Understanding the molecular mechanism of citrus canker promotion mediated by the transcription factor CsLOB1

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Thi Thu Trang Phan

Aus Da Nang, Vietnam

Tübingen

2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	07.02.2024
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Thomas Lahaye
2. Berichterstatter:	Prof. Dr. Ulrike Zentgraf

TABLE OF CONTENTS

PUBLICATION

Kyrylo Schenstnyi, Annett Strauß, Angela Dressel, Robert Morbitzer, Markus Wunderlich, Ana Gabriela Andrade, Trang-Thi-Thu Phan, Paloma de los Angeles Aguilera, Caterina Brancato, Kenneth Wayne Berendzen and Thomas Lahaye

The tomato resistance gene *Bs4* suppresses leaf water soaking phenotypes induced by AvrHah1, a transcription activator-like effector from tomato-pathogenic xanthomonads

New Phytologist, published 2022. https://doi.org/10.1111/nph.18456

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Transcription activator-like effectors (TALEs) play a key role in the virulence of various *Xanthomonas* species by inducing host susceptibility (S) genes. Our study focused on elucidating the disease-promoting mechanism of the TALE PthA4 from *Xanthomonas citri pv. citri* (*Xcc*), responsible for citrus canker disease. *Xcc* lacking PthA4 exhibits significantly reduced growth *in-planta* and fails to induce disease symptoms. Previous research identified the citrus gene *CsLOB1*, a member of the lateral organ boundary (LOB) transcription factor (TF) family, as the key S gene activated by PthA4. It is conceivable that CsLOB1 promotes disease by activating or repressing host target genes. Consequently, our project aimed to identify CsLOB1-regulated host genes and its binding sites, giving insights into the molecular mechanism of disease promotion by CsLOB1 target genes. Combining RNA-seq and ChIP-seq studies, we identified ~100 target genes, many of which are associated with cell wall loosening/degradation. We identified a 15bp CsLOB1 binding motif which was validated *in-vitro* via electrophoretic mobility shift assays (EMSA) and *in-planta* via promoter-reporter assays. Furthermore, the promoter-reporter assay identified the activation of CsLOB1 binding motif by a LOB homolog in tomato, SlLOB1. SlLOB1 is a regulator of fruit softening that exhibits tissue-specific expression, prompting us to investigate the native expression profile of *CsLOB1*. Similarly, *CsLOB1* is highly expressed in mature fruits and not in other tissues. Additionally, mass spectrometry data showed that the apoplastic fluid obtained from *Xcc*-infected leaves had increased glucose, fructose and mannose content, suggesting that *Xcc* hijacked the fruit ripening program to produce sugars. To identify the carbon source, we studied transcriptomic changes in *Xcc in-planta* using RNA-seq and observed upregulation of genes linked to xylan utilisation, a hemicellulose component of plant cell wall. We hypothesise that PthA4 induces the host's cell wall-loosening mechanism by activating the fruit softening regulator, CsLOB1, thereby enabling bacterial enzymes to access and degrade plant xylan as a carbon source. Given that there are other TALEs activating TFs that also induce cell wall degradation to promote disease, our research suggests a novel strategy of TALEs, hijacking the host's cell wall-loosening mechanism to access cell wall polysaccharides for nutrient acquisition.

ZUSAMMENFASSUNG

Transkriptionsaktivator-ähnliche Effektoren (TALEs) spielen eine maßgebliche Rolle für die Virulenz verschiedener *Xanthomonas*-Arten, indem sie Suszeptibilitätsgene (S) des Wirts induzieren. Unsere Studie zielte darauf ab, den Mechanismus zu klären, durch welchen der TALE PthA4 aus *Xanthomonas citri ssp. citri* (*Xcc*) die Krankheit Zitruskrebs verursacht. *Xcc* ohne PthA4 zeigen ein reduziertes Wachstum in der Pflanze und lösen keine Krankheitssymptome aus. Bisherige Untersuchungen identifizierten das Citrus-Gen *CsLOB1*, ein Mitglied der Transkriptionsfaktor-Familie (TF) der Lateral Organ Boundary (LOB), als das entscheidende durch PthA4 aktivierte S-Gen. Denkbar ist, dass CsLOB1 die Krankheit durch Aktivierung oder Unterdrückung von Zielgenen auslöst. Ziel unseres Projektes war es daher, CsLOB1-regulierte Wirtsgene und deren Bindungsstellen zu identifizieren, um Aufschluss über den molekularen Mechanismus der Pathogenese durch CsLOB1-Zielgene zu erhalten. Durch die Kombination von RNA-seq- und ChIP-seq-Studien konnten wir rund 100 Zielgene identifizieren, von denen viele mit der Aufweichung/Degradation der Zellwand in Verbindung stehen. Wir identifizierten ein 15bp CsLOB1-Bindungsmotiv, das in-vitro über Electrophoretic Mobility-Shift-Assays (EMSA) und in-planta über Promoter-Reporter-Assays validiert wurde. Darüber hinaus identifizierte der Promoter-Reporter-Assay die Aktivierung des CsLOB1-Bindungsmotivs durch ein LOB-Homolog der Tomate, SlLOB1. SlLOB1 reguliert die Erweichung der Frucht und weist eine gewebespezifische Expression auf, was uns zu einer Untersuchung des nativen Expressionsprofils von *CsLOB1* veranlasste. Auch *CsLOB1* wird in reifen Früchten stark exprimiert, in anderen Geweben hingegen nicht. Darüber hinaus zeigten Massenspektrometriedaten einen erhöhten Glucose-, Fructose- und Mannosegehalt in der apoplastischen Flüssigkeit von Xcc-infizierten Blättern, was darauf hindeutet, dass Xcc das Reifungsprogramm der Früchte zur Produktion von Zuckern übernimmt. Um die Kohlenstoffquelle zu identifizieren, untersuchten wir transkriptomische Veränderungen in *Xcc* in-planta mit Hilfe von RNA-seq und beobachteten eine Hochregulierung von Genen, die mit der Verwertung von Xylan, einem Hemicellulose-Bestandteil der Pflanzenzellwand, in Verbindung stehen. Unsere Hypothese ist, dass PthA4 die Lockerung der Zellwand auslöst, indem es den Fruchterweichungsregulator *CsLOB1* aktiviert, wodurch bakterielle Enzyme Zugang zu Xylan als Kohlenstoffquelle erhalten und dieses abbauen können. In Anbetracht der Tatsache, dass es weitere TALEs gibt, welche TFs

aktivieren, die den Zellwandabbau einleiten und Krankheiten auslösen, deuten unsere Forschungsergebnisse auf eine neuartige Strategie der TALEs hin, den Zellwandauflockerungsmechanismus des Wirts auszunutzen, um zur Nährstoffgewinnung an Polysaccharide aus den Zellwänden zu gelangen.

1. INTRODUCTION

1.1 TALEs are virulence-promoting effectors

Bacterial species of the genus *Xanthomonas* cause disease in a wide range of plant species, including monocot and dicot plant species. Infection by *Xanthomonas* is promoted by the warm and humid climate of tropical and subtropical regions (Leyns et al., 1984). Pathogenic xanthomonads inject cocktails of effectors into host cells via the type III secretion system (T3SS) to manipulate host metabolism to promote disease (Büttner and Bonas, 2002). Transcription activator-like effectors (TALEs) account for a large group of effectors that are delivered to host cells via type III secretion (T3S). TALEs have a modular composition of functional domains comprising an N-terminal T3S signal, a C-terminal acidic activation domain, and a DNA binding domain that is located in the centre of the polypeptide chain (Van den Ackerveken et al., 1996, Zhu et al., 1998) (Figure 1). A unique structural feature of TALEs is the DNA-binding domain located in the centre of the polypeptide chain that binds in a sequence-specific fashion to a matching *effector binding element* (*EBE*) in a given target promoter. The TALE DNA binding domain is composed of 1.5 to 33.5 tandemly arranged 33-35 amino acid modules that are nearly identical in sequence and are commonly referred to as repeats. Each TALE repeat pairs with one nucleotide in a given *EBE*. The 12th and 13th amino acids of each TALE repeat, which are highly variable between repeats and are therefore referred to as the repeat-variable diresidue (RVD), determine the base preference of a given repeat. The known correlation between the RVD amino acids and their base preference is known as the TALE code. The TALE code facilitates the prediction of TALE target sites in host genomes as well as the construction of TALEs with user-defined target motifs (Figure 1). TALEs that bind to *EBEs* in host promoters recruit the RNA polymerase II complex and transcriptionally activate the downstream host gene. If TALE-induced activation results in elevated virulence of the TALE-delivering bacterial strain the activated host genes are termed susceptibility (*S*) gene since their activation increases susceptibility of the host plant.

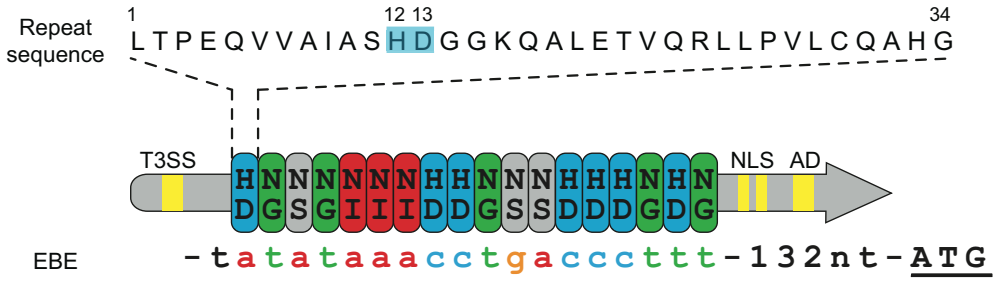


Figure 1: Schematic illustration of a Transcription factor-like effector (TALE) based on AvrBs3 structure.
The TALE protein AvrBs3 is depicted as a bold grey arrow with coloured ovals representing repeats. Repeat variable diresidues (RVDs) are depicted in single letter code (uppercase) and are aligned with the targeted bases (lowercase) located upstream of the pepper (*Capsicum annuum*) *UPA20* ATG start codon (underlined). The colour of the TALE RVDs indicates their base preference (specifically: NI-a [green], HD-c [blue], NG-t [red], NH-g (orange). N* [blue/red] preferentially targets both c and t, NS [grey] can target all 4 nucleotides). The amino acid sequence of the first repeat is shown with the RVD at the 12th and 13th positions highlighted in blue at the top. AD: activation domain, NLS: nuclear localisation signal, T3SS: type III secretion signal ,. Figure adapted from Boch et al. (2009).

1.1.1 TALEs activate host genes to modulate host developmental functions to promote disease.

Although the mode of action of how TALEs bind DNA and transcriptionally activate downstream host genes is well understood, the mechanism of how expressional activation of host genes promotes disease remains mostly unclear. Molecular analysis of non-TALE T3S effectors from *Xanthomonas* revealed that most of them suppress or manipulate defence responses (reviewed in Timilsina et al. (2020)). For example, the *Xanthomonas* T3S effector XopL has been shown to inhibit defence responses triggered by the *Xanthomonas* PAMPs (pathogen-associated molecular patterns) elf18 or flg22 (Singer et al., 2013). Similarly, XopR and AvrBs2 from *X. axonopodis* pv. *manihotis* (now known as *X. phaseoli* pv. *manihotis, Xpm*) were shown to suppress PAMP-triggered immunity (Medina et al., 2018).

While *Xanthomonas* non-TAL effectors typically suppress plant immunity *Xanthomonas* TAL effectors have been shown to manipulate diverse host functions (reviewed by Teper et al. (2022)). Specifically, some TALEs modulate host hormone pathways such as abscisic acid and salicylic acid, which have been shown to be associated with water-soaked leaf tissue and/or regulation of defence responses, respectively (Peng et al., 2019, Wu et al., 2022). A large group of TALEs from *X. oryzae* pv. *oryzae* (*Xoo*), *Xpm*, and *X. citri* pv. *malvacearum* (*Xcm*), which infect rice (*Oryza sativa*), cassava (*Manihot esculenta*) and cotton (*Gossypium spp.*),

respectively, have been shown to convergently activate members of a clade III sugar efflux transporter termed SWEETs (Sugar Will Eventually be Exported Transporters) (Chen et al., 2012, Cohn et al., 2014, Cox et al., 2017, Streubel et al., 2013). It has been suggested that expressional activation of SWEETs causes the efflux of sucrose from the cytoplasm of host cells into the apoplast. Since *Xanthomonas* pathogens proliferate in the apoplast, the increased sucrose levels are likely to provide a carbon source for *Xanthomonas* thereby promoting the proliferation of the bacterial pathogen. While this simple model is appealing, there is evidence for it to be challenged. Mutants of *Xpm* that do not contain a functional copy of the sucrose importer *suxC* were not altered with respect to *in-planta* growth suggesting that sucrose does not serve as a carbon source for *Xpm* (Cohn et al., 2014). It has been speculated that increased sucrose concentration causes an increase in osmotic pressure, resulting in an aqueous apoplast that favours bacterial growth. However, this hypothesis still requires experimental validation (Teper et al., 2022). While it is unclear how increased SWEET expression promotes *in-planta* growth, it has been demonstrated that transcriptional activation of *SWEETs* is essential for the virulence of *Xoo*, *Xpm* and *Xcm* strains. Similarly, the TALE Tal2g, as a major virulence factor of *X. oryzae* pv. *oryzicola* (*Xoc*), was shown to promote disease via the transcriptional activation of a sulfate transporter (*OsSULTR3*). Yet, while the functional relevance of OsSULTR3 for the virulence of *Xoc* has been demonstrated, it remains unclear how up-regulation of the sulfate transporter promotes disease (Cernadas et al., 2014). Although the functional relevance of i) creating a habitable aqueous environment, ii) suppressing plant immunity, or iii) activating SWEETs to induce elevated apoplastic sugar levels has not been demonstrated, it seems at least conceivable that these TALE-induced changes would promote bacterial growth. In contrast, there is currently no working hypothesis as to how TALE-dependent transcriptional activation of plant transcription factors could favour *Xanthomonas* growth in the host plant.

1.1.2 The activation of transcription factors by TALEs to modulate host physiology could reveal a novel disease-promoting mechanism.

The most well-studied transcription factor (TF) targeting TALEs are the TALE AvrBs3 from *X. euvesicatoria* (*Xe*) and *AvrHah1* from *X. gardneri,* which have been shown to activate *C. annuum UPA20* and *Solanum lycopersicum bHLH3* and *bHLH6*, orthologous members of the basic helix-loop-helix (bHLH) TF family (Schwartz et al., 2017, Schenstnyi et al., 2022). These

bHLHs activated by AvrBs3/AvrHah1 were able to activate the same pectate lyase that could promote water soaking via the host cell wall modification (Schwartz et al., 2017, Schenstnyi et al., 2022). Although highly conserved among *Xanthomonas* strains infecting pepper and tomato (Potnis et al., 2011, Schwartz et al., 2015, Subedi et al., 2023), AvrBs3 and AvrHah1 do not contribute to bacterial growth *in-planta* (Bonas et al., 1989, Schwartz et al., 2017). By contrast, the TALE PthA4 of the citrus pathogen *Xcc* activates the lateral organ boundary (LOB) TF *CsLOB1* causing 10-100-fold increased levels of bacterial growth (Hu et al., 2014). Transcriptional profiling revealed that the PthA4-induced CsLOB1 TF and the AvrBs3/AvrHah1-induced bHLH both induce transcription of genes encoding cell-wall degrading enzymes, suggesting that inducing the host cell wall degrading machinery might be an important disease mechanism of these xanthomonads.

1.1.3 Activation of the plant host cell wall degradation machinery could be a novel disease mechanism of Xanthomonas

The plant cell wall is the first interface of the plant-microbe interaction. It is composed of three main building blocks: cellulose, hemicellulose, and pectin (reviewed in (Cosgrove, 2005, Srivastava et al., 2017). Cellulose is a linear polysaccharide consisting of monomeric D-glucose linked via glycosidic bonds. Cellulose polymers connect to each other via hydrogen bonds to form a rigid insoluble bundle called microfibrils which are often not easily targeted by microbes. Hemicellulose (mainly xylan, xyloglucan, mannan, and β-glucan), on the other hand, has more diverse structures and chemical compositions whose backbones can be glucose, xylose, or mannose. It is often decorated with sidechain sugars such as arabinose and galactose or non-sugar sidechains (extensively reviewed in Scheller and Ulvskov (2010)). Hemicellulose and cellulose can interact and form strong hydrogen bonds. The strength of the interaction is determined by the degree of sidechain decoration in the hemicellulose molecule (Simmons et al., 2016). Pectin is a rather complex polymer with a backbone made of galacturonic acids, and it can be split into three groups: homogalacturonan, rhamnogalacturonan I (RG-I), and RG-II depending on the substitutes on the galacturonic acid molecules (Willats et al., 2001, Atmodjo et al., 2013). Pectin can interact with other polymers and with lignin, cutin, or suberin. Depending on cell wall types, the contents of cell wall polysaccharides are largely different. There are three types of plant cell walls: primary cell wall, secondary cell wall and middle lamella. The primary cell wall of monocots has

higher xylan but lower xyloglucan and pectin contents compared to dicots (Vogel, 2008). The secondary cell wall in general has more xylan and lignin but minor pectin contents compared to the primary cell wall, while the middle lamella is rich in pectin and acts as the mediator of cell-to-cell adhesion (Vogel, 2008). The degree of methyl-esterification of pectin affects how fluid it is, which ultimately determines the cell-to-cell adhesion and the stiffness of the tissue (Daher and Braybrook, 2015).

To acquire nutrients from the plant hosts, pathogenic microbes are equipped with cell wall degrading machinery to break through the cell walls (reviewed by Ishida and Noutoshi (2022)). *Xanthomonas* has been known to contain a large set of genes for cell-wall degrading enzymes (CWDEs), also known as Carbohydrate-Activate enZymes (CAZymes), that degrade pectin, xylan, xyloglucan and, to a lesser extent, cellulose (reviewed by Giuseppe et al. (2023)). The secretion of these CAZymes is mainly via the Type II secretion system (T2SS) or via less well-described outer membrane vesicles (often known as Type zero secretion system) (Ryan et al., 2011, Sole et al., 2015). CAZymes were shown to be important in xanthomonad virulence and adaptation to the plant host (Giuseppe et al., 2023). For example, the presence or absence of a single cellobiohydrolase encoded by *cbsA,* that degrades cellulose, determines whether the *Xanthomonas* species has a vascular or non-vascular lifestyle (Gluck-Thaler et al., 2020). While the mutations in pectin-degrading enzymes did not affect *Xoo* virulence in the rice host as it has lower pectin content (Tayi et al., 2016), the defect of polygalacturonases had a great impact on *X. campestris sp. campestris* when it infects the dicot *Arabidopsis* plant which is known to have higher pectin content (Wang et al., 2008, Giuseppe et al., 2023). As *Xanthomonas* is already equipped with a large set of CAZymes, it is interesting to know that *Xe* uses the TALEs AvrBs3/AvrHah1 to activate host *bHLH* TFs which then activate a host CWDE-pectate lyase, which in turn promotes water soaking (Schwartz et al., 2017, Schenstnyi et al., 2022). *Xcc* was also shown to use PthA4 to induce the *LOB* TF which then also activates many citrus CWDEs (Hu et al., 2014). This suggests a common disease mechanism of these xanthomonads via inducing host TFs, which are likely to be the mediator of the host cell wall degradation machinery. Therefore, to clarify how TALE-induced expression of plant TFs promotes disease, the first step would be to uncover the TF target genes. We chose to study *CsLOB1* rather than *bHLH*

proteins since CsLOB1 induces not only very clear disease symptoms but also promotes *in-planta* growth.

1.2 *Xanthomonas citri* promote citrus canker by activating a lateral organ boundary transcription factor

Citrus is an economically very important crop, accounting for the second most produced fruit in the world. The most important citrus fruits are oranges, tangerines, limes, lemons, pomelos, and grapefruits (Pereira Gonzatto and Scherer Santos, 2023). China, Brazil, and Italy were the top citrus-producing countries in 2021 (Pereira Gonzatto and Scherer Santos, 2023).

1.2.1 Citrus canker, a devastating disease

Citrus canker, a devastating disease to citrus plants, originated in Southeast Asia and is now a major threat to citrus production in Asia, the Americas, and Africa. Recent outbreaks of citrus canker have been reported in Darwin, Australia in 2018 (Australian Department of Agriculture, 2022), Saudi Arabia in 2020 (Ibrahim et al., 2023) and are still occurring at high incidence in southern Brazil since 2019 (Behlau, 2021), in Florida, Texas and Louisiana, USA since 2020 (Gochez et al., 2020). The disease is characterised by raised corky lesions (pustules) surrounded by yellow-halo water-soaked margins (water-soaking) on fruits, leaves, and stems. Infected plants show defoliation, fruit drop, and dieback (Brunings and Gabriel, 2003). To date, there are no effective measures to eradicate this disease. The most commonly used control measures are copper spraying, eradication of infected plants, establishment of quarantine zones with windbreaks, and quarantine measures in regions where citrus canker is not present or has been eradicated (Graham et al., 2004, Behlau et al., 2010). To date, the disease has had a severe economic impact due to loss of fruit yield and marketability, and the high cost of quarantine and eradication programmes. Many commercial citrus cultivars are susceptible to citrus canker. The most susceptible citrus varieties include grapefruit and lime, while sweet orange can be tolerant to certain pathovars (Gottwald et al., 2002). Understanding the pathogens and their mechanism of disease promotion is important in developing effective crop protection strategies.

1.2.2 Causal agents of citrus canker

Citrus canker is caused by two distinct *Xanthomonas* pathovars, *Xcc* and *X*. *fuscans* ssp. *aurantifolii* (*Xfa*), which produce similar symptoms including pustules and water soaking (Ference et al., 2018). They infect through openings such as stomata and wounds and spread mainly by wind and rain (Gottwald et al., 2002). They are able to survive on plant debris for up to two months, potentially increasing the chance of spreading (Gottwald et al., 2002, Graham et al., 2004). *Xcc* is the most widespread group infecting a broad range of citrus cultivars and originates from Asia (hence the name "Asiatic group") (Gottwald et al., 2002). *Xcc* can be classified into pathotypes A, A^w and A^*, with type A infecting most citrus species, while types A^w and A^* are restricted to Key lime (*Citrus aurantiifolia*) (Rybak et al., 2009, Sun et al., 2004). In contrast, *Xfa* is mainly distributed in South America and consists of two types, type B and type C. Type B can cause severe disease in lemon and Key lime (also known as Mexican lime) and less severe symptoms in orange, whereas type C is limited in Mexican lime and causes a hypersensitive response (HR) in most other citrus species (Brunings and Gabriel, 2003, Fonseca et al., 2019).

1.2.3 The contribution of T3S effectors to citrus canker

To increase its virulence, *Xanthomonas* injects a cocktail of about 20 distinct effector proteins into host cells via the T3SS to manipulate plant processes and promote disease. Seven effectors that are common to all *Xanthomonas* genomes, AvrBs2, XopK, XopL, XopQ, XopR, XopX and XopZ, and all have immunity-suppressing activity (Dalio et al., 2017) (Section 1.1.1). This may explain the suppression of immunity-regulating genes early in infection by *Xcc* (Cervera et al., 2008, Hu et al., 2016). In addition, the host restriction for *Xcc*-A^w, *Xcc*-A^* and *Xfa*-C in Key lime was due to the presence of the avirulent T3S homologous XopAG effectors, AvrGf1 (in *Xcc*-A^w, *Xcc*-A^*) and AvrGf2 (in *Xfa*-C), which induce HR in most citrus species except Key lime (Rybak et al., 2009, Ngoc et al., 2009, Gochez et al., 2015). The exact mechanism promoting HR and the host executors responsible for it remain unclear, although the mode of action of AvrGf2 has been shown to be related to the protein folding catalyst cyclophilin (Gochez et al., 2017). Nevertheless, AvrGf1 and AvrGf2 were not found in *Xcc*-A strains, which may explain their high virulence and wide host range compared to other citrus canker-causing xanthomonads.

While AvrGf1 and AvrGf2 are responsible for host restriction, the other T3S effector family, AvrBs3/PthA, classified as TAL effectors or TALEs, are responsible for the typical symptoms of citrus canker, pustules and water soaking, and for promoting xanthomonad growth in the host. Mutational studies show that PthA from *Xcc* confers the ability to promote citrus canker (Swarup et al., 1991, Swarup et al., 1992). Strains of *Xcc*-A^{w}, *Xcc*-A* and *Xfa*-B, *Xfa*-C were found to contain at least one PthA homolog, namely PthA^{w}, PthA*, PthB, and PthC, respectively, all of which contribute to pustule formation and water soaking in citrus (Al-Saadi et al., 2007). There are four *PthA* genes in *Xcc*-A, namely *PthA1*, *PthA2*, *PthA3,* and *PthA4*. However, only *PthA4* has been consistently and exclusively shown to be essential for citrus canker disease. *Xcc*-A lacking PthA4 was unable to induce pustules, water soaking and had a reduced bacterial population (Swarup et al., 1992, Yan and Wang, 2012). In contrast, the virulence contribution of *PthA1*, *PthA2,* and *PthA3* was somewhat inconsistent and dependent on host cultivars (Yan and Wang, 2012, Abe and Benedetti, 2016). *Xcc*-A strain 306 has its genome fully sequenced and functionally annotated (da Silva et al., 2002, Laia et al., 2009, Jalan et al., 2013). Therefore, we chose *Xcc*-A strain 306 (hereafter referred to as *Xcc*) as a working model to study the TALE PthA4 and its disease mechanism.

1.2.4 *Xcc* PthA4 promotes Citrus canker symptoms and *Xcc* growth by activating the lateral organ boundary transcription factor *CsLOB1*

The combination of microarray-based transcriptomic analysis searching for PthA4-induced genes and *in silico* prediction of PthA4 *EBEs* revealed 15 citrus genes that are potential targets of PthA4 and promoting citrus canker (Hu et al., 2014). Designer TALEs (dTALEs) with user-defined DNA binding specificity (Figure 1) that target individual candidate genes showed that only activation of *CsLOB1* rescued the PthA4-mutant *Xcc* (*XccΔpthA4*) in promoting disease symptoms (pustules and water soaking) and *in-planta* bacterial growth, revealing that *CsLOB1* is the PthA-induced susceptibility gene in citrus that promotes canker disease (Hu et al., 2014). Follow-up analysis showed that PthA^{w}, PthA*, PthB and PthC from *Xcc* relatives (Section 1.2.2) also targeted the *CsLOB1* promoter (Hu et al., 2014), suggesting a convergent evolutionary scheme of different *Xanthomonas* species to promote citrus canker. *CsSWEET1*, a citrus homolog of the *SWEET* genes that promote bacterial blight disease caused by xanthomonads in rice, cassava, and cotton (Chen et al., 2012, Cohn et al., 2014, Cox et al., 2017) is also induced by PthA4. However, dTALEs activating *CsSWEET1*

promoted neither bacterial growth of *Xcc* nor pustule and water-soaking symptoms (Hu et al., 2014), suggesting that *CsLOB1*-activating disease promotion by PthA4 may be distinct from SWEET-induced bacterial blight. PthA4-induced genes are overrepresented in cell wall-related functions, as seen in several independent studies, but the exact mechanism remains elusive (Hu et al., 2016, Pereira et al., 2014, Zhang et al., 2017a). Therefore, investigation of the host genes regulated by CsLOB1 is key to understanding this novel PthA4-induced disease mechanism.

1.2.5 CsLOB1 belong to a LOB-domain containing transcription factor family

CsLOB1 is a member of the plant-specific LOB-domain (LBD) containing transcription factor family. The majority of LBD family members, including CsLOB1, belong to Class I, which is known for the characteristic LOB domain consisting of a cysteine-rich $CX_2CX_6CX_3C$ DNA-binding motif (known as zinc finger), followed by a conserved glycine-alanine-serine motif in which the glycine residue is invariant, and a C-terminal leucine zipper-like coiled-coil motif $LX_6LX_3LX_6L$ for protein-protein dimerisation (Shuai et al., 2002, Chen et al., 2019). Class II LOB proteins lack the $LX_6LX_3LX_6L$ domain, and their members were shown to have functions enriched in nitrogen metabolism and anthocyanin biosynthesis, as reviewed by Zhang et al. (2020). In contrast, the Class I LOB members have been shown to participate in more diverse developmental pathways such as leaf polarisation, hormone signalling, lateral root development and tissue architecture (Xu et al., 2016, Zhang et al., 2020).

1.2.5.1 LOB proteins have diverse functions

The closest CsLOB1 homologs in *Arabidopsis thaliana* (Arabidopsis hereafter), AtLBD1 and AtLBD11, have been shown to be involved in root secondary growth and cambial stem cell maintenance (Ye et al., 2021), with AtLBD11 regulating reactive oxygen species (ROS) metabolism genes that in turn regulate cambial cell proliferation (Dang et al., 2023). PtXPaLOB1, another close homolog from a poplar hybrid *Populus tremula* x *Populus alba* shown by Hu et al. (2014), has been shown to be involved in poplar secondary growth (Yordanov et al., 2010). Another closely related homolog is SlLOB1 (Hu et al., 2014) which was shown to regulate fruit softening in tomato (Shi et al., 2021). Other examples include AtLBD18, which acts in root development by directly upregulating the expansin expression (Berckmans et al., 2011, Lee et al., 2013), and the Maize homolog ZmLBD30, which controls

the pollen development (Hou et al., 2023). Interestingly, several LOB homologs have been implicated in disease susceptibility. AtLBD20 is involved in susceptibility to the fungal pathogen *Fusarium oxysporum* (Thatcher et al., 2012), while AtLBD16 contributes to gall formation as a feeding tissue for the root-knot nematode (Cabrera et al., 2015). Beet necrotic yellow vein virus, which causes rhizomania disease in sugar beet (*Beta vulgaris*), was shown to hijack the auxin signalling pathway to activate several LBDs to promote massive lateral root and root hair formation (Gil et al., 2018). In citrus, of the 34 LOB homologs, the closest CsLOB1 homologs, CsLOB2 and CsLOB3, were shown to be capable of promoting citrus canker when induced by dTALEs delivered by *XccΔpthA4* (Zhang et al., 2017a). This suggests that activation of LBDs may be important for disease promotion. It is noteworthy that the native function of CsLOB1 has not been fully elucidated, despite being known as the susceptibility gene in citrus canker disease. Understanding the native function of CsLOB1 may provide insight into how the physiological changes regulated by the TF are beneficial to the pathogen.

1.2.5.2 LOB-targeted DNA binding motifs

One way to determine if a TF-regulated gene is a direct TF target gene is to screen for the presence of TF binding sites in its promoter sequence. The LOB domain of the Arabidopsis LOB homolog, AtLOB, was previously shown to bind to a hexameric GCGGCG motif by combining two *in-vitro* assays, the selection and amplification binding (SAAB) assay and the electrophoretic mobility shift assay (EMSA) (Husbands et al., 2007). A high throughput *in vitro* protein binding microarray showed that AtLBD16 also binds to a hexamer TCCGGA (Franco-Zorrilla et al., 2014). Using the GCGGCG motif to search for possible candidate genes among PthA4-induced genes, a predicted binding site was found in the promoter of the gene *Cs2g20600,* encoding a putative zinc finger E3 ligase (Duan et al., 2018). Oligonucleotides containing this binding site were able to interact with CsLOB1 in EMSA, however, the functional contribution of the gene to citrus canker remains unclear (Duan et al., 2018). Subsequently, chromatin immunoprecipitation sequencing (ChIP-seq) was performed on a stable citrus transgenic line constitutively expressing CsLOB1 and identified also a hexamer GCGGCG consensus motif (Zou et al., 2021). The motif, however, was not present in the promoter of the *Cs2g20600* gene found by Duan et al. (2018) (Zou et al., 2021). More importantly, both works from Duan et al. (2018) and Zou et al. (2021) lack *in-planta*

validation for binding between CsLOB1 and the identified motifs. The inconsistency in studies of CsLOB1 binding motifs and the lack of proper experimental validation brings the credibility of the identified motifs into questions. Meanwhile, the crystalised structure of a LOB family member was shown to form dimers via a leucine zipper-like domain when binding to the DNA (Chen et al., 2019). The homodimeric structure was found to maintain a strict distance between two DNA-binding zinc finger motifs, each recognising a similar GCGGCG motif separated by a defined number of nucleotides (Chen et al., 2019). Interestingly, it is the number of nucleotides separating the two motifs, but not the sequence of the separator, that was critical for functionality, which ultimately suggests that the motif should be longer than the canonical hexamer. Indeed, a genome-wide DNA affinity purification sequencing to discover the binding sites of Arabidopsis TFs revealed a rather different motif pattern, in which eight LOB homologs appear to target a nearly palindromic sequence between 15 to 21-bp in length, where the 5’ and 3’ end bases are GC-rich and highly conserved, connected by a less conserved sequence in between (Figure 2) (Plant Cistrome database, (O'Malley et al., 2016)). Notably, the canonical GCGGCG motif is weakly present in the nearly palindromic sequence. In addition, a recent ChIP-seq experiment using Maize ZmLBD30 also showed a similar pattern (Hou et al., 2023). Taken together, we suspect that the motif of LOB family members may have a more diverse motif pattern. We propose re-examination of the binding motif of CsLOB1 using a different approach from previous studies and including experimental validation, which will be discussed in detail in Section 1.3.

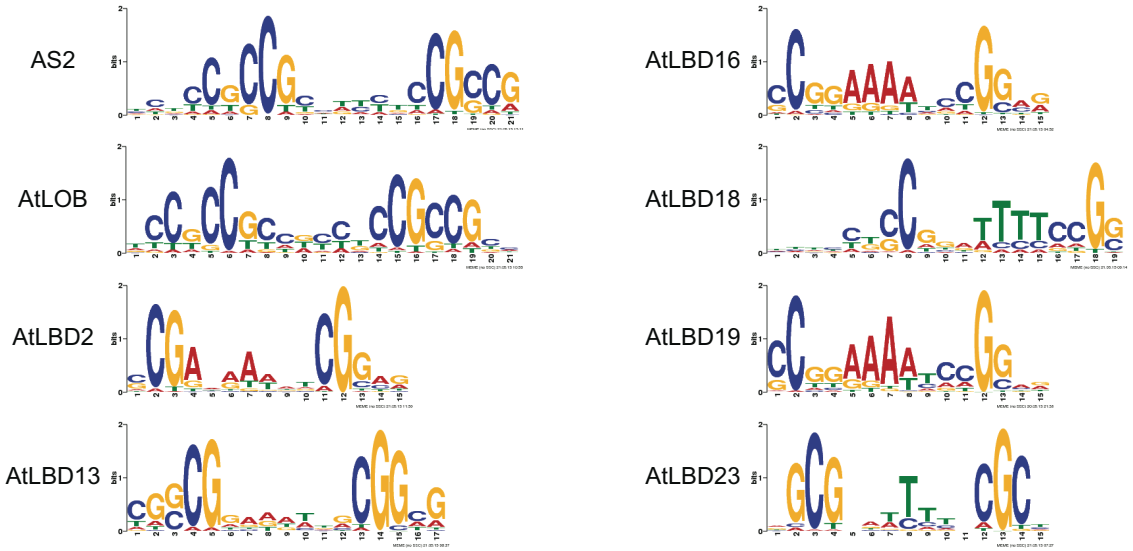


Figure 2: The binding motifs of Arabidopsis LOB homologs show a nearly palindromic pattern.

The binding motifs of Arabidopsis LOB homologs obtained from genome-wide DNA affinity purification sequencing (Plant Cistrome database, (O'Malley et al., 2016)). Each logo is a position-dependent base probability matrix represented as letters with proportional size corresponding to the base probability in the motif sequence.

1.3 Aim of this project

Xcc uses PthA4 to activate *CsLOB1* to promote pustules, water soaking and bacterial growth (Hu et al., 2014). As CsLOB1 is a member of the LOB TF family, it is conceivable that genes regulated by CsLOB1, but not CsLOB1 itself, induce physiological changes in the host that benefit *Xcc.* Therefore, we aim to identify the genes directly regulated by CsLOB1 as a first step towards understanding its disease-promoting mechanism.

We will use *Xcc*-delivered PthA4 and functionally equivalent dTALEs to induce *CsLOB1* expression. Transcript profiling by RNA sequencing (RNA-seq) will identify citrus genes whose expression is CsLOB1-dependent while ChIP-seq with CsLOB1-specific antibodies will identify genes whose promoters have CsLOB1 binding sites. The combination of ChIP-seq and RNA-seq data will uncover direct CsLOB1 targets. The predicted function of direct CsLOB1 target genes will give an indication of what physiological changes benefit the bacteria. In addition, ChIP-seq has the potential to uncover a putative CsLOB1 binding motif that eventually will be validated by *in-vitro* and *in-planta* studies

While *Xcc* injects PthA4 into host cells to activate *CsLOB1* and manipulate host metabolism, *Xcc* remains in the apoplast, suggesting that CsLOB1 expression in host cells would

ultimately translate into changes in the apoplast, where *Xcc* resides. Therefore, we will also perform metabolite profiling by gas chromatography-mass spectrometry (GC-MS) on citrus apoplastic fluids upon *CsLOB1* activation. Taken together with what is observed with respect to CsLOB1-regulated genes/transcripts, we could uncover how CsLOB1 activation translates into downstream gene expression, and how in turn it could affect the compositions of Citrus apoplastic fluids that promote *Xcc* growth.

To support rapid *Xcc* growth, the pathogen must acquire a carbon source from the host. While changes in host metabolism would provide indications of which metabolites are available to *Xcc*, it will not inform which metabolites are in fact used by *Xcc*. As bacteria rapidly adapt the set of expressed catabolic enzymes to the available carbon sources, we will also examine the transcriptomic changes in *Xcc* when the host experiences pustules and water soaking compared to when infected with the isogenic *XccΔpthA4* mutant. We envision that integration of CsLOB1-induced metabolic changes in the host and transcriptomic changes in the parasite, will elucidate how PthA4 by activation of a TF promotes disease. Moreover, we envision that the investigation of this TALE-induced TF will provide clues on how other TF-inducing *Xanthomonas* TALEs for example AvrBs3/AvrHah1 promote bacterial disease.

2. RESULTS

2.1 CsLOB1-dependent transcriptome analysis reveals up-regulation of genes involved in cell wall degradation and modification

CsLOB1 belongs to the lateral organ boundary (LOB) transcription factor family and is convergently targeted by PthA4 and related TALEs from distinct *Xcc* strains (Hu et al., 2014). Investigation of host genes regulated by CsLOB1 is the key to understanding citrus canker disease development. This chapter describes the approach to analyse CsLOB1-dependent transcriptome using designer TALEs targeting *CsLOB1* promoter in a time-course experiment.

2.1.1 Designer TALEs targeting *CsLOB1* for the elimination of PthA4 off-targets

The first goal was to identify genes directly activated or repressed by CsLOB1. We used PthA4-containing *Xcc* was used as a tool to activate *CsLOB1* and thereby *CsLOB1* downstream targets. To overcome the problem of off-targets, two dTALEs were cloned by Dr. Robert Morbitzer, namely dTALE2 and dTALE4, with user-defined DNA-binding specificity (Morbitzer et al., 2010) that target *CsLOB1* promoter at 26bp and 15bp upstream of PthA4 *EBE*, respectively (Figure 3A). These dTALEs target different *EBEs* within the *CsLOB1* 5' upstream region and are therefore likely to have different off-target genes compared to PthA4's. Identification of differentially expressed genes (DEGs) shared by dTALE2, dTALE4, and PthA4 would eliminate off-target genes of PthA4 and dTALE2/4 and thus identify CsLOB1-dependent DEGs.

The dTALEs were introduced into the non-pustule-promoting *XccΔpthA4* strain and were able to functionally complement the loss-of-canker phenotype of the mutant strain (Figure 3B), consistent with their ability to induce *CsLOB1* gene expression at relatively similar levels to *Xcc* wild type at 36 hours post infection (hpi) (Figure 3C).

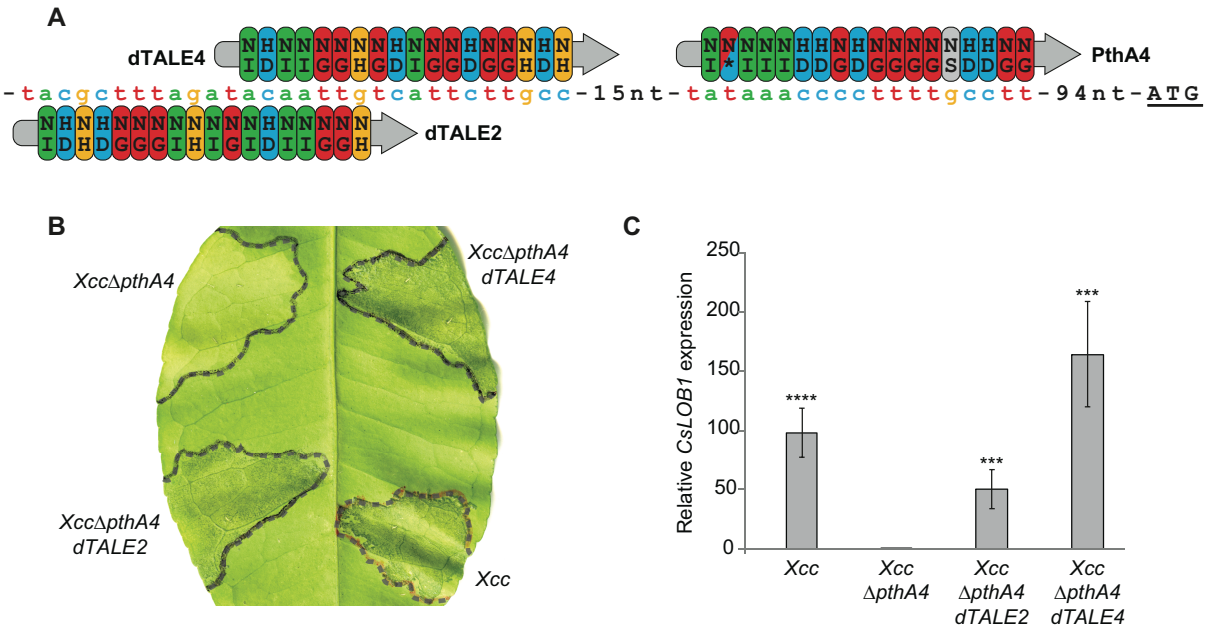


Figure 3: A TALE-based concept for the discovery of CsLOB1-dependent transcriptomic changes

A: Schematic illustration of *CsLOB1*-activating TALEs. The *CsLOB1*-activating *Xcc* TALE protein PthA4 and dTALE2 and dTALE4 cloned by Dr. Robert Morbitzer are depicted as bold grey arrows with coloured ovals representing TALE repeat variable diresidues (RVDs) in single letter code (uppercase) that are aligned with the targeted bases (lowercase) located upstream of the *CsLOB1* ATG start codon (underlined). The colour of the TALE RVDs indicates their base preference (specifically: NI-a [green], HD-c [blue], NG-t [red], NH-g (orange). N* [blue/red] preferentially targets both c and t, NS [grey] can target all 4 nucleotides). **B: dTALEs targeting *CsLOB1* rescue Citrus canker phenotype of *Xcc* lacking PthA4 in grapefruit cultivar (cv.) Duncan.** Black dashed lines indicate areas syringe infiltrated with depicted *Xcc* strains. Photo was taken 7 days post infiltration. **C: dTALEs restore *CsLOB1* expression to similar levels as PthA4 does.** Duncan leaves were infiltrated with *Xanthomonas* strain shown in B. Total RNA was extracted at 36 hpi. Transcript levels were quantified by reverse-transcription real-time PCR (RT-qPCR). *CsLOB1* expression was normalised to the housekeeping gene *EF1α* (*Cs7g27470*). Bar graph represents the mean ± SD (standard deviation) of 6 biological replicates. Student's t-test was used to calculate the significant difference between depicted samples and the *XccΔpthA4* data points. ***: p ≤ 0.001, ****: p ≤ 0.0001. dTALE: designer TALE, RVD: repeat variable diresidues.

High-throughput Illumina RNA sequencing (RNA-seq) was then performed on young fully-expanded citrus leaves (*C. paradisi*, known as Duncan grapefruit) harvested at 36hpi infiltrated with *Xcc, XccΔpthA4* (negative control), *XccΔpthA4::dTALE2* or *XccΔpthA4::dTALE4*. Host genes with expression fold change of ≥2 or ≤-2 (FDR < 0. 05) when compared to expression of leaves infected with *XccΔpthA4* were considered as up- or down-regulated genes, respectively. Integration of differential transcript yielded 942 CsLOB1-dependent Up-regulated at 36hpi (CU36) and 465 down-regulated (CD36) genes that we consistent across three comparisons: *Xcc* vs *XccΔpthA4*, *XccΔpthA4::dTALE2* vs *XccΔpthA4 and XccΔpthA4::dTALE4* vs *XccΔpthA4* (Figure 4D). Data integration also uncovered a

significant number of PthA4's off targets, accounting for approx. 500 up- and 700 down-regulated genes (Figure 4).

Annotation of CU36 and CD36 genes displayed that many upregulated genes are related to signalling, cell wall, and stress response, whereas downregulated genes are related to metabolism such as hormone, lipid, amino acid, nucleotide, N- and secondary metabolism (Supplemental Figure 5).

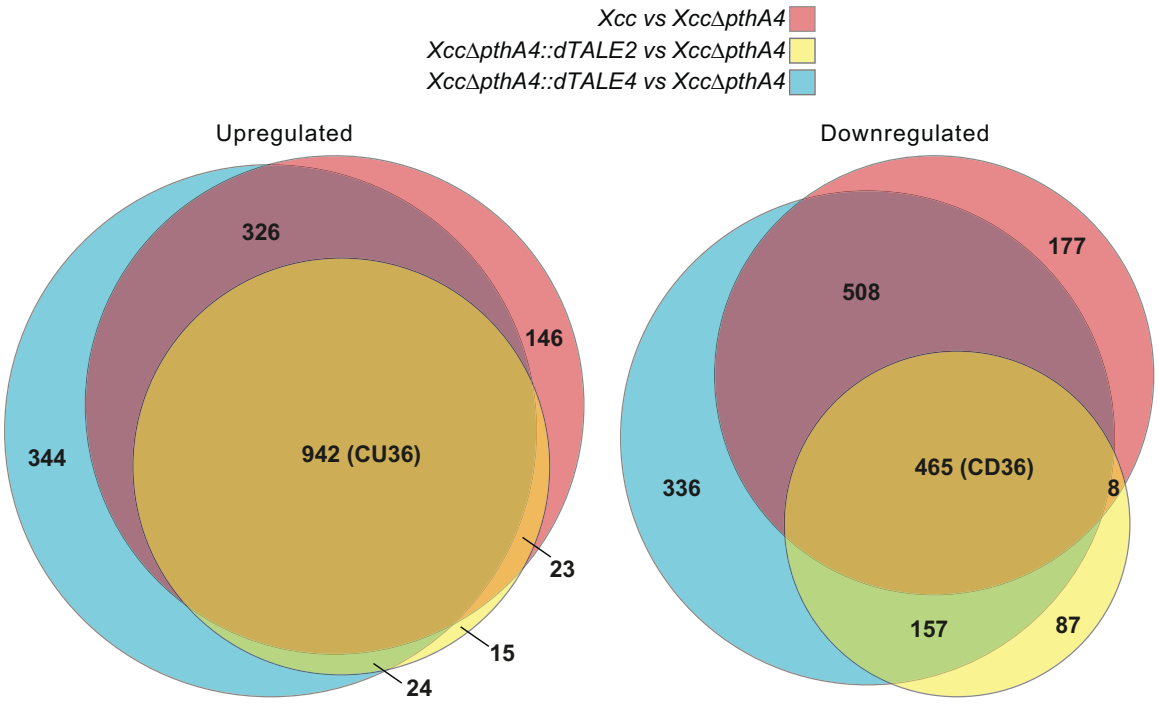


Figure 4: RNA-Seq-based identification of TALE-induced genes facilitates discovery of CsLOB1-dependent transcripts.
Size-proportional Venn diagrams show citrus differentially expressed genes (DEGs) upon infection with *Xcc* containing PthA4, *XccΔpthA4* containing dTALE2 or dTALE4 relative to *XccΔpthA4,* as depicted. Genes that at 36hpi are up- or down-regulated (fold change ≥ 2 or ≤ -2 (FDR <0.05), respectively) consistently by all three TALEs are designated as CsLOB1-UP-36 (CU36, dark orange) or CsLOB1-DOWN-36 (CD36, dark orange) regulon, respectively. The number of genes in each overlap is in bold black. RNA-seq was performed with 4 biological replicates.

2.1.2 Early-induced CsLOB1-dependent host genes encode for cell-wall degrading enzymes

CsLOB1-dependent transcripts can be formally divided into two distinct functional classes: mRNAs of genes that have a CsLOB1 binding site (direct CsLOB1 target genes) or mRNAs whose corresponding genes do not have a CsLOB1 binding site, but whose transcription depends on the translation of direct CsLOB1 target genes. To reduce the number of "indirect

targets" in the RNA-Seq studies, an early time point after infection when the indirect targets were not yet transcribed was sought after. To identify a suitable time point, the transcript levels of a number of CU36 genes at 12, 14 and 16hpi were quantified by RT-qPCR (Figure 5). At these early time points, transcripts that are elevated in abundance were able to be detected (such as *orange1.1t00187*) and therefore may be direct CsLOB1 target genes (Figure 5).

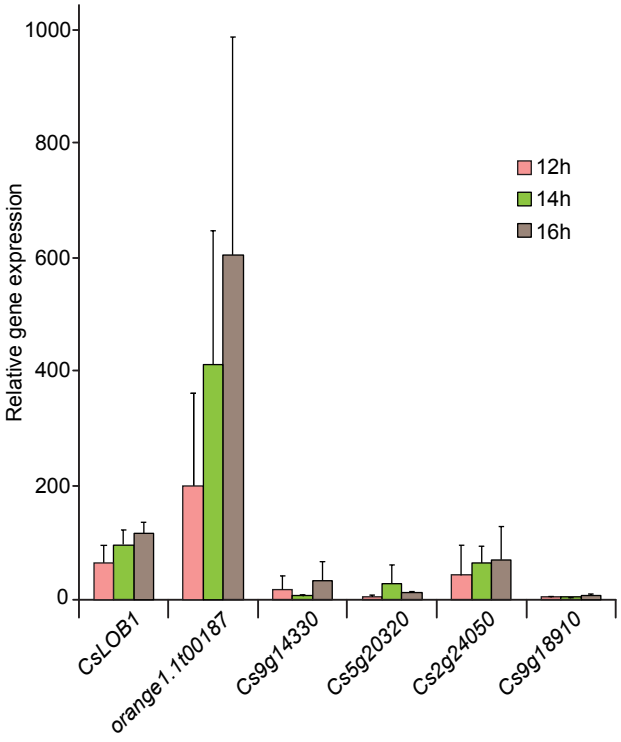


Figure 5: Time course analysis shows that some CsLOB1-induced genes, initially identified at 36 hpi were already induced at 12hpi upon *Xcc* infection.
Total RNA was extracted from Duncan leaves infiltrated with either *Xcc* (containing PthA4) or *XccΔpthA4* at 12, 14 or 16hpi, as depicted. Transcript levels were quantified by RT-qPCR. Relative expression was obtained from double normalisation of the gene of interest to the housekeeping gene *FBOX* (*Cs6g19880*) and to *XccΔpthA4*. Bar graph represents the mean + SD of 3 biological replicates.

We then performed RNA-seq at 12hpi on tissue infected with either *Xcc* or *XccΔpthA4.* Comparison of tissue harvested at 12hpi uncovered 390 up-regulated DEGs and 200 down-regulated DEGs of *Xcc* compared to *XccΔpthA4*. The DEGs identified at 12 hpi were then compared with those identified at 36 hpi (CU36 and CD36), revealing 216 up- and 76 down-regulated genes in common (Figure 6A).

Gene ontology characterisation of these genes using the EggNOG gene ontology database (Huerta-Cepas et al., 2019) revealed that a large proportion of the up-regulated genes have

biological functions in cell wall organisation or biogenesis, cell wall polysaccharide/macromolecule catabolic processes and cell wall loosening (Figure 6B). Numerous genes are also predicted to have hydrolase, transferase, transferring hexosyl group, mannosyltransferase, beta-galactosidase, and glucuronosyltransferase activity (Figure 6C). On the other hand, down-regulated genes are enriched in protein activation cascade (Figure 6B) with down-regulated activity of some peptidase, agmatine N4-coumaroyltransferase, ATPase-coupled transmembrane transporter, and primary active transmembrane transporter (Figure 6C).

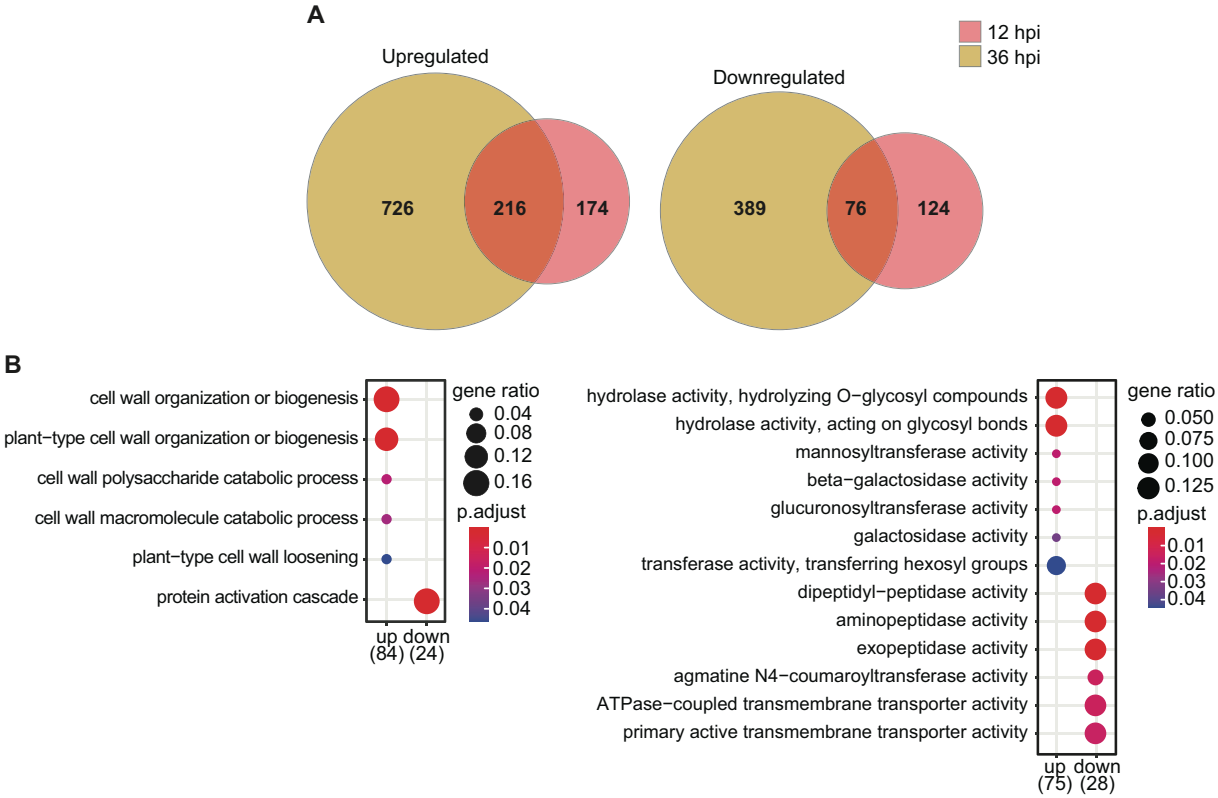


Figure 6: A high percentage of early-induced CsLOB1-dependent genes encode for cell-wall degrading enzymes

A: RNA-seq of leaf tissue harvested at 12hpi with *Xcc* reduces the number of CsLOB1-dependent host genes. Size-proportional Venn diagrams illustrate the overlap between CU36 or CD36 (dark orange; see also Figure 1D) and DEGs of *Xcc* relative to *XccΔpthA4* at 12hpi. Depicted DEGs with fold change ≥ 2 or ≤ -2 (FDR <0.05) are classified as up- or down-regulated genes. The number of genes in each overlap is in bold black. RNA-seq was performed with 4 biological replicates. **B: Biological function (left) and molecular assignment (right) of the DEGs shown in Figure 2B indicate that many genes are involved in cell wall degradation.** Functional enrichment analysis was performed by Dr. Paulo Teixeira in R using the EggNOG gene ontology database (Huerta-Cepas et al., 2019). Circle sizes are proportional to the number of genes belonging to a given category. Numbers (in brackets) of either up-regulated (up) or down-regulated (down) genes with functional annotation based on the EggNOG database are indicated. The circle colour scale indicates the adjusted p-value (ranging from 0-0.05).

2.2 ChIP-seq using CsLOB1-specific antibody reveals CsLOB1 direct target genes that mainly participate in cell wall degradation

RNA-seq has identified host genes that are activated in a CsLOB1-dependent fashion. However, RNA-seq studies do not differentiate between host genes preceded by a CsLOB1 binding site (direct CsLOB1 target genes) and genes transcribed due to translation of these direct targets, that are themselves indirect targets. While it seems plausible that early induced CsLOB1-dependent genes have a high proportion of direct vs. indirect CsLOB1 target genes, the direct or indirect nature of these target genes remains uncertain. To distinguish between direct and indirect targets, chromatin immunoprecipitation sequencing (ChIP-seq), an approach that enables us to identify genes whose promoters physically interact with CsLOB1 was employed.

Initially, two approaches were pursued simultaneously: generation of a transgenic Citrus line encoding an epitope-tagged CsLOB1 protein and generation of a CsLOB1-specific antibody. However, due to time constraints, we focused on the latter approach using the CsLOB1-specific antibody.

2.2.1 ChIP-seq in *Xcc*- infected tissue using anti-CsLOB1 antibody showed that CsLOB1 binds preferably around transcription start sites

For recombinant expression of CsLOB1, the *CsLOB1* gene was optimised with *E. coli* codon usage with 6 N-terminal histidines and was driven by a T7 promoter (Figure 7A). The recombinant *CsLOB1* was expressed by the arabinose- and IPTG (Isopropyl β-D-1-thiogalactopyranoside)-inducible *E. coli BL2AI* strain in high yield but in an insoluble state (Figure 7B). For antibody production, the CsLOB1 protein was purified under denaturing conditions with 4M urea. The purified product was dialysed against the desired buffer with 1M urea to maintain solubility and sent for antibody production (Section 4.2.3.3).

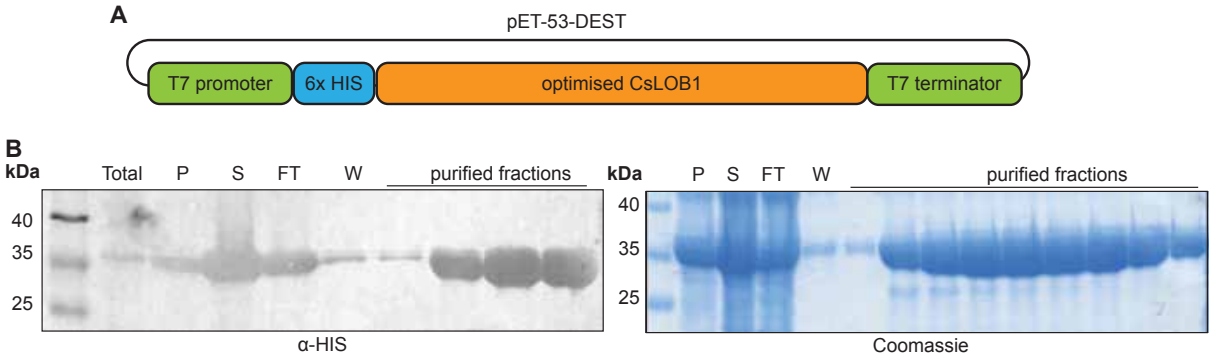


Figure 7: Recombinant CsLOB1 with 6x histidine tag was purified under denaturing conditions and subjected to polyclonal antibody production

A: Schematic representation of a gene construct for overexpression of recombinant His-CsLOB1 in *E. coli*. T7 promoter, terminator: recognised by T7 polymerase from T7 bacteriophage; HIS: histidine; optimised CsLOB1: *CsLOB1* gene with *E. coli* codon usage; pET-53-DEST: vector backbone. **B: Western blot was performed on representative purified fractions from HIS affinity purification** (Section 4.2.3.1). Total bacterial protein extract (Total), pellet fraction (P), soluble fraction (S), flow through after affinity purification (FT) and wash fraction (W) using anti-HIS antibody. Coomassie staining shows the purity level of CsLOB1 in the purified fractions. The predicted molecular weight of the HIS-CsLOB1 fusion protein is 28.84 kDa.

The specificity of the newly generated CsLOB1 antibody was then analysed. To do this, *CsLOB1* was expressed in *Nicotiana benthamiana* using *Agrobacterium tumefaciens* that delivered p35S-*driven CsLOB1-HA*-GFP on a transfer DNA (T-DNA) into plant cells. Then either the generated CsLOB1 antibody or a commercial GFP antibody was used to immunoprecipitate the recombinant CsLOB1 at 2dpi. Western blot using α-HA antibody showed that CsLOB1 antibody precipitated proteins at the same molecular weight as the one pulled down by anti-GFP antibody, indicating that the newly raised antibody is indeed CsLOB1 specific (Figure 8A).

Following this, the newly raised antibody was tested if it could detect PthA4-induced increases in CsLOB1 protein in the context of *Xcc*-infected Duncan grapefruit leaves. To do this, Duncan leaves infected with either the PthA4-containing *Xcc* strain or the isogenic *XccΔpthA4* mutant were examined. *CsLOB1* transcripts are known to be detectable by RT-qPCR after 12hpi and increase dramatically thereafter. On a time-course experiment and subsequent Western blot analysis on total leaf extract, the antibody was able to detect CsLOB1 in *Xcc*-infected leaves as early as 1.5dpi and the detected signal rapidly accumulated over time. By contrast, leaf tissue infected with *XccΔpthA4* showed no signal with the newly raised antibody (Figure 8B). Given that the antibody was able to immunoprecipitate CsLOB1 protein with high specificity only from *Xcc*- but not from *XccΔpthA4*-infected tissue (Figure

8C), it indicated that the raised antibody was indeed CsLOB1-specific and suggested that the antibody was suitable to proceed with ChIP sequencing on *Xcc*-infected leaf tissue.

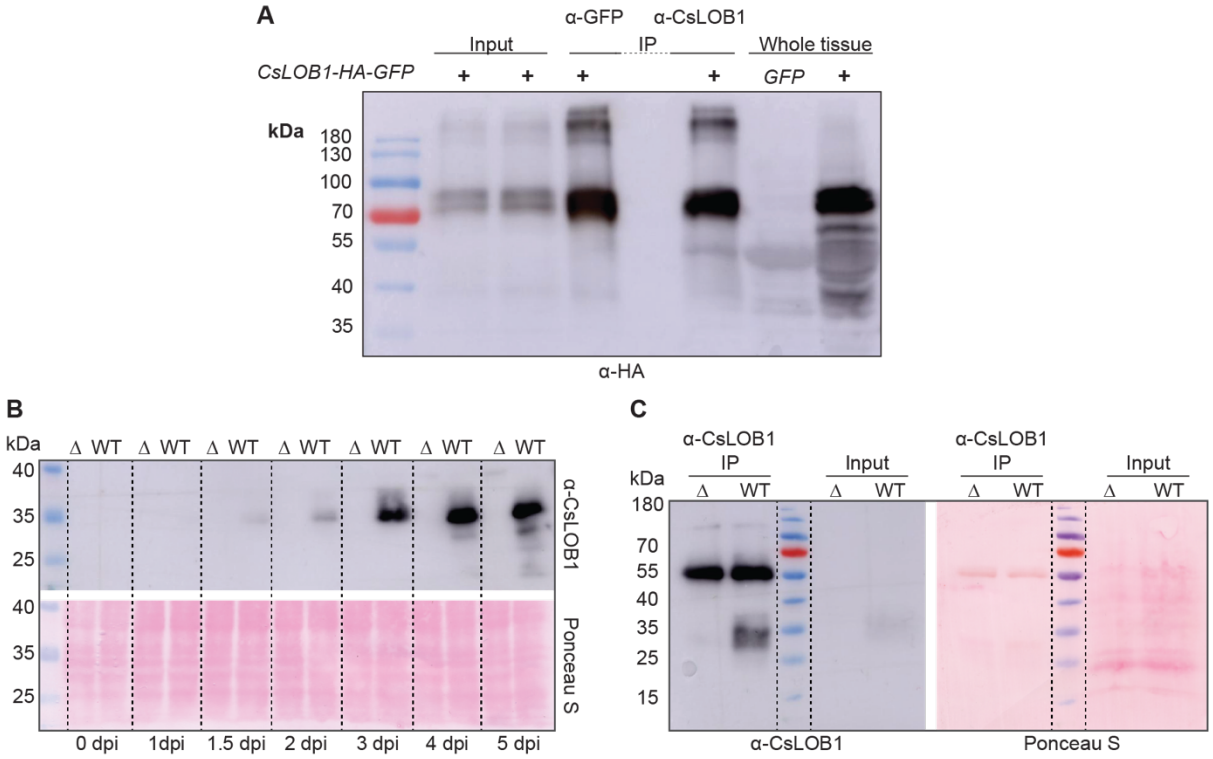


Figure 8: An antibody raised against recombinantly expressed CsLOB1 facilitates the detection and immunoprecipitation of CsLOB1 from *Xcc*-infected Duncan grapefruit leaves with high specificity

A: An antibody raised against recombinantly expressed CsLOB1 allows specific purification of a CsLOB1-GFP fusion protein from *N. benthamiana* leaf extracts. *Agrobacterium*-mediated transient overexpression of *CsLOB1* (*35S-driven CsLOB1-HA-GFP,* +) in *N. benthamiana* was subjected to immunoprecipitation. Horseradish peroxidase (HRP)-conjugated α-HA Western blot was performed on: Input: soluble nuclear fraction; IP: immunoprecipitated fractions using either commercial α-GFP or α-CsLOB1 antibodies; Whole tissue: total protein extracted from *N. benthamiana* leaves expressing p35S-*driven GFP* (*GFP*) or p35S-*driven CsLOB1-HA-GFP*. Leaf samples were collected at 2dpi. The molecular weight of CsLOB1-HA-GFP is 59.56 kDa. **B: An antibody raised against recombinantly expressed CsLOB1 produced a signal only in protein extracts from Duncan leaves infected with CsLOB1-inducing *Xcc* strains (WT) but not in extracts from leaves infected with the *XccΔpthA4* strain (*Δ*).** Total protein extracted from Duncan leaves infected with either *XccΔpthA4* (Δ) or *Xcc* (WT) strains for 5 days was subjected to Western blot with α-CsLOB1 primary antibody and HRP-linked α-rabbit secondary antibody visualised by CCD camera (top). Loading control by Ponceau S staining (bottom). The molecular weight of CsLOB1 is 26kDa. **C: Antibody against CsLOB1 effectively immunoprecipitated CsLOB1 from leaf tissue infected with *Xcc* (WT) strains.** Soluble nuclear fraction of Duncan leaves infected with either *XccΔpthA4* (Δ) or *Xcc* (WT) strains at 40hpi was immunoprecipitated with α-CsLOB1 antibody. Western blot was performed on either immunoprecipitated samples (IP) or soluble nuclear fractions prior to IP (input) using α-CsLOB1 primary antibody and HRP-linked α-rabbit secondary antibody visualised by CCD camera (left). The 50kDa bands detected in the IP samples are likely to be rabbit IgG heavy chain.

ChIP-seq was then performed on *Xcc*-infected Duncan leaves at 12hpi and 36hpi using the generated antibody as described in Liu et al. (2016) with adaptations. Briefly, tissue was

harvested and immediately fixed with paraformaldehyde (PFA) to cross-link protein and chromatin prior to nuclei isolation and chromatin fragmentation. Different concentrations of PFA were tested and 0.5% PFA for 10 minutes with optimised sonication was shown to give expected sizes of chromatin fragments (200-700bp) (Section 4.2.3.6). The sonicated soluble nuclear fractions were immunoprecipitated (IP) with either α-CsLOB1 or IgG as a negative control. IgG from the same species as the CsLOB1 antibody would share the same non-specific binding regions, but not the antigen-specific binding sites, with α-CsLOB1. This allows detection of non-specific protein-protein interaction during immunoprecipitation. ChIP-seq analysis was done following the pipeline from the Galaxy project (Ostrovsky et al., 2022). Reads were aligned to *Citrus sinensis* reference genome (Liu et al., 2022) and binding peaks were defined by MACS2 (Zhang et al., 2008) in comparison between α-CsLOB1 and IgG control samples (Section 4.2.3.6).

A pilot sequencing was performed on replicate 1 at 36hpi. The preliminary analysis was then used to designed ChIP-qPCR primers targeting enriched versus non-enriched regions in two putative target genes, *Cs8g11330* and *orange1.1t00187* (Figure 9). ChIP was then performed on the second replicate at 36hpi and two replicates at 12hpi. ChIP-qPCR primers were used to validate the enrichment of the bound regions in these samples prior to library preparation and sequencing. ChIP-qPCR showed that the peak regions were indeed enriched compared to the control negative regions (Figure 9).

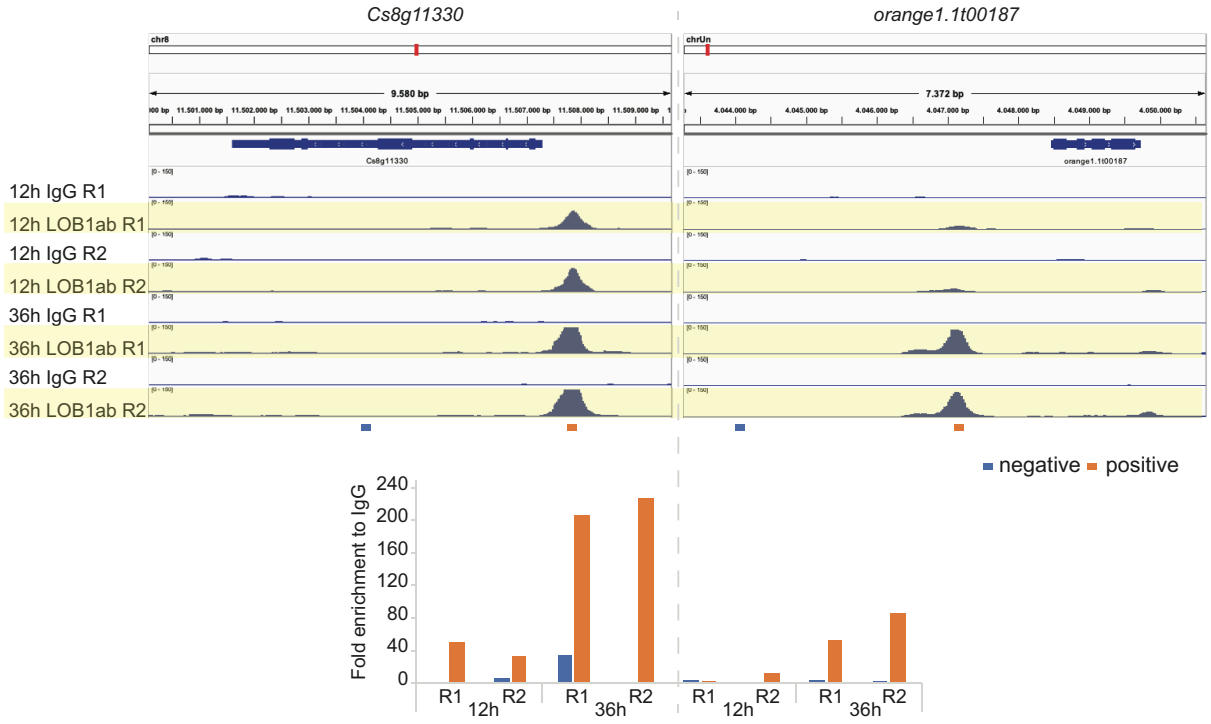


Figure 9: ChIP-qPCR confirms enrichment of CsLOB1-bound regions in the promoters of two representative candidate genes.
Top: The Integrative Genomic Viewer zoom-in snapshots of the two selected genes with up- and down-stream regions (thick blue box: exon, thin blue box: intron/untranslated regions, white arrow: gene direction) and peak enrichment in ChIP-seq samples at 12hpi (12h) and 36hpi (36h) in *Xcc*-infected Citrus leaves using either α-CsLOB1 antibody (LOB) or IgG (IgG) with two independent replicates (R1, R2). **Bottom:** ChIP-qPCR on DNA samples prior to sequencing. 100bp non-enriched regions where no significant peaks were seen on sample 36LOBab R1 are called "negative" while enriched regions with peaks are "positive" (also depicted in the snapshots). Fold enrichment was obtained by normalising the qPCR ct values of the α-CsLOB1 ChIP sample to the IgG control sample.

Consensus CsLOB1 binding peaks between two replicates of each time points were extracted using PreciseTAD package (Stilianoudakis et al., 2022). Both replicates for each time point show high reproducibility of the identified peaks, accounting for 62% and 80% of the total called peaks at 12hpi and 36hpi, respectively (Figure 10).

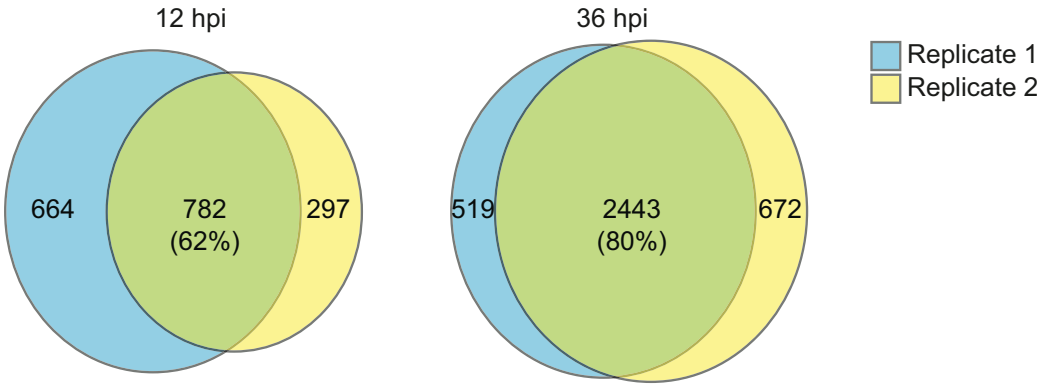


Figure 10: Independent ChIP-seq experiments identified largely identical CsLOB1 binding regions.
Size-proportional Venn diagrams show the overlap of CsLOB1 binding regions (percentage number in parentheses) from independent ChIP-seq experiments performed on *Xcc*-infected Duncan leaves per time point, at 12hpi and 36hpi. CsLOB1 binding regions were obtained from MACS2 (Zhang et al., 2008) by comparing peak counts between α-CsLOB1 and IgG ChIP-seq data. Shared peaks were analysed using the PreciseTAD R package (Stilianoudakis et al., 2022). The number of binding regions in each overlap is shown in bold black.

Peak distribution analysis showed that peaks were predominantly located in the promoter region (upstream of the transcription start site or TSS), with more than 70% located within 3kb region at both 12 and 36hpi (Figure 11A). Peaks overlapping a 4kb window upstream (-4kb) and downstream (+4kb) of the TSS were extracted and the average number of peaks enriched within this window was calculated (Section 4.2.3.6). The analysis showed that peaks were highly enriched around the TSS, extending the confluence to 3kb upstream (-3kb) and 200bp downstream (+200bp) of the TSS (Figure 11B).

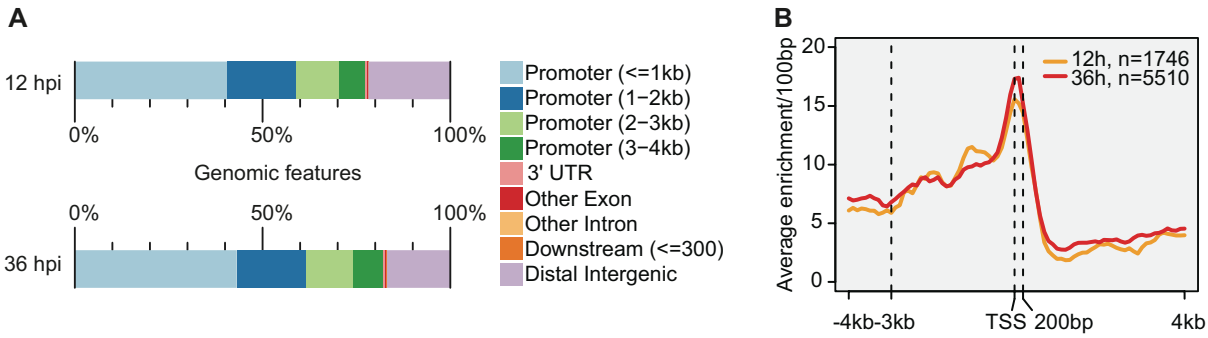


Figure 11: ChIP-seq with CsLOB1-specific antibody uncovers CsLOB1 preferential binding sites near transcription start sites (TSS).
A: CsLOB1 binds predominantly in promoters. Size-proportional bar graphs show the percentage of peaks, representing CsLOB1-binding sites, relative to the location of genes at 12hpi and 36hpi, as depicted. **B: CsLOB1 preferentially binds around the TSS.** The average enrichment (peak counts) was done by Prof. Dr. Chang Liu, counting peaks in a 100bp window in the region -4kb to +4kb of the TSS at 12hpi and 36hpi as plotted against their distance from the TSS. The number (n) of peaks included in the calculation at 12 and 36hpi is shown.

2.2.2 Integration of RNA-seq and ChIP-seq uncovers the association of CsLOB1 binding and upregulated DEGs, many of which target cell wall degradation

Based on the distance-to-TSS analysis (Figure 11B), genes with narrow peaks within -3kb to +200bp of the gene TSS were defined as potential candidates. 438 and 1432 candidate genes at 12hpi and 36hpi, respectively, were found with 418 genes common to both time points (Figure 12). These candidate targets were then compared with RNA-seq datasets and defined 96 genes that were both: transcriptionally up-regulated by CsLOB1 based on RNA-seq and bound by CsLOB1 based on ChIP-seq (Figure 12). Notably, none of the down-regulated genes at 12 and 36 hpi had CsLOB1 binding sites based on ChIP-Seq data, indicating that binding of CsLOB1 generally correlates with elevated rather than reduced transcription.

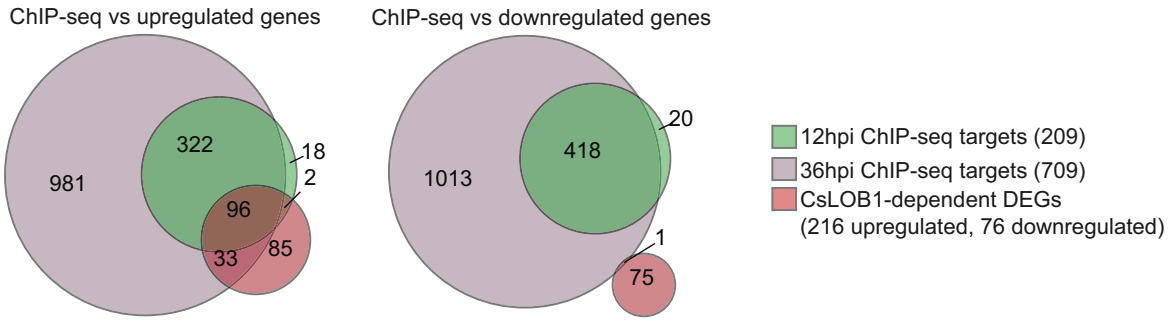


Figure 12: Integration of ChIP-seq and RNA-seq revealed that CsLOB1 binding correlates exclusively with transcription activation, but not repression

Size-proportional Venn diagrams show the overlap of genes interacting with CsLOB1 in the region from -3kb to +200bp relative to the TSS in ChIP-seq at 12hpi and 36hpi compared to CsLOB1-dependent DEGs found in RNA-seq in Figure 4. The total number of genes in ChIP-seq at each time point and in RNA-seq are in parentheses. The number of genes in in each overlap of the size-proportional Venn diagrams are in black.

In the group of 96 common genes, 35 are predicted to encode CWDEs such as pectinesterase, expansin, pectate lyase and other proteins involved in cell wall synthesis and modification (Figure 13). Another functional group also abundant among these 96 genes has cell elongation and cell division related functions such as formin, microtubule-associate protein and patellin (Supplemental Table 3).

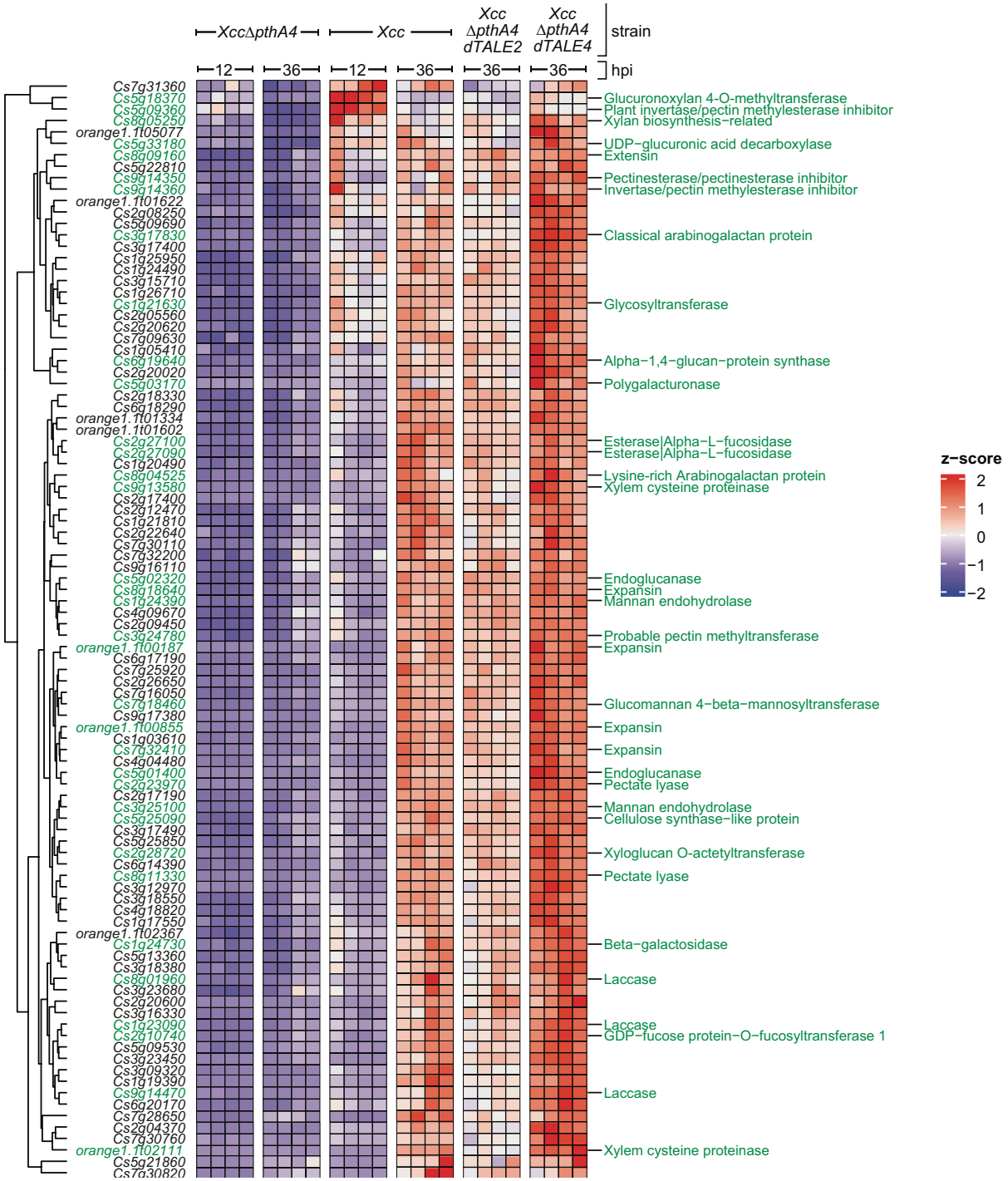


Figure 13: CsLOB1 target genes are mainly involved in cell wall degradation.
The heat map shows the transcript levels of 96 genes in Figure 4A in RNA-seq upon infection with the depicted strains at 12 or 36hpi. All read counts per million (CPM) values of the genes were converted to Z scores using ComplexHeatmap Bioconductor package (Gu et al., 2016) in R. The R script was adapted based Dr. Paulo Teixeira's R script (Appendix 1.5). Genes with similar expression profiles were clustered into subclades. Each coloured square represents the Z-score of each gene in a single biological replicate. Green: cell-wall related genes, black: other function.

Since many of CsLOB1 direct targets are cell wall related enzymes, an attempt was made to detect them at the protein/enzymatic level. Due to the large number of functionally distinct

enzymes, enzyme activity assays would be tedious and time consuming to set up. We therefore decided to look at the abundance of the protein in leaf tissue using a proteomic study. We infected Duncan leaves with either *Xcc* or *XccΔpthA4* mutant, each with 5 biological replicates, and extracted total leaf protein at 5dpi. The extracted total protein samples were subjected to quantitative mass spectrometry using label-free quantification (LFQ)method. We detected 45 out of 96 proteins of our candidates, 24 of which are cell wall enzymes. Importantly, we found 17 of them present in the wild-type *Xcc-,* but not or very little present in the *XccΔpthA4*-infected tissues (Figure 14A).

However, we were unable to detect CsLOB1 protein in these datasets, although we had previously detected CsLOB1 in Duncan leaves using a specific antibody (Figure 14B). A possible explanation is that CsLOB1, as a transcription factor, is generally present at low levels in the cell. This means that the current LFQ setting might not be sensitive enough to detect low levels of CsLOB1 protein in the total protein extract. This could also explain the absence of proteins encoded by other CsLOB1 target genes in the proteomics datasets (Figure 14A). We then attempted to detect CsLOB1 from the total protein extract by fractionating the sample based on size using SDS-PAGE (Sodium dodecyl-sulfate polyacrylamide gel electrophoresis). We then gel extracted the region where CsLOB1 was approximately located based on previous detection with CsLOB1 antibody. The protein was digested with trypsin and subjected to LC-MS/MS coupled to multiple reaction monitoring (MRM) mode. We were able to detect two specific peptides of CsLOB1 using this method, suggesting that the proteomics protocol needs to be optimised for low abundance proteins.

Nevertheless, the differentially abundant proteins such as expansin, pectate lyase and endoglucanase are worthy of further investigation as they might have a more profound effect on the cell wall and disease development (Figure 14A).

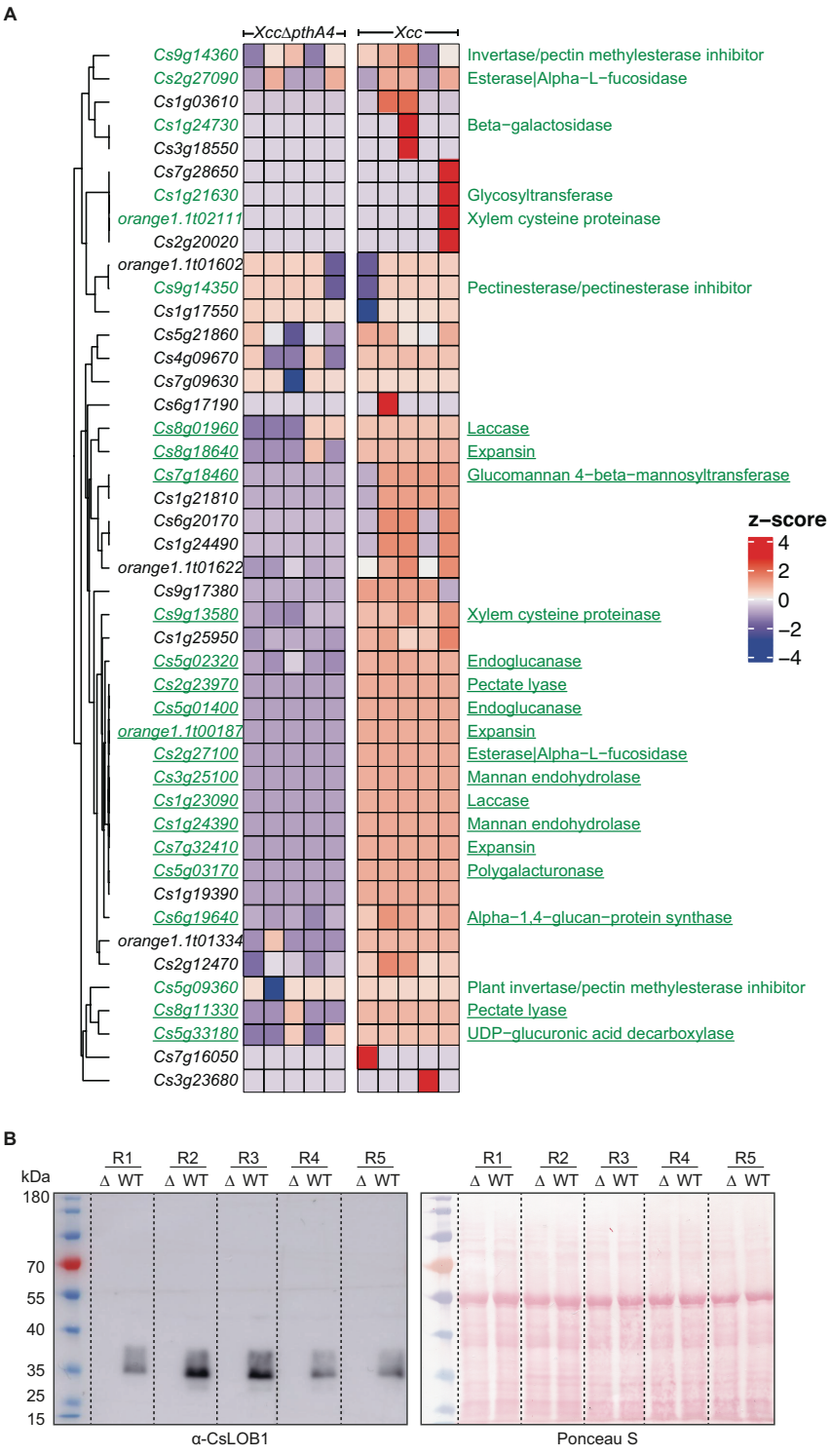


Figure 14: Proteins encoded by CsLOB1 direct target genes were enriched in cell-wall related functions

A: The heat map shows the label-free quantification (LFQ) intensities of 45 proteins as direct targets of CsLOB1 in Figure 13 in RNA-seq upon infection with the depicted strains at 5 days post infection. All LFQ intensity values were converted to Z scores using the ComplexHeatmap Bioconductor package (Gu et al., 2016). The R script was adapted based Dr. Paulo Teixeira's R script (Appendix 1.5). Proteins with similar abundance profiles were clustered into subclades. Each coloured square represents the Z-score of each protein in a single biological replicate. Green: cell-wall related genes, black: other function. Underlined:

proteins differentially abundant in *Xcc* compared to *XccΔpthA4* infected samples. **B: CsLOB1 was detected by Western blot in the *Xcc*- (WT) but not *XccΔpthA4*-(Δ)infected leaf samples sent for proteomic studies.** Western blot was done using α-CsLOB1 primary antibody and HRP-linked α-rabbit secondary antibody visualised by CCD camera (top). Loading control by Ponceau S staining (bottom). The molecular weight of CsLOB1 is 26kDa. R: replicate.

2.3 CsLOB1 activates genes via a 15-bp motif

2.3.1 ChIP-seq identified a conserved 15-bp putative CsLOB1 binding motif

A 300-bp sequence around the ChIP-seq peak summits was extracted and searched for potential CsLOB1 binding motifs using the MEME tool (Bailey and Elkan, 1994). A 15-bp motif was found in all ChIP-seq replicates at both 12 and 36hpi, consisting of two GC-rich highly consensual ends with an almost palindromic pattern C/GCG/A and C/TGG/C (hereafter called C2 and G12 groups, respectively) and a less conserved AT-rich middle region (AT spacer) (Figure 15A). The shared peaks between independent replicates (R1 & R2) or between two time points (12 & 36hpi) resulted in a consistent 13bp motif with a higher occurrence percentage (Figure 15A). This indicates that the identified motif is highly conserved and reproducible. For simplicity, we have used the logo from R1&R2 at 12hpi for later illustrations.

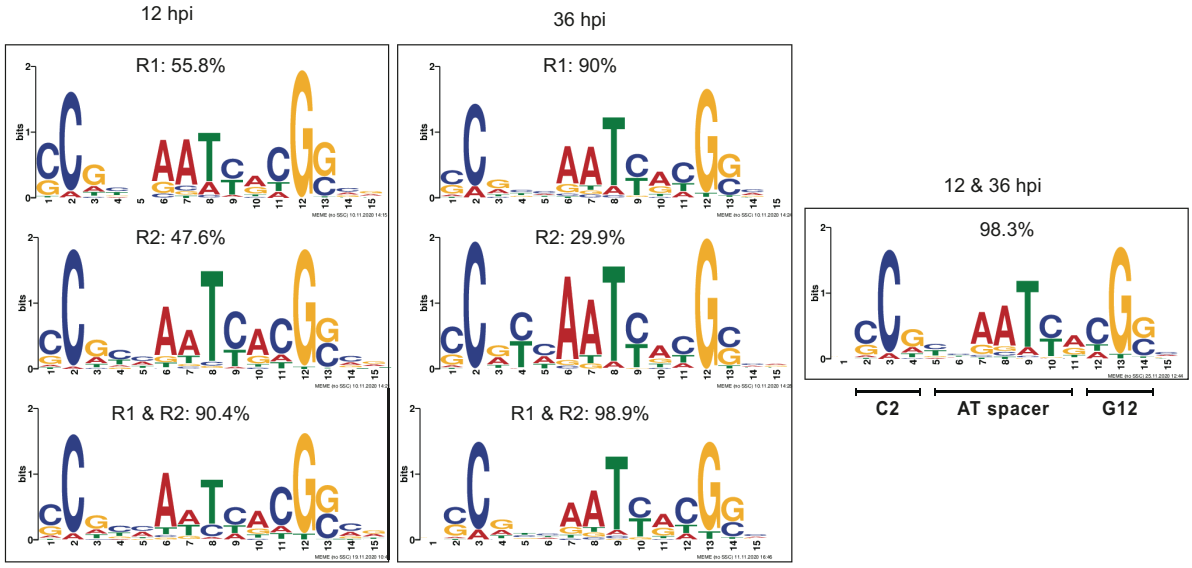


Figure 15: A reproducible 15-bp DNA-binding motif was enriched across different independent ChIP-seq replicates and shows a nearly palindromic pattern.
Logos show conserved motifs derived by the MEME-suite tool (Bailey and Elkan, 1994). Percentage show how many peaks containing the predicted motif in the total number of peaks. R1, R2: replicate 1, replicate 2. R1 & R2, 12&36hpi: shared peaks between two replicates or two time points, respectively. C2, G12: GC rich highly conserved ends, AT spacer: AT rich, less conserved spacer. Each logo is a position-dependent base-probability matrix represented as letters with proportional size corresponding to the base probability in the motif sequence.

2.3.2 CsLOB1 binds to the putative motif *in vitro*

The motif has an interesting consensus palindromic pattern in which only a few positions are highly conserved, specifically position 2 with cytosine (C2) and position 12 with guanine (G12), while the rest has intermediate or much lower specification, especially in the spacer region enriched in adenine (A) and thymine (T). To validate the binding specificity of CsLOB1 to the predicted motif, we tested the importance of the highly conserved and less conserved bases for CsLOB1 binding in an in vitro DNA-protein binding assay called EMSA (electrophoretic mobility shift assay, Section 4.2.5.1). We selected a putative binding site from one of the candidate promoters in our 96-gene list - *Cs2g20600* (wt, Figure 16A). To test how changes in the motif consensus (shown in Figure 16A) affect CsLOB1 binding, we generated different motif variants. Briefly, these include mutations in a GC-rich end (m1), the two highly conserved C2 and G12 (m2), the single C2 (m3) or G12 (m4), the single base at position 6 in the spacer (m5), the 3 bases in the spacer (m6), 1bp insertion (in) and 1bp deletion (d1) (Figure 16A). The flanking bases of the 15bp motif were also replaced with completely different types of bases (flk) (Figure 16A) to examine whether such changes affect CsLOB1 binding.

Due to the insolubility of the histidine-tagged CsLOB1 protein (Figure 7), we designed a new recombinant CsLOB1 construct with a C-terminal maltose-binding protein (MBP) that could be expressed in *E. coli* and purified in a soluble state (Figure 16B). EMSA was performed with the purified protein and a Cy5-labelled 33-bp double-stranded wt DNA probe and showed a band shift relative to samples without added CsLOB1 protein, indicating that CsLOB1 indeed binds to the Cy5-labelled wt probe (Figure 16C, top left gel, first two lanes). When unlabelled wt probe was added in excess, we observed a dramatic decrease in bound probe, indicating that the unlabelled wt probe can outcompete the Cy5-wt probe (Figure 16C, top left gel, last three lanes). In contrast, an unlabelled competitor with an unrelated DNA sequence could not outcompete the Cy5-labelled wt probe, indicating that the competitor DNA had a lower affinity for CsLOB1 than the Cy5-labelled wt probe (Figure 16C, top right gel). Taken together, this indicates that: 1) the recombinant CsLOB1-MBP fusion protein is capable of binding to double-stranded DNA and 2) that CsLOB1 binds to the predicted binding site in a sequence-specific fashion.

We then performed competitive EMSA with competitors containing the different mutated sites mentioned above and shown in Figure 16A. The mutation in a GC-rich end (m1) abolished its functionality as a competitor, indicating that a mutation of this base reduces the affinity of the probe for CsLOB1 (Figure 16C). Interestingly, the loss of affinity to CsLOB1 was also observed for mutations of two highly conserved bases at positions 2 (C2) and 12 (G2) (m2), suggesting that both C2 and G12 are crucial for interaction with CsLOB1 (Figure 16C). Surprisingly, a separate C to T mutation at either position 2 (m3) or a G to A mutation at position 12 (m4) both abolished binding of competitor DNA to CsLOB1 (Figure 16C), highlighting the importance of these highly conserved nucleotides for interaction with CsLOB1. On the other hand, a competitor carrying a mutation at position 6 in the spacer, where A was mutated to G (m5), was able to compete with the Cy5-wt probe (Figure 16C) and thus this nucleotide had little or no impact on the capacity to interact with CsLOB1. However, substitution of the moderate consensus 3 bases in the spacer (AAA) to GCG (m6) eliminated CsLOB1 binding to the competitor, suggesting a role of the AT-rich region in binding affinity (Figure 16C).

We also studied motif derivatives in which the size of the spacer between the two GC-rich triplet motifs was altered, since in DNA-binding motifs not only the base composition but also the length of the motif spacer is often important. We found that a single base pair deletion or insertion in the spacer region (in or d1, respectively) significantly reduce binding affinity for CsLOB1 (Figure 16C), highlighting the importance of the 9-bp distance between C2 and G12. In contrast, a mutant probe in which the motif-flanking sequences were replaced by different bases (flk) did not show different binding compared to the WT probe (Figure 16C), indicating that the 15-bp motif is self-sufficient for *in vitro* CsLOB1 binding and appears to function in a context-independent fashion. Based on the importance of the positions C2 and G12 and the non-negotiable 9-bp AT-rich spacing between the two, we hereafter refer the newly-identified motif as the C9G motif for simplicity.

The C9G CsLOB1 binding motif that we identified and experimentally validated by EMSA shows similarity to the binding motifs of *Arabidopsis* and Maize LOB homologs (Figure 2) (O'Malley et al., 2016, Hou et al., 2023). A previous study also reported on two putative 10- and 14-nucleotide GC-rich CsLOB1 binding motifs (GCCGCGGCGG and CGGCGGCGGCGGCG), but these were not functionally validated (Zou et al., 2021). This previously reported GC-rich

binding motif shows no obvious similarity to the CsLOB1 binding motif we identified. To functionally study this GC-rich putative CsLOB1 binding motif, we generated a competitor probe containing this GC-rich motif instead of the C9G motif present in our Cy5-labelled wt probe (Z1, Z2, Figure 16D). As a control, we also generate the optimised probe (opt) containing the sequence closest to our predicted motif (Figure 16A). While both the wt and opt probes outcompeted the Cy5-wt probe even at the lowest concentration (Figure 16C), Z1 and Z2 did not outcompete even at the highest concentration, suggesting that they have a much lower binding affinity to CsLOB1 compared to our identified site (Figure 16D).

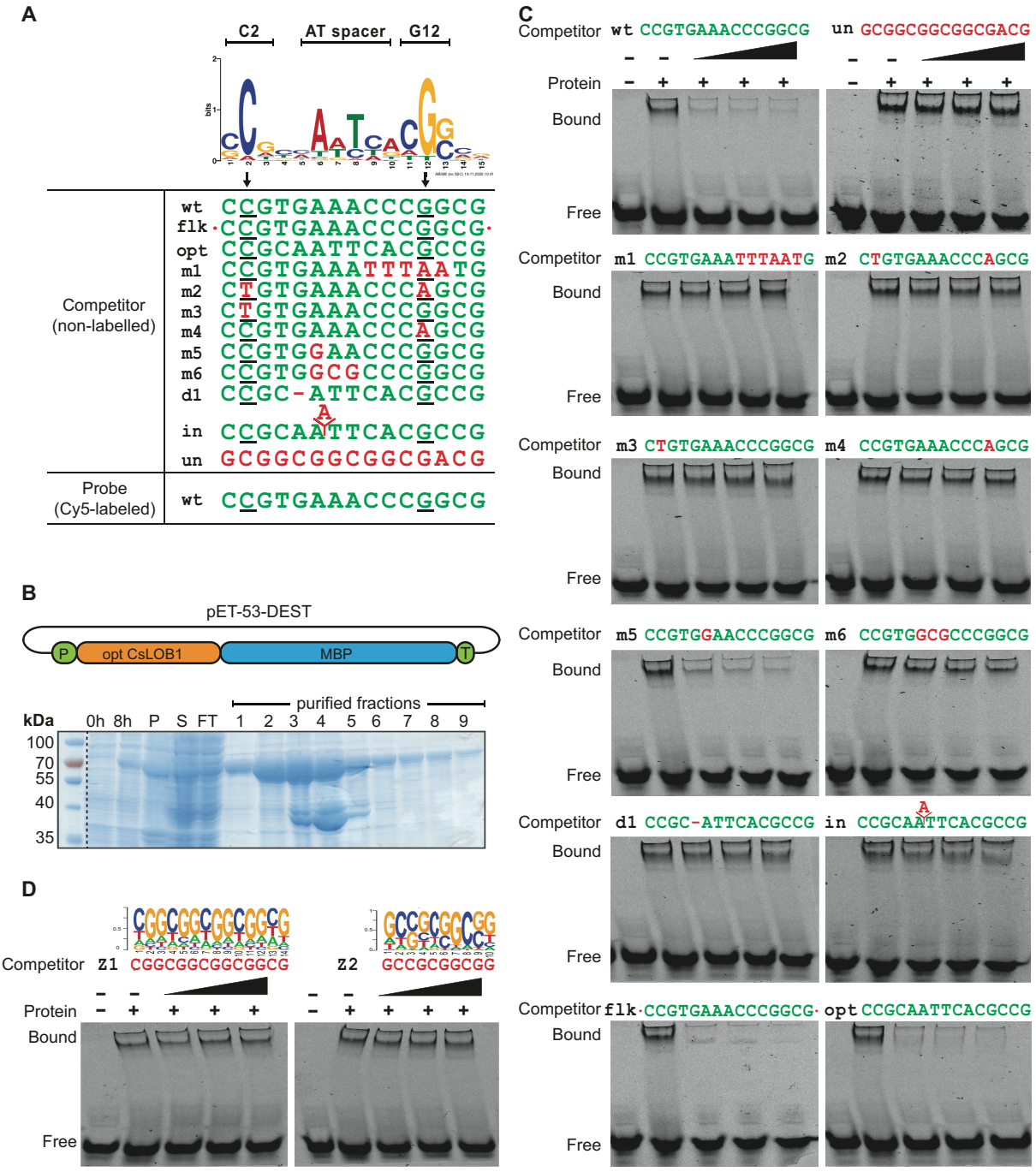


Figure 16: Competitive Electrophoretic Mobility Shift Assays (EMSA) show that mutations in the C9G motif reduce the affinity for CsLOB1

A: Representative predicted binding site of CsLOB1 and its derivatives. The 33bp predicted binding sequence in the CsLOB1 target gene promoter *Cs2g20600* (wt) was used as a Cy5-labelled probe (only the C9G motif sequence is shown in green). The affinity of the depicted unlabelled probes (green: wt bases, red: mutations, deletion or insertion) was studied in a competitive EMSA. Black arrows and underlines indicate the position of highly conserved bases. C2, G12: highly conserved bases at the 2nd and 12th positions of the depicted motif; AT spacer: less conserved sequence and AT-rich in motif centre region. **B/C: EMSA show the binding specificity of CsLOB1 to the predicted motif but not to the mutant derivatives.** B. Schematic showing *CsLOB1-MBP* gene construct for recombinant expression in *E. coli.* opt: CsLOB1 coding sequence with codons optimised for expression in *E. coli*. All components are assembled in the pET-53-DEST vector. The construct was expressed in the arabinose- and IPTG-inducible *E. coli BL21-AI* strain and purified on an amylose resin column (Coomassie stained gel, Section 4.2.3.1). 0h: bacterial

lysate before induction, 8h: bacterial lysate 8 hours post induction, P: pellet, S: supernatant fraction for purification, FT: flow-through fraction, W: wash fraction. Purified fractions number 1,2,6,7,8,9 were pooled and dialysed against EMSA buffer. CsLOB1-MBP molecular weight is approx. 70kDa. C/D. EMSA with purified CsLOB1-MBP at 0µM (-) or 2µM (+) and 200nM Cy5-labelled 33-bp wt probe with unlabelled competitors of the depicted sequence at the concentrations of 0 (-), 25x, 50x, 100x molar excess relative to Cy5-probe (triangle). D: The two GC-rich motifs found in (Zou et al., 2021) replace the 15-bp C9G CsLOB1 motif sequence in the 33bp competitor probe (Z1, Z2, sequence in red font, motif logos shown). Gels were imaged using Typhoon™ (Cytiva) with the default Cy5 fluorophore setting.

To quantify the affinity of CsLOB1 to the predicted binding site and mutant derivatives, we measured the equilibrium dissociation constant (K_D) between CsLOB1-MBP and Cy5-labelled probes using MicroScale Thermophoresis (MST, Section 4.2.5.2) (Figure 17A). The K_D values of the mutant derivatives m1, m2, m3, m4 and m6 appeared to be at least 10-fold higher than those of wt and the m5 probe, indicating at least 10-fold lower affinity of the mutated motifs to CsLOB1 (Figure 17B&C). This is in agreement with our competitive EMSA where the m5 probe with one base mutation in the spacer showed a binding capacity to CsLOB1 similar to the wt probe, whereas the other mutants were no longer able to bind to CsLOB1 (Figure 16C).

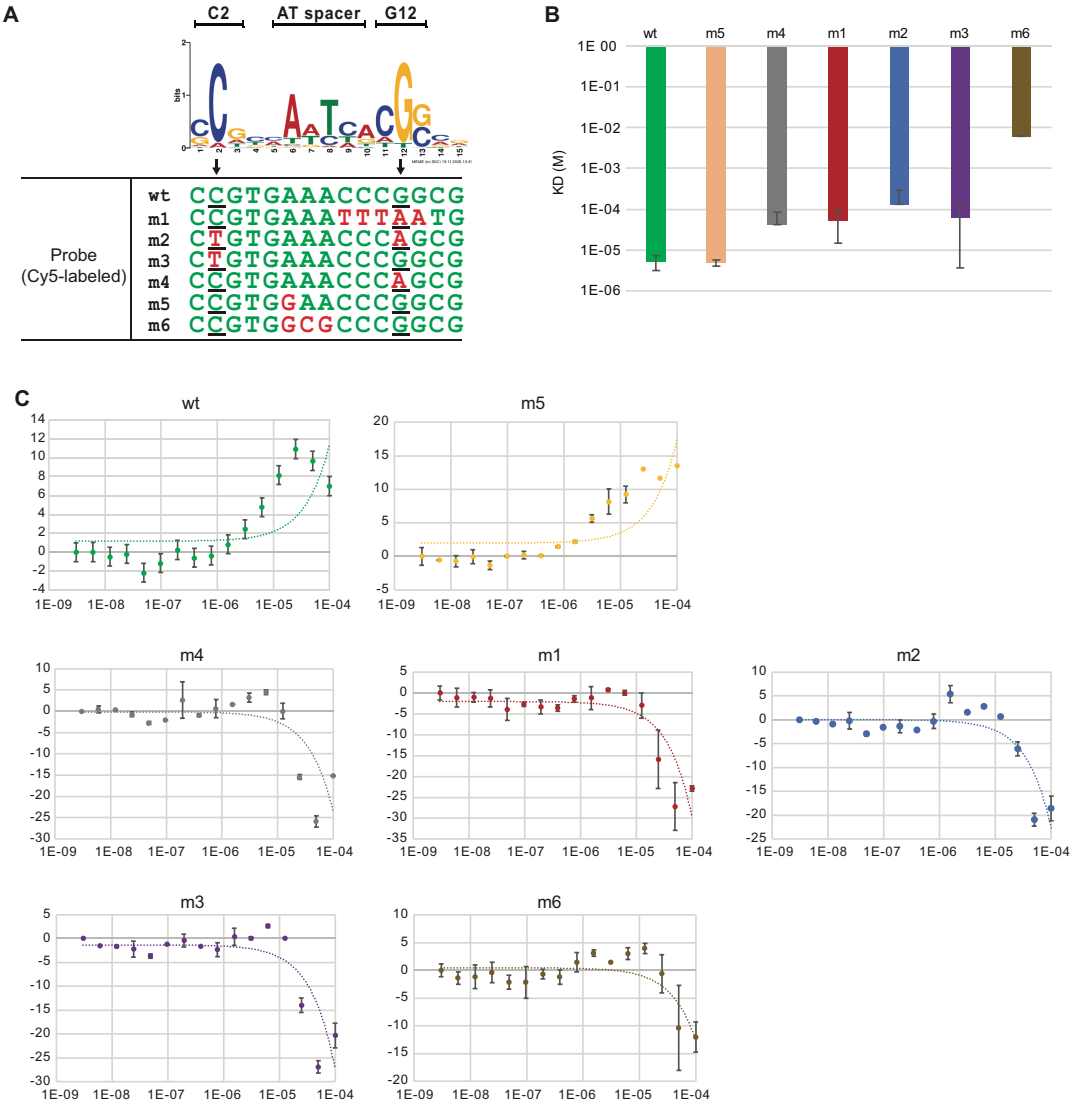


Figure 17: MicroScale Thermophoresis (MST) shows that the affinity of the C9G motif to CsLOB1 decreases >10-fold if conserved bases of the motif are mutated

A: **Sequences of probes previously used in EMSA (Figure 15A) labelled with Cy5.** wt: wild-type sequence, m: mutated. Green: wild-type bases, red: mutated bases. Black arrows and underlines indicate the position of the highly conserved bases, C2 and G12. **B: Probes with wt and a single base mutation in the AT spacer (m5) have equilibrium dissociation constants (K_D, in molar) more than 10 times lower than mutations the conserved bases** (average values from of 2-3 independent replicates). K_D value was extracted from MST analysis using the MO. Affinity Analysis v2.3 (Section 4.2.5.2). **C: Only probes with wt and a single base mutation in the AT spacer have CsLOB1-dependent increase in DNA-protein complex formation.** The graph shows the delta normalised fluorescence (ΔF_{norm}, y axis) of Cy5 probe in a serial dilution of CsLOB1 protein (x-axis). The higher the value, the more protein-probe complexes are formed. Each data point represents an average of 2-3 independent replicates with standard deviation bar. Each probe type is colour coded as depicted.

In brief, the EMSA and MST assays have validated the binding of CsLOB1 to the predicted C9G motif in a highly specific manner. Furthermore, these two *in-vitro* DNA-protein binding assays suggest that CsLOB1 does not require other accessory proteins for binding to this sequence. As LOB domain containing GAS block and leucine rich domain is able to form

dimers with other proteins (Chen et al., 2019), the *in-vitro* assays indicate that CsLOB1 form homodimers when binding to the C9G bp motif, supported by the nearly palindromic motif sequence.

2.3.3 *In-planta* loss-of-function and gain-of-function confirmed CsLOB1 C9G binding motif

Recombinant CsLOB1 was shown to bind to the representative C9G motif *in vitro* with high sequence specificity (Section 2.3.2). To validate the interaction between CsLOB1 and the predicted motif *in-planta,* we used promoter-reporter assays in *N. benthamiana* leaves. We cloned the 1.4kb upstream of the *ATG* of the CsLOB1-inducible gene *Cs2g20600,* which contains the predicted CsLOB1 binding site used in EMSA (1.2kb upstream of *Cs2g20600 ATG*), upstream of the β-glucuronidase (*uidA*) gene encoding a GUS reporter into a T-DNA vector. We also generated a promoter derivative in which the predicted CsLOB1 binding site was mutated at either one of the two GC-rich ends (m1-15) or at both GC-rich ends (mX-15) (Figure 18A). The promoter-reporter constructs were delivered into *N. benthamiana* leaves together with either *p35S-driven CsLOB1-GFP* or an empty vector (EV) control. The GUS assay showed that CsLOB1-GFP, but not the EV control, could activate the wild-type *Cs2g20600* promoter (wt-15). In contrast, CsLOB1-GFP did not activate the promoter derivatives in which either one GC-rich end (m1-15) or both GC-rich ends (mX-15) were mutated (Figure 18A).

EMSA showed that the motif is self-sufficient (flk, Figure 16C), so we investigated whether introducing the motif into a minimal promoter would result in a CsLOB1-inducible promoter. We chose the 343bp sequence upstream of the ATG of the pepper (*Capsicum annum*) *Bs3* executor resistance gene, which has been shown to be transcribed exclusively in the presence of the matching TALE protein AvrBs3 (Romer et al., 2007). We introduced into the Bs3 promoter either the 15-bp C9G motif sequence (wt-15), its mutant derivative (m1-15), or the entire 33-bp EMSA probe sequences (wt-33 & m1-33) at 160-bp upstream of the ATG of *pBs3*. These *Bs3* promoter derivatives were cloned to drive a downstream RUBY reporter consisting of three tandem-arranged betalain biosynthetic genes separated by a translational self-cleaving T2A peptide (Figure 18B) (He et al., 2020). Each T-DNA construct was co-delivered with either p35S-*driven CsLOB1-GFP* or empty vector into *N. benthamiana* leaves

via *Agrobacterium*. These assays showed that CsLOB1 was able to activate Bs3p derivatives containing the wt-15- or wt-33 motifs. In contrast, CsLOB1 did not activate the m1-15- and m1-33- containing *pBs3* mutant derivatives (Figure 18B). These data suggest that the identified C9G motif binds to CsLOB1 in a context-independent fashion. In conclusion, the loss- and gain-of-function promoter-reporter assays demonstrate that the identified C9G motif mediates CsLOB1 binding *in-planta* and are consistent with the results of the *in-vitro* assays on the CsLOB1-binding motif (Figure 16, Figure 17) that was originally identified by ChIP-seq (Figure 15).

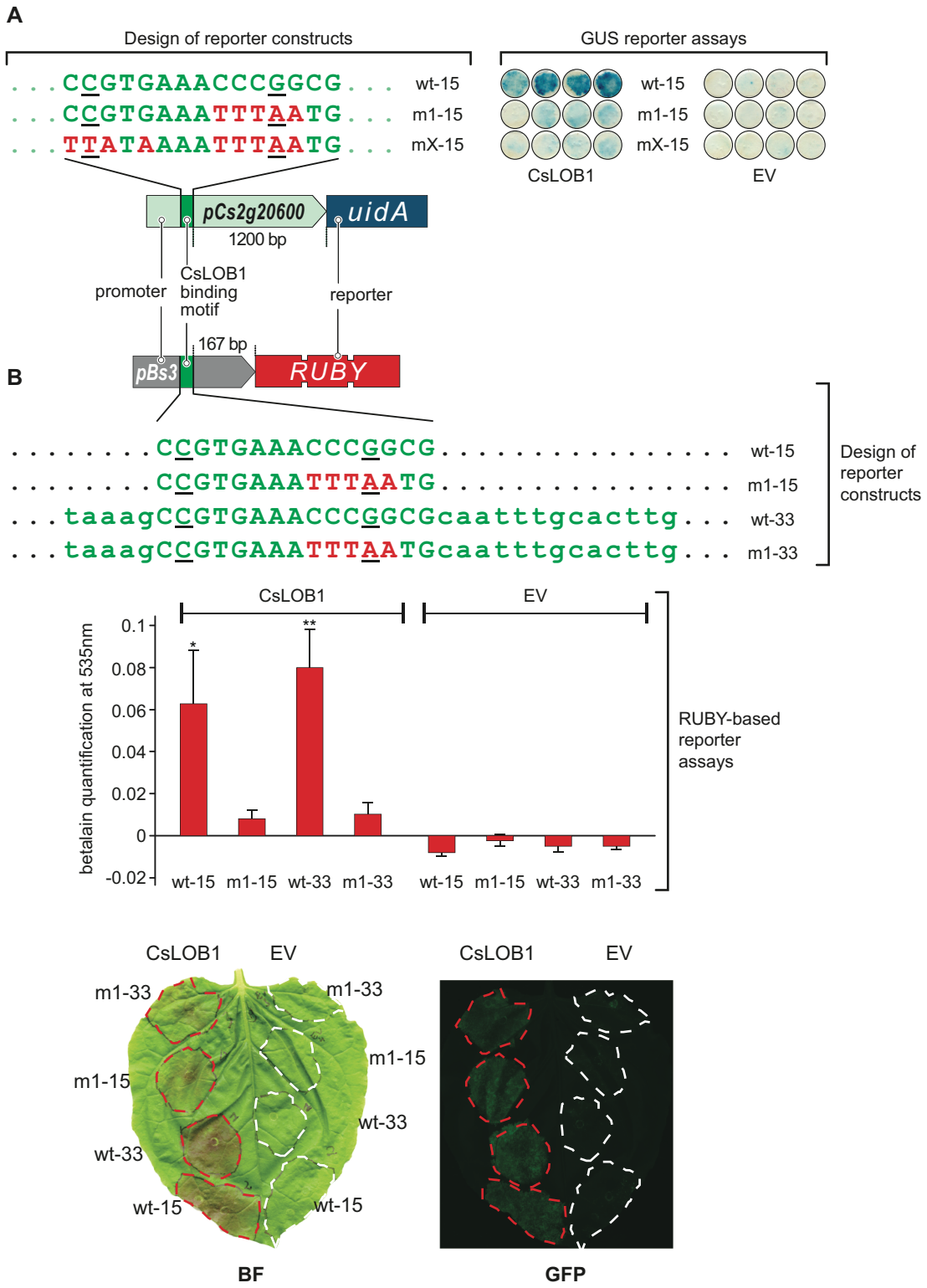


Figure 18: Loss- and gain-of-function assays confirm that the newly identified C9G DNA motif mediates CsLOB1-dependent promoter activation *in-planta*

A. Mutations within the predicted CsLOB1 binding site of a CsLOB1-inducible target promoter abolish CsLOB1-mediated activation in *N. benthamiana*. Schematic depiction of the promoter-reporter constructs in which *uidA*, encoding a GUS reporter is driven by the upstream sequence (p: promoter, 1.4kb) of the CsLOB1-inducible gene *Cs2g20600* gene containing the predicted CsLOB1 binding motif (wt-15) or its mutant derivatives (m1-15 & mX-15). The depicted reporter genes were introduced into *N. benthamiana* leaves by *Agrobacterium*, together with either p35S-promoter-driven *CsLOB1* (CsLOB1) or

empty vector control (EV). Inoculated leaf discs were collected at 2dpi for GUS staining (Section 4.2.4.4). **B. Insertion of the predicted CsLOB1 binding site into the *Bs3* promoter (*pBs3*) confers CsLOB1 inducibility.** Schematic depiction of the reporter constructs used for the gain-of-function assay. The grey arrow indicates *pBs3* driving a downstream RUBY reporter consisting of three distinct tandem-arranged betalain biosynthesis genes separated by T2A peptides (He et al., 2020). Either the predicted C9G CsLOB1 binding motif (wt-15) or the 33bp EMSA probe containing the motif-flanking sequence (wt-33) and their mutant derivatives (m1-15 & m1-33) were incorporated 160-bp upstream of the ATG. Reporter constructs shown were delivered into *N. benthamiana* leaves by *Agrobacterium* together with either p35S-promoter-driven *CsLOB1-GFP* (CsLOB1) or empty vector control (EV). Samples were harvested at 42hpi. Extracted betalain was quantified via photometric measurement at 535nm (Section 4.2.4.5). Error bars represent the standard deviation from 4 biological replicates. Photos were taken at 42hpi in a bright field (BF) for RUBY observation. CsLOB1-GFP expression was examined using Typhoon™ (Cytiva) with default GFP fluorophore setting (GFP). Green letters: wild-type bases in the 15-bp binding site. Underline: C2 and G12. Red letters: mutated bases. Lowercase green letters: motif-flanking nucleotides that were part of the EMSA probes. Student's t-test was used to calculate the significant difference between wt-15 and m1-15, and between wt-33 and m1-33 in the presence of CsLOB1. *: $p \leq 0.05$, **: $p \leq 0.01$.

2.3.4 CsLOB1 motif exhibits a synergetic mode of action

Using the CsLOB1 motif and the FIMO tool (Grant et al., 2011), potential binding sites are able to be identified within a given promoter. However, this process typically results in multiple hits in the narrow peak region interacting with CsLOB1 but have different predicted binding scores ($p<0.01$) (Figure 19A). We dissected the relative distance of the predicted motifs to each other in the promoters of seven candidate promoters (four shown in Figure 19A as an illustration) and observed that the motifs have a mean distance of about 90-bp (Figure 19B).

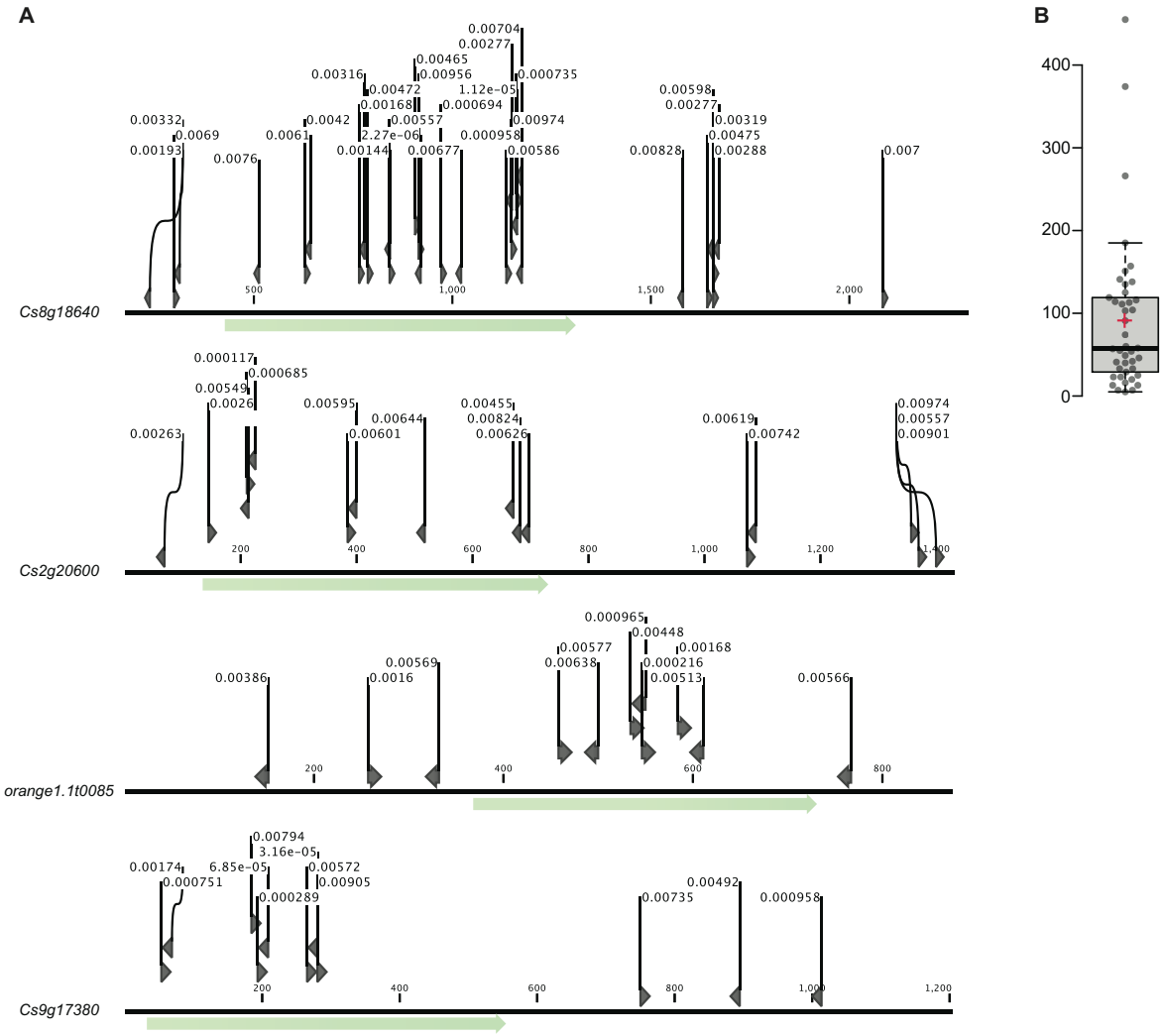


Figure 19: Predicted motifs are enriched in candidate promoters

A: Predicted motifs are enriched in the narrow peak regions (green arrow) obtained from ChIP-seq. Upstream regions (from the ATG) of four direct target genes of CsLOB1 that contain the narrow peaks shown in ChIP-seq to interact with CsLOB1 (black lines). Black arrow heads: predicted CsLOB1 motifs from FIMO motif scanning (Grant et al., 2011). Numbers represent the p value (<0.01). Green arrow: narrow peak regions found in ChIP-seq. **B: Predicted motifs (p<0.01) in seven candidate promoters (four genes annotated in A, the other three are *Cs1g19390*, *Cs2g23970*, *Cs5g20320*) have an average distance to each other approx. 90-bp.** Box plots show distance between two closest predicted motifs in bp from seven candidate promoters. Motifs overlapping each other were counted as one. Centre lines show the medians; box limits show the 25th and 75th percentiles as determined by R; whiskers extend 1.5 times the interquartile range of the 25th and 75th percentiles; grey dots indicate individual data points; red sign: mean value.

We questioned whether the presence of multiple motifs would enhance CsLOB1 activation. We tested whether the integration of several CsLOB1-binding motifs into the *Bs3* minimal promoter could enhance CsLOB1-mediated activation. Specifically, we cloned two tandem copies of the 33-bp motif sequence, separated by a sequence of approx. 90 base pairs from the same *Cs2g20600* promoter, into the same minimal promoter-RUBY reporter system. As a

control, we created variants using the mutant derivative of the CsLOB1 binding motif (m1 & mX) instead of the WT (wt) (Figure 20A). We then introduced these promoter-reporter variants into *N. benthamiana* via *Agrobacterium* together with either p35S-*driven CsLOB1* or EV. As expected, the promoters containing the wt motif sequence were activated by CsLOB1, whereas the mutant sequence was not (Figure 20B). Quantification of the RUBY reporter showed that the construct containing two motifs had approx. two times higher CsLOB1-dependent reporter activation than the construct containing one motif. By contrast, constructs containing mutant derivatives (mX) did not produce significant reporter activity regardless of the number of motifs in the promoter (Figure 20C). These finding support our hypothesis that there is a direct correlation between the number of CsLOB1 binding elements and the level of CsLOB1-dependent promoter activation.

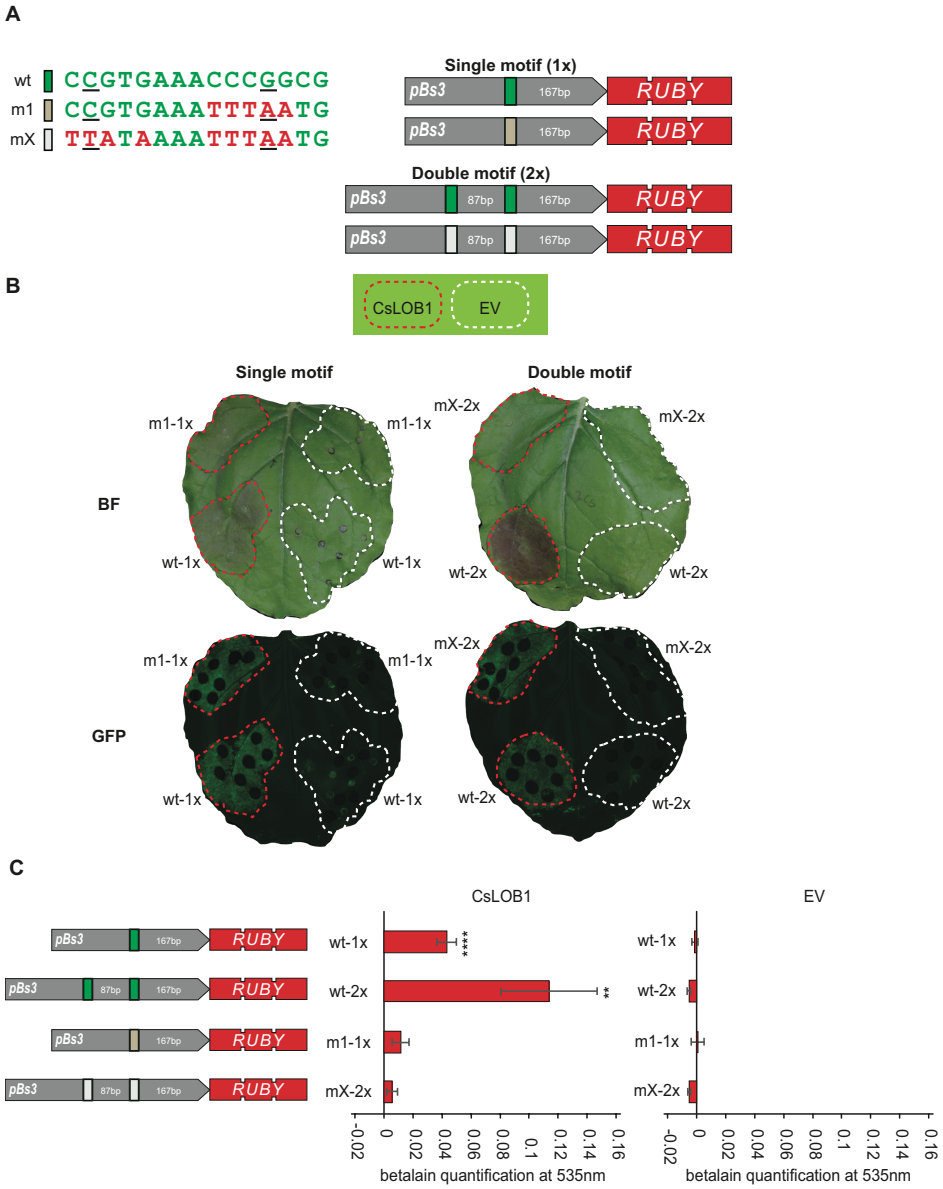


Figure 20: Duplicate motifs enhance CsLOB1-mediated promoter activation

A: Schematic representation of the reporter construct used for the gain-of-function oligo-motif assay. The grey arrow indicates *pBs3* driving a downstream RUBY reporter used in We incorporated either once (single motif -1x) or twice (double motif -2x) the 33bp EMSA probe containing motif-flanking sequence (wt-only 15bp motif sequence shown, see Figure 18) and mutant derivatives (m1 & mX). The distance between the motifs or between the motif and the start codon ATG is shown in white. The nucleotide sequence of the predicted C9G CsLOB1 binding motif is shown in green. Mutations are shown in red. **B/C: RUBY signal was elevated in the double-motif containing promoter twice as much as in the single-motif containing promoter.** B: Reporter genes shown were delivered into *N. benthamiana* leaves by *Agrobacterium* together with either p35S-promoter-driven *CsLOB1-GFP* (CsLOB1) or an empty vector control (EV) as depicted. Photos were taken at 42hpi in bright field (BF) for RUBY observation. CsLOB1-GFP expression was examined using Typhoon™ (Cytiva) with default GFP fluorophore setting (GFP). C: Samples were harvested at 42hpi. Extracted betalain was quantified by photometric measurement at 535nm. Error bars represent the standard deviation of 4 biological replicates. Student's t-test was used to calculate the significant difference between wt-15 and m1-15, and between wt-33 and m1-33 in the presence of CsLOB1. **: $p \leq 0.01$, ****: $p \leq 0.0001$

2.3.5 *CsLOB1* closest homologs in Citrus and *Arabidopsis thaliana* activate CsLOB1 binding motif

We took the advantage of the promoter-RUBY construct containing two CsLOB1 motifs that has enhanced activation to investigate some of the closest homologs of CsLOB1 (Figure 20). In citrus, *CsLOB2 (Cs7g27620)* and *CsLOB3 (Cs8g17160)* were shown to be the closest homologs of *CsLOB1* and are able to induce water soaking and pustules when activated by dTALEs in the background of *XccΔpthA4* (Zhang et al., 2017a). They also shared a highly similar LOB domain (Figure 21) (Zhang et al., 2017a); therefore, it is likely that they also activate similar motifs. In contrast, another Citrus LOB homolog, CsLOB4 (*Cs7g30620*) is more distant from CsLOB1 (Figure 21) and was shown by Zhang and colleagues not to restore Citrus canker symptoms.

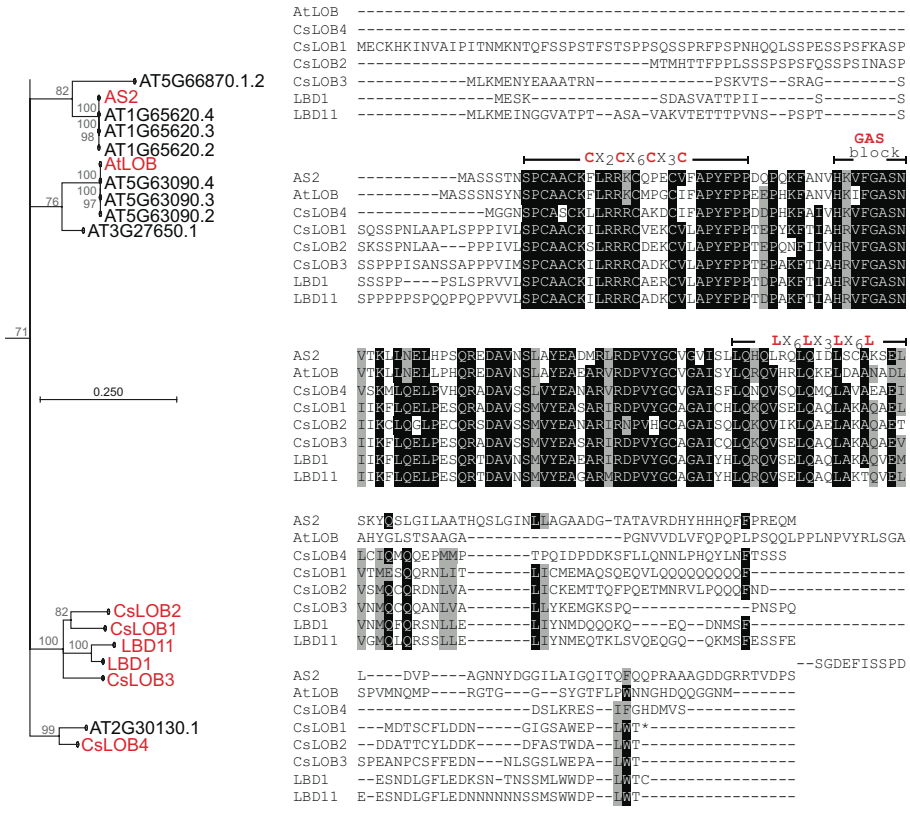


Figure 21: CsLOB2, CsLOB3, Arabidopsis LBD1 and LBD11 are the closest homologs of CsLOB1.

Left: Simplified phylogenetic tree of Arabidopsis LOB proteins related to CsLOB1, CsLOB2 (Cs7g27620), CsLOB3 (Cs8g17160) and CsLOB4 (Cs7g30620) (CsLOB1-4). CsLOB1-4, LBD1 (AT1G07900) and LBD11 (At2g28500), AS2 (At1g65620) and AtLOB (At5g63090) are shown in red. Bootstrap values are in grey. The phylogenetic tree was constructed using Qiagen CLC Main Workbench with default settings. The complete phylogenetic tree of 43 Arabidopsis LOB proteins and CsLOB1-4 is in Supplemental Figure 7. **Right: Full**

length protein alignment of the close homologs CsLOB1-4, LBD1 and LBD11 in comparison to more distant homologs, AS2 and AtLOB. LOB motif indicated with a cysteine-rich DNA-binding domain (CX2CX6CX3C), and protein dimerization domain: GAS block and leucine zipper-like domain (LX6LX3LX6L). LOB-domain motif features are shown below with conserved amino acids in red, X letter with number indicates how many non-conserved amino acids occur in between. Letters highlighted in black indicate identical amino acids. Non-highlighted letters are non-identical amino acids. Hyphens show gaps.

We cloned *CsLOB2* and *CsLOB3* coding sequences driven by *p35S* and tagged with a C-terminal GFP in the same backbone as CsLOB1 construct used in Figure 18 & Figure 20. We co-expressed the RUBY construct driven by the *pBs3* derivative containing two CsLOB1 wt motifs with either *p35S-driven CsLOB1-GFP*, *p35S-driven CsLOB2-GFP*, *p35S-driven CsLOB3-GFP* or EV control into *N. benthamiana* via *Agrobacterium*. As expected, we observed RUBY activation in the co-expression of CsLOB1, CsLOB2 or CsLOB3 but not in the EV control (Figure 22, left). In contrast, we did not observe visible RUBY signal when co-expressing CsLOB1-3 with the RUBY construct driven by the *pBs3* derivative containing two mutated CsLOB1 motifs (mX) (Figure 22, right). The ability to activate CsLOB1-motif-containg promoter suggests that CsLOB2 and CsLOB3 might share very similar downstream regulons as CsLOB1, which explains why they are able to promote water soaking and pustules in Citrus (Zhang et al., 2017a).

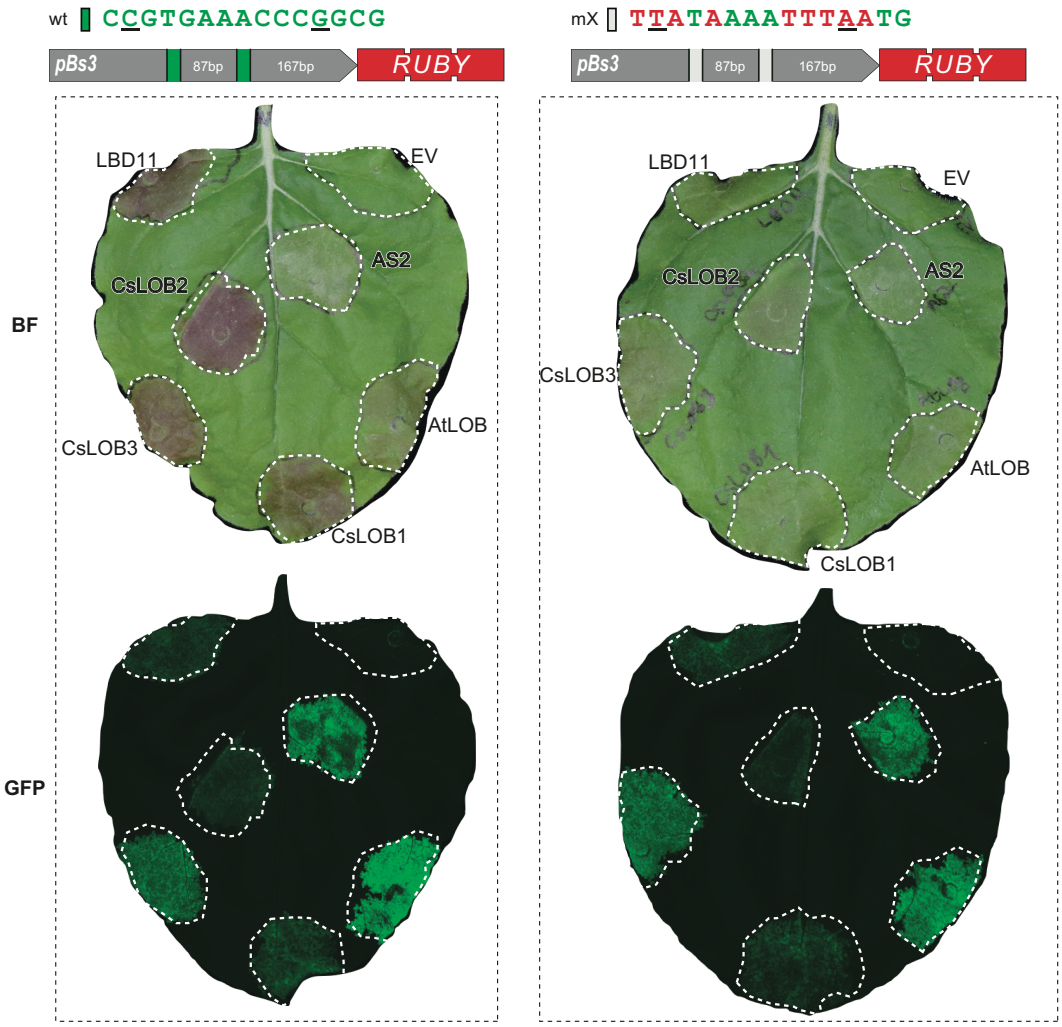


Figure 22: Homologs of CsLOB1 can activated minimal *Bs3* promoter in *N. benthamiana*.
Top panel: Schematic representation of a promoter-reporter construct using the RUBY reporter consisting of three distinct tandem-arranged betalain biosynthesis genes separated by T2A peptides (He et al., 2020). RUBY was driven by the minimal promoter *Bs3* containing a duplicate of CsLOB1 motif (wt) or its mutated variant (mX). Wild-type nucleotides are shown in green; mutations are in red. **Bottom panel: Expressions of LOB homologs activated RUBY reporter driven by *pBs3* promoter containing CsLOB1 wt motif.** Representative *N. benthamiana* leaves co-infiltrated with *Agrobacterium* delivering wt or motif-containing *pBs3* driving the RUBY reporter and p35S-*driven CsLOB1, CsLOB2, CsLOB3*, *Arabidopsis LBD11, AS2, AtLOB* or empty vector (EV) marked by white-dashed lines. Photos were taken at 42hpi in bright field (BF) for RUBY observation. All transcription factors were tagged with C-terminal GFP. Expression of transcription factors was examined using Typhoon™ (Cytiva) with the default GFP fluorophore setting (bottom).

In contrast, CsLOB4 is an interesting case. Zhang et al. (2016) showed that a dTALE targeting *CsLOB4* in the background of *XccΔpthA4* did not restore the ability to induce pustules and water soaking in citrus. However, when using the two different dTALEs targeting *CsLOB4*, we observed the development of citrus canker, though the first appearance was delayed at 14dpi as compared to 5-6dpi when *CsLOB1*, *CsLOB2* and *CsLOB3* were activated (Figure 23).

This suggests that CsLOB4 might be able to induce similar downstream genes required for citrus canker development, though the induction is slower. The difference in the LOB domain of CsLOB4 compared to that of CsLOB1 (Figure 21) could suggest that it might have different binding affinity to CsLOB1 motif. Due to time limitation, we have not addressed this point in the scope of the thesis.

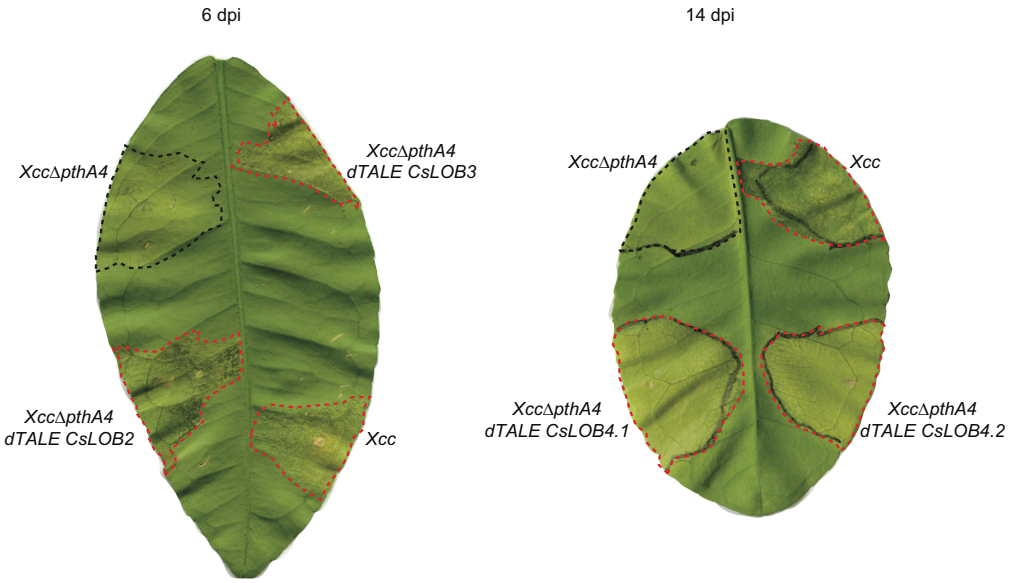


Figure 23: dTALEs targeting *CsLOB1* homologs can promote citrus canker in Duncan grapefruit

Black dashed lines indicate areas inoculated with *Xcc* (*Xcc306* wild type), *XccΔpthA4* (*Xcc306* mutant lacking TALE PthA4) and *XccΔpthA4* harbouring dTALEs targeting *CsLOB2* (*dTALE CsLOB2*), *CsLOB3* (*dTALE CsLOB3*) and *CsLOB4* (*dTALE CsLOB4.1* and *dTALE CsLOB4.2*) promoters as depicted. dTALEs used wre cloned by Dr. Robert Morbitzer. Photos were taken 6 dpi with dTALEs targeting *CsLOB2* and *CsLOB3*, 14 dpi with dTALEs targeting *CsLOB4*.

We then expanded our investigation to LOB homologs in the model plant Arabidopsis. It has been known that Arabidopsis LBD1 (*AT1G07900*) and LBD11 (*At2g28500*) are the closest homologs of CsLOB1, but functionality analysis has not been done so far (Hu et al., 2014). When comparing to the 43 Arabidopsis homologs, we indeed observed that LBD1 and LBD11 share the highest homology to CsLOB1, CsLOB2 and CsLOB3, particularly in the LOB domain that includes the cysteine-rich CX2CX6CX3C motif for DNA binding, the GAS block and leucine-rich LX6LX3LX6L motif for protein dimerisation (Figure 21, Supplemental Figure 6) In contrast, the well-known LOB member, AS2 (*At1g65620*), that regulates leaf polarisation (Semiarti et al., 2001) is relatively distant from CsLOB1 (Figure 21). Another member of Arabidopsis LOB family, AtLOB (*At5g63090*) also belongs to the same subclade as AS2. AtLOB

was shown in a Selection and amplification binding assay and in EMSA to bind to a hexamer GCGGCG motif (Husbands et al., 2007), which is rather different from CsLOB1 motif. To test if the Arabidopsis homologs close and distant from CsLOB1 are able to activate CsLOB1 motif, we cloned the coding sequence of *LBD11*, *AS2* and *AtLOB* into a *p35S*-driven overexpression construct and tagged with a C-terminal GFP tag in the same backbone that was used for *CsLOB1* overexpression construct. Due to time limitation, *LBD1* construct was not included. We then co-expressed the T-DNA constructs of those Arabidopsis LOBs or the EV control together with the RUBY reporter driven by *pBs3* containing either two CsLOB1 wild type motifs (wt) or two mutated motifs (mX). LBD11 was able to activate RUBY driven by wt-motif but not mx-motif containing *pBs3,* as expected (Figure 22). In contrast, AS2 and AtLOB failed to induce significant RUBY signal from both wt- and mX-containing *pBs3* constructs, despite the fact that they were expressed (Figure 22).

2.4 *CsLOB1* and its homologs are fruit softening regulator

The promoter-RUBY assays in *N. benthamiana* have shown that the closest homologs of CsLOB1 in both citrus and Arabidopsis with highly similar LOB domains are able to activate promoters containing CsLOB1 binding motif (Figure 22). We noticed from Hu et al. (2014) that LOB homologs from other crops are closely related to CsLOB1. These include a *Solanum lycopersicum* LOB homolog, SlLOB1 (*Solyc11g072470*), and LOB members from *Vitis vinifera* (common grape vine), *Glycine max* (soybean) and *Cucumis sativus* (cucumber). Interestingly, SlLOB1 was shown recently to also regulate expression of cell wall degrading enzymes (CWDEs) and interestingly, induce tomato fruit softening (Shi et al., 2021). This chapter describes the functional comparison between CsLOB1 and SlLOB1*,* the characterisation of native tissue expression pattern that expands to other Citrus *LOB* homologs.

2.4.1 CsLOB1 and SlLOB1 are closely related and expressed during fruit ripening

We first confirmed the phylogenetic relationship between CsLOB1 and SlLOB1. We compared the full length CsLOB1 protein with the entire tomato (*Solanum lycopersicum*) LOB family, which consists of 47 members (Jin et al., 2017). Phylogenetic analysis showed that CsLOB1 is indeed closely related to SlLOB1, together with two other tomato LOBs (*Solyc06g075330* and *Solyc06g005090*) (Figure 24A, Supplemental Figure 7). In particular, SlLOB1 and CsLOB1 share the highly conserved LOB domain region including the

CX2CX6CX3C motif responsible for DNA binding, the GAS block and the LX6LX3LX6L motif for protein dimerisation (Figure 24C, Supplemental Figure 8).

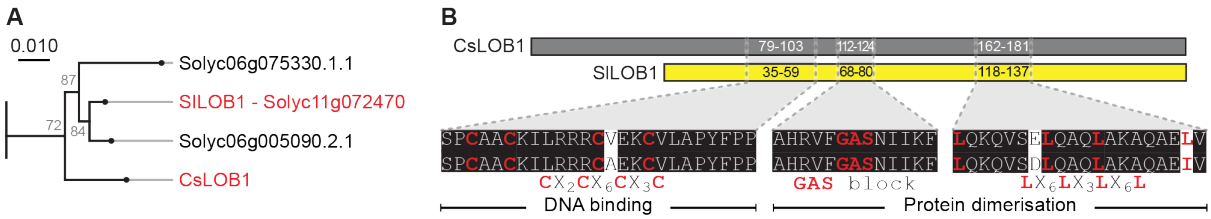


Figure 24: CsLOB1 and SlLOB1 are closely related and share an identical LOB domain
A: Phylogenetic tree of Tomato LOB proteins related to CsLOB1. SlLOB1 (Solyc11g072470) and CsLOB1 are shown in red. Bootstrap values are in grey. The phylogenetic tree was constructed using Qiagen CLC Main Workbench with default settings. The complete phylogenetic tree of 47 Tomato LOB proteins and CsLOB1 is in Supplemental Figure 7. **B: LOB domains of CsLOB1 and SlLOB1 are almost identical.** LOB domain sequence alignment of CsLOB1 and SlLOB1 in the cysteine-rich zinc finger-like DNA-binding domain (CX2CX6CX3C), GAS block and leucine zipper-like domain (LX6LX3LX6L). Letters highlighted in black indicate identical amino acids. Red letters on black background indicate characteristic conserved amino acids. The grey and yellow boxes indicate the conserved domain composition of CsLOB1 and SlLOB1, with dashed lines boxes indicating the domain borders. Numbers indicate the coordinators of each structure. LOB-domain motif features are shown below with conserved amino acids in red, X letter with number indicates how many non-conserved amino acids occur in between. The full-length alignment is shown Supplemental Figure 8.

Interestingly, Shi and colleagues showed that SlLOB1 activates a number of CWDEs and has >300-fold elevated transcript levels in the fruit at the breaker stage, when the fruit begins to ripen, compared to the expression in leaves and other organs (Shi et al., 2021). In contrast, the two other tomato LOBs (*Solyc06g075330* and *Solyc06g005090*) do not show fruit expression pattern. This inspired us to look into CsLOB1 native expression profile. We used the Citrus HZAU database (Liu et al., 2022) to study spatiotemporal expression patterns of *CsLOB1.* We observed that expression levels of *CsLOB1* are relatively low in leaf, root, seed and young fruits of Sweet Orange (SWO) *Citrus sinensis* cv. Valencia (below 10 RPKM). By contrast, *CsLOB1* transcript levels are high in mature SWO fruits (around 150 RPKM) (Figure 25) (Ke et al., 2019). When we compared CsLOB1 transcript levels in ripe fruits to those observed in the *Xcc*-infected Duncan leaf (defined by RNA-seq), we found similar high *CsLOB1* levels in the range of 100 RPKM, while the expression level leaf tissue infected with the PthA4-lacking strain *XccΔpthA4* was similar as in non-infected leaf tissue (Figure 25), demonstrating that the high *CsLOB1* level in *Xcc*-infected leaves are PthA4-dependent.

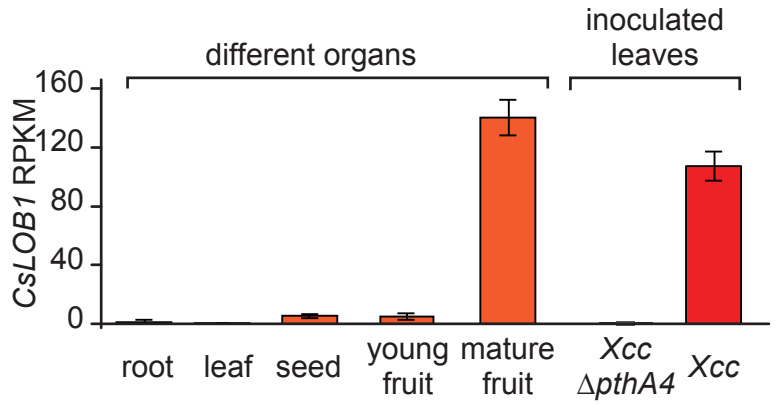


Figure 25: *CsLOB1* transcript levels are highly in mature fruit and *Xcc*-infected leaves
CsLOB1 RPKM (reads per kilobase per million mapped reads) extracted from the Citrus HZAU database for depicted tissues (110 days after flowering-DAF) and mature fruit (312DAF) of Sweet Orange (SWO) *Citrus sinensis* cv. Valencia (Liu et al., 2022) and our RNA-seq dataset at 36hpi.

Intrigued by the specific high *CsLOB1* transcript levels in mature citrus fruit, we inspected *CsLOB1* expression in different fruit tissues. To do so we extracted *CsLOB1* RPKMs from a transcriptomic study of different fruit tissues of *C. sinensis* L. Osbeck cv. Navel, another SWO cultivar, across six developmental stages (Feng et al., 2021). Except for albedo (spongy white tissue), *CsLOB1* expression increased during fruit ripening from 50 to 220 days after flowering (DAF) in other tissues including the epicarp (outer layer of the pericarp-orange peel), juice sacs and segment membrane. The expression is highest in the juice sacs and epicarp (Figure 26). Therefore, similar to SlLOB1, CsLOB1 might play a role in fruit softening.

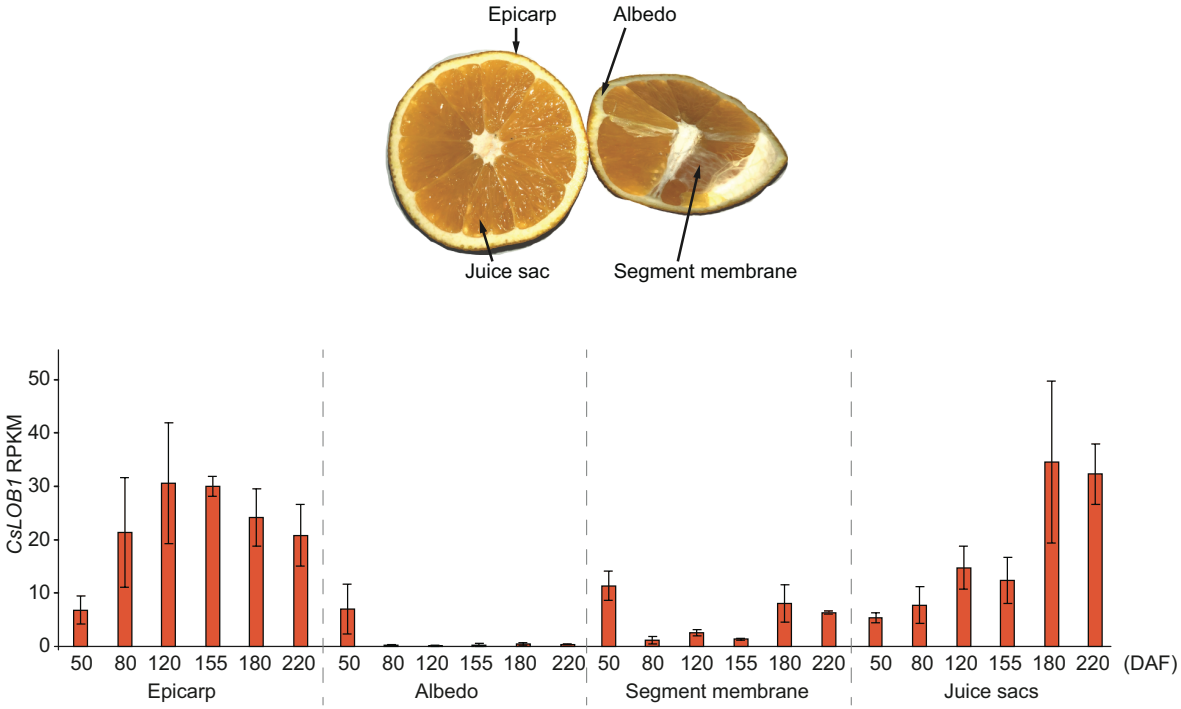


Figure 26: *CsLOB1* is highly expressed in juice sacs and epicarp when the fruit ripens
CsLOB1 RPKM from the epicarp, albedo, segment membrane and juice sacs (as illustrated) extracted from the RNA-seq dataset performed by (Feng et al., 2021) on *C. sinensis* L. Osbeck cv. Navel at 50, 80, 120,

155, 180 and 220 days after flowering (DAF) as depicted. Error bars represent the standard deviation of three RNA-seq replicates.

2.4.2 CsLOB1 and SlLOB1 might share a similar binding motif

This motivated us to investigate first whether SlLOB1 and CsLOB1 have a similar mode of action. Because of the identical LOB domain, which is essential for dimerisation and DNA binding, it is plausible that CsLOB1 and SlLOB1 have similar or identical DNA binding preferences. If this so, they would be able to activate each other's target genes. It has been shown that SlLOB1 highly activated *Expansin 1* (*EXP1*) (Shi et al., 2021). We cloned a 1kb 5' upstream region of the tomato *EXPANSIN1* gene (*pSlEXP1*) and placed it upstream of a GUS reporter gene. Similarly, we cloned a 1kb 5' region upstream of a predicted endoglucanase (*pCs5g20320*), one of the genes highly up-regulated by CsLOB1 at 36hpi, and showed minimal background expression when transiently expressed in *N. benthamiana*. The promoter-GUS reporter constructs were co-delivered into *N. benthamiana* with either p35S-*driven CsLOB1*, p35S-*driven SlLOB1* or empty vector via *Agrobacterium*. GUS assays at 2dpi show that CsLOB1 and SlLOB1 were able to not only activate their own target gene promoters, but also strongly cross-activated each other's target promoters (Figure 27), highly suggesting that they have similar DNA-binding specificity. Not only activating one of CsLOB1 target genes, SlLOB1 was also shown to activate several cell-wall softening enzymes such as endoglucanase, expansin and pectate lyase (Shi et al., 2021), which is very similar to CsLOB1 downstream targets (Figure 13, Supplemental Table 3). This suggests that SlLOB1 might also target CsLOB1 motif.

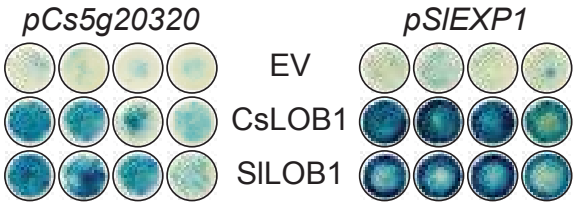


Figure 27: CsLOB1 and SlLOB1 both cross-activate corresponding tomato and citrus target genes

1kb sequence upstream of the ATG of citrus endoglucanase 9 (*Cs5g20320*) and tomato expansin 1 (*SlEXP1-Solyc06g051800*) was cloned to drive a driving *GUS* reporter gene. These promoter-reporter constructs were co-expressed with either p35S-*driven CsLOB1* or p35S-*driven SlLOB1* or empty vector (EV) in *N. benthamiana* by *Agrobacterium* with 4 biological replicates. Leaf discs were GUS-stained at 2dpi (Section 4.2.4.4).

To clarify if SlLOB1 binds to the 15-bp CsLOB1 target motif, we used *pBs3-driven RUBY* constructs, containing two tandem-arranged wildtype (wt) or mutated (mX) CsLOB1 motives (Figure 20). We co-delivered the promoter-reporter constructs with either p35S-*driven SlLOB1* (SlLOB1) or an empty vector (EV) into *N. benthamiana* leaves via *Agrobacterium*-mediated transient transformation. In these assays, SlLOB1 was found to induce RUBY expression in the wt-motif reporter by more than 20 -fold compared to the mX version (Figure 28). This indicates that SlLOB1 is a true ortholog of CsLOB1 and suggests that SlLOB1 and CsLOB1 have similar or even identical DNA binding specificity.

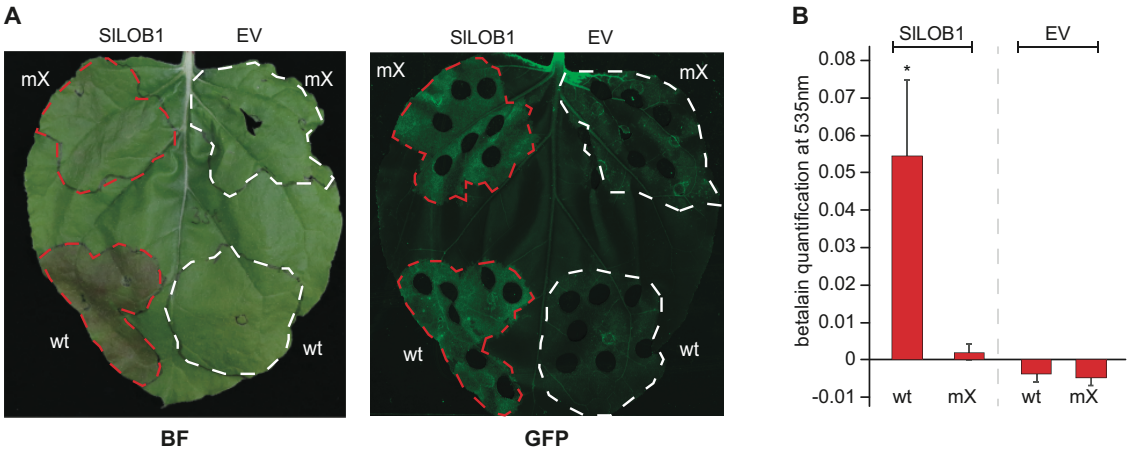


Figure 28: A&B SlLOB1 activates CsLOB1-motif containing promoter
A: RUBY reporter genes driven by the *Bs3* promoter containing a duplicate of CsLOB1 binding motif (wt) or its mutated version (mX) were delivered into *N. benthamiana* leaves by *Agrobacterium* together with either p35S-promoter-driven *SlLOB1* (SlLOB1-red dashes) or an empty vector control (EV- white dashes). Photos were taken at 42hpi in bright field for RUBY observation (left). SlLOB1 was tagged with C-terminal GFP and the expression was examined using Typhoon™ (Cytiva) with the default GFP fluorophore setting. **B:** At the same time point, samples were harvested for betalain extraction, each with 4 different biological replicates. Extracted betalain was quantified by photometric measurement at 535 nm (Section 4.2.4.5). Student's t-test was used to calculate the significant difference between wt and mX activated by SlLOB1. *: $p \leq 0.05$. Assays were done with the help from Paloma Aguilera.

2.4.3 *LOB* homologs are highly expressed during fruit ripening

As SlLOB1 was able to activate CsLOB1 binding motif and both *CsLOB1* and *SlLOB1* are highly expressed during fruit ripening, we speculated whether other closest *LOB* homologs shown to activate CsLOB1 binding motif (Figure 21) also have fruit expression pattern. We looked into the transcripts of citrus *LOB* homologs, *CsLOB2*, *CsLOB3* and *CsLOB4* (*CsLOB2-4*) in citrus HZAU database (Liu et al., 2022). We observed that *CsLOB2-4* also had low or no expression in root, leaf, seed, and young fruit (Figure 29). Similar to *CsLOB1*, the transcript levels of *CsLOB2* and *CsLOB4* in mature fruit are significantly higher than other tissues (Figure 29).

Notably, *CsLOB3* was expressed but at a much lower level in mature fruit. However, looking at the fruit transcriptomic data of another SWO cultivar, called Tarocco or Blood Orange (BO) (Huang et al., 2020), we could see that *CsLOB3* was expressed on average more than 150 RPKM in the pericarp of the mature BO fruit, while *CsLOB2* was expressed more than 200 RPKM in the flesh of BO fruit. Both *CsLOB1* and *CsLOB4* were expressed in mature BO pericarp and flesh, but at lower levels compared to *CsLOB2* and *CsLOB3* (Figure 29A). To gain a better resolution of which fruit tissues express *CsLOB2-4*, we also extracted their read counts from the transcriptomic study of different fruit tissues of *C. sinensis* L. Osbeck cv. Navel (Feng et al., 2021). Here we only examined the 220 DAF when the fruit is ripe. Similarly, we observed a high expression of *CsLOB2* in the juice sacs while *CsLOB3* was more expressed in the epicarp and segment membrane. *CsLOB4* was most highly expressed in the juice sacs compared to other tissues (Figure 29B).

Although there is variation in native expression levels of *CsLOB1*, *CsLOB2*, *CsLOB3* and *CsLOB4* (*CsLOB1-4*) in different citrus cultivars; overall, these four *LOB* homologs were preferentially expressed in fruits compared to other tissues. *CsLOB2* and *CsLOB3* could promote *Xcc* growth in citrus, suggesting that their downstream regulons might be similar to the CsLOB1 regulon. As *CsLOB4* was also expressed during fruit ripening and promoted canker symptoms when targeted by dTALEs in citrus, albeit at a rather late time point, it is likely that it would also share a similar downstream regulon as CsLOB1 and target CsLOB1 binding motif. Due to time constraints, we were unable to address these points within the scope of this thesis.

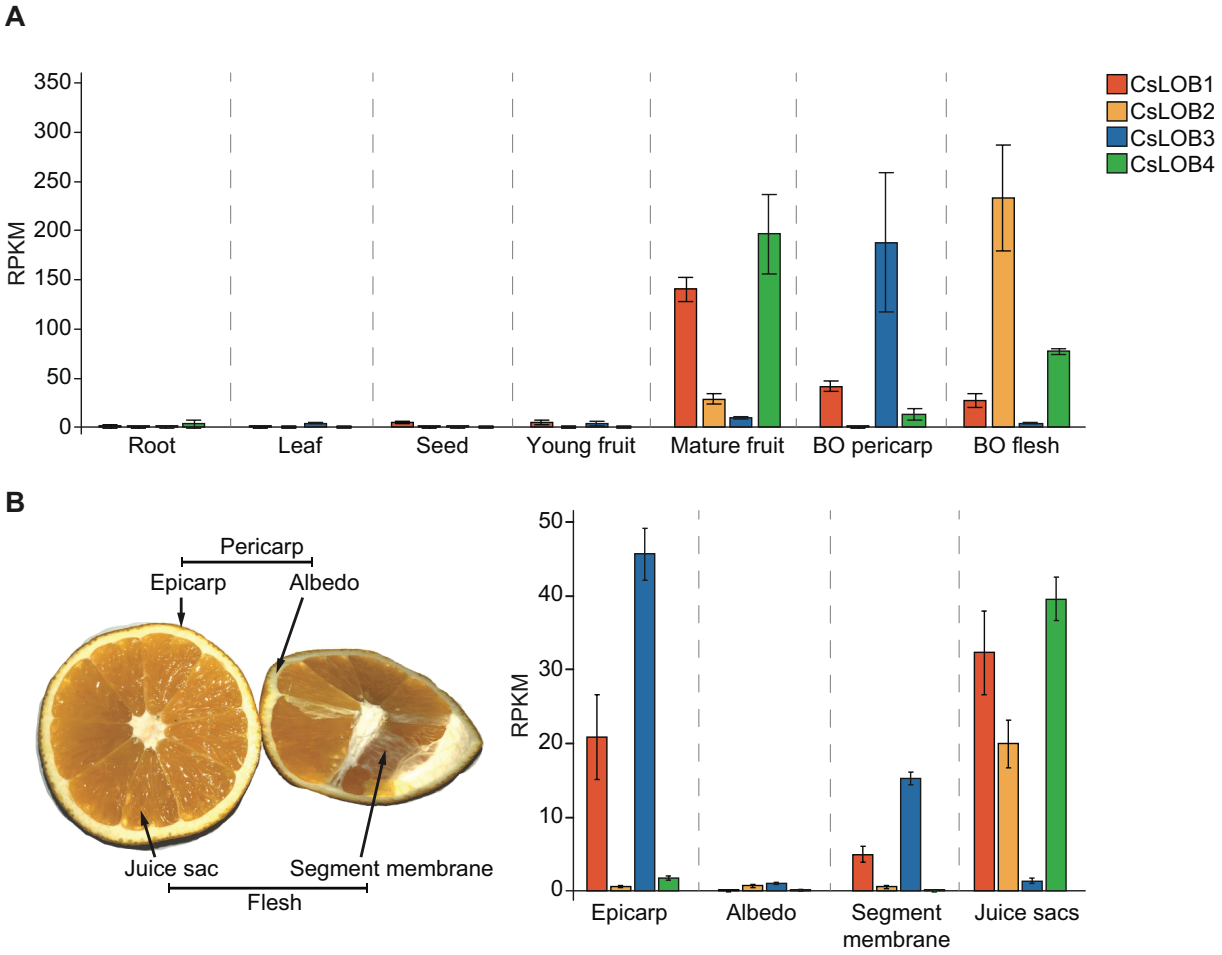
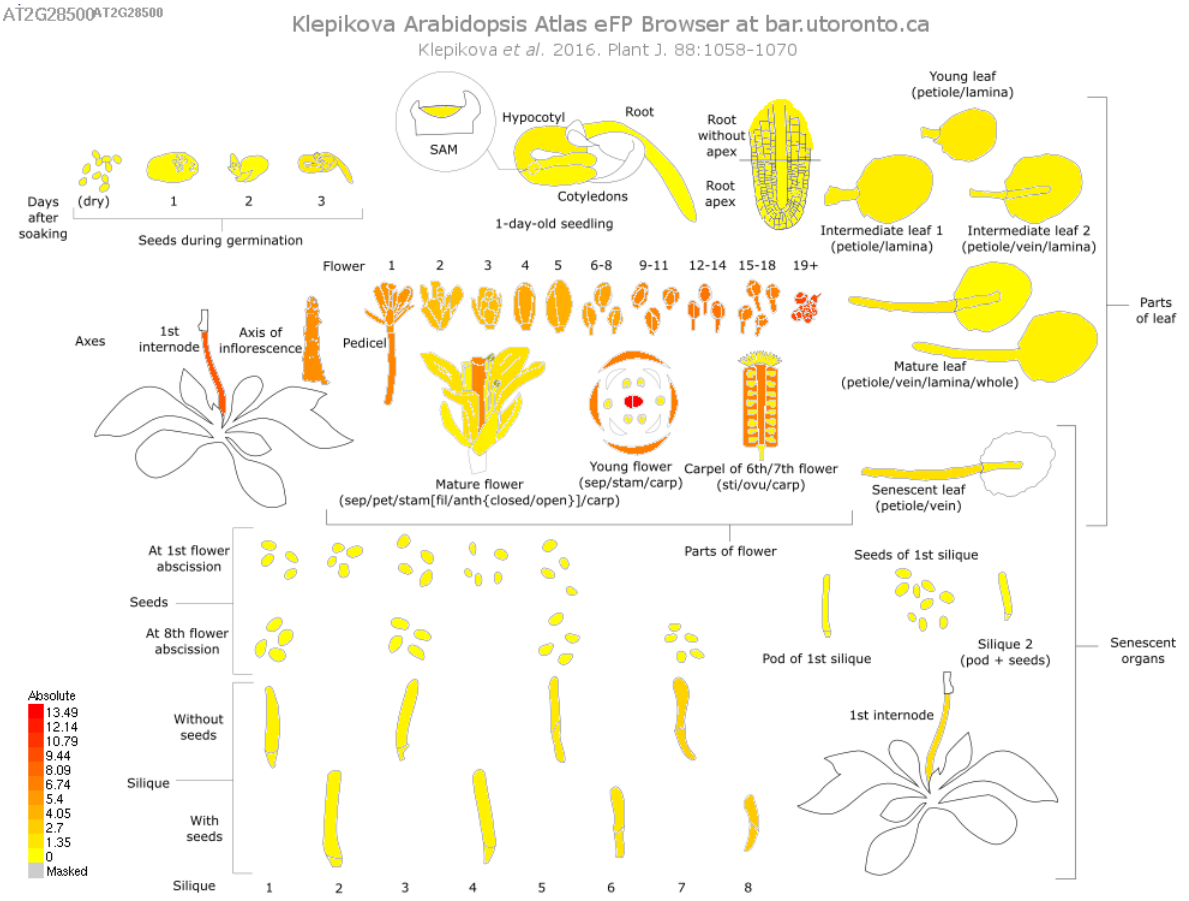


Figure 29: *CsLOB1* homologs are expressed in ripening fruit

A: *CsLOB1, CsLOB2, CsLOB3, CsLOB4* (*CsLOB1-4*) are highly expressed in fruits but not in other tissues such as root, leaf and seed. RPKM of *CsLOB1-4* extracted from the Citrus HZAU database for root, leaf, young fruit (110DAF) and mature fruit (312DAF) SWO cv. Valencia and the pericarp and flesh of another citrus sinensis variety called Blood Orange (BO) cv. Tarocco (Huang et al., 2020) (Liu et al., 2022). Error bars represent standard deviations of 2-3 available RNA-seq replicates. **B: *CsLOB1-4* expression in different fruit tissues.** RPKM of *CsLOB1-4* from the epicarp, albedo, segment membrane and juice sacs (as depicted) extracted from the RNA-seq dataset performed by (Feng et al., 2021) on *C. sinensis* L. Osbeck cv. Navel at 220 days after flowering (DAF). Error bars represent the standard deviation of 3 RNA-seq replicates.

Among Arabidopsis LOB homologs, LBD1 and LBD11 are closest to CsLOB1 (Figure 21). LBD11 was shown to activate CsLOB1-motif-containing promoter (Figure 22). Interestingly, *LBD11* also has a high expression in the carpel of the 6th and 7th flower which contains the ovules (Figure 30). It has been shown carpel tissue is an equivalent fruit tissue in dry fruits to fleshy fruits (Roeder and Yanofsky, 2006), suggesting that LBD11 might play a role Arabidopsis fruit development. However, tissue expression of *LBD11* needs to be first experimentally validated via alternative methods such as qPCR.



Data from A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling: Klepikova et al., 2016, Plant J. 88:1058-1070. Total RNA was extracted with RNeasy Plant Kit and Illumina cDNA libraries were generated using the respective manufacturer's protocols. cDNA was then sequenced using Illumina HiSeq2000 with a 50bp read length. The read data are publicly available in NCBI's Sequence Read Archive under the BioProject ID 314076 (accession: PRJNA314076). Reads were aligned to the reference TAIR10 genome (Lamesch et al., 2012) using TopHat (Trapnell et al., 2009). Default TopHat settings and job resource parameters were used, with read groups unspecified. Reads per gene were counted with an in-house Python script using functions from the HTSeq package (Anders et al., 2015). Reads were filtered so that only uninterrupted reads corresponding to a region within exactly one gene were used for RPKM calculation. If a gene's expression level is not displayed, this indicates the reads for this gene did not pass the filtering criteria. RPKM values were compiled using an in-house R script.

Figure 30: *LBD11* transcripts were enriched in flower carpel
Expression atlas for *LBD11* extracted from the Arabidopsis expression atlas project (Klepikova et al., 2016) from The Arabidopsis Information Resource (Bailey and Elkan, 1994). Expression levels are shown as heatmap as depicted.

2.5 Homologs of CsLOB1 promote *Xanthomonas* growth

Members of the LOB transcription factor family have a wide range of diverse functions in plant development, and many are expressed in different organs such as leaf, roots, pollen, embryo sac (Zhang et al., 2020). Up to now, there have been no detailed studies on the native function of CsLOB1. We have shown that *CsLOB1* and their closest homologs were highly expressed during fruit softening (Figure 29) and *CsLOB1-4* were shown to rescue canker symptoms when ectopically activated by PthA4 or dTALEs in *XccΔpthA4* background (Figure 23). This suggests that the fruit softening activation might translate into host metabolic changes that promote the disease and benefit the bacteria. We reason that equivalent to *CsLOB1* activation that promote *Xanthomonas* growth, the expression of those

homologs will also promote bacterial growth. Indeed, the activation of *CsLOB2* via dTALE delivered by *Xanthomonas* strain lacking PthA4 was shown to restore *Xcc* growth (Zhang et al., 2017a). This chapter describes the characterisation of another close homolog of CsLOB1 in citrus, CsLOB3 and the Tomato homolog, SlLOB1, as another two examples for the benefit of fruit softening CsLOB1 homologs in *Xanthomonas* growth promotion.

2.5.1 CsLOB2 and CsLOB3 promote *Xcc* growth in citrus

To confirm the growth promotion by CsLOB2 and investigate whether CsLOB3 could also promote *Xcc* growth, dTALEs targeting CsLOB2 and CsLOB3 were cloned by Dr. Robert Morbitzer (Figure 31A). The *XccΔpthA4* mutant were then complemented with the dTALEs and was infiltrated into Duncan leaves. The quantification of transcript levels of *CsLOB1*, *CsLOB2* and *CsLOB3* at 12hpi by RT-qPCR shows that *CsLOB1* was upregulated by PthA4 as expected, but not by dTALEs targeting *CsLOB2* and *CsLOB3* (Figure 31B). Similarly, *CsLOB2* and *CsLOB3* were specifically activated by the *XccΔpthA4* strains complemented with dTALEs targeting *CsLOB2* and *CsLOB3*, respectively, but not by the strain complemented with PthA4 (Figure 31B).

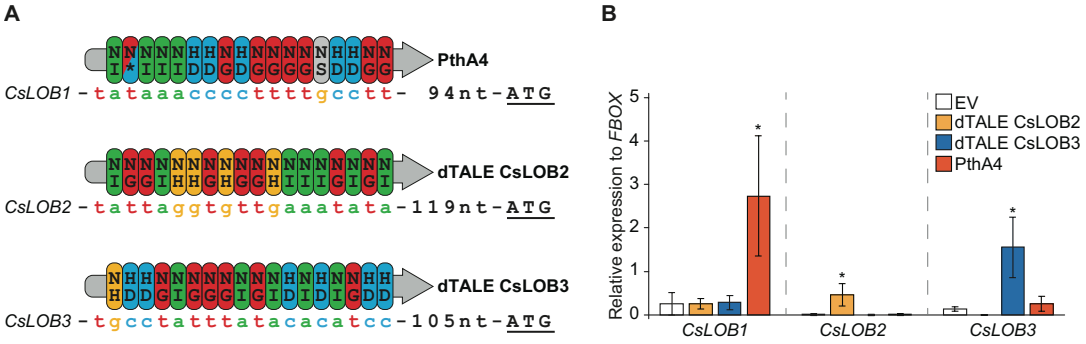


Figure 31: *XccΔpthA4* containing dTALEs targeting *CsLOB2* and *CsLOB3* activates specifically *CsLOB2* and *CsLOB3*, respectively

A: Schematic illustration of dTALEs targeting *CsLOB2* and *CsLOB3*. The *CsLOB1*-activating *Xcc* TALE protein PthA4 and *CsLOB2*- and *CsLOB3*- activating dTALEs are depicted as bold grey arrows with coloured ovals representing TALE repeat variable diresidues (RVDs) in single letter code (uppercase) that are aligned with the targeted bases (lowercase) located upstream of the ATG start codon (underlined). The colour of the TALE RVDs indicates their base preference (specifically: NI-a [green], HD-c [blue], NG-t [red], NH-g (orange). N* [blue/red] preferentially targets both c and t, NS [grey] can target all 4 nucleotides). **B: Specific activation of *CsLOB1*, *CsLOB2* and *CsLOB3* by *Xcc* PthA4, dTALE CsLOB2 and dTALE CsLOB3, respectively.** Total RNA was extracted from Duncan leaves infiltrated with either *Xcc306ΔpthA4* complemented with depicted TALEs or *Xcc306ΔpthA4* containing empty vector (EV) at 12hpi. Transcript levels of *CsLOB1*, *CsLOB2* and *CsLOB3* were quantified by RT-qPCR. Relative expression was obtained by normalising the ct value of the gene of interest to the housekeeping gene *FBOX* (*Cs6g19880*). Bar graph represents the mean ± SD of 3 biological replicates. Student's t-test was used for statistical analysis. *: $p \leq 0.05$

We then examined *Xcc* growth levels when these homologs were activated and confirmed that CsLOB2 could enhance bacterial growth compared to the *XccΔpthA4* mutant. As expected, *CsLOB3* activation also promoted *Xcc* growth, though the growth promotion levels by CsLOB2, CsLOB3 and CsLOB1 by *Xcc* wild type as a control, are notably different (Figure 32). While the strains complemented with dTALEs targeting *CsLOB2* and *CsLOB3* have relatively similar growth level compared with the *XccΔpthA4* mutant strain, *Xcc* wild type strain showed the highest level of growth promotion, with almost 2 $\log_{10}$ fold change (100-fold) compared to the mutant (Figure 32).

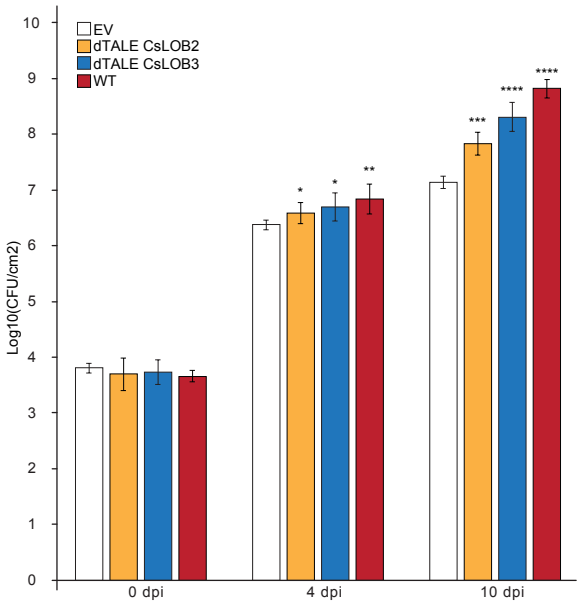


Figure 32: The dTALE-based interventions targeting *CsLOB2* and *CsLOB3* gained *in-planta* growth promotion for *XccΔpthA4*

Xcc, XccΔpthA4 complemented with dTALEs targeting *CsLOB2* and *CsLOB3* and *XccΔpthA4* (with empty vector, EV) were inoculated into Duncan leaves. Leaf discs were harvested at 0, 4 and 10 dpi and subjected to bacterial counting (colony formed units (CFU) per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 6 biological replicates. Student t-test was used to calculate the significant difference between depicted strains versus EV control. *: $p \leq 0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$.

The difference in growth promotion between *Xcc* (inducing *CsLOB1* by PthA4) and the complemented *XccΔpthA4* strains activating *CsLOB2* and *CsLOB3* by dTALEs could be due to the different expressions of the *TALE/dTALEs* in the bacteria as the *dTALEs* were driven by a different promoter in a different genotypic background as compared to *PthA4*. To minimise this effect, we used the *XccΔpthA4* strain complemented with *PthA4* in the same vector backbone as the *dTALEs* targeting *CsLOB2* and *CsLOB3*, which was shown to activate *CsLOB1* (Figure 31). After 10dpi, all complemented strains had outgrown the *XccΔpthA4* mutant

(Figure 33). Interestingly, the *CsLOB3*-activating *XccΔpthA4* grew more than half a $\log_{10}$ (approx. 3-fold) compared to the *CsLOB2*-activating or *PthA4*-containing strains. In addition, while the *XccΔpthA4* targeting *CsLOB2* and *CsLOB3* continued to grow over the course of 12 days, the PthA4-containing strain surprisingly started to plateau on day 14 and reduced in growth on day 22 (Figure 33). None of the dTALEs and PthA4 showed cross activation (Figure 31), indicating that the growth promotion was specific to the activation of *CsLOB2* and *CsLOB3* rather than *CsLOB1* cross-activation. Interestingly, although the level of gene activation by *CsLOB2* and *CsLOB3* dTALEs is lower compared to PthA4 at an early time point (Figure 31), they promoted a higher bacterial growth level later on (Figure 33).

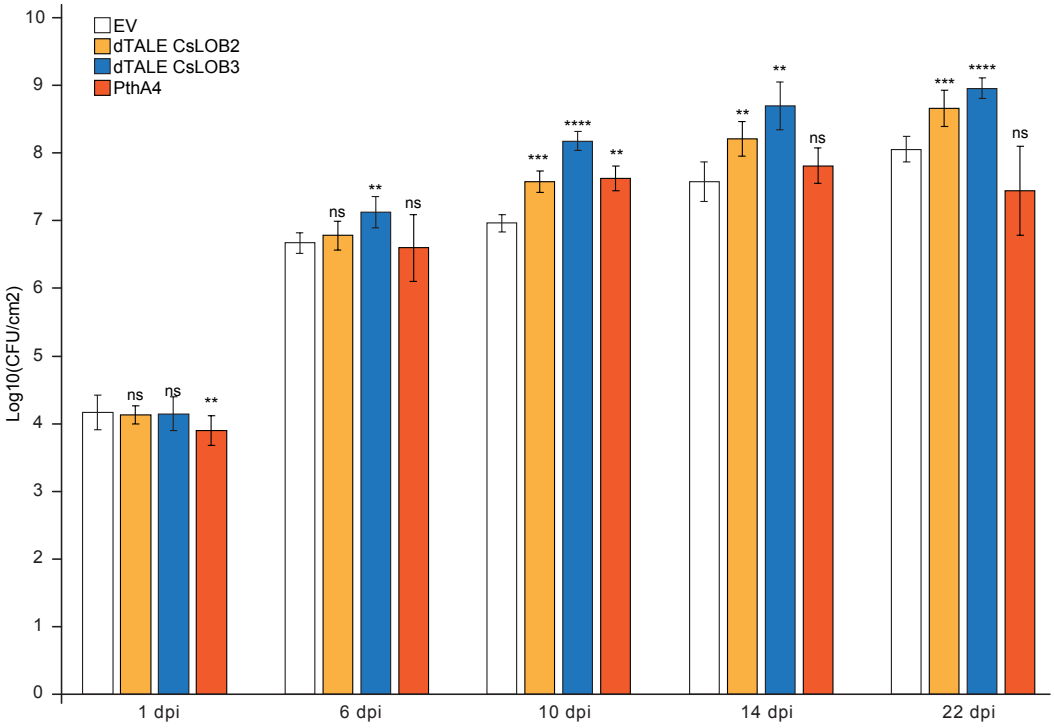


Figure 33: The dTALE-based interventions targeting *CsLOB2* and *CsLOB3* exhibited a significantly greater *Xcc* growth promotion compared to *CsLOB1*-targeting PthA4

Xcc, XccΔpthA4 complemented with dTALEs targeting *CsLOB2* and *CsLOB3* and *XccΔpthA4* (with empty vector, EV) were inoculated into Duncan leaves. Leaf discs were harvested at 0, 4 and 10 dpi and subjected to bacterial counting (CFU per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 6 biological replicates. Student's t-test was used for statistical analysis. *: $p \leq 0.05$, **: $p<0.01$, ***: $p<0.001$, ns: not significant.

2.5.2 CsLOB1 promotes *Xanthomonas* growth in *N. benthamiana*

The fact that CsLOB2 and CsLOB3 could activate CsLOB1 motif and could promote water soaking and pustules and *Xcc* growth in citrus suggest that some common downstream genes activated by them are important for the disease. In addition, we have shown that CsLOB1 and its closest homologs in citrus and tomato have high expression during fruit softening while the Arabidopsis homolog *LBD11* transcripts are enriched in the carpel tissue which is an equivalent fruit tissue in dry fruits to fleshy fruits (Section 2.4.1 and 2.4.3) (Roeder and Yanofsky, 2006). It is likely that the cellular changes during fruit softening are required for bacterial growth and disease development.

We hypothesised that the activation of these LOB homologs would also promote *Xanthomonas* growth. We examined SlLOB1 first as citrus and tomato are both fleshy fruits and possibly have more conserved downstream elements. Ideally, one would need a system that has both *CsLOB1* and *SlLOB1* in order to compare. Although citrus is an economically important crop, the difficulty of producing transgenic citrus plants and the very long generation time of about five years limit the possibilities for genetic studies in citrus. To overcome these limitations, we considered studies of CsLOB1 and other LOB homolog downstream mechanism in *N. benthamiana*, the model plant species for transient, *A. tumefaciens* mediated transformation.

The pepper/tomato *Xanthomonas* strain *euvesicatoria* (*Xe*) is known to induce a hypersensitive response (HR) in *N. benthamiana* due to the presence of XopQ effector, which is recognised by the XopQ-sensing resistance protein ROQ1. The *N. benthamiana* mutants lacking ROQ1 or the central immune signalling regulator EDS1 do not produce HR in the presence of *Xanthomonas* and support *in-planta* growth of tomato/pepper-pathogenic *Xanthomonas* strains (Schultink et al., 2017). To clarify if *N. benthamiana roq1* mutants would also support *in-planta* growth of citrus pathogenic *Xcc,* we infiltrated *Xcc* and *Xe* into leaves of WT, *roq1* and *eds1 N. benthamiana* genotypes. As expected, *Xe* grew in all genotypes, while bacterial counts were more than 10 -fold higher in *roq1* and *eds1* mutants compared as compared to WT *N. benthamiana* (Figure 34). By contrast, *Xcc* was not detectable at 6dpi in *N. benthamiana* wild type (WT). Yet bacterial counts were about 1000-fold increased at 6dpi in both *roq1* and *eds1 N. benthamiana* mutants compared to the

bacterial counts observed at 0 dpi (Figure 34). These findings suggest, that the *N. benthamiana*, *roq1* and *eds1* mutants could be used instead of citrus plants to study *in-planta Xcc* growth.

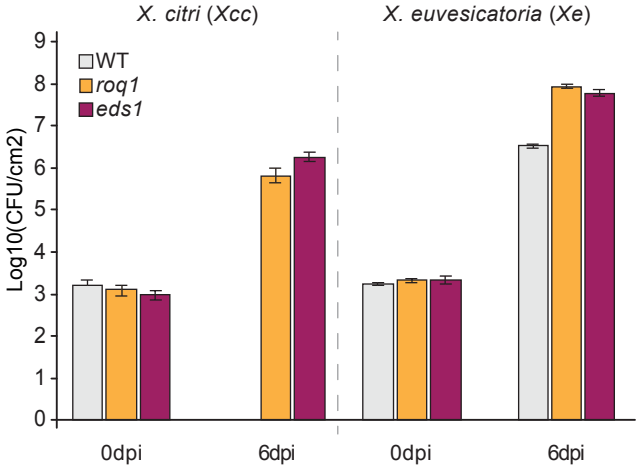


Figure 34: *Xcc* can grow in *N. benthamiana roq1* and *eds1* mutants
Xcc and *Xe* were infiltrated into leaves of *N. benthamiana* wild type (WT), *roq1* and *eds1* mutants. *Xcc* in WT *N. benthamiana* is not shown in the graph at 6dpi because the amount was not detectable at the dilution of 30CFU/cm^2 (Log_{10}<1.5). Leaf discs were harvested at 0 and 6dpi and subjected to bacterial counting (CFU per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 4 biological replicates.

In the context of citrus leaves, *Xcc* delivered PthA4 was able to transcriptionally activate *CsLOB1,* allowing it to grow 10-15 times better than the PthA4-deficient mutant, *XccΔpthA4* (Hu et al., 2014). In *N. benthamiana*, it is not known whether PthA4 binds upstream of and transcriptionally activates *CsLOB1* related genes. When we infiltrated *roq1 N. benthamiana* with either *Xcc* or *XccΔpthA4* strains, we observed no significant difference of bacterial growth *in-planta* at 6dpi, suggesting that PthA4-does not transcriptionally activate *CsLOB*-related genes and therefore does not promote growth of *Xcc* in *N. benthamiana* (Figure 35).

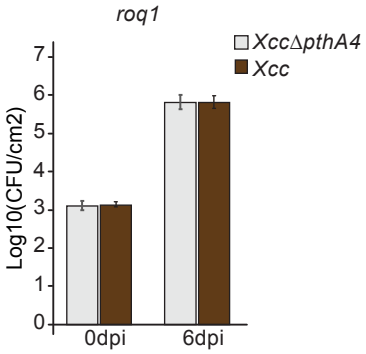


Figure 35: PthA4 has no growth contribution to *Xcc* in *N. benthamiana roq1* mutant
Xcc and *XccΔpthA4* were infiltrated into *N. benthamiana roq1* mutant. Leaf discs were harvested at 0 and 6dpi and subjected to bacterial counting (CFU per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 4 biological replicates.

To functionally replace TALE-mediated activation of *CsLOB1* as it is observed in citrus, we delivered a *p35S-driven CsLOB1* T-DNA into *N. benthamiana* leaves via *Agrobacterium* and

co-infected the plant with *Xanthomonas* strains to see the impact of the T-DNA on *in-planta* growth promotion of *Xanthomonas*. We first tested the *Xe* strain in the leaves of the *eds1* mutant, as it showed better growth in all *N. benthamiana* compared to *Xcc* (Figure 34). At 6dpi, we observed a significant increase of almost 3-fold in *Xe* growth in the presence of CsLOB1, but not in the empty vector. In contrast, *Xcc* showed no significant increase in growth (Figure 36).

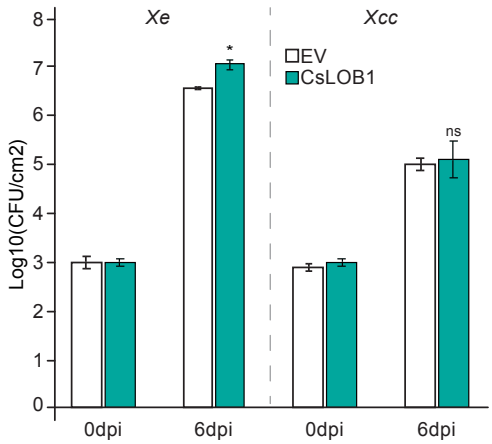


Figure 36: Transient expression of *CsLOB1* promotes *Xe* growth in *N. benthamiana eds1*
Xe or *Xcc* was co-infiltrated with either *p35S-driven CsLOB1* (CsLOB1) or empty vector (EV into *N. benthamiana eds1* via *Agrobacterium*. Leaf discs were harvested at 0 and 6dpi and subjected to bacterial counting (CFU per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 4 biological replicates. Student's t-test was used to calculate the significant difference between CsLOB1 and EV data points. *: $p \leq 0.05$, ns: not significant.

As PthA4-induced *CsLOB1* expression in citrus can induce a more than 10-fold increase in *Xcc* growth upon (Hu et al., 2014), we attempted to optimise the growth curve assay in *N. benthamiana*. We observed a reduction in CsLOB1 protein accumulation over time when expressed in *N. benthamiana* (Figure 37), which presumably was due to activation of host silencing mechanisms. To overcome this and to maintain *CsLOB1* expression over a longer period, we co-infected our *p35S-driven CsLOB1* T-DNA with *Agrobacterium*-delivering the viral gene-silencing inhibitor *p19*. We found that *p19* expression maintained CsLOB1 accumulation until 6dpi as compared to the non-p19 control sample (Figure 37).

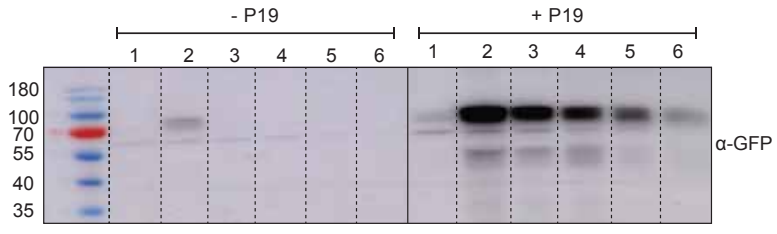


Figure 37: The viral silencing suppressor p19 induces elevated *CsLOB1* expression over several days
N. benthamiana leaves were infiltrated with *p35S-driven CsLOB1* tagged with *C-terminal GFP* with or without p35S-*driven P19* via *Agrobacterium*. Leaf discs were harvested over the course of 6 days as

depicted. Western blot was performed with primary antibody against GFP antibody and HRP-conjugated α-rabbit secondary antibody visualised by CCD camera.

We then co-infiltrated *Xcc*, *Agrobacterium* delivering *p35S-driven CsLOB1* and *p35S-driven p19* into *N. benthamiana eds1* leaves. At 6 dpi, *Xcc* growth was indeed significantly higher with almost 10-fold in *CsLOB1* overexpression tissue compared to EV (Figure 38), indicating that *Agrobacterium*-mediated transient expression of *CsLOB1* could promote growth not only of the pepper/tomato *Xe* strain (Figure 36), but also of the citrus *Xcc* strain in the *N. benthamiana eds1* mutant. This suggests that 1) CsLOB1 target sites are conserved between Citrus and *N. benthamiana* and that 2) *N. benthamiana* can be used as an alternative system to study how CsLOB1 promotes *in-planta* growth of *Xcc* or *Xe*.

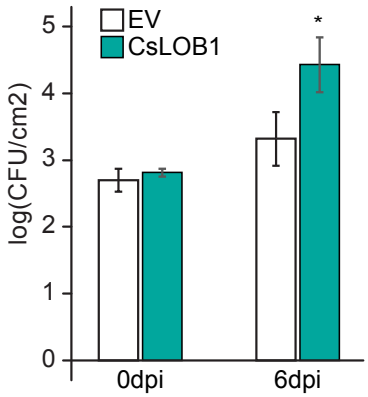


Figure 38: Maintained transient expression of *CsLOB1* increased *Xcc* growth by 10-fold
Xcc was co-infiltrated with *p35S-driven p19* and either *p35S-driven CsLOB1* (CsLOB1) or empty vector (EV) into *N. benthamiana eds1* via *Agrobacterium*. Leaf discs were harvested at 0 and 6dpi and subjected to bacterial counting (colony formed units (CFU) per leaf cm^2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 4 biological replicates. Student's t-test was used to calculate the significant difference between CsLOB1 and EV data points, *: $p \leq 0.05$.

2.5.3 SlLOB1 promotes *Xanthomonas* growth in the *N. benthamiana eds1* mutant

SlLOB1 has been shown to have a similar spatiotemporal expression pattern during fruit ripening (Shi et al., 2021) as *CsLOB1* , a nearly identical LOB domain (Figure 24), and the ability to activate promoters containing a CsLOB1-motif (Figure 28). It therefore seemed likely that transient expression of *SlLOB1* in *N. benthamiana* would also promote bacterial growth like *CsLOB1* did.

Due to the complications of working with *Xcc* as it is a quarantined pathogen, we investigated SlLOB1-dependent growth promotion using the pepper/tomato strain *Xe,* as this strain showed elevated *in-planta* growth when *CsLOB1* was expressed (Figure 36). We co-infiltrated *Xe* with *Agrobacterium* delivering *p35S-driven p19* in combination with either *p35S-driven SlLOB1*, *p35S-driven CsLOB1* or empty vector into the *N. benthamiana eds1* mutant. At 6dpi, CsLOB1 was shown to indeed promote *Xe* growth up to a 10-fold increase compared to EV, similar to what has been seen with *Xcc* (Figure 38). Interestingly, transient

expression of *SlLOB1* also resulted in a 10-fold increase in *Xe* growth compared to EV, supporting our hypothesis that CsLOB1 and SlLOB1 share downstream elements that are conserved between citrus and Solanaceae plants and have growth-promoting benefits for *Xanthomonas* bacteria (Figure 39).

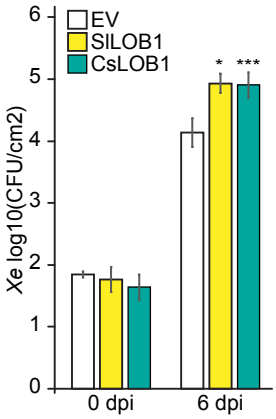


Figure 39: Transient overexpression of *CsLOB1* and *SlLOB1* promotes growth of *Xanthomonas euvesicatoria* (*Xe*) strain *85-10* in *N. benthamiana eds1*

Xe was co-infiltrated with *p35S-driven P19* and either *p35S-driven CsLOB1* (CsLOB1) *p35S-driven SlLOB1* (SlLOB1) or empty vector (EV) into *N. benthamiana eds1* via *Agrobacterium*. Leaf discs were harvested at 0 and 6dpi and subjected to bacterial counting (colony formed units (CFU) per leaf cm2 on a base 10 logarithmic scale). Error bars represent the standard deviation of 4 biological replicates. Student's t-test was used to calculate the significant difference between CsLOB1 or SlLOB1 treatment versus EV control. *: $p \leq 0.05$, ***: $p<0.001$

To obtain a genome-wide view on CsLOB*1*- and SlLOB1-activated genes in *N. benthamiana,* we performed RNA-seq on wild type *N. benthamiana* leaves upon expression of *p35S-driven CsLOB1*, *p35S-driven SlLOB1* or delivery of an empty vector via *Agrobacterium*. Both LOB constructs have a C-terminal GFP tag and were detectable at 42hpi (Figure 40A). Note that although both *CsLOB1* and *SlLOB1* were cloned into the same vector backbone and are transcriptionally controlled by same *p35S* upstream sequence, *SlLOB1* expression was consistently lower than *CsLOB1* (Figure 40A). Nevertheless, *SlLOB1* expression was sufficiently high to activate the promoter containing CsLOB1 motif (Figure 28) and to promote *Xanthomonas* growth in *N. benthamiana* (Figure 39).

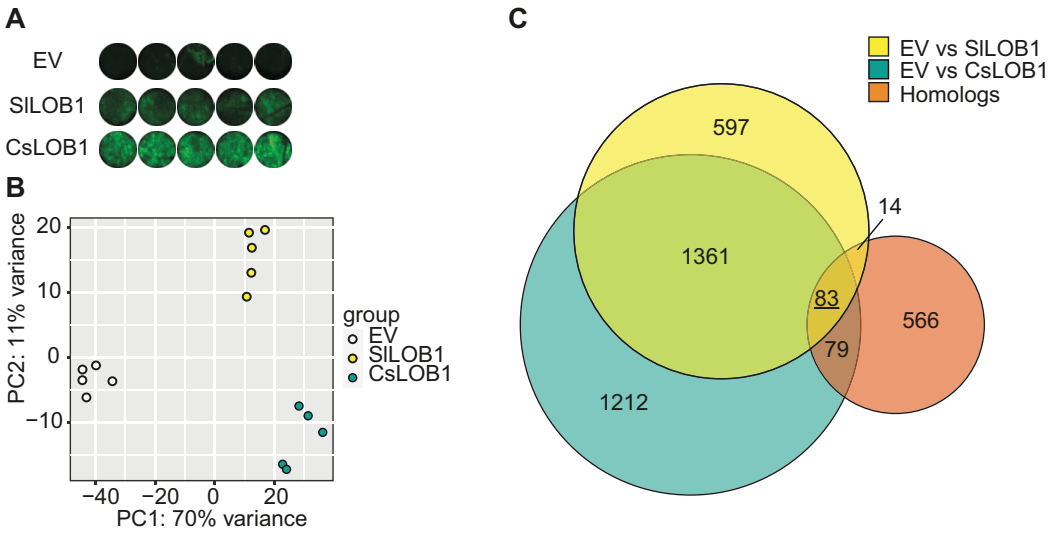


Figure 40: CsLOB1 and SlLOB1 induced similar sets of genes, 83 of which are homologs of the Citrus direct target genes of CsLOB1 when transiently expressed in *N. benthamiana*.
A: Leaf discs expressing either empty vector (EV), *p35S-driven SlLOB1-GFP* (SlLOB1) and *p35S-driven CsLOB1-GFP* (CsLOB1) at 42hpi in *N. benthamiana* wild type. GFP signal was detected using Typhoon™ (Cytiva) with default GFP fluorophore setting. **B: Principal Component Analysis (PCA) of RNA-seq data of the 3 treatments (EV, SlLOB1, CsLOB1) in A), each with 5 replicates.** All sample variances were plotted against the two main components, PC1 and PC2, which account for 70% and 11% of the variance, respectively**. C: DEGs upregulated by CsLOB1 and SlLOB1 show a large overlap, with 83 genes being the homologs of Citrus genes that are direct targets of CsLOB1.** Size-proportional Venn diagram between the DEGs upregulated by CsLOB1 (2735 genes), by SlLOB1 (2055 genes) and predicted homologs of CsLOB1 direct target genes in *N. benthamiana* (742 genes). The number of genes in individual intersections of the size-proportional Venn diagrams are shown in black. The underlined number indicates the genes common to three datasets. The analysis was performed by Dr. Jan Grau (Martin-Luther-Universität Halle-Wittenberg, Germany).

RNA-seq showed that all biological replicates for a given treatment clustered together, indicating a good reproducibility (Figure 40B). DEG analyses were performed between *CsLOB1*- or *SlLOB1*-expressing tissues and the EV control based on the *N. benthamiana* draft genome and annotation v2.6.1 (Bombarely et al., 2012). We found 1444 DEGs that were commonly up-regulated by CsLOB1 and SlLOB1 (Figure 40C), suggesting that their *N. benthamiana* regulons have a large overlap. We examined whether any of these *N. benthamiana* DEGs were the homologs of CsLOB1 direct target genes identified in citrus. For this, a homology-based gene prediction based on the protein sequences of these 96 direct target genes of CsLOB1 in citrus was performed by Dr. Jan Grau (Figure 13) and 742 *N. benthamiana* genes were found to have significant homology to CsLOB1 direct targets. Overlap analysis showed that 83 homologous *N. benthamiana* genes are up-regulated by both CsLOB1 and SlLOB1. These 83 genes are predicted homologs of 50 citrus genes that are

direct targets of CsLOB1 in Figure 13. As some citrus genes are homologous to each other, they were predicted to be the homologs of the same *N. benthamiana* gene.

Our transcriptome studies demonstrated that CsLOB1 and SlLOB1 transcriptionally activate homologous genes in citrus and *N. benthamiana*. Given that *N. benthamiana* and tomato are phylogenetically closely related, it seemed likely that ectopic expression of *SlLOB1* in tomato will also cause elevated *in-planta* growth of the *Xe* strain in tomato.

Given that *Agrobacterium*-mediated transient expression is less efficient in comato as compared to *N. benthamiana*, we decided to induce ectopic expression of *SlLOB1* via *SlLOB1*-activating dTALEs delivered by *Xe.* However, wildtype tomato plants contain the NLR R protein Bs4 which mediates recognition of TALE proteins (Schenstnyi et al., 2022). Yet, our lab has knocked-out the *Bs4* gene by CRISPR/Cas9-mutagenesis and in these *Bs4*-KO lines, dTALE-mediated gene activation is feasible (Schenstnyi et al., 2022). Taking advantage of the Golden Gate compatible TALE assembly kit (Morbitzer et al., 2011), three dTALEs targeting the *SlLOB1* promoter were cloned (Figure 41A). The dTALEs were introduced into the pepper/tomato *Xe* strain 85-10 used in our previous *N. benthamiana* growth curve assays.

To test the dTALE functionality, a GUS reporter construct driven by the 600bp 5' *SlLOB1* upstream region was generated. We co-infected the pepper/tomato *Xe* strain containing SlLOB1 dTALEs together with *Agrobacterium* delivering *SlLOB1 promoter-driven GUS* construct into leaves of the *N. benthamiana roq1* mutant, which no longer recognises *Xanthomonas* XopQ effector and does not produce HR (Schultink et al., 2017). GUS assays showed that all three dTALEs strongly activated the *SlLOB1* promoter compared to the empty vector control (Figure 41B).

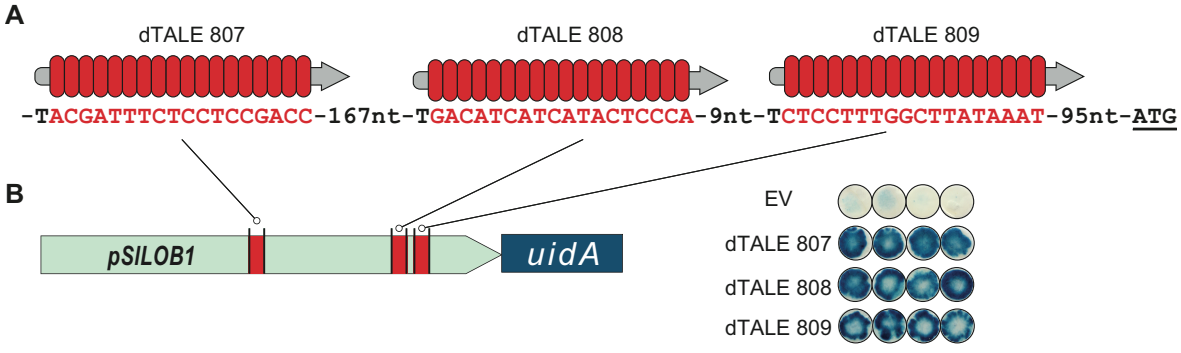


Figure 41: dTALEs can activate *SlLOB1* promoter-GUS in *N. benthamiana*

A: Schematic representation of *SlLOB1*-activating TALEs. Three *SlLOB1*-activating dTALEs (807, 808, and 809) are depicted as bold grey arrows with red ovals representing TALE repeat variable diresidues (RVDs) aligned with the target bases (red font) located upstream of the *SlLOB1* ATG start codon (underlined). The dTALEs were introduced into *Xe* strain 85-10. The distance between the dTALEs *EBEs* and to the ATG is

depicted. ATG: start codon. **B: Schematic representation of the promoter-reporter constructs where *uidA*, encoding a GUS reporter, is driven by the upstream sequence (promoter; *p,* 600bp) of the *SlLOB1* gene.** The *pSlLOB1* sequence contained the EBEs of the dTALEs in A) marked as red box. The reporter construct depicted was delivered into the leaves of *N. benthamiana roq1* mutant by *Agrobacterium* together with either *Xe* containing the dTALEs in A) or an empty vector control (EV). Inoculated leaf discs were harvested at 2dpi for GUS staining.

Next, we examined if the pepper/tomato-pathogenic *Xe* strain that delivers the *SlLOB1*-targeting dTALEs could transcriptionally activate *SlLOB1* in the *bs4* tomato line. We infiltrated *Xe* strains containing either the dTALEs or an empty into 5-week-old *bs4* tomato leaflets and harvested inoculated leaf tissue at 24hpi. RT-qPCR studies showed that dTALE 808 and 809 activated *SlLOB1* by 25-fold and 15-fold, respectively, compared to the EV control. By contrast dTALE 807 failed to transcriptionally activate *SlLOB1* (Figure 42). Due to the higher gene activation ability, we selected dTALE 808 for our subsequent experiments to study SlLOB1, including its transcriptomic analysis and *Xe* growth-promoting ability in Tomato.

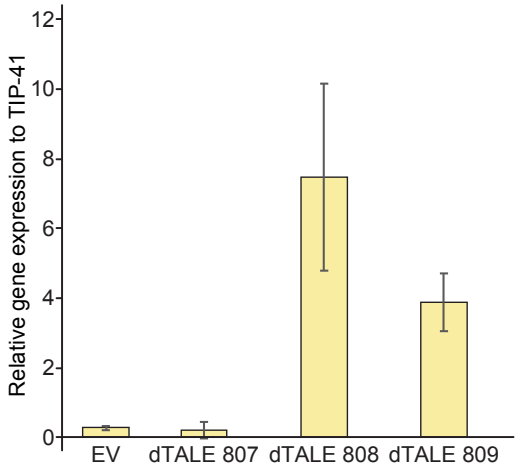


Figure 42: dTALE 808 and 809 delivered by *Xe* strain *85-10* activate *SlLOB1 (Solyc11g072470)* in the Tomato *bs4* mutant but not by dTALE 807or EV.
Leaves of Tomato cv. Money Maker with *Bs4* gene knocked out by CRISPR Cas9 (*bs4* mutant) (Schenstnyi et al., 2022) were infiltrated with *Xe* containing either empty vector (EV) or dTALE 807, 808 or 809 as indicated in Figure 41. Total RNA was extracted after 24hpi. Transcript levels were quantified by RT-qPCR. *SlLOB1* expression was normalised to the housekeeping gene *TIP-41* (*Solyc10g049850*) from *Solanum lycopersicum*. Bar graph represents the mean ± SD of 3 biological replicates.

2.6 CsLOB1 changes the apoplastic fluid composition

As CsLOB1 and their closely related homologs in citrus and tomato regulate fruit softening, it is likely that *Xcc* hijacked this development process to create a favourable condition for bacterial proliferation. Notably, *Xcc* injects PthA4 into the plant host cell to modulate this developmental program, while it remains and multiplies in the apoplastic region. Therefore,

it is plausible that the intracellular changes are translated into extracellular metabolic changes that benefit the bacteria. This chapter describes the analysis of apoplastic fluids upon infection with *Xcc* and its non-pustule-inducing mutant *XccΔpthA4*.

2.6.1 Overexpression of *CsLOB1* increased ion leakage in *N. benthamiana eds1*

Our studies suggest that like SlLOB1, CsLOB1 is a regulator of fruit ripening, a process that is known to affect cell wall integrity (CWI) (Jia et al., 2023). Accordingly, we wondered if *CsLOB1* expression in *N. benthamiana* would have any obvious effects on CWI. Overexpression of *CsLOB1* in *N. benthamiana* leaves did not result in pustule formation that was known to be characteristic for expression of *CsLOB1* in citrus leaves. However, the *N. benthamiana* leaf tissue in which we expressed *CsLOB1* appeared to be soft and chlorotic. The degree of tissue chlorosis was visibly higher when *CsLOB1* expression was maintained in the presence of *p19,* which appears to mimic water soaking phenotypes observed in citrus (Figure 43).

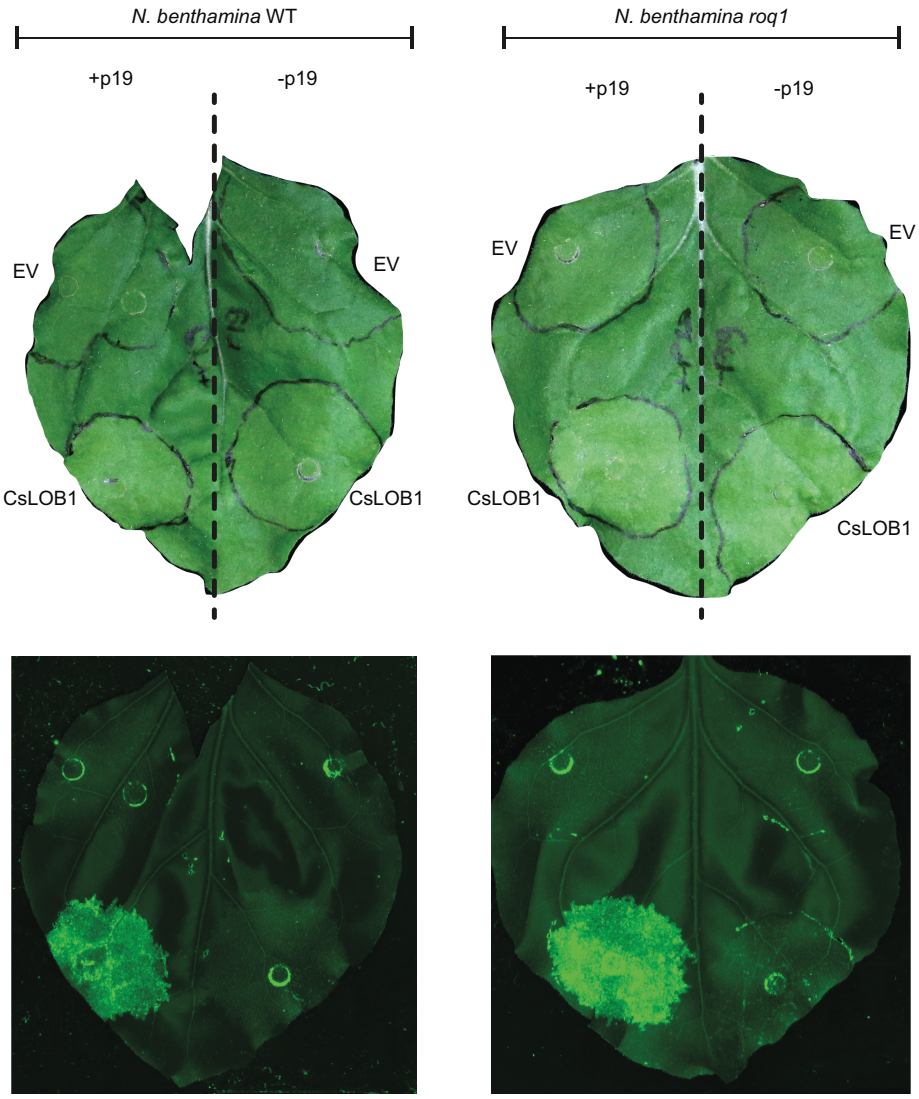


Figure 43: Transient overexpression of *CsLOB1* in *N. benthamiana* led to chlorotic tissue.
p35S-driven CsLOB1-GFP or empty vector (EV) together with or without *p35S-driven P19* (+p19 or -p19, respectively) was delivered to *N. benthamiana* wild type (WT). The Infiltrated area is marked in black. Expression of CsLOB1 was examined by GFP fluorescence using Typhoon™ (Cytiva) with default GFP fluorophore setting at 5dpi.

We performed an electrolyte leakage assay to investigate whether there are changes in overall CWI upon *Agrobacterium*-delivery of either p35S-*driven CsLOB1* or empty vector in *N. benthamiana* leaves (Section 4.2.6.1). We observed an approx. 3-fold increase in conductivity in the tissue that has *CsLOB1* expression (Figure 44). Interestingly, the presence of *Xcc* in the tissue did not significantly affect the level of ion leakage (Figure 44). This suggests that *CsLOB1* expression is the reason for the change in CWI.

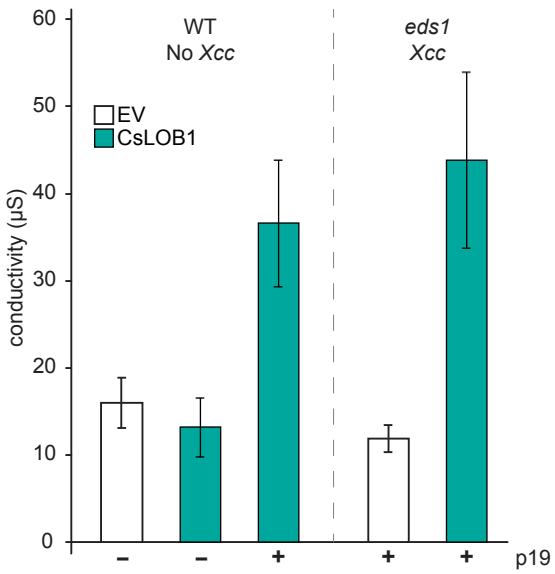


Figure 44: *CsLOB1* expression increases tissue conductance

p35S-driven CsLOB1 or empty vector (EV) was delivered into *N. benthamiana* wild type (WT) or *eds1* mutant by *Agrobacterium*. *Xcc* was co-infiltrated with the *Agrobacterium* in *eds1*. The presence (+) or absence (-) of *p35S-driven P19* was indicated. Leaf discs were harvested at 6dpi and subjected to electrolyte measurement. Bar graph represents the mean ± SD of 4 biological replicates.

2.6.2 *Xcc*-infected Citrus apoplastic fluid contained higher amount of sugar and promoted in vitro *Xcc* growth

We recovered the apoplastic fluid or intercellular washing fluid (IWF) in a centrifuge-based method from leaves infected with either wild type *Xcc*, the *XccΔpthA4* mutant or water. Subsequently, recovered IWFs were filter-sterilised pre-washed *Xcc* cells were added and *in-vitro* growth of bacteria was monitored over a period of 8 hours. Due to the limited amount of IWF that was obtained from citrus leaves (about 300µl/gram of leaf), we mixed the IWF with the minimal media (MME) for the growth assays. The MME medium contains minimal amount of salt and the amino acids including casamino acids and sodium-L-Glutamate that maintain the survival of *Xcc* without significant growth increase (Table 3) (Brich, 2017, master's thesis). Therefore, its diluting effect to IWF will be less tremendous to the bacterial growth compared to using water as the diluting solution and has been used as minimal medium for *Xanthomonas in vitro* growth assays (Turner et al., 1984, Turner et al., 1985). As a result, we observed that *Xcc* in the IWF from *Xcc* wild type-infected leaves doubled its optical density at 600nm (OD_{600}) already after 6 hours, whereas in the same time period IWFs from *XccΔpthA4*- or water-infiltrated leaves, *Xcc* only increased by about 35% compared to the initial OD_{600} (Figure 45).

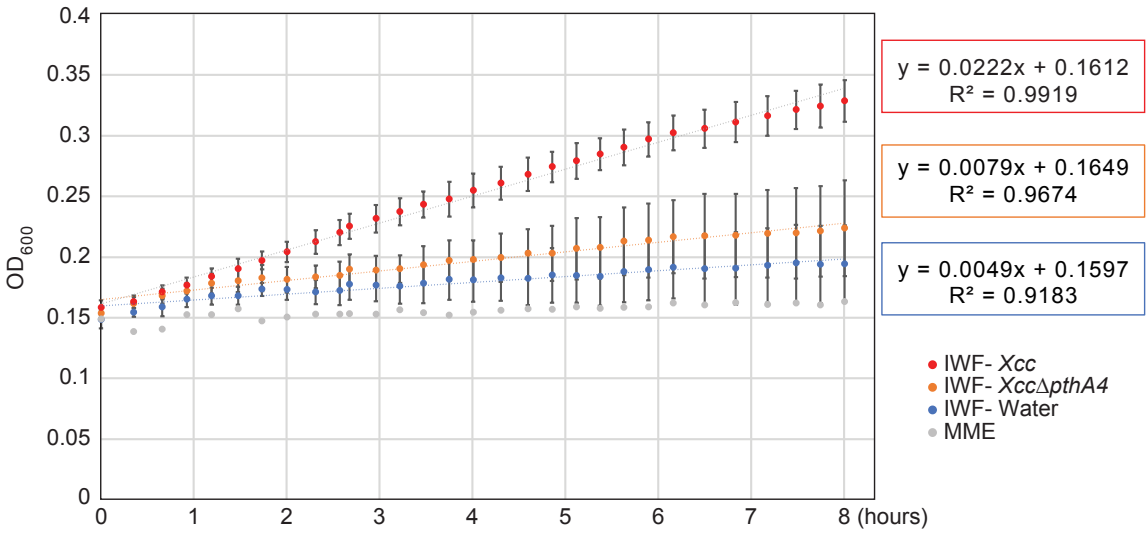


Figure 45: Apoplastic fluid extracted from *Xcc* infected tissue promotes *Xcc* growth *in-vitro*
Apoplastic fluid or intercellular washing fluid (IWF) was centrifuge-based extracted from leaves infiltrated with water, *XccΔpthA4* or *Xcc* wild type at 7dpi in three replicates, diluted 1:1 (v, v) ratio with the minimal media MME and mixed with prewashed *Xcc*. Each data point represents the average of the optical density (OD) of *Xcc* mix at 600nm in approx. 20-minute increments over the course of 8 hours. Error bars represent the standard deviation of the three biological replicates except for the MME control with one replicate. The overall growth pattern was assessed using a linear regression model, with the equation and coefficient of determination depicted.

We reasoned that there may be metabolites that are differentially accumulated in the IWF of *Xcc*- infected leaves and promote *Xcc in-vitro* growth compared to the IWF of *XccΔpthA4*-infected or water-infiltrated leaves. We hypothesised that these metabolites were released due to the action of cell wall loosening proteins and CWDEs induced by CsLOB1. Sugars are likely to be one metabolite type released as a result of the cell wall degradation and eventually serves the bacteria as a carbon resource. Using anion-exchange chromatography coupled to ion exchange chromatography-pulsed amperometric detection (IEC-PAD), sugars in the IWF were quantified, which showed that glucose and fructose increased approximately 3-fold, while sucrose decreased after *Xcc* infection, in contrast to the PthA4-deficient mutant (Brich, 2017, master's thesis). Although the 3-fold increase in glucose and fructose was not sufficient to explain why the presence of PthA4 resulted in at least a 10-fold increase in bacterial growth, it must be taken into account that at the beginning of the project in 2017, unfortunately we did not have a suitable greenhouse temperature and humidity for citrus plants and *Xcc* infection. As a result, infection assays did not result in consistent disease development, which means it is likely that sugar measurement was performed on leaves that did not show symptoms. In addition, the plant model used at that

time was Sweet Orange (*C. sinensis*), while we used Duncan grapefruit for our RNA-seq and ChIP-seq analyses due to faster disease symptom development and easier infiltration practice. Therefore, the sugar measurement was repeated with IWF extracted from Duncan leaves infected with *Xcc*, *XccΔpthA4* or water as a control using the currently available in-house gas chromatography-mass spectrometry (GC-MS) system run by Dr. Joachim Kilian. The GC-MS was also able to target some common mono- and disaccharides.

Our first attempts to measure sugars in the IWF by GC-MS showed no significant change in the sugars of interest. We then decided to harvest the leaves for IWF extraction at the time point of first symptoms. We observed that the levels of glucose, fructose and mannose in the *Xcc*-infected IWF increased by 5 to 6-fold compared to the *XccΔpthA4* sample, while remaining quite low in the water control (Figure 46). Isomaltose and xylose also showed an increasing trend, but the total amount, although shown in arbitrary units, was much lower compared to glucose, fructose and mannose, possibly suggesting that they may not act as a major nutrient source. In contrast, galacturonic acid and arabinose showed no significant difference between the *Xcc*-infected and *XccΔpthA4*-infected samples. As these two monosaccharides are the building block of pectin, a polysaccharide predicted to be targeted by CsLOB1-induced cell wall enzymes such as pectate lyase, pectinesterase, xyloglucanase, it was unexpected that their levels in the IWF was not upregulated in the presence of *Xcc*. One possible explanation was that, as *Xcc* that contain PthA4 grow faster than the mutant, they might quickly deplete the carbon source, making it harder to measure the difference. Sucrose showed inconsistent patterns between different experiments, with none showing a significant increase in *Xcc*-infected IWFs compared to *XccΔpthA4*-infected IWFs (Figure 46).

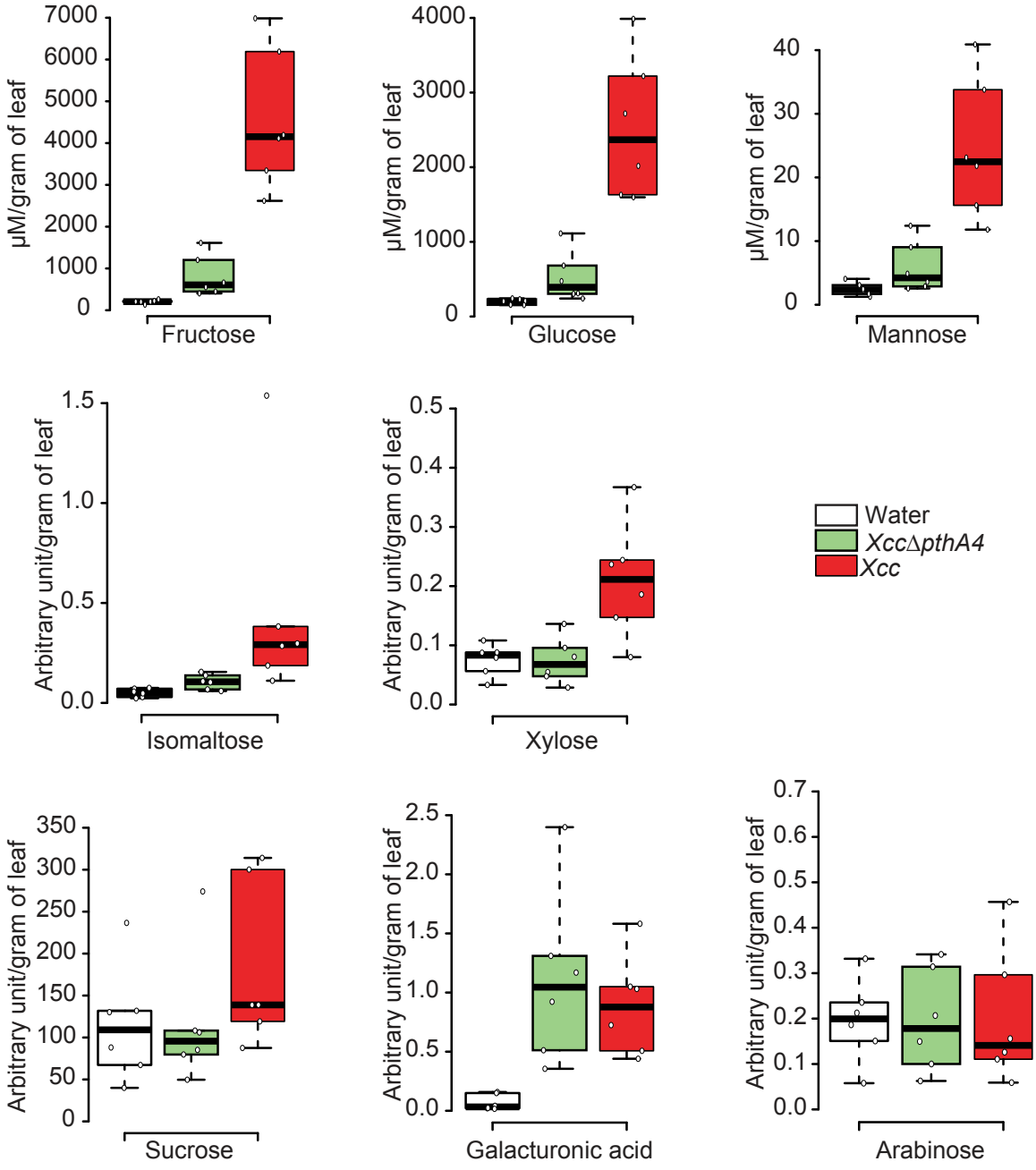


Figure 46: Apoplastic fluid increases in fructose, glucose and mannose when infected with *Xcc* but not with the mutant lacking PthA4 (*XccΔpthA4*) or water control.
Apoplastic fluid was centrifuge-based extracted from Duncan leaves infiltrated with water, *XccΔpthA4* or *Xcc* wild type at 7dpi and measured by gas chromatography mass spectrometry (GC-MS) at µM per gram of leaf for fructose, glucose, mannose and arbitrary units for the others due to the missing calibration curves. The signal was normalised with the internal factors 3' O-Methyl-D-Glucose. Box plots show sugars measured in µM/arbitrary unit per gram of leaf from six biological replicates. Centre lines show the medians; box limits show the 25th and 75th percentiles as determined by R; whiskers extend 1.5 times the interquartile range of the 25th and 75th percentiles; white dots indicate individual data points. GC-MS was done by Dr. Joachim Kilian.

However, we noticed that *Xcc*- infected IWFs appeared to have greenish colour compared to *XccΔpthA4*- infected IWFs and the water-infiltrated leaf IWFs at 7dpi when pustule and water

soaking phenotypes are already visible (around 7dpi). The colour change may indicate more cytoplasmic contents that have leaked into the apoplastic fluid. Due to the limited time and sample volume, we did not examine the level of cytoplasmic contamination in the IWFs in Figure 46. However, for the dataset coming from *Xcc*-infected but not yet symptomatic IWFs, a preliminary examination of cytoplasmic contamination was done by quantifying Rubisco protein using in-solution trypsin digestion and liquid chromatography-mass spectrometry (LC-MS) with the help from Dr. Edda von Roepenack- Lahaye. This method has been shown to be the most sensitive compared to the one measuring the activity of glucose-6-phosphate dehydrogenase (G6PDH), an abundant cytosolic enzyme that converts NADP+ (Nicotinamide adenine dinucleotide phosphate) to NADPH (Brich, 2017, master's thesis). We observed an approx. 1.5-fold increase in the mean level of Rubisco in the *Xcc*-infected IWF compared to the *XccΔpthA4* or water control samples (Figure 47). In this dataset, we observed a non-significant change, although with a tendency to increase, in the levels of fructose, glucose and mannose in the *Xcc*-infected IWF as compared to *XccΔpthA4* and water control (Figure 47). In contrast, sucrose levels were relatively lower in the *Xcc*-infected sample compared to the other controls. In summary, our findings suggest that PthA4-mediated activation of CsLOB1 by *Xcc* affects CWI. The release of cellular contents of some cells in the leaf will eventually induce higher level cytoplasmic proteins in the IWF before the tissues have expressed visible disease symptoms.

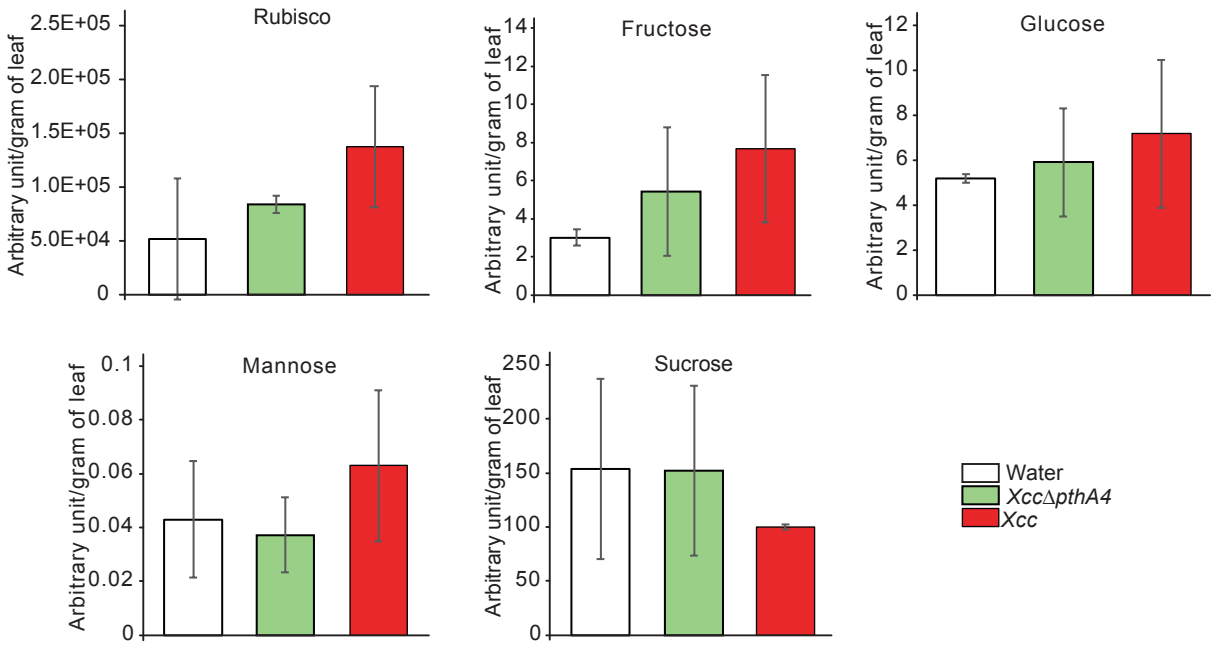


Figure 47: *Xcc*-infected tissues have more cytoplasmic contamination
Apoplastic fluid was centrifuge-based extracted from Duncan leaves infiltrated with water, *XccΔpthA4* or *Xcc* at 7dpi and measured by gas chromatography mass spectrometry (GC-MS) at arbitrary unit per gram of leaf for fructose, glucose, mannose and sucrose. Rubisco was measured by LC-MS (Dr. Edda von Roepenack-Lahaye). The signal was normalised with internal factors (trypsin peptides in Rubisco LC-MS and 3' O-Methyl-D-Glucose in sugar GC-MS) and presented in arbitrary units against total leaf weight. Error bars represent the standard deviation of 3 biological replicates.

We have seen that *CsLOB1* overexpression has increased electrolyte leakage in *N. benthamiana*, indicating weakened CWI (Figure 44), which suggests there might be higher cytoplasmic contamination upon physical application during IWF extraction. In pectin-stained *Xcc*-infected tissues at 5dpi, spongy mesophyll cells expanded into the apoplastic region in a distorted manner as compared to tissues infected with *XccΔpthA4* or water control (Figure 48). The mesophyll junction sites in the water and *XccΔpthA4*-infected tissues showed strong pectin staining by ruthenium red (Figure 48). In comparison, the signal in the *Xcc*-infected samples was less pronounced, although this observation needs to be confirmed by quantification. Junction sites with the pectin-rich middle lamella are known to hold mesophyll cells in space and maintain the integrity of the leaf tissue (Daher and Braybrook, 2015). Therefore, it is likely that the overall CWI was compromised, resulting in high cell damage during physical application, leading to higher cytoplasmic contamination.

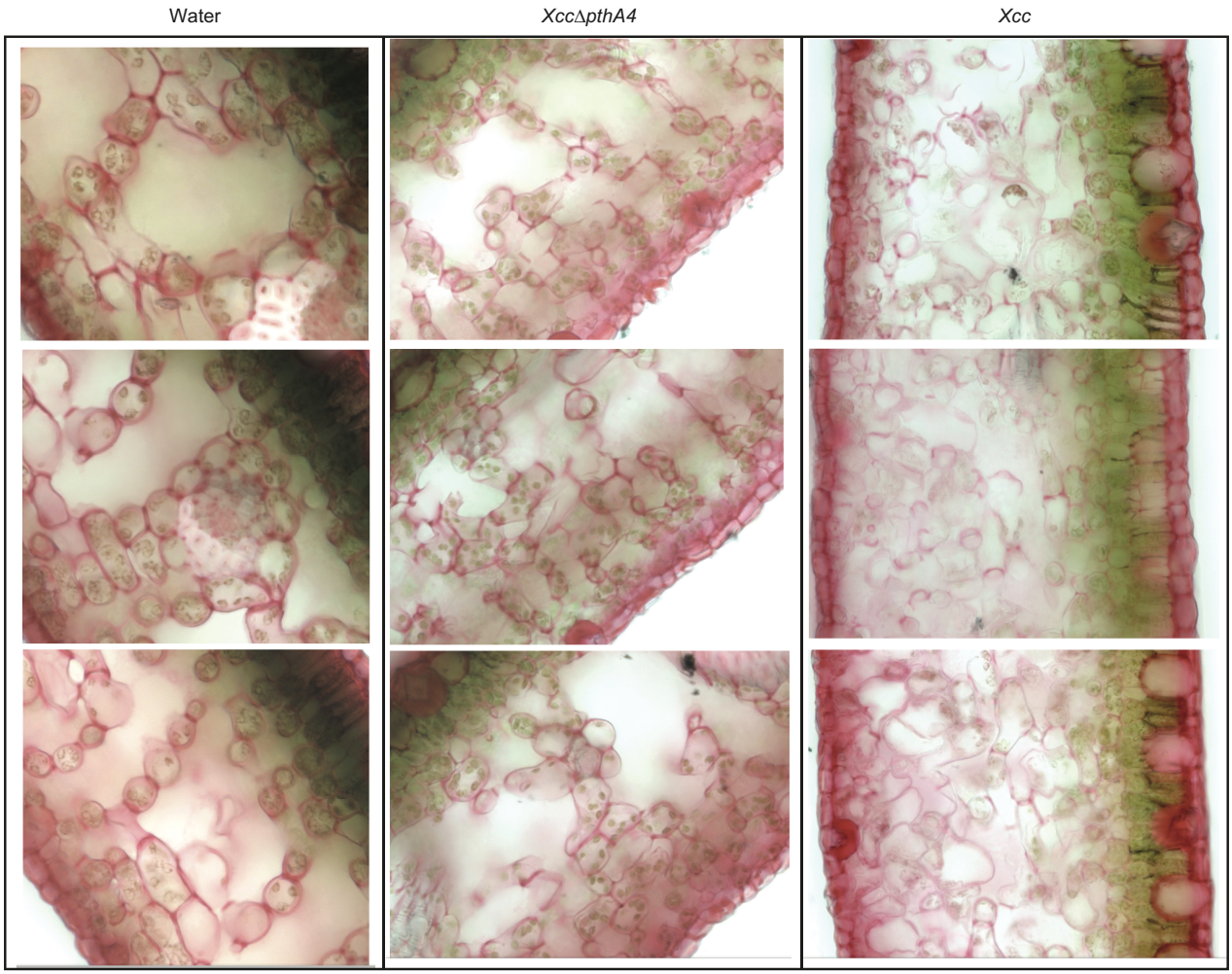


Figure 48: Pectin-stained cross sections of Duncan leaves infiltrated with water, *XccΔpthA4* or *Xcc* show distorted cell shape of *Xcc*-infected leaf tissue

Duncan leaves were infected with water, *XccΔpthA4* or *Xcc* and harvested after 5dpi. Leaf pieces were fixed in formaldehyde and sectioned with a vibratome before staining with ruthenium red dye. Observations and photos were taken under bright field microscope by Dr. Sandra Richter and her Team.

Pectin is known to be particularly enriched in the middle lamella at the cell-to-cell contact sites (reviewed by Daher and Braybrook (2015)). The degree of methylesterification of pectin building block, homogalacturonan (HG), determines whether pectin can be targeted by degraded by enzymes. The activity of pectinesterase, also known as pectin methylesterase, removes the methyl group from HG. Depending on whether the dimethylesterification is random or blockwise, the demethylesterified HG can be targeted by pectate lyase or pectinesterase or crossed link with calcium ions to form pectin gels and increase cell wall stiffness (Wormit and Usadel, 2018). As a result, pectin is degraded, and the cell wall is loosened, allowing the cell to expand. We have shown that CsLOB1 directly activates CWDEs, including pectinesterase and pectate lyase. The presence of these enzymes may alter the

degree of methylesterification of cell wall pectin. However, our preliminary test with an antibody against the highly esterified HG showed no obvious difference between the *Xcc* and *XccΔpthA4*-infected tissues. Due to the complexity of different esterified HGs, there is a wide range of pectin antibodies to be tested, which we have not covered within the scope of this thesis.

2.7 *Xcc* degrades xylan as a carbon nutrient source

Metabolic changes in the apoplastic fluids upon host gene activation might give us a clue of what carbon source can be utilised by *Xcc*. However, this does not directly reflect what *Xcc* consumes. At the same time, due to *Xcc* fast growth, the carbon source can be depleted quickly before the measurement. In addition, cytoplasmic contamination during IWF extraction can have a great impact on the metabolite contents, especially the level will be greater when the tissue is experiencing cell wall degradation and expansion, reducing CWI (Figure 48). Bacteria quickly adapt their metabolism to the available carbon source in which enzymes required for the breakdown, uptake and metabolism of the carbon substrates will be quickly produced, which reflects in the expression of corresponding genes (Eisenreich et al., 2010). Therefore, we investigated *Xcc* transcriptomic changes when they infect citrus and promote water soaking and pustules at 5dpi in comparison to the non-symptomatic *XccΔpthA4* mutant.

To reduce the contamination of plant RNA, which is much abundant compared to bacterial RNA, we did not extract total leaf RNA. Instead, IWFs containing the bacteria at 5dpi were extracted first and bacterial cells were quickly fixed, harvested and subject to RNA extraction (Section 4.2.2.2). Due to the lower yield of IWFs and bacteria, we infiltrated either *Xcc* or *XccΔpthA4* into three leaves on the same plant and pooled the IWFs together as one biological replicate. To make sure that the changes are not intrinsic due to the knockout of *PthA4* but are due to the interaction with the plant host in a PthA4-dependent fashion, we included the control of *Xcc* and *XccΔpthA4* that grew on the synthetic NYG media. Bacterial and plant rRNAs were depleted before cDNA library preparation and sequencing.

Dr. Mario Murakami and colleagues kindly processed the sequencing outputs for us. Strand-specific reads were aligned to the genome of *Xcc* strain 306. PCA analysis showed that while media-growing strains well clustered, the *in-planta* strains were not distinctly separated

(Figure 49A), indicating that the *in-planta* global transcriptomes of the wild type and PthA4 mutant are not so much different as expected, as they are isogenic. Differential expression analysis shows that there are 120 upregulated- and 115 downregulated- DEGs in *Xcc* compared to *XccΔpthA4 in-planta.* There are also a considerable number of DEGs between *Xcc* and *XccΔpthA4* in NYG media (Figure 49B). However, the genes that show upregulation or downregulation *in-planta* but not when grown in media were most focused on as the changes in their expression were likely to be dependent on CsLOB1-mediated host changes. Overlapping analysis shows that among the DEGs found in leaf, 104 and 96 DEGs were specifically upregulated and downregulate, respectively (Figure 49B). While upregulated DEGs are enriched in transmembrane transport and response to chemical, downregulated genes are overrepresented in the response to oxidative stress (Figure 49C).

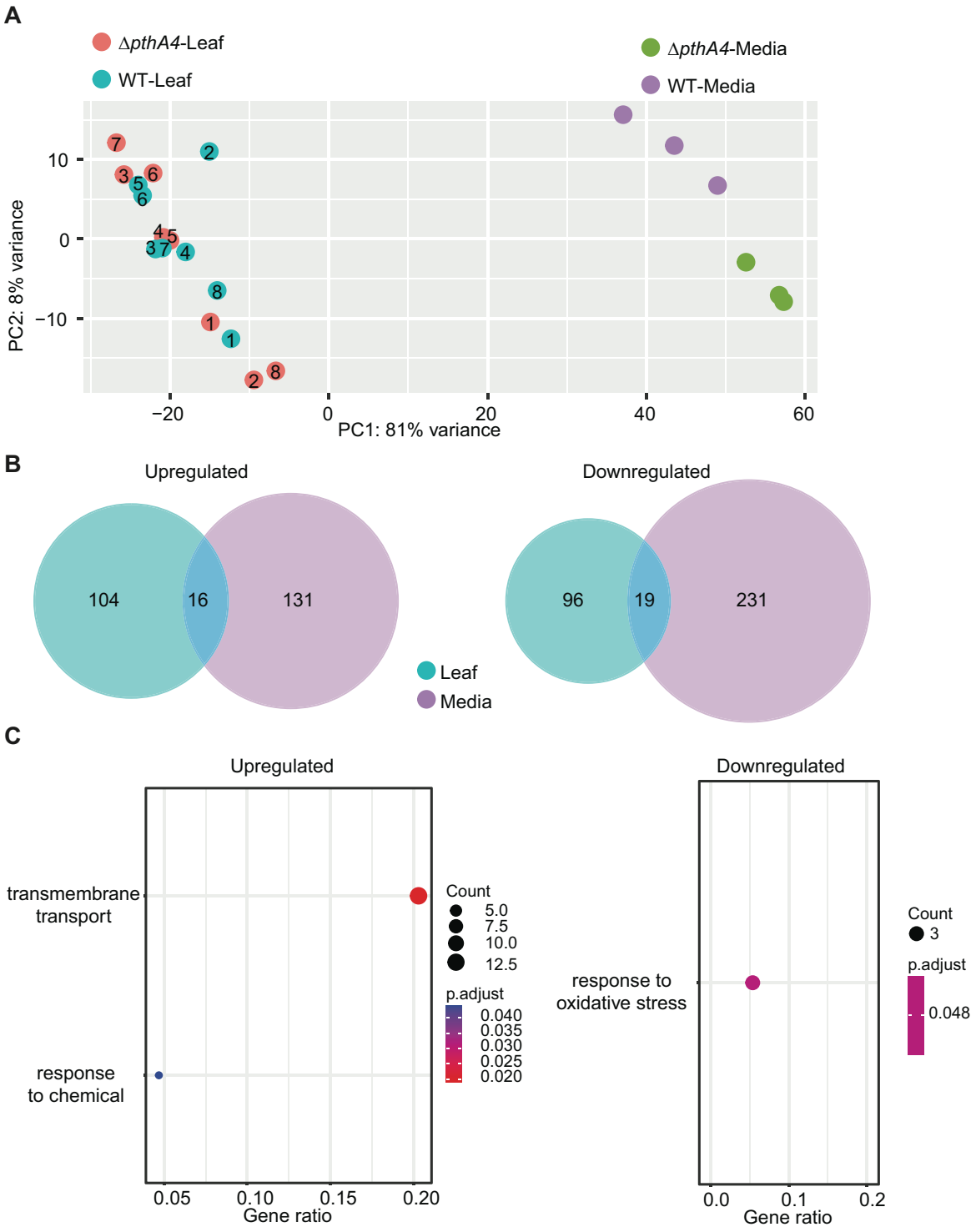


Figure 49: RNA-seq revealed genes differentially expressed in PthA4 containing *Xcc* compared to *XccΔpthA4*

A: The transcriptomes of *Xcc* and its isogenic *ΔpthA4* mutant extracted from Duncan leaves at 5dpi show low variation among each other but well separated from *Xcc* and *XccΔpthA4* growing in NYG media. *Xcc* and *XccΔpthA4* growing were syringe-infiltrated on different leaves on the same plant as one biological replicate (replicate numbers in the circle). As a control for non-plant-interaction changes, media-growing strains were also sequenced, each with three replicates (colour depicted). **B: 104-upregulated and 96-downregulated DEGs when comparing *Xcc* to *XccΔpthA4* are specific to bacteria-leaf interactions.** Size-proportional Venn diagrams show differentially expressed genes (DEGs) (fold

change ≥ 2 or ≤ -2 (FDR <0.05)) in *Xcc* (containing PthA4) relative to *XccΔpthA4* extracted from 5dpi Duncan leaves (Leaf) and from NYG media (Media). The number of genes in each overlap is in bold black. **C: Upregulated DEGs are enriched in transmembrane transport and response to chemical while downregulated DEGs are enriched in response to oxidative stress.** Functional enrichment analysis was performed with the R package clusterProfiler (v4.2.2) (Yu et al., 2012) using the *Xcc* strain 306 gene ontology term from PANNZER2 server (Toronen et al., 2018). Circle sizes are proportional to the number of genes belonging to a given category. Gene ratio axis show the percentage of the genes among total given genes. The circle colour scale indicates the adjusted p-value (ranging from 0-0.05). Analysis was done by Dr. Mario Tyago Murakami's lab, Centro Nacional de Pesquisa em Energia e Materiais, São Paulo, Brazil.

Excitingly, many genes that belong to the xylan utilisation system are co-expressed in *Xcc* compared to *XccΔpthA4* in leaf samples but not differentially expressed in when *Xcc* grew on the synthetic media, indicating that *Xcc* particularly metabolises xylan upon infection when PthA4 is present (Figure 50). *Xanthomonas* has a large repertoire of CAZymes (Lombard et al., 2014), many of which are co-localised and likely to be co-regulated in a multienzymatic Carbohydrate Utilisation system associated with TonB-Dependent Transporters (CUT system). The CUT system for xylan has been studied extensively in *Xcc* strain 306 (Santos et al., 2014, Chow et al., 2015). One cluster of the xylan CUT system was upregulated in *Xcc* growing *in-planta* as compared to *XccΔpthA4* (Figure 50). These include xylanase that catalyses the first step in xylan breakdown, xylosidases, glucosidases that remove side chain sugars and transmembrane transporters (Figure 50A). In contrast, their transcript levels in the strains from media did not significantly differ (Figure 50B). Interestingly, though citrus pectin degrading enzymes were highly upregulated, none of pectin degrading genes were upregulated in *Xcc* compared to *XccΔpthA4* mutant. Similarly, enzymes targeting other cell wall polysaccharides such as xyloglucan and cellulose were not enriched. Our finding was the first to show that *Xcc* specifically degrades and utilises xylan as a carbon source in a PthA4/CsLOB1 dependent fashion.

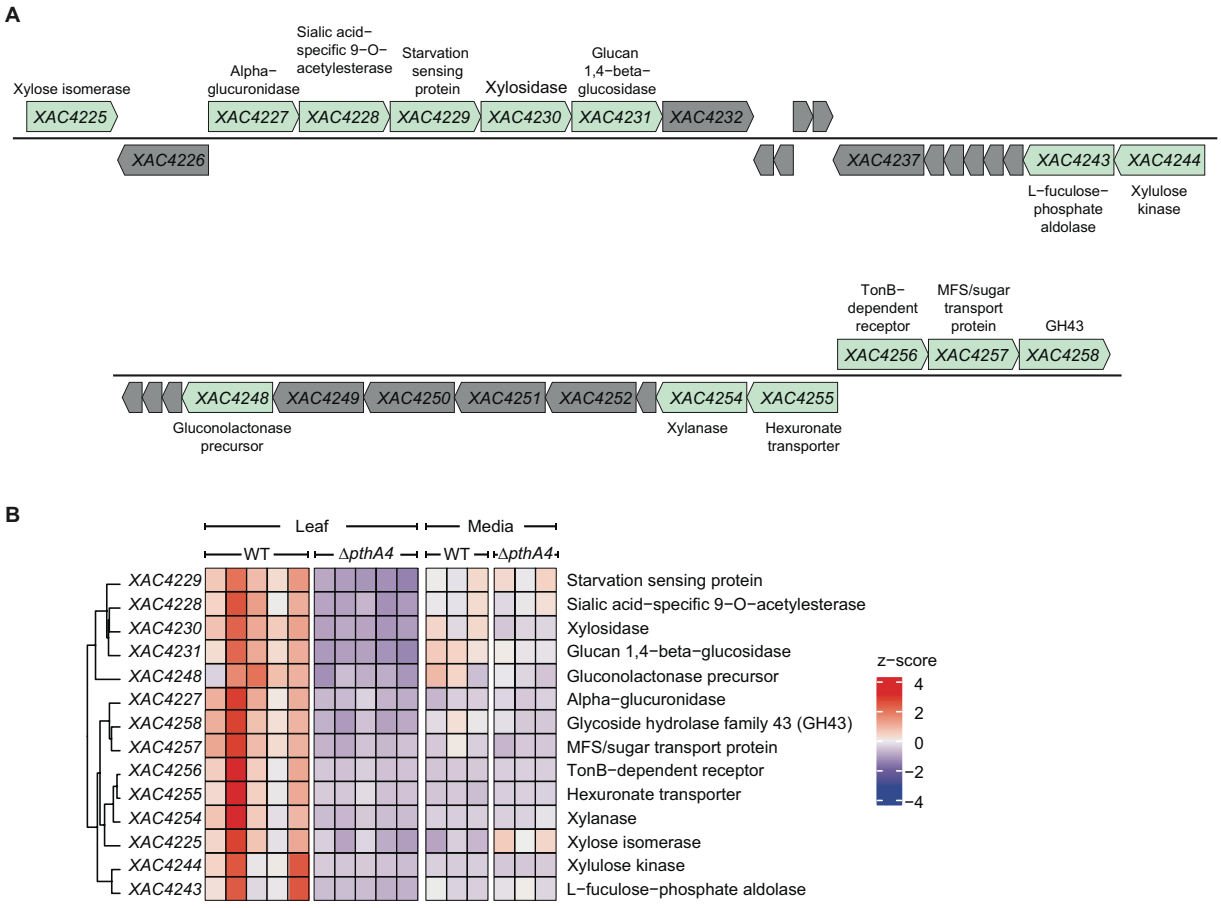


Figure 50: Xylan utilisation system was activated during Citrus canker in a PthA4-dependent fashion

A: Schematic illustration of the xylan CUT (Carbohydrate Utilisation system associated with TonB-Dependent Transporters) system in *Xcc* strain 306 genome, modified from (Chow et al., 2015) and (Giuseppe et al., 2023). Arrow direction shows gene direction. Blue arrows are gene upregulated during *Xcc* infection when compared to *XccΔpthA4*. Grey arrows are non-differentiated genes. Hypothetical genes are in grey arrows without gene annotation. **B: 14 xylan CUT genes are highly expressed only in *Xcc* when they infect citrus, but not in *XccΔpthA4* or media growing strains.** All transcripts per million (TPM) values of the genes were converted to Z scores using ComplexHeatmap Bioconductor package (Gu et al., 2016). The R script was adapted based Dr. Paulo Teixeira's R script (Appendix 1.5). Genes with similar expression profiles were clustered into subclades. Each coloured square represents the Z-score of each gene in a single biological replicate. Only replicates 1, 2, 5, 6, 8 for leaf samples were included in the analysis. Green: cell-wall related genes, black: other function.

2.8 Common mechanisms for TALEs targeting host transcription factors?

Since being discovered 30 years ago, the understanding in TALE mechanistic function has been advanced tremendously. The basic function of TALEs is to activate host target genes, known in rice, citrus, pepper, tomato, cassava and cotton (reviewed in Teper et al. (2022)). Amongst TALE targets, the activation of *SWEET* genes was known to be absolutely required for *Xanthomonas* full virulence in rice, cassava and cotton (Chen et al., 2012, Cohn et al.,

2014, Cox et al., 2017). The upregulation of *SWEETs* has been hypothesised to promote sucrose efflux into the vascular tissues; however, the knockout of a sucrose importer, *SuxC*, in *Xpm* did not affect *Xpm in-planta* virulence, questioning the direct benefit of sucrose availability to the bacteria (Cohn et al., 2014). At the same time, host invertase, which breaks down sucrose into glucose and fructose, was upregulated during *Xpm* infection, it is possible that the *XpmΔsuxC* mutant was still able to benefit from the increase in apoplastic sucrose upon *SWEET* activation, but this requires experimental evidence (Cohn et al., 2014). Similar to SWEETs, *CsLOB1* activation is essential for the full virulence of *Xcc* (Hu et al., 2014). We also observed an increase of simple sugars in the host apoplast upon the infection with the mesophyll-colonizing pathogen *Xcc* (Figure 46). In contrast, Citrus *SWEETs* do not play a role in citrus canker disease as we did not identify any *SWEET* homologs regulated by CsLOB1 in our RNA-seq (Supplemental Table 4), which was also seen by Hu and colleagues (Hu et al., 2014). Our research has revealed that *Xcc* uses PthA4 to activate *CsLOB1* to hijack the fruit softening process to loosen host cell wall to gain access to xylan as their carbon source. This is a novel strategy and has not been described in other *Xanthomonas* strains. Interesting, another group of TALEs from the pepper/tomato pathogenic *Xanthomonas* species, AvrBs3 and AvrHah1, convergently target TFs that belong to the *bHLH* family. These TFs, including *UPA20* in pepper and *bHLH3* and *bHLH6* in tomato, were shown to activate the same pectate lyase that promotes water soaking phenotype in pepper and tomato (Schenstnyi et al., 2022, Schwartz et al., 2017). We question whether *bHLHs* have been hijacked by *Xanthomonas* for a similar purpose as *CsLOB1*. This chapter describes the preliminary analysis of the similarity between the CsLOB1 and bHLH regulons involved in cell wall degradation.

2.8.1 CsLOB1 activate downstream target of bHLHs

We examined whether CsLOB1 could activate known targets of bHLHs, the tomato pectinesterase (*SlPE*, *Solyc11g019910*) and pectate lyase (*SlPL*, *Solyc05g014000*). To do so, we co-infiltrated *N. benthamiana* leaves with *p35S-driven bHLH3*, *UPA20*, or *CsLOB1* or an empty vector control in combination with a GUS reporter that is under transcriptional control of the *SlPE* or *SlPL* 5' upstream sequences. We found that bHLH3 and UPA20 activated tomato *PL* and *PE* promoters, which agrees with previous finding (Schenstnyi et al.,

2022). Excitingly, *CsLOB1* also activated the promoters bHLH3/UPA20 target promoters *SlPL* and *SlPE* (Figure 51) (Aguilera, 2023, bachelor's thesis).

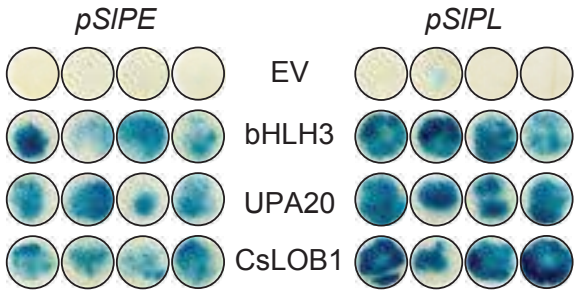


Figure 51: CsLOB1 activates Tomato pectinesterase and pectate lyase, the targets of bHLH3/6 and UPA20

1kb upstream of the ATG of tomato pectinesterase (*pSlPE*, *Solyc11g019910)* and pectate lyase (*pSlPL*, *Solyc05g014000*) were cloned to drive the *uidA* coding for the GUS reporter. The depicted reporter constructs were delivered into the leaves of *N. benthamiana* by *Agrobacterium* together with either p35S-*driven bHLH3* (*Solyc03g097820*), *UPA20* (*CA03g22700*), *CsLOB1* or an empty vector control (EV). The inoculated leaf discs were collected at 2dpi for GUS staining. Figure was adapted from Aguilera's bachelor thesis (2023).

The activation of *SlPL* and *SlPE* by CsLOB1 suggests that CsLOB1 and bHLHs may have common downstream targets. In tomato, *SlLOB1* was shown to activate another pectate lyase, *Solyc03g111690,* hereafter referred to as *SlPL03* (Shi et al., 2021). We compared the protein sequence of SlPL, SlPL03 and Cs2g23970 and Cs8g11330-the two pectate lyases found to be the direct targets of CsLOB1 (Figure 13). Protein alignment revealed a very high homology between the citrus and tomato pectate lyases (Figure 52). SlPL03 shows 80% protein sequence identity with SLPL and 76% and 70% identity with Cs2g23970 and Cs8g11330, respectively. This suggests that their mode of action in pectin degradation may be very similar.

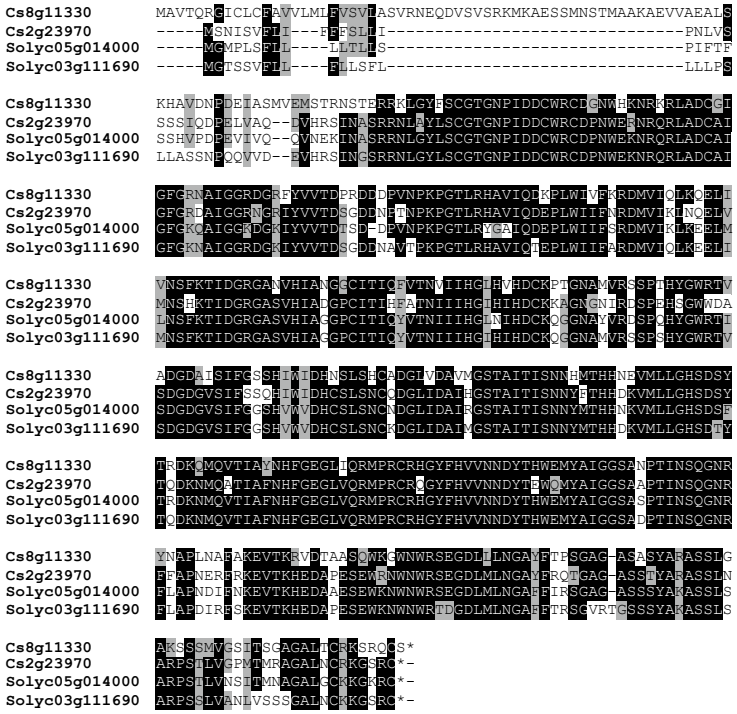


Figure 52: Tomato pectate lyases have high homology to Citrus pectate lyases
Full length protein sequence alignment of citrus pectate lyases, Cs2g23970 and Cs8g11330 and tomato pectate lyases, Solyc05g014000 (SlPL) and Solyc03g111690 (SlPL03). Letters highlighted in black indicate identical amino acids. Red letters highlighted in black are characteristic conserved amino acids. Grey highlighted letters indicate amino acids with similar properties. Non-highlighted letters are non-identical amino acids.

We explored whether there are conserved cis elements in the upstream regions of these pectate lyases that could be targeted by the two TF family members. However, promoter comparison did not reveal any significant conservation. Next, we searched for predicted CsLOB1 binding sites in the 3kb upstream and 200bp downstream of the *SLPL* TSS using the FIMO tool (Grant et al., 2011) and CsLOB1 motif matrix found in ChIP-seq. We found that the best match site to the CsLOB1 motif located 150bp upstream of the TSS with a p-value of 0.00172. Similarly, the *SlPL03* promoter, which was activated by SlLOB1 in *N. benthamiana,* also contains predicted CsLOB1 binding sites 2.3kb and 572bp upstream of the TSS with p-values of 0.000477 and 0.000913, respectively. In contrast, the predicted binding sites in citrus pectate lyases, *Cs8g11330* and *Cs2g23970,* have p-values of 5.55e-07 and 6.046e-05, respectively (Figure 53).

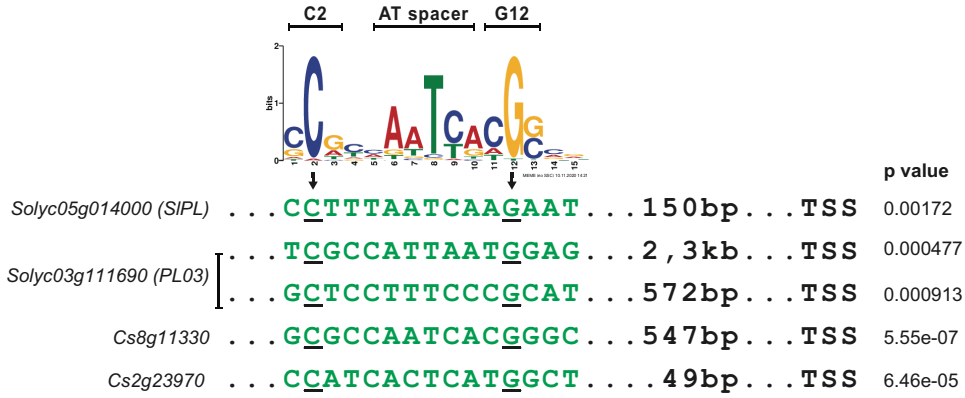


Figure 53: Predicted CsLOB1 binding sites in tomato and citrus pectate lyases
3kb upstream and 200bp downstream of the TSS of *SlPL* (*Solyc05g014000)* and *PL03* (*Solyc03g111690*) were subjected to CsLOB1 motif scanning using FIMO (Grant et al., 2011). The best matching sequences were shown in green, with the distance to the TSS in black. The second site of *PL03* was the best match within the 1-kb *PL03* promoter region reported to be activated by SlLOB1 (Shi et al., 2021). Predicted CsLOB1 binding sites of citrus pectate lyases, *Cs8g11330* and *Cs2g23970* were found within ChIP-seq peaks. The p-value of a motif occurrence represents the probability of a random sequence of the same length as the motif matching that position of the sequence with an equal or better score.

Whether the predicted binding motifs are able to interact with CsLOB1 remains to be experimentally validated, however, the p-value, which indicates sequence relatedness of a sequence to a given motif, makes an interaction between CsLOB1 and the given site a plausible scenario. The EMSA and MST assays showed that the mutated variants of CsLOB1 binding motif have different affinities to CsLOB1. While the m5 mutant with a single A to G mutation in the low conserved spacer still retained the binding capacity (Figure 16, Figure 17) and has a p-value of 0.000952, the other mutants showed a higher p-value (>0.003) and were unable to bind to CsLOB1 (Figure 54). In contrast, the promoters of *SlPL* and *SlPL03* have best predicted binding site p-values of 0.00172 and 0.000477, which are lower than those of the mutants used in EMSA. Therefore, it is plausible that these sites are capable of binding to CsLOB1/SlLOB1.

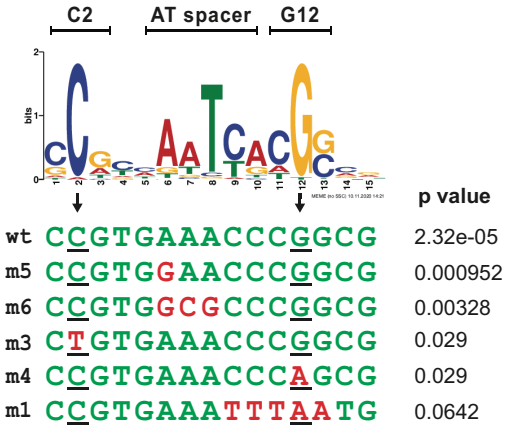


Figure 54: Motif occurrence probability of probes used in EMSA
1.2kb upstream of ATG Cs2g20600 containing the predicted CsLOB1 binding site (wt, green letter) was subjected to motif scanning together with variants consisting of exactly the same sequence except for the CsLOB1 binding site which was replaced by mutant variants (m1, m3, m4, m5, m6). The p-value was only available for the depicted mutants because the other mutations had p-value > 0.1 and were eliminated from the analysis. Green: wild-type bases, red: mutated bases.

In short, CsLOB1, as a target gene of the *Xcc* TALE PthA4, was able to activate the downstream targets of the bHLH transcription factors, the targets of the *Xe* TALEs AvrBs3/AvrHah1. This suggests that the *Xcc* TALE PthA4 and *Xe* TALEs AvrBs3/AvrHah1 manipulate host cells in a similar fashion.

2.8.2 bHLH regulons might partially differ from CsLOB1/SlLOB1 regulons

Considering the observed activation of the bHLH3/6's target genes, *SLPL* and *SlPE,* by CsLOB1, we wondered if reciprocally CsLOB1 target genes are activated by bHLHs. Specifically, we tested whether bHLH3 and bHLH6 could activate the CsLOB1/SlLOB1 target genes *pCs5g20320,* encoding an endoglucanase from citrus, and *pSlEXP1,* encoding an expansin from tomato (Figure 27). We co-expressed either *p35S-driven bHLH3*, *bHLH6*, *CsLOB1* or empty vectors with *pCs5g20320* or *pSlEXP1* driving a GUS reporter gene. As a control, we included the promoter of *SlPL*, which has been shown to be activated by all the three transcription factors (Figure 51). As expected, *pSlPL* can be activated by bHLH3, bHLH6 and CsLOB1, confirming that the transcription factors were expressed. bHLH6 appeared to activate the citrus promoter *pCs5g20320,* but to a much lesser extent compared to CsLOB1. By contrast, the tomato promoter *pSlEXP1* was not significantly activated by bHLH3 and bHLH6 compared to the empty vector, whereas CsLOB1 strongly induced the promoter (Figure 55). This suggests that the regulons targeted by CsLOB1/SlLOB1 may differ to some extent from those targeted by bHLHs.

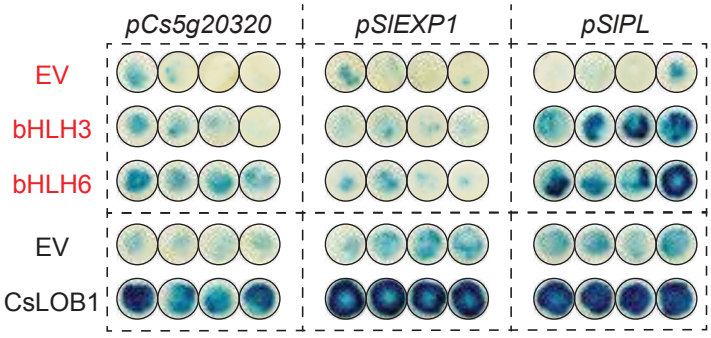


Figure 55: bHLH6 weakly activates CsLOB1-targeted promoter, *pCs5g20320*

1kb upstream of the ATG of tomato pectinesterase (*pSlPE)* and pectate lyase (*pSlPL*) 1kb upstream of the ATG of citrus endoglucanase 9 (*Cs5g20320*), tomato expansin 1 (*SlEXP1*), driving *GUS* reporter gene were co-expressed with either *p35S-driven bHLH3* (bHLH3)*, p35S-driven bHLH6* (bHLH6, *Solyc06g072520*) in *LIIαF1-2* backbone (red) or their empty vector (EV-red) or *p35S-driven CsLOB1* (black) in *pICH47732* backbone or its empty vector (EV-black) in *N. benthamiana* via *Agrobacterium* infiltration with four biological replicates. Leaf discs were harvested at 2dpi for GUS staining.

3. DISCUSSION

3.1 dTALEs eliminate PthA4's off targets

PthA4 was used as an activator of *CsLOB1* to study the downstream regulon of *CsLOB1*. However, due to the nature of TALEs, which bind to their *EBE* via the central repeat domain, PthA4 is likely to bind to and activate genes whose promoters contain an *EBE* related to the *EBE* present upstream of the *CsLOB1*. In addition, it is known that some TALE RVDs bind to more than one type of nucleotide, thereby increasing the chance of binding potential off-targets (Moscou and Bogdanove, 2009, Boch et al., 2009). An example of a TALE binding to an EBE-related sequence and activating a non-intended gene (or off-target gene) was described by Morbitzer et al. (2010). A dTALE designed to target *EGL3*, a *bHLH* transcription factor involved in root hair and trichome formation in Arabidopsis, was shown to activate another gene with a promoter containing a predicted TALE binding site that differs in two bases from the predicted target *EBE* (Morbitzer et al., 2010). This raised a concern about using PthA4 to study the CsLOB1 regulon as PthA4 off-targets would result in identification of genes that are not CsLOB1-dependent and would interfere with the elucidation of the *CsLOB1*-dependent regulon.

To overcome this, we generated two functional dTALEs, dTALE2 and dTALE4, targeting different *EBEs* in the *CsLOB1* promoter and generated corresponding transconjugants in *XccΔpthA4*. The dTALEs activate *CsLOB1* expression at a a similar level to PthA4. They also induce citrus canker in a PthA4-mutant background, indicating that Citrus canker is promoted via the shared genes activated by the dTALEs and PthA4, particularly CsLOB1 and its regulon, not by the off-target genes. This independently confirmed the observation made by Hu and colleagues that not the off-targets but CsLOB1 is the sole responsible susceptibility gene for citrus canker, and not the off targets (Hu et al., 2014).

Differential overlap analysis of DEGs induced by these dTALEs and PthA4 with the TALE-lacking strain *XccΔpthA4* revealed 942 shared *CsLOB1*-dependent genes, but surprisingly 495 off targets generated by PthA4, accounting for 34% of the total DEGs that were induced (Figure 4). The number of off-target genes induced by dTALE4 is also substantial, with 368 DEGs (42%) and dTALE2 with fewer number of off targets, 62 DEGs (6%). A similarly high number of off targets were observed by Zhao and colleagues when using two different

dTALEs to activate a MYB transcription factor to study its downstream genes in Arabidopsis (Zhao et al., 2023). The high number of off targets produced by PthA4 justifies the concern about using solely PthA4 in the study of the CsLOB1 regulon. This observation highlights the importance of using multiple dTALEs to activate a gene of interest at different *EBEs* to study downstream transcriptomic changes, as the influence of off-targets can be profound, as was observed with PthA4 and dTALE4 in the RNA-seq analysis.

3.2 *CsSWEETs* are not involved in the development of Citrus canker

SWEET genes encode sucrose efflux transporters, which actively export sucrose and glucose to the apoplast, where they are taken up by sucrose transporters (SUTs) for phloem loading (Chen et al., 2010, Chen et al., 2012). *SWEET* homologs from the same clade III in rice, cassava and cotton are targeted by the vascular *Xanthomonas* pathogens *Xoo*, *Xpm* and *Xcm*, respectively, for their full virulence (Table 1) (Chen et al., 2012, Cohn et al., 2014, Cox et al., 2017). It was hypothesised that the co-operation of SWEETs and SUTs resulted in increased levels of sugars in the phloem, which *Xanthomonas* used as a carbon source for growth (Chen et al., 2010). However, a knockout in *Xpm* sucrose transporter S*uxC* did not hinder bacterial growth and virulence in cassava, indicating that the direct usage of sucrose due to the enhanced sucrose efflux does not contribute to *Xanthomonas* virulence (Cohn et al., 2014). Invertase was shown to upregulated in the presence of *Xpm*, leading to an alternative model in which sucrose exported by SWEETs is quickly broken down into glucose and fructose by invertase and up-taken by the bacteria (Cohn et al., 2014), though this requires further experimental evidence. Nevertheless, *SWEETs* are essential conserved *S* genes for *Xanthomonas* virulence across phylotypes.

Table 1: *SWEET* homologs targeted by TALEs

Pathogen	TALE	Host	Target gene	Reference
Xoo	PthXo1	Rice	OsSWEET11	(Yang et al., 2006, Romer et al., 2010, Yuan et al., 2011)
Xoo	PthXo2	Rice	OsSWEET13	(Zhou et al., 2015)
Xoo	PthXo3/AvrXa7/TalC/Tal5	Rice	OsSWEET14	(Antony et al., 2010, Romer et al., 2010, Yu et

				al., 2011, Streubel et al., 2013)
Xpm	TAL20Xam668	Cassava	MeSWEET10a	(Cohn et al., 2014)
Xcm	AvrB6	Cotton	GhSWEET10	(Cox et al., 2017)

Xoo: *X. oryzae* pv. *oryzae, Xpm: X. phaseoli* pv. *manihotis, Xcm: X. citri* pv. *malvacearum*.

As *SWEETs* are sugar efflux transporters that pump sucrose and glucose into the apoplast, it is likely that their functions could also benefit mesophyllic pathogens. Hu and colleagues found the *CsSWEET1* gene (*Cs3g20720*), which contains the predicted PthA4 *EBE* and is activated upon *Xcc* infection. However, the dTALE targeting *CsSWEET1* did not rescue citrus canker phenotypes in the *XccΔpthA4* background (Hu et al., 2014). As the *EBE* prediction was performed on the differentially expressed genes found in a microarray setup, it is likely that many genes were not included due to the limited number of gene probes available on the Affymetrix GeneChip Citrus Array. In this case, RNA-seq provides less biased detection of transcripts with higher sensitivity for low abundance transcripts and higher specificity for transcript isoforms compared to microarray (Wang et al., 2014). Surprisingly, the RNA-seq data using dTALEs for off-target elimination showed that none of the predicted *CsSWEETs* were up-regulated by all TAL effectors. This indicates that 1) the *CsSWEET1* gene was an off target of PthA4, 2) none of CsSWEET members are CsLOB1-dependent and do not play a role in citrus canker development as is the case for rice bacterial blight and 3) LOB-mediated susceptibility by mesophyllic *Xcc* in citrus is distinct from *SWEETs*-mediated disease promotion by vascular *Xanthomonas* in rice, cassava and cotton.

3.3 CsLOB1 native function is a fruit softening regulator

To understand the disease-promoting mechanism regulated by CsLOB1, it is important to understand the native function of CsLOB1. Yet the native function of CsLOB1 in fruit tissue has not been uncovered as of yet. We have shown that *CsLOB1* is natively expressed during fruit softening (Section 2.4.1) and is exploited by *Xcc* as a tool to ectopically activate CWDEs and cell wall loosening proteins so that *Xcc* can get access to xylan as a carbon nutrient (Section 2.2.2 and 2.7). *CsLOB1* native expression profile which is high in mature fruits and

low in other organs suggest an insignificant role of CsLOB1 before fruit ripening. Indeed, the knock-out or knockdown of *CsLOB1* coding sequence did not show developmental changes before flowering compared to the wild type plant (Jia et al., 2017, Su et al., 2023, Zou et al., 2021), consistent with the low native expression of *CsLOB1* in tissues other than fruits (Figure 25). Similarly in tomato, when *SlLOB1*- the ortholog of CsLOB1, was silenced by RNAi, there were no obvious developmental changes before fruit setting (Shi et al., 2021), which is unlike other reported *LOB* homolog mutants (Semiarti et al., 2001, Berckmans et al., 2011, Feng et al., 2022). However, tomato fruits with *SlLOB1* knockdown showed an increase in firmness and a longer shelf life after harvest (Shi et al., 2021). In citrus, due to a long juvenile stage (before flowering), the knockout and knockdown plants did not have fruits by the time of the study for characterisation (Su et al., 2023, Zou et al., 2021). However, based on the phenotype of *SlLOB1* knockdown fruits, we predict that *CsLOB1* knockout/silenced lines may also exhibit increased fruit firmness. In contrast to *SlLOB1* knockdown fruits, *SlLOB1* overexpressing fruits show early fruit softening, consistent with higher expression levels of expansins, pectate lyases and other CWDEs (Shi et al., 2021). The fruits of *CsLOB1* overexpression lines were not developed yet by the time of Zou and colleagues' study (Zou et al., 2021), but we anticipate that they also undergo early fruit softening.

Studies by Duan et al. (2018) and Zou et al. (2021), however, showed that *CsLOB1* was highly expressed in mature leaves, compared to young leaves, stems, flowers and roots. Duan and colleagues quantified *CsLOB1* transcripts in these tissues via qPCR and used root transcript level as a normalisation control, which showed that leaf expression level was 30 times higher than that of the root (Duan et al., 2018). We observed that *CsLOB1* RPKM was very low or almost absent in citrus roots, leaves and stem, while it was extremely high in mature fruit (about 150 RPKM) (Figure 25). As qPCR is very sensitive and the level of *CsLOB1* transcripts on the control root sample is rather low, normalisation to root transcript may exaggerate the level of *CsLOB1* in other tissues, therefore the message might not be meaningful. On the other hand, Zou and colleagues attempted to reveal *CsLOB1* native expression by generating a GUS reporter line driven by the *CsLOB1* 497bp upstream region of the ATG. GUS staining showed that *CsLOB1* was expressed in leaf, root and stem (Zou et al., 2021). As GUS is also a highly sensitive reporter and considering that they used a rather short upstream region and that there might be a positional effect of the T-DNA insertion in the transgenic lines, the

reporter line might not reflect the native expression of *CsLOB1* as we have seen with the RNA-seq data. Taken together, both qPCR and GUS reporter lines to track *CsLOB1* native expression might not be precisely informative. In contrast, the fact that *CsLOB1* knockout/knockdown did not have an impact on the development of the plants support our findings that CsLOB1 expression is restricted in fruits, rather than leaf, root and stem as reported by Duan et al. (2018) and Zou et al. (2021)

3.4 Pectin, cellulose and xyloglucan degradation did not contribute to *Xcc in-planta* PthA4-dependent growth promotion

Using a combination of RNA-seq and ChIP-seq, we were able to uncover the direct target genes of CsLOB1 during citrus canker development (Figure 13, Supplemental Table 3). It is possible that not all developmental processes regulated by CsLOB1 are beneficial for *Xcc* during citrus canker. The increase in some simple sugars in the citrus apoplast upon *Xcc* infection led to our current hypothesis that PthA4 hijacks CsLOB1 to modify and degrade the cell wall, releasing sugars as a carbon source for *Xcc*. It remains to be determined which sugars play this role. Based on the enrichment of genes for several pectin-modifying enzymes such as pectin methyltransferase (*Cs3g24780*), pectate lyase (*Cs2g23970*, *Cs8g11330*), polygalacturonases (*Cs5g03170*), pectinesterase/pectinesterase inhibitor (*Cs9g14350*, *Cs9g14360*, *Cs5g09360*), we would assume that the products of pectin degradation would be more abundant in the apoplast and readily consumed by *Xcc*, resulting in its high growth promotion. In contrast, the *XccΔpthA4* mutant would not have access to such a carbon source and would therefore have reduced growth. However, even though the release of galacturonic acid was expected we have not observed any changes in the level of galacturonic acid, the monomeric unit of the pectin backbone (Figure 46).

Although host gene induction would give an indication of the possible nutrient sources for the pathogen, it would not provide any information on what is actually consumed by *Xcc* delivering PthA4. We therefore shifted our focus to bacterial metabolic adaptation in order to identify enzymes in *Xcc* that increase upon activation of *CsLOB1*. We envisioned that expressional changes in *Xcc* would provide an indication of which metabolites are consumed by *Xcc*. *Xanthomonas* contains about 160 genes for CWDEs, or CAZymes (Lombard et al., 2014) and are typically expressed as operons together with genes encoding for

polysaccharide binding and uptake in a polysaccharide utilisation locus (PUL). CAZymes are specific for the type of carbohydrate substrate and the glycosidic linkage, which means that the composition of CAZymes in the loci reflects the structure of the target polysaccharide (Grondin et al., 2017, Lapebie et al., 2019). Therefore, the presence of a PUL in the genome reflects the bacterial ability to digest the corresponding polysaccharide. For example, a study of the PULs of deep-sea Bacteroidetes from the Mariana Trench shows an enrichment of genes targeting hemicellulose and pectin, which are complex and difficult to degrade. This led to the hypothesis that they have adapted to the available carbon source at the bottom of the ocean, which happens to be the "leftover" from sinking cell wall debris, as simpler carbohydrates would be quickly depleted by bacterial communities in the upper layer of the ocean (Zhu et al., 2023b). In addition, bacteria rapidly adapt their metabolism to the carbon source available to them, which is reflected in the expression of genes for appropriate transporters and enzymes (Eisenreich et al., 2010). Therefore, we not only looked into the PULs of *Xanthomonas* genome, but also studied the transcriptomic changes in *Xcc* by analysing the carbohydrate degradation pathways that are up-regulated in *Xcc* during citrus canker development.

In the genus *Xanthomonas*, CAZymes are organised and co-regulated in multi-enzymatic systems and co-localised with TonB-dependent transporters (TBDTs), known as CUT systems (Giuseppe et al., 2023). These organisations are known to orchestrate the rapid and efficient degradation of the polysaccharide and allow the efficient vertical and horizontal transfer of the PULs to other bacteria, which has been extensively studied in mammalian gut Bacteroidetes (Terrapon et al., 2015). To date, CUT systems targeting xylan, xyloglucan, N-glycan and starch have been discovered in the *Xanthomonas* genome (Giuseppe et al., 2023, Bonfim et al., 2023). Interestingly, there is no known pectin CUT system in the *Xanthomonas* genus. *Xcc* strain 306 contains six pectate lyases, but they are located in different genomic regions and are not co-localised with other CAZymes. Whether they are co-regulated in pectin utilisation has not yet been characterised (Giuseppe et al., 2023). However, the fact that there is no known PUL targeting pectin suggests that the polysaccharide may not play as important a role in nutrient supply as other polysaccharides such as starch, xylan, and xyloglucan. In addition, genes related to pectin degradation and utilisation are not up

regulated in *Xcc* compared to the *XccΔpthA4* mutant, suggesting that pectin is unlikely to be a nutrient source responsible for the *Xcc* growth promotion induced by PthA4.

Although cellulose is the most abundant polysaccharide on Earth, it is extremely resistant to depolymerisation, mainly due to its crystalline state. The species that degrade cellulose and use it as a carbon source often contain genes for lytic polysaccharide monooxygenases (LPMOs), which are important for the degradation of crystalline cellulose microfibrils (Vaaje-Kolstad et al., 2010). LPMOs are significantly enriched in the bacterial and fungal kingdoms, but no LPMO has been found in any plant genome as of yet. They can access the crystalline part and efficiently cleave cellulose via oxidative reactions (Vaaje-Kolstad et al., 2010, Horn et al., 2012, Cragg et al., 2015). The oxidised cellulose oligosaccharides derived from LPMO-assisted cellulose degradation were able to induce a defence response in Arabidopsis (Zarattini et al., 2021). Although the *Xcc* genome contains several genes encoding cellulases, it does not contain any LPMOs (Giuseppe et al., 2023). In our *Xcc* RNA-seq data, we observed an up-regulation of a glycosyl hydrolase 5 (GH5) - a degenerated cellulase (*XAC0346*) in the *Xcc* wild-type, but not in the *XccΔpthA4* mutant *in-planta*. However, Xia and colleagues did not observe extracellular endoglucanase activity of this GH5 cellulase in an *E. coli* enzymatic assay (Xia et al., 2016); hence, whether it has a role in cellulose degradation is questionable. In addition, cellobiose and oligomers as a product of cellulose degradation were shown to elicit host recognition and defence responses (Souza et al., 2017). Disruption of the cellulose degradation machinery in the fungal pathogen *Fusarium oxysporum*, while severely affecting the saprophytic lifestyle, increased virulence and delayed the host response during initial colonisation. In contrast, when external cellobiose, a product of cellulose degradation, was applied, host recognition for the mutant was restored to levels similar as the wild-type strain (Gamez-Arjona et al., 2022). This reinforces the reason why most biotrophic pathogens, including *Xanthomonas*, do not possess a strong cellulose degradation system (Giuseppe et al., 2023). More importantly, this indicates that cellulose is not used as a carbon source for PthA4-dependent *Xcc* growth in citrus.

In contrast to pectin and cellulose degradation machinery, the genus *Xanthomonas* contains highly conserved PULs that target xyloglucans (XyGs) in their genomes, whether they are mesophyllic or vascular (Vieira et al., 2021). However, the deletion of XyGs transporters and an endo-β-1,4-xyloglucanase did not alter the *in-planta* virulence of *Xcc* (Vieira et al., 2021),

suggesting that XyGs are not the only carbon source for *Xcc* during infection or that XyGs are not required. CsLOB1 directly activates two alpha-L-fucosidases (*Cs2g27090* and *Cs2g27100* (Figure 13, Supplemental Table 3)), which participate in xyloglucan degradation by removing the fucosyl residue from xyloglucan and facilitate cell wall expansion and cell elongation (Gunl et al., 2011). CsLOB1 also indirectly targets xyloglucan endotransglucosylase/hydrolase (*Cs2g14920*, *Cs8g12020* and *orange1.1t02385*), which remodel the arrangement of xyloglucan and promote cell wall loosening (Ishida and Yokoyama, 2022). Nevertheless, the products of CsLOB1-mediate xyloglucan targeting enzymes seem not to be the nutrient source for *Xcc,* as the xyloglucan CUT system was not upregulated in PthA4-containing *Xcc* (Section 2.7). Only one of the two TBDTs, *XAC1768*, was up-regulated in the presence of PthA4, but this transporter was shown not to have high activity in up-taking xyloglucan oligosaccharides (Vieira et al., 2021), so the upregulation is not necessarily correlated with xyloglucan oligosaccharide presence. Although *Xcc* possesses CUT system for xyloglucan degradation, we therefore negate that xyloglucan can act as a carbon source for *Xcc* growth in a PthA4-dependent fashion.

3.5 Host cell wall loosening enables xylan degradation

Xylan is the second most abundant cell wall polysaccharide after cellulose and is targeted as a carbon source by many microorganisms, so-called xylanolytic microbes (Scheller and Ulvskov, 2010, Saha, 2003, Dodd et al., 2010). The xylan CUT system is present in *Xanthomonas* and has been extensively studied and reviewed (Santos et al., 2014, Chow et al., 2015, Giuseppe et al., 2023). Our *Xcc* RNA-seq was the first to show that *Xcc* specifically upregulates xylan CUT system upon citrus canker promotion by PthA4 effector, indicating that *Xcc* uses xylan but not cellulose, pectin, or other hemicelluloses as a carbon source during citrus canker promotion (Section 2.7 and 3.4). Furthermore, the mixture of monosaccharides: xylose, galactose, and glucose, as known products of xylan degradation, was able to activate the T3SS (Vieira et al., 2021), indicating that xylan acts as both a nutrient source and virulence elicitor. The xylan CUT system is also expressed in *XccΔpthA4 in-planta* and in both NYG-growing *Xcc* and *XccΔpthA4* but at lower non-differential levels, indicating that the up-regulation of the xylan CUT system is specific to the PthA4-dependent host-pathogen interaction.

Interestingly, the supply of xylan polymers supplied in a synthetic medium did not promote *Xcc in-vitro* growth, but the oligoxylosides derived from xylan depolymerisation demonstrated the ability to foster *Xcc* growth (Chow et al., 2015). This indicates a role of the host enzymatic activities in the proper xylan utilisation by *Xcc*. During citrus canker, host cell wall loosening proteins and CWDEs were activated by CsLOB1 (Figure 13, Supplemental Table 3). We hypothesise that this will facilitate the expansion of cell wall, allowing bacterial CWDEs access and the ability to digest the polysaccharides of interest. We observed that four expansins (*orange1.1t00187*, *orange1.1t00855*, *Cs7g32410* and *Cs8g18640*) were found to be directly activated by CsLOB1 (Figure 13), and three of them were detected in the whole tissue proteomics study (Figure 14). Expansin breaks the hydrogen bonds between cellulose microfibrils and hemicelluloses, leading to cell wall expansion (Wang et al., 2013). In biofuel production from plant cell wall mixture, the combination use of expansin with cellulases clearly increases cell wall breakdown by creating more enzyme-accessible regions in the cell wall complex (Georgelis et al., 2015, Seki et al., 2015). It is conceivable that cell wall loosening also facilitates the access of xylan degrading enzymes. The enzyme groups xyloglucan endotransglucosylase/hydrolase proteins (Cs2g14920, Cs8g12020 and orange1.1t02385) and mannan endohydrolase (Cs1g24390 and Cs3g25100) that are also direct targets of CsLOB1 (Figure 13) are also known to promote cell wall loosening (Ishida and Yokoyama, 2022, Schröder et al., 2009) and increases access of other CWDEs. Similarly, CWDEs targeting pectin increase cell wall porosity, promotes cell wall swelling, and allows for the access of other CWDEs to active sites for degradation (Redgwell et al., 1997, Karr and Albersheim, 1970, Bauer et al., 1977). CsLOB1 directly activates pectate lyases, pectinesterases and other related pectin-degrading enzymes (Figure 13), increasing cell porosity, swelling, and exposing xylan to degradation. The pectin layer provides a significant mechanical barrier to the penetration of plant pathogens and are preferably targeted by many pathogens (Chen et al., 2021). Citrus, particularly, has a high pectin content and has been used as a commercial source of pectin (Chandel et al., 2022). Although the *Xcc* genome contains several genes for pectin degradation (Giuseppe et al., 2023), their role in citrus canker appears to be PthA4-independent and was discussed in detail in Section 3.4. This indicates that their own pectin degrading enzymes are not sufficient for their virulence in the absence of PthA4. Therefore, revealing the targets of CsLOB1 helps us understand why the activation of *CsLOB1* by PthA4 is essential for bacterial virulence and growth. CsLOB1

induces a large set of cell wall loosening and CWDEs that target not only pectin but also hemicelluloses, breaking down the linkage between cell wall polymers, which, as a result, has a profound benefit as it not only allows better bacterial colonisation by weaking pectin layers, but also increases the access of bacterial CWDEs to host cell wall polysaccharides, most importantly xylan, to digest as a source of nutrition and virulence elicitor.

3.6 Why *Xcc* target a transcription factor?

Plant pathogens are equipped with a wide range of effectors to overcome host immunity and gain advantage. The mechanisms of host immunity suppression have been well studied and reviewed (Deslandes and Rivas, 2012, Wang et al., 2022, Kong and Yang, 2023). Less attention has been paid to effectors that manipulate host metabolism for nutrient mobilisation, though this would logically be the ultimate aim of the pathogen in order to thrive in the host. Some examples of host transcriptional reprogramming have been shown, including the study of *Ustilago maydis*, a fungal pathogen that causes common smut in maize, which promotes hyperplasia in tumour cells by interacting with a host transcriptional activator (Zuo et al., 2023); the Aster Yellows phytoplasma, an obligate intracellular bacterium that causes Witches' Broom in Arabidopsis, which uses various effectors to degrade host TFs involved in flower development to induce proliferation of leaves and sterile shoots as a food source for bacterial growth and reproduction (MacLean et al., 2011, Sugio et al., 2011, Correa Marrero et al., 2023). Although there is correlation between host developmental manipulation and pathogen virulence and growth, the exact mechanism of how the pathogens directly benefit from these changes is not fully understood.

Recently, manipulation of host metabolites has attracted more attention and potentially helps understand further how the pathogens benefit from the host. A sugar transporter protein STP13, involved in the sugar-mediated defence response in wheat, increases the uptake of monosaccharides from the apoplast into leaf cells to deprive pathogens of a nutrient source (Yamada et al., 2016). This defence response is hijacked by the obligate biotrophic wheat stripe rust *Puccinia striiformis f.* sp. *tritici* (*Pst*), which feeds on living cells instead of the apoplast. This is thought to increase the uptake of sugars from the apoplast into the living cells for *Pst* sugar acquisition (Huai et al., 2020). More recently, a study of an oomycete pathogen called *Phytophthora sojae* showed that a soybean trehalose-6-

phosphate synthase 6 was hijacked by the pathogen effector PsAvh413 to promote trehalose production, which was then directly acquired by the pathogen as a carbon source during soybean root and stem rot (Zhu et al., 2023a).

The manipulation of host metabolites by bacterial plant pathogens is rarely reported. *Ralstonia solanacearum* was shown to use the effector RipI to enhance the production of gamma-aminobutyric acid for use as a nutrient source (Xian et al., 2020). A well-studied target of TALEs is the SWEET gene, whose homologues in rice, cassava and cotton have been shown to be convergently targeted by different TALEs to promote disease (Chen et al., 2010, Cohn et al., 2014, Cox et al., 2017), although the direct usage of elevated sucrose remains to be verified. Interestingly, none of the Citrus *SWEET* genes have been shown to play a role in citrus canker development following *Xcc* infection (Section 3.2) (Hu et al., 2014), suggesting a completely different strategy adopted by *Xcc* using TALEs to activate a TF rather than a sugar transporter. Our study shows that the susceptibility gene *CsLOB1* TF regulates a number of CWDEs and cell wall loosening proteins that allow bacterial CWDEs to access and degrade xylan in the cell wall complex for nutrition (Section 2.2.2 and 2.7). Interestingly, *Xanthomonas* strains targeting pepper and tomato use the TALEs AvrBs3 and AvrHah1 to target *UPA20* and *bHLH3*/*bHLH6*, respectively, which belong to a different TF family, but also activate CWDEs (Schwartz et al., 2017, Schenstnyi et al., 2022), potentially following the same strategy as *Xcc* to acquire nutrients. We argue that there are several advantages for the pathogen to target transcription factors that regulate cell wall degradation.

The cell wall is a complex structure with pectin and hemicelluloses cross-linked to the crystalline cellulose microfibrils (Ha et al., 1998). The enzymatic reactions of cell wall degradation often involve a cascade of different enzymes to degrade the complex polymers with different side-chain components (Daher and Braybrook, 2015, Hayashi and Kaida, 2011). Targeting a single step in an enzymatic cascade may not necessarily lead to an increase in the final products, as other enzymatic steps may be limiting factors. For example, overexpression of a phytoene synthase in carotenoid synthesis pathway from *Erwinia uredovora* in tomato, increased its activity from 6- to 14-fold but resulted in a disproportionate increase in carotenoid content of about 2- to 4-fold, suggesting that there were other limiting factors (Fraser et al., 2002). Similarly, to produce high levels of beta-carotene in golden rice, it was necessary to overexpress not only phytoene synthase but also

three other enzymes that catalyse precursors for the phytoene synthase step (Ye et al., 2000). In this case, targeting the transcription factor that regulates the pathway would overcome the limitations of targeting individual pathway genes. Our study has shown that CsLOB1 directly activates 35 genes involved in cell wall loosening, modification, and degradation (Figure 13), many of which may need to be expressed together to sufficiently degrade a particular polysaccharide.

In addition, a review by Broun (2004) suggested that the limitation may be due to tissue-specific expression of other relevant enzymes in the pathway. When chalcone isomerase, an enzyme involved in the earlier steps of the flavonoid biosynthetic pathway, was constitutively expressed in tomato fruit, flavonoid production was increased only in the peel (Muir et al., 2001). The restriction of metabolite accumulation to one tissue could be explained by the lack of activation of other necessary genes in other tissues, thus blocking the biosynthesis of the metabolite in these tissues (Broun, 2004). In contrast, when two anthocyanin regulators were expressed under a fruit-specific promoter in tomato, flavonol levels were increased throughout the fruit tissue (Bovy et al., 2002). This suggests that activation of a transcription factor would have a more global effect, which is necessary for bacterial virulence. Indeed, *Xcc* is able to promote citrus canker in various tissues including leaves, stems, thorns, and fruits (Brunings and Gabriel, 2003) which could be explained by the fact that it targets a TF.

Another hypothesis is that CsLOB1 enables the activation of genes with epigenetic marks. A transcription factor can often cooperate with methyl-CpG-binding domain-containing proteins or bind directly to methylated regions (Zhu et al., 2016). In contrast, the binding of TALEs to DNA is affected by DNA methylation (Zhang et al., 2017b, Deng et al., 2012), which will restrict the activation of target genes to specific tissues with appropriate epigenetic marks, thereby reducing virulence globally. In addition, CsLOB1 regulates cell wall loosening during fruit ripening, a process known to involve epigenetic changes (Huang et al., 2019). In contrast to tomato and strawberry, where DNA methylation is reduced during fruit ripening (Zhong et al., 2013, Cheng et al., 2018), global DNA methylation in citrus is increased during fruit ripening, and inhibition of DNA methylation resulted in delayed fruit ripening (Huang et al., 2019). Interestingly, the data showed that this hypermethylation correlated with the upregulation of several hundred genes involved in citrus fruit ripening (Huang et al., 2019),

suggesting that these genes require hypermethylation to be activated, but the mechanism requires further investigation. Interestingly, at least six direct target genes of CsLOB1 were shown to be hypermethylated and upregulated during fruit ripening in the work of Huang and colleagues. This suggests that the activation of *CsLOB1* may potentially overcome the hurdle for *Xcc* of activating genes with epigenetic marks.

In short, we discussed that the promotion of loosening of cell wall requires many enzymes, and it is therefore conceivable why *Xcc* targets a fruit softening regulator, in order to overcome 1) the need of activating many enzymes at the same time 2) across different tissues with 3) possible different epigenetic marks.

3.7 Does *Xanthomonas* strains infecting pepper/tomato share the same strategy as *Xcc*?

We observed growth promotion of *Xe* in *N. benthamiana* when co-infiltrated with *Agrobacterium* delivering p35S-*CsLOB1* or p35S-*SlLOB1*, suggesting that both *Xe* and *Xcc* may derive similar benefits from cellular changes upon CsLOB1/SlLOB1 overexpression. *Xe, X. perforans,* and *X. gardneri* are known to infect pepper and tomato and use the TALE AvrBs3/AvrHah1 to activate host *bHLH* transcription factors that convergently induce tomato pectate lyase and pectinesterase (Schwartz et al., 2017, Schenstnyi et al., 2022). In addition, CsLOB1 and SlLOB1 could also activate these tomato pectate lyase and pectinesterase in a transactivation assay, suggesting a common target for both groups of effectors, although they activate two different transcription factor families. Activating pectate lyase and pectinesterase could help loosen the host cell wall via breaking pectin junctions and allow CWDEs access to cell wall polysaccharides such as xylan. It is therefore conceivable that the pepper/tomato strains also target xylan as a carbon source. Knockouts of genes in *Xe* xylan CUT system could be used to confirm the xylan use. Alternatively, promoters of *Xcc* genes that are specific for the degradation of certain polysaccharides can be used to drive a reporter gene such as LUX (Jutras et al., 2021) to monitor the expression of the genes when they digest such polysaccharides *in-planta*. An example of a xylan-degradation-specific gene is *XAC4254*, which encodes an endo-β-1,4-xylanase that initiates xylan depolymerisation (Giuseppe et al., 2023). These reporters can also be transferred to other *Xanthomonas* species to monitor the metabolic usage of the bacteria in the presence or absence of the TAL

effector of interest. Furthermore, a high throughput transcription analysis of *Xe* strains containing or lacking AvrBs3/AvrHah1 when they infect plant host will reveal a more complete picture of their carbon usage and whether AvrBs3/AvrHah1 has a role in nutrient acquisition as PthA4 does.

Despite activating pectin-degrading enzymes, AvrBs3 and AvrHah1 did not affect bacterial *in-planta* growth, in contrast to PthA4. The lack of *in-planta* growth promotion in tomato and pepper could suggest that there are other effectors that loosen the host cell wall and compensate for the loss of AvrBs3/AvrHah1. However, this does not seem plausible as AvrBs3 and AvrHah1 are strictly maintained in the *Xanthomonas* population (Potnis et al., 2011, Schwartz et al., 2015, Subedi et al., 2023), implying that they must play an essential role in the *Xanthomonas* life cycle. Interestingly, *CsLOB1*/*SlLOB1* could activate pectate lyase and pectinesterase, which are the downstream genes of *bHLHs*, but the latter could not significantly activate citrus endoglucanase and tomato expansin, which are the direct targets of *CsLOB1* and *SlLOB1*, respectively (Figure 55). This suggests a certain difference in the downstream mechanism regulated by these two families of TFs, which in turn may indicate the different strategies exploited by *Xanthomonas* species containing PthA4 versus those containing AvrBs3/AvrHah1.

To compare these two strategies in detail, we use tomato as a working model because of the presence of both transcription factor members: *bHLHs* as the target of AvrBs3/AvrHah1 and *SlLOB1* as the ortholog of *CsLOB1*-target of PthA4. We used a tomato line in which the NLR *Bs4* gene is knocked out, allowing the delivery of TALEs containing 33 to 34 amino acid repeats such as AvrBs3. Delivery of AvrBs3 via *Xe* into this *bs4* mutant was shown to activate *bHLH3* and *bHLH6* which then induce water soaking (Schenstnyi et al., 2022). We will use a similar approach to dissect the direct targets of bHLH3, bHLH6, and SlLOB1 by RNA-seq and ChIP-seq as we did with CsLOB1. dTALEs designed to specifically target *SlLOB1* were shown to be functional and could be used to uncover genes upregulated by SlLOB1 via RNA-seq (Figure 41), in future studies we will use a dTALE to target the AvrBs3 *EBE* in the promoters of bHLH3 and bHLH6 and study the downstream genes of them via RNA-seq. In addition, we plan to generate tomato transgenic lines in the *bs4* mutant background containing *bHLH3*, *bHLH6,* or *SlLOB1* coding sequences translationally fused to a C-terminal FLAG to perform ChIP-seq. All three constructs are driven by their own 3kb native promoters rather than

constitutive or chemically inducible promoters, allowing the use of dTALEs delivered by *Xanthomonas* to activate the genes; therefore, ChIP-seq will also be performed in an infection context. As both TF families will be studied in the same plant host system via *Xanthomonas* infection, we expect that the comparison between the two mechanisms will be more direct and meaningful. In combination with *Xe* transcriptomic analysis, the described RNA-seq and ChIP-seq approach will help reveal whether *Xanthomonas* strains infecting pepper/tomato using AvrBs3/AvrHah1 share the same strategy as *Xcc* using PthA4 and explain why one TALE group does not promote bacterial growth *in planta* and the other does.

3.8 AvrBs3/AvrHah1 could be important for bacterial entry and dispersal

Both AvrBs3 and AvrHah1 induce bHLH3 and bHLH6 in tomato and promote water soaking (Schenstnyi et al., 2022). Water soaking is known to be a key virulence mechanism of many plant pathogens, such as *Xcc* using PthA4 to promote water soaking and bacterial growth (Hu et al., 2014), *Pseudomonas*, which infects Arabidopsis using the T3S effector AvrE/HopM1 to promote an aqueous environment and enhance bacterial growth (Hu et al., 2022, Xin et al., 2016). Similarly, $TAL20_{xam668}$ and $TaL14_{xam668}$ from X. *axonopodis* pv. *manihotis*, which infects cassava, a dicot crop, also induce water soaking via activation of the sugar transporter *MeSWEET10a*, and deletion of individual or both TALEs reduces water soaking and bacterial growth (Cohn et al., 2014).

In contrast, the absence of AvrBs3/AvrHah1 abolished the water-soaking phenotype but had no effect on bacterial growth (Bonas et al., 1989, Schwartz et al., 2017). Instead, AvrHah1-dependent water soaking was shown to facilitate the entry and survival of *Salmonella enterica* (Potnis et al., 2015, Cowles et al., 2022), while AvrBs3 could play a role in transmission and the ability to survive outside of the host in the field (Wichmann and Bergelson, 2004). We hypothesise that the activation of pectate lyase by AvrBs3/AvrHah1 (Schwartz et al., 2017, Schenstnyi et al., 2022) could contribute to xanthomonad entry and/or dispersal. The study of pectin-degrading enzymes in *Xanthomonas* infecting dicots suggested that pectin degradation may be particularly important in the early stages of infection. As shown in *X. campestris* pv. *campestris*, single or double knockouts of two polygalacturonases induced weak wilt symptoms, but the complemented strains restored

symptoms similar to the wild-type strain in Arabidopsis ecotype Kendle (Wang et al., 2008). Importantly, Wang and colleagues observed that the polygalacturonases were important for infection by spray and dipping methods, whereas the mutants and wild type did not differ significantly in virulence by infiltration inoculation (Wang et al., 2008). They suggested that these pectin-degrading enzymes play an important role in the early stage of infection to overcome the host pectin barrier. On the other hand, pectin degradation was not significantly essential for *Xoo* during the infection of rice, a monocot with a lower pectin content. The mutation of polygalacturonase (*pglA*), pectin methyl esterase (*pmt*), and two pectate lyases (*pel* and *pelL*) did not significantly affect the virulence of *Xoo* (Tayi et al., 2016). A plausible hypothesis is that *Xoo* and *X. campestris* infect two different hosts with different cell wall pectin compositions, with the lower pectin content in rice suggesting a lower requirement for pectin degradation during the early stage of infection (Giuseppe et al., 2023). This is also consistent with a higher number of genes encoding pectate lyase and polygalacturonase (*GH28*) in *Xanthomonas* strains infecting dicots compared to those infecting monocots, whether vascular or mesophyllic (Potnis et al., 2011, Giuseppe et al., 2023). This suggests that pectin degradation is required for successful colonisation of dicot hosts.

Therefore, our hypothesis for the function of AvrBs3/AvrHah1 focuses towards the correlation between pectin degradation through induction of *PE* and *PL* and bacterial colonisation. The strains containing the effectors or dTALEs that activate *bHLHs* and *PL* could promote water uptake that allows other bacteria to colonise the apoplast (Schwartz et al., 2017). Water soaking was shown to be positively correlated with the survival and multiplication of *S. enterica* in tomato leaves (Potnis et al., 2015). Cowles and colleagues observed a similar benefit of *S. enterica* in the presence of water-soaking-promoting *Xanthomonas* strains and stated that this benefit is AvrHah1-dependent, although there was no evidence for AvrHah1's target gene activation. However, their conclusion can be supported by the fact that the presence of AvrHah1 in *X. gardneri* is always linked with *bHLH3/bHLH6* activation and water soaking promotion (Schwartz et al., 2017, Schenstnyi et al., 2022), suggesting that AvrHah1 has an advantage for a commensal species to enter and survive *in planta* (Cowles et al., 2022). Experimental evidence of AvrBs3/AvrHah1-mediated bacterial entry has not been reported, most likely because infection assays have always been

performed using the infiltration method with high bacterial inoculum, which may not reflect natural infection where bacteria must attach to the leaf surface and move through natural openings such as stomata, hydathode, or wounds. Our hypothesis is that after successful infection of a few *Xanthomonas* through natural openings, AvrBs3/AvrHah1 are injected into the plant cells, activating the orthologous *bHLH3* and *bHLH6* TFs, which directly or indirectly activate cell wall degrading enzymes, including *PL*. This leads to pectin degradation, cell wall loosening and further degradation, increased water potential, allowing water from rain or irrigation to be absorbed, increasing the chance that more bacteria can move from the leaf surface into the apoplastic region.

To test this hypothesis, we first look at an efficient infection method that mimics the native infection scenario. As syringe infiltration bypasses the stomatal immunity barrier, spraying and leaf dipping have been used as the closest to natural infection, where dipping has been shown to be more robust (Tornero and Dangl, 2001). Obviously, the development of symptoms is often slower and less homogeneous compared to syringe infiltration. In addition, as the number of host cells in contact with bacteria and receiving effectors is lower, we predicted that the observable water-soaking phenotype would be more difficult to identify. Therefore, our current goal is to generate a sensitive reporter construct that gives a signal upon AvrBs3/AvrHah1 injection. We have constructed the transgenic tomato lines mentioned in section 3.5 containing 3kb native promoters of *bHLH3* and *bHLH6* driving the TF coding sequence translationally tagged with a C-terminal FLAG. The FLAG tag will be fused to a GUS gene separated by a translational self-cleaving peptide called T2A, which minimises the effect of the reporter protein on the functionality of the TF. As GUS staining is very sensitive, a low level of gene induction may still be visible upon delivery of AvrBs3/AvrHah1. This will not only show whether AvrBs3/AvrHah1 has an effect on bacterial entry into the host, but also reveal the spatial information of *Xanthomonas* infection.

On the other hand, it cannot be excluded that AvrBs3/AvrHah1 may play a role in bacterial dispersal. Many *Xanthomonas* strains are known to disperse by water splashing and air turbulence (Ryan et al., 2011, Bock et al., 2012). A known TAL effector that has been shown to play an important role in *Xanthomonas* release is Avrb6 from the strain *Xcm*, which causes bacterial blight in cotton, a dicot crop. Avrb6 was responsible for the induction of water soaking and increased the release of bacteria to the leaf surface 240-fold compared to the

strain lacking Avrb6 but had no effect on *in-planta* growth (Yang et al., 1994). Although activation of cotton *SWEET10* by Avrb6 alone was sufficient to restore the water-soaking phenotype, it was not known whether this activation alone restored bacterial release. Interestingly, a putative *bHLH* (*Gorai.009G327300.1*) containing a predicted Avrb6 *EBE*, a number of LOB-containing proteins, expansins, and pectin lyases were found to be up-regulated in the presence of Avrb6, whose role in bacterial release deserves further investigation (Cox et al., 2017). Surprisingly, although the targets of AvrBs3 and AvrHah1 have been discovered and studied for a long time, no explicit experiments have been performed to study the release of bacteria to the leaf surface in the presence of AvrBs3/AvrHah1. We proposed a similar experiment to that done by Yang and colleagues, using leaf surface washing fluid to recover *Xanthomonas* released from the inner leaf to the surface.

In conclusion, we have uncovered the mechanism underlining citrus canker promoted by *Xcc* using PthA4 to activate the fruit softening transcription factor CsLOB1 that loosens host cell wall and enables the bacteria to gain access to xylan as the nutrient source. This is the first step to understanding of why bacteria manipulate host developmental pathways via targeting TFs. These findings also establish a foundation for investigating the elusive role of the long-researched TALE group AvrBs3/AvrHah1 that also target TFs regulating cell wall loosening.

4. MATERIALS AND METHODS

4.1 Materials

4.1.1 Plant materials and growing conditions

Grapefruit cv. Duncan (*Citrus* x *paradisi*) seeds were kindly provided by Dr. Concetta Licciardello (CREA, Research Centre for Olive, Fruit and Citrus Crops, Acireale, Italy). Seeds were germinated and grown *in vitro* for 5-6 weeks before transfer to soil. Seedlings were grown at 28 ±1°C and 60-70% air humidity (16 h: 8 h, light: dark). Infection assays were performed on young and fully expanded leaves on plants aged approx. 1 year or older.

Tomato cv. Money Maker *bs4* knockout from the work of Schenstnyi et al. (2022) and *N. benthamiana* WT, *roq1* and *eds1* (Gantner et al., 2019) mutants are grown at 20 ±2°C and 30-50% air humidity (16 h: 8 h, light: dark).

4.1.2 Bacterial materials

Table 2: Bacterial strains

Name	Strain	Genotype	Reference
E. coli	Top10	*F*-mcrA *Δ(* mrr-hsd*RMS*-mcr*BC) Φ80*LacZ*ΔM15 Δ* Lac*X74* rec*A1* ara*D139 Δ(* araleu*) 7697* gal*U* gal*K* rps*L (StrR)* end*A1* nup*G*	Invitrogen™
E. coli	ccd*B Survival*™	F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R) *end*A1 *nup*G *fhu* A::*IS2*	Invitrogen™
E. coli	BL21-AI^{TM}	F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm araB::T7RNAP-tetA*	Invitrogen™
A. tumefaciens	GV3101	*C58 (rif R) Ti pMP90 (pTiC58DT-DNA) (gentR) Nopaline*	Holsters et al. (1980)
A. tumefaciens	EHA101	*C58 (rif R) Ti pEHA101 (pTiBo542 D T-DNA) (kan R) Nopaline.*	Hood et al. (1986)
Xcc	306	WT and *ΔpthA4*	Hu et al. (2014)
Xcc	306	*ΔpthA4::dTALE2* (targeting CsLOB1)	This study
Xcc	306	*ΔpthA4::dTALE4* (targeting CsLOB1)	This study
Xcc	306	*ΔpthA4::dTALE CsLOB2*	This study
Xcc	306	*ΔpthA4::dTALE CsLOB3*	This study
Xe	85-10	TALE-less strain	Minsavage et al. (1990)
Xe	85-10	dTALE807 (targeting SlLOB1)	This study
Xe	85-10	dTALE808 (targeting SlLOB1)	This study
Xe	85-10	dTALE809 (targeting SlLOB1)	This study

Table 3: Media compositions

Name	Bacteria species	Content
Luria-Bertani (LB)	*E. coli*	10g/l tryptone, 10g/l NaCl, 5g/l yeast extract
Yeast-Extract-	*A.*	5g/l beef extract, 5g/l peptone, 5g/l sucrose, 1g/l yeast

Beef (YEB)	*tumefaciens*	extract, 2mM MgSO4
Nutrient-Yeast-Glycerol (NYG)	*Xanthomonas*	5g/l peptone, 3g/l yeast extract, 0.2% (v/v) glycerol
Minimal media MME	*Xanthomonas*	20mM Sodium-L-glutamate, 0.15g/l Casamino acids, 7.57mM $(NH_4)_2SO_4$, 1mM $MgSO_4.7H_2O$, 60.34mM Na_2HPO_4, 33.07mM NaH_2PO_4
1.5% (w/v) agar added for solidified media		

Table 4: Antibiotics

Name	**Working concentration**
Ampicillin	100mg/l (in plate), 50mg/l (in liquid)
Carbenicillin	50mg/l (*E. coli*), 100mg/l (*A. tumefaciens*)
Chloramphenicol	15mg/l
Gentamycin	15mg/l
Kanamycin	25mg/l (*E. coli*); 100mg/l (*A. tumefaciens*)
Rifampicin	100mg/l
Spectinomycin	100mg/l
Tetracycline	10mg/l

4.2 Methods

4.2.1 Cloning work

4.2.1.1 Cloning

All constructs were cloned using Golden Gate tool kit adapted from Binder et al. (2014). Briefly, inserts were amplified with primers containing additional BsaI sites and appropriate overhangs (Table). *In-silico* cloning was done in Qiagen CLC Main Workbench. General PCR mix components are in Table 5 and the PCR reaction followed the setting in Table 6. The inserts were subcloned to a high copy vector via blunt-end restriction to generate a Level I module. Different components of gene structure (promoter, coding sequence, terminator, tags) were assembled into Level II module via multiple cut-ligation steps using BsaI and T4 DNA ligase. All constructs were confirmed by Sanger sequencing and analysed in Qiagen CLC Main Workbench.

Table 5: General PCR mix for insert amplification

Component	**Volume (µl)**
Phusion polymerase	0.5
2mM dNTPs	2
5x HF buffer (NEB)	4
5x PRECES I*	4

10µM forward primer	0.5
10µM reverse primer	0.5
Diluted DNA template	1
Double distilled H_2O	to make to 20µl
*: 5X concentration consists of 4M betaine, 16mM DTT, 16% (v/v) DMSO and 83µg/ml bovine serum albumin (known as BSA) (Ralser et al., 2006)	

Table 6: PCR setting

Temperature	Time	Number of cycles
98˚ C	3min	1
98˚ C	20s	35
56˚ C*	20s	35
72˚ C**	1min/kb	35
72˚ C**	3-10min	1
*: temperature depends on the primer annealing temperature **: 68˚ C was used when PrimeSTAR® GXL DNA polymerase was used instead of Phusion kb: 1000-bp All runs were done in SensoQuest Labcycler.		

4.2.1.2 Bacterial transformation

100-150ng of DNA was added to 25-50µl *E. coli* chemical competent cells and left on ice for 15-30min. The DNA-cell mixture went through heat shock at 42˚ C for 45s and immediately went back on ice for 1min. LB media was added to make up the volume to 300µl and the bacterial suspension was incubated at 37˚ C with shaking at 300rpm (rounds per minute) for 1h. 100-150µl of bacterial suspension was plated on LB agar plate with appropriate antibiotics (Table 3, Table 4) and incubate in a 37˚ C incubator for 1 day.

For *E. coli* BL21-AI™, *A. tumefaciens* and *Xanthomonas*, 100-150ng of DNA was added to 50µl of electro-competent cells and kept on ice for 15min. The DNA-cell mixture was transferred to a 2mm-gapped sterile electroporation cuvette and applied with an electrical pulse at 2.5kV for 6ms using the MicroPulser from Bio-Rad. 250µl of 2xYEB/NYG was added and the cell suspension was incubated at 28˚ C, 300rpm for 1h and plated on appropriate agar media with antibiotics and grown in a 28˚ C incubator for 2 days.

4.2.1.3 Plasmid extraction

Cells were inoculated in 5-6ml of LB media with appropriate antibiotics and let grown overnight at 37˚ C with shaking at 200rpm. Bacterial cells were precipitated from overnight culture by centrifugation at maximum speed for 1min. Multiple rounds of precipitation were used for low-copy plasmids. The plasmid of interest was extracted using the NucleoSpin Plasmid Mini kit from Macherey-Nagel following the manufacturer's protocol. DNA was eluted in double-distilled water and the quantity and purity was assessed by NanoDrop® ND-1000 UV-Vis Spectrophotometer from ThermoFisher Scientific.

4.2.2 RNA work

4.2.2.1 Plant RNA extraction

4-6 leaf discs were punched out using the 8mm-diameter cork borer and immediately flash-frozen by liquid nitrogen in a 2mm Eppendorf tube containing one 6.2-6.8mm SiLibeads® ceramic bead from SiLi. The tissue was homogenised into powder using Qiagen RetschMill at 22Hz for 1min in an adaptor previously kept at -20˚ C. Citrus RNA was extracted using the Qiagen RNeasy Plant Mini kit following the manufacturer's protocol with some adaptation. 450µl of buffer RLT supplemented with 40mM of DTT (dithiothreitol) was added to the ground tissue and immediately homogenised with the RetschMill at room temperature (RT) for 30s at 20Hz. The lysate was centrifuged for 30s at max speed to precipitate large debris. The supernatant was transferred to a lilac QIAshredder spin column and centrifuged for 2min at full speed for thorough sample homogenisation. The supernatant was transferred to a clean 1.5ml Eppendorf tube and mixed with 0.5 volume of 96-100% ethanol by pipetting. The mixture was transferred to a pink RNeasy Mini spin column and centrifuged for 15s at 13000rpm (>15000 x g) and the flow-through was discarded. The membrane was then washed one time with 700µl of the washing buffer RW1 and two times with 500µl of buffer RPE, each for 15s at 13000rpm (>15000 x g). After the last washing step, the column was placed on a fresh 2ml collection tube and centrifuged for 2min at max speed to eliminate residuals of ethanol. 30-50µl of RNase-free water was added directly to the membrane, incubated for 3min at RT. RNA was eluted into a clean 1.5ml Eppendorf by centrifugation for 1min at max speed. The quantity and purity of RNA was assessed by NanoDrop® ND-1000

UV-Vis Spectrophotometer from ThermoFisher Scientific. RNA was proceeded with DNase treatment or stored at -80˚ C.

RNA from Tomato leaf tissue was extracted using the GeneMATRIX Universal RNA Purification Kit from EURx® from Roboklon following the manufacturer's protocol.

4.2.2.2 Bacterial RNA extraction

Bacterial cell pellets obtained from IWFs (Section 4.2.6.2) or from cell suspension in liquid media was resuspended in 500µl of ice cold ddH_2O and immediately mixed with 1ml of RNAprotect Reagents (Qiagen) by vortexing for 5s. The mixture was incubated for 5min at RT and then centrifuged for 10min at 5000xg. The supernatant was carefully decanted and the remaining liquid was removed by gently dapping the inverted tube once onto a paper towel (about 10s). At this point, the cell pellet might not be visible but still yielded sufficient amount of RNA later on. The pellets could be stored at -80˚ C for about a month. For cell lysis, the pellets were mixed using 100µl of 10mg/ml lysozyme in TE buffer by vortexing for 10s. The cell suspension was incubated for 5min at 22˚ C at 300rpm with regular 10s vortexing every 2min. After that, RNA was extracted using the Qiagen RNeasy Mini Kit with an adapted protocol. 350µl of RLT buffer already mixed with 1% beta mercaptoethanol (v/v) was added to the lysate and vortexed vigorously. 250µl of 96-100% ethanol was then added and mixed gently by pipetting. The lysate was next transferred to a RNeasy Mini spin column supplied with a 2ml collection tube. The column was centrifuged for 15s at 13000rpm and the flow through was discarded. 700µl of RW1 buffer was added to the spin column and washed with the same centrifugation setting. The flow through and the collection tube were then discarded, and the column was placed onto a new 2ml collection tube. Next, the column was washed with 500µl of RPE buffer with the same centrifugation setting, and then washed again with 500µl RPE buffer by centrifugation for 2min at 13000rpm. The column was then placed onto a new collection tube and centrifuge for 1min at full speed to remove residual ethanol. After that, the column was placed onto a new RNase-free 1.5 tube and the membrane was incubated with 30µl of RNase-free water for 3min at RT. RNA was then eluted by centrifugation for 1min at full speed. DNase treatment was performed using the Qiagen RNase-Free DNase Set following the steps in Section 4.2.2.4.

4.2.2.3 DNase treatment and reverse transcription

In-solution DNase treatment was done following the RevertAid First Strand cDNA Synthesis Kit, with DNase I from ThermoFisher Scientific with adaptation. Reaction mixture for DNase treatment was shown in Table 7.

Table 7: DNase treatment reaction mixture

Component	**Volume/amount**
RNA	1µg
10x Reaction buffer with $MgCl_2$	1µl
RNase-free DNase I (1U/µl)	1µl
Nuclease-free water	to 10µl

The reaction mixture in Table 7 was incubated for 30min at 37˚ C. 1µl of 50mM EDTA was added to the mixture and incubated for 10min at 65˚ C to inactivate the DNase I enzyme. 0.5µl of the mixture was used in PCR to check for gDNA contamination. The rest was used for first-strand cDNA reverse transcription. The mixture was mixed with 1µl of Oligo $(dT)_{18}$ primer and incubated for 5min at 65˚ C. The reaction was chilled on ice and spun down before adding with the components listed in the **Table 8**. The reverse transcription was done for 60min at 42˚ C followed by 5min heating at 70˚ C to terminate the reaction.

Table 8: The reaction mixture for first-strand cDNA synthesis

Component	**Volume/amount**
DNase-treated RNA	~10µl
5x Reaction buffer	4µl
RiboLock RNase Inhibitor (20 U/µl)	1µl
10 mM dNTP Mix	2µl
RevertAid RT (200U/µl)	1µl
Total volume	**20µl**

4.2.2.4 RNA sequencing and analysis

For RNA sequencing, RNA extracted from Step 4.2.2.1 was treated with DNase I using the Qiagen RNase-Free DNase Set following the manufacturer's protocol in combination with Qiagen RNeasy® MinElute® Cleanup kit with adaption. Up to 87.5µl of RNA solution was mixed with 10µl of buffer RDD and 2.5µl of DNase I stock solution. The reaction was incubated for 15min at RT and proceeded with RNA clean-up. 350µl of buffer RLT containing 40mM of DTT was mixed with the RNA solution, then 250µl of 96-100% ethanol was added and mixed by pipetting before being transferred to a RNeasy MinElute spin column. The tube was centrifuged for 15s at 13000rpm and the flow through was discarded. The RNA-bound membrane was washed first with 500µl of buffer RPE by centrifugation for 15s at 13000rpm and then with 500µl of 80% ethanol for 2min at 13000rpm. After that, the spin column was placed in a new collection tube and spun for 5min at full speed with the lid open to eliminate ethanol residual. The RNA was then eluted with 14µl of RNase-free water into a 1.5ml Eppendorf by centrifugation for 1min at full speed. Genomic DNA contamination was checked via PCR using the eluted RNA as the template. RNA quantity and quality were assessed using NanoDrop and Agilent 2100 Bioanalyzer.

Directional cDNA library was prepared using the NEBNext Ultra II kit. The sequencing was done using Illumina NovaSeq 6000 to produce 30 million pair-end clusters. RNA-seq reads in Citrus samples at both 12 and 36hpi were mapped against the *C. sinensis* reference genome SWO.v1.0 from Citrus HZAU database (http://citrus.hzau.edu.cn, (Xu et al., 2013)) using the STAR tool version 2.7 (Dobin et al., 2013). Read counting was done using Subread version 1.6.4 (Liao et al., 2019). Differential expression analysis was done with edgeR version 3.24.3 (Robinson et al., 2010). Overlapping analyses and proportional Venn diagrams were generated by DeepVenn tool (Hulsen, 2022) (https://www.deepvenn.com). Heat maps were generated using the R scrip 3 (Appendix 1.5).

4.2.2.5 Bacterial RNA sequencing and analysis

Xcc RNA was subjected to strand-specific meta-transcriptome library preparation with the depletion of both eukaryotic and prokaryotic rRNA (Novogene). The library was sequenced with 150-bp paired-end reads using Illumina Sequencing platform. RNA-seq data was kindly analysed by the lab of Dr. Mario Tyago Murakami. Briefly, reads were quality assessed using

Fastp (v0.23.2) to remove adapter sequences and low-quality bases (Q<30). rRNA sequences were then removed using sortmeRNA (v4.3.6) in the sensitive mode. FastQC was used to analyse the result. The reads were then aligned to *Xcc306* genome sequence (RefSeq GCF_000007165.1) using HISAT2 (v2.2.1) in "very sensitive" mode and not tolerable to discordant read pair alignments. The alignment was converted to .bam format using Samtools (v1.17). A .bed file containing gene coordinates obtained from the gtf file of *Xcc306* genome was used in combination with the .bam file to assess the strandedness of the reads using RSeQC (v5.0.1). Read counts for each gene was done using subread's featureCounts (subread v2.0.5) which only counted paired reads and ignored single reads or paired reads with unmatched alignments. Differential gene expression analysis was done using the R package DESeq2 (v1.34.0) with an adjusted p value < 0.05. PCA map was generated using the rlog-normalised read pair counts in DESeq2. GO enrichment was done using the R package clusterProfiler (v4.2.2) with the "enricher" setting (pAdjustMethod= "BH") using the *Xcc306* gene ontology terms obtained from Pannzer2 server.

4.2.2.6 qRT-PCR

cDNA was diluted 1:4 ratio and 1µl was used for an 8-µl qPCR reaction performed in a 384-well CFX384™ Touch qRT-PCR C1000 Thermocycler Real-Time System from BioRad. The compositions of qPCR mix were: 4µl of 2x MESA BLUE qPCR Master Mix Plus for SYBR® Assay - No ROX from Eurogentec, 1µl of forward and 1µl of reverse primers (10µM concentration of each), double distilled water to 8µl. List of qPCR primers is in Appendix 1.2. The standard qPCR program in Table 6 was used. Three technical replicates were used for each cDNA sample. Mean Ct value was used as the Ct value for each biological replicate. Citrus *FBOX* (*Cs6g19880*) gene was used as the housekeeping gene for normalisation of most qPCR runs except for the qPCR in Figure 3C where *EF1α* (*Cs7g27470*) was used. Tomato *TIP41-Like* (*Solyc10g049850)* gene was used as the reference gene for Tomato cDNAs.

4.2.3 Protein work

4.2.3.1 Recombinant CsLOB1 protein expression

E. coli codon optimised *CsLOB1* coding sequence (*CsLOB1opt*) was synthesised by Eurofins Genomics in a high-copy pEX plasmid. *CsLOB1opt* was PCR amplified with primers TPF4 and TPR5 to introduce BsaI sites that generate the overhang CACC at the 5' end and AAGG at the

3’ end. The digested CsLOB1opt sequence was then cloned into a lab-modified pENTR plasmid with compatible BsaI sites via Golden Gate assembly. The cloned plasmid had *att*-sites flanking the *CsLOB1opt* sequence which allowed the insertion of *CsLOB1opt* into the expression vector pET-53-DEST (Novagen) that contains a T7 promoter and a 6 N-terminal histidines via Gateway LR reaction including 0.75µl of vector (100ng/µl), 0.75µl of insert plasmid (50ng/µl), 0.5µl of LR mix, 0.5µl of TE pH=8. The reaction was incubated for 1h at RT. All reaction was transformed to Top10 cells.

For soluble CsLOB1opt expression, *CsLOB1opt* from *pENTR CsLOB1op1* was cloned into the Golden Gate compatible *pET-53-DEST* together with a C-terminal tag consisting of a flexible glycine-serine linker, a canonical cleavage site ENLYFQG for TEV (Tobacco Etch Virus) protease, 7 Histidines followed by a MBP (maltose-binding protein) sequence to generate a *pET-53-DEST CsLOB1opt-MBP* construct.

For protein expression in *E. coli*, the recombinant *CsLOB1opt* constructs were freshly transformed to the electrocompetent BL21-AI™ (Invitrogen™) via electroporation. On the next day, multiple cells were inoculated into 5ml of liquid LB media containing appropriate Tetracycline and Ampicillin concentration (Table 4) and let grow overnight at 37˚ C, 200rpm. The whole 5ml starting culture was inoculated to 250ml (1:50 ratio) of fresh LB containing appropriate antibiotics with the starting OD_{600}= 0.05-0.1. The cell suspension was grown at 37˚ C, 200rpm for 3-4 hours until the OD_{600} reached 0.8-1.0. The culture was cooled down on ice for 5-10min before being added with 1mM IPTG and 0.2% L-Arabinose (w/v) to induce *CsLOB1* expression. The induced culture was grown for 18-24h at 18˚ C. Cells were harvested by centrifugation at 4˚ C and 4600rpm (Heraeus Multifuge 3SR+ Centrifuge- ThermoFisher). 1ml of the culture was pelleted and subject to western blot to confirm protein expression. The cell pellet could be stored at -20˚ C for not more than 4 weeks before extraction.

Up to 5g of wet cell pellet was suspended in 30ml of Lysis buffer (**Table *9***). Cell suspension was sonicated with 50% amplitude, 5s on, 10s off for 15min in cold adaptor using Active Motif® sonicator. Lysed cell suspension was subject to ultracentrifugation for 30min at 16000x g at 4˚ C using Sorvall RC 6 Plus Centrifuge from Thermo Scientific rotor type SA-600. The supernatant was carefully decanted into a clean 50ml Falcon tube and proceeded with purification. All purification steps were performed on the ÄKTA pure™ chromatography system from Cytiva. pET53-DEST CsLOB1opt was purified in a denatured condition using

HisTrap™ FF 5ml column from Cytiva, while MBP tagged CsLOB1opt was purified with MBPTrap™ column in a native condition (Table 9) following recommended protocols. Briefly, an appropriate purification column was washed with 5 column volume (CV) of ddH_2O to rinse out ethanol residue. The column was then equilibrated with 5CV of Washing buffer before the injection of cell lysis supernatant at low speed (0.5ml/min) for better protein binding and the flow through was discarded. Protein-bound column was washed intensively with up to 20CV of Washing buffer at 1ml/min rate or until total protein reached below 5mAU (absorbance unit) at 280nm. The bound protein was eluted in 1ml fractions using by 5CV of Elution buffer at 0.5ml/min rate. 30µl of each fraction was mixed with 10µl of 4xSDS loading buffer for Coomassie staining and Western blot.

Fractions with correct protein elution were pooled and subjected to dialysis using Slide-A-Lyzer® Dialysis cassette with 20,000 MWCO (molecular weight cut-off) 0.5-3.0ml capacity from Thermo Scientific following the manufacturer's instructions. Briefly, protein fractions were injected into the pre-hydrated cassette via clean syringe and needle through one of the corner ports. The filled cassette was attached to a float buoy and dialyzed at 4˚ C in a flask containing 900ml (at least 200-500 times of the sample volume) of appropriate dialysis buffer (Table 9) and stirred using a magnetic stirring bar to facilitate buffer exchange. The buffer was changed after 2 hours for two times to a fresh buffer. After the third buffer exchange, the dialysis was continued overnight. The dialyzed sample was withdrawn from a different corner port with a fresh syringe and needle into a fresh 2ml Eppendorf. Protein concentration was measured with Bradford assay.

Table 9: Buffers for CsLOB1 purification

Buffer	**Denaturing purification**	**Native purification**
Lysis buffer	50mM Na-P (1M Na-P consists of 0.39M NaH_2PO_4 and 0.61M of Na_2HPO_4, pH=7); 300mM NaCl; 8% (v/v) glycerol, 4M Urea; 1 tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Sigma-Aldrich)	20mM Tris-HCl, pH 8; 200mM NaCl; 10% glycerol (v/v); 1mM DTT; 1 tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Sigma-Aldrich)

Washing buffer	Same as lysis buffer, without proteinase inhibitor cocktail	
Elution buffer	50mM Na-P, pH=7; 300mM NaCl; 4M Urea; 35mM Imidazole	20mM Tris-HCl, pH 8; 200mM NaCl; 10% glycerol (v/v); 1mM DTT; 0.5% maltose (w/v),
Dialysis buffer	50mM Na-P, pH=7; 300mM NaCl; 3M, 2M and 1M Urea for the 1st, 2nd and 3rd buffers, respectively	25 mM Tris–HCl, pH 8.8; 50 mM KCl; 1mM DTT; 2mM EDTA; 10mM MgCl2; 20% glycerol (v/v); 0.5% NP-40(v/v) (EMSA binding buffer (Husbands et al., 2007))

4.2.3.2 Bradford assay:

The Bio-Rad Protein Assay Dye Reagent Concentrate was dilute in 1:5 ratio as the working concentration. Protein was mixed at different concentration with the working solution and incubated at RT for 5min. The absorbance of protein-reagent was measured at 600nm. Calibration curve was prepared with 5, 10, 20, 40 and 60µg/ml of BSA (Bovine Serum Albumin Fraction V) in triplicates. A linear regression, with **a** as the slope, **b** as the intercept on the y-axis, was used to calculate the approximate **x** protein concentration in Bradford solution with a **y** readout from 600nm absorbance following the formula: $\mathbf{x} = \frac{y-b}{a}$ (µg/ml). Dialysed protein was concentrated using the Amicon Ultra-15 Centrifugal Filter Unit from Merck following the manufacturer's instructions.

4.2.3.3 CsLOB1 antibody generation

The purified denatured CsLOB1 at the concentration of around 1mg/ml was sent to BioGenes® for polyclonal antibody production in rabbits. Two animals were selected for immunisation based on the low background signal in a Western blot done on citrus and *N. benthamiana* tissues using their pre-immunised antisera. The rabbits were injected with the purified protein (first immunisation) two times, each was one week apart from the other, and the first antiserum was harvested two weeks after that to be checked for specificity by Western blot. Two additional boosts were performed in the following four months. 30ml of the final antisera were purified using CsLOB1-bound columns and were sent in Tris-Glycine buffer, pH=7.5, 250mM NaCl and 0.1% ProClin 300. The antibody solutions were

concentrated using the Amicon Ultra-15 Centrifugal Filter Unit from Merck and preserved in 20% glycerol (v/v) final concentration at -80˚ C (antibody concentration was 0.6mg/ml).

4.2.3.4 Western blot

1-2 leaf discs (8mm in diameter) were harvested,and flash frozen in a 2ml Eppendorf containing a ceramic bead. The frozen leaf tissue was ground into fine powder using Qiagen RetschMill at 22Hz for 1min in an adaptor previously kept at -20˚ C. 120µl of freshly prepared denaturing extraction buffer (10mM Tris-Cl, pH=8.0, 100mM NaH_2PO_4, 8M Urea) was mixed with the powder by vigorous vortexing for every 10s up to 2min to homogenise. The lysate was centrifuged for 10mins at 13000rpm at 4˚ C. The supernatant was transferred to a clean 1.5ml Eppendorf and mixed with the 4x SDS buffer (200mM Tris-Cl, pH=6.8, 400mM DTT, 8% SDS (w/v), 40% glycerol, 0.4% bromophenol blue (w/v)). The protein mix was heated at 100˚ C for 5min and proceeded with SDS-PAGE.

Extracted protein in SDS buffer was loaded on a 2-layer SDS gel consisting of a short stacking layer (125mM Tris-Cl, pH=6.8, 0.1% SDS (w/v), 0.1% ammonium persulphate (w/v), 4% acrylamide (v/v), 0.1% TEMED (v/v)) and a resolving layer (375mM Tris-Cl, pH=8.8, 0.1% SDS, 0.1% ammonium persulphate, 10% acrylamide, 0.1% TEMED). The gel was run at 120V in SDS running buffer (25mM Tris, pH= 8.3, 192mM glycine, 0.1% SDS) in the Mini-PROTEAN Tetra Vertical Electrophoresis Cell from Bio-Rad until the front dye line reached the bottom of the gel. The proteins were then blotted onto a 0.45µm pore-size PVDF-membrane from Immobilon®-P using a Bio-Rad TurboTransfer system at 25V and 1.0mA for 30min. Next, the membrane was blocked with 3% of BSA in TBST buffer (50mM Tris-Cl, pH=7.5, 150mM NaCl, 0.05% Tween-20 (v/v)) for 1 hour. The membrane was then incubated in an appropriate primary antibody solution diluted in TBST containing 0.5% of BSA overnight at 4˚ C. After that, the membrane was washed three times with TBST, each for 15min, followed by a 30min incubation in 5% milk in TBST. Appropriate HRP-conjugated secondary antibody was added and incubated with the membrane for another 30min. The membrane was then washed three times with TBST, each for 15min, then one time with TBS without Tween-20. The membrane was quickly (<2min) incubated with the Clarity ECL substrate from Bio-Rad and then visualized using the Amersham™ Imager 600. Total protein was visualised by the staining with Ponceau S solution (5% glacial acetic acid (v/v), 0.1% Ponceau S (w/v)).

4.2.3.5 Protein immunoprecipitation

The protocol was adapted from Liu et al. (2016) without formaldehyde cross-linking and DNA purification steps. The components of buffers used are listed in Table 10. Specifically, 4-5 infected leaves (about 5g) per treatment were flash frozen and ground in liquid nitrogen that yielded about 20ml of tissue powder. The tissue was added with regular mixing to a clean 100ml conical flask on ice that already contained 20ml of NIB buffer. Additional 10ml could be added to facilitate the mixing. The slurry solution was filtered through a double Mira-cloth layer placed in a conical filter funnel that collected the flow through into a 50ml Falcon tube. Additional 5-10ml of NIB was used to recover tissue from the flask. A clean 1ml pipette tip was used to squeeze the folded filter paper to increase nuclei recovery. Nuclei suspension was centrifuged for 15min at 4000xg at 4˚ C and the supernatant was carefully discarded by pipetting. The pellet was gently resuspended in 4ml of NIB by pipetting up and down using a cut 1ml pipette tip. The resuspension was centrifuged for 5min at 4000xg at 4˚ C and the supernatant was discarded. The nuclei pellet was washed additional two more times in NIB and could be kept resuspended in NIB at -80˚ C or proceeded with one washing step with PBS buffer containing 0.1mM PMSF. During the last centrifugation step with PBS buffer, NLB was freshly prepared. The nuclei pellet was then resuspended in 1.6ml of the Nuclei lysis buffer and transferred to two Covaris® milliTUBE, each with 800µl of nuclei suspension. The samples were then sonicated in the pre-cooled E220*evolution* Focused-ultrasonicator from Covaris® with the setting: Burst: 200, Duty factor: 21, Power:140, Duration: 240s. Samples were transferred to a 2ml Eppendorf and centrifuged for 5mins at 20000xg at 4˚ C. 45µl of supernatant was saved as nuclei soluble fraction and the rest of the supernatant (roughly 1.1ml) was diluted in 1:10 ratio in ChIP diluting buffer () in a 15ml Falcon tube. 10µl of CsLOB1 antibody (#26166) in 20% glycerol was added and mixed by rotation at 4˚ C for 3 hours using the LABINCO LD-79 rotator at 11rpm. In the last 1 hour of rotation, Pierce™ protein A/G magnetic beads were equilibrated in ChIP diluting buffer as followed. 20µl of beads for one sample was diluted in 60µl of ChIP diluting buffer in a 2ml Eppendorf and briefly mixed by gentle vortexing. Beads were collected using a magnetic rack for 2min and the supernatant was discarded by pipetting. Beads were then washed for another two times with the same buffer before being transferred to the nuclei fraction sample using a cut-end tip and rotated for 3 hours. Protein-antibody bound beads were

reclaimed using a magnetic rack and was washed with 1ml of Low salt wash buffer by mixing with finger flicking, rotating for 3min at 8rpm at 4˚ C (Helmut Saur Laborbedarf Rotator with 2ml tube adapter) before reclaimed using the magnetic rack. The beads were then washed with 1ml of the High salt wash buffer, then 1ml of LiCl wash buffer, and two times with 1ml of TE buffer, each washing step was done with 5min rotation at 8rpm at 4˚ C using the same rotator and bead claiming using the magnetic rack. Recovered beads were resuspended 35µl of 2xSDS buffer and heated at 100˚ C for 5min. 10µl of the sample was loaded on SDS gel and proceeded with Western blot protocol (Section 4.2.3.4) using CsLOB1 antibody as the primary antibody.

Protein immunoprecipitation in *N. benthamiana* was done with similar steps, except the initial material was 1g of leaf tissue using appropriate antibodies specified in individual experiments.

Table 10: Buffers for protein immunoprecipitation

Buffer	**Components**
NIB (Nuclei isolation buffer)	20mM HEPES, pH7.5; 1mM $MgCl_2$; 5mM KCl; 250mM sucrose; 40% glycerol (v/v); 0.25% Triton X-100 (v/v); 0.1mM PMSF (added shortly before the extraction) 0.1% β-mercaptoethanol (v/v); 1 tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Sigma-Aldrich) per 30ml
PBS buffer	137mM NaCl; 2.7mM KCl; 10mM Na_2HPO_4; 1.8mM KH_2PO_4; Adjusted to pH 7.4 with HCl
Nuclei lysis buffer	50mM Tris-Cl, pH=8.0; 10mM EDTA; 1% SDS; 1 tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Sigma-Aldrich); 0.1mM PMSF
ChIP diluting buffer	16.7mM Tris-Cl, pH=8.0; 167mM NaCl; 1.2mM EDTA; 1.1% Triton X-100; 2 tablets of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack

	Protease Inhibitor Cocktail (Sigma-Aldrich) per 30ml; 0.1mM PMSF
Low salt wash buffer	20mM Tris-Cl, pH=8.0; 150mM NaCl; 2mM EDTA; 0.5% Triton X-100; 0.2% SDS
High salt wash buffer	20mM Tris-Cl, pH=8.0; 500mM NaCl; 2mM EDTA; 0.5% Triton X-100; 0.2% SDS
LiCl wash buffer	10mM Tris-Cl, pH=8.0; 250mM LiCl; 1mM EDTA; 0.5% NP-40; 0.5% Sodium deoxycholate
TE buffer	10mM Tris-Cl, pH=8.0; 1mM EDTA

4.2.3.6 ChIP sequencing

ChIP-seq protocol was extended from the protein immunoprecipitation protocol (Section 4.2.3.5) and adapted from Liu et al. (2016). In particular, freshly harvested leaves were shortly rinsed in ice-cold ddH_2O and tapped dry. Then, they were cut into 5-10mm width pieces in ice-cold PBS buffer and transferred immediately to a 50ml Falcon tube containing 37.5ml of ice-cold PBS added with 0.5% of formaldehyde (Stock: 37% Formaldehyde stabilised with 5-15% methanol from ThermoFisher). The tube was vacuum infiltrated for 10min at around 120mbar (several rounds of vacuuming and releasing could be applied until the tissue became dark). 2ml of 2M glycine was added to the buffer and the vacuum infiltration was continued for 5min to stop the crosslinking. Fixed tissue was rinsed in ice-cold ddH_2O and was tapped dry to remove excess water before being flash frozen in liquid nitrogen. Nuclei isolation, sonication and immunoprecipitation were performed following the protocol in Section 4.2.3.5.

ChIP DNA isolation

After nuclei sonication, 35µl of the nuclei soluble fraction was saved as the ChIP input sample. After the bead washing with TE buffer, beads were resuspended in 200µl of Elution buffer added with 8µl of 5M NaCl (IP sample). 165ml Elution buffer and 8µl of 5M NaCl were also added to the 35µl ChIP input sample. Tubes were taped with parafilm and incubated

overnight at 65˚ C to elute and reverse-crosslink the protein-DNA complex. Beads of IP sample were collected using the magnetic rack and the supernatant was transferred into a clean 1.5ml Eppendorf. Two additional elution was performed by adding 100µl of Elution buffer to the beads and incubating at 65˚ C for 15min. Proteinase K was added to the IP and input samples to the final concentration of 44ng/µl and incubated at 45˚ C for 1 hour and proceeded with DNA purification. The sample was added with an equal volume of phenol:chloroform:isoamyl alcohol and vortexed vigorously. The mixture was centrifuged for 30min at 10,000xg at 4˚ C and the upper liquid layer was carefully transferred to a new 1.5ml Eppendorf. An equal volume of chloroform was then added to the sample, vortexed vigorously and centrifuged for 30min at 10,000xg at 4˚ C. The upper layer was then transferred to a new 1.5ml tube and DNA was next mixed with 2.5 volume of 100% ethanol, 1/10 volume of 3M NaAc (pH=5.2) and 2µl of 20mg/ml glycogen. The solution was kept for at least 1 hour at -80˚ C before being centrifuged for 30min at 10,000xg at 4˚ C. DNA pellet was vortexed with 500µl of 70% ethanol and precipitated for 10min at 10,000xg at 4˚ C. The DNA pellet was dried with open lid at RT overnight and was resuspended in 22µl of 10mM Tris-Cl pH=8.0.

End repair and adaptor ligation

10µl of ChIP DNA was used for generating ChIP-seq library using the NEBNext Ultra II DNA Library Prep Kit for Illumina kindly provided by Prof. Chang Liu, University of Hohenheim, Germany. The following components were added to a sterile nuclease-free PCR tube for the End repairing step. The components were carefully mixed by pipetting at least 10 times without introducing air bubbles, then quickly spun down and incubated in a thermocycler with preheated lid (95˚ C) and precool tube holders (20˚ C) following the program: 30min at 20˚ C, 30min at 65˚ C, hold at 4˚ C. Meanwhile, 0.5µl of NEBNext Adaptor for Illumina was diluted in 4.5µl of ddH_2O. Then, the following components were added sequentially into the End-repaired reaction and mixed gently by pipetting without introducing air bubbles.

Components	**Volume**
End-repaired reaction	24µl
Ultra II Ligation Master Mix	12µl
Ligation Enhancer	0.4µl

Diluted NEBNext Adaptor	1µl
Total volume	37.4µl

The reaction was then incubated for 15min at 20˚ C in a thermocycler and hold at 4˚ C before the next step. 1.2µl of USER enzyme was mixed with the reaction and incubated at 37˚ C for 15min.

Desalting and size selection

The purified magnetic SeraPure beads kindly provided by Prof. Chang Liu following the protocol from Rohland and Reich (2012) were used to purify adaptor ligated DNA. 40µl of beads was added to 37µl of the previous reaction and vortexed immediately for 5-10s. The mixture was left for 10min at RT and beads were recovered using a magnetic rack for 0.2ml PCR tubes, while the supernatant was discarded by pipetting. 170µl of 80% ethanol was added, beads were collected in the magnetic rack and the supernatant was discarded. The tube was then left at RT with open lid for 10min to air dry and 20µl of 10mM Tris-cl pH=8.0 was added. The beads were resuspended by flicking, vortexing then were collected by a quick centrifugation and 1-2min on the magnetic rack. DNA solution was then transferred to a new PCR tube. 20µl of fresh beads was added to the DNA solution (1:1 ratio) and mixed by vortexing. The mixture was incubated for 10min at RT and the beads were recovered and washed again with 170µl of ethanol. The beads were air dried for 10min at RT with open lid and DNA was eluted in 10µl of 10mM Tris-Cl pH=8.0.

PCR enrichment of Adaptor-ligated DNA

A trial PCR was done to determine the number of PCR cycles necessary. 0.5µl of the eluted DNA was mixed with 5µl of 2x NEBNext Ultra II Q5 Master mix, 0.5µl of Universal Primer, 0.5µl of Index Primer and 3.5µl of ddH_2O. The mixture was then run in a thermocycler using the following setting:

Cycle Step	**Temp**	**Time**	**Cycles**
Initial Denaturation	98°C	30s	1
Denaturation	98°C	10s	18
Annealing/Extension	65°C	75s	18
Final extension	65°C	5 min	1
Hold	4°C	∞	

The PCR product was analysed on 1% agarose gel stained with ethidium bromide. 18 cycles of PCR were shown to be sufficient, therefore final PCR was done with the same setting using 4.5µl of eluted DNA, 10µl of 2x NEBNext Ultra II Q5 Master mix, 1µl of Universal Primer, 1µl of Index Primer and 3.5µl of ddH_2O.

The amplified product was then desalted using the purified magnetic SeraPure beads using 30µl of beads for 20µl of DNA. DNA was then eluted in 20µl of 10mM Tris-Cl pH=8.0. After that, fragments above 700bp were filtered out by mixing the 20µl of DNA solution in 10µl of beads (0.5 beads: DNA (v/v) ratio). The beads were discarded, and the supernatant was transferred to a new PCR tube. 1.5 volume of fresh beads was then mixed with the supernatant. The beads were then recovered and washed with 170µl of 80% ethanol before DNA elution in 10µl of 10mM Tris-Cl pH=8.0. DNA concentration was measured by Qubit™ dsDNA quantification kit using the high sensitivity protocol recommended by the manufacturer.

ChIP-seq and raw data processing

50-100ng of DNA was sent for NextSeq™ 500 Mid Output sequencing with 100bp paired-end reads to produce minimum 100 million paired-end clusters. ChIP-seq analysis was done following the pipeline from the Galaxy project (Ostrovsky et al., 2022). Specifically, Trim Galore (Krueger et al., 2023) was used to remove adaptors and low-quality reads. Reads were aligned to *C. sinensis* reference genome (Liu et al., 2022) using BWA tool (Li and Durbin, 2009). Low-quality mapped reads was filtered using the Samtools from Li et al. (2009). bamCoverage was used to generate bigWig file format from mapped reads for the visualisation of read coverage. Binding peaks were then defined by MACS2 (Zhang et al., 2008) in comparison between anti-CsLOB1 and IgG control samples. Peaks with -log10 of q value lower than 10 were discarded using the Filter tool using simple expressions on Galaxy. Integrative Genomics Viewers (Robinson et al., 2011) was used to visualise the distribution of mapped reads and peaks in Citrus reference genome.

ChIP-seq data mining

300bp around the peak submits were extracted using R script 1 (Appendix 1.5). The sequences were subjected to the MEME motif discovery (Bailey and Elkan, 1994) in the MEME Suite platform (Bailey et al., 2015). The motif scanning tool FIMO was used to search

for the occurrence of a motif in a given DNA sequence (Grant et al., 2011). ChIPseeker was used for the analysis of peak annotation with different genomic features (Figure 11A) with plotAnnoBar function (Yu et al., 2015). Peaks that are in common between two replicates were filtered by bedtools Intersect intervals with the default overlap of 1bp, A-intersect-B mode (Quinlan and Hall, 2010). The analysis of the distance of peaks to TSS was kindly done by Prof. Chang Liu using R. Briefly, all gene loci were transformed to align at the TSS and the -4kb to +4kb regions were split into 100bp segments. Then the common peaks with the coordinators overlapping with the region of -4kb and +4kb of the TSS were selected. Each 100bp segment that overlaps with the peak region was counted as one hit (R script 2, Appendix 1.5). A gene that has a peak whose summit coordinators locate within the -3kb to +200bp around TSS of the gene was defined as a ChIP-seq target gene (R script 3, Appendix 1.5).

4.2.3.7 Proteomics

2 leaf discs (8mm in diameter) of Citrus leaf infected with *Xcc* and *XccΔpthA4* at 40hpi were ground in liquid nitrogen and mixed vigorously with 200µl of denaturing buffer (Section 4.2.3.4) supplemented with 10mM EDTA, 1mM PMSF and Proteinase inhibitor cocktail. The samples were centrifuged for 10min at 13,000rpm at 4˚ C. Total protein in the supernatant was used for Bradford measurement. 30µl of the supernatant was used to check for CsLOB1 protein by Western blot using CsLOB1 antibody before sending to the Proteomic Centre Tübingen and handled by Irina Droste-Borel. Briefly, the samples were run and extracted SDS gel. LC-MS/MS run was performed on a Proxeon Easy-nLC coupled to Q Exactive™ HF (ThermoScientific) with 130min run. Data was processed using MaxQuant software version 1.6.7.0 with integrated Andromeda Peptid search engine. The spectra were searched against two *C. sinensis* databases (PRJNA86123 from NCBI and *Citrus sinensis* v1.0 from Citrus HZAU database (Liu et al., 2022)) and *X. citri* database (*Xanthomonas citri* NCBI).

4.2.4 Plant assays

4.2.4.1 Bacterial infection assays

Xanthomonas or *A. tumefaciens* were streaked out on a NYG or YEB agar plate, respectively, with appropriate antibiotics and let grown for 2 days in a 28˚ C incubator. Bacterial cells were inoculated in 5ml of corresponding liquid media containing antibiotics and grow overnight at

28˚ C in a 200rpm shaker. The next morning, cells were harvested by centrifugation for 15min at 4000xg at RT. Cells were then resuspended to the final OD of 0.4 in ddH_2O for *Xanthomonas* and in the buffer containing 10mM MgCl2, 10mM MES pH=5.6 and 150µM of acetosyringone for *Agrobacterium*. Cell suspension was infiltrated to the lower side of the leaf using needless syringe. For infection on Citrus leaf, a small scratch was made on the cuticle by a needle to facilitate the infiltration. For transient expression in *N. benthamiana*, leaf tissue was harvested at 2dpi, unless otherwise stated. For Citrus infection, leaf tissue was harvested at 5dpi when water soaking and pustules developed, unless otherwise stated.

4.2.4.2 *Xanthomonas* i*n-planta* growth assay in Citrus

Xcc was grown for 2 days on NYG agar plate with appropriate antibiotics. Bacterial cells were then inoculated into 5ml NYG media and grown overnight at 28˚ C at 200rpm. On the next day, the whole culture was diluted 1:10 in fresh NYG media and let grown for 5-6 hours until the OD_{600}reached between 0.8-1.0 to reach the exponential phase. The cells were harvested by centrifugation for 15min at RT at 4000xg. The cells were then resuspended in double-distil H_2O at OD_{600}=0.004 (10^5 CFU ml^{-1}). Bacteria were syringe-infiltrated into leaves as described in Section 4.2.4.1. For bacterial counting on day 0, one leaf disc (8mm in diameter) was harvested after 30min- 1 hour of infiltration and saved in a 2ml Eppendorf containing one 6.2-6.8mm SiLibeads® ceramic bead and 200µl of 10mM $MgCl_2$. Leaf tissue was homogenised at RT using Qiagen RetschMill at 22Hz for 1min. 10µl of the mixture was immediately plated on NYG agar containing appropriate antibiotics. For bacterial counting at later timepoints after infiltration, the leaf disc was homogenised in 100µl of 10mM $MgCl_2$. 900µl of 10mM $MgCl_2$ was mixed to produce 10^{-1} dilution of bacterial suspension. Serial dilution was made in 1:10 ratio and 10µl of appropriate dilutions were plated on NYG agar with antibiotics.

4.2.4.3 Xanthomonas in-planta growth assay coupled with *Agrobacterium*-mediated transient expression in *N. benthamiana*

Agrobacterium EHA101 and Xanthomonas strains were grown on appropriate agar plates with antibiotics for 2 days. Bacterial cells were inoculated in a 5ml liquid media and grew overnight. The cell suspension was sub-cultured in 1:10 dilution into a 50ml fresh media and grew for 5-6 hours until the OD_{600} reached 0.8-1.0. Cells were harvested by centrifugation for 15min at RT at 4000xg. Cells were washed with 10mM $MgCl_2$ and resuspended in the

same buffer. *Agrobacterium* and *Xanthomonas* were then mixed to the final OD_{600} of 0.2 for *Agrobacterium* and 0.0001 for *Xanthomonas*. The bacterial mixture was infiltrated to the lower side of the leaf of *N. benthamiana roq1* or *eds1* mutants. Bacterial cells were extracted on day 0 and day 6 as described in the Section 4.2.4.2. Cells were plated on NYG agar plates with antibiotics for *Xanthomonas* selection.

4.2.4.4 GUS staining

3-4 leaf discs (6mm in diameter) (one per biological replicate) were harvested and vacuum infiltrated in 300µl of GUS staining solution containing 0.1M sodium phosphate pH=7.0 (1M contains 390mM NaH_2PO_4 and 610mM Na_2HPO_4), 5mM EDTA, 1mM $K_4[Fe(CN)_6]$, 1mM $K_3[Fe(CN)_6]$, 0.1% Triton X-100 (v/v), 0.05% X-Gluc (w/v). The leaf discs were incubated overnight at 37˚ C then destained in 70-80% ethanol at >80˚ C for 1 hour. The destaining could be done twice for better results. Leaf discs were placed on a plastic bag and scanned with Epson Perfection V700 with default settings.

4.2.4.5 RUBY quantification

7 leaf discs (6mm in diameter) were harvested and place in a 2ml tube containing 1ml of 15% ethanol. The tube was incubated for 1 hour at 20˚ C at 300rpm. 800µl of the solution was then transferred to a 10x4x45mm polystyrene cuvette from Sarstedt and measured at 535nm using the GENESYS™ 10S UV-Vis Biophotometer. The solution from leaf discs punched out from non-infiltrated leaves were used as a blank.

4.2.5 Molecular assays

4.2.5.1 EMSA (Electrophoretic mobility shift assay)

Purified CsLOB1-MBP protein (Section 4.2.3.1) was used for EMSA. List of Cy5-labelled and nonlabelled probes were shown in Table 11. 40pmol protein was mixed with nonlabelled competitors at 0, 100, 200 or 400pmol and incubated for 30min at 25˚ C in dark. 4pmol of Cy5-labelled probe was mixed gently by finger flicking to make a 20µl reaction, then was shortly spun down. The mixture was then incubated for another 30min in the same condition. The mixture was added with 2µl of native loading buffer (25mM Tris, 192mM glycine, 50% glycerol, 0.4% Orange G (w/v)). Meanwhile, a 6% native tris-glycine gel (25mM Tris, 192mM glycine, 0.05% ammonium persulphate, 6% acrylamide, 0.05% TEMED) was pre-

electrophoresed in Tris-glycine buffer pH=8.8 at 4˚ C in dark for 1 hour to eliminate the residual non-polymerised acrylamides. Next, the reaction was loaded on the native gel and run for about 1 hour in the same condition. The gel was transferred carefully onto a transparent plastic bag and visualised using Typhoon™ (Cytiva) with the default Cy5 fluorophore setting.

Table 11: Sequences of probes used in EMSA and MST

Probe name	Sequence
wt	TAAAGCCGTGAAACCCGGCGCAATTTGCACTTG
m1	TAAAGCCGTGAAATTTAATGCAATTTGCACTTG
m2	TAAAGCTGTGAAACCCAGCGCAATTTGCACTTG
m3	TAAAGCTGTGAAACCCGGCGCAATTTGCACTTG
m4	TAAAGCCGTGAAACCCAGCGCAATTTGCACTTG
m5	TAAAGCCGTGGAACCCGGCGCAATTTGCACTTG
m6	TAAAGCCGTGGCGCCCGGCGCAATTTGCACTTG
flk	ATTTTCCGTGAAACCCGGCGATTCCCTATAAAT
opt	TAAAGCCGCAATTCACGCCGCAATTTGCACTTG
d1	TTAAAGCCGC-ATTCACGCCGCAATTTGCACTTG
in	TAAAGCCGCAAATTCACGCCGCAATTTGCACTTG
z1	TAAAGCGGCGGCGGCGGCGCAATTTGCACTTG
z2	TAAAGGCCGCGGCGGCAATTTGCACTTG
un	CGTTTTTCCATTGCGGCGGCGGCGACGGTGTC

Highlighted region: CsLOB1 predicted binding motif; letters in red: mutation or insertion; hyphen: deletion. Abbreviations are explained in

4.2.5.2 MST (MicroScale Thermophoresis)

16 dilutions of purified CsLOB1-MBP were prepared from the starting concentration of 204µM in EMSA buffer as in Figure 56 in a PCR tube. 10µl of the 10nM Cy5-labelled probe in EMSA buffer was added to the 10µl of the protein solution and mixed gently by pipetting without introducing air bubbles. The reactions were incubated for 30min on ice. Next, a Monolith NT.115 standard capillary from NanoTemper was dipped at one end into each reaction to allow the solution to enter and fill up the capillary. The capillaries were then put into the MST sample tray from position 1 to 16. The MST program was run with Cold region

start: -1s, Cold region end: 0s, Hot region start: 9s. Hot region end: 10s. Data analysis was done using the MO. Affinity Analysis v2.3. Each type of interaction was repeated three times.

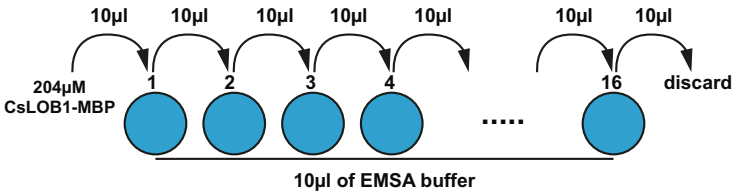


Figure 56: The preparation of CsLOB1 serial dilution.
MST 16 dilution tubes were first added with 10µl of EMSA buffer. 10µl of 204µM of CsLOB1-MBP was mixed with Tube 1, and 10µl of the mixture was transferred to Tube 2, mixed well and then being transferred to Tube 3 and so on. After mixing in Tube 16, 10µl was discarded. Later on, 10µl of 10nM Cy5-labelled probe was added to each tube. The figure was adapted from (Huang and Zhang, 2021)

4.2.6 Metabolic studies

4.2.6.1 Electrolyte leakage

One leaf disc (8mm in diameter) was added to 1ml of ddH_2O in a 2ml Eppendorf and incubated in a heat block at 25˚ C for 1 hour. The liquid was carefully transferred to the sample tray of the CM100-2 Conductivity Meter from Reid & Associates for measurements following the manufacturer's manual. Average of 6 constitutive measurements on the same sample was calculated as the final conductivity readout.

4.2.6.2 Apoplastic fluid extraction

Fully expanded Citrus leaves were infiltrated completely with either *Xcc, XccΔpthA4* or ddH_2O following the infiltration protocol for *in-planta* growth (Section 4.2.4.2). At 7dpi, leaves were removed at the base, rinsed in ice-cold ddH_2O and tap-dried with clean tissue paper. Several scratches on the lower side of the leaf were made using a pair of tweezers. The leaves were then submerged in ice-cold ddH2O with the lower side facing upwards and were vacuum infiltrated with 120mBar vacuum pressure until the whole leaf turned dark green. Each leaf was then wrapped around a 10ml pipetting tip using parafilm to keep the leaf in place. The tip was then inserted into a 50ml Falcon tube with the larger end downwards. The tube was then centrifuged for 30min at 750xg at 4˚ C. Each leaf yielded around 100-300µl of apoplastic fluid or IWF. The fluid was transferred to a 1.5ml Eppendorf and centrifuged for 45min at 20,000xg at 4˚ C. The supernatant was then filter sterilised

through the Ultrafree CL columns (Merk) if the samples were later used for GC-MS or *in-vitro* growth assays. For bacterial RNA extraction, pellets were recovered, flash frozen and stored at -80˚ C.

4.2.6.3 GC-MS

30µl of IWF was mixed with 2000pmol of 3' O-methyl-α-D-glucopyranoside (3' OMG) from Sigma in a 1.5ml Eppendorf. The reaction was collected at the bottom of the tube and gently flash frozen in liquid nitrogen. The tube lid was left opened and an extra lid with punched holes was used to close the tube. Tubes were left at -80˚ C in a metal 1ml tube holder for at least 30min for remaining liquid nitrogen to evaporate and then transferred on the metal tube holder to the Christ Alpha 2-4 LSC lyophilisation system for freeze drying overnight. The samples were then proceeded with GC-MS by Dr. Joachim Kilian. Briefly, samples went through methoxamination in which 50µl of Methoxamine in Pyridine 20mg/ml was added freshly to each dried sample, sonicated for 10min in Ultrasonic bath and incubated for 90min at 30˚ C and 1400rpm. Next, silylation was performed by adding 20µl of MSTFA (Sigma) to the sample and incubated for 60min at 40˚ C and 1200rpm. The sample was then left at RT for 2 hours before being centrifuged for 10min at 13000rpm. 60µl of the 120µl reaction was transferred to a 2ml GC-vial equipped with 200µl inserts for measurement. The measurement was then performed on a GC-MS system model TQ8040 from Shimadzu. In a GC set, 1 µl of the reaction was injected at 280˚ C in a split mode (1:10). The compounds were then separated by a glass capillary column from Restek (Rxi-5SIL-MS) with 0.25mm in diameter, 0.25µm in film thickness and 30m in length. The carries gas was Helium and was run at 1.1ml/min column flow at a controlled linear velocity of 39cm/s. The oven program started at 100˚ C and was hold for 4min before being heated with a rate of 10˚ K/min to a temperature of 320˚ C and was hold for 11min. Next, the interfaces to the mass spectrometer (MS) and the ion source were set to 280˚ C and 200˚ C, respectively. MRM and SIM were used as the acquisition mode.

The signal of 3' OMG was used as an internal normalisation control for derivatisation and GC-MS. Calibration curves for fructose, glucose and mannose were included with 8 different amounts (pmol):

Standard (S) ID	Glucose	Fructose	Mannose

S1	1,000	1,000	20
S2	2,500	2,500	40
S3	5,000	5,000	60
S4	7,500	7,500	80
S5	10,000	10,000	100
S6	15,000	15,000	200
S7	20,000	20,000	300
S8	30,000	30,000	400

Leaf weight before and after vacuum infiltration was recorded. The correction factor for air space difference between leaves was calculated based on the formula suggested by Gentzel et al. (2019) as leaf weight increase was negatively correlated with apoplast hydration level. Therefore, we calculated the normalised sugar level based on the following formula:

Correction factor (or leaf weight increase) = $\frac{\text{Leaf weight after infiltration} - \text{Leaf weight before infiltration}}{\text{Leaf weight before infiltration}}$.

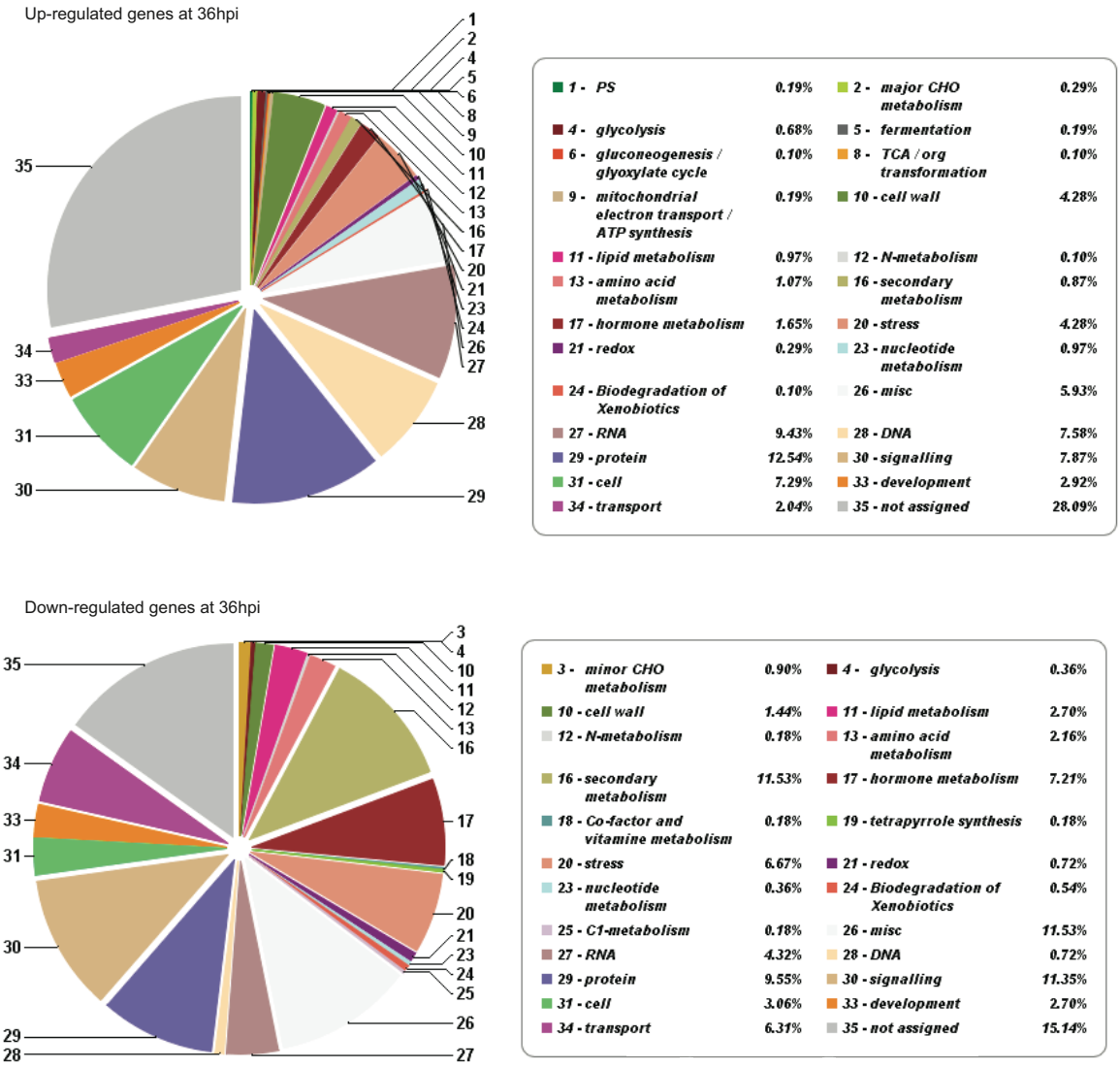
Normalised sugar level = $\frac{\text{GCMS Readout of the sugar of interest}}{\text{GCMS Readout of 3'OMG} * \text{Correction factor}}$/(leaf weight)

4.2.6.4 *In-vitro Xcc* growth

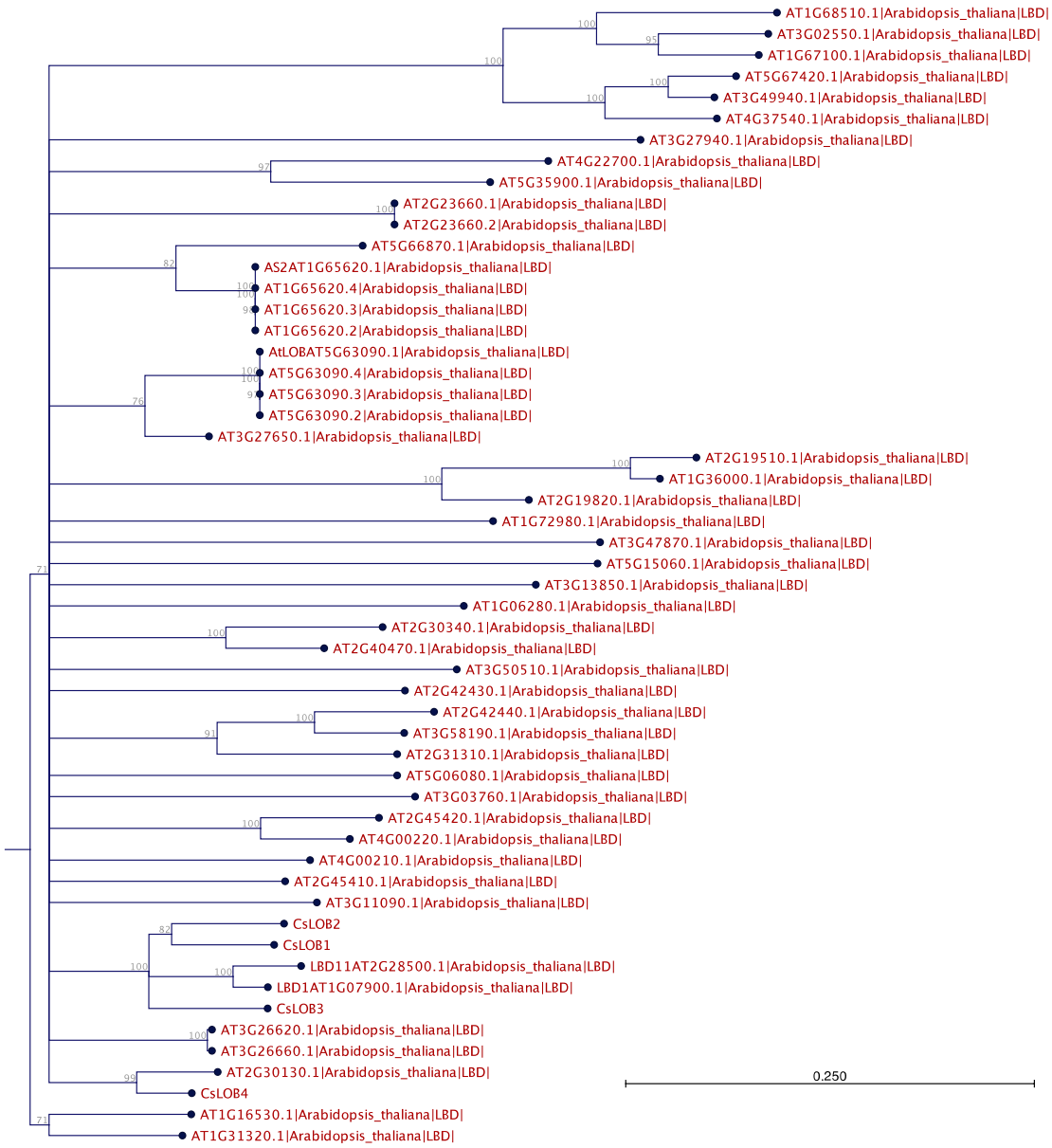
The protocol for *in-vitro* growth assays was adapted from Brich, master thesis (2017). *Xcc* cells growing in exponential phase was prepared as in Section 4.2.4.2. Cells were washed three times in MME media at RT. Cells were then resuspended in MME media to the OD_{600}=0.4. 50µl of cell suspension was added to 50µl of IWF already mixed with MME media in a 60:40 ratio and transferred to a 96-well plate. Bacterial population was determined using the Tecan Safire plate reader using the following setting:

Temperature target	28˚ C
Fixed measure wavelength	600nm
Number of reads per interval	10
Number of kinetic cycles	50
Measurement interval	20min
Shaking mode	Orbital
Intensity	Normal
Shaking duration between intervals	1000s

SUPPLEMENTAL DATA

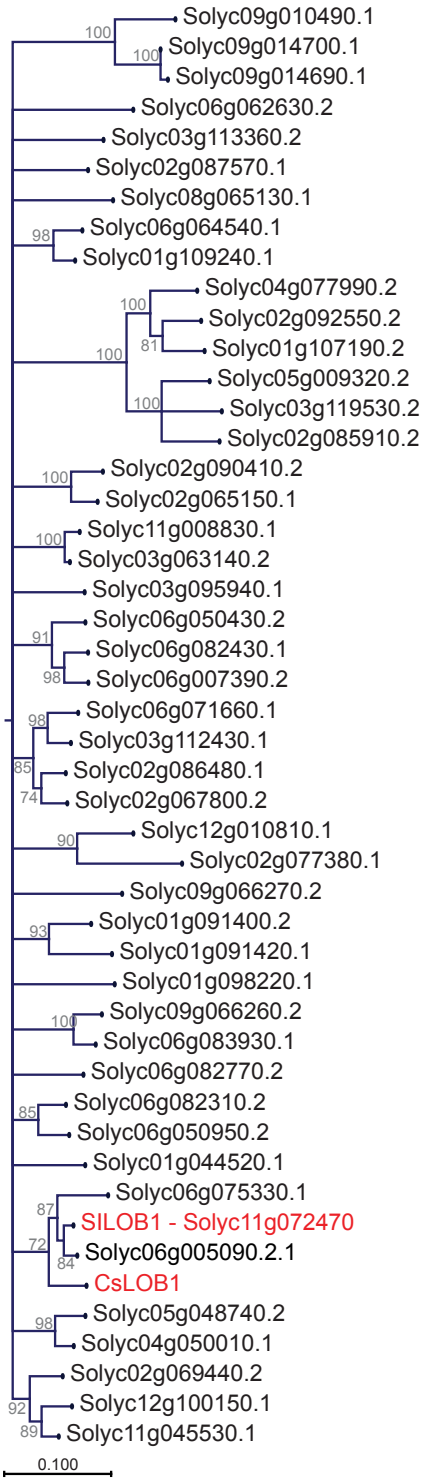


Supplemental Figure 5: Gene Ontology analysis of CsLOB1-dependent DEGs at 36hpi
Gene Ontology analysis was done using MapMan software (Thimm et al., 2004).



Supplemental Figure 6: Phylogenetic tree of Arabidopsis thaliana LOB proteins related to CsLOB1-4

Bootstrap values are in grey. The phylogenetic tree was constructed using Qiagen CLC Main Workbench with default settings.



Supplemental Figure 7: Phylogenetic tree of Tomato LOB proteins related to CsLOB1. SlLOB1 (Solyc11g072470) and CsLOB1 are shown in red. Bootstrap values are in grey. The phylogenetic tree was constructed using Qiagen CLC Main Workbench with default settings.

Supplemental Figure 8: Full length protein alignment of CsLOB1 and SlLOB1
Letters highlighted in black indicate identical amino acids. Red letters highlighted in black indicate characteristic conserved amino acids. Non-highlighted letters are non-identical amino acids. Hyphens show gaps.

Supplemental Table 3: The full list of CsLOB1 direct target genes with putative functions

Gene ID	Putative function	Log_2FC
orange1.1t00187	Putative expansin-B2	9.76
Cs5g01400	Endoglucanase 1	6.84
Cs6g17190	Protein RSI-1|Gibberellin-regulated protein 4	6.55
orange1.1t00855	Putative uncharacterized protein PtrEXPB3 (Fragment)|Expansin-B3	5.79
Cs2g23970	Probable pectate lyase P18	5.58
Cs9g17380	Putative PAR1 protein (bHLH family)	5.54
Cs6g18290	Probable inactive leucine-rich repeat receptor-like protein kinase At1g66830|LRR receptor-like serine/threonine-protein kinase GSO1	4.87
Cs9g14350	Probable pectinesterase/pectinesterase inhibitor 61	4.71
Cs2g20600	Putative E3 ligase	4.69
Cs2g26650	Protein E6	4.69
Cs7g32410	Expansin-A4	4.61
Cs2g12470	Leucine-rich repeat transmembrane protein kinase, putative, expressed|LRR receptor-like serine/threonine-protein kinase GSO1	4.58
Cs5g02320	Endoglucanase 6	4.34
Cs5g25850	Putative uncharacterized protein At5g18310	4.25
Cs8g11330	Probable pectate lyase	4.21
Cs2g27100	GDSL esterase/lipase At5g14450|Esterase|Alpha-L-fucosidase 3	4.19
Cs1g19390	Patellin	4.18
Cs9g13580	Xylem cysteine proteinase 1	4.00
Cs3g09320	Probable calcium-binding protein CML11	3.87
Cs8g09160	Putative uncharacterized protein, PROLINE-RICH PROTEIN 2, EXTENSIN, PROLINE-RICH PROTEIN	3.84
Cs3g17400	Pleckstriny domain-containing protein, putative, expressed|Myosin-IXb	3.75
Cs8g01960	L-ascorbate oxidase homolog|Monocopper oxidase-like protein SKU5|Laccase-2|Laccase-3|Laccase-1|Laccase-23	3.69
Cs6g14390	Indole-3-acetic acid-induced protein ARG7|Auxin-induced protein 15A|Indole-3-acetic acid-induced protein ARG7, putative	3.66
Cs7g18460	Glucomannan 4-beta-mannosyltransferase 2|Probable mannan synthase 9|Mannan	3.66

	synthase 1	
Cs3g23450	Putative remorin	3.66
Cs2g27090	GDSL esterase/lipase At5g14450|Esterase|Alpha-L-fucosidase 3	3.64
Cs1g21630	Putative glycosyltransferase 7|Glycosyltransferase 6|Galactomannan galactosyltransferase 1|Xyloglucan 6-xylosyltransferase	3.61
Cs2g05560	Protein E6	3.44
Cs2g17190	Dynein light chain, cytoplasmic, cell motility	3.36
Cs1g24490	Probable mitochondrial chaperone bcs1	3.34
Cs2g10740	GDP-fucose protein-O-fucosyltransferase 1 (Fragment)	3.30
Cs6g20170	Peroxidase 63|Cationic peroxidase 2|Peroxidase C1C	3.29
Cs8g18640	Putative uncharacterized protein Ptr+E2EXPA15|Expansin-A4	3.23
Cs1g24390	(1-4)-beta-mannan endohydrolase (Mannan endo-1,4-beta-mannosidase), putative|Mannan endo-1,4-beta-mannosidase 7	3.21
Cs1g23090	L-ascorbate oxidase homolog|Monocopper oxidase-like protein SKU5|Laccase-3|Laccase-2|Laccase-1|Laccase-5	3.02
Cs3g15710	Putative uncharacterized protein Sb01g014910, LIPASE	2.87
Cs2g20020	Rho GDP-dissociation inhibitor 1|Putative rho GDP-dissociation inhibitor 1	2.83
Cs5g03170	Probable polygalacturonase|Polygalacturonase	2.78
Cs5g22810	Calmodulin-like protein 3	2.70
Cs1g20490	(RAP Annotation release2) Protein kinase-like domain containing protein|Receptor-like cytosolic serine/threonine-protein kinase RBK1	2.70
Cs2g09450	Remorin, C-terminal region family protein	2.64
Cs3g16330	Transcription factor bHLH74|Transcription factor BPE	2.59
Cs7g25920	WD repeat and HMG-box DNA-binding protein 1|Probable serine/threonine-protein kinase pkwA	2.57
Cs9g14470	Monocopper oxidase-like protein SKS1|L-ascorbate oxidase homolog|Laccase-2	2.48
Cs2g17400	Probable calcium-binding protein CML25	2.45
orange1.1t01602	Heat shock protein Hsp20 domain-containing protein	2.44
Cs3g25100	(1-4)-beta-mannan endohydrolase (Mannan endo-1,4-beta-mannosidase), putative	2.42
Cs8g04525	LYSINE-RICH ARABINOGALACTAN PROTEIN 17-RELATED, Proline rich extensin signature, CLASSICAL ARABINOGALACTAN PROTEIN 9	2.28
Cs7g30760	Putative Myb family transcription factor At1g14600|Probable transcription factor KAN4	2.26
Cs4g18820	short transcript	2.16
orange1.1t02111	Putative uncharacterized protein Sb06g005410|Xylem cysteine proteinase 1	2.11
Cs5g21860	Chitinase 4|Endochitinase A|Endochitinase B (Fragment)	2.05
Cs1g17550	Adenosine diphosphate glucose pyrophosphatase (Precursor)|Auxin-binding protein ABP19a	2.02
Cs2g18330	Formin-like protein 1|Inverted formin-2	1.99
Cs3g17830	CLASSICAL ARABINOGALACTAN PROTEIN 10-RELATED, CLASSICAL ARABINOGALACTAN PROTEIN 9	1.98
Cs1g03610	Malate dehydrogenase, cytoplasmic|Malate dehydrogenase	1.97
Cs5g13360	Mechanosensitive ion channel domain-containing protein-like	1.97
Cs4g09670	Microtubule-associated protein MAP65-1a	1.94
orange1.1t01622	Glutamate decarboxylase, putative, expressed|Glutamate decarboxylase 1	1.91
Cs5g09690	ACT domain containing protein, putative, uridylyltransferase	1.88
Cs5g25090	Cellulose synthase-like protein D3|Putative cellulose synthase-like protein D5	1.83

Cs5g18370	GLUCURONOXYLAN 4-O-METHYLTRANSFERASE 2-RELATED, Polysaccharide biosynthesis	1.81
Cs2g20620	Armadillo/beta-catenin-like repeat family protein, expressed|Vacuolar protein 8	1.80
Cs3g18550	Pollen-specific protein C13|Olee1-like protein|Anther-specific protein LAT52	1.73
Cs7g16050	MLO-like protein 1|MLO protein homolog 1|Protein MLO	1.72
Cs3g17490	Cyclin-D4-1|Cyclin-D4-2	1.69
Cs1g26710	MLO-like protein 8|MLO protein homolog 1	1.68
Cs9g14360	Invertase/pectin methylesterase inhibitor family protein / DC 1.2-like protein|21 kDa protein|Pectinesterase 2|Pectinesterase/pectinesterase inhibitor 3	1.67
Cs2g04370	Aquaporin NIP1-2|Nodulin-26|Putative aquaporin NIP4-1	1.64
Cs8g05250	Plant-specific domain TIGR01627 family protein	1.63
Cs6g19640	Alpha-1,4-glucan-protein synthase [UDP-forming], putative|UDP-arabinopyranose mutase 1	1.62
Cs3g23680	Probable protein phosphatase 2C 12|Protein phosphatase 1K, mitochondrial	1.60
orange1.1t02367	Transcription factor bHLH62|Transcription factor BPE	1.55
Cs1g25950	Chloroplast nucleoid DNA-binding protein-like protein|Aspartic proteinase-like protein 1	1.50
Cs7g28650	Aquaporin TIP1-3|Probable aquaporin TIP1-2|Probable aquaporin TIP-type	1.48
orange1.1t01334	PROTEIN MANNAN SYNTHESIS-RELATED 1	1.42
Cs3g12970	Serine/threonine-protein kinase Nek1|Probable serine/threonine-protein kinase nek3	1.40
Cs7g30110	Putative uncharacterized protein, chloroplast movement	1.39
Cs7g09630	Cytochrome P450 71B34|4-hydroxyphenylacetaldehyde oxime monooxygenase|2-methylbutanal oxime monooxygenase	1.37
Cs2g08250	MYOSIN HEAVY CHAIN-LIKE PROTEIN	1.37
Cs4g04480	UNCHARACTERIZED NODULIN-LIKE PROTEIN, MFS general substrate transporter	1.36
Cs3g18380	Putative uncharacterized protein Sb02g039760, N-LYSINE METHYLTRANSFERASE	1.31
Cs1g21810	UPF0553 protein|UPF0553 protein C9orf64 homolog, QUEUOSINE SALVAGE PROTEIN	1.31
Cs7g30820	Hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyltransferase-like protein (Fragment)|Taxadien-5-alpha-ol O-acetyltransferase	1.29
orange1.1t05077	Armadillo/beta-catenin repeat family protein|U-box domain-containing protein 6|E3 ubiquitin-protein ligase SPL11	1.25
Cs1g24730	Beta-galactosidase 3|Putative beta-galactosidase	1.24
Cs2g22640	Putative cyclin-D6-1	1.23
Cs7g31360	Putative nematode resistance protein Hs1pro-1|Nematode resistance protein-like HSPRO2	1.19
Cs9g16110	Cell cycle regulated microtubule associated protein	1.15
Cs2g28720	Putative XYLOGLUCAN O-ACETYLTRANSFERASE 3	1.12
Cs1g05410	Uncharacterized WD repeat-containing protein C3H5.08c|WD repeat-containing protein 44	1.11
Cs7g32200	TIP41-LIKE PROTEIN	1.10
Cs3g24780	Probable methyltransferase PMT28|Probable pectin methyltransferase QUA2	1.10
Cs5g09360	Plant invertase/pectin methylesterase inhibitor domain-containing protein	1.08
Cs5g09530	Protein NSP-INTERACTING KINASE 2|Leucine-rich repeat receptor-like serine/threonine-protein kinase At1g17230	1.02
Cs5g33180	UDP-glucuronic acid decarboxylase 1|Putative UDP-glucose 4-epimerase	1.01

Supplemental Table 4: Genes activated by CsLOB1 at both 12 and 36hpi

Gene ID	Putative function	Log_2FC
orange1.1t00187	Putative uncharacterized protein Sb01g013560|Putative expansin-B2	9.76
Cs4g02980	Pathogenesis-related protein PR-1	9.35
Cs2g27810	Xylem cysteine proteinase 1	7.36
Cs5g01400	Endoglucanase 1	6.84
Cs6g17190	Protein RSI-1|Gibberellin-regulated protein 4	6.55
orange1.1t05518	Mandelonitrile glucosyltransferase UGT85A19	6.31
Cs9g12620	Putative uncharacterized protein T8G24.2 (Fragment)	6.17
Cs7g27640	CsLOB1	6.17
orange1.1t00855	Putative uncharacterized protein PtrEXPB3 (Fragment)	5.79
Cs2g12620	Adenylate isopentenyltransferase	5.73
Cs9g18910	S-type anion channel SLAH3|Guard cell S-type anion channel SLAC1|C4-dicarboxylate transporter/malic acid transport family protein	5.70
Cs2g23970	Putative uncharacterized protein|Probable pectate lyase P18	5.58
Cs9g17380	Putative uncharacterized protein Sb09g022145 (Fragment)	5.54
orange1.1t03118	Chitinase 4|Endochitinase A	5.39
Cs3g19990	Putative uncharacterized protein	5.34
Cs3g18090	UDP-glycosyltransferase 85A2	5.23
Cs5g06600	Putative uncharacterized protein|Probable pectate lyase 5|Pectate lyase	5.17
Cs7g01530	Putative uncharacterized protein T8G24.2 (Fragment)	4.93
Cs6g18290	Probable inactive leucine-rich repeat receptor-like protein kinase At1g66830|LRR receptor-like serine/threonine-protein kinase GSO1	4.87
Cs9g14350	21 kDa protein|Probable pectinesterase/pectinesterase inhibitor 61|Pectinesterase 2	4.71
Cs2g20600	Putative uncharacterized protein Sb02g029170	4.69
Cs2g26650	Protein E6|Putative uncharacterized protein	4.69
Cs7g32410	Expansin-A4	4.61
Cs2g12470	Leucine-rich repeat transmembrane protein kinase, putative, expressed|LRR receptor-like serine/threonine-protein kinase GSO1	4.58
Cs2g20750	Endoglucanase 11|Endoglucanase A|Putative uncharacterized protein Sb02g024050	4.45
Cs6g22000	Probable serine/threonine-protein kinase At1g54610|Cyclin-dependent kinase C-2	4.40
Cs5g02320	Endoglucanase 6	4.34
Cs5g25850	Putative uncharacterized protein At5g18310	4.25
orange1.1t05099	Xylem cysteine proteinase 1|Ananain	4.25
Cs3g18100	UDP-glycosyltransferase 85A2|Cyanohydrin beta-glucosyltransferase	4.24
Cs8g11330	Putative uncharacterized protein|Probable pectate lyase 1	4.21
Cs2g27100	GDSL esterase/lipase At5g14450|Esterase	4.19
Cs1g19390	Putative uncharacterized protein Sb03g041370|Patellin-5	4.18
Cs7g12860	Di-glucose binding protein with Leucine-rich repeat domain	4.11
Cs6g18720	Cytochrome P450, family 709, subfamily B, polypeptide 2	4.09
Cs3g11340	Putative uncharacterized protein Sb01g004400	4.05
Cs9g13580	Xylem cysteine proteinase 1|Ananain	4.00

Cs1g07650	Uncharacterized basic helix-loop-helix protein At1g06150|Serine/threonine-protein kinase WNK (With No Lysine)-related protein	3.98
Cs5g34710	Probable pectate lyase 8	3.93
Cs3g09320	Probable calcium-binding protein CML11	3.87
Cs8g09160	Putative uncharacterized protein	3.84
Cs3g17400	Pleckstriny domain-containing protein, putative, expressed|Myosin-IXb	3.75
orange1.1t03127	Vignain|Low-temperature-induced cysteine proteinase (Fragment)|Cysteine proteinase RD21a	3.71
Cs7g13770	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin-like protein	3.70
Cs8g01960	Putative uncharacterized protein Sb02g034420|L-ascorbate oxidase homolog	3.69
Cs6g14390	Indole-3-acetic acid-induced protein ARG7|Auxin-induced protein 15A	3.66
Cs7g18460	Glucomannan 4-beta-mannosyltransferase 2	3.66
Cs3g23450	Putative uncharacterized protein Sb04g034210|Uncharacterized protein At3g61260|Remorin	3.66
Cs2g27090	GDSL esterase/lipase At5g14450	3.64
Cs1g21630	Putative glycosyltransferase 7	3.61
Cs1g15020	Whole genome shotgun sequence assembly, scaffold_25, strain Mel28|Probable ATP-dependent RNA helicase DDX11	3.56
Cs3g27500	Glucan endo-1,3-beta-glucosidase 12	3.56
Cs2g05560	Protein E6|Putative uncharacterized protein	3.44
Cs2g17750	Putative uncharacterized protein Sb04g004830	3.44
Cs2g17190	Putative uncharacterized protein Sb03g024870|Dynein light chain, cytoplasmic	3.36
Cs1g24490	Probable mitochondrial chaperone bcs1|Mitochondrial chaperone BCS1	3.34
Cs4g07340	Subtilisin-like protease|Cucumisin	3.33
Cs2g10740	GDP-fucose protein-O-fucosyltransferase 1 (Fragment)	3.30
Cs6g20170	Peroxidase 63	3.29
Cs8g18640	Putative uncharacterized protein Ptr+E2EXPA15|Expansin-A4	3.23
orange1.1t01283	Protein ABIL2	3.22
Cs1g24390	(1-4)-beta-mannan endohydrolase (Mannan endo-1,4-beta-mannosidase), putative	3.21
Cs3g08920	NAC domain-containing protein 102	3.09
Cs1g23090	Putative uncharacterized protein Sb02g034420|L-ascorbate oxidase homolog	3.02
Cs3g15710	Putative uncharacterized protein Sb01g014910	2.87
Cs2g12000	DNA replication licensing factor mcm6|Zygotic DNA replication licensing factor mcm6	2.85
Cs2g20020	Rho GDP-dissociation inhibitor 1	2.83
Cs5g03170	Probable polygalacturonase	2.78
Cs5g31250	Putative uncharacterized protein Sb01g006140	2.75
Cs8g17250	Putative uncharacterized protein Sb03g035030|Interactor of constitutive active ROPs 3	2.74
Cs5g07710	Putative uncharacterized protein	2.71
Cs6g07280	Putative uncharacterized protein Sb03g039920	2.70
Cs5g22810	Calmodulin-like protein 3	2.70
Cs1g20490	(RAP Annotation release2) Protein kinase-like domain containing protein|Receptor-like cytosolic serine/threonine-protein kinase RBK1	2.70
Cs3g26900	Putative uncharacterized protein Sb02g034420|L-ascorbate oxidase homolog	2.67
Cs2g09450	Remorin, C-terminal region family protein, expressed	2.64
orange1.1t02550	Auxin-induced protein 10A5|Indole-3-acetic acid-induced protein ARG7	2.61

Cs3g16330	Transcription factor bHLH74|Transcription factor BPE	2.59
Cs7g25920	WD repeat and HMG-box DNA-binding protein 1|Probable serine/threonine-protein kinase pkwA	2.57
Cs2g18980	Membrane protein involved in ER to golgi transport, putative|Protein transport protein sft2	2.49
Cs9g14470	Monocopper oxidase-like protein SKS1	2.48
Cs3g02660	Domain found in Dishevelled, Egl-10, and Pleckstrin family protein, expressed	2.46
Cs9g14320	21 kDa protein|Pectinesterase 1	2.45
Cs2g17400	Probable calcium-binding protein CML25	2.45
orange1.1t01602	Heat shock protein Hsp20 domain-containing protein	2.44
Cs1g06980	Putative uncharacterized protein Sb05g005260|MLO-like protein 11	2.44
Cs3g25100	(1-4)-beta-mannan endohydrolase (Mannan endo-1,4-beta-mannosidase), putative	2.42
Cs5g20320	Putative uncharacterized protein Sb03g012840|Endoglucanase 9	2.40
Cs5g27660	Putative uncharacterized protein	2.39
Cs2g10070	FAD-binding and BBE domain-containing protein|Reticuline oxidase-like protein|Reticuline oxidase	2.36
orange1.1t01428	Probable leucine-rich repeat receptor-like protein kinase At1g35710	2.34
Cs8g04525		2.28
Cs7g30760	Putative Myb family transcription factor At1g14600|Probable transcription factor KAN4	2.26
Cs4g18820		2.16
Cs2g06560	Putative uncharacterized protein	2.13
orange1.1t02111	Putative uncharacterized protein Sb06g005410|Xylem cysteine proteinase 1	2.11
Cs1g05510	Putative uncharacterized protein Sb06g017280|Probable pectinesterase/pectinesterase inhibitor 51	2.11
Cs5g21860	Chitinase 4|Endochitinase A	2.05
Cs1g17550	Adenosine diphosphate glucose pyrophosphatase (Precursor)	2.02
Cs2g17090	Endoglucanase 8	1.99
Cs2g18330	Putative uncharacterized protein Sb01g043600|Formin-like protein 1	1.99
Cs3g17830		1.98
Cs1g03610	Malate dehydrogenase, cytoplasmic	1.97
Cs5g13360	Mechanosensitive ion channel domain-containing protein-like	1.97
Cs5g09230	Multi-copper oxidase type I family protein, putative	1.95
Cs4g09670	65-kDa microtubule-associated protein 1	1.94
orange1.1t04784	Armadillo/beta-catenin repeat family protein	1.92
orange1.1t01622	Glutamate decarboxylase, putative, expressed	1.91
Cs4g17450	Probable serine/threonine-protein kinase DDB_G0267514	1.90
Cs5g34900	Putative cysteine-rich receptor-like protein kinase 20	1.90
Cs9g14190	Phosphate-responsive 1 family protein (Fragment)	1.89
Cs7g30800	Pentatricopeptide repeat-containing protein At3g56030	1.88
Cs5g09690	Putative uncharacterized protein Sb01g032575 (Fragment)|ACT domain containing protein, putative, expressed|[Protein-PII] uridylyltransferase	1.88
Cs7g11920	Putative uncharacterized protein Sb01g032575 (Fragment)|[Protein-PII] uridylyltransferase	1.88
Cs8g20310	Omega-3 fatty acid desaturase, endoplasmic reticulum	1.87
Cs5g21760	DNA replication licensing factor mcm2	1.86
Cs8g16690		1.84

Cs1g07210		1.84
Cs5g25090	Cellulose synthase-like protein D3	1.83
Cs5g18370	Putative uncharacterized protein At4g09990 (Fragment)	1.81
Cs6g10060	Leucine-rich repeat transmembrane protein kinase, putative, expressed	1.81
Cs2g20620	Armadillo/beta-catenin-like repeat family protein, expressed|Vacuolar protein 8	1.80
Cs5g07600	Helicase superfamily C-terminal domain containing protein	1.80
Cs7g02380	Putative calcium-transporting ATPase 11, plasma membrane-type	1.80
Cs3g16190	1-O-acylglucose:anthocyanin-O-acyltransferase-like protein	1.79
Cs2g12080	CTP synthase 2|CTP synthase 1 (Fragment)|Glutamine amidotransferase, class I, active site	1.76
Cs2g03160	Putative uncharacterized protein P0010D04.11	1.75
Cs7g28720	Putative uncharacterized protein At5g13670 (Fragment)|Auxin-induced protein 5NG4	1.74
Cs3g24430	Calcium uniporter protein, mitochondrial	1.74
Cs3g18550	Pollen-specific protein C13|Olee1-like protein	1.73
Cs9g02950	Glucan endo-1,3-beta-glucosidase 4	1.73
Cs7g16050	MLO-like protein 1	1.72
Cs2g14920	Probable xyloglucan endotransglucosylase/hydrolase protein 5	1.72
Cs9g03230	Receptor-like protein 12|LRR receptor-like serine/threonine-protein kinase FLS2	1.72
Cs9g06400	Protein FANTASTIC FOUR 3|Putative uncharacterized protein Sb01g040410	1.71
Cs3g27390	Putative uncharacterized protein WOX7-1 (Fragment)|WUSCHEL-related homeobox 13	1.70
Cs3g17490	Cyclin-D4-1	1.69
Cs4g15930	Proline-rich receptor-like protein kinase PERK1	1.69
Cs1g26710	MLO-like protein 8	1.68
Cs9g14360	Invertase/pectin methylesterase inhibitor family protein / DC 1.2-like protein	1.67
Cs8g13800	Glycoside hydrolase, family 3, N-terminal|Glycoside hydrolase, family 3, C-terminal	1.67
Cs7g14960	Putative uncharacterized protein Sb06g032680|BTB/POZ domain-containing protein At3g19850	1.66
Cs2g04370	Aquaporin NIP1-2|Nodulin-26	1.64
Cs8g05250	Plant-specific domain TIGR01627 family protein	1.63
Cs6g19640	Alpha-1,4-glucan-protein synthase [UDP-forming], putative|UDP-arabinopyranose mutase 1	1.62
Cs1g24630	Putative uncharacterized protein P0469E09.26-1	1.61
Cs3g23680	Probable protein phosphatase 2C 12	1.60
Cs8g15700	Putative uncharacterized protein OSJNBb0070O09.5	1.60
Cs2g20730	Glycine-rich RNA-binding protein 2, mitochondrial	1.57
Cs7g18150	Leucine-rich repeat receptor protein kinase EXS, putative	1.55
orange1.1t02367	Transcription factor bHLH62	1.55
Cs3g09000	Cellulose synthase-like D1-2, glycosyltransferase family 2 protein	1.51
Cs1g25950	Chloroplast nucleoid DNA-binding protein-like protein|Aspartic proteinase-like protein 1	1.50
Cs9g04170	Glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a	1.50
Cs7g28650	Putative uncharacterized protein At3g26520 (Fragment)|Aquaporin TIP1-3	1.48
Cs1g08380	Putative uncharacterized protein Sb03g041450|Beta-galactosidase 10	1.48
orange1.1t04716	Probable serine/threonine-protein kinase At1g54610	1.45
Cs6g15650	Leucine-rich repeat receptor-like protein kinase (Fragment)	1.43

orange1.1t01334	Putative uncharacterized protein Sb02g028940	1.42
Cs1g16670	Putative uncharacterized protein Sb01g007310|Subtilisin-like protease	1.41
Cs3g12970	Serine/threonine-protein kinase Nek1	1.40
Cs7g30110	Putative uncharacterized protein Sb01g036010	1.39
Cs7g09630	Cytochrome P450 71B34|4-hydroxyphenylacetaldehyde oxime monooxygenase	1.37
Cs2g12180	Endoglucanase 12|Endo-1,4-beta-glucanase Cel1, putative, expressed	1.37
Cs2g08250	Putative uncharacterized protein	1.37
Cs4g04480	Putative uncharacterized protein Sb06g021760	1.36
Cs1g01410	Putative uncharacterized protein P0524E08.125	1.35
orange1.1t00264	Plastocyanin-like domain containing protein, expressed	1.32
orange1.1t03544	Putative uncharacterized protein Sb03g013025 (Fragment)	1.31
Cs3g18380	Putative uncharacterized protein Sb01g033740|Putative uncharacterized protein Sb02g039760	1.31
Cs1g21810	Chromosome undetermined scaffold_99, whole genome shotgun sequence|UPF0553 protein	1.31
Cs6g16890	Proline-rich family protein, putative, expressed	1.31
Cs2g17840	Probable receptor protein kinase TMK1	1.30
Cs1g09840	Probable methyltransferase PMT26|Probable pectin methyltransferase QUA2	1.29
Cs5g08760	Putative uncharacterized protein At4g17310	1.29
Cs7g30820	Hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyltransferase-like protein (Fragment	1.29
Cs5g04130	Putative uncharacterized protein (Fragment)|Protein IQ-DOMAIN 31|Myosin-VIIb	1.28
Cs3g25940	Endoglucanase 10|Putative uncharacterized protein Sb02g030990	1.25
orange1.1t05077	Armadillo/beta-catenin repeat family protein	1.25
Cs2g14810	Helicase, C-terminal|Haem peroxidase, plant/fungal/bacterial (Fragment)	1.25
Cs1g24730	Putative uncharacterized protein|Beta-galactosidase 3	1.24
Cs2g05650		1.24
Cs2g22640	Putative cyclin-D6-1	1.23
Cs7g17520	60S ribosomal protein L18a|Putative uncharacterized protein Sb03g030600	1.23
Cs5g13710	Zn-finger in Ran binding protein and others, putative	1.23
Cs7g24570	Armadillo repeat-containing kinesin-like protein 1	1.22
Cs8g16800	Beta-galactosidase 8	1.21
Cs2g04390	Brassinosteroid signaling positive regulator-related protein (Fragment)	1.20
Cs7g31360	Putative nematode resistance protein Hs1pro-1|Nematode resistance protein-like HSPRO2	1.19
orange1.1t02348	Probable sugar phosphate/phosphate translocator At3g10290	1.17
Cs7g29820	Transcription factor RF2b	1.17
Cs9g16110	Cell cycle regulated microtubule associated protein	1.15
Cs7g10300	ATP-dependent DNA helicase Q-like 5	1.15
Cs8g03680	Tubulin alpha chain|Tubulin alpha-2 chain|Putative uncharacterized protein	1.14
Cs2g28720	Putative uncharacterized protein Sb03g029740	1.12
Cs2g14550	Putative uncharacterized protein Sb07g019863 (Fragment)|Histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH4	1.11
Cs1g06650	Receptor-like protein kinase THESEUS 1	1.11
Cs1g05410	Uncharacterized WD repeat-containing protein C3H5.08c|WD repeat-containing protein 44	1.11

Cs7g22000	Putative uncharacterized protein OSJNBb0070O09.5|Putative uncharacterized protein Sb01g004940	1.10
Cs7g32200	Putative uncharacterized protein	1.10
Cs3g24780	Probable methyltransferase PMT28	1.10
Cs9g17330	Putative uncharacterized protein Sb01g050310|Tubulin beta-1 chain	1.10
Cs6g07480	Tubulin beta-2/beta-3 chain	1.09
Cs3g21610	Alba-like protein C9orf23 homolog|Alba-like protein C9orf23	1.09
Cs5g09360	Plant invertase/pectin methylesterase inhibitor domain-containing protein	1.08
Cs6g21140	Uncharacterized protein At5g39865|Glutaredoxin domain-containing cysteine-rich protein CG12206	1.08
Cs7g13250	Putative uncharacterized protein Sb03g007230|Alpha-xylosidase	1.08
Cs9g07060	Pentatricopeptide repeat-containing protein At4g14850	1.03
Cs5g09530	Protein NSP-INTERACTING KINASE 2	1.02
Cs5g08550	Putative uncharacterized protein OSJNBa0042H09.5	1.02
Cs1g21050	Tubulin alpha chain|Tubulin alpha-2 chain|Putative uncharacterized protein	1.01
Cs2g20030	Fasciclin-like arabinogalactan protein 7|Putative uncharacterized protein Sb06g019725 (Fragment)	1.01
orange1.1t04699	Non-specific lipid-transfer protein-like protein	1.01
Cs6g22090	Biotin carboxyl carrier protein subunit of of Het-ACCase (BCCP1)	1.01
Cs6g08210	Putative uncharacterized protein Sb02g011100	1.01
Cs5g33180	UDP-glucuronic acid decarboxylase 1	1.01

REFERENCES

ABE, V. Y. & BENEDETTI, C. E. 2016. Additive roles of PthAs in bacterial growth and pathogenicity associated with nucleotide polymorphisms in effector-binding elements of citrus canker susceptibility genes. *Molecular Plant Pathology,* 17**,** 1223-1236.

AL-SAADI, A., REDDY, J. D., DUAN, Y. P., BRUNINGS, A. M., YUAN, Q. & GABRIEL, D. W. 2007. All five host-range variants of Xanthomonas citri carry one pthA homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. *Mol Plant Microbe Interact,* 20**,** 934-43.

ANTONY, G., ZHOU, J. H., HUANG, S., LI, T., LIU, B., WHITE, F. & YANG, B. 2010. Rice xa13 Recessive Resistance to Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene Os-11N3. *Plant Cell,* 22**,** 3864-3876.

ATMODJO, M. A., HAO, Z. & MOHNEN, D. 2013. Evolving views of pectin biosynthesis. *Annu Rev Plant Biol,* 64**,** 747-79.

AUSTRALIAN DEPARTMENT OF AGRICULTURE, F. A. F. 2022. *Citrus canker* [Online]. Available: https://www.agriculture.gov.au/biosecurity-trade/pests-diseases-weeds/plant/citrus-canker [Accessed 02 November 2023].

BAILEY, T. L. & ELKAN, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol,* 2**,** 28-36.

BAILEY, T. L., JOHNSON, J., GRANT, C. E. & NOBLE, W. S. 2015. The MEME Suite. *Nucleic Acids Res,* 43**,** W39-49.

BAUER, W. D., BATEMAN, D. F. & WHALEN, C. H. 1977. Purification of an Endo-Beta-1,4 Galactanase Produced by Sclerotinia-Sclerotiorum - Effects on Isolated Plant-Cell Walls and Potato Tissue. *Phytopathology,* 67**,** 862-868.

BEHLAU, F. 2021. An overview of citrus canker in Brazil. *Tropical Plant Pathology,* 46**,** 1-12.

BEHLAU, F., BELASQUE, J., GRAHAM, J. H. & LEITE, R. P. 2010. Effect of frequency of copper applications on control of citrus canker and the yield of young bearing sweet orange trees. *Crop Protection,* 29**,** 300-305.

BERCKMANS, B., VASSILEVA, V., SCHMID, S. P. C., MAES, S., PARIZOT, B., NARAMOTO, S., MAGYAR, Z., KAMEI, C. L. A., KONCZ, C., BOGRE, L., PERSIAU, G., DE JAEGER, G., FRIML, J., SIMON, R., BEECKMAN, T. & DE VEYLDER, L. 2011. Auxin-Dependent Cell Cycle Reactivation through Transcriptional Regulation of Arabidopsis E2Fa by Lateral Organ Boundary Proteins. *Plant Cell,* 23**,** 3671-3683.

BINDER, A., LAMBERT, J., MORBITZER, R., POPP, C., OTT, T., LAHAYE, T. & PARNISKE, M. 2014. A Modular Plasmid Assembly Kit for Multigene Expression, Gene Silencing and Silencing Rescue in Plants. *Plos One,* 9.

BOCH, J., SCHOLZE, H., SCHORNACK, S., LANDGRAF, A., HAHN, S., KAY, S., LAHAYE, T., NICKSTADT, A. & BONAS, U. 2009. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science,* 326**,** 1509-1512.

BOCK, C. H., COOK, A. Z., PARKER, P. E., GOTTWALD, T. R. & GRAHAM, J. H. 2012. Short-distance dispersal of splashed bacteria of Xanthomonas citri subsp citri from canker-infected grapefruit tree canopies in turbulent wind. *Plant Pathology,* 61**,** 829-836.

BOMBARELY, A., ROSLI, H. G., VREBALOV, J., MOFFETT, P., MUELLER, L. A. & MARTIN, G. B. 2012. A Draft Genome Sequence of Nicotiana benthamiana to Enhance Molecular Plant-Microbe Biology Research. *Molecular Plant-Microbe Interactions,* 25**,** 1523-1530.

BONAS, U., STALL, R. E. & STASKAWICZ, B. 1989. Genetic and Structural Characterization of the Avirulence Gene Avrbs3 from Xanthomonas-Campestris Pv Vesicatoria. *Molecular & General Genetics,* 218**,** 127-136.

BONFIM, I. M., PAIXAO, D. A., ANDRADE MD, O., JUNIOR, J. M., PERSINOTI, G. F., GIUSEPPE, P. & MURAKAMI, M. T. 2023. Plant structural and storage glucans trigger distinct transcriptional

responses that modulate the motility of Xanthomonas pathogens. *Microbiol Spectr*, e0228023.
BOVY, A., DE VOS, R., KEMPER, M., SCHIJLEN, E., PERTEJO, M. A., MUIR, S., COLLINS, G., ROBINSON, S., VERHOEYEN, M., HUGHES, S., SANTOS-BUELGA, C. & VAN TUNEN, A. 2002. High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. *Plant Cell,* 14**,** 2509-2526.
BROUN, P. 2004. Transcription factors as tools for metabolic engineering in plants. *Current Opinion in Plant Biology,* 7**,** 202-209.
BRUNINGS, A. M. & GABRIEL, D. W. 2003. Xanthomonas citri: breaking the surface. *Mol Plant Pathol,* 4**,** 141-57.
BÜTTNER, D. & BONAS, U. 2002. Getting across -: bacterial type III effector proteins on their way to the plant cell. *Embo Journal,* 21**,** 5313-5322.
CABRERA, J., FENOLL, C. & ESCOBAR, C. 2015. Genes co-regulated with LBD16 in nematode feeding sites inferred from in silico analysis show similarities to regulatory circuits mediated by the auxin/cytokinin balance in Arabidopsis. *Plant Signaling & Behavior,* 10.
CERNADAS, R. A., DOYLE, E. L., NIÑO-LIU, D. O., WILKINS, K. E., BANCROFT, T., WANG, L., SCHMIDT, C. L., CALDO, R., YANG, B., WHITE, F. F., NETTLETON, D., WISE, R. P. & BOGDANOVE, A. J. 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.
CERVERA, M., NAVARRO, A., NAVARRO, L. & PENA, L. 2008. Production of transgenic adult plants from clementine mandarin by enhancing cell competence for transformation and regeneration. *Tree Physiology,* 28**,** 55-66.
CHANDEL, V., BISWAS, D., ROY, S., VAIDYA, D., VERMA, A. & GUPTA, A. 2022. Current Advancements in Pectin: Extraction, Properties and Multifunctional Applications. *Foods,* 11.
CHEN, L. H., KRACUN, S. K., NISSEN, K. S., MRAVEC, J., JORGENSEN, B., LABAVITCH, J. & STERGIOPOULOS, I. 2021. A diverse member of the fungal Avr4 effector family interacts with de-esterified pectin in plant cell walls to disrupt their integrity. *Science Advances,* 7.
CHEN, L. Q., HOU, B. H., LALONDE, S., TAKANAGA, H., HARTUNG, M. L., QU, X. Q., GUO, W. J., KIM, J. G., UNDERWOOD, W., CHAUDHURI, B., CHERMAK, D., ANTONY, G., WHITE, F. F., SOMERVILLE, S. C., MUDGETT, M. B. & FROMMER, W. B. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature,* 468**,** 527-U199.
CHEN, L. Q., QU, X. Q., HOU, B. H., SOSSO, D., OSORIO, S., FERNIE, A. R. & FROMMER, W. B. 2012. Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport. *Science,* 335**,** 207-211.
CHEN, W. F., WEI, X. B., RETY, S., HUANG, L. Y., LIU, N. N., DOU, S. X. & XI, X. G. 2019. Structural analysis reveals a "molecular calipers" mechanism for a LATERAL ORGAN BOUNDARIES DOMAIN transcription factor protein from wheat. *Journal of Biological Chemistry,* 294**,** 142-156.
CHENG, J. F., NIU, Q. F., ZHANG, B., CHEN, K. S., YANG, R. H., ZHU, J. K., ZHANG, Y. J. & LANG, Z. B. 2018. Downregulation of RdDM during strawberry fruit ripening. *Genome Biology,* 19.
CHOW, V., SHANTHARAJ, D., GUO, Y., NONG, G., MINSAVAGE, G. V., JONES, J. B. & PRESTON, J. F. 2015. Xylan Utilization Regulon in Xanthomonas citri pv. citri Strain 306: Gene Expression and Utilization of Oligoxylosides. *Applied and Environmental Microbiology,* 81**,** 2163-2172.
COHN, M., BART, R. S., SHYBUT, M., DAHLBECK, D., GOMEZ, M., MORBITZER, R., HOU, B. H., FROMMER, W. B., LAHAYE, T. & STASKAWICZ, B. J. 2014. Xanthomonas axonopodis Virulence Is Promoted by a Transcription Activator-Like Effector Mediated Induction of a SWEET Sugar Transporter in Cassava. *Molecular Plant-Microbe Interactions,* 27**,** 1186-1198.
CORREA MARRERO, M., CAPDEVIELLE, S., HUANG, W., AL-SUBHI, A. M., BUSSCHER, M., BUSSCHER-LANGE, J., VAN DER WAL, F., DE RIDDER, D., VAN DIJK, A. D. J., HOGENHOUT, S. A. & IMMINK, R. G. H. 2023. Protein interaction mapping reveals widespread targeting of development-related host transcription factors by phytoplasma effectors. *bioRxiv*.

COSGROVE, D. J. 2005. Growth of the plant cell wall. *Nat Rev Mol Cell Biol,* 6**,** 850-61.
COWLES, K. N., BLOCK, A. K. & BARAK, J. D. 2022. Xanthomonas hortorum pv. gardneri TAL effector AvrHah1 is necessary and sufficient for increased persistence of Salmonella enterica on tomato leaves. *Sci Rep,* 12**,** 7313.
COX, K. L., MENG, F. H., WILKINS, K. E., LI, F. J., WANG, P., BOOHER, N. J., CARPENTER, S. C. D., CHEN, L. Q., ZHENG, H., GAO, X. Q., ZHENG, Y., FEI, Z. J., YU, J. Z., ISAKEIT, T., WHEELER, T., FROMMER, W. B., HE, P., BOGDANOVE, A. J. & SHAN, L. 2017. TAL effector driven induction of a SWEET gene confers susceptibility to bacterial blight of cotton. *Nature Communications,* 8.
CRAGG, S. M., BECKHAM, G. T., BRUCE, N. C., BUGG, T. D. H., DISTEL, D. L., DUPREE, P., ETXABE, A. G., GOODELL, B. S., JELLISON, J., MCGEEHAN, J. E., MCQUEEN-MASON, S. J., SCHNORR, K., WALTON, P. H., WATTS, J. E. M. & ZIMMER, M. 2015. Lignocellulose degradation mechanisms across the Tree of Life. *Current Opinion in Chemical Biology,* 29**,** 108-119.
DA SILVA, A. C., FERRO, J. A., REINACH, F. C., FARAH, C. S., FURLAN, L. R., QUAGGIO, R. B., MONTEIRO-VITORELLO, C. B., VAN SLUYS, M. A., ALMEIDA, N. F., ALVES, L. M., DO AMARAL, A. M., BERTOLINI, M. C., CAMARGO, L. E., CAMAROTTE, G., CANNAVAN, F., CARDOZO, J., CHAMBERGO, F., CIAPINA, L. P., CICARELLI, R. M., COUTINHO, L. L., CURSINO-SANTOS, J. R., EL-DORRY, H., FARIA, J. B., FERREIRA, A. J., FERREIRA, R. C., FERRO, M. I., FORMIGHIERI, E. F., FRANCO, M. C., GREGGIO, C. C., GRUBER, A., KATSUYAMA, A. M., KISHI, L. T., LEITE, R. P., LEMOS, E. G., LEMOS, M. V., LOCALI, E. C., MACHADO, M. A., MADEIRA, A. M., MARTINEZ-ROSSI, N. M., MARTINS, E. C., MEIDANIS, J., MENCK, C. F., MIYAKI, C. Y., MOON, D. H., MOREIRA, L. M., NOVO, M. T., OKURA, V. K., OLIVEIRA, M. C., OLIVEIRA, V. R., PEREIRA, H. A., ROSSI, A., SENA, J. A., SILVA, C., DE SOUZA, R. F., SPINOLA, L. A., TAKITA, M. A., TAMURA, R. E., TEIXEIRA, E. C., TEZZA, R. I., TRINDADE DOS SANTOS, M., TRUFFI, D., TSAI, S. M., WHITE, F. F., SETUBAL, J. C. & KITAJIMA, J. P. 2002. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. *Nature,* 417**,** 459-63.
DAHER, F. B. & BRAYBROOK, S. A. 2015. How to let go: pectin and plant cell adhesion. *Front Plant Sci,* 6**,** 523.
DALIO, R. J. D., MAGALHAES, D. M., RODRIGUES, C. M., ARENA, G. D., OLIVEIRA, T. S., SOUZA-NETO, R. R., PICCHI, S. C., MARTINS, P. M. M., SANTOS, P. J. C., MAXIMO, H. J., PACHECO, I. S., DE SOUZA, A. A. & MACHADO, M. A. 2017. PAMPs, PRRs, effectors and R-genes associated with citrus-pathogen interactions. *Annals of Botany,* 119**,** 749-774.
DANG, T. V. T., LEE, S., CHO, H., CHOI, K. & HWANG, I. 2023. The LBD11-ROS feedback regulatory loop modulates vascular cambium proliferation and secondary growth in Arabidopsis. *Mol Plant,* 16**,** 1131-1145.
DENG, D., YIN, P., YAN, C. Y., PAN, X. J., GONG, X. Q., QI, S., XIE, T., MAHFOUZ, M., ZHU, J. K., YAN, N. & SHI, Y. G. 2012. Recognition of methylated DNA by TAL effectors. *Cell Research,* 22**,** 1502-1504.
DESLANDES, L. & RIVAS, S. 2012. Catch me if you can: bacterial effectors and plant targets. *Trends in Plant Science,* 17**,** 644-655.
DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.
DODD, D., MOON, Y. H., SWAMINATHAN, K., MACKIE, R. I. & CANN, I. K. O. 2010. Transcriptomic Analyses of Xylan Degradation by Prevotella bryantii and Insights into Energy Acquisition by Xylanolytic Bacteroidetes. *Journal of Biological Chemistry,* 285**,** 30261-30273.
DUAN, S., JIA, H. G., PANG, Z. Q., TEPER, D., WHITE, F., JONES, J., ZHOU, C. Y. & WANG, N. 2018. Functional characterization of the citrus canker susceptibility gene *CsLOB1*. *Molecular Plant Pathology,* 19**,** 1908-1916.
EISENREICH, W., DANDEKAR, T., HEESEMANN, J. & GOEBEL, W. 2010. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nature Reviews Microbiology,* 8**,** 401-412.

FENG, G., WU, J., XU, Y., LU, L. & YI, H. 2021. High-spatiotemporal-resolution transcriptomes provide insights into fruit development and ripening in Citrus sinensis. *Plant Biotechnol J,* 19**,** 1337-1353.

FENG, X. J., XIONG, J., ZHANG, W. X., GUAN, H. R., ZHENG, D., XIONG, H., JIA, L., HU, Y., ZHOU, H. M., WEN, Y., ZHANG, X. M., WU, F. K., WANG, Q. J., XU, J. & LU, Y. L. 2022. ZmLBD5, a class-II LBD gene, negatively regulates drought tolerance by impairing abscisic acid synthesis. *Plant Journal,* 112**,** 1364-1376.

FERENCE, C. M., GOCHEZ, A. M., BEHLAU, F., WANG, N., GRAHAM, J. H. & JONES, J. B. 2018. Recent advances in the understanding of Xanthomonas citri ssp citri pathogenesis and citrus canker disease management. *Molecular Plant Pathology,* 19**,** 1302-1318.

FONSECA, N. P., PATANE, J. S. L., VARANI, A. M., FELESTRINO, E. B., CANESCHI, W. L., SANCHEZ, A. B., CORDEIRO, I. F., LEMES, C. G. C., ASSIS, R. A. B., GARCIA, C. C. M., BELASQUE, J., JR., MARTINS, J., JR., FACINCANI, A. P., FERREIRA, R. M., JACIANI, F. J., DE ALMEIDA, N. F., FERRO, J. A., MOREIRA, L. M. & SETUBAL, J. C. 2019. Analyses of Seven New Genomes of Xanthomonas citri pv. aurantifolii Strains, Causative Agents of Citrus Canker B and C, Show a Reduced Repertoire of Pathogenicity-Related Genes. *Front Microbiol,* 10**,** 2361.

FRANCO-ZORRILLA, J. M., LOPEZ-VIDRIERO, I., CARRASCO, J. L., GODOY, M., VERA, P. & SOLANO, R. 2014. DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 2367-2372.

FRASER, P. D., ROMER, S., SHIPTON, C. A., MILLS, P. B., KIANO, J. W., MISAWA, N., DRAKE, R. G., SCHUCH, W. & BRAMLEY, P. M. 2002. Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 1092-1097.

GAMEZ-ARJONA, F. M., VITALE, S., VOXEUR, A., DORA, S., MULLER, S., SANCHO-ANDRES, G., MONTESINOS, J. C., DI PIETRO, A. & SANCHEZ-RODRIGUEZ, C. 2022. Impairment of the cellulose degradation machinery enhances Fusarium oxysporum virulence but limits its reproductive fitness. *Science Advances,* 8.

GANTNER, J., ORDON, J., KRETSCHMER, C., GUEROIS, R. & STUTTMANN, J. 2019. An EDS1-SAG101 Complex Is Essential for TNL-Mediated Immunity in. *Plant Cell,* 31**,** 2456-2474.

GENTZEL, I., GIESE, L., ZHAO, W. Y., ALONSO, A. P. & MACKEY, D. 2019. A Simple Method for Measuring Apoplast Hydration and Collecting Apoplast Contents. *Plant Physiology,* 179**,** 1265-1272.

GEORGELIS, N., NIKOLAIDIS, N. & COSGROVE, D. J. 2015. Bacterial expansins and related proteins from the world of microbes. *Applied Microbiology and Biotechnology,* 99**,** 3807-3823.

GIL, J. F., LIEBE, S., THIEL, H., LENNEFORS, B. L., KRAFT, T., GILMER, D., MAISS, E., VARRELMANN, M. & SAVENKOV, E. I. 2018. Massive up-regulation of LBD transcription factors and EXPANSINs highlights the regulatory programs of rhizomania disease. *Molecular Plant Pathology,* 19**,** 2333-2348.

GIUSEPPE, P. O., BONFIM, I. M. & MURAKAMI, M. T. 2023. Enzymatic systems for carbohydrate utilization and biosynthesis in Xanthomonas and their role in pathogenesis and tissue specificity. *Essays Biochem*.

GLUCK-THALER, E., CERUTTI, A., PEREZ-QUINTERO, A. L., BUTCHACAS, J., ROMAN-REYNA, V., MADHAVAN, V. N., SHANTHARAJ, D., MERFA, M. V., PESCE, C., JAUNEAU, A., VANCHEVA, T., LANG, J. M., ALLEN, C., VERDIER, V., GAGNEVIN, L., SZUREK, B., BECKHAM, G. T., DE LA FUENTE, L., PATEL, H. K., SONTI, R. V., BRAGARD, C., LEACH, J. E., NOEL, L. D., SLOT, J. C., KOEBNIK, R. & JACOBS, J. M. 2020. Repeated gain and loss of a single gene modulates the evolution of vascular plant pathogen lifestyles. *Sci Adv,* 6.

GOCHEZ, A. M., BEHLAU, F., SINGH, R., ONG, K., WHILBY, L. & JONES, J. B. 2020. Panorama of citrus canker in the United States. *Tropical Plant Pathology,* 45**,** 192-199.

GOCHEZ, A. M., MINSAVAGE, G. V., POTNIS, N., CANTEROS, B. I., STALL, R. E. & JONES, J. B. 2015. A functional XopAG homologue in Xanthomonas fuscans pv. aurantifolii strain C limits host range. *Plant Pathology,* 64**,** 1207-1214.

GOCHEZ, A. M., SHANTHARAJ, D., POTNIS, N., ZHOU, X., MINSAVAGE, G. V., WHITE, F. F., WANG, N., HURLBERT, J. C. & JONES, J. B. 2017. Molecular characterization of XopAG effector AvrGf2 from Xanthomonas fuscans ssp. aurantifolii in grapefruit. *Mol Plant Pathol,* 18**,** 405-419.

GOTTWALD, T. R., GRAHAM, J. H. & SCHUBERT, T. S. 2002. Citrus Canker: The Pathogen and Its Impact. *Plant Health Progress,* 3.

GRAHAM, J. H., GOTTWALD, T. R., CUBERO, J. & ACHOR, D. S. 2004. Xanthomonas axonopodis pv. citri: factors affecting successful eradication of citrus canker. *Molecular Plant Pathology,* 5**,** 1-15.

GRANT, C. E., BAILEY, T. L. & NOBLE, W. S. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics,* 27**,** 1017-1018.

GRONDIN, J. M., TAMURA, K., DEJEAN, G., ABBOTT, D. W. & BRUMER, H. 2017. Polysaccharide Utilization Loci: Fueling Microbial Communities. *Journal of Bacteriology,* 199.

GU, Z. G., EILS, R. & SCHLESNER, M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics,* 32**,** 2847-2849.

GUNL, M., NEUMETZLER, L., KRAEMER, F., DE SOUZA, A., SCHULTINK, A., PENA, M., YORK, W. S. & PAULY, M. 2011. AXY8 Encodes an alpha-Fucosidase, Underscoring the Importance of Apoplastic Metabolism on the Fine Structure of Arabidopsis Cell Wall Polysaccharides. *Plant Cell,* 23**,** 4025-4040.

HA, M. A., APPERLEY, D. C., EVANS, B. W., HUXHAM, M., JARDINE, W. G., VIETOR, R. J., REIS, D., VIAN, B. & JARVIS, M. C. 1998. Fine structure in cellulose microfibrils: NMR evidence from onion and quince. *Plant Journal,* 16**,** 183-190.

HAYASHI, T. & KAIDA, R. 2011. Functions of Xyloglucan in Plant Cells. *Molecular Plant,* 4**,** 17-24.

HE, Y., ZHANG, T., SUN, H., ZHAN, H. & ZHAO, Y. 2020. A reporter for noninvasively monitoring gene expression and plant transformation. *Hortic Res,* 7**,** 152.

HOLSTERS, M., SILVA, B., VAN VLIET, F., GENETELLO, C., DE BLOCK, M., DHAESE, P., DEPICKER, A., INZE, D., ENGLER, G., VILLARROEL, R. & ET AL. 1980. The functional organization of the nopaline A. tumefaciens plasmid pTiC58. *Plasmid,* 3**,** 212-30.

HOOD, E. E., HELMER, G. L., FRALEY, R. T. & CHILTON, M. D. 1986. The hypervirulence of Agrobacterium tumefaciens A281 is encoded in a region of pTiBo542 outside of T-DNA. *J Bacteriol,* 168**,** 1291-301.

HORN, S. J., VAAJE-KOLSTAD, G., WESTERENG, B. & EIJSINK, V. G. H. 2012. Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels,* 5.

HOU, Q., AN, X., MA, B., WU, S., WEI, X., YAN, T., ZHOU, Y., ZHU, T., XIE, K., ZHANG, D., LI, Z., ZHAO, L., NIU, C., LONG, Y., LIU, C., ZHAO, W., NI, F., LI, J., FU, D., YANG, Z. N. & WAN, X. 2023. ZmMS1/ZmLBD30-orchestrated transcriptional regulatory networks precisely control pollen exine development. *Mol Plant,* 16**,** 1321-1338.

HU, Y., DUAN, S., ZHANG, Y. Z., SHANTHARAJ, D., JONES, J. B. & WANG, N. 2016. Temporal Transcription Profiling of Sweet Orange in Response to PthA4-Mediated Xanthomonas citri subsp citri Infection. *Phytopathology,* 106**,** 442-451.

HU, Y., ZHANG, J. L., JIA, H. G., SOSSO, D., LI, T., FROMMER, W. B., YANG, B., WHITE, F. F., WANG, N. A. & JONES, J. B. 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** E521-E529.

HU, Y. Z., DING, Y. X., CAI, B. Y., QIN, X. H., WU, J. N., YUAN, M. H., WAN, S., ZHAO, Y. & XIN, X. F. 2022. Bacterial effectors manipulate plant abscisic acid signaling for creation of an aqueous apoplast. *Cell Host & Microbe,* 30**,** 518-+.

HUAI, B. Y., YANG, Q., WEI, X. B., PAN, Q. L., KANG, Z. S. & LIU, J. 2020. TaSTP13 contributes to wheat susceptibility to stripe rust possibly by increasing cytoplasmic hexose concentration. *Bmc Plant Biology,* 20.

HUANG, D., TANG, Z. Z., FU, J. L., YUAN, Y., DENG, X. X. & XU, Q. 2020. CsMYB3 and CsRuby1 form an 'Activator-and-Repressor' Loop for the Regulation of Anthocyanin Biosynthesis in Citrus. *Plant and Cell Physiology,* 61**,** 318-330.

HUANG, H., LIU, R. E., NIU, Q. F., TANG, K., ZHANG, B., ZHANG, H., CHEN, K. S., ZHU, J. K. & LANG, Z. B. 2019. Global increase in DNA methylation during orange fruit development and ripening. *Proceedings of the National Academy of Sciences of the United States of America,* 116**,** 1430-1436.

HUANG, L. & ZHANG, C. H. 2021. Microscale Thermophoresis (MST) to Detect the Interaction Between Purified Protein and Small Molecule. *Plant Chemical Genomics, 2 Edition,* 2213**,** 187-193.

HUERTA-CEPAS, J., SZKLARCZYK, D., HELLER, D., HERNANDEZ-PLAZA, A., FORSLUND, S. K., COOK, H., MENDE, D. R., LETUNIC, I., RATTEI, T., JENSEN, L. J., VON MERING, C. & BORK, P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research,* 47**,** D309-D314.

HULSEN, T. 2022. DeepVenn - a web application for the creation of area-proportional Venn diagrams using the deep learning framework Tensorflow. *arXiv*.

HUSBANDS, A., BELL, E. M., SHUAI, B., SMITH, H. M. S. & SPRINGER, P. S. 2007. LATERAL ORGAN BOUNDARIES defines a new family of DNA-binding transcription factors and can interact with specific bHLH proteins. *Nucleic Acids Research,* 35**,** 6663-6671.

IBRAHIM, Y. E., WIDYAWAN, A., SHARAFADDIN, A. H., PRUVOST, O., BOYER, K. & AL-SALEH, M. A. 2023. Re-occurrence of Asiatic citrus canker in the Makkah Province of Saudi Arabia and characterization of the causal agent *Xanthomonas citri* pv. *citri* isolated from Mexican lime. *Journal of Plant Pathology,* 105**,** 805-815.

ISHIDA, K. & NOUTOSHI, Y. 2022. The function of the plant cell wall in plant-microbe interactions. *Plant Physiol Biochem,* 192**,** 273-284.

ISHIDA, K. & YOKOYAMA, R. 2022. Reconsidering the function of the xyloglucan endotransglucosylase/hydrolase family. *Journal of Plant Research,* 135**,** 145-156.

JALAN, N., KUMAR, D., ANDRADE, M. O., YU, F., JONES, J. B., GRAHAM, J. H., WHITE, F. F., SETUBAL, J. C. & WANG, N. 2013. Comparative genomic and transcriptome analyses of pathotypes of Xanthomonas citri subsp. citri provide insights into mechanisms of bacterial virulence and host range. *BMC Genomics,* 14**,** 551.

JIA, H. G., ZHANG, Y. Z., ORBOVIC, V., XU, J., WHITE, F. F., JONES, J. B. & WANG, N. 2017. Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant Biotechnology Journal,* 15**,** 817-823.

JIA, K. A., WANG, W., ZHANG, Q. & JIA, W. S. 2023. Cell Wall Integrity Signaling in Fruit Ripening. *International Journal of Molecular Sciences,* 24.

JIN, J. P., TIAN, F., YANG, D. C., MENG, Y. Q., KONG, L., LUO, J. C. & GAO, G. 2017. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research,* 45**,** D1040-D1045.

JUTRAS, P. V., SOLDAN, R., DODDS, I., SCHUSTER, M., PRESTON, G. M. & VAN DER HOORN, R. A. L. 2021. AgroLux: bioluminescent Agrobacterium to improve molecular pharming and study plant immunity. *Plant Journal,* 108**,** 600-612.

KARR, A. L. & ALBERSHEIM, P. 1970. Polysaccharide-Degrading Enzymes Are Unable to Attack Plant Cell Walls without Prior Action by a Wall-Modifying-Enzyme. *Plant Physiology,* 46**,** 69-+.

KE, L. L., ZHOU, Z. W., XU, X. W., WANG, X., LIU, Y. L., XU, Y. T., HUANG, Y., WANG, S. T., DENG, X. X., CHEN, L. L. & XU, Q. 2019. Evolutionary dynamics of lincRNA transcription in nine citrus species. *Plant Journal,* 98**,** 912-927.

KLEPIKOVA, A. V., KASIANOV, A. S., GERASIMOV, E. S., LOGACHEVA, M. D. & PENIN, A. A. 2016. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant Journal,* 88**,** 1058-1070.

KONG, F. & YANG, L. 2023. Pathogen-triggered changes in plant development: Virulence strategies or host defense mechanism? *Front Microbiol,* 14**,** 1122947.

KRUEGER, F., JAMES, F., EWELS, P., AFYOUNIAN, E., WEINSTEIN, M., SCHUSTER-BOECKLER, B. & GERT HULSELMANS, G. 2023. FelixKrueger/TrimGalore: v0.6.10 - add default decompression path (0.6.10). .

LAIA, M. L., MOREIRA, L. M., DEZAJACOMO, J., BRIGATI, J. B., FERREIRA, C. B., FERRO, M. I., SILVA, A. C., FERRO, J. A. & OLIVEIRA, J. C. 2009. New genes of Xanthomonas citri subsp. citri involved in pathogenesis and adaptation revealed by a transposon-based mutant library. *BMC Microbiol,* 9**,** 12.

LAPEBIE, P., LOMBARD, V., DRULA, E., TERRAPON, N. & HENRISSAT, B. 2019. Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nature Communications,* 10.

LEE, H. W., KIM, M. J., KIM, N. Y., LEE, S. H. & KIM, J. 2013. LBD18 acts as a transcriptional activator that directly binds to the EXPANSIN14 promoter in promoting lateral root emergence of Arabidopsis. *Plant Journal,* 73**,** 212-224.

LEYNS, F., DECLEENE, M., SWINGS, J. G. & DELEY, J. 1984. The Host Range of the Genus Xanthomonas. *Botanical Review,* 50**,** 308-356.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25**,** 1754-1760.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., DURBIN, R. & PROC, G. P. D. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.

LIAO, Y., SMYTH, G. K. & SHI, W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res,* 47**,** e47.

LIU, C., WANG, C. M., WANG, G., BECKER, C., ZAIDEM, M. & WEIGEL, D. 2016. Genome-wide analysis of chromatin packing in Arabidopsis thaliana at single-gene resolution. *Genome Research,* 26**,** 1057-1068.

LIU, H. M. Z., WANG, X., LIU, S. J., HUANG, Y., GUO, Y. X., XIE, W. Z., LIU, H., UL QAMAR, M. T., XU, Q. & CHEN, L. L. 2022. Citrus Pan-Genome to Breeding Database (CPBD): A comprehensive genome database for citrus breeding. *Molecular Plant,* 15**,** 1503-1505.

LOMBARD, V., RAMULU, H. G., DRULA, E., COUTINHO, P. M. & HENRISSAT, B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research,* 42**,** D490-D495.

MACLEAN, A. M., SUGIO, A., MAKAROVA, O. V., FINDLAY, K. C., GRIEVE, V. M., TOTH, R., NICOLAISEN, M. & HOGENHOUT, S. A. 2011. Phytoplasma Effector SAP54 Induces Indeterminate Leaf-Like Flower Development in Arabidopsis Plants. *Plant Physiology,* 157**,** 831-841.

MEDINA, C. A., REYES, P. A., TRUJILLO, C. A., GONZALEZ, J. L., BEJARANO, D. A., MONTENEGRO, N. A., JACOBS, J. M., JOE, A., RESTREPO, S., ALFANO, J. R. & BERNAL, A. 2018. The role of type III effectors from Xanthomonas axonopodis pv. manihotis in virulence and suppression of plant immunity. *Molecular Plant Pathology,* 19**,** 593-606.

MINSAVAGE, G. V., DAHLBECK, D., WHALEN, M. C., KEARNEY, B., BONAS, U., STASKAWICZ, B. J. & STALL, R. E. 1990. Gene-for-Gene Relationships Specifying Disease Resistance in Xanthomonas-Campestris Pv Vesicatoria - Pepper Interactions. *Molecular Plant-Microbe Interactions,* 3**,** 41-47.

MORBITZER, R., ELSAESSER, J., HAUSNER, J. & LAHAYE, T. 2011. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Research,* 39**,** 5790-5799.

MORBITZER, R., ROMER, P., BOCH, J. & LAHAYE, T. 2010. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 21617-21622.

MOSCOU, M. J. & BOGDANOVE, A. J. 2009. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science,* 326**,** 1501-1501.

MUIR, S. R., COLLINS, G. J., ROBINSON, S., HUGHES, S., BOVY, A., DE VOS, C. H. R., VAN TUNEN, A. J. & VERHOEYEN, M. E. 2001. Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nature Biotechnology,* 19**,** 470-474.

NGOC, L. B., VERNIERE, C., VITAL, K., GUERIN, F., GAGNEVIN, L., BRISSE, S., AH-YOU, N. & PRUVOST, O. 2009. Development of 14 minisatellite markers for the citrus canker bacterium, Xanthomonas citri pv. citri. *Mol Ecol Resour,* 9**,** 125-7.

O'MALLEY, R. C., HUANG, S. S. C., SONG, L., LEWSEY, M. G., BARTLETT, A., NERY, J. R., GALLI, M., GALLAVOTTI, A. & ECKER, J. R. 2016. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell,* 165**,** 1280-1292.

OSTROVSKY, A. E., MAHMOUD, A., LONIE, A. J., SYME, A., FOUILLOUX, A., BRETAUDEAU, A., NEKRUTENKO, A., KUMAR, A., ESCHENLAUER, A. C., DESANTO, A. D., GUERLER, A., SERRANO-SOLANO, B., BATUT, B., GRUNING, B. A., LANGHORST, B. W., CARR, B., RAUBENOLT, B. A., HYDE, C. J., BROMHEAD, C. J., BARNETT, C. B., ROYAUX, C., GALLARDO, C., BLANKENBERG, D., FORNIKA, D. J., BAKER, D., BOUVIER, D., CLEMENTS, D., MORAIS, D. A. D., TABERNERO, D. L., LARIVIERE, D., NASR, E., AFGAN, E., ZAMBELLI, F., HEYL, F., PSOMOPOULOS, F., COPPENS, F., PRICE, G. R., CUCCURU, G., LE CORGUILLE, G., VON KUSTER, G., AKBULUT, G. G., RASCHE, H., HANS-RUDOLF, H., EGUINOA, I., MAKUNIN, I., RANAWAKA, I. J., TAYLOR, J. P., JOSHI, J., HILLMAN-JACKSON, J., GOCKS, J., CHILTON, J. M., KAMALI, K., SUDERMAN, K., POTERLOWICZ, K., YVAN, L., LOPEZ-DELISLE, L., SARGENT, L., BASSETTI, M. E., TANGARO, M. A., VAN DEN BEEK, M., CECH, M., BERNT, M., FAHRNER, M., TEKMAN, M., FOLL, M. C., SCHATZ, M. C., CRUSOE, M. R., RONCORONI, M., KUCHER, N., CORAOR, N., STOLER, N., RHODES, N., SORANZO, N., PINTER, N., GOONASEKERA, N. A., MORENO, P. A., VIDEM, P., MELANIE, P., MANDREOLI, P., JAGTAP, P. D., GU, Q., WEBER, R. J. M., LAZARUS, R., VORDERMAN, R. H. P., HILTEMANN, S., GOLITSYNSKIY, S., GARG, S., BRAY, S. A., GLADMAN, S. L., LEO, S., MEHTA, S. P., GRIFFIN, T. J., JALILI, V., YVES, V., WEN, V., NAGAMPALLI, V. K., BACON, W. A., DE KONING, W., MAIER, W., BRIGGS, P. J., et al. 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research,* 50**,** W345-W351.

PENG, Z., HU, Y., ZHANG, J., HUGUET-TAPIA, J. C., BLOCK, A. K., PARK, S., SAPKOTA, S., LIU, Z., LIU, S. & WHITE, F. F. 2019. Xanthomonas translucens commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. *Proc Natl Acad Sci U S A,* 116**,** 20938-20946.

PEREIRA, A. L. A., CARAZZOLLE, M. F., ABE, V. Y., DE OLIVEIRA, M. L. P., DOMINGUES, M. N., SILVA, J. C., CERNADAS, R. A. & BENEDETTI, C. E. 2014. Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. *Bmc Genomics,* 15.

PEREIRA GONZATTO, M. & SCHERER SANTOS, J. 2023. Introductory Chapter: World Citrus Production and Research. *Citrus Research - Horticultural and Human Health Aspects.*

POTNIS, N., COLEE, J., JONES, J. B. & BARAK, J. D. 2015. Plant pathogen-induced water-soaking promotes Salmonella enterica growth on tomato leaves. *Appl Environ Microbiol,* 81**,** 8126-34.

POTNIS, N., KRASILEVA, K., CHOW, V., ALMEIDA, N. F., PATIL, P. B., RYAN, R. P., SHARLACH, M., BEHLAU, F., DOW, J. M., MOMOL, M. T., WHITE, F. F., PRESTON, J. F., VINATZER, B. A., KOEBNIK, R., SETUBAL, J. C., NORMAN, D. J., STASKAWICZ, B. J. & JONES, J. B. 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *Bmc Genomics,* 12.

QUINLAN, A. R. & HALL, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26**,** 841-2.

RALSER, M., QUERFURTH, R., WARNATZ, H. J., LEHRACH, H., YASPO, M. L. & KROBITSCH, S. 2006. An efficient and economic enhancer mix for PCR. *Biochemical and Biophysical Research Communications,* 347**,** 747-751.

REDGWELL, R. J., MACRAE, E., HALLETT, I., FISCHER, M., PERRY, J. & HARKER, R. 1997. In vivo and in vitro swelling of cell walls during fruit ripening. *Planta,* 203**,** 162-173.
ROBINSON, J. T., THORVALDSDÓTTIR, H., WINCKLER, W., GUTTMAN, M., LANDER, E. S., GETZ, G. & MESIROV, J. P. 2011. Integrative genomics viewer. *Nature Biotechnology,* 29**,** 24-26.
ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.
ROEDER, A. H. & YANOFSKY, M. F. 2006. Fruit development in Arabidopsis. *Arabidopsis Book,* 4**,** e0075.
ROHLAND, N. & REICH, D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research,* 22**,** 939-946.
ROMER, P., HAHN, S., JORDAN, T., STRAUSS, T., BONAS, U. & LAHAYE, T. 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science,* 318**,** 645-648.
ROMER, P., RECHT, S., STRAUSS, T., ELSAESSER, J., SCHORNACK, S., BOCH, J., WANG, S. & LAHAYE, T. 2010. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, Xanthomonas oryzae pv. oryzae. *New Phytol,* 187**,** 1048-1057.
RYAN, R. P., VORHOLTER, F. J., POTNIS, N., JONES, J. B., VAN SLUYS, M. A., BOGDANOVE, A. J. & DOW, J. M. 2011. Pathogenomics of Xanthomonas: understanding bacterium-plant interactions. *Nature Reviews Microbiology,* 9**,** 344-355.
RYBAK, M., MINSAVAGE, G. V., STALL, R. E. & JONES, J. B. 2009. Identification of Xanthomonas citri ssp. citri host specificity genes in a heterologous expression host. *Mol Plant Pathol,* 10**,** 249-62.
SAHA, B. C. 2003. Hemicellulose bioconversion. *Journal of Industrial Microbiology & Biotechnology,* 30**,** 279-291.
SANTOS, C. R., HOFFMAM, Z. B., MARTINS, V. P. D., ZANPHORLIN, L. M., ASSIS, L. H. D., HONORATO, R. V., DE OLIVEIRA, P. S. L., RULLER, R. & MURAKAMI, M. T. 2014. Molecular Mechanisms Associated with Xylan Degradation by Xanthomonas Plant Pathogens. *Journal of Biological Chemistry,* 289**,** 32186-32200.
SCHELLER, H. V. & ULVSKOV, P. 2010. Hemicelluloses. *Annual Review of Plant Biology, Vol 61,* 61**,** 263-289.
SCHENSTNYI, K., STRAUSS, A., DRESSEL, A., MORBITZER, R., WUNDERLICH, M., ANDRADE, A. G., PHAN, T. T., AGUILERA, P. L. A., BRANCATO, C., BERENDZEN, K. W. & LAHAYE, T. 2022. The tomato resistance gene Bs4 suppresses leaf watersoaking phenotypes induced by AvrHah1, a transcription activator-like effector from tomato-pathogenic xanthomonads. *New Phytol,* 236**,** 1856-1870.
SCHRÖDER, R., ATKINSON, R. G. & REDGWELL, R. J. 2009. Re-interpreting the role of endo-β-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Annals of Botany,* 104**,** 197-204.
SCHULTINK, A., QI, T. C., LEE, A., STEINBRENNER, A. D. & STASKAWICZ, B. 2017. Roq1 mediates recognition of the Xanthomonas and Pseudomonas effector proteins XopQ and HopQ1. *Plant Journal,* 92**,** 787-795.
SCHWARTZ, A. R., MORBITZER, R., LAHAYE, T. & STASKAWICZ, B. J. 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 of the United States of America,* 114**,** E897-E903.
SCHWARTZ, A. R., POTNIS, N., TIMILSINA, S., WILSON, M., PATANE, J., MARTINS, J., JR., MINSAVAGE, G. V., DAHLBECK, D., AKHUNOVA, A., ALMEIDA, N., VALLAD, G. E., BARAK, J. D., WHITE, F. F., MILLER, S. A., RITCHIE, D., GOSS, E., BART, R. S., SETUBAL, J. C., JONES, J. B. & STASKAWICZ, B. J. 2015. Phylogenomics of Xanthomonas field strains infecting pepper and tomato reveals

diversity in effector repertoires and identifies determinants of host specificity. *Front Microbiol,* 6**,** 535.

SEKI, Y., KIKUCHI, Y., YOSHIMOTO, R., ABURAI, K., KANAI, Y., RUIKE, T., IWABATA, K., GOITSUKA, R., SUGAWARA, F., ABE, M. & SAKAGUCHI, K. 2015. Promotion of crystalline cellulose degradation by expansins from Oryza sativa. *Planta,* 241**,** 83-93.

SEMIARTI, E., UENO, Y., TSUKAYA, H., IWAKAWA, H., MACHIDA, C. & MACHIDA, Y. 2001. The asymmetric leaves2 gene of Arabidopsis thaliana regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. *Development,* 128**,** 1771-1783.

SHI, Y. N., VREBALOV, J., ZHENG, H., XU, Y. M., YIN, X. R., LIU, W. L., LIU, Z. M., SORENSEN, I., SU, G. Q., MA, Q. Y., EVANICH, D., ROSE, J. K. C., FEI, Z. J., VAN ECK, J., THANNHAUSER, T., CHEN, K. S. & GIOVANNONI, J. J. 2021. A tomato LATERAL ORGAN BOUNDARIES transcription factor, SlLOB1, predominantly regulates cell wall and softening components of ripening. *Proceedings of the National Academy of Sciences of the United States of America,* 118.

SHUAI, B., REYNAGA-PENA, C. G. & SPRINGER, P. S. 2002. The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant Physiol,* 129**,** 747-61.

SIMMONS, T. J., MORTIMER, J. C., BERNARDINELLI, O. D., POPPLER, A. C., BROWN, S. P., DEAZEVEDO, E. R., DUPREE, R. & DUPREE, P. 2016. Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat Commun,* 7**,** 13902.

SINGER, A. U., SCHULZE, S., SKARINA, T., XU, X. H., CUI, H., ESCHEN-LIPPOLD, L., EGLER, M., SRIKUMAR, T., RAUGHT, B., LEE, J., SCHEEL, D., SAVCHENKO, A. & BONAS, U. 2013. A Pathogen Type III Effector with a Novel E3 Ubiquitin Ligase Architecture. *Plos Pathogens,* 9.

SOLE, M., SCHEIBNER, F., HOFFMEISTER, A. K., HARTMANN, N., HAUSE, G., ROTHER, A., JORDAN, M., LAUTIER, M., ARLAT, M. & BUTTNER, D. 2015. Xanthomonas campestris pv. vesicatoria Secretes Proteases and Xylanases via the Xps Type II Secretion System and Outer Membrane Vesicles. *J Bacteriol,* 197**,** 2879-93.

SOUZA, C. D., LI, S. D., LIN, A. Z., BOUTROT, F., GROSSMANN, G., ZIPFEL, C. & SOMERVILLE, S. C. 2017. Cellulose-Derived Oligomers Act as Damage-Associated Molecular Patterns and Trigger Defense-Like Responses. *Plant Physiology,* 173**,** 2383-2398.

SRIVASTAVA, V., MCKEE, L. S. & BULONE, V. 2017. Plant Cell Walls. *eLS.*

STILIANOUDAKIS, S. C., MARSHALL, M. A. & DOZMOROV, M. G. 2022. preciseTAD: a transfer learning framework for 3D domain boundary prediction at base-pair resolution. *Bioinformatics,* 38**,** 621-630.

STREUBEL, J., PESCE, C., HUTIN, M., KOEBNIK, R., BOCH, J. & SZUREK, B. 2013. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to Xanthomonas oryzae pv. oryzae. *New Phytologist,* 200**,** 808-819.

SU, H., WANG, Y., XU, J., OMAR, A. A., GROSSER, J. W., CALOVIC, M., ZHANG, L., FENG, Y., VAKULSKAS, C. A. & WANG, N. 2023. Generation of the transgene-free canker-resistant Citrus sinensis using Cas12a/crRNA ribonucleoprotein in the T0 generation. *Nat Commun,* 14**,** 3957.

SUBEDI, A., BARRERA, L. B. T., LEWIS IVEY, M., EGEL, D., KEBEDE, M., KARA, S., AYSAN, Y., MINSAVAGE, G. V., ROBERTS, P., JONES, J. B. & GOSS, E. M. 2023. Population genomics reveals an emerging lineage of Xanthomonas perforans on pepper. *Phytopathology*.

SUGIO, A., KINGDOM, H. N., MACLEAN, A. M., GRIEVE, V. M. & HOGENHOUT, S. A. 2011. Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** E1254-E1263.

SUN, X., STALL, R. E., JONES, J. B., CUBERO, J., GOTTWALD, T. R., GRAHAM, J. H., DIXON, W. N., SCHUBERT, T. S., CHALOUX, P. H., STROMBERG, V. K., LACY, G. H. & SUTTON, B. D. 2004. Detection and Characterization of a New Strain of Citrus Canker Bacteria from Key/Mexican Lime and Alemow in South Florida. *Plant Dis,* 88**,** 1179-1188.

SWARUP, S., DEFEYTER, R., BRLANSKY, R. H. & GABRIEL, D. W. 1991. A Pathogenicity Locus from Xanthomonas-Citri Enables Strains from Several Pathovars of Xanthomonas-Campestris to Elicit Cankerlike Lesions on Citrus. *Phytopathology,* 81**,** 802-809.

SWARUP, S., YANG, Y., KINGSLEY, M. T. & GABRIEL, D. W. 1992. An Xanthomonas citri pathogenicity gene, pthA, pleiotropically encodes gratuitous avirulence on nonhosts. *Mol Plant Microbe Interact,* 5**,** 204-13.

TAYI, L., MAKU, R. V., PATEL, H. K. & SONTI, R. V. 2016. Identification of Pectin Degrading Enzymes Secreted by Xanthomonas oryzae pv. oryzae and Determination of Their Role in Virulence on Rice. *Plos One,* 11.

TEPER, D., WHITE, F. F. & WANG, N. 2022. The Dynamic Transcription Activator-Like Effector Family of Xanthomonas. *Phytopathology*.

TERRAPON, N., LOMBARD, V., GILBERT, H. J. & HENRISSAT, B. 2015. Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics,* 31**,** 647-655.

THATCHER, L. F., POWELL, J. J., AITKEN, E. A. B., KAZAN, K. & MANNERS, J. M. 2012. The Lateral Organ Boundaries Domain Transcription Factor LBD20 Functions in Fusarium Wilt Susceptibility and Jasmonate Signaling in Arabidopsis. *Plant Physiology,* 160**,** 407-418.

THIMM, O., BLASING, O., GIBON, Y., NAGEL, A., MEYER, S., KRUGER, P., SELBIG, J., MULLER, L. A., RHEE, S. Y. & STITT, M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J,* 37**,** 914-39.

TIMILSINA, S., POTNIS, N., NEWBERRY, E. A., LIYANAPATHIRANAGE, P., IRUEGAS-BOCARDO, F., WHITE, F. F., GOSS, E. M. & JONES, J. B. 2020. Xanthomonas diversity, virulence and plant-pathogen interactions. *Nat Rev Microbiol,* 18**,** 415-427.

TORNERO, P. & DANGL, J. L. 2001. A high-throughput method for quantifying growth of phytopathogenic bacteria in Arabidopsis thaliana. *Plant Journal,* 28**,** 475-481.

TORONEN, P., MEDLAR, A. & HOLM, L. 2018. PANNZER2: a rapid functional annotation web server. *Nucleic Acids Research,* 46**,** W84-W88.

TURNER, P., BARBER, C. & DANIELS, M. 1984. Behavior of the Transposons Tn5 and Tn7 in Xanthomonas-Campestris Pv Campestris. *Molecular & General Genetics,* 195**,** 101-107.

TURNER, P., BARBER, C. & DANIELS, M. 1985. Evidence for Clustered Pathogenicity Genes in Xanthomonas-Campestris Pv Campestris. *Molecular and General Genetics,* 199**,** 338-343.

VAAJE-KOLSTAD, G., WESTERENG, B., HORN, S. J., LIU, Z. L., ZHAI, H., SORLIE, M. & EIJSINK, V. G. H. 2010. An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science,* 330**,** 219-222.

VAN DEN ACKERVEKEN, G., MAROIS, E. & BONAS, U. 1996. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell,* 87**,** 1307-16.

VIEIRA, P. S., BONFIM, I. M., ARAUJO, E. A., MELO, R. R., LIMA, A. R., FESSEL, M. R., PAIXAO, D. A. A., PERSINOTI, G. F., ROCCO, S. A., LIMA, T. B., PIROLLA, R. A. S., MORAIS, M. A. B., CORREA, J. B. L., ZANPHORLIN, L. M., DIOGO, J. A., LIMA, E. A., GRANDIS, A., BUCKERIDGE, M. S., GOZZO, F. C., BENEDETTI, C. E., POLIKARPOV, I., GIUSEPPE, P. O. & MURAKAMI, M. T. 2021. Xyloglucan processing machinery in Xanthomonas pathogens and its role in the transcriptional activation of virulence factors. *Nat Commun,* 12**,** 4049.

VOGEL, J. 2008. Unique aspects of the grass cell wall. *Curr Opin Plant Biol,* 11**,** 301-7.

WANG, C., GONG, B. S., BUSHEL, P. R., THIERRY-MIEG, J., THIERRY-MIEG, D., XU, J. S., FANG, H., HONG, H. X., SHEN, J., SU, Z. Q., MEEHAN, J., LI, X. J., YANG, L., LI, H. Q., LABAJ, P. P., KREIL, D. P., MEGHERBI, D., GAJ, S., CAIMENT, F., VAN DELFT, J., KLEINJANS, J., SCHERER, A., DEVANARAYAN, V., WANG, J., YANG, Y., QIAN, H. R., LANCASHIRE, L. J., BESSARABOVA, M., NIKOLSKY, Y., FURLANELLO, C., CHIERICI, M., ALBANESE, D., JURMAN, G., RICCADONNA, S., FILOSI, M., VISINTAINER, R., ZHANG, K. K., LI, J. Y., HSIEH, J. H., SVOBODA, D. L., FUSCOE, J. C., DENG, Y. D., SHI, L. M., PAULES, R. S., AUERBACH, S. S. & TONG, W. D. 2014. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nature Biotechnology,* 32**,** 926-932.

WANG, L. F., RONG, W. & HE, C. Z. 2008. Two Xanthomonas extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. *Molecular Plant-Microbe Interactions,* 21**,** 555-563.

WANG, T., PARK, Y. B., CAPORINI, M. A., ROSAY, M., ZHONG, L. H., COSGROVE, D. J. & HONG, M. 2013. Sensitivity-enhanced solid-state NMR detection of expansin's target in plant cell walls. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 16444-16449.

WANG, Y., PRUITT, R. N., NURNBERGER, T. & WANG, Y. C. 2022. Evasion of plant immunity by microbial pathogens. *Nature Reviews Microbiology,* 20**,** 449-464.

WICHMANN, G. & BERGELSON, J. 2004. Effector genes of Xanthomonas axonopodis pv. vesicatoria promote transmission and enhance other fitness traits in the field. *Genetics,* 166**,** 693-706.

WILLATS, W. G., MCCARTNEY, L., MACKIE, W. & KNOX, J. P. 2001. Pectin: cell biology and prospects for functional analysis. *Plant Mol Biol,* 47**,** 9-27.

WORMIT, A. & USADEL, B. 2018. The Multifaceted Role of Pectin Methylesterase Inhibitors (PMEIs). *International Journal of Molecular Sciences,* 19.

WU, T., ZHANG, H., YUAN, B., LIU, H., KONG, L., CHU, Z. & DING, X. 2022. Tal2b targets and activates the expression of OsF3H(03g) to hijack OsUGT74H4 and synergistically interfere with rice immunity. *New Phytol,* 233**,** 1864-1880.

XIA, T., LI, Y. J., SUN, D. L., ZHUO, T., FAN, X. J. & ZOU, H. S. 2016. Identification of an Extracellular Endoglucanase That Is Required for Full Virulence in Xanthomonas citri subsp citri. *Plos One,* 11.

XIAN, L., YU, G., WEI, Y., RUFIAN, J. S., LI, Y., ZHUANG, H., XUE, H., MORCILLO, R. J. L. & MACHO, A. P. 2020. A Bacterial Effector Protein Hijacks Plant Metabolism to Support Pathogen Nutrition. *Cell Host Microbe,* 28**,** 548-557 e7.

XIN, X. F., NOMURA, K., AUNG, K., VELASQUEZ, A. C., YAO, J., BOUTROT, F., CHANG, J. H., ZIPFEL, C. & HE, S. Y. 2016. Bacteria establish an aqueous living space in plants crucial for virulence. *Nature,* 539**,** 524-+.

XU, C., LUO, F. & HOCHHOLDINGER, F. 2016. LOB Domain Proteins: Beyond Lateral Organ Boundaries. *Trends in Plant Science,* 21**,** 159-167.

XU, Q., CHEN, L. L., RUAN, X. A., CHEN, D. J., ZHU, A. D., CHEN, C. L., BERTRAND, D., JIAO, W. B., HAO, B. H., LYON, M. P., CHEN, J. J., GAO, S., XING, F., LAN, H., CHANG, J. W., GE, X. H., LEI, Y., HU, Q., MIAO, Y., WANG, L., XIAO, S. X., BISWAS, M. K., ZENG, W. F., GUO, F., CAO, H. B., YANG, X. M., XU, X. W., CHENG, Y. J., XU, J., LIU, J. H., LUO, O. J., TANG, Z. H., GUO, W. W., KUANG, H. H., ZHANG, H. Y., ROOSE, M. L., NAGARAJAN, N., DENG, X. X. & RUAN, Y. J. 2013. The draft genome of sweet orange (Citrus sinensis). *Nature Genetics,* 45**,** 59-U92.

YAMADA, K., SAIJO, Y., NAKAGAMI, H. & TAKANO, Y. 2016. Regulation of sugar transporter activity for antibacterial defense in Arabidopsis. *Science,* 354**,** 1427-1430.

YAN, Q. & WANG, N. 2012. High-Throughput Screening and Analysis of Genes of Xanthomonas citri subsp citri Involved in Citrus Canker Symptom Development. *Molecular Plant-Microbe Interactions,* 25**,** 69-84.

YANG, B., SUGIO, A. & WHITE, F. F. 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 10503-10508.

YANG, Y. N., DEFEYTER, R. & GABRIEL, D. W. 1994. Host-Specific Symptoms and Increased Release of Xanthomonas-Citri and Xanthomonas-Campestris Pv Malvacearum from Leaves Are Determined by the 102-Bp Tandem Repeats of Ptha and Avrb6, Respectively. *Molecular Plant-Microbe Interactions,* 7**,** 345-355.

YE, L. L., WANG, X., LYU, M., SILIGATO, R., ESWARAN, G., VAINIO, L., BLOMSTER, T., ZHANG, J. & MAHONEN, A. P. 2021. Cytokinins initiate secondary growth in the Arabidopsis root through a set of LBD genes. *Current Biology,* 31**,** 3365-+.

YE, X. D., AL-BABILI, S., KLOTI, A., ZHANG, J., LUCCA, P., BEYER, P. & POTRYKUS, I. 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science,* 287**,** 303-305.

YORDANOV, Y. S., REGAN, S. & BUSOV, V. 2010. Members of the LATERAL ORGAN BOUNDARIES DOMAIN transcription factor family are involved in the regulation of secondary growth in Populus. *Plant Cell,* 22**,** 3662-77.

YU, G., WANG, L. G. & HE, Q. Y. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics,* 31**,** 2382-3.

YU, G. C., WANG, L. G., HAN, Y. Y. & HE, Q. Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology,* 16**,** 284-287.

YU, Y. H., STREUBEL, J., BALZERGUE, S., CHAMPION, A., BOCH, J., KOEBNIK, R., FENG, J. X., VERDIER, V. & SZUREK, B. 2011. Colonization of Rice Leaf Blades by an African Strain of Xanthomonas oryzae pv. oryzae Depends on a New TAL Effector That Induces the Rice Nodulin-3 Os11N3 Gene. *Molecular Plant-Microbe Interactions,* 24**,** 1102-1113.

YUAN, T., LI, X., XIAO, J. & WANG, S. 2011. Characterization of Xanthomonas oryzae-responsive cis-acting element in the promoter of rice race-specific susceptibility gene Xa13. *Mol Plant,* 4**,** 300-9.

ZARATTINI, M., CORSO, M., KADOWAKI, M. A., MONCLARO, A., MAGRI, S., MILANESE, I., JOLIVET, S., DE GODOY, M. O., HERMANS, C., FAGARD, M. & CANNELLA, D. 2021. LPMO-oxidized cellulose oligosaccharides evoke immunity in Arabidopsis conferring resistance towards necrotrophic fungus B. cinerea. *Communications Biology,* 4.

ZHANG, J., TAPIA, J., HU, Y., JONES, J., WANG, N., LIU, S. & FRANK, W. 2016. Homologs of CsLOB1 can function as disease susceptibility genes in citrus canker. *Phytopathology,* 106**,** 147-148.

ZHANG, J. L., HUGUET-TAPIA, J. C., HU, Y., JONES, J., WANG, N., LIU, S. Z. & WHITE, F. F. 2017a. Homologues of CsLOB1 in citrus function as disease susceptibility genes in citrus canker. *Molecular Plant Pathology,* 18**,** 798-810.

ZHANG, Y., LIU, L. L., GUO, S. J., SONG, J. H., ZHU, C. X., YUE, Z. W., WEI, W. S. & YI, C. Q. 2017b. Deciphering TAL effectors for 5-methylcytosine and 5-hydroxymethylcytosine recognition. *Nature Communications,* 8.

ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol,* 9**,** R137.

ZHANG, Y. W., LI, Z. W., MA, B., HOU, Q. C. & WAN, X. Y. 2020. Phylogeny and Functions of LOB Domain Proteins in Plants. *International Journal of Molecular Sciences,* 21.

ZHAO, M., PENG, Z., QIN, Y., TAMANG, T. M., ZHANG, L., TIAN, B., CHEN, Y., LIU, Y., ZHANG, J., LIN, G., ZHENG, H., HE, C., LV, K., KLAUS, A., MARCON, C., HOCHHOLDINGER, F., TRICK, H. N., LIU, Y., CHO, M. J., PARK, S., WEI, H., ZHENG, J., WHITE, F. F. & LIU, S. 2023. Bacterium-enabled transient gene activation by artificial transcription factors for resolving gene regulation in maize. *Plant Cell*.

ZHONG, S. L., FEI, Z. J., CHEN, Y. R., ZHENG, Y., HUANG, M. Y., VREBALOV, J., MCQUINN, R., GAPPER, N., LIU, B., XIANG, J., SHAO, Y. & GIOVANNONI, J. J. 2013. Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology,* 31**,** 154-159.

ZHOU, J. H., PENG, Z., LONG, J. Y., SOSSO, D., LIU, B., EOM, J. S., HUANG, S., LIU, S. Z., CRUZ, C. V., FROMMER, W. B., WHITE, F. F. & YANG, B. 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant Journal,* 82**,** 632-643.

ZHU, H., WANG, G. H. & QIAN, J. 2016. Transcription factors as readers and effectors of DNA methylation. *Nature Reviews Genetics,* 17**,** 551-565.

ZHU, W. G., YANG, B., CHITTOOR, J. M., JOHNSON, L. B. & WHITE, F. F. 1998. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Molecular Plant-Microbe Interactions,* 11**,** 824-832.

ZHU, X., FANG, D., LI, D., ZHANG, J., JIANG, H., GUO, L., HE, Q., ZHANG, T., MACHO, A. P., WANG, E., SHEN, Q. H., WANG, Y., ZHOU, J. M., MA, W. & QIAO, Y. 2023a. Phytophthora sojae boosts host trehalose accumulation to acquire carbon and initiate infection. *Nat Microbiol*.

ZHU, X. Y., LI, Y., XUE, C. X., LIDBURY, I., TODD, J. D., LEA-SMITH, D. J., TIAN, J., ZHANG, X. H. & LIU, J. 2023b. Deep-sea Bacteroidetes from the Mariana Trench specialize in hemicellulose and pectin degradation typically associated with terrestrial systems. *Microbiome,* 11**,** 175.

ZOU, X. P., DU, M. X., LIU, Y. N., WU, L., XU, L. Z., LONG, Q., PENG, A. H., HE, Y. R., ANDRADE, M. & CHEN, S. C. 2021. CsLOB1 regulates susceptibility to citrus canker through promoting cell proliferation in citrus. *Plant Journal,* 106**,** 1039-1057.

ZUO, W., DEPOTTER, J. R. L., STOLZE, S. C., NAKAGAMI, H. & DOEHLEMANN, G. 2023. A transcriptional activator effector of Ustilago maydis regulates hyperplasia in maize during pathogen-induced tumor formation. *Nat Commun,* 14**,** 6722.

ABBREVIATION

ΔF_{norm}	Delta normalised fluorescence
AD	activation domain
ATG	Start codon
bHLH	basic Helix-Loop-Helix
BO	Blood Orange
BSA	Bovine Serum Albumin
C9G	CsLOB1 binding motif
CAZymes	Carbohydrate-Activate enZymes
CCD camera	charge-coupled device camera
CD36	CsLOB1-dependent Down-regulated at 36hpi
CFU	Colony formed units
ChIP-seq	Chromatin immunoprecipitation sequencing
CPM	read counts per million
CU36	CsLOB1-dependent Up-regulated at 36hpi
CUT	Carbohydrate Utilisation system associated with TonB-Dependent Transporters
CWDEs	cell-wall degrading enzymes
CWI	Cell wall integrity
DAF	Days after flowering
DEGs	Differentially expressed genes
dpi	days post infection
dTALE CsLOB2	dTALE targeting CsLOB2 promoter
dTALE CsLOB3	dTALE targeting CsLOB3 promoter
dTALE CsLOB4	dTALE targeting CsLOB4 promoter
dTALEs	Designer TALEs
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EBE	Effector binding element
EMSA	Electrophoretic mobility shift assays
EXP1	Expansin 1
G6PDH	Glucose-6-phosphate dehydrogenase
GC-MS	Gas chromatography-mass spectrometry
GFP	Green fluorescent protein
GH	Glycosyl hydrolase
HA	Human influenza hemagglutinin
HIS	Histidine
hpi	hours post infection
HR	Hypersensitive response
HRP	Horseradish peroxidase
IEC-PAD	Ion exchange chromatography-pulsed amperometric detection
IgG	Immunoglobulin G
IP	Immunoprecipitated
IPTG	Isopropyl β-D-1-thiogalactopyranoside

IWF	Intercellular washing fluid
K_D	Equilibrium dissociation constant
KO	Knock out
LBD	LOB domain
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometr
LFQ	Label-free quantification
LOB	lateral organ boundary
LPMOs	Lytic polysaccharide monooxygenases
MACS2	Model-based Analysis for ChIP-seq
MBP	Maltose-binding protein
Me	*Manihot esculenta*
MME	minimal media
MRM	Multiple reaction monitoring
MST	MicroScale Thermophoresis
NADP+	Nicotinamide adenine dinucleotide phosphate
N.benthamiana	*Nicotiana benthamiana*
NLR	nucleotide-binding domain leucine-rich repeat
NLS	nuclear localisation signal
p35S	promoter of Cauliflower mosaic virus
PAMP	Pathogen-associated molecular pattern
PFA	Paraformaldehyde
Pst	*Puccinia striiformis f.* sp. *tritici*
PtXPa	*Populus tremula* x *Populus alba*
PUL	Polysaccharide utilisation locus
RG	rhamnogalacturonan
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
RPKM	Reads per kilobase per million mapped reads
RT-qPCR	reverse-transcription real-time PCR
RVD	Repeat-variable diresidue
S gene	susceptibility gene
SAAB	Selection and amplification binding
SD	Standard deviation
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SlPE	tomato pectinesterase
SlPL	tomato pecate lyase
STP	Sugar transporter protein
SUTs	Sucrose transporters
SWEETs	Sugar Will Eventually be Exported Transporters
SWO	Sweet Orange
T2A	Thosea asigna virus 2A self-cleaving peptides
T2SS	Type II secretion system
T3S	Type III secretion
T3SS	Type III secretion system

TALE	Transcription activator-like effector
TBDTs	TonB-dependent transporters
TEV	Tobacco Etch Virus
TF	transcription factor
TPM	Transcripts per million
TSS	Transcription start site
uidA or GUS	β-glucuronidase
wt	wild type
X.	*Xanthomonas*
Xcc	*Xanthomonas citri* pv*. citri*
Xcm	*X. citri* pv. *malvacearum*
Xe	*X. euvesicatoria*
Xfa	*X*. *fuscans* pv. *aurantifolii*
Xoc	*X. oryzae* pv. *oryzicola*
Xoo	*X*. *oryzae* pv. *oryzae*
Xpm	*X. phaseoli* pv*. manihotis*
XyGs	Xyloglucans

APPENDIX

1.1 List of plasmids

Plasmid	Resistance	Source
LII p35S:bHLH3-GFP	Spectinomycin	This work; Paloma Aguilera
LII p35S:bHLH6-GFP	Spectinomycin	This work; Paloma Aguilera
LII p35S:UPA20-GFP	Spectinomycin	This work; Paloma Aguilera
LII pBs3-m1-15:RUBY	Spectinomycin	This work; Trang Phan
LII pBs3-m1-33:RUBY	Spectinomycin	This work; Trang Phan
LII pBs3-mX-2x:RUBY	Spectinomycin	This work; Paloma Aguilera
LII pBs3-wt-15:RUBY	Spectinomycin	This work; Trang Phan
LII pBs3-wt-2x:RUBY	Spectinomycin	This work; Paloma Aguilera
LII pBs3-wt-33:RUBY	Spectinomycin	This work; Trang Phan
LII pCs2g20600:uidA	Spectinomycin	This work; Trang Phan
LII pCs2g20600m1-15:uidA	Spectinomycin	This work; Trang Phan
LII pCs2g20600mX-15:uidA	Spectinomycin	This work; Trang Phan
LII pCs5g20320:uidA	Spectinomycin	This work; Niels Gallas
LII pSlEXP1:uidA	Spectinomycin	This work; Trang Phan
LII pSlPE:uidA	Spectinomycin	This work; Paloma Aguilera
LII pSlPL:uidA	Spectinomycin	This work; Paloma Aguilera
pDSK602 dTALE CsLOB2	Spectinomycin	This work; Robert Morbitzer
pDSK602 dTALE CsLOB3	Spectinomycin	This work; Robert Morbitzer
pDSK602 dTALE CsLOB4.1	Spectinomycin	This work; Robert Morbitzer
pDSK602 dTALE CsLOB4.2	Spectinomycin	This work; Robert Morbitzer
pDSK602 dTALE2 (targeting CsLOB1)	Spectinomycin	This work; Robert Morbitzer
pDSK602 dTALE4 (targeting CsLOB1)	Spectinomycin	This work; Robert Morbitzer
pDSK602 pthA4	Spectinomycin	This work; Trang Phan
pET-53-DEST optimised CsLOB1	Ampicillin	This work; Trang Phan
pET-53-DEST optimised CsLOB1-MBP	Ampicillin	This work; Trang Phan
pICH47732 p35S:AS2-HA-GFP	Ampicillin	This work; Paloma Aguilera
pICH47732 p35S:AtLOB-HA-GFP	Ampicillin	This work; Paloma Aguilera
pICH47732 p35S:CsLOB1-HA-GFP	Ampicillin	This work; Robert Morbitzer
pICH47732 p35S:CsLOB2-HA-GFP	Ampicillin	This work; Paloma Aguilera
pICH47732 p35S:CsLOB3-HA-GFP	Ampicillin	This work; Paloma Aguilera
pICH47732 p35S:LBD11-HA-GFP	Ampicillin	This work; Paloma Aguilera
pICH47732 p35S:SlLOB1-HA-GFP	Ampicillin	This work; Paloma Aguilera
pSKX1 dTALE807 (targeting SlLOB1)	Gentamycin	This work; Trang Phan
pSKX1 dTALE808 (targeting SlLOB1)	Gentamycin	This work; Trang Phan
pSKX1 dTALE809 (targeting SlLOB1)	Gentamycin	This work; Trang Phan

1.2 List of primers

Name	Sequence (5' to 3')	Purpose	Reference
TPF6 LOB RT F-CsLOB1	CTCACTAACTACTACAACCCAACAGT	qPCR	Hu et al (2014)
TPR6 LOB RT R CsLOB1	GCATGGAGAAGATTGAGAAGGAGG	qPCR	Hu et al (2014)
TPF7 45080 RP F-CsLOB2	CCCTCTTTCTTCTTCTCCTTCTC	qPCR	Hu et al (2014)
TPR7 45080 RP R-CsLOB2	AAGCTGCGCATGGACTAA	qPCR	Hu et al (2014)
TPF8 47491 RP F-CsLOB3	GCCACAAGAAATCCTAGCAAAG	qPCR	Hu et al (2014)
TPR8 47491 RP R-CsLOB3	CAGCACAAGGGCTCATGATA	qPCR	Hu et al (2014)
TPF11 CsEF1aF	GTAACCAAGTCTGCTGCCAAG	qPCR	Hu et al (2014)
TPR11 CsEF1aR	GACCCAAACACCCAACACATT	qPCR	Hu et al (2014)
TPF13 7877F ExpansinB2	ACAGATTCAGCACAGAAGAGTT	qPCR	Hu et al (2014)
TPR13 7877R ExpansinB2	GAAGCAAGGTCACCGTCAC	qPCR	Hu et al (2014)
TP P69 Cs2g24050 qPCR fwd	CAAAAGAGAGAAAGTGGATGTG	qPCR	This work
TP P70 Cs2g24050 qPCR rev	GTCTCTAACACCAACCATGT	qPCR	This work
TP P71 Cs5g20320 qPCR fwd	CTTCCCAAGGACCAACAAAT	qPCR	This work
TP P72 Cs5g20320 qPCR rev	TTCTTGCCGTATTCGAGTG	qPCR	This work
TP P117qPCR Cs9g14330 fwd	CCCATCATGTTCACgtCTA	qPCR	This work
TP P118qPCR Cs9g14330 rev	GCGTgCGATGAGAGATAC	qPCR	This work
TP P121qPCR Cs9g18910 fwd	GAAAGTCAAGGGAAAGCGG	qPCR	This work
TP P122qPCR Cs9g18910 rev	AGCACCATCTCCTCTGAAG	qPCR	This work
TP_P137_FBOX_F	TTGGAAACTCTTTCGCCACT	qPCR	Mafra et al (2012)
TP_P138_FBOX_R	CAGCAACAAAATACCCGTCT	qPCR	Mafra et al (2012)
TPF4 CsLOB1 CACC	tttGGTCTCTCACCATGGAATGCAAGCACAAGATCAACG	cloning optimised CsLOB1	This work
TPR5 FLCsLOB1 AAGG	tttGGTCTCTCCTTcTAGGTCCACAGCGGTTCCCAAGC	cloning optimised CsLOB2	This work
TP_P235-CsLOB1opt_D_rev	aaaGGTCTCTCCTTGGTCCACAGCGGTTCCCA	cloning optimised CsLOB3	This work
TP_P224_MBP_E_rev	aaGGTCTCTGATTTCACTGCGCGTCTTTCAGGGCTTCATC	cloning HIS-MBP	This work
TP_P238-TEVHISMBP-addBsaI-fwd	tttGGTCTCAAAGGGAggcggtggaagtggtggcggaggTAGCgaaaatctatacttccaaggaCATCATCATCATCATCATCA TGAAGAAGGTAAACTGGTAA	cloning MBP	This work
TP_P327_Cs8g11330_neg2_f	CGCTTCTATGTTGTCACTG	ChIPqPCR	This work
TP_P328_Cs8g11330_neg2_r	TGAAAACAATCCAGAGAGGC	ChIPqPCR	This work
TP_P329_Cs8g11330_pos1_f	GCTTGTTCGTGTTTGCGA	ChIPqPCR	This work
TP_P330_Cs8g11330_pos1_r	TTTTTCTTTTTCGGTCGAGG	ChIPqPCR	This work
TP_P335_orange1.1t00187_neg2_f	TGAGCCCTCCTAATTCTT	ChIPqPCR	This work
TP_P336_orange1.1t00187_neg2_r	AAAAAATCGATAGGAGGGATGG	ChIPqPCR	This work
TP_P339_orange1.1t00187_pos2_f	AATGGAATGAGTTCAACGGT	ChIPqPCR	This work
TP_P340_orange1.1t00187_pos2_r	ATGACCATGAGAGCGCGA	ChIPqPCR	This work
TP_E302_wt_fwd	CAAGTGCAAATTGCGCCGGGTTTCACGGCTTTA	EMSA	This work
TP_E303_wt_rev	TAAAGCCGTGAAACCCGGCGCAATTTGCACTTG	EMSA	This work
TP_E305_m1_fwd	CAAGTGCAAATTGCATTAAATTTCACGGCTTTA	EMSA	This work
TP_E306_m1_rev	TAAAGCCGTGAAATTTAATGCAATTTGCACTTG	EMSA	This work
TP_E353_m2_fwd	TAAAGCTGTGAAACCCAGCGCAATTTGCACTTG	EMSA	This work
TP_E354_m2_rev	CAAGTGCAAATTGCGCTGGGTTTCACAGCTTTA	EMSA	This work
TP_E343_m3_fwd	TAAAGCTGTGAAACCCGGCGCAATTTGCACTTG	EMSA	This work
TP_E344_m3_rev	CAAGTGCAAATTGCGCCGGGTTTCACAGCTTTA	EMSA	This work
TP_E347_m4_fwd	TAAAGCCGTGAAACCCAGCGCAATTTGCACTTG	EMSA	This work
TP_E348_m4_rev	CAAGTGCAAATTGCGCTGGGTTTCACGGCTTTA	EMSA	This work
TP_E351_m5_fwd	TAAAGCCGTGGAACCCGGCGCAATTTGCACTTG	EMSA	This work
TP_E352_m5_rev	CAAGTGCAAATTGCGCCGGGTTCCACGGCTTTA	EMSA	This work
TP_E357_m6_fwd	TAAAGCCGTGGCGCCCGGCGCAATTTGCACTTG	EMSA	This work
TP_E358_m6_rev	CAAGTGCAAATTGCGCCGGGCGCCACGGCTTTA	EMSA	This work
TP_E389_opt_f	TAAAGCCGCAATTCACGCCGCAATTTGCACTTG	EMSA	This work
TP_E390_opt_r	CAAGTGCAAATTGCGGCGTGAATTGCGGCTTTA	EMSA	This work
TP_P499_Z1_f	TAAAGCGGCGGCGGCGGCGCAATTTGCACTTG	EMSA	This work
TP_P500_Z1_r	CAAGTGCAAATTGCGCCGCCGGCGGCTTTA	EMSA	This work
TP_P501_Z2_f	TAAAGGCCGCGGCGGCAATTTGCACTTG	EMSA	This work
TP_P502_Z2_r	CAAGTGCAAATTGCCGCCGCGGCCTTTA	EMSA	This work
TP_P505_un_f	CGTTTTTCCATTGCGGCGGCGGCGACGGTGTC	EMSA	This work
TP_P506_un_r	GACACCGTCGCCGCCGCCGCAATGGAAAAACG	EMSA	This work
TP_E513_d1_F	TTAAAGCCGCATTCACGCCGCAATTTGCACTTG	EMSA	This work
TP_E514_d1_R	CAAGTGCAAATTGCGGCGTGAATGCGGCTTTAA	EMSA	This work
TP_E525_in_F	TAAAGCCGCAAATTCACGCCGCAATTTGCACTTG	EMSA	This work
TP_E526_in_R	CAAGTGCAAATTGCGGCGTGAATTTGCGGCTTTA	EMSA	This work
TP_E341_flk_fwd	ATTTTCCGTGAAACCCGGCGATTCCCTATAAAT	EMSA	This work
TP_E342_flk_rev	ATTTATAGGGAATCGCCGGGTTTCACGGAAAAT	EMSA	This work
CBP3-Cand 0-BasI-A-f	ttGGTCTCTGCGGGAACCAGAAAGGGTAAAAAG	cloning pCs2g20600	This work
TP P43 cs2g20600pro rev	GGTCTCACAGACATTCCCTTCAATCGCAGC	cloning pCs2g20601	This work
CB13_pC0_Posphate_rev	GCAATTTGCACTTGCAAATG	SDM to mutate CsLOB1 binding site	This work
CB14_pC0_half_mutated_fwd	attaaaTTTCACGGCTTTACGAGAGTAAATTGAATCATTATAT	SDM to mutate CsLOB1 binding site	This work
P605-Ecand0WTmotif-CACC-f	CACCCAAGTGCAAATTGCGCCGGGTTTCACGGCTTTA	cloning for *in-planta* gain of function	This work
P606-Ecand0WTmotif-CCAG-r	CCAGTAAAGCCGTGAAACCCGGCGCAATTTGCACTTG	cloning for *in-planta* gain of function	This work
P607-cand0WTmotif-CACC-f	CACCCGCCGGGTTTCACGG	cloning for *in-planta* gain of function	This work
P608-cand0WTmotif-CCAG-r	CCAGCCGTGAAACCCGGCG	cloning for *in-planta* gain of function	This work
P609-Ecand0m0motif-CACC-f	CACCCAAGTGCAAATTGCATTAAATTTCACGGCTTTA	cloning for *in-planta* gain of function	This work
P610-Ecand0m0motif-CCAG-r	CCAGTAAAGCCGTGAAATTTAATGCAATTTGCACTTG	cloning for *in-planta* gain of function	This work
P611-cand0m0motif-CACC-f	CACCCATTAAATTTCACGG	cloning for *in-planta* gain of function	This work
P612-cand0m0motif-CCAG-r	CCAGCCGTGAAATTTAATG	cloning for *in-planta* gain of function	This work
P623_CACC_E0WT_ATTG_f	CGTCTCACACCCAAGTGCAAATTGCGCCGGGTTTCACGGCTTTACGAGAGTAAATTGTGAGACG	cloning 2x CsLOB1 motifs	This work
P624_E0WT_ATTG_1_r	CGTCTCACAATTTACTCTCGTAAAGCCGTGAAACCCGGCGCAATTTGCACTTGGGTGTGAGACG	cloning 2x CsLOB1 motifs	This work
P625_ATTG_E0WT__2_f	CGTCTCAATTGAATCATTATATATTCTCATAACCAATGACCTTTTTCTATCAACAATAGATGAGACG	cloning 2x CsLOB1 motifs	This work
P626_TAGA_E0WT__2_r	CGTCTCATCTATTGTTGATAGAAAAAGGTCATTGGTTATGAGAATATATAATGATTCAATTGAGACG	cloning 2x CsLOB1 motifs	This work
P627_TAGA_E0WT__3_f	CGTCTCATAGATGTGAGATGAGATGGAGGCGAGCAGCAAGTGCAAATTGCGCCGGGTTTCACGGCTTTACTGGTGAGACG	cloning 2x CsLOB1 motifs	This work
P628_CTGG_E0WT__3_r	CGTCTCACCAGTAAAGCCGTGAAACCCGGCGCAATTTGCACTTGCTGCTCGCCTCCATCTCATCTCACATCTATGAGACG	cloning 2x CsLOB1 motifs	This work
P629_CACC_E0mX_1_f	CGTCTCACACCCAAGTGCAAATTGCGTTAAATTTTATAATTTTACGAGAGTAAATTGTGAGACG	cloning 2x CsLOB1 mutated motifs	This work
P630_E0mX_ATTG_1_r	CGTCTCACAATTTACTCTCGTAAAATTATAAAATTTAACGCAATTTGCACTTGGGTGTGAGACG	cloning 2x CsLOB1 mutated motifs	This work
P631_TAGA_E0mX__3_f	CGTCTCATAGATGTGAGATGAGATGGAGGCGAGCAGCAAGTGCAAATTGCGTTAAATTTTATAATTTTACTGGTGAGACG	cloning 2x CsLOB1 mutated motifs	This work
P632_CTGG_E0mX__3_r	CGTCTCACCAGTAAAATTATAAAATTTAACGCAATTTGCACTTGCTGCTCGCCTCCATCTCATCTCACATCTATGAGACG	cloning 2x CsLOB1 mutated motifs	This work
TP_P493_LOB2exBsaI_f	tttGGTCTCCAATGCACACTACATTTCCCC	cloning CsLOB2 CDS	This work
P693_LOB2ex2_BsaIpICH_r	aaaGGTCTCaCACCAGTCCAAAGAGCATCCCAAG	cloning CsLOB2 CDS	This work
TP_P495_LOB3ex1_BsaI_f	tttGGTCTCCAATGCTAAAAATGGAGAACTATGA	cloning CsLOB3 CDS	This work
P694_LOB3ex2_BsaIpICH_r	aaaGGTCTCaCACCTGTCCAAAGAGCTGGTTCCC	cloning CsLOB3 CDS	This work
NG56 GCGG BsaI	ttGGTCTCTGCGGGGGTAATTTTGGCTTTCGGCCTG	cloning pCs5g20320	This work
NG57 TCTG BsaI	aaGGTCTCACAGATTTTTTTTGGTACTTAGCTGCC	cloning pCs5g20320	This work
P568_SlEXP1_A-bsaI-f	ttGGTCTCTGCGGGGCCACATCCATGATTTGT	cloning pSlEXP1	This work
P569_SlEXP1_BbsaI-r	aaGGTCTCACAGATCTTAAGATTTAATGGAGTTGA	cloning pSlEXP1	This work
P596_SlLOB1_AATGbsaI_TACGbpiI	ttGAAGACATTACGGTCTCAAATGGAATCTACAATATCTTCATCT	cloning SlLOB1 CDS	This work
P597_SlLOB1_GGTGbsaI_TCTGbpiI	aaGAAGACAAcagaGGTCTCaCACCTGTCCATAGGGTTTCCCAT	cloning SlLOB1 CDS	This work
P594_SlLOB1ex2_qPCRf	GCAATGTCAACAAGCAAATC	qPCR	This work
P595_SlLOB1ex2_qPCRr	TCCCATGATCCAATGTTATC	qPCR	This work
PA_1_fwr_PL_BsaI_A	TTTGGTCTCTGCGGTGATTCATTAGCCAAACCCCA	cloning pSIPL	This work
PA_2_rv_PL_BsaI_B	AAAGGTCTCACAGATTGAAATCCCCTTTGTGTTGTT	cloning pSIPL	This work
PA_3_fwd_PE_BsaI_A	TTTGGTCTCTGCGGTTTTGTTGCAGTAGGTCTTC	cloning pSIPE	This work
PA_4_rv_PE_BsaI_B	AAAGGTCTCACAGATGCTTTATCTTTGTGATATTGG	cloning pSIPE	This work

1.3 Important gene sequences (grey: coding sequence; bold: Start or Stop codon; yellow: EMSA probe)

>CsLOB1 mRNA

AAGCAGCTCCTCCTCATCCCTTACTGTCTTTGCTTTCTCACTAACTACTACAACCCAACAGTTTTCTTCTCTCAAAA**ATG**GAA
TGCAAACACAAAATTAATGTAGCAATCCCAATCACTAATATGAAGAACACTCAATTCTCATCTCCATCTACTTTCTCTACTTC
TCCTCCTTCTCAATCTTCTCCACGCTTCCCTTCTCCTAATCATCAACAATTGTCTTCTCCAGAATCTTCTCCAAGCTTTAAAG
CTTCTCCTTCACAATCCTCTCCAAATCTTGCAGCTCCCCTCTCTCCGCCGCCTATAGTTCTTAGCCCTTGTGCTGCTTGCAAA
ATCCTCCGCCGCAGATGCGTCGAGAAATGTGTTTTAGCTCCATATTTTCCACCAACCGAACCATACAAGTTCACCATTGCTCA
TAGAGTCTTCGGTGCTAGCAATATCATCAAGTTCTTGCAGGAACTGCCAGAATCTCAACGAGCAGATGCAGTGAGCAGCATGG
TCTATGAAGCAAGTGCCAGAATCCGGGATCCTGTTTACGGCTGCGCCGGGGCTATTTGCCATCTCCAGAAACAAGTCAGTGAG
CTTCAGGCTCAGTTAGCCAAGGCACAGGCTGAGCTTGTCACCATGGAAAGCCAGCAACGCAATTTAATAACTCTAATTTGCAT
GGAAATGGCACAATCTCAAGAACAAGTCTTGCAGCAGCAGCAGCAGCAGCAACAGTTCATGGATACTAGCTGTTTTTTGG
ATGACAATGGTATTGGATCAGCTTGGGAGCCTCTGTGGACA**TGA**TCAAGAGAAATTAAAGCAAGATTGTTGAAATTTTAACCT
TTTAAGAGATTATTTACATAAAGCTAAACATACTTAATTATAAAAGTTTCTGATCAATAATTAAAGTTATTTGCTGCGCGGTA
GATGGGAGTGATTATTATGTGCTTTAATTTTCATTAGTCTTGTTGACAAAAAGGAATCTTTGAACCATCTGGAGAAGTCCTTT
GTTAACGGTTCGAGATTAATTATTAGTTTATCTTTATTTACATTTAGTGAAATTTTGTTTTTAACTAATTTTATAGACATAAA
TAACCCAACCAAGATGGGAATTCAGTGCT

>CsLOB1 optimised codon for E. coli expression

ATGGAATGCAAGCACAAGATCAACGTGGCGATTCCCATCACGAACATGAAGAACACGCAGTTCAGCAGCCCTTCCACCTTCTC
GACCAGTCCGCCCAGCCAGAGCAGTCCCCGCTTCCCCAGTCCGAACCATCAGCAGCTGAGCAGCCCGGAATCGAGTCCGTCTT
TCAAAGCCTCCCCGTCGCAGTCCTCTCCGAATCTGGCTGCGCCGTTGAGCCCGCCGCCGATCGTTCTGAGCCCGTGTGCAGCG
TGCAAAATCCTCCGCCGTCGCTGCGTGGAGAAATGCGTGTTGGCCCCGTATTTTCCGCCGACCGAACCGTACAAGTTCACCAT
TGCGCATCGGGTGTTTGGGGCCAGCAACATCATCAAGTTCCTGCAGGAGTTGCCGGAATCCCAACGCGCTGATGCGGTTAGCT
CCATGGTCTATGAGGCGAGCGCCCGCATTCGTGACCCGGTCTATGGCTGTGCGGGCGCAATTTGCCACCTGCAAAAGCAGGTC
AGCGAGCTGCAGGCGCAGCTGGCCAAAGCCCAGGCGGAACTGGTGACCATGGAGAGCCAGCAGCGGAATCTGATCACGCTGAT
CTGCATGGAGATGGCCCAATCCCAGGAACAAGTGCTGCAACAGCAACAACAGCAGCAGCAACAGTTCATGGATACCAGCTGCT
TTCTGGACGACAATGGCATTGGCAGCGCTTGGGAACCGCTGTGGACC

>CsLOB2 mRNA

GCATTATTATATTATCAATTTCATGACC**ATG**CACACTACATTTCCCCCTCTTTCTTCTTCTCCTTCTCCTTCTTTTCAATCTT
CTCCAAGCATTAATGCCTCTCCTTCAAAATCTTCTCCAAATCTTGCTGCCCCACCGCCTATTGTTCTTAGTCCATGCGCAGCT
TGCAAAAGCCTTCGCCGGAGATGCGACGAGAAGTGTGTTTTAGCTCCATATTTTCCACCAACTGAACCACAAAATTTCATCAT
TGTTCATAGAGTCTTCGGAGCTAGCAACATCATCAAGTGCTTGCAGGGACTGCCAGAGTGTCAACGATCAGATGCAGTGAGCA
GCATGGTCTACGAAGCAAACGCTAGAATCCGAAATCCGGTGCATGGCTGCGCGGGTGCAATTAGTCAACTCCAGAAACAAGTG
ATTAAGCTTCAAGCAGAGTTGGCCAAGGCACAAGCCGAGACGGTCAGCATGCAGTGCCAACGAGACAATTTGGTTGCCCTAAT
TTGCAAGGAAATGACGACACAATTTCCTCAAGAAACCATGAATAGGGTCTTGCCCCAACAGCAATTCAATGACGACGACGCCA
CCACATGTTATTTAGATGACAAGGATTTTGCCTCCACTTGGGATGCTCTTTGGACT**TAA**AATTAGAGTTAATTTACATGAAGT
TGATGATTGATGAAGAAGAAGAAGAAGAGAGATATTAGTAATAATGAATTAGACAGAAAAGTTTCATTTGAAGTATACTATGG
GTTGTA

>CsLOB3 mRNA

CCTCTCTCTCAGGTCTCACTCTCAGCTAATCCCTCTTAAATAAAAATATATCAATATTAATAATTACCAAACCAT**ATG**CTAAA
AATGGAGAACTATGAGGCCGCTGCCACAAGAAATCCTAGCAAAGTTACCAGTAGCCGTGCTGGTTCGTCATCTCCTCCTCCAA
TATCAGCTAATTCTTCTGCTCCGCCGCCAGTTATCATGAGCCCTTGTGCTGCATGCAAGATTCTGAGACGTCGATGTGCTGAC
AAATGTGTTTTGGCTCCTTATTTTCCTCCAACCGAGCCTGCCAAGTTCACTATTGCTCATCGTGTTTTCGGTGCCAGCAATAT
CATCAAGTTTCTTCAGGAACTTCCAGAGTCTCAGAGAGCAGATGCTGTGAGCAGCATGGTTTACGAAGCAAGTGCAAGAATCC
GGGACCCAGTTTACGGATGTGCGGGAGCAATTTGTCAGCTACAGAAGCAAGTGAGTGAGCTGCAGGCACAGCTAGCTAAAGCA
CAAGCCGAGGTTGTGAACATGCAATGCCAGCAAGCAAACCTTGTGGCCTTGCTCTACAAGGAAATGGGAAAATCGCCGCAGCC
CAATTCGCCACAATCTGTTGACCACTTTATTACAAGCCCAGAAAGCCCTGAAGCAAATCCCTGCAGCTTCTTTGAGGATAACA
ACTTATCTGGGTCATTATGGGAACCAGCTCTTTGGACA**TGA**TCTTAATTAAAGTATTAGTTAGTTATTTAATTATAATTAGAT
AATAATTGAGAACCATATTTAATGGTACGCGTCCGACGAAATTCTCCTAACTCAAAGCCAATGGGAAGGAGGAGTTGGAGGAA
ATTAAACGAGCAGTCTCTAGCTAGCGGAAAATCCTCAGAGTGTTCCAAGGAAACATTGCTCTTTTTCCATTTACTCCTTGTTA
TGTGTGTAACTTAACTAGTTATGATCACATGTATTTACATGTGGTAAATAGTAAATCGATAAGCAGCAGTGCTACCGGTGCTC
CTAAGTAGCTAGACTGATTTAGTCAATATTTAACTTAGTTTTGTTGGAGGAAACCGAAATTTGATTAGGATCTAACTAATCTC
TTTGTTAGCTTGGCGAAGTAGATTATGGGTCTTTTCTTTCTTTTCTTCCTTCAATAAATATATGTACGGTTGAAAAAGGGGTT
AATTTTTTTGTTTAAGTCCAATATTTGTAAGCTTTATTATTGAATCAGAAACTACTTTTTGGTCACC

>CsLOB4 mRNA

TACTATAAATACACAAACCCCCTATCTCATTTCTCTTCAAGCCCTTAAAAGCTAGCTCTGTTGTATCTCTATAGCCATCTTCT
CCTTCCTAGAAGCACTTTTTTTTTTTTTTCTTTTTTCACAGAGGAAA**ATG**GGTGGAAATTCTCCATGTGCCTCTTGCAAGTTGC
TGCGTCGCCGTTGCGCCAAAGACTGCATCTTTGCTCCTTACTTCCCTCCTGATGATCCCCACAAGTTTGCCATTGTTCATAAG
GTCTTTGGTGCTAGCAACGTTAGCAAAATGTTGCAGGAGCTTCCGGTGCATCAAAGGGCAGATGCAGTGAGCAGTTTGGTTTA
TGAAGCAAACGCCAGAGTGAGAGACCCAGTATATGGCTGTGTTGGTGCCATATCTTTTTTGCAAAACCAAGTGTCTCAGCTGC
AAATGCAACTAGCAGTGGCTGAAGCAGAGATATTGTGCATCCAGATGCAGCAAGAACCGATGATGCCAACTCCACAAATAGAC
CCAGATGACAAGTCGTTTCTTCTCCAAAATAACCTCCCCCATCAATACCTCAATTTTACTTCATCGAGCAATAATGTAATTCA
TGACTCCCTCAAGAGAGAGAGCATCTTTGGACATGACATGGTTTCT**TAA**TTAAAATCTGCCCTCTGTACGTTTTTCAA

>SlLOB1 mRNA

AACACACCAAGTAAGCACTATTTTCTTTTTATACACATTTCTACAAAGATCAAAAAAATTGAAAAAAAA**ATG**GAATCTACAAT
ATCTTCATCTTCATCTCGATCCGTATCGCCTTCGTCTTCATCTTCACCTAACTCACCACCACCCCCGTCAATGACAGTGGTGG

TGGTGAGTCCTTGCGCCGCCTGTAAAATATTACGGCGGCGCTGTGCTGAGAAATGTGTGTTGGCACCTTATTTTCCTCCAAAT
GATCCTATTAAGTTCACTACGGCTCATCGTGTGTTTGGTGCTAGCAATATTATCAAGTTTTTACAGGAATTACCAGAATCTCA
AAGGGCAGATGCTGTAAGCAGCATGGTATATGAAGCAAATGCAAGACTAAGAGATCCAGTATATGGTTGTGCTGGTTCAATTT
GTCAATTACAAAAGCAAGTTAGTGATCTTCAAGCACAATTAGCTAAAGCACAAGCTGAAATTGTCAACATGCAATGTCAACAA
GCAAATCTCATGGCACTAATTTGTATGGAAATGGGCCAATCTAATCCACAACCAATATCACCACCACAACAATCTTTAGAAAA
TTTCCCTATGAATTACTTAGATGATAACATTGGATCATGGGAAACCCTATGGACA**TGA**
TCGAGTCGATTTAAGGAAAAAACGCTTTCTATACCAATAGATTCATCAGTTCTAACTCAGGTCCTAGAAAATTTCATCGAATA
CTCAAGATAAATTTTGAGTAAATTATATAGTATTTAAAAAAAAAAAAAGAATCGCGAATAAAATTTATGTAATATTATTTGAT
TTCTTAGTGAAAGGAAAAATGAAAATTGATGCAAATCTTTAAAGATTTGTGTCAATGTCTAGTCTAGTAATTTTTATTTTTAT
ATATATATATATATATAATATGATTCTAGTCAGAAAAA

>LBD11 CDS synthesised

ATGCTAAAGATGGAGATTAACGGTGGGGTCGCTACACCTACTGCTTCCGCCGTCGCCAAGGTGACAGAAACCACCACTCCCGT
TAACTCTCCTTCTCCCACTTCTTCACCTCCGCCGCCTCCTTCGCCGCAACAGCCGCCACAACCACCGGTTGTACTAAGCCCTT
GTGCGGCTTGCAAGATTTTGCGGCGGCGGTGCGCCGATAAATGTGTTTTGGCACCGTACTTTCCTCCTACGGACCCGGCGAAG
TTCACAATCGCTCACCGTGTCTTTGGAGCTAGCAACATTATTAAGTTCTTGCAGGAACTTCCAGAATCGCAAAGAACAGATGC
TGTTAATAGCATGGTCTATGAAGCCGGAGCTAGGATGAGAGATCCGGTTTACGGATGTGCGGGTGCAATTTATCATTTGCAGA
GACAAGTGAGTGAGCTTCAAGCACAACTTGCAAAAACTCAAGTAGAGTTAGTAGGCATGCAACTCCAAAGATCAAGTCTACTA
GAATTGATATATAATATGGAGCAAACAAAGTTGTCAGTACAAGAGCAAGGACAACAGAAGATGTCCTTTGAGAGTTCATTTGA
AAGCGGCGATGAGTTCATTAGTAGCCCGGACGAAGAGAGCAATGATTTGGGGTTCCTTGAGGACAACAACAACAACAATAATT
CATCAATGTCGTGGTGGGATCCTCTTTGGACA

>AtLOB synthesised

ATGGCGTCGTCATCAAACTCATACAACTCACCATGCGCGGCGTGCAAGTTTCTTCGCCGGAAATGCATGCCGGGATGCATATT
CGCGCCATATTTTCCACCGGAGGAGCCACACAAATTCGCCAACGTCCACAAAATCTTTGGAGCAAGCAACGTCACGAAGCTCC
TCAACGAGCTCCTCCCTCACCAACGTGAGGATGCGGTCAACTCCTTAGCCTATGAGGCCGAGGCACGTGTCCGTGACCCCGTC
TATGGCTGCGTCGGAGCCATCTCTTATCTCCAGAGACAAGTCCATAGGCTTCAGAAGGAGCTCGACGCAGCTAATGCTGACTT
GGCACATTATGGTTTGTCCACATCAGCCGCCGGAGCACCTGGCAACGTCGTGGACTTGGTTTTTCAGCCTCAGCCGCTTCCGT
CGCAGCAACTGCCTCCCTTGAACCCGGTTTATAGGCTTTCGGGGGCAAGTCCGGTGATGAATCAGATGCCACGTGGCACCGGA
GGGTCGTATGGGACTTTTCTTCCTTGGAACAATGGTCATGATCAGCAAGGAGGTAACATG

>AS2 CDS synthesised

ATGGCATCTTCTTCAACAAACTCACCATGCGCCGCTTGCAAATTCCTCCGGCGAAAATGTCAACCGGAATGTGTATTCGCGCC
CTATTTCCCACCGGACCAGCCACAAAAATTCGCAAACGTTCACAAAGTGTTTGGAGCAAGTAACGTGACAAAGCTCCTCAACG
AGCTTCACCCTTCACAACGTGAGGATGCCGTGAACTCTTTGGCCTATGAAGCCGACATGCGCCTCCGTGACCCTGTCTACGGC
TGCGTCGGCGTCATCTCTCTCCTCCAACATCAGCTTCGTCAGCTTCAGATAGATCTCAGCTGTGCTAAATCTGAGCTCTCTAA
GTACCAAAGCCTCGGTATCCTCGCCGCCACTCATCAGAGTCTTGGCATCAACTTACTCGCCGGAGCAGCAGATGGAACAGCCA
CCGCCGTGCGTGATCATTATCACCACCACCAGTTTTTTCCTAGAGAACAAATGTTTGGTGGCTTGGATGTTCCGGCCGGTAAC
AACTACGACGGTGGGATTCTTGCCATTGGACAGATCACTCAGTTTCAGCAGCCGAGAGCCGCCGCTGGAGATGATGGTCGCCG
TACTGTTGATCCGTCT

>UPA20 CDS synthesised

ATGTCTACTTTTTCATCATACCAATTACAACAACACAATACCCCTTCTCTTCTTGATTCAGTATTTTTGCCAAGTTCACCTCT
AAAAATGTCTGGCTTTTTTGAGGAACCAAACAATTCTTGTGTTGTCCAAAATTGTTTTTCTCAATTTTACCAACCAGAATTCC
CTACCAATAATGTTAATGTTAATGAAAATAGCTATTGCCTTGACCAAAACACAAATGTTACTAACAAAAGTAGTAGCAGTATA
AGCTTGGATATGGACTCTTCTTCTGTTACTGATAAAATGGAAAGTGGAAATAATAAGCCTAATGTTACTAGTCCTATGGACAA
GAAAAGAAAATCCAGAGAAGGGTCTTCTTCTATGACTTCTGCAAATTCTAAGAATGTAACACAAGGTGATAATGGGAAAAAGA
ACAAGAGCAATAGCAAATTAGTAGCCAAAGATGAGAAAAAAGCTAATGAAGAGGCACCAACAGGCTACATTCATGTTAGAGCA
AGAAGAGGCCAAGCAACTGACAGCCATAGTCTTGCTGAAAGGGTGAGGAGAGAGAAAATAAGTGAAAGGATGAAAATGCTGCA
ATCTCTTGTTCCTGGTTGTGACAAGGTAACTGGGAAGGCCCTTATGTTGGATGAGATAATCAATTATGTCCAATCTTTGCAAA
ATCAAGTTGAGAATTTGAGTGGCTTGGAGACACAACTACCAAATATTCAGCAAGCTAGCACTACTACATCACAAGCAGCTGAA
GTTATTCCTAACACTAATAGTGGCTATCCTTTTTTGGATAATTCAGCATCACTTATGTTTCAACAAGCCCATTTCCCTAATTC
CATTCATCAGGGAATTGGACAGCTCTTATGGGGTGCAGAAGACCAAAGACAAAAAATAATAAATCAGTCAGGATTCAGCAACA
ACTTTTGTTCTTTCCAT

>bHLH3 CDS synthesised

ATGGCTGCTTTTTCATCACACCAATTACAACACAATAACCCATTTCTTCTTGATTCAGTATTTTTGCCAACTTCTCCTATTAA
AATGTCTGGCTTTTTTGAGGAACCAAACAATTCTTGTATAGTACAACAGTTTTACCAACAAGAATTCCCTTCCAATTTAATTT
CTCATGAAAATAGCTTTTGCCTTGACCCTAAAAGTAGCAGCAGTATAAGCTTAGATATGGATGCTTCCTCTGTTACTGATAAA
ATTGAAAGTGGAATTAATAATAATAAGGCTAATGTTAGTCCTTTGGATAAGAAAAGAAAATCTAGTGAAGGGTCTTCTTCTAT
GACTTCTGCTCATTCTAAGAATGAGAAACAGGGTGATAATGGGAAAAAGAAGAAAATTATCAGCAAGTTAGTAGCCAAAGATG
AAAAGAAAGCTAATGAAGAAGCACCAACAGGGTACATTCATGTTAGAGCAAGAAGGGGCCAGGCAACTGATAGCCATAGTCTT
GCTGAAAGGGTGAGGAGAGAGAAAATAAGTGAAAGGATGAAGATATTGCAATCTCTTGTTCCTGGTTGTGACAAGGTAACTGG
AAAGGCCCTTATGTTGGATGAGATAATTAATTATGTCCAATCTTTGCAAAACCAAGTTGAGTTTCTCTCAATGAAACTTACTT
CTTTGAATCCAATGTACTATGACTTTGGAATGGACTTAGATGCACTCATGGTCAGACCTGATGACCAGAGTTTAAGTGGCTTG
GAGACACAAATGGCAAATATTCAGCAAGGTAGCACAACTACTACATCACAGGCAGCTGAAGTTATTGCTAACACTAATAGTGG
CTACCAATTTTTGGATAATTCAACATCACTCATGTTTCAACAATCCCATTTCCCTAATTCTATTCCTCAGGGTATTGGACAGC
TCTTATGGGGTGCAGATGAGCAAACACAAAAAATAATAAATCAGTCTGGATTTAGCAACAACTTTTGTTCTTTCCAT

>bHLH6 CDS synthesised

ATGGAAGCTAATAGTAACTCTTTTCATGTAGATTCTGTTTTTCATGTGCCCATTAAGATGTCTGGTTTTTTTGAGGAACCAAA
CAATAATATAACAAGTAGTAGTACACTACCAAATTGTGTTTCTCAATTTTATTTGCAAGAGCTTTCTGTCAATATGAGTAATA
ATGTTCATGAAATTAGCCATAATGAACCTTCTCATGTGACAAACAAAACCAATTCTTCCTCCCTCTGCTCTACTCAATCTAAG
AATGTAAGAGACGGTGATGATGGGAAAGGGCAAAAGAAAAGAAATGGTAATGTAAAAAGAGAGAAAAAAACAAAGGAAAATAA

GAAGAAAGCCCCTGAAGAGGCCCCTACAGGATATGTTCATGTTAGAGCAAGGAGGGGCCAGGCAACTGACAGTCACAGTCTTG
CTGAAAGGGTGAGGAGAGAGAAAATAAGTGAAAGGATGAAGATATTGCAAGCACTTGTTCCAGGTTGTGACAAGGTAACTGGG
AAGGCCCTTATGTTGGATGAGATAATCAACTATGTCCAGTCTTTGCAAAACCAAGTTGAGTTTTTATCCATGAAGCTTGCTTC
TTTGAACCCCATGTACTATGACTTTGGCATGGACTTAGATGCTCTCATGGTCAAACCTGATCAGAGTTGGAGTGGCTTGGAAG
GACCATTATTAGAGAACACAACTAGTAACTACCCTCACTTGGATAGTTCAACATCACTTATGTTTCAACAATTGCATCTACCA
AATTCCGTTTCTCAGGGTAGTGGACATGTATTATGGAGTGTAGATGACCAAAGACAAAAGATGATTATTAATCATTCAGAATT
GATCAGTAATAACAACAACTTGAGTGTCCCTTTCCAT
>pCs2g20600
GAACCAGAAAGGGTAAAAAGTTCAGCTGTCATAATTTATAACCAACGACCCTGGCAGTAACAAATTTAATTCAGTAATGTGCT
TCCTGTGTGATTCGCTTTGAGGGAGCCCTCATTTCGTTAAATTAAGCTGCCATAAAATGGGAAACAACCCTAAATTATCTACC
CAACTGGGACACATTTGCAAGTGCAAATTGCGCCGGGTTTCACGGCTTTACGAGAGTAAATTGAATCATTATATATTCTCATA
ACCAATGACCTTTTTCTATCAACAATAGATGTGAGATGAGATGGAGGCGAGCAGCTTTATATTACAGGGCCTTTTTCTTTTCT
TTTTTTTTACCCCCCCTTGGTTACAAAAAGTGCTTCTGACGCAACATTGATGGGGTCCTAGTAGGGTTGTTCATGCTGATACC
GTTCGATCTTCAATGTTTAGTTAGTTAGTTGACATGAATTATTTTTCTGGGTATTTCATTTCATTACAAACAAAACAATAAAT
TAAGTTGATTAGGCAAGGGGATAGTGCAACAACAGTATAGCTAAGCTCTTCGAGGAGTTGGTTCAAACTCCAAATTCAAATTA
ATTAATCCCAATGAGGAAACCCATGCCATATATTGCTTGCTCCTCATTTAAGATAATAATTGAATTAAATTAAGCCCTCCTTC
ACTGTTCAGTCATTCGTTCAAAGAGGAAAGGAAGATGAATCCAATCAAATCAGTTACAGAGAGATTTTAAAGTCACTGAAAGC
AAATTGTTTTAAAGGGTTGCGTTGGGACCCTACTTATAGAATACACATTATATATGTATTCCCCCCGCTGCATGCATTGAAAA
CTCTCTGTGTAATTAGACAAACGTGATTACGCGTTTTATCTAATCTGTAGTTCACTGTAACAGAAGATTGACAAAGTAATCAA
GTAGCTAAATATTTTCTCACTTGCCCTTTTTTTTTTTTCCGTCCCTTTTATAAGCCCTATCTAGATGGCTAGCCTTTAATTTA
AACTTTGAAGTCCCCAACTTCTACTCCCAAGTTGCGATTGCCAGTTCAATTTGCTTATTCCTCTGGCTCTGCTTCTGGCCTCC
TCGAGAGAACCTAAATCTACAAGAAACTCACAAGATAAGACTCCCAATTCTTCCCCAAGTAAGTGAATATTGTGAAGCTTAAT
TAATATATCCATAGCATTAGCTGGTTCATTACTTGCACTAATATAATAATATCATTATGCAAGAATCACATTATCATTAATGC
AACTTTTTAAAGCAGGGGGTAGAGTGTGATCACTGAATTTCCTTTCTTGTTTTTATTTTATTTTATTTTATTGATGTTTTATT
TAAAAGAAATGAACGGCATAGAGAGGAGGACGAGAGTGTGGCAGAATCTAAAGCAACGGCTGCGATTGAAGGGA**ATG**
>pCs5g20320
GGTAATTTTGGCTTTCGGCCTGTCTAGAAAAAATTCGGACAAAAATGTATGTCAAAACACTGCTTTTGGCCTTCAACTTAGAA
TTGAGAACATTTAGCTCATTGTTTTACCTCTCTTAATTCGAGATTGTAGAACATACAGTTGATATCACATCTAAGGCGGGTTT
GAATTGTTATGTTTTACTAATTCAACTAAATAATACTCACAAATTAAATTGTTGTGTTCATAAAATAGATGAAAATTAGTAAT
TGCTGTTTTATTTAAGTCATAGTGAAGTAACCAAACAAATTTTTTCTTTTTTTCTTTTTCTTTATAAATCTTTGAAAAGAATG
AAAATCTTCATCTACTAAATACCAACTTTATACTATTTGATACCAACTTTATACTATTTGTTATTAGTAGATCTATCGAGTTA
CACTAGGAGGCCTATCCAATTCCTATCCAATTTTGTTCTATAATTATAATTTATCCAAAATTCAACTCTGAACTTTGTCTCTA
AATTTCTTATGAAAATATTTTTTAACAAAACATAGTGCTAAAATTTAACCAGTATAGTGATGTATGGTGATTATTTTTAACCT
CAAAATAATCATAATGAACTATATGATATTTTTTTTTTTGGGTTAAAGAAAAAAGAGTATATGGTTGATGTAGCTAAATTATA
GAATAGTGAGAATAATGTTAATTATATTTAAGAAAAGTTAATAATAAAACTATAAGGTTCAATGTTAAGGGTATTTGAGAGTC
ACCTGACAAGTCAACTCAGTTGTAATTTTGACAGTTAATTATCTCCTTTTCCAATATGTAAACCCTAAAACGACGGACAAGAT
GTTGAATTTGGACACATAAAAACGGCGCCATTTTATAATGACAACTAAACAAATTAACAAAACAAGTCTATGCATGGAAAATC
AAAAGCCTAATATTGCGGCATTGTGATAGCAAGGGCCTTGCTCTTTTTTGCTCTCATTGATTCCTCTATAAAAGGCATAGCCA
TGCTTGACATAAACCCATTAACAGAAACTAATCAAAATTAATTAAAATTAGCCAAAATATTGGTACTTTGGCAGCTAAGTACC
AAAAAAAA
>pBs3 native
CTGGTTAAACAATGAACACGTTTGCCTGACCAATTTTATTATATAAACCTAACCATCCTCACAACTTCAAGTTATCATCCCCT
TTCTCTTTTCTCCTCTTGTTCTTGTCACCCGCTAAATCTATCAAAACACAAGTAGTCCTAGTTGCACATATATTTCTCATAGT
CAAGCTAACGAAACTTATGCAAGGGAAATATGAAATTAGTATGCAAGTAAACTCAAAGAACTAATCATTGAACTGAAAGATCA
ATATATCAAAAAAAAAAAAAAACAATAAAACCGTTTAACCGATAGATTAACCATTTCTGGTTCAGTTTATGGGTTAAACCACA
ATTTGCA
>pSlEXP1
GCCACATCCATGATTTGTGAGACATTCAACTGCACTGATGCTTCAATTTAAAGAAAATAAAGAAAAAAGGGAGTTTGACCAAA
GTGAGATATGTAGGGCGTAATATATGGGCCGATGATTGCGGTGGATAGTTTACCTCAAATTTGGCCACTTTCCATATTGTCAA
ACACACTATATTATTAATTTATTTTTCAAAACTTCACAATTCCAGCAGCTAAGTTGGCAAAAAGTGAAACCACTAATCAACTA
ATAGACAAAAATAAAATAAAATAAAAACTAGTACTATTATTATGCTCAAAGTTTTGGGCAAAAACATGATATGCTTTAAGACA
AAGAGAATTTTGTAGCATAATGAGATGAGACTGTTATTATTATTTTCGTACCTACTTTCAATTAATTATGCAACGTAACAACA
ATTTTCAAACACAAGCAACTTAGGTAAATATGGTTCATTTTAGGCTCATAATATGATCTCTCTGTACCACCTAATTGTAATAA
TTTTACAGTCTTAATATTTATATTATATATTTTTTTTCACATGAGTATTTTAGATAACTTTCAAATGAATTTTTTTTTTCTCT
TCTGCATTCAAATACATATTACTGAAATTTTAGCTAATTTTAACACGTATTCAATTTTGTATGATTGAATGTAGACATAATGG
AAAAATTGAAAGGGTGTTAAATAAGTTTATTAATTATAAAACATTTGTATTTCAAATTTAGTTAAATAGTGGGAGCGGAAAAT
TCGGGAGTAATTAATGAAAATGTGTTAAAAAGTGTCAGTAGATTGACGGAATTAAAGATAAAAAGCGGCTGCAAGTGTTCCAT
GGGAAACCTTTAATGTGATTTGGAACTTTTTTTTTTCTTTCTTTTTTTAATCCCATTCCATTCCAATCAAATCAAATTAAAAC
AAAACCAAAAGATTCAGTATTATGATTGCTTCATTCATAATCAGAGATTTTATATTTGAGTTTTGAATTTTGTTTTTGAAAGT
TCTACTTTCGAATGAGCATTGCAGTATGCGATTCATATTTAGTTGAACCTCAAATGCGAGTGAAATTTGAAAATACAAACACA
ACGCATCTAAACACACATTTAAGACATCTAAAATCAATAGACTGTACCAACCTATTATGTCAATAGTGTTCAAAATATGTTAT
TTAATCACTTTGTATATTCAATATTTTTTTCCGACGATTGGACACCTAAAAATAAGACAAAACAAAAAACAAAATAAATATTA
TTTCTTTTTATGTTTTTGTTCCCGCCAATATGTCCTTTGATTAAGGAGGACCAAACTTAGGTATACACACATAATTTCTCTGT
TATAAGTAGTGCTTCTCGTTGCATCATTTTCTTCAACAACCTCAACTCCATTAAATCTTAAGA
>pSlLOB1
AAAATAGTAAATAAATAAAAATGAACAGTAGAAGTACATTTTAAGACATAAATAAAGATAAAATATCAGTATTTCATCCTTAA
TAAGTAATATTTTCTTTGTAAACGTGTAAGCCGTGTAATTAATTTGAGATGAATGGAGTATTTACTAGTACTAGTGAAAATAG
AAAATGGCCACAACCACAACAAAATAAAAATCAGGTAAAAATTTCGAGACAGAAATCAAACTCTTAATAATAAATTAATTTTT

ATGTTTATATAGCCAACTGAACTATTACGATTTCTCCTCCGACCAAAAAAACTAAAAACCAAAAGAAATTACAAACAATAACA
ACAACAAACTCATTTTCACTTTTTCCTAGTTTCAAATTATTATATCATAATACTATATGGAAAAAAAAAGTTCTTTTTTAAAA
AATAAAAATGACAAAAAGGTTAGTGATGTCATTTTTCAAATTTATGACATCATCATACTCCCACTTATTTTTCTCCTTTGGCT
TATAAATAATACCCATTGTTCACTCCATCTTGGAACACACCAAGTAAGCACTATTTTCTTTTTATACACATTTCTACAAAGAT
CAAAAAAATTGAAAAAAAA
>pSlPL
TGATTCATTAGCCAAACCCCAAAAACAAAGTTCCATGAAAATTGATTTCTTATTTATTTTCCTGTGGATACACATAACAAACA
TTAAATACAAGAAACAAGAAAACCGCTTTTGTTCCTTCCAATGTCTCCTCTTATTATTTTCTTCTACCTTTTTTTGTCAAAGA
CATCGTACAATATTTTCATACAATTAGACTAAGTATCAGTAATCTTAATTAGGTGACATAAAATTACAGTACATAATTTTCTA
TAATTAGTTAAGGCACATTAATTGTGTAAGAAGATTATTTCTATTCAGTATTTGTATTGAGACTCGACTAAATGTAGATTTGC
TTCCCAATATGACTTTTACACAACAGACTCAAACTCGTGACCTTTAATTAAGAATGAAAAAATACTTATCGCTTCTTTAGTCA
AGAATGAAGAAGTATTTACCACTCAATATTCAATATTCACATTGATACTCGTTAAATTCATATTTGCTCCCCAATGCCCCTAT
TAGAAGTGACTTAGACCCAGCAGACTCAAGCTCGAGATCTTTAATGAAGAATGAAAAAGTACTAAAAGTATTCACATCGGGAG
CACGGTTAGATTCATATTTTCTTCCCAATGCCCTTATCAAAAAGCAACTTTGTATCGAGCAAACTCAAACTTGAAATTTTTAA
TCAAGAATACAAAGTACTTTAGATTCCGCATTAGAGTTTAGTTGAATTCAAATATGCTTCCCAATGCCCCTATCAAAAGCGAT
TTTGTACCGAACAGACTCAAACTCGTGACCTTTAATCAAGAATAAAAAAGTACATATCACTCCGTAAGAAGGGAAGTAAAATC
AAATTTGTGAAATAGTAGATGCATGTAGGTCCAAGTGGTGTTTGGTTGGTTCATAACTCCCTTCATCTTCTCTTCCCCTGTAT
ATAAAACCCCTCCAAAAGTCCCACATCCCCTTCATTCCAAACTCCATTATTTTCTCTTCCCTTAACAACACAAAGGGGATTTC
AA
>pSlPE
TTTTGTTGCAGTAGGTCTTCCTATTTCCTATTTGTATGTCGATAGACATAATGTATATATAATTAGCATATAGTAGAAAATTA
ATTGTTTTGACATCGGGACTTTAACTATTCTATCTTGATTTAAATATATTTTTGTGTGATTGGTTAATGAATGAAATTTCTTG
TAATTTCTTTTGAATATTTCTATATATATCGGGTGTAAGTATGATATAATTGAGATGTTTCTTTAAAATTCAAAGTCATAACT
CATAAGTACGATACAATTAAGATGTTTCTTAAAATATTGAAGTCATAAGAATAATACAATCGAGATGCTTCTTGAAAAGTCAA
AGTCATAATTCTCTACTATATTAATAATCATAATTAGTTAGATATTATGTTACATATATTAAGCTAATACCAAATAAATTATG
ATTACCAACTAACAATATCATAAATATTGAATCTAACTAATTTTTCATCCATTTTATTGCAGGTTGAAAATTGGTTAGCAAAT
TGGGAGGTTGAATACAAGTTTAGTATGAACAGGTTAATTAGTCAGGTCAACAAAAATATGGTTTTATTTCCCTCACTAATACT
CCCTTAATTTGTTCCATTATTATTGACTTTGACACACATTTTTAAAAACAATAAATGACAGAGGTAATATTATTATATTATTT
TTTTAAATATATTAAATTTTGATATTTCGTAAAATATTTTAGATAATTGAATAACATTTAACAATAGAAAAATAAAATAGACA
TGAAAAGATTAATTATCTCTTAATTTTTCAAAGTGAACAAGTATTATAAATACAAATTCTTAGTAAAATAAATAATTAAACAT
GAACAAAAGAAATAGCAGACATTGGGTGAAATAGTTGAAATAGACGCAAATTAATCCCAACACCACAGGTCTAGTCCACCTGA
CATTGAGCTCTATATATATCTACACTTTTGTGTAAAACTCTAAAACATTCCAATATCACAAAGATAAAGCA

1.4 Important protein sequences

>CsLOB1
MECKHKINVAIPITNMKNTQFSSPSTFSTSPPSQSSPRFPSPNHQQLSSPESSPSFKASPSQSSPNLAAPLSPPPIVLSPCAA
CKILRRRCVEKCVLAPYFPPTEPYKFTIAHRVFGASNIIKFLQELPESQRADAVSSMVYEASARIRDPVYGCAGAICHLQKQV
SELQAQLAKAQAELVTMESQQRNLITLICMEMAQSQEQVLQQQQQQQQQFMDTSCFLDDNGIGSAWEPLWT*
>CsLOB2
MTMHTTFPPLSSSPSPSFQSSPSINASPSKSSPNLAAPPPIVLSPCAACKSLRRRCDEKCVLAPYFPPTEPQNFIIVHRVFGA
SNIIKCLQGLPECQRSDAVSSMVYEANARIRNPVHGCAGAISQLQKQVIKLQAELAKAQAETVSMQCQRDNLVALICKEMTTQ
FPQETMNRVLPQQQFNDDDATTCYLDDKDFASTWDALWT*
>CsLOB3
MLKMENYEAAATRNPSKVTSSRAGSSSPPPISANSSAPPPVIMSPCAACKILRRRCADKCVLAPYFPPTEPAKFTIAHRVFGA
SNIIKFLQELPESQRADAVSSMVYEASARIRDPVYGCAGAICQLQKQVSELQAQLAKAQAEVVNMQCQQANLVALLYKEMGKS
PQPNSPQSVDHFITSPESPEANPCSFFEDNNLSGSLWEPALWT*
>CsLOB4
MGGNSPCASCKLLRRRCAKDCIFAPYFPPDDPHKFAIVHKVFGASNVSKMLQELPVHQRADAVSSLVYEANARVRDPVYGCVG
AISFLQNQVSQLQMQLAVAEAEILCIQMQQEPMMPTPQIDPDDKSFLLQNNLPHQYLNFTSSSNNVIHDSLKRESIFGHDMVS
*
>LBD1
MESKSDASVATTPIISSSSSPPPSLSPRVVLSPCAACKILRRRCAERCVLAPYFPPTDPAKFTIAHRVFGASNIIKFLQELPE
SQRTDAVNSMVYEAEARIRDPVYGCAGAIYHLQRQVSELQAQLAKAQVEMVNMQFQRSNLLELIYNMDQQQKQEQDNMSFESN
DLGFLEDKSNTNSSMLWWDPLWTC*
>LBD11
MLKMEINGGVATPTASAVAKVTETTTPVNSPSPTSSPPPPPSPQQPPQPPVVLSPCAACKILRRRCADKCVLAPYFPPTDPAK
FTIAHRVFGASNIIKFLQELPESQRTDAVNSMVYEAGARMRDPVYGCAGAIYHLQRQVSELQAQLAKTQVELVGMQLQRSSLL
ELIYNMEQTKLSVQEQGQQKMSFESSFESGDEFISSPDEESNDLGFLEDNNNNNNSSMSWWDPLWT*
>AS2
MASSSTNSPCAACKFLRRKCQPECVFAPYFPPDQPQKFANVHKVFGASNVTKLLNELHPSQREDAVNSLAYEADMRLRDPVYG
CVGVISLLQHQLRQLQIDLSCAKSELSKYQSLGILAATHQSLGINLLAGAADGTATAVRDHYHHHQFFPREQMFGGLDVPAGN
NYDGGILAIGQITQFQQPRAAAGDDGRRTVDPS*
>AtLOB
MASSSNSYNSPCAACKFLRRKCMPGCIFAPYFPPEEPHKFANVHKIFGASNVTKLLNELLPHQREDAVNSLAYEAEARVRDPV
YGCVGAISYLQRQVHRLQKELDAANADLAHYGLSTSAAGAPGNVVDLVFQPQPLPSQQLPPLNPVYRLSGASPVMNQMPRGTG
GSYGTFLPWNNGHDQQGGNM*
>SlLOB1

```
MESTISSSSSRSVSPSSSSSPNSPPPPSMTVVVVSPCAACKILRRRCAEKCVLAPYFPPNDPIKFTTAHRVFGASNIIKFLQE
LPESQRADAVSSMVYEANARLRDPVYGCAGSICQLQKQVSDLQAQLAKAQAEIVNMQCQQANLMALICMEMGQSNPQPISPPQ
QSLENFPMNYLDDNIGSWETLWT*
>Solyc05g014000
MGMPLSFLLLLTLLSPIFTFSSHVPDPEVIVQQVNEKINASRRNLGYLSCGTGNPIDDCWRCDPNWEKNRQRLADCAIGFGKQ
AIGGKDGKIYVVTDTSDDPVNPKPGTLRYGAIQDEPLWIIFSRDMVIKLKEELMLNSFKTIDGRGASVHIAGGPCITIQYVTN
IIIHGLNIHDCKQGGNAYVRDSPQHYGWRTISDGDGVSIFGGSHVWVDHCSLSNCNDGLIDAIRGSTAITISNNYMTHHNKVM
LLGHSDSFTRDKNMQVTIAFNHFGEGLVQRMPRCRHGYFHVVNNDYTHWEMYAIGGSASPTINSQGNRFLAPNDIFNKEVTKH
EDAAESEWKNWNWRSEGDLMLNGAFFIRSGAGASSSYAKASSLSARPSTLVNSITMNAGALGCKKGKRC*
>Cs8g11330.1
MAVTQRGICLCFAVVLMLFVSVLASVRNEQDVSVSRKMKAESSMNSTMAAKAEVVAEALSKHAVDNPDEIASMVEMSTRNSTE
RRKLGYFSCGTGNPIDDCWRCDGNWHKNRKRLADCGIGFGRNAIGGRDGRFYVVTDPRDDDPVNPKPGTLRHAVIQDKPLWIV
FKRDMVIQLKQELIVNSFKTIDGRGANVHIANGGCITIQFVTNVIIHGLHVHDCKPTGNAMVRSSPTHYGWRTVADGDAISIF
GSSHIWIDHNSLSHCADGLVDAVMGSTAITISNNHMTHHNEVMLLGHSDSYTRDKQMQVTIAYNHFGEGLIQRMPRCRHGYFH
VVNNDYTHWEMYAIGGSANPTINSQGNRYNAPLNAFAKEVTKRVDTAASQWKGWNWRSEGDLLLNGAYFTPSGAGASASYARA
SSLGAKSSSMVGSITSGAGALTCRKSRQCS*
>Cs2g23970.1
MSNISVFLIFFFSLLIPNLVSSSSIQDPELVAQDVHRSINASRRNLAYLSCGTGNPIDDCWRCDPNWERNRQRLADCAIGFGR
DAIGGRNGRIYVVTDSGDDNPTNPKPGTLRHAVIQDEPLWIIFNRDMVIKLNQELVMNSHKTIDGRGASVHIADGPCITIHFA
TNIIIHGIHIHDCKKAGNGNIRDSPEHSGWWDASDGDGVSIFSSQHIWIDHCSLSNCQDGLIDAIHGSTAITISNNYFTHHDK
VMLLGHSDSYTQDKNMQATIAFNHFGEGLVQRMPRCRQGYFHVVNNDYTEWQMYAIGGSAAPTINSQGNRFFAPNERFRKEVT
KHEDAPESEWRNWNWRSEGDLMLNGAYFRQTGAGASSTYARASSLNARPSTLVGPMTMRAGALNCRKGSRC*
>Solyc03g111690
MGTSSVFLLFLLSFLLLLPSLLASSNPQQVVDEVHRSINGSRRNLGYLSCGTGNPIDDCWRCDPNWEKNRQRLADCAIGFGKN
AIGGRDGKIYVVTDSGDDNAVTPKPGTLRHAVIQTEPLWIIFARDMVIQLKEELIMNSFKTIDGRGASVHIAGGPCITIQYVT
NIIIHGIHIHDCKQGGNAMVRSSPSHYGWRTVSDGDGVSIFGGSHVWVDHCSLSNCKDGLIDAIMGSTAITISNNYMTHHDKV
MLLGHSDTYTQDKNMQVTIAFNHFGEGLVQRMPRCRHGYFHVVNNDYTHWEMYAIGGSADPTINSQGNRFLAPDIRFSKEVTK
HEDAPESEWKNWNWRTDGDLMLNGAFFTRSGVRTGSSSYAKASSLSARPSSLVANLVSSSGALNCKKGSRC*
```

1.5 R scripts

R script 1: Extraction of 300bp around the ChIP-seq peak summits

```
options(stringsAsFactors=F, scipen=0)
library(seqinr)
chr_seq <-
read.fasta(file='/Users/trang/Desktop/Results/ChIP/all_samples/MACS_callpeak/Filter_q>10_datamining/
csi.chromosome.masked.fa', seqtype="DNA", as.string=T)
summit <- read.table("12h_R1_summit_Filter_196.bed",F)
flank <- 150
seq_list <- list()
for(i in 1:nrow(summit)){
  chr <- summit[i,1]; pt <- summit[i,2]; peak_name <- sub("1.*BAM_", "", summit[i,4])
  sequence <- substr(chr_seq[[chr]][[1]][1], pt-flank, pt+flank)
  seq_list <- c(seq_list,list(sequence)); names(seq_list)[i] <- peak_name
}
write.fasta(sequences = seq_list, names(seq_list), file.out = "12hR1summit_300bp.fa")
```

R script 2: Distance to TSS

```
options(stringsAsFactors=F, scipen=0)
setwd("~/Documents/ZMBP/collaborations_and_visit/Thomas Lahaye/for_Trang/files_2022_02/")
targets <- read.table("target_genes_4kb_flank.txt",T)
peaks_12h <- read.table("12hconsensus.bed",F)
peaks_36h <- read.table("36hconsensus.bed",F)

library(IRanges)
summits_IR_list_12h <- summits_IR_list_36h <- vector("list",10)
names(summits_IR_list_12h) <- names(summits_IR_list_36h) <- names(table(peaks_12h[,1]))

for(i in names(summits_IR_list_12h)){
```

```
  summits_IR_list_12h[[i]] <- IRanges(peaks_12h[peaks_12h[,1]==i,2], peaks_12h[peaks_12h[,1]==i,3])
  summits_IR_list_36h[[i]] <- IRanges(peaks_36h[peaks_36h[,1]==i,2], peaks_36h[peaks_36h[,1]==i,3])
}

expand_feature <- function(vec, n){
  #vec: vector with features
  #n: number of elements after expanding, NOTE: I assume here that n > length(vec)
  vec_len <- length(vec); result <- rep(0, n)
  if(n>=vec_len){
    for(i in 1:vec_len){
      result[ (floor(n*(i-1)/vec_len)+1) :floor(n*i/vec_len) ] <- vec[i]
    }
    return(result)
  }
  if(n<vec_len){
    vec_temp <- rep(vec, each=n)
    for(i in 1:n){
      result[i] <- mean(vec_temp[((i-1)*vec_len+1):(i*vec_len)])
    }
    return(result)
  }
}

#approximat each gene locus to a segment of 100bp windows. As the scan was made via including 1kb
flanking regions, I will take the same region for each locus
bin <- 100
result_mat_12h <- result_mat_36h <-c()
for(i in 1:nrow(targets)){
  chr_name <- targets[i,1]
  left_bin <- ceiling(targets[i,2]/bin)-4000/bin; right_bin <- ceiling(targets[i,3]/bin)+4000/bin
  ref_nt <- (left_bin-1)*bin

  ###!!!!!!! we only check -4kb to 4kb downstream of TSS!!
  #12h
  vector_ovalap <- rep(0, 500)
  overlap_peaks <-  findOverlaps(IRanges(targets[i,2]-4000,targets[i,3]+4000),
summits_IR_list_12h[[chr_name]])
  if(length(overlap_peaks)>0){individual_peaks <- summits_IR_list_12h[[chr_name]][overlap_peaks@to,]
  mat_gene <- cbind(0,rep(left_bin:right_bin, each=100))
  for(j in 1:length(individual_peaks)){
    start_nt <- max(individual_peaks[j,]@start-ref_nt+1, 1); end_nt <-
min(individual_peaks[j,]@start+individual_peaks[j,]@width-ref_nt+1, dim(mat_gene)[1])
    if(start_nt>dim(mat_gene)[1]|end_nt<1){print(i);next}
    mat_gene[start_nt:end_nt,1] <- mat_gene[start_nt:end_nt,1]+1
  }
  gene_vector <- aggregate(mat_gene[,1]~mat_gene[,2], mat_gene, sum)[,2]
  ###!!!!!!! we only check -4kb to 4kb downstream of TSS!!
  if(targets[i,"strand"]=="+"){vector_ovalap <- gene_vector[1:80]
}else{
    vector_ovalap <- gene_vector[(length(gene_vector)-79):length(gene_vector)]
    vector_ovalap <- rev(vector_ovalap)
  }
  result_mat_12h <- rbind(result_mat_12h, vector_ovalap)
  }
```

```
 #36h
 vector_ovalap <- rep(0, 500)
 overlap_peaks <-  findOverlaps(IRanges(targets[i,2]-4000,targets[i,3]+4000),
summits_IR_list_36h[[chr_name]])
 if(length(overlap_peaks)>0){individual_peaks <- summits_IR_list_36h[[chr_name]][overlap_peaks@to,]
 mat_gene <- cbind(0,rep(left_bin:right_bin, each=100))
 for(j in 1:length(individual_peaks)){
  start_nt <- max(individual_peaks[j,]@start-ref_nt+1, 1); end_nt <-
min(individual_peaks[j,]@start+individual_peaks[j,]@width-ref_nt+1, dim(mat_gene)[1])
  if(start_nt>dim(mat_gene)[1]|end_nt<1){print(i);next}
  mat_gene[start_nt:end_nt,1] <- mat_gene[start_nt:end_nt,1]+1
 }
 gene_vector <- aggregate(mat_gene[,1]~mat_gene[,2], mat_gene, sum)[,2]
 ###!!!!!!! we only check -4kb to 4kb downstream of TSS!!
 if(targets[i,"strand"]=="+"){vector_ovalap <- gene_vector[1:80]
}else{
  vector_ovalap <- gene_vector[(length(gene_vector)-79):length(gene_vector)]
  vector_ovalap <- rev(vector_ovalap)
 }
 result_mat_36h <- rbind(result_mat_36h, vector_ovalap)
 }

 if(i%%1000==0){print(i)
}
}

par(mar=c(4,4,4,6))
plot(apply(result_mat_12h,2,mean), type="l", xaxt="n", ylab="Average enrichment per 100bp", ylim=c(0,
20),
   col="orange", lwd=2)
lines(apply(result_mat_36h,2,mean), type="l", lwd=2, col="red")

axis(1, at=c(1,30,40,42,80), labels = c("4kb", "", "","","4kb downstream TSS"))
axis(1, at=c(26,36,46), labels = c("1kb","TSS", "200bp"), tick=F)

abline(v=c(30,40,42), lty=2 )
legend("topleft", legend=c(paste0("12h, ", "n=", nrow(result_mat_12h)),paste0("36h, ",
"n=",nrow(result_mat_36h))),
     lwd=2, col=c("orange","red"), bty="n")
```

R script 3: Extraction of genes that have ChIP-seq peaks in the -3kb to +200bp region of their TSS

```
#target genes with ChIP peak summits in -3kb to +200bp around TSS region
options(stringsAsFactors=F, scipen=0)
#SECTION A: to identify genes overlapping with ChIP-seq peaks
setwd("/Users/trang/Desktop/Results/ChIP/all_samples/MACS_callpeak/Filter_q>10_datamining")
library(IRanges)


gene_anno <- read.table("gene_anno_simple.txt",T)
summits <- read.table("12h_R1_summit_Filter_196.bed",F)
```

```
summits[,4] <- gsub(".*BAM_", "", summits[,4])

summits_IR_list <- vector("list",10); names(summits_IR_list) <- names(table(summits[,1]))
for(i in names(summits_IR_list)){
 summits_IR_list[[i]] <- IRanges(summits[summits[,1]==i,2], width=2, names=summits[summits[,1]==i,4])
}

gene_anno <- data.frame(gene_anno[,1:5], corepro=NA)
#find overlap with corepromoter:3kb upstream TSS and 200bp downstream TSS
for(i in 1:nrow(gene_anno)){
 chr <- gene_anno[i,1]
 if(gene_anno[i,"strand"]=="+"){
  result <- findOverlaps(IRanges(gene_anno[i,2]-3000,gene_anno[i,2]+200), summits_IR_list[[chr]])
  if(length(result)>0){gene_anno[i,"corepro"]<-
paste(summits_IR_list[[chr]][result@to,]@NAMES,collapse=";")}
 }else{
  result <- findOverlaps(IRanges(gene_anno[i,3]-200,gene_anno[i,3]+3000), summits_IR_list[[chr]])
  if(length(result)>0){gene_anno[i,"corepro"]<-
paste(summits_IR_list[[chr]][result@to,]@NAMES,collapse=";")}

 }
}

write.table(gene_anno[!is.na(gene_anno[,6]),], file="12hR1_target_genes_corepro3kb.txt", sep="\t",
row.names=F, quote=F)
```

R script 4: Draw heatmaps

```
#Heatmap of citrus genes
#load packages
BiocManager::install("ComplexHeatmap")
library(ComplexHeatmap)
BiocManager::install("circlize")
library(circlize)
library("RColorBrewer")
library(ggplot2)

setwd("/Users/trang/Dropbox/lahaye's project/Writing/RNAseq")

rpkm <- read.csv("96genes3.csv",header = TRUE,row.names = 1)
str(rpkm)

data <- t(scale(t(rpkm)))

#read groups
samples <- read.csv("groups.csv",header = TRUE,row.names = 1)
#samples$Group <- factor(samples$Group, levels = c("Xcc12h", "Xcc36h",
# "mutant12h", "mutant36h","dTALE2","dTALE4"))
samples$Group <- factor(samples$Group, levels = c("mutant12h", "mutant36h","Xcc12h",
                                    "Xcc36h","dTALE2","dTALE4"))

HM = Heatmap(data, show_row_names = T, show_column_names = F,
          clustering_distance_rows = "euclidean",
          clustering_method_rows = "complete",
```

```
        name = "z-score",
        column_title = NULL,
        row_names_gp = gpar(fontsize = 7),
        row_names_side = "left",
        cluster_columns=F,
        column_split=samples$Group,
        row_title = NULL,
        border = TRUE,
        cluster_column_slices = F,
        use_raster = TRUE
)

draw(HM)
```

ACKNOWLEDGEMENT

I would like to express my deep gratitude to my PhD supervisor, Prof. Dr. Thomas Lahaye, for giving me the opportunity to work in his lab and for this tremendous support me throughout my PhD. Without his patience, teaching, supervision and willingness to discuss the experiments and follow up research ideas, this PhD thesis would not be possible.

Thank you to Prof. Dr. Ulrike Zentgraf for her suggestions throughout the PhD, for being a supportive committee member and for reviewing my thesis. Many thanks for Prof. Dr. Marjia Timmermans for her valuable insights into the research as a committee member and for her participation in the examination panel. Thanks to Prof. Dr. Rosa Lozano-Durán for being on my examination panel.

Many thanks to Prof. Dr. Chang Liu, his lab members: Dr. Ezgi Karaaslan and Dr. Nan Wang for their massive help with the ChIP-seq experiments. Many thanks to Dr. Joachim Kilian and Dr. Edda von Roepenack-Lahaye for their help in GC-MS and LC-MS experiments. Thanks to our collaborators, the labs of Dr. Murakami, Dr. Concetta Licciardello, Dr. Paulo Teixeira, Prof. Dr. Jeff Jones and Dr. Sandra Richter for their continuous support and valuable contributions.

I am very grateful for the help, support and kindness of all the members of AG Lahaye, especially Dr. Annett Strauß and Dr. Robert Morbitzer for their fruitful discussions and suggestions on research data and experimental troubleshooting. Thanks to Annett for proofreading the thesis. I am also very grateful to Paloma Aguilera for working with me for the last two years and helping me so much in the lab. I would like to thank Dr. Danalyn Holmes for her support in proofreading this thesis and for always giving me supportive talks. Thank you to Stefan, Angela, Charley, Erin, Niels, Kyrylo, Kevin, Markus, Dousheng, Christina and Nikolas for being awesome colleagues and creating such a great working environment! Many thanks to Johanna Schröter and the ZMBP Plant Cultivation Team for their amazing work in taking care of the plants. Without their support my PhD would be very difficult.

Thanks to Dr. Khoa Nguyen, Van Pham and Nga Pham for their help and their friendship. Thank you to my dear friend Binh Nguyen who has given me endless support through the ups and downs of my PhD. Thank you to JB for proofreading my German abstract and sharing a great time with me.

I would like to dedicate this thesis to my parents, brother, sister-in-law, dear little niece and nephew for their unconditional love and for always believing in me.