He SY (2016) Bacteria establish an aqueous living space in plants crucial for virulence. Nature 539: 524–529

3.8 Supplemental Information

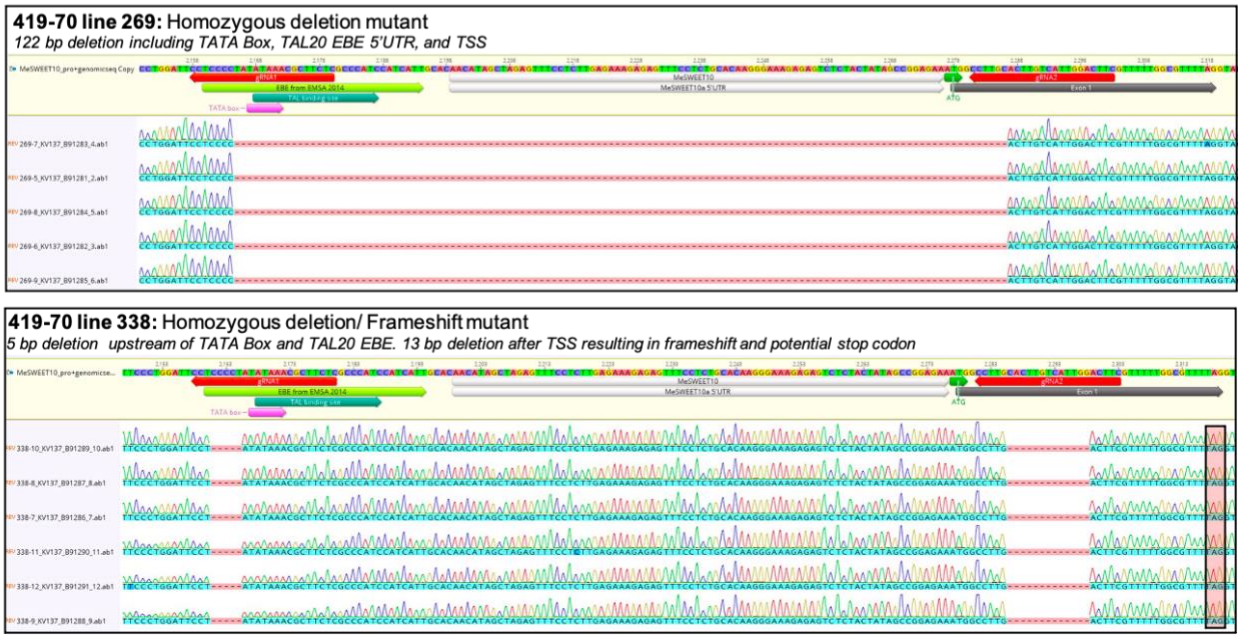


Supplemental Figure 3.1: Construct 108 Transgenic Lines Restriction Digest Screen
A) Overview of HaeIII restriction digest strategy for *MeSWEET10a* region of interest in WT (top) and potential mutant with the EBE template repair (bottom). WT-like sequences were expected to remain uncut. However, mutants with integration of the repair template were expected to have an abolished HaeIII site and remain uncut. Blue arrows point to the HaeIII cut site and potential gRNA2 repair site from the template integration by homology-directed repair. **B)** Example of an HaeIII digest performed on sixteen transgenic lines. White arrows point to digest patterns unlike the WT digest pattern. The expected digest pattern for WT is bands at 895, 384, 270, 271, 195, and 192 bp. Bolded line numbers (2, 269, and 338) were moved forward for Sanger sequencing.

A.


419-70 line 2: WT-like transgenic line:
419-70 line 243: TATA/Box and TAL20 EBE deletion mutant:
Allele 1: 156 bp deletion
Allele 2: 5 bp deletion
419-70 line 323: Biallelic silent mutant:
Allele 1: 2 bp deletion upstream of TATA Box and TAL20 EBE
Allele 2: 992 bp deletion including TATA Box, TAL20 EBE, 5'UTR, and TSS

B.


419-70 line 269: Homozygous deletion mutant
122 bp deletion including TATA Box, TAL20 EBE 5'UTR, and TSS
419-70 line 338: Homozygous deletion/ Frameshift mutant
5 bp deletion upstream of TATA Box and TAL20 EBE. 13 bp deletion after TSS resulting in frameshift and potential stop codon

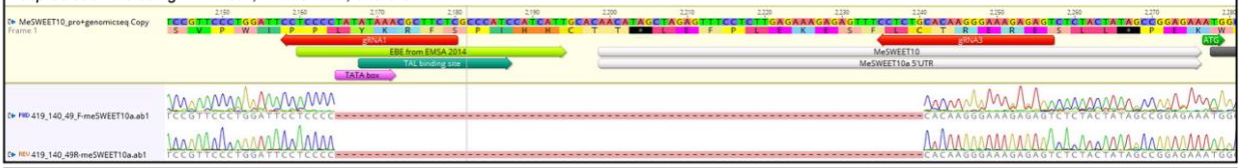
Supplemental Figure 3.2: Transgenic Line Sequencing from The First Construct 108 Transformation
A) Geneious screenshots showing Sanger-sequencing results for lines 2, 243, and 323. Mutation types and INDELs are described in text. **B)** Geneious screenshots of Sanger-sequencing results of *E.coli* clones containing *MeSWEET10a* gDNA from lines 269 and 338 plants.

A.


419-115 line 69B: Silent mutant
1 bp deletion upstream of TATA box and TAL20 EBE, potential 1bp deletion after TSS but not well resolved by sanger-seq alone
419-136 line 81: WT like transgenic or silent mutant
No edits detected
419-136 line 241*: TAL20 EBE mutant
16 bp deletion in TAL20 EBE with intact TATA box
419-140 line 11: TATA box mutant, predicted non functional
1 bp deletion in the TATA box
419-140 line 18: TATA box mutant, predicted non functional
4bp deletion in TATA box upstream of TAL20 EBE and 6bp deletion downstream of ATG
419-140 line 21: TATA box mutant, predicted non functional
10bp deletion in TATA box upstream of TAL20 EBE
419-140 line 35: TATA box, TAL20 EBE, and 5'UTR mutant, predicted non functional
76bp deletion including TATA box, TAL20 EBE, and 5'UTR
419-140 line 47: Silent mutant
2bp insertion upstream of TATA box and 1bp deletion in 5'UTR

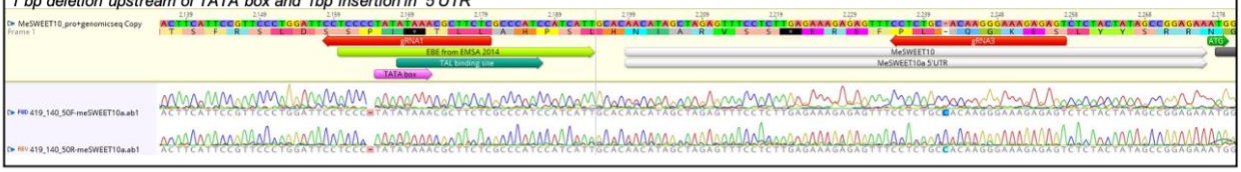
419-140 line 49: TATA box and TAL20 EBE mutant

75bp deletion including TATA box, TAL20 EBE, and 5'UTR



419-140 line 50: Silent mutant

1 bp deletion upstream of TATA box and 1bp insertion in 5'UTR

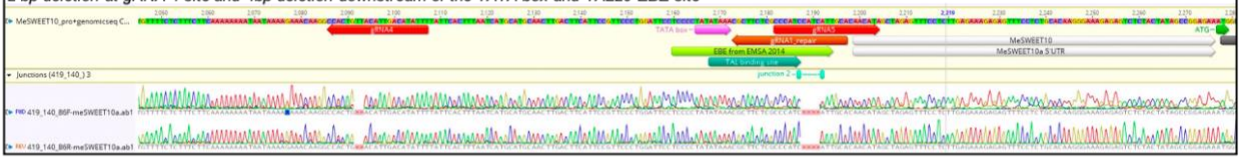


419-140 line 71: Silent mutant

1 bp deletion downstream of TATA box and possible deletions at gRNA 4 site

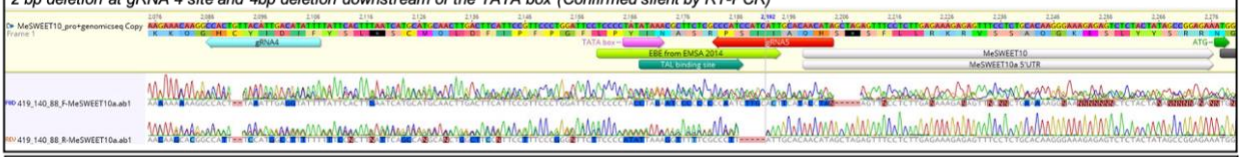
419-140 line 86: Silent mutant

2 bp deletion at gRNA 4 site and 4bp deletion downstream of the TATA box and TAL20 EBE site



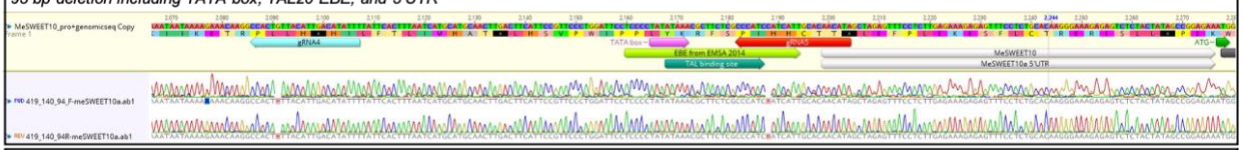
419-140 line 88: Silent mutant

2 bp deletion at gRNA 4 site and 4bp deletion downstream of the TATA box (Confirmed silent by RT-PCR)



419-140 line 95: TATA Box and TAL20 EBE Mutant

95 bp deletion including TATA box, TAL20 EBE, and 5'UTR

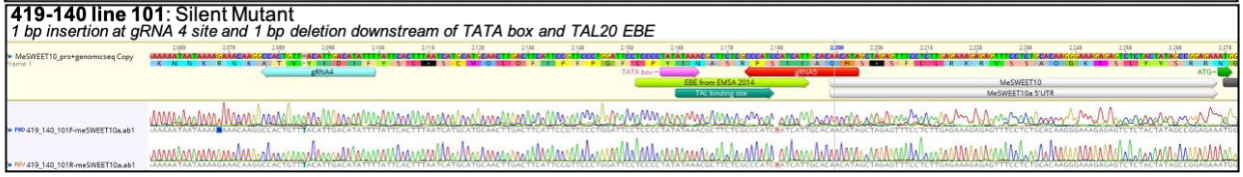


419-140 line 99: Silent Mutant

11 bp deletion at gRNA 4 site and 1 bp deletion downstream of TATA box and TAL20 EBE

419-140 line 101: Silent Mutant

1 bp insertion at gRNA 4 site and 1 bp deletion downstream of TATA box and TAL20 EBE



Supplemental Figure 3.3: Constructs 108, 249, and 250 Sanger-sequencing data from four replicate transformations

A) Geneious screenshots showing Sanger-sequencing results for all lines obtained by additional transformation with constructs 108, 249, and 250. Mutation types and INDELs are described in

text. '*' Denotes a line with EBE specific mutants that did not survive in tissue culture. **B)** Geneious screenshots of Sanger-sequencing results of *E. coli* clones containing *MeSWEET10a* gDNA from lines 27, 30, 41, 54, and 69A plants.

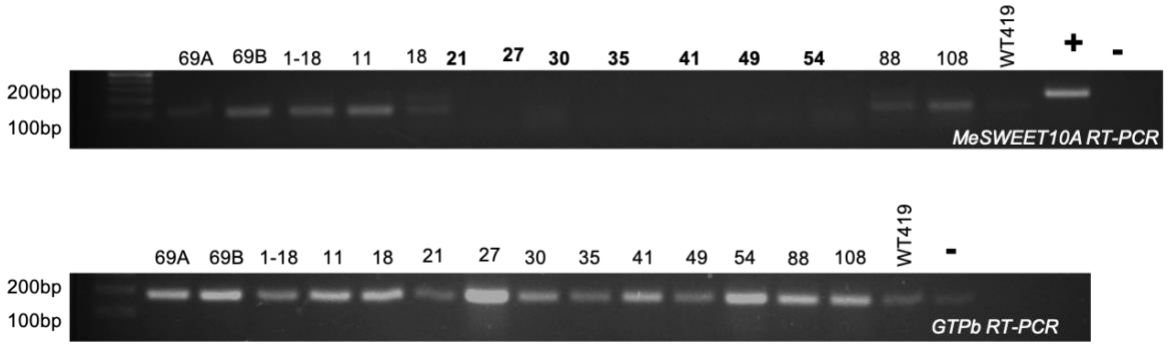
Construct:	Callus selected (Stage 2):	Mature lines recovered (MS2):	Lines characterized:	Lines with confirmed edits:
108 (First generation)	344	30	5	4
108	206	7	7	6
249	264	8	8	8
250	324	9	9	9

Supplemental Table 3.1: Transformation Results Overview
Table with a summary of transgenic line recovered from all rounds of transformations separated by construct type. Callus selected represents the number of lines that survived selection and made it stage 2 plates for the cassava transformation pipeline. Mature lines recovered represent transgenic lines that survived as plantlets up to the first MS2 plate stage of cassava transformation as described by previously (Chauhan et al, 2015). Select lines were characterized and the number of lines with edits are reported.

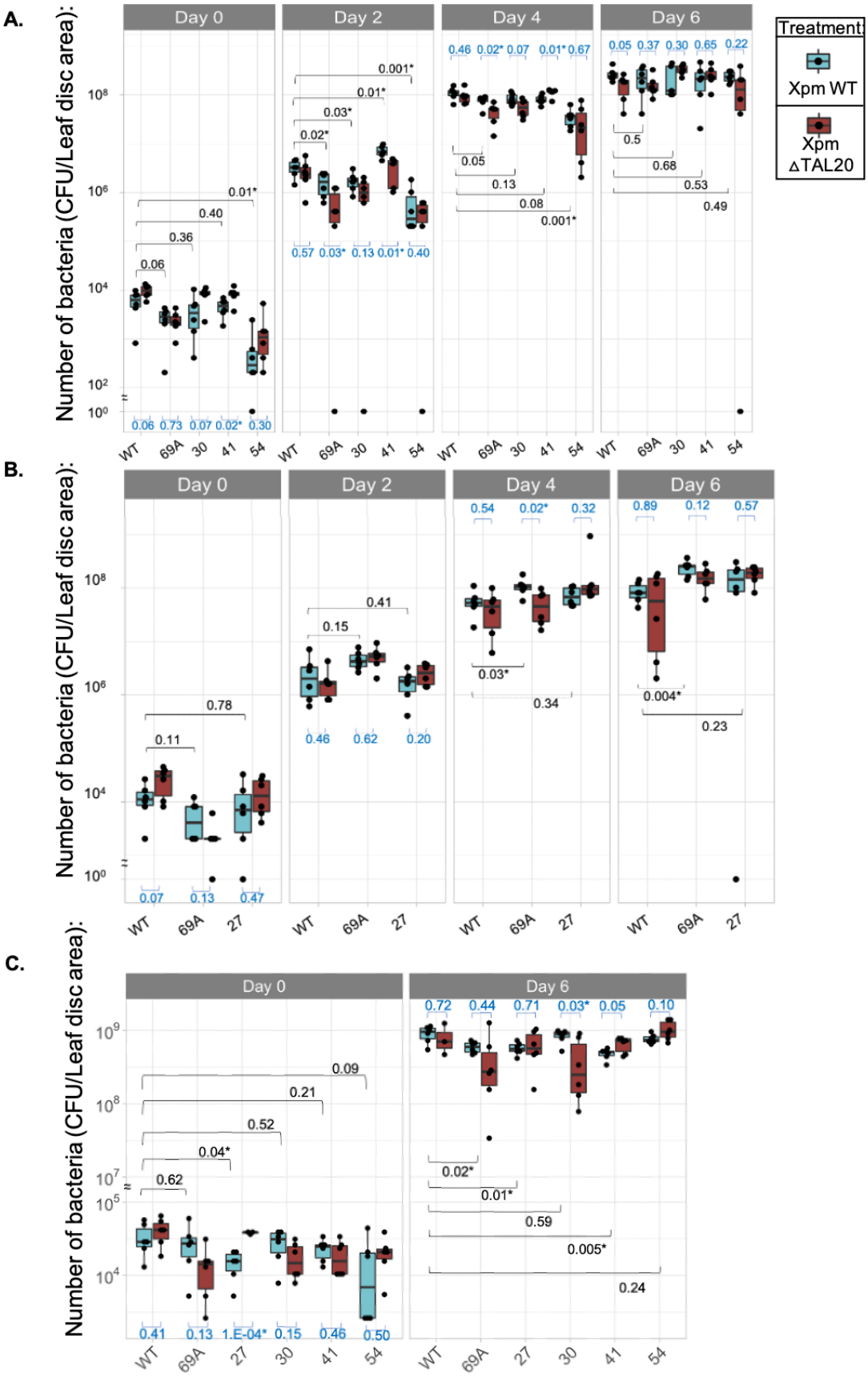
Plant background:	WT419:	Line #27:	Line #30:	Line #41:	Line #54:	Line #69A:
Average Plant height (cm)	44.87 ±1.44	47.15 ±3.18	54.25 ±3.44	46.89 ±3.42	41.70 ±1.32	46.11 ±7.31
T.test (WT vs Line)		0.35	0.03	0.42	0.05	0.85
Average number of nodes	34.33 ±0.58	32.33 ±3.21	34.00 ±2.31	30.33 ±2.08	35.67 ±0.58	31.00 ±1.41
T.test (WT vs Line)		0.39	0.34	0.07	0.05	0.16
Average Internode length (above the woody transition) (cm)	0.42 ±0.09	0.53 ±0.09	0.40 ±0.08	0.74 ±0.17	0.56 ±0.08	1.00 ±0.29
T.test (WT vs Line)		0.23	0.72	0.06	0.13	0.20
Average Internode length (below the woody transition) (cm)	0.26 ±0.09	0.45 ±0.17	0.24 ±0.00	0.56 ±0.14	0.45 ±0.12	0.75 ±0.06
T.test (WT vs Line)		0.18	0.67	0.05	0.11	0.01
Leaf lobe # average	6.78 ±0.67	6.11 ±0.93	6.00 ±0.71	6.89 ±0.33	7.00 ±0.00	5.33 ±0.52
T.test (WT vs Line)		0.08	0.07	0.69	0.42	0.01
Petiole length average (cm)	13.86 ±2.67	15.45 ±1.71	15.79 ±3.30	13.67 ±2.71	13.40 ±1.88	14.87 ±0.74
T.test (WT vs Line)		0.44	0.51	0.94	0.81	0.57
Central lobe length: 3 representatives per plant average (cm)	14.18 ±0.94	15.38 ±1.14	17.03 ±1.65	15.41 ±1.12	12.87 ±0.83	13.20 ±0.64
T.test (WT vs Line)		0.02	0.03	0.19	0.03	0.06
Central lobe width: 3 representatives per plant Average (cm)	3.62 ±0.24	4.04 ±0.25	4.10 ±0.39	4.42 ±0.38	3.31 ±0.23	3.57 ±0.28
T.test (WT vs Line)		0.05	0.03	0.04	0.03	0.75
Whole leaf width: 3 representatives per plant average (cm)	22.08 ±2.23	23.42 ±1.77	24.41 ±2.88	23.14 ±2.12	20.03 ±1.90	17.61 ±0.45
T.test (WT vs Line)		0.31	0.13	0.57	0.16	0.04

Supplemental Table 3.2: Plant Morphology Measurements

Measurements of cassava morphology traits including: plant height, node numbers, internode lengths above and below the woody transition, leaf lobe number, central lobe length and width, and whole length width for 3 representative individuals each of WT419 and lines 27, 30, 41, and 54 plants. Line 69A had 2 individual plants. The average of each trait measurement was taken along with standard deviation (measurement ± standard deviation). The calculated p-value from unpaired student T-tests with unequal variance is shown for each measurement comparing WT and each mutant line.



Supplemental Figure 3.4: Xpm Infected Detached Leaf RT-PCR
RT-PCR of wildtype (WT419) cassava and *MeSWEET10a* mutant lines detached leaves from plantlets infected with Xpm WT. The top gel shows results of RT-PCR with primers amplifying *MeSWEET10a* with an expected product size of 123 bp. The bottom gel shows results of RT-PCR with primers amplifying the housekeeping gene GTPB as a control for sample loading with an expected product size of 184 bp.'+' denotes a WT419 gDNA positive control. '-'denotes a negative water control. The bold text represents mutants that lack TAL20-mediated induction of *MeSWEET10a*.

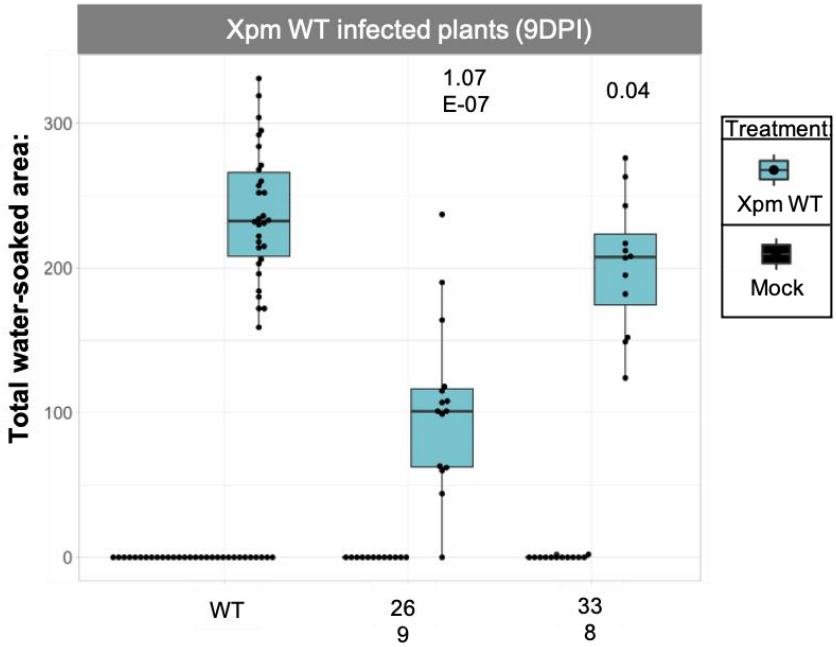

A.
B.
C.
Number of bacteria (CFU/Leaf disc area):
Day 0
Day 2
Day 4
Day 6
Treatment:
Xpm WT
Xpm ΔTAL20
WT
69A
30
41
54
27

Supplemental Figure 3.5: Additional Replicates of Bacterial Growth Assays
A) Replicate 1 bacterial growth assay for WT, line 69A, 30, 41, and 54 plants. Number of bacteria in cassava leaves measured at 0, 2, 4, and 6DPI post syringe infiltration with Xpm (blue) and Xpm△TAL20 (red) treatments. **B)** Replicate 1 bacterial growth assay for mutant line 27. Along with WT and 69A controls. Number of bacteria in cassava leaves measured at 0, 2, 4, and 6DPI post syringe infiltration with Xpm (blue) and Xpm△TAL20 (red) treatments. **C)** Replicate 2 bacterial growth assay for WT, line 69A, 27, 30, 41, and 54 plants. Number of bacteria in cassava leaves measured at 0 (left) and 6DPI (right) post syringe infiltration with Xpm (blue) and Xpm△TAL20 (red) treatments. For all box plots, Colony Forming Units (CFU/cm2, Y-axis) are plotted by plant genotype (X-axis) tested (wildtype or mutant). Black dots represent technical replicates from one independent bacterial growth experiment. Results of statistical analyses (Unpaired student's t-test with unequal variance) comparing the difference between Xpm WT growth across wild-type and mutant cassava genotypes infected with Xpm WT. P-values are shown above or below brackets indicating the comparison types for statistical analyses. Black dots represent individual water-soaked lesions from three independent water-soaking assay experiments combined. In all boxplots, the calculated *p*-values (Unpaired Student's T-test with unequal variance) are shown above or below each box plot. Black text represents Xam WT comparisons between WT and mutant infected plants. Blue text represents comparisons between Xam WT and Xpm△TAL20 within each genotype. Dots outside whiskers represent outliers based on default settings of the R package ggplot2. The horizontal line within the box represents the median sample value. The ends of the boxes represent the 3rd (Q3) and 1st (Q1) quartiles. The whiskers show values that are 1.5 times interquartile range (1.5xIQR) above and below Q1 and Q3.

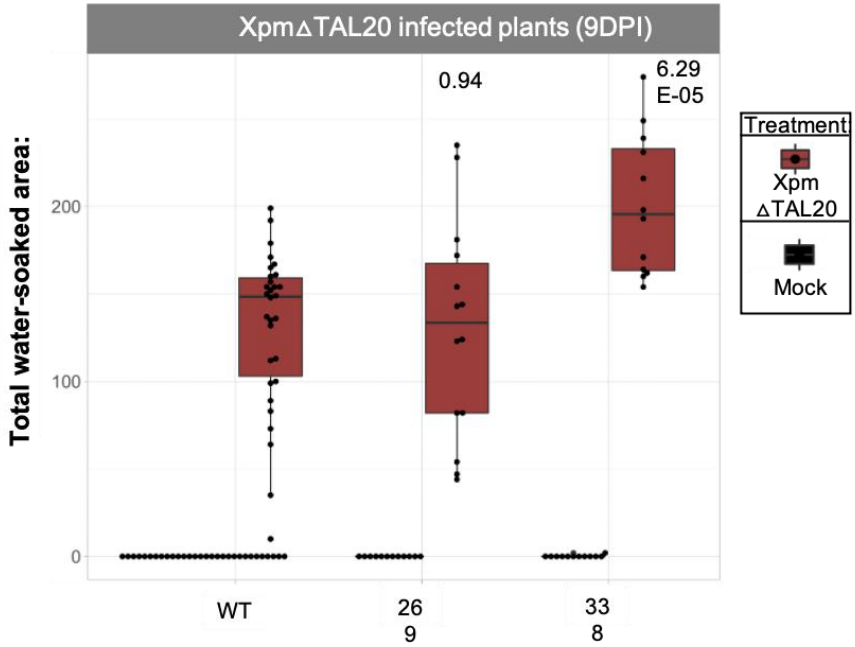
A.


WT419
338

B.

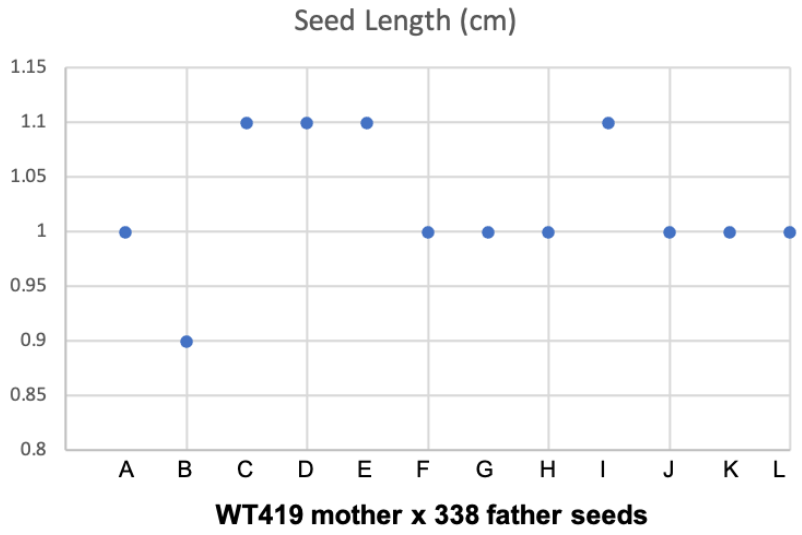

Xpm WT infected plants (9DPI)
Total water-soaked area:
1.07
E-07
0.04
300
200
100
0
WT
26
9
33
8
Treatment
Xpm WT
Mock

C.

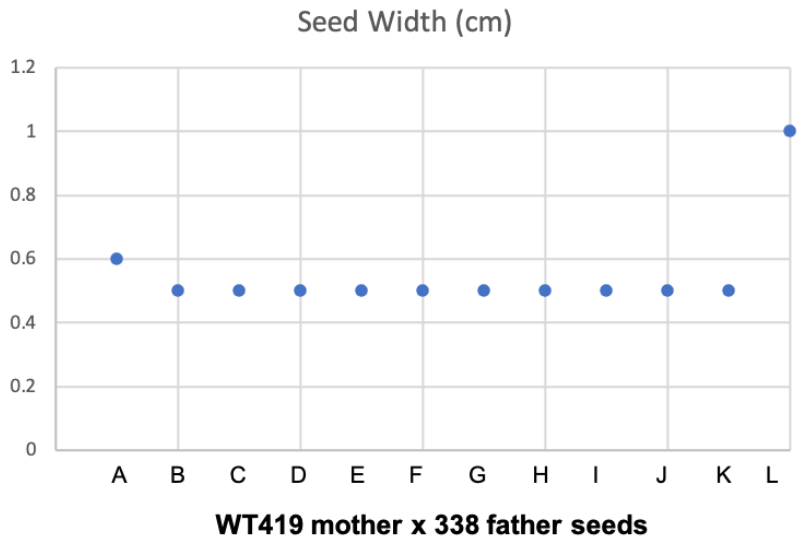

XpmΔTAL20 infected plants (9DPI)
Total water-soaked area:
0.94
6.29
E-05
200
100
0
WT
26
9
33
8
Treatment
Xpm
ΔTAL20
Mock

Supplemental Figure 3.6: Lines 269 and 338 Water-soaked Lesion Assay
A) Representative images of infected wildtype (left) and mutant line 338 (right) cassava leaves detached from the plant and imaged at 6DPI. X= Xpm WT, T=Xpm△TAL20, and M=Mock. Scale bar = 1cm. **B)** Total water-soaked area (pixels, y-axis) of mock and Xpm WT infected plants (genotypes, x-axis) at 9DPI. **C)** Total water-soaked area (pixels, y-axis) of mock and Xpm△TAL20 infected plants (genotypes, x-axis) at 9DPI. Black dots represent individual water-soaked lesions from three independent water-soaking assay experiments combined. Calculated *p*-values (Unpaired Student's T-test with unequal variance) comparing mutant line to wildtype water-soaked area shown above each box plot. For all box plots, dots outside whiskers represent outliers. The horizontal line within the box represents the median sample value. The ends of the boxes represent the 3rd (Q3) and 1st (Q1) quartiles. The whiskers show values that are 1.5 times interquartile range (1.5xIQR) above and below Q1 and Q3.

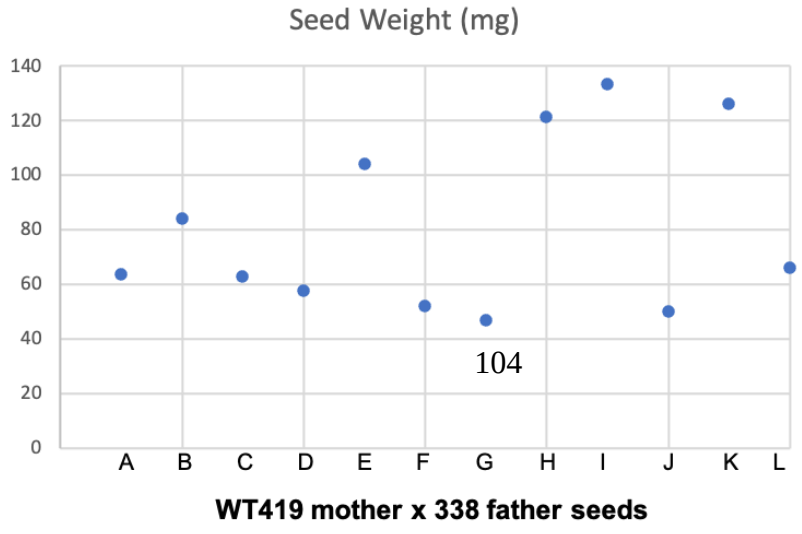
A.


Seed Length (cm)
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
A B C D E F G H I J K L
WT419 mother x 338 father seeds

B.


Seed Width (cm)
1.2
1
0.8
0.6
0.4
0.2
0
A B C D E F G H I J K L
WT419 mother x 338 father seeds

C.


Seed Weight (mg)
140
120
100
80
60
40
20
0
A B C D E F G H I J K L
WT419 mother x 338 father seeds

Supplemental Figure 3.7: WT419 x 338 F1 Seed Traits
The measurements from twelve F1 seeds (A-L) recovered from crosses of WT419 mother by 338 father flowers. Measurements include **A)** Seed length (cm) **B)** seed width (cm), and **C)** seed weight (mg).

Seed ID	Letter ID	Date of cross	Mother	Father	Seed length (cm)	Seed width (cm)	Weight (mg)	Sink(s) Float(f)	Germination date
P3760	A	10/10/22	WT419	338	1	0.6	63.6	f	N/A
P3761	B	10/10/22	WT419	338	0.9	0.5	84.1	f	3/1/23
P3788	C	10/13/22	WT419	338	1.1	0.5	63	f	N/A
P3789	D	10/13/22	WT419	338	1.1	0.5	57.5	f	N/A
P3790	E	10/13/22	WT419	338	1.1	0.5	103.9	f	4/12/23
P3791	F	10/13/22	WT419	338	1	0.5	52.2	f	N/A
P3792	G	10/13/22	WT419	338	1	0.5	46.7	f	N/A
P3793	H	10/13/22	WT419	338	1	0.5	121.1	s	2/21/23
P3794	I	10/13/22	WT419	338	1.1	0.5	133.1	s	2/26/23
P3795	J	10/13/22	WT419	338	1	0.5	50.1	f	N/A
P3825	K	10/14/22	WT419	338	1	0.5	126.2	s	3/16/23
P3826	L	10/14/22	WT419	338	1	1	66	f	N/A
P3711	N/A	Spring 22	WT419	open pollination	1.0	0.6	Not recorded	f	N/A
P3712	N/A	Spring 22	WT419	open pollination	0.9	0.6	Not recorded	s	12/25/22
P3723	N/A	Spring 22	WT419	open pollination	1.0	0.6	Not recorded	s	12/25/22
P3724	N/A	Spring 22	WT419	open pollination	0.9	0.6	Not recorded	f	12/28/22
P3736	N/A	Spring 22	WT419	open pollination	0.9	0.6	Not recorded	s	12/28/22
P3737	N/A	Spring 22	WT419	open pollination	0.9	0.5	Not recorded	f	N/A

Supplemental Table 3.3: Recovered Seed Information
Table of recorded information for WT419 by 338 F1 seed and WT419 open pollinated seed recovered from the field and tested for germination. Seed ID, Letter ID, Date of Cross, Mother, Father, Seed length (cm), width (cm), weight (mg), Float test results, and germination data are reported. Results for seed float test denoted as ‘s’ for sink and ‘f’ for float.


Line 269
Insertion 1
hap2–chromosomeI
T-DNA Insertion
Manes.01G010150
Manes.01G010100
Manes.01G010200
Manes.01G010300
3280000
3290000
3300000
3310000
Insertion 2
hap1–chromosomeXII
T-DNA Insertion
Manes.12G015800
Manes.12G015600
Manes.12G015700
Manes.12G016000
1475000
1480000
1485000
1490000
1495000
1500000
Line 338
Insertion 1
hap1–chromosomeVIII
T-DNA Insertion
Manes.15G176704
Manes.02G145651
5205000
5210000
5215000
5220000
5225000
Insertion 2
hap2–chromosomeXVIII
T-DNA Insertion
32090000
32095000
32100000
32105000
32110000
Line 27
Insertion 1
hap2–chromosomeXIV
T-DNA Insertion
Manes.14G047600
Manes.14G047700
4275000
4280000
4285000
4290000
Insertion 2
hap1–chromosomeXI
T-DNA Insertion
Manes.11G152174
Manes.11G152042
32900000
32925000
32950000
32975000
33000000
Line 30
Insertion 1a
hap1–chromosomeVI
T-DNA Insertion
Manes.06G060600
Manes.06G060700
20320000
20325000
20330000
20335000
20340000
20345000
Insertion 1b
hap1–chromosomeVI
T-DNA Insertion
Manes.06G168900
Manes.06G169000
Manes.06G169100
30485000
30490000
30495000
30500000
Insertion 2
hap2–chromosomeIV
T-DNA Insertion
Manes.04G111300
Manes.04G111400
Manes.04G111500
Manes.04G111600
31310000
31320000
31330000
Insertion 3
hap1–chromosomeXV
T-DNA Insertion
Manes.15G051025
Manes.15G050825
Manes.15G050900
Manes.15G051150
4160000
4170000
4180000
4190000
Line 41
Insertion 1
hap1–chromosomeII
T-DNA Insertion
Manes.02G157450
Manes.02G157300
Manes.02G157400
Manes.02G157500
Manes.02G157700
12610000
12615000
12620000
12625000
12630000
Insertion 2
hap2–chromosomeVII
T-DNA Insertion
Manes.07G059533
Manes.16G057275
9255000
9260000
9265000
9270000
9275000
Insertion 3a
hap1–chromosomeIII
T-DNA Insertion
Manes.03G023800
Manes.03G023900
Manes.03G024000
Manes.03G024100
1965000
1970000
1975000
1980000
1985000
1990000
Insertion 3b
hap1–chromosomeIII
T-DNA Insertion
Manes.03G072850
12070000
12080000
12090000
Line 54
Insertion 1
hap1–chromosomeXVIII
T-DNA Insertion
Manes.18G002800
Manes.18G003000
Manes.18G002500
Manes.18G002601
Manes.18G002700
Manes.18G002900
610000
615000
620000
625000
630000
Insertion 2
hap1–chromosomeIX
NA Insertion
Manes.09G002700
960000
970000
980000
990000
Insertion 3
hap2–chromosomeXIII
T-DNA Insertion
5910000
5915000
5920000
5925000
5930000

Supplemental Figure 3.8: Transgene Insertion Number and Location

Graphic depicting transgene insertion number and location in lines 269, 338, 27, 30, 41, and 54. For each line, the total number of insertions is listed (left-hand side). For each insertion, the haplotype and chromosome location for the site of insertion is noted. Red triangles and brackets depict the site of transgene insertion. The bracket indicates a larger distance between insertion coordinates. Nearby genes are annotated with their Manes ID.

Primer Number:	Sequence: (5'-3')	Description:
MeSWEET10a outer F (150)	GGGAATGAGAGTGTGTGGTTAG	Genotyping primer avoiding EBE repair
MeSWEET10a outer R (151)	GCAGTCAAGGTGAGGAAGAAT	Genotyping primer EBE repair
MeSWEET10a inner F (21)	GTACCCTTCCTCAACACCC	Genotyping primer
MeSWEET10a inner R (22)	GCTGCAAAACCATCGAGCAA	Genotyping primer
MeSWEET10a RT-PCR F (200)	CTTACTGTGTACCTTATCTATGCCACAAAGAAG	RT-PCR *MeSWEET10a* expression primer
MeSWEET10a RT-PCR R (201)	GCCATGTGTAAGGAAAAGAGTTAAGATAGCG	RT-PCR *MeSWEET10a* expression primer
Actin RT-PCR F (37)	ACAGTGTCTGGATCGGAGGATC	RT-PCR Actin housekeeping gene expression primer
Actin RT-PCR R (38)	GAAGCACTTCCTGTGGACGATG	RT-PCR Actin housekeeping gene expression primer
GTPb RT-PCR F (109)	CCTCAAAGGCTGAGCCACAGA	RT-PCR GTPb housekeeping gene expression primer
GTPb RT-PCR R (110)	GGGAGAAACAATACAGGCACCAATCAC	RT-PCR GTPb housekeeping gene expression primer
A_CmYLCV-F (113)	TGCTCTTCGCGCTGGCAGACATACTGTCCCAC	CmYLCV primer for CRISPR/Cas9 construct assembly.
B_10gRNA1-R (114)	TCGTCTCCTAAACGCTTCTCCTGCCTATACGGCAGTGAACCTG	gRNA1 R primer for CRISPR/Cas9 construct assembly.
C_10gRNA1-F (115)	TCGTCTCATTTATATAGGGGGTTTTAGAGCTAGAAATAGC	gRNA1 F primer for CRISPR/Cas9 construct assembly.
D_10gRNA2-R (116)	TCGTCTCATCATTGGACTTCCTGCCTATACGGCAGTGAAC	gRNA2 R primer for CRISPR/Cas9 construct assembly.
C_10gRNA2-F (117)	TCGTCTCAATGACAAGTGCAGTTTTAGAGCTAGAAATAGC	gRNA2 F primer for CRISPR/Cas9 construct assembly.
E_35Sterm-R (118)	TGCTCTTCTGACCTGCCTATACGGCAGTGAAC	35S terminator primer for CRISPR/Cas9 construct assembly
B_10gRNA4-R (236)	TCGTCTCCTTGACATATTTTCTGCCTATACGGCAGTGAACCTG	gRNA4 R primer for CRISPR/Cas9 construct assembly.
C_10gRNA4-F (237)	TCGTCTCATCAATGTAACAGGTTTTAGAGCTAGAAATAGC	gRNA4 F primer for CRISPR/Cas9 construct assembly.
D_10gRNA5-R (238)	TCGTCTCATTGCACAACATACTGCCTATACGGCAGTGAAC	gRNA5 R primer for CRISPR/Cas9 construct assembly.
C_10gRNA5-F	TCGTCTCAGCAATGATGGATGTTTTAGAGCTAGAAATAGC	gRNA5 F primer for CRISPR/Cas9 construct assembly.

Supplemental Table 3.4: *MeSWEET10a* Primer List
Table of all primers used in this study. The primer name/stock number (left), sequence (middle), and description (right) are provided.

Chapter 4: Applying CRISPR/Cas9 Genome Editing to Investigate the Role of Cassava Pectate Lyases in Promoting Bacterial Blight

4.1 Introduction

XpmΔTAL14 mutants (suicide vector knockout) were reported to have lower bacterial titer *in planta* compared to Xpm WT post midvein inoculations. (Cohn and Bart et al., 2014). Two of the candidate S genes (Gene IDs: Manes.15G048700/ Cassava4.1_007516 and Manes.03G152600/ Cassava4.1_007568) were identified as putative cassava pectate lyase-like (PLL) proteins. Pectate lyases are enzymes that degrade a key structural component of the plant cell wall called pectate. (Uluisik and Seymour, 2020). The cell wall functions in several vital roles for plants such as providing mechanical strength, maintaining cell integrity, and limiting the size and kinds of molecules that go into the cell. Additionally, the cell wall acts as a barrier and defense mechanism to resist pathogen invasion and spread (Taiz et al., 2015, Chapter 14). The cell wall is an extracellular matrix comprised of proteins and polysaccharides. There are three types of polysaccharides including cellulose, hemicellulose, and pectin. Cellulose provides tensile strength, hemicellulose binds to cellulose and aids in microfibril assembly, and pectin is considered the glue of the plant cell wall. Pectin forms a hydrophilic gel layer where both cellulose and hemicellulose are embedded and prevents the collapse of cellulose (Weaver, 2012). Pectin is a family of different polysaccharides such as homogalacturonan, rhamnogalacturonan-I, rhamnogalacturonan-II, and

xylogalacturonan (Mohnen, 2008). Pectate is a demethylated polygalacturonic acid (PGA) involved in crosslinking with different molecules (Muller et al., 2007). Pectate lyases are able to degrade pectate by cleaving α-1,4-PGA links between individual pectate groups (Uluisik and Seymour, 2020). Pectate lyases have known roles in pollen and flower development, seed germination, secondary cell wall formation, leaf senescence, and fruit peeling (Muller et al., 2013, Leng et al., 2017, Prakash et al., 2017, and Uluisik and Seymour, 2020).

Pectate lyases are also known to contribute to plant disease. They can either be produced and secreted by pathogens and/or induced in the plant host during pathogen infection. For example, soft rot-inducing bacteria in the species *Erwinia* and the pathogen responsible for citrus canker, *X. citri* subsp. *citri* both use pectate lyases to promote pathogenesis (Brencic and Winans, 2005 and Chang et al., 2016). Several fungal and nematode phytopathogens are known to secrete pectate lyases during infection (Uluisik and Seymour, 2020). Research is sparse on host-derived pectate lyases that are expressed due to a pathogen. However, a study of the *Arabidopsis* susceptibility locus, *PMR6* which encodes a pectin lyase-like protein found that *PMR6* mutants are less susceptible to the causal agent of powdery mildew, *Erysiphe cichoracearum* (Vogel et al., 2002).

Cell wall degrading enzymes like pectate lyases are used by pathogens to aid in tissue invasion and to gain access to plant cell nutrients (Yang et al., 2018). In cassava, there is evidence of pectin degradation occurring post-Xpm infection (Boher et al., 1995). The density of pectin in healthy and infected cassava parenchyma and xylem cells was quantified and in both cell types, infected tissues had significantly reduced pectin density compared to healthy uninfected tissue (Boher et al., 1995). In some plants, pectinaceous blockages called tyloses are produced in the

xylem as a defense mechanism to prevent pathogen spread through the vasculature (Yadeta and J. Thomma, 2013). It is possible that pectate lyases can be used to prevent such blockages and aid in pathogen dispersal. We hypothesize that TAL14-mediated expression of *MePLLs* aids Xpm colonization of cassava either by degrading pectate at host cell surfaces or tyloses in the vasculature. Various studies have demonstrated that editing TAL effector S genes is a viable strategy for decreasing plant susceptibility to pathogen-induced diseases (Koseoglou et al., 2021).

In this study, we compared the sequence of both putative *MePLL*s to previously validated pectate lyases protein sequences, used RT-PCR and bacterial growth assays to characterize virulence-related phenotypes related to TAL14 in the TME419 cassava cultivar, and used a dual gRNA CRISPR/Cas9 strategy to generate *MePLL* mutants with edits to the EBE site and coding sequence.

4.2 Results

4.2.1 Validation of Putative Cassava Pectate Lyase-Like Genes

Two of the genes upregulated post Xpm infection were annotated as putative *MePLL*s. However, gene annotation or classification quality and accuracy can vary for numerous reasons such as genome assembly contiguity, genome complexity, and the caliber of the tool used for annotation (Yandell and Ence, 2012). Furthermore, gene homology is often used to annotate genes across species, and annotations based on experimental evidence are limited (Bolger, 2018). These along with other complications can lead to incorrect gene annotation. Therefore, a literature search was conducted to identify pectate lyases that were previously validated in the literature by

experimental evidence from pectate lyase cleavage assays and/or characterization of transcript and protein sequences for conserved domains. The validated pectate lyases protein sequences were then compared to the putative *MePLL*s. Four validated pectate lyases were selected from different plants including *Arabidopsis thaliana* (Gene ID: AT3G54920), *Solanum lycopersicum* (Gene ID: NM_001309386.1), *Gossypium hirsutum* (NC_030080.1), and *Oryzae sativia* (Gene ID: XM_015759304.2) (Palusa et al., 2007, Wing et al., 1990, Wang et al., 2010, and Zheng et al., 2018). Additionally, a well-characterized bacterial pectate lyase, from *Erwinia chrysanthemi* (Gene ID: Y13340.1), was chosen as an outgroup (Keen et al., 1984). Information for each pectate lyase gene is provided in **Table 4.1**.

Gene name	Organism	Gene ID	Gene Size (bp)	Protein length (AA)
At_PMR6	*Arabidopsis thaliana*	AT3G54920	3.8kb	501
Ec_PELI	*Erwinia chrysanthemi*	Y13340.1	1.03kb	344
Sl_LAT56	*Solanum lycopersicum*	NM_001309386.1	1.6kb	398
Gh_PLL	*Gossypium hirsutum*	NC_030080.1	1.7kb	411
Os_DEL1	*Oryzae sativia*	XM_015759304.2	2.79kb	608
MePLL_7516	*Manihot esculenta*	cassava4.1_007516	4.2kb	449
MePLL_7568	*Manihot esculenta*	cassava4.1_007568	4.0kb	448

Table 4.1: Pectate Lyases Used for Protein Sequence Comparison
Gene information from five validated pectate lyase amino acid sequences identified in the literature and compared to putative cassava pectate lyases (*MePLL*s).

The transcript and protein sequences for all genes were downloaded from either the Phytozome or National Center for Biotechnology Information (NCBI) databases. Using the software program Geneious, the validated pectate lyase and putative MePLL amino acid sequences

were aligned and regions of sequence identity were noted. The plant pectate lyases shared 55 percent sequence identity (**Figure 4.1A**). As expected, including the bacterial PLL in the alignment confirmed divergence across plants and bacteria. However, when each PL protein sequence was analyzed with the database, pfam, to identify conserved motifs or domains, we found that all pectate lyases had a conserved ~195 AA motif at the C-terminal end called pectate lyase C (Pel C) (Thurn and Chatterjee, 1987 and Marin-Rodriguez et al., 2002). Pel C was previously isolated from *Bacillus subtilis,* cloned into *Escherichia coli,* and enzymatic activity assays confirmed it cleaved pectate (Soriano et al., 2006). The resolved 3D structure of Pel C from *E. chrysanthemi* determined that Pel C contains a parallel beta-helix folding motif (Yoder et al., 1993). Additionally, the Pel C domain has been found in other plant and bacterial PLLs. The Pel C domain alone was aligned for all plant pectate lyases and 66.3 percent pairwise identity was observed. When the two MePLL Pel C domain sequence alone was aligned, we found they shared 96.1 percent sequence identity (**Figure 4.2**).

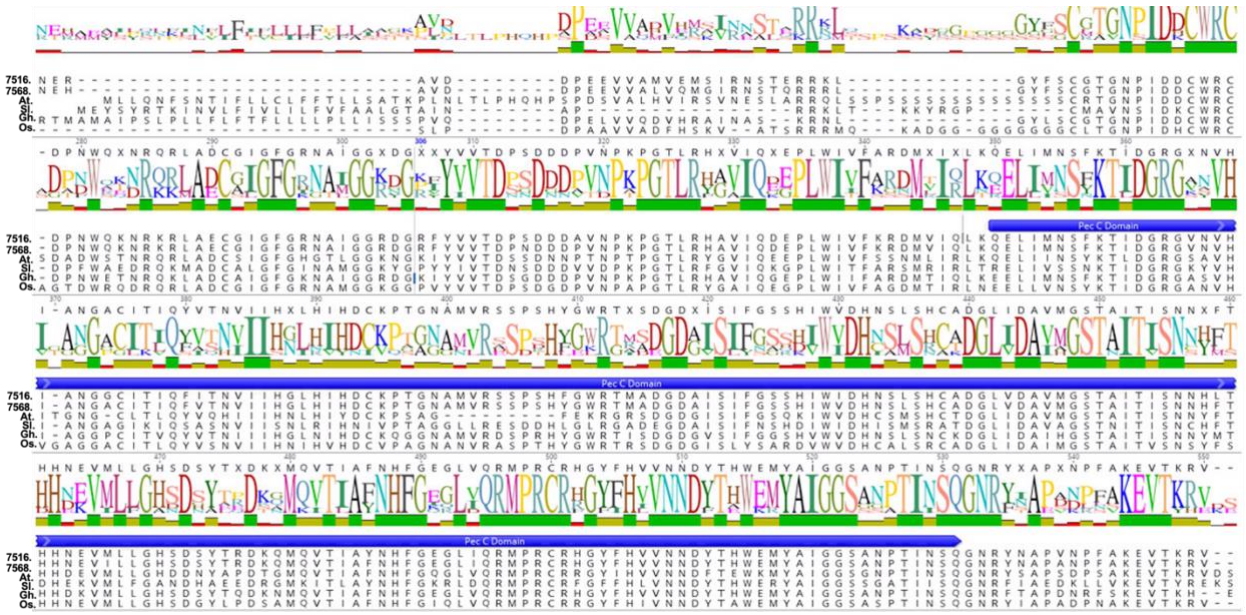


Figure 4.1: Plant Pectate Lyase Alignment

A) The plant pectate lyase amino acid sequence alignment results for MePLL_7516, MePLL_7568, and the validated plant pectate lyases from *Arabidopsis thaliana* (At.), *Solanum lycopersicum* (Sl.), *Gossypium hirsutum (*Gh.), and *Oryzae sativia* (Os.). The consensus sequence, sequence logo, and graphic of amino acid similarity are shown above the alignments. Green highlights areas of high sequence similarity, yellow is medium similarity, and red low similarity.

Figure 4.2: Cassava Pectate Lyase Pel C Domain Alignment.

A) An alignment of the MePLLs Pel C domain alone. Red boxes highlight areas of polymorphisms. The consensus sequence, sequence logo, and graph of amino acid identity are shown above the alignments. Green highlights areas of high sequence similarity.

Next in our study, the protein sequences from 26 AtPLLs, 12 OsPLLs and the validated PLLs from cotton, tomato, and *E. chrysanthemi* were obtained and a neighbor-joining phylogeny analysis was performed (**Figure 4.3**). As expected, *E. chrysanthemi* clustered as an outgroup away from plant PLLs. Further, we found that the MePLLs most closely clustered with other PLLs from PLL subfamily 1 including AtPLL26. Transcriptome data shows *AtPLL26* has high expression in mature flowers (Klepikova et al., 2016). Moreover, spatial and temporal analysis of *ATPLL26* fused with GUS showed it is expressed at the abscission zones of *Arabidopsis* sepals, petals, and stamen (Sun and van Nocker, 2010). The abscission zone is an area of cell separation where plants can shed organs like fruit leaves or flowers and other pectate lyases genes have previously been shown to have expression at plant abscission zones (Merelo et al., 2017). Together, these data support the annotation of the putative *MePLLs* as pectate lyases.

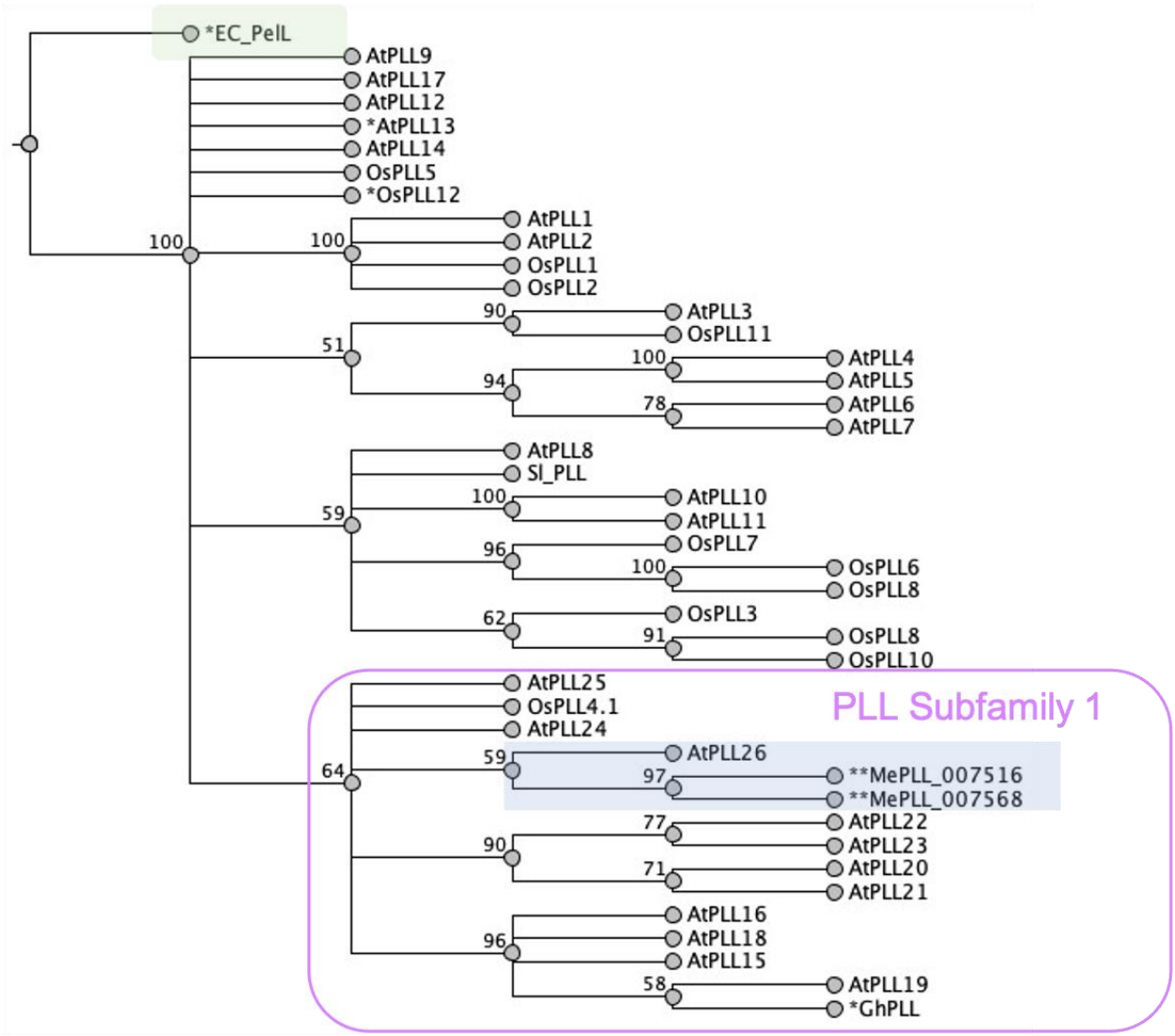


Figure 4.3: Pectate Lyases Phylogenetic Tree

A) A Neighbor-joining phylogenetic tree of pectate lyase protein sequences from *Arabidopsis thaliana* (26 AtPLLs), *Oryza sativa* (12 OsPLLs), the validated pectate lyases (denoted by *), and the cassava pectate lyases (denoted by **). The green box highlights the bacterial *E. chrysanthemi* pectate lyase used as an outgroup for phylogenetic analysis. The blue box highlights the MePLLs and the next most closely clustered plant pectate lyase, AtPLL26. The purple box outlines pectate lyase gene subfamily 1.

4.2.2 Characterizing Disease Phenotypes and TAL14-Mediated Induction of *MePLL*s in Xpm infected WT419 Plants

To validate the disease phenotypes associated with TAL14 in the cassava cultivar TME419 (WT419), wildtype plants were infected with mock, Xpm WT, and XpmΔTAL14 treatments. At 48HPI, samples were collected for RNA extraction and semi-quantitative reverse transcriptase (RT**)**-PCR experiments. All samples were analyzed for expression of *MePLL_7516* and *MePLL_7568.* Results for samples amplified with *MePLL_7516* RT-PCR primers (Cohn and Bart et al., 2014) showed that in mock and XpmΔTAL14 treated samples, there was a band indicating some basal expression of *MePLL_7516* in cassava leaves. In Xpm WT-treated samples, there was a brighter band indicating TAL14-mediated overexpression of *MePLL_7516*. Results for samples amplified with *MePLL_7568* RT-PCR primers showed that in mock and XpmΔTAL14 treated samples, there was no expression of *MePLL_7568.* However, there was RT-PCR product in Xpm WT treated samples signifying TAL14 is able to induce *MePLL_7568* expression (**Figure 4.4A**). Next, wildtype plants were infected for bacterial growth assays with Xpm WT and XpmΔTAL14 treatments. At 0 and 9 DPI, there was no significant difference between Xpm WT and XpmΔTAL14 colony-forming units (CFUs). However, at 4 and 6 DPI, XpmΔTAL14 showed reduced growth *in planta* compared to Xpm WT (**Figure 4.4B**). In additional bacterial growth assays**,** the difference between Xpm WT or XpmΔTAL14 bacterial titer varied as seen in Xpm WT and XpmΔTAL20 growth assays. These results along with published literature are consistent with the hypothesis the *MePLL* may play a role in Xpm virulence in TME419.

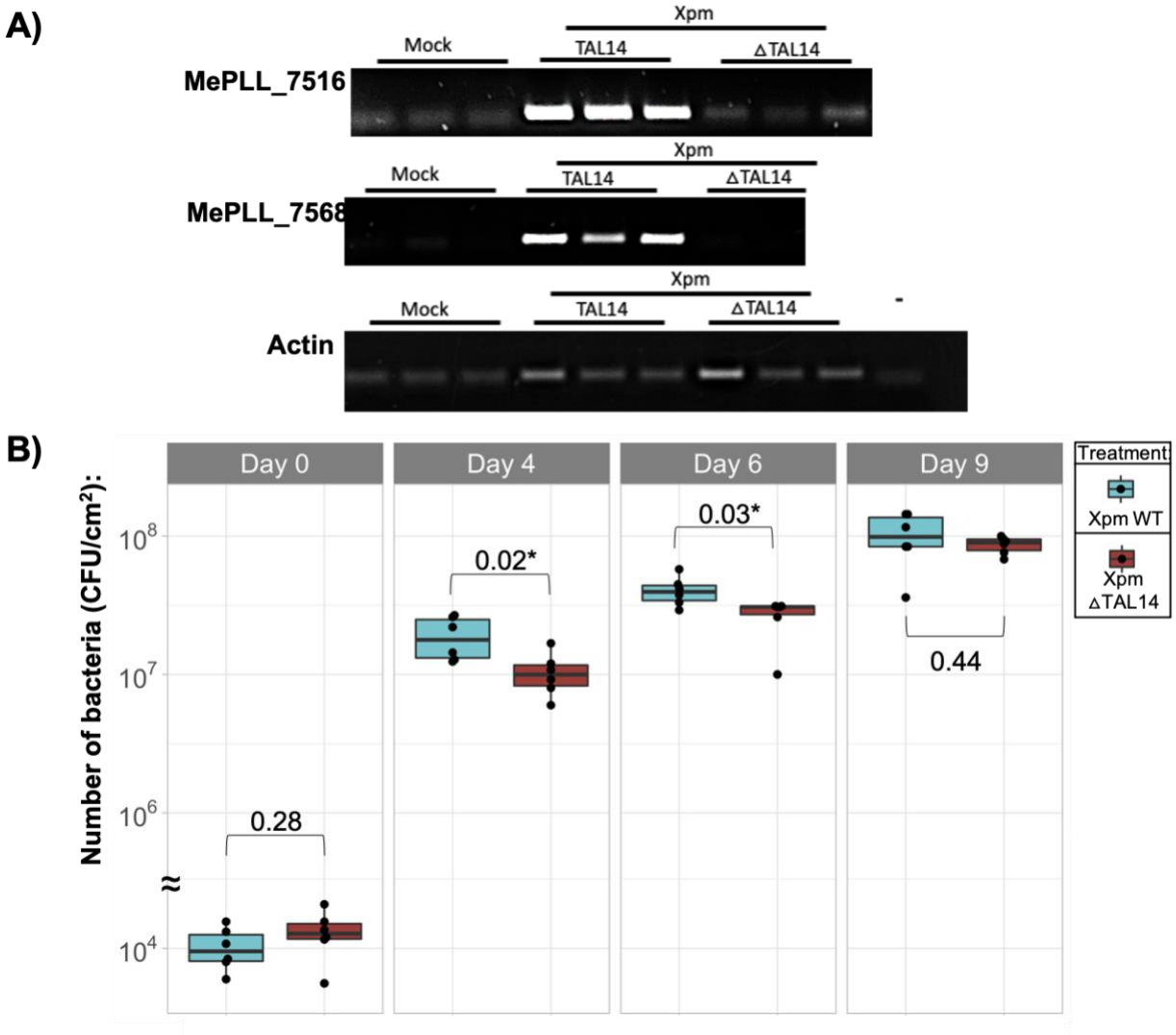


Figure 4.4: TAL14-Mediated Induction of *MePLL*s and Bacterial Growth Assay
A) RT-PCR of wild type cassava plants infected with mock, Xpm WT, Xpm△TAL14. Top gel shows results of RT-PCR with primers amplifying MePLL_7516 cDNA. The middle gel shows results of RT-PCR with primers amplifying MePLL_7568. The bottom gel shows results of RT-PCR with primers amplifying the housekeeping gene, Actin. '-'denotes a negative water control. **B)** Number of bacteria in cassava leaves measured at 0, 4, 6, or 9 DPI post syringe infiltration with Xpm (blue) and Xpm△TAL14 (red) treatments (X-axis, OD600=0.01). Colony forming units (CFU/cm2, Y-axis) are plotted. Black dots (N=6) represent technical replicates from one

independent bacterial growth experiment. Results of statistical analyses (Unpaired student's t-test with unequal variance) comparing the difference between Xpm WT and Xpm△TAL14. P-values are shown above or below brackets indicating the comparison types for statistical analyses.

4.2.3 Generation of CRISPR/Cas9 Constructs Targeting *MePLL*

Next, CRISPR/Cas9 constructs were designed with gRNAs targeting each *MePLL* EBE site, 5'UTR, and translation start site (ATG). The *MePLL* gDNA sequence was downloaded from Phytozome (Reference genome AM560-2). The *MePLL* sequences were blasted against a separate cassava reference genome made by our lab, TME7_v0.5.3a, to identify the sequencing scaffolds where each *MePLL* was located. The genes were then annotated with promoter regulator elements of interest including the TAL14 EBE site (Cohn 2016). The Geneious software "find CRISPR sites" function was used to identify all potential targets and Cas9 *(S. pyogenes*)-specific PAM sites (sequence: 5′-NGG-3′). Candidate gRNA were identified for target sites at the *MePLL* EBE, 5'UTR, and ATG. The gRNA targeting the TAL14 EBE was designed to target both *MePLLs*. Therefore, it could potentially make edits to both genes. All candidate gRNA sequences were blasted against the cassava genome to identify potential off-targets. gRNAs with BLAST results specific to each respective *MePLL* scaffold, and with no strong off-target candidates were selected for further assessment. The genome browser JBROWSE was used to search for SNPs in candidate gRNAs using whole genome re-sequencing data from TME419, the cultivar used for eventual cassava transformations, mapped to the TME7 reference genome. No SNPs expected to impact gRNA binding to a target site were identified and all final gRNAs were found to be suitable for construct design.

In total, four dual gRNA CRISPR/Cas9 constructs were assembled using a multiple gRNA spacer Csy4 array as previously described (Cermak et al 2017). Construct 7516 set 1 contains gRNAs4 and 7, targeting both *MePLL* EBE sites and the 7516 5'UTR. Construct 7516 set 2 contains gRNAs4 and 15, targeting both *MePLL* EBE sites and the 7516 ATG. Construct 7568 set 1 contains gRNAs2 and 9 targeting both *MePLL EBE* sites and the 7568 5'UTR. Construct 7568 set 2 contains gRNAs2 and 14 targeting both *MePLL EBE* sites and the 7568 ATG. All constructs were cloned in the pTRANS_220D backbone and have a kanamycin resistance cassette. To ensure proper construct assembly, all constructs were verified using colony PCR and Sanger sequencing (**Figure 4.5**).

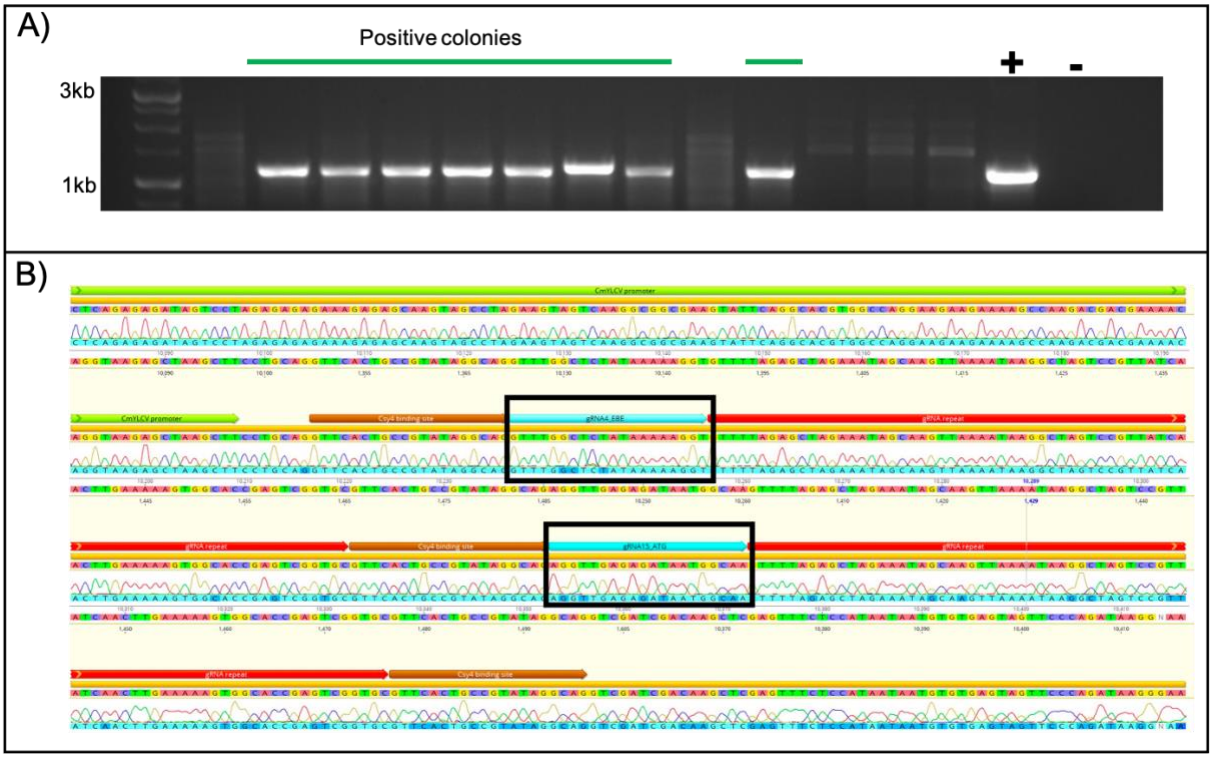


Figure 4.5: Generating *MePLL* CRISPR/Cas9 Constructs
A) Colony PCR results from stellar competent *E. coli* colonies with the *MePLL_7516* gRNAs4 and 15 sequences cloned into the vector, pTRANS_220D. The expected product size for the target

area is 1.2 kb. The green lines highlight colonies positive for the expected PCR product. + denotes a positive control and '–'denotes a negative water control. B) Sanger sequencing results of a positive clone carrying *MePLL_7516* gRNAs4 and 1. Black boxes outline the presence of expected gRNA sequences.

4.2.4 Recovery and Characterization of Transgenic Lines

Five rounds of cassava transformations were completed with the *MePLL* constructs. In total, forty-five transgenic lines were recovered. Fourteen lines were recovered from *MePLL_7516* set 1. Five lines were recovered from *MePLL_7516* set 2. Fourteen lines were recovered from *MePLL_7568* set 1 and twelve lines were recovered from *MePLL_7568* set 2 (**Table 4.2**). Leaf tissue was sampled from each line at the plantlet stage in tissue culture for genomic DNA (gDNA) extraction. Since the EBE gRNAs are potentially able to edit both *MePLL* TAL14 binding sites, promoter-specific forward primers and reverse primers localizing downstream of the ATG were designed and used to amplify each target site. Plants were genotyped by Sanger sequencing and a selection of ten transgenic lines (#2, 6, 9, 10, 63, 66, 94, 121, and 142) were characterized for mutation type and location (**Table 4.3**). In all ten lines, we identified edits at one or both TAL14 EBE sites. This confirmed the ability of the EBE targeting gRNA to induce edits at both *MePLL* sites. Additionally, we observed mutants with edits to the *MePLL*s 5'UTR and frameshift mutants with edits after the ATG. Sanger sequencing results for three of the mutants are shown (**Figure 4.6**). These data suggest that editing the *MePLL*s is non-lethal to the plant and that gRNAs selected for construct design resulted in a mix of desired mutation types.

	1st gRNA target	2nd gRNA target	Callus selected (Stage 2):	Mature lines recovered:
***MePLL 7516* Set 1 (gRNAs 4+7)**	*7516* and *7568* EBE sites	*7516* 5' UTR	355	14
***MePLL 7516* Set 2 (gRNAs4+15)**	7516 and 7568 EBE sites	*7516* ATG	263	5
***MePLL 7568* Set 1 (gRNAs2+9)**	*7516* and *7568* EBE sites	*7568* 5' UTR	245	14
***MePLL* 7568 Set 2 (gRNAs2+14)**	*7516* and *7568* EBE sites	*7568* ATG	339	12
			Total lines recovered:	**45**

Table 4.2: *MePLL* Transformation Overview
Summary table of transgenic lines recovered from each *MePLL* targeting construct at both the callus and mature line stages of cassava transformation.

Line #:	Construct	*MePLL_7516* Edits	*MePLL_7568* Edits
2	*7516* set 1	102 bp deletion including EBE and 5'UTR	2 bp deletion at EBE
6	*7516* set 1	301 bp deletion including EBE and 5'UTR	4 bp deletion at EBE
9	*7568* set 1	2 bp deletion at EBE	1 bp deletion at EBE
10	*7516* set 1	1 bp deletion at EBE	14 bp deletion at EBE
63	*7568* set 1	2 bp deletion at EBE	3 bp deletion at EBE
66	*7568* set 1	2bp deletion at EBE	4 bp deletion upstream of EBE (silent)
94	*7516* set 2	3 bp bp deletion at EBE and 9 bp deletion after ATG (frameshift)	1 bp deletion at EBE
121	*7516* set 1	No edits (WT-like or biallelic silent mutant)	1 bp deletion at EBE
142	*7516* set 1	3 bp deletion in 5'UTR (silent)	2 bp deletion at EBE

Table 4.3: Ten *MePLL* Mutant Genotypes
Summary table of genotypes mutation types recovered for *MePLL 7516* and *7568*.


A)
Line 6: MePLL_7516 Set 1
7516: 301 bp deletion including TAL14 EBE and 5'UTR
gRNA4
gRNA7
EBE
MePLL_7516 5'UTR
7568: 4 bp deletion at TAL14 EBE
gRNA4
MePLL_7568 5'UTR
B)
Line 94: MePLL_7516 Set 2
7516: 3 bp deletion at TAL14 EBE and 9 bp deletion after TSS
gRNA4
gRNA15
EBE
MePLL_7516 5'UTR
Exon 1
7568: 1 bp deletion at TAL14 EBE
gRNA4
MePLL_7568 5'UTR
C)
Line 66: MePLL_7568 Set 1
7568: 3 bp deletion at TAL14 EBE
gRNA2
gRNA9
EBE
MePLL_7568 5'UTR
7516: 1 bp deletion at TAL14 EBE
gRNA2
MePLL_7516 5'UTR
D)
WT419
4/2
419-149
3/15
WT419
Line 6

Figure 4.6: Sanger Sequencing Results for Select *MePLL* Mutants
Sanger sequencing results for **A)** line 6, **B)** line 94, and **C)** line 66. Orange box denotes the TAL14 EBE site, red denotes gRNAs, and gray boxes denote *MePLL* 5'UTR. **D)** Representative images of WT419 and *MePLL* mutant line 6 plantlets on MS2 plates.

4.3 Discussion and Conclusion

This study investigated putative cassava pectate lyases as susceptibility gene candidates for the Xpm effector TAL14. Previous work identified two cassava pectate lyase genes as direct susceptibility targets of Xpm TAL14 (Cohn and Bart et al., 2014 and Cohn et al., 2016). These genes were annotated as pectate lyases based on automated annotation which is not always accurate. Using pectate lyases validated in literature, we conducted sequence homology analysis to compare the *MePLL*s to other pectate lyases. We found that the MePLL protein sequences were similar to other plant pectate lyases and that both MePLLs contained the pectate lyase domain, Pel C. A phylogenetic analysis of the *MePLL*s with other pectate lyases showed that both *MePLL*s clustered closely with members of the pectate lyase subfamily 1. The *MePLL*s were further examined for a connection to TAL14-mediated induction and disease phenotypes in Xpm-infected WT419 plants through RT-PCR and bacterial growth analyses.

Additionally, we wanted to determine if a CRISPR/Cas9 gene editing approach to generate *MePLL* mutant lines was a viable strategy for improving cassava resistance to bacterial blight. Plant cell walls play several vital functions and therefore editing a cell wall factor could be lethal or detrimental to plant health or development. We developed CRISPR/Cas9 constructs with gRNAs targeting each *MePLL* EBE, 5'UTR, and/or ATG. We performed several rounds of

transformation to determine if recovery of mutant lines with edits to the *MePLLs* EBE site and/or coding sequences was possible. Forty-five transgenic lines were generated and Sanger-sequencing from a select number of lines confirmed edits to the *MePLL* sites of interest.

The exact role that *MePLL*s play in promoting CBB susceptibility and Xpm virulence remains unknown. We hypothesize that the *MePLL* mutants generated in this study will have decreased susceptibility to cassava bacterial blight. Future work using RT-PCR to validate that Xpm effector TAL14 cannot induce *MePLL* induction is required. Once a number of *MePLL* mutants are selected based on genotype and RT-PCR results, these mutants can be characterized for disease susceptibility following Xpm infection. Based on bacterial growth assays performed on wildtype cassava infected with Xpm WT and XpmΔTAL14, there may be a subtle phenotype of decreased Xpm WT growth in *MePLL* mutants compared to infected wildtype plants.

Previous work reported that loss of TAL14 in Xpm did not decrease water-soaking symptoms, which are the first indicator of cassava bacterial blight (Cohn and Bart et al., 2014). Therefore, we did not expect to see an early-stage visible reduction of CBB disease symptoms in *MePLL* mutants and this was confirmed by eye post-Xpm WT and XpmΔTAL14 infection in wildtype cassava. Additional disease phenotyping strategies may be required to capture disease severity differences between *MePLL* mutants and wildtype cassava infected with Xpm. We hypothesize that the *MePLL*s contribute to pathogen virulence by degrading pectate at host cell surfaces or tyloses in the vasculature. Therefore, measuring pectate or pectin concentrations in the plant cell walls at sites of Xpm infection with techniques such as immunolabelling may be a better measure of CBB disease severity for infected *MePLL* mutants. Host-derived pectate lyase gene

expression induced by bacterial pathogens has previously been understudied. *MePLL* mutants generated in this study will serve as a valuable resource and provide a novel opportunity to investigate the role of pectate lyases in Xpm pathogen virulence. Furthermore, if *MePLL* mutants exhibit decreased susceptibility to bacterial blight, cassava with stronger tolerance to CBB can be generated by stacking edits at the *MePLL*s and additional S gene targets such as *MeSWEET10a*.

4.4 Materials and Methods

4.4.1 *MePLL* Construct Design and Cloning:

The *MePLL* gDNA sequences (Gene IDs: Manes.15G048700/ Cassava4.1_007516 and Manes.03G152600/ Cassava4.1_007568) were downloaded from Phytozome (Reference genome AM560-2, *Manihot esculenta* genome v6.1) and uploaded to the software Geneious. Notable promoter regulatory elements were annotated as previously reported by Cohn et al 2016 including the effector binding element (EBE) site where TAL14 binds. The reported EBE sequence was confirmed using the TALEnt target finder reference. The Geneious "find CRISPR sites" function was used to find all potential targets and Cas9 *(S. pyogenes*)-specific PAM sites (sequence: 5′-NGG-3′). Candidate guide RNA (gRNAs) target sequences were selected based on those whose targets were near the TAL14 EBE and the translational start site of MePLL or within the 5'UTR. Candidate gRNAs were further analyzed by comparing the candidate gRNAs against the cassava genome to identify potential off-targets using NCBI-BLAST. Candidate gRNAs without strong potential off-target hits were moved forward for construct design. All constructs were assembled using a multiple gRNA spacer Csy4 array as previously described (Čermák et al., 2017). Three

constructs were used for this study. All constructs were designed to carry two gRNAs, were cloned in the pTRANS_220D backbone, and have a kanamycin resistance cassette. Construct 7516 set 1 carries gRNA4 (GTTTGGCTCTATAAAAAGGTCGG) which targets both *MePLL* TAL14 EBE sites and gRNA7 (ACAGCAACATTACAATAGAGAGG) targeting the *MePLL_7516* translation start site (ATG). Construct *7516* set 2 was designed to carry gRNA4 and gRNA15 (AGGTTGAGAGATAATGGCAATGG) which targets the *MePLL_7516* 5'UTR. Construct 7568 set 1 was designed to carry gRNA 2 (GTTTGGCTCTATAAAAAGGTCGG) which targets both *MePLL* TAL14 EBE sites and gRNA 9 (ACATCTACATTACTACACAGAGG) which target the *MePLL_7568* 5'UTR. Construct *7568* set 2 was designed to carry gRNA2 and gRNA 14 (GAGGGAATCGGGGTAAATGGCGG) which target the *MePLL_7568* translation start site (ATG). Construct assembly was confirmed through colony PCR and Sanger sequencing. Constructs were transformed into LBA4404 Agrobacterium cells for cassava transformations. All construct sequences, maps, and Illumina reads are available in supplementary data.

4.4.2 Plant Materials and Growing Conditions:

Transgenic cassava lines expressing the CRISPR/Cas9 machinery and gRNAs were generated in the cassava cultivar TME419 through *Agrobacterium*-mediated transformation as described (Chauhun et al., 2015). Transgenic FEC cells were selected for resistance using 100mM paramomycin (275uL/L) on spread plates. 100mM paramomycin (450uL/L) was used to further select for resistant transgenic cells at stage 1, 2, and 3 plates. Transgenic FEC cells and eventual transgenic plantlets were maintained in tissue culture in conditions set to 28°C +/- 1°C, 75 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light; 16 hrs light / 8 hrs dark. Wildtype TME419 plants used for RT-PCR and bacterial growth

assays were clonally propagated from the woody stem of mature stock plants. Plants were kept in greenhouse conditions set to 28°C; 50% humidity; 16 hrs light / 8 hrs dark and 1000 W light fixtures that supplemented natural light levels below 400 W / m^2. Following bacterial infection assays, plants were kept in a post-treatment growth chamber with conditions set to 27°C; 50% humidity and 12 hrs light / 12 hrs dark or in a post-inoculation room set to 50% humidity, ambient room temperature, 12 hrs light / 12hr dark and 32 W light fixtures.

4.4.3 DNA Extraction and Transgenic Line Genotyping:

Leaf lobe samples from transgenic lines were collected from 2-3 individual plantlets and pooled into 2mL Eppendorf Safelock tubes with three disposable 3 mm Propper solid glass beads. The sample tubes were flash frozen in liquid nitrogen and ground to a fine powder using a QIAGEN TissueLyser II machine at 30hz for 3 minutes until the sample was fully homogenized. Genomic DNA was extracted using the Sigma GenElute Plant Genomic DNA Miniprep Kit. The region of interests for *MePLL_7516* or *7568* were amplified using primers designed to amplify each target specifically. The expected product size for *MePLL_7516* is 2.1 kb and for *MePLL_7568* it is 1.1 kb. All primers used in this study are provided as **Supplemental Table 4.1**. The PCR product was purified using the Qiagen QIAquick PCR Purification kit. The samples were sent for Sanger sequencing. Transgenic line trace files were compared to wildtype TME419 trace files and edits within and across each gRNA were identified using the Geneious bioinformatics tool.

4.4.4 Bacterial Inoculations:

Xanthomonas strains were struck from glycerol stocks onto NYG agar plates containing appropriate antibiotics. The strains used were Xpm668 (rifampicin 50 µg/ml) or Xpm668ΔTAL14 (suicide vector knockout, tetracycline 5 µg/ml, rifampicin 50 µg/ml) (Cohn and Bart et al., 2014). *Xanthomonas* strains were grown in a 30°C incubator for 2-3 days. Inoculum for each strain was made by transferring bacteria from plates into 10mM $MgCl_2$ using inoculation loops and brought up to a concentration of OD600 = 0.01 for bacterial growth and water-soaked lesion assays and OD600 = 1.0 for RT-PCR. Leaves on cassava plants were inoculated using a 1.0 mL needleless syringe. For each replicate assay, two cassava wildtype TME419 plants were used for inoculations, and four leaves were inoculated on each plant. On a single leaf lobe, one bacterial strain suspended in 10mM $MgCl_2$, was inoculated into three distinct spots along with a single lobe of three distinct mock inoculations of 10mM $MgCl_2$ alone). Injections were made into three leaf lobes per leaf so in total there were nine infiltrated sites per leaf.

4.4.5 RNA Extraction and RT-PCR Analysis:

RNA extractions and RT-PCR were performed on wildtype TME419 plants grown from stakes in the greenhouse. Cassava leaves were infiltrated with either 10 mM $MgCl_2$ (mock control) or *Xanthomonas* (*Xpm668* strains +/- TAL14) suspended in 10 mM $MgCl_2$ One leaf was selected from three individual plants and syringe infected with 3 infiltrated sites per treatment on separate leaf lobes. At 48HPI, leaf punches from the infiltrated sites were collected using a size 7 mm core borer, and technical reps were pooled into Eppendorf safelock tubes with 3mm glass beads. Samples were flash frozen in liquid nitrogen and ground using TissueLyser settings described

above. Total RNA was extracted from each sample using the Sigma Spectrum Plant Total RNA Kit. 0.5-1 µg of RNA was DNase treated using Promega RQ1 DNase enzyme and reverse transcribed into cDNA using Thermo Fisher Scientific SuperScript III Reverse Transcriptase. RT-PCR was performed on each sample using primers specific to each *MePLL* and to cassava actin as a constitutively expressed control. RT-PCR results were analyzed for TAL14-mediated induction or over expression of each *MePLL*.

4.4.6 Bacterial Growth Assay:

Cassava leaves were infiltrated with either 10 mM $MgCl_2$ (mock control) or *Xanthomonas* (*Xpm668* strains +/- TAL14) suspended in 10 mM $MgCl_2$ as described above. Leaf punch samples were taken at the site of infiltration using a 0.5 cm core borer (size 2) at 0- , 2-, 4-, 6- , and 9 days post-inoculation. For day-0 samples, infiltrated spots were allowed to dry down prior to processing. Individual leaf punches were transferred to 2 mL Eppendorf Safelock tubes with 200 uL of 10 mM $MgCl_2$ and three disposable 3 mm Propper solid glass beads. Samples were ground with a Qiagen Tissuelyzer at 28 hZ for 3 minutes. 200 uL of the ground sample was transferred to the first column of a labeled 96-well plate. Serial dilutions were performed by transferring 20 uL of the non-dilute sample (10^{1}) to the next well containing 180 ul of 10 mM $MgCl_2$. Samples were serially diluted to 10^{4} for day 0, 10^{6} for day 2 , 4, and 6, and to 10^{8} for day 8. Ten ul of each serial dilution was pipetted and spread onto labeled quadrants of an NYG plate with cycloheximide, and the appropriate antibiotics for the infiltrated bacterial strain. Plates were incubated at 30°C for 2-3 days, and the number of colonies were counted. Colony Forming Units (CFU) reported in this manuscript were transformed by leaf disc sample area where area = 1.96^{-01}.

4.5 References

1. Boher B, Kpémoua KE, Nicole M, Luisetti J, Geiger J (1995) Ultrastructure of interactions between Cassava and Xanthomonas campestris pv. manihotis : cytochemistry of cellulose and pectin degradation in a susceptible cultivar. doi: 10.1094/phyto-85-777
2. Brencic A, Winans SC (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. Microbiol Mol Biol Rev 69: 155–194
3. Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, et al (2017) A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. Plant Cell 29: 1196–1217
4. Chang S-C, Deng W-L, Huang H-C, Chung K-R, Tzeng K-C (2016) Differential expression of pectolytic enzyme genes in Xanthomonas citri subsp. citri and demonstration that pectate lyase Pel3 is required for the formation of citrus canker. Microbiological Research 192: 1–10
5. Chauhan RD, Beyene G, Kalyaeva M, Fauquet CM, Taylor N (2015) Improvements in Agrobacterium-mediated transformation of cassava (Manihot esculenta Crantz) for large-scale production of transgenic plants. Plant Cell, Tissue and Organ Culture (PCTOC) 121: 591–603
6. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T, Staskawicz BJ (2014) Xanthomonas axonopodis virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. Mol Plant Microbe Interact 27: 1186–1198
7. Cohn M, Morbitzer R, Lahaye T, Staskawicz BJ (2016) Comparison of gene activation by two TAL effectors from Xanthomonas axonopodis pv. manihotis reveals candidate host susceptibility genes in cassava. Molecular Plant Pathology 17: 875–889
8. Keen NT, Dahlbeck D, Staskawicz B, Belser W (1984) Molecular cloning of pectate lyase genes from Erwinia chrysanthemi and their expression in Escherichia coli. J Bacteriol 159: 825–831
9. Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA (2016) A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. Plant J 88: 1058–1070
10. Koseoglou E, van der Wolf JM, Visser RGF, Bai Y (2021) Susceptibility reversed: modified plant susceptibility genes for resistance to bacteria. Trends in Plant Science S1360138521002053
11. Leng Y, Yang Y, Ren D, Huang L, Dai L, Wang Y, Chen L, Tu Z, Gao Y, Li X, et al (2017) A Rice PECTATE LYASE-LIKE Gene Is Required for Plant Growth and Leaf Senescence1[OPEN]. Plant Physiol 174: 1151–1166
12. Marın-Rodrıguez MC, Smith DL, Manning K, Orchard J, Seymour GB Pectate lyase gene expression and enzyme activity in ripening banana fruit. 7

13. Merelo P, Agustí J, Arbona V, Costa ML, Estornell LH, Gómez-Cadenas A, Coimbra S, Gómez MD, Pérez-Amador MA, Domingo C, et al (2017) Cell Wall Remodeling in Abscission Zone Cells during Ethylene-Promoted Fruit Abscission in Citrus. Frontiers in Plant Science 8:
14. Mohnen D (2008) Pectin structure and biosynthesis. Current Opinion in Plant Biology 11: 266–277
15. Müller C, Kües U, Schöpper C, Kharazipour A (2007) Natural Binders. pp 347–381
16. Müller K, Levesque-Tremblay G, Bartels S, Weitbrecht K, Wormit A, Usadel B, Haughn G, Kermode AR (2013) Demethylesterification of Cell Wall Pectins in Arabidopsis Plays a Role in Seed Germination. Plant Physiology 161: 305–316
17. Palusa SG, Golovkin M, Shin S-B, Richardson DN, Reddy ASN (2007) Organ-specific, developmental, hormonal and stress regulation of expression of putative pectate lyase genes in Arabidopsis. New Phytologist 174: 537–550
18. Prakash R, Hallett IC, Wong SF, Johnston SL, O'Donoghue EM, McAtee PA, Seal AG, Atkinson RG, Schröder R (2017) Cell separation in kiwifruit without development of a specialised detachment zone. BMC Plant Biol 17: 86
19. Robert F. Weaver (2012) Molecular Biology, 5th ed. McGraw-Hill
20. Soriano M, Diaz P, Pastor FIJ (2006) Pectate lyase C from Bacillus subtilis: a novel endo-cleaving enzyme with activity on highly methylated pectin. Microbiology (Reading) 152: 617–625
21. Sun L, van Nocker S (2010) Analysis of promoter activity of members of the PECTATE LYASE-LIKE (PLL) gene family in cell separation in Arabidopsis. BMC Plant Biol 10: 152
22. Taiz, L., Zeiger, E, Moller, I.M, Murphy, A. (2015) Plant Physiology and Development, 6th ed. Sinauer Associates, Sunderland, CT.
23. Thurn KK, Chatterjee AK (1987) PEL-C is the Major Pectate Lyase Produced by Erwinia Chrysanthemi (EC16) in Vitro and in Plant Tissue. In EL Civerolo, A Collmer, RE Davis, AG Gillaspie, eds, Plant Pathogenic Bacteria: Proceedings of the Sixth International Conference on Plant Pathogenic Bacteria, Maryland, June 2–7, 1985. Springer Netherlands, Dordrecht, pp 257–257
24. Uluisik S, Seymour GB (2020) Pectate lyases: Their role in plants and importance in fruit ripening. Food Chemistry 309: 125559
25. Vogel JP, Raab TK, Schiff C, Somerville SC (2002) PMR6, a Pectate Lyase–Like Gene Required for Powdery Mildew Susceptibility in Arabidopsis. The Plant Cell 14: 2095–2106
26. Wang H, Guo Y, Lv F, Zhu H, Wu S, Jiang Y, Li F, Zhou B, Guo W, Zhang T (2010) The essential role of GhPEL gene, encoding a pectate lyase, in cell wall loosening by depolymerization of the de-esterified pectin during fiber elongation in cotton. Plant Mol Biol 72: 397–406

27. Wing RA, Yamaguchi J, Larabell SK, Ursin VM, McCormick S (1990) Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen Erwinia. Plant Mol Biol 14: 17–28
28. Yadeta KA, J. Thomma BPH (2013) The xylem as battleground for plant hosts and vascular wilt pathogens. Front Plant Sci. doi: 10.3389/fpls.2013.00097
29. Yandell M, Ence D (2012) A beginner's guide to eukaryotic genome annotation. Nat Rev Genet 13: 329–342
30. Yang Y, Zhang Y, Li B, Yang X, Dong Y, Qiu D (2018) A Verticillium dahliae Pectate Lyase Induces Plant Immune Responses and Contributes to Virulence. Frontiers in Plant Science 9:
31. Yoder MD, Keen NT, Jurnak F (1993) New Domain Motif: the Structure of Pectate Lyase C, a Secreted Plant Virulence Factor. Science 260: 1503–1507
32. Zheng Y, Yan J, Wang S, Xu M, Huang K, Chen G, Ding Y (2018) Genome-wide identification of the pectate lyase-like (PLL) gene family and functional analysis of two PLL genes in rice. Molecular Genetics and Genomics 293: 1317–1331

4.6 Supplementary Information

Primer Number:	Sequence: (5'-3')	Description:
MePLL_7516 F genotyping	AACAAAGGAACAAGGGGGCA	F primer targeting the *MePLL_7516* promoter
MePLL_7516 R genotyping	TGCATACGCACAAGCCAGTA	R primer targeting downstream of *MePLL_7516 exon 1*
MePLL_7568 F genotyping	CGGACCTGACTTGACCCAAA	F primer targeting the *MePLL_7568* promoter
MePLL_7568 R genotyping	ACAGGGGAAGCTCTGTTTCG	R primer targeting downstream of *MePLL_7568 exon 1*
MePLL_7516 RT-PCR F	AGACAGCTGCAAGCCAATG	RT-PCR F *MePLL_7516* expression primer
MePLL_7516 RT-PCR R	GGCCAAATTTAATTGATTAGTAAG	RT-PCR R *MePLL_7516* expression primer
MePLL_7568 RT-PCR F	GGTGGAAACAAGTACAGGCAT	RT-PCR F *MePLL_7568* expression primer
MePLL_7568 RT-PCR R	TTGGCCAAATCTAATTCATCAG	RT-PCR R *MePLL_7568* expression primer
GTPb RT-PCR F	CCTCAAAGGCTGAGCCACAGA	RT-PCR GTPb housekeeping gene expression primer
GTPb RT-PCR R	GGGAGAAACAATACAGGCACCAATCAC	RT-PCR GTPb housekeeping gene expression primer
A_CmYLCV-F	TGCTCTTCGCGCTGGCAGACATACTGTCCCAC	CmYLCV primer for CRISPR/Cas9 construct assembly.
B_gRNA4-R	TCGTCTCCATAGAGCCAAACCTGCCTATACGGCAGTGAACCTG	gRNA4 R primer for CRISPR/Cas9 construct assembly.
C_gRNA4-F	TCGTCTCACTATAAAAAGGTGTTTTAGAGCTAGAAATAGC	gRNA4 F primer for CRISPR/Cas9 construct assembly.
D_gRNA7-R	TCGTCTCATAATGTTGCTGTCTGCCTATACGGCAGTGAAC	gRNA7 R primer for CRISPR/Cas9 construct assembly. .
C_gRNA7-F	TCGTCTCAATTACAATAGAGGTTTTAGAGCTAGAAATAGC	gRNA7 F primer for CRISPR/Cas9 construct assembly.
E_35Sterm-R	TGCTCTTCTGACCTGCCTATACGGCAGTGAAC	35S terminator primer for CRISPR/Cas9 construct assembly
D_gRNA15-R	TCGTCTCAATCTCTCAACCTCTGCCTATACGGCAGTGAAC	gRNA15 R primer for CRISPR/Cas9 construct assembly.
C_gRNA 15-F	TCGTCTCAAGATAATGGCAAGTTTTAGAGCTAGAAATAGC	gRNA15 F primer for CRISPR/Cas9 construct assembly.
B_gRNA2-R	TCGTCTCCATAGAGCCAAACCTGCCTATACGGCAGTGAACCTG	gRNA2 R primer for CRISPR/Cas9 construct assembly.
C_gRNA2-F	TCGTCTCACTATAAAAAGGTGTTTTAGAGCTAGAAATAGC	gRNA2 F primer for CRISPR/Cas9 construct assembly.
D_gRNA9-R	TCGTCTCATAATGTAGATGTCTGCCTATACGGCAGTGAAC	gRNA9 R primer for CRISPR/Cas9 construct assembly.
C_gRNA9-F	TCGTCTCAATTACTACACAGGTTTTAGAGCTAGAAATAGC	gRNA9 F primer for CRISPR/Cas9 construct assembly.
D_gRNA14-R	TCGTCTCACCCGATTCCCTCCTGCCTATACGGCAGTGAAC	gRNA14 R primer for CRISPR/Cas9 construct assembly.
C_gRNA14-F	TCGTCTCACGGGGTAAATGGGTTTTAGAGCTAGAAATAGC	gRNA14 F primer for CRISPR/Cas9 construct assembly.

Supplemental Table 4.1: *MePLL* Primer List
Table of all primers used in this study. The primer name/stock number (left), sequence (middle), and description (right) are provided.

Chapter 5: Conclusions and Future Directions

Phytopathogens have evolved various strategies to invade plant hosts, evade plant defenses, and capture resources needed to proliferate. Some pathogens in the *Xanthomonas* and *Ralstonia* genera have TAL effectors that commandeer the expression of host susceptibility genes and manipulate the plant cell to benefit pathogen proliferation. There are several unanswered questions about how susceptibility genes contribute to pathogen virulence. In my dissertation, I investigated cassava susceptibility to *Xanthomonas* induced bacterial blight (CBB) and had two main objectives. Previous work showed that XpmΔTAL20 mutants resulted in visibly reduced water-soaking symptoms post-infection compared to Xpm WT (Cohn and Bart, 2014). However, a more quantitative method was needed to delineate differences in water-soaked lesions. Therefore, I first developed a more rigorous method for assessing water-soaked lesion symptoms as a CBB disease severity measure. Additionally, I generated several mutants with edits to TAL20 and TAL14 EBE sites and/or S gene targets via CRISPR/Cas9 to determine if such modifications could reduce host disease vulnerability. This was a complementary approach to prior research which demonstrated that removal of these effectors reduced Xpm virulence *in planta* (Cohn and Bart, 2014). Overall, this research culminated in novel cassava mutants with reduced CBB susceptibility and the development of valuable tools to further study host-pathogen interaction in a vital food security crop. In this chapter, I will present the final conclusions and potential future directions for the work presented in this dissertation.

5.1 Water-soaking Analysis is a Reliable and Robust Measure of CBB Disease Severity

In chapter two, we developed image analysis tools to measure water-soaked lesions as a measure of cassava bacterial blight severity. Water-soaked lesions are a common symptom of bacterial foliar pathogens (Aung 2018). It was previously reported that an Xpm mutant with loss of the effector TAL20 exhibited visibly decreased water-soaking symptoms (Cohn and Bart et al., 2014). We developed a water-soaking analysis method in which cassava leaves were infected with Xpm or mock treatments, collected at different time-points, and imaged. The infected leaf images were processed using either the ImageJ or support vector machine (SVM) learning based-methods to quantify water-soaked lesion area and color intensity. We found that both image analysis methods allowed for accurate segmentation and differentiation between Xpm WT, XpmΔTAL20, and mock treatments. Additionally, we found that XpmΔTAL20 infected sites had water-soaked lesion areas and grayscale values that were significantly reduced compared to Xpm WT infected sites.

Future directions:

Our work resulted in successful methods to quantify cassava bacterial blight symptoms. In the future, this method can be used to screen for natural CBB disease resistance in cassava wild species or to evaluate disease severity in novel cassava mutants cultivated for improved tolerance to CBB. Work to better the machine learning classifier file and workflow could be beneficial. The

classifier file developed in this dissertation was created using the cassava cultivar TME419. For cultivars with differing leaf traits, a new classifier file that captures a new range of water-soaking phenotypes would be necessary. Additionally, water-soaked lesions are commonly seen in other foliar diseases caused by Xanthomonads and additional phytopathogens. The general framework for the machine learning or ImageJ based image analysis tools could be applied to other pathosystems. Lastly, while the machine learning tool allows for faster analysis of water-soaked lesions compared to the ImageJ method which requires manual segmentation of lesions in every image, the tool can be further automated to increase processing speed.

5.2 Editing *MeSWEET10a* reduces cassava susceptibility to bacterial blight

Prior work found that manipulating TAL effectors on the Xpm side reduced pathogen virulence *in planta* (Cohn and Bart et al., 2014). In chapter three, we used a dual gRNA CRISPR/Cas9 strategy to generate *MeSWEET10a* mutants and quantified disease severity following mutant plant infection with Xpm. One objective of this work was to determine if restricting S gene access on the host side could reduce cassava susceptibility and Xpm virulence. We recovered several *MeSWEET10a* mutant lines and used Sanger-sequencing to genotype them. We identified mutants with edits to the TAL20 EBE/TATA box and/or *MeSWEET10a* coding sequence. Mutants that lacked TAL20-mediated induction post-Xpm infection were identified using RT-PCR and select mutants were moved forward for disease phenotyping using bacterial growth and water-soaking analyses. As previously reported, we found that bacterial growth assays

were not always sensitive enough to show a robust difference in Xpm bacterial titer (Veley 2023). However, *MeSWEET10a* EBE, TATA box, or coding sequence mutants exhibited significantly reduced water-soaked lesions compared to wildtype cassava when infected with Xpm WT. This indicated that *MeSWEET10a* mutants have reduced CBB susceptibility compared to wildtype cassava.

We were also interested in determining how editing *MeSWEET10a* would impact normal plant development. In the leaves, *MeSWEET10a* is not normally expressed but is ectopically induced by TAL20. However, there is endogenous expression of the gene in cassava flowers (Veley et al. 2020). Therefore, we grew a *MeSWEET10a* mutant line in a tropical field environment similar to native cassava growing conditions and examined flower development. From qualitative analysis, we found that the *MeSWEET10a* mutant produced flowers that were phenotypically similar to wildtype plants. Furthermore, viable seed was produced from *MeSWEET10a* mutants crossed with wildtype cassava. Overall, these results show that editing the *MeSWEET10a* S gene and/or TAL20 EBE site is a viable strategy to reduce susceptibility to CBB. Moreover, this work produced tools that can be used to probe additional questions about the role of *MeSWEET10a* as an Xpm targeted S gene.

Future directions:

While SWEET genes are commonly targeted susceptibility genes in various plant-Xanthomonad pathosystems (Gupta et al., 2021), the exact function of SWEETs in pathogen virulence remains unknown. It is hypothesized that sugars exported into the apoplast following *MeSWEET10a* induction by TAL20 is used by Xpm as a carbon source or osmolyte. However,

further work is needed to validate either hypothesis. One plausible approach to investigating the role of sugar in Xpm virulence is to supplement Xpm inoculation with sucrose, glucose, or known osmolytes ($NaCl_2$, $MgCl_2$, mannitol, etc.) in *MeSWEET10a* mutant plants. I hypothesize that if exported sugars are used as an osmolyte, supplementing Xpm infection with sugar or other osmolytes would restore the water-soaking phenotype. To investigate the role of these sugars as a carbon source for Xpm, isotope tracing to measure the metabolic flux of ^{13}C labeled sucrose or glucose isotopes between Xpm and cassava could be explored.

Another outstanding question is if *MeSWEET10a* mutants have reduced disease susceptibility at later stages of the Xpm infection cycle. In my research, I focused on initial Xpm infection and used the early indicator of CBB, water-soaking, as a disease severity measure. However, Xpm is a vascular pathogen that eventually spreads throughout the plant. Additional work looking at downstream disease symptoms such as stem browning in Xpm infected *MeSWEET10a* mutants would determine if Xpm pathogenesis is impacted only in early stages or throughout all stages of infection. For example, sugar efflux may be needed for spread of Xpm *in planta* and perhaps in *MeSWEET10a* mutant cassava, Xpm is unable to spread throughout the vasculature. During my thesis, I completed preliminary work investigating this line of thinking by tracking bioluminescent-tagged Xpm:pLUX movement post infection in wildtype cassava, however additional assay optimization is needed (**Appendix 4**).

Further, while we did not identify obvious defects in *MeSWEET10a* mutant flowers, more examination is required to determine if there are any subtle impacts on flower development or function. For instance, microscopy comparing the structure of *MeSWEET10a* mutant and wildtype

cassava flowers could uncover differences in flower development, not visible by macro imaging alone. Additionally, a full field trial with more crosses of the *MeSWEET10a* mutants is needed to better quantify mutant plant ability to produce viable fruit and seed. Lastly, we recognize that editing one S gene alone may not be enough to significantly reduce cassava susceptibility to CBB in a field setting. Thus, the development of stacked edits at different S gene targets can be produced to further reduce CBB susceptibility.

5.3 *MePLLs* are Additional Targets for Reducing CBB Susceptibility

In chapter four, we studied the contribution of putative *MePLLs* in promoting cassava bacterial blight. Our first objective was to confirm that putative *MePLLs* shared sequence similarity with validated pectate lyases from the literature. Using multi-sequence alignment and the protein database, pfam, we found that MePLL sequences had the conserved domain pectate lyase C (Pel C). Next, we used RT-PCR and bacterial growth analyses to characterize virulence related phenotypes in the farmer preferred cassava cultivar, TME419. Results from these assays suggested that the *MePLLs* are involved in Xpm virulence and therefore, we used CRISPR/Cas9 to edit the *MePLLs* and develop mutants with potentially reduced CBB susceptibility. We recovered several *MePLL* transgenic lines and using Sanger-sequencing found mutants with edits to *MePLLs* coding sequence and/or TAL14 EBE site. These novel *MePLL*s mutants provide an opportunity to examine how host pectate lyases are hijacked by pathogens for virulence.

Future directions:

To uncover the impact of *MePLLs* in promoting CBB susceptibility and Xpm pathogenesis, additional work is required. First, qRT-PCR to confirm that Xpm effector TAL14 cannot induce *MePLL* expression in mutants is needed. Furthermore, mutants must be characterized for disease susceptibility following Xpm infection. It is possible that a difference in bacterial growth may not be captured between Xpm WT-infected *MePLL* mutants and wildtype plants. Therefore, additional disease phenotyping strategies may be required to identify differences between *MePLL* mutants and wildtype cassava infected with Xpm.

One hypothesis is that the *MePLLs* contribute to pathogen virulence by degrading pectate at host cell surfaces. By degrading the cell surface, Xpm would gain access to plant cell nutrients and thus continue disease progression. Therefore, measuring the amount of pectate or pectin concentrations in the plant cell walls at sites of Xpm infection may reveal differences between *MePLL* mutants and wildtype cassava. There are various methods available to examine and measure cell wall composition. For example, immunofluorescence microcopy can be used to visualize pectate in Xpm infected cassava with pectate specific antibodies such as JIM5 or LM19 (Knox et al, 2021 and Verhertbruggen et al.,2009). In wildtype cassava infected with XpmΔTAL14 or mock treatments, I would expect reduced immunofluorescence signal at mesophyll cell walls compared to wildtype cassava infected with Xpm WT. In the case of *MePLL* mutants infected with Xpm WT, I expect immunofluorescence signal similar to wildtype cassava infected with XpmΔTAL14 or mock treatments. One method to compare the concentration of pectate in Xpm infected *MePLL* mutants and wildtype cassava is through a dot blot analysis in which proteins from crude extracts can be semi-quantitatively measured. A previous study reported a dot blot

assay protocol to measure different types of pectin using a crude extraction method and varying antibodies (Barany et al., 2009). If such an approach were applied to *MePLL* mutant and wildtype cassava, I would expect decreased concentrations of pectate in wildtype cassava infected with Xpm WT compared to wildtype cassava infected with XpmΔTAL14 or mock treatments. However, I would predict similar pectate levels between Xpm WT-infected *MePLL* mutant cassava and wildtype cassava infected with XpmΔTAL14 or mock treatments.

Another hypothesis is that pectate lyases could degrade pectinaceous blockages that occur as a plant defense response in the vasculature. To investigate this hypothesis, the pectate lyase activity of *MePLLs* could be functionally validated using *in vitro* techniques. For instance, *MePLLs* could be cloned and expressed in *E. coli* or yeast cells. Crude extracts of the *MePLLs* from cell cultures could then be tested for pectate lyase activity using sodium polygalacturonate as a substrate. If PLL cleavage of sodium polygalacturonate occurs, it results in oligogalacturonic acid which can be measured at 235 nm on a plate reader (Keen 1984). I would expect pectate cleavage to occur with both *MePLLs* indicating functional pectate lyase activity. To determine if *MePLLs* degrade pectin *in vivo*, the amount of pectate could be measured post Xpm infection in *MePLL* mutants and wildtype cassava in the vasculature. Moreover, if *MePLLs* degrade pectin blockages in vasculature to allow for pathogen spread, this could be examined using *in planta* tracking assays of Luciferase-luciferin tagged Xpm with or without TAL14 infected in *MePLL* mutants and wildtype cassava. In the case where *MePLLs* expression cannot be induced because Xpm lacks TAL14 or because of edits to the *MePLLs*, I anticipate restricted pathogen spread in the vasculature compared to wildtype cassava infected with Xpm WT.

5.4 References

1. Aung K, Jiang Y, He SY (2018) The role of water in plant–microbe interactions. The Plant Journal 93: 771–780
2. Bárány I, Fadón B, Risueño MC, Testillano PS (2010) Cell wall components and pectin esterification levels as markers of proliferation and differentiation events during pollen development and pollen embryogenesis in Capsicum annuum L. J Exp Bot 61: 1159–1175
3. Cohn M, Bart RS, Shybut M, Dahlbeck D, Gomez M, Morbitzer R, Hou B-H, Frommer WB, Lahaye T, Staskawicz BJ (2014) Xanthomonas axonopodis virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. Mol Plant Microbe Interact 27: 1186–1198
4. Gupta PK, Balyan HS, Gautam T (2021) SWEET genes and TAL effectors for disease resistance in plants: Present status and future prospects. Mol Plant Pathol 22: 1014–1026
5. Keen NT, Dahlbeck D, Staskawicz B, Belser W (1984) Molecular cloning of pectate lyase genes from Erwinia chrysanthemi and their expression in Escherichia coli. J Bacteriol 159: 825–831
6. Knox JP, Linstead PaulJ, King J, Cooper C, Roberts K (1990) Pectin esterification is spatially regulated both within cell walls and between developing tissues of root apices. Planta. doi: 10.1007/BF00193004
7. Veley KM, Okwuonu I, Jensen G, Yoder M, Taylor NJ, Meyers BC, Bart RS (2021) Gene tagging via CRISPR-mediated homology-directed repair in cassava. G3 (Bethesda) 11: jkab028
8. Verhertbruggen Y, Marcus SE, Haeger A, Ordaz-Ortiz JJ, Knox JP (2009) An extended set of monoclonal antibodies to pectic homogalacturonan. Carbohydrate Research 344: 1858–1862

Appendix

A.1 Beyond the Bench: A Word on Graduate Student Development and Creating Supportive Environments in Academia for Black Scientists to Thrive

A1.1:

Through traditional graduate school training, students gain many valuable skills. For example, a trainee can hone problem-solving, critical thinking, and technical skills. I personally became a graduate student to gain the expertise, experiences, and training needed to further develop my skillset as a scientist and to conduct research that would ultimately improve crop production and human lives. In addition to independent lab research, I intentionally sought experiences in collaboration, communication, outreach, and mentorship which were both professionally and personally cherished. In this appendix section, I would like to highlight some of the activities and additional roles I have served during my time in graduate school because I believe they are equally important training opportunities.

Collaborating with others is a critical skill for any scientist because science is a joint effort. One way I was able to do this was by managing cassava transformation needs for researchers in my lab during the COVID-19 shutdown from mid-May through August 2020. I completed work transforming cassava with constructs for four different projects and in total, I helped generate 75 different transformed lines. During this time, I improved my organizational and technical skills and provided regular written and oral updates to other researchers about transformation progress and results.

Secondly, I have a vested interest in science communication to share the experiences of researchers, establish an understanding among scientists and non-scientists, and build a scientifically literate society. During my time in graduate school, I served as the Co-founder and Executive Producer of a student-led podcast called GradCast. With the GradCast team, I produced monthly themed episodes that shared the stories, research, and life experiences of graduate and professional students from diverse areas of study. Additionally, I served as the Science Communication Director for a student group that Promotes Science Policy, Education, and Research (ProSPER). With ProSPER, I organized science communication workshop and events. I was also selected as a Danforth Plant Science Committee for Scientific Training and mentoring Co-Chair. In this role, I was able to advocate for the needs of trainees at the Danforth center and organize professional development and training opportunities including R coding workshops and career networking events. These experiences allowed me to enhance my leadership and communication skills while also helping to provide others with similar training opportunities.

Additionally, I prioritized outreach and mentorship during my graduate career. As an African American woman and first-generation college graduate, I am especially invested in mentoring underrepresented students because I recognize the added barriers that students from such backgrounds experience. Thus, I volunteered in multiple outreach programs with a specific focus on reaching students from underrepresented backgrounds including Black girls do stem (BGDS), the young scientists' program (YSP), and the science education partnership award program (SEPA). As a mentor, I trained lab technicians and rotating graduate students on different techniques including tissue culture maintenance, cassava infection, and molecular bench work. I

also served as the primary mentor to both an undergraduate REU student and High School intern. Overall, these opportunities allowed me to share the knowledge I have gained as a graduate student, train other scientists, and enhance my mentorship skills. Lastly, I believe it is imperative that we improve diversity and inclusion in science and create spaces in which people from all backgrounds can thrive in academic spaces. Along with other Black Scientists, I shared my perspective on this in a Letter to the Plant Cell editor highlighted below.

A1.2: Plant Cell Letter to the Editor

The complete article for work presented this section is available in the plant cell journal:

Kevin L. Cox, **Kiona R. Elliott,** Taylor M. Harris, Creating supportive environments in academia for Black scientists to thrive, *The Plant Cell*, Volume 33, Issue 7, July 2021, Pages 2112–2115, https://doi-org.libproxy.wustl.edu/10.1093/plcell/koab125

Dear Editor,

Plants are resilient organisms that manage to grow despite the dangers of diseases and stresses. One reason they are able to do so is their complex signaling pathways (Lamers et al., 2020). However, even their genetic compositions will not be enough if plants are not grown in an optimal environment. Additionally, there is no "universal environment," as each plant species has its own optimal environment that allows them to grow and thrive. This is something well-known in plant biology. But something that is not as well-known is how this principle also applies to people. For people to thrive and realize their full potential, they also need to be in a fostering environment/culture that supports them and allows them to grow (Montgomery, 2020b, 2021).

Unfortunately, this very often is not the case with Black students/scientists studying in the plant sciences. Black students/scientists are not only underrepresented in this area of study, but also the lack of an appropriate support system in universities/institutions discourages them from remaining in the sciences since they may feel that "they do not belong" (Hinton et al., 2020). Thus, this "leaky Science, Technology, Engineering and Mathematics (STEM) pipeline" appears to be a result of the lack of diversity inclusion (Asai, 2020). While events such as #BlackBotantistsWeek and #ShutDownSTEM have amplified the voices of Black plant scientists and helped people of noncolor to educate themselves on the awareness of racial discrimination in STEM (Mallenbaum, 2020), more actions need to be implemented by universities/institutions to create supportive environments to keep Black scientists in STEM fields. Here we discuss some of the problematic areas with current environments at academic institutions and provide actions that can be implemented to improve academic cultures across different levels of early career stages; undergraduate, graduate school, and postdoctoral tenure. We share some of our personal experiences as early career Black scientists and highlight some of the positive experiences that have benefited us in our academic journeys. Through this letter, we will highlight the importance of creating communities to prevent social isolation, promote active mentoring, and develop nurturing environments for Black scientists to thrive in academia and to develop successful independent careers.

Undergraduate perspective written by Taylor Harris:

Undergraduate years are an important time for student exposure to opportunities in science. The experience students gain from intense courses and research can make or break their interests and will shape their outlook on their future. While students of many different backgrounds are equally likely to pursue a major in STEM degrees, Black and Latinx students are more likely to leave STEM majors compared to their peers (Riegle-Crumb et al., 2019). Reasons for this difference may involve, but are not limited to, availability of resources and the difficulty of establishing a sense of identity and belonging within science as a person of color, considering the lack of representation. Thus, creating environments that support student-belonging and retention in science is crucial at the undergraduate level.

Having attended a Historically Black College or University (HBCU) as an undergraduate, I developed an intense curiosity for science in a welcoming environment that helped me build identity and confidence to pursue a STEM career. Though HBCUs make up only 3% of the country's colleges and universities, they contribute to 16% of Black STEM degrees and 25% of Black science and engineering doctorate recipients graduated from HBCUs (National Center for Science and Engineering Statistics, 2019; National Science Board, 2019). As an undergraduate, I had mentors who created a safe, embracing culture for my professional development; in these safe spaces, I never needed to minimize who I was and could bring my full-self to endeavors in science. I also benefited from having access to peers and scientists from diverse backgrounds which helped me envision myself pursuing STEM. In addition to learning in a welcoming and embracing culture, I participated in undergraduate STEM training programs designed to support and retain students from underrepresented backgrounds in science such as the NIH Maximizing Access to Research

Careers Undergraduate Student Training in Academic Research program and the NSF Tennessee Louis Stokes Alliance for Minority Participation program. Other than being fully submerged in Black culture, there are some simple, intimate, and practical approaches to fostering inclusive environments that can be achieved in other spaces and within the plant science community.

Institutions should consider shifting their learning cultures to one that encourages students to learn and engage with others about science fearlessly. This might involve being creative and rethinking how we typically approach teaching and practicing science. This has the potential to build students' confidence and allows them better see how they can *fit within* and belong in science, rather than trying to minimize who they are in an effort to *fit into* science. For example, if you have new students of color join your lab, be open to welcoming unique/alternative ways for them to give presentations at lab meetings and/or encourage the students' own ideas that might help them (and others) draw better understanding on specific topics. Besides having lab meetings only in the format of a formal seminar, other formats that encourage individual participation could be explored. This might include discussing publications with interesting findings or new methods that can be used for their experiments, having small group discussion on research progress/troubleshooting, or any other alternative method that better supports student needs. Second, institutions should consider programs that support, train, and provide opportunities for students of color to develop their skills. These can be internally designed and orchestrated programs or externally funded programs. With programs centered on supporting students of color, students can be exposed to different disciplines, including plant science, and be provided unique learning and training opportunities to both encourage and increase their participation. Lastly, it

would serve universities well to build a pipeline that connects scientists and trainees of color in STEM to undergraduates. For example, if an institution has a STEM training program in place, there can be opportunities for scientists who are further along in their career to engage with students and serve as role models. This too can help students visualize themselves in the scientific space long-term and help create a sense of belonging in science.

Graduate School perspective written by Kiona Elliott:

Graduate school is a time dedicated to expanding general knowledge in one's field and further developing technical and professional skills. While being a graduate student is a wonderful and exploratory experience, it can also be a difficult one. A survey of over two thousand graduate students found that "students were more than six times as likely to report experiences of depression and anxiety compared to the general population" (Evans et al., 2018). The hardships of graduate school are often further amplified for Black graduate students due to the unique barriers they may face. These barriers can include experiences such as racial and cultural aggressions and isolation due to lack of diversity and inclusion in academic institutions (Blanchard, 2018). Academic leaders often recognize and accept the need for resources, such as mental health counselors, for student success, well-being, and retention in graduate education. Similarly, it is critical that we recognize the needs of Black graduate students and invest in resources that contribute to their advancement.

When I began my search for graduate schools, I looked not only for interesting research opportunities, but also for programs that had tangible support systems in place for underrepresented minorities (URMs) in STEM. The support structures I sought included:

Diversity, Inclusion, and Equity (DEI) centers, URM training programs, and associations dedicated to celebrating student identities. DEI centers can help institutions promote and increase diversity by analyzing the needs related to DEI within the institution, and by developing/executing strategic plans that address them. For example, DEI centers can cultivate training sessions that equip faculty to identify and address microaggressions and train them to provide culturally sensitive mentorship for students. It is important to note that DEI centers should be adequately funded, staffed, and assessed in order for them to advance beyond performative allyship and to ensure they provide meaningful support for Black graduate students and other trainees (U.S. Department of Education, 2016). Additionally, training programs specifically designed to help URM students can provide meaningful advising, mentorship, and professional development opportunities. One example of such a training program is the Initiative to Maximize Student Development (IMSD) program. IMSD programs are NIH funded, can be found at various institutions across the nation, and strive to increase the number of scientists earning a Ph.D. who identify as URM (National Institute of General Medical Sciences, 2021). Last, student-led associations (e.g., Black graduate student associations) that celebrate student cultural identities are valuable because they provide community for students to connect with others who share similar backgrounds and experiences. Furthermore, such associations can serve to increase institutional awareness of the needs of students by serving as a unified platform for students to communicate concerns to leadership. These support structures will not only be beneficial for Black students, but they will also provide important resources for non-Black students/faculty to learn about the needs and views of their Black colleagues and a community to interact and connect with.

As a current Ph.D. candidate, I have benefited from the previously described resources and they have aided in my training by providing me with tools, mentorship, and community. However, this is sadly not true for all Black scientists pursuing graduate education. In order for institutions to create an environment that supports *and* retains Black graduate scientists the first step is to listen to their needs, and the second is to establish well-designed programs that nurture their talents.

Postdoctoral fellow perspective written by Kevin Cox:

Ph.D. students that graduate and transition into postdoctoral positions enter a critical phase in their early careers. In this period, postdoctoral scientists take their specialized knowledge and experience to fully craft themselves as independent researchers. However, since generally there is a lack of cohorts, organizations, or offices tailored specifically for postdocs, it is easy to fall into a sense of isolation. This feeling is greatly enhanced for Black postdoctoral trainees, as they are very likely to be the only Black postdoc in their mentor's lab or even their department (National Science Foundation, 2019). After completing graduate school, I feared that I would succumb to this isolation upon starting my postdoctoral training and that my productivity would be affected. While I have been fortunate and thankful to be in a welcoming lab with an active postdoctoral mentor and an institution that has a supportive environment that gives me a sense of "belonging" and allows me to develop my career, I am well aware that this experience is not common. This needs to change if we're going to have more Black postdoctoral fellows become faculty members or independent researchers in industry.

Active mentoring could play a significant role in increasing the number of Black postdoctoral associates. Postdoctoral mentors need to have more of a "groundskeeping" mentality and less "gatekeeping" (Montgomery, 2020a). One activity to which this applies is recruiting. Mentors with a gatekeeping recruitment approach, for example, hire candidates that meet their requirements solely or principally based on the candidate's publication record. An example of a groundskeeping recruitment approach would involve mentors assisting interested candidates in securing funding, such as fellowships, to provide them with an opportunity to have a postdoctoral tenure in their lab. Mentors being intentional in assisting Black scientists to find postdoctoral funding is key, as it will encourage them to remain in the sciences to achieve their career goals. Another stage when the mentor can apply more groundskeeping actions is during the early training/development of their mentees. Because of the diverse programs in graduate school, postdoctoral trainees may have deficiencies in some fields of science (i.e., bioinformatics, genetics, etc.) that are needed for their projects. A typical gatekeeping mentality would be to just wait for the trainee to solve this problem on their own ("sink or swim"). Instead, mentors can help their trainees address their deficiencies by suggesting relevant workshops to attend or collaborators to work with. Giving trainees these resources will allow them to think critically without limitations and to remain motivated on their projects. To summarize, active mentors with groundskeeping mentalities can help to remove a number of obstacles and roadblocks for Black postdoctoral fellows so they can have fruitful postdoctoral tenures.

Institutions can also change their environments to make them more welcoming for Black postdocs. One approach is to ensure that there is a type of postdoctoral association in each

department/institution. This allows new postdoctoral trainees to interact immediately with their colleagues and have a community they can turn to for advice on experiments, campus/town life, etc. Another change that institutions can implement is creating and funding annual training grants that are geared specifically for individuals that are racially underrepresented in STEM (Hinton et al., 2020). Such training grants can provide key professional training, from scientific and grant writing to practicing for faculty interviews, which helps prepare early career investigators for their next career stage. Additionally, these training grants can encourage postdoctoral associates to form committees to oversee and advise their research. Changes in mentorship and infrastructure at research institutions will be critical to create a supportive environment for Black postdoctoral associates.

Conclusion:

It is clear that the current environments in academia are not working for many Black scientists in plant biology as well as other STEM fields. As described in this letter, recruiting more Black scientists into these early career stages is only part of the solution; the environment needs to be improved as well. Additionally, it is critical for institutions to actively engage in conversations with their Black students and scientists to understand how they are interacting with their home institutions and to identify the problematic areas in their academic environments. As a collective, we reflected on our experiences and recognized common themes both in the resources that have aided us in our career journeys and the resources we wish existed. To summarize, these themes are:

1. *Community building:* Creating a social network to combat isolation and provide trainees with a sense of belonging and support.
2. *Mentors and leadership*: Creating cognizant faculty and staff members through consistent training and awareness of trainee concerns.
3. *Nurturing environments*: Creating opportunities for trainee professional development through access to workshops, seminars, fellowships, and training programs.

If institutions actively incorporate these themes into structures, we can create fostering and inclusive environments that welcome, support, and embrace Black scientists. Only then can more Black scientists be retained in the plant sciences, allowing them to grow and thrive into successful careers.

A1.3: References

1. Asai, D.J. (2020). Race Matters. Cell 181: 754–757.
2. Blanchard, S.J. (2018). Black Graduate Students' Experiences of Stress and Coping.
3. Evans, T.M., Bira, L., Gastelum, J.B., Weiss, L.T., and Vanderford, N.L. (2018). Evidence for a mental health crisis in graduate education. Nat. Biotechnol. 36: 282–284.
4. Graduate Students and Postdoctorates in Science and Engineering - Public Use Data Files
5. Hinton, A.O., Jr et al. (2020). Patching the Leaks: Revitalizing and Reimagining the STEM Pipeline. Cell 183: 568–575.
6. Initiative for maximizing student development (IMSD) program (T32)
7. Lamers, J., van der Meer, T., and Testerink, C. (2020). How Plants Sense and Respond to Stressful Environments. Plant Physiol. 182: 1624–1635.

8. Mallenbaum, C. (2020). BlackBirdersWeek,# BlackInNeuro: Black scientists, physicians are using hashtags to uplift. USA Today. https://www. usatoday. com/story/life/2020/08/04/blackinneuro-blackinchem-can-hashtags-help-black-scientists-build-community-spotlight-excellence/5541431002.

9. Montgomery, B.L. (2020a). Academic Leadership: Gatekeeping or Groundskeeping? The Journal of Values-Based Leadership 13: 16.

10. Montgomery, B.L. (2020b). Planting Equity: Using What We Know to Cultivate Growth as a Plant Biology Community. Plant Cell 32: 3372–3375.

11. National Center for Science and Engineering Statistics (2019). Bachelor's degrees awarded by all institutions and by HBCUs to black U.S. citizens and permanent residents, by field: 2006–16.

12. National Science Board (2019). Science & Engineering Indicators Report.

13. Riegle-Crumb, C., King, B., Irizarry, Y. (2019). Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields. Educational Researcher 48:133-144.

14. U.S. Department of Education (2016). Advancing diversity and inclusion in higher education.

A.2 Results and Methods of Bacterial Growth and Water-soaking Analyses from the "Improving Cassava Bacterial Blight Resistance by Editing the Epigenome" Manuscript

The complete manuscript for work presented this section is available in the journal Nature Communications:

Veley KM, **Elliott K**, Jensen G, Zhong Z, Feng S, Yoder M, et al. (2022) Improving cassava bacterial blight resistance by editing the epigenome. *Nat Commun* **14**, 85 (2023). https://doi.org/10.1038/s41467-022-35675-7

A2: Results

For water-soaking analysis, DMS3-ZF lines and control plants were infected with Xam WT, Xam ΔTAL20, and mock (10mM $MgCl_2$) treatments and infected leaf samples were collected at 0 and 4 DPI. ImageJ based analysis of the water-soaked lesions was completed. Results showed that DMS3-ZF plants infected with Xam WT had significantly reduced water-soaked area and gray scale intensity compared to wildtype and ZF-only infected plants (**Figure A2.1**). For bacterial growth assays, DMS3-ZF and control plants were infected with Xam WT, XamΔTAL20, and mock (10mM $MgCl_2$) treatments and at 0 and 4 DPI, leaf punches were collected. No difference in bacterial growth was detected between DMS3-ZF plants compared to control plants (**Figure A2.2)**.

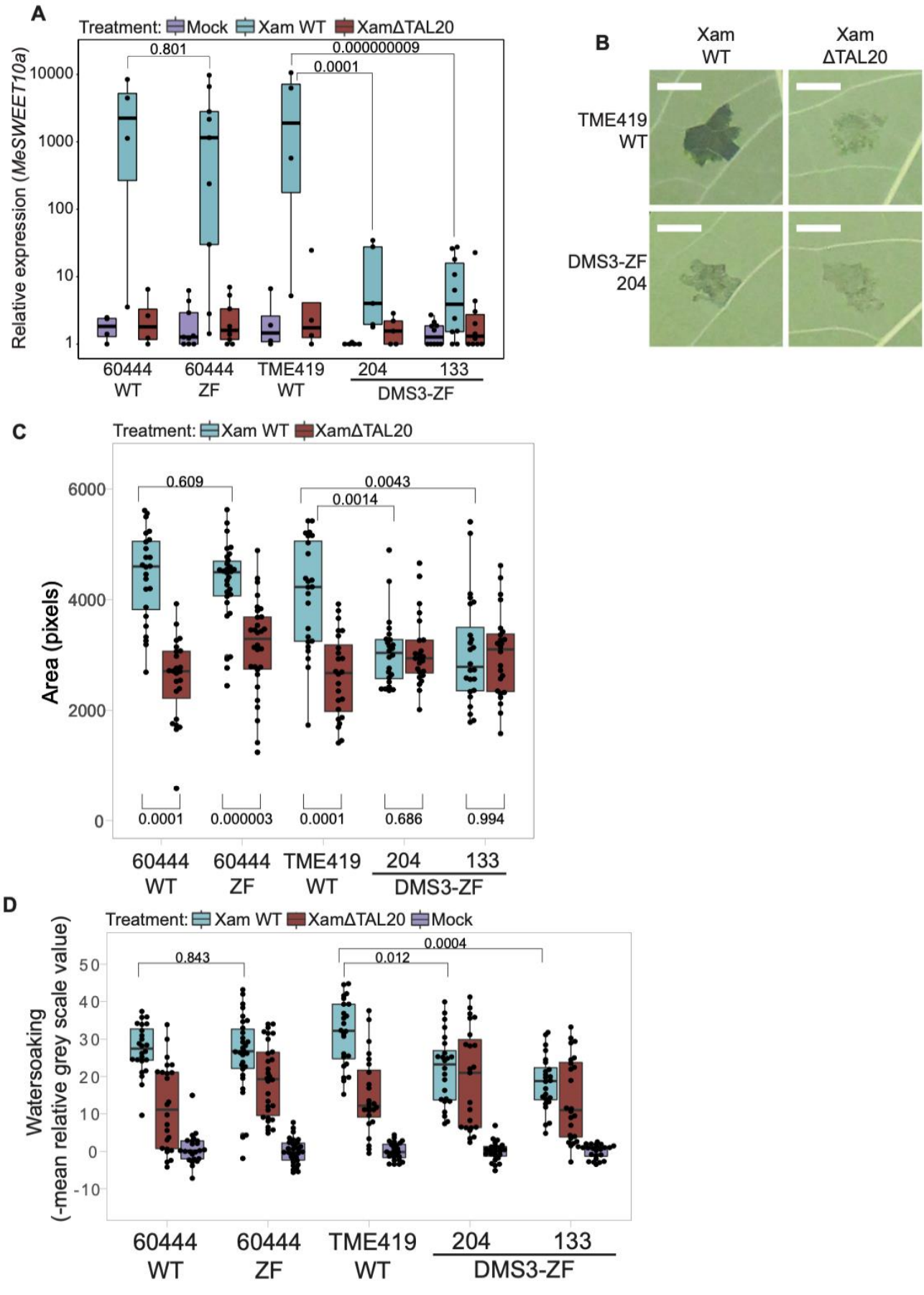

A
Treatment: Mock Xam WT XamΔTAL20
Relative expression (MeSWEET10a)
10000
1000
100
10
1
0.801
0.000000009
0.0001
60444 WT
60444 ZF
TME419 WT
204
133
DMS3-ZF
B
Xam WT
Xam ΔTAL20
TME419 WT
DMS3-ZF 204
C
Treatment: Xam WT XamΔTAL20
Area (pixels)
6000
4000
2000
0
0.609
0.0043
0.0014
0.0001
0.000003
0.0001
0.686
0.994
60444 WT
60444 ZF
TME419 WT
204
133
DMS3-ZF
D
Treatment: Xam WT XamΔTAL20 Mock
Watersoaking (-mean relative grey scale value)
50
40
30
20
10
0
-10
0.843
0.0004
0.012
60444 WT
60444 ZF
TME419 WT
204
133
DMS3-ZF

Figure A2.1 (Manuscript Figure 4): Effect of Methylation on CBB Disease Phenotypes in Cassava

MeSWEET10a expression (y-axis, log10 scale) in WT and transgenic lines as determined by RT-qPCR. The cassava genes GTPb (Manes.09G086600) and PP2A4 (Manes.09G039900) were used as internal controls. Boxes are colored according to Xanthomonas treatment. Biological replicates (black dots) included in each background (x-axis) are as follows: n = 4, 9, 4, 5, 10 examined over 4 independent experiments. The n included in each treatment group for each background are consistent. Horizontal black line within boxes indicates the value of the median while the box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range (1.5xIQR) from the 25th and 75th percentiles. Two-sided Welch's *t*-test *p*-values are noted above brackets within plot. (**B**) Representative images of water-soaking phenotype of leaves from TME419 WT and DMS3-ZF-expressing plants. Images were taken 4 days post-infection with either *Xam668* (XamWT) or a *Xam668* TAL20 deletion mutant (XamΔTAL20). Scale bar = 0.5 cm. (**C**) Observed area (pixels, y-axis) of water-soaking from images of Xam-infiltrated leaves (backgrounds, x-axis) 4 days post-infiltration. Images from (**B**) and **Supplementary Fig. 11** are included in dataset. Calculated *p*-values (Kolmogorov-Smirnov test) are shown above brackets within plot. (**D**) Intensity of water-soaking phenotype (y-axis) of region measured in panel **C**. The negative mean grey-scale value for the water-soaked region relative to the average of the mock-treated samples within the same leaf is reported (see methods for details). Calculated *p*-values (two-sided Kolmogorov-Smirnov test) are shown above brackets within plot. Both box plots: Biological replicates (black dots) included in each background (x-axis) are as follows: n = 24, 30, 24, 24, 24 examined over three independent experiments. Horizontal black line within boxes indicates the value of the median while the box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range (1.5xIQR) from the 25th and 75th percentiles.

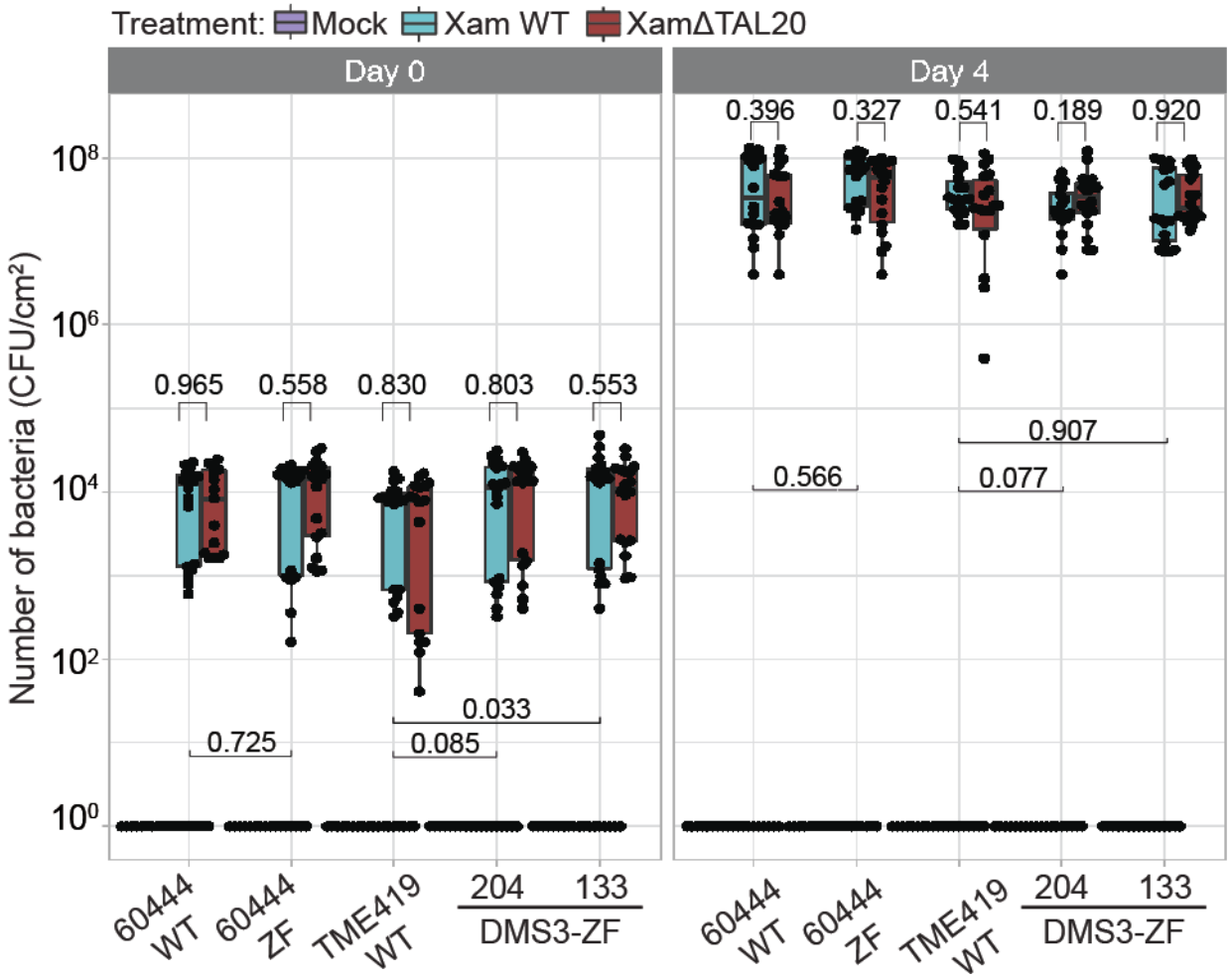


Figure A.2.2 (Manuscript Supplementary Figure 13): Bacterial Growth is Unaffected by Methylation at the EBE

Bacterial populations in leaves measured at day-0 (left) and day-4 (right) post-infiltration with Xam (treatments listed above plot). Mean colony forming units (CFU/cm^2, y-axis) are plotted per background tested (x-axis). Individual data points are represented as black dots analyzed across 3 independent experiments. For each background from left to right, Day 0 Mock, Xam WT, and XamΔTAL20 n = (18, 18, 15), n = (16,18,18), n = (18,18,17), n = (18,18,18), and n = (18,18,18). For each background from left to right, Day 4 Mock, Xam WT, and XamΔTAL20 n = (18,18,18), n = (18,18,18), n = (18,18,18), n = (18,16,18), and n = (16,18,18). Dots outside whiskers represent outliers. The horizontal line within the box represents the median sample value. The ends of the boxes represent the 3rd (Q3) and 1st (Q1) quartiles. The whiskers show values that are 1.5 times interquartile range (1.5xIQR) above and below Q1 and Q3. Results of statistical analyses (*p*-values, Student's *t*-test) comparing the difference between treatments within each background

(black text, above boxes) and the difference between Xam WT growth across different backgrounds (below boxes) are shown.

A2: Methodology

Bacterial inoculations

Xanthomonas strains were grown on plates containing necessary antibiotics for 2-3 days at 30°C. The *Xanthomonas* strains used and their antibiotic resistances are as follows: 1. *Xam668* (rifampicin 50 µg / ml) and 2. *Xam668ΔTAL20* (suicide vector knockout, tetracycline 5 µg / ml, rifampicin 50 µg / ml). Bacteria were scraped from the plates into 10 mM $MgCl_2$ to either OD_{600} = 1.0 (RT-qPCR) or OD_{600} = 0.01 (water-soaking) concentrations according to established lab protocols. Leaves (2-3 weeks after plants were transferred to a greenhouse) were inoculated with a 1.0 ml needleless syringe using one bacterial strain per lobe with three injection points on each of 2-3 leaves. After inoculation, plants were kept under fluorescent light (12 hr day / night light cycle) at 27°C for either 48 hours (RT-qPCR and ampBS-seq) or 4 days (water-soaking).

Water-soaking area and intensity quantification, image analysis

Cassava leaves were infiltrated with either 10 mM $MgCl_2$ alone (mock) or containing *Xanthomonas* (*Xam668* strains with and without TAL20) as described above and imaged using a Canon EOS Rebel T5i camera with a 15-85mm lens at 0- and 4-days post-inoculation. Images were grey balanced using a X-Rite Passport by estimating the saturation of the six grey chips and lowering the brightness accordingly. Gray corrected images were uploaded to FIJI[54] and

duplicated. In order to define the infected area, visible water-soaking spots were manually outlined on the duplicate image using the pencil tool (color: #ff00b6 and size 2). The outlined images were converted from RGB to LAB and split to obtain the A color channel. The A channel images were thresholded, converted to a mask and the mask for each spot was added to the ROI manager using the analyze particle tool. The ROI masks were applied to the original RGB grey corrected images. In order to define the background color of a given leaf, mock-infiltrated spots (no water-soaking) were added to the ROI manager using an arbitrarily sized rectangle selection tool consistently set to a W = 51 and H = 61. Area and grey scale mean data were obtained for each infiltrated spot using the FIJI measure tool. Grey balancing using the X-Rite Passport is a coarse adjustment and finer standardization is performed post-image processing using statistical methods. Grey value standardization was achieved by estimating the grand mean of all grey values in each image and centering those values to the grand mean of all images by using model residuals. Further, the residuals within each image were re-centered with respect to the mock treatment by subtracting the mean grey value of mock from all other features.

Bacterial growth assay

Cassava leaves were infiltrated with either 10 mM $MgCl_2$ alone (mock) or containing *Xanthomonas* (*Xam668* strains with and without TAL20) as described above. Leaf punch samples were taken at the site of infiltration using a 5 mm cork borer at 0- and 4-days post-inoculation. For day-0 samples, infiltrated spots were allowed to dry down prior to processing. Individual leaf punches were transferred to 2 mL Eppendorf Safelock tubes with 200 ul of 10 mM $MgCl_2$ and

three disposable 3 mm Propper solid glass beads. Samples were ground with a Qiagen Tissuelyzer at 28 hZ for 3 minutes. 200 ul of the ground sample were transferred to the first column of a labelled 96-well plate. Serial dilutions were performed by transferring 20 ul of the non-dilute sample (10^1) to the next well containing 180 ul of 10 mM $MgCl_2$. Samples were serially diluted to 10^4 for day-0 and 10^6 for day-4. 10 ul of each serial dilution was spread onto labelled quadrants of a NYG plate with cycloheximide and the appropriate antibiotics for the infiltrated bacterial strain. Plates were incubated at 30°C for 2-3 days, and the number of colonies were counted. Colony forming Units (CFU) reported in this manuscript were transformed by sample area (CFU/cm^2 where $cm^2 = 0.52$).

A.3 Testing Various Conditions for Bacterial Growth Assays

When comparing Xpm WT and XpmΔTAL20 growth over time *in planta*, we expected to see a significant decrease in XpmΔTAL20 bacterial titer compared to Xpm WT. In some growth assays, a subtle difference was observed and XpmΔTAL20 titer trended downward compared to Xpm WT. However, in several assays, there was no observed difference between Xpm WT and XpmΔTAL20 Colony Forming Units (CFUs). A study by Xin et al. found that different humidity levels impacted *P. syringae* growth *in planta*. We considered that the initial settings used in our growth assays did not allow for robust differences in Xpm WT and XpmΔTAL20 to be captured. Therefore, over the course of my thesis work, I tested various experimental conditions for bacterial growth assays in an attempt to enhance potential differences between Xpm WT and XpmΔTAL20 growth in wildtype cassava. This appendix section outlines the results of numerous growth assays with varied experimental conditions including different bacterial inoculum concentrations, humidity settings, and time-points sampled post inoculation with *Xanthonomas* strains.

Figure A.3.1: Growth Assay 1:

Mock, XpmΔTAL20, and Xpm WT treatments were infected in WT419 cassava at an OD600 of 0.01. Samples (N=3 at Day 0, N=6 at Day 4 and 6) were collected at 0, 4, and 6 DPI. No significant difference was observed between bacterial CFU/leaf disc area.

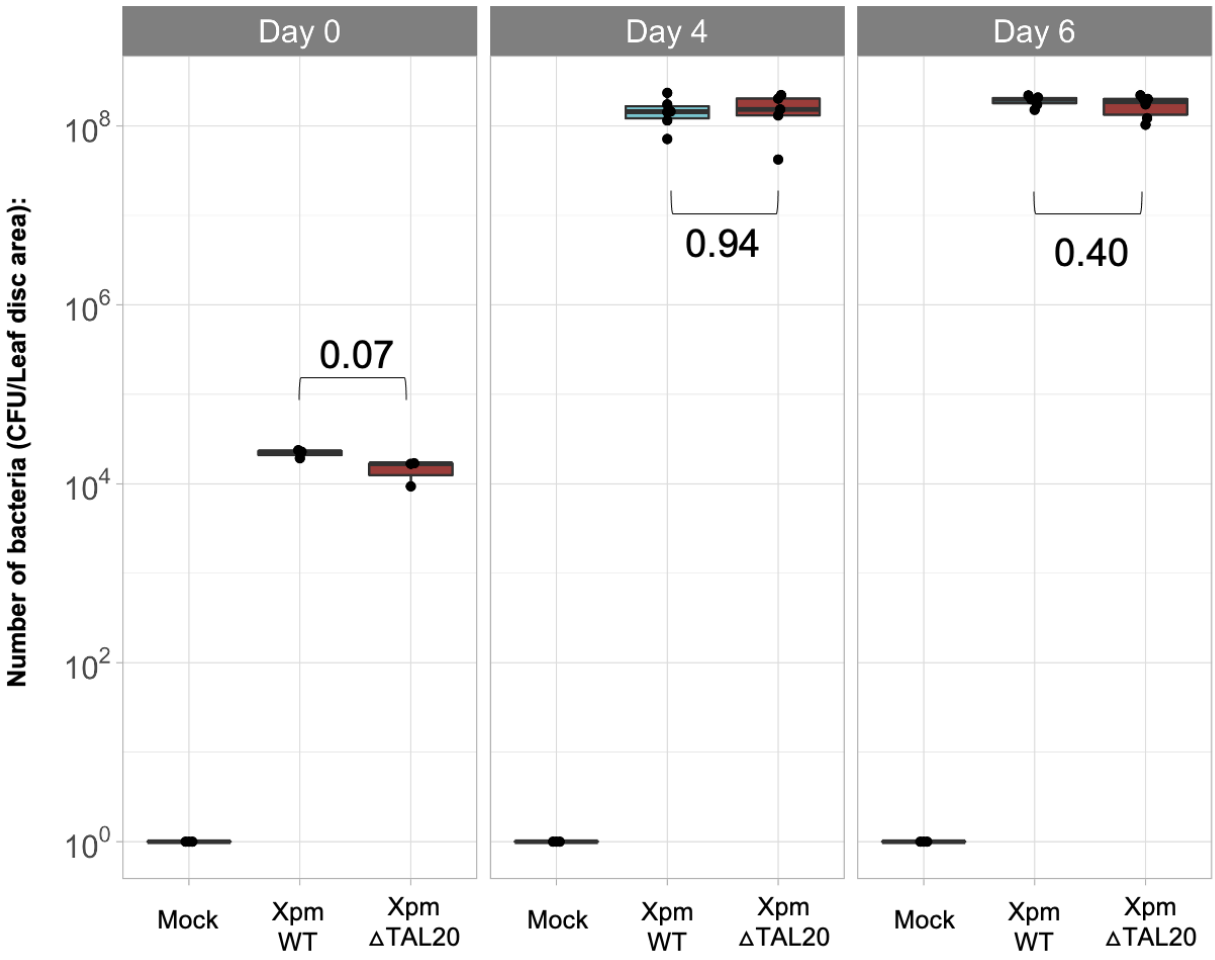
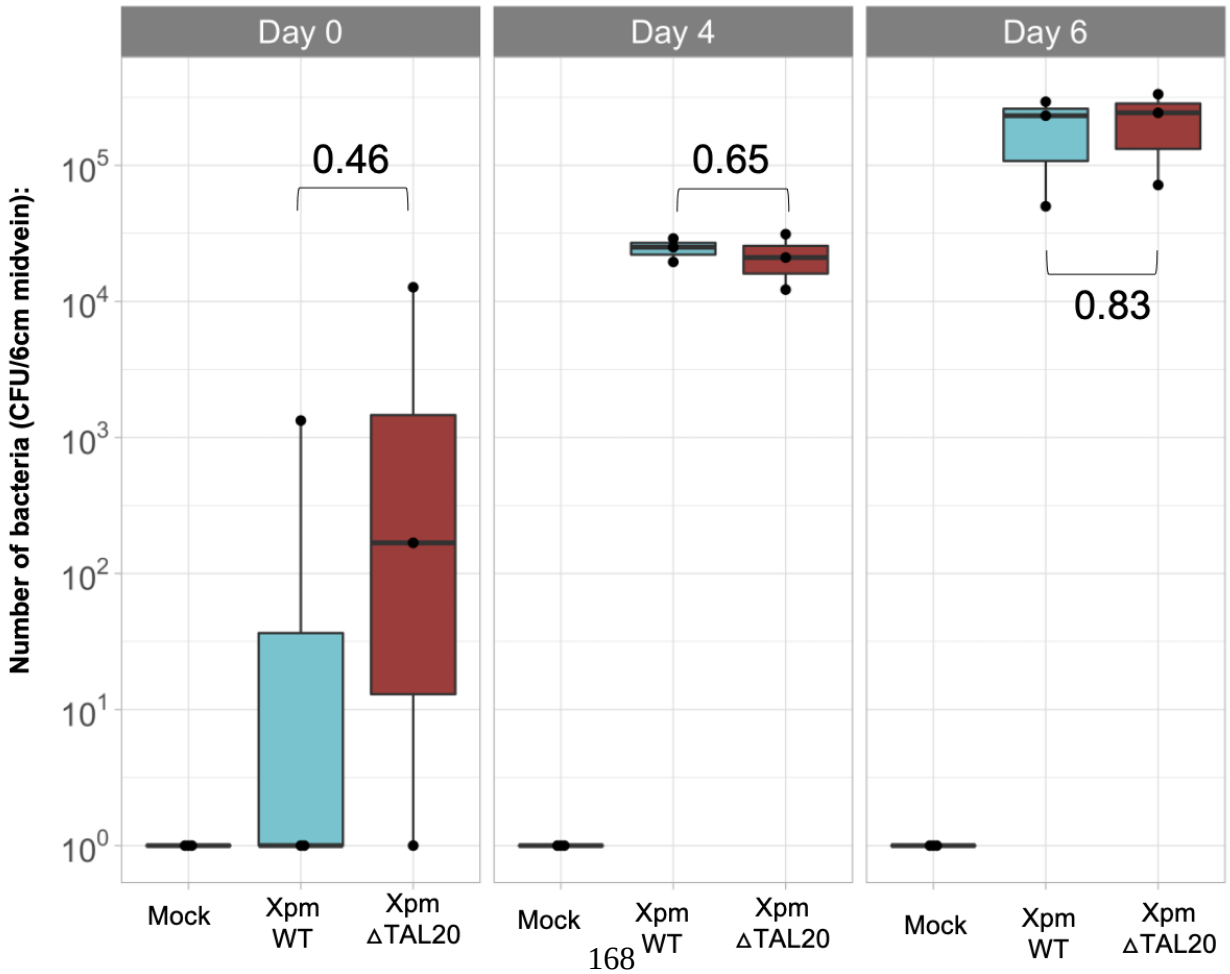

Day 0
Day 4
Day 6
Number of bacteria (CFU/Leaf disc area):
10^8
10^6
10^4
10^2
10^0
0.07
0.94
0.40
Mock
Xpm WT
Xpm ΔTAL20

Figure A.3.2: Growth Assay 2:

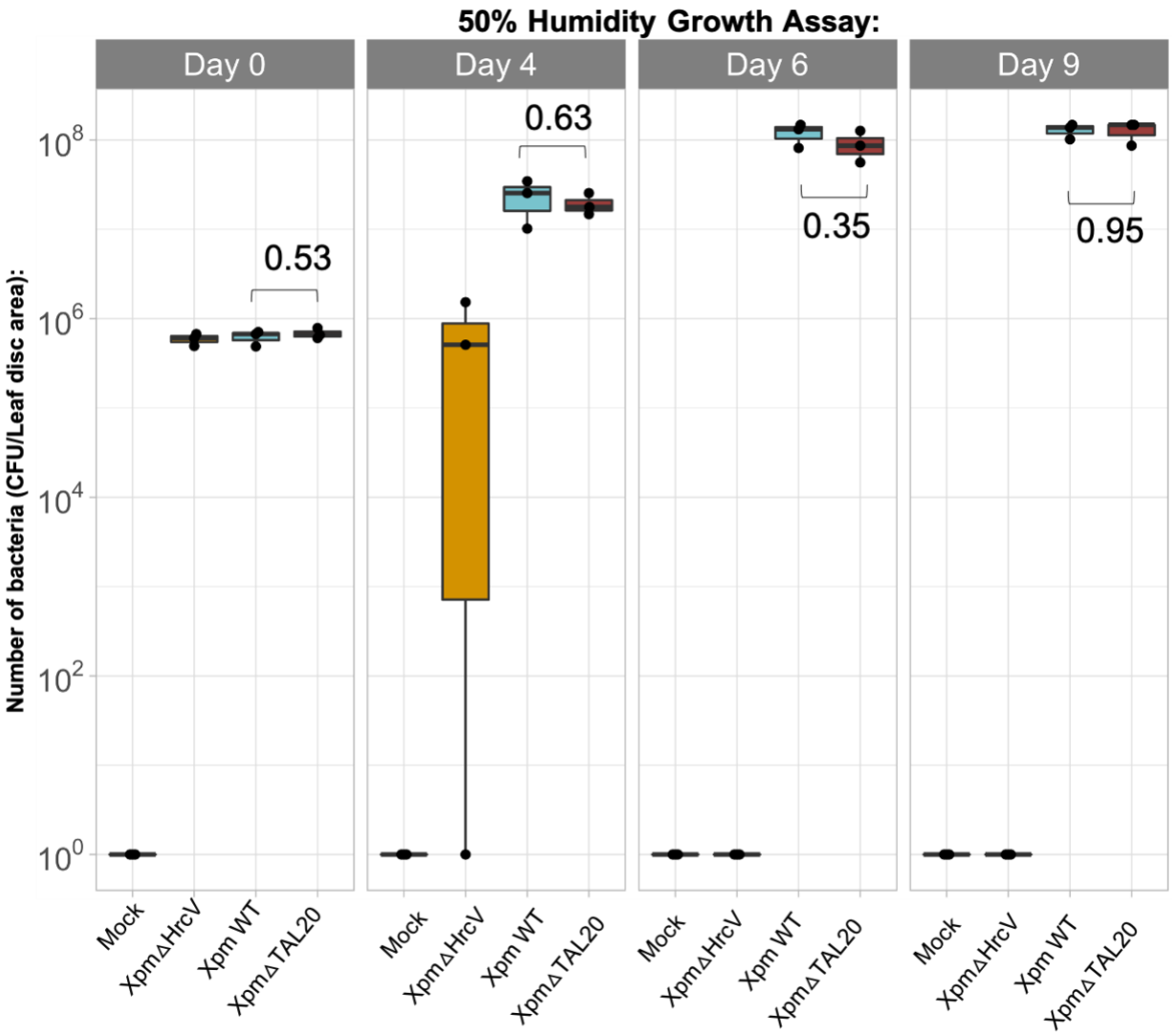
Mock, XpmΔTAL20, and Xpm WT treatments were infected into wildtype cassava using a midvein inoculation assay previously described (Cohn and Bart et al., 2014 and Cohn et al., 2016). 2mm holes were made at the midveins on the underside of the cassava leaves. Five ul drops of bacteria suspended in 10mM $MgCl_2$ (OD600=0.2) were added to hole and allowed to dry down for fifteen mins. At 0, 4, and 6 DPI, the cm midvein sections were collected, two 3cm sections were pooled per replicate for a total of N=3 per treatment. The samples ground in 200ul of 10mM



$MgCl_2$ and serial dilutions were performed. No significant difference was observed in bacterial titer.

Figure A.3.3: Growth Assay 3:

Next, I compared the impact of different humidity conditions on bacterial growth. An CFU assay comparing plants infected with mock, Xpm WT, and XpmΔTAL20 (OD600=0.01) treatments and stored in 50 percent (usual conditions) or 25 percent humidity was conducted. XpmΔHrcV (A T3SS mutant) was included as a control as decreased growth was expected in XpmΔHrcV versus Xpm WT infected plants. At 0, 4, 6, and 9 DPI, samples were collected (N=3 per treatment) and CFU/leaf disc area was calculated. In both 50 and 25 percent humidity conditions, no significant difference between Xpm WT and XpmΔTAL20 growth was observed. However, reduced XpmΔHrcV growth was observed as expected.


50% Humidity Growth Assay:
Day 0
Day 4
Day 6
Day 9
Number of bacteria (CFU/Leaf disc area):
10^8
10^6
10^4
10^2
10^0
0.53
0.63
0.35
0.95
Mock
XpmΔHrcV
Xpm WT
XpmΔTAL20

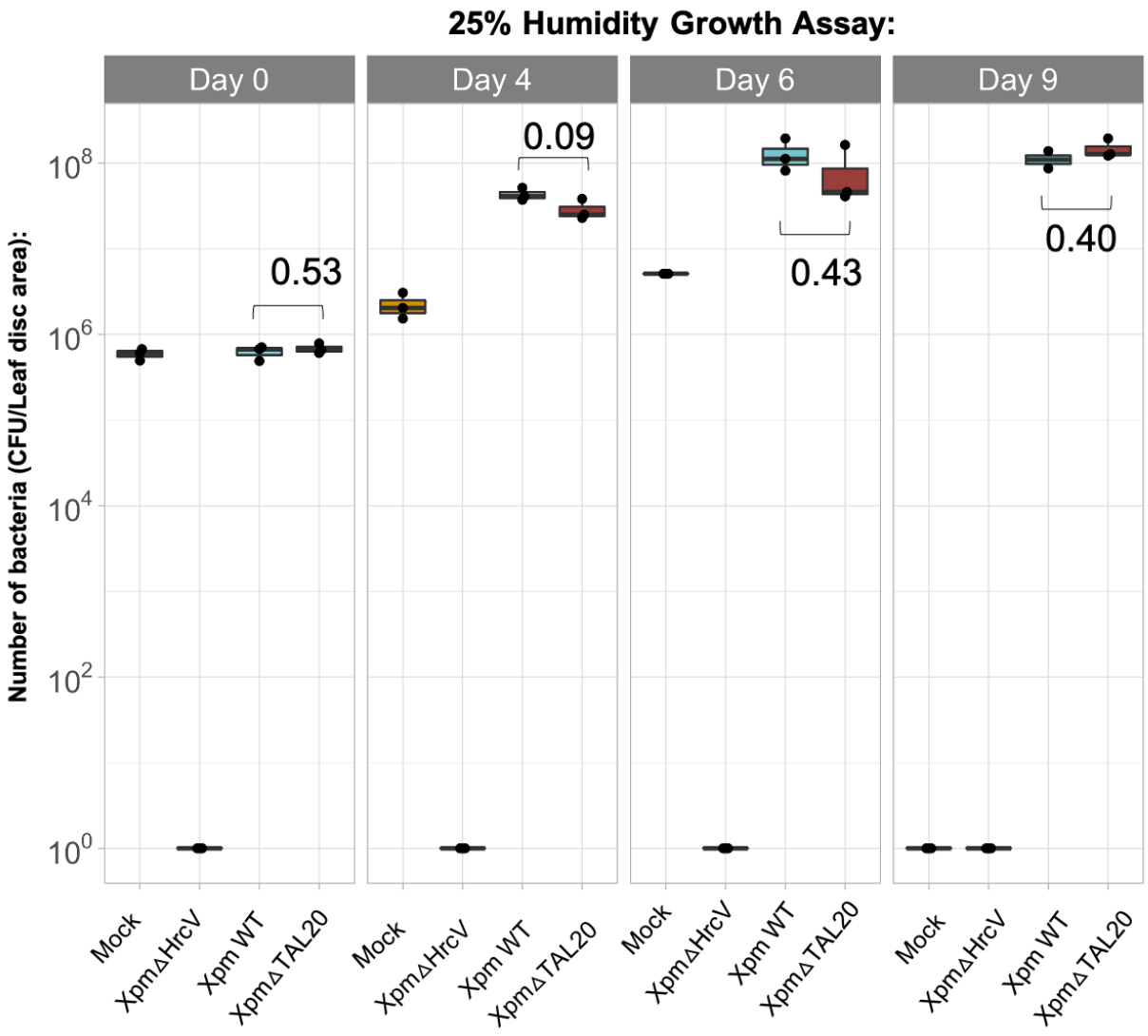
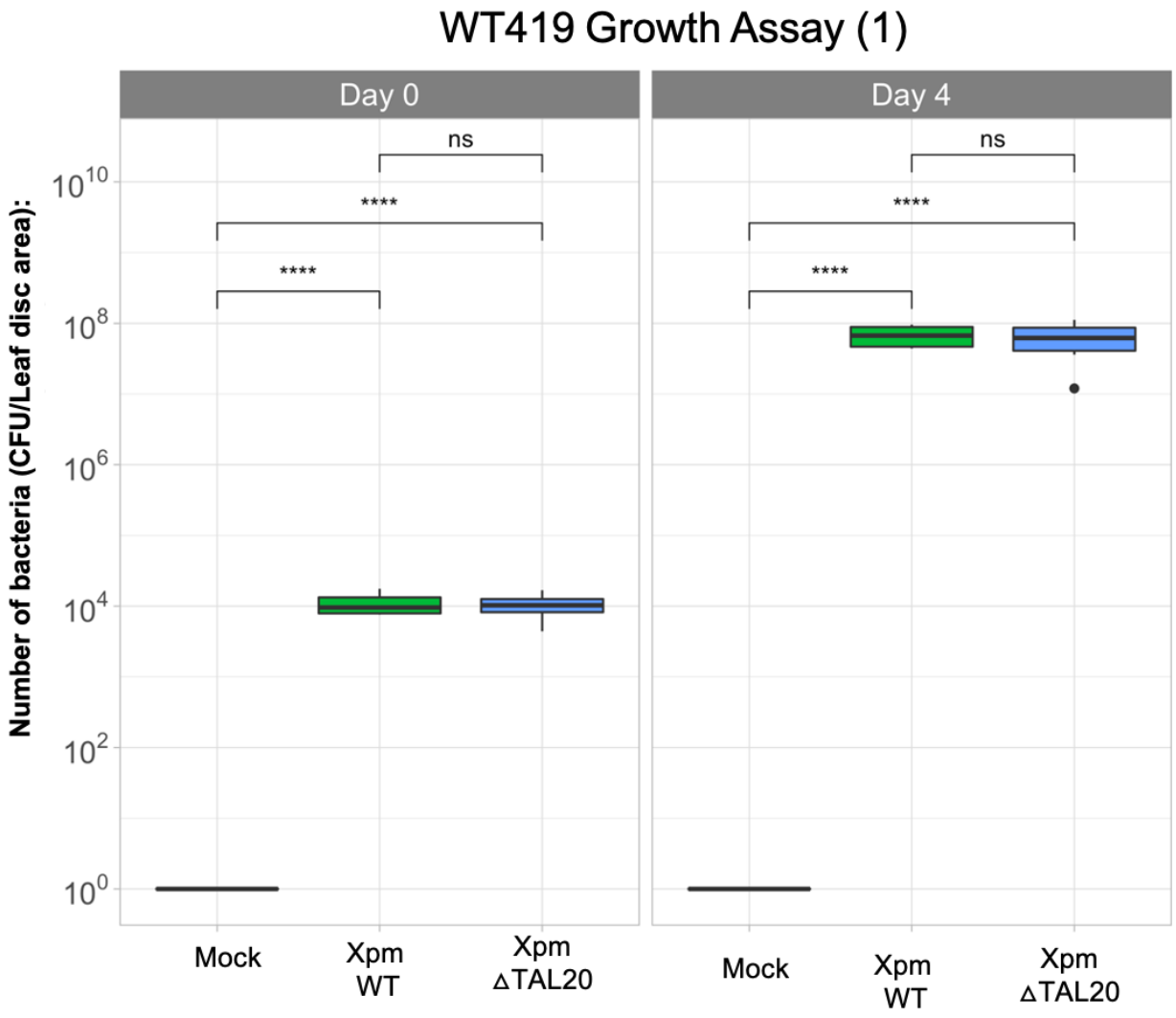
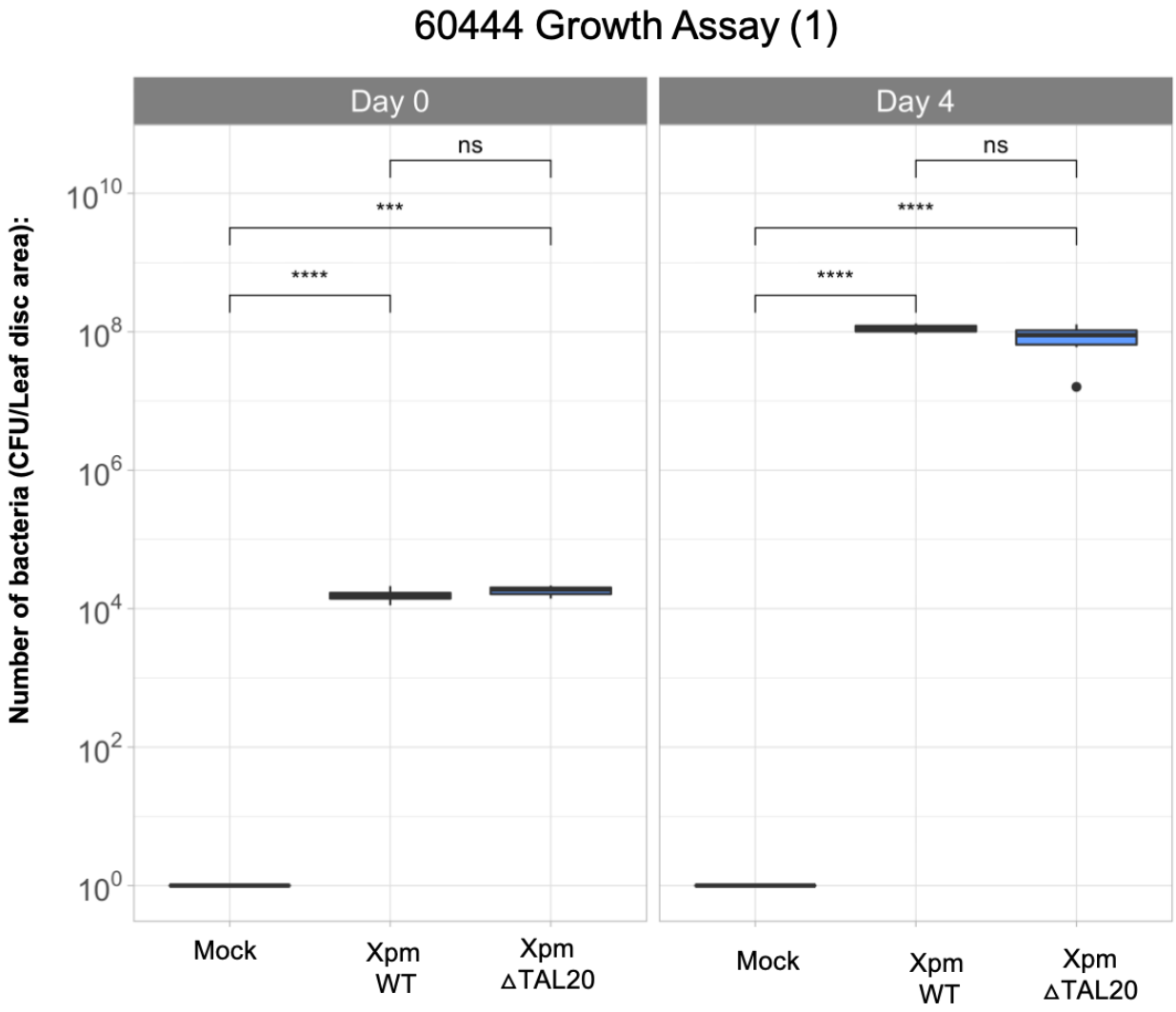
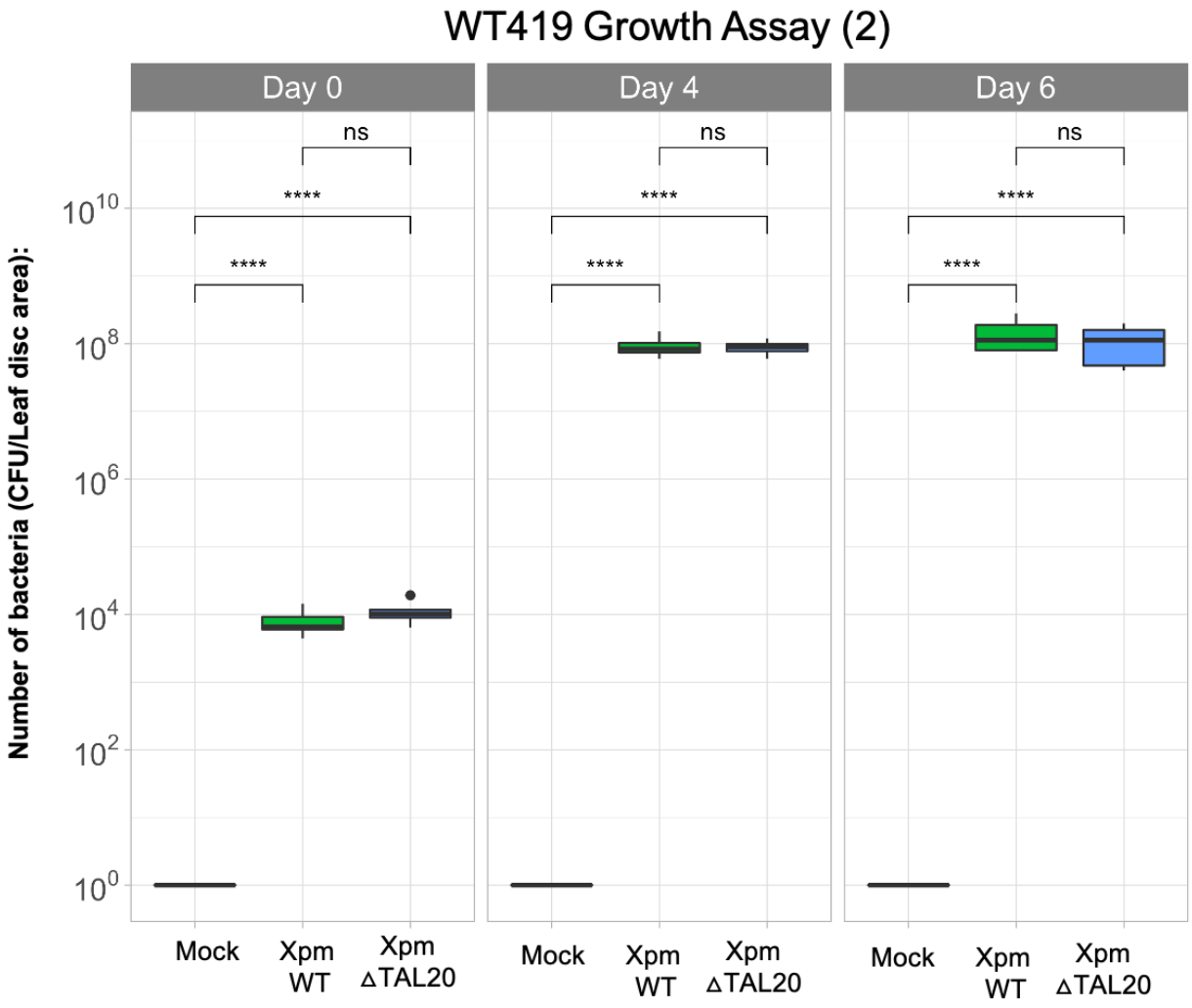

25% Humidity Growth Assay:
Day 0
Day 4
Day 6
Day 9
Number of bacteria (CFU/Leaf disc area):
10^8
10^6
10^4
10^2
10^0
0.53
0.09
0.43
0.40
Mock
XpmΔHrcV
Xpm WT
XpmΔTAL20

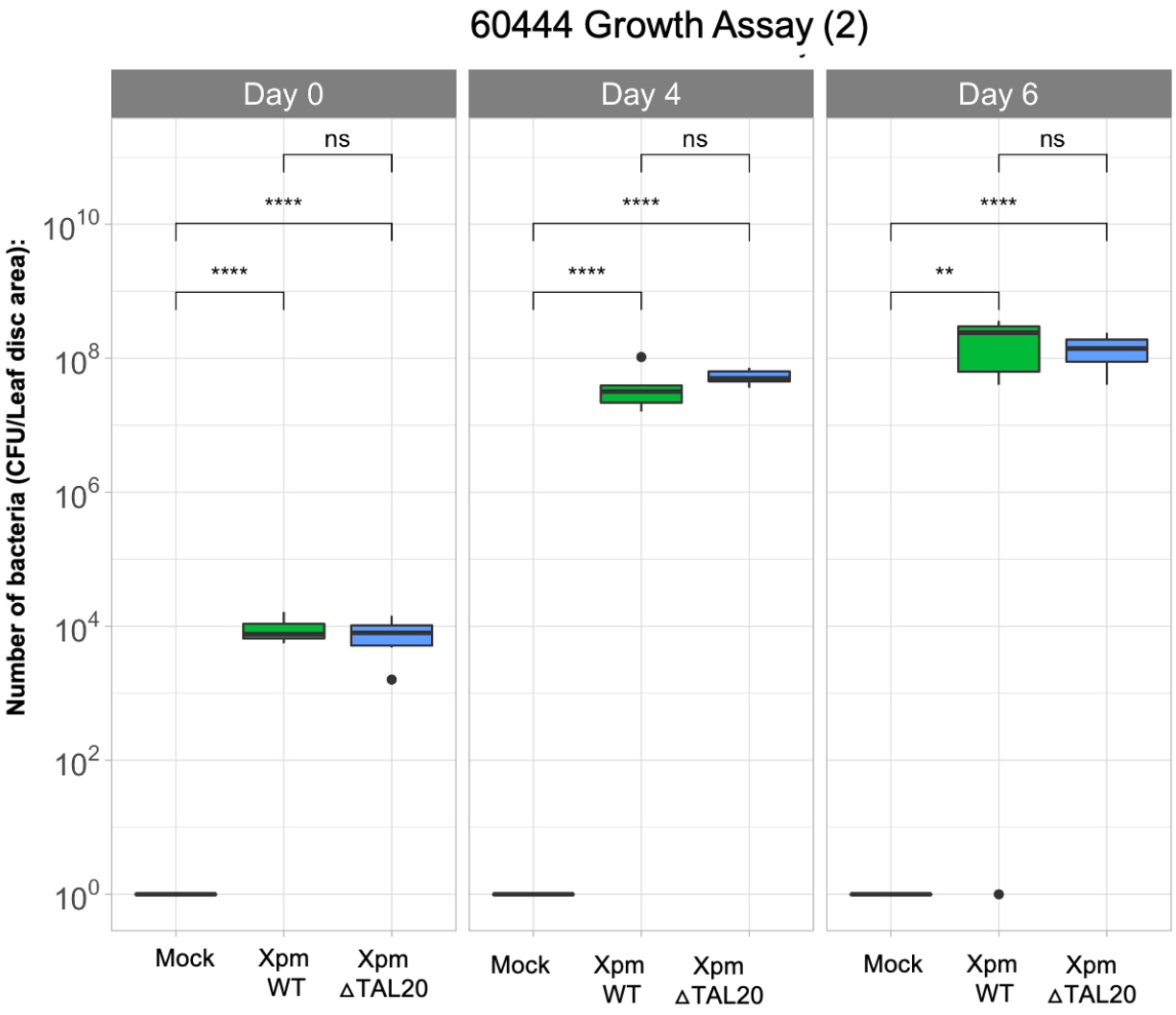
Figure A.3.4: Growth Assay 4:

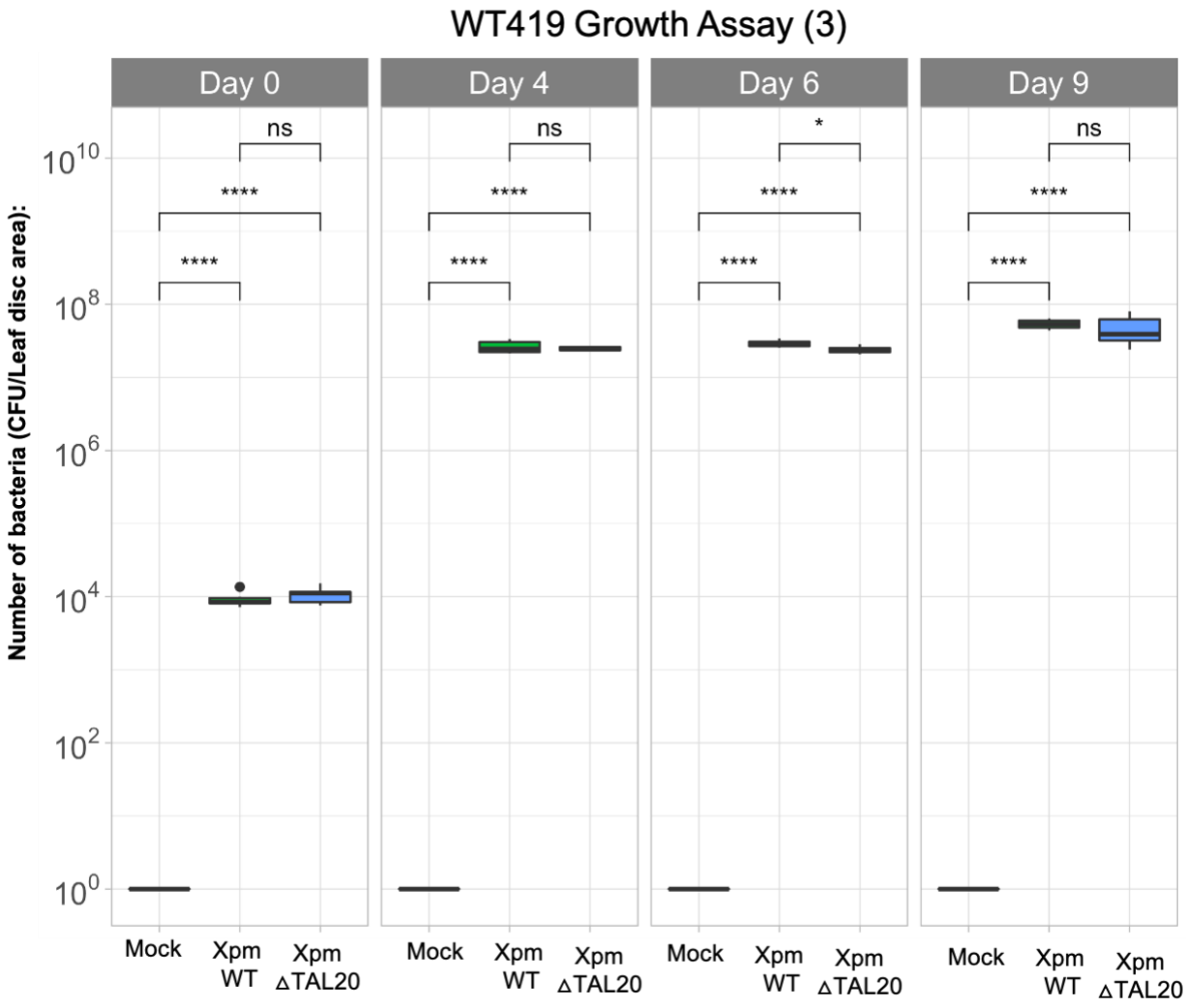
In previous work published reporting differences in Xpm WT, and XpmΔTAL20 growth overtime, the cassava cultivar 60444 was used while in my thesis I focused on the farmer-preferred cultivar of cassava TME419 or WT419 (Cohn and Bart, 2014). We considered that there may be cultivar genotype differences that could impact *in planta* bacterial growth. Therefore, I completed three rounds of bacterial growth assays in the WT419 and 60444 treated with mock, Xpm WT, and XpmΔTAL20 (OD600=0.01). At 0, 4, 6, and/or 9 DPI, samples were collected for CFU assays (N=6 per treatment). For both WT419 and 60444 growth assays one and two, no difference between Xpm WT and XpmΔTAL20 was captured at any time-point. In growth assay three XpmΔTAL20 infected in WT419 plants had significantly decreased growth compared to Xpm WT. However, by 9 DPI, no difference in bacterial CFUs was observed and in 60444 Xpm infected plants no growth difference was seen. (T-Test Significance Key – ns: $p > 0.05$, *: $p <= 0.05$, ** $p <= 0.01$, ***: $p <= 0.001$, and ****: $p <= 0.0001$).


WT419 Growth Assay (1)
Day 0
Day 4
ns
****
****
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
Mock
Xpm WT
Xpm ΔTAL20


60444 Growth Assay (1)
Day 0
Day 4
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
ns
***
****
ns
****
****
Mock
Xpm WT
Xpm ΔTAL20
Mock
Xpm WT
Xpm ΔTAL20


WT419 Growth Assay (2)
Day 0
Day 4
Day 6
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
ns
****
****
Mock
Xpm WT
Xpm ΔTAL20


60444 Growth Assay (2)
Day 0
Day 4
Day 6
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
ns
****
****
ns
****
****
ns
****
**
Mock
Xpm WT
Xpm ΔTAL20
Mock
Xpm WT
Xpm ΔTAL20
Mock
Xpm WT
Xpm ΔTAL20


WT419 Growth Assay (3)
Day 0
Day 4
Day 6
Day 9
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
ns
****
****
ns
****
****
*
****
****
ns
****
****
Mock
Xpm WT
Xpm ΔTAL20

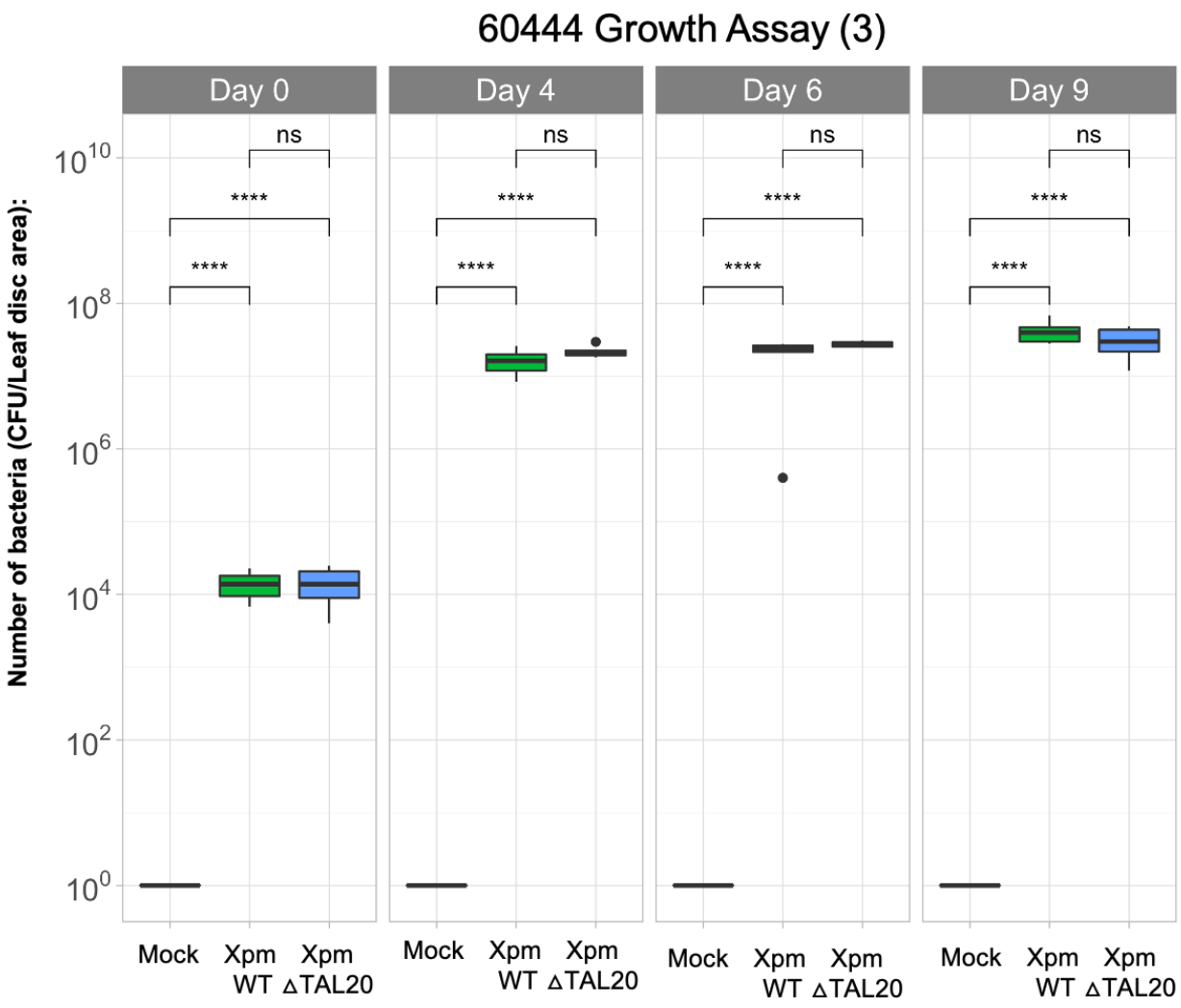
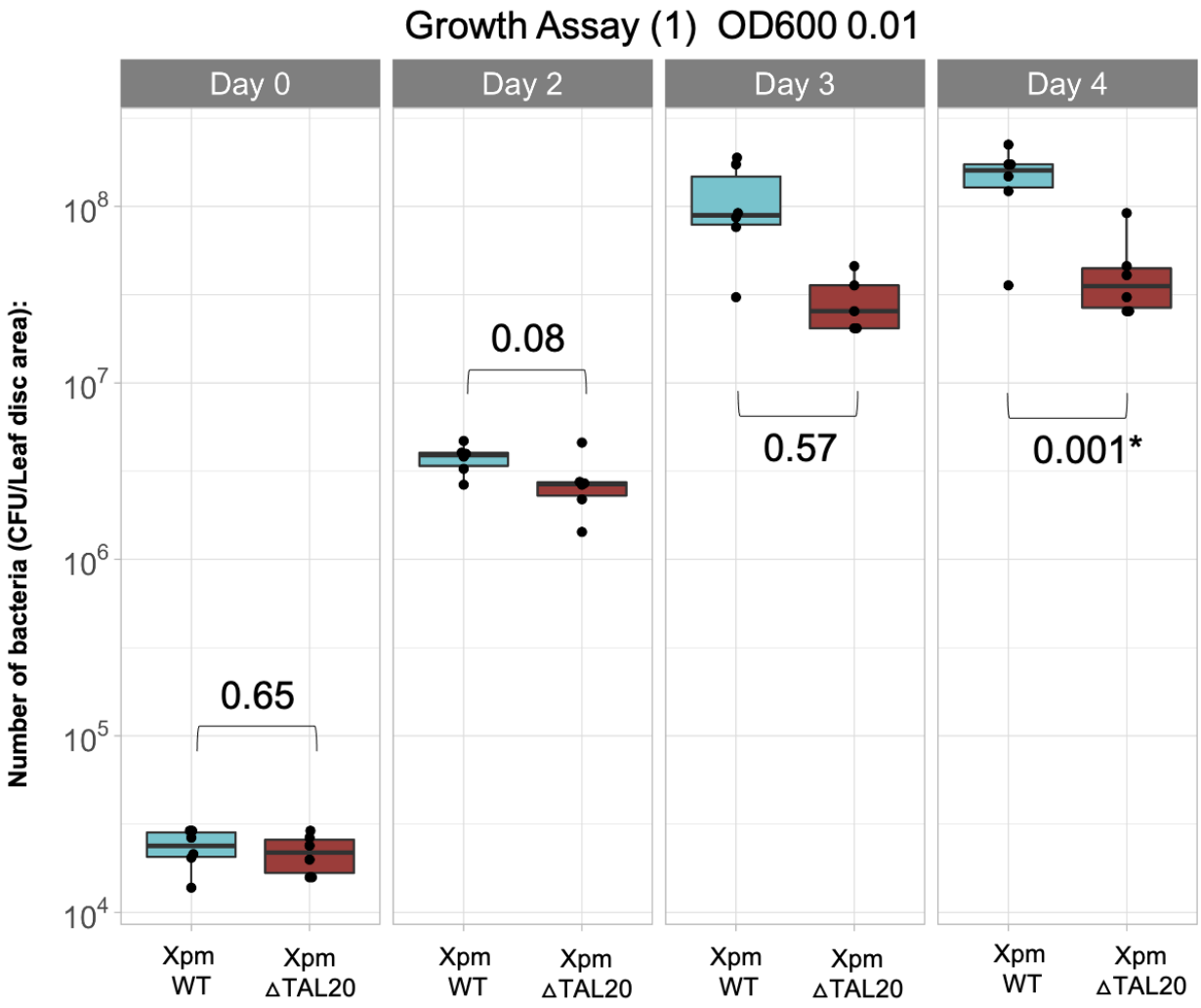
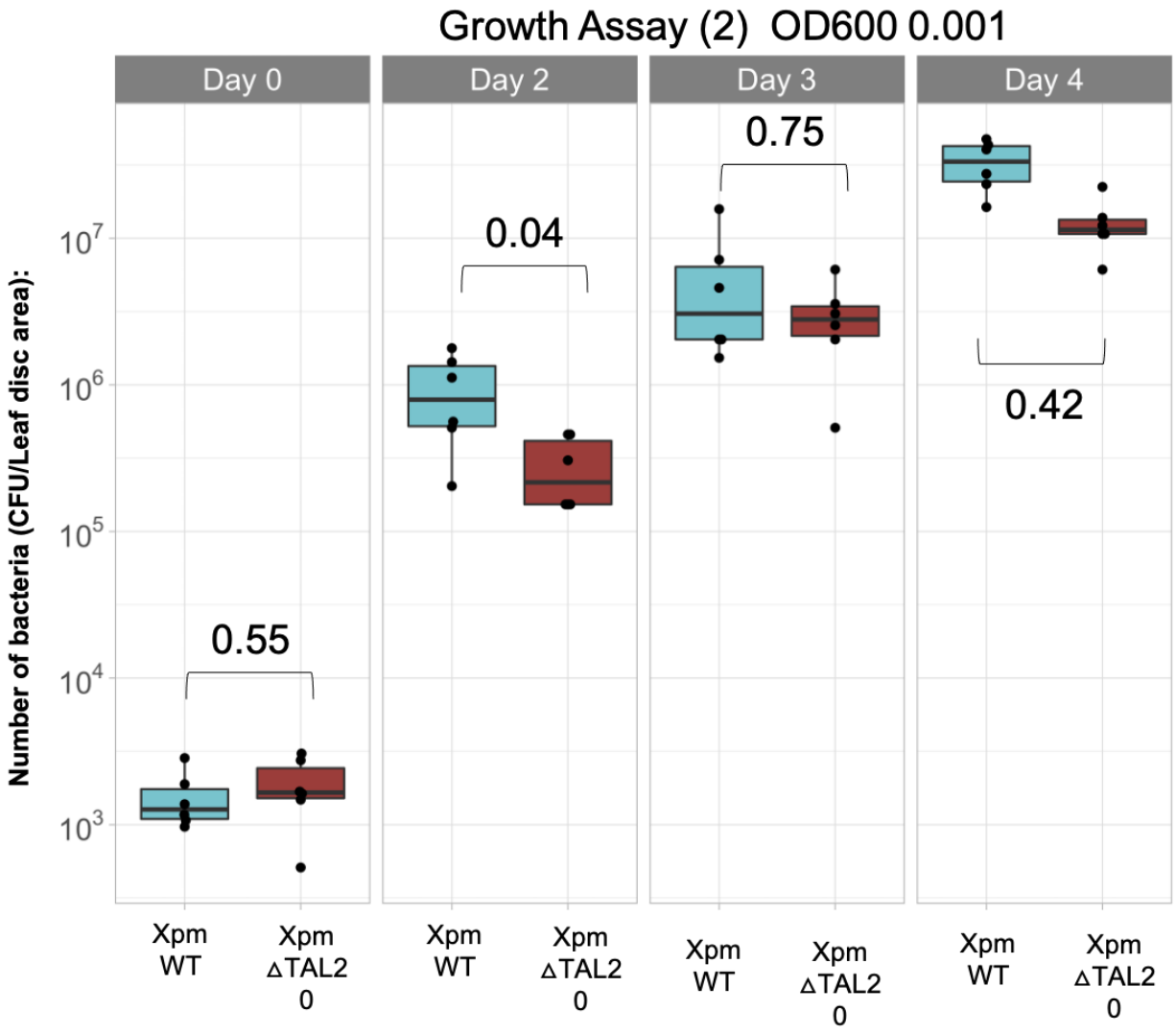

60444 Growth Assay (3)
Day 0
Day 4
Day 6
Day 9
Number of bacteria (CFU/Leaf disc area):
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
ns
****
****
Mock
Xpm WT
Xpm ΔTAL20

Figure A3.5: Growth Assay 5:

Bacterial growth begins in a lag phase, transitions to an exponential log phase, and eventually enters a stationary phase before cells die. It is possible that at the time-points sampled and with the optimal density (OD600=0.01) used for growth assays we are missing Xpm growth during the log phase. Therefore, I completed Xpm WT and XpmΔTAL20 bacterial growth assays with earlier sample collection times and with a lower concentration of inoculum. In growth assay one, an OD600 of 0.01 was used and in growth assay 2, an OD600 of 0.001 was used. At 0, 2, 3, and 4 DPI samples were collected (N=6) and the CFUs were calculated. In growth assay 1, at 2 and 3 DPI no significant difference in Xpm WT and XpmΔTAL20 CFUs was seen. However, at Day 4, XpmΔTAL20 did have decreased growth compared to Xpm WT. In growth assay 2, a significant difference was observed at 2 and 4 DPI but not at 3DPI. We found that these results were still varied and did not seem to fully capture virulence differences between Xpm WT and XpmΔTAL20. However, we did notice that the earlier time points allowed for a more gradual curve of bacterial growth suggesting we were capturing the log phase of Xpm growth.


Growth Assay (1) OD600 0.01
Day 0
Day 2
Day 3
Day 4
Number of bacteria (CFU/Leaf disc area):
10^8
10^7
10^6
10^5
10^4
0.65
0.08
0.57
0.001*
Xpm WT
Xpm ΔTAL20


Growth Assay (2) OD600 0.001
Day 0
Day 2
Day 3
Day 4
Number of bacteria (CFU/Leaf disc area):
10^7
10^6
10^5
10^4
10^3
0.55
0.04
0.75
0.42
Xpm WT
Xpm ΔTAL20

A.4 Tracking Xpm:pLUX spread *in planta*

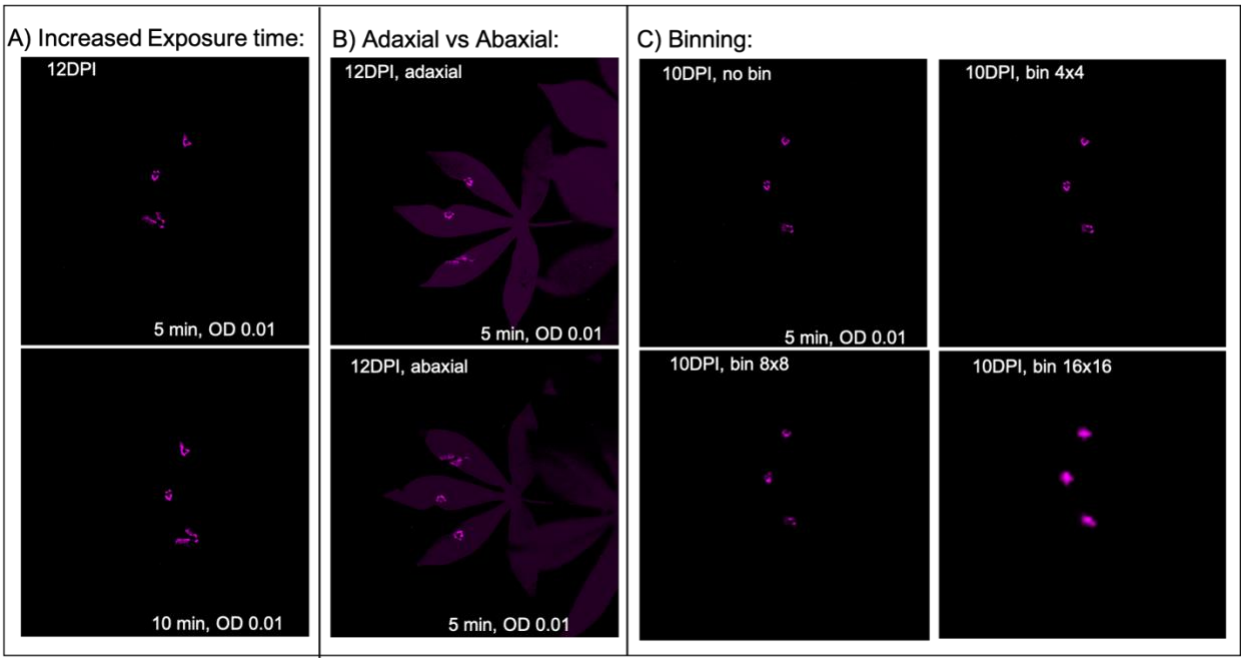
A4.1: Overview

Previous work tracking the spread of luciferin-luciferase tagged (Xpm:pLUX) in the cassava cultivar, 60444, validated the dispersal of Xpm in the cassava vasculature (Mutka 2016). One hypothesis is that sugar efflux may be needed for the spread of Xpm *in planta*. It is possible that in *MeSWEET10a* mutant cassava, Xpm is unable to spread throughout the vasculature. However, a reliable method for examining Xpm spread in cassava must be developed before this hypothesis can be tested. I sought to do this by infecting wildtype cassava with Xpm:pLUX and using an iKon-M 934 low noise charge-coupled device (CCD) camera to visualize the movement of Xpm throughout the cassava vasculature in real-time. WT419 plants were inoculated with Xam:pLUX at varying optical densities. At different timepoints, infected leaves were imaged using the CCD camera and examined for Xpm:pLUX movement *in planta*. In this appendix section, some initial findings and suggestions for future work and assay optimization will be presented.

A4.2: Trouble Shooting CCD camera settings

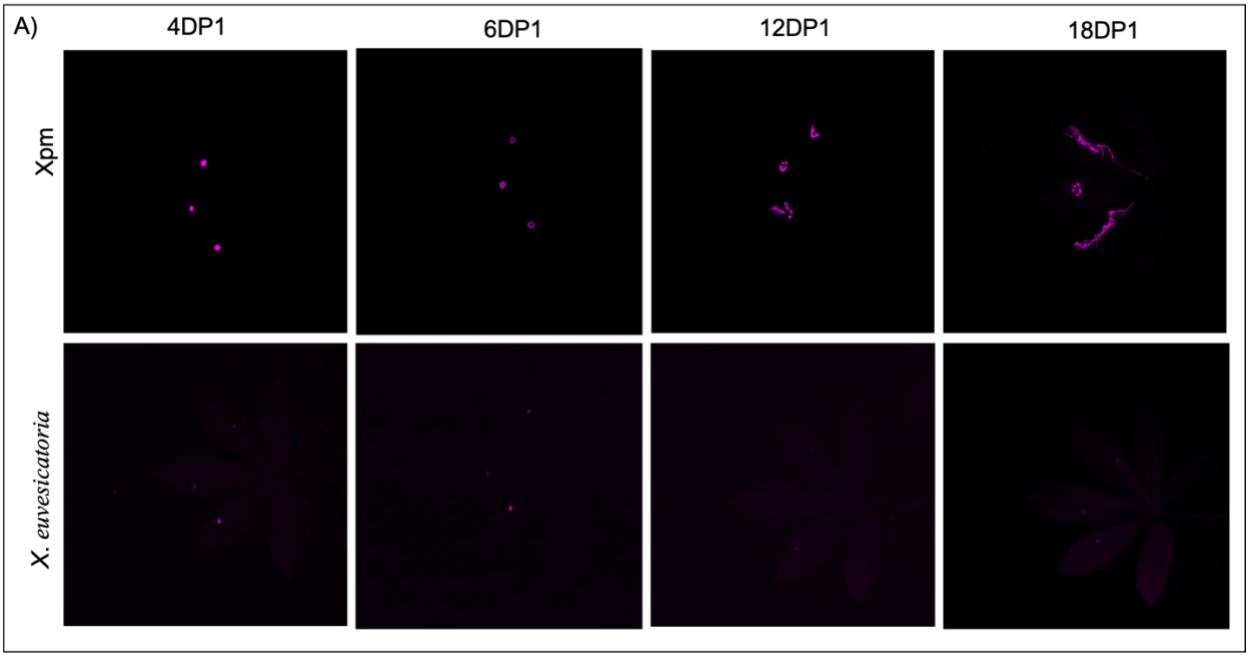
To ensure adequate visualization of Xpm:pLUX, we tested different exposure times, imaging setups, and binning settings during image capture. We saw there was no significant improvement in the amount of Xpm:pLUX emitted signal observed between 5 and 10 minutes. Additionally, the amount of signal observed was qualitatively similar when leaves were imaged from the adaxial or abaxial signal. We also tested how binning or grouping pixels together to reduce noise would impact the signal. We observed no significant difference in infected site signal

between no bin and 4x4 binned images. However, as binning increased, the resolution decreased but did not seem to enlarge the area of signal observed.



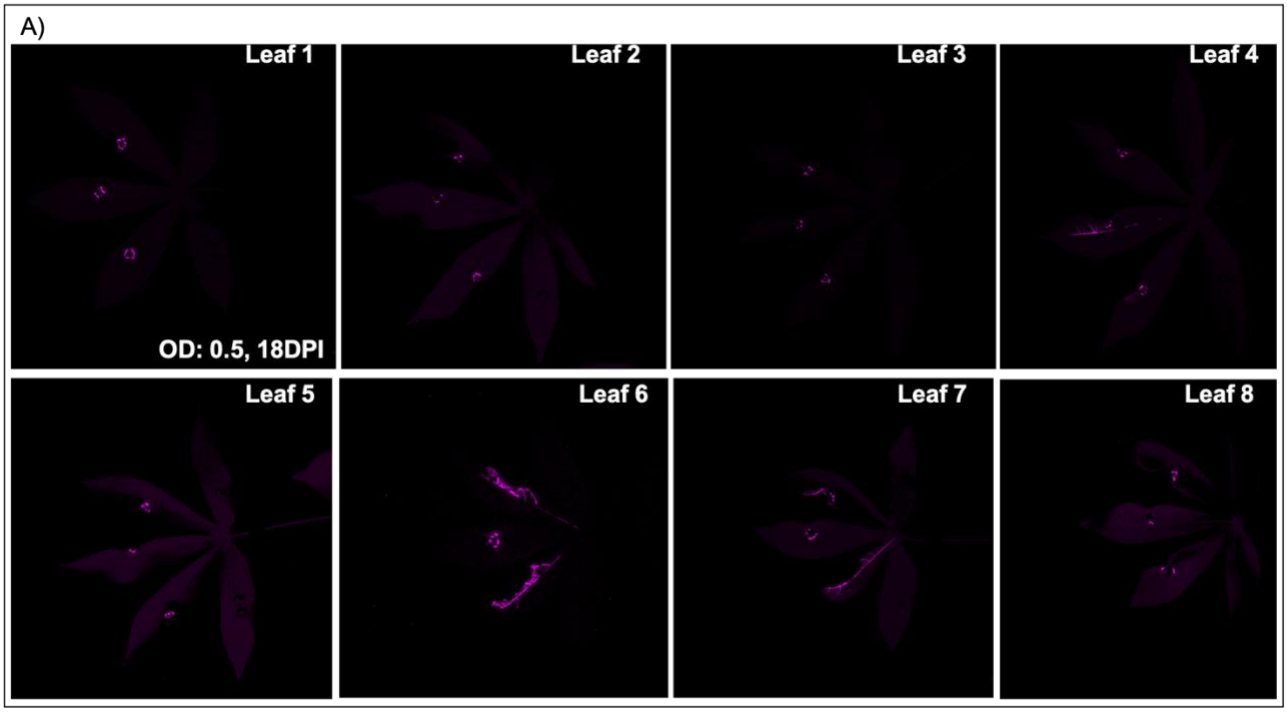
A4.2: Comparing Xpm and *X. euvesicatoria* spread *in planta*

Next, we compared the spread of Xpm to *X. euvesicatoria* carrying pLUX *in planta.* We expected that Xpm would be able to spread beyond the infection site into the vasculature while *X. euvesicatoria* (a nonpathogenic Xanthomonad in cassava) would not. As expected, we saw that Xpm eventually spread into the vasculature. While *X. euvesicatoria* signal was visible at the infection site at 4 and 6DPI, but it was not visible at later timepoints suggesting possible cell death.



A4.3: Observed inconsistent spread of Xpm:pLUX spread into the vasculature

Next, we surveyed the number of times Xpm movement into the vasculature was observed in eight WT419 leaves infected with Xpm:pLUX (OD600 = 0.5). In only 5 out of 24 infected sites was Xpm:pLUX signal seen beyond initial sites of infection in the cassava leaf veins.



Additional assay optimization is required before a bioluminescent tracking assay can be applied to *MeSWEET10a* mutants. Potential routes to optimize this assay include testing different humidity conditions to see if alternative post infection settings can provide more consistent spread of Xpm:pLUX. Similarly, a range of bacterial inoculum concentrations can be tested to find a more ideal optical density. Once conditions allowing for consistent spread of Xpm *in planta* have been identified, XpmΔTAL20 mutants tagged with pLUX will be needed to determine if there is a measurable difference between Xpm WT and XpmΔTAL20. Once a suitable assay has been designed, bioluminescent tracking assay can be applied to *MeSWEET10a* mutant and the potential role of sucrose in Xpm movement *in planta* can be assessed.