CONTRIBUTION OF TAL EFFECTORS IN *Xanthomonas* TO DISEASES OF RICE AND WHEAT

by

ZHAO PENG

B.A., China Agricultural University, 2006
M.A., China Agricultural University, 2009

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Plant Pathology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Abstract

Rice and wheat are two major crops that suffer losses from the diseases of bacterial blight and bacterial leaf streak, which are caused by *Xanthomonas oryzae* pv. *oryzae* (Xoo) and *Xanthomonas translucens* pv. *undulosa* (Xtu), respectively. Transcriptional-Activator Like (TAL) effectors, a special family of type III effector proteins from *Xanthomonas,* have been demonstrated as critical virulence factors that act by inducing corresponding susceptibility (S) genes in several disease complexes of plants. In this study, I analyzed the contributions of TAL effectors from Xoo and Xtu to virulence and in modulating host gene expression to enhance susceptibility. Specifically, the TalC effector from the African Xoo strain AXO1947 was identified as a critical virulence factor, which functions by promoting expression of the gene *OsSWEET14* in rice. TalC is interchangeable with other major TAL effectors from Asian strains of Xoo on the basis of functional complementation. The TAL effector PthXo2 from the Asian Xoo strain JXO1 is a major virulence factor and contains 21.5 repeats in the central repetitive region that targets *OsSWEET13* in *indica* rice varieties but not in *japonica* rice varieties. A one repeat deletion in the PthXo2 effector enabled effector specificity to switch from *indica* rice to *japonica* rice. TAL effector genes from a genomic analysis of the Xtu strain XT4699 and related strains were characterized with regards to their involvement in virulence and the modulation of host gene expression in the Chinese Spring wheat cultivar. The identification of TAL effectors with virulence contributions and their target S genes is important for understanding the virulence mechanisms of *Xanthomonas* bacteria and promises to provide new strategies for disease control.

CONTRIBUTION OF TAL EFFECTORS IN *Xanthomonas* TO DISEASES OF RICE AND WHEAT

by

ZHAO PENG

B.A., China Agricultural University, 2006
M.A., China Agricultural University, 2009

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Plant Pathology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Dr. Frank F. White

Copyright

Abstract

Rice and wheat are two major crops that suffer losses from the diseases of bacterial blight and bacterial leaf streak, which are caused by *Xanthomonas oryzae* pv. *oryzae* (Xoo) and *Xanthomonas translucens* pv. *undulosa* (Xtu), respectively. Transcriptional-Activator Like (TAL) effectors, a special family of type III effector proteins from *Xanthomonas,* have been demonstrated as critical virulence factors that act by inducing corresponding susceptibility (S) genes in several disease complexes of plants. In this study, I analyzed the contributions of TAL effectors from Xoo and Xtu to virulence and in modulating host gene expression to enhance susceptibility. Specifically, the TalC effector from the African Xoo strain AXO1947 was identified as a critical virulence factor, which functions by promoting expression of the gene *OsSWEET14* in rice. TalC is interchangeable with other major TAL effectors from Asian strains of Xoo on the basis of functional complementation. The TAL effector PthXo2 from the Asian Xoo strain JXO1 is a major virulence factor and contains 21.5 repeats in the central repetitive region that targets *OsSWEET13* in *indica* rice varieties but not in *japonica* rice varieties. A one repeat deletion in the PthXo2 effector enabled effector specificity to switch from *indica* rice to *japonica* rice. TAL effector genes from a genomic analysis of the Xtu strain XT4699 and related strains were characterized with regards to their involvement in virulence and the modulation of host gene expression in the Chinese Spring wheat cultivar. The identification of TAL effectors with virulence contributions and their target S genes is important for understanding the virulence mechanisms of *Xanthomonas* bacteria and promises to provide new strategies for disease control.

Table of Contents

Contents

List of Figures

Chapter Two

Chapter Three

List of Tables

Chapter One

Chapter Two

Chapter Three

Acknowledgements

First of all, I would like to express my deepest appreciation to my major advisor Dr. Frank F. White. His broad knowledge, inspiring perspectives and critical thinking in the research field pave the way for my Ph. D. study. His patience and encouragement is always supporting me to overcome any obstacles. He is a truly role model in science and research. I am also very grateful to my co-major advisor Dr. Sanzhen Liu for his inspiring guidance and passionate support. He opens the door for me in the genomics study. He is easy-communicating in whatever research and life, and helps me to release the pressure in the difficulty times. Additionally, I would like to show my appreciation to my other committee members, Dr. Barbara Valent, Dr. Eduard Akhunov, Dr. Sunghun Park and Dr. Sherry Fleming for their critical comments and valuable suggestions. I also learned a lot from enrolling their excellent courses.

I would like to thank all my colleagues in White lab, Liu lab and Park lab. I benefit a lot from working with them and enjoy the atmosphere of helping each other. I thank all the people from the Department of Plant Pathology, a friendly and wonderful family. Specially, I am grateful to Kim Park, Dr. Harold Trick and Dr. Gill for their wheat seeds. I would like to thank Dr. Alina Akhunova, Hanquan, Yanni and Yuye in the Genomic Facility for their excellent services. I am also thankful to my collaborator Dr. Zhaohui Liu at NDSU for the experimental help from his lab.

I am very grateful to my wife Ying, my lovely son Lionel, my parents, brother and other relatives in China for their love, care and support in my life.

Chapter One

Discovery, Function and Application of TAL effectors

Abstract

Transcriptional-Activator Like (TAL) effectors are a special family of type III effector proteins from *Xanthomonas* and *Ralstonia* spp. of plant pathogenic bacteria. TAL effectors possess nuclear localization motifs and transcriptional activation domain in the C-terminus, type III secretion signal in the N-terminus, and a variable central repetitive region that moderates DNA binding specificity. TAL effectors contribute to bacterial virulence by transcriptionally activating plant susceptibility genes. On the other hand, plants have evolved resistance genes with promoters that recognize specific TAL effectors to trigger active resistance. Plants have also evolved resistance to specific TAL effectors through modified susceptibility gene promoter elements, which no longer bind the cognate TAL effectors. The presence of a code for the DNA binding specificity by TAL effectors has opened the door for scientists to design artificial TAL effector genes for genome editing and transcriptional modulation in a wide range of organisms. At the same time, natural TAL effector genes are present in many uncharacterized strains of *Xanthomonas,* and the role played by most TAL effectors in nature is unknown. Next generation sequencing platforms are facilitating the rapid characterization of new TAL effector genes in *Xanthomonas* and *Ralstonia* species. Combined with RNA-seq and microarray technologies, these genomic tools will aid in identifying new host susceptibility and TAL effector-dependent R genes.

Introduction

Xanthomonas is a genus of plant pathogenic γ-proteobacteria, which can cause disease on more than 300 different species of plants, including important crops, such as rice, wheat, barley, citrus, cotton, and soybean (Boch & Bonas, 2010, White et al., 2009). Many xanthomonads rely on the type III secretion system to inject effectors into plant host cells for pathogenicity (Cornelis, 2006). The type III effectors target different host subcellular compartments and manipulate a variety of cellular components to suppress plant immunity and gain nutritional and virulence benefits for bacteria (Block et al., 2008, Deslandes & Rivas, 2012). As one family of type III effectors, the Transcriptional-Activator Like (TAL) effectors represent the largest-membered family within the genus of *Xanthomonas* (White et al., 2009). TAL effectors play critical roles in some disease complexes, such as bacterial blight of rice, cassava, and citrus canker (Cohn et al., 2014, Hu et al., 2014, Yang et al., 2006, Yang & White, 2004).

TAL effectors are a structurally and functionally distinct family of proteins, which have highly conserved N-terminal, C-terminal and central repetitive regions. The N-terminal amino acid residues are involved in the type III secretion, while the C-terminal region is required for eukaryotic nuclear localization and acidic transcription activation activity (Yang et al., 2000, Zhu et al., 1998). The central repetitive regions consist of a variable number of near identical 34-35 amino acid repeats with the exception of the 12^{th} and 13^{th} amino acid residues, which are variable and called repeat variable di-residues (RVD) (Boch et al., 2009, Moscou & Bogdanove, 2009, Yang & White, 2004). The central repetitive regions determine the specificity of TAL effectors and function as DNA-binding domain in a manner of one RVD optimally matching up with one nucleotide (Boch et al., 2009, Moscou & Bogdanove, 2009).

The Discovery of TAL effectors

The first member of TAL effectors to be identified was AvrBs3 from *Xanthomonas campestris* pv. *vesicatoria*, which was found to harbor avirulence activity on pepper with the *Bs3* resistance gene (Bonas et al., 1989). The homologous effector genes *avrXa7* and *avrXa10* were subsequently characterized in *Xanthomonas oryzae* pv. *oryzae* on the basis of the corresponding avirulence activities (Hopkins et al., 1992). Additional homologs *pthA* and *avrb6* from *Xanthomonas citri* pv. *citri* and *X. campestris* pv. *malvacearum* were identified as virulence factors (Duan et al., 1999, Yang et al., 1994). AvrXa7 was also demonstrated to have virulence activity (Bai et al., 2000, Yang et al., 2000). TAL effectors PthXo1, PthXo2, and PthXo3, in different *Xanthomonas oryzae* pv. *oryzae* strains, played major roles in the disease process of bacterial blight similar to AvrXa7 (Yang & White, 2004). As the sequencing and other high-throughput technologies improved, more and more TAL effector genes have been identified from a wide range of plant pathogenic bacteria within the genus of *Xanthomonas* and *Ralstonia* (Bogdanove et al., 2011, Heuer et al., 2007, Salanoubat et al., 2002, Salzberg et al., 2008).

At the same time, research on the structural features of TAL effectors provided insight into TAL effector function. The nuclear localization signal (NLS) motifs of TAL effector AvrBs3 was determined to be required for AvrBs3-dependent resistance in *Bs3* containing plants (Van den Ackerveken et al., 1996). Similarly, an acidic transcriptional activation domain (TAD) of TAL effector AvrXa10 was first demonstrated to activate gene expression, when fused with Gal4 DNA binding domain in yeast and *Arabidopsis* and was shown to be required for *Xa10*-mediated resistance in rice (Zhu et al., 1998). Furthermore, it was shown that the TAD of AvrXa10 could be replaced by the heterologous, VP16 TAD domain from herpex simplex virus (Zhu et al., 1999). Later, all three activities, nuclear localization activity, transcriptional

activation, double strand DNA binding activity were demonstrated for the TAL effector AvrXa7 and AvrBs3 (Yang et al., 2000, Szurek et al., 2001). The specific induction of the rice *OsSWEET11/Os8N3* associated with the TAL effector PthXo1 prompted renaming AvrBs3/PthA family protein to TAL effector (Yang et al., 2006).

The central repetitive regions of TAL effectors were also shown to have an essential function in activity. Deletion derivatives of AvrBs3, which were missing repeats from the central region, were no longer recognized by *Bs3*, while one particular deletion derivative, AvrBs3Δrep16 (missing the 11^{th}-14^{th} repeats) triggered a hypersensitive response in another isogenic pepper line (Herbers et al., 1992). The exchange of the repetitive regions of the TAL effectors AvrXa10 and AvrXa7 from Xoo strain PXO86, rendered the compatible strain $PXO99^{A}$ incompatible on rice plants containing *Xa10* and *Xa7*, respectively, depending on the source of the repetitive region (Hopkins et al., 1992). Sequencing of other TAL effector genes *pthA* and *avrb6* further implicated the repetitive region in the biological specificity of each TAL effector (Yang et al. 1994). Repeat deletion derivatives of AvrXa7, PthXo4 and PthXo5, lost *Xa7*-mediated elicitor capacity but retained virulence, while another derivative, *avrXa7Δ38*, maintained avirulence activity but lost virulence activity (Yang et al., 2005). The specific binding of a TAL effector to the cognate promoter of the host target gene was first demonstrated for AvrBs3, which could bind to the promoter elements of *Bs3* and the gene *upa20* (Kay et al., 2007, Römer et al., 2007). Subsequently, a binding code was deciphered that governs the specific recognition of DNA in correspondence to specific RVDs (Boch et al., 2009, Moscou & Bogdanove, 2009). In this way, TAL proteins with desired arrangement of RVDs could be designed to specifically bind any DNA nucleotide bases in target. This great discovery in the

TAL effector history has brought about revolution in genome editing and targeted modulation of gene expression (Li et al., 2013, Li et al., 2012, Miller et al., 2011, Zhang et al., 2011).

TAL effectors function in plant disease

In contrast to other type III effectors, TAL effectors target specific promoter elements of host genes and directly activate host gene transcription, which can subsequently lead to enhanced host susceptibility. Host transcriptome profiling studies on the xanthomonad-plant interactions has revealed a series of susceptibility genes and upregulated genes by microarray or RNA-seq analyses and further functional analysis (Antony et al., 2010, Cernadas et al., 2014, Cohn et al., 2014, Hu et al., 2014, Kay et al., 2007, Sugio et al., 2007, Yang et al., 2006). A list of TAL effector-dependent plant susceptibility genes identified is shown in Table 1-1.

On the basis of the virulence contributions, TAL effectors are sometimes classified into three levels: major, moderate and undetectable (Boch & Bonas, 2010, White & Yang, 2009). PthXo1, PthXo2, PthXo3, AvrXa7, TalC and Tal5 from Xoo strains are considered major TAL effectors in rice bacterial blight since mutation of any one of them from strains that harbor only one major TAL gene leads to severe loss of virulence (Streubel et al., 2013, Yang & White, 2004, Yang et al., 2000, Yu et al., 2011). The *pthXo1* mutant of PXO99^{A} ME2 (hereafter ME2) could be complemented by any one of the major TAL effector genes *pthXo2*, *pthXo3* and *avrXa7* (Yang & White, 2004). The major virulence contribution results from the specific induction of critical susceptibility genes, the so-called clade III SWEET sugar transporter genes (also known as MtN3 genes), in a gene-for-gene manner (Antony et al., 2010, Chen et al., 2010, Yang et al., 2006). PthXo1 from Xoo strain PXO99^{A} specifically targets the promoter element and activates the gene expression of the rice gene *Os8N3/OsSWEET11* (Chu et al., 2006, Yang et al., 2006). In *xa13* and *Os8N3*-silenced plants, the absence of elevated expression of *Os8N3/OsSWEET11* was

correlated with resistance to PXO99^{A}. However, the plants remained susceptible to other strains that relied on the major TAL effectors PthXo2, PthXo3 and AvrXa7 to turn on alternative SWEET genes (Antony et al., 2010, Yang et al., 2006). It had been proposed that activated expression of SWEET sugar transporter genes leads to nutrient efflux to the vascular systems for the benefit of bacterial propagation (Chen et al., 2010). Induction of any one of the five genes in the clade III of the SWEET gene family in rice could render the plants susceptible, whether targeted by native or artificial TAL effectors (Li et al., 2013, Liu et al., 2011, Streubel et al., 2013, Yu et al., 2011, Zhou et al., 2015). The AvrXa7, PthXo3, TalC and Tal5 from different strains recognized different effector binding elements (EBEs) of *Os11N3/OsSWEET14* gene, second identified SWEET gene in rice, and activated gene expression for disease (Antony et al., 2010; Yu et al., 2011; Streubel et al., 2013, Chapter 2). The third identified SWEET gene in rice was the *OsSWEET13* gene, which was targeted by unknown effector from PXO339 strain (Liu et al., 2011). In the chapter 2, data is presented demonstrating the role played by the TAL effector PthXo2 from JXO1 strain in targeting the EBE of *OsSWEET13* gene in the rice line IR24 but not in *japonica* rice lines. Although no native TAL effectors had been identified for *OsSWEET12* and *OsSWEET15* genes, the induction of either one of them by artificial TAL effectors could confer disease symptoms (Li et al., 2013, Streubel et al., 2013). Specific induction of *OsSWEET* genes in other clades did not render the rice plants susceptible to pathogens (Streubel et al., 2013). The transfer of major TAL effectors into US *X. oryzae* strains, which are weakly virulent and devoid of TAL effector genes, enabled switching on the SWEET gene expression and enhanced disease on a variety of rice cultivars (Verdier et al., 2012). Furthermore, the transcriptional activation of SWEET sugar transporter gene by TAL effector is not unique to rice-Xoo system. Bacterial blight of cassava was demonstrated to have the model of gene-for-

gene susceptibility by TAL effector. Elevated gene expression of *MeSWEET10a* by TAL effector $TAL20_{Xam668}$ contributed to bacterial virulence in promoting *in planta* bacterial growth and water soaking symptoms (Cohn et al., 2014).

Two additional distinct host susceptibility genes *OsTFX1* and *OsTFIIAγ1* are specifically induced by PthXo6 and PthXo7, respectively, and both make contributions to virulence of Xoo strains on rice (Sugio et al., 2007). *OsTFX1* is a member of the bZip family of transcriptional factor genes and is induced by a wide range of Xoo strains from different geographic locations (Sugio et al., 2007, Chapter 2). The *pthXo6* mutant of $PXO99^A$ ME1 had a 35% loss of virulence in the lesion length measurement assays, and the ectopic of *OsTFX1* abrogated the requirement of PthXo6 effector to ME1 for full virulence (Sugio et al., 2007). In contrast to the induction of *OsTFX1, OsTFIIAγ1*, encoding a small subunit of transcription factor IIA (TFIIA) on chromosome 1, was only upregulated by $PXO99^A$ and was specifically targeted by PthXo7. The heterologous transfer of *pthXo7* effector gene enabled PXO86 become compatible on IRBB5 rice plants, containing the recessive resistance gene *xa5*, a mutant allele of small subunit of TFIIA on chromosome 5 (Sugio et al., 2007). This may reflect an adaptation of $PXO99^A$ strain in plants containing *xa5* genotype, but could not completely explain the virulence of $PXO99^A$ strain on *xa5* plants since *pthXo7* mutant of $PXO99^A$ was still highly virulent on *xa5* plants (Sugio et al., 2007). In total, targeted induction of *OsTFX1* and *OsTFIIAγ1* by TAL effectors indicates the ability of Xoo strains to modulate host transcription system to promote bacterial growth and lesion development and adaptation to resistance genotype *xa5* (Sugio et al., 2007).

Although it had been known that *pthA* and its homologs contribute to virulence and pustule formation for *X. citri* strains in citrus bacterial canker (Duan et al., 1999, Yang et al., 1994), the effectors' targets were unknown until recently (Hu et al., 2014, Li et al., 2014).

Through microarray analysis and qRT-PCR confirmation, two candidate host genes *CsSWEET1*, a sugar transporter gene, and *CsLOB1*, a member of the LOB family of transcriptional factor genes, had been identified as targets for enhanced expression by PthA4 (or PthAw) (Hu et al., 2014). The *CsLOB1* gene was subsequently demonstrated as the gene responsible for canker formation and enhanced bacterial populations. Designer TAL effectors that specifically targeted the promoter element of *CsLOB1* restored the virulence loss of the *pthA4* mutant, while the designer TAL effectors corresponding to *CsSWEET1* failed to do so. Furthermore, the effectors PthB and PthC, from other *Xanthomonas* strains, targeted *CsLOB1* but not *CsSWEET1* and restored pustule formation for *pthA4* mutant. Although the exact function of *CsLOB1* in citrus canker is not known, it was proposed to promote expression of downstream genes that are associated with the cell wall remodeling process (Hu et al., 2014).

Transcription profiling studies of bacterial blight and bacterial leaf streak in rice revealed two different sets of upregulated genes. The *OsSWEET* and *OsTFX1* genes are typically induced by Xoo strains, while neither of them was upregulated by *X. oryzae* pv. *oryzycola* (Xoc) strains, the causal agent of bacterial leaf streak (Cernadas et al., 2014). Targeted mutagenesis of TAL effector genes in Xoc strain BLS256 and screening of TAL effector mutants based on the lesion expansion phenotype indicated that the Tal2g effector was a contributor to virulence in bacterial leaf streak, while loss of other TAL effector genes had no detectible phenotype. The *tal2g* mutant had shorter expanded lesion and less bacterial exudation in leaf surface compared to WT inoculations. Complementation of the mutant with *tal2g* and designer TAL effector genes revealed a novel susceptibility gene named *OsSULTR3;6*, which is not induced in rice by Xoo strains. The function of the sulfate transporter gene *OsSULTR3;6* is currently unknown (Cernadas et al., 2014). Working similarly as Tal2g from Xoc strain, AvrBs3, the first identified

TAL effector from *X. campestris* pv. *vesicotoria*, did not affect the bacterial population growth but played roles in bacterial spreading, resulting from tissue hypertrophy and mesophyll cells enlargement in infected sites (Boch & Bonas, 2010, Kay et al., 2007). In contrast to the Tal2g target, the susceptibility gene directly targeted by AvrBs3 is *upa20* gene, encoding a member of the basic helix-loop-helix (bHLH) family of transcriptional factors. The *upa20* gene was proposed to function as an inducer of cell enlargement by activating a series of downstream genes, one of which was *upa7* encoding an α-expansin (Kay et al., 2007).

TAL effectors function in plant resistance

Despite the fact that pathogens have developed a wide range of virulence effectors to facilitate colonization in plant tissue, plants have evolved different resistance (R) genes accordingly in the co-evolutionary arm race. Here, an atypical group of resistance (R) genes (also called terminator or executor R genes) function somewhat as a promoter traps for specific TAL effectors. A list of plant executor R genes identified so far is shown in Table 1-2.

The first characterized executor type R gene was *Xa27* in rice. The same coding sequence of *Xa27* was found in resistance and susceptible cultivars, but only the *Xa27* gene in resistance plants was induced by AvrXa27 effector from Xoo strains. Detailed characterization of the promoter sequence revealed that only the *Xa27* in resistance cultivars harbored an effector binding element for AvrXa27 (Gu et al., 2005, Römer et al., 2009). Heterologous expression of *avrXa27* in rice plants with inducible *Xa27* or ectopic constitutive expression of *Xa27* could confer increased resistance to otherwise compatible *X. oryzae* strains that are lack of *avrXa27* effector gene (Gu et al., 2005, Tian & Yin, 2009). The precise function of Xa27 is currently unknown. The N-terminal signal-anchor-like sequence was shown to be essential for localization to apoplasts and resistance (Wu et al., 2008).

In contrast to *Xa27*, the expression of the second executor R gene *Xa10* has been shown to elicit strong hypersensitive response in rice and *Nicotiana benthamiana* and apoptosis in mammalian HeLa cells. The programmed cell death triggered by expression of *Xa10* was proposed to be a consequence of disruption of the endoplasmic reticulum (ER) and ER Ca^{2+} homeostasis (Tian et al., 2014). The natural *Xa10* rice plants only confer race-specific resistance to Xoo strains that harbor *avrXa10* or homologs. *Xa10* has been engineered for broad resistance by modifying the promoter regions to include multiple EBEs for trapping different TAL effectors from Xoo strains (Zeng et al., 2015). Functioning similarly as *Xa10*, the expression of the third executor R gene *Xa23*, encoding a protein with 50% identity as Xa10, also could trigger strong hypersensitive response. The induction of *Xa23* was dependent on AvrXa23, which is a widely present TAL effector in many extant Asian Xoo strains (Wang et al., 2015).

In addition to rice, pepper (*Capsicum annuum*) also has evolved executor R genes, namely, *Bs3* and *Bs4C-R*, against the pathogen *X. campestris* pv. *vesicatoria* (*Xcv*) strains that harbor the TAL effector AvrBs3 and AvrBs4, respectively (Römer et al., 2007, Strauß et al., 2012). AvrBs3 was shown to interacte directly with the promoter element of *Bs3*, leading to expression of *Bs3* and hypersensitive response at the site of infection (Römer et al., 2007). The *Bs3* product is related to the flavin-dependent mono-oxygenases (FMOs) family of proteins, including the YUCCA family, which, in turn, are believed to be associated with indole acetic acid biosynthesis in *Arabidopsis*. The mechanism for *Bs3*-promoted hypersensitive response is currently unknown (Römer et al., 2007). The second executor R gene from pepper, *Bs4C-R*, was isolated by transcriptome profiling *via* RNA-seq technology. The *Bs4C-R* gene, encoding a structurally unique protein, was only expressed in the presence of AvrBs4 from the pathogens while it was otherwise tightly regulated (Strauß et al., 2012). In addition to interaction with

Bs4C-R, AvrBs4 also interacts with the R gene *Bs4*, which is predicted to encode a nucleotide-binding leucine-rich repeat (NB-LRR) protein, from tomato (*Lycopersicon esculentum*). However, the *Bs4*-dependent response was independent of nuclear localization signal motifs (NLS) and acidic activation domain (AAD) in the C-terminus of AvrBs4. It is interesting to note that overexpression of *avrBs3* also triggered *Bs4*-mediated hypersensitive response (Schornack et al., 2004).

Rice has also evolved recessive resistance to TAL effector-mediated virulence through sequence polymorphisms at the promoters of the susceptibility genes to avoid binding of specific TAL effectors. Resistance rice plants in *xa13* is due to a ~200bp DNA fragment insertion in the effector binding element (EBE) of *OsSWEET11* gene to avoid the transcriptional activation by PthXo1 effector (Yang et al., 2006). A number of rice plants of *japonica* species have a 1-bp deletion in the EBE of *OsSWEET13* gene to abolish the PthXo2 effector-mediated gene induction (Liu et al., 2011, Zhou et al., 2015). A natural mutation in EBE region of *OsSWEET14* has yet to be identified. Rice plants containing the *xa5* resistance gene, which is a mutant allele on *OsTFIIAγ5* encoding small subunit of transcription components, may reflect an adaptation to reduce the efficiency of transcriptional activation of susceptibility genes by TAL effectors and therefore to evade infection by Xoo strains (Bogdanove et al., 2010, Iyer & McCouch, 2004, Sugio et al., 2007, White & Yang, 2009, Antony, 2010).

The application of TAL effectors

TAL effectors would remain restricted to the field of plant pathology or plant biology if not for the deciphering of the DNA recognition code. The code that repeat variable di-residues (RVD) of the repetitive region of TAL effectors specifically recognize the DNA in one RVD to one nucleotide correspondence is unparalleled in simplicity (Boch et al., 2009, Moscou &

Bogdanove, 2009). . This code sheds an exceptional light for the scientists to design artificial binding specificity of TAL proteins with desirable rearranged repeat RVDs in order to target any DNA nucleotides. The break-through leads to an outburst of genome editing and targeted gene transcription in many organisms (Miller et al., 2011; Zhang et al., 2011; Li et al., 2012; Li et al., 2013). The strong impact brought by the application of TAL effectors in the field of plant pathology or other plant sciences is reviewed here.

Engineered promoter elements of executor R genes that are recognized by a number of common TAL effectors could be applied for broad spectrum and durable resistance in plants (Hummel et al., 2012, Zeng et al., 2015). 16-bp EBE of AvrXa27 TAL effector transferred to the promoter region of *Bs3* gene could confer both AvrBs3 and AvrXa27 mediated HR reaction in tobacco plants (Römer et al., 2009). Five effector binding elements, corresponding to AvrXa10, AvrXa27, PthXo1, PthXo6 and PthXo7, were connected in tandem in the promoter region of Xa10 and enable the transgenic rice plants to be resistant to 27 of 28 strains collected worldwide (Zeng et al., 2015). In theory, concatenation of the predicted EBEs of conserved TAL effectors, such as AvrXa23, and those virulence effectors from different strains, such as AvrXa7, PthXo2, PthXo3 and TalC, in the promoter region of *Xa27*, *Xa10* or *Xa23* gene may provide another version of highly durable resistance rice plants. However, caution is still needed to avoid weak or leaky expression of the R gene, which is deleterious to plant growth and development (Zeng et al., 2015).

Targeted genome editing of specific promoter sequences of susceptibility genes in rice might impede the virulence contribution from the major TAL effectors and thereby thwart the disease development of Xoo strains (Li et al., 2012). *OsSWEET14* (also called *Os11N3*) is sugar efflux transporter gene targeted by AvrXa7, PthXo3, TalC and Tal5 from different strains for

virulence (Antony et al., 2010; Yu et al., 2011; Streubel et al., 2013; Chapter 2). The EBE of PthXo3 and AvrXa7 in the promoter region of *OsSWEET14* can be modified by a TAL effector nuclease (TALENs), which was comprised of a DNA binding domain (from designer TAL effector protein) and the DNA nuclease domain (from the FokI endonuclease protein) (Li et al., 2012). The specific rice plants with disruption of the EBE for AvrXa7 and PthXo3 were resistant to strains that depended on the two effectors for virulence. The marker and the TALEN genes can be removed by genetic crossing, resulting in modified rice plants with only desired mutations but free of marker and TALEN genes (Li et al., 2012). TALEN-mediated targeted genome editing in all three homoeoalleles of *MLO* genes in hexaploid wheat created heritable resistance to powdery mildew fungi (Wang et al., 2014).

Targeted transcriptional activation of host susceptibility or executor R genes results in either disease or resistance. The TAL effector DNA binding code enables scientist to design special TAL effectors to transcriptionally induce host genes of interest (Li et al., 2013; Streubel et al., 2013; Hu et al., 2014; Cernadas et al., 2014). The *xa27* gene in the susceptible rice plant IR24 was targeted by artificially designed TAL effector (dTALe), resulting in resistance to strains expressing the corresponding dTALe (Li et al., 2013) The promoter element of recessive resistance gene *xa13* was also targeted by a specific dTALe, and the plant became susceptible to the strain harboring the dTALe (Li et al., 2013). The application of dTALes also facilitates distinguishing between candidate susceptibility or executor R genes in crop plants that have been identified by transcription profiles. Seventeen paralogs of SWEET genes in rice could be individually targeted by dTALes, and only induction of one of five SWEET genes belonging to the clade III conferred susceptibility to infecting strains (Streubel et al., 2013). The targeted transcriptional induction of *CsLOB1* by strains expressing corresponding designer TAL

effectors, which having different EBE from native TAL effector, restored the virulence to the TAL mutants while the designer TAL of *CsSWEET1* could not. This demonstrated the *CsLOB1* as an important susceptibility gene in citrus canker disease (Hu et al., 2014). The same strategy was applied for identifying the *OsSULTR3;6* and *Bs4C-R*, which are a susceptibility gene in bacterial leaf streak disease of rice and an executor R gene in bacterial leaf spot of pepper, respectively (Cernadas et al., 2014, Strauß et al., 2012)

Conclusions and Objectives

Since the isolation of the first TAL effector gene, *avrBs3*, in 1989, twenty-years of exciting research on TAL effectors has changed our knowledge and perspectives on this special family of type III effector. TAL effectors are not just avirulence effectors interacting with host R genes, but also have critical virulence contributions by transcriptionally activating host susceptibility genes. TAL effectors not only have the nuclear localization motifs and transcription activation domains, but also possesses special DNA binding activity under a simple code that governs the DNA recognition by the RVDs of the repetitive regions of TAL effectors. TAL effectors not only function as transcriptional factors inside plant cells, but also can work as a modulators of gene expression in cells of other organisms. TAL effectors have also facilitated genome editing, when fused with nuclease or methylase at specific DNA target in the field of biotechnology (Miller et al., 2011).

A number of plant susceptibility genes and executor R genes that are targeted by TAL effectors have been identified and characterized (Table 1-1 and 1-2). Most of the genes are derived from the model systems of rice-*Xanthomonas oryzae* interaction and pepper-*Xanthomonas campestris* interaction. However, the research on identification of plant susceptibility genes is still incomplete, and also the virulence and avirulence roles of TAL

effectors and corresponding host gene targets in other plant-*Xanthomonas* interaction systems are largely unknown.

The bacterial blight disease is becoming severe in recent years due to the increasing rice cultivation without proper management. Fluorescent amplified fragment-length polymorphism (FAFLP) studies showed African Xoo strains were different from Asian Xoo strains. Southern blot analysis also revealed a reduced number of TAL effector genes (8-10) present in Africa strains compared to Asian strains, which typically contain 15-20 TAL genes. However, pathogenicity assays showed that African Xoo strains were also highly virulent on Asian rice varieties (Gonzalez et al., 2007). **The first objective of this study is to identify the susceptibility genes targeted by TAL effectors from African strain AXO1947.**

In addition, the PthXo2 effector from JXO1 strain of Xoo had been demonstrated as a major TAL effector that could complement the virulence loss of *pthXo1* mutant of $PXO99^{A}$ strain in IR24 rice (Yang & White, 2004). This finding indicated that PthXo2 may target the one of the SWEET gene in rice. However, the strain expressing the PthXo2 effector was not virulent on Nipponbare rice. **The second objective of this study is to identify the host target gene of PthXo2 effector and to elucidate the resistance from the Nipponbare rice plants compared to *indica* cultivar IR24.**

Furthermore, bacterial leaf streak (BLS) and black chaff disease of wheat, caused by *X. translucens* pv. *undulosa* (Xtu), is becoming prevalent in the northern Great Plains in recent years (Adhikari et al., 2012). Preliminary data of Southern blot analysis has revealed these strains harbor 7-8 TAL effector genes. To date, roles of TAL effectors in BLS of wheat are unknown. **The third objective of this study is to characterize the virulence contribution of**

TAL effectors in Xtu strain XT4699 in plant disease and identify wheat host genes targeted by specific TAL effectors.

References

Adhikari TB, Gurung S, Hansen JM, Bonman JM, 2012. Pathogenic and Genetic Diversity of Xanthomonas translucens pv. undulosa in North Dakota. *Phytopathology* **102**, 390-402.

Antony G, 2010. *Molecular basis of avrXa7 mediated virulence in bacterial blight of rice*: Kansas State University.

Antony G, Zhou J, Huang S, *et al.*, 2010. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell* **22**, 3864-76.

Bai J, Choi S-H, Ponciano G, Leung H, Leach JE, 2000. Xanthomonas oryzae pv. oryzae avirulence genes contribute differently and specifically to pathogen aggressiveness. *Molecular plant-microbe interactions* **13**, 1322-9.

Block A, Li G, Fu ZQ, Alfano JR, 2008. Phytopathogen type III effector weaponry and their plant targets. *Current opinion in plant biology* **11**, 396-403.

Boch J, Bonas U, 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology* **48**, 419-36.

Boch J, Scholze H, Schornack S, *et al.*, 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-12.

Bogdanove AJ, Koebnik R, Lu H, *et al.*, 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic Xanthomonas spp. *Journal of bacteriology* **193**, 5450-64.

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

Bonas U, Stall RE, Staskawicz B, 1989. Genetic and structural characterization of the avirulence gene avrBs3 from Xanthomonas campestris pv. vesicatoria. *Molecular and General Genetics MGG* **218**, 127-36.

Cernadas RA, Doyle EL, Niño-Liu DO, *et al.*, 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathog* **10**, e1003972.

Chen L-Q, Hou B-H, Lalonde S, *et al.*, 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527-32.

Chu Z, Yuan M, Yao J, *et al.*, 2006. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes & development* **20**, 1250-5.

Cohn M, Bart RS, Shybut M, *et al.*, 2014. Xanthomonas axonopodis Virulence Is Promoted by a Transcription Activator-Like Effector-Mediated Induction of a SWEET Sugar Transporter in Cassava. *Molecular plant-microbe interactions* **27**, 1186-98.

Cornelis GR, 2006. The type III secretion injectisome. *Nature Reviews Microbiology* **4**, 811-25.

Deslandes L, Rivas S, 2012. Catch me if you can: bacterial effectors and plant targets. *Trends in plant science* **17**, 644-55.

Duan Y, Castaneda A, Zhao G, Erdos G, Gabriel D, 1999. Expression of a single, host-specific, bacterial pathogenicity gene in plant cells elicits division, enlargement, and cell death. *Molecular plant-microbe interactions* **12**, 556-60.

Gonzalez C, Szurek B, Manceau C, Mathieu T, Séré Y, Verdier V, 2007. Molecular and pathotypic characterization of new Xanthomonas oryzae strains from West Africa. *Molecular plant-microbe interactions* **20**, 534-46.

Gu K, Yang B, Tian D, *et al.*, 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122-5.

Heuer H, Yin Y-N, Xue Q-Y, Smalla K, Guo J-H, 2007. Repeat domain diversity of avrBs3-like genes in Ralstonia solanacearum strains and association with host preferences in the field. *Applied and environmental microbiology* **73**, 4379-84.

Hopkins CM, White F, Choi S, Guo A, Leach J, 1992. Identification of a family of avirulence genes from Xanthomonas oryzae pv. oryzae. *Mol. Plant-Microbe Interact* **5**, 451-9.

Hu Y, Zhang J, Jia H, *et al.*, 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences* **111**, E521-E9.

Hummel AW, Doyle EL, Bogdanove AJ, 2012. Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**, 883-93.

Iyer AS, Mccouch SR, 2004. The rice bacterial blight resistance gene xa5 encodes a novel form of disease resistance. *Molecular plant-microbe interactions* **17**, 1348-54.

Kay S, Hahn S, Marois E, Hause G, Bonas U, 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-51.

Li T, Huang S, Zhou J, Yang B, 2013. Designer TAL effectors induce disease susceptibility and resistance to Xanthomonas oryzae pv. oryzae in rice. *Molecular plant* **6**, 781-9.

Li T, Liu B, Spalding MH, Weeks DP, Yang B, 2012. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature biotechnology* **30**, 390-2.

Li Z, Zou L, Ye G, *et al.*, 2014. A Potential disease susceptibility gene CsLOB of citrus is targeted by a major virulence effector PthA of Xanthomonas citri subsp. citri. *Molecular plant* **7**, 912-5.

Liu Q, Yuan M, Zhou Y, Li X, Xiao J, Wang S, 2011. A paralog of the MtN3/saliva family recessively confers race-specific resistance to Xanthomonas oryzae in rice. *Plant, cell & environment* **34**, 1958-69.

Miller JC, Tan S, Qiao G, *et al.*, 2011. A TALE nuclease architecture for efficient genome editing. *Nature biotechnology* **29**, 143-8.

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

Römer P, Hahn S, Jordan T, Strauß T, Bonas U, Lahaye T, 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**, 645-8.

Römer P, Recht S, Lahaye T, 2009. A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proceedings of the National Academy of Sciences* **106**, 20526-31.

Salanoubat M, Genin S, Artiguenave F, *et al.*, 2002. Genome sequence of the plant pathogen Ralstonia solanacearum. *Nature* **415**, 497-502.

Salzberg SL, Sommer DD, Schatz MC, *et al.*, 2008. Genome sequence and rapid evolution of the rice pathogen Xanthomonas oryzae pv. oryzae PXO99A. *BMC genomics* **9**, 204.

Schornack S, Ballvora A, Gürlebeck D, *et al.*, 2004. The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *The Plant Journal* **37**, 46-60.

Strauß T, Van Poecke RM, Strauß A, *et al.*, 2012. RNA-seq pinpoints a Xanthomonas TAL-effector activated resistance gene in a large-crop genome. *Proceedings of the National Academy of Sciences* **109**, 19480-5.

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**, 808-19.

Sugio A, Yang B, Zhu T, White FF, 2007. Two type III effector genes of Xanthomonas oryzae pv. oryzae control the induction of the host genes OsTFIIAγ1 and OsTFX1 during bacterial blight of rice. *Proceedings of the National Academy of Sciences* **104**, 10720-5.

Szurek B, Marois E, Bonas U, Van Den Ackerveken G, 2001. Eukaryotic features of the Xanthomonas type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. *The Plant Journal* **26**, 523-34.

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

Tian D, Yin Z, 2009. Constitutive heterologous expression of avrXa27 in rice containing the R gene Xa27 confers enhanced resistance to compatible Xanthomonas oryzae strains. *Molecular Plant Pathology* **10**, 29-39.

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.

Verdier V, Triplett LR, Hummel AW, *et al.*, 2012. Transcription activator-like (TAL) effectors targeting OsSWEET genes enhance virulence on diverse rice (Oryza sativa) varieties when expressed individually in a TAL effector-deficient strain of Xanthomonas oryzae. *New Phytologist* **196**, 1197-207.

Wang C, Zhang X, Fan Y, *et al.*, 2015. XA23 is an executor R protein and confers broad-spectrum disease resistance in rice. *Molecular plant* **8**, 290-302.

Wang Y, Cheng X, Shan Q, *et al.*, 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature biotechnology* **32**, 947-51.

White FF, Potnis N, Jones JB, Koebnik R, 2009. The type III effectors of Xanthomonas. *Molecular Plant Pathology* **10**, 749-66.

White FF, Yang B, 2009. Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction. *Plant Physiology* **150**, 1677-86.

Wu L, Goh ML, Sreekala C, Yin Z, 2008. XA27 depends on an amino-terminal signal-anchor-like sequence to localize to the apoplast for resistance to Xanthomonas oryzae pv oryzae. *Plant Physiology* **148**, 1497-509.

Yang B, Sugio A, White FF, 2005. Avoidance of host recognition by alterations in the repetitive and C-terminal regions of AvrXa7, a type III effector of Xanthomonas oryzae pv. oryzae. *Molecular plant-microbe interactions* **18**, 142-9.

Yang B, Sugio A, White FF, 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* **103**, 10503-8.

Yang B, White FF, 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular plant-microbe interactions* **17**, 1192-200.

Yang B, Zhu W, Johnson LB, White FF, 2000. The virulence factor AvrXa7 of Xanthomonas oryzae pv. oryzae is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proceedings of the National Academy of Sciences* **97**, 9807-12.

Yang Y, De Feyter R, Gabriel DW, 1994. Host-specific symptoms and increased release of Xanthomonas citri and X. campestris pv. malvacearum from leaves are determined by the 102-bp tandem repeats of pthA and avrb6, respectively. *MPMI-Molecular Plant Microbe Interactions* **7**, 345-55.

Yu Y, Streubel J, Balzergue S, *et al.*, 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**, 1102-13.

Zeng X, Tian D, Gu K, *et al.*, 2015. Genetic engineering of the Xa10 promoter for broad-spectrum and durable resistance to Xanthomonas oryzae pv. oryzae. *Plant biotechnology journal*.

Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P, 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature biotechnology* **29**, 149-53.

Zhou J, Peng Z, Long J, *et al.*, 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant Journal* **82**, 632-43.

Zhu W, Yang B, Chittoor JM, Johnson LB, White FF, 1998. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Molecular plant-microbe interactions* **11**, 824-32.

Zhu W, Yang B, Wills N, Johnson LB, White FF, 1999. The C terminus of AvrXa10 can be replaced by the transcriptional activation domain of VP16 from the herpes simplex virus. *The Plant Cell* **11**, 1665-74.

Table 1-1. Cloned susceptibility genes targeted by TAL effectors in plants.

Susceptibility genes	Gene functions	Comments	Plant sources	Corresponding TAL effector	Reference
OsSWEET11	Sugar efflux transporter	Also named *Os8N3/Xa13*, with promoter variants in xa13 recessive resistance	Rice	PthXo1	Chu et al., 2006; Yang et al., 2006
OsSWEET14	Sugar efflux transporter	Also named *Os11N3*, no promoter variants identified with recessive resistance	Rice	PthXo3, AvrXa7, TalC, Tal5	Antony et al., 2010; Yu et al., 2011; Streubel et al., 2013; Chapter 2
OsSWEET13	Sugar efflux transporter	Also named *Os12N3*, with promoter variants in *xa25* recessive resistance	Rice	PthXo2	Chapter 2; Zhou et al., 2015
OsSWEET12	Sugar efflux transporter	Induction by designer TAL effector results in susceptibility	Rice	No natural TAL effector	Li et al., 2013

OsSWEET15	Sugar efflux transporter	Induction by designer TAL effector results in susceptibility	Rice	No natural TAL effector	Streubel et al., 2013
OsTFX1	bZip family transcription factor	Loss of *OsTFX1* induction by *pthXo6* mutant results in 35% virulence loss	Rice	PthXo6	Sugio et al., 2007
OsTFIIAγ1	Subunit of transcription factor IIA	Induction by PthXo7 enables PXO86 strain to overcome *xa5* resistance	Rice	PthXo7	Sugio et al., 2007
OsSULTR3;6[1]	sulfate transporter	Induction of gene promotes bacterial spreading and exudation in bacterial leaf streak disease	Rice	Tal2g	Cernadas et al., 2014

MeSWEET10a	Sugar efflux transporter	Induction of gene enhances bacterial growth and water soaking symptoms	Cassava	$TAL20_{Xam668}$	Cohn et al., 2014
CsLOB1	LOB family transcription factor	Induction of gene enhances bacterial growth and pustule development	Citrus	PthA4, PthB, PthC, pthA*, $PthA^{w}$	Hu et al., 2014; Li et al., 2014
UPA20	bHLH family transcription factor	Induction of gene promotes tissue hypertrophy and cell enlargement	Pepper	AvrBs3, AvrHah1	Kay et al., 2007; Schornack et al., 2008

1, only one susceptibility gene identified in rice to bacterial leaf streak disease, other susceptibility genes in rice derived from bacterial blight disease.

Table 1-2. Cloned executor R genes mousetrapping TAL effectors in plants

Executor R gene	Comments	Plant sources	Corresponding TAL effector	References
Xa27	Low strength of HR, no homology to other R proteins	Rice	AvrXa27	Gu et al., 2005
Xa10	Strong HR, 50% identity with Xa23	Rice	AvrXa10	Tian et al., 2014
Xa23	Strong HR, 50% identity with Xa10	Rice	AvrXa23	Wang et al., 2014
Bs3	Strong HR, flavin-dependent mono-oxygenases protein	Pepper	AvrBs3	Romer et al., 2007
Bs4C-R	Strong HR, no homology to other R proteins	Pepper	AvrBs4	Strauβ et al., 2012

Chapter Two

Identification of Susceptibility Genes in Bacterial Blight of Rice

Abstract

Xanthomonas oryzae pv. *oryzae* (Xoo) is the causal agent of bacterial blight of rice in Asia and West Africa. One of major virulence mechanisms of Xoo is to transcriptionally activate the rice susceptibility (S) genes by transcription-activator like (TAL) effectors *via* the type III secretion system. Preliminary results showed that PthXo2 was a major TAL effector and induces the gene expression of *OsSWEET13* in an *indica* rice cultivar IR24 but not in Nipponbare, a *japonica* rice. Prediction of the effector binding element of PthXo2 reveals a 1-bp polymorphism between *indica* and *japonica* rice varieties in the promoter region of this *OsSWEET13* gene. PthXo2 effector virulence and *OsSWEET13* induction ability is altered by the fifth repeat deletion. Repeat deletion of *pthXo2* enables the switch of *OsSWEET13* induction and allelic resistance/disease susceptibility between *indica* and *japonica* rice varieties. The identification of rice susceptibility genes induced by African strain AXO1947, harboring a reduced number of TAL effector genes in comparison to Asian strains, was also attempted. Targeted mutagenesis of TAL effector genes was performed, and 14 different TAL effector mutants were obtained. Analysis of a mutant AME2-3 indicated a loss of the effector gene *talC*, and virulence assays of the mutant revealed a reduction in disease symptoms on rice plants. qRT-PCR of candidate host susceptibility genes showed reduced expression of *OsSWEET14* in comparison to expression upon challenge by AXO1947. The introduction of *talC* enables PXO99^{A} to defeat *xa13* resistance of rice. TAL effector genes *pthxo1* and *avrXa7* from Asian strains could complement the AME2-3 mutant for virulence. A second mutant AME8 of AXO1947, which can induce *Os11N3* and harbors an intact copy of *talC*, also has reduced virulence. AME8 loses the ability to induce the host genes *OsERF* and *OsTFX1*. PthXo6 effector from PXO99^{A} strain could restore the gene expression of *OsTFX1* and partially complement the virulence loss for AME8. The results

revealed the presence of cryptic recessive resistance in *japonica*-related rice cultivars against PthXo2-expressing Xoo strains and that African strains of Xoo target functionally similar host S genes with unique TAL effectors.

(Portions of this research were published in Zhou et al., 2015)

Introduction

Rice is a major staple crop that feeds much of populations throughout the world. Bacterial blight (BB) of rice is one of the most devastating diseases, especially in Asia and West Africa, which is also present in Australia, Latin America and the Caribbean (Mew et al., 1993). BB in some areas can cause yield loss as much as 50% in the appropriate conditions (Mew, 1987). BB is caused by the γ-proteobacterium *Xanthomonas oryzae* pv. *oryzae* (Xoo). Xoo invades rice plants through water pores or wounds and then moves systemically in the xylem tissue (Ou, 1985, Salzberg et al., 2008).

The Xoo-rice system is very important not only due to BB threats to world food security but also as a model system for research about plant-microbe interactions (NIÑO-LIU et al., 2006, White & Yang, 2009). The best-characterized virulence mechanism of Xoo is called the type III secretion systems (TTSS). The secreted effector proteins *via* TTSS are called type III effectors (T3E), which play major roles in pathogenicity (Block et al., 2008, Boch & Bonas, 2010, Bogdanove et al., 2010, Cornelis, 2006, White et al., 2009). As one group of T3E, Transcription-activator like (TAL) effectors are a structurally and functionally distinct family of proteins, which have DNA-binding activity, eukaryotic nuclear localization signal motifs and a potent acidic transcription activation domain in the C-terminal portion (Van den Ackerveken et al., 1996, Yang et al., 2000, Zhu et al., 1998, Zhu et al., 1999). TAL effectors have highly conserved N-terminal, C-terminal and central repetitive regions. The repetitive regions consist of 34-35 amino acid repeats ranging from 12.5 to 28.5 units in a single protein (White et al., 2009). All the repeats are almost identical with the exception of the 12^{th} and 13^{th} positions, which, together, are called the repeat variable di-residue (RVD) of each repeat. The RVD sequence of a given TAL effector determines the DNA binding specificity (Boch et al., 2009, Moscou & Bogdanove,

2009). Currently, TAL effectors are only found within the genus of *Xanthomonas* and *Ralstonia*, comprising the largest family of T3E in bacteria (White et al., 2009).

The contribution of TAL effectors to bacterial virulence, proliferation, and dissemination is conditioned by inducing expression of the host susceptibility genes (Antony et al., 2010, Cernadas et al., 2014, Cohn et al., 2014, Hu et al., 2014, Kay et al., 2007, Yang et al., 2006, Yu et al., 2011, Zhou et al., 2015). The first identified susceptibility gene in BB is the *OsSWEET11* (or *Os8N3*), which is targeted and induced by major virulence TAL effector PthXo1 (Chen et al., 2010, Yang et al., 2006). Natural mutations in the effector binding element (EBE) of PthXo1 in rice cultivars with the *xa13* gene impart resistance to Xoo strains that only rely on PthXo1 as the major virulence TAL effector (Chu et al., 2006, Yang et al., 2006). Second, the gene *OsSWEET14* (or *Os11N3*) is targeted by other major TAL effectors AvrXa7, PthXo3, Tal5 and TalC for susceptibility (Antony et al., 2010, Streubel et al., 2013, Yu et al., 2011). Furthermore, induced expression of any one of the five SWEET genes in the clade III in rice by designer TAL effectors brings about susceptibility, while elevated expression of other SWEET genes which belong to other clades does not (Streubel et al., 2013). Promoted expression of SWEET gene by TAL effector is also involved in cassava bacterial blight (Cohn et al., 2014).

Besides to the SWEET genes found to be targeted by TAL effectors, other susceptibility genes have been characterized. Two susceptibility genes *OsTFX1* and *OsTFIIAγ1*, which are targeted by TAL effector PthXo6 and PthXo7, respectively, are also involved in the BB disease. *OsTFX1* is a bZip family transcriptional factor gene that is up-regulated by most Xoo strains. Loss of induction of *OsTFX1* due to a *pthXo6* mutation causes 35% loss of virulence in lesion length development (Sugio et al., 2007). The *OsTFIIAγ1*, a gene encoding small subunit of the transcriptional factor IIA, is specifically induced by PXO99^{A} and the transfer of *pthXo7* enables

PXO86 strain to partially overcome the *xa5* resistance (Sugio et al., 2007). In bacterial leaf streak disease of rice, a TAL effector Tal2g specially targets the sulfate transporter gene and causes susceptibility by promoting lesion expansion and bacterial exudation (Cernadas et al., 2014). In citrus bacterial canker disease, the *lateral organ boundaries* 1 gene is specifically activated by PthA or functional equivalent TAL effectors for susceptibility, making critical contributions in bacterial growth and pustule formation (Hu et al., 2014, Li et al., 2014).

The disease susceptibility in BB is conditioned by induction of at least one of five SWEET sucrose transporter genes by TAL effectors from Xoo strains. The first two SWEET genes *OsSWEET11* and *OsSWEET14* have been reported to be targeted by natural TAL effectors (Antony et al., 2010, Streubel et al., 2013, Yang et al., 2006, Yu et al., 2011). Failure of induction of the third SWEET gene *OsSWEET13/Os12N3* is implicated in the *xa25* resistance in Minghui63 to strain PXO339 (Liu et al., 2011). The corresponding natural TAL effector to this SWEET gene is yet to be found. Although the difference in the promoter element and coding sequence on *xa25/Xa25* gene has been described (Liu et al. 2011), the resistance from the critical polymorphism has not been pinpointed. PthXo2, one of the major TAL effectors from JXO1, can restore the virulent loss of PXO99ME2 (mutant of major TAL gene *pthxo1*, hereafter ME2) and is the principle candidate as the cognate effector for *OsSWEET13/Os12N3* (Yang & White, 2004).

With increasing cultivation of rice in Africa, BB is becoming an emerging disease in some regions. The significant recurrence of BB in Africa maybe correlated with the expansion of rice crops without strict and effective quarantine control (Verdier et al., 2012). Fluorescent amplified fragment-length polymorphism (FAFLP) study reveals African Xoo strains are distinguishable from Asian Xoo strains. Southern blot analysis shows a reduced number of TAL

effectors present in Africa Xoo strains than in Asian strains (Gonzalez et al., 2007). It has been speculated that the African Xoo strains may be endemic to the continent (Verdier et al., 2012). Based on the pathogenicity assays on nearly isogenic lines of rice plants with different resistance genes, African Xoo strains have been divided into 3 races. BAI3 and AXO1947, strains used in this study, belong to the race 1 (Gonzalez et al., 2007). TalC from BAI3 is a major TAL effector that activates the *OsSWEET14* for disease susceptibility (Yu et al., 2011). Interestingly, TalC, as reported, could not enable the PXO99 strain to overcome the *xa13* resistance, while another major TAL effector AvrXa7 does (Yu et al., 2011; Antony et al., 2010). To gain insight to the virulence mechanism of African Xoo strains, another race 1 strain AXO1947 was used in this study. The possibility to discover other alternative major TAL effector and susceptibility genes is still intriguing.

Materials and Methods

Plant materials, bacterial strains, plasmids and growth conditions

Rice varieties Nipponbare (*Oryza sativa ssp. japonica*), IR24 and IRBB13 (*Oryza sativa ssp. indica*) were used in this study. *X. oryzae* pv. *oryzae* strains, *E. coli* strains and plasmids are listed in Table 2-1. All rice plants were grown in growth chamber with the temperature of 28 °C, relative humidity of 85% and a photoperiod of 12 hours. *E. coli* strains were grown in Luria-Bertani medium. *X. oryzae* pv. *oryzae* strains were grown in tryptone sucrose agar (TSA) medium (tryptone, 10 g/L, sucrose, 10 g/L, glutamic acid, 1 g/L) at 28°C. Antibiotics used in this study included carbenicillin (100 mg/L), kanamycin (50 mg/L), and spectinomycin (100 mg/L).

Construction of the fifth repeat deletion mutant of *pthXo2*

pZW-PthXo2 plasmid was first linearized by the restriction endonuclease *Hind*III, then partially digested by *Nco*I (NEB). The approximate 4 kb region of the digestion product was gel purified (Qiagen) and re-ligated by T4 ligase (NEB). The ligation product was then transformed into *E. coli*. The desired pZW-PthXo2D plasmid with the 4^{th} or 5^{th} repeat deletion was obtained by sequencing the cloned inserts from the resulting Ampicillin resistant colonies. The pZW-PthXo2D plasmid was linearized by *Hind*III and cloned into the pHM1 vector. The pHM1-PthXo2D plasmid was transformed into ME2 cells by electroporation (Bio-rad).

Microarray analysis and real time quantitative RT-PCR

Fourteen-day-old rice plants were inoculated with the indicated bacterial inoculum (OD=0.5 under 600nm) by the needleless syringe method. The inoculated leaves were harvested 24 h post inoculation, and RNA was extracted by using Trizol reagent (Invitrogen), following the protocols provided by the supplier. RNA quality was tested in Agilent chip. The Affymetrix Rice GeneChip was used for hybridization. Microarray analysis was performed by the Integrated Genomics Facility lab at Kanas State University. For real time quantitative PCR (qRT-PCR) assays, 1 microgram of RNA was pretreated with DNAseI (Invitrogen) and treated with reverse transcriptase for cDNA (Bio-rad). The resulting cDNA was diluted ten times and then used for real time qRT-PCR in the CFX-96 PCR machine (Bio-rad). Relative gene expression level was calculated with $2^{-\Delta\Delta Ct}$ method ((Livak & Schmittgen, 2001). Due to the polymorphisms of *OsSWEET13* between IR24 and Nipponbare, two different gene-specific primers were used in qRT-PCR assay. Primer pair IN-RT12N3-F & IN-RT12N3-R were used in IR24 while Nip-RT12N3-F &Nip-RT-12N3-R were applied for Nipponbare. Both primer pairs were tested with 100% efficiency in the primer efficiency assay. *OsTFIIAγ5* was used as internal control. Internal

control primer pairs RT-TF2-5F & RT-TF2-5R were also tested and had 100% efficiency with both Nipponbare and IR24 cDNA. Primer sequences are shown in Table 2-2.

Virulence assays

Four-week-old rice plants with expanded leaves were used for leaf tip-clipping assay (Kauffman et al., 1973). The bacteria was scraped off the TSA culture plates and suspended in sterilized water. The inoculum was adjusted to OD=0.5 under 600nm in Genesys20 spectrometer (thermo scientific). The leaf tip was clipped by scissors dipped in bacterial inoculum. The lesion length was measured several days after inoculation.

Construction of bacterial cosmid library and subcloning of TAL effector genes

The genomic DNA of AXO147 was partially digested by *Sau3*AI (NEB) for 10 minutes at 37^0C. The desired 20~30 kb fragments was harvested by gel electrophoresis, purified by electro-elution (Millipore) and ligated into pHC79, which was pretreated with *Bam*HI digestion and rSAP dephosphorylation (NEB). The ligation product was transferred by electroporation into DH5α competent cells (Invitrogen). The positive cosmid clones with TAL effector genes were identified by PCR with specific primers NTAL-F_{Xoo} and NTAL-R_{Xoo} (shown in Table 2-2). Cosmids with different TAL effector genes were differentiated by sequencing the repetitive region with the primer SeqF (shown in Table 2-2). The *Sph*I fragment of each TAL effector gene was subcloned into the pZW vector (Zhu et al., 1998), and the correct orientation of the *Sph*I fragment was determined by sequencing with primer SeqF.

Transposon mutagenesis and sequencing of repetitive regions of TAL effector genes

After the cloning of *Sph*I fragments of individual TAL effector genes of into pZW, the resulting plasmids were subjected to transposon mutagenesis using the EZ-Tn5 <Kan-2> *in vitro* transposon kit (Epicenter). The desired plasmids with transposon inserted in the repetitive

regions were identified by enzyme digestion. The primers Kan2 FP-1 and RP-1 located at both ends of transposon were applied for sequencing through the repetitive regions.

Southern blot analysis

The genomic DNA of AXO1947 or cosmid DNA was completely digested by *Bam*HI (NEB) and separated in 1% agarose gel by gel electrophoresis at 40 volt in cold room overnight. The gel portion with 2-7 kb DNA fragment was selected for transferring DNA to nylon membrane (Amersham). *Sph*I fragment of AvrXa7 was used for probe. The DNA labeling, hybridization and detection were following the protocols provided by GE healthcare ECL Kit.

Targeted mutagenesis and functional complementation of TAL effector genes

pK89-1 plasmid was obtained by transposon mutagenesis of pZWTAL8-9 and identified as 10^{th} repeat with single insertion. pK89-1 was electroporated into AXO1947 competent cells and positive clones were selected on TSA+Kanamycin plates. The resulting candidate mutants were tested on IR24 for virulence assay and subjected to Southern blot analysis for detection of TAL gene mutation. Candidate TAL effector genes in the pZW vector were then cloned into pHM1 vector and transformed into corresponding TAL mutants by electroporation or conjugation for complementation.

Results

Virulence function of PthXo2 is associated with *OsSWEET13* in rice cultivar IR24 not in Nipponbare

The gene *pthXo2*, from JXO1 strain, was previously cloned by functional complementation for ME2 to restore bacterial virulence in the rice cultivar IR24 (Yang and White, 2004). Further testing revealed that ME2/*pthXo2* was not compatible on the cultivar

Nipponbare (Figure 2-1A). Microarray analysis of gene expression by ME2/*pthXo2* relative to ME2 showed no induction of *OsSWEET13* in the presence of PthXo2 in Nipponbare, while *OsSWEET13*, as represented by probe set Os.5491.1.A1_s_at, was induced approximately 34 fold by ME2/*pthXo2* compared to ME2 in IR24 (Table 2-3). Consistently, qRT-PCR revealed *OsSWEET13* was induced by PthXo2-expressing strain in IR24 but not in Kitaake, another *japonica* cultivar with similar genetic background as Nipponbare (Figure 2-1B). We speculated that the failure in induction of *OsSWEET13* by ME2/*pthXo2* in Nipponbare or Kitaake was attributable to variation in the promoter region. As predicted from the code of DNA binding specificity by RVDs of TAL effector, there was one base pair difference on the putative effector binding element of PthXo2 between IR24 and Nipponbare (Figure 2-1C).

Deletion of the fifth repeat of PthXo2 enables allelic switch of *OsSWEET13* induction for resistance and disease susceptibility

The deduced OsSWEET13/Os12N3 proteins from Nipponbare and IR24 differ by eight amino acids in the C-terminal (Figure 2-2). To determine whether *OsSWEET13* in Nipponbare or Kitaake can function as a potential bacterial blight susceptibility gene for *X*. *oryzae* pv. *oryzae* strains, the *pthXo2* mutant *pthXo2D* was constructed wherein the fifth repeat of the central repetitive region was deleted, leading to a new arrangement of RVDs that matched the EBE in Nipponbare (or Kitaake) but not that in IR24 (Figure 2-3). The introduction of the TAL effector gene *pthXo2* into ME2 could restore virulence on IR24, while the fifth repeat deletion from *PthXo2* resulted in the loss of virulence on IR24 (Figure 2-4A). However, the Nipponbare leaves were curled, a state of disease susceptibility, by the inoculation of ME2/*pthxo2D*. In contrast, the ME2/*pthXo2*, like ME2/pHM1, is almost non-virulent on Nipponbare leaves. (Figure 2-4B). Further measurement of the lesion length also revealed the switch of resistance and disease

susceptibility (Figure 2-4C). Consistently, the repeat deletion derivative *pthXo2D* gained the ability to induce the gene expression of *OsSWEET13* in Nipponbare but lost such ability in IR24 when introduced into ME2 (Figure 2-5).

AXO1947 harbors eight TAL effector genes

Twelve cosmids, containing TAL effector genes, from an approximately 1200-member library of AXO1947 were characterized by PCR screening and Southern hybridization analysis. Eight TAL effector genes were recovered by Southern hybridization analysis (Figure 2-6 A), which is consistent with previous results on other African strains (Gonzalez et al., 2007). The RVD sequence of one TAL gene that was subcloned from cosmid 12 was identical in sequence to *talC* from strain BAI3 (Yu et al., 2011). The RVD sequence of other TAL effector genes is also shown (Figure 2-6 B). Considering the AXO1947 was incompatible with IRBB7, it is interesting to know there was no TAL effector gene having similar RVD as *avrXa7*.

Mutation of TAL effector genes in AXO1947 results in reduction of virulence

Fourteen different TAL effector gene mutants were obtained by targeted mutagenesis. The Southern blot analysis revealed the loss of fragments corresponding to individual TAL effector genes in comparison to the pattern for the wild type AXO1947 (Figure 2-7). Virulence assays of the 14 mutants showed AME2-3 had a severe reduction in lesion length development, while AME8 had around 50% loss of virulence (Figure 2-8). From the Southern blot analysis, AME2-3 had lost the third band corresponding the internal *SphI* fragments (Figure 2-7A, from top to bottom), which corresponds in size to the *talC*. While there were loss of two TAL genes in AME8, *TAL8-9* and *TAL8-2*, the gene *talC* was intact.

TalC of AXO1947 targets the SWEET gene *OsSWEET14/Os11N3* in rice

Microarray data indicated that the expression level of *OsSWEET14/Os11N3* gene in rice was highly elevated upon the inoculation of AXO1947 compared to Asian strain PXO99^{A} (Table 2-4). The loss of *talC* effector gene in AME2-3 was associated with the loss in the ability to induce *OsSWEET14/Os11N3*, while the other TAL mutants, AME2-1, AME2-2, AME8 and WT, harboring *talC*, strongly induced *OsSWEET14/Os11N3* expression (Figure 2-9). This observation is consistent with previous semi-quantitative RT-PCR results (Yu et al. 2011).

TalC effector enables PXO99^{A} to overcome the *xa13* resistance and *pthxo1* and *avrXa7* effector genes complement the loss of *talC* in mutant AME2-3

Virulence loss of ME2 can be complemented by any one of several major TAL effectors that target the same or related SWEET genes (Yang and White, 2004; Antony et al., 2010). *avrXa7* and *pthXo3*, major TAL effector genes, enable PXO99^{A} to defeat the *xa13* resistance (Antony et al., 2010). *talC*, major TAL effector gene from AXO1947, also enabled PXO99^{A} to defeat the *xa13* resistance (Figure 2-10 A), which is consistent with the observations mentioned above. Transformation of *talC* into ME2 also enabled ME2 to become more virulent on IR24 plants (Figure 2-10 B). Transfer of *pthXo1* and *avrXa7*, two major TAL effector genes from Asian strains, also restored the virulence loss of AME2-3 while the complementation of *talC* partially restored the loss (Figure 2-10 C).

AME8 loses the ability to induce *OsERF* and *OsTFX1*

AME8 lost approximately 50% of virulence in the lesion length measurement assays in comparison to the wild type strain (Figure 2-8), while retaining the *OsSWEET14/Os11N3* gene expression (Figure 2-9 and Figure 2-11). To investigate the host genes that are targeted by the mutated TAL effector genes in the AME8, microarray analysis was performed on leaves

inoculated with AXO1947 and AME8. The top 10 differently induced genes (>10 fold) between WT and AME8 are listed in Table 2-5. Expression of *OsTFX1* was elevated in leaves inoculated by AXO1947 but not in leaves treated with AME8. *OsTFX1* is a bZip transcriptional factor family gene and is targeted by PthXo6 from the Asian strain PXO99^{A} for disease susceptibility (Sugio et al. 2007). From the microarray data shown in Table 2-4, *OsERF* was the most highly differentially induced gene by PXO99 and AXO1947. Compared to wild type inoculation, AME8 lost the ability to induce the *OsERF* gene expression. The difference of in gene expression induced by AXO1947 vs. AME8 was corroborated by the qRT-PCR data (Figure 2-11). AME2-3, which is missing *talC*, still retained the ability to induce the *OsTFX1* and *OsERF*.

Loss of virulence in AME8 mutant can be partially complemented by *pthXo6* of Asian Xoo strain

Previous results had showed that *pthXo6* mutant of Asian strain PXO99A ME1 (hereafter ME1) lost the ability to induce the *OsTFX1* and had a 35% virulence loss in terms of lesion length development assays (Sugio et al., 2007). Considering that AME8 also failed to trigger *OsTFX1* gene expression, the transformation of *pthXo6* into AME8 was performed. In the Southern blot analysis, AME8 had the missing bands from both *TAL8-9* and *TAL8-2* effector genes (Figure 2-7). Accordingly, *TAL8-9* was cloned into pHM1 and transferred into AME8 and ME1, respectively, to restore the gene expression level of *OsTFX1*. Only when *pthXo6* was transferred to AME8 (strains AME8/pthXo6#7 and AME8/pthXo6#10) enhanced the expression of *OsTFX1*, while the *TAL8-9*, when transferred into AME8 and ME1, failed to enhance *OsTFX1* expression (Figure 2-12 A). The two *pthXo6* transformants of AME8 also became more virulent on Nipponbare rice than the AME8 mutant (Figure 2-12 B).

Discussion

Through this study, the sugar transporter gene *OsSWEET13* in rice has been identified to be specifically targeted by natural TAL effector PthXo2 from *X. oryzae* pv. *oryzae* JXO1 strain. The transcriptional induction of *OsSWEET13* allele in *indica* species IR24 but not in *japonica* species results from the single nucleotide deletion in the EBE of PthXo2 in *japonica* rice plants. Although polymorphisms of promoter elements and coding regions were identified between *Xa25* and *xa25* rice plants, the critical 1 bp deletion in the EBE of *xa25* rice plant was not pinpointed (Liu et al., 2011). According to the code deciphering the DNA recognition specificity, the 1 bp deletion brings about the improper alignment between RVDs of PthXo2 and DNA nucleotides in the promoter region of *OsSWEET13*, and, as a result, induction of this susceptibility gene in *japonica* rice species is lost. The results also demonstrate the critical 1 bp deletion accounts for the recessive resistance. First, transcriptional activation of the *OsSWEET13* gene in Nipponbare by PthXo2 derivatives PthXo2D leads to host susceptibility, indicating that induction of this sugar transporter gene in Nipponbare cultivar is sufficient for disease. Second, loss of a single RVD in PthXo2D, corresponding to the single 1 bp nucleotide deletion in the EBE region of *OsSWEET13,* converts the effector to a virulence factor in *japonica* rice cultivars with the concomitant loss of function in IR24. A number of *japonica* rice plants had been tested with the 1 bp deletion in the EBE of PthXo2, indicating the *xa25* recessive resistance gene is not rare in rice cultivars (Zhou et al., 2015). The hitherto unknown presence of the resistance allele in *japonica* varieties may reflect selection of cultivars that were resistant to BB in early breeding efforts.

Sequence of all the repetitive fragments from the cloned TAL effector genes in AXO1947 shows there is an identical TAL effector gene as *talC* from African strain BAI3 (Yu et

al., 2011). The result that *talC* mutant AME2-3 fails to induce the expression of *OsSWEET14/Os11N3* and thereby has severe virulence loss is consistent with observations by Yu et al. (2011) and Antony et al. (2010). However, our data indicates the Asian and African major TAL effector genes are interchangeable in virulence contribution, whereas Yu et al. (2011) were unable to restore virulence with other major TAL effector genes. Major TAL genes *pthXo1* and *avrXa7* from Asian strains can complement the virulence loss of *talC* mutant and, conversely, the *talC* from AXO1947 enables the PXO99^{A} strain to overcome the *xa13* resistance and restores the virulence loss of *pthXo1* mutant. The reason why the *pthXo1* and *avrXa7* could not complement the *talC* mutant of BAI3 of Yu et al. (2011) is probably a technical problem. Only 2 of 20 transformants of AME2-3 with pHM1-*avrXa7* were positive and virulent on IR24 rice plants (data not shown). We speculate the restriction and modification systems in AXO1947 affect the transfer of foreign plasmids, especially for clones with highly repetitive sequences, and may cause rearrangements and deletions in the transfer process (data not shown).

In the lesion length measurement assay, we identified another TAL effector gene mutant AME8 that had a virulence defect compared to the wild type strain. According to the microarray data, AME8 failed to induce expression of a number of genes, including *OsERF* and *OsTFX1*. *OsTFX1* is a susceptibility genes targeted by PthXo6 TAL effector from the Asian strain PXO99^{A} (Sugio et al., 2007). Failure to activate gene expression of *OsTFX1* by the *pthXo6* mutant lead to a 35% loss of virulence in lesion development assays (Sugio et al., 2007). Our observations that *pthXo6* partially restored the virulence loss and the gene expression of *OsTFX1* after transfer to AME8 indicates that the effectors missing in AME8 directly or indirectly enhance *OsTFX1* expression. According to the Southern blot analysis, AME8 lost the *TAL8-9* and *TAL8-2* genes compared to the WT. However, complementation of either one into *pthXo6*

mutant ME1 or AME8 failed to elevate the gene expression of *OsERF* and *OsTFX1*. Furthermore, no candidate EBEs for *TAL8-9* and *TAL8-2* could be found in the promoter regions of *OsERF* and *OsTFX1* (Doyle et al., 2012). One possible reason that may account for the failure in complementation and DNA targeting prediction, is that AME8 may have mutations in other TAL effector genes that are not detectible in the Southern hybridization analysis. It is not rare to observe the multiple TAL gene mutations in one mutant using the homologous recombination strategy. The ME1 mutant of PXO99^{A} harbored the mutations of at least two TAL effector genes, namely *avrXa27* and *pthXo6* (Gu et al., 2005, Sugio et al., 2007). Similar observations were made during the process of knocking out TAL genes in *X. oryzae* pv. *oryzicola* (Cernadas et al., 2014).

The heterogeneous transfer of *pthXo6* only partially restored the reduced virulence of AME8, which brings about one interesting question. Is *OsERF* a susceptibility gene in bacterial blight? The gene annotation indicates that *OsERF* encodes an ethylene responsive transcription factor, which are predicted to be involved in regulation process of pathogenesis (http://signal.salk.edu/cgi-bin/RiceGE). Using a strategy that has been applied to identify susceptibility genes in citrus canker and bacterial leaf streak disease (Cernadas et al., 2014, Hu et al., 2014), specific artificial TAL effector genes corresponding to the promoter region of *OsERF* could be designed and tested for their virulence contribution when transformed into AME8.

References

Antony G, Zhou J, Huang S, *et al.*, 2010. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell* **22**, 3864-76.

Block A, Li G, Fu ZQ, Alfano JR, 2008. Phytopathogen type III effector weaponry and their plant targets. *Current opinion in plant biology* **11**, 396-403.

Boch J, Bonas U, 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology* **48**, 419-36.

Boch J, Scholze H, Schornack S, *et al.*, 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-12.

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

Cernadas RA, Doyle EL, Niño-Liu DO, *et al.*, 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathog* **10**, e1003972.

Chen L-Q, Hou B-H, Lalonde S, *et al.*, 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527-32.

Chu Z, Yuan M, Yao J, *et al.*, 2006. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes & development* **20**, 1250-5.

Cohn M, Bart RS, Shybut M, *et al.*, 2014. Xanthomonas axonopodis Virulence Is Promoted by a Transcription Activator-Like Effector-Mediated Induction of a SWEET Sugar Transporter in Cassava. *Molecular plant-microbe interactions* **27**, 1186-98.

Cornelis GR, 2006. The type III secretion injectisome. *Nature Reviews Microbiology* **4**, 811-25.
Doyle EL, Booher NJ, Standage DS, *et al.*, 2012. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic acids research* **40**, W117-W22.

Gonzalez C, Szurek B, Manceau C, Mathieu T, Séré Y, Verdier V, 2007. Molecular and pathotypic characterization of new Xanthomonas oryzae strains from West Africa. *Molecular plant-microbe interactions* **20**, 534-46.

Gu K, Yang B, Tian D, *et al.*, 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122-5.

Hopkins CM, White F, Choi S, Guo A, Leach J, 1992. Identification of a family of avirulence genes from Xanthomonas oryzae pv. oryzae. *Mol. Plant-Microbe Interact* **5**, 451-9.

Hu Y, Zhang J, Jia H, *et al.*, 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences* **111**, E521-E9.

Kauffman H, Reddy A, Hsieh S, Merca S, 1973. Improved technique for evaluating resistance of rice varieties to Xanthomonas oryzae. *Plant Disease Reporter*.

Kay S, Hahn S, Marois E, Hause G, Bonas U, 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-51.

Li Z, Zou L, Ye G, *et al.*, 2014. A Potential disease susceptibility gene CsLOB of citrus is targeted by a major virulence effector PthA of Xanthomonas citri subsp. citri. *Molecular plant* **7**, 912-5.

Liu Q, Yuan M, Zhou Y, Li X, Xiao J, Wang S, 2011. A paralog of the MtN3/saliva family recessively confers race-specific resistance to Xanthomonas oryzae in rice. *Plant, cell & environment* **34**, 1958-69.

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods* **25**, 402-8.

Mew T, 1987. Current status and future prospects of research on bacterial blight of rice. *Annual review of phytopathology* **25**, 359-82.

Mew T, Alvarez A, Leach J, Swings J, 1993. Focus on bacterial blight of rice. *Plant disease* **77**, 5-12.

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

Niño-Liu DO, Ronald PC, Bogdanove AJ, 2006. Xanthomonas oryzae pathovars: model pathogens of a model crop. *Molecular Plant Pathology* **7**, 303-24.

Ou SH, 1985. *Rice diseases*. IRRI.

Salzberg SL, Sommer DD, Schatz MC, *et al.*, 2008. Genome sequence and rapid evolution of the rice pathogen Xanthomonas oryzae pv. oryzae PXO99A. *BMC genomics* **9**, 204.

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**, 808-19.

Sugio A, Yang B, Zhu T, White FF, 2007. Two type III effector genes of Xanthomonas oryzae pv. oryzae control the induction of the host genes OsTFIIAγ1 and OsTFX1 during bacterial blight of rice. *Proceedings of the National Academy of Sciences* **104**, 10720-5.

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.

Verdier V, Cruz CV, Leach JE, 2012. Controlling rice bacterial blight in Africa: needs and prospects. *Journal of biotechnology* **159**, 320-8.

White FF, Potnis N, Jones JB, Koebnik R, 2009. The type III effectors of Xanthomonas. *Molecular Plant Pathology* **10**, 749-66.

White FF, Yang B, 2009. Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction. *Plant Physiology* **150**, 1677-86.

Yang B, Sugio A, White FF, 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* **103**, 10503-8.

Yang B, White FF, 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular plant-microbe interactions* **17**, 1192-200.

Yang B, Zhu W, Johnson LB, White FF, 2000. The virulence factor AvrXa7 of Xanthomonas oryzae pv. oryzae is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proceedings of the National Academy of Sciences* **97**, 9807-12.

Yu Y, Streubel J, Balzergue S, *et al.*, 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**, 1102-13.

Zhou J, Peng Z, Long J, *et al.*, 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant Journal* **82**, 632-43.

Zhu W, Yang B, Chittoor JM, Johnson LB, White FF, 1998. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Molecular plant-microbe interactions* **11**, 824-32.

Zhu W, Yang B, Wills N, Johnson LB, White FF, 1999. The C terminus of AvrXa10 can be replaced by the transcriptional activation domain of VP16 from the herpes simplex virus. *The Plant Cell* **11**, 1665-74.

Table 2-1. Strains and plasmids used in this study

Strain or plasmid	Comment	Reference or sources
Xoo strains		
PXO99[A]	5-Azacytidine resistant, race 6	(Hopkins et al., 1992)
ME2	pthXo1 mutant of PXO99[A]	Yang and White, 2004
ME1	pthXo6 mutant of PXO99[A]	Sugio et al., 2007
ME2/PthXo2	ME2 with pZWpthXo2 in pHM1	Yang and White, 2004
ME2/PthXo2D	ME2 with pZWpthXo2D in pHM1	This study
AXO1947	African Xoo strains, race 1	Gonzalez et al., 2007
AME8	TAL effector mutant of AXO1947	This study
AME2-3	TAL effector mutant of AXO1947	This study
AME8/PthXo6	AME8 with pZWpthXo6 in pHM1	This study
AME2-3/TalC	AME2-3 with pZWtalC in pHM1	This study
AME2-3/PthXo1	AME2-3 with pZWpthXo1 in pHM1	This study
AME2-3/AvrXa7	AME2-3 with pZWavrXa7 in pHM1	This study
***E. coli* strains and plasmids**		
DH5α	F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ-	Invitrogen
pHC79	Resistance to carbenicillin (Cb^r)	Yang and White, 2004
pHM1	Broad host range vector, resistance to spectinomycin, polylinker, cos site	Hopkins et al., 1992
pZW	Repeat deletion by sphI digeston of pZWavrXa10 vector, Cb^r	Zhu et al., 1998

Table 2-2. Primer sequences used in this study

NTAL-F_{Xoo}	5’-ATCTGCTCCGTCAGTTCGATCC
NTAL-R_{Xoo}	5’-AACTGTAACGGCGGACCTCTC
SeqF	5’-AGTGCATGCATGGCGCAATGC
SeqR	5’-CGAGATAACTGGGCAACAATGCTC
Kan2 FP-1	5’-ACCTACAACAAAGCTCTCATCAACC
Kan2 RP-1	5’-GCAATGTAACATCAGAGATTTTGAG
IN-RT12N3-F	5’- AGACAACAGCTACCACCAA
IN-RT12N3-R	5’- GTGATTGATTGATAAGAGGAGACA
Nip-RT12N3-F	5’-AGCAACTTCTTCACCTATATTCTC
Nip-RT12N3-R	5’-GTTGGGATTGATTGATAAGAGGAG
RT-TF2-5F	5′-GGGTTTGCCTGGTATTTGTTAG-3’
RT-TF2-5R	5′-GTTGCTGCTGTGATATACTCTG-3′
RT-47ERF-F1	5’-GCTCTTCATCTTCCTCTTCCG
RT-47ERF-R1	5’- CTTGTTGGTGGTGTGGTCTC
RT-11N3-F	5’- ATCAAGCCTTCAAGCAAAGC
RT-11N3-R	5’- CTAGGAGACCAAAGGCGAAG
RT-TFX1-F3	5’-CTTCTGGCCCTCTAAGAGTTCTCC
RT-TFX1-R3	5’-GGTGATCATGTTGCTGTGGTAGTG

Table 2-3. Microarray analysis of gene expression from rice leaves inoculated with ME2 or ME2(*pthXo2*).

		ME2[a]	ME2(*pthXo2*)[a]	Fold change[b]
OsSWEET13	Nip	1024.8	1462.2	1.4
	IR24	313.8	10629.4	33.9
OsTFIIY5	Nip	4958.9	5275.6	1.1
	IR24	6007.0	5597.8	0.9

a, The score is the average of three samples.
b, Fold change is calculated by dividing the average score of ME2(*pthXo2*) by that of ME2.

Table 2-4. Microarray analysis of rice gene expression induced differently by strains

Affymetrix ID	Mutant rice line	Comments	1947 Fold‡	PXO99^{A} Fold ‡	1947 Score*	PXO99^{A} Score *
1. OsAffx-6536-1	PFG_1A-21411	Ethyl. Resp. Factor	49	1.0	19725	400
2. Os-4974-1-S1	PFG_3D-03008	Os11N3 (N3)	35	1.0	13932	400
3. Os-56924-1-S1	RdSpm1307_3	nodulin	34	1.0	13711	400
4. Os-6863-1-S1	PFG_3A-08573;	beta-gluc. agg	32	0.1	12844	400
5. Os-4974-1-S1	PFG_3D-03008	Os11N3 (N3)	30	1.0	12140	400
6. Os-5808-1-S1	RMD_03Z11GC34	delta-6-desat.	28	1.0	11350	400
7. Os-323-1-S1	NONE	PRB1-3	28	1.0	11119	400
8. Os-9417-1-S1	PFG_2B-40423	Ribonuc. T2	28	0.1	18913	400
9. Os-15830-1-S1	DAG3C01;	Ala-glyoxylate	26	2.8	10564	1116
10.Os-6863-1-S1	PFG_3A-08573;	beta-gluc agg.	25	0.1	10101	400

*, all the scores of rice gene expression induced by AXO1947, PXO99^{A} and PXO99^{A}T3S^{-} were from three replicates and the lowest scores were normalized to 400

‡, the fold change was calculated by dividing the scores of indicated strains with that of PXO99^{A}T3S^{-}.

Table 2-5. Microarray analysis reveals *OsERF* and *OsTFX1* were not induced by AME8 compared to WT inoculation

Probe ID	1947 Ave [a]	AME8 Ave [a]	1947/AME8 Fold [b]	Annotation
Os.49201.1.S1_at	4617.5	34.6	35	*OsTFX1*
OsAffx.6536.1.S1_at	11991.9	253.8	34.3	*OsERF*
Os.43896.1.S1_at	6651.8	209.9	21.9	Putative ripening regulated protein
Os.5025.1.S1_at	3100.2	114.2	14.9	Nitrate-induced NOI family protein
Os.10507.1.S1_at	3603.2	158.4	14.3	dehydration response related protein
Os.36393.1.S1_at	2683.7	116.3	12.9	retrotransposon protein
Os.7530.1.S1_at	1343.4	17.5	12.3	beta-expansin precursor
Os.8667.1.S1_at	9533.5	819.5	10.5	KH domain-containing protein
Os.39994.1.S1_at	5244.3	426.5	10.1	retrotransposon protein
Os.14187.1.S1_at	4417.1	360.1	9.8	erythronate-4-phosphate dehydrogenase domain containing protein

a, all the scores of rice gene expression were averaged from scores of 3 biological replicates;
b, the fold is calculated by the fomula Fold=Scores of AXO1947+100/Scores of AME8+100

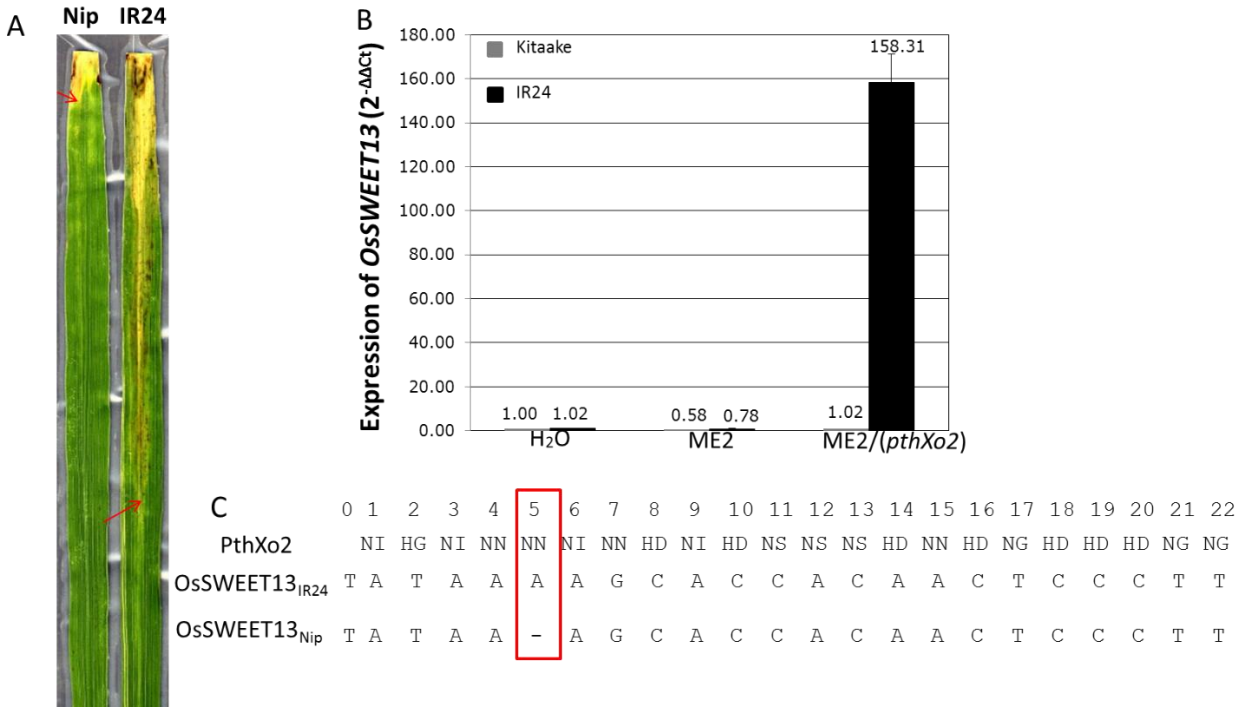


Figure 2-1. Association of blight susceptibility and resistance in IR24 and Nipponbare with PthXo2-expressing strains. A, Phenotypes of blight resistance (left leaf, Nip) and disease (right leaf, IR24) in Nipponbare and IR24 as indicated at the right side of leaves. Arrows present the edges of disease progression. Nip, Nipponbare. B, *OsSWEET13* is induced by PthXo2-expressing strain. Total leaf RNA from 3-week old rice cultivars Kitaake and IR24 inoculated with respective strains were used for real-time quantitative RT-PCR with gene-specific primers. Kitaake has the same *OsSWEET13* promoter region as Nipponbare (data not shown). The fold changes relative to water treatment are shown by using the $2^{-\Delta\Delta Ct}$ method. C, Alignment of PthXo2 RVDs with the predicted *OsSWEET13* promoter EBEs from IR24 ($OsSWEET13_{IR24}$) and Nipponbare ($OsSWEET13_{Nip}$). The dashed line in $OsSWEET13_{Nip}$ denotes the 1-bp deletion compared to $OsSWEET13_{IR24}$.

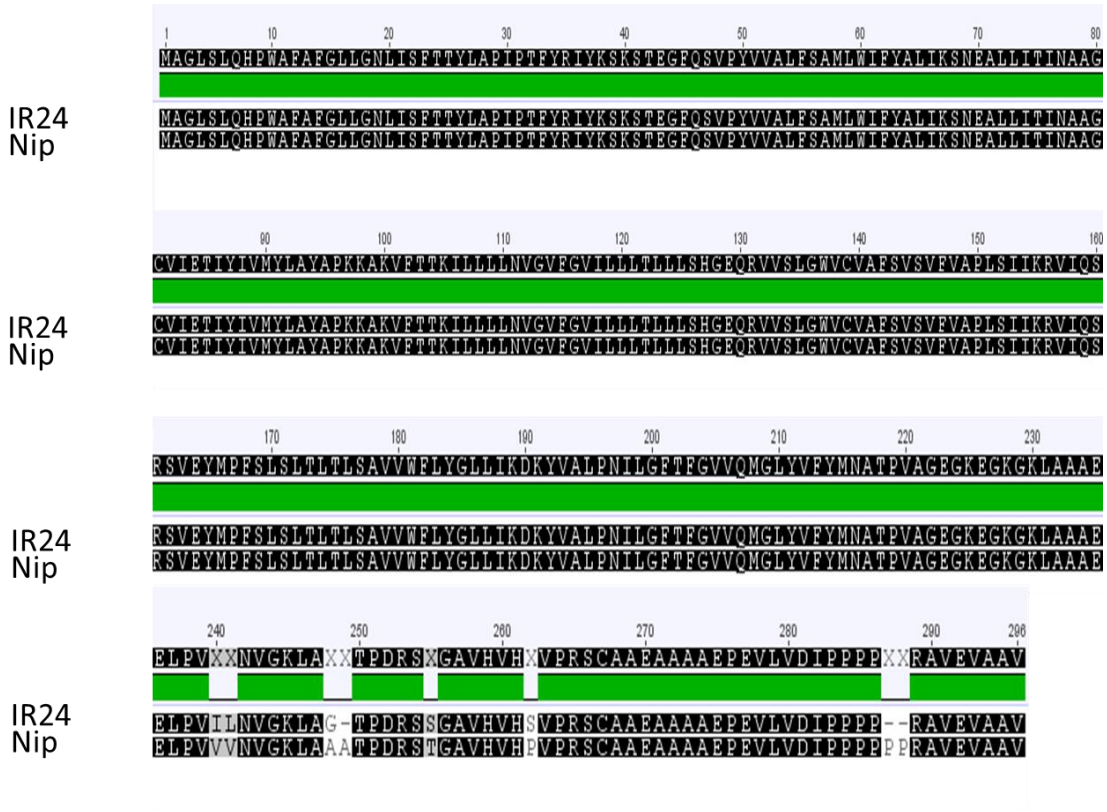


Figure 2-2. Predicted amino acid sequence alignment of OsSWEET13 between IR24 and Nipponbare. Eight amino acid residues, all in the C-terminal portion, are different between IR24 and Nipponbare.

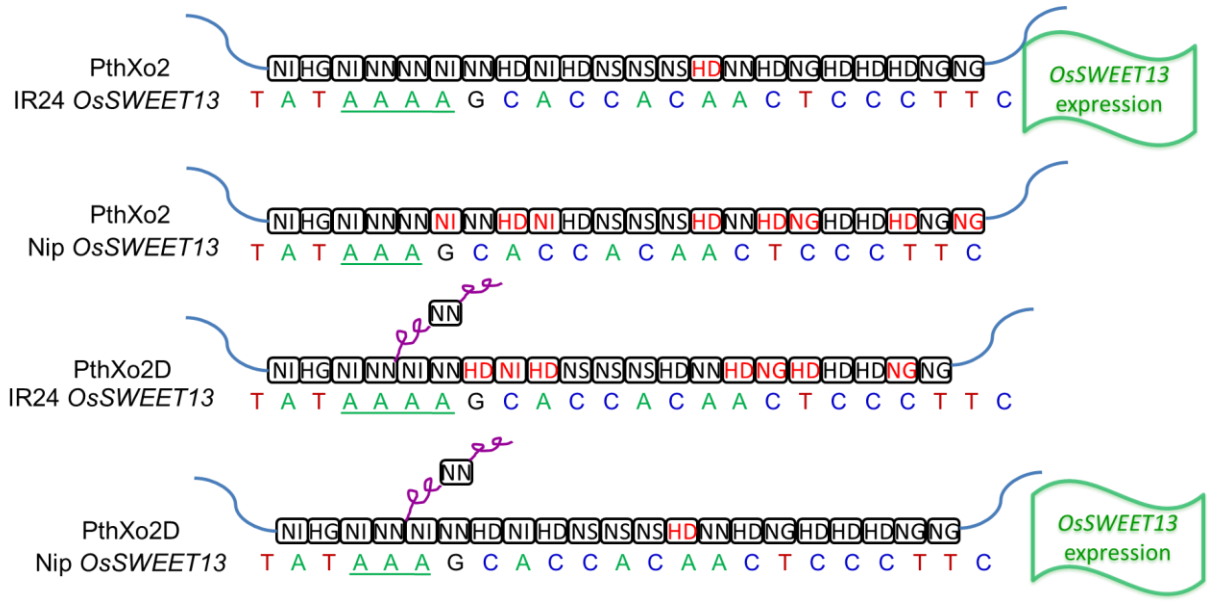
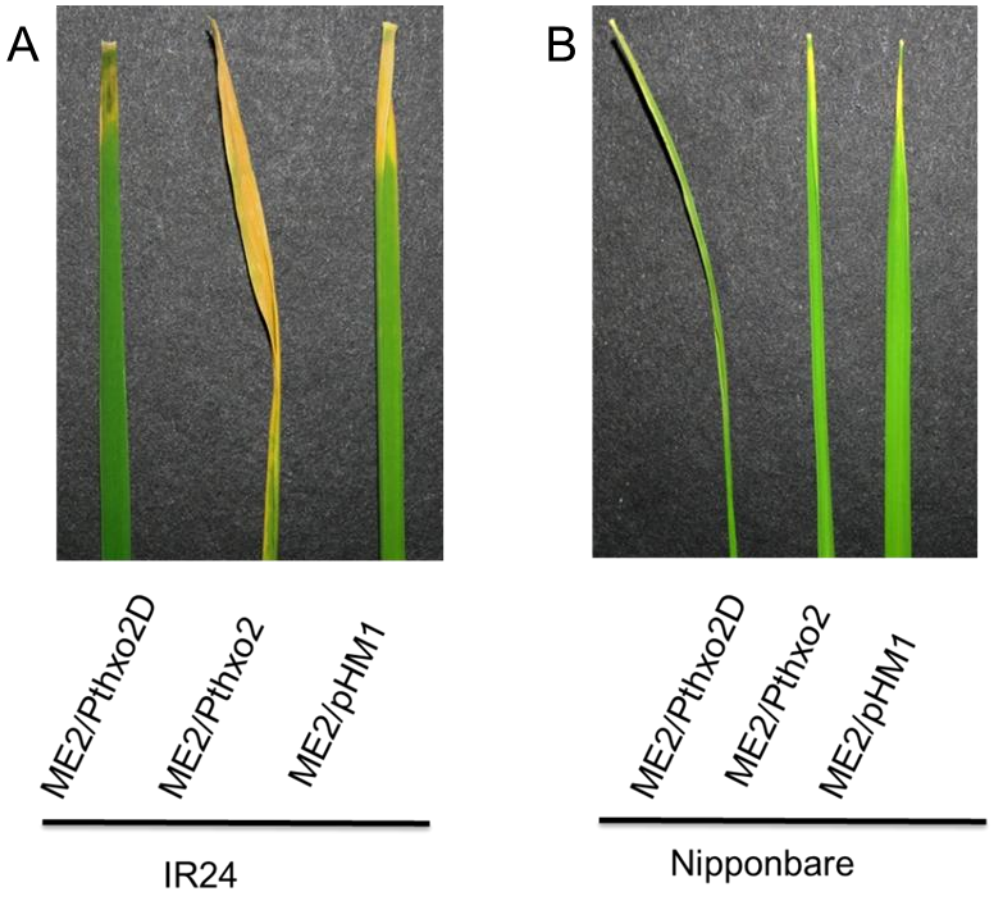


Figure 2-3. Alignments between RVDs of PthXo2 or PthXo2D effector and the EBE in the promoter of *OsSWEET13/Os12N3* in IR24 and Nipponbare.

The nucleotides underlined indicate the single nucleotide differences of promoter element of *Os12N3*/*OsSWEET13* between IR24 and Nipponbare. RVDs with predicted low affinity with the corresponding nucleotide are represented in red. In the PthXo2D, the fifth repeat with the RVD of NN is deleted compared to PthXo2. Nip=Nipponbare.


A
B
ME2/Pthxo2D
ME2/Pthxo2
ME2/pHM1
IR24
ME2/Pthxo2D
ME2/Pthxo2
ME2/pHM1
Nipponbare

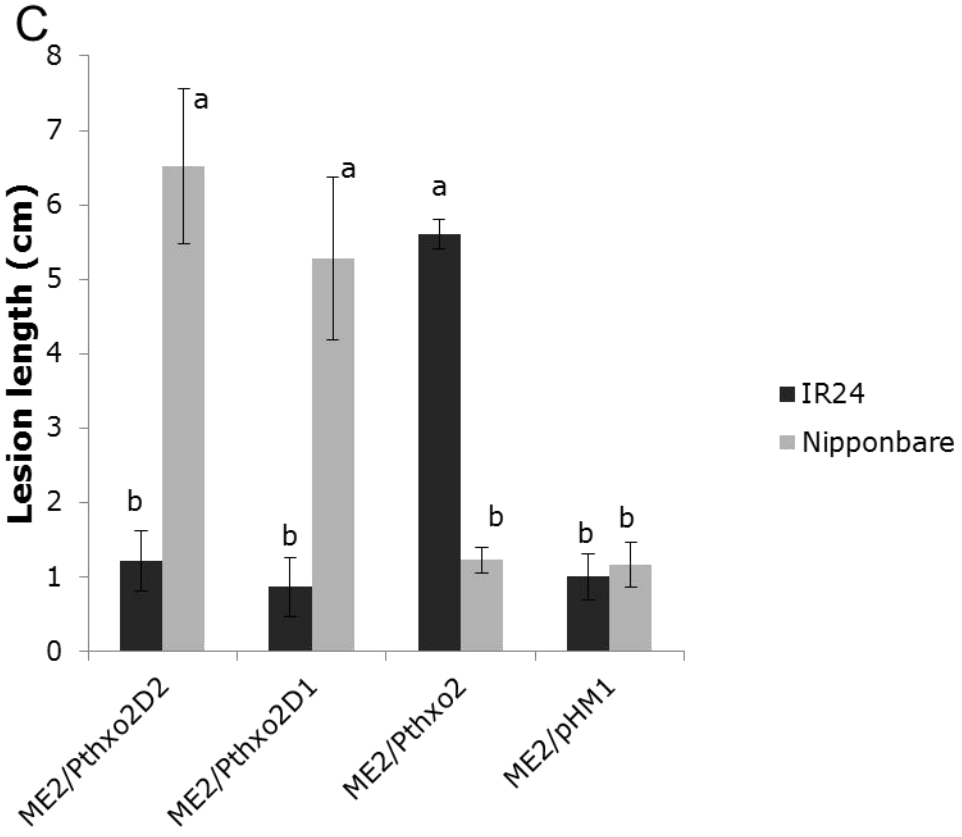

C
Lesion length (cm)
0
1
2
3
4
5
6
7
8
a
a
a
b
b
b
b
b
IR24
Nipponbare
ME2/Pthxo2D2
ME2/Pthxo2D1
ME2/Pthxo2
ME2/pHM1

Figure 2-4. **Repeat deletion mutation of *pthXo*2 leads to switch of resistance and disease susceptibility between IR24 and Nipponbare.** A and B, 20-day-old IR24 and Nipponbare plants were inoculated by clipping with ME2/*pthxo2D*, ME2/*pthxo2* and ME2/pHM1, respectively. The concentration of all bacteria suspensions were adjusted to OD=0.5 under 600nm. The pictures were taken 12 days post inoculation. C, the lesion length represents the average from seven rice plants 9 days post inoculation. Two different transformants ME2/*Pthxo2D1* and ME2/*Pthxo2D2* were tested in the experiment. Values with the same lowercase letter are not significantly different at the P-value<0.05 with Turkey statistics and ANOVA analysis.

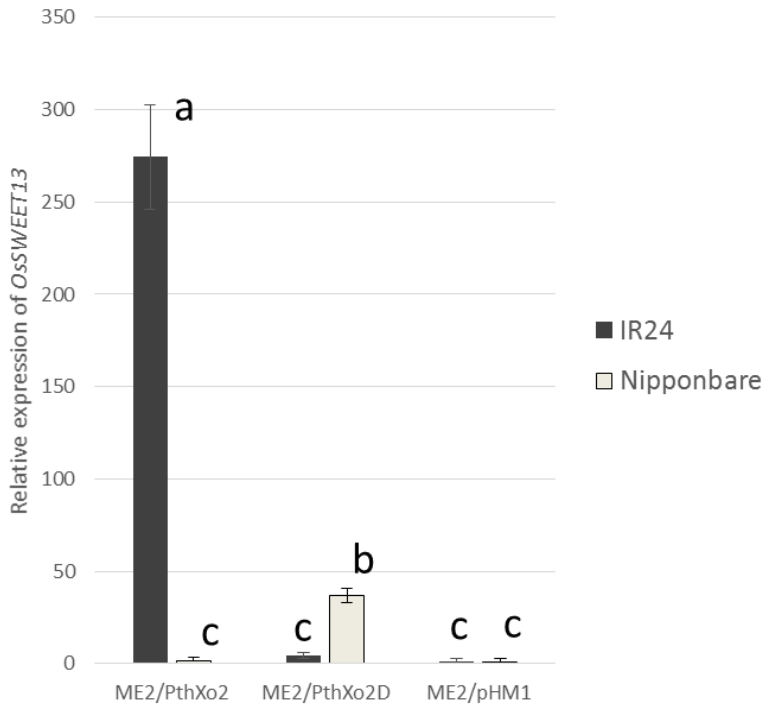
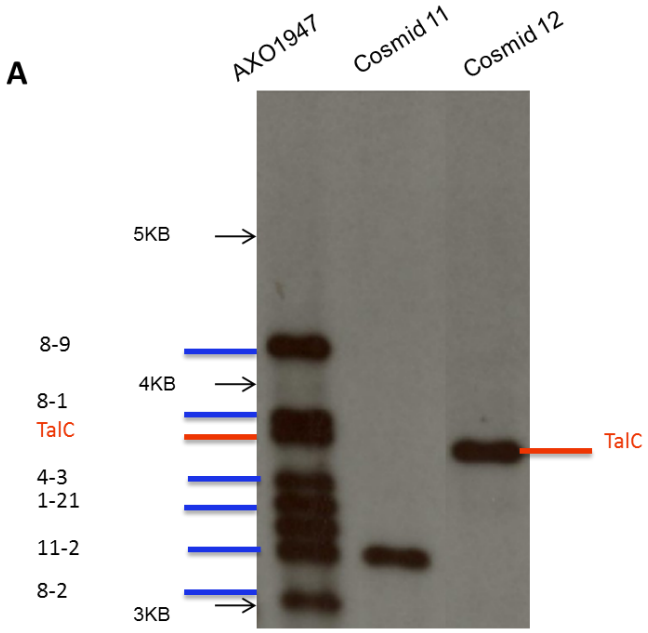


Figure 2-5. **A single repeat deletion of *pthXo2* switches *OsSWEET13* gene allele expression in Nipponbare and IR24.** qRT-PCR of RNA from inoculated rice leaves harvested 24 hours post inoculation. All the relative *OsSWEET13* gene expression level calculated by the $2^{\Delta\Delta Ct}$ methods and in comparison to ME2/pHM1 inoculation. The values with different lowercase letter indicate significant difference at the P-value<0.05 with ANOVA statistic tests.



B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
8-9	NN	NG	NN	HD	HD	NI	N*	NG	HD	NI	NG	NN	HD	NI	NG	NI	NG	NN	NG	HD	NI	NI	NG	HD	NN	NG
8-1	NN	ND	NN	NI	NK	NN	HD	NN	NG	NG	N*	HD	N*	HD	NI	NN	HD	NG	HD	HD	HD	NG	NN	HD	HD	NG
TalC	NS	NG	NS	HD	NI	NG	NN	NG	HD	NI	NN	N*	NI	NN	HD	NG	NI	NN	N*	HD	NN	NG				
2-1	NI	HD	NN	NS	NN	NG	HD	NG	HD	NG	NN	NG	HD	NS	HD	NI	NG	HD	HD	NN	HD	NN				
4-3	NN	HD	HD	NN	NN	NG	NG	HD	NG	HD	HD	NG	HD	HD	HD	HD	NG	HD	NG							
1-21	NN	NN	NN	HD	NI	NN	HD	HD	HD	NI	NN	NN	HD	HD	N*	NG	HD	NI								
11-2	NN	HD	NV	HD	NI	NG	NI	NN	NS	HD	HD	NI	NG	NI	NG	NI										
8-2	NN	HD	NI	NN	HD	NG	HD	HD	NG	NG	NI	NG	NI	NG												

Figure 2-6. Characterization of the TAL effector genes of AXO1947. A, Southern blot analysis of genomic and cosmid DNA, digested by *Bam*HI and separated in 1% agarose gel. *Sph*I fragment of *avrXa7* was used as the probe. The bands corresponding to the TAL genes cloned are marked with blue bar in the left while the *talC* band is marked as red bar; B, Eight TAL genes were cloned and sequenced. The RVDs of TalC effector is marked as red.

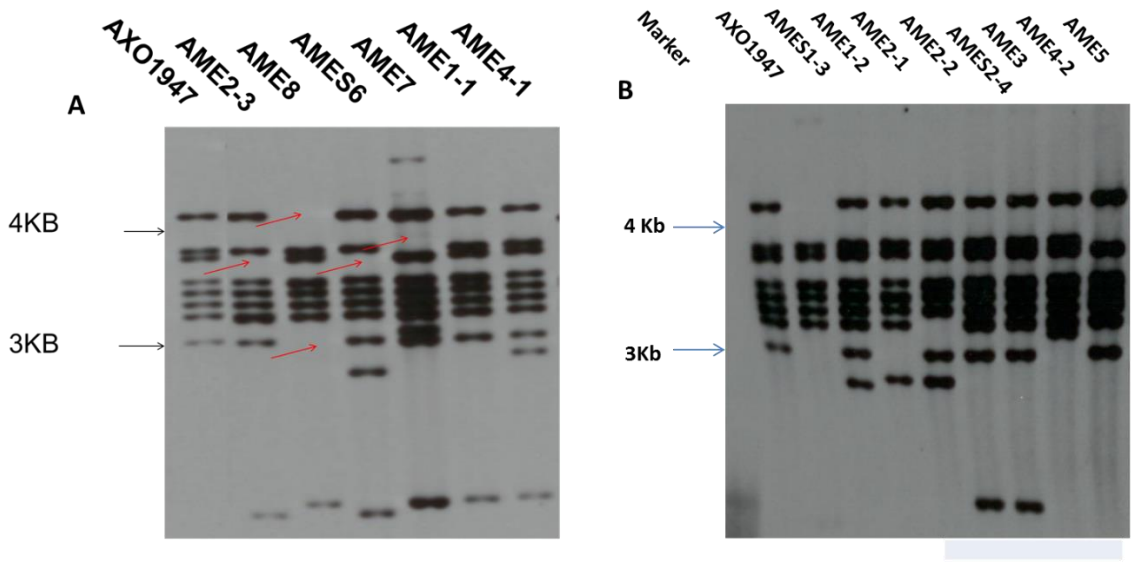


Figure 2-7. Southern hybridization analysis of 14 TAL mutants of AXO1947. Genomic DNA was digested by *Bam*HI and separated in 1% agarose gel. The central *Sph*I fragment of TAL gene *avrXa7* was used as probe. A, blot analysis of wild type strain and six TAL mutants. The red arrows indicate the loss of TAL effector genes compared to WT. B, blot analysis of WT strain and the other eight TAL mutants.

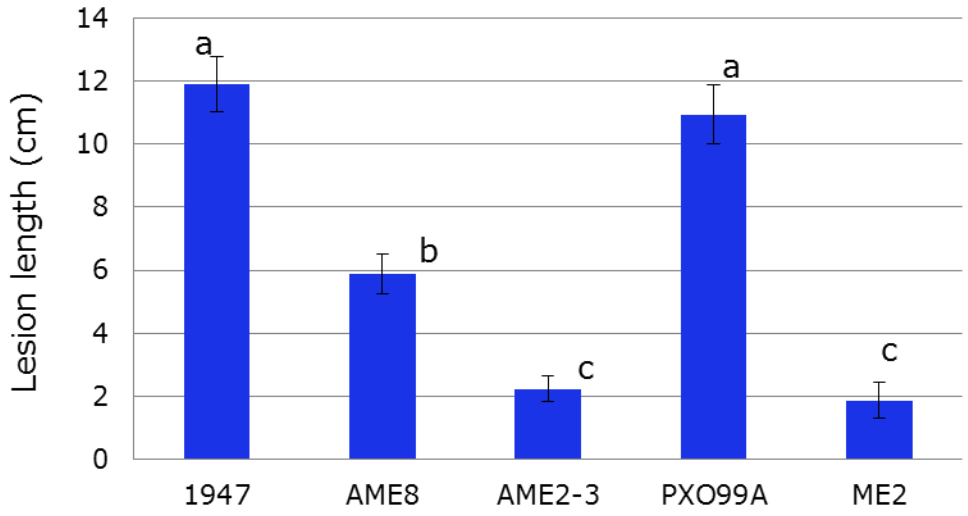


Figure 2-8. Virulence assays of TAL mutants of AXO1947 and PXO99^{A}. 20-day-old Nipponbare plants were inoculated with bacterial suspensions (OD=0.5 under 600 nm) of indicated strains by leaf tip-clipping method. AME8 and AME2-3 are TAL mutants of AXO1947 while ME2 is *pthXo1* mutant of PXO99^{A}. The lesion length was measured 7 days post inoculation. At least 8 leaves were used for each treatment. Values with the same lowercase letter are not significantly different at the P-value<0.05 with Turkey statistics and ANOVA analysis.

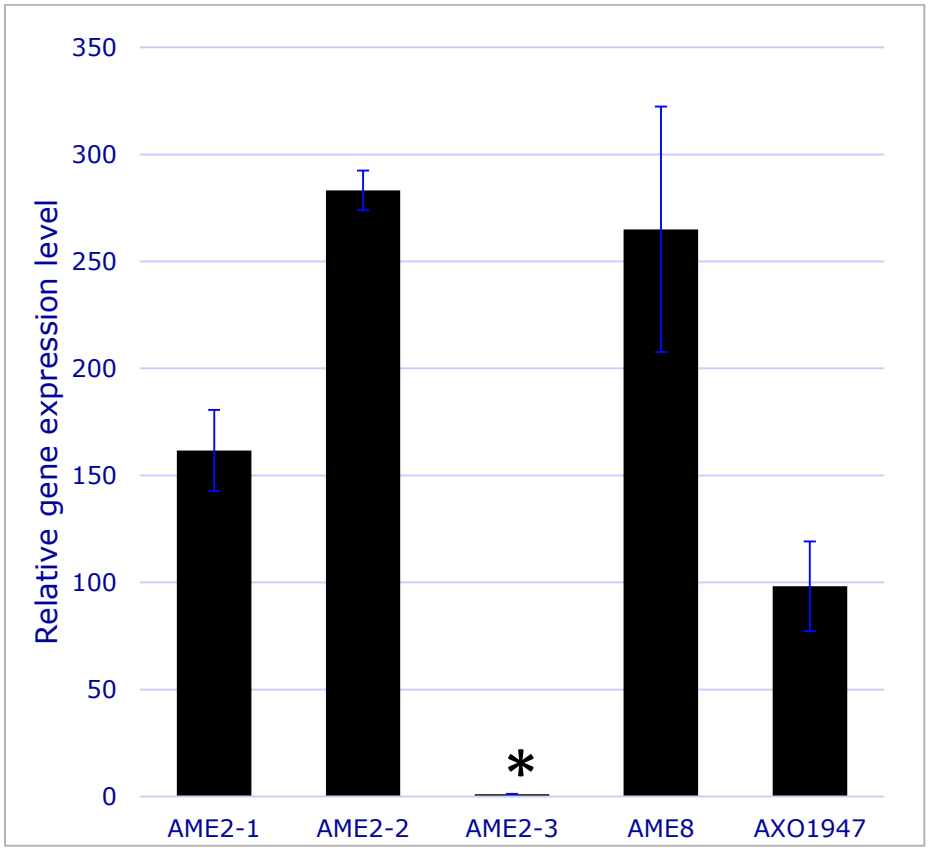


Figure 2-9. *talC* mutant AME2-3 fails to induce *OsSWEET14/Os11N3* gene expression. Leaves of 14-day-old rice plants were inoculated with bacterial suspensions (OD=0.5 under 600 nm) of indicated strains by needleless syringe inoculation method. The RNA extraction was performed 24 hours post inoculation. All the relative *OsSWEET14* gene expression level are calculated by the $2^{\Delta\Delta Ct}$ methods with the comparison of AME2-3 inoculation. AME2-1, 2-2, 2-3 and AME8 are TAL mutants of AXO1947. * indicates significant difference of *OsSWEET14* expression level by AME2-3 compared to other strains.

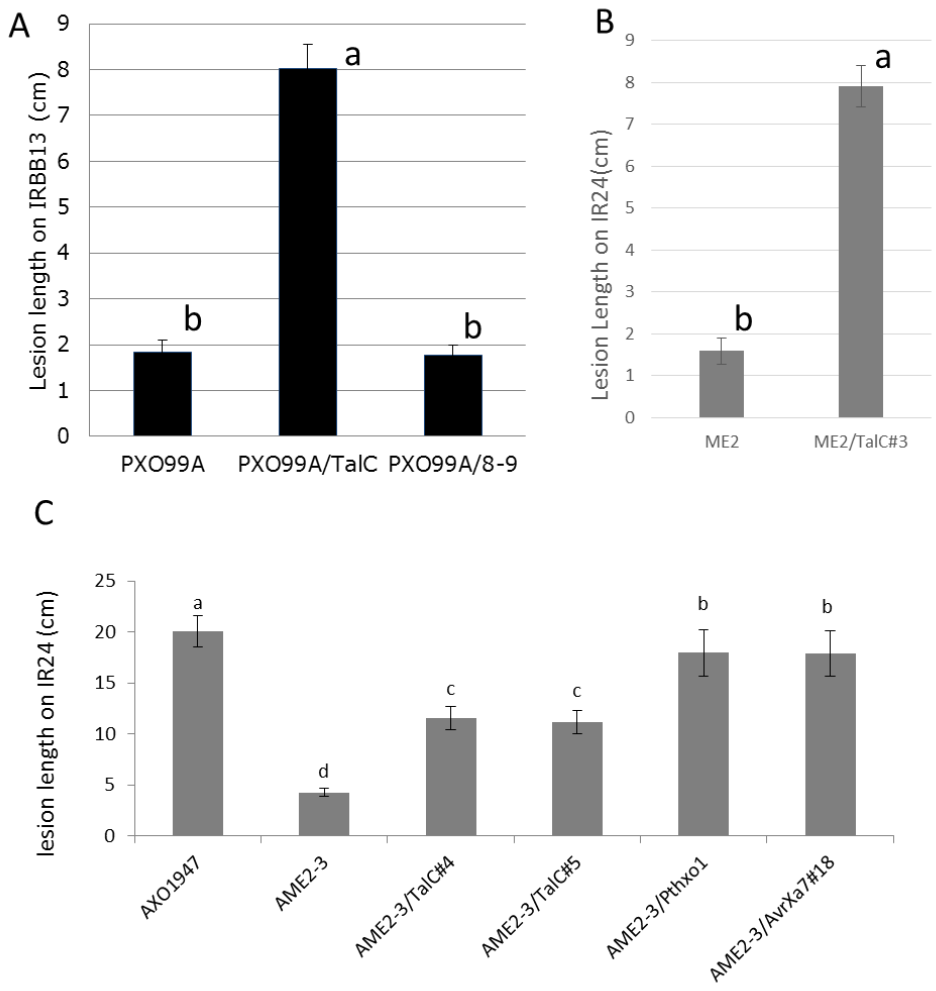


Figure 2-10. **Effects of heterologous TAL effector gene transfer on virulence in rice.** A, *talC* enables PXO99^{A} to defeat *xa13* resistance. PXO99^{A} with *tal8-9* TAL effector gene from AXO1947 behave similarly as PXO99^{A}. B, TalC effector enables ME2 to be more virulent on IR24. C, *talC*, *pthXo1* and *avrXa7* complement the low virulence of AME2-3. TalC is the major virulence TAL effector from AXO1947. PthXo1 is one major virulence effector from PXO99^{A}, which targets *Os8N3*/*OsSWEET11* gene; AvrXa7 is the major virulence TAL effector from PXO86 and targets *Os11N3/SWEET14*. 20-day-old rice plants were inoculated with indicated bacterial suspensions of OD600=0.5.The different letters indicate the significant difference under P-value<0.05 in the ANOVA statistics analysis.

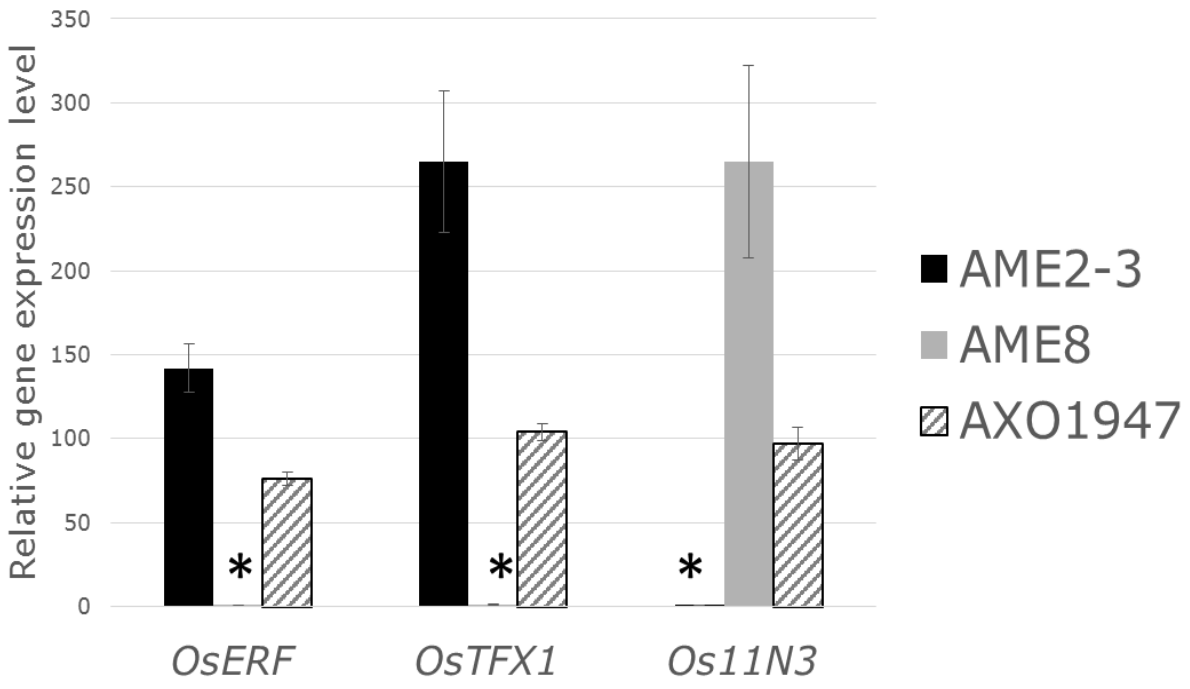


Figure 2-11. Effect of TAL effector gene mutation on *OsERF*, *OsTFX1* and *Os11N3* gene expression. Leaves of 14-day-old rice plant were inoculated with bacterial suspensions (OD=0.5 under 600 nm) by needleless syringe infiltration. The RNA extraction was performed 24 hours post inoculation. AME2-3 and AME8 are TAL mutants of AXO1947.The relative gene expression level of *Os11N3* is calculated by the $2^{\Delta\Delta Ct}$ methods with the comparison of AME2-3 inoculation and the * indicates the significant difference in expression level of *Os11N3* by AME2-3 compared to AME8 and WT with P-value<0.05 in the ANOVA statistics analysis; The gene expression of *OsERF* and *OsTFX1* is calculated in comparison with AME8 inoculation and the * suggests the significant difference in expression level of *OsERF* and *OsTFX1* by AME8 relative to WT and AME2-3 in similar statistics analysis.

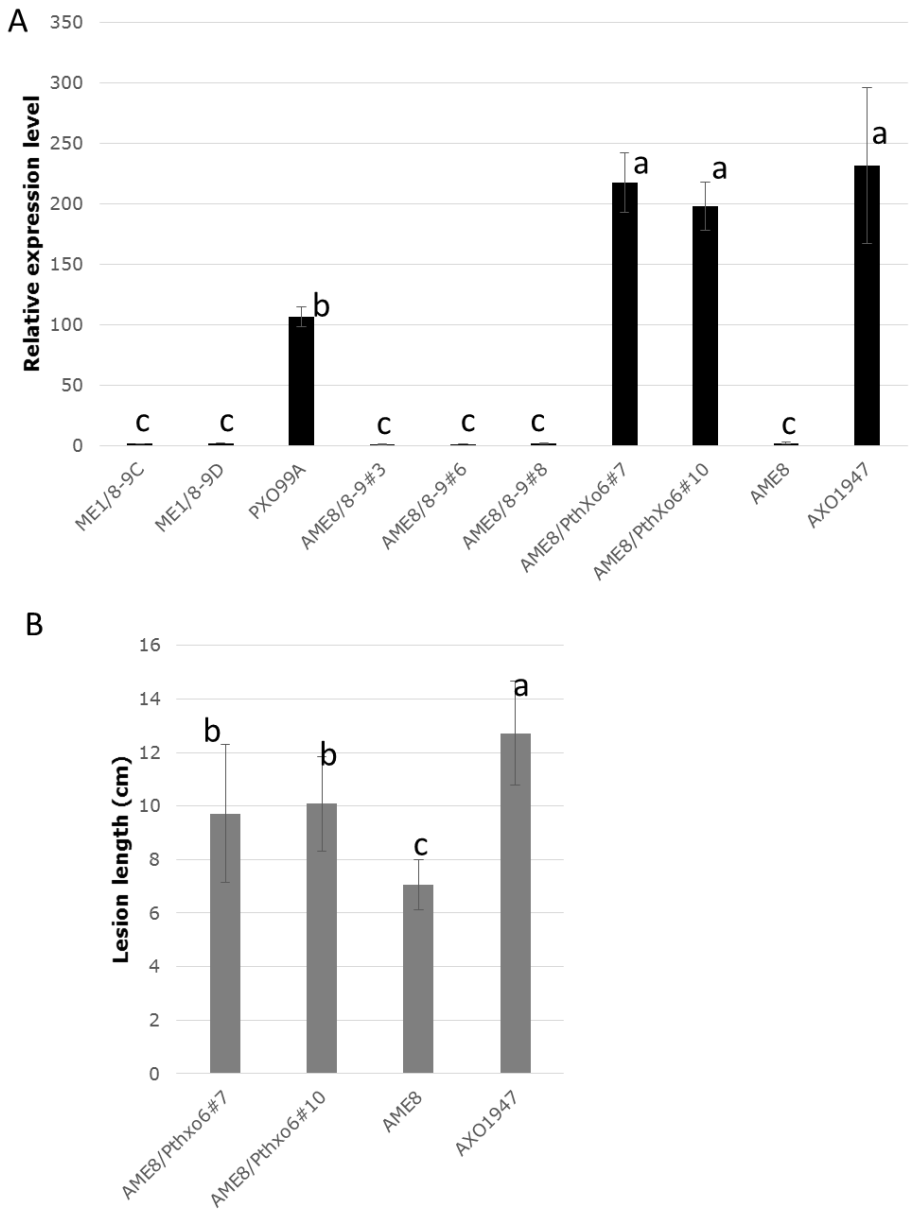


Figure 2-12. ***pthXo6*** **restores the** ***OsTFX1*** **gene expression of AME8 and partially complements loss of virulence in AME8.** A, The relative gene expression level of *OsTFX1* is calculated by the $2^{\Delta\Delta Ct}$ methods with the comparison of AME8 inoculation. *OsTFIIA**γ**5* gene was used as internal control. B, leaves of three week old rice plants were used in the lesion length measurement assay. The lesion length was measured 10 days post inoculation. Two individual complementation clones were tested for the virulence on Nipponbare plants. Values with the same lowercase letter are not significantly different at the P-value<0.05 with Turkey statistics and ANOVA analysis.

Chapter Three

Genomic Analysis of *Xanthomonas translucens*

Abstract

Xanthomonas translucens is a group of bacterial strains that cause diseases and yield loss on many important cereal crops. Three pathovars, with the designation of *X. translucens* pv. *undulosa* (Xtu), *X. translucens* pv. *translucens* and *X. translucens* pv.*cerealis*, are the major causal agents. However, no complete genome sequence for a strain of this species is currently available. A complete genome sequence of Xtu strain XT4699 was obtained by assembling PacBio long-read sequences. Whole genome structure and content comparison between XT4699 and ICMP11055, another Xtu strain with a recent full genome assembly, reveals occurrence of two major rearrangements and missing of genomic regions with potential virulence factors in XT4699. Full genome sequences indicate the presence of 8 and 7 Transcription-Activator Like (TAL) effector genes in XT4699 and ICMP11055, respectively. In addition, draft genome sequences of 19 other *X. translucens* strains collected from different areas and at different times were generated by Illumina sequencing. Phylogenetic relationship analysis among different *Xanthomonas* strains reveals *X. translucens* is distant from *X. oryzae* and *X. campestris* and also indicates that the three pathovars of this species represent distinct clades. Comparative analysis indicates three major factors, the absence and presence, copy numbers and frameshift mutations, influence the non-TAL type III effector repertories among different pathovars or within one pathovar. Knockout mutation of type III secretion system of XT4699 eliminates the ability to cause water-soaking symptoms and results in a significant reduction in growth on wheat compared to wild type. Furthermore, microarray and qRT-PCR analyses reveal that the strain specific TAL effector XT4699-Tal6 contributes in wheat gene expression modulation.

Introduction

Bacterial pathogens of the genus *Xanthomonas* cause disease symptoms in a wide range of plant species, including many economical important cereal crops (White et al., 2009). The species *X. translucens* are a group of strains that are pathogenic on various *poaceae*, including wheat, barley, oat, rye and other grass species. Bacterial leaf streak (BLS) and black chaff symptoms in the grain spikes on wheat are caused by *X. tranlucens* pv. *undulosa* strains. Outbreaks of BLS occur sporadically in central Great Plains and are often associated with relatively warm and humid conditions, although the disease has been prevalent in recurrent years in the northern Great Plains (Adhikari et al., 2012).

The *X. translucens* strains are classified by pathogenicity types and appropriate DNA fingerprinting technologies (Bragard et al., 1997). Strains causing disease symptoms on barley and wheat are named as *X. translucens* pv. *undulosa*, while strains only pathogenic on barley are called *X. translucens* pv. *translucens* (Bragard et al., 1995). However, phylogenic analyses of various *X. translucens*, particularly, the small grains infecting strains does not align with the pathovar designations, and clarifications await genomic analyses on larger strain collections. In addition, many strains, whether isolated from wheat or not, often have been reported to cause disease symptoms on wheat (Bragard et al., 1997).

Currently, a few draft genome sequences have been assembled for strains of the *X. translucens* group, to which Illumina or Roche 454 shotgun platforms were used. Next-generation sequencing technologies have made transformational change over the Sanger sequencing by improving throughput and reducing cost (Bentley, 2006). The draft genome sequences provide valuable information on major genome contents and enable genome

comparison among interested strains (Gardiner et al., 2014, Wichmann et al., 2013). However, the genome assemblies based on Illumina and Roche 454 sequencing, due to relatively short reads and sequencing biases, are fragmented and most assemblies failed to assemble complex repetitive sequence, including multiple transcription activator-like (TAL) effector genes. TAL effector genes typically have highly conserved N- and C-terminal sequences, and harbor 12.5-28.5 units of 102 or 105 bp repeats in the central regions (White et al., 2009). Recently, a single-molecule real-time (SMRT) sequencing technology was developed by Pacific Bioscience (PacBio), which typically produces long reads with no obvious sequencing biases. Due to a high error rate of PacBio reads, a high sequencing depth (*e.g.*, 50x or higher) is usually required for a high-quality *de novo* assembly (Bashir et al., 2012, Chin et al., 2013, Koren et al., 2012).

In this study, the complete genome sequence of *X. translucens* strains, XT4699, isolated from wheat in Kansas in 1999, was generated by using high-depth PacBio data. This study showed that a genome with multiple copies of TAL effector genes can be accurately assembled with long PacBio reads. The genome was compared to draft genomic sequences of 19 other different *X. translucens* strains assembled with Illumina data. Genome comparison on structure and contents between XT4699 and ICMP11055, another *X. translucens* strain with complete genome assembly (Charkhabi et al., unpublished) was performed. Phylogenetic relationship and comparison on Type III effector repertories and CRISPR region of all *translucens* strains were examined. Comparisons of TAL effectors from four *X. translucens* strains and their relationship with TAL effectors from other *Xanthomonas* species were also performed. Candidate wheat host target genes by TAL effectors of XT4699 were identified *via* microarray analysis and qRT-PCR.

Materials and Methods

Bacterial strains and genomic DNA extraction

The *X. translucens* strains used in this study are listed in Table 3-1. *Xanthomonas* strains were grown on tryptone sucrose agar (tryptone, 10 g/liter; sucrose, 10 g/liter; glutamic acid, 1 g/liter; Difco Bacto agar, 15 g/liter) medium at 28 ^{o}C (Yang & White, 2004). Bacterial strains were stored in nutrient broth with 20% glycerol at -80^{o}C. For genomic DNA extraction, fresh grown bacteria from medium was washed first by sterilized water and be treated with 1% SDS and proteinase K (1mg/ml) at 37^{o}C for 20-30 minutes. Samples were then treated with 0.5M NaCl and incubated at 65^{o}C for 30 minutes for cell lysis. Phenol and Chloroform extraction and 2.5 times 95% ethanol precipitation were applied to isolate DNA. Samples were treated with RNAse A (0.1 mg/ml) and incubated at 37^{o}C for 30 minutes for RNA removal.

Genome sequencing and assembly

In this study, Illumina 2x250 bp MiSeq data for all sequenced strains were generated at the Genomic Facility at Kansas State University. Reads were subjected to adaptor and quality trimming and assembled *via* CLC Genomics Workbench software. For the whole genome assembly for XT4699 strain, two SMRT cells of PacBio RS II sequencing data were generated at the Interdisciplinary Center for Biotechnology Research (ICBR) at University of Florida. These reads were assembled into three contigs using an optimized PacBio pipeline, HGAP2 (Chin et al., 2013). As suggested by a PacBio HGAP instruction for bacterial assemblies, the parameter of "Target Coverage" was changed from 30 to 15. One short contig was discarded due to low coverage of PacBio reads. Two remaining contigs were merged into a single contig with Illumina assembled contigs using the minimus2 module in AMOS (sourceforge.net/apps/mediawiki/amos

). The contig was further circularized through removing the overlap at two ends. To improve the quality of the draft assembly at the ends of the contig, a standard PacBio resequencing pipeline with PacBio reads was used for additional error correction. Consequently, a circular assembly genome sequence was obtained.

Alignment of XT4699 Illumina reads to the assembly genome to assess the assembly quality

Trimmed Illumina reads were aligned to the assembly sequences with BWA-MEM (Li & Durbin, 2010). The alignment was subjected to stringent filtering criteria (minimum mapping score: 40; minimum overlap: 100 bp; minimum identity: 97%; and minimum read coverage: 98%) to obtain a set of uniquely and confidently mapped reads coverage. To identify mismatches between Illumina sequencing data and the assembly genome, GATK was applied to recalibrate base quality scores, perform INDEL local realignment, and discover SNPs and INDELs (DePristo et al., 2011, McKenna et al., 2010). A set of criteria for polymorphism filtering, including a minimum polymorphic site coverage set as five reads and minimum percentage of reads of the polymorphic allele set at 90%, were used as the filter to identify the mismatches between Illumina sequences and the assembly genome.

Genome annotation

The finished assembled genome of XT4699 and ICMP11055 were annotated using the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) (ncbi.nlm.nih.gov/genome/annotation_prok). The complete and draft genomic sequences were also annotated using the RAST (Rapid Annotation using Subsystem Technology)

(http://rast.nmpdr.org) (Overbeek et al., 2014). The sequence feature of TTCGN16TTCGN, where N represents A, T, G, and C, was used to search the 500-bp promoter region before each annotated gene to identify genes with the plant inducible promoter (PIP) element. IS elements of XT4699 and ICMP11055 genomes were annotated by ISfinder (www-is.biotoul.fr) (Siguier et al., 2006)

Phylogenetic tree construction

Multilocus sequence analysis (MLSA) was used to build phylogenetic trees in *Xanthomonas* strains (Young et al., 2008). To generate the phylogenetic relationship tree, concatenated sequences of four conserved loci (*dnaK*, *gyrB*, *groEL* and *recA*) were assembled. The tree was constructed with Tamura-Nei genetic distance model and Neighbor-joining method, with *Stenotrophomonas maltophilia* K279a as outgroup.

Whole-genome discovery of single nucleotide polymorphisms (SNPs) was performed by using the assembly draft or finished assembly sequences with MUMmer 3.0 (Kurtz et al., 2004). The modules of nucmer, delta-filter, and show-snps were sequentially run to identify SNPs. At least 100 bp and 90% identity of one-to-one match was required for the alignment between each assembled sequence and the XT4699 reference genome. After identifying all the SNPs of each strain, the XT4699 reference genome was modified at all the SNP sites to polymorphic bases. The resulting modified XT4699 sequence was used as the reference for the second run of the SNP discovery. Consequently, a matrix of the genotyping result of all the strains was obtained. The genotyping data at the SNP sites with no missing data were used to construct a phylogenetic tree using an R package APE (Paradis et al., 2004).

Identification of type III effector in assembly genomes

A combined resource of Type III effector (T3E) (Gardiner et al., 2014, Potnis et al., 2011, White et al., 2009, Wichmann et al., 2013) in the genus of *Xanthomonas*, also called Xop effectors, were used as query to do blastp against the NCBI and RAST annotation protein database of XT4699. The cutoff e-value was set as $1e^{-10}$. For other *X. translucens* strains, the same query was applied using blastp against the RAST annotation database of each strain. In the second round of identification of T3Es, the same query was applied against genome sequence of each strains using tblastn to avoid the missing of T3Es, which may be overlooked by annotation process. This round of tblastn was also used to confirm the presence, absence and number of copies and frameshift mutation of T3Es in the genome. Subject proteins in contigs with less than 5x sequencing coverage were not considered as true. Frameshift or premature termination codon in T3Es was considered when encountering pieces of subject proteins within one contig that aligned well with same query protein but had bad overlapping ends and the average coverage of contig was above 100. The missing proteins overlooked by the annotation but identified by tblastn were re-evaluated by doing blastp in the NCBI protein database.

Southern blot analysis

Genomic DNA manipulation was performed according to standard protocols (Ausubel et al., 1994). Genomic DNA of *X. translucens* strains was completely digested by *Bam*HI (New England Biolabs, MA). The digested DNA was separated in 0.9% agarose gel *via* electrophoresis at 4 ^{o}C at 40 Volt overnight. The *sphI* fragment of TAL effector gene in Strain AXO1947 was used as probe. AlkPhos Direct Labeling and Detection System with CDP-Star Kit (GE healthcare) was applied for the probe labeling, hybridization and detection procedures.

Mutagenesis and mutant validation in XT4699

A *hrcC*$^-$ mutant of XT4699 was generated by gene transfer and homologous recombination of a mutant copy. The partial fragment of *hrcC* gene were amplified by PCR with primers XTThrcC-F and XTThrcC-R (Table 3-2) and cloned into suicide vector pKNOCK-Km vector (Alexeyev, 1999). The suicide vector with cloned fragment was transformed into *E. Coli* S17-1 pir strain for bacterial conjugation. The mixture of XT4699 and S17-1 strains was plated on NA at 28 ^{o}C for 24hrs, then transferred to TSA plates with 20 ug/ml Cephalexin and 50 ug/ml Kamaycin for selection of XT4699*hrcC*$^-$ mutants (Hopkins et al., 1992). The mutants were validated by PCR with primers XTThrcC-Out and 07KM-Val. The TAL mutants of XT4699 were generated in similar way. Partial fragment of N-terminal region of each TAL effector gene in XT4699 was individually amplified by PCR with primers 4699N-TAL-F and 4699N-TAL-R. The partial fragment was cloned into pKNOCK-Km vector for conjugation. Mutants were validated by PCR with specific primers located upstream of each TAL gene and the reverse primer Forall-Val in the vector (Table 3-2).

Pathogenicity assays

Two days old *Xanthomonas* strains were scraped off the TSA medium plates and resuspended in sterilized water. For pathogenicity type assays, bacterial suspensions (OD600=0.2) were infiltrated onto one-month-old wheat and barley plants by needleless syringe. Hexaploid wheat cultivars ‘Chinese Spring’, ‘Jagger’, ‘Hope’ and ‘Canthatch’, one *Triticum turgidum* wheat (accession number 107 in WGRC at KSU), ‘KS Southeast’ and ‘Morex’ barleys were applied. The symptoms were observed and photographed at 4DPI. For dip inoculation assays, second leaves of 14-day-old wheat plants were dipped into the bacterial suspensions (OD600=0.2) coated with 0.02% Silwet L-77. The water soaking symptoms appeared at 4-5DPI

and were photographed at 8DPI. For bacterial growth population assays, inoculum of 4 x 10^4 dilutions of original bacterial suspension (OD600=0.2) was infiltrated on 2nd leaves of 3-week-old Chinese Spring wheat plants by needleless syringe. In each treatment, three biological replicates of inoculated leaves (3 cm) were pooled and ground together at 6DPI. Ground samples were serially diluted and 100 µl of diluted samples were added onto TSA agar plates for colony counting three days later. 10-100 colonies per plate is considered as optimal and three plates were applied for calculating colony forming units (CFU).

Microarray and qRT-PCR

First leaves of 10-day-old wheat plants were inoculated with bacterial suspensions (OD600=0.5) for 24 hours. RNA samples were isolated from three independently inoculated leaves using Trizol reagents (Invitrogen). Three biological replicates were applied for the microarray experiments, which were performed at Genomic Facility at Kansas State University. For qRT-PCR, 1 µg RNA samples were treated with DNAseI following the protocols provided by Invitrogen and subjected to reverse transcriptase reaction using kits provided by Biorad. 10x diluted samples of the cDNA products were applied for real-time PCR in the CFX-96 machines at Genomic Facility at Kansas State University. Relative gene expression level was calculated with $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001). The EF-1α gene was set as an internal control gene. Primers were shown in the Table 3-2.

Results

Complete genome assembly of XT4699

Both Illumina and Pacbio reads were generated for the XT4699 strain. Illumina 2x250 bp MiSeq data provided about 60x coverage, comprising 267.5 Mbp total assembled bases. The

assembly via SOAPdenov2 (Li et al., 2010) resulted in 547 contigs of at least 400 bp in length. The total assembly length was 4.4 Mb and the N50 equals to 15,114 bp. Two SMRT cells of PacBio data for the same strain were generated as well. In total, 114,394 PacBio raw reads consisted of 446 Mb of bases were generated. The average length of reads was 3,898 bp. The reads were assembled into two long contigs using an optimized assembly pipeline, HGAP2 (Chin et al., 2013). The longest contig is 4,357,621 bp in length. Two contigs were merged into a single contig *via* combining with Illumina assembled contigs using the minimus2 module in AMOS (sourceforge.net/apps/mediawiki/amos). Two ends of the resulting contig shared 8,439 bp with 99.33% identity. The contig was circularized after removing the sharing sequence at one end. The resulting circularized assembly draft was subjected to additional two rounds of error correction using the PacBio resequencing pipeline that includes the error correction module (Chin et al., 2013). Consequently, a single finished genome sequence (N=4,561,137 bp) was obtained.

To assess the assembly quality, Illumina reads were mapped to the assembled contig. 99.53% reads can be mapped and 95.46% are uniquely mapped with the stringent mapping criteria (see Methods). Respectively, 99.99% and 99.44% of the assembled genome was covered by mapping reads and uniquely mapped reads. Based on the alignments, no mismatches were identified between reads and the assembly genome (see Methods), indicating the high quality of the assembly sequence. The Illumina-only assembly contigs were also aligned *via* nucmer (Kurtz et al., 2004) to the final assembly. Of the 547 contigs, 98.54% (539/547), accounting for 99.77% total contig sequence, can be perfectly or nearly-perfectly aligned (see Methods). Approximately, 97% of the final assembly was covered by the Illumina contigs. Only four single nucleotide variants, including two substitutions and two INDELs (insertion-and-deletions), were identified

between the Illumina assembly contigs and the final assembly (Table 3-3). To assess the quality of the assembly on the repetitive regions, especially the regions of great interest, TAL effector genes, six of them have been fully cloned by high fidelity PCR and Sanger sequenced. All the Sanger sequences were perfectly matched to the assembled genome sequence with 100% identity and coverage (Table 3-4). Collectively, this assembly is a finished high-quality assembly by mainly using PacBio, with a final step of a gap closing with Illumina-only assembled contigs (Figure 3-1).

XT4699 genome content and the influence of repetitive sequences and GC contents on the assembly

The XT4699 genome is a single circular chromosome (4,561,137 bp) with an overall G+C content of 68.1%. The complete genome was annotated using the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP). 3,528 CDS genes, 94 pseudogenes, and 54 genes with frameshifts were annotated. XT4699 has 2 ribosomal RNA operons and 54 tRNA genes. It was predicted to have 74 complete Insertion Sequence (IS) elements and 56 partial ones via ISfinder (Siguier et al., 2006). Annotation *via* RAST shows that it encodes three long nonribosomal peptide synthesis (NRPS) proteins of 4827, 7451, and 6419 amino acids, in a cluster of 80 kb regions flanked by IS elements (Figure 3-1). Examination of promoters of each predicted CDS reveals plant inducible promoter (PIP) element is present in 49 genes, including some Hrp genes and type III effector genes (Figure 3-1).

The XT4699 genome consists of 4.3%, or approximately 0.2 Mb, of repetitive sequence. Among 252 genomic repetitive regions, 163 repeats are less than 1 kb and 89 repeats are longer than 1 kb. The longest repeat is approximately 6 kb with the 99.98% identity between two copies. Eleven pairs of tandem repeats were identified, ranging in size from 107 bp to 537 bp.

In terms of copy number, 42.5% of the repeats have two copies and 57.5% repeats have three or more copies.

Repetitive sequences increase the difficulty of the assembly, especially when only using short reads (e.g., Illumina sequences) for the assembly. In the Illumina-only assembly, only 21.8% (55/252) repetitive regions were resolved. The vast majority (51/55) of these resolved repeats are of small size (<400 bp). However, not all small repeats are resolvable using Illumina data only. Indeed, 63.5% small repeats (<400 bp) were not resolved in our Illumina-only assembly. The result from the further inspection suggests that tandem repeats and high copy (>2) of repeats increase the complexity of the assembly. More than 90% of the tandem repeats were occurred at assembly gaps, even when the tandem repeats are small (115-375 bp). The results also suggest that high-copy repeats (>2 copies) greatly increase the assembly difficulty (Figure 3-2). In fact, the vast majority of high-copy repeats were not resolved. Except for one repeat, none of repetitive regions with greater than 500 bp were resolved in the Illumina-only assembly. Overall, about 40% gaps in the Illumina-only assembly are located at repetitive regions.

The overall GC percentage of the XT4699 final assembly is 68.1%, while the overall GC percentage of the Illumina-only assembly is 67.2%, indicating some GC-rich regions might not be represented in the Illumina-only assembly. To understand how the Illumina and PacBio sequences cover the GC-rich regions, the final assembly genome was scanned with non-overlapping 200 bp windows and the GC% and sequencing coverage of both Illumina and PacBio sequences were determined at each window. The dot plots showed an obvious GC bias in sequencing coverage of Illumina but not PacBio, which at least partially explains high variation of Illumina sequencing coverage but fairly uniform distribution of PacBio sequencing reads across the genome (Figure 3-3). As a result, the vast majority (50/55) of extremely high GC

regions that exhibit > 80% GC content overlap with assembly gaps in the Illumina-only assembly. Most gaps on GC rich regions are very small (<100 bp). The flanking contigs of many of these gaps shares tails but the tails are not long enough to join two adjacent contigs. On the contrary, the gaps caused by repeats are relatively longer. The sizes of repeat-induced gaps are actually linearly correlated with repeat sizes (Figure 3-2). Regardless of gaps in the Illumina-only assembly, the contig sequences exhibit high accuracy when compared with the final assembly, which was almost entirely attributed to PacBio reads (Figure 3-3).

Genome comparison between XT4699 and ICMP11055

Using PacBio long-read sequencing, a complete genome of ICMP11055 (4,761,583 bp) was also obtained (Charkhabi et al., unpublished). Whole genome structure comparison reveals there are at least two major inverted and rearranged genomic regions between XT4699 and ICMP11055 (Figure 3-4). Transposable elements were identified in the terminal ends of the rearranged regions, indicating the rearrangements might be induced by transposon elements. One TAL effector gene, XT4699-*tal3*, which is flanked by several transposon elements and located in the terminal end of one rearranged region, is absent in the genome of ICMP11055. Comparison of RVDs of TAL effector genes between two strains shows that five TAL genes are shared in both strains, in which three are identical and two are highly similar (Figure 3-4, Figure 3-5). Interestingly, the XT4699-Tal6 and ICMP11055-Tal4a have the same RVDs but the arrangement of 34- and 35-amino-acid repeats is different (Figure 3-5).

In genome size, ICMP11055 strain is 200 kb larger than XT4699. The genome content comparison between two strains is listed in Table 3-5. Comparative genomic analysis between two strains reveals that there are nine extra genomic regions (>10 kb) found in ICMP11055

while only one genomic region (>10 kb) is uniquely identified in XT4699, which is a gene cluster predicted to encode the type III restriction and modification system (Region J). Prophage elements (Region D, F, G, H), CRISPR gene cluster (Region C), Vgr (valine-glycine repeats)-related genes of Type VI secretion system (Region A, F, I), genes encoding cyclolysin secretion ATP-binding protein, hemolysin secretion protein D and alkaline phosphatase (Region E) and gene cluster for Type II/IV secretion system (Region B) are the main components for the nine extra genomic regions of ICMP11055 over XT4699 (Figure 3-4). Prophage, CRISPR, gene cluster for Type II/IV secretion system and genes in Region E are unique for ICMP11055. Although Vgr-related genes are not unique for ICMP11055, there are three more genomic regions harboring Vgr-related genes than XT4699. The IS elements are often located in the ends of the extra genomic regions in ICMP11055, indicating transposon activity may be involved in loss and acquisition of these regions.

Phylogenetic relationship within *Xanthomonas* strains

Genome assembly using Illumina 2x250 bp reads for other *X. transluens* strains was performed, with contigs having 70-200X coverage. The N50, assembled genome size, maximum contig length, and contig number of each assembly are shown in Table 3-1. Consistent with the classification results from previous studies (Gardiner et al., 2014, Wichmann et al., 2013), the MLSA revealed the *X. translucens* strains were closer to *X. albilineans* in xanthomonad group 1 than *X. oryzae* and *X. campestris* (Figure 3-6A). The *X. translucens* pv. *undulosa* strains, XT4699, XT-Rocky, LB10, P3, ICMP11055, LG48, and DAR61454, which are virulent on both wheat and barley, were separated from the *X. translucens* pv. *transluencs* strains (XT8, B1, B2 and DSM18974) that are only virulent on barley. In addition, the above strains were

distinguishable from the *X. translucens* pv. *cerealis* CFBP2541, XT123, and *X. translucens* pv. *graminis* ART-Xtg29, which is consistent with previous phylogenetic classifications based on AFLP analysis (Bragard et al., 1997). The phylogenetic tree developed by a whole genome SNP comparison further corroborated MLSA results and revealed higher resolution of differences among very closely related strains (Figure 3-6B, Table 3-6). LG54, XT5523, XT-Rocky, CS2, CS22, LB10, and DAR61454 strains are highly related strains, sharing almost identical housekeeping genes and only having 2-17 SNP difference in a total of 9836 SNPs (Table 3-6), although they were isolated from different geographic origins and times (Table 3-1). In addition, XT4699 and LG48 are another pair of highly close strains in both the MLSA and the whole genome SNP comparison. International or countrywide germplasm exchanges or trades may be related with this high similarity among geographically distinct strains.

Distinct pathogenicity types of *X. translucens* strains

Based on the pathogenicity tests on 'Alondra' wheat and 'Corona' barley, *X. translucens* strains had been grouped into type A, virulent on both wheat and barley, type B, only virulent on barley, and type C, non-virulent on either (Bragard et al., 1997). Though strains from *X. translucens* pv. *undulosa* and pv. *cerealis* are classified as pathogenicity type A, they are distant from each other in the phylogenetic relationship (Figure 3-6). Here, different wheat and barley cultivars other than 'Alondra' wheat and 'Corona' barley were tested to determine possible potential pathogenicity differences between strains from each subgroup. Three representative strains, XT4699 (pv. *undulosa*), XT8 (pv. *translucens*) and XT123 (pv. *cerealis*), from each subgroup were selected and inoculated on 'Chinese Spring', 'Jagger', 'Hope' and 'Canthatch' hexaploid wheat, one *Triticum turgidum* wheat (accession number 107 in WGRC at KSU), 'KS

Southeast' and 'Morex' barley. XT4699 induced water-soaking symptoms on 'Chinese Spring', 'Jagger', 'Hope', 'Canthatch' wheat and 'KS Southeast' barley, while inducing mixed symptoms of chlorosis and water soaking on 'Morex' barley. XT123 only triggered chlorosis symptoms at 4 DPI on 'Chinese Spring', 'Jagger', 'Hope' wheat and 'KS Southeast', and 'Morex' barley (Figure 3-7). Though XT123 could cause water-soaking on 'Canthatch' and *Triticum turgidum* wheat cultivars, the strain exhibited different lesion symptoms from XT4699. XT8 caused strong water soaking symptoms on 'KS Southeast' barley and less severe water soaking lesions on 'Morex' barley, while only inducing chlorosis on all the wheat cultivars (Figure 3-7).

Type III secretion system in *X. translucens*

Comparative genomics have previously revealed extensive divergence of the type III secretion system (T3SS) between group 1 and group 2 of *Xanthomonas* species. *X. albilineans* does not possess the Hrp T3SS that is usually present and crucial for pathogenicity in group 2 *Xanthomonas* species (Pieretti et al., 2009). The T3SS of *X. translucens* strain ART-Xtg29 is shown to be divergent from the sequenced genome of *Xanthomonas* species in group 2 (Wichmann et al., 2013). Disruption of the T3SS in ART-Xtg29 strain did not completely reduce the disease symptoms. Furthermore, the survival of T3SS mutants inside of plant tissue was not significantly affected compared to the wild type strain within 14 DPI (Wichmann et al., 2013). The genomic comparison in the T3SS cluster between XT4699 and ART-Xtg29 strains revealed they share the same synteny in the organization of Hrp genes. To test whether the disruption of T3SS in XT4699 affects virulence, an insertion mutant on the *hrcC* gene was generated. The type III secretion mutant *hrcC*$^-$ resulted in loss of virulence on 'Chinese Spring' wheat on the basis that water-soaking lesions were not visible by 10 DPI, while the WT strain XT4699 caused

extensive water-soaked lesion starting at 3 DPI. The *hrcC* gene fully complemented the virulence loss for the mutant (Figure 3-8A). Using a dip inoculation assay, no disease was present at 10 DPI for *hrcC*$^-$ mutant inoculation while the disease spots appeared at 4-5 DPI for the WT inoculation (Figure 3-8B). The ability to cause disease symptoms on barley and non-host tobacco plants was also lost in the *hrcC*$^-$ mutant. Two different barley cultivars, 'KS Southeast' and 'Morex', were used in the inoculation assays. XT4699 formed sharp water soaking lesions and a mixed symptom of chlorosis and water soaking at 3 DPI on 'KS Southeast' and 'Morex', respectively, while the *hrcC*$^-$ mutant failed to develop any symptoms on both cultivars (Figure 3-7B and 3-7C). In addition, the *hrcC*$^-$ failed to trigger HR and serious chlorosis on non-host KY14 tobacco plants, while the WT and complementation strains induced a HR (Figure 3-8C). Bacterial population assays in 'Chinese Spring' wheat indicated that the *hrcC*$^-$ strain population grew much slower, about 450 fold reduction, than the WT strain at 6DPI (Figure 3-9). Altogether, the type III secretion system of *X. translucens* was demonstrated to play a major role in developing pathogenicity in host plants and inducing symptoms in non-host tobacco plants.

Type III effector content

Thirty-nine putative T3Es were predicted, not including TAL effectors, from diverse *X. translucens* strains (Figure 3-10). Among them, 23 T3Es are conserved among the *X. translucens* pv. *undulosa*, *X. translucens* pv. *translucens*, and *X. translucens* pv. *cerealis*. The core set of T3Es, which are present in a wide range of *Xanthomonas* strains, includes AvrBs2, XopF, XopK, XopL, XopN, XopP, XopQ, XopR, XopX and XopZ, with the exception that the two *cerealis* strains have frameshifts in XopR. Multiple copies of *avrBs2*, *xopF*, *xopX*, *xopL* and *xopP* are found in all three *X. translucens* pathovars, while strains from other *Xanthomonas* species only harbor one.

Effector composition varies between different pathovars, even within one pathovar. Presence and absence, copy number difference, and frameshift in coding sequences of T3Es are the three types of variation that were observed (Figure 3-10). The gene *avrBs1* is present in only *undulosa* strain CS4 and both *cerealis* strains. The *xopJ1* effector gene is present in XT5523, XT-Rocky, XT5770, and DAR61454, but XT4699, ICMP11055, LW16, and XT5791. The gene *xopJ1* is present in all four *translucens* strains but not in either of *cerealis* strains. The *xopAH* and *xopAK* genes are present in all *undulosa* and *translucens* strains but not in either of *cerealis* strains. Five different *xopE* effector genes are present in the four pathovars. All *undulosa* strains have *xopE1* and *xopE5*, and *translucens* strains harbor *xopE2* and *xopE3*, while *cerealis* strains have *xopE1*, *xopE2* and *xopE5*. The *graminis* strain Xtg29 has *xopE1*, *xopE2* and *xopE4*. Besides the difference on the presence and absence of T3Es, variability in gene copy number exists among pathovars or within a pathovar. Two *xopAF* in *undulosa* strains DAR61454, XT-Rocky, XT5523, and XT4699, but only one copy in *undulosa* strains ICMP11055, LW16, CS4, and XT5791. All four *translucens* strains have two copies of *xopAF*, while *cerealis* strains harbor one. Two copies of *xopL* are present in all *undulosa* and *translucens* strains, while there are four copies in *cerealis* strains. Two copies of *xopP* in *undulosa* strains, and three copies in *translucens* and *cerealis* strains with the exception that *translucens* strain B1 only has two *xopP*. Four *undulosa* strains have frameshift mutations in one *xopP*. Consistent with a previous study on T3Es in *X. axonopodis* pv. *manihotis* (Bart et al., 2012), frameshift mutations of T3Es is not rare in *X. translucens* strains. Frameshifts in the coding sequence of one *avrBs2* is present in *undulosa* strain P3. Two *cerealis* strains have frameshift mutations in *xopR* compared to *undulosa* and *translucens* strains. All four *translucens* strains harbor frameshift mutations in *xopAP* compared to *undulosa* and *cerealis* strains.

CRISPR clusters are present in most *X. translucens* strains

Clustered regularly interspaced short palindromic repeat (CRISPR) and a set of CRISPR-associated genes (*Cas*) comprise an adaptive immune system against foreign DNA, which has been recently identified in numerous bacterial genomes (Barrangou et al., 2007). The spacer sequences between repeats in the CRISPR locus usually correspond to a homologous sequence, called proto-spacer, in a foreign DNA element and subsequently leading to the cleavage of the DNA element with the involvement of Cas proteins (Garneau et al., 2010). The sequencing and monitoring of repeat and spacer array of CRISPR in strains may provide critical insights into the coevolution relationship between prokaryotic strains and invader phage DNA, and, to some extent, may record the DNA elements of past invaders that the strains had ever been encountered (Horvath et al., 2008, Marraffini & Sontheimer, 2010).

Genomic sequence data of the *X. translucens* strains in this study and previously sequenced strains revealed CRISPR loci present in most *X. translucens* strains, except for XT4699, LG48, and XT123. If present, only one CRISPR locus is present in each strain. The CRISPR locus is often flanked by transposon gene elements, which may be one factor accounting for hypervariable nature. Annotation of CRISPR *Cas* genes of all *X. translucens* strains and the *X. oryzae* pv. *oryzae* PXO99 strain indicates they belong to the Type I-C (Dvulg or CASS1) in the classification of CRISPR-Cas systems, based on the architecture of Cas operons (Makarova et al., 2011). The phylogenetic relationship, based on the nucleotide sequences of *Cas* gene clusters, among these strains is somewhat different from the MLSA and whole genome SNP comparison results (Figure 3-11; Figure 3-6), suggesting the evolution of CRISPR-Cas system is distinct from genome evolution. LW16 and CS4 strains are special

variants in this subtype, harboring two *Cas3* helicase genes, while other strains only have one, which may account for the their distance from other strains in the phylogenetic trees of *Cas* genes. The direct repeat sequence in the CRISPR loci is conserved among PXO99 and *X. translucens* strains, although they are quite distinct in the CRISPR *Cas* loci.

The comparison study of spacer elements among strains reveals *undulosa* and *translucens* pathovars could share a large number of identical spacers. B2 (*X. translucens* pv. *translucens*) and ICMP11055 (*X. translucens* pv. *undulosa*) share 49 identical spacers, XT8 (*X. translucens* pv. *translucens*) shares 44 identical spacers with ICMP11055 and the two *translucens* pathovars XT8 and B2 share 55 identical spacers (Figure 3-12A). This indicates these three strains may experience an overlapping history of fighting against similar invader DNAs or acquisition of CRISPR elements from the same ancestor by lateral gene transfer although ICMP11055 is quite different from other two strains based on either pathogenicity types or the MLSA result. The highly close relationship among the *Cas* genes content among ICMP11055, B2 and XT8 is consistent with great sharing of spacer elements (Figure 3-11). In addition, strains B1 (*X. translucens* pv. *translucens*) and P3 (*X. translucens* pv. *undulosa*) share twenty-three identical spacers, while B1 only shares three identical spacers with other three *translucens* pathovars DSM18974, B2, XT8 (Figure 3-12A). LW16 and CS4, which have close relationship on CRISPR *Cas* loci (Figure 3-11), share thirty-two identical spacers, the maximum sharing observed in the panel B of Figure 3-12. Although there is some varying degree of sharing with regards to the older spacer elements for the *X. translucens* strains, the most recently acquired spacer sequences are different from each other (Figure 3-12). However, nine strains, XT-Rocky (collected from KS in 2009), XT5523 (collected from Canada in 1966), LG54, CR31, CS2, CS22, LB5, LB10 (collected from ND in 2009) and DAR61454 (collected from Australia

in 1988), share identical elements of *Cas* cluster, repeats, and spacer array (Figure 3-12B), implying that CRISPR *Cas* loci can be maintained over decades. The spacer sequences of *X. translucens* pv. *cerealis* CFBP2541 and *X. translucens* pv. *graminis* ART-Xtg29 strains were not detected in any other *X. translucens* strains.

TAL effector gene content

The repetitive regions often cause assembly gaps in whole genome assemblies based on the Illumina or other short reads data. To date, the assemblies of high copies of TAL effector genes in *Xanthomonas* genomes are still intractable using short reads. With the PacBio long reads, we are able to resolve TAL effector genes in both XT4699 and ICMP11055 genomes. Eight and seven TAL effector genes were identified in our final assembly genome of XT4699 and ICMP11055, respectively. The number and lengths of TAL effector genes of XT4699 were confirmed by DNA Southern hybridization, in which the size of *Bam*HI fragments with TAL effector genes was consistent with the size of blotted bands (Figure 3-13). Moreover, Sanger sequences of six TAL effector genes cloned by PCR with high fidelity polymerase were identical to the sequences of corresponding TAL genes in the assembly genome of XT4699 (Table 3-4). Additionally, *Bam*HI fragments of five TAL effector genes in XT4699 were subcloned using restriction enzyme digestion method and sequence was consistent with the assembly genome (data not shown).

Besides TAL effector genes in XT4699 and ICMP11055, four TAL genes from XT-Rocky strain were subcloned and sequenced. Two different TAL effector genes from *cerealis* strain CFBP2541 were also known (Pesce et al., 2015). TAL effector gene is not present in the genome of *X. translucens* pv. *graminis* ART-Xtg29 strain (Wichmann et al., 2013). The sequence of all RVDs (repeat variable di-residues) available from these four strains reveals

some, such as YD, YK, QD, KG, Y*, NF, KI and GI, are unusual RVDs, rarely found in *X. oryzae* and *X. campestris* strains (Figure 3-5 and 3-14). Another special feature is that most of TAL effector genes in *X. translucens* strains possess 34- and 35-amino-acid repeats in a single gene, except the XT4699-tal4 and ICMP11055-tal5, which are identical in RVDs and only having 35-amino-acid repeats (Figure 3-5 and 3-14). These two TAL genes are closely related to each other but distant from other TAL effectors in the *X. translucens* based on the phylogenetic relationship of N- and C-teriminal amino acid sequences (Figure 3-15). TAL effectors in *X. translucens* strains are much closer to AvrBs3 and PthXo1 than Brg11 in *Ralstonia Solanacearum* GMI1000 (Figure 3-15 and 3-16). The two TAL effectors from *cerealis* strain CFBP2541 are quite distinct from TAL effectors of *undulosa* strains in both RVD alignment and homology of N- and C-teriminal amino acid (Figure 3-15 and 3-14). The XT4699 genome annotation shows that most TAL effector genes are flanked by transposon elements, indicating their potential to undergo rapid evolution (Figure 3-17).

Comparison of TAL effector RVD revealed *tal7* was conserved among *undulosa* strains XT4699, XT-Rocky and ICMP11055. To check this TAL effector gene was conserved, the primer pairs used for amplification of *tal7* genes in XT4699 were applied for 10 other ND *undulosa* strains. PCR bands with similar size as *tal7* of XT4699 were present in 8 of 12 strains in total (Figure 3-18). Sequencing of four of these PCR fragments revealed all of them were TAL genes, possessing exactly the same RVDs as XT4699-Tal7 except for one amino acid variation in corresponding TALs of both CS4 and LB5. This TAL gene is also conserved in the ICMP11055 strain except for one amino acid difference in 11th RVD (Figure 3-5)

TAL effectors from XT4699 modulate transcriptional profiling of wheat

To identify the host genes targeted by TAL effectors from *X. translucens* strains, microarray analysis of wheat leaf samples that were inoculated with strains XT4699, XT-Rocky, or 4699*hrcC*$^-$ were performed. Wheat gene expression profiles upon inoculation of XT4699 and 4699*hrcC*$^-$ mutant were found to be distinct from each other, revealing approximately 250 genes with at least a 10-fold difference of expression (unpublished data).

To investigate if host gene expression associated with specific TAL effectors, mutations were generated in TAL effector genes of XT4699 (Figure 3-19). Expression analysis of each TAL effector gene mutant revealed three differentially induced genes, corresponding with Affymetrix probes Ta.7291.1.S1_s_at, Ta.14824.1.S1_at and Ta.9765.1.S1_at, from the microarray analysis of XT4699 vs XT4699*hrcC*$^-$ comparison (Table 3-7), that were not induced significantly by the infections of mutants M2 (*tal6*$^-$), M3 (*tal7*$^-$) and M4 (*tal1*$^-$), respectively (Figure 3-20). The probes Ta.7291.1.S1_s_at, Ta.14824.1.S1_at and Ta.9765.1.S1_at corresponded to coding regions that were predicted to encode a succinate dehydrogenase subunit, a choline transporter related protein and a cell wall invertase, respectively.

Two genes from the microarray analysis, corresponding with to probes Ta.7291.1.S1_s_at and Ta.14164.1.S1_x_at, were induced by infection with XT4699 strain but not by XT-Rocky (Table 3-7). As qRT-PCR assays revealed, M2, mutant of *tal6*, failed to induce Ta.7291.1.S1_s_at (Figure 3-20 and Figure 3-21A) and also lost the ability to induce the gene corresponding to Ta.14164.1.S1_x_at, which is predicted to encode a bHLH family transcriptional factor (Figure 3-21B). Comparison of RVDs of TAL effectors between XT4699 and XT-Rocky revealed there was one TAL effector, XT-Rocky-3E3, highly similar to XT4699-Tal6 except for variation in the last 5 RVDs (Figure 3-5). The qRT-PCR assays indicated these two genes could also not be induced by *X. translucens* pv. *translucens* XT8, *X. translucens* pv.

cerealis XT123 or *X. translucens* pv. *undulosa* XT5523 (the latter being very close to XT-Rocky in phylogeny), while the two genes were induced by three other *X. translucens* pv. *undulosa* strains (Figure 3-22). Complementation of the M2 mutant with *tal6* resulted in restored expression of both genes (Figure 3-21A and 3-21B). The transcriptional induction of the two genes by XT4699 occurred in the presence of the eukaryotic translation inhibitor cycloheximide (CHX) (Figure 3-21C and 3-21D), which had been used for distinguishing direct or indirect target of TAL effectors from *Xanthomonas* strains in previous study, indicating both genes are likely the direct targets of the XT4699-Tal6 effector (Hu et al., 2014, Strauß et al., 2012).

Discussion

Genome assembly of a pathogen is a foundation for further genetic analyses into the pathogenicity mechanisms. The advances of genome sequencing technologies, especially long-molecular technologies such as PacBio sequencing, and better engineering of assembly algorithms facilitate genome assembly (Koren & Phillippy, 2015). In this study, complete genome sequence were generated for *X. translucens* pv. *undulosa* strain XT4699. Our results demonstrated that the PacBio data was sufficient to achieve a high-accuracy assembly of the relatively repetitive *Xanthomonas* genome. Intact sequences of TAL effector genes can be used to predict the potential targets in host genomes, thereby facilitating the identification of possible resistant or susceptible genes (Boch et al., 2009, Moscou & Bogdanove, 2009). However, given a relatively high sequencing cost, it is still not realistic to sequence every bacterial strain with PacBio to achieve a finished assembly. To obtain most genomic content of a bacterial genome, Illumina sequencing is more cost-efficient. We, therefore, generated Illumina sequences for 20

strains, including XT4699, to determine the relatedness and compare effector contents among them. The comparison between the XT4699 Illumina-only assembly and the complete XT4699 assembly shows that approximately 97% the complete genome is covered by the Illumina-only assembled contigs. The characterization of assembly gaps on the Illumina-only assembly indicates that gaps are largely located at repetitive regions and high GC contents. The lengths of gaps at repetitive regions tend to be linearly correlated to repeat lengths while the gaps at high GC regions are typically small. Gaps due to a high GC content may, in the future, be reduced by using a new Illumina PCR-free library preparation protocol, which is expected to ameliorate amplification biases associated with high GC contents. To resolve repeats, especially long and complex repeats, PacBio or other emerging long read sequencing technologies, such as Nanopore (Ashton et al., 2015, Loman & Watson, 2015), are needed. The improvement of these long read sequencing technologies should dramatically reduce the sequencing cost per bacterial strain in the near future. The ability to sequence and completely assemble a large number of independent strains would greatly accelerate genome comparison and the identification of bacterial virulent factors.

The comparison of Type III effector (T3E) content in *X. tranlucens* strains identified both conservative (N=23) and variable T3Es among three pathvors, *undulosa, translucens* and *cerealis*. Effector triggered immunity (ETI) is well known for some T3Es, such as AvrBs2, when the corresponding R genes are present in the host plants (Jones & Dangl, 2006, Tai et al., 1999). These highly conserved T3Es provide an opportunity to engineer R genes-mediated resistance in crop breeding controlling the bacterial leaf streak disease although there is currently no clue of R genes recognizing the effectors or other conserved components from *X. translucens* strains. Presence and absence variation, copy number variation and frameshift mutations are the three

major large-effect variations. In addition, variable TAL effector repertories may also be another source. The *X. tranlucens* pv. *cerealis* strain CFBP2541 only contains two TAL effectors, sharing no similarity with TAL effectors from *X. tranlucens* pv. *undulosa* strains based on the RVD sequences. Variation in T3Es may be one factor accounting for pathogenic diversity and host range determinants of *X. tranlucens* strains. It has been shown that the pathogenic and genetic diversity was present in *X. tranlucens* strains collected from five different locations in North Dakota (Adhikari et al., 2012). There are reports of host specific virulence genes acted as host range determinants (Mellano & Cooksey, 1988, Waney et al., 1991). CFBP2541 (*X. tranlucens* pv. *cerealis*), a phylogenetically close strain to XT123 (*X. tranlucens* pv. *cerealis*), was reported to cause severe symptoms on 'Morex' barley (Pesce et al., 2015). XT123 was only induced chlorosis on 'Morex' barley. However, pathogenicity diversity among different *X. translucens* strains may also derive from variation in other genomic islands related with bacterial fitness. Genomic comparison between XT4699 and ICMP11055 reveals the latter harbors extra genomic regions, including prophage elements and specific gene cluster of *tadA*, *tadB* and *tadC* for Type II/IV secretion systems, which are reported to be potential virulence factors (Planet et al., 2003, Rice et al., 2009, Wagner & Waldor, 2002). Functional mutation and complementation in genes related with virulence is needed for detecting the virulence contribution of genes and explaining their roles leading to difference in pathogenicity types.

The TAL effector triggered susceptibility has been demonstrated in four plant species. The TAL effectors are a large family of closely related type III effector proteins, which transcriptionally activate host gene expression by directly interacting with promoter elements of host genes and have a varying degree of contribution in bacterial virulence, proliferation and other disease symptomatology (Boch & Bonas, 2010, White & Yang, 2009). So far, TAL

effectors with virulence contribution have been detected in strains of *X. oryzae* pv. *oryzae*, *X. citri* pv. *citri*, *X. campestris* pv. *vesicatoria*, *X. campestris* pv. *malvacearum* and *X. oryzae* pv. *oryzicola* (Antony et al., 2010, Cernadas et al., 2014, Hu et al., 2014, Kay et al., 2007, Li et al., 2014, Wichmann & Bergelson, 2004, Yang et al., 2006, Yang et al., 1994, Zhou et al., 2015). However, it is still an open question whether TAL effector play roles in triggering susceptibility in bacterial leaf streak on wheat. In this study, we are confirmed that TAL effectors do modulate the transcriptional profiling of wheat host. The XT4699-Tal6 effector affect transcriptional level of two candidate target genes, encoding a bHLH transcriptional factor and a succinate dehydrogenase subunit, respectively. It is still unknown whether or not they are directly or indirectly targeted by Tal6. Searching in the promoter regions of the two genes using the RVD sequence of Tal6 did not give positive EBE above the cutoff value (https://tale-nt.cac.cornell.edu/node/add/talef-off). We also failed to detect the direct interaction between Tal6 and 400 bp promoter sequence upstream of translation start site of bHLH gene in the assays of Agrobacteria-mediated transient expression in *Nicotiana tabacum* KY-14 (data not shown). In this study, no TAL effector mutants suffered an obvious loss of virulence when compared to the WT strain in the bacterial growth population assays (Figure 3-9). It is not known if TAL effectors contribute to bacterial virulence in other ways as AvrBs3 from *X. campestris* pv. *vesicatoria* and Tal2g from *X. oryzae* pv. *oryzicola* (Cernadas et al., 2014, Kay et al., 2007). Hence, other approaches to quantitatively compare the bacterial virulence between TAL effector mutants and WT strain may be needed to further characterize the role of the TAL effectors in bacterial leaf streak disease.

Reference

Adhikari TB, Gurung S, Hansen JM, Bonman JM, 2012. Pathogenic and Genetic Diversity of Xanthomonas translucens pv. undulosa in North Dakota. *Phytopathology* **102**, 390-402.

Alexeyev M, 1999. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. *Biotechniques* **26**, 824-6, 8.

Antony G, Zhou J, Huang S, *et al.*, 2010. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell* **22**, 3864-76.

Ashton PM, Nair S, Dallman T, *et al.*, 2015. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nature Biotechnology* **33**, 296-300.

Ausubel F, Brent R, Kingston R, *et al.*, 1994. Current Protocols in Molecular Biology John Wiley & Sons New York NY.

Barrangou R, Fremaux C, Deveau H, *et al.*, 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-12.

Bart R, Cohn M, Kassen A, *et al.*, 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences* **109**, E1972-E9.

Bashir A, Klammer AA, Robins WP, *et al.*, 2012. A hybrid approach for the automated finishing of bacterial genomes. *Nature biotechnology* **30**, 701-7.

Bentley DR, 2006. Whole-genome re-sequencing. *Current opinion in genetics & development* **16**, 545-52.

Boch J, Bonas U, 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology* **48**, 419-36.

Boch J, Scholze H, Schornack S, *et al.*, 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-12.

Bragard C, Singer E, Alizadeh A, Vauterin L, Maraite H, Swings J, 1997. Xanthomonas translucens from small grains: diversity and phytopathological relevance. *Phytopathology* **87**, 1111-7.

Bragard C, Verdier V, Maraite H, 1995. Genetic diversity among Xanthomonas campestris strains pathogenic for small grains. *Applied and environmental microbiology* **61**, 1020-6.

Cernadas RA, Doyle EL, Niño-Liu DO, *et al.*, 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathog* **10**, e1003972.

Chin C-S, Alexander DH, Marks P, *et al.*, 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature methods* **10**, 563-9.

Depristo MA, Banks E, Poplin R, *et al.*, 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics* **43**, 491-8.

Gardiner DM, Upadhyaya NM, Stiller J, *et al.*, 2014. Genomic analysis of Xanthomonas translucens pathogenic on wheat and barley reveals cross-kingdom gene transfer events and diverse protein delivery systems. *PLoS ONE* **9**.

Garneau JE, Dupuis M-È, Villion M, *et al.*, 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67-71.

Hopkins CM, White F, Choi S, Guo A, Leach J, 1992. Identification of a family of avirulence genes from Xanthomonas oryzae pv. oryzae. *Mol. Plant-Microbe Interact* **5**, 451-9.

Horvath P, Romero DA, Coûté-Monvoisin A-C, *et al.*, 2008. Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *Journal of bacteriology* **190**, 1401-12.

Hu Y, Zhang J, Jia H, *et al.*, 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences* **111**, E521-E9.

Jones JD, Dangl JL, 2006. The plant immune system. *Nature* **444**, 323-9.

Kay S, Hahn S, Marois E, Hause G, Bonas U, 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-51.

Koren S, Phillippy AM, 2015. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr Opin Microbiol* **23**, 110-20.

Koren S, Schatz MC, Walenz BP, *et al.*, 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nature biotechnology* **30**, 693-700.

Kurtz S, Phillippy A, Delcher AL, *et al.*, 2004. Versatile and open software for comparing large genomes. *Genome biology* **5**, R12.

Li H, Durbin R, 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589-95.

Li R, Zhu H, Ruan J, *et al.*, 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome research* **20**, 265-72.

Li Z, Zou L, Ye G, *et al.*, 2014. A Potential disease susceptibility gene CsLOB of citrus is targeted by a major virulence effector PthA of Xanthomonas citri subsp. citri. *Molecular plant* **7**, 912-5.

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods* **25**, 402-8.

Loman NJ, Watson M, 2015. Successful test launch for nanopore sequencing. *Nat Methods* **12**, 303-4.

Makarova KS, Haft DH, Barrangou R, *et al.*, 2011. Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* **9**, 467-77.

Marraffini LA, Sontheimer EJ, 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics* **11**, 181-90.

Mckenna A, Hanna M, Banks E, *et al.*, 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-303.

Mellano VJ, Cooksey DA, 1988. Development of host range mutants of Xanthomonas campestris pv. translucens. *Applied and environmental microbiology* **54**, 884-9.

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

Overbeek R, Olson R, Pusch GD, *et al.*, 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic acids research* **42**, D206-D14.

Paradis E, Claude J, Strimmer K, 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-90.

Pesce C, Bolot S, Cunnac S, *et al.*, 2015. High-quality draft genome sequence of the Xanthomonas translucens pv. cerealis pathotype strain CFBP 2541. *Genome announcements* **3**, e01574-14.

Pieretti I, Royer M, Barbe V, *et al.*, 2009. The complete genome sequence of Xanthomonas albilineans provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. *BMC genomics* **10**, 616.

Planet PJ, Kachlany SC, Fine DH, Desalle R, Figurski DH, 2003. The widespread colonization island of Actinobacillus actinomycetemcomitans. *Nature genetics* **34**, 193-8.

Potnis N, Krasileva K, Chow V, *et al.*, 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC genomics* **12**, 146.

Rice SA, Tan CH, Mikkelsen PJ, *et al.*, 2009. The biofilm life cycle and virulence of Pseudomonas aeruginosa are dependent on a filamentous prophage. *The ISME journal* **3**, 271-82.

Siguier P, Pérochon J, Lestrade L, Mahillon J, Chandler M, 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic acids research* **34**, D32-D6.

Strauß T, Van Poecke RM, Strauß A, *et al.*, 2012. RNA-seq pinpoints a Xanthomonas TAL-effector activated resistance gene in a large-crop genome. *Proceedings of the National Academy of Sciences* **109**, 19480-5.

Tai TH, Dahlbeck D, Clark ET, *et al.*, 1999. Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proceedings of the National Academy of Sciences* **96**, 14153-8.

Wagner PL, Waldor MK, 2002. Bacteriophage control of bacterial virulence. *Infection and Immunity* **70**, 3985-93.

Waney V, Kingsley M, Gabriel D, 1991. Xanthomonas campestris pv. translucens genes determining host-specific virulence and general virulence on cereals identified by Tn5-gusA insertion mutagenesis. *Mol. Plant-Microbe Interact* **4**, 623-7.

White FF, Potnis N, Jones JB, Koebnik R, 2009. The type III effectors of Xanthomonas. *Molecular Plant Pathology* **10**, 749-66.

White FF, Yang B, 2009. Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction. *Plant Physiology* **150**, 1677-86.

Wichmann F, Vorhölter FJ, Hersemann L, *et al.*, 2013. The noncanonical type III secretion system of Xanthomonas translucens pv. graminis is essential for forage grass infection. *Molecular Plant Pathology* **14**, 576-88.

Wichmann G, Bergelson J, 2004. Effector genes of Xanthamonas axonopodis pv. vesicatoria promote transmission and enhance other fitness traits in the field. *Genetics* **166**, 693-706.

Yang B, Sugio A, White FF, 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* **103**, 10503-8.

Yang B, White FF, 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Molecular plant-microbe interactions* **17**, 1192-200.

Yang Y, De Feyter R, Gabriel DW, 1994. Host-specific symptoms and increased release of Xanthomonas citri and X. campestris pv. malvacearum from leaves are determined by the 102-bp tandem repeats of pthA and avrb6, respectively. *MPMI-Molecular Plant Microbe Interactions* **7**, 345-55.

Young J, Park D-C, Shearman H, Fargier E, 2008. A multilocus sequence analysis of the genus Xanthomonas. *Systematic and Applied Microbiology* **31**, 366-77.

Zhou J, Peng Z, Long J, *et al.*, 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant Journal* **82**, 632-43.

Table 3-1. Bacterial isolation information, pathogenicity type and genome assembly statistics for *X. translucens* strains

Strain name	Isolation			Pathogenicity type ±	Assembly statistics			
	Year	Host origin	Place		Contig number	Genome size (bp)	N50 (bp)	Maximum contig length (bp)
Strains used in this study								
XT-4699ǂ	1999	Wheat	KS,USA	A	1	4,561,137	NA	NA
ICMP11055ǂ	1983	Wheat	Iran	A	1	4,761,583	NA	NA
XT-Rocky	2009	Wheat	KS,USA	A	521	4,459,068	21,643	85,632
XT8	1942	Barley	Canada	B	152	4,617,556	79,563	279,321
XT123	1952	Barley	Canada	A	344	4,284,749	28,687	135,676
XT130	1939	NA	Canada	A	329	4,654,290	95,229	230,109
XT5523*	1966	Wheat	Canada	A	330	4,665,768	111,158	270,930
XT5770	NA	NA	Canada	A	533	4,617,837	27,152	149,044
XT5791	1969	Wheat	Canada	A	738	4,719,363	62,926	185,369
B1	2009	Barley	ND,USA	B	642	4,824,098	28,186	96,565
B2	2009	Barley	ND,USA	B	283	4,503,259	38,929	154,569
P3	2009	Wheat	ND,USA	A	161	4,522,131	78,928	240,099
LW16	2009	Wheat	ND,USA	A	325	4,600,125	33,769	145,918
LB5	2009	Wheat	ND,USA	A	434	4,766,161	111,172	269,877
LB10	2009	Wheat	ND,USA	A	157	4,543,985	98,403	211,728
LG48	2009	Wheat	ND,USA	A	293	4,486,555	38,250	111,281
LG54	2009	Wheat	ND,USA	A	213	4,623,672	123,210	274,463
CS2	2009	Wheat	ND,USA	A	285	4,722,832	85,714	207,734
CS22	2009	Wheat	ND,USA	A	129	4,605,395	95,219	274,395
CR31	2009	Wheat	ND,USA	A	386	4,720,715	62,461	158,913
CS4	2009	Wheat	ND,USA	A	498	4,779,534	37,460	107,554
Strains from references								
CFBP2541*	1941	Bromegrass	USA	A	31	4,515,938	1,399,657	1,809,000
DSM18974*	1933	Barley	MN,USA	B	551	4,463,577	13,041	55,714
DAR61454	1988	Wheat	Australia	A	404	4,452,091	27,210	121,856
ART-Xtg29	NA	Forage grass	Switzerland	C	788	4,100,864	8,376	37,754

ǂ indicates genome assembly with PacBio long-read and Illumina Miseq sequencing.

* Type strains was marked with *.

± Strains virulent on wheat and barley are classified as pathogenicity type A, strains virulent only on barley are considered as type B and strains virulent on forage grass are assigned to type C.

NA means not available or not applicable.

Table 3-2. Primers used in TAL mutant validation, cloning of TAL effector genes and qRT-PCR.

Tal1-F:	caccGAATTCGACATGAACACCGACCACAC
Tal1-R:	atatAAGCTTCCGATCCTATGATGCCAAAG
Tal2-F:	caccGAATTCACGCTTCTTCACGTCCAATC
Tal2-R:	atatAAGCTTGTTCTTCAACGCGTGGATCT
Tal3-F:	caccGAATTCTTACCCCTTGGCAAAGTACG
Tal3-R:	atatAAGCTTGCTCAACTGGATCGACTGGT
Tal4-F:	caccGAATTCGACCCTTGAAGCTGAAGCAG
Tal4-R:	atatAAGCTTGCACCATCTGGTCCATCTCT
Tal5-F:	caccGAATTCCTGCTAGTGCACTCGGATCA
Tal5-R:	atatAAGCTTGGGGATTTAGAGCGGCTAAC
Tal6-F:	atatAAGCTTGTTCAGCCGACAGACGTGTA
Tal6-R:	caccGAATTCATGTCTCGCTCTTCGGTGAC
Tal7-F: *	caccGAATTCCCCTCAGTAGCTGGGGTGTA
Tal7-F2:*	caccGAATTCCACATGGCTGGTCACATACG
Tal7-R:	atatAAGCTTTGCCAAGACACTGAAGCAAG
Tal8-F:	caccGAATTCCTTCATATCCCCGAGCAAAC
Tal8-R:	atatAAGCTTTCAGGTAGTCTCCGACTGAG
Forall-Val	AACAAGCCAGGGATGTAACG
XTThrcC-F	GCAGAAGTTCCTGGACACGCTGG
XTThrcC-R	ATCGATCTGCAGCAGTTGCGGAC
XTTHrcC-Out	CGACACCCGCGTCATCAGAAAAC
07KM-Val	TTCTGCGGACTGGCTTTCTACG

4699N-TAL-F	aataGGTACCATGAGGTGCAATCGGSTCT
4699N-TAL-R	atatCTCGAGCCAACTGTAACGGTGGAGCT
Succinate Dehy-F1	ACCAGGTGACAAGCAGAA
Succinate Dehy-R1	ATCTTGGAAGTTGGCAATGAA
EF1A-F1	CAGATTGGCAACGGCTACG
EF1A-R1	CGGACAGCAAAACGACCAAG
bHLH-F	TCTTTTCCTTCCTGCTCCC
bHLH-R	GCATAAATGAAACACAGTGGAGTA

*Tal7-F was used for identifying TAL mutant, while the Tal7-F2 was used for cloning TAL gene.

Table 3-3. The sequence difference between the reference genome and the Illumina assembly

Reference	Position on reference	Bases of reference	Bases of the Illumina assembly
CP008714	693502	.	C
CP008714	2210565	A	C
CP008714	3303063	T	C
CP008714	3499763	.	C

Table 3-4. Summary of alignments of Sanger reads from TAL genes to the assembly genome

Query	Identity (%)	Overlap (bp)	Coverage (%)	Query_start	Query_end	Assembly_start	Assembly_end	E-value
Tal5-pBS-T3_R	100	506	506	1	506	2,856,687	2,857,192	0
Tal5-pBS-T7	100	918	918	1	918	2,860,925	2,860,008	0
Tal5-pBS-TALclone-F1	100	869	869	1	869	2,860,418	2,859,550	0
Tal5-pBS-TALclone-F2	100	746	746	1	746	2,859,667	2,858,922	0
Tal5-pBS-TALclone-R2	100	817	817	1	817	2,857,531	2,858,347	0
Tal3-pBS-T3	100	874	874	1	874	1,991,073	1,991,946	0
Tal3-pBS-T7	100	555	555	1	555	1,995,163	1,994,609	0
Tal3-pBS-TALclone-F1	100	788	788	1	788	1,994,634	1,993,847	0
Tal3-pBS-TALclone-R2	100	704	704	1	704	1,991,844	1,992,547	0
Tal2-pBS-T3	100	677	677	1	677	608,874	608,198	0
Tal2-pBS-T7	100	680	680	1	680	604,409	605,088	0
Tal2-pBS-TALclone-F1	100	823	823	1	823	605,076	605,898	0
Tal2-pBS-TALclone-F2	100	745	745	1	745	605,841	606,585	0
Tal2-pBS-TALclone-R2	100	684	684	1	684	608,167	607,484	0
Tal4-pBS-T3	100	833	833	1	833	2,673,556	2,672,724	0
Tal4-pBS-T7	100	885	885	1	885	2,669,478	2,670,362	0
Tal4-pBS-TALclone-F2	100	815	815	1	815	2,670,781	2,671,595	0
Tal4-pBS-TALclone-R2	100	812	812	1	812	2,672,917	2,672,106	0
Tal6-pBS-T3	100	762	762	1	762	2,865,507	2,864,746	0
Tal6-pBS-TALclone-F1	100	824	824	1	824	2,865,014	2,864,191	0
Tal6-pBS-TALclone-R2	100	801	801	1	801	2,862,257	2,863,057	0
Tal6-pBS-TALclone-F2	100	948	948	1	948	2,864,262	2,863,315	0
pBS-tal1-TALclone-F1	100	751	751	1	751	584,355	585,105	0
pBS-tal1-TALclone-F2	100	708	708	1	708	585,112	585,819	0
pBS-tal1-TALclone-R2	100	814	814	1	814	587,241	586,428	0
pBS-tal1-T3	100	867	867	1	867	587,876	587,010	0
pBS-tal1-T7	100	802	802	1	802	583,743	584,544	0
pBS-tal1- tal1-F	100	750	750	1	750	583,795	584,544	0
pBS-tal1- tal1-R	100	863	863	1	863	587,822	586,960	0

Table 3-5. Genome content comparison between XT4699 and ICMP11055

	XT4699	ICMP11055
General features		
Genome size (bp)	4,561,137	4,761,583
GC content (%)	68.1	67.8
Number of predicted CDS genes	3,528	3,656
Pseudogenes	94	236
Genes with frameshift	54	83
TAL effector genes	8	7
Non-TAL T3E genes	32	32
rRNA operons	2	2
tRNA	54	54
CRISPR array	Not detected	1
PIP genes	49	53
Insertion sequence elements (complete/partial)	74/56	83/58

Table 3-6. SNP difference count among *Xanthomonas* strains in a total of 9836 SNPs.

	XT4699	B1	B2	CR31	CS2	CS22	CS4	ICMP11055	LB10	LB5	LG48	LW16	P3	Rocky	X.albiline	DAR6145	Xtg29	DSM1897	XT123	KACC103	MAFF311	PXO99A	XT8
XT4699	0	663	697	115	113	119	111	157	119	148	18	129	147	119	4705	118	1156	769	1354	5810	5814	5810	702
B1	663	0	318	637	627	633	637	616	633	618	649	637	597	629	4464	630	917	252	1150	5564	5568	5562	331
B2	697	318	0	671	665	667	671	641	667	636	683	681	655	663	4479	664	941	283	1218	5593	5597	5591	35
CR31	115	637	671	0	30	24	97	136	24	53	99	133	129	28	4681	29	1122	741	1326	5782	5786	5782	682
CS2	113	627	665	30	0	8	103	134	8	37	97	143	129	16	4680	17	1120	739	1331	5785	5789	5785	676
CS22	119	633	667	24	8	0	107	134	2	31	103	147	129	10	4681	11	1122	739	1329	5789	5793	5789	678
CS4	111	637	671	97	103	107	0	144	107	130	95	87	140	105	4666	106	1106	745	1302	5769	5773	5768	679
ICMP11055	157	616	641	136	134	134	144	0	134	156	141	150	128	130	4674	131	1099	715	1306	5784	5788	5784	647
LB10	119	633	667	24	8	2	107	134	0	31	103	147	128	10	4681	11	1122	739	1329	5789	5793	5789	678
LB5	148	618	636	53	37	31	130	156	31	0	132	174	156	39	4668	40	1113	712	1323	5779	5783	5779	647
LG48	18	649	683	99	97	103	95	141	103	132	0	113	131	103	4707	102	1143	755	1341	5808	5812	5808	686
LW16	129	637	681	133	143	147	87	150	147	174	113	0	133	145	4675	146	1106	747	1303	5776	5780	5776	681
P3	147	597	655	129	129	129	140	128	128	156	131	133	0	129	4682	130	1114	702	1303	5780	5784	5780	661
Rocky	119	629	663	28	16	10	105	130	10	39	103	145	129	0	4678	3	1117	735	1324	5787	5791	5787	674
X.albilineans PC73	4705	4464	4479	4681	4680	4681	4666	4674	4681	4668	4707	4675	4682	4678	0	4679	4286	4471	4427	5857	5856	5858	4482
Xt_DAR61454	118	630	664	29	17	11	106	131	11	40	102	146	130	3	4679	0	1118	736	1325	5788	5792	5858	675
ART-Xtg29	1156	917	941	1122	1120	1122	1106	1099	1122	1113	1143	1106	1114	1117	4286	1118	0	931	1007	5403	5407	5398	937
Xt DSM18974	769	252	283	741	739	739	745	715	739	712	755	747	702	735	4471	736	931	0	1206	5590	5594	5588	290
XT123	1354	1150	1218	1326	1331	1329	1302	1306	1329	1323	1341	1303	1303	1324	4427	1325	1007	1206	0	5541	5545	5541	1217
Xoo_KACC10331	5810	5564	5593	5782	5785	5789	5769	5784	5789	5779	5808	5776	5780	5787	5857	5788	5403	5590	5541	0	39	55	5599
Xoo MAFF311018	5814	5568	5597	5786	5789	5793	5773	5788	5793	5783	5812	5780	5784	5791	5856	5792	5407	5594	5545	39	0	60	5603
Xoo_PXO99A	5810	5562	5591	5782	5785	5789	5768	5784	5789	5779	5808	5776	5780	5787	5858	5788	5398	5588	5541	55	60	0	5597
XT8	702	331	35	682	676	678	679	647	678	647	686	681	661	674	4482	675	937	290	1217	5599	5603	5597	0

Table 3-7. Microarray analysis of gene expression of inoculated wheat leaves by XT4699, XT-Rocky and XT4699*hrcC*$^{-}$

Affymetrix ID	Annotation	4699/*hrcC* cor	Rocky/*hrcC* cor	4699 Ave	Rocky Ave	*hrcC*$^{-}$ Ave
Ta.14164.1.S1_x_at	bHLH transcriptional factor	110.5	1.0	12741.4	15.7	16.2
Ta.14164.1.S1_s_at	bHLH transcriptional factor	107.5	1.0	12523.1	8.7	17.4
Ta.9765.1.S1_at	Cell wall invertase	15.9	15.7	33720.8	33383.4	2025.1
Ta.7814.1.S1_at	Putative nodule-specific protein	15.6	16.7	27728.1	29670.9	1679.5
Ta.7291.1.S1_s_at	succinate dehydrogenase flavoprotein subunit	15.0	1.0	7120.9	393.3	381.1
Ta.14824.1.S1_at	choline transporter-related	12.2	11.3	2763.0	2566.5	134.9

Note: all the scores of wheat gene expression were averaged from scores of 3 biological replicates; The corrected fold is calculated by the fomula Fold=Scores of XT4699(or XTRocky)+100/Scores of *hrcC*+100

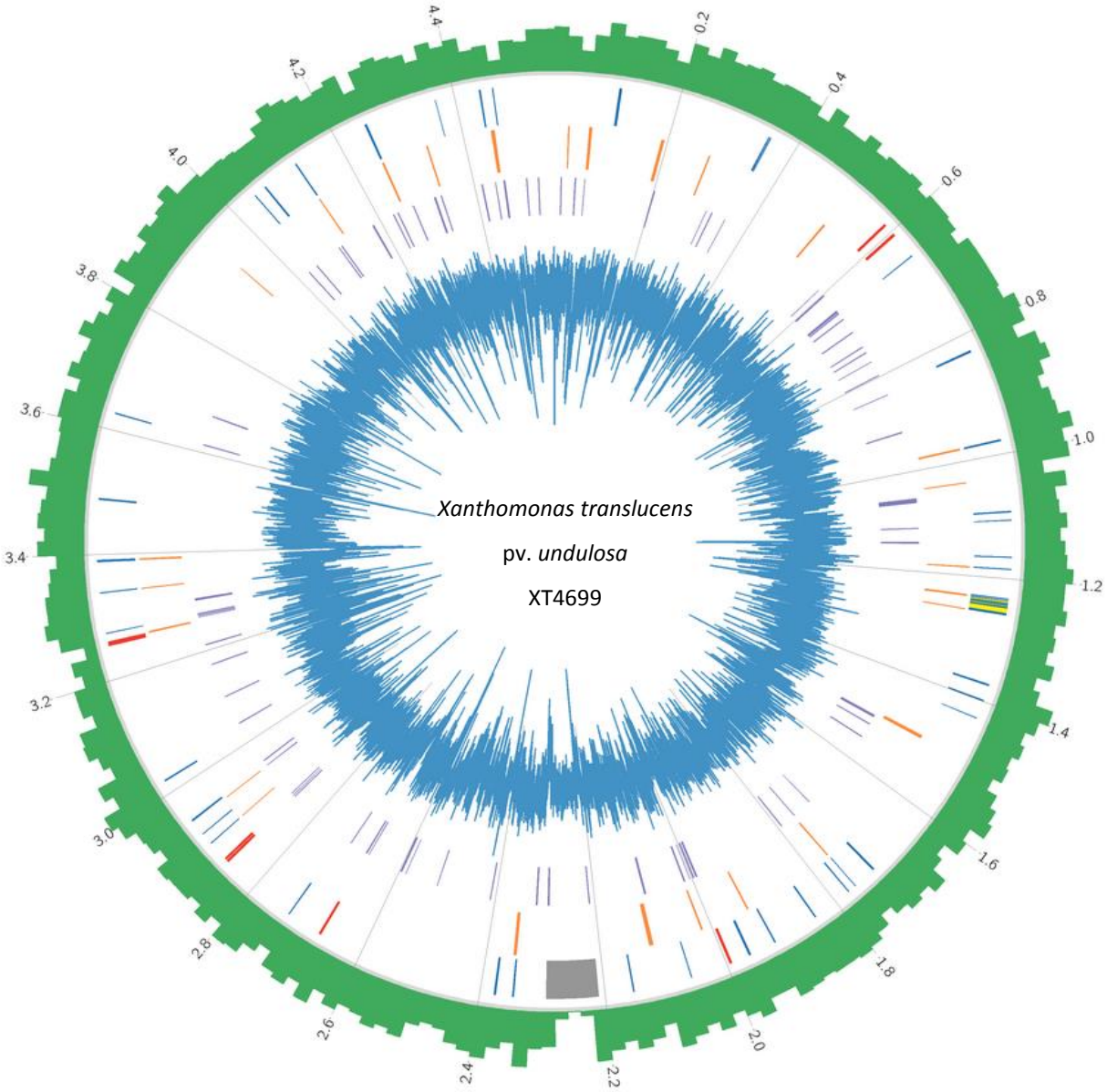


Figure 3-1. Genome features of *Xanthomonas translucens* pv. *undulosa* strain XT4699. Outer histogram (green) indicates gene density while the inter histogram (blue) shows the GC content. From the outside to inside, there are three layers. In the layer 1, the NRPS genes are indicated by grey, the Hrp gene cluster is marked by yellow, the genes with PIPs are marked with blue and the TAL effector genes are dictated by red color; In the layer 2, all non-TAL type III effector genes are marked by orange color; in the layer 3, the IS elements are shown as purple.

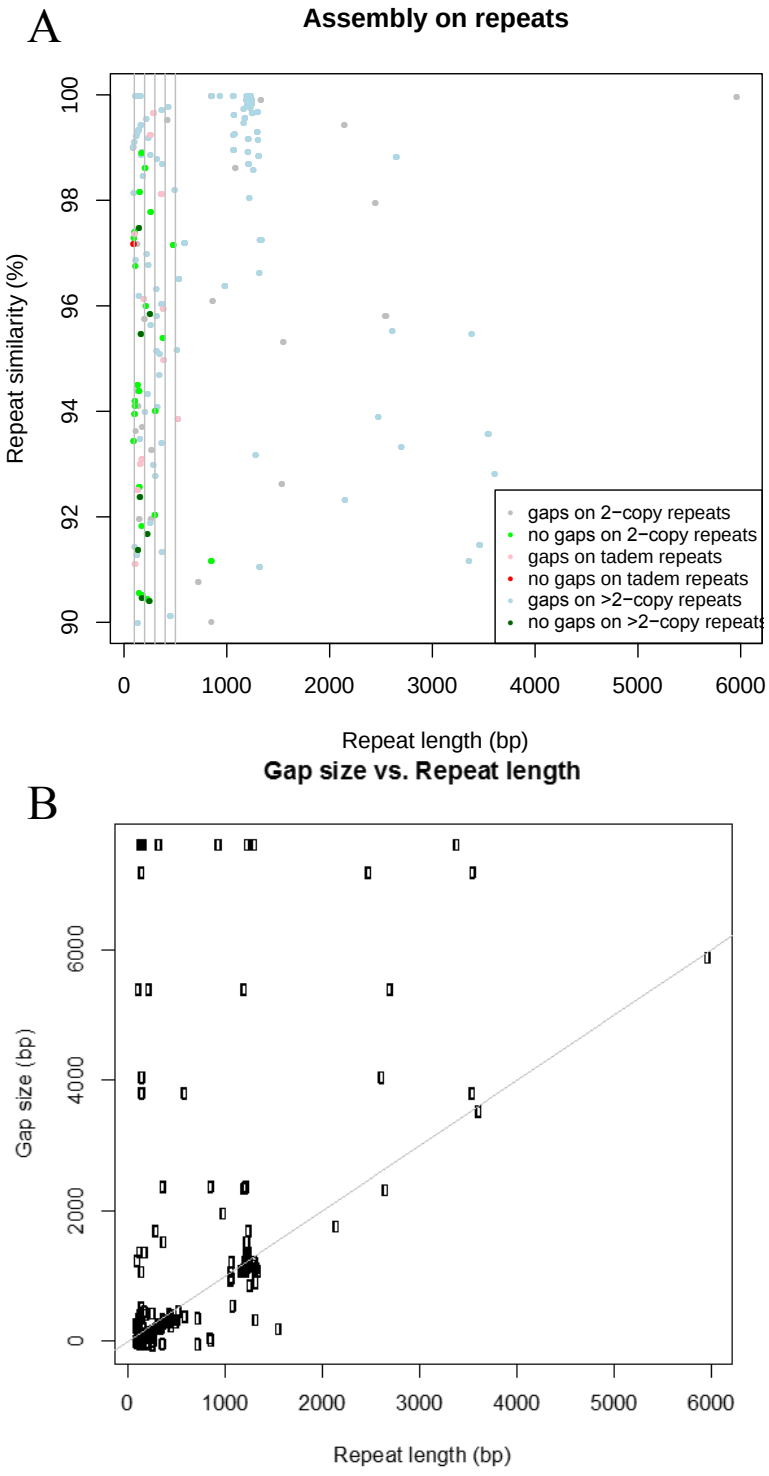


Figure 3-2. The relationship between genome repeats and assembled gaps using XT4699 Illumina data. A) The repeat features (length, similarity between repeated copies, and copy number) of assembled repeats and unassembled repeats. B) The relationship between assembled gaps and repeat lengths.

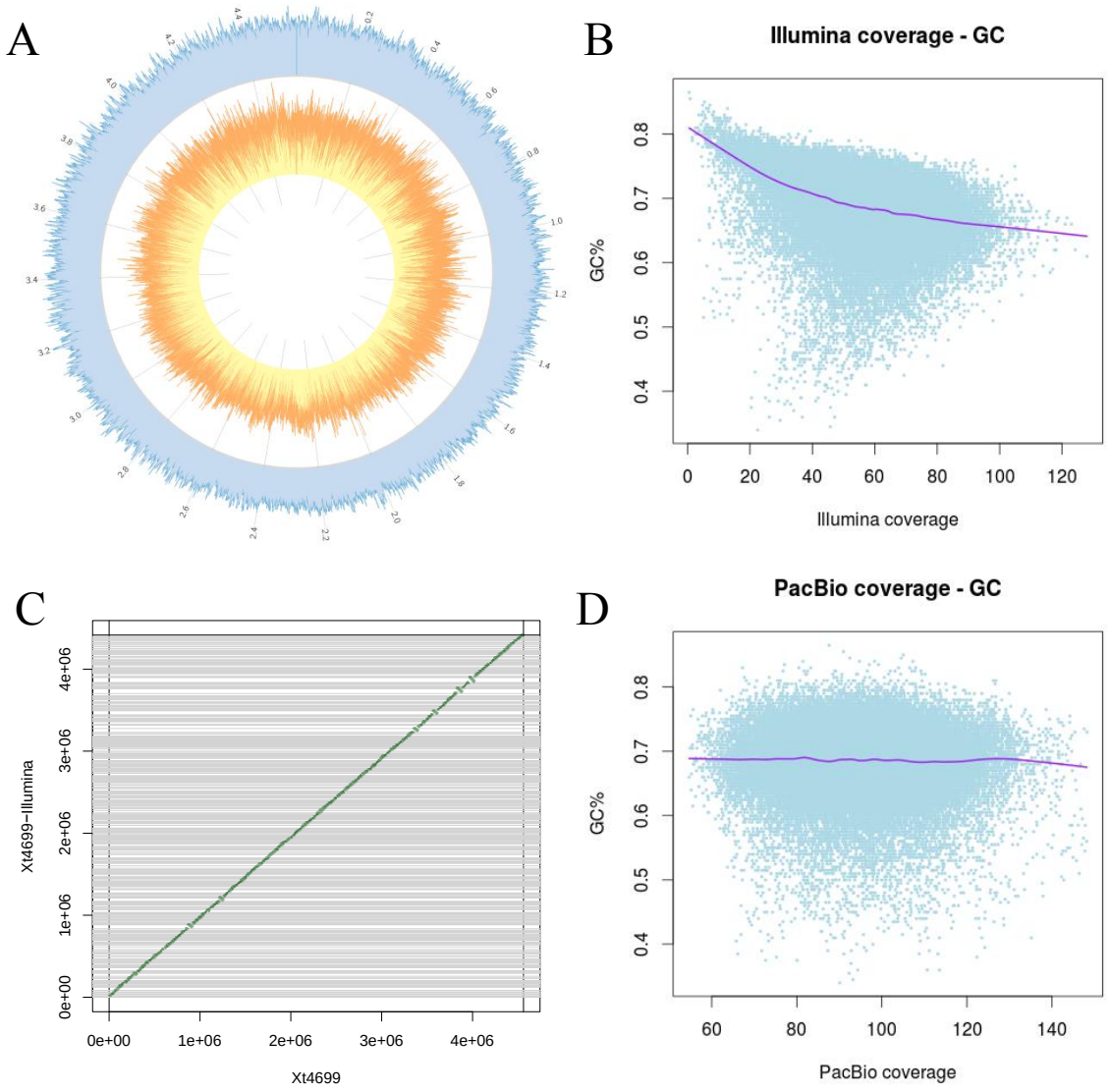


Figure 3-3. Data coverage of XT4699 PacBio and Illumina sequences. A) Coverage of PacBio sequences (light blue) and Illumina sequences (orange) across the genome. B) The relationship between Illumina sequencing coverage and GC% of non-overlapping 200 bp windows. C) Visualization of the alignment of Illumina assembled contigs on the XT4699 reference genome. D) The relationship between PacBio sequencing coverage and GC% of non-overlapping 200 bp windows.

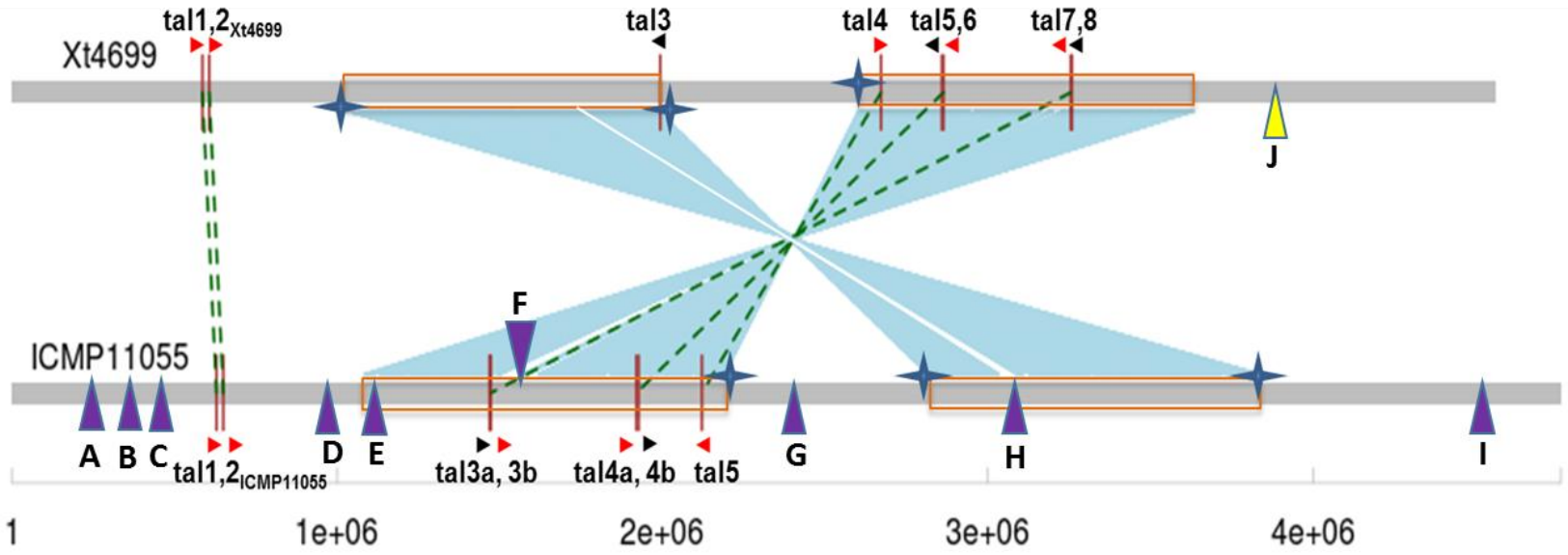

tal1,2Xt4699
tal3
tal4
tal5,6
tal7,8
Xt4699
J
F
ICMP11055
A
B
C
tal1,2ICMP11055
D
E
tal3a, 3b
tal4a, 4b
tal5
G
H
I
1
1e+06
2e+06
3e+06
4e+06

Figure 3-4. Genomic comparison between XT4699 an ICMP11055. Two major inverted and rearranged regions are marked with orange rectangles, which are associated together with light blue. The terminal ends of the rearranged regions with transposase elements are marked with dark blue stars. The TAL effector genes in each chromosome are marked with brown vertical bars. The related TAL effector genes between two strains are linked with green dash lines and marked with red triangles underneath or above the bars. Unrelated TAL genes are marked with black triangles. Purple filled triangles are used for indicating extra genomic regions of ICMP11055. The yellow triangle (J) indicates strain specific region of XT4699. A, 10 kb genomic region, encoding putative VGR related proteins for Type VI secretion; B. 14 kb genomic region, with gene cluster, *TadA*, *TadB* and *TadC*, comprising Type II/IV secretion systems; C, 15kb region, encoding CRISPR gene cluster; D, 12 kb region, harboring phage related integrase gene; E, 11 kb region, having CDS for cyclolysin secretion ATP-binding protein, hemolysin secretion protein D and alkaline phosphatase; F, putative VGR related proteins for Type VI secretion and phage related proteins; G, 15 kb region, encoding phage related proteins; H, 54 kb region, encoding phage related proteins; I, 15 kb region encoding putative VGR related proteins for Type VI secretion; J, 10 kb region with genes for type III restriction and modification systems.

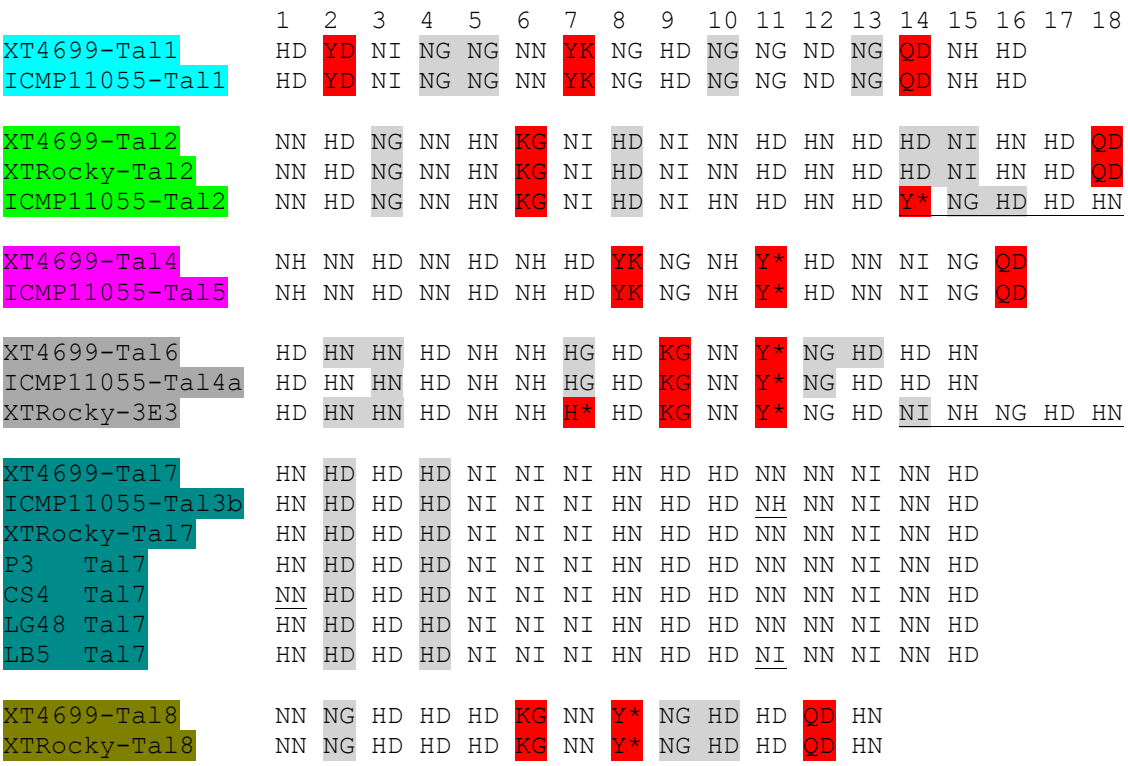
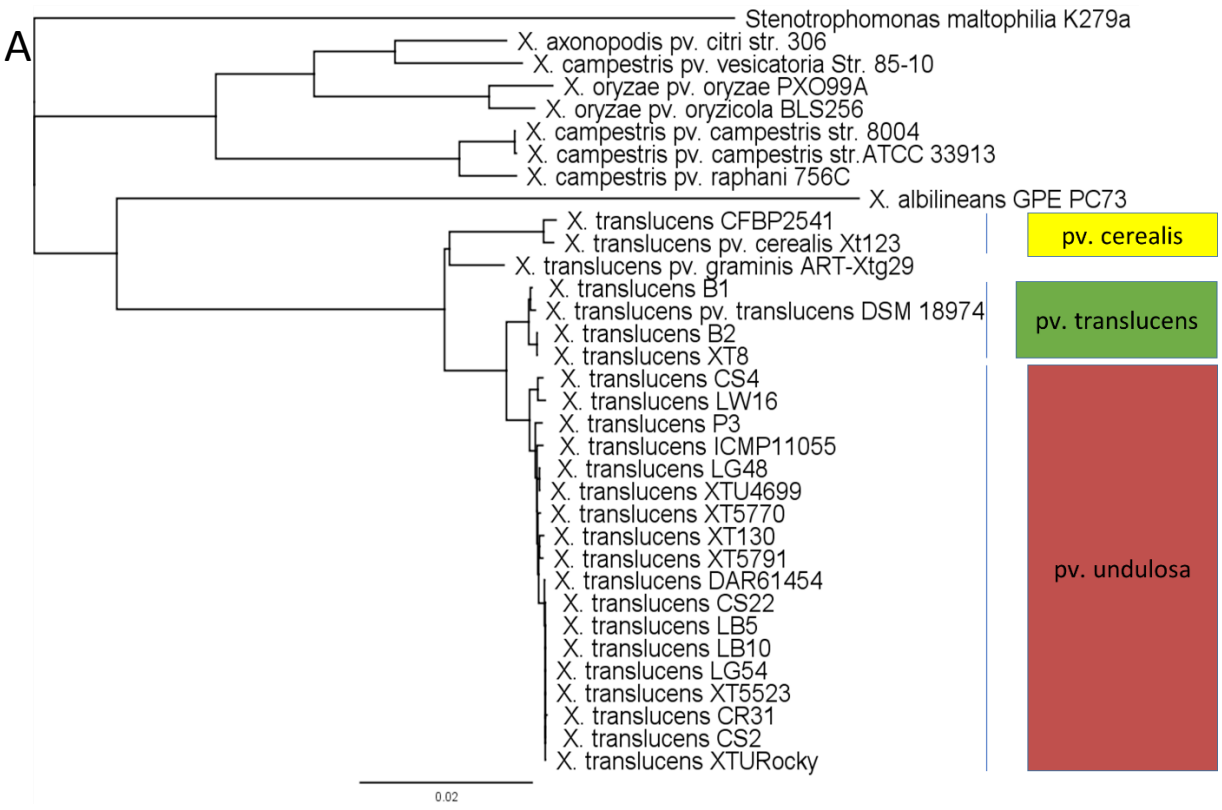


Figure 3-5. Highly similar TAL effectors present in different *X. translucens* pv. *undulosa* strains. XT4699 and ICMP11055 share five highly related TAL effectors except the Tal8 shown above. The names of highly similar TAL effectors are marked with the same color. The unusual RVDs are marked with red color. The repeat with 34 amino acids are marked by gray color. The missing 13th amino acid in the repeat is represented by *. The variation of RVDs among highly similar TAL effectors is underlined.


A
Stenotrophomonas maltophilia K279a
X. axonopodis pv. citri str. 306
X. campestris pv. vesicatoria Str. 85-10
X. oryzae pv. oryzae PXO99A
X. oryzae pv. oryzicola BLS256
X. campestris pv. campestris str. 8004
X. campestris pv. campestris str.ATCC 33913
X. campestris pv. raphani 756C
X. albilineans GPE PC73
X. translucens CFBP2541
X. translucens pv. cerealis Xt123
X. translucens pv. graminis ART-Xtg29
X. translucens B1
X. translucens pv. translucens DSM 18974
X. translucens B2
X. translucens XT8
X. translucens CS4
X. translucens LW16
X. translucens P3
X. translucens ICMP11055
X. translucens LG48
X. translucens XTU4699
X. translucens XT5770
X. translucens XT130
X. translucens XT5791
X. translucens DAR61454
X. translucens CS22
X. translucens LB5
X. translucens LB10
X. translucens LG54
X. translucens XT5523
X. translucens CR31
X. translucens CS2
X. translucens XTURocky
0.02
pv. cerealis
pv. translucens
pv. undulosa

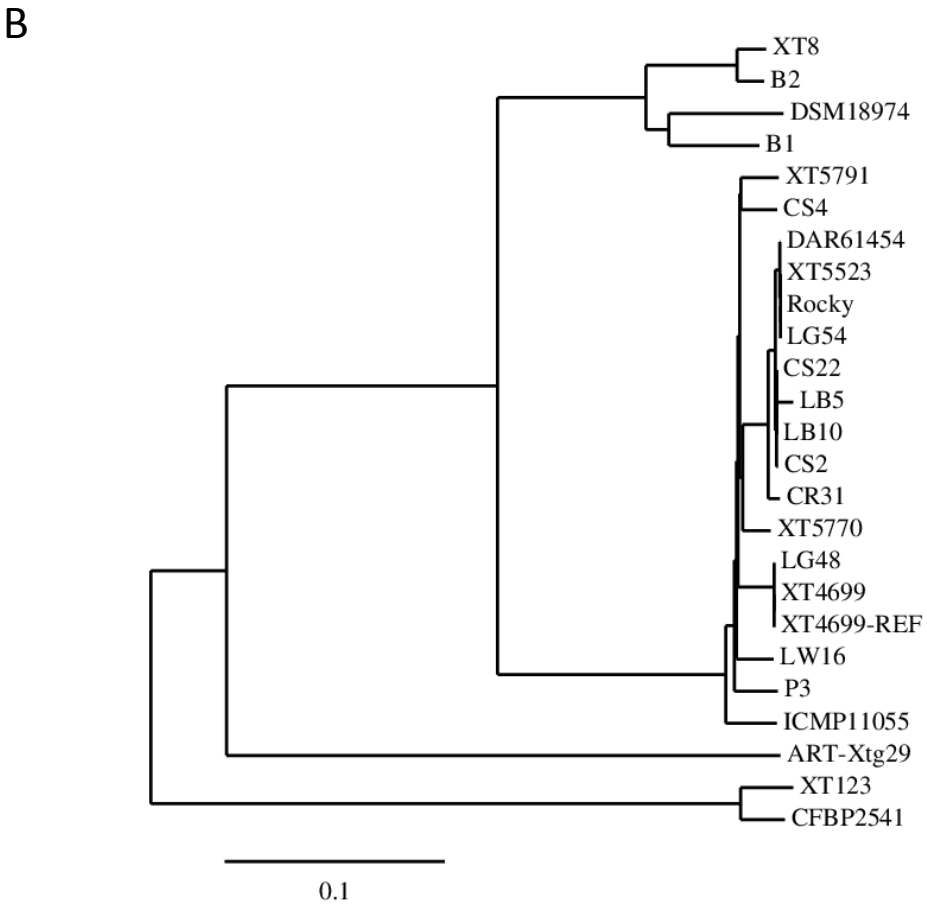

B
XT8
B2
DSM18974
B1
XT5791
CS4
DAR61454
XT5523
Rocky
LG54
CS22
LB5
LB10
CS2
CR31
XT5770
LG48
XT4699
XT4699-REF
LW16
P3
ICMP11055
ART-Xtg29
XT123
CFBP2541
0.1

Figure 3-6. Phylogenetic trees. A, Phylogenetic tree in the genus of *Xanthomonas* using MLSA. Nucleotide sequence of 4 concatenated housekeeping genes (*DnaK*, *gyrB*, *GroEL* and *RecA*) was extracted from the draft or complete genomes. The tree was constructed by applying all the concatenated sequences into Geneious software Version 6 with Tamura-Nei genetic distance model and Neighbor-joining method with *Stenotrophomonas maltophilia* K279a as outgroup. Bar indicates number of nucleotide substitutions per site. The dark red represents *undulosa* pathovars, the green color indicates the *translucens* pathovar and the yellow color represents the *cerealis* pathovars; B, Phylogenetic tree of *X. translucens* strains using whole-genome discovery of SNP. Bar indicates number of nucleotide substitutions per site. Detail is described in methods.

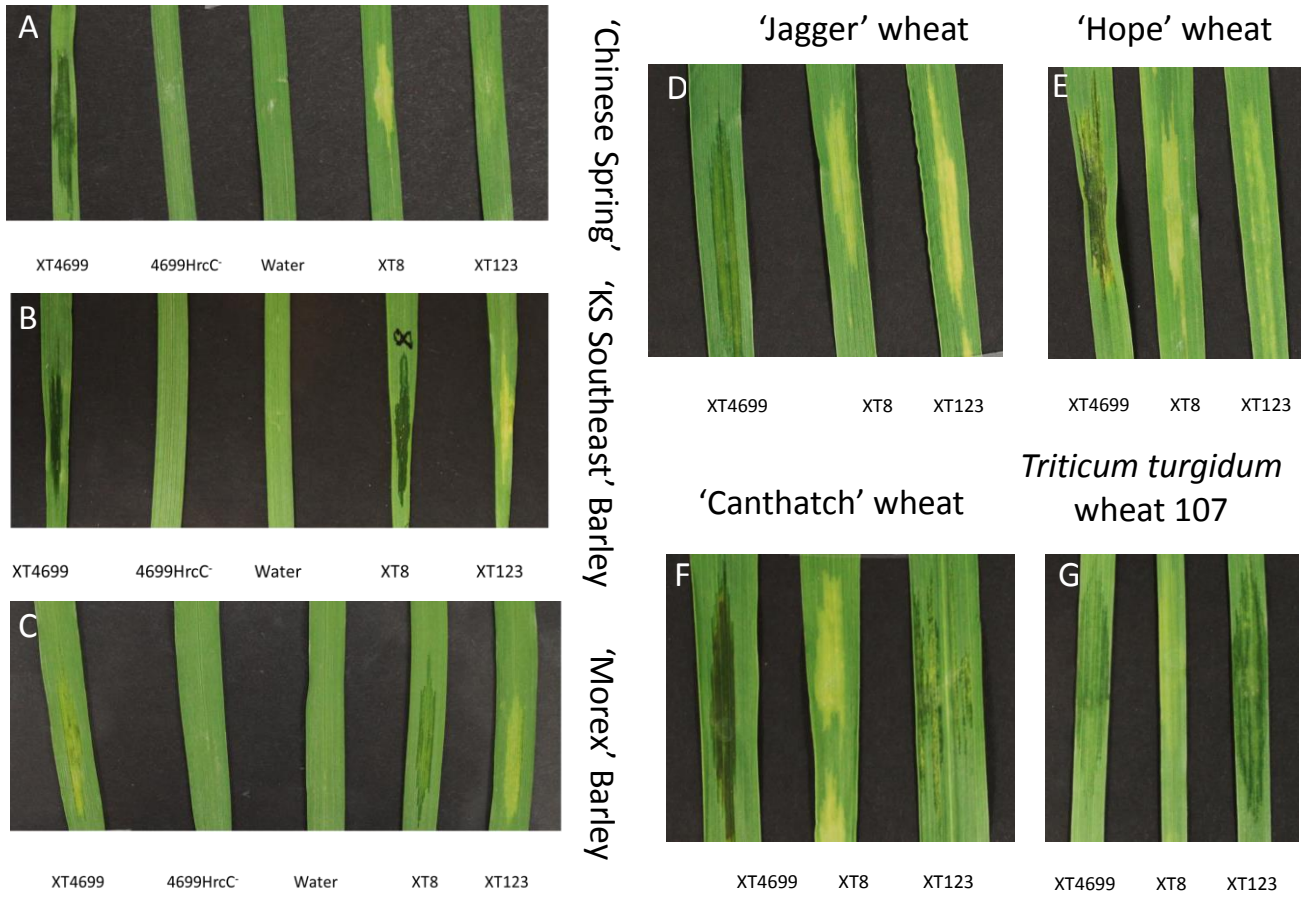


Figure 3-7. Different disease symptoms induced by pathovars of *X. translucens*. In A-G panels, different wheat and barley cultivars were inoculated. A, 'Chinese Spring' wheat; B, 'KS Southeast' barley; C, 'Morex' barley; D, 'Jagger' wheat; E, 'Hope' wheat; F, 'Canthatch' wheat; G, *Triticum turgidum* wheat #107. In A, B, C, plants are 3 weeks old and second leaves were inoculated. In D,E, F,G, leaves with similar age from 50-day-old plants were inoculated.

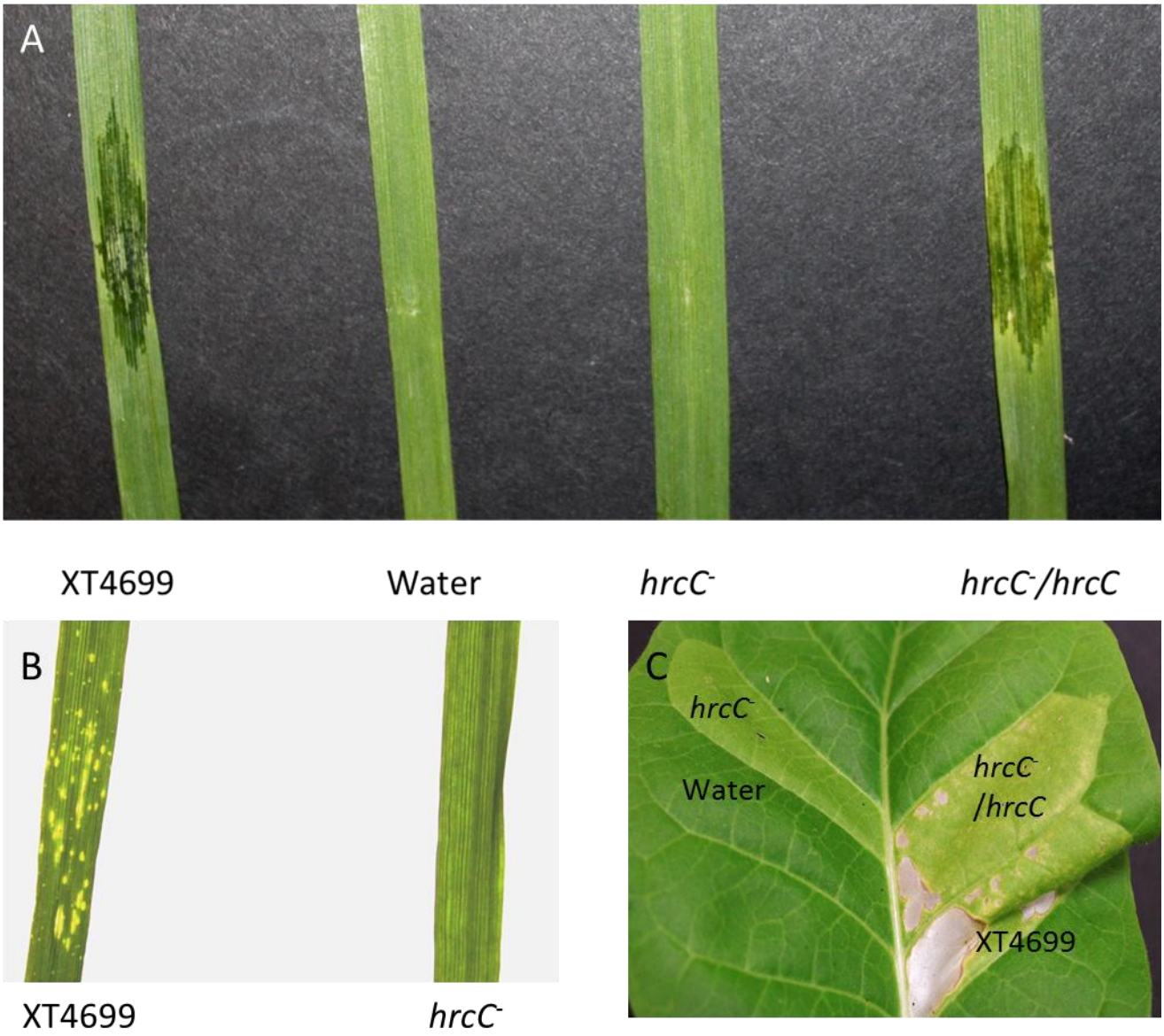


Figure 3-8. **Phenotypic differences between XT4699 and the *hrcC*- mutant**. A, third leaves of one-month-old Chinese Spring wheat plants were inoculated with XT4699, water, *hrcC*- and complementation strain of *hrcC*-, respectively, by needleless syringe. The picture was taken at 4DPI; B, WT and *hrcC*- (OD600=0.2) strains were coated with 0.02% Silwet L-77 and applied for the dip inoculation assays. Second leaves of 14-day-old Chinese Spring wheat plants were used for experiment. Picture was taken 8DPI. C, the same inoculum, as in panel A, was infiltrated into leaves of two-month-old KY-14 tobacco plants. It was photographed at 8DPI;

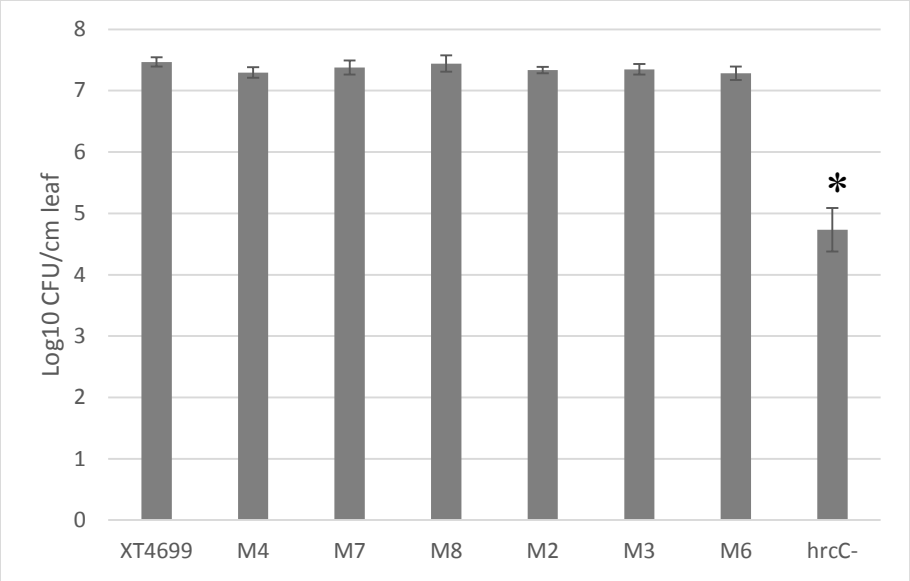


Figure 3-9. Bacterial population assay. Equal amount of $4x10^4$ dilution of bacterial inoculum of different *tal* mutants, *hrcC* mutant and WT were inoculated on 3-week-old Chinese Spring wheat by needleless syringe. At 6DPI, three inoculated leaves (3 cm) were pooled and ground together in each treatment. Samples were serially diluted and 100 µl of diluted samples were added to TSA plates. Plates were incubated at 28^0C for colony formation. Data represents the mean of Log10 CFU/cm leaf ± standard deviation. The * indicates the significant difference of bacterial population of *hrcC*$^-$ mutant compared to other *tal* mutants or WT at 6DPI under p-value <0.001 in the ANOVA statistics analysis.

Effector	DAR 61454	CR31	LB10	LB5	CS2	CS22	LG54	XT 5523	XTU-Rocky	LG48	XT 4699	ICMP 11055	XT 5791	XT 130	P3	XT 5770	LW16	CS4	DSM XT8	DSM 18974	DSM B1	DSM B2	XT 123	CFBP 2541	ART-Xtg29
hpaH(HopP1)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopB	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopC2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopF	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
XopG	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopK	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F
XopN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopQ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopV	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopX	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++
XopY	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopZ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopAA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	ND
XopAD	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
XopAM	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AvrBs1	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	ND	ND	ND	+	+	ND
AvrBs2	++	++	++	++	++	++	++	++	++	++	++	++	++	++	F,+	++	++	++	++	++	++	++	++	++	+
XopE1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	ND	ND	ND	ND	+	+	+
XopE2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	+	+	+	+	+	+
XopE3	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	+	+	+	ND	ND	ND
XopE4	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+
XopE5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	ND	ND	ND	ND	+	+	ND
XopI	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+
XopJ1	+	+	+	+	+	+	+	+	+	ND	ND	ND	ND	ND	ND	+	ND	ND	+	+	+	+	ND	ND	ND
XopJ2(AvrRxv)	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+
XopL	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++++	++++	+
XopP	++	++	F,+	F,+	F,+	F,+	++	++	++	++	++	++	++	++	++	++	++	++	+++	+++	++	+++	+++	+++	++
XopR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F	F	+
XopAF (AvrXv3)	++	++	++	++	++	++	++	++	++	++	++	+	+	+	+	+	+	+	++	++	++	++	+	+	+
XopAH(AvrB3)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	ND	ND	ND
XopAK(HopK1)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	ND	ND	+
XopAP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	F	F	F	F	+	+	+
TALE									≥7		8	7					≥4						2	2	ND

Figure 3-10. Comparison of Type III effector repertories among four pathovars of *X. translucens*. 15 effectors (from hapH to XopAM), marked with gray color, are predicted to be conserved among *undulosa*, *translucens* and *cerealis* pathovars while the rest of 17 effectors (from AvrBs1 to XopAP), marked with blue color, are shown to be variable among strains. Single copy of effector is indicated by '+' and multiple copies are shown as multiple '+'. If the effecotr is not detected in the genome, it is shown as 'ND'. The frameshift mutation of effector is marked with 'F' and shown out with red rectangle frame. The number of TAL effectors in each strain is predicted by genome sequence if full genome or high quality draft genome available, or by Southern blot analysis. The blank shows it is not known for the number of TAL effector genes.

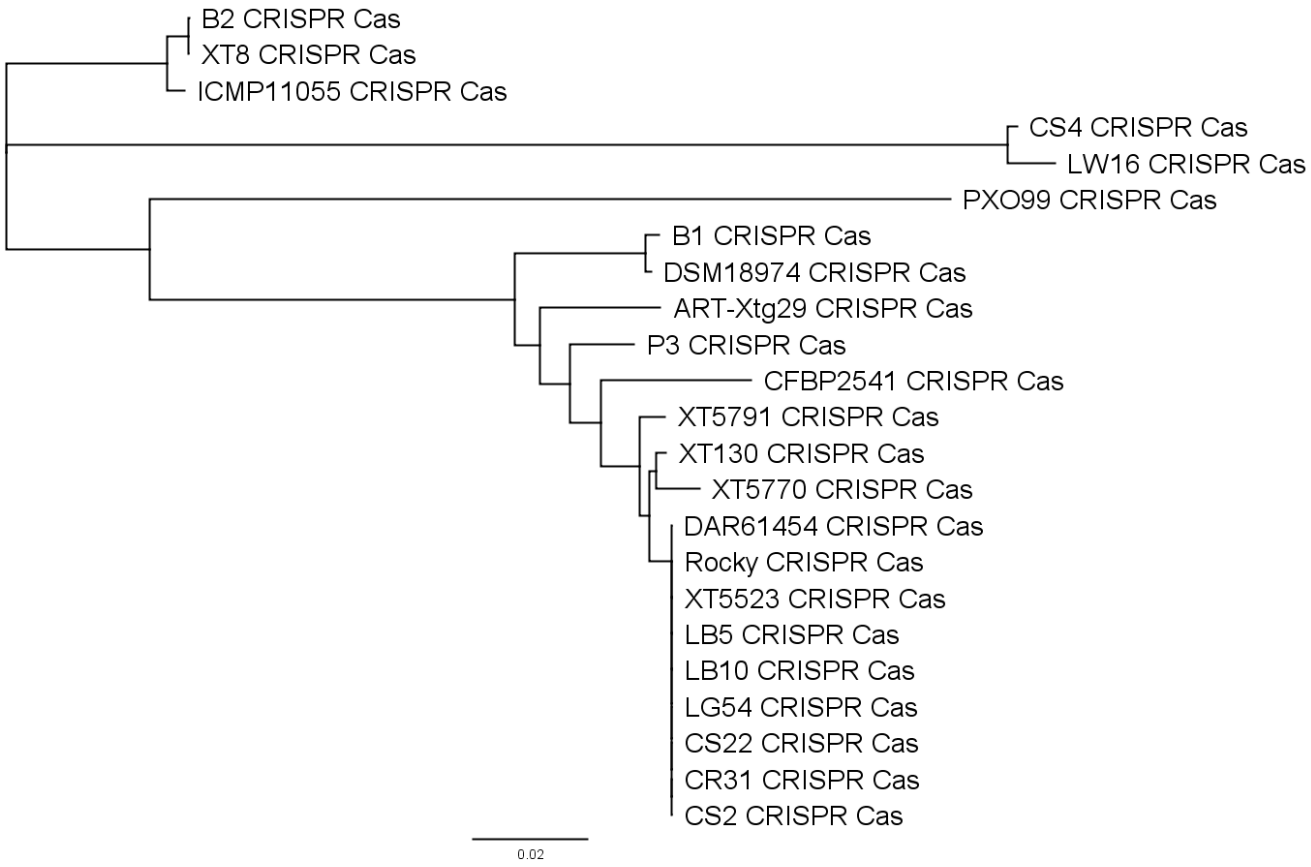


Figure 3-11. Phylogenetic tree based on CRISPR *Cas* loci of *X. translucens* strains. The sequence of CRISPR *Cas* genes is annotated in RAST website http://rast.nmpdr.org.The phylogenetic tree was developed by applying the *Cas* gene sequences into Geneious software Version 6 with Tamura-Nei genetic distance model and Neighbor-joining method, with PXO99 CRISPR *Cas* genes as outgroup. Bar indicates number of nucleotide substitutions per site.

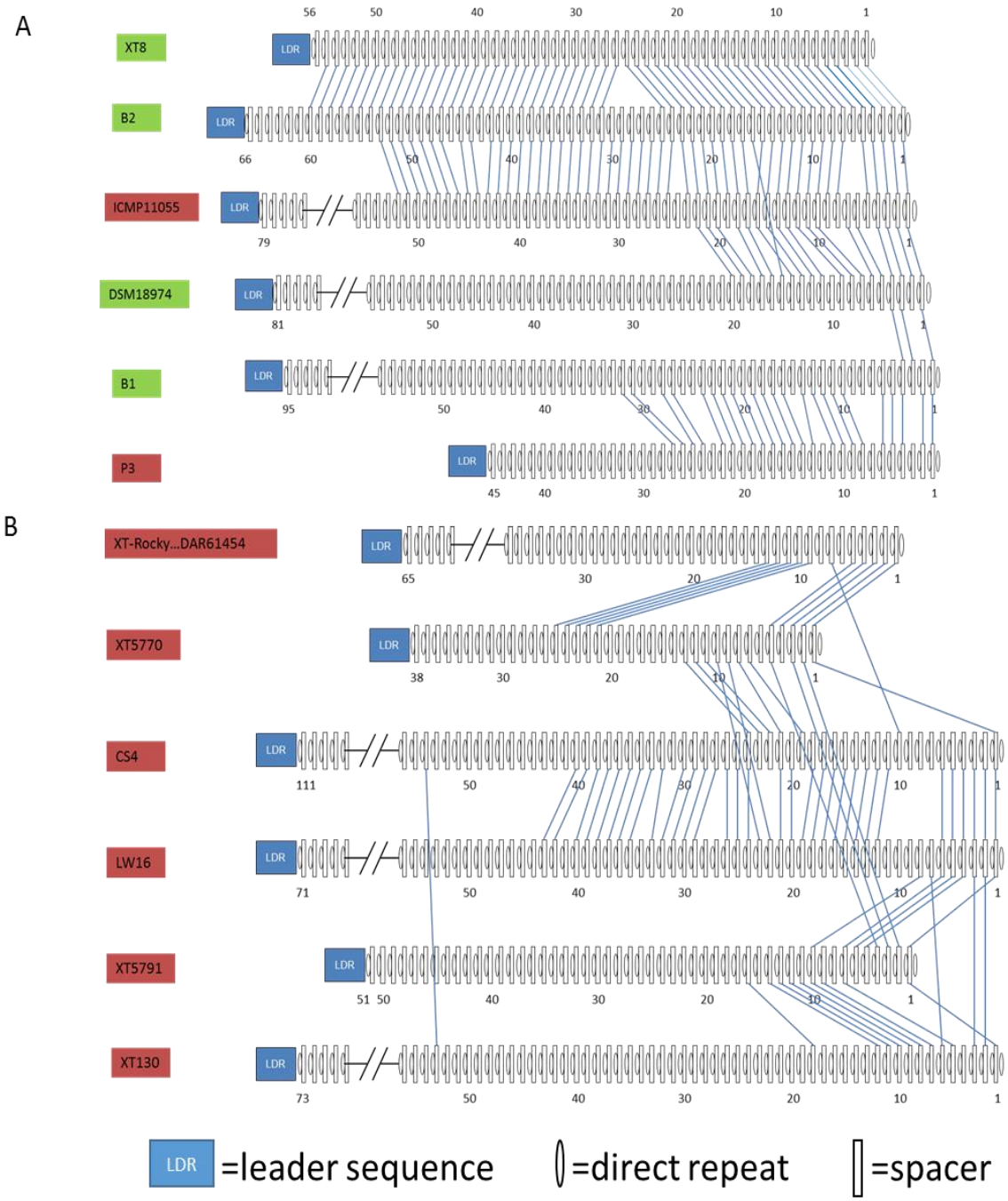


Figure 3-12. **Spacer comparison among different *X. translucens* strains.** The leader sequence, spacer and direct repeat sequence are analyzed at http://crispr.u-psud.fr/Server/. The identical spacer elements are linked together by blue lines. The strain names marked with dark red suggest *undulosa* pathovar and the green color indicates the *translucens* pathovar. There is no identical spacer between panel A and B. The XT-Rocky…DAR61454 represents 9 strains of LG54, XT5523, XT-Rocky, CR31, CS2, CS22, LB5, LB10 and DAR61454, which share the same *Cas* genes, leader sequence, direct repeat and spacer arrays.

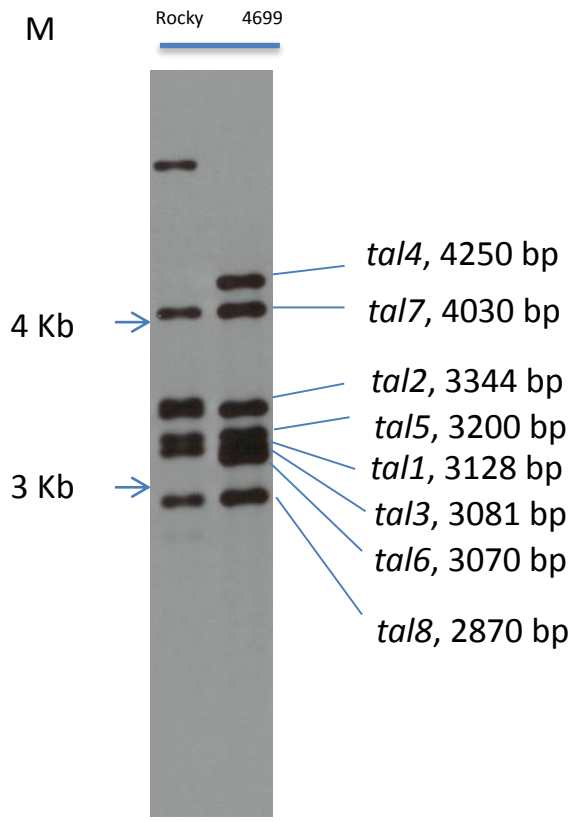


Figure 3-13. **Southern blot analysis of two *X. translucens* pv. *undulosa* strains from Kansas.** Genomic DNA digested by *Bam*HI and hybridized by *sph*I fragment of TAL effector gene in AXO1947. The size of blotted bands is consistent with assembly genome of XT4699.

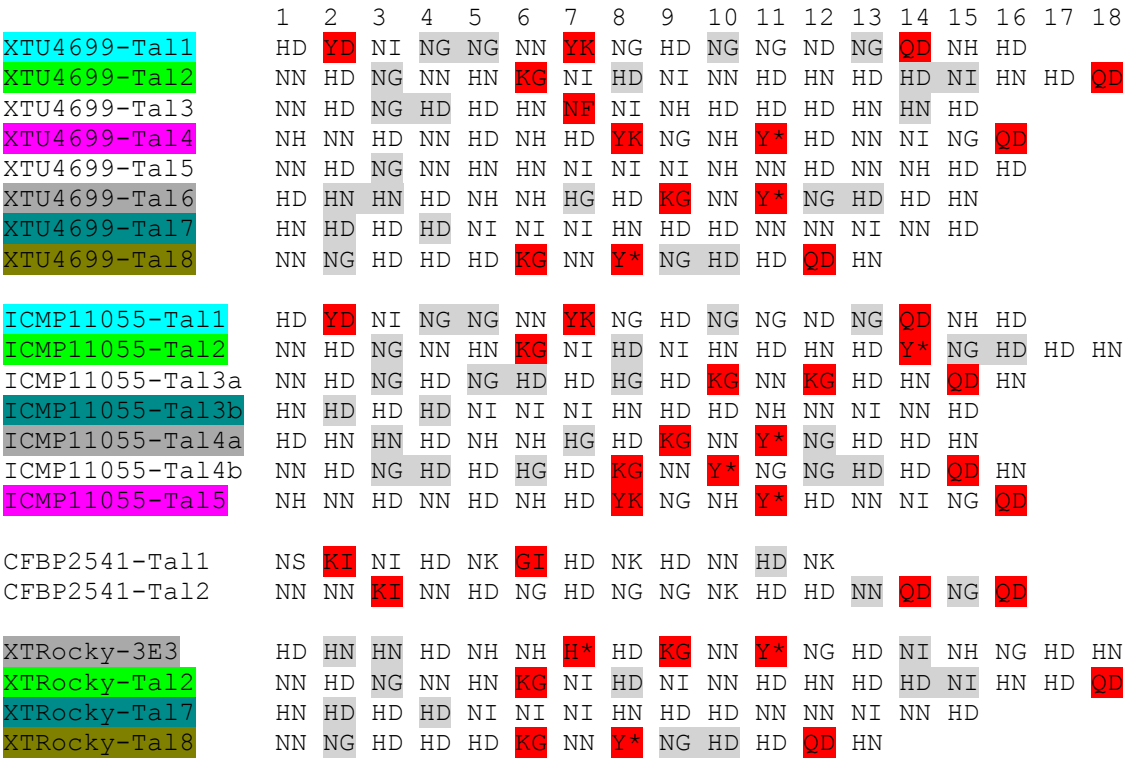


Figure 3-14. Alignment of RVDs of TAL effectors among XT4699, ICMP11055, CFBP2541 and XT-Rocky strains. The unusual RVDs are marked with red color. The repeat with 34 amino acids are marked with gray color. The missing 13th amino acid in the repeat is represented by *. Closely related TAL effector are marked with same color. Highly similar TAL effector genes from different strains are marked by the same color on the names.

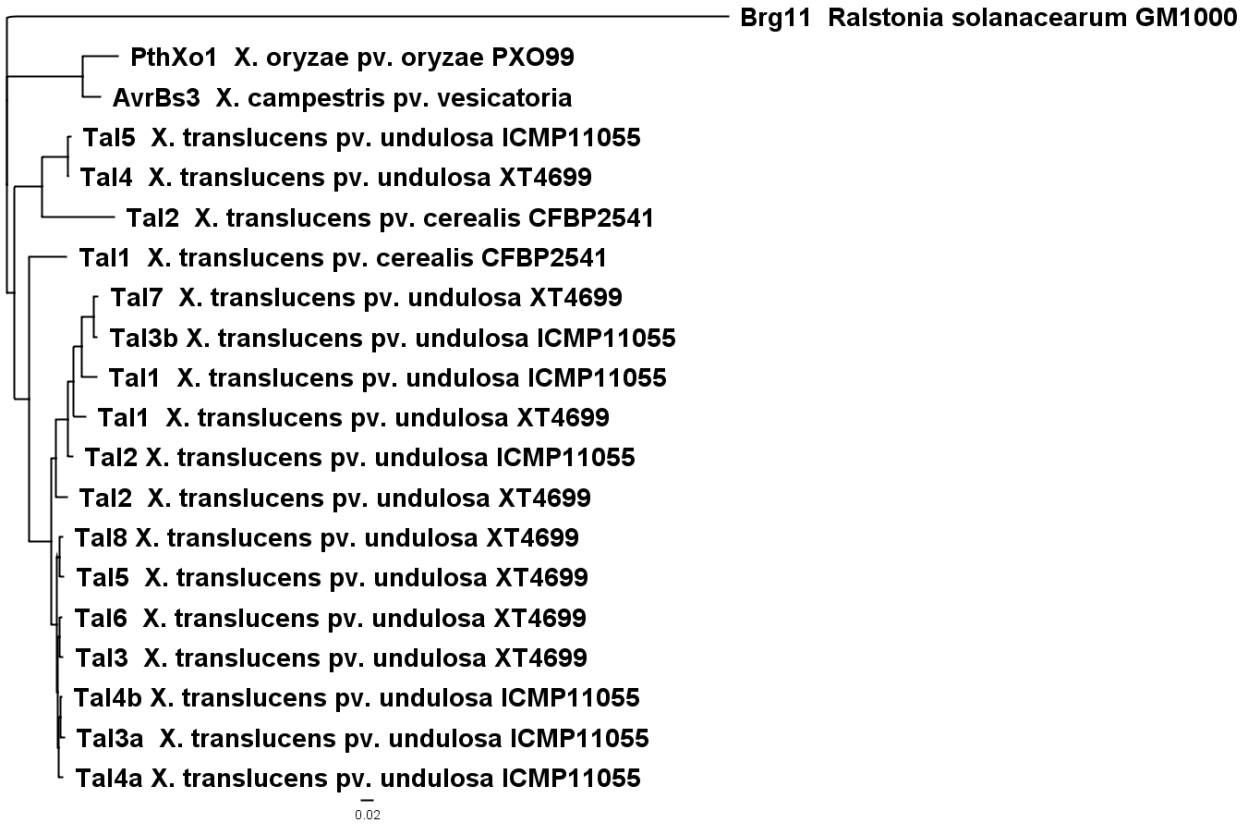


Figure 3-15. **Phylogenetic tree of the concatenated N-terminal and C-terminal amino acid sequences of TAL effectors.** The tree was constructed with Tamura-Nei genetic distance model and Neighbor-joining method, with Brg11 TAL effector from *Ralstonia Solanacearum* GMI1000 as outgroup, according to Geneious software. The bar with 0.02 indicates number of substitutions per site.

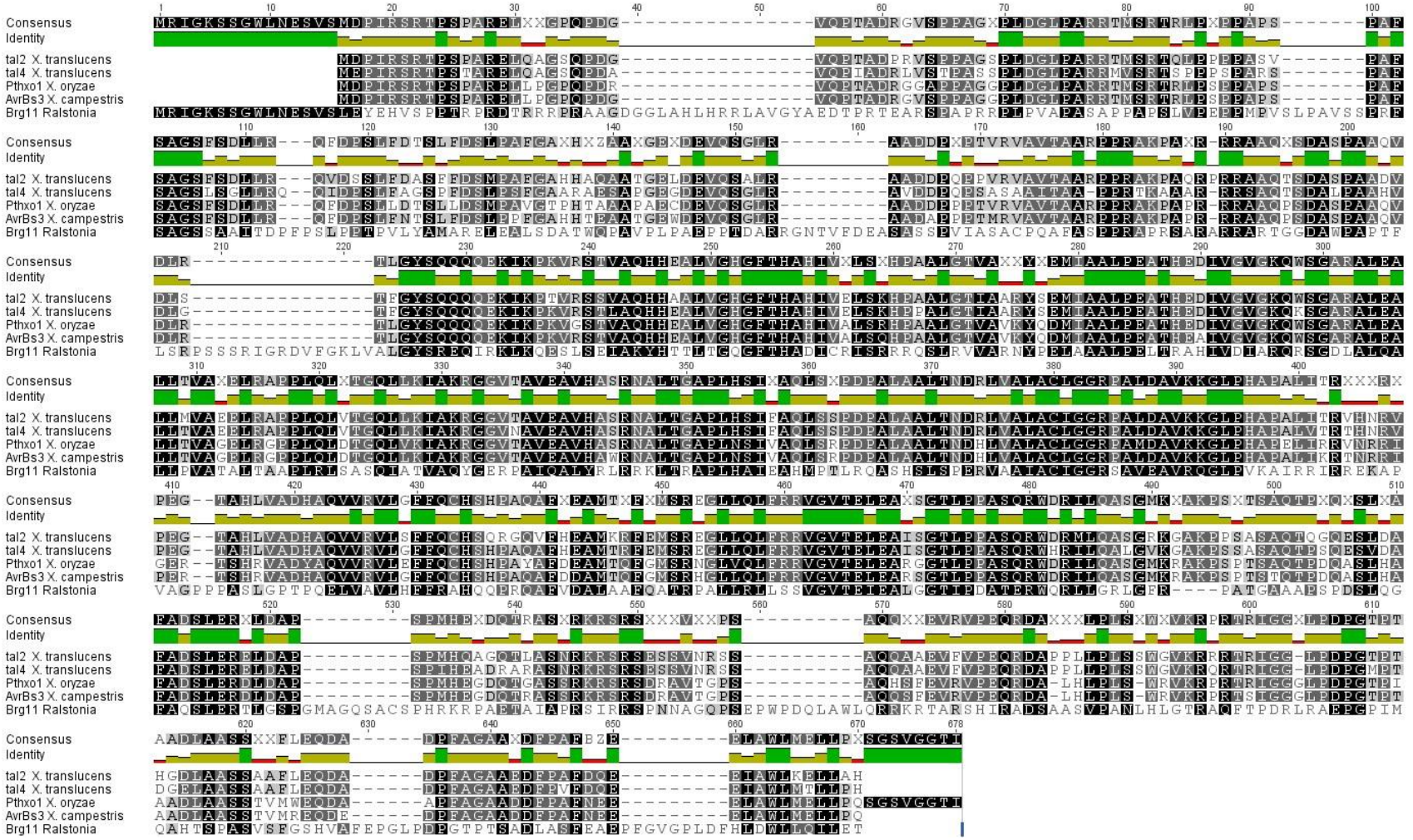

Consensus
Identity
tal2 X. translucens
tal4 X. translucens
Pthxo1 X. oryzae
AvrBs3 X. campestris
Brg11 Ralstonia

Figure 3-16. Alignment of concatenated N-terminal and C-terminal amino acid sequences of TAL effectors. The sequences of Tal2 and Tal4 of *X. translucens* XT4699, PthXo1 of *X. oryzae* PXO99, AvrBs3 of *X. campestris* 85-10 and Brg11 of *Ralstonia Solanacearum* GMI1000 are applied for alignment, which is generated by Geneious software.

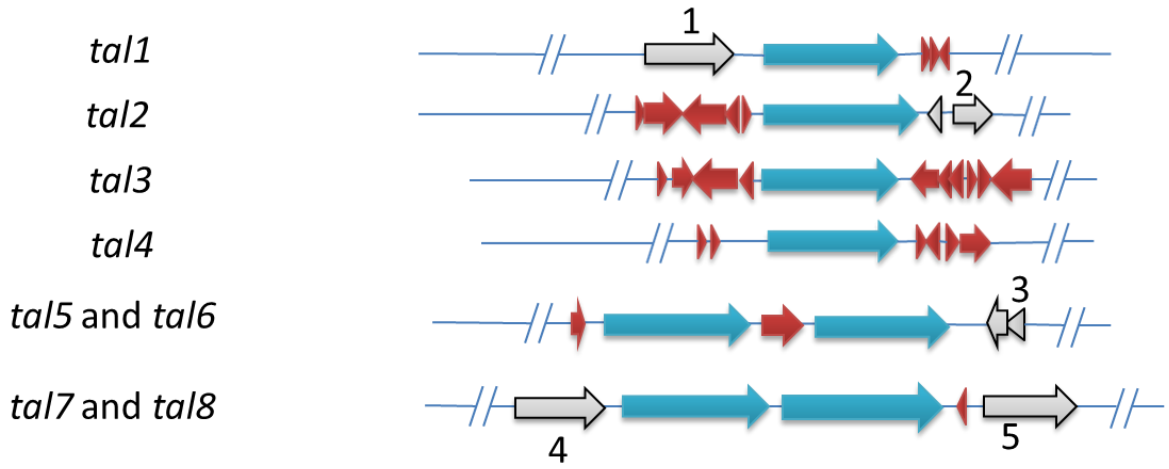


Figure 3-17. Most TAL effector genes are flanked by transposon elements. The blue arrows indicate the TAL effector genes and the dark red arrows show the transposase genes; the gray arrows indicate other genes. The genes are listed as following: 1, *RpoD*, encoding RNA polymerase subunit sigma-70; 2, Hypothetical genes; 3, Acetyltransferase and Hypothetical gene; 4, Type III effector gene *xopL*; 5, *vgrG*, encoding type IV secretion protein.

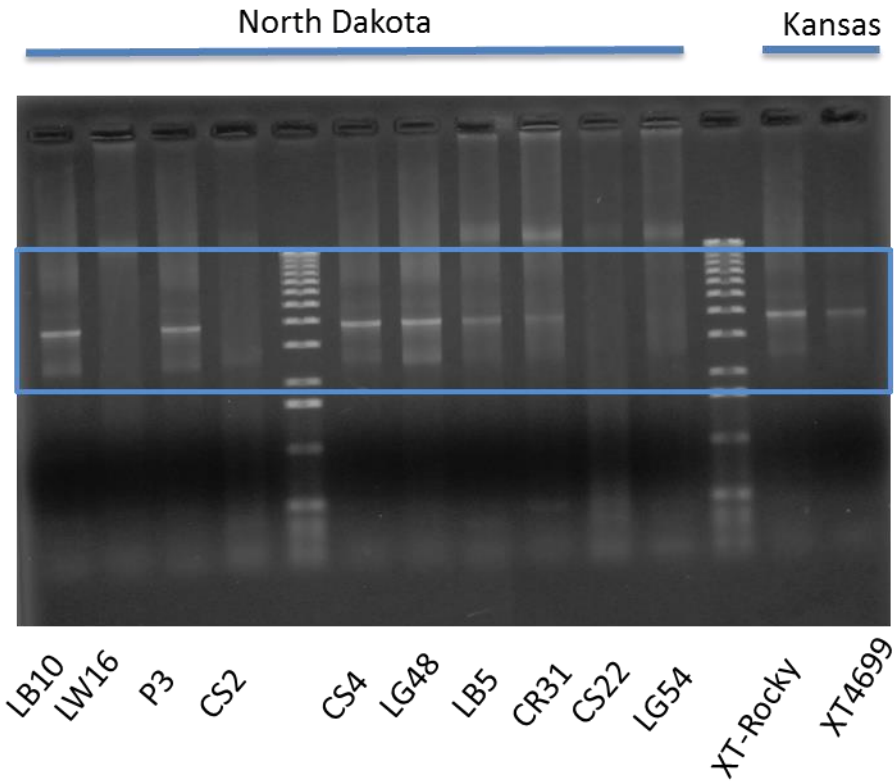
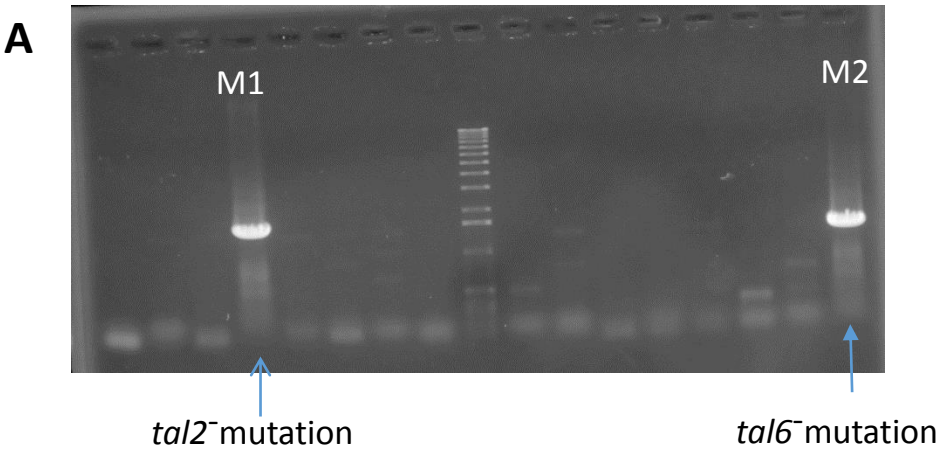


Figure 3-18. XT4699-*tal7* is a conserved TAL effector gene. Primer pair used for cloning *tal7* gene in XT4699 was applied in PCR with template DNA of other strains from North Dakota and Kansas. PCR bands with similar size as XT4699-*tal7* were shown in the rectangle frame.



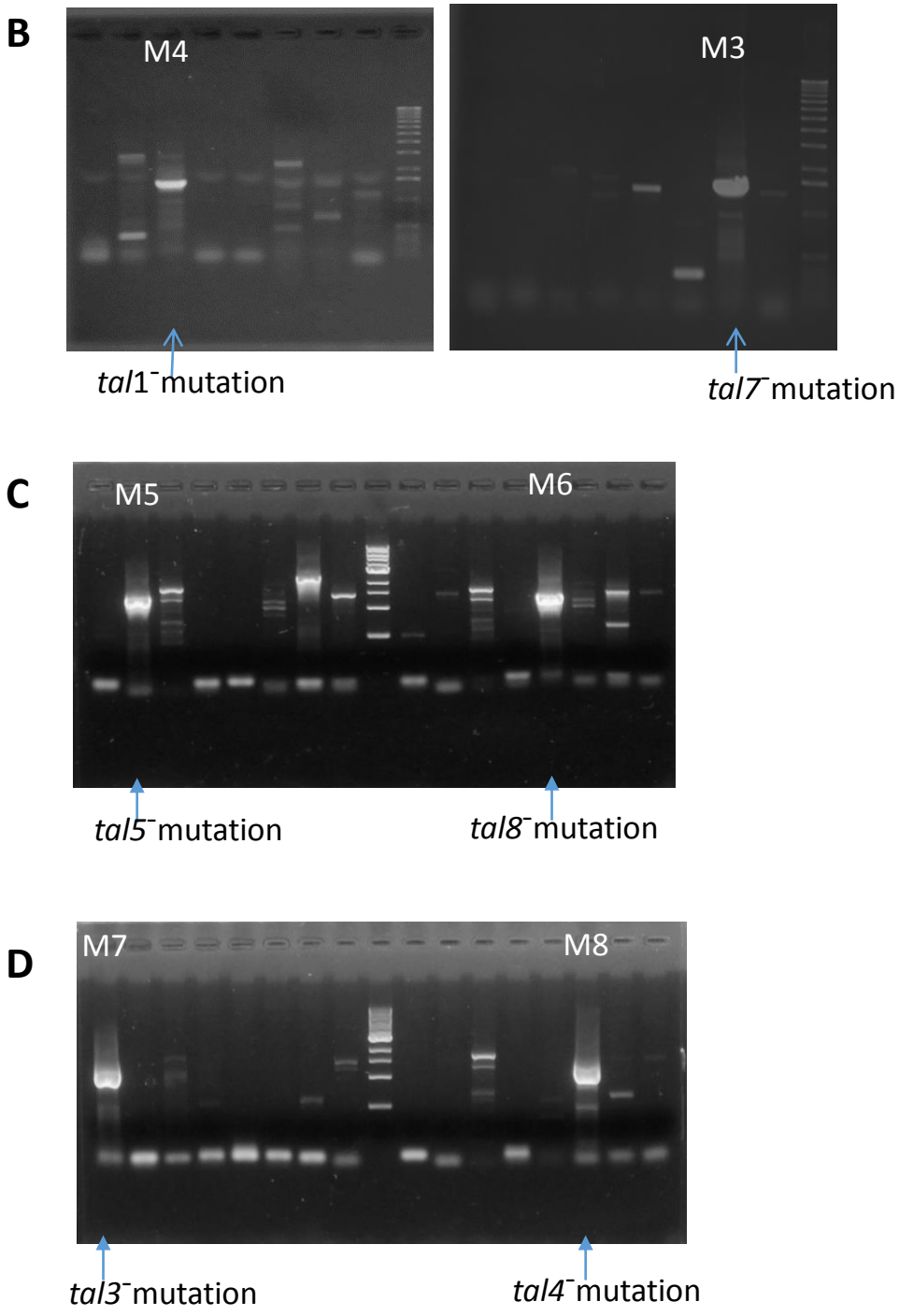


Figure 3-19. Validation of all eight TAL mutants in XT4699 strain by PCR. The sequence of specific primers for each TAL gene and primer from the vector are provided in Table 3-2. The desired PCR product size is around 1.5 kb. The corresponding name of TAL genes are shown for all 8 mutants. A, M1 (*tal2* mutant), M2 (*tal6* mutant); B, M3 (*tal7* mutant), M4 (*tal1* mutant); C, M5 (*tal5* mutant), M6 (*tal8* mutant); D, M7 (*tal3* mutant), M8 (*tal4* mutant).

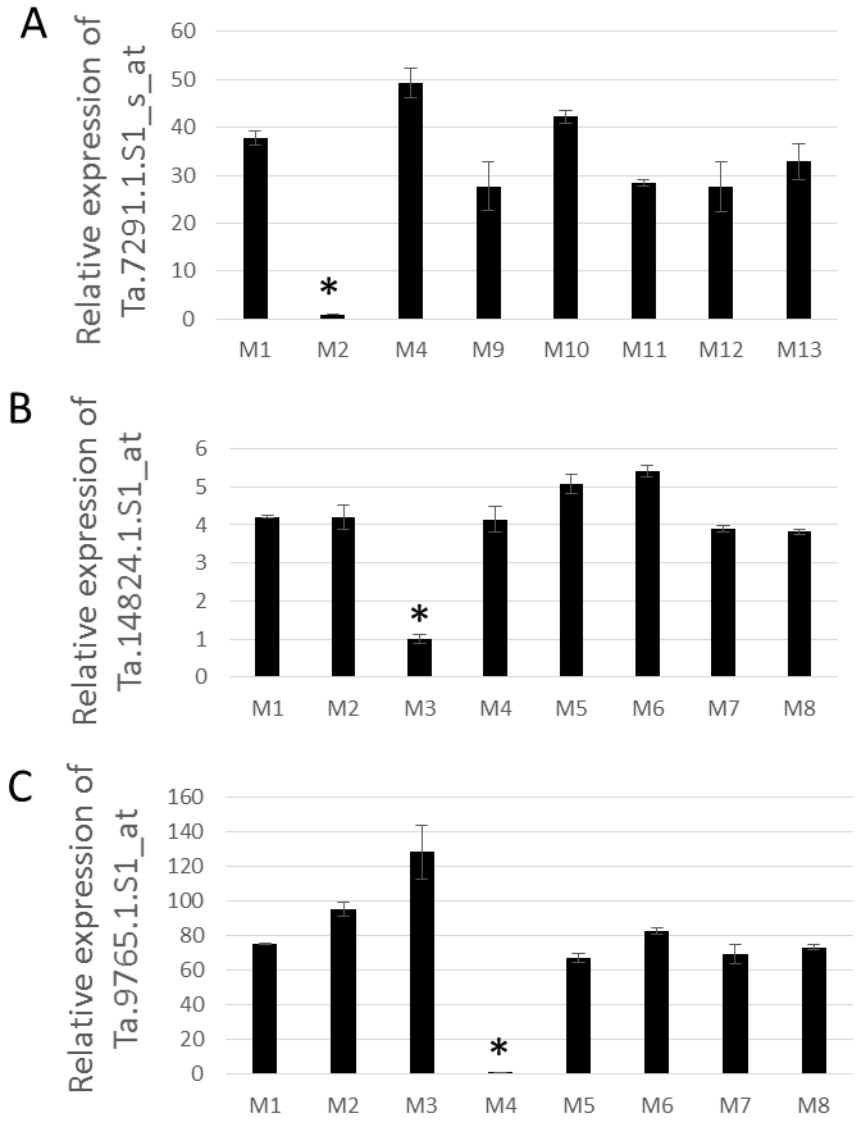


Figure 3-20. Induction of host genes is impaired by mutation of specific TAL genes.

M2 is lack of *tal6*, M3 is lack of *tal7* and M4 is the mutant of *tal1*. A, relative expression level of Ta.7291.1.S1_at, corresponding to succinate dehydrogenase subunit gene, was calculated compared to M2 treatment; B, relative expression level of Ta.14824.1.S1_at, corresponding to choline transporter related gene, was calculated compared to M3 treatment; C, relative expression level of Ta.9765.1.S1_at, corresponding to cell wall invertase gene, was calculated compared to M4 treatment. $2^{-\Delta\Delta Ct}$ method was applied in calculation. The * indicates significant difference compared to other TAL mutants treated samples with P-value<0.05 in the ANOVA statistics analysis.

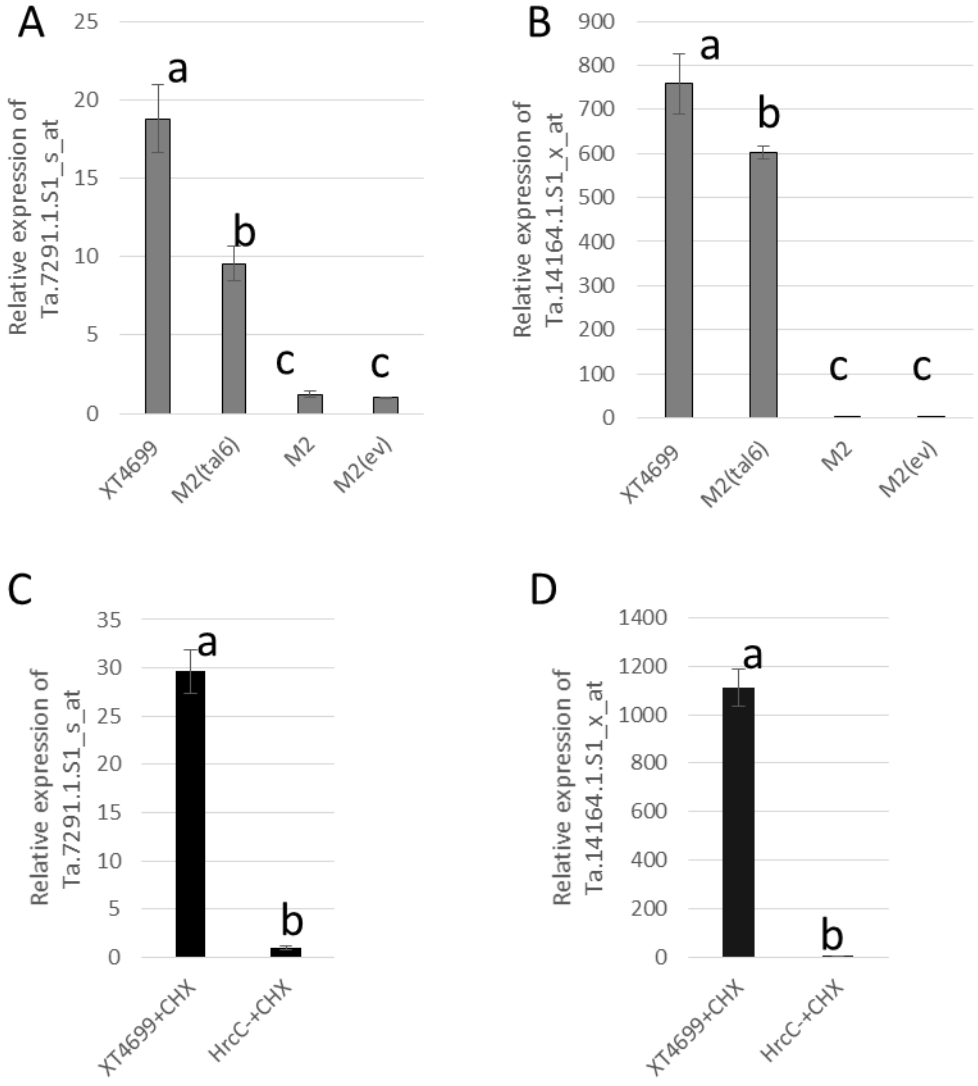


Figure 3-21. XT4699-Tal6 effector modulates gene expression in Chinese Spring wheat. First leaves of 10-day-old Chinese Spring wheat plants were inoculated with XT4699, M2, M2(empty vector) and M2(tal6), respectively. Inoculated leaves were collected 24hr post inoculation. Gene specific primers were applied for qRT-PCR in each panel. A and C, relative expression of Ta.7291.1.S1_s_at was calculated; B and D, relative expression level of Ta.14164.1.S1_x_at was calculated; A and B, the fold change relative to M2(ev) treatment was calculated with $2^{-\Delta\Delta Ct}$ method. C and D, bacterial strains mixed with 100 µm cycloheximide were applied for inoculation. The fold change relative to *hrcC*$^{-}$+CHX treatment was calculated with $2^{-\Delta\Delta Ct}$ method. The different lowercase letters indicate significant difference with P-value<0.05 in the ANOVA statistics analysis. Two biological replicates were applied and the experiments were repeated twice with consistent results. CHX=cycloheximide, ev= empty vector.

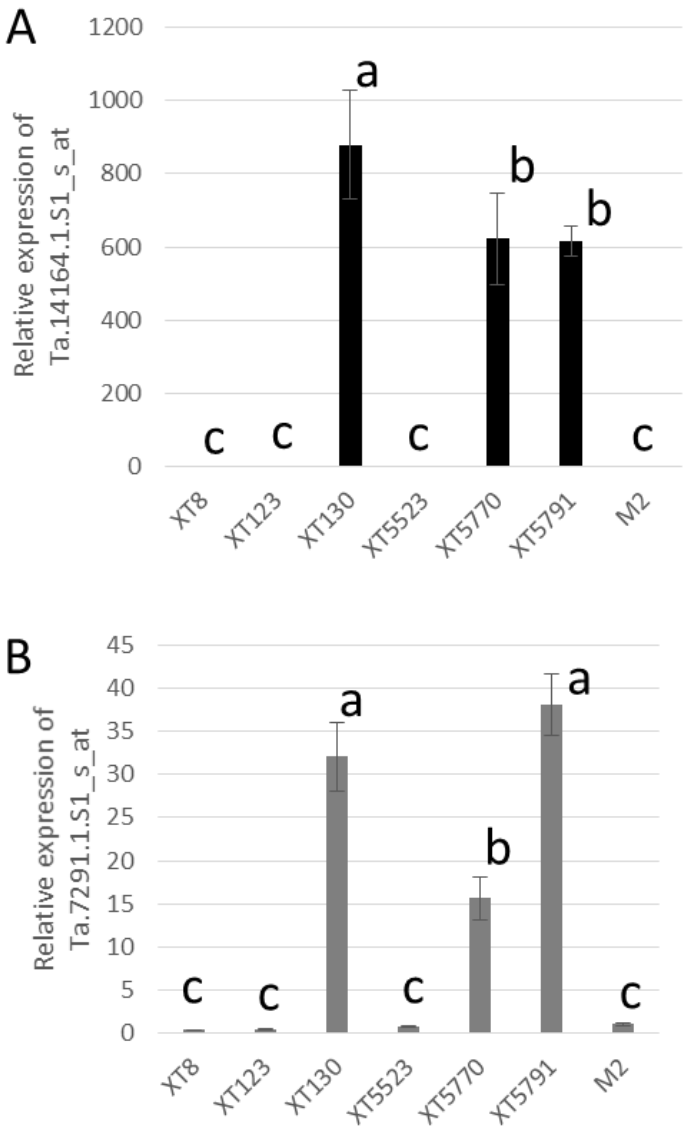


Figure 3-22. Induction of Ta.14164.1.S1_s_at and Ta.7291.1.S1_s_at is strain specific. XT8 is *X. translucens* pv. *translucens*, XT123 is *X. translucens* pv. *cerealis*, XT130, XT5523, XT5770 and XT5791 are strains of *X. translucens* pv. *undulosa*. M2 is the mutant of XT4699 lack of *tal6*. The relative expression level is calculated relative to M2 treatment by with $2^{-\Delta\Delta Ct}$ method. A, the relative expression of Ta.14164.1.S1_s_at, corresponding to bHLH gene, was calculated; B, the relative expression of Ta.7291.1.S1_s_at, corresponding to succinate dehydrogenase gene, was calculated. The different lowercase letters indicate significant difference with P-value<0.05 in the ANOVA statistics analysis.

Chapter Four

Conclusions and Prospects

The genus of *Xanthomonas* encompasses diverse groups of plant pathogenic bacteria, causing diseases on a wide range of plants (White et al., 2009). Transcription-Activator Like (TAL) effectors are important weapons used by *Xanthomonas* bacteria, turning on expression of corresponding host susceptibility genes and resulting in virulence contributions. TAL effector triggered susceptibility is well documented in bacterial blight of rice, which is caused by *X. oryzae* pv. *oryzae.* The disease susceptibility is mainly dependent on the transcriptional activation of sugar transporter *OsSWEET* genes by the major TAL effectors, PthXo1, PthXo2, PthXo3, AvrXa7, TalC and Tal5 from *X. oryzae* pv. *oryzae* strains (Antony et al., 2010, Streubel et al., 2013, Yang et al., 2006, Yu et al., 2011, Zhou et al., 2015).

In chapter 2, the major TAL effector PthXo2 was shown to function in a similar manner to previously characterized major TAL effectors in *X. oryzae* pv. *oryzae* in that each targets a member of the clade III sweet genes of rice. In the process, a cryptic recessive resistance gene was identified in bacterial blight of rice, indicating that selection for recessive resistance may have been occurring prior to formal organized breeding efforts and the rise of formal R gene nomenclature. PthXo2, from *X. oryzae* pv. *oryzae* Asian strain JXO1, transcriptionally activates the clade III gene *OsSWEET13*, which joins the two previously identified genes *OsSWEET11* (targeted by PthXo1) and *OsSWEET14* (targeted by multiple TAL effectors including AvrXa7) as the three genes that are known to be targeted in the field. Interestingly, PthXo2 was found to only function in *indica* rice but not in *japonica* rice. Analysis of the DNA binding specificity of PthXo2 revealed a critical one base pair deletion in the effector binding element (EBE) of PthXo2 in *japonica* cultivars and an *indica* cultivar harboring the recessive resistance gene *xa25*. Deletion of fifth repeat in the gene

for PthXo2, corresponding with the one base pair deletion in the EBE region of *OsSWEET13* in *japonica* rice, enabled the modified effector PthXo2D to function in *japonica* rice varieties and confer virulence to strains containing the gene on *japonica* rice varieties. However, the change brought about a concomitant loss of function in *indica* rice cultivars. The presence of the one base pair deletion in the EBE of PthXo2 effector in many *japonica* rice cultivars suggests that the *xa25* recessive resistance is not rare (Zhou et al., 2015).

In chapter 2, I identified TalC from *X. oryzae* pv. *oryzae* African strain AXO1947 as the major virulence effector and the host susceptibility gene *OsSWEET14* targeted by TalC, which was consistent with previous experiment data of TalC from another African strain BAI3 (Yu et al., 2011). My data also showed the *talC* effector gene was interchangeable with other major TAL effector genes, *pthXo1* and *avrXa7*, from Asian strains in virulence contribution, which was failed to be demonstrated in Yu et al., 2011. In addition, I also identified another potential TAL mutant AME8 that lost the ability to induce *OsTFX1* and *OsERF* genes and suffered 50% virulence loss compared to wild type although it contained intact *talC* effector gene. The virulence loss of AME8 could be partially complemented by TAL effector gene *pthXo6* of Asian strain, which had been demonstrated as virulence effector by inducing expression of susceptibility gene *OsTFX1* in rice (Sugio et al., 2007). My experiment data suggests TAL effectors from both Asian and African *X. oryzae* pv. *oryzae* strains target similar host susceptibility genes in the bacterial blight diseases. The results indicate that despite independent lineages the Asian and African strains have convergently evolved the same strategy in terms of targeted genes for achieving virulence in rice.

Based on the presence of TAL effectors in many species of *Xanthomonas*, TAL effector-mediated enhanced susceptibility likely occurs in many *Xanthomonas* associated diseases (White et al., 2009). However, the TAL effector dependent host susceptibility genes are only known in a

few cases. Besides to bacterial blight of rice, TAL effector triggered susceptibility was also observed in three other disease complexes, including citrus bacterial canker by *X. citri* pv. *citri*, bacterial leaf streak of rice by *X. oryzae* pv. *oryzicola* and bacterial blight of cassava by *X. axonopodis* pv. *manihotis* (Cernadas et al., 2014, Cohn et al., 2014, Hu et al., 2014). In addition, the biochemical basis is unknown for any TAL effector based enhanced susceptibility. Therefore it behooves us to examine TAL effector-dependent susceptibility in a wider set of diseases. In chapter 3, genomic analysis of *X. translucens* strains, the causal agents of bacterial leaf streak in cereal crops, was performed with the eventual goal of determining the roles played by TAL effectors in *X. translucens*. A complete genome sequence of *X. translucens* pv. *undulosa* strain XT4699 was obtained by PacBio long read, single molecule, real time (SMRT) DNA sequences. The complete genome sequence revealed the presence of eight members of TAL effector genes in XT4699. Phylogenetic analyses revealed TAL effectors of *X. translucens* strains (group 1 of *Xanthomonas*) were distant from TAL effectors from *Xanthomonas* strains of group 2. Microarray and qRT-PCR analyses revealed the modulation of wheat host gene expression by specific TAL effectors of XT4699. However, mutations of any TAL effector gene in XT4699 did not reduce the bacterial population growth significantly in 'Chinese Spring' wheat cultivar.

The possibility of virulence contributions of TAL effectors from *X. translucens* strains to diseases on cereal crops could not be precluded even if the bacterial growth of XT4699 in 'Chinese Spring' wheat is not impaired by TAL mutations. In previous studies, AvrBs3 from *X. campestris* pv. *vesicatoria* contributed to bacterial virulence by inducing cell hypotrophy in pepper, Tal2g of *X. oryzae* pv. *oryzicola* promoted lesion extension and bacterial exudation in rice and Avrb6 from *X. campestris* pv. *malvacearum* functioned in causing watersoaking and bacterial exudation in cotton, but none of them played roles in bacterial growth in their cognate plant tissues (Cernadas

et al., 2014, Kay et al., 2007, Yang et al., 1994). Therefore, other experimental strategies might need to be applied for detecting the phenotype difference between WT and TAL mutants. The potential strategies are: (1) screen different wheat or barley cultivars for inoculations; (2) screen plants at different growth stages; (3) try different *X. translucens* pv. *undulosa* strains or other pathovars for detecting the virulence contributions of TAL effector genes; (4) try different inoculation techniques; and (5) introduce TAL effector genes into other weak virulent strains for gain-of-function studies.

In addition, XT4699-*tal7* was identified as a conserved TAL effector gene among 9 of 13 strains in chapter 3. Cloning and sequencing of TAL effector genes in larger collection of *X. translucens* strains is needed to get more reliable information regarding with conserved TAL effector genes. The insertion of DNA elements, specifically designed to be recognized by RVDs of conserved TAL effectors, in the promoter regions of effective R genes was applied for controlling bacterial blight and bacterial leaf streak of rice (Hummel et al., 2012, Zeng et al., 2015). Therefore, for controlling the bacterial leaf streak on cereal crops, use of the RVDs of conserved *X. translucens* TAL effectors may allow the engineering of specific super terminator R genes in cereal crops.

Reference

Antony G, Zhou J, Huang S, *et al.*, 2010. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell* **22**, 3864-76.

Cernadas RA, Doyle EL, Niño-Liu DO, *et al.*, 2014. Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathog* **10**, e1003972.

Cohn M, Bart RS, Shybut M, *et al.*, 2014. Xanthomonas axonopodis Virulence Is Promoted by a Transcription Activator-Like Effector-Mediated Induction of a SWEET Sugar Transporter in Cassava. *Molecular Plant-Microbe Interactions* **27**, 1186-98.

Hu Y, Zhang J, Jia H, *et al.*, 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proceedings of the National Academy of Sciences* **111**, E521-E9.

Hummel AW, Doyle EL, Bogdanove AJ, 2012. Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist* **195**, 883-93.

Kay S, Hahn S, Marois E, Hause G, Bonas U, 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-51.

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**, 808-19.

Sugio A, Yang B, Zhu T, White FF, 2007. Two type III effector genes of Xanthomonas oryzae pv. oryzae control the induction of the host genes OsTFIIAγ1 and OsTFX1 during bacterial blight of rice. *Proceedings of the National Academy of Sciences* **104**, 10720-5.

White FF, Potnis N, Jones JB, Koebnik R, 2009. The type III effectors of Xanthomonas. *Molecular Plant Pathology* **10**, 749-66.

Yang B, Sugio A, White FF, 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* **103**, 10503-8.

Yang Y, De Feyter R, Gabriel DW, 1994. Host-specific symptoms and increased release of Xanthomonas citri and X. campestris pv. malvacearum from leaves are determined by the 102-bp tandem repeats of pthA and avrb6, respectively. *MPMI-Molecular Plant Microbe Interactions* **7**, 345-55.

Yu Y, Streubel J, Balzergue S, *et al.*, 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**, 1102-13.

Zeng X, Tian D, Gu K, *et al.*, 2015. Genetic engineering of the Xa10 promoter for broad-spectrum and durable resistance to Xanthomonas oryzae pv. oryzae. *Plant biotechnology journal*.

Zhou J, Peng Z, Long J, *et al.*, 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *The Plant Journal* **82**, 632-43.