DISSERTATION

BIOLOGY, COMPARATIVE GENOMICS AND MOLECULAR DIAGNOSTICS OF *XANTHOMONAS* SPECIES INFECTING RICE AND CORN

Submitted by

Jillian M. Lang

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2017

Doctoral Committee:

Advisor: Jan E. Leach

Daniel Bush
Anireddy Reddy
Valérie Verdier

ABSTRACT

BIOLOGY, COMPARATIVE GENOMICS AND MOLECULAR DIAGNOSTICS OF *XANTHOMONAS* SPECIES INFECTING RICE AND CORN

Emerging bacterial diseases on staple and economically important crops can pose critical threats to food security. Accurate identification of bacterial plant pathogens is the foundation of effective management for growers. This work advances the application of genomics to identify and characterize bacterial plant pathogens in the genus *Xanthomonas* that can cause destructive diseases on most agricultural crops, including rice and corn. In this thesis, taxonomy, host range, disease phenotypes and basic biology of the following pathogens were established: *X. oryzae* pv. *oryzae, X. o.* pv. *oryzicola, X. o.* pv. *leersiae* and *X. vasicola* pv. *vasculorum. X. o.* pv. *oryzae* and *X. o.* pv. *oryzicola* infect rice and cause bacterial blight and bacterial leaf streak, respectively. *X. o.* pv. *leersiae* infects cutgrass (*Leersia* sp.), weedy grasses that can serve as alternative hosts to *X. oryzae* and are endemic in all rice growing regions. *X. vasicola* pv. *vasculorum* was identified as the causal agent of bacterial leaf streak of corn, an emerging and now wide-spread disease in the United States, that was reported for the first time in 2017. This work established that *X. vasicola* pv. *vasculorum* can also infect sorghum and sugarcane and that the US strain is 99% similar to strains isolated over 20 years ago in S. Africa.

To develop robust molecular diagnostic tools for these pathogens, unique features needed to be first identified. Using comparative genomics that included closely related bacteria and distant relatives, PCR-based diagnostic tools were developed, then validated using isolated cultures and field grown plant materials. Comparative genomics also contributed to elucidation of the taxonomy and phylogeny of *X. o.* pv. *leersiae* and *X. v.* pv. *vasculorum.* Characterization of *X. o.* pv. *leersiae* revealed adaptations to both the weedy grass hosts and rice. These features include virulence proteins that target homologous host genes (transcription activator like effectors, TALEs) to influence host gene expression. I conclude that *X. oryzae* is a complex that includes *X. oryzae* pv. *oryzae, X. o.* pv. *oryzicola* and *X. o.* pv. *leersiae* and that

this complex can provide a unique window into pathogen evolution. By better understanding how pathogens adapt to their environments including new hosts, growers can manage surrounding ecosystems more effectively to minimize yield losses and therefore, contribute to food security.

ACKNOWLEDGEMENTS

‘Alone, all alone
Nobody, but nobody
Can make it out here alone.’
- From *‘Alone’* in *Oh Pray My Wings Are Gonna Fit Me Well*, Maya Angelou, 1975

This long journey has been alongside a village of generous, kind humans. I would not have come this far alone. Below is a small attempt to acknowledge them, though no amount of words can fully articulate the gratitude and humility I feel.

Dr. Jan Leach: your role in my life is so multi-faceted. You are my mentor, advisor, boss, friend and my family. I love you dearly. You inspire me with your brilliance, strength and immense generosity. Thank you for supporting me, challenging me and giving me more opportunities than I ever thought possible in my life. You have let me see the world and subsequently, grow as a scientist and a person. Dr. Valérie Verdier: Thank you for your generous support both while Owen and I were in France and for the invigorating science while you were in the US. I admire your perseverance, breadth of knowledge and ability to manage and balance so many things. I am incredibly fortunate to have two women who are amazing scientists, leaders and friends. My remaining committee members, Dr. Dan Bush and Dr. Anireddy Reddy: Thank you for serving on my committee and for having patience as my research shifted, turned, and ultimately culminated in this dissertation. Thank you for challenging me with questions of science and career, and the time you spent with me.

There are far too many people to thank in the lab for being fun, charismatic co-workers and integral members of my research. I thank you all past and present for your patience as I tackled this challenge while balancing my job duties. Particularly, these last two months as I have driven to completion I know my absence was significant. Thank you for giving me the support and space to complete this massive endeavor. A few people specifically I would like to acknowledge here: Emily Luna and John Long – you both accepted and relieved me of duties while I traveled abroad and as I turned my focus to this document, I am indebted to your willingness to help our group and myself. Your quiet generosity and

kindness is inspiring. Elysa DuCharme – I still argue you have the ability to read my mind. I am so proud of the scientist you have become and thankful for all the time you worked on all my crazy ideas with grace and calm. This data would not have been generated without you. Paul Langlois – your contributions to optimizing LAMP were no small feat and I admire your organization and patience. Robin Mom and Rémi Pelissier – two excellent interns who traveled from abroad to join me in research. It was a privilege to mentor you. Thank you for your excitement, flexibility and dedication to helping through several projects and making your summers in the lab fun and productive. Drs. Kirk Broders, Tamra Jackson-Ziems, Kevin Korus, Jeff Hoy and Terra Hartman: thank you for coordinating corn and sugarcane samples and facilitating the challenging progress of reporting bacterial leaf streak of corn in the USA for the first time. Lastly, to all the wonderful undergraduate and high school students who have contributed in some form to this work from growing countless plants in the greenhouse to washing dishes to pouring mountains of media, I thank each of you for your bright energy, hard work, excitement about plant pathology and smiles. Funding for this work and many miles traveled came from the Embassy of France, USDA, C2B2, BecA, IS-MPMI and APS.

My dear friends and colleagues at IRD (past and present) – Drs. Alvaro Pérez-Quintero, Ralf Koebnik, Mathilde Hutin, Céline Pesce, Tu Tran and Boris Szurek – thank you for hosting me and continuing to engage in collaborative science and friendship, I am grateful to know such an intelligent, lovely group of people. Drs. Ned Tisserat and Howard Schwartz while you have both retired, you hold a dear place in my scientific journey. I admire your immense knowledge of plant pathology and dedication to it. Thank you for mentoring me and giving me the foundation to build this degree on.

Drs. Scott Fulbright and Stephen Chisholm– you provided me with so much encouragement when I was deciding to complete this degree and throughout its duration. Thank you for having faith in me and pushing me to take this on. Dr. Jonathan Jacobs – you are my scientific compliment and have become a beloved friend. Thank you for your neverending enthusiasm, support at the bench and beyond and so much laughter. Dr. Federico Martin – you hold a special place in my heart. Thank you for always

encouraging and supporting me. You taught me to be paitient and confident and to strive for balance when I needed it most.

To Kate and Mitch Schneider – you are such dear, generous friends. I am incredibly grateful to you for the shelter, love and support these past ten years, you are my family. To Paul Covey – our ability to tackle and balance the diverse challenges of life humbles and amazes me. Thank you for your support scientifically and personally over so many years. To my mother (Kath): you have donated hundreds of hours caring for my son, my dog, preparing meals, offering refuge and support endlessly and selflessly. You have been my biggest cheerleader my entire life and I never take that for granted. Jamie, my dear sister and lifelong friend, you have been my rock and I know the three letters that will soon follow my name are because of you both. Thank you, I love you dearly. Dr. Christine Battista – you are my beacon of light, laughter, wisdom, strength and comfort. You hold such a large place in my life and I am certain I would not have survived this journey without you.

Finally, I dedicate this work to my son, Owen Jude Covey. You are my favorite human. You have given me a reason to pursue science for a greater good and have been so patient through this big adventure. I hope you can see that hard work and dedication can culminate in a bountiful basket of friends and family, not just results and pride. My love for you is infinite.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 RICE AND CORN: STAPLE CROPS AND MODEL PHYTOPATHOLOGICAL SYSTEMS

Staple food crops are defined as those foods that regularly consumed in large quantities, and that form the basis of traditional diets and serve as a major source of energy and nutrients for the consumers. Both rice (*Oryza sativa*) and corn (syn. maize, *Zea mays*) are staple foods. As of 2014, over 165 and 180 million hectares of land was used for the production of rice and corn, respectively, worldwide. Approximately 90% of that rice production occurred in Asia and 53% of corn production occurred in the Americas (1). In many Asian countries, rice is consumed in at least three meals a day. While rice is primarily a food crop, corn is produced for livestock feed, processed into starch, sweeteners, corn oil, beverage and industrial alcohol, and, in addition to food, fuel ethanol. The United States is a major contributor to the world corn trade market, with between 10 and 20 percent of the corn crop exported to other countries (2). These crops are monocotyledonous and members of the *Poaceae* family that is comprised of other agriculturally important grass crops, such as wheat, barley and millet, as well as weedy grasses. A weedy species of interest to this work is the genus *Leersia*. *Leersia* spp., commonly called cutgrass, are pan-tropically distributed members of the *Oryzeae* tribe in *Poaceae* and the most closely related genus to *Oryza* (3). These genera branched from remaining genera in this family c. 20 mya and diverged from each other c. 14 mya (4).

Rice and corn are considered model systems for examining biological questions of genetics, molecular breeding, bioenergy, molecular plant-microbe interactions and agricultural improvements in yield, quality and resilience to climate change. For these reasons, significant resources have been developed to support research on these plants and their environments including genomes and transcriptomic, proteomic and metabolomic data sets. Rice has a relatively small, diploid genome (430 Mb) that has been fully sequenced (5). Moreover, 3000 additional genomes were sequenced (6) fueling

the identification of SNPs across the immense diversity in the species (7). The maize genome is much larger (2500 Mb) and more complicated in part due to the presence of highly repetitive regions and transposable elements (8). Both of these staple crops have highly virulent pathogens that can infect, spread and devastate paddies or fields significantly, impacting yields and possibly food security in developing countries where they depend upon these crops for nutrition and income. The combination of the critical importance of these crops, the genetic and genomic resources available and the complexity of their phytobiomes, make them model, translational phytopathological systems.

1.2 *XANTHOMONAS* SPECIES: DIVERSE AND SUCCESSFUL PATHOGENS

The genus *Xanthomonas* is part of a large order of Gram-negative bacterial plant pathogens within the class γ-Proteobacteria that can cause diseases on at least 124 monocots and 268 dicots (9). Almost every agronomically important crop is infected by at least one *Xanthomonas* spp. Plant pathogenic bacteria are further classified beyond species into pathovars. The term *pathovar* is used to refer to a strain or set of strains with the same or similar characteristics, differentiated at an infra-subspecific level from other strains of the same species or subspecies on the basis of distinctive pathogenicity to one or more plant hosts (10). More simply, pathovars differentiate these organisms based on host and tissue specificity. These intricate details in taxonomy are important for growers and researchers because sustainable management of plant diseases can depend on accurate identification of the causal agents and what sources of resistance may be available.

Valuable examples of this scenario are *X. oryzae* pv. *oryzae* (Xoo) and *X. o.* pv. *oryzicola* (Xoc), that are the causal agents of bacterial blight and bacterial leaf streak of rice, respectively. These diseases continue to threaten major rice growing regions of Asia and Africa because the of the potential for significant yield loss (11, 12). There are reports of *X. oryzae* (Xo), with no pathovar designation, in the United States, but it is distinct from Xoo and Xoc (13). Currently, all *X. oryzae* are considered select agents (https://www.selectagents.gov/) by the United States Department of Agriculture according to the Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (Public Law, 107-188, June 12, 2002). The US strains are weakly virulent and divergent from highly virulent African and Asian

lineages (14). All *X. oryzae* are physiologically, morphologically and genetically similar yet, they cause distinct diseases of rice. Xoo elegantly enters leaves through wounds or hydathode water pores and colonizes xylem vessels (15, 16). Over 40 resistance (R) genes to Xoo have been identified, and nine have been cloned: *Xa1, Xa3/Xa26, xa5, Xa10, xa13, Xa21, Xa23, xa25,* and *Xa27* (11, 12, 17–25). Xoc on the other hand, is restricted to the apoplast and there are only two potential resistance genes reported from rice that are not yet employed in the field (26, 27). Interestingly, one additional R gene from corn, *Rxo1*, confers stable resistance to Xoc strains containing the effector, *avrRxo1* (syn. *xopAJ*) (28, 29). Unfortunately, deployment of this potentially very effective R gene requires transgenic approaches that are heavily regulated or even illegal in many rice growing countries.

The longstanding evolutionary battle between plants and microbes has produced novel and impressive mechanisms of defense and virulence. Probably the most impressive cross-kingdom influences are the suite of effectors bacteria produce and inject into a plant cell via the type-three secretion system. Transcription activator like effectors (TALEs) are one group of these proteins that can directly and precisely bind host target promoter sequence or effector binding elements (EBEs) to influence gene expression, mimicking a eukaryotic transcription factor (Fig. 1.1, 30, 31). These targeted genes, or

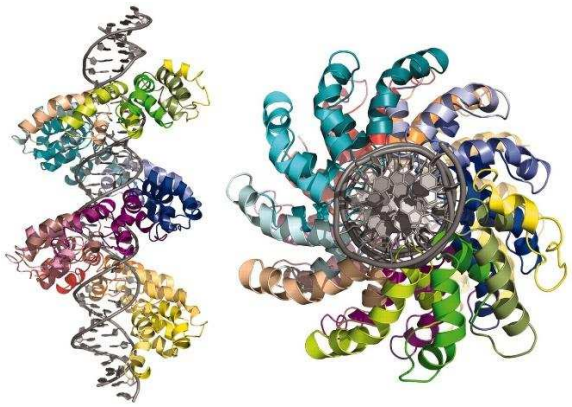

Fig. 1.1. Crystal structure of DNA binding region of TAL effector pthXo1 (*X. oryzae* pv. *oryzae)* bound to its DNA target (32).

susceptibility genes, can create a conducive environment for bacterial fitness thereby promoting disease. Establishing libraries of the TAL effector genes, called TALomes, of *Xanthomonas* species has triggered

an immense new field of research from taxonomy, evolution, functional biology and engineering of host resistance (33).

Corn bacterial leaf streak disease was initially reported in South Africa but has become a concerning new disease in the United States caused by *X. vasicola* (34). Little is known about the etiology and biology of this disease. Known *X. vasicola* pathovars include *vasculorum, holcicola* and *musacearum.* The host ranges for these pathovars overlap, and include corn, sugarcane, sorghum, and banana, but not all pathovars can infect all hosts. This complex of bacteria is convoluted and genome sequence has provided some pieces of information about what differentiates these organisms possibly allowing them to adapt to their hosts and cause disease. Chapter 4 of this dissertation unravels the taxonomy of *X. vasicola* using genomics, and reports a preliminary disease phenotyping of this group of organisms (**Fig 1.2**).

Fig. 1.2. Field symptoms of corn bacterial leaf streak caused by *X. vasicola* pv. *vasculorum.*
(Photo credit – T. Jackson-Ziems)

1.3 GENOMICS OF *XANTHOMONAS* INFORM MECHANISMS AND EVOLUTION OF PATHOGENICITY

As of this writing, there are 591 publicly available *Xanthomonas* genomes representing almost every species in the genus (https://www.ncbi.nlm.nih.gov/genome). Of these, approximately 50 are closed. High-throughput sequencing technologies have advanced rapidly in recent years. Early genomes were assembled from shotgun Sanger reads or 454 pyrosequencing (Roche, Branford, CT). Next generation sequencing by Illumina has contributed the highest number of *Xanthomonas* spp. genomes by its 'sequencing by synthesis' approach and continues to be the standard for draft bacterial genome sequence. Finally, the most promising technologies yet, include single molecule real time long read (SMRT, Pacific Biosciences, Menlo Park, CA) and the MinION (Oxford Nanopore). SMRT sequencing generates significantly longer reads (~ 10-60 Kb) that can cover highly repetitive or complicated modified regions (methylation, e.g.). This means, that with a single run of SMRT sequencing, an entire bacterial genome can be assembled to completion. Of interest to this work, highly repetitive TALE sequences can be captured immediately with this type of sequencing. Whereas, just a few years ago the resolution of TALE was laborious, error prone and costly. For this reason, SMRT sequencing has already become the a standard in *Xanthomonas* genomics (35–39). Arguably, the next revolutionary innovation is the MinION (Oxford Nanopore) portable sequencer because it can simply run on a standard computer via USB (40). The option to select read length and the ability to generate 5-10 Gb of data in a single cell at low cost will certainly advance the field of genomics in the near future.

The concept of a species in prokaryotes can be convoluted. Particularly with evolutionary pressures such as environment, increasing climatic temperatures and low agricultural diversity combined with their ability to rapidly exchange genetic material. Monocultures not only shape soils but directly dictate biological pests and predators. Ideally, a combination of multi-locus sequence alignment (MLSA), comparison of whole genome sequence and ecology are integrated to define a prokaryotic species (41). DNA-DNA hybridization combined with restriction fragment length polymorphism were historically used to differentiate bacterial species. Average nucleotide identity (ANI), based on whole genome alignment

has now replaced this once complicated laboratory procedure with a simple program that can be completed in less than a day. A cut-off of 95% or greater delineates two organisms belonging to the same species (42–44).

As sequencing costs and error rates continue to decline, application of whole genome comparisons from single organism identification and taxonomic placements to whole population monitoring could rapidly enable epidemiological surveys and ultimately, crop disease management. Beyond this, genomics is facilitating population biology and epidemiology by allowing the precise monitoring of strains of organisms or differentiation of pathogens that are often misdiagnosed or present in mixed infections. Genomics can even predict the center(s) of pathogen diversity, which could be the basis for a network of phenotyping centers to analyze germplasm resistance (45, 46).

1.4 MOLECULAR DIAGNOSTICS FACILITATE DISEASE MANAGEMENT

Diagnosis of a plant disease is the first step towards deciding effective management strategies that can reduce crop losses. Growers, extension agents, federal regulatory agencies and private production companies rely on fast, accurate tools to identify threatening pathogens, particularly emerging diseases that may be difficult to visually diagnose. Even for commonly seen diseases, diagnoses made primarily on the basis of symptoms and knowledge of previous host-pathogen relationships and foregoing isolation then biochemical and morphological pathogen identification may lead to a missed opportunity to discover new pathogens or observe changing pathogen populations (46). Application of molecular diagnostics in plant pathology has improved the speed and accuracy of identifying pathogens from seed to post-harvest. However, the emergence of new or reoccurring diseases requires continually improved efficiency in surveillance techniques as well as expanded libraries of tools specific to these new diseases. Widely accepted and currently implemented molecular approaches improved the capacity to respond to new threats as they emerge, but they can be costly and time-consuming (47). Immunodetection by ELISA and conventional or multiplex PCR are still used in many labs for detection. Quantitative real-time PCR emerged as a more sensitive standard for detection and quantification of pathogens, particularly obligate organisms such as viruses, phytoplasmas and unculturable bacteria, such as the devastating *Liberobacter*

spp. However, these assays require expensive reagents and equipment that are not field appropriate or available in developing countries.

Certainly, the most revolutionary advancement in diagnostics has been genomics. The fundamental basis for a specific molecular diagnostic assay is discovering a unique sequence feature of an organism. Whole draft genome comparisons can quickly identify polymorphic regions on which design can be based (48–50). Validation of assays not only *in silico* using abundant, publicly available databases, but also against a large, diverse panel of closely related organisms in real time is essential to establish confidence in an assay's specificity. While genome sequence requires a lab environment to achieve pure cultures and quality DNA, employment of a pipeline from genomics to in-field diagnosis has begun. One such application is loop mediated isothermal amplification (LAMP) which incorporates six primers surrounding a unique target locus amplified by a displacing polymerase (e.g., Bst DNA polymerase) functioning at one temperature (51). This isothermal feature removes the need for cycling equipment thereby enabling field application with incubation in something as simple as a thermos of warm water. Furthermore, this technique has inherently higher specificity than most conventional assays due to incorporation of six primers surrounding a unique region as opposed to a single primer pair. LAMP is also less sensitive to inhibitors (52) that can complicate results and cause time and monetary losses.

Disease diagnosis requires intelligent field and laboratory observations as well as accurate identification of the pathogen. Plant pathogenic bacteria, which are enormously diverse in the environment, often require multiple complementary tests for a definitive identification (53) and leveraging genomics for this task will continue to prove substantially informative, increase accuracy and speed to management decisions.

1.5 SCOPE OF DISSERTATION

This dissertation aims to demonstrate the powerful applications of genomics in phytopathology. These applications include molecular diagnostics, informing pathogen identification and taxonomic placement, and providing insights into evolutionary adaptation of pathogens to hosts in agroecosystems. In Chapter 2, I demonstrate the translation of an existing molecular tool to differentiate *X. oryzae*

pathovars in a field applicable assay (LAMP) based on unique loci identified through comparative genomics. We optimized assay conditions and characterized sensitivity and specificity of this assay with different specific primers. Chapter 3 examines an organism that is closely related to *X. oryzae, X. o.* pv. *leersiae*. My goal was to begin the process of identifying genes involved in adaptation to its weedy host (*Leersia* spp.) and to characterize its virulence mechanisms. I clarified the taxonomy of this organism and identified its unique suite of TALEs using SMRT sequencing technologies. My collaborators and I used host draft genome sequence to predict virulence targets in *Leersia perrieri,* a sequenced *Leersia* spp., and inferred relationships with rice. This information was used to determine similarities and differences in the parallel *X. oryzae* – rice pathosystem.

Lastly, in Chapter 4, we report on an emerging disease of corn in the United States, bacterial leaf streak, caused by *X. vasicola* pv. *vasculorum*. Little is known about this organism and how it has spread so rapidly. Therefore, I used MLSA and draft genome sequence to confirm identity and propose the pathovar name. Further, I developed a diagnostic assay that is now widely used across the United States in academic and regulatory institutions to diagnose and monitor the pathogens presence and spread. I collected comprehensive phenotypic data to better understand the pathogen's host range. In summary, this thesis reports the host range and genome characterization of *Xanthomonas* from three different pathosystems. My results have contributed to (1) clarification of the taxonomic classification of these important pathogens, 2) insights into the biology and evolution of bacterial pathogenicity, and (3) the development and deployment of validated diagnostics for epidemiologic studies, quarantine applications, and disease control decisions.

REFERENCES

1. **Food and Agriculture Organization of the United Nations**. 2017. FAOSTAT.
2. **United States Department of Agriculture Economic Research Service**. 2017. Corn and other feed grains.
3. **Kellogg EA**. 2001. Evolutionary history of the grasses. Plant Physiol **125**:1198–1205.
4. **Guo Y-L**, **Ge S**. 2005. Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. Am J Bot **92**:1548–58.
5. **Kawahara Y**, **de la Bastide M**, **Hamilton JP**, **Kanamori H**, **McCombie WR**, **Ouyang S**, **Schwartz DC**, **Tanaka T**, **Wu J**, **Zhou S**, **Childs KL**, **Davidson RM**, **Lin H**, **Quesada-Ocampo L**, **Vaillancourt B**, **Sakai H**, **Lee SS**, **Kim J**, **Numa H**, **Itoh T**, **Buell CR**, **Matsumoto T**. 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice **6**:4.
6. **3,000 rice genomes project**. 2014. The 3,000 rice genomes project. Gigascience **3**:7.
7. **Alexandrov N**, **Tai S**, **Wang W**, **Mansueto L**, **Palis K**, **Fuentes RR**, **Ulat VJ**, **Chebotarov D**, **Zhang G**, **Li Z**, **Mauleon R**, **Hamilton RS**, **McNally KL**. 2015. SNP-Seek database of SNPs derived from 3000 rice genomes. Nucleic Acids Res **43**:D1023–D1027.
8. **Schnable PS**, **Ware D**, **Fulton RS**, **Stein JC**, **Wei F**, **Pasternak S**, **Liang C**, **Zhang J**, **Fulton L**, **Graves TA**, **Minx P**, **Reily AD**, **Courtney L**, **Kruchowski SS**, **Tomlinson C**, **Strong C**, **Delehaunty K**, **Fronick C**, **Courtney B**, **Rock SM**, **Belter E**, **Du F**, **Kim K**, **Abbott RM**, **Cotton M**, **Levy A**, **Marchetto P**, **Ochoa K**, **Jackson SM**, **Gillam B**, **Chen W**, **Yan L**, **Higginbotham J**, **Cardenas M**, **Waligorski J**, **Applebaum E**, **Phelps L**, **Falcone J**, **Kanchi K**, **Thane T**, **Scimone A**, **Thane N**, **Henke J**, **Wang T**, **Ruppert J**, **Shah N**, **Rotter K**, **Hodges J**, **Ingenthron E**, **Cordes M**, **Kohlberg S**, **Sgro J**, **Delgado B**, **Mead K**, **Chinwalla A**, **Leonard S**, **Crouse K**, **Collura K**, **Kudrna D**, **Currie J**, **He R**, **Angelova A**, **Rajasekar S**, **Mueller T**, **Lomeli R**, **Scara G**, **Ko A**, **Delaney K**, **Wissotski M**, **Lopez G**, **Campos D**, **Braidotti M**, **Ashley E**, **Golser W**, **Kim H**, **Lee S**, **Lin J**, **Dujmic Z**, **Kim W**, **Talag J**, **Zuccolo A**, **Fan C**, **Sebastian A**, **Kramer M**, **Spiegel L**, **Nascimento L**, **Zutavern T**, **Miller B**, **Ambroise C**, **Muller S**, **Spooner W**, **Narechania A**, **Ren L**, **Wei S**, **Kumari S**, **Faga B**, **Levy MJ**, **McMahan L**, **Van Buren P**, **Vaughn MW**, **Ying K**, **Yeh C-T**, **Emrich SJ**, **Jia Y**, **Kalyanaraman A**, **Hsia A-P**, **Barbazuk WB**, **Baucom RS**, **Brutnell TP**, **Carpita NC**, **Chaparro C**, **Chia J-M**, **Deragon J-M**, **Estill JC**, **Fu Y**, **Jeddeloh JA**, **Han Y**, **Lee H**, **Li P**, **Lisch DR**, **Liu S**, **Liu Z**, **Nagel DH**, **McCann MC**, **SanMiguel P**, **Myers AM**, **Nettleton D**, **Nguyen J**, **Penning BW**, **Ponnala L**, **Schneider KL**, **Schwartz DC**, **Sharma A**, **Soderlund C**, **Springer NM**, **Sun Q**, **Wang H**, **Waterman M**, **Westerman R**, **Wolfgruber TK**, **Yang L**, **Yu Y**, **Zhang L**, **Zhou S**, **Zhu Q**, **Bennetzen JL**, **Dawe RK**, **Jiang J**, **Jiang N**, **Presting GG**, **Wessler SR**, **Aluru S**, **Martienssen RA**, **Clifton SW**, **McCombie WR**, **Wing RA**, **Wilson RK**. 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science (80-) **326**:1112–1115.
9. **Hayward AC**. 1993. The hosts of Xanthomonas, p. 1–17. *In* Swings, JG, Civerolo, EL (eds.), Xanthomonas. Chapman and Hall, London.
10. **Dye DW**, **Bradbury JF**, **Goto M**, **Hayward AC**, **Lelliott R.**, **Schroth MN**. 1980. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. Rev Pl Pathol **59**:153–168.
11. **Verdier V**, **Vera Cruz C**, **Leach JE**. 2012. Controlling rice bacterial blight in Africa: Needs and prospects. J Biotechnol **159**:320–328.
12. **Nino-Liu DO**, **Ronald PC**, **Bogdanove AJ**. 2006. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. Mol Plant Pathol **7**:303–324.
13. **Jones RK**, **Barnes LW**, **Gonzalez CF**, **Leach JE**, **Alvarez AM**, **Benedict AA**. 1989. Identification of low virulence strains of *Xanthomonas campestris* pv. *oryzae* from rice in the

United States. Phytopathology **79**:984–990.
14. **Triplett LR**, **Hamilton JP**, **Buell CR**, **Tisserat NA**, **Verdier V**, **Zink F**, **Leach JE**. 2011. Genomic analysis of *Xanthomonas oryzae* isolates from rice grown in the United States reveals substantial divergence from known *X. oryzae* pathovars. Appl Environ Microbiol **77**:3930–3937.
15. **Guo A**, **Leach JE**. 1989. Examination of rice hydathode water pores exposed to *Xanthomonas campestris* pv. *oryzae*. Phytopathology **79**:433–436.
16. **Mew TW**. 1984. Scanning electron microscopy of virulent and avirulent strains of *Xanthomonas campestris* pv. *oryzae* on rice leaves. Phytopathology **74**:635.
17. **Yoshimura S**, **Yamanouchi U**, **Katayose Y**, **Toki S**, **Wang Z-X**, **Kono I**, **Yano M**, **Iwata N**, **Sasaki T**. 1998. Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc Natl Acad Sci USA **95**:1663–1668.
18. **Gu K**, **Yang B**, **Tian D**, **Wu L**, **Wang D**, **Sreekala C**, **Yang F**, **Chu Z**, **Wang GL**, **White FF**, **Yin Z**. 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. Nature **435**:1122–1125.
19. **Song W-Y**, **Wang G-L**, **Chen L-L**, **Kim H-S**, **Pi L-Y**, **Holsten T**, **Gardner J**, **Wang B**, **Zhai W-X**, **Zhu L-H**, **Fauquet C**, **Ronald P**. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. Science (80-) **270**:1804–1806.
20. **Iyer AS**, **McCouch SR**. 2004. The rice bacterial blight resistance gene xa5 encodes a novel form of disease resistance. Mol Plant-Microbe Interact **17**:1348–1354.
21. **Blair MW**, **Garris AJ**, **Iyer AS**, **Chapman B**, **Kresovich S**, **McCouch SR**. 2003. High resolution genetic mapping and candidate gene identification at the xa5 locus for bacterial blight resistance in rice (Oryza sativa L.). Theor Appl Genet **107**:62–73.
22. **Wang GL**, **Ruan DL**, **Song WY**, **Sideris S**, **Chen L**, **Pi LY**, **Zhang S**, **Zhang Z**, **Fauquet C**, **Gaut BS**, **Whalen MC**, **Ronald PC**. 1998. Xa21D encodes a receptor-like molecular with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell **10**:765–780.
23. **Xiang Y**, **Cao Y**, **Xu C**, **Li X**, **Wang S**. 2006. Xa3, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as Xa26. Theor Appl Genet **113**:1347–1355.
24. **Khush GS**, **Angeles ER**. 1999. A new gene for resistance to race 6 of bacterial blight in rice, Oryza sativa L. Rice Genet Newsl **16**:92–93.
25. **Ogawa T**, **Lin L**, **Tabien RE**, **Khush GS**. 1987. A new recessive gene for resistance to bacterial blight of rice. Rice Genet Newsl **4**:98–100.
26. **He WA**, **Huang DH**, **Li RB**, **Qiu YF**, **Song JD**, **Yang HN**, **Zheng JX**, **Huang YY**, **Li XQ**, **Liu C**, **Zhang YX**, **Ma ZF**, **Yang Y**. 2012. Identification of a resistance gene bls1 to bacterial leaf streak in wild rice Oryza rufipogon Griff. J Integr Agric **11**:962–969.
27. **Triplett LR**, **Cohen SP**, **Heffelfinger C**, **Schmidt CL**, **Huerta AI**, **Tekete C**, **Verdier V**, **Bogdanove AJ**, **Leach JE**. 2016. A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. Plant J **87**:472–483.
28. **Zhao BY**, **Ardales E**, **Brasset E**, **Claflin LE**, **Leach JE**, **Hulbert SH**. 2004. The *Rxo1/ Rba1* locus of maize controls resistance reactions to pathogenic and non-host bacteria. Theor Appl Genet **109**:71–79.
29. **Zhao BY**, **Ardales EY**, **Raymundo A**, **Bai JF**, **Trick HN**, **Leach JE**, **Hulbert SH**. 2004. The *avrRxo1* gene from the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* confers a nonhost defense reaction on maize with resistance gene *Rxo1*. Mol Plant-Microbe Interact **17**:771–779.
30. **Boch J**, **Scholze H**, **Schornack S**, **Landgraf A**, **Hahn S**, **Kay S**, **Lahaye T**, **Nickstadt A**, **Bonas U**. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science (80-) **326**:1509–1512.
31. **Moscou MJ**, **Bogdanove AJ**. 2009. A simple cipher governs DNA recognition by TAL effectors. Science (80-) **326**:1501.

32. **Mak AN-S**, **Bradley P**, **Cernadas RA**, **Bogdanove AJ**, **Stoddard BL**. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science **335**:716–9.
33. **Hutin M**, **Pérez-Quintero AL**, **Lopez C**, **Szurek B**. 2015. MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. Front Plant Sci **6**:535.
34. **Korus K**, **Lang JM**, **Adesemoye AO**, **Block CC**, **Pal N**, **Leach JE**, **Jackson-Ziems TA**. 2017. First report of *Xanthomonas vasicola* causing bacterial leaf streak on corn in the United States. Plant Dis PDIS-10-16-1426-PDN.
35. **Peng Z**, **Hu Y**, **Xie J**, **Potnis N**, **Akhunova A**, **Jones J**, **Liu Z**, **White FF**, **Liu S**. 2016. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. BMC Genomics **17**:21.
36. **Wilkins KE**, **Booher NJ**, **Wang L**, **Bogdanove AJ**. 2015. TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. Front Plant Sci **6**:536.
37. **Booher NJ**, **Carpenter SCD**, **Sebra RP**, **Wang L**, **Salzberg SL**, **Leach JE**, **Bogdanove AJ**. 2015. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. Microb Genomics **1**.
38. **Quibod IL**, **Perez-Quintero A**, **Booher NJ**, **Dossa GS**, **Grande G**, **Szurek B**, **Vera Cruz C**, **Bogdanove AJ**, **Oliva R**. 2016. Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. Sci Rep **6**:34137.
39. **Bart R**, **Cohn M**, **Kassen A**, **McCallum EJ**, **Shybut M**, **Petriello A**, **Krasileva K**, **Dahlbeck D**, **Medina C**, **Alicai T**, **Kumar L**, **Moreira LM**, **Neto JR**, **Verdier V**, **Santana MA**, **Kositcharoenkul N**, **Vanderschuren H**, **Gruissem W**, **Bernal A**, **Staskawicz BJ**. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proc Natl Acad Sci USA **109**:E1972–E1979.
40. **Loman NJ**, **Constantinidou C**, **Chan JZM**, **Halachev M**, **Sergeant M**, **Penn CW**, **Robinson ER**, **Pallen MJ**. 2012. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. Nat Rev Microbiol **9**:599–606.
41. **Gevers D**, **Cohan FM**, **Lawrence JG**, **Spratt BG**, **Coenye T**, **Feil EJ**, **Stackebrandt E**, **de Peer Y Van**, **Vandamme P**, **Thompson FL**, **Swings J**. 2005. Opinion: Re-evaluating prokaryotic species. Nat Rev Microbiol **3**:733–739.
42. **Richter M**, **Rosselló-Móra R**. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A **106**:19126–31.
43. **Konstantinidis KT**, **Tiedje JM**. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A **102**:2567–72.
44. **Goris J**, **Konstantinidis KT**, **Klappenbach JA**, **Coenye T**, **Vandamme P**, **Tiedje JM**. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol **57**:81–91.
45. **Michelmore RW**, **Coaker G**, **Bart R**, **Beattie GA**, **Bent A**, **Bruce T**, **Cameron D**, **Dangl J**, **Dinesh-Kumar S**, **Edwards R**, **Eves-van den Akker S**, **Gassmann W**, **Greenberg J**, **Harrison R**, **He P**, **Harvey J**, **Huffaker A**, **Hulbert S**, **Innes R**, **Jones JD**, **Kaloshian I**, **Kamoun S**, **Katagiri F**, **Leach JE**, **Ma W**, **McDowell JM**, **Medford J**, **Meyers B**, **Nelson R**, **Oliver RP**, **Qi Y**, **Saunders D**, **Shaw M**, **Subudhi P**, **Torrance L**, **Tyler BM**, **Walsh J**. 2017. Foundational and translational research opportunities to improve plant health. Mol Plant-Microbe Interact MPMI-01-17-0010-CR.
46. **Bull CT**, **Koike ST**. 2015. Practical benefits of knowing the enemy: Modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. Annu Rev Phytopathol **53**:157–180.
47. **Vincelli P**, **Tisserat N**. 2008. Nucleic Acid–Based Pathogen Detection in Applied Plant Pathology. Plant Dis **92**:660–669.

48. **Langlois PA**, **Snelling J**, **Hamilton JP**, **Bragard C**, **Koebnik R**, **Verdier V**, **Triplett LR**, **Blom J**, **Tisserat NA**, **Leach JE**. 2017. Characterization of the *Xanthomonas translucens* Complex Using Draft Genomes, Comparative Genomics, Phylogenetic Analysis, and Diagnostic LAMP Assays. Phytopathology **107**:519–527.
49. **Ash GJ**, **Lang JM**, **Triplett LR**, **Stodart BJ**, **Verdier V**, **Cruz CV**, **Rott P**, **Leach JE**. 2014. Development of a genomics-based LAMP (Loop-Mediated Isothermal Amplification) assay for detection of *Pseudomonas fuscovaginae* from rice. Plant Dis **98**:909–915.
50. **Lang JM**, **Hamilton JP**, **Diaz MGQ**, **Van Sluys MA**, **Burgos MRG**, **Vera Cruz CM**, **Buell CR**, **Tisserat NA**, **Leach JE**. 2010. Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola.* Plant Dis **94**:311–319.
51. **Tomita N**, **Mori Y**, **Kanda H**, **Notomi T**. 2008. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nat Protoc **3**:877–82.
52. **Wilson IG**. 1997. Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol **63**:3741–51.
53. **Alvarez AM**. 2004. Integrated approaches for detection of plant pathogenic bacteria and diagnosis of bacterial diseases. Annu Rev Phytopathol **42**:339–366.

CHAPTER 2

SENSITIVE DETECTION OF *XANTHOMONAS ORYZAE PV. ORYZAE* AND *X. ORYZAE* PV. *ORYZICOLA* BY LOOP MEDIATED ISOTHERMAL AMPLIFICATION[1,2]

2.1 INTRODUCTION

Severe rice diseases, such as bacterial leaf streak (BLS) caused by *Xanthomonas oryzae* pv. *oryzicola* and bacterial blight (BB) caused by *X. oryzae* pv. *oryzae*, are increasing in prevalence in parts of Asia and sub-Saharan Africa and can cause average yield losses of 20 or 50%, respectively (1). Increased incidences of BLS and BB are considered to be the result of the introduction of new susceptible rice varieties, the intensification of cultivation, the absence of adequate phytosanitary controls, and environmental changes such as rising global temperatures (2, 3). Losses incurred by these diseases could jeopardize global food security.

Documenting the extent and distribution of BB and BLS is invaluable to understanding the severity of their threat on rice production. Seedborne dissemination of *X. oryzae* pv. *oryzicola* is a problem in parts of Asia and presumably in Africa (4). While clean seed and quarantine programs are prevalent in Asia, these are not yet developed in Africa. *X. o.* pv. *oryzae* has been detected in seed, but whether or not this form of transmission is important is still controversial (5–10).

High quality genome sequences of four strains of *X. oryzae* pv. *oryzae* and two strains of *X. oryzae* pv. *oryzicola* are publicly available (11–14; Genbank accession numbers PRJNA228925 and

[1] Published as "Sensitive Detection of *Xanthomonas oryzae* Pathovars oryzae and oryzicola by Loop-Mediated Isothermal Amplification" in Applied and Environmental Microbiology, 2014, 80(15) 4519-4530 by **J.M. Lang**, P. Langlois, M.H.R. Nguyen, L.R. Triplett, L. Purdie, T.A. Holton, A. Djikeng, C.M. Vera Cruz, V. Verdier and J.E. Leach.

[2] Contributions by J.M. Lang: Design of experiments; design and validation of primers; optimizing all assay conditions; wrote manuscript

PRJNA228927). These resources, along with draft genome sequences of another nine *X. oryzae* strains, provided insights into the genetic diversity among strains within this species, including a unique group of weakly pathogenic *X. oryzae* isolated in the United States ((13) and V. Verdier, personal communication).

In a previous study, we used a comparative genomics approach to develop diagnostic primers that distinguished strains by pathovar (*X. oryzae, X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola)*, and differentiated certain groups of strains based on their geographic origin (13, 15). Multi-locus sequence and RFLP analysis have shown that *X. oryzae* pv. oryzae is composed of two major genetic groups, the Asian and African lineages (16, 17). Pathovar-specific primers have been adopted for identification of *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* from field-collected leaf samples (4), and from seed samples (IRRI Seed Health Unit, personal communication). However, the adoption of these primers for field-level surveys or for routine screens of seed samples by quarantine officials has been limited largely due to the high costs and requirements for sophisticated laboratories to perform the available diagnostic assays.

A recent advance for molecular diagnostics is the adaptation of the loop mediated isothermal amplification (LAMP) method for rapid, specific amplification of target DNA sequences at a single temperature (18). Incubation can be accomplished using a simple water bath without the need for expensive equipment (19). LAMP can be more sensitive and less prone to inhibitors in test samples than PCR, and it can be adapted to a simple visual discrimination of the test result without requiring electrophoresis or other equipment (20). LAMP assays have been developed for phytoplasma, viral, bacterial and fungal plant pathogens as well as the detection of genetically modified crops (21–28). Visual assays in particular are ideally suited for deployment in non-specialized laboratories with limited equipment and resources, or for incorporation into a simple-to-use diagnostic test for use in the field. The increased sensitivity of the LAMP assay coupled with a closed tube system where no addition of DNA intercalating dye is necessary post reaction, is attractive for regulatory labs. LAMP can be used in epidemiological surveys, to support microbial forensic investigations for quarantine officials.

The intent of this project was to develop and evaluate LAMP assays for *X. oryzae* pathovars to enable surveillance activities in rice fields and testing of traded materials (seeds) in regional quarantine offices. We focused on genomic regions unique for *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* (15) to develop pathovar-specific LAMP primers that detect and differentiate strains of each pathovar. We show the effectiveness of these assays in detecting the pathovars in diverse sample preparations such as DNA, heat-killed cells or crude preparations from plant tissue. In addition, we used draft genomic comparisons to develop LAMP assays that distinguish African and Asian lineages of *X. oryzae* pv. *oryzae*.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial Strains, DNA and Plant Samples

The bacterial strains used in this study are listed in Table 2.1. Strains of *X. oryzae,* selected to represent the genetic and geographic diversity of the pathovars, were screened to determine assay specificity. These included 45 strains of *X. oryzae* pv. *oryzae*, 40 strains of *X. oryzae* pv. *oryzicola* and seven strains of a distinct group of *X. oryzae* isolated in the United States (13, 29). An additional 31 strains representing other plant pathogenic species and unknown bacteria isolated from rice tissue and seed were tested. Heat-killed cells, genomic DNA or crude plant exudate were used as template in LAMP reactions. Genomic DNA was isolated using either the Easy-DNA Kit (Life Technologies, Grand Island, NY) following the manufacturer's recommendations, or the DNeasy Blood and Tissue Kit (Qiagen, Inc., Valencia, CA) following the manufacturer's recommendations, except that DNA was eluted in 30 µl of water in the final step. All samples were diluted to 20 ng μl^{-1} in sterile water. Heat-killed cells were prepared from cultures grown for 24 h on peptone-sucrose agar (PSA) (30) at 28°C, diluted in sterile water to appropriate concentrations and incubated at 95°C for 10 min. Plant tissue was collected from rice plants at 0, 24, 48 and 72 h post inoculation (hpi) by syringe infiltration with either *X. oryzae* pv. *oryzae* $PXO99^A$ or *X. oryzae* pv. *oryzicola* BLS256 or MAI10 as previously reported (31). Each inoculum was adjusted to 0.2 OD_{600} diluted in distilled water (about 10^8 CFU ml^{-1}) prior to inoculation. Tissues were individually ground in a TissueLyser II (Qiagen, Inc., Valencia, CA) in one ml of distilled water and were

serially diluted. Diluted ground tissue was sampled from three independent leaves for testing in each appropriate assay and the experiment was repeated at least twice.

2.2.2 Primer Design and Screening

Loci PXO_00080 (conserved hypothetical protein) and Xoryp_010100019045 (putative glycosyltransferase) are unique for *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*, respectively (15). These loci were used to develop LAMP primers that amplify all isolates within each pathovar. In addition, loci that distinguish isolates of *X. oryzae* pv. *oryzae* by geographic origin were identified by analyzing draft genomic sequence of African strains of *X. oryzae* pv. *oryzae* (GenBank accession number PRJNA228925) (Fig. S1). The locus specific to Asian populations is PXO_03925 (conserved hypothetical protein; putative lipase). Primers were designed based on all of these unique sequences using either LAMP Designer Version 1.02 (Premier Biosoft, Palo Alto, CA) or PrimerExplorer (Eiken Chemical Company, https://primerexplorer.jp/e/) and synthesized by Integrated DNA Technologies (Coralville, IA). Four primers (external primers, F3 and B3; internal primers, FIP and BIP) were designed for each assay. Loop primers were also designed for pathovar-specific *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae* assays. All oligonucleotide sequences are listed in Table 2.

2.2.3 Loop Mediated Isothermal Amplification

LAMP reactions (final volume of 12 µl) were performed in a CFX Connect Real-Time System (BioRad, Hercules, CA) or a Genie II (Optigene, Sussex, United Kingdom). The reaction contained 7.2 µl Isothermal Master Mix (Optigene, Sussex, United Kingdom), 32 nM outer primers (F3 and B3) and 0.32 µM inner primers (FIP and BIP). Pathovar specific assays included 0.16 µM loop primers (LoopF and LoopB). Lastly, 1 µl of template was added that was either genomic DNA (20 ng μl^{-1}), heat-killed bacterial cells, or serially diluted, ground, inoculated tissue (as described above). African and Asian *X. oryzae* pv. *oryzae* assays did not include loop primers and the remaining volume was substituted with water. LAMP reactions in the CFX Connect were incubated for 60 min at 65°C, followed by melt curve analysis from 65°C to 95°C. Incubations on the Genie II were 30 or 60 min at 65°C. All LAMP assays for

screening purposes were replicated at least twice and all experiments included no template controls (water and no template DNA).

2.2.4 Assay Specificity and Sensitivity

Assay specificities were established using a pooling strategy to screen large collections of negative controls after initial specificity with positive control strains were confirmed. Positive controls were strains used to derive the published genome sequences (PXO99^{A}, MAFF311018 and KACC10331) for *X. oryzae* pv. *oryzae* and BLS256 for *X. oryzae* pv. *oryzicola* (11, 12, 14, 32). Non-target bacterial DNAs were pooled in equimolar concentrations, 10 strains per pool. Each negative pool was separately spiked with 1µl of positive control genomic DNA to validate detection in a mixed sample. Sensitivity of each assay was determined using serial dilutions of both genomic DNA (10 ng to 1 fg) and heat killed cells (10^{8} to 10^{1} CFU ml^{-1}). Initial *X. oryzae* pv. *oryzicola* assay development included loop primers, but subsequent testing for specificity and sensitivities removed these primers for greater consistency and to reduce the incidences of false positives. Volumes in each reaction were replaced with water.

2.2.5 Seed Detection

A lot of clean (known to be free of *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*) *Oryza sativa* cv. IR24 seeds was disinfected using 70% ethanol, then rinsed thrice with sterile distilled water and dried in a laminar flow hood. Subsamples of this lot were artificially inoculated by soaking the seeds in bacterial suspensions of *X. oryzae* pv. *oryzae* PXO99 and *X. oryzae* pv. *oryzicola* BLS256 for 2 h at room temperature, 2 h at 4 °C, and placed in a laminar flow hood until dry. PXO99 is a Philippine strain of *X. oryzae* pv. *oryzae*, and was used as a control strain in all experiments completed at the International Rice Research Institute; in experiments completed at Colorado State University, PXO99^{A}, a 5-azacytidine-resistant derivative of PXO99 (33), was used as a control strain. The remaining clean seeds were subdivided into 5 g seedlots (approximately 200 seeds). Cell counts were estimated to be 1.1 x 10^{4} CFU seed^{-1} (PXO99) and 4.6 x 10^{4} CFU seed^{-1} (BLS256). To test sensitivity of the pathovar specific LAMP assays to detect a 0.5% contamination of a 5 g seedlot (1 contaminated seed in 200 seeds), a single contaminated seed from the pool of PXO99 or BLS256 contaminated seeds was added to 5 g of clean

seed. Thirty samples each of PXO99-contaminated and BLS256-contaminated seedlots were prepared and processed using an extraction protocol for bacteria from rice seeds by sonication (Nguyen et al, unpublished). Seed extracts were stored at 0ºC. LAMP reactions were carried out as described above with 1 µl aliquots from seed extracts as template DNA. Each run included one positive DNA control, four non-target DNA controls, and one no template control in which 1 µl of water was added to the reaction mix. In experiments screening for *X oryzae* pv. *oryzae*, non-target controls included *X. oryzae* pv. *oryzicola* BLS175, *X. oryzae* pv. *oryzicola* BLS256, *Acidovorax avenae* BPJ4821, and an uncharacterized yellow non-pathogenic seed-associated bacterium named SHU199. *X oryzae* pv. *oryzae* PXO99^{A} genomic DNA was used as a positive control. Experiments using the *X. oryzae* pv. *oryzicola* pathovar specific primers included a positive control – *X. oryzae* pv. *oryzicola* BLS256, and four non-target controls – *X. oryzae* pv. *oryzae* PXO99, X. *oryzae* pv. *oryzae* PXO349, *Acidovorax avenae* BPJ4821 and SHU199. All DNA controls were normalized to 20 ng ul^{-1}. In analyzing the *X oryzae* pv. *oryzae* contaminated seedlots, pathovar loop primers (LoopF and LoopB) were used, but not in *X. oryzae* pv. *oryzicola* tests. Reactions were incubated in the Genie II (Optigene, Sussex, United Kingdom) at 65ºC for 60 min. All LAMP tests were conducted in triplicate. Sensitivity and specificity values were computed using the formulas discussed by Armitage et al (2002).

2.2.6 Visual Detection

A visual LAMP detection protocol was adapted for detection and identification of the *X. oryzae* pathovars to reduce cost and requirement for sophisticated equipment. Assays were performed in conventional thermal cyclers or a water bath at 65°C for 60 min. The 25 µl reaction mix contained 2.5 µl 10x Isothermal Amplification Buffer (New England Biolabs, Ipswich, MA), 1.4 mM dNTPs, 6 mM additional $MgSO_4$ for a final concentration of 8 mM (New England Biolabs, Ipswich, MA), 0.8 M Betaine (Sigma Aldrich, St. Louis, MO), 4 U Bst DNA polymerase large fragment or Bst DNA Polymerase 2.0 (New England Biolabs, Ipswich, MA), 0.32 µM of FIP and BIP, 32 nM of F3 and B3, and 0.16 µM LoopF and LoopB (loop primers in pathovar-specific assays only), with 1 µl of 20 ng µl^{-1} DNA, heat killed cells, or plant extract. Mineral oil (EMD Millipore, Darmstadt, Germany) was added on top of the

reaction mixture (20 µl) to minimize introduction of aerosolized product in workspaces. Amplification was terminated by heat inactivation at 80°C for 3 min. Post incubation, tubes were individually opened in a separate lab and 0.5 - 1µl of Quant-IT™ Pico Green® Reagent (Invitrogen, Carlsbad, CA, USA) was added. Reactions were incubated at room temperature for 5 min and then observed under normal and UV light for either a color change from orange to green or for fluorescence.

2.3 RESULTS

2.3.1 Primer design, specificity and sensitivity of LAMP assays

At least five different primer sets were predicted for each unique sequence, and were used to develop specific LAMP assays for each *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* (designated 'pathovar' primers), and for geographically distinct lineages of *X. oryzae* pv. *oryzae* (African vs Asian, designated as 'geographic' primers). After initial screening with control DNAs (*X. oryzae* pv. *oryzae* PXO99^{A}, BAI3, *X. oryzae* pv. *oryzicola* BLS256), primer sets listed in Table 2 were used for testing. Ratios of primer concentrations were based on previous reports and consultation with colleagues (2,4,6, Bühlmann, personal communication) and did not require optimization.

The pathovar and geographic assays were tested for specificity and efficiency with a panel comprising 44 *X. oryzae* pv. *oryzae,* 38 *X. oryzae* pv. *oryzicola,* seven *X. oryzae,* 11 *Xanthomonas* sp. (species unknown, but determined not to be *X. oryzae* by multiplex PCR) (15) and multiple sequence alignment (Cottyn et al, unpublished) and 19 strains representing eight different bacterial genera using the Isothermal Master Mix (Optigene, Sussex, United Kingdom) in a CFX Connect Real-Time System (BioRad, Hercules, CA) or a Genie II (Optigene, Sussex, United Kingdom) (Table 2.1).
Genomic DNA, diluted to concentrations ranging from 10 ng to 1 fg, was used to establish sensitivities of each assay. A no template control (water) was included in each experimental replication. Thresholds of detection were 10 pg for pathovar specific *X. oryzae* pv. *oryzae* (Fig. 2.1 A), 1 fg for pathovar-specific *X. oryzicola* pv. *oryzicola* (Fig. 2.1 B), 1 ng for African *X. oryzae* pv. *oryzae* (Fig. 2.1 D) and 1 pg for Asian *X. oryzae* pv. *oryzae* (Fig. 2.1 C). As few as 10^5 CFU ml^{-1} bacterial cells were detected using both pathovar-specific assays and the African *X. oryzae* pv. *oryzae* assay (Fig. 2.2 A, B, C). The Asian *X.*

oryzae pv. *oryzae* assay detected as few as 10^4 CFU/ml, though there was more variation in the technical replicates at the lower concentrations (Fig. 2.2 D). There were no false positives for any assay in the no template control tests confirming specificity of the target region and the primers. Loop primers designed for *X. oryzae* pv. *oryzicola* contributed to amplification of specific targets, however, false positive results were identified more often when they were included during initial validations. To address this concern in subsequent testing, loop primers were removed and the volume was replaced with water. This change in protocol reduced the sensitivity and therefore the incidence of false positive results. So, while these primers are reported in Table 2.2, we recommend conducting this assay without them. Overall, the adapted primers were specific and sensitive in the LAMP assay. In addition, using a comparative genomics approach with draft genome sequences, we identified loci that differentiated *X. oryzae* pv. *oryzae* lineages based on geographic origin.

2.3.2. Detection of X. oryzae pv. oryzae and X. oryzae pv. oryzicola from seed and crude plant extracts

Artificially inoculated seedlots were used to assess the capacity of the pathovar-specific LAMP assays to detect bacteria at a 0.5% contamination level in 5 g seed samples (Table 3). Using the *X. oryzae* pv. *oryzicola* pathovar-specific LAMP primers, BLS256 DNA was detected in 30 out of 30 samples, giving a sensitivity of 100%. None of the 30 seed samples contaminated with *X. oryzae* pv. *oryzae* PXO99 amplified with the *X. oryzae* pv. *oryzicola* pathovar LAMP primers (0% false positive detection). The sensitivity of the *X. oryzae* pv. *oryzae* pathovar LAMP primers was 93.3%, with the target strain PXO99 detected in 28 out of 30 contaminated seedlots. Five of 30 seed samples contaminated with *X. oryzae* pv. *oryzicola* BLS256 were detected as positive by the *X. oryzae* pv. *oryzae* LAMP primers (83.3% specificity), i.e., the non-target organism was detected in at least two of three technical replications. The trials were done with little or no optimization needed, but due to high sensitivity and robustness of the primers designed for the pathovar specific *X. oryzae* pv. *oryzicola*, the loop primers were excluded from the reaction mix to prevent random false positives. Representative amplification curves for each pathovar specific assay and appropriate controls are shown in Fig. 2.3 A and B for *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae,* respectively. Pairwise inoculations were used to

demonstrate consistent specificity in detection of the presence of either organism. As described above, a single contaminated seed of either *X. oryzae* pv. *oryzae* PXO99 or *X. oryzae* pv. *oryzicola* was added to 5 g of clean seed. The *X. oryzae* pv. *oryzae* PXO99 contaminated lots were tested with the *X. oryzae* pv. *oryzicola* pathovar specific assay and conversely, the *X. oryzae* pv. *oryzicola* BLS256 contaminated seed lots were tested with the *X. oryzae* pv. *oryzae* pathovar specific assay. Non-target DNAs from *Acidovorax avenae* (BPJ4821) or *Xanthomonas* sp. (SHU199) were included in each experiment as negative controls alongside no template controls and did not amplify with either assay (Table 2.1).

Amplifications in seed testing were delayed compared to when pure DNA was used as a template in specificity tests. The pathovar-specific *X. oryzae* pv. *oryzae* fluorescence values (RFU) detecting amplification were also higher compared to specificity and sensitivity assays. Seed detection experiments were conducted with a Genie II (Optigene, Sussex, United Kingdom) while sensitivity tests were completed on a CFX Connect Real-Time System (BioRad, Hercules, CA). We attribute the variation in these results to the two platforms used for detecting amplification as well as the nature of sample. Seed extracts may contain contaminants that slow the amplification reaction as was described in a watermelon system detecting cucumber mottle mosaic virus (35), and possibly influence fluorescence detection capabilities. However, data were evaluated for presence or absence of an exponential amplification as compared to negative controls, and for these tests, the amplifications were specific for primer/DNA combinations, and were consistent with all previous results.

Crude extracts from inoculated rice leaf tissue also served as viable templates for all of the reported LAMP assays (Table 2.4). Representative data from the pathovar-specific *X. oryzae* pv. *oryzicola* LAMP assay are illustrated in Fig. 2.4 and correlate with the threshold detected when using heat killed cells (starting concentration of 10^7 CFU ml^{-1}).

Interestingly, although viable bacteria could not be recovered, *X. oryzae* pv. *oryzae* strain $PXO99^A$ was correctly detected by both the *X. oryzae* pv. *oryzae* pathovar-specific and *X. oryzae* pv. *oryzae* Asian lineage assays in leaf samples that were inoculated 23 years ago and stored at room temperature. These samples did not amplify with the *X. oryzae* pv. *oryzicola* pathovar-specific or African

lineage *X. oryzae* pv. *oryzae* primers confirming that the assays are robust and can detect target bacteria in diverse sample preparations (data not shown).

2.3.3 Visual detection of LAMP products

A visual detection protocol was adapted and tested for all LAMP primers. Chemistries, including hydroxynapthol blue, Gel Red (Biotium, In., Hayward, CA), and ethidium bromide (data not shown) did not perform as reliably or clearly as the SYBR stain Quant-IT™ Pico Green® Reagent (Life Technologies, Grand Island, NY, USA) added post incubation. SYBR green was able to detect DNA directly in heat killed cells to the same threshold as the Isothermal Master Mix using a thermal cycler. SYBR green stained reactions are shown for the *X. oryzae* pv. *oryzae* pathovar specific assay in Fig. 2.5. A water bath was successfully used for incubation and crude inoculated plant extract amplified in each specific primer set designed (data not shown).

2.4 DISCUSSION

Adaptation of previously designed specific conventional PCR primers to LAMP resulted in a reliable, sensitive, specific and robust test to detect and differentiate *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. The primers and LAMP assays were validated on a wide diversity of bacterial strains, including a large collection of both *X. oryzae* pathovars as well as other *Xanthomonas* species and other genera of bacteria, to demonstrate primer specificity and assay reliability. The pathovar-specific *X. oryzae* pv. *oryzicola* primers provided the most sensitive assay. The Asian *X. oryzae* pv. *oryzae* primers, pathovar-specific *X. oryzae* pv. *oryzae* primers, and the African *X. oryzae* pv. *oryzae* primers were slightly less sensitive but still detected 1ng of genomic DNA. Differences in assay sensitivity are not likely due to copy number of the target, because, where sequences are known, all loci are present in single copies. Therefore, we predict that the inherent efficiency of each primer set in annealing causes this variation. Regardless, sensitivity thresholds among the four assays developed were consistent with those previously reported for other plant pathogenic bacteria ranging from 10 fg to 0.01 ng genomic DNA and 10^3 to 10^4 CFU ml^{-1} (22, 36–38) and correlate to the equivalent range of 10^3 to 10^6 genome copies based on the 5.2 Mbp *X. orzyae* pv. *oryzae* $PXO99^A$ and the 4.8 Mbp *X. orzyae* pv. *oryzicola* genomes (11, 12).

We previously used comparative genomic approaches with whole genome (15) and draft genome (13) sequences for the design of pathovar-specific primers. In this study, we mirrored this approach, and took advantage of compared draft genomes (authors, unpublished) that identified sequences differentiating geographically distinct populations of *X. oryzae pv. oryzae*, i.e., from Africa or Asia. The ability to differentiate geographic origin of a strain could facilitate epidemiological and surveillance studies along with the monitoring of imported seed at quarantine stations. The specificity for all primers after adaptation to LAMP-based assays was confirmed and no cross reactivity to other Xanthomonads or bacterial genera was detected (Table 2.1). Furthermore, no false positive amplification occurred in water (no template) controls. These results confirm the utility of draft sequence for development of unique primers and the streamlined adaptation of conventional PCR to LAMP. Because of their specificity and sensitivity, the primers and assays will be useful as forensic tools in quarantine offices, epidemiological studies and seed certification.

Using available public databases such as GenBank (https://www.ncbi.nlm.nih.gov/genbank/), the Pathosystems Resource Integration Center (PATRIC) (https://www.patricbrc.org/), and the Comprehensive Phytopathogen Genomics Resource (CPGR) (http://cpgr.plantbiology.msu.edu/), we searched for relationships of loci used for primer development with annotated genes in existing curated genomes. The *X. oryzae* pv. *oryzicola* specific locus (Xoryp_010100019045) was the only gene with available annotation. It encodes a predicted glucosyltransferase with sequence similarity to rhamnose-glucose polysaccharide assembly protein F50 (RgpF) of *Streptococcus mutans* and shows high similarity to a recently reported *X. oryzae* pv. *oryzicola* virulence factor, *wxocB* (39). This locus is flanked by a methyltransferase on one side and a cluster of ABC O-antigen lipopolysaccharide (LPS) transporters on the other side, consistent with clustering of LPS gene islands (40).

PXO_00080, unique to all *X. oryzae* pv. *oryzae* and absent in *X. oryzae* pv. *oryzicola*, encodes an uncharacterized protein. The *X. oryzae* pv. *oryzae* African locus was identified in draft sequence of NAI8, a strain originally from Niger (16), and bears no similarity to any publicly available sequence. Its presence was confirmed in draft genome sequence of MAI1 and BAI3. The *X. oryzae* pv. *oryzae* Asian

locus also is annotated as a hypothetical protein, but has similarity to the lipase family of proteins. *X. oryzae* pv. *oryzae* loci PXO_00080 have sequence similarity to plasmids found in two related species of *Xanthomonas* including *X. arboricola* pv. *pruni* CFBP 5530 plasmid pXap41 (FR875157) and *X. axonopodis* pv. *citri* 306 plasmid pXAC33 (AE008924). For example, four mismatches occur between locus PXO_00080 and the plasmid sequences in question, and two of our primers include those mismatches. While the presence of mismatches does not prove the primer would not amplify the non-target sequences, and we did not test the specific strains containing these plasmids, our use of a highly diverse and comprehensive panel of negative control strains for testing primers increases our confidence that the assays are specific for *X. oryzae* pv. *oryzae*. Furthermore, the likelihood of *X. arboricola* pv. *pruni* and *X. axonopodis* pv. *citri* being present on rice tissue or seed is unlikely.

Out of 90 assays (30 independently inoculated seedlots tested in triplicate), only five false positive results were generated using pathovar-specific *X. oryzae* pv. *oryzae* primers. This 83.3% error rate is consistent with previous reports for LAMP assays detecting bacteria (22). These data further emphasize the importance of reliable positive controls and replicated experiments to validate all results. For critical samples, further testing to confirm results using a multiplex PCR or individual PCRs reported in Lang et al 2010 (15) targeting different unique loci is recommended. False positives were detected more often when using the loop primers designed for the pathovar-specific *X. oryzae* pv. *oryzicola* assay. While in some assays loop primers play an important role in enhancing sensitivity, their removal can reduce the false positive rates, presumably by reducing the assay sensitivity, resulting in a more reliable and robust assay.

An advantage of LAMP assays over conventional PCR assays is that sophisticated equipment is not required for the LAMP protocols. We detected *X. oryzae* pv. *oryzicola* and *X. oryzae* pv. *oryzae* directly from leaf tissue ground in water (data not shown) using the SYBR green detection system demonstrating the feasibility of using the assay on field samples. Because the reaction is isothermal, a water bath or thermos with hot water could be used for reactions, simplifying the process for use in the field. We estimate the current cost for a single 12 µl reaction is 1 USD. The availability of improved

DNA polymerases was a key factor in improving the efficiency and accuracy of LAMP technology (41). New formulations of the enzyme, Bst 2.0, including one requiring a warm start (Bst 2.0 Warm Start, New England Biolabs, Ipswich, MA) have improved efficiency and multiplexing capabilities in custom developed reaction mixes. Because of the sensitivity of the LAMP assays, extreme care must be taken to avoid aerosolization of the amplified products. This is particularly important when using post reaction detection methods. We added mineral oil to the top of each reaction to avoid contamination when tubes are opened to add the detecting dye. Liang et al (42) recently developed a technology to separate DNA intercalating dye from the samples during incubation using a temperature sensitive wax. We did not test this approach, but it could be very useful for field-based LAMP assays.

In agriculture, the ability to work with crude extracts from plants and minimal equipment enhance the value of LAMP technology for rapid detection for etiological or epidemiological studies or for regulatory purposes. In this study, we demonstrate the utility with leaf and seed extracts. Most recently, Gurinder et al (25) reported detection of genetically modified organisms using LAMP based on promoter sequences. Researchers will be able to quickly determine sources of contaminated material or outbreaks in the field using LAMP. In future work, surveys conducted using the reported LAMP assays will provide a reliable picture of the presence of these diseases across a region or country.

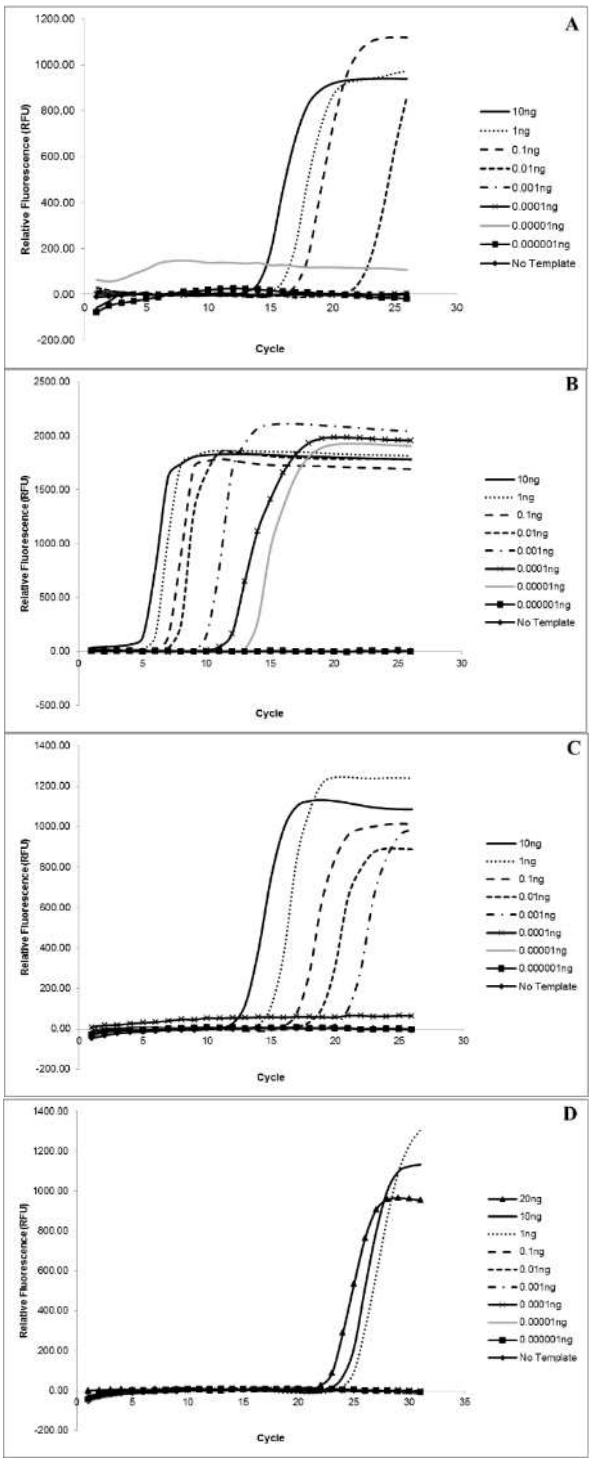

Figure 2.1 Pathovar-specific *X. oryzae* pv. *oryzae* (A), pathovar-specific *X. oryzae* pv. *oryzicola* (B), African *X. oryzae* pv. *oryzae* (C) and Asian *X. oryzae* pv. *oryzae* (D) LAMP assay sensitivity tests. Each diluted DNA was tested at least three times.

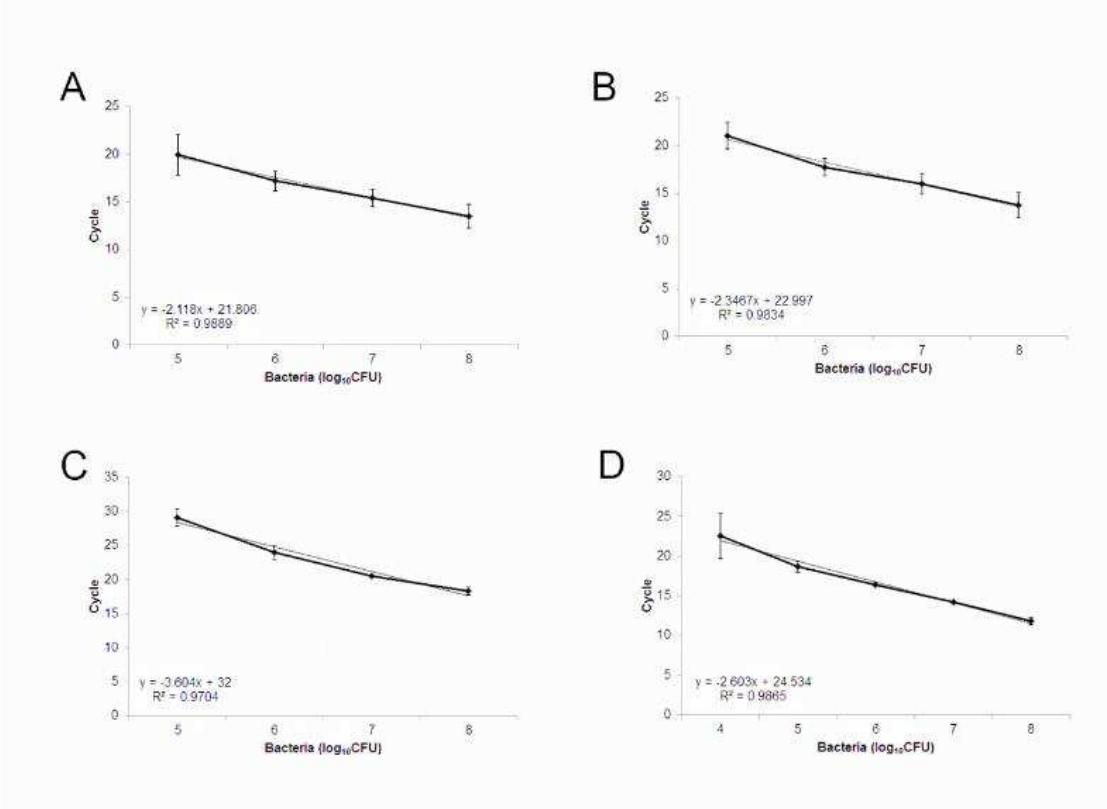

Figure 2.2 Standard curves showing sensitivity of detection by LAMP using dilutions of heat killed cells. Pathovar-specific *X. oryzae* pv. *oryzae* primers with *X. oryzae* pv. *oryzae* PXO99^{A} heat killed cells (A), pathovar-specific *X. oryzae* pv. *oryzicola* primers with *X. oryzae* pv. *oryzicola* BLS256 cells (B), African *X. oryzae* pv. *oryzae* primers using *X. oryzae* pv. *oryzae* BAI3 cells (C) and Asian *X. oryzae* pv. *oryzae* using *X. oryzae* pv. *oryzae* PXO99^{A} cells (D). Each dilution was tested three times. Bars represent +/- standard deviation of the mean and associated R^2 value after linear regression analysis.

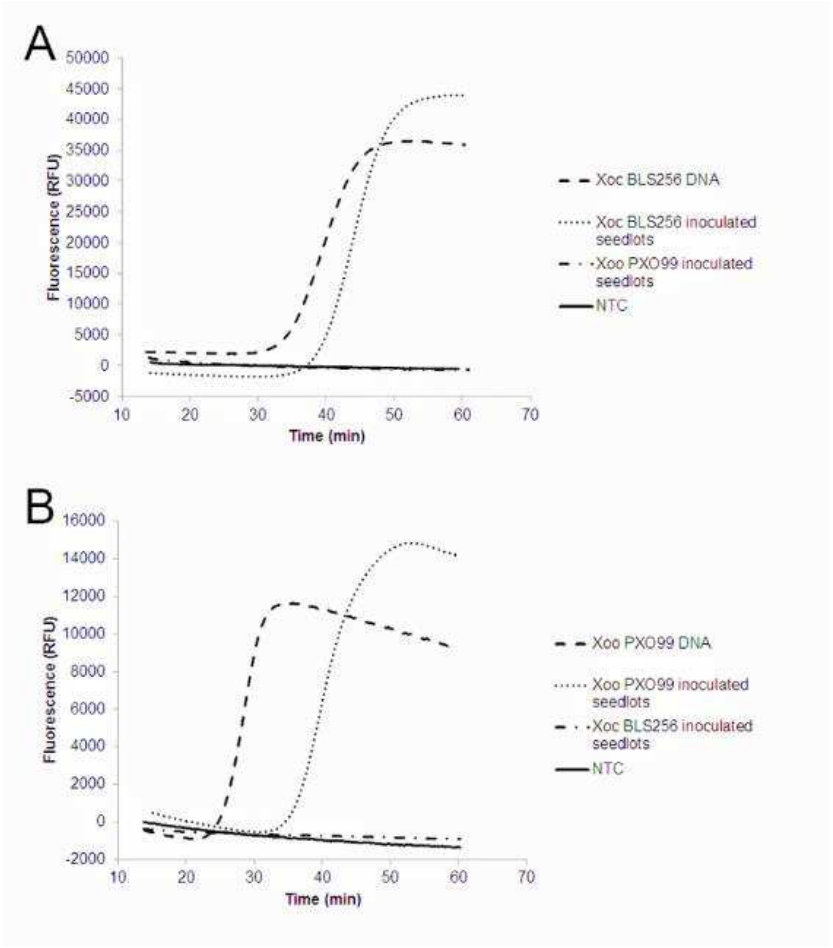

Figure 2.3 LAMP detection of *X. oryzae* pv. *oryzicola* (A) and *X. oryzae* pv. *oryzae* (B) using pathovar specific assays and seed extracts of 5 g seedlots with 0.5% contamination (1 seed carrying 10^4 CFU out of the 200 seeds) and from DNA extracts (20 ng μl^{-1} in TE buffer). Mean fluorescence data for seed extracts were obtained from the amplification curves of 10 seedlots per isolate against each primer set. *X. oryzae* pv. *oryzae* (*Xoo*) PXO99 DNA and *X. oryzae* pv. *oryzicola* (*Xoc*) BLS256 DNA, no template controls (NTC) consisting of water and non-target DNAs served as negative controls and failed to amplify in any run with either primer set in at least three independent runs. Data were collected on a Genie II (Optigene, Sussex, United Kingdom) and normalized to background fluorescence.

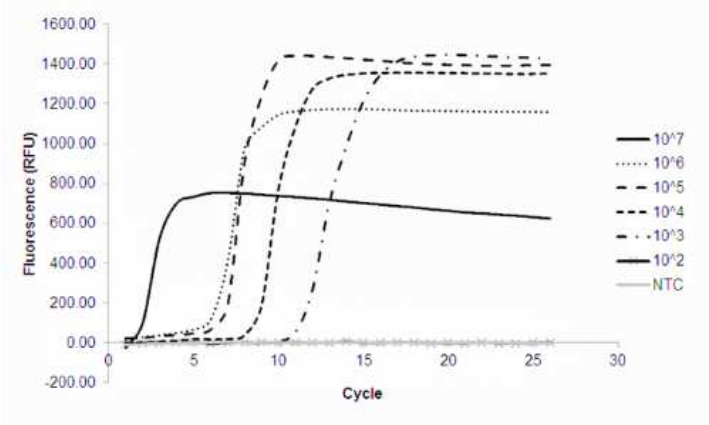

Figure 2.4 Sensitivity of pathovar-specific *X. oryzae* pv. *oryzicola* primers in a LAMP assay using ground tissue inoculated with *X. oryzae* pv. *oryzicola* MAI10, sampled at 48 hpi and then serially diluted.

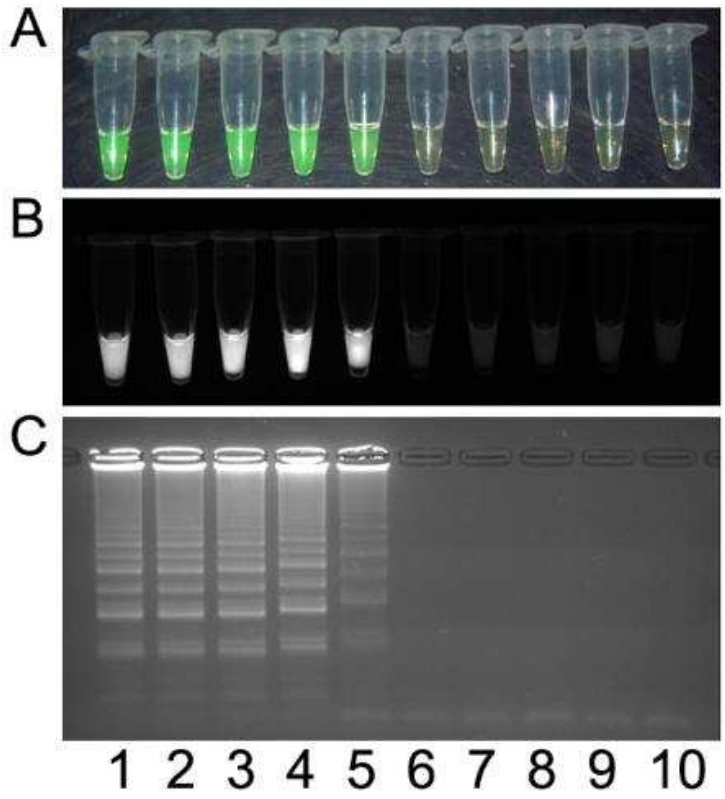

Figure 2.5 Visual detection of *X. oryzae* pv. *oryzae* PXO99^{A} heat killed cells using the pathovar-specific LAMP assay. A dilution series was tested consisting of 10^{8} (1), 10^{7} (2), 10^{6} (3), 10^{5} (4), 10^{4} (5), 10^{3} (6), 10^{2} (7), 10^{1} CFU ml^{-1} (8), *X. oryzae* pv. *oryzicola* BLS256 (9) and a no template control (10). Products were detected using 1µl Quant-IT™ Pico Green® Reagent (Life Technologies, Grand Island, NY, USA) under visual light (A), where a positive result changes from orange to green; ultra-violet light (B) where a positive result fluoresces or by 1.5% agarose gel electrophoresis (C), where a positive result is a laddered product.

Table 2.1. Bacterial strains used in this study to validate specificity of each assay[a,b,c].

Species	Strain	Origin	Host	Source	Xoc LAMP	Xoo LAMP	African Xoo	Asian Xoo
Xanthomonas oryzae pv. *oryzae*	R-3	Australia	*Oryza sativa*	I. Buddenhagen	-	+	-	+
X. o. pv. *oryzae*	BAI3	Burkina Faso	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	BAI4	Burkina Faso	*O. glaberrima*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	BAI1	Burkina Faso	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	BAI2	Burkina Faso	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	CFBP1948	Cameroon	*O. sativa*	J.L. Notteghem	-	+	+	-
X. o. pv. *oryzae*	CFBP1952	Mali	*O. sativa*	J.L. Notteghem	-	+	+	-
X. o. pv. *oryzae*	94 (35)	China	*O. sativa*		-	+	-	+
X. o. pv. *oryzae*	A3857	India	*O. sativa*	J.E. Leach	-	+	-	+
X. o. pv. *oryzae*	IXO16	Indonesia	*O. sativa*		-	+	-	+
X. o. pv. *oryzae*	MAFF311018	Japan	*O. sativa*	A. Bogdanove	-	+	-	+
X. o. pv. *oryzae*	KACC10331	Korea	*O. sativa*		-	+	-	+
X. o. pv. *oryzae*	Xoo197	Korea	*O. sativa*	S.H. Choi	-	+	-	+
X. o. pv. *oryzae*	Xoo199	Korea	*O. sativa*	S.H. Choi	-	+	-	+
X. o. pv. *oryzae*	MXO90	Malaysia	*O. sativa*		-	+	-	+
X. o. pv. *oryzae*	MXO92	Malaysia	*O. sativa*		-	+	-	+
X. o. pv. *oryzae*	CFBP1949	Mali	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	CFBP1951	Mali	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	MAI1	Mali	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	MAI9	Mali	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	R-13	Myanmar	*O. sativa*	I. Buddenhagen	-	+	-	+
X. o. pv. *oryzae*	NXO537	Nepal	*O. sativa*	Tika Adhikari	-	+	-	+
X. o. pv. *oryzae*	NXO544	Nepal	*O. sativa*	Tika Adhikari	-	+	-	+
X. o. pv. *oryzae*	NXO588	Nepal	*O. sativa*	Tika Adhikari	-	+	-	+
X. o. pv. *oryzae*	NAI2	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	NAI5	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	NAI6	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	NAI7	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	NAI8	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	NAI9	Niger	*O. sativa*	V. Verdier	-	+	+	-
X. o. pv. *oryzae*	PXO69	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO99A	Philippines	*O. sativa*	A. Bogdanove	-	+	-	+
X. o. pv. *oryzae*	PXO86	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO111	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO121	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO130	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO132	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO172	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO183	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+
X. o. pv. *oryzae*	PXO344	Philippines	*O. sativa*	C.M. Vera Cruz	-	+	-	+

				I.				
X. o. pv. *oryzae*	CL-1	Sri Lanka	*O. sativa*	Buddenhagen	-	+	-	+
X. o. pv. *oryzae*	Xoo1	Thailand	*O. sativa*	J.E. Leach	-	+	-	+
X. o. pv. *oryzae*	Xoo3	Thailand	*O. sativa*	J.E. Leach	-	+	-	+
X. o. pv. *oryzae*	RS61	China	*O. sativa*	J.S. Wang	-	+	-	+
X. o. pv. *oryzicola*	RS85	China	*O. sativa*	J.S. Wang	+	-	-	x
X. o. pv. *oryzicola*	CFBP2286	Malaysia	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI4	Mali	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI5	Mali	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI6	Mali	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI7	Mali	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI8	Mali	*O. sativa*	V. Verdier	+	-	-	-
X. o. pv. *oryzicola*	MAI10	Mali	*O. sativa*	V. Verdier	+	-	+	-
X. o. pv. *oryzicola*	BLS46	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS96	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS98	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	-
X. o. pv. *oryzicola*	BLS105	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS106	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS111	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS114	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS123	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS125	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS170	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS175	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS179	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS220	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS256	Philippines	*O. sativa*	A. Bogdanove	+	-	-	x
X. o. pv. *oryzicola*	BLS276	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS281	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS291	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS294	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS333	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS346	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-

X. o. pv. *oryzicola*	BLS354	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS356	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS377	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	-
X. o. pv. *oryzicola*	BLS404	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	-
X. o. pv. *oryzicola*	BLS413	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS415	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. o. pv. *oryzicola*	BLS417	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	-
X. o. pv. *oryzicola*	BLS420	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	-
X. o. pv. *oryzicola*	BLS421	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	x	x
X. o. pv. *oryzicola*	BLS468	Philippines	*O. sativa*	C.M. Vera Cruz	+	-	-	x
X. oryzae	X1-8	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	X8-1A	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	X211-2	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	X4-2C	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	RU87-17	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	X11-5A	USA	*O. sativa*	C. Gonzalez	-	+	-	-
X. oryzae	X1-10	USA	*O. sativa*	C. Gonzalez	-	+	-	-
Xanthomonas sp.	97M	Philippines	*O. sativa*	V. Verdier	-	-	-	-
Xanthomonas sp.	M136	Mali	*O. sativa*	V. Verdier	-	-	-	-
Xanthomonas sp.	SHU36	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU50	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU100	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU147	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU178	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU199	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	x	x
Xanthomonas sp.	SHU202	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU222	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU268	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	-	-
Xanthomonas sp.	SHU303	Philippines	*O. sativa,*seed	C.M. Vera Cruz	-	-	x	-
Acidovorax avenae pv. *avenae*	NCPPB3112	Brazil	*Canna indica*	NCPPB[d]	-	-	-	x
A. avenae	BPJ4821	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	x	x
A. a. pv. *citrulli*	94-21	USA	*Citrullus lanatus*	R. Walcott	-	-	-	-
Burkholderia andropogonis	3549	USA	*Saccharum officinarum*	L.E. Claflin	-	-	-	-
B. gladioli	O187	USA	*Allium cepa*	H.F. Schwartz	-	-	-	-
Curtobacterium flaccumfaciens	B473	USA	*Phaseolous* sp.	H.F. Schwartz	-	-	-	-

Escherichia coli	DH5α	USA	NA	Life Technologies	-	-	-	-
Enterobacter sp.	O121	USA	*Allium cepa*	H.F. Schwartz	-	-	-	-
Enterobacter sp.	O174	USA	*Allium cepa*	H.F. Schwartz	-	-	-	-
Pseudomonas marginalis	ATCC10844	USA	*Lactuca sativa*	H.F. Schwartz	-	-	-	-
P. syringae pv. *syringae*	M72	USA	*Capsicum annuum*	H.F. Schwartz	-	-	-	-
P. syringae pv. *syringae*	M108	USA	*Solanum lycopersicum*	H.F. Schwartz	-	-	x	-
P. viridiflava	ATCC13223	USA	*Phaseolus coccineus*	H.F. Schwartz	-	-	x	-
Ralstonia solanacearum	K60	USA	*S. lycopersicum*	J.E. Leach	-	-	-	x
X. axonopodis pv. *vesicatoria*	85-10	USA	*Capsicum* sp.	A. Bogdanove	-	-	x	-
X. a. pv. *sojense*	4455	USA	*Glycine max*	L.E. Claflin	-	-	-	x
X. campestris pv. *campestris*	X1910	USA	*Brassica oleracea*	N. Dunlop	-	-	x	-
X. c. pv. carotae	NCPPB1422	New Zealand	*Daucus carota*	L.E. Claflin	-	-	-	x
X. translucens pv. cerealis	NCPPB1836	USA	*Secale cereale*	L.E. Claflin	-	-	-	-
Unknown	Ven	Venezuela	*O. sativa,*seed	V. Verdier	-	-	-	-

[a]+ represents a positive amplification

[b]– represents absence of amplification

[c]x designates strain not tested

[d]National Collection of Plant Pathogenic Bacteria

Table 2.2. LAMP primers for detection of *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*

Target	Primer	Sequence (5'-3')
X. o. pv. *oryzae* PXO_00080[a]	F3	CTTCAAGGCCAAGGACATC
	B3	CACGATCTTGCAAGGGAT
	FIP	CGGTGCCGGACTGGATTTGCTAGGAATGAGCAATGCA
	BIP	GTAGTTGCCGACGGCTACCAGAAGCGTCCTCGTCTAA
	LoopF	TTTGAGGTCCCTTTCCACG
	LoopB	GTTTGTGCGCCGTCTATC
X. o. pv. *oryzicola* Xoryp_010100019045[b]	F3	GGATCACAGTGATCGTGC
	B3	CACTTATCGTCCAGTACGC
	FIP	CGATGCCGCCTTGATCGAGTTGTACTCCTACGATGAGC
	BIP	ACCGAGTCGTTGCAGGTCTCTTGCGAAACACAAGGAA
	LoopF	TTGTGACCACGCTGTCATT
	LoopB	TCGCCATCTCCAGTCCTAT
African *X. o.* pv. *oryzae* Hypothetical protein, draft sequence[c]	F3	TATTGGGTGCTGCCGATGA
	B3	GGCAACCTCACTTCCGTAAG
	FIP	ATGTAGCCATCATGCCCGCCTTTTCCCAGATTTGCGAGTCCTT
	BIP	GCGCTCTTCGGATGGTAGTGATTTTTGCCATGCTTGTTTTGTGCA
Asian *X. o.* pv. *oryzae* PXO_03925[a]	F3	GGTGGTCAGCGCATCGA
	B3	ACTGCTGCTGTTCCAACG
	FIP	ATGCTGACGCGCAGCTGCTTTTAGCCCAGATGCCGCAC
	BIP	TGACGGCAATGAATACCCCGCTTTTGACCGACTGGCGTGCT

[a]Design based on *X. oryzae* pv. *oryzae* PXO99A (PRJNA28127) genome sequence
[b]Design based on *X. oryzae* pv. *oryzicola* BLS256 (PRJNA54411) genome sequence
[c]Design based on *X. oryzae* pv. *oryzae* NAI8 (PRJNA228925) and BAI3 (Verdier et al, unpublished) draft genome sequences

Table 2.3. LAMP detection of *X. oryzae* pv. *oyrzae* and *X. oryzae* pv. *oryzicola* using pathovar-specific assays in seedlots (200 seeds) inoculated with one seed carrying 10^4 CFU of *X. oryzae* pv. *oryzae* (*Xoo* PXO99) or *X. oryzae* pv. *oryzicola* (*Xoc* BLS256) (0.5% contamination). Each sample was analyzed in triplicate against each primer set where presence of an amplification curve at least two out of three runs was interpreted as a positive result.

Samples	Amplification from seed with pathovar-specific assays			
	Pathovar-specific *X. oryzae* pv. *oryzicola* LAMP assay		**Pathovar-specific *X. oryzae* pv. *oryzae* LAMP assay**	
	Xoc BLS256	*Xoo* PXO99	*Xoo* PXO99	*Xoc* BLS256
1	3/3	0/3	2/3	2/3
2	3/3	0/3	3/3	0/3
3	3/3	0/3	3/3	0/3
4	3/3	0/3	3/3	0/3
5	3/3	0/3	3/3	2/3
6	3/3	0/3	2/3	0/3
7	3/3	0/3	1/3	1/3
8	3/3	0/3	3/3	0/3
9	3/3	0/3	0/3	1/3
10	3/3	0/3	2/3	0/3
11	3/3	0/3	3/3	0/3
12	3/3	0/3	3/3	2/3
13	3/3	0/3	3/3	0/3
14	3/3	0/3	3/3	3/3
15	3/3	0/3	3/3	0/3
16	3/3	0/3	3/3	0/3
17	3/3	0/3	3/3	0/3
18	3/3	0/3	3/3	0/3
19	3/3	0/3	3/3	0/3
20	3/3	0/3	3/3	0/3
21	3/3	0/3	3/3	0/3
22	3/3	1/3	2/3	0/3
23	3/3	0/3	2/3	0/3
24	3/3	0/3	3/3	3/3
25	3/3	0/3	3/3	0/3
26	3/3	1/3	3/3	0/3
27	3/3	0/3	2/3	1/3
28	3/3	1/3	3/3	0/3
29	3/3	0/3	2/3	0/3
30	3/3	0/3	3/3	0/3
BPJ4821[a]	0/9	0/9	0/9	0/9
SHU199[a]	0/9	0/9	0/9	0/9

Water (NTC)	0/9	1/9	0/9	0/9
Total positives[b]	30	0	28	5
Total negatives[b]	0	30	2	25
Sensitivity	100.00		93.33	
Specificity	100.00		83.33	

[a]Genomic DNA from bacteria isolated from seeds (strains BPJ4829 and SHU199) were included as negative controls and are referenced in Table 1.

[b]Total positives and negatives refer to inoculated seed preparations, negative and no template controls were not included in this sum or for sensitivity and specificity calculations.

Table 2.4. LAMP detection of *X. oryzae* pv. *oyrzae* and *X. oryzae* pv. *oryzicola* in leaf tissue inoculated by syringe infiltration 72 hpi. Three independently inoculated leaves were analyzed at least twice.

Strain[a]	**Number of positives/total samples amplified with primer set:**			
	Pathovar-specific *X. oryzae* pv. *oryzae*	Pathovar-specific *X. oryzae* pv. *oryzicola*	African lineage *X. oryzae* pv. *oryzae*	Asian lineage *X. oryzae* pv. *oryzae*
PXO99[A]	3/3	0/3	0/3	3/3
MAI1	3/3	0/3	3/3	0/3
BLS256	0/3	3/3	0/3	0/3
Uninoculated leaf	0/3	0/3	0/3	0/3
No template[b]	-	-	-	-

[a]Strains are referenced in Table 1.

[b]No template controls contained water and were included in each reaction. A "–" represents the absence of any amplification signal.

REFERENCES

1. **Ou SH**. 1985. Rice disease. Ed 2 The Cambrian News Ltd, UK.
2. **Verdier V**, **Vera Cruz C**, **Leach JE**. 2012. Controlling rice bacterial blight in Africa: Needs and prospects. J Biotechnol **159**:320–328.
3. **Garrett K**, **Dendy S**, **Frank E**, **Rouse M**, **Travers S**. 2006. Climate change effects on plant disease: genomes to ecosystems. Annu Rev Phytopathol **44**:489–509.
4. **Wonni I**, **Cottyn B**, **Detemmerman L**, **Dao S**, **Ouedraogo L**, **Sarra S**, **Tekete C**, **Poussier S**, **Corral R**, **Triplett L**, **Koita O**, **Koebnik R**, **Leach J**, **Szurek B**, **Maes M**, **Verdier V**. 2013. Analysis of *Xanthomonas oryzae* pv. *oryzicola* population in Mali and Burkina Faso reveals a high level of genetic and pathogenic diversity. Phytopathology doi:10.1094/PHYTO-07-13-0213-R.
5. **Sakthivel N**, **Mortensen C**, **Mathur S**. 2001. Detection of *Xanthomonas oryzae* pv. *oryzae* in artificially inoculated and naturally infected rice seeds and plants by molecular techniques. Appl Microbiol Biotechnol **56**:435–441.
6. **Kauffman HE**, **Reddy APK**. 1975. Seed transmission studies of *Xanthomonas oryzae* in rice. Phytopathology **65**:663–666.
7. **Mew TW**, **Misra JK**. 1994. A Manual of Rice Seed Health Testing. International Rice Research Institute, Los Banos.
8. **Devadath S**, **Thri Murty VS**. 1984. Role of seed in survival and transmission of *Xanthomonas campestris* pv. *oryzae* causing bacterial blight of rice. J Phytopathol **110**:15–19.
9. **Gnanamanickam SS**, **Shigaki T**, **Medalla ES**, **Mew TW**, **Alvarez AM**. 1994. Problems in detection of *Xanthomonas oryzae* pv. *oryzae* in rice seed and potential for improvement using monoclonal antibodies. Plant Dis **78**:173–178.
10. **Xie GL**, **Mew TW**. 1998. A leaf inoculation method for detection of *Xanthomonas oryzae* pv. *oryzicola* from rice seed. Plant Dis **82**:1007–1011.
11. **Salzberg SL**, **Sommer DD**, **Schatz MC**, **Phillippy AM**, **Rabinowicz PD**, **Tsuge S**, **Furutani A**, **Ochiai H**, **Delcher AL**, **Kelley D**, **Madupu R**, **Puiu D**, **Radune D**, **Shumway M**, **Trapnell C**, **Aparna G**, **Jha G**, **Pandey A**, **Patil PB**, **Ishihara H**, **Meyer DF**, **Szurek B**, **Verdier V**, **Koebnik R**, **Dow JM**, **Ryan RP**, **Hirata H**, **Tsuyumu S**, **Won Lee S**, **Seo Y-S**, **Sriariyanum M**, **Ronald PC**, **Sonti R V**, **Van Sluys M-A**, **Leach JE**, **White FF**, **Bogdanove AJ**. 2008. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics **9**:204.
12. **Bogdanove AJ**, **Koebnik R**, **Lu H**, **Furutani A**, **Angiuoli S V**, **Patil PB**, **Van Sluys MA**, **Ryan RP**, **Meyer DF**, **Han SW**, **Aparna G**, **Rajaram M**, **Delcher AL**, **Phillippy AM**, **Puiu D**, **Schatz MC**, **Shumway M**, **Sommer DD**, **Trapnell C**, **Benahmed F**, **Dimitrov G**, **Madupu R**, **Radune D**, **Sullivan S**, **Jha G**, **Ishihara H**, **Lee SW**, **Pandey A**, **Sharma V**, **Sriariyanun M**, **Szurek B**, **Vera-Cruz CM**, **Dorman KS**, **Ronald PC**, **Verdier V**, **Dow JM**, **Sonti R V**, **Tsuge S**, **Brendel VP**, **Rabinowicz PD**, **Leach JE**, **White FF**, **Salzberg SL**. 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. J Bacteriol2011/07/26. **193**:5450–5464.
13. **Triplett LR**, **Hamilton JP**, **Buell CR**, **Tisserat NA**, **Verdier V**, **Zink F**, **Leach JE**. 2011. Genomic analysis of *Xanthomonas oryzae* isolates from rice grown in the United States reveals substantial divergence from known *X. oryzae* pathovars. Appl Environ Microbiol **77**:3930–3937.
14. **Lee B-M**, **Park Y-J**, **Park D-S**, **Kang H-W**, **Kim J-G**, **Song E-S**, **Park I-C**, **Yoon U-H**, **Hahn J-H**, **Koo B-S**, **Lee G-B**, **Kim H**, **Park H-S**, **Yoon K-O**, **Kim J-H**, **Jung C**, **Koh N-H**, **Seo J-S**, **Go S-J**. 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res **33**:577–586.
15. **Lang JM**, **Hamilton JP**, **Diaz MGQ**, **Van Sluys MA**, **Burgos MRG**, **Vera Cruz CM**, **Buell CR**, **Tisserat NA**, **Leach JE**. 2010. Genomics-based diagnostic marker development for *Xanthomonas*

oryzae pv. *oryzae* and *X. oryzae* pv. *oryzicola*. Plant Dis **94**:311–319.
16. **Gonzalez C**, **Szurek B**, **Manceau C**, **Mathieu T**, **Sere Y**, **Verdier V**. 2007. Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. Mol Plant-Microbe Interact2007/05/18. **20**:534–546.
17. **Hajri A**, **Brin C**, **Zhao S**, **David P**, **Feng J**, **Koebnik R**, **Szurek B**, **Verdier V**, **Boureau T**, **Poussier S**. 2012. Multilocus sequence analysis and type III effector repertoire mining provide new insights into the evolutionary history and virulence of *Xanthomonas oryzae*. Mol Plant Pathol **13**:288–302.
18. **Notomi T**, **Okayama H**, **Masubuchi H**, **Yonekawa T**, **Watanabe K**, **Amino N**, **Hase T**. 2000. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res **28**:e63.
19. **Tomlinson J**. 2013. In-Field Diagnostics Using Loop-Mediated Isothermal Amplification, p. 291–300. *In* Dickinson, M, Hodgetts, J (eds.), Phytoplasma. Humana Press.
20. **Kaneko H**, **Kawana T**, **Fukushima E**, **Suzutani T**. 2007. Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. J Biochem Biophys Methods **70**:499–501.
21. **Bühlmann A**, **Pothier JF**, **Tomlinson JA**, **Frey JE**, **Boonham N**, **Smits THM**, **Duffy B**. 2013. Genomics-informed design of loop-mediated isothermal amplification for detection of phytopathogenic *Xanthomonas arboricola* pv. *pruni* at the intraspecific level. Plant Pathol **62**:475–484.
22. **Bühlmann A**, **Pothier JF**, **Rezzonico F**, **Smits THM**, **Andreou M**, **Boonham N**, **Duffy B**, **Frey JE**. 2013. *Erwinia amylovora* loop-mediated isothermal amplification (LAMP) assay for rapid pathogen detection and on-site diagnosis of fire blight. J Microbiol Methods **92**:332–9.
23. **Villari C**, **Tomlinson JA**, **Battisti A**, **Boonham N**, **Capretti P**, **Faccoli M**. 2013. Use of loop-mediated isothermal amplification for detection of *Ophiostoma clavatum*, the primary blue stain fungus associated with *Ips acuminatus*. Appl Environ Microbiol **79**:2527–33.
24. **Wei Q-W**, **Yu C**, **Zhang S-Y**, **Yang C-Y**, **Miriam K**, **Zhang W-N**, **Dou D-L**, **Tao X-R**. 2012. One-step detection of Bean pod mottle virus in soybean seeds by the reverse-transcription loop-mediated isothermal amplification. Virol J **9**:187.
25. **Randhawa GJ**, **Singh M**, **Morisset D**, **Sood P**, **Zel J**. 2013. Loop-mediated isothermal amplification: Rapid visual and real-time methods for detection of genetically modified crops. J Agric Food Chem **61**:11338–11346.
26. **Dharmapuri S**, **Yashitola J**, **Vishnupriya MR**, **Sonti R V**. 2001. Novel genomic locus with atypical G+C content that is required for extracellular polysaccharide production and virulence in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant-Microbe Interact **14**:1335–1339.
27. **Yasuhara-Bell J**, **Kubota R**, **Jenkins DM**, **Alvarez AM**. 2013. Loop-Mediated amplification of the *Clavibacter michiganensis* subsp. *michiganensis* micA gene is highly specific. Phytopathology **103**:1220–1226.
28. **Ash GJ**, **Lang JM**, **Triplett LR**, **Stodart BJ**, **Verdier V**, **Cruz CV**, **Rott P**, **Leach JE**. 2014. Development of a genomics-based LAMP (Loop-Mediated Isothermal Amplification) assay for detection of *Pseudomonas fuscovaginae* from rice. Plant Dis **98**:909–915.
29. **Jones RK**, **Barnes LW**, **Gonzalez CF**, **Leach JE**, **Alvarez AM**, **Benedict AA**. 1989. Identification of low virulence strains of *Xanthomonas campestris* pv. *oryzae* from rice in the United States. Phytopathology **79**:984–990.
30. **Karganilla A**, **Paris-Natural M**, **Ou SH**. 1973. A comparative study of culture media for *Xanthomonas oryzae*. PhilippAgric **57**:141–152.
31. **Verdier V**, **Triplett LR**, **Hummel AW**, **Corral R**, **Cernadas RA**, **Schmidt CL**, **Bogdanove AJ**, **Leach JE**. 2012. Transcription activator-like (TAL) effectors targeting *OsSWEET* genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. New Phytol2012/10/20. **196**:1197–1207.
32. **Ochiai H**, **Inoue V**, **Takeya M**, **Sasaki A**, **Kaku H**. 2005. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences

to its race diversity. Japan Agric Res Q **39**:275–287.
33. **Mew TW**, **Vera C. M. C**, **Medalla ES**. 1992. Changes in race frequency of *Xanthomonas oryzae* pv. *oryzae* in response to rice cultivars planted in the Philippines. Plant Dis **76**:1029.
34. **Gosch C**, **Gottsberger RA**, **Stich K**, **Fischer TC**. 2012. Blue EaLAMP—a specific and sensitive method for visual detection of genomic *Erwinia amylovora* DNA. Eur J Plant Pathol **134**:835–845.
35. **Li J**, **Wei Q**, **Liu Y**, **Tan X**, **Zhang W**, **Wu J**, **Charimbu MK**, **Hu B**, **Cheng Z**, **Yu C**, **Tao X**. 2013. One-step reverse transcription loop-mediated isothermal amplification for the rapid detection of cucumber green mottle mosaic virus. J Virol Methods **193**:583–8.
36. **Li X**, **Nie J**, **Ward L**, **Madani M**, **Hsiang T**, **Zhao Y**, **De Boer SH**. 2009. Comparative genomics-guided loop-mediated isothermal amplification for characterization of *Pseudomonas syringae* pv. *phaseolicola*. J Appl Microbiol **107**:717–26.
37. **Rigano L a**, **Marano MR**, **Castagnaro AP**, **Do Amaral AM**, **Vojnov A a**. 2010. Rapid and sensitive detection of Citrus Bacterial Canker by loop-mediated isothermal amplification combined with simple visual evaluation methods. BMC Microbiol **10**:176.
38. **Kubota R**, **Vine BG**, **Alvarez a M**, **Jenkins DM**. 2008. Detection of *Ralstonia solanacearum* by loop-mediated isothermal amplification. Phytopathology **98**:1045–51.
39. **Yamashita Y**, **Tsukioka Y**, **Tomihisa K**, **Nakano Y**, **Koga T**. 1998. Genes involved in cell wall localization and side chain formation of rhamnose-glucose polysaccharide in *Streptococcus mutans*. J Bacteriol **180**:5803–5807.
40. **Wang L**, **Vinogradov E V**, **Bogdanove AJ**. 2013. Requirement of the lipopolysaccharide O-chain biosynthesis gene wxocB for type III secretion and virulence of *Xanthomonas oryzae* pv. *oryzicola*. J Bacteriol **195**:1959–69.
41. **Mori Y**, **Kanda H**, **Notomi T**. 2013. Loop-mediated isothermal amplification (LAMP): recent progress in research and development. J Infect Chemother 1–8.
42. **Liang C**, **Cheng S**, **Chu Y**, **Wu H**, **Zou B**, **Huang H**, **Xi T**, **Zhou G**. 2013. A closed-tube detection of loop-mediated isothermal amplification (LAMP) products using a wax-sealed fluorescent intercalator. J Nanosci Nanotechnol **13**:3999–4005.

CHAPTER 3

TWO COMPLETE GENOME SEQUENCES OF A NEW PATHOVAR OF *XANTHOMONAS ORYZAE* INFECTING WILD GRASSES PROVIDE INSIGHT INTO THE EVOLUTION OF PATHOGENICITY[3,4]

3.1 INTRODUCTION

Rice is a staple crop for more than half the world. Severe rice diseases, such as bacterial leaf streak (BLS) caused by *Xanthomonas oryzae* pv. *oryzicola* and bacterial blight (BB), caused by *X. o.* pv. *oryzae*, are increasing in prevalence in parts of Asia and sub-Saharan Africa and cause significant yield losses. In Asia, perennial weeds are considered an important source of primary pathogen inoculum (1, 2).

Southern cutgrass (*Leersia hexandra* Swartz), is a common grass found in the southern United States, South America, Africa and Asia. It is a member of the *Poaceae* family of grasses and is closely related to rice, but diverged from *Oryza* approximately 14 mya (3). *L. hexandra* is an invasive species that frequently grows along rivers and canals surrounding rice paddies. Because of its close relationship to *Oryza* spp., *Leersia* spp. are included as outgroups in phylogenetic studies. Recent genome investigations of *L. perrieri,* a cutgrass found in Madagascar, were done to compare repetitive elements and transposable elements among *Oryza* sp. and to uncover orthologs of the important submergence tolerance gene, *SUB1* (4–6).

Early reports of phytoremediation by *L. hexandra* showed this grass' capacity to sequester Cr, Cu and Ni, and it has now been proposed as a tool in wastewater treatment (7–9). Interestingly, this and other

[3] Chapter to be submitted for publication with co-authors: Ralf Koebnik, Alvaro Pérez-Quintero, Jonathan Jacobs, Elysa DuCharme, Boris Szurek, Valérie Verdier and Jan E. Leach

[4] Contributions by J.M. Lang: Design of experiments; all pathogenicity screening; bacterial quantification; molecular diagnostics; draft genome comparisons; type 3 effector survey (with guidance from R. Koebnik), FuncTAL and Talvez processing (with guidance from A. Perez-Quintero); data analysis; wrote manuscript

grasses in this genus: *L. sayanuka, L. oryzoides* and *L. japonica* are susceptible to *X. o.* pv. *oryzae* (10) and can serve as reservoirs for inoculum. *X. o.* pv. *oryzae* strains isolated from symptomless *L. hexandra* cause bacterial blight symptoms in rice, and, in artificially inoculated weed plants, *X. o.* pv. *oryzae* multiplied without evidence of disease (11) implicating this grass as an alternative host for the pathogen. These and other findings reinforce that effective integrated management of crop diseases must incorporate knowledge of pathogen interactions with weedy species.

The species *X. oryzae* is highly diverse, and is represented by distinct lineages for *X. o.* pv. *oryzae* from Asia and Africa, *X. o.* pv. *oryzicola* from Asia and Africa, and undesignated pathovars, *X. oryzae* from the United States (12–16). Poulin et al (2015) used multi-locus variable-number tandem-repeat analysis (MLVA) to investigate genetic structures of microbial populations (14, 17), and suggested that *X. o.* pv. *oryzae* and *X. o.* pv. *oryzicola* from Africa had a common Asian ancestor; this conclusion was based on the fact that the allelic richness, or number of allelles, was less in these populations. However, further analyses on an extensively sampled set of isolates are needed to confirm this ancestral hypothesis.

X. campestris pv. *leersiae,* a pathogen of *L. hexandra*, was previously shown to group distinctly from *X. oryzae* by host range and phylogenetic analysis (18, 19). Using a multi-locus sequence alignment (MLSA) analysis, however, Triplett et al (2011) showed that *X. c.* pv. *leersiae* strain NCPPB4346, which was isolated from southern cutgrass in China, groups within the *X. oryzae* cluster, yet it could not be assigned to any described pathovar. A more recently isolated strain, BAI23 from southern cutgrass in Burkina Faso, showed high sequence similarity with *X. c.* pv. *leersiae* NCPPB4346 based on a MLSA analysis (20). Together, the two strains form a distinct genetic cluster within *X. oryzae.* Further taxonomic and phenotypic investigation is needed to uncover if *X. c.* pv. *leersiae* evolved from the rice pathogens *X. o.* pvs. *oryzae* or *oryzicola* or if these weed pathogens served as their progenitor.

Xanthomonas spp. inject effector proteins into plant host cells to elicit disease via a type-three secretion system (21). These proteins can confer pathogenicity and/or dictate host specificity (22). *Xanthomonas* spp. are most notable for production of transcription activator-like effectors (TALEs). TALEs influence host gene expression by directly binding to specific sequences (effector binding

elements, EBEs) in the target promoter as dictated by repeat-variable di-residues (RVDs) at the 12 and 13 amino acid position in the central repeat region (CRR) (21, 23–25). The CRR contains different numbers of repeats, each with 33-35 amino acids. TALEs may enhance diseases by targeting susceptibility (S) genes, or may trigger a resistance response through activation of an 'executor' resistance gene (R) expression (26, 27).

The presence of TALE effectors is variable in the genus (22). *X. o.* pv. *oryzae* contain nine to 20 TALEs while *X. o.* pv. *oryzicola* can contain up to 29. US *X. oryzae* do not contain any TALEs but have been employed to study TALE effector biology in rice (28, 29). New sequencing technologies and predictive algorithms have accelerated the characterization of TALEs and their host gene targets. Long read, single molecule real time sequencing (SMRT, Pacific Biosciences, Menlo Park, CA) has enabled the rapid assembly of TALE sequences that are otherwise laborious to capture due to their highly repetitive structure (30–35).

Prediction algorithms based on the TALE's specific RVD pattern and their corresponding degenerate DNA code has facilitated identification of plant target genes whose promoters contain EBEs for TALE binding (35–40). Many *S* genes are transporters (sugar or sulfate) or transcription factors and upon induction facilitate bacterial colonization and symptom development (27). The SWEET family of sugar-efflux transporters are the best characterized susceptibility targets of *X. o.* pv. *oryzae,* but to date, none have been demonstrated to be targets for *X. o.* pv. *oryzicola* TALEs (41–43). As a defense mechanism, plants evolved mutations in EBEs of TALE targets to prevent binding or evolved an EBE in the promoters of a resistance gene, also called executor R genes, to activate defenses upon TALE binding (27, 44–46).

Although a large body of work is available on TALEs from *X. o.* pvs. *oryzae* and *oryzicola,* no information has been reported on the TALEs from *X. c.* pv. *leersiae,* how they compare to those in other *X. oryzae,* and the nature of their predicted targets within *Leersia* spp. In this study, we used comparative genomics, identification of T3E and TALE repertoires, and disease phenotyping to characterize *X. c.* pv. *leersiae*. We used gene target prediction algorithms to identify potential *X. c.* pv. *leersiae* TALE gene

targets in draft *Leersia* genome sequences. Finally, we provide evidence to support renaming *X. c.* pv. *leersiae* to *X. oryzae* pv. *leersiae* (Fang 1947) sp. nov.

3.2 MATERIALS AND METHODS

3.2.1 Bacterial Strains and Plant Varieties

Bacterial strains included in this study are listed in Table 3.1. Bacteria were cultured on peptone sucrose agar (PSA) at 28°C for plant inoculations. Genomic DNA for sequencing was isolated from *X. o.* pv. *leersiae* strains NCPPB4346 and BAI23 grown for 48 h on nutrient agar at 28°C (20, 47). Barley (*Hordeum vulgare* L. cultivar Morex), wheat (*Triticum aestivum* cv. Chinese Spring) and tobacco (*Nicotiana tabacum*) were grown in a growth chamber at 22°C, 50% humidity and 16 h of light. Rice (*Oryza sativa*) varieties included in pathogenicity assays were Azucena, Carolina Gold, Cypress, IR64 and Nipponbare.

Southern cutgrass (*Leersia hexandra)* was collected in Texas, USA and seed was propagated at Colorado State University. The seed was scarified then germinated in porous ceramic silica (Greens Grade, Profile Products, LLC, Buffalo Grove, IL, USA) and 0.5x Hoagland's solution with the following modifications: 2.5 mM KNO_3, 1mM $MgSO_4$, 3 mM KH_2PO_4, 2.5 $Ca(NO_3)_2$, 0.05 mM $FeSO_4$, and 0.1 mM $(Na)_2EDTA$ (48). Seed was incubated in a petri dish in the dark for 8 days, then in the light for 30 days at 28°C. Germinated seeds were transplanted into 1-gallon pots with equal parts Greens Grade and ProMix BX (ProMix, Quakertown, PA, USA), and grown in a greenhouse (27 ± 1°C, 16 h day length, and 80-85% relative humidity). Additional plants were obtained from the two mother plants via rhizome propagation. Propagation began 30 days after planting with subsequent propagation every week to allow time for new growth. To promote root growth, plants were placed in the dark for 24 h.

3.2.2 Molecular Diagnostics

To test relationships of *X. c.* pv. *leersiae* to *X. oryzae* (12, 20), a diagnostic multiplex and LAMP PCR were used (20, 47, 49). Previously described universal US *X. orzyae* primers were also tested to differentiate *X. c.* pv. *leersiae* from this novel US clade within the species (12).

3.2.3 Pathogenicity Assays

Rice varieties were inoculated at four to five weeks-old with suspensions (10^8 CFU mL^{-1}) of bacterial strains listed in Table 3.1. Bacterial suspensions were both infiltrated into the intercellular spaces of rice leaves on either side of the main vein with a needleless syringe and inoculated by leaf clipping as described (50, 51). Two leaves were inoculated on each of three to six separate plants; water was included as a negative control. The entire experiment was conducted twice. Lesions were measured at 12 days post inoculation (dpi), and bacterial numbers *in planta* were quantified as previously described (28).

3.3.4 Genome Sequencing and Assembly

Draft genome sequence was obtained (Illumina, San Diego, CA) for *X. c.* pv. *leersiae* strains NCPPB4346 (NZ_LHUK01000461) and BAI23 and genomes were assembled using SPAdes v 3.1.1 (52). In addition, long read, single molecule real time sequencing (SMRT, PacBio, Menlo Park, CA) was used to facilitate assembly of TALE (31). DNA for SMRT sequencing was isolated by culturing strains on nutrient agar for 48 h then using the Genomic DNA buffer set and Genomic-tips according manufacturer instructions (Qiagen, Valencia, CA). SMRT sequence was assembled using HGAP v3 (PacBio, Menlo Park, CA). T3E repertoires were computationally determined as previously described (53).

To capture full-length TALE sequences we analyzed SMRT-sequences using AnnoTALE and conserved N and C-termini *in silico* (35). TALE targets were predicted using the QueTAL suite of tools (38, 54). FuncTAL, a program designed to find TALEs with similar DNA-binding capabilities, was used to build trees based on these correlations, and DisTAL was used to infer phylogenetic relationships between genes by considering each repeat as a unit and calculating distances between arrays of repeats (54). In this study 563 TALEs from 29 strains (16 *X. o.* pv. *oryzae*, 11 *X. o.* pv. *oryzicola* and two *X. c.* pv. *leersiae*) were used as input in DisTAL. Effector binding elements (EBE) were predicted against promoter sequences (1kb upstream from the translation start site) for annotated genes from the *L. perrieri* genome (v1.4.30) (4) using Talvez against both strands (38). Talvez scans input sequences for EBEs based on input RVDs. Next, the closest orthologs of the predicted targets from *L. perrieri* were identified from rice by a reciprocal BLAST search against the genome of rice cv. Nipponbare genome v7 (55).

3.3 RESULTS

3.3.1 Comparative genomics elucidate taxonomic placement of X. campestris pv. leersiae

Amplification of *X. c.* pv. *leersiae* DNA with primers specific for *X. oryzae* but not *X. o.* pv. *oryzicola, X. o.* pv. *oryzae* or US *X. oryzae* suggested that *X. c.* pv. *leersiae* were related to *X. oryzae,* but distinct from other *X. oryzae* pathovars (12, 49) (Fig. A.2.1a, b, c). To better understand the relationships of these *Xanthomonas* species, we generated draft genomes for *X. c.* pv. *leersiae* strains NCPPB436 and BAI23. The genomes were 4.88 and 4.59 Mb, respectively, and GC content for both was 63%, consistent with other *Xanthomonas* species. Whole draft genome assemblies were compared to calculate percent average nucleotide identity (ANI) using JSpecies v1.2.1 (56, 57) and the Genome-based Distance Matrix Calculator (http://enve-omics.ce.gatech.edu/g-matrix/index) with similar results (58, 59). At a 95% cut off for species delineation, *X. c.* pv. *leersiae* strains were 99% identical to one another, and were 97-99% identical to US *X. oryzae, X. o.* pv. *oryzicola* and *X. o.* pv. *oryzae*. *X. c.* pv. *leersiae* were most similar (99%) to the African lineage of *X. o.* pv. *oryzicola.* They were 76-91% similar to other *Xanthomonas* species (Table 3.2, Figure 3.1, Figure A.2.2). *X. vasciola* was the next most similar *Xanthomonas* species to *oryzae*, sharing 91% ANI with *X. c.* pv. *leersiae* and all *X. oryzae*. These data suggest that strains NCPPB4346 and BAI23 are more closely related to *X. oryzae* than other *Xanthomonas* species and combined with previously reported MLSA data (12, 20), we recommend that the taxonomic placement of *X. c.* pv. *leersiae* be in the species '*oryzae'*. Further biological support for this shift from host range and effector repertoires is described below. Hereafter, we will refer to *X. c.* pv. *leersiae* as *X. o.* pv. *leersiae*.

3.3.2 X. o. pv. leersiae is pathogenic on rice and southern cutgrass

To better understand the biology of *X. o.* pv. *leersiae*, we established the host range by screening for pathogenicity on rice and southern cutgrass (*L. hexandra*) using different inoculation techniques. Relative to the virulent Philippine *X. o.* pv. *oryzicola* strain BLS256, the *X. o.* pv. *leersiae* strains were less aggressive to rice, and caused less expansive water-soaked leaf streaking on several tested rice varieties (Figs. 3.2a and c). *X. o.* pv. *leersiae* BAI23 was more aggressive than NCPPB4346 on rice varieties Cypress and IR64, and caused more disease than the US *X. oryzae* strain X11-5A on Azucena

(Fig. 3.2c). Rice variety Carolina Gold was resistant to *X. o.* pv. *leersiae*, exhibiting a hypersensitive-like response (Fig. 3.2a) and no lesion expansion after infiltration (Fig. 3.2c). Both *X. o.* pv. *leersiae* strains caused longer lesions on southern cutgrass than *X. oryzae*, but were not as aggressive as *X. o.* pv. *oryzicola* BLS256 on rice (Figs 3.2a and b). *X. oryzae* produced water-soaked spots at the point of infiltration that did not expand on southern cutgrass (Figs. 3.2c).

After clip inoculations, lesions caused by *X. o.* pv. *leersiae* did not expand on rice or southern cutgrass, unlike those caused by the vascular pathogen *X. o.* pv. *oryzae* (Fig. 3.3). When infiltrated into leaves, populations of *X. o.* pv. *leersiae* were equivalent to *X. o.* pv. *oryzicola* BLS256 and *X. oryzae* X11-5A on rice cvs. Nipponbare or Azucena after 72hpi (Fig. 3.4). But, on southern cutgrass, their native host, *X. o.* pv. *leersiae* grew to a significantly higher population than *X. oryzae* X11-5A. *X. o.* pv. *leersiae* did not cause disease on wheat or barley (Fig. A.2.3, A.2.4). *X. o.* pv. *leersiae* caused minor chlorosis, but not water soaking or a hypersensitive response, when infiltrated into *N. tabacum*, similar to the phenotype caused by *X. o.* pv. *oryzae* PXO99A and *X. o.* pv. *oryzicola* BLS256 (Fig. A.2.5). As described before (15), *X. o.* pv. *oryzae* BAI3 and *X. oryzae* US 11-5A caused a typical hypersensitive response at the site of infiltration on *N. tabacum* (15). Phenotyping *Nicotiana* species can serve as a screen for the ability of microbes, particularly *Xanthomonas* spp., to elicit a non-host resistance response (15). Like *X. o.* pvs. *oryzae* and *oryzicola*, both *X. o.* pv. *leersiae* strains caused only minor chlorosis when infiltrated into leaves. Intriguingly, *X. oryzae* from the USA and *X. o.* pv. *oryzae* BAI3, isolated in Burkina Faso, caused a strong hypersensitive response at 2 dpi (Fig. A.2.5).

Our studies show that *X. o.* pv. *leersiae* is pathogenic to southern cutgrass and mildly pathogenic on diverse varieties of rice, but does not cause disease on barley or wheat, and that some rice varieties exhibit resistance to *X. o.* pv. *leersiae.* After artificial inoculation, *X. o.* pv. *leersiae* causes symptoms on rice that are most similar to *X. o.* pv. *oryzicola*, i.e., expanding lesions when introduced into the intercellular spaces (leaf infiltration) and no spreading lesions when introduced into the xylem vessels (clipping).

3.3.3 T3E repertoires of X. o. pv. leersiae are similar to X. o. pvs. oryzae and oryzicola

Computational prediction of *X. o.* pv. *leersiae* T3E repertoires revealed characteristics that are both unique and shared with *X. oryzae,* with 23/56 T3Es shared among *X. o.* pv. *leersiae* and other *X. oryzae* pathovars (Table 3.3). *xopG* and *xopAF* are unique to *X. o.* pv. *oryzae*, and *xopO* and *xopAJ* are present only in *X. o.* pv. *oryzicola;* none of these are present in *X. o.* pv. *leersiae. xopAK* is absent in *X. o.* pv. *oryzicola* but present in *X. o.* pv. *oryzae* and *X. o.* pv. *leersiae. xopE2* is found only in *X. o.* pv. *leersiae* BAI23 and not NCPPB4346. *xopW* is interrupted by an insertion sequence in African strains of *X. o.* pv. *oryzicola* and *X. o.* pv. *oryzae;* this insertion was only observed in the African *X. o.* pv. *leersiae* BAI23 (16, 20). Both *X. o.* pv. *leersiae* strains possess a *xopD* gene that is absent in all other *X. oryzae* genomes. The US strains of *X. oryzae* do not contain *xopU*, a T3E that is present in all *X. o.* pv. *oryzicola, X. o.* pv. *oryzae,* and *X. o.* pv. *leersiae*.

3.3.4 SMRT sequencing uncovers novel X. o. pv. *leersiae TALEs*

Previous Southern blot analyses using conserved TALE probes predicted that *X. o.* pv. *leersiae* BAI23 and NCPPB4346 had five and four TALEs, respectively (20). However, Southern blot analysis cannot resolve TALEs that are close in size. Thus, we used SMRT sequencing to deduce the complement of TALEs, referred to as the TALome, of *X. o.* pv. *leersiae.* Indeed, the long read sequences from SMRT sequencing revealed that *X. o.* pv. *leersiae* NCPPB4346 and BAI23 contain 12 and 13 different TALEs, respectively. RVD sequences for each strain are listed in Table 3.4. Note that TALEs are named according to their sequential location in the genome from the origin of replication and alphabetically according to their location in that locus starting at the 5' end of the locus, though new naming schemes have been proposed for further clarity (32, 35, 60, 61).

Analysis in AnnoTALE (35) revealed that the *X. o.* pv. *leersiae* TALEs are unique relative to other *Xanthomonas* species. *X. o.* pv. *leersiae* BAI23 possesses one TALE (Tal1) that contains an RVD (TI) not found in *X. o.* pv. *leersiae* NCPPB4346 or any other *X. oryzae*. FuncTAL (54) grouped the BAI23 TALEs into five distinct clades and the NCPPB4346 TALEs into four clades (Fig. 3.5a and b). Comparing these two TALomes to one another in FuncTAL showed distinct pairing of TALEs from both

strains (Fig 3.5c). Tal9a and 6c from NCPPB4347 and Tal4 from BAI23 were the only TALEs to not group with a complement from the other strain. Based on a comprehensive DisTAL analysis (54), *X. o.* pv. *leersiae* TALEs group most closely to African *X. o.* pv. *oryzae*, and are in a subgroup that contains both Asian and African *X. o.* pv. *oryzicola* strains.

3.3.5. X. o. pv. *leersiae TALE targets in L. perrieri are orthologous to rice genes targeted by X. o.* pvs. *oryzae* and *oryzicola*

TALVEZ (38) was used to predict host targets from the completed TALE RVD sequence and the corresponding effector binding elements in the host promoterome. To predict gene function in the list of *X. o.* pv. *leersiae* BAI23 TALE targets from the *L. perrieri* genome, we performed a reciprocal BLAST analysis against the rice Nipponbare reference genome v7 (55). Over 5000 genes were identified in *L. perrieri* with promoters containing EBEs that match the 13 TALEs of BAI23. Based on relevant, currently available information about *X. oryzae* TALE targets, a selected, relevant subset of *X. o.* pv. *leersiae* TALE targets and their corresponding ortholog in rice are curated in Table 3.5. Of the diverse targets identified, several were *SWEET* gene orthologs. Tal8 and Tal9a both target an *OsSWEET1b* ortholog (*LOC_Os05g35140*), but this is expected due to the sequence similarity between these two TALEs as they group together into their own clade (Figure 3.5b). Tal9b targets LPERR09G12040, an ortholog of rice *OsSWEET11* (*Os8N3, Xa13*), and a well-studied virulence target of the *X. o.* pv. *oryzae* TALE, *pthXo1* (62). Tal9b and PthXo1 share 90% amino acid identity, but *pthXo1* has 24.5 repeats while Tal9b has 19.5.

The two *X. o.* pv. *leersiae* strains have fewer TALEs (12-13) than *X. o.* pv. *oryzicola* (22-29) but homology was discovered in *S* gene targets (32). The confirmed primary target of Tal2g from *X. o.* pv. *oryzicola* is a sulfate transporter, *OsSULTR3;6* (LOC_Os01g52130) (60), but no ortholog of *OsSULTR3;6* was identified as a *X. o.* pv. *leersiae* TALE target in *L. perrieri.* However, Tal5a and Tal6b are predicted (ranked 287 and 230, respectively), to target another orthologous member of this family, *OsSULTR3;1.* Of the selected target gene orthologs (Table 3.5), Tal6a, is predicted to bind the promoter of *LOC_Os06g46500*, a monocopper oxidase, that is also targeted by the *X. o.* pv. *oryzicola* TALE Tal2g.

Previous studies showed upregulation upon binding, but no virulence contribution when individually complemented with a designer TALE (60), therefore, it is unclear if this potential *S* gene is a significant target for *X. o.* pv. *leersiae* but is further evidence of their convergent virulence strategies. Tal6b additionally targets an ortholog of a bHLH family protein most recently shown as a virulence target of AvrHah1 from *X. gardneri* on tomato (63). In summary, *X. o.* pv. *leersiae* TALEs resemble those found in *X. oryzae* pathovars and exhibit congruency in their predicted host targets. While their TALomes are unique relative to other bacteria, their predicted host target strategy is comparable with the other members of this species.

3.4 DISCUSSION

X. o. pv. *leersiae*, which has been isolated from the pervasive weed species *L. hexandra* surrounding rice paddies (19, 20), was historically grouped as a distinct species and pathovar (12, 20). We compared pathogenicity on diverse cereal hosts and draft genomes of two *X. o.* pv. *leersiae* strains, one from Africa and one from Asia, to demonstrate that *X. o.* pv. *leersiae* is closely related to *X. oryzae* that are pathogens of rice. *X. o.* pv. *leersiae,* like *X. o.* pvs. *oryzae* and *oryzicola*, caused water soaked lesions on rice and *L. hexandra*, but were not virulent to wheat and barley. In phylogenetic analyses, the *X. o.* pv. *leersiae* strains grouped more closely with *X. oryzae* pathovars than other members of this genus, and in ANI analyses, were above a species delineation threshold. Therefore we propose naming these strains *X. oryzae* pv. *leersiae* (Fang 1947) sp. nov.

X. o. pv. *leersiae* colonize and cause water-soaking on rice, but the southern cutgrass isolates are not as aggressive as *X. oryzae* isolated from rice. The lesions caused by *X. o.* pv. *leersiae* were phenotypically similar to rice bacterial leaf streak caused by *X. o.* pv. *oryzicola* and these strains did not cause disease when introduced into rice or southern cutgrass by leaf-clipping. Thus, we suggest that *X. o.* pv. *leersiae* are non-systemic pathogens, and are more like *X. o.* pv. *oryzicola* than the systemic *X. o.* pv. *oryzae*.

Previous MLSA analyses established that *X. o.* pv. *leersiae* are closely related to other *X. oryzae* (12, 20), but these analyses were not sufficient to propose a taxonomic change to this species. In this

work, we used whole genome comparisons to support renaming the species *X. oryzae* because the strains from southern cutgrass demonstrated greater than 95% similarity to all other *X. oryzae*. Interestingly, *X. o.* pv. *leersiae* share 99% nucleotide similarity with *X. o.* pv. *oryzicola* strains from Africa, but only 98% similarity with *X. o.* pv. *oryzicola* from Asia (China and Philippines) and *X. o.* pv. *oryzae* from Asia or Africa. This is further evidence for an African origin of these strains.

T3Es, which are delivered through a type III secretion system, are important contributors to bacterial pathogenicity to plants because they can act as virulence factors by promoting disease or as avirulence factors by triggering host defenses depending on the host genotype (64–66). These proteins may define host range, and can also inform lineages and evolutionary relationships among different populations of related bacteria (16, 20, 53, 67). A recent comprehensive study of *X. o.* pv. *oryzae* genomes in Asia revealed three distinct lineages shaped by T3E repertoires and phenotypic adaptation to their agroecosystems (34). Likewise, this study sought to uncover similarities between *X. o.* pv. *leersiae* and *X. oryzae* effectomes that could further inform their evolutionary lineage. MLSA and T3E analysis demonstrated that Asian *X. o.* pv. *oryzae* harbor T3Es with the highest genetic diversity and defined a core set of 14 T3Es present in all *X. oryzae* pathovars screened (16). This core set consists of *avrBs2, avrBs3* (TALEs)*, xopL, xopN, xopP, xopQ, xopV, xopW, xopY, xopAA, xopAB, xopAE* and *xopF1*. Both *X. o.* pv. *leersiae* strains contain all of the core effectors except *xopW*, which is absent in NCPPB4346, and, although present in BAI23, it contains a large insertion sequence. This insertion sequence most likely prevented its amplification in previous studies (20). Interestingly, this same insertion sequence was identified in other African *X. o.* pv. *oryzicola* strains, consistent with a shared evolutionary origin with *X. o.* pv. *leersiae* (16).

xopD is present in *X. o.* pv. *leersiae* but absent in *X. o.* pvs. *oryzae* and *oryzicola.* XopD is a SUMO protease mimic that suppresses host defense responses during *X. euvesicatoria* infection (68, 69). In addition to *X. o.* pv. *leersiae, xopD* is predicted to be present in *X. vasicola* pv. *vasculorum* (70), *X. campestris* pv. *campestris, X. euvesicatoria* and *Acidovorax citrulli* (www.xanthomonas.org). Absence in other *X. oryzae* suggests an independent acquisition by *X. o.* pv. *leersiae.* Conversely, loss of this effector

at some time by *X. o.* pv. *oryzae* or *X. o.* pv. *oryzicola* could have occurred, but further validation of these hypotheses are needed. Since it is not known what level of virulence, and/or host specificity any T3E conveys for *X. o.* pv. *leersiae;* future work should include functional validations.

Capturing complete TALE sequences enabled several important analyses of their homology to one another and to other species, predicted host targets and ultimately, the basis for future work. While phenotypically and by ANI *X. o.* pv. *leersiae* most closely resemble *X. o.* pv. *oryzicola*, their TALomes place them within a subgroup of African *X. o.* pv. *oryzae.* Gonzalez et al (2007) proposed that due to the distinct features of African *X. o.* pv. *oryzae* that these strains are endemic to Africa and perhaps were originally isolated from *O. glaberrima*, the African cultivated rice progenitor. They also propose that African *X. o.* pv. *oryzicola* were possibly introduced to Africa from Asia. An interesting question is whether *X. o.* pv. *leersiae*, with its intermediate number of TALE, represents a transition between African *X. o.* pv. *oryzae* (with a relatively small group of conserved TALEs), Asian *X. o.* pv. *oryzae* and all lineages of *X. o.* pv. *oryzicola* (with a larger number of TALEs). Interestingly, one unique RVD was uncovered in Tal1 from *X. o.* pv. *leersiae* BAI23: TI. This RVD has not been previously reported in any species.

Rice and southern cutgrass are closely related members of the *Poaceae. Leersia* species are often used as an outgroup in phylogenetic and most recently, genomic investigations (4). Evidence of genome duplication events and defense response genes shared by rice and *Leersia* species has been reported (71, 72), but the only available genome in the genus *Leersia* is *L. perrieri*. Therefore, this genome served as proxy for our EBE target predictions by *X. o.* pv. *leersiae* TALEs, even though they were isolated from a different species (*L. hexandra*). Talvez creates a RVD-nucleotide association matrix to convert a sequence of RVDs for a given TALE into a positional weight matrix (PWM). Gene targets are then ranked based on the probability of a nucleotide corresponding to the PWM versus the probabilities of finding this nucleotide in any position according to the nucleotide composition (38). Over 5000 gene target predictions in the *L. perrieri* genome were made using the TALEs from *X. o.* pv. *leersiae* BAI23 and the predictions for their function are overall diverse.

While *X. o.* pv. *leersiae* TALEs are unique, our reciprocal BLAST against the rice genome revealed four *OsSWEET* orthologs as putative targets in *L. perrieri* for five different TALEs, and all but one rank above 150. The rice SWEET-*nodulin* family, composed of 21 members segregating into four clades (43) are among the most extensively studied virulence targets of TALEs. Members of this family localize to the plasma membrane and mediate efflux of sucrose and glucose from the phloem parenchyma to the apoplast allowing pathogens access to these valuable molecules for greater fitness. SWEET genes appear to be conserved targets among many *Xanthomonas* spp. (73). *MeSWEET10a* and *CsSWEET1* are, respectively, induced by TAL20 from *X. axonopodis* pv. *manihotis* in cassava and the PthA series from *X. citri* ssp. *citri* in citrus (27, 74, 75). *OsSWEET11* (*Os8N3, xa13*) has been shown as a host-susceptibility gene in rice specifically targeted by the *X. o.* pv. *oryzae* TALE pthXo1 (62). We hypothesize that *Leersia* spp. have a family of SWEET genes not yet characterized that are susceptibility targets of *X. o.* pv. *leersiae* for virulence. Redundancy in TALE targets has been reported and is not surprising that an overlap exists in the example of SWEET orthologs by *X. o.* pv. *leersiae* (43, 60). The ortholog of *OsSWEET1b* contains three different EBEs targeted by two different TALEs all with high ranking. This provides further evidence that the evolution of TALEs can converge on an important family of susceptibility factors in plant hosts. These data are however, predictions; functional validation of these targets in southern cutgrass is necessary to determine if these TALEs are indeed influencing predicted target gene expression. A comprehensive analysis of sucrose transporters in *Leersia* spp. would be informative for not only understanding plant-microbe virulence activity but also evolutionary physiology in higher grasses.

One of the top targets of the *X. o.* pv. *leersiae,* Tal6b, is a bHLH transcription factor family member (LPERR03G04710). This is an interesting and logical TALE target, as activation of these genes may improve colonization in the apoplast by promoting water intake through activation of two pectin modification genes, as was elegantly demonstrated recently in tomato by the *X. gardneri* TALE avrHah1 (63). Our studies suggest that *X. o.* pv. *leersiae* is a non-systemic pathogen, likely multiplying in the extracellular spaces, and this target prediction is consistent with a mechanism of how these organisms

may colonize this tissue space. The highest ranked selected TALE target of this set (by Tal7) was a predicted dehydrogenase ortholog, potentially interfering with reactive oxygen species during a defense response.

The rice genome encodes 14 sulfate transporters divided into five groups (76). Eight members of this family were predicted as orthologous targets of seven different *X. o.* pv. *leersiae* TALEs. The primary virulence target of Tal2g from *X. o.* pv. *oryzicola* is *OsSULTR3;6* which is a member of the sulfate transporter family 3 (60, 77). Tal5a and Tal6b from *X. o.* pv. *leersiae* BAI23 share a predicted target from another member of this family: *OsSULTR3;1* with high ranking (Table 3.5). It is feasible that this TALE mirrors the virulence function of Tal2g in *X. o.* pv. *leersiae* and is particularly intriguing since two independent TALEs could target this gene.

One proposed explanation for multiple TALEs converging on one target in *X. oryzae* pathovars is that these geographically diverse strains evolved targets with functional redundancy but not total sequence homology (38) or that certain TALEs are necessary on certain genotypes. The conservation of TALE gene targets is quite high within this genus and provides a baseline for what *X. o.* pv. *leersiae* has evolved to target over time. Furthermore, access to the *L. hexandra* genome would complement this work and future phylogenetic studies in the *Poaceae*. Another limitation of this study is the number of strains available. Pathogenicity and identification of additional strains from southern cutgrass (*L. hexandra),* other *Leersia* species as well as other weeds could support further phylogenomic analyses to better understand the population ecology of these organisms. With such a small sample number, population dynamics could not be completed, however, valuable fundamental information was acquired through these phenotypic and genomic approaches.

It is plausible that based on their genetic and evolutionary relatedness that *Oryza* and *Leersia* have co-evolved their respective pathogens and that would appear evident here with the information to date. But as Jacques *et al.*, 2016 suggest in their recent review, this system most likely represents a sympatric scenario where *X. o.* pv. *oryzae, X. o.* pv. *oryzicola, X. oryzae* and *X. o.* pv. *leersiae* have all lived in the same habitat and have had the opportunity to exchange genetic material, then adapted to their

respective host while maintaining basic homology. It does not appear that these organisms are demonstrating a recent host-shift, but rather have simply co-evolved after a long adaptation to southern cutgrass. Host jumps seldom occur, suggesting host defense responses are difficult to overcome (78). However, in this scenario, the hosts are quite similar making it more feasible for *X. o.* pv. *oryzicola,* for example, to colonize *L. hexandra* and adapt quickly over time. This information could be inferred through forced evolution experiments, though these approaches can be problematic with many variables to consider.

X. o. pv. *leersiae* could also be a bridge between *X. o.* pvs. *oryzae* and *oryzicola* or were once *X. o.* pv. *oryzicola* that colonized southern cutgrass due to their similar infection biology. The T3E repertoire is very similar to the other members of *X. oryzae* suggesting that *X. o.* pv. *leersiae* evolved from *X. oryzae* and then adapted its unique TALome to southern cutgrass. It would be advantageous for the rice pathogens to be able to survive and colonize other hosts in the same agro-ecosystem to prolong fitness and it would take very little time in theory, for *X. o.* pv. *leersiae* to quickly and advantageously acquire more pathogenicity factors, such as a rice specific SWEET-inducing TALE, or evolve simple changes to adapt to southern cutgrass, in order to become a virulent rice pathogen. We hypothesize that any number of other TALEs or T3Es would result in the same gain of virulence.

In summary, we propose *X. o.* pv. *leersiae*, which was isolated from southern cutgrass (*L. hexandra*), a weedy grass closely related to rice, as a new member of the *oryzae* species. Genomic analysis and disease phenotyping on various hosts demonstrated the close relationship of *X. o.* pv. *leersiae* to the rice pathogens *X. o.* pvs. *oryzae* and *oryzicola.* T3E and TALE content of the *X. o.* pv. *leersiae* indicated that this group of organisms uses similar virulence mechanisms to the rice pathogens. While weeds such as southern cutgrass are not agronomic crops, as competitors for resources and potential reservoirs of pathogen inoculum, they are important in management considerations for rice growers. The fact that they harbor a pathogen group that can also impact rice emphasizes that, in general, more attention should be focused to the surrounding ecosystem in rice production. Research contributing

towards understanding the *X. o.* pv. *leersiae*/rice/southern cutgrass pathosystem will be significant for all rice-producing countries.

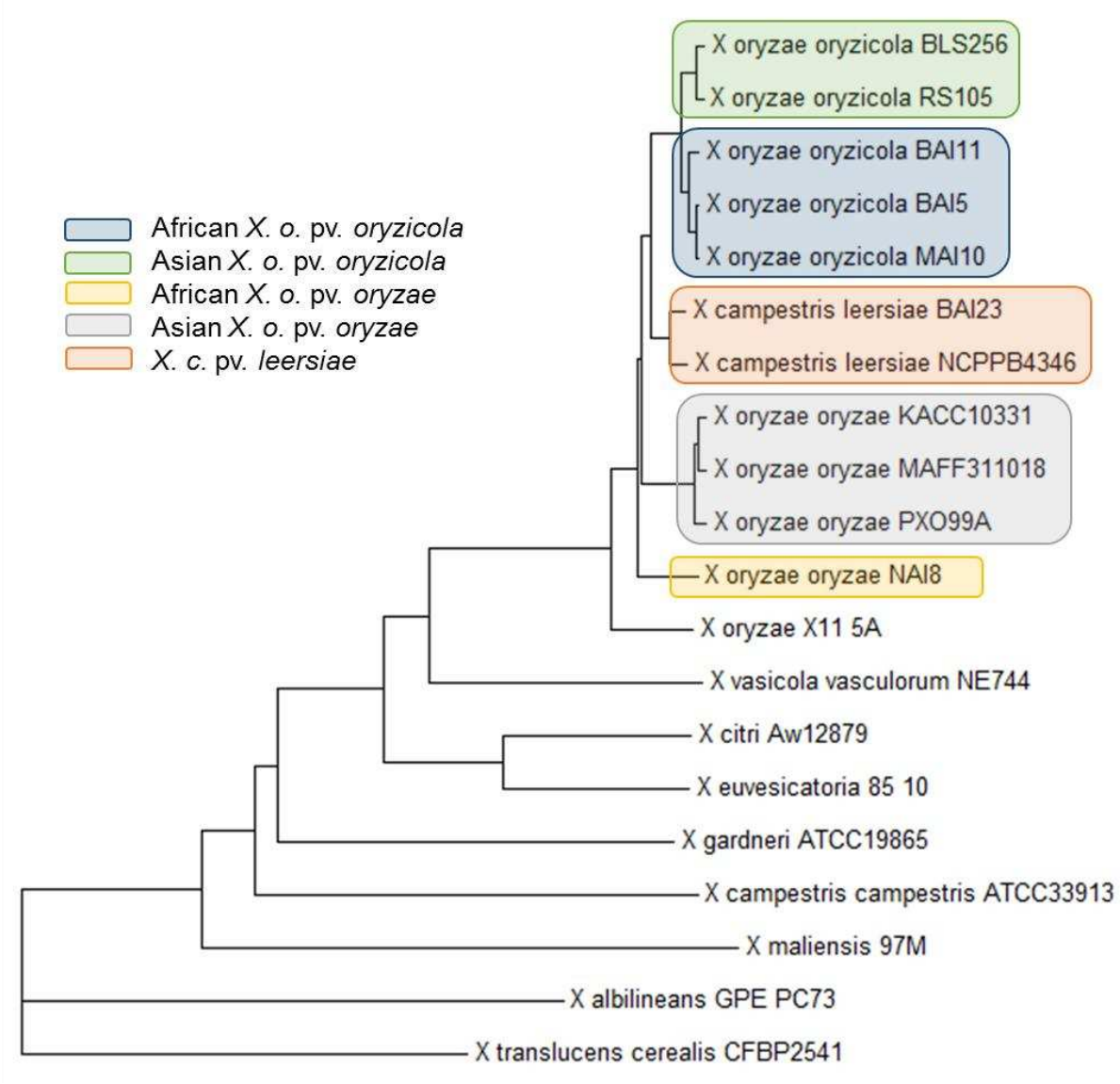

Figure 3.1. Hierarchical clustering of *X. campestris* pv. *leersiae* whole genomes relative to other *Xanthomonas* species based on average nucleotide identity (ANI). Values and tree generated using the Average Nucleotide Genome Matrix Calculator http://enve-omics.ce.gatech.edu/g-matrix/ (58, 59). Clustering by balanced minimal evolution and rooted with *X. vasicola* pv. *vasculorum* NE744. All assembled genomes were collected from NCBI Genome database except *X. c.* pv. *leersiae* BAI23 which was generated in this study.

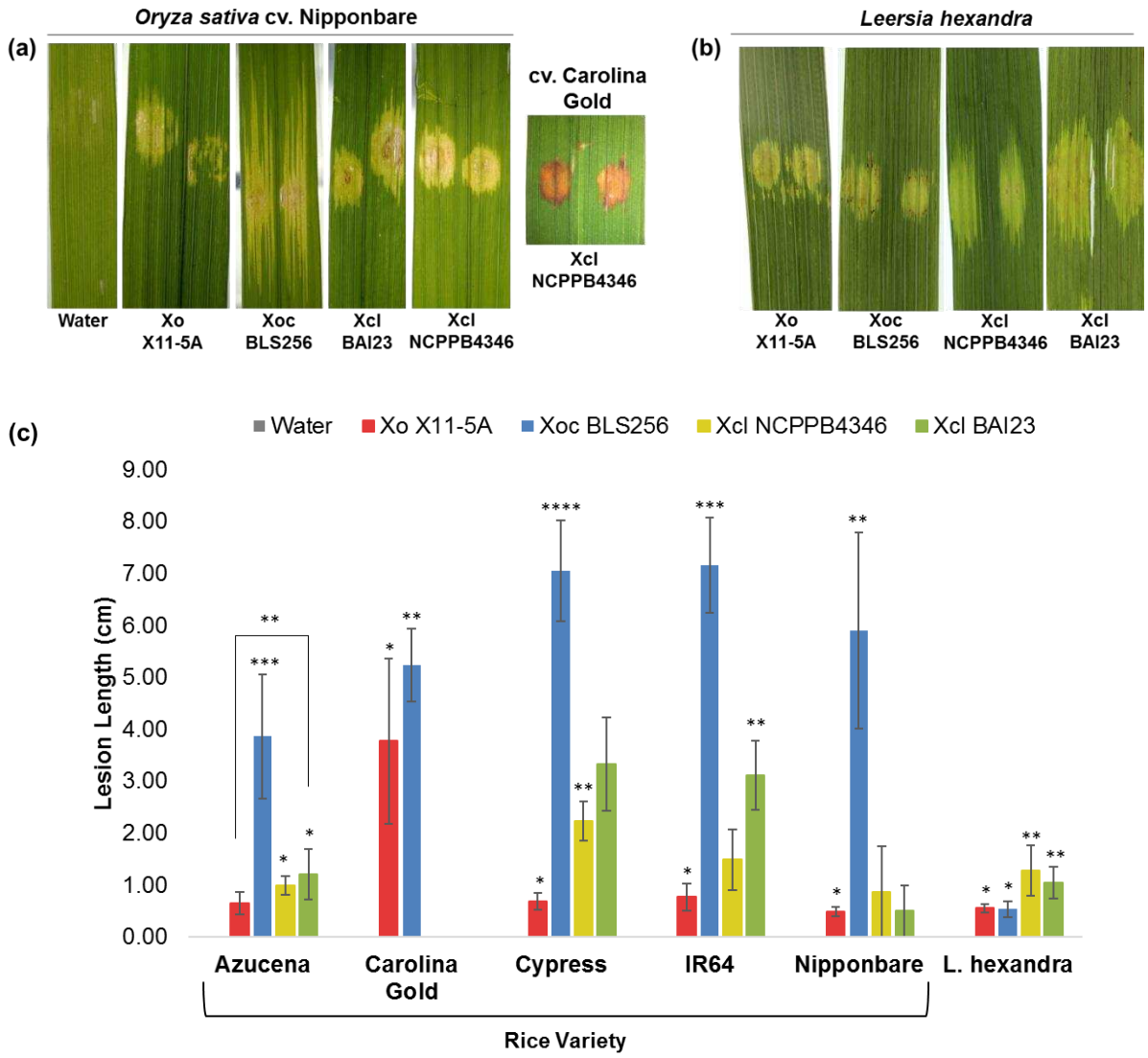

Figure 3.2. *X. oryzae* (*campestris*) pv. *leersiae* (Xcl) and *X. oryzae* cause water soaking on reciprocal hosts, but virulence is highest on original host. Phenotypes and quantitative lesion lengths caused by *Xanthomonas oryzae* (Xo) X11-5A, *X. o.* pv. *oryzicola* (Xoc) BLS256 and *X. o.* pv. *leersiae* NCPPB4346 on (a, c) diverse rice varieties and (b, c) wild *Leersia hexandra* were measured 12 days post infiltration inoculation. An asterisk denotes a significant difference between strains on each variety in a Student's T-test ($p \leq 0.05$). Error bars represent ± SD.

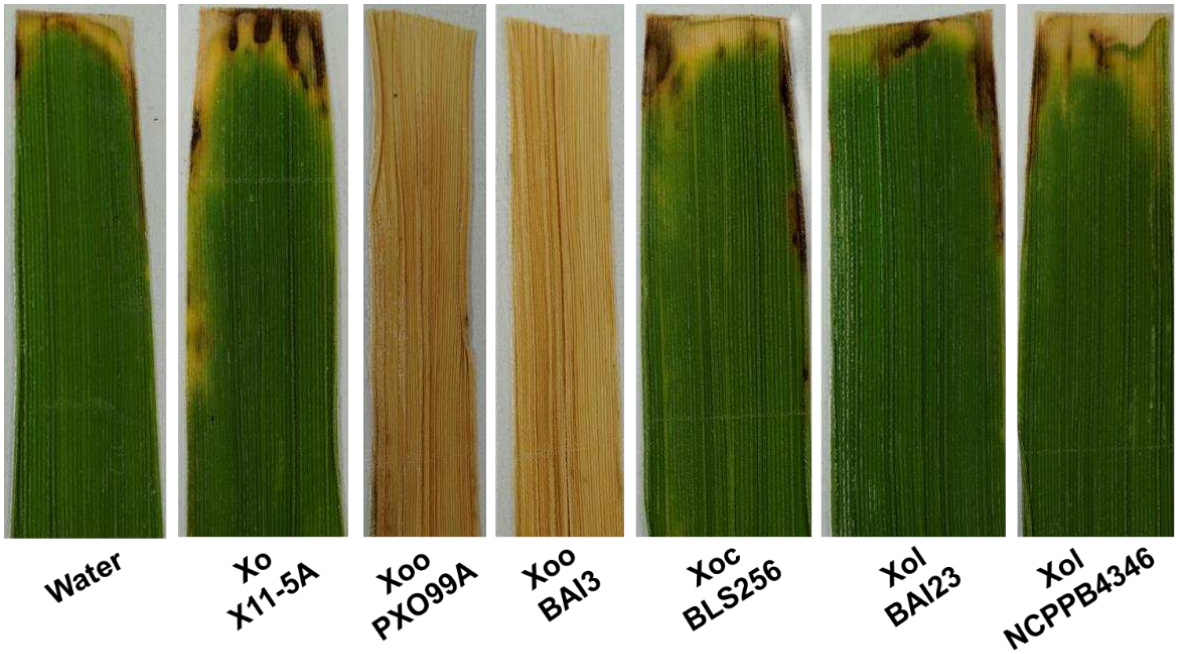

Figure 3.3. *X. oryzae* pv. *leersiae* (Xol)*, X. o.* pv. *oryzicola* (Xoc), and US *X. oryzae* (Xo*)* strains do not induce spreading lesions after leaf clip inoculation on rice cv. Nipponbare 14dpi, whereas *X. o.* pv. *oryzae* (Xoo) strains do.

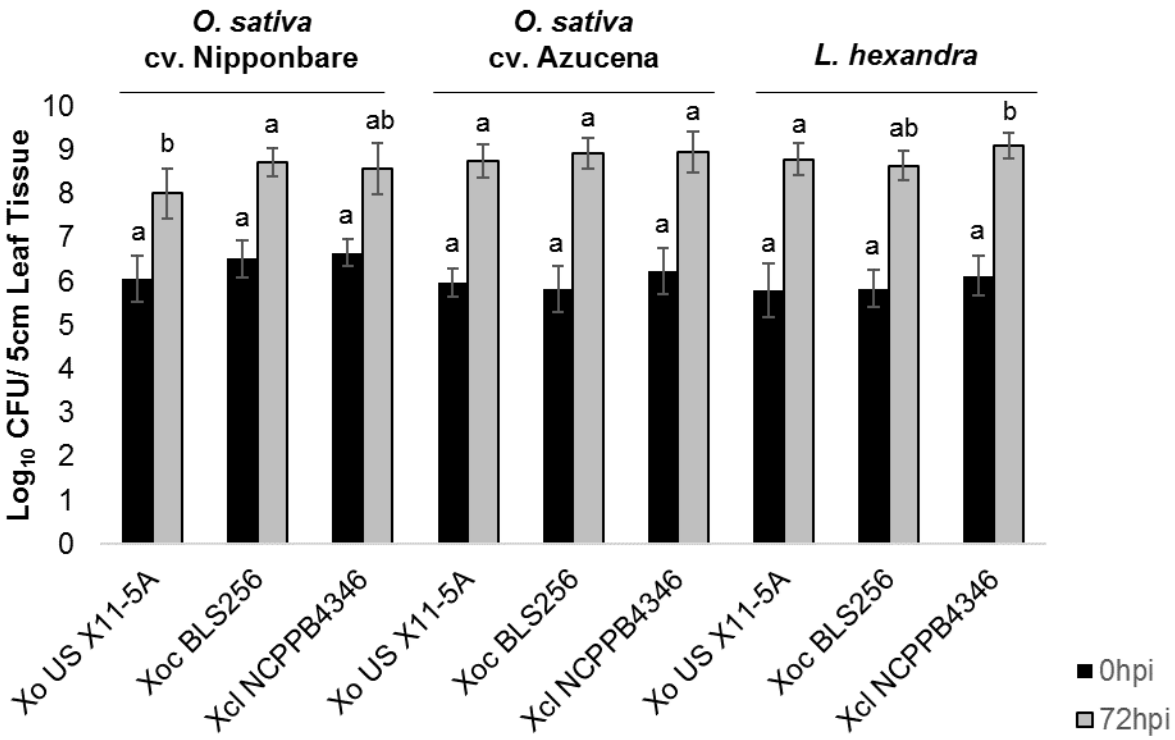

Figure 3.4. *X. oryzae* pv. *leersiae* can grow and colonize rice as effectively as rice pathogens and rice pathogens can colonize *Leersia hexandra* as effectively as *X. oryzae* pv. *leersiae*. Bacterial population growth in leaves of rice varieties Nipponbare and Azucena and *Leersia hexandra* inoculated with *X. oryzae* X11-5A, *X. oryzae* pv. *oryzicola* BLS256 and *X. oryzae* pv. *leersiae* NCPPB4346 were quantified at 0 and 72 hours post inoculation (hpi). Population sizes were measured in a 5 cm leaf segment infiltrated with each strain. Error bars represent ± SD of six independent leaves, and letters denote treatments significantly different from one another on each variety ($p \leq 0.05$).

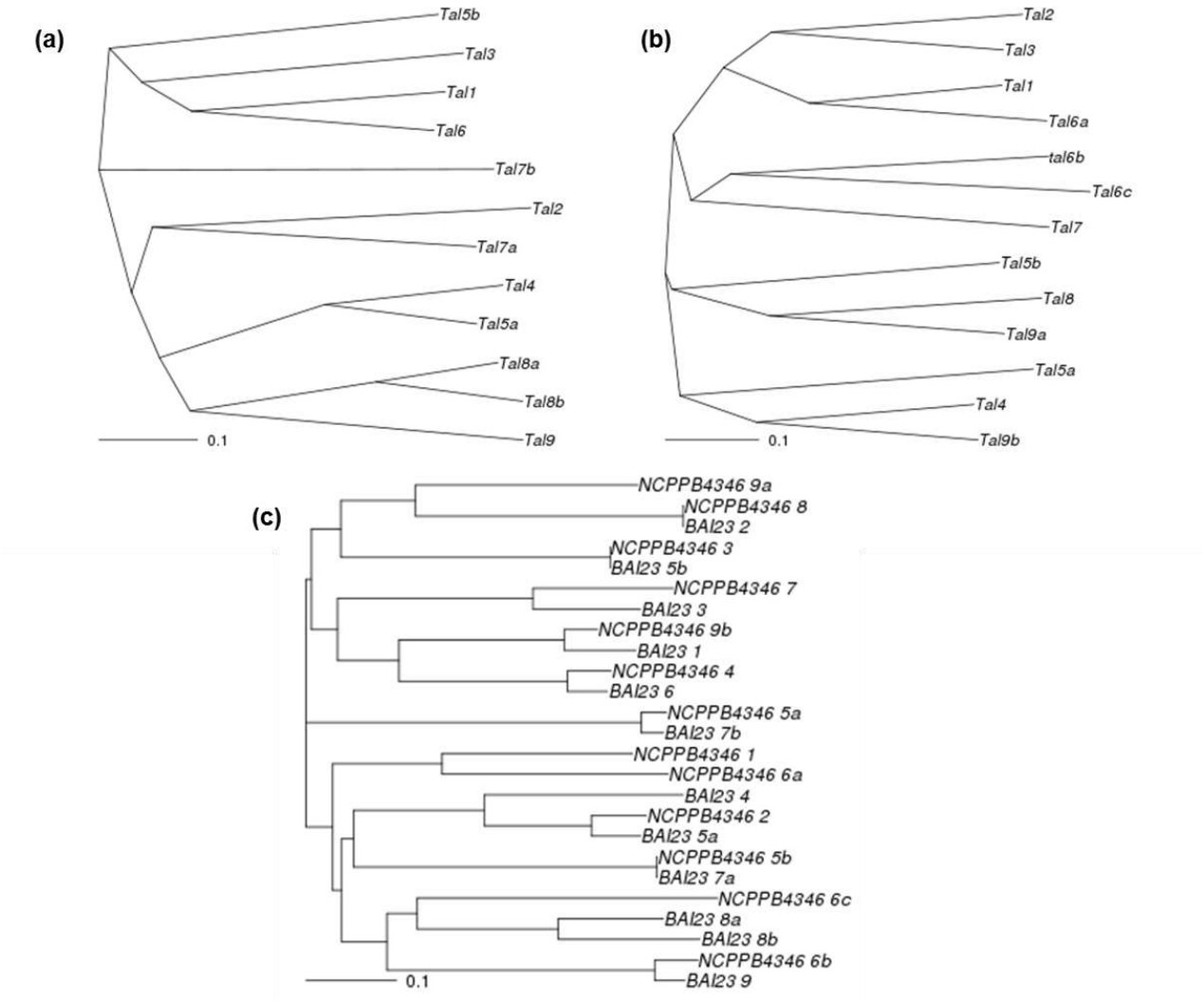

Figure 3.5 Cladograms and phylogram of functionally related TALEs for *X. oryzae* pv. *leersiae* strains NCPPB436 (A) and BAI23 (B) and both strains combined (C) based on predicted DNA binding similarities from full effector sequence. Analysis and figures developed in FuncTAL of the QueTAL suite of tools (54).

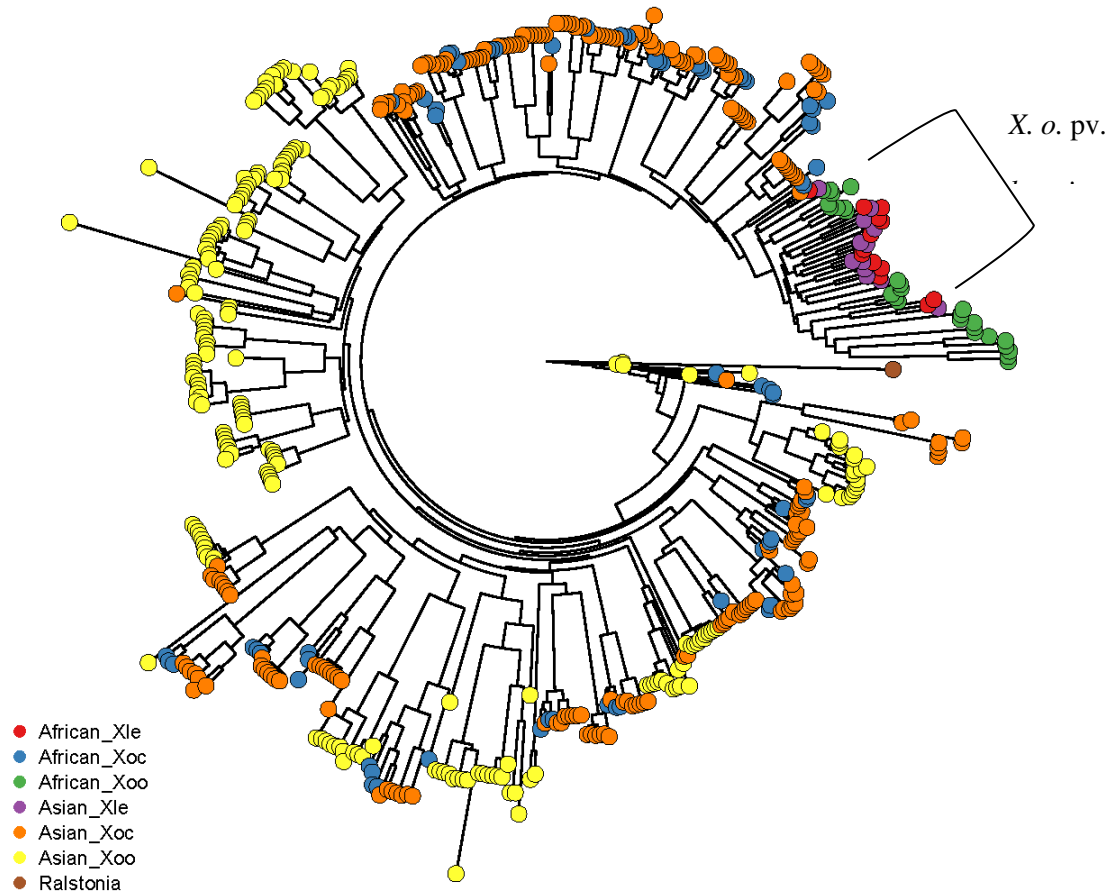

Figure 3.6 *X. o.* pv. *leersiae* TALE phylogeny within *X. oryzae* as determined by DisTAL (54). African Xle = *X. o.* pv. *leersiae* BAI23, Asian Xle = *X. o.* pv. *leersiae* NCPPB4346. Total data set was composed of 563 TALEs from 29 strains (16 *X. o.* pv. *oryzae*, 11 *X. o.* pv. *oryzicola* and two *X. o.* pv. *leersiae*).

Table 3.1. Bacterial strains used in phenotyping experiments.

Species	Strain	Origin	Host	Reference or Source
X. oryzae	X11-5A	USA	*Oryza sativa*	Triplett et al 2011
X. o. pv. *leersiae*	BAI23	Burkina Faso	Weeds	V. Verdier, Wonni et al 2014
X. o. pv. *leersiae*	NCPPB4346	China	*Leersia hexandra*	Triplett et al 2011, Wonni et al 2014
X. o. pv. *oryzae*	BAI3	Burkina Faso	*O. sativa*	V. Verdier
X. o. pv. *oryzae*	PXO99[A]	Philippines	*O. sativa*	J.E. Leach
X. o. pv. *oryzicola*	BLS256	Philippines	*O. sativa*	A. Bogdanove
X. o. pv. *oryzicola*	MAI10	Mali	*O. sativa*	V. Verdier

Table 3.2. Percent average nucleotide identity (ANI) calculated with JSpecies between *X. o.* pv. *leersiae* strains NCPPB4346, BAI23 relative to other representative members of the genus *Xanthomonas*

Species	Strain	Host	Origin	NCBI Accession	% ANI with NCPPB4346	% ANI with BAI23
Xanthomonas oryzae pv. *leersiae*	NCPPB4346	*Leersia hexandra*	China	GCA_001276975.1	-	99.2
X. o. pv. *leersiae*	BAI23	Weeds	Burkina Faso	NA	99.4	-
X. oryzae	X11-5A	*O. sativa*	USA	AFHK00000000	97.0	97.4
	X8-1A	*O. sativa*	USA	AFHL00000000	96.8	97.4
X. o. pv. *oryzae*	KACC10331	*O. sativa*	Korea	NC_006834	97.7	98.0
	MAFF311018	*O. sativa*	Japan	NC_007705	97.7	98.0
	NAI8	*O. sativa*	Niger	NZ_AYSX01000000	97.8	98.2
	PXO86	*O. sativa*	Philippines	NZ_CP007166	97.7	97.9
	PXO99^{A}	*O. sativa*	Philippines	NC_010717	97.6	97.9
X. o. pv. *oryzicola*	BLS256	*O. sativa*	Philippines	NC_017267	98.1	98.3
	CFBP7342	*O. sativa*	Burkina Faso	NZ_CP007221	98.2	98.6
X. albilineans	GPE PC73	*Saccharum* sp.	Guadeloupe	GCA_000087965.1	76.2	76.4
X. axonopodis pv. *manihotis*	CIO-1	*Manihot esculenta*	Colombia	AKCZ00000000	89.3	89.5
X. axonopodis pv. *vasculorum*	NCPPB900	*Saccharum* sp.	La Réunion	JPHD00000000	89.5	89.5
X. campestris pv. *campestris*	ATCC33913	*Brassica oleracea* var. *gemmifera*	United Kingdom	GCA_000007145.1	84.8	84.7
X. citri	A^{w}12879	*Citrus* spp.	USA	GCA_000349225.1	89.9	89.9
X. euvesicatoria	85-10	*Capsicum* sp.	USA	AM039952.1	89.9	90.0
X. fuscans pv. *fuscans*	CFBP4884	*Phaseolus vulgaris*	France	JPHG00000000	89.7	89.8
X. vasicola pv. *vasculorum*	NCPPB206	*Zea mays*	South Africa	AKBM00000000	91.2	91.3

Table 3.3. Type III effector content in *X. o.* pv. *oryzae*, *X. o.* pv. *oryzicola* and *X. o.* pv. *leersiae* where numbers represent copies of each effector gene, 'x' represents absence, '>' or 'IS' indicate a rearrangement or an IS element, respectively, in the gene.

	X. oryzae* pv. *oryzae	***X. oryzae* pv. *oryzicola***	***X. oryzae* pv. *leersiae***	
Type III Effector	PXO99[A]	BLS256	BAI23	NCPPB4346
avrBs1	x	x	x	x
avrBs2	1	1	1	1
avrBs3 (TALE)	19	29	13	12
xopA	1	1	1	1
xopB	x	x	x	x
xopC1	x	x	x	x
xopC2	1	1	1	1
xopD	x	x	1	1
xopE1	x	x	x	1
xopE2	x	x	1	x
xopE3	x	x	x	x
xopF1	1	1	1	1
xopF2	x	x	x	x
xopG1	1	x	x	x
xopG2	x	x	x	x
xopG3	x	x	x	x
xopH1	x	x	x	x
xopH2	x	x	x	x
xopI1	2	1	1	1
xopJ1	x	x	x	x
xopJ2	x	x	x	x
xopJ3	x	x	x	x
xopJ4	x	x	x	x
xopJ5	x	x	x	x
xopK	1	1	1	1
xopL1	1	1	1	1
xopN	1	1	1	2
xopO	x	1	x	x
xopP1	1	1	1	1
xopQ	1	1	1	1
xopR1	1	1	1	1
xopT	x	x	x	x
xopU	1	1	1	1

xopV	1	1	1	1
xopW	1	1	1, IS	x
xopX	1	1	1	1
xopY	1	1	1, >	1, >
xopZ1	2	1	1	1
xopAA	1	1	1	1
xopAB	1	1	1	1
xopAC	x	x	x	x
xopAD	1	1	1	1
xopAE	1	1	1	1
xopAF1	1	x	x	x
xopAG	x	x	x	x
xopAH1	1	1	1	1
xopAI	x	x	x	x
xopAJ	x	1	x	x
xopAK	1	x	1	1
xopAL1	x	x	x	x
xopAL2	x	x	x	x
xopAM	x	x	x	x
hpaA	1	1	1	1
hrpW	x	x	x	x
avrXccA1	x	x	x	x
avrXccA2	x	x	x	x

Table 3.4. Repeat number and repeat-variable di-residue (RVD) composition of the TALEs in *X. oryzae* pv. *leersiae.*

		RVDs
TALE[a]	**Repeats**	**BAI23**
1	16.5	NN-HD-HD-NH-NN-TI-NI-NG-HD-NI-NG-HD-HD-NG-NI-NG-NI
2	16.5	NN-HD-HD-NH-HD-NG-NI-NG-NI-NI-NN-HD-HD-HD-HD-HD-NG
3	17.5	NN-HD-NH-HD-NI-HD-NI-HD-NI-NN-HD-NS-HD-HD-HD-NG-HD-NG
4	16.5	NN-NN-HD-NI-HD-NI-HD-HD-NG-HD-HD-NG-HD-NI-HD-HD-NG
5a	13.5	NI-NN-HD-NG-HD-NN-HD-NG-NG-HD-NN-HD-NN-NG
5b	19.5	NN-HD-HD-NI-NN-HD-NI-NG-NN-HD-NI-NN-NG-HD-HD-NN-NG-NI-NN
6a	19.5	NN-HD-HD-NN-NN-NG-NI-NG-HD-NI-NI-NG-NS-NG-NI-HD-NI-HD-NG
6b	17.5	NN-HD-NI-HD-NG-NI-NG-NI-HD-NG-NH-HD-NI-NI-NN-NN-NG
6c	18.5	NS-HD-NH-NH-NG-NN-NN-NI-NI-NG-HD-NI-NI-NI-NN-NH-NH-NG
7	19.5	NN-HD-NI-HD-NN-NI-HD-HD-NI-NI-NG-NN-NI-NG-NH-NG-NG-NH-NG
8	22.5	NN-HD-NI-NG-NI-HD-HD-NH-NN-HD-HD-NS-NN-NG-NN-HD-NH-NG-HD-NN-HD-NG
9a	22.5	NS-HD-NI-HD-NI-HD-NI-HD-NN-HD-NI-HD-NN-NG-HD-NH-HD-NG-HD-NN-HD-NG
9b	19.5	NS-HD-NI-HD-NN-HD-HD-NI-NI-NG-HD-HD-NG-HD-NI-NN-NG-NG-NN
		NCPPB4346
1	18.5	NS-HD-NI-HD-NN-HD-HD-NI-NI-NG-HD-HD-NG-HD-NI-NN-NG-HD-NH
2	21.5	NN-HD-NI-NG-NI-HD-HD-NH-NN-HD-HD-NS-NN-NG-NN-HD-NH-NG-HD-NN-HD-NG
3	16.5	NN-HD-NI-HD-NN-NI-NI-HD-NI-HD-NG-NN-NI-NG-NN-NG-NG
4	16.5	NN-HD-NI-NH-HD-NG-NI-HD-NG-NI-NI-NN-HD-NS-HD-NH-NG
5a	16.5	NN-HD-HD-NH-HD-NG-NI-NG-NI-NI-NN-HD-HD-HD-HD-NH-NG
5b	17.5	NN-HD-NH-HD-NI-HD-NI-HD-NI-NN-HD-NS-HD-HD-HD-NG-HD-NG
6	16.5	NN-NN-HD-NS-HD-NI-NI-HD-NG-HD-HD-NG-HD-NI-HD-HD-NG
7a	18.5	NN-HD-HD-NI-NN-HD-NI-NG-NN-HD-NI-NN-NG-HD-HD-NN-NG-NI-NN
7b	14.5	NN-NI-NN-HD-NG-HD-NN-HD-NG-NG-HD-NN-HD-NN-NG
8a	14.5	NN-HD-HD-NG-NI-HD-NI-NG-HD-NI-NI-NN-NN-NH-NG
8b	16.5	NS-NH-HD-NN-NN-NG-NI-HD-NI-NG-HD-NI-NI-NN-NN-NH-NG
9	16.5	NN-HD-NI-HD-NG-NI-NG-NI-HD-NG-NN-HD-NI-NI-NN-NH-NG

[a]TALE nomenclature according to (61)

Table 3.5. Selected EBE content and gene orthologs from *L. perrieri* that are putative targets of *X. o.* pv. *leersiae* BAI23 TALEs.

TALE ID	EBE[a]	*L. perrieri* gene target[b]	Rank[c]	*O. sativa* ortholog[d]	*O. sativa* gene name
Tal5a	TAGCTCCCTTCGCCT	LPERR01G18870	14	LOC_Os01g42090	OsSWEET6b
Tal5a	TCGATCGATTCGCAT	LPERR10G07460	287	LOC_Os10g28440	sulfate transporter 3.1, putative, expressed
Tal5b	TGCCAACCAACCAACCATAG	LPERR07G11940	241	LOC_Os07g29750	glycosyl hydrolases family 16, putative, expressed
Tal5b	TGCAGACATGCAGTGCACAG	LPERR09G05320	166	LOC_Os09g20220	glutathione S-transferase, putative, expressed
Tal6a	TACCACCACCACTGTACTCT	LPERR06G20600	79	LOC_Os06g46500	monocopper oxidase, putative, expressed
Tal6a	TGTCGATATCAATCCAAACC	LPERR09G04350	91	LOC_Os09g32100	expressed protein
Tal6b	TGGGCTATACTGAAAAAT	LPERR03G02350	59	LOC_Os03g05370	expressed protein
Tal6b	TGGGCTCTACTGCCAAGT	LPERR03G04710	237	LOC_Os03g07540	bHLH family protein, putative, expressed
Tal6b	TTCACTATCCTGGAAAAT	LPERR01G15230	210	LOC_Os05g12450	hydroquinone glucosyltransferase, putative, expressed
Tal6b	TACCTTATCCTGAAAAGT	LPERR10G07460	230	LOC_Os10g28440	sulfate transporter 3.1, putative, expressed
Tal6c	TGCGGGAGAAGAAAATGGT	LPERR05G05960	361	LOC_Os05g12320	OsSWEET3a
Tal7	TACAAGAAAAATGATGTGTT	LPERR09G06570	377	LOC_Os09g23560	dehydrogenase, putative, expressed
Tal8	TGCACACACCCACGAGCGTCGCT	LPERR06G21810	236	LOC_Os05g35140	OsSWEET1b
Tal9a	CACACCCACGAGCGTCGCTTGAT	LPERR06G21810	217	LOC_Os05g35140	OsSWEET1b
Tal9a	TGCACACACCCACGAGCGTCGCT	LPERR06G21810	194	LOC_Os05g35140	OsSWEET1b
Tal9b	TGCACACAAACAGTAAATTG	LPERR09G12040	203	LOC_Os08g42350	OsSWEET11, Os8N3, Xa13

[a]EBE = effector binding element as determined by RVD sequence of *X. o.* pv. *leersiae* TALE in Talvez (54)
[b]TALE targets predicted in promoterome of *L. perrieri* genome v1.4, accession ALNV02000000 (4)
[c]Rank values calculated by a positional weight matrix (PWM) according to Talvez (38, 79)
[d]*O. sativa* cv. Nipponbare MSU v7 (55)

REFERENCES

1. **Ou SH**. 1985. Rice Diseases, 2nded. Association Applied Biology, Surrey, England.
2. **Mew TW**. 1987. Current status and future prospects of research on bacterial blight of rice. Annu Rev Phytopathol **25**:359–382.
3. **Guo Y-L**, **Ge S**. 2005. Molecular phylogeny of *Oryzeae (Poaceae)* based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. Am J Bot **92**:1548–58.
4. **Copetti D**, **Zhang J**, **El Baidouri M**, **Gao D**, **Wang J**, **Barghini E**, **Cossu RM**, **Angelova A**, **Maldonado L CE**, **Roffler S**, **Ohyanagi H**, **Wicker T**, **Fan C**, **Zuccolo A**, **Chen M**, **Costa de Oliveira A**, **Han B**, **Henry R**, **Hsing Y-I**, **Kurata N**, **Wang W**, **Jackson SA**, **Panaud O**, **Wing RA**. 2015. RiTE database: a resource database for genus-wide rice genomics and evolutionary biology. BMC Genomics **16**:538.
5. **Copetti D**. 2013. Genomic resources for Leersia perrieri: An outgroup species for the genus OryzaPlant and Animal Genome VIII. Plant and Animal Genome.
6. **dos Santos RS**, **Farias D da R**, **Pegoraro C**, **Rombaldi CV**, **Fukao T**, **Wing RA**, **de Oliveira AC**. 2017. Evolutionary analysis of the SUB1 locus across the *Oryza* genomes. Rice **10**:4.
7. **You S-H**, **Zhang X-H**, **Liu J**, **Zhu Y-N**, **Gu C**. 2013. Feasibility of constructed wetland planted with *Leersia hexandra* Swartz for removing Cr, Cu and Ni from electroplating wastewater. Environ Technol **35**:187–94.
8. **Liu J**, **Duan C**, **Zhang X**, **Zhu Y**, **Lu X**. 2011. Potential of *Leersia hexandra* Swartz for phytoextraction of Cr from soil. J Hazard Mater **188**:85–91.
9. **Wang D**, **Zhang X**, **Liu J**, **Zhu Y**, **Zhang H**, **Zhang A**, **Jin X**. 2012. Oxalic acid enhances Cr tolerance in the accumulating plant *Leersia hexandra* Swartz. Int J Phytoremediation **14**:966–77.
10. **Noda T**, **Yamamoto T**. 2008. Reaction of *Leersia* grasses to *Xanthomonas oryzae* pv. *oryzae* collected from Japan and Asian countries. J Gen Plant Pathol **74**:395–401.
11. **Gonzalez C**, **Xu G**, **Li H**, **Cosper JW**. 1991. *Leersia hexandra*, an alternate host for *Xanthomonas campestris* pv. *oryzae* in Texas. Plant Dis **75**:159–162.
12. **Triplett LR**, **Hamilton JP**, **Buell CR**, **Tisserat NA**, **Verdier V**, **Zink F**, **Leach JE**. 2011. Genomic analysis of *Xanthomonas oryzae* isolates from rice grown in the United States reveals substantial divergence from known *X. oryzae* pathovars. Appl Environ Microbiol **77**:3930–3937.
13. **Triplett L**, **Leach J**. 2016. Host mechanisms for resistance to TAL effectors: Thinking outside the UPT box. Physiol Mol Plant Pathol **95**:66–69.
14. **Poulin L**, **Grygiel P**, **Magne M**, **Gagnevin L**, **Rodriguez-R LM**, **Forero Serna N**, **Zhao S**, **El Rafii M**, **Dao S**, **Tekete C**, **Wonni I**, **Koita O**, **Pruvost O**, **Verdier V**, **Vernière C**, **Koebnik R**. 2015. New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. Appl Environ Microbiol **81**:688–98.
15. **Gonzalez C**, **Szurek B**, **Manceau C**, **Mathieu T**, **Sere Y**, **Verdier V**. 2007. Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. Mol Plant-Microbe Interact2007/05/18. **20**:534–546.
16. **Hajri A**, **Brin C**, **Zhao S**, **David P**, **Feng J**, **Koebnik R**, **Szurek B**, **Verdier V**, **Boureau T**, **Poussier S**. 2012. Multilocus sequence analysis and type III effector repertoire mining provide new insights into the evolutionary history and virulence of *Xanthomonas oryzae*. Mol Plant Pathol **13**:288–302.
17. **Zhao S**, **Poulin L**, **Rodriguez RL**, **Serna NF**, **Liu SY**, **Wonni I**, **Szurek B**, **Verdier V**, **Leach JE**, **He YQ**, **Feng JX**, **Koebnik R**. 2012. Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology2012/09/11. **102**:948–956.
18. **Parkinson N**, **Aritua V**, **Heeney J**, **Cowie C**, **Bew J**, **Stead D**. 2007. Phylogenetic analysis of

Xanthomonas species by comparison of partial gyrase B gene sequences. Int J Syst Evol Microbiol **57**:2881–2887.

19. **Fang CT**, **Ren HC**, **Chen TK**, **Chu YK**, **Faan HC**, **Wu SC**. 1957. A comparison of rice bacterial leaf blight organism with the bacterial leaf streak organism of rice and *Leersia hexandra* Swartz. RevApplMycol **37**:475.
20. **Wonni I**, **Cottyn B**, **Detemmerman L**, **Dao S**, **Ouedraogo L**, **Sarra S**, **Tekete C**, **Poussier S**, **Corral R**, **Triplett L**, **Koita O**, **Koebnik R**, **Leach J**, **Szurek B**, **Maes M**, **Verdier V**. 2014. Analysis of *Xanthomonas oryzae* pv. *oryzicola* population in Mali and Burkina Faso reveals a high level of genetic and pathogenic diversity. Phytopathology **104**:520–31.
21. **White FF**, **Potnis N**, **Jones JB**, **Koebnik R**. 2009. The type III effectors of *Xanthomonas*. Mol Plant Pathol **10**:749–766.
22. **Jacques M-A**, **Arlat M**, **Boulanger A**, **Boureau T**, **Carrère S**, **Cesbron S**, **Chen NWG**, **Cociancich S**, **Darrasse A**, **Denancé N**, **Fischer-Le Saux M**, **Gagnevin L**, **Koebnik R**, **Lauber E**, **Noël LD**, **Pieretti I**, **Portier P**, **Pruvost O**, **Rieux A**, **Robène I**, **Royer M**, **Szurek B**, **Verdier V**, **Vernière C**. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. Annu Rev Phytopathol **54**:163–187.
23. **Boch J**, **Scholze H**, **Schornack S**, **Landgraf A**, **Hahn S**, **Kay S**, **Lahaye T**, **Nickstadt A**, **Bonas U**. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science (80-) **326**:1509–1512.
24. **Bogdanove AJ**, **Voytas DF**. 2011. TAL effectors: customizable proteins for DNA targeting. Science (80-) **333**:1843–1846.
25. **Moscou MJ**, **Bogdanove AJ**. 2009. A simple cipher governs DNA recognition by TAL effectors. Science (80-) **326**:1501.
26. **Boch J**, **Bonas U**, **Lahaye T**. 2014. TAL effectors - pathogen strategies and plant resistance engineering. New Phytol **204**:823–832.
27. **Hutin M**, **Pérez-Quintero AL**, **Lopez C**, **Szurek B**. 2015. MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. Front Plant Sci **6**:535.
28. **Verdier V**, **Triplett LR**, **Hummel AW**, **Corral R**, **Cernadas RA**, **Schmidt CL**, **Bogdanove AJ**, **Leach JE**. 2012. Transcription activator-like (TAL) effectors targeting *OsSWEET* genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. New Phytol2012/10/20. **196**:1197–1207.
29. **Ryba-White M**, **Notteghem JL**, **Leach JE**. 1995. Comparison of *Xanthomonas oryzae* pv. *oryzae* strains from Africa, North America, and Asia by restriction fragment length polymorphism analysis. Int Rice Res Newsl **20**:25–26.
30. **Peng Z**, **Hu Y**, **Xie J**, **Potnis N**, **Akhunova A**, **Jones J**, **Liu Z**, **White FF**, **Liu S**. 2016. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. BMC Genomics **17**:21.
31. **Eid J**, **Fehr A**, **Gray J**, **Luong K**, **Lyle J**, **Otto G**, **Peluso P**, **Rank D**, **Baybayan P**, **Bettman B**, **Bibillo A**, **Bjornson K**, **Chaudhuri B**, **Christians F**, **Cicero R**, **Clark S**, **Dalal R**, **Dewinter A**, **Dixon J**, **Foquet M**, **Gaertner A**, **Hardenbol P**, **Heiner C**, **Hester K**, **Holden D**, **Kearns G**, **Kong X**, **Kuse R**, **Lacroix Y**, **Lin S**, **Lundquist P**, **Ma C**, **Marks P**, **Maxham M**, **Murphy D**, **Park I**, **Pham T**, **Phillips M**, **Roy J**, **Sebra R**, **Shen G**, **Sorenson J**, **Tomaney A**, **Travers K**, **Trulson M**, **Vieceli J**, **Wegener J**, **Wu D**, **Yang A**, **Zaccarin D**, **Zhao P**, **Zhong F**, **Korlach J**, **Turner S**. 2009. Real-time DNA sequencing from single polymerase molecules. Science **323**:133–8.
32. **Wilkins KE**, **Booher NJ**, **Wang L**, **Bogdanove AJ**. 2015. TAL effectors and activation of predicted host targets distinguish Asian from African strains of the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance of five TAL effectors. Front Plant Sci **6**:536.
33. **Booher NJ**, **Carpenter SCD**, **Sebra RP**, **Wang L**, **Salzberg SL**, **Leach JE**, **Bogdanove AJ**. 2015. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic

structure and complex TAL (transcription activator-like) effector gene relationships. Microb Genomics **1**.

34. **Quibod IL**, **Perez-Quintero A**, **Booher NJ**, **Dossa GS**, **Grande G**, **Szurek B**, **Vera Cruz C**, **Bogdanove AJ**, **Oliva R**. 2016. Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. Sci Rep **6**:34137.
35. **Grau J**, **Reschke M**, **Erkes A**, **Streubel J**, **Morgan RD**, **Wilson GG**, **Koebnik R**, **Boch J**. 2016. AnnoTALE: bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. Sci Rep **6**:21077.
36. **Grau J**, **Wolf A**, **Reschke M**, **Bonas U**, **Posch S**, **Boch J**. 2013. Computational predictions provide insights into the biology of TAL effector target sites. PLoS Comput Biol **9**:e1002962.
37. **Doyle EL**, **Booher NJ**, **Standage DS**, **Voytas DF**, **Brendel VP**, **VanDyk JK**, **Bogdanove AJ**. 2012. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res **40**:W117–W122.
38. **Pérez-Quintero AL**, **Rodriguez-R LM**, **Dereeper A**, **López C**, **Koebnik R**, **Szurek B**, **Cunnac S**. 2013. An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. PLoS One **8**:e68464.
39. **Booher NJ**, **Bogdanove AJ**. 2014. Tools for TAL effector design and target prediction. Methods **69**:121–7.
40. **Yang J**, **Zhang Y**, **Yuan P**, **Zhou Y**, **Cai C**, **Ren Q**, **Wen D**, **Chu C**, **Qi H**, **Wei W**. 2014. Complete decoding of TAL effectors for DNA recognition. Cell Res **24**:628–631.
41. **Yang B**, **Sugio A**, **White FF**. 2006. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci USA **103**:10503–10508.
42. **Li T**, **Huang S**, **Zhou J**, **Yang B**. 2013. Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. Mol Plant.
43. **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 Phytol **200**:808–19.
44. **Wang C**, **Zhang X**, **Fan Y**, **Gao Y**, **Zhu Q**, **Zheng C**, **Qin T**, **Li Y**, **Che J**, **Zhang M**, **Yang B**, **Liu Y**, **Zhao K**. 2015. XA23 is an executor R protein and confers broad-spectrum disease resistance in rice. Mol Plant **8**:290–302.
45. **Gu K**, **Yang B**, **Tian D**, **Wu L**, **Wang D**, **Sreekala C**, **Yang F**, **Chu Z**, **Wang GL**, **White FF**, **Yin Z**. 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. Nature **435**:1122–1125.
46. **Tian D**, **Wang J**, **Zeng X**, **Gu K**, **Qiu C**, **Yang X**, **Zhou Z**, **Goh M**, **Luo Y**, **Murata-Hori M**, **White FF**, **Yin Z**. 2014. The rice TAL effector–dependent resistance protein Xa10 triggers cell death and calcium depletion in the endoplasmic reticulum. Plant Cell **26**:497–515.
47. **Lang JM**, **Langlois P**, **Nguyen MHR**, **Triplett LR**, **Purdie L**, **Holton TA**, **Djikeng A**, **Vera Cruz CM**, **Verdier V**, **Leach JE**. 2014. Sensitive detection of *Xanthomonas oryzae* pathovars oryzae and oryzicola by loop-mediated isothermal amplification. Appl Environ Microbiol **80**:4519–30.
48. **Hoagland DR**, **Arnon DI**. 1950. The water-culture method for growing plants without soil. Berkeley, Calif. : College of Agriculture, University of California, Berkeley, CA.
49. **Lang JM**, **Hamilton JP**, **Diaz MGQ**, **Van Sluys MA**, **Burgos MRG**, **Vera Cruz CM**, **Buell CR**, **Tisserat NA**, **Leach JE**. 2010. Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. Plant Dis **94**:311–319.
50. **Reimers PJ**, **Leach JE**. 1991. Race-specific resistance to *Xanthomonas oryzae* pv. *oryzae* conferred by bacterial blight resistance gene *Xa-10* in rice *Oryza sativa* involves accumulation of a lignin-like substance in host tissues. Physiol Mol Plant Pathol **38**:39–55.
51. **Kauffman HE**, **Reddy APK**, **Hsieh SPY**, **Merca SD**. 1973. Improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis Report **57**:537–541.
52. **Bankevich A**, **Nurk S**, **Antipov D**, **Gurevich AA**, **Dvorkin M**, **Kulikov AS**, **Lesin VM**,

Nikolenko SI, **Pham S**, **Prjibelski AD**, **Pyshkin A V.**, **Sirotkin A V.**, **Vyahhi N**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455.

53. **Barak JD**, **Vancheva T**, **Lefeuvre P**, **Jones JB**, **Timilsina S**, **Minsavage G V.**, **Vallad GE**, **Koebnik R**. 2016. Whole-genome sequences of *Xanthomonas euvesicatoria* strains clarify taxonomy and reveal a stepwise erosion of type 3 effectors. Front Plant Sci **7**:1805.
54. **Pérez-Quintero AL**, **Lamy L**, **Gordon J**, **Escalon A**, **Cunnac S**, **Szurek B**, **Gagnevin L**. 2015. QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. Front Plant Sci **6**.
55. **Kawahara Y**, **de la Bastide M**, **Hamilton JP**, **Kanamori H**, **McCombie WR**, **Ouyang S**, **Schwartz DC**, **Tanaka T**, **Wu J**, **Zhou S**, **Childs KL**, **Davidson RM**, **Lin H**, **Quesada-Ocampo L**, **Vaillancourt B**, **Sakai H**, **Lee SS**, **Kim J**, **Numa H**, **Itoh T**, **Buell CR**, **Matsumoto T**. 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice **6**:4.
56. **Richter M**, **Rosselló-Móra R**. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A **106**:19126–31.
57. **Konstantinidis KT**, **Tiedje JM**. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci USA **102**.
58. **Goris J**, **Konstantinidis KT**, **Klappenbach JA**, **Coenye T**, **Vandamme P**, **Tiedje JM**. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol **57**:81–91.
59. **Rodriguez-r LM**, **Konstantinidis KT**. 2016. The enveomics collection : a toolbox for specialized analyses of microbial genomes and metagenomes. Peer J Prepr.
60. **Cernadas RA**, **Doyle EL**, **Niño-Liu DO**, **Wilkins KE**, **Bancroft T**, **Wang L**, **Schmidt CL**, **Caldo R**, **Yang B**, **White FF**, **Nettleton D**, **Wise RP**, **Bogdanove AJ**. 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.
61. **Salzberg SL**, **Sommer DD**, **Schatz MC**, **Phillippy AM**, **Rabinowicz PD**, **Tsuge S**, **Furutani A**, **Ochiai H**, **Delcher AL**, **Kelley D**, **Madupu R**, **Puiu D**, **Radune D**, **Shumway M**, **Trapnell C**, **Aparna G**, **Jha G**, **Pandey A**, **Patil PB**, **Ishihara H**, **Meyer DF**, **Szurek B**, **Verdier V**, **Koebnik R**, **Dow JM**, **Ryan RP**, **Hirata H**, **Tsuyumu S**, **Won Lee S**, **Seo Y-S**, **Sriariyanum M**, **Ronald PC**, **Sonti R V**, **Van Sluys M-A**, **Leach JE**, **White FF**, **Bogdanove AJ**. 2008. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics **9**:204.
62. **Yang B**, **Sugio A**, **White FF**. 2006. Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci U S A **103**:10503–8.
63. **Schwartz AR**, **Morbitzer R**, **Lahaye T**, **Staskawicz BJ**. 2017. TALE-induced bHLH transcription factors that activate a pectate lyase contribute to water soaking in bacterial spot of tomato. Proc Natl Acad Sci U S A **114**:E897–E903.
64. **Arlat M**, **Gough C**, **Barber C**, **Boucher C**, **Daniels M**. 1991. *Xanthomonas campestris* contains a cluster of hrp genes related to larger hrp cluster of *Pseudomonas solancearum*. MolPlant-Microbe Interact **4**:593–601.
65. **Grant SR**, **Fisher EJ**, **Chang JH**, **Mole BM**, **Dangl JL**. 2006. Subterfuge and Manipulation: Type III Effector Proteins of Phytopathogenic Bacteria. Annu Rev Microbiol **60**:425–449.
66. **Bogdanove AJ**, **Koebnik R**, **Lu H**, **Furutani A**, **Angiuoli S V**, **Patil PB**, **Van Sluys MA**, **Ryan RP**, **Meyer DF**, **Han SW**, **Aparna G**, **Rajaram M**, **Delcher AL**, **Phillippy AM**, **Puiu D**, **Schatz MC**, **Shumway M**, **Sommer DD**, **Trapnell C**, **Benahmed F**, **Dimitrov G**, **Madupu R**, **Radune D**, **Sullivan S**, **Jha G**, **Ishihara H**, **Lee SW**, **Pandey A**, **Sharma V**, **Sriariyanun M**, **Szurek B**, **Vera-Cruz CM**, **Dorman KS**, **Ronald PC**, **Verdier V**, **Dow JM**, **Sonti R V**, **Tsuge S**, **Brendel VP**, **Rabinowicz PD**, **Leach JE**, **White FF**, **Salzberg SL**. 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. J

Bacteriol2011/07/26. **193**:5450–5464.
67. **Schwartz AR**, **Potnis N**, **Timilsina S**, **Wilson M**, **Patané J**, **Martins J**, **Minsavage G V**, **Dahlbeck D**, **Akhunova A**, **Almeida N**, **Vallad GE**, **Barak JD**, **White FF**, **Miller SA**, **Ritchie D**, **Goss E**, **Bart RS**, **Setubal JC**, **Jones JB**, **Staskawicz BJ**. 2015. Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. Front Microbiol **6**:535.
68. **Kim J-G**, **Stork W**, **Mudgett MB**. 2013. Xanthomonas type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. Cell Host Microbe **13**:143–54.
69. **Kim J-G**, **Taylor KW**, **Hotson A**, **Keegan M**, **Schmelz EA**, **Mudgett MB**. 2008. XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in *Xanthomonas*-infected tomato leaves. Plant Cell **20**:1915–29.
70. **Harrison J**, **Studholme DJ**. 2014. Draft genome sequence of *Xanthomonas axonopodis* pathovar *vasculorum* NCPPB 900. FEMS Microbiol Lett **360**:113–116.
71. **Jacquemin J**, **Laudié M**, **Cooke R**. 2009. A recent duplication revisited: phylogenetic analysis reveals an ancestral duplication highly-conserved throughout the *Oryza* genus and beyond. BMC Plant Biol **9**:146.
72. **Xiao W**, **Liu H**, **Li Y**, **Li X**, **Xu C**, **Long M**, **Wang S**. 2009. A rice gene of de novo origin negatively regulates pathogen-induced defense response. PLoS One **4**:e4603.
73. **Chen L-Q**, **Hou B-H**, **Lalonde S**, **Takanaga H**, **Hartung ML**, **Qu X-Q**, **Guo W-J**, **Kim J-G**, **Underwood W**, **Chaudhuri B**, **Chermak D**, **Antony G**, **White FF**, **Somerville SC**, **Mudgett MB**, **Frommer WB**. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. Nature **468**:527–532.
74. **Cohn M**, **Bart RS**, **Shybut M**, **Dahlbeck D**, **Gomez M**, **Morbitzer R**, **Hou B-H**, **Frommer WB**, **Lahaye T**, **Staskawicz BJ**. 2014. *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector mediated induction of a SWEET sugar transporter in cassava. Mol Plant-Microbe Interact **27**:1186–1198.
75. **Hu Y**, **Zhang J**, **Jia H**, **Sosso D**, **Li T**, **Frommer WB**, **Yang B**, **White FF**, **Wang N**, **Jones JB**. 2014. Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. Proc Natl Acad Sci U S A **111**:E521-9.
76. **Kumar S**, **Asif MH**, **Chakrabarty D**, **Tripathi RD**, **Trivedi PK**. 2011. Differential expression and alternative splicing of rice sulphate transporter family members regulate sulphur status during plant growth, development and stress conditions. Funct Integr Genomics **11**:259–273.
77. **Takahashi H**, **Buchner P**, **Yoshimoto N**, **Hawkesford MJ**, **Shiu S-H**. 2011. Evolutionary relationships and functional diversity of plant sulfate transporters. Front Plant Sci **2**:119.
78. **Hajri A**, **Brin C**, **Hunault G**, **Lardeux F**, **Lemaire C**, **Manceau C**, **Boureau T**, **Poussier S**. 2009. A “Repertoire for Repertoire” hypothesis: Repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. PLoS One **4**:e6632.
79. **Megraw M**, **Hatzigeorgiou AG**. 2010. MicroRNA promoter analysis.Methods in molecular biology (Clifton, N.J.).

CHAPTER 4

DETECTION AND CHARACTERIZATION OF *XANTHOMONAS VASICOLA* PV. *VASCULORUM* COMB. NOV. (COBB 1894) CAUSING BACTERIAL LEAF STREAK OF CORN IN THE UNITED STATES[5,6]

4.1 INTRODUCTION

Corn (*Zea mays*) is a staple crop worldwide and is the most widely produced feed grain in the United States. In 2014, symptoms of bacterial leaf streak disease were first observed on corn in Nebraska and by 2016, the disease was reported in Colorado, Iowa, Kansas, Minnesota, Oklahoma, South Dakota and Texas (1). Given similarity of symptoms to those caused by other corn pathogens, it is not known how long the disease has been present in the USA. Bacterial leaf streak was first described in 1949 on corn in South Africa (2), but prior to 2017 it had not been documented in the USA (1). Symptoms occur on leaves of field (dent), sweet corn, seed corn and popcorn crops, appearing as dark, water-soaked, linear lesions with wavy margins confined to the interveinal spaces. Due to the importance of corn in the USA, and the implications of the emergence and spread of a new disease, accurate identification of the causal agent and determination of its relationship to the strains from South Africa are of critical importance to the corn industry. Impacts on yield loss due to this disease are unknown.

The pathogen causing corn bacterial leaf streak was first named *Xanthomonas campestris* pv. *vasculorum* (3), a species and pathovar that also included bacteria causing gumming disease of sugarcane and palm. In early reports, *X. c.* pv. *vasculorum* strains isolated from sugarcane and palm were pathogenic

[5] Accepted for publication as 'Detection and characterization of *Xanthomonas vasicola* pv. *vasculorum* (Cobb 1894) comb. nov. causing bacterial leaf streak of corn in the United States' by **Lang, J.M**., E. DuCharme, J. Ibarra Caballero, E. Luna, T. Hartman, M. Ortiz-Castro, K. Korus, J. Rascoe, T.A. Jackson-Ziems, K. Broders, and J.E. Leach in *Phytopathology* https://doi.org/10.1094/PHYTO-05-17-0168-R

[6] Contributions by J.M. Lang: Design of experiments; all pathogenicity screening; molecular diagnostics; MLSA and phylogenetic analysis; draft genome comparisons; statistical analysis; wrote manuscript

on corn, sorghum and sugarcane (4), while South African isolates from corn were virulent to corn only and not sorghum or sugarcane (5). The South African corn bacterial leaf streak isolates were further distinguished from *X. c.* pv. *holcicola* (causal agent of sorghum bacterial leaf streak) and sugarcane isolates of *X. c.* pv. *vasculorum* by restriction fragment length polymorphism (RFLP), and based on these differences, Qhobela et al. (1990) and Coutinho (1988) proposed renaming the corn isolates to *X. c.* pv. *zeae*, to distinguish them from sugarcane *X. c.* pv. *vasculorum* isolates and from *X. c.* pv. *holcicola*. Later, based on DNA-DNA hybridization (6) and fatty acid profiling (7), the species *X. vasicola* was proposed for *X. c.* pv. *zeae*, *X. c.* pv. *vasculorum* and *X. c.* pv. *holcicola*, with some sugarcane isolates being separated into a second species, *X. axonopodis*. More recent reports demonstrated that the corn bacterial leaf streak strain NCPPB 206 caused disease on both corn and sugarcane (8). This and other studies that included phylogeny based on multilocus sequence analyses (MLSA) proposed designation of the causal agents of gumming disease of sugarcane and bacterial leaf streak of corn as *X. vasicola* pv. *vasculorum* (6–14). While there is consensus among the reports, the pathovar naming was not proposed according to the rules of the ISPP-CTPPB 'International Standards for Naming Pathovars of Plant Pathogenic Bacteria' (15, 16). Thus, nomenclature of the *X. vasicola* strains that cause bacterial leaf streak of corn and gumming disease of sugarcane were still unresolved.

MLSA and whole genome comparisons are now a regular accessory for classification of bacteria (17–20). Calculations of average nucleotide identity (ANI) from draft genomes are a widely accepted baseline for taxonomic placement of prokaryotes (21). While these tools are useful for placement of organisms into a common species (22, 23), the assignment of plant pathogens as pathovars still requires determining an organism's capacity to cause disease on reciprocal hosts compared with pathotype strains of the various pathovars (24). This is complicated if a pathogen can infect multiple hosts, as in the case of the *X. vasicola* complex.

In addition to resolving nomenclature, reliable and robust tools for accurate and rapid identification of the corn bacterial leaf streak pathogen were needed to confirm presence of the pathogen, monitor its spread, and develop management practices. PCR-based diagnostic assays are available to

identify *X. vasicola* to species level, but these tests do not differentiate *X. v.* pv. *vasculorum* from *X. v.* pv. *holcicola* (25, 26). Using the rich genome sequence data available for members of the genus *Xanthomonas*, unique sequences identified through comparative genomic approaches have enabled development of diagnostic assays for various *Xanthomonas* species and, in some cases, even pathovars (20, 27–30). Thus, leveraging genomics is a powerful approach to developing diagnostic tools for the rapid and accurate disease diagnosis needed to inform disease mitigation strategies and regulatory entities.

In this study, we address the identity of the pathogen causing the newly found bacterial leaf streak of corn in the USA. We used MLSA and comparative genomic approaches to compare corn bacterial leaf streak isolates from the USA and South Africa to determine phylogenetic relationships to other *Xanthomonas* species and pathovars. We performed greenhouse inoculations of corn, sugarcane, and sorghum to determine host range of the USA strains. Based on this work, we propose the USA strains causing bacterial leaf streak of corn, formerly named *X. c.* pv. *vasculorum, X. c.* pv. *zeae* or *X. vasicola*, be named *X. vasicola* pv. *vasculorum* (Cobb 1894) comb. nov. Finally, we developed primers for PCR-based diagnostic assays that distinguish the corn bacterial leaf streak pathogen, *X. v.* pv. *vasculorum,* and the sorghum bacterial leaf streak pathogen, *X. v.* pv. *holcicola,* from each other and from other *Xanthomonas* species.

4.2 MATERIALS AND METHODS

4.2.1 Multi-Locus Sequence Analysis

Bacterial strains (Table 4.1) were grown overnight at 28°C on nutrient agar (Becton, Dickinson and Company, Franklin Lakes, New Jersey) to reduce production of extracellular polysaccharides that impact quality of DNA extractions. Genomic DNA for all strains in this study was prepared using the Easy DNA kit (Life Technologies, Grand Island, NY) according to manufacturer's recommendations except the final product was recovered in 50 µl sterile molecular grade water. If public genome or partial sequence were not available (Tables 4.1 and 4.2), six genes (*atpD, dnaK, gyrB, fusA, lepA* and *rpoD*) were amplified from genomic DNAs using previously described primers (17, 31, 32). PCR products were purified using the Qiaquick PCR Purification kit (Qiagen, Valencia, CA) according to manufacturer's

instructions, but eluted in 30 µl sterile water, then directly sequenced at Genewiz (San Francisco, CA). Generated sequences were aligned, trimmed and concatenated in MEGA7 (33). Partial sequences totaling 3,927 bp were used for multi-locus sequence analyses (MLSA).

4.2.2 Phylogenetic Analysis

Evolutionary history was inferred by using the Maximum Likelihood method based on the Equal Input model (34). Bootstrap values were generated from 1,000 replicates. Phylogenetic trees were constructed using the Neighbor Joining method (35). All phylogenetic analyses were performed in MEGA7 (33). Comprehensive sequences and phylogenetic data were submitted to TreeBase under study number # S20566.

4.2.3 Genome sequencing, assembly and comparison

Draft genome sequences were generated for *X. v.* pv. *vasculorum* strains NE744 (Holt, Co., Nebraska, USA) and the historic South African strain *X. v.* pv. *vasculorum* SAM119 (5) using DNA extracted as described above. Genomic libraries were prepared at Michigan State University's Research Technology Core Facility (East Lansing, MI) using the Illumina TruSeq Nano DNA Library Preparation kit. Completed libraries were quality checked and quantified using a combination of Qubit dsDNA HS, Caliper LabChipGX HS DNA and Kapa Illumina Library Quantification qPCR assays. Libraries were pooled and loaded on an Illumina MiSeq standard v2 flow cell; sequencing was performed in a 2 x 75 bp paired end format for NE744 and a 2 x 250 bp paired end format for SAM119 using a v2 500 cycle MiSeq reagent cartridge. Base calling was done by Illumina Real Time Analysis (RTA) v1.18.64 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v1.8.4. Genomes were assembled using SPAdes v3.9.0 (36). QUAST was used to assess assembly qualities (37). Assembled genomes were uploaded to PATRIC (38) for preliminary annotation using a RASTtk pipeline (39).

4.2.4 Disease phenotyping

Corn (hybrid DKC 61-88) and sorghum (cv. Mycogen IG588) were grown in a 1:1 mix of Promix-BX Mycorrhizae (Quakertown, PA) and Greens Grade (Profile Products, LLC, Buffalo Grove,

IL), then inoculated four weeks after planting with eight selected bacterial strains (Table 1). Sugarcane (L-99-226) nodes were grown in a 1:1 mix of Promix-BX Mycorrhizae and sand for two weeks then transplanted to 100% Promix-BX Mycorrhizae. Each bacterial strain was cultured in peptone sucrose agar (PSA) (40) for 24 h at 28°C and then suspended to 10^8 CFU ml^{-1} in sterile, distilled water. Bacterial suspensions were infiltrated into the intercellular spaces of corn, sorghum, and sugarcane leaves on either side of the abaxial main vein with a needleless 1 cc syringe (41) and by leaf clipping as previously described (42). For stem inoculations, 500 µl of a bacterial suspension was injected into stems using a 21-gauge needle (43). Distilled water was included as a negative control in all inoculations. At least two leaves were inoculated on three to six individual plants. All inoculations with each isolate to host combination were repeated at least two times. Plants were maintained in a greenhouse (27 ± 1°C, 16 h day length, and 80% relative humidity). At 7 days post inoculation (dpi), lesions on infiltrated and stab-inoculated plants were measured; the rating scale proposed by (44) was applied for stab-inoculated plants.

4.2.5 Molecular diagnostic assay development

Draft genomes were collected from NCBI's Genome database for all strains listed in Table 2 with the exception of two sequenced genomes generated in this study. Unique loci that were conserved in all *X. v.* pv. *vasculorum* genomes, but either not present in *X. v.* pv. *holcicola* or *X. c.* pv. *musacearum* or were polymorphic in these strains relative to *X. v.* pv. *vasculorum*, were targeted for development of diagnostic primers. Design was done using an in-house genome alignment-based computational pipeline: UniqPrimer (Triplett et al, in preparation, code available upon request). UniqPrimer runs a comparative analysis to identify unique loci in the included genomes that are absent in excluded genomes and then design primers for these regions. Draft genome sequences of all publicly available *X. v.* pv. *vasculorum* and genomes generated in this study were set as 'include' genomes in UniqPrimer while *X. v.* pv. *holcicola*, *X. c.* pv. *musacearum*, other xanthomonads as well as phylogenetic outliers were set as 'exclude' genomes. Default settings were used in this design and yielded over 50 potential diagnostic primer sets; however, outputs are not ranked nor curated. Specificity of primer targets was validated *in silico* by Primer-BLAST (45) against the NCBI Whole Genome Shotgun (wgs) database where any non-

X. v. pv. *vasculorum* hits were discarded. Next, a robust set of primers located across the genome were synthesized (IDT, Coralville, IA) and screened in the lab first with a small panel of DNA from positive and negative control strains. Primer sets that successfully amplified only positive controls were then evaluated on the remaining *X. v.* pv. *vasculorum* isolates from diverse hosts and a set of negative control isolates comprised of several different *Xanthomonas* species and other genera of phytobacteria by conventional PCR as previously described (Lang et al. 2010, 2014). Negative controls pools were also separately spiked with a positive control DNA to ensure detection. Additionally, genome comparisons revealed variation among *X. v.* pv. *vasculorum* isolates in the presence/absence of the *xopAF* gene that encodes the bacterial effector XopAF (11, 13). Primers were designed using Geneious 8.0.5 (http://www.geneious.com/) to amplify a 362 bp fragment of *xopAF* by aligning *xopAF* sequences of five *X. v.* pv. *vasculorum* isolates (NCPPB accessions 702, 890, 895, 1326, and 1381) retrieved from the GenBank wgs database. Each 25 µl reaction included 1 µl of each primer at 10 µM (Table 3), 15.7 µl water, 5 µl 5x GoTaq reaction buffer, 0.75 µl 25mM $MgCl_2$, 0.5 µl 10mM dNTP, and 0.05 µl GoTaq DNA polymerase (Promega, Madison, WI). Optimized cycling conditions were an initial denaturation at 94ºC for 3 min followed by 30 cycles of 94ºC for 30 sec, annealing temperature specific to each primer (Table 3) for 30 sec, 72ºC for 1 min, with a final elongation at 72ºC for 10 min.

4.2.6 PCR-based detection of X. v. pv. vasculorum in corn leaves

To detect the presence of *X. v.* pv. *vasculorum* in corn leaf tissue, a small section (~3 cm) with a characteristic lesion was excised to include tissue beyond the lesion margin. The leaf tissue was surface disinfested in 5 ml of fresh 10% bleach for 30 sec, with vigorous shaking. Samples were rinsed three times with 10-15 ml sterile distilled water for 30 sec. Using flame sterilized forceps, the tissue was placed in 1 ml of sterile distilled water in a sterile 1.5 ml microcentrifuge tube. The tissue was cut several times to promote release of bacteria with flame-sterilized scissors, then incubated at room temperature for at least 1 h. For bacterial isolation, one loop-full (10 µl) of solution was spread onto nutrient agar and incubated at 28°C for two days. Single characteristic yellow colonies were selected, restreaked for isolated colonies on nutrient agar, and then a single colony was selected, suspended in water for use in

colony PCR. Alternatively, *X. v.* pv. *vasculorum* was directly detected using 1 µl of leachate from the cut tissue as the DNA template for direct PCR.

4.3 RESULTS

4.3.1 MLSA

DNA gyrase B is an accepted benchmark to delineate *Xanthomonas* species (46, 47). A preliminary alignment of the DNA gyrase B gene (*gyrB*) of *Xanthomonas* strains representing over 25 different described species and *Stenotrophomonas maltophilia* K279a as an outgroup was used to evaluate the placement of recent isolates of *X. v.* pv. *vasculorum* from corn in the genus *Xanthomonas* (Supplementary Fig. 1). Strains NE744 and NE181 isolated in 2015 from dent corn in two different counties of NE, USA (Holt and Cedar, respectively) grouped with *X. v.* pv. *holcicola* and *X. v.* pv. *vasculorum* from corn and sugarcane as well as *X. c.* pv. *musacearum* from banana, consistent with its identification as a *X. vasicola* (1).

MLSA using six different housekeeping genes (*atpD, dnaK, gyrB, fusA, lepA* and *rpoD*) established genotypic relationships of the USA corn strains with other *Xanthomonas. Stenotrophomonas maltophilia* K279A and *X. albilineans* GPEPC73 (causal agent of leaf scald of sugarcane) were included in the analysis as outgroups (Fig. 4.1). *X. v.* pv. *vasculorum* and *X. c.* pv. *musacearum* group together in one clade, with *X. v.* pv. *holcicola* next in proximity (Fig. 4.1). The sequences for the six genes were identical for the South African strain SAM119 and all other *X. v.* pv. *vasculorum* strains; thus, they grouped into a single clade. Three single nucleotide polymorphisms (SNPs) were identified in the six concatenated genes between the *X. v.* pv. *vasculorum* and *X. c.* pv. *musacearum* strains. Fifteen SNPs differentiated *X. v.* pv. *holcicola* strains from *X. v.* pv. *vasculorum* and *X. c.* pv. *musacearum.* Curated sequences and alignments are in TreeBASE under study # S20566.

4.3.2 Genomics

Draft genomes of *X. v.* pv. *vasculorum* NE744 and SAM119 were assembled using SPAdes v3.9.0 (36) yielding 104 and 84 contigs ≥ 500 bp (N_{50} between 139,700 and 169,893), for total lengths of 4,869,712 and 4,856,397 bp for *X. v.* pv. *vasculorum* NE744 and SAM119, respectively. Assembled

contigs and raw reads for *X. v.* pv. *vasculorum* NE744 and SAM119 were deposited to Genbank and the Sequence Read Archive with accession numbers MVYW00000000 and MVYX00000000. GC content for these strains were 63.25 and 63.22%, which is consistent with other members of the genus *Xanthomonas*.

Genome assemblies were used to calculate average nucleotide identity (ANI) to delineate species relative to other *Xanthomonas* at a 95% cut off using JSpecies v1.2.1 (22, 23) (Table 4.2). The genomes of the causal agent of corn bacterial leaf streak isolated in NE, USA are 99% identical to *X. v.* pv. *vasculorum* from Madagascar, South Africa and Zimbabwe, whether isolated from corn or sugarcane. The highest ANI percentage was between *X. v.* pv. *vasculorum* NE744 and SAM119. Both genomes generated in this study were on average 98% similar to *X. c.* pv. *musacearum* from banana (Uganda or Ethiopia) and *X. v.* pv. *holcicola* isolated from sorghum or *Holcus* sp. (USA, Australia or New Zealand), including the *X. vasicola* type strain NCPPB2417.

Evidence of type III and type IV secretion systems was found in preliminary annotations of the USA (NE744) and South African (SAM119) genomes (data not shown), similar to predictions from other *X. v.* pv. *vasculorum* genomes (13, 48). A TBLASTN search for transcription activator like (TAL) effector gene sequences using conserved features (N and C termini) of this protein family from *X. oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* did not reveal TAL effectors in the *X. vasicola* draft genomes. Thus, these strains do not appear to contain genes encoding TAL effectors. However, the highly repetitive sequence of TAL effectors cannot be fully resolved by draft sequence alone and further investigation by DNA hybridization or long-read sequencing is necessary.

4.3.3 Pathogenicity

To determine pathogenicity, a panel of historical and recently collected strains (Table 4.1) were inoculated to corn, sorghum and sugarcane using two previously reported assays, infiltration into leaves or stab injection into the stems (5, 8, 14, 44, 49, 50). Phenotypes after stem injection across corn and sorghum were highly variable (Supplemental Fig. S2, S3), therefore, we focused our analysis on results from quantitative leaf infiltration assays (Fig. 4.2). When introduced into corn leaves (cv. hybrid DKC 61-88) by infiltration, *X. v.* pv. *holcicola* and *X. v.* pv. *vasculorum* strains were all moderate to highly

virulent (lesions ranging 5.6 -10.6 cm) with two exceptions, *X. v.* pv. *vasculorum* strains NCPPB206 (corn) and NCPPB1326 (sugarcane), which were weakly virulent (average lesion lengths ≤ 1.6 cm). However, *X. v.* pv. *vasculorum* isolates caused less water soaking and shorter lesions on sorghum than on corn. Prolific bacterial exudate was observed after inoculation of either *X. v.* pv. *vasculorum* or *X. v.* pv. *holcicola* to corn, or *X. v.* pv. *holcicola* to sorghum regardless of the inoculation technique used (Supplementary Fig. S4).

Sugarcane was inoculated by leaf infiltration only, but similar to corn and sorghum, lesion lengths were variable between plants and even on the same leaf (Supplementary Table 1). Representative phenotypes at 4 dpi with selected strains are shown in Fig. 4.3. Strains of *X. v.* pv. *vasculorum* and *X. v.* pv. *holcicola* produced a deep purple response in sugarcane leaves over time, but *X. v.* pv. *holcicola* caused the earliest reaction, starting at 3 dpi. Regardless of original host, *X. v.* pv. *vasculorum* strains caused water soaking at 4 dpi; after 5 dpi, infiltration sites turned deep purple and all lesions continued to expand. Although not as abundant as on corn or sorghum, bacterial exudate was observed on sugarcane after inoculation by all *X. v.* pv. *vasculorum* strains. *X. c.* pv. *musacearum* was not pathogenic and caused no water soaking on sugarcane and again, *X. v.* pv. *vasculorum* NCPPB 1326 was not highly virulent.

Based on our pathogenicity tests, *X. v.* pv. *vasculorum* strains from either corn or sugarcane cause disease on corn, sorghum, or sugarcane, but they are usually most aggressive to the host from which they were originally isolated. The disease phenotypes caused by the USA corn isolate on all three hosts are most similar to those caused by *X. v.* pv. *vasculorum* from South Africa. *X. v.* pv. *holcicola* strains are virulent to all three hosts, but are more aggressive to sorghum than the sugarcane and corn isolates.

4.3.4 Diagnostics

Four primer sets (Xvv 3, Xvv5, Xvv7 and Xvv8) were generated that specifically amplify *X. v.* pv. *vasculorum* from isolated DNA, heat killed cells, or infected, macerated tissues, and that did not amplify any non-target strains. Gene targets, primer sequences and recommended annealing temperatures are listed in Table 3. In total, 17 strains positively amplified with the four sets of *X. v.* pv. *vasculorum* primers, and were differentiated from the *X. v.* pv. *holcicola, X. c.* pv. *musacearum* and other non-target

strains by these assays. No false positives, i.e., amplification of non-target strains, were detected. Furthermore, if the primers amplified DNA from plant exudates, if leaf tissue was in tact and not exceptionally decomposed, *X. v.* pv. *vasculorum* was isolated from the leaves (9/13 samples). *X. v.* pv. *vasculorum* NCPPB206 did not amplify with primers Xvv7 or Xvv8. Primers specific for *X. v.* pv. *holcicola* were also generated and tests with the same sets of target and non-target strains revealed consistent amplification of target strains and no false positives (Table 4.3, Table 4.1). Predicted gene function of loci identified for primer design from draft genome sequences are included in Table 4.3.

4.3.5 XopAF

The effector gene *xopAF* was reported in *X. euvesicatoria, X. translucens* pv. *translucens, X. citri* subsp. *citri* and *X. v.* pv. *vasculorum* from sugarcane, but is absent in the only other strain of *X. v.* pv. *vasculorum* publicly available from corn, NCPPB206, and also absent in *X. c.* pv. *musacearum* (11, 13, 51). To determine if newly isolated corn strains from the USA contained this effector or close relatives, we tested for the effector gene using *xopAF*-specific primers (Table 4.3). No strain of *X. v.* pv. *vasculorum* isolated from corn in the USA, no *X. v.* pv. *holcicola* nor *X. c.* pv. *musacearum* contain *xopAF,* while all sugarcane strains of *X. v.* pv. *vasculorum* did possess this gene regardless of geographic region of isolation (Fig. 4.4).

4.4 DISCUSSION

The causal agent of the bacterial leaf streak that recently emerged on corn in the USA was reported as *X. vasicola* (1). In this study, we performed phylogenetic analyses and compared genome sequences and host ranges of *X. vasicola* isolated from corn, sugarcane and sorghum to refine the taxonomic designation of the USA corn isolates. We confirm that the pathogen causing bacterial streak of corn in the USA is *X. vasicola*, and propose the designation *Xanthomonas vasicola* pv. *vasculorum* (Cobb 1893) comb. nov. Further, we propose strain SAM119 (5) as the pathotype strain. Below, we summarize our results and rationale for these conclusions.

X. campestris pv. *vasculorum* groups A and B were initially proposed based on SDS-PAGE of proteins, gas chromatography of fatty acid methyl-esters and DNA-DNA hybridization (52). Vauterin et

al (1995) then proposed renaming and reclassification of the species based on DNA-DNA hybridization; this report separated group A, composed of *X. axonopodis* pv. *vasculorum* from sugarcane, and group B, composed of *X. vasicola* pv. *vasculorum* from corn or sugarcane and *X. v.* pv. *holcicola* from sorghum. Our MLSA supports placement of corn strains from the USA and the previously reported South African strain SAM119 into *X. vasicola* group B (Fig. 4.1). The corn strains form a distinct clade from *X. v.* pv. *holcicola,* but interestingly, do not separate from *X. c.* pv. *musacearum*. Using comparisons of genomic similarities as measured by ANI to provide further taxonomic context (53), we found that the corn and sugarcane isolates that clustered in MLSA phylogenetic trees (Fig. 1) had ANI of greater than 99.3% (Table 4.2). In addition, ANI values for *X. v.* pv. *holcicola* and *X. v.* pv. *vasculorum* from corn in the USA indicated about 98.6% genome similarity, and phylogenetic analyses showed these organisms branching into two distinct groups.

Based on MLSA, *X. c.* pv. *musacearum* and *X. v.* pv. *vasculorum* were indistinguishable, but previous whole genome comparisons (13, 48), the ANI values, and the diagnostic primers in this study (Table 4.2, 4.3) did differentiate the two pathovars. Since MLSA compares only variation in housekeeping genes while ANI detects variation across the entire genome sequence, differences in predicted relationships might be expected. Therefore, host range studies are necessary to tease apart this complex. We note that neither pathogenicity of *X. v.* pv. *vasculorum* to banana or *X. c.* pv. *musacearum* to corn were assessed in this study. However, in previous studies, *X. v.* pv. *vasculorum* from corn or sugarcane were not pathogenic to banana, while *X. c.* pv. *musacearum* did cause symptoms on corn (8, 10). The African strain of *X. c.* pv. *musacearum* (NCBBP4381) is not pathogenic to sugarcane.

Ideally, a combination of MLSA, comparison of whole genome sequence, and ecology are integrated to define a prokaryotic species (54, 55). For plant pathogenic bacteria, pathovars are distinguished based on differences in host range, although differences in symptomology on the same plant species can also warrant separate pathovar designations (19, 24). Previous reports showed that corn *X. v.* pv. *vasculorum* isolates were pathogenic to corn when reintroduced but they did not cause disease on sorghum or sugarcane (5, 56). These authors also indicated that *X. v.* pv. *holcicola* strains were

pathogenic to both corn and sorghum, while a sugarcane isolate of *X. v.* pv. *vasculorum* caused disease on all three hosts. However, our results and those from other pathogenicity studies (Karamura et al. 2015), which compare some of the same isolates, conflicted with these results, possibly due to the use of different cultivars or experimental conditions. In our studies using established protocols, *X. v.* pv. *vasculorum* strain NE744 from corn in Nebraska was pathogenic to corn, sorghum and sugarcane, and the symptoms (water soaking with spreading lesions) were most similar to those caused by corn and sugarcane isolates of *X. v.* pv. *vasculorum* from other countries. Based on our pathogenicity tests, we recommend that the USA and South African corn strains be included in *X. v.* pv. *vasculorum.*

X. v. pv. *holcicola* strains were also pathogenic to all three hosts, but were more aggressive to sorghum than the corn and sugarcane isolates. We note that when bacterial leaf streak was first observed in corn in the USA, it was thought to be caused by *X. v.* pv. *holcicola* because, in early literature based on greenhouse inoculations, *X. v.* pv. *holcicola* caused symptoms on corn (3, 49). However, to date, *X. v.* pv. *holcicola* has never been associated with bacterial leaf streak symptoms on corn in the field. The lack of evidence for occurrence of this pathovar on corn in the field as well as the genomic differences between the two pathovars support the current separate pathovar designation for these members of *X. vasicola*. Further phenotyping with more strains, more hosts and different host varieties would improve resolution of the pathovar designation.

Both the sugarcane isolates of *X. a.* pv. *vasculorum* and *X. v.* pv. *vasculorum* are reported to cause gumming disease of sugarcane (3, 6, 7), creating confusion for pathologists. Unfortunately, little information is available on the etiology of sugarcane gumming disease, and resolution of this issue will require comparisons of host range and symptomology that includes both species in optimized disease assays. A plausible hypothesis is that *X. a. pv. vasculorum* is a vascular pathogen and causal agent of gumming disease while *X. v.* pv. *vasculorum* is non-vascular and causal agent of leaf streak.

The emergence of corn bacterial leaf streak and its spread throughout the midwestern USA created an urgent need for diagnostic tools that could be used to accurately identify the causal agent and to track its distribution. By comparing the draft genomes generated as part of this study with the large

number of *Xanthomonas* genomes now available, we identified unique regions and used these to develop sets of diagnostic primers that distinguish *X. v.* pv. *vasculorum* (four primer pairs) from *X. v.* pv. *holcicola* (three primer pairs) and from other bacteria. The primers Xvv3, Xvv5, Xvh2 and Xvh3 have been tested in our labs and in collaboration with colleagues in Iowa (C. Block, A. Robertson, G. Munkvold, Iowa State University) and Kansas (D. Jardine, Kansas State University). The primers are currently being used for disease diagnosis and for epidemiological surveys across the USA in private, public and federal institutions, including USDA-APHIS-PPQ. Future integration of multiple loci variable number of tandem repeat analysis (MLVA) to specifically type populations of the *X. vasicola* complex from corn and sorghum in the USA would be valuable for epidemiological surveillance and could also help determine geographic lineages of these organisms (57–59).

Consistent with previous reports, of the *X. v.* pv. *vasculorum,* only those from sugarcane contain the gene *xopAF,* an effector that may contribute to host range and symptomology (11, 13). No evidence of TAL effectors was found in the draft genomes of *X. v.* pv. *vasculorum* strains NE744 and SAM119, but Harrison and Studholme (2014) did predict their presence in *X. a.* pv. *vasculorum* NCPPB900. TAL effectors are found in diverse *Xanthomonas* spp. but are not present in all (Jacques et al. 2016). In many systems, these proteins contribute to pathogenicity. Given the absence of *xopAF* and genes for TAL effectors in the corn *X. v.* pv. *vasculorum,* these organisms must rely on a distinct set of effectors for virulence.

In the phylogenetic tree, the most closely related *Xanthomonas* species to *X. vasicola* is *X. oryzae,* which includes two pathovars, *oryzae* and *oryzicola,* causal agents of bacterial blight and bacterial leaf streak of rice, respectively. An interesting possibility is that these xanthomonads diverged from an ancestral group adapted to monocots. Rice and sugarcane are grown in proximity to one another in some tropical areas, which may have historically fostered evolution of these distinct groups. Although *X. v.* pv. *vasculorum* was previously thought to only infect monocots, including palms, it was recently identified on *Eucalyptus* in association with *Pantoea ananatis* (14). An outbreak of brown stalk rot on corn caused by *P. ananatis* was also reported in South Africa in 2004 (50). Intriguingly, we frequently isolated *P.*

ananatis from corn samples in the USA that exhibited bacterial leaf streak symptoms; however, when these *P. ananatis* were inoculated to corn, they did not cause disease (Jillian M. Lang, unpublished results). It is unknown if a synergistic relationship exists between *X. v.* pv. *vasculorum* and *P. ananatis*, and if so, if this interaction is related to the emergence of bacterial leaf streak in the USA.

The origin of the USA corn bacterial leaf streak pathogen is unknown. Based on ANI values, *X. v.* pv. *vasculorum* strain NE744 is highly similar to strain SAM119 (99.9%) that was isolated from corn in South Africa over 20 years ago (5). While it is tempting to speculate that *X. v.* pv. *vasculorum* was introduced via international germplasm movement, given its distribution in at least seven states, pinpointing if, when, or how it was introduced would be difficult.

It is not known if *X. v.* pv. *vasculorum* is seed transmitted, or how it is moved within and across fields. As the disease has been confirmed on various hybrids of popcorn, seed, dent and sweet corn (T.A. Jackson-Ziems, K. Broders, personal communication), it is unlikely that the rapid spread is due to widespread planting of one or a few susceptible varieties. The disease occurs in both irrigated and non-irrigated areas, and in a variety of corn production systems, including various tillage and crop rotation regimes (T.A. Jackson-Ziems, K. Broders, personal communication). A deeper examination of the *X. v.* pv. *vasculorum* genomes from geographically diverse locations and hosts as well as epidemiological investigations are important to understanding how and why this recent epidemic in the USA has occurred.

In conclusion, the phylogenetic analyses, genome comparisons and pathogenicity studies reported here support classification of the USA corn bacterial leaf streak pathogen as *X. v.* pv. *vasculorum* (Cobb 1893) comb. nov. (Korus et al. 2017). This classification is consistent with previous groupings of the South African corn isolates as *X. v.* pv. *vasculorum* (6, 9–11, 13). We propose a new pathotype strain, *X. v.* pv. *vasculorum* SAM119, based on historical precedence and the availability of genome sequence data (5, 55). Unlike NCPPB206, the oldest (1949) recorded publicly available strain isolated from corn, SAM119 has 99% genomic identity with NCPPB206 and recent isolates from the USA, and remains highly virulent on corn.

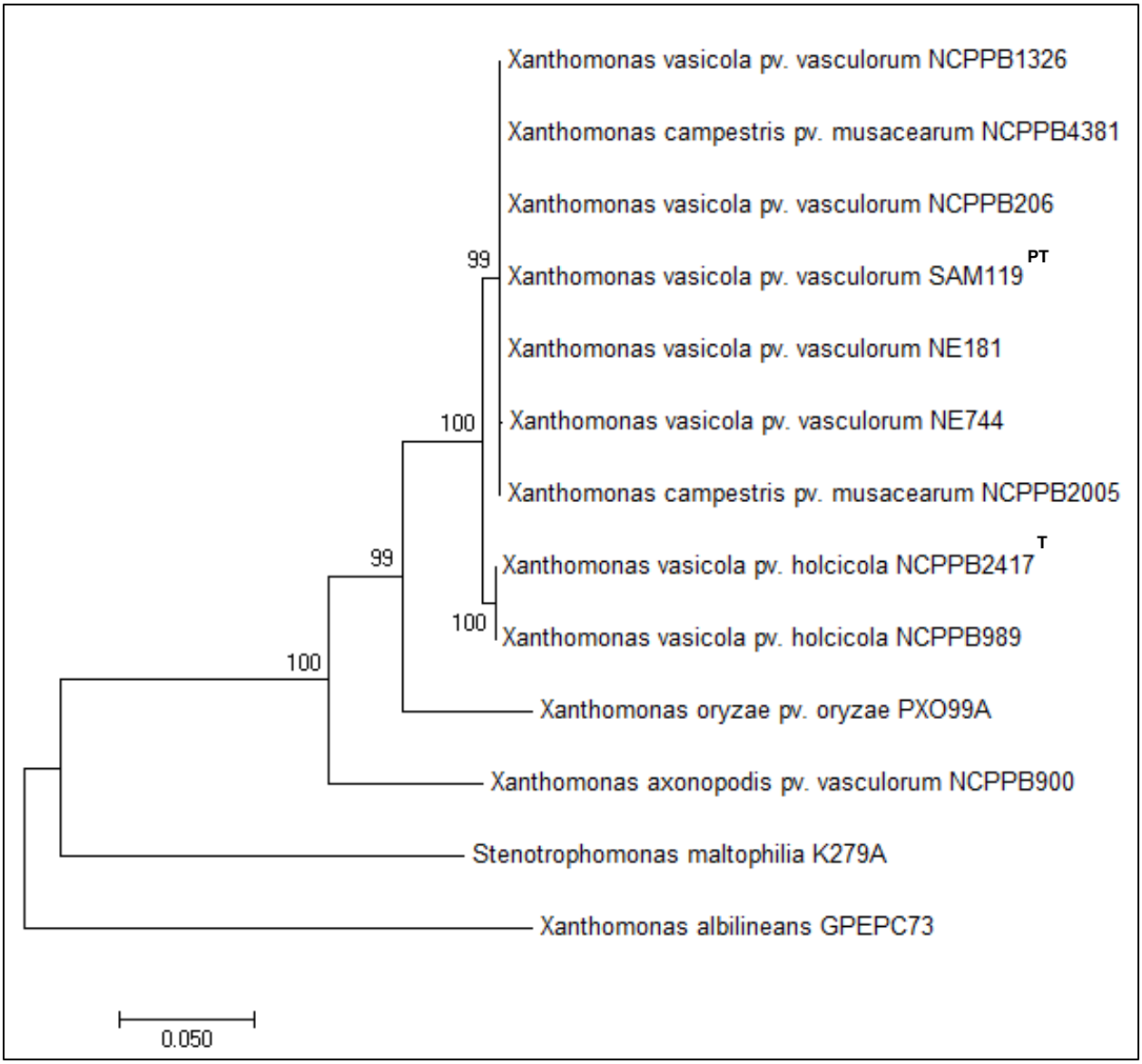

Figure 4.1. Phylogeny based on partial *atpD, dnaK, fusA, gyrB, lepA* and *rpoD* sequence alignment. The evolutionary history was inferred by using the Maximum Likelihood method based on the Equal Input model in MEGA7 (33). Bootstrap values generated from 1000 replicates are shown at nodes. Branch lengths measure in the number of substitutions per site. There were a total of 3967 bp in the final dataset. T = type strain, PT = pathotype strain.

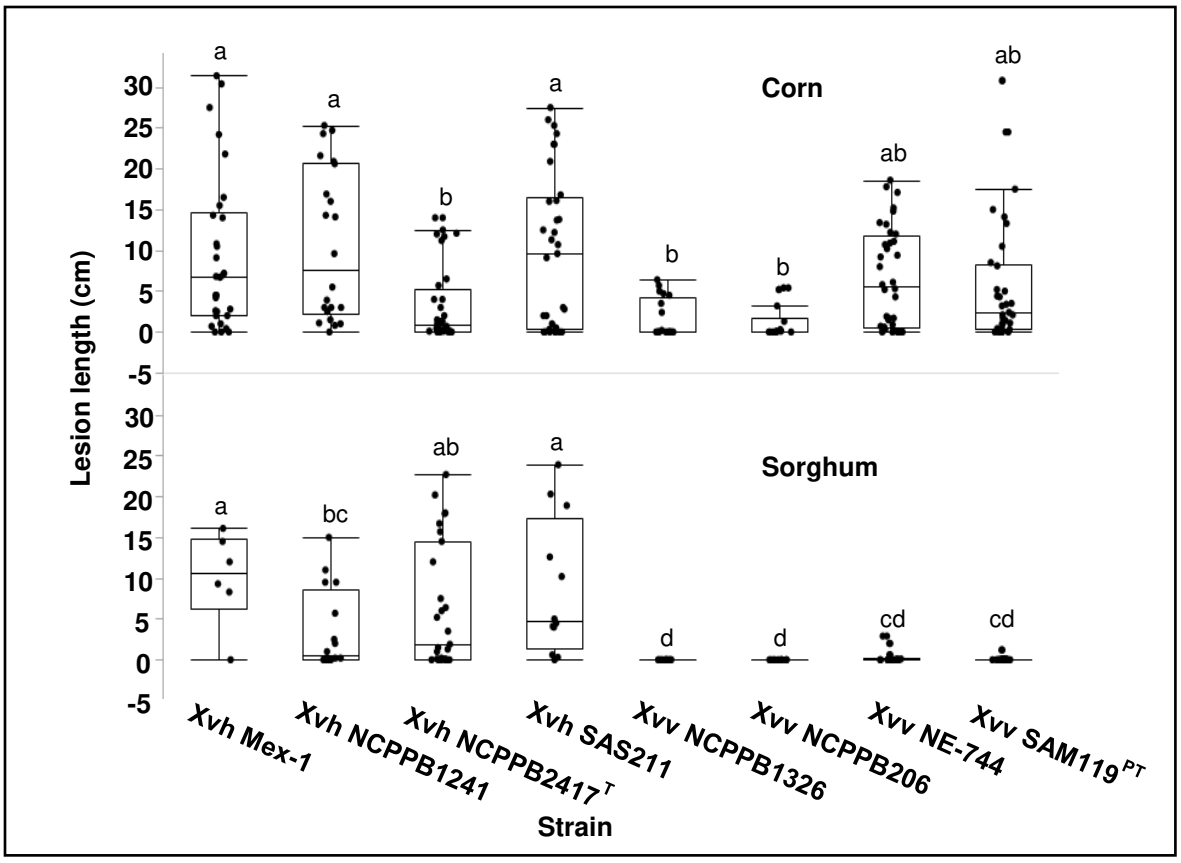

Figure 4.2. Disease caused by *X. vasicola* pv. *vasculorum* (Xvv) and *X. v.* pv. *holcicola* (Xvh) on corn (cv. hybrid DKC 61-88) and sorghum (cv. Mycogen IG588). Four week old plants were infiltrated with 10^8 CFU ml^{-1} of each strain, and disease was assessed at 7 days post inoculation (dpi). Lesion lengths indicate expansion beyond the infiltration site. The entire experiment was replicated four times and combined data from all replications is shown here. Letters designate significance at $P < 0.0001$. T = type strain, PT = pathotype strain.

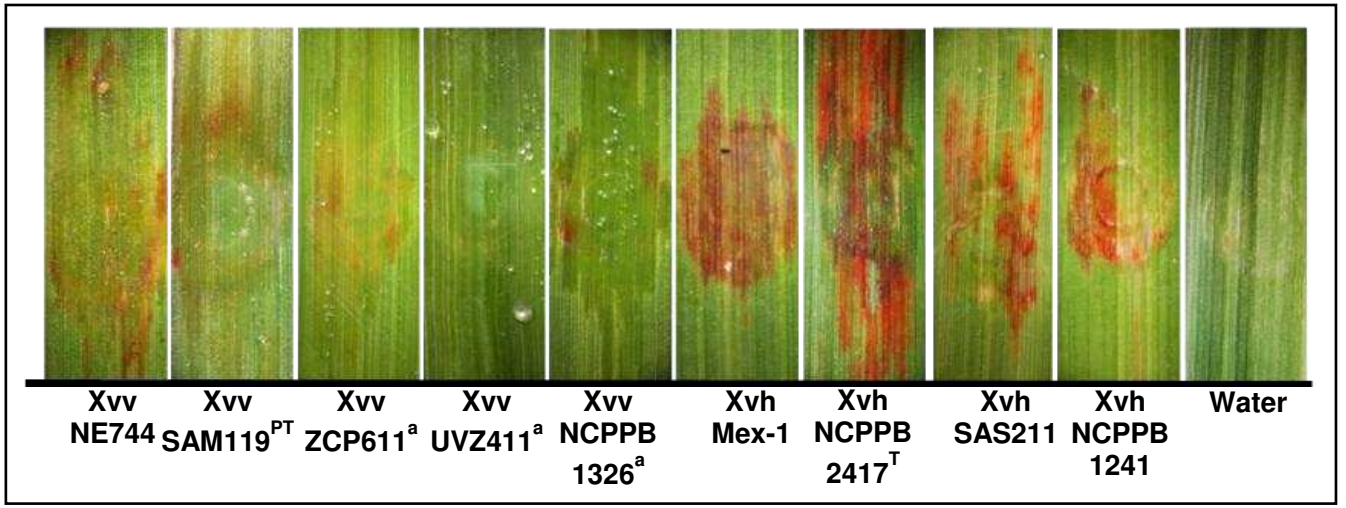

Figure 4.3. Qualitative disease phenotyping of *X. v.* pv. *vasculorum* (Xvv) and *X. v.* pv. *holcicola* (Xvh) on sugarcane (cv.L-99-226). Six week old plants were infiltrated with 10^8 CFU ml^{-1} of each strain. Reactions were recorded 4 dpi. [a] Strains isolated from sugarcane, T = type strain, PT = pathotype strain.

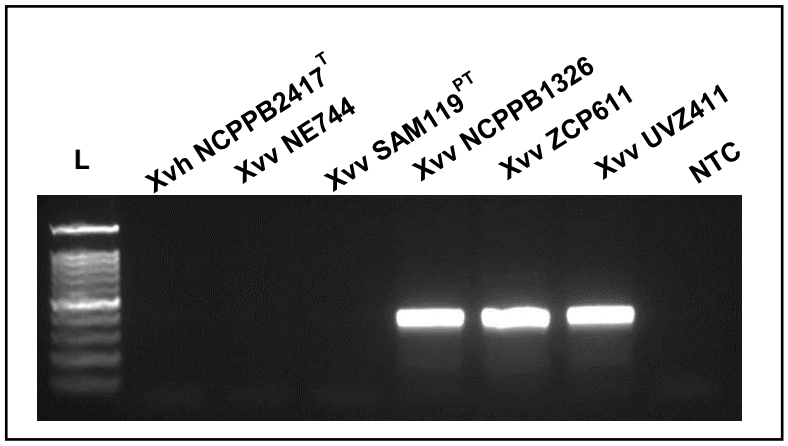

Figure 4.4. Targeted amplification of *xopAF* in *X. v.* pv. *vasculorum* (Xvv) and *X. v.* pv. *holcicola* (Xvh). Lane L: 100bp DNA ladder (GoldBio, St. Louis, MO); NTC: no template control. T = type strain, PT = pathotype strain.

Table 4.1. Bacterial strains used in this study and confirmed specificity of diagnostic primers for *X. v.* pv. *holcicola* (Xvh) and *X. v.* pv. *vasculorum* (Xvv).

					Xvv[e]				Xvh[e]		
Species	**Strain**	**Origin**	**Host**	**Source**	**3**	**5**	**7**	**8**	**1**	**2**	**3**
Burkholderia andropogonis	3549		*Zea mays*	L.E. Claflin	-	-	-	-	-	-	-
Clavibacter michiganensis* pv. *nebraskensis	CO428	USA	*Z. mays*	K. Broders	-	-	-	-	-	-	-
***Clavibacter* sp.**	CO-4	USA	*Z. mays*	K. Broders	-	-	-	-	-	-	-
***Enterobacter* sp.**	CO-3	USA	*Z. mays*	K. Broders	-	-	-	-	-	-	-
***Enterobacter* sp.**	CO-22	USA	*Z. mays*	K. Broders	-	-	-	-	-	-	-
Escherichia coli	DH5α				-	-	-	-	-	-	-
***Pantoea* sp.**	CO-2	USA	*Z. mays*	K. Broders	-	-	-	-	-	-	-
Pantoea agglomerans	B55	USA	*T. aestivum*	N. Tisserat	-	-	-	-	-	-	-
Pseudomonas fuscovaginae	SE-1	Philippines	*O. sativa*	G. Ash	-	-	-	-	-	-	-
P. syringae* pv. *syringae	M108	USA	*Solanum lycopersicum*	H.F. Schwartz	-	-	-	-	-	-	-
***Xanthomonas* sp.**	M136	Mali	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
***Xanthomonas* sp.**	SHU100	Philippines	*O. sativa* seed	C.M. Vera Cruz	-	-	-	-	-	-	-
X. axonopodis* pv. *allii	0177	USA	*Allium cepa*	H.F. Schwartz	-	-	-	-	-	-	-
X. campestris* pv. *alfalfae	KX-1	USA		L.E. Claflin	-	-	-	-	-	-	-
X. c.* pv. *campestris	X1910	USA	*Brassica oleracea*	N. Dunlop	-	-	-	-	-	-	-
X. c.* pv. *leersiae	NCPPB4346	China	*Leersia hexandra*	V. Verdier	-	-	-	-	-	-	-
X. c.* pv. *pelargonii	X5	USA	*Geranium* sp.	L.E. Claflin	-	-	-	-	-	-	-
X. c.* pv. *pennamericanium	ATCC 49152	Nigeria	*Pennisetum glaucum*	L.E. Claflin	-	-	-	-	-	-	-
X. c. pennisetum	PMS91	Senegal	*P. glaucum*	L.E. Claflin	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB2005[b]	Ethiopia	*Ensete ventricosum*	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB2251	Ethiopia	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4378	Uganda	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4381[c]	Uganda	*Musa* sp.	A. Bogdanove	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4386	Uganda	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-

X. c.* pv. *musacearum	NCPPB4387	D.R. Congo	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4388	D.R. Congo	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4389	Rwanda	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4390	Rwanda	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4393	Tanzania	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4394	Tanzania	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4433	Burundi	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c.* pv. *musacearum	NCPPB4434	Kenya	*Musa* sp.	E.Wicker	-	-	-	-	-	-	-
X. c. pv. zeae/X. vasicola* pv. *vasculorum	SAM 113	S. Africa	*Z. mays*	L.E. Claflin	+	+	+	+	-	-	-
X. c. pv. zeae/X. v.* pv. *vasculorum	SAM118	S. Africa	*Z. mays*	L.E. Claflin	+	+	+	+	-	-	-
X. c. pv. zeae/X. v.* pv. *vasculorum	SAM119^{abcPT}	S. Africa	*Z. mays*	L.E. Claflin	+	+	+	+	-	-	-
X. euvesicatoria	85-10	USA	*Capsicum frutescens*	A. Bogdanove	-	-	-	-	-	-	-
X. oryzae* pv. *oryzae	A3857	India	*Oryza sativa*	J.E. Leach	-	-	-	-	-	-	-
X. o.* pv. *oryzae	BAI3	Burkina Faso	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
X. o.* pv. *oryzae	R3	Australia	*O. sativa*	J.E. Leach	-	-	-	-	-	-	-
X. o.* pv. *oryzae	Xoo4	Thailand	*O. sativa*	J.E. Leach	-	-	-	-	-	-	-
X. o.* pv. *oryzae	MAI1	Mali	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
X. o.* pv. *oryzae	NAI8	Niger	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
X. o.* pv. *oryzae	PXO86	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	-	-	-	-	-
X. o.* pv. *oryzae	PXO99A[a]	Philippines	*O. sativa*	J.E. Leach	-	-	-	-	-	-	-
X. o.* pv. *oryzae	X11-5A	USA	*O. sativa*	C. Gonzalez	-	-	-	-	-	-	-
X. o.* pv. *oryzae	Xoo199	Korea	*O. sativa*	S.H. Choi	-	-	-	-	-	-	-
X. o.* pv. *oryzicola	BLS98	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	-	-	-	-	-

X. o.* pv. *oryzicola	BLS105	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	-	-	-	-	-
X. o.* pv. *oryzicola	BLS256	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	-	-	-	-	-
X. o.* pv. *oryzicola	BLS305	Philippines	*O. sativa*	C.M. Vera Cruz	-	-	-	-	-	-	-
X. o.* pv. *oryzicola	MAI4	Mali	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
X. o.* pv. *oryzicola	MAI10	Mali	*O. sativa*	V. Verdier	-	-	-	-	-	-	-
X. translucens* pv. *cerealis	NCPPB1943	USA	*Triticum aestivum*	L.E. Claflin	-	-	-	-	-	-	-
X. t.* pv. *cerealis	NCPPB1944	USA	*Bromus inermis*	V. Verdier	-	-	-	-	-	-	-
X. t.* pv. *phleipratensis	ICMP5744	USA	*Phleum pretense*	L.E. Claflin	-	-	-	-	-	-	-
X. t.* pv. *translucens	B76	USA	*H. vulgare*	N. Tisserat	-	-	-	-	-	-	-
X. t.* pv. *translucens	NCPPB2389	India	*H. vulgare*	C. Bragard	-	-	-	-	-	-	-
X. t.* pv. *translucens	UPB787	Paraguay	*H. vulgare*	C. Bragard	-	-	-	-	-	-	-
X. t.* pv. *undulosa	UPB513	Mexico	*Triticum aestivum*	C. Bragard	-	-	-	-	-	-	-
X. v.* pv. *holcicola	66	KS, USA	*Sorghum bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	86	KS, USA	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	93	KS, USA	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	107	Lesotho	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	114	Lesotho	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	123	Lesotho	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	124	Lesotho	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	Mex-1[bc]	Mexico	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	NCPPB989[a]	TX, USA	*Holcus* sp.	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	NCPPB1241[bc]	Australia	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	NCPPB2417[abcT]	New Zealand	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	SAS211[bc]	South Africa	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+
X. v.* pv. *holcicola	Z-5	Zimbabwe	*S. bicolor*	L.E. Claflin	-	-	-	-	+	+	+

X. v.* pv. *vasculorum	201600017x	NE, USA	*Z. mays*	T. Jackson-Ziems	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	201600018x	NE, USA	*Z. mays*	T. Jackson-Ziems	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	201600039x	NE, USA	*Z. mays*	T. Jackson-Ziems	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	201600068x	NE, USA	*Z. mays*	T. Jackson-Ziems	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	NE744[abc]	NE, USA	*Z. mays*	K. Korus	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	NE181[a]	NE, USA	*Z. mays*	K. Korus	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	CO-5[c]	CO, USA	*Z. mays*	K. Broders	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	KS444	KS, USA	*Z. mays*	J. Chaky	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	NCPPB206[ab]	S. Africa	*Z. mays*	L.E. Claflin	+	+	-	-	-	-	-
X. v.* pv. *vasculorum	NCPPB1326[abc]	Zimbabwe	*Saccharum officinarum*	L.E. Claflin	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	NE429	NE, USA	*Z. mays*	T. Jackson-Ziems	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	NE442	NE, USA	*Z. mays*	L. Appel	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	UVZ 411[c]	S. Africa	*S. officinarum*	R. A. Bailey	+	+	+	+	-	-	-
X. v.* pv. *vasculorum	ZCP 611[c]	Zimbabwe	*S. officinarum*	P. Sinai	+	+	+	+	-	-	-

[a]Strains used in MLSA. PXO99A is a 5-azacytidine resistant strain of race 6 PXO99 (61, 62)
[b]Strains tested for pathogenicity to corn and sorghum
[c]Strains tested for pathogenicity to sugarcane
[e]Diagnostic primers listed in Table 3
[PT]Proposed pathotype strain
[T]Type strain
NCPPB = National Collection of Plant Pathogenic Bacteria (http://ncppb.fera.defra.gov.uk/)
ICMP = International Collection of Micro-organisms from Plants (http://www.landcareresearch.co.nz/resources/collections/icmp)
UPB = Unité de Phytopathologie Bacterial, Université Catholique de Louvain

Table 4.2. Average nucleotide identity (%ANI) values calculated by whole genome comparison using draft genomes.

Species	Strain	Origin	Host	GenBank Accession	Reference	% ANI with NE744[a]	% ANI with SAM119[a]
Stenotrophomonas maltophilia	K279a	United Kingdom	*Homo sapiens*	NC_010943	Crossman et al. 2008	75.68	75.71
Xanthomonas albilineans	GPE PC73	Guadeloupe	*Saccharum officinarium*	GCA_000087965.1	Pieretti et al. 2009	79.17	79.17
X. axonopodis* pv. *vasculorum	NCPPB900	La Reunion	*S. officinarium*	GCA_000724905.2	Harrison and Studholme 2014	89.28	89.17
X. campestris* pv. *musacearum	NCPPB 4381	Uganda	*Musa* sp.	ACHT00000000	Aritua et al 2007; Studholme et al 2010	98.62	98.63
X. c.* pv. *musacearum	NCPPB 2005	Ethiopia	*Ensete ventricosum*	AKBE01000000	Wasukira et al 2014	98.81	98.81
X. oryzae* pv. *oryzae	PXO99[A]	Philippines	*Oryza sativa*	NC_010717	Salzberg et al 2008; Booher et al 2015	90.96	90.9
X. vasicola* pv. *holcicola	NCPPB 989	USA	*Holcus* sp.	JSCA01000000		98.58	98.55
X. v.* pv. *holcicola	NCPPB 1241	Australia	*Sorghum vulgare*	JSBV01000000		98.58	98.58
X. v.* pv. *holcicola	NCPPB 2417	New Zealand	*S. vulgare*	JSBW02000000		98.61	98.57
X. v.* pv. *vasculorum	NCPPB 702	Zimbabwe	*S. officinarium*	ACHS00000000.1	Studholme et al 2010	99.38	99.38
X. v.* pv. *vasculorum	NCPPB 206	South Africa	*Zea mays*	AKBM00000000	Studholme et al 2010; Wasukira et al 2012	99.50	99.47
X. v.* pv. *vasculorum	NCPPB 890	South Africa	*S. officinarium*	AKBN01000000	Wasukira et al 2014	99.53	99.52
X. v.* pv. *vasculorum	NCPPB 895	Madagascar	*S. officinarium*	AKBO01000000	Wasukira et al 2014	99.51	99.53
X. v.* pv. *vasculorum	NCPPB 1326	Zimbabwe	*S. officinarium*	AKBK01000000	Wasukira et al 2012	99.34	99.36

X. v.* pv. *vasculorum	NCPPB 1381	Zimbabwe	*S. officinarium*	AKBL 00000000.1	Wasukira et al 2012	99.40	99.42
X. v.* pv. *vasculorum	NE744	Holt Co., Nebraska, USA	*Z. mays*	MVYW00000000	This study	n.a.	99.98
X. c.* pv. *zeae/* *X. v.* pv. *vasculorum	SAM119	Klerksdorp, South Africa	*Z. mays*	MVYX00000000	Qhobela et al. 1990	99.98	n.a.

[a] % ANI calculated using draft genome sequence in JSpecies V1.2.1 relative to strain NE744 or SAM119.
NCPPB = National Collection of Plant Pathogenic Bacteria (http://ncppb.fera.defra.gov.uk/).

Table 4.3. Primers developed with specificity to *X. v.* pv. *vasculorum*, *X. v.* pv. *holcicola* and the effector XopAF.

Target	Name	Sequence (5'-3')	Product size (bp)	Annealing Temp.
X. v.* pv. *vasculorum				
	Xvv3_F	CAAGCAGAGCATGGCAAAC	207	55ºC
	Xvv3_R	CACGTAGAACCGGTCTTTGG		
	Xvv5_F	CCGTCGAAATGGTCTCAACT	200	55ºC
	Xvv5_R	CGGAAGAGTTGGAAGACAGC		
	Xvv7_F	CTACTACGCCCAGCGACTTC	205	53ºC
	Xvv7_R	ACGTCGAGCCATTCTGAAAC		
	Xvv8_F	GGGTTATTGACGGCACTCTC	206	53ºC
	Xvv8_R	GGGCAGCCTGTAACGAATTA		
X. v.* pv. *holcicola				
	Xvh1_F	GCAGATTGTCAGCATCAGGA	201	55ºC
	Xvh1_R	GATCTTACGCACAGCACCAA		
	Xvh2_F	CGAATTTTGTGTGACCAGGA	200	53ºC
	Xvh2_R	GAATTCACCAAATGGGCATC		
	Xvh3_F	ATCCCATGGGTCTGAGTCTG	200	53ºC
	Xvh3_R	AGTCCATTGCAGGAGTTTG		
	XopAF_F	CCATTGCCATTGCTAGCACC	362	60ºC
	XopAF_R	TATTCGACGGTTCCCACTGC		

REFERENCES

1. **Korus K**, **Lang JM**, **Adesemoye AO**, **Block CC**, **Pal N**, **Leach JE**, **Jackson-Ziems TA**. 2017. First report of *Xanthomonas vasicola* causing bacterial leaf streak on corn in the United States. Plant Dis PDIS-10-16-1426-PDN.
2. **Dyer RA**. 1949. Botanical surveys and control of plant diseases. Farming in South Africa. Annu Rep Dep Agric South Africa **275**:119–121.
3. **Bradbury JF**. 1986. Guide to plant pathogenic bacteria. CAB International.
4. **Qhobela M**, **Claflin LE**. 1992. Eastern and southern African strains of *Xanthomonas campestris* pv. *vasculorum* are distinguishable by restriction fragment length polymorphism of DNA and polyacrylamide gel electrophoresis of membrane proteins. Plant Pathol **41**:113–121.
5. **Qhobela M**, **Claflin LE**, **Nowell DC**. 1990. Evidence that *Xanthomonas campestris* pv. *zeae* can be distinguished from other pathovars capable of infecting maize by restriction fragment length polymorphism of genomic DNA. Can J Plant Pathol **12**:183–186.
6. **Vauterin L**, **Hoste B**, **Kersters K**, **Swings J**. 1995. Reclassification of *Xanthomonas*. Int J Syst Bacteriol **45**:472–489.
7. **Dookun A**, **Stead DE**, **Autrey LJ-C**. 2000. Variation among strains of *Xanthomonas campestris* pv. *vasculorum* from Mauritius and other countries based on fatty acid analysis. Syst Appl Microbiol **23**:148–55.
8. **Karamura G**, **Smith J**, **Studholme D**, **Kubiriba J**, **Karamura E**. 2015. Comparative pathogenicity studies of the *Xanthomonas vasicola* species on maize, sugarcane and banana. African J Plant Sci **9**:385–400.
9. **Rademaker JLW**, **Louws FJ**, **Schultz MH**, **Rossbach U**, **Vauterin L**, **Swings J**, **de Bruijn FJ**. 2005. A comprehensive species to strain taxonomic framework for Xanthomonas. Phytopathology **95**:1098–1111.
10. **Aritua V**, **Parkinson N**, **Thwaites R**, **Heeney J V.**, **Jones DR**, **Tushemereirwe W**, **Crozier J**, **Reeder R**, **Stead DE**, **Smith J**. 2008. Characterization of the *Xanthomonas* sp. causing wilt of enset and banana and its proposed reclassification as a strain of *X. vasicola*. Plant Pathol **57**:170–177.
11. **Studholme DJ**, **Kemen E**, **MacLean D**, **Schornack S**, **Aritua V**, **Thwaites R**, **Grant M**, **Smith J**, **Jones JDG**. 2010. Genome-wide sequencing data reveals virulence factors implicated in banana Xanthomonas wilt. FEMS Microbiol Lett **310**:182–92.
12. **Harrison J**, **Studholme DJ**. 2014. Draft genome sequence of *Xanthomonas axonopodis* pathovar *vasculorum* NCPPB 900. FEMS Microbiol Lett **360**:113–116.
13. **Wasukira A**, **Coulter M**, **Al-Sowayeh N**, **Thwaites R**, **Paszkiewicz K**, **Kubiriba J**, **Smith J**, **Grant M**, **Studholme DJ**. 2014. Genome sequencing of *Xanthomonas vasicola* pv. *vasculorum* reveals variation in plasmids and genes encoding lipopolysaccharide synthesis, type-IV pilus and type-III secretion effectors. Pathogens **3**:211–37.
14. **Coutinho TA**, **van der Westhuizen L**, **Roux J**, **McFarlane SA**, **Venter SN**. 2015. Significant host jump of *Xanthomonas vasicola* from sugarcane to a *Eucalyptus grandis* clone in South Africa. Plant Pathol **64**:576–581.
15. **Lapage S**, **Sneath P**, **Lessel E**, **Skerman V**, **Seeliger H**, **Clark W**. 1992. International code of nomenclature of bacteriaInternational Code of Nomenclature of Bacteria: Bacteriological Code, 1990 Revision. ASM Press.
16. **Bull CT**, **De Boer SH**, **Denny TP**, **Firrao G**, **Saux MF-L**, **Saddler GS**, **Scortichini M**, **Stead DE**, **Takikawa Y**. 2008. Demystifying the nomenclature of bacterial plant pathogens. J Plant Pathol **90**:403–417.
17. **Young JM**, **Park DC**, **Shearman HM**, **Fargier E**. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. Syst Appl Microbiol **31**:366–377.

18. **Almeida NF**, **Yan S**, **Cai R**, **Clarke CR**, **Morris CE**, **Schaad NW**, **Schuenzel EL**, **Lacy GH**, **Sun X**, **Jones JB**, **Castillo JA**, **Bull CT**, **Leman S**, **Guttman DS**, **Setubal JC**, **Vinatzer BA**. 2010. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. Phytopathology **100**:208–15.
19. **Jacques M-A**, **Arlat M**, **Boulanger A**, **Boureau T**, **Carrère S**, **Cesbron S**, **Chen NWG**, **Cociancich S**, **Darrasse A**, **Denancé N**, **Fischer-Le Saux M**, **Gagnevin L**, **Koebnik R**, **Lauber E**, **Noël LD**, **Pieretti I**, **Portier P**, **Pruvost O**, **Rieux A**, **Robène I**, **Royer M**, **Szurek B**, **Verdier V**, **Vernière C**. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. Annu Rev Phytopathol **54**:163–187.
20. **Langlois PA**, **Snelling J**, **Hamilton J**, **Bragard C**, **Koebnik R**, **Verdier V**, **Triplett L**, **Blom J**, **Tisserat N**, **Leach J**. 2017. Characterization of the Xanthomonas translucens complex using draft genomes, comparative genomics, phylogenetic analysis, and diagnostic LAMP assays. Phytopathology **107**:519–527.
21. **Bull CT**, **Koike ST**. 2015. Practical benefits of knowing the enemy: Modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. Annu Rev Phytopathol **53**:157–180.
22. **Richter M**, **Rosselló-Móra R**. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A **106**:19126–31.
23. **Konstantinidis KT**, **Tiedje JM**. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci USA **102**.
24. **Young JM**, **Bull CT**, **Boer SH De**, **Firrao G**, **Gardan L**, **Saddler GE**, **Stead DE**, **Takikawa Y**. 2001. International standards for naming pathovars of phytopathogenic bacteria.
25. **Adriko J**, **Aritua V**, **Mortensen CN**, **Tushemereirwe WK**, **Kubiriba J**, **Lund OS**. 2012. Multiplex PCR for specific and robust detection of *Xanthomonas campestris* pv. *musacearum* in pure culture and infected plant material. Plant Pathol **61**:489–497.
26. **Lewis Ivey ML**, **Tusiime G**, **Miller SA**. 2010. A polymerase chain reaction assay for the detection of *Xanthomonas campestris* pv. *musacearum* in banana. Plant Dis **94**:109–114.
27. **Ash GJ**, **Lang JM**, **Triplett LR**, **Stodart BJ**, **Verdier V**, **Cruz CV**, **Rott P**, **Leach JE**. 2014. Development of a genomics-based LAMP (Loop-Mediated Isothermal Amplification) assay for detection of *Pseudomonas fuscovaginae* from rice. Plant Dis **98**:909–915.
28. **Lang JM**, **Hamilton JP**, **Diaz MGQ**, **Van Sluys MA**, **Burgos MRG**, **Vera Cruz CM**, **Buell CR**, **Tisserat NA**, **Leach JE**. 2010. Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. Plant Dis **94**:311–319.
29. **Lang JM**, **Langlois P**, **Nguyen MHR**, **Triplett LR**, **Purdie L**, **Holton TA**, **Djikeng A**, **Vera Cruz CM**, **Verdier V**, **Leach JE**. 2014. Sensitive detection of *Xanthomonas oryzae* pathovars oryzae and oryzicola by loop-mediated isothermal amplification. Appl Environ Microbiol **80**:4519–30.
30. **Verdier V**, **Triplett LR**, **Corral R**, **Leach JE**. 2011. Genome-enabled primer design to distinguish geographic origin of *Xanthomonas oryzae* pvs. *oryzicola* and *oryzae*. Phytopathology **101**:S183.
31. **Triplett LR**, **Verdier V**, **Campillo T**, **Van Malderghem C**, **Cleenwerck I**, **Maes M**, **Deblais L**, **Corral R**, **Koita O**, **Cottyn B**, **Leach JE**. 2015. Characterization of a novel clade of *Xanthomonas* isolated from rice leaves in Mali and proposal of *Xanthomonas maliensis* sp. nov. Antonie Van Leeuwenhoek **107**:869–81.
32. **Moreira LM**, **Almeida NF**, **Potnis N**. 2010. Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of Xanthomonas fuscans subsp. aurantifolii. BMC Genomics **11**.
33. **Kumar S**, **Stecher G**, **Tamura K**. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol **33**:1870–4.
34. **Tajima F**, **Nei M**. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol Biol Evol **1**:269–285.

35. **Saitou N**, **Nei M**. 1987. The neighboring-joining method: a new method for reconstructing phylogenetic trees. MolBiolEvol **4**:406–425.
36. **Bankevich A**, **Nurk S**, **Antipov D**, **Gurevich AA**, **Dvorkin M**, **Kulikov AS**, **Lesin VM**, **Nikolenko SI**, **Pham S**, **Prjibelski AD**, **Pyshkin A V.**, **Sirotkin A V.**, **Vyahhi N**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455.
37. **Gurevich A**, **Saveliev V**, **Vyahhi N**, **Tesler G**. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics **29**:1072.
38. **Wattam AR**, **Abraham D**, **Dalay O**, **Disz TL**, **Driscoll T**, **Gabbard JL**, **Gillespie JJ**, **Gough R**, **Hix D**, **Kenyon R**, **Machi D**, **Mao C**, **Nordberg EK**, **Olson R**, **Overbeek R**, **Pusch GD**, **Shukla M**, **Schulman J**, **Stevens RL**, **Sullivan DE**, **Vonstein V**, **Warren A**, **Will R**, **Wilson MJC**, **Yoo HS**, **Zhang C**, **Zhang Y**, **Sobral BW**. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res **42**:D581-91.
39. **Brettin T**, **Davis JJ**, **Disz T**, **Edwards RA**, **Gerdes S**, **Olsen GJ**, **Olson R**, **Overbeek R**, **Parrello B**, **Pusch GD**, **Shukla M**, **Thomason JA**, **Stevens R**, **Vonstein V**, **Wattam AR**, **Xia F**, **Xia F**. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep **5**:8365.
40. **Karganilla A**, **Paris-Natural M**, **Ou SH**. 1973. A comparative study of culture media for *Xanthomonas oryzae*. PhilippAgric **57**:141–152.
41. **Reimers PJ**, **Leach JE**. 1991. Race-specific resistance to *Xanthomonas oryzae* pv. *oryzae* conferred by bacterial blight resistance gene *Xa-10* in rice *Oryza sativa* involves accumulation of a lignin-like substance in host tissues. Physiol Mol Plant Pathol **38**:39–55.
42. **Kauffman HE**, **Reddy APK**, **Hsieh SPY**, **Merca SD**. 1973. Improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis Report **57**:537–541.
43. **Vidaver AK**. 1977. Maintenance of viability and virulence of *Corynebacterium nebraskense*. Phytopathology **67**:825–827.
44. **Coutinho T**. 1988. Bacterial leaf streak of maize: A study on host specificity and cultivar susceptibility. Thesis. University of Natal, Pietermaritzburg, South Africa. University of Natal, Pietermaritzburg, South Africa.
45. **Ye J**, **Coulouris G**, **Zaretskaya I**, **Cutcutache I**, **Rozen S**, **Madden TL**. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics **13**:134.
46. **Parkinson N**, **Cowie C**, **Heeney J**, **Stead D**. 2009. Phylogenetic structure of *Xanthomonas* determined by comparison of gyrB sequences. Int J Syst Evol Microbiol **59**:264–274.
47. **Parkinson N**, **Aritua V**, **Heeney J**, **Cowie C**, **Bew J**, **Stead D**. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. Int J Syst Evol Microbiol **57**:2881–7.
48. **Wasukira A**, **Tayebwa J**, **Thwaites R**, **Paszkiewicz K**, **Aritua V**, **Kubiriba J**, **Smith J**, **Grant M**, **Studholme DJ**. 2012. Genome-wide sequencing reveals two major sub-lineages in the genetically monomorphic pathogen *Xanthomonas campestris* pathovar *musacearum*. Genes (Basel) **3**:361–77.
49. **Qhobela M**, **Claflin LE**. 1988. Characterization of *Xanthomonas campestris* pv. *pennamericanum* pv. nov., Causal Agent of Bacterial Leaf Streak of Pearl MilletT. Int J Syst Bacteriol **38**:362–366.
50. **Goszczynska T**, **Botha WJ**, **Venter SN**, **Coutinho TA**. 2007. Isolation and identification of the causal agent of brown stalk rot, a new disease of maize in South Africa. Plant Dis **91**:711–718.
51. **Jalan N**, **Kumar D**, **Andrade MO**, **Yu F**, **Jones JB**, **Graham JH**, **White FF**, **Setubal JC**, **Wang N**. 2013. Comparative genomic and transcriptome analyses of pathotypes of *Xanthomonas citri* subsp. *citri* provide insights into mechanisms of bacterial virulence and host range. BMC Genomics **14**:551.
52. **Vauterin L**, **Yang P**, **Hoste B**, **Swings J**, **Kersters K**. 1992. Taxonomy of xanthomonads from cereals and grasses based on SDS-PAGE of proteins, fatty acid analysis and DNA hybridization. J

Gen Microbiol **138**:1467–1477.
53. **Vinatzer BA**, **Weisberg AJ**, **Monteil CL**, **Elmarakeby HA**, **Sheppard SK**, **Heath LS**. 2017. A proposal for a genome similarity-based taxonomy for plant-pathogenic bacteria that is sufficiently precise to reflect phylogeny, host range, and outbreak affiliation applied to *Pseudomonas syringae sensu lato* as a proof of concept. Phytopathology **107**:18–28.
54. **Gevers D**, **Cohan FM**, **Lawrence JG**, **Spratt BG**, **Coenye T**, **Feil EJ**, **Stackebrandt E**, **de Peer Y Van**, **Vandamme P**, **Thompson FL**, **Swings J**. 2005. Opinion: Re-evaluating prokaryotic species. Nat Rev Microbiol **3**:733–739.
55. **Whitman WB**. 2015. Genome sequences as the type material for taxonomic descriptions of prokaryotes. Syst Appl Microbiol **38**:217–222.
56. **Coutinho TA**, **Wallis FM**. 1990. Bacterial streak disease of maize (*Zea mays* L.) in South Africa. J Phytopathol **133**:112–112.
57. **Zhao S**, **Poulin L**, **Rodriguez RL**, **Serna NF**, **Liu SY**, **Wonni I**, **Szurek B**, **Verdier V**, **Leach JE**, **He YQ**, **Feng JX**, **Koebnik R**. 2012. Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology2012/09/11. **102**:948–956.
58. **Pruvost O**, **Magne M**, **Boyer K**, **Leduc A**, **Tourterel C**, **Drevet C**, **Ravigné V**, **Gagnevin L**, **Guérin F**, **Chiroleu F**, **Koebnik R**, **Verdier V**, **Vernière C**. 2014. A MLVA genotyping scheme for global surveillance of the citrus pathogen Xanthomonas citri pv. citri suggests a worldwide geographical expansion of a single genetic lineage. PLoS One **9**:e98129.
59. **Poulin L**, **Grygiel P**, **Magne M**, **Gagnevin L**, **Rodriguez-R LM**, **Forero Serna N**, **Zhao S**, **El Rafii M**, **Dao S**, **Tekete C**, **Wonni I**, **Koita O**, **Pruvost O**, **Verdier V**, **Vernière C**, **Koebnik R**. 2015. New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by Xanthomonas oryzae. Appl Environ Microbiol **81**:688–98.
60. **Boch J**, **Bonas U**. 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol2009/04/30. **48**:419–436.
61. **Mew TW**, **Vera C. M. C**, **Medalla ES**. 1992. Changes in race frequency of *Xanthomonas oryzae* pv. *oryzae* in response to rice cultivars planted in the Philippines. Plant Dis **76**:1029.
62. **Salzberg SL**, **Sommer DD**, **Schatz MC**, **Phillippy AM**, **Rabinowicz PD**, **Tsuge S**, **Furutani A**, **Ochiai H**, **Delcher AL**, **Kelley D**, **Madupu R**, **Puiu D**, **Radune D**, **Shumway M**, **Trapnell C**, **Aparna G**, **Jha G**, **Pandey A**, **Patil PB**, **Ishihara H**, **Meyer DF**, **Szurek B**, **Verdier V**, **Koebnik R**, **Dow JM**, **Ryan RP**, **Hirata H**, **Tsuyumu S**, **Won Lee S**, **Seo Y-S**, **Sriariyanum M**, **Ronald PC**, **Sonti R V**, **Van Sluys M-A**, **Leach JE**, **White FF**, **Bogdanove AJ**. 2008. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics **9**:204.
63. **Crossman LC**, **Gould VC**, **Dow JM**, **Vernikos GS**, **Okazaki A**, **Sebaihia M**, **Saunders D**, **Arrowsmith C**, **Carver T**, **Peters N**, **Adlem E**, **Kerhornou A**, **Lord A**, **Murphy L**, **Seeger K**, **Squares R**, **Rutter S**, **Quail MA**, **Rajandream M-A**, **Harris D**, **Churcher C**, **Bentley SD**, **Parkhill J**, **Thomson NR**, **Avison MB**. 2008. The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. Genome Biol **9**:R74.
64. **Pieretti I**, **Royer M**, **Barbe V**, **Carrere S**, **Koebnik R**, **Cociancich S**, **Couloux A**, **Darrasse A**, **Gouzy J**, **Jacques MA**, **Lauber E**, **Manceau C**, **Mangenot S**, **Poussier S**, **Segurens B**, **Szurek B**, **Verdier V**, **Arlat M**, **Rott P**. 2009. The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited *Xanthomonadaceae*. BMC Genomics2009/12/19. **10**:616.

CHAPTER 5

CONCLUSION

The imminent increase in our human population coupled with documented changes in our climate are creating an intense challenge for agriculture. Food security requires creative use of resources and innovative management strategies for pathogens and pests. To apply effective management strategies, an accurate diagnosis of crop diseases is necessary. Identification of bacterial plant pathogens in particular, can be difficult because the pathogens are diverse, they cause a wide range of symptoms, and the diseases they cause can be highly influenced by environmental pressures. As a result, rapid evolution in the pathogen can occur, making existing assays potentially obsolete. This dynamic evolution of pathogenicity is not well understood but successful diagnostic assays and management of diseases and pests need to account for this. Emerging diseases in important cereal crops, such as corn and rice, threaten yields and therefore, food security, global economies and trade. This dissertation examines (1) the application of comparative genomics to accurately identify bacterial plant pathogens and elucidate complex taxonomy for important *Xanthomonas* pathogens, (2) the analysis of environmental strains to understand how the evolution of pathogenicity in agroecosystems occurs, and lastly, (3) the characterization of host range and genomic information for an emerging disease on corn in the United States.

Chapter 2 builds on previous work leveraging comparative genomics to identify unique loci for PCR-based detection and differentiation of *Xanthomonas oryzae* pv. *oryzae* and *X. o.* pv *oryzicola,* causal agents of bacterial blight and bacterial leaf streak of rice, respectively. These assays were adapted to loop-mediated isothermal amplification (LAMP) to enable the detection of these organisms in the field without the need for expensive and sophisticated equipment. Additional loci were identified to differentiate lineages (African vs. Asian) of these organisms. LAMP is advantageous because it targets multiple locations on a single locus (six vs. two in conventional PCR), therefore inherently creating more

sensitivity and specificity. A molecular assay based on a single locus for example, could no longer be a valuable marker for an organism due to rapidly evolving populations, even within a few cropping seasons.

While significantly more sensitive than conventional PCR, LAMP also brings some difficulties. Due to the extreme efficiency of the reaction, aerosolized DNA can quickly contaminate workspaces and generate false-positive results. Closed-tube reactions are necessary to avoid this potential setback. LAMP could have a permanent place in the field and in quarantine facilities for precise diagnostics if costs go down for fluorescent isothermal mixes and GPS enabled equipment; or stabilized master mixes of reagents that can be easily transported and incubated on site in a simple water bath. A long term vision for the success of this approach to diagnostics and an ideal scenario for extension agents and even growers would be a 'LAMP library' of publicly available, validated assays for multiple pathogens. The challenge for the success of such a broad collection of diagnostic tools would be full validation of assays which requires a curated collection of pure and diverse isolates to test against. This is no simple task. It would require coordinated sharing of DNAs and strict protocols for confident validations.

Advocacy for culture collections and publicly available sequence data should be a priority for policy makers and funding agencies. That said, in the near future it is possible that LAMP will also become obsolete with the advent of advanced, low-cost sequencing capabilities because rather than run an assay on an organism of interest, the genome could simply be captured and quickly compared to public databases for identification. An even simpler approach that is becoming increasingly attractive and feasible for field diagnostics are lateral flow devices using immunochromatographic strip tests. Where a piece of suspected infected tissue can be incubated then run with the antisera/antibody-infused strip for genus or species specific results in less than an hour. I believe future of plant diagnostics will be in a more true integration of advances in genomics and field ready applications. Safeguarding crops begins with knowing what is causing losses and if as scientists, we can put the power of genomics, immunology and technology in the hands of a grower in a developing country; we have made a significant contribution to protecting agriculture.

Chapter 3 is an investigation of *X. o.* pv. *leersiae*, a close relative of *X. o.* pvs. *oryzae* and *oryzicola.* The strains that are the focus of the study were isolated from grass hosts in the genus *Leersia* adjacent to rice production sites in China and more recently in Burkina Faso, however, the pathogen has been isolated from *Leersia* from many rice producing countries, including the USA. This pathosystem is a unique window into the evolution of pathogenicity because it involves two closely related hosts, rice and southern cutgrass. My work addressed first if these organisms were actually *X. o.* pv. *oryzae* or *X. o.* pv. *oryzicola* that were colonizing southern cutgrass as an alternative host. Using comparative genomics and disease host phenotyping, I demonstrated that the *Leersia* strains are closest to *X. oryzae* species, but are distinct from *X. o.* pv. *oryzae* or *X. o.* pv. *oryzicola*. Thus, I propose that *X. campestris* pv. *leersiae* be renamed to *X. oryzae* pv. *leersiae.* This taxonomic work lays the foundation for developing diagnostic assays to differentiate *X. o.* pv. *leersiae* from the other two rice pathogens. Misidentification of *X. oryzae* can have deep implications in trade and quarantine, especially in the United States where they are select agents and therefore, highly regulated.

During the in-depth analysis of the *X. oryzae* genomes, we found the *X. o.* pv. *leersiae* had a mostly conserved repertoire of type III effectors but a unique set of transcription activator like effectors (TALEs). Through a comprehensive examination of their biology and behavior on different hosts including multiple rice varieties, I hoped to address the question "Are *X. o.* pv. *leersiae* evolving parallel to the two important rice pathogens, or were they once a rice pathogen that adapted to a new host?"

To answer this question, I compared *X. o.* pv. *leersiae* and *X. oryzae* TALEs. Beyond the novelty of identifying new TALEs, I used predictive software programs to (1) compare the repeat variable domains (RVDs) of these TALEs to each other and to other TALEs from other bacteria, (2) determine the specific sequence (effector binding element, EBE) that the *X. o.* pv. *leersiae* TALE would bind in the host promoterome, and finally, (3) predict what genes contained these EBEs and infer homology in rice. This allowed us to predict if they could target similar genes in *Leersia* spp. following a strategy much like rice pathogens. Our predictions showed that the *X. o.* pv. *leersiae* TALE indeed target similar genes to other *X. oryzae*, including the well-known family of OsSWEET sugar transporters and a mono-copper oxidase

orthologue. This suggests that the *X. oryzae* use similar virulence strategies to infect their hosts. Next steps should include functional validation of the predicted gene targets of the *X. o.* pv. *leersiae* TALEs in *Leersia* and *Oryza* spp.

Our alternative hypothesis is that *X. o.* pv. *leersiae* are just simply 'pathogens in waiting.' Meaning, they are a snapshot into the evolution of a rice pathogen that began as a pathogen of a closely related grass. It is possible these weakly virulent strains could simply acquire or adapt a rice-specific TALE and become a more virulent rice pathogen. Based on the information we gathered so far, I believe it actually might be that these were once rice pathogens that were dispersed into the surrounding ecosystem, adapted to a new grass host and maintained the structural ability to deliver effectors and survive in these environments, but over time evolved to capture the orthologous gene targets in *Leersia* spp. It is understood that TALEs can evolve as they are often flanked by transposases to allow for rapid movement, but how quickly is unknown. Because we do not have a finished *Leersia* genome with complete annotation or any functional data to demonstrate upregulation of predicted targets in rice or *Leersia* spp., this is speculation. Sequencing another member of this genus, preferably *L. hexandra* from which our strains were isolated, would not only inform this work but would contribute to broad analysis of members in the *Poaceae.* The inferences made using the *L. perrieri* genome may not be are not accurate for this system, but are at least informative for moving forward.

An interesting study would be to swap TALEs from *X. o.* pvs. *oryzicola* and *oryzae* into *X. o.* pv. *leersiae* and vice versa, then monitor successful upregulation of targets by qPCR. Conversely, if resistance is triggered via an executor R gene, this could demonstrate how the pathogens were forced to adapt to their new host. On a larger scale, unraveling these mechanisms of virulence (or avoidance of resistance) can inform targets for breeders to combat yield-threating diseases. Any further investigation of this unique pathosystem should include more strains. A major limitation of this work was that we only had two strains available to characterize. A large scale survey for more isolates from *Leersia* spp. and other weed hosts would be important for future studies enabled by a validated diagnostic assay for this pathovar.

The environments surrounding agriculture are an abundant source of information that often does not garner the attention that cultivated crops do, especially if these environments are composed of non-invasive organisms and are well-managed. This is economically understandable, but I argue that plants surrounding crops, including weeds, should be considered in the scope of plant pathology, i.e., an understanding of the system or phytobiome. Alternative hosts for plant pathogens may not simply be harboring pathogenic microorgansims, but they could be a window into the current evolution and adaptation of bacteria, fungi and viruses, particularly in changing climates. It is hypothesized that agro-ecosystems will continue to select for new pathogens unless they are diversified and I agree with the need to support and enhance biodiversity in all ecosystems.

The final research chapter (4) of this dissertation describes *X. vasicola* pv. *vasculorum,* the causal agent of bacterial leaf streak of corn. This year was the first report of this disease in the United States and it is considered an emerging pathogen that could reach epidemic proportions. Almost nothing, however, is known about the pathogen's etiology or how or when it was introduced. The United States is a leading exporter of corn and this industry reaches up to $6 billion on average annually. After the diagnosis was completed, we continued to investigate more about the host range of this bacteria, elucidate the taxonomy of this organism, and generate a much needed diagnostic assay to support extension, industry, state and federal regulatory agencies to monitor its spread. We sequenced several isolates, and reported the sequences of two representative strains, one collected in 1988 in S. Africa and one from the recent epidemic in Nebraska. Both were highly virulent to corn and sugarcane but weakly virulent to sorghum and shared 99% sequence identity. The diagnosis that the new disease observed in the US is bacterial streak of maize and that the causal agent is *X. v.* pv. *vasculorum* is significant because of the threat to export, since this disease was never previously reported in the US. We systematically screened many isolates, optimized an isolation protocol, developed genomics based molecular diagnostics (now widely distributed and validated) and settled the complicated taxonomy of this organism with full genome comparisons and multi-locus sequence alignments. Interestingly, while this organism is highly virulent, it

does not contain TALEs. An investigation into what virulence mechanisms (perhaps prolific cell wall degrading enzymes) it uses to cause such significant disease so quickly would be very interesting.

We have only begun to unravel this disease complex. Going forward, there is much to uncover about the basic biology of *X. v.* pv *vasculorum.* Priorities should include: (1) population biology *in planta* and broadly across corn growing regions in the US; (2) determining yield loss due to this disease; (3) understanding how it is transmitted; (4) screening for possible resistance in existing germplasm; (5) whole genome SNP comparisons of *X. v.* pv. *vasculorum* to geographically diverse strains and closely related members of the *X. vasicola* complex (4) collecting new isolates from other corn growing countries to help determine if this pathogen is a recent introduction, or if it has been present; (5) determining the role of climatic factors on the emergence of the pathogen.

X. v. pv. *vasculorum* are highly virulent to certain corn varieties and are capable of moving rapidly through leaf tissues; yet, they do not contain TALEs. This is a powerful infection strategy by the organism to quickly reproduce then disseminate through wind and/or water. In corn, overhead irrigation is common in the United States, creating a perfect mechanism for the bacteria to spread. Information about the bacteria's behavior *in planta,* the environmental conditions conducive for disease, the epidemiology of the disease, and how the bacteria are surviving from season to season will inform how to reduce potential yield losses and determine the best management strategies for growers in the future.

In summary, *Xanthomonas* genomics has advanced many fields of research, not just diagnostics. Given the optimization of new technologies, such as portable sequencers like the Oxford MinION, I think the use of genomics for diagnostics will only increase. The potential to capture an entire genome of an emerging bacterial pathogen in the field in one day is in the reasonable future. The application of 'phylogenomics' of emerging and existing diseases will create a baseline of understanding how these microorganisms adapt to new hosts and environments. Careful consideration of sampling techniques and observing not just causal agents of disease, but rather the phytobiome of a plant will allow us to unravel complex interactions occurring both beneficially and detrimentally to crops.

Appendix A is a collection of the supplemental data complimenting chapters 2-4. Appendix B is a summary of my activities beyond advanced research. It is focused on outreach and education in our local school system and my mentorship of diverse students and peers. In my non-traditional role as a graduate student, I have been fortunate to have these opportunities that have broadened my overall experience.

APPENDIX A

SUPPLEMENTARY MATERIALS

A.1 CHAPTER 2 - SUPPLEMENTARY MATERIALS

Figure A.1.1. *Xanthomonas oryzae* pv. *oryzae* sequence unique to African populations

***X. oryzae* pv. *oryzae*, NAI8, Niger, PRJNA228925**

TTGAAGCCCGTGATCGAGGGGGCGAACCAGACTTGGATGGTGGAGGCCATTGAAAGCAAGTTGATCTGGT
TGGACTGCCTAGACGAGGGTCCGGAGGGAATGAATGCACGCAAGCATGCGCAGAAGCAGCTTAAAGAATT
GAAAGGCGTTGTTGCAGAGAATAATTCACTCAGATTCGTTTACTTTATCTGCTCTCGCGATCGCGTTCGT
ATCTCTACAGTTAAAAATTCAATATATTCGAAATGGCGGGATAGGATGTATCTCTGGCTTGAGATTGGGG
GATCCAGAGAGCTTGAGCAGATCCACTTCCCTTGCCCTATAAGGGGGCTTGATGGTGCTCCATTGAAAGT
TCCTTTCAGAGTTACCGACAAGTTTATAACTTTCGATGTTTATGATTTTGGCCTTGGCGGGAAAAAGAAA
GTGAGTATGGATTTGCACTATATTTTAGAAATGCAGGGTGTCGACCTGGGTCATGTGTCGCGAGTTCAAT
ATGTCGGCACAACGTTTGACCCAGATTTGCGAGTCCTTCCCGCCACAAGTCCTGGCGCAAACGGCGGGCA
TGATGGCTACATTCGCAGTCGAGAAGCGCGCTCTTCGGATGGTAGTGATCTTTTTTTCTATTCCCATTCC
TTGAAGTCTAATATAATATCTGCACAAAACAAGCATGGCCTACTGTTTGCTGGAACAAATTATGCTGCGG
GAATATTGGGTGCTGCCGATGAAGGGGTTCTCATGGAAAAGGCTCTCATTGCTTATTTTGAGCCTAAACA
AGCCGGGAACCTTGAGGCGGAAAAGTCTCGCCTAGTCAACATGGCGAGAAAGAATAATGTTACCGAGATC
AGGTTGGGCCTTGATTTTGTAGCGCCTGGAAGCTATTTCACTTACGGAAGTGAGGTTGCCCCTGCTAGTC
CCAGTATTTTGGCGACAATAAATTTTACCGATCGAGGAATTTTGATTGATCGCCATCTGGACTCGTCGCT
TGAAGAGCTTATGTCTTCGAAGTGGTCCATTCATAACGACGCGGCTCCGGAGACCTCCTCTCAGGGCTAA

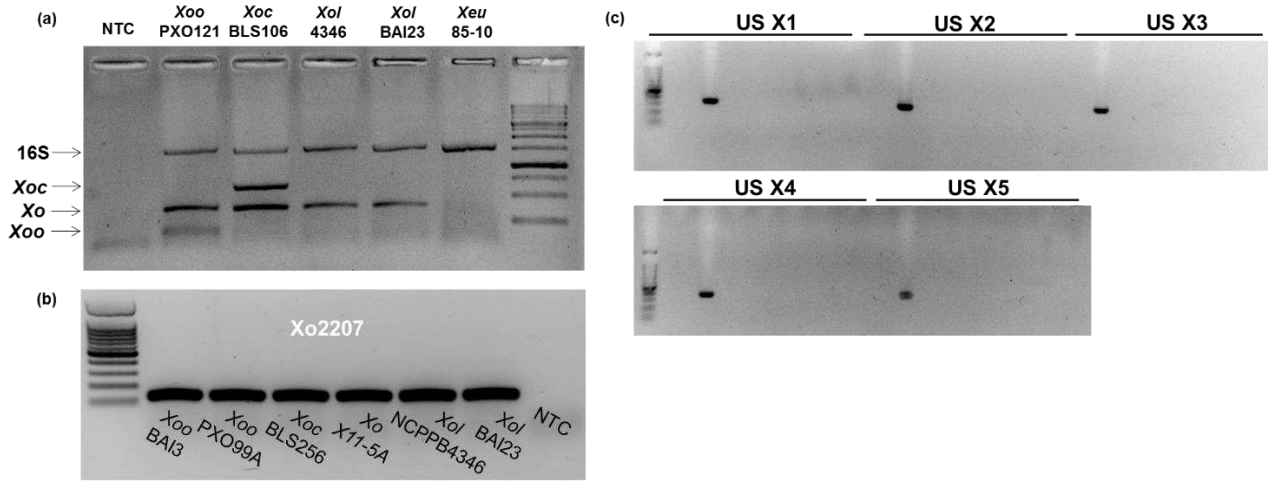

Figure A.2.1. Amplification of *X. o.* pv. *leersiae* with PCR primers specific for *X. oryzae.* The primers differentiated *X. o.* pv. *leersiae* from *X. orzyae* pvs. *oryzae* and *oryzicola* (a) but demonstrated a shared locus among them. *X. o.* pv. *leersiae* strains and all other *X. oryzae* amplify with primers designed to an additional *X. oryzae* specific locus (Xo2207). (b). US Xo-specific primers do not amplify *X. o.* pv. *leersiae* (c). Lane 1: 1 Kb Plus Ladder (Gold Biotechnology, St. Louis, MO), Lane 2: no template control (NTC), Lane 3: US Xo X11-5A, Lane 4: *X. orzyae* pv. *oryzae* PXO99A, Lane 5: *X. orzyae* pv. *oryzicola* BLS256, Lane 6: *X. o.* pv. *leersiae* NCPPB4346, Lane 7: *X. o.* pv. *leersiae* BAI23, Lane 8: *X. translucens* LH2-1. Multiplex primers and Xo2207 primers (a,b: Lang et al 2010); US Xo-specific primers (c: Triplett et al 2011)

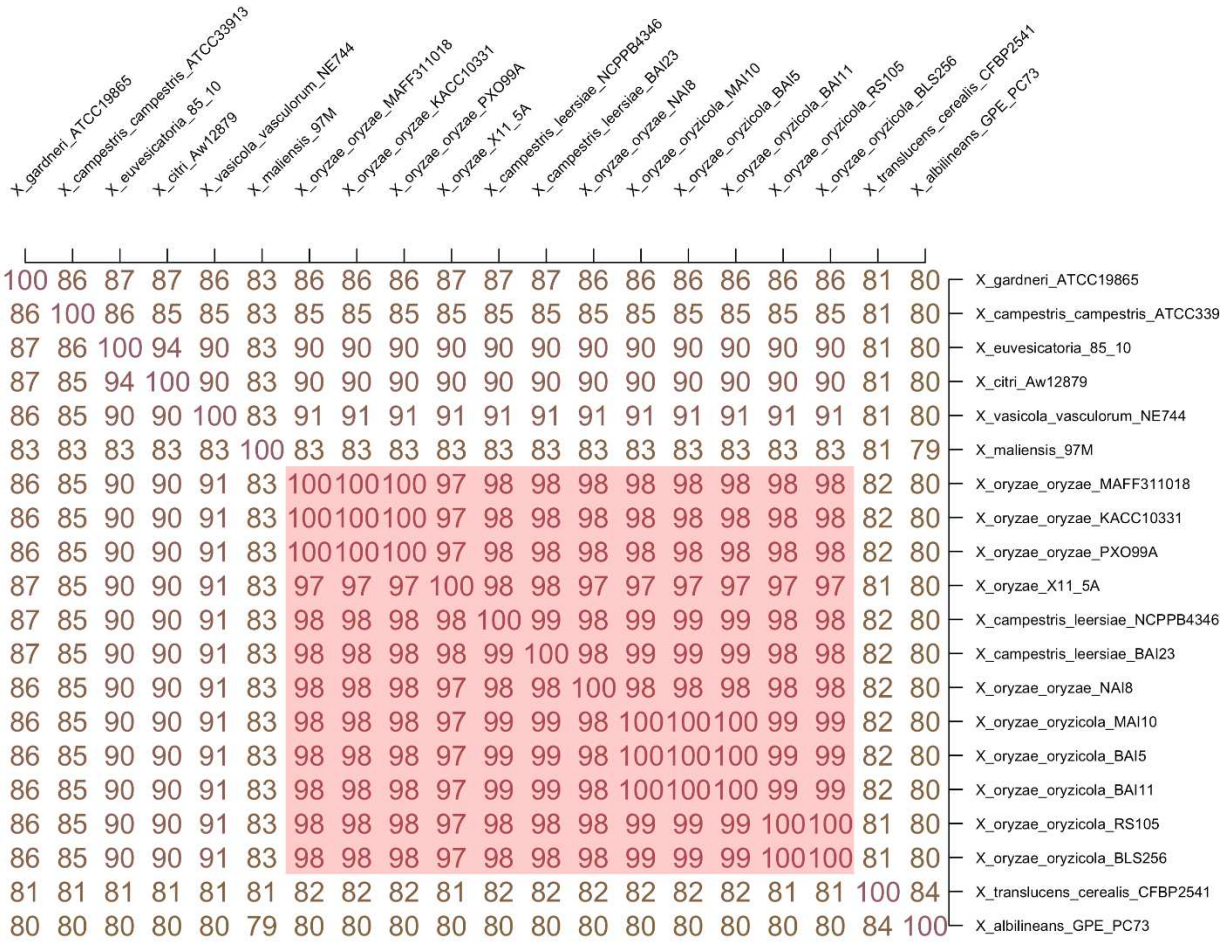

Figure A.2.2. Average nucleotide identity (ANI) genome matrix comparing *X. c.* pv. *leersiae* whole draft genomes with other *Xanthomonas* species. Red box designates strains grouping within the same species at a >95% similarity. Values calculated using Average Nucleotide Genome Matrix Calculator http://enve-omics.ce.gatech.edu/g-matrix/. All assembled genomes were collected from NCBI Genome database except *X. c.* pv. *leersiae* BAI23 which was generated in this study.

Figure A.2.3. *X. oryzae* strains, including *X. oryzae* pv. *leersiae* (Xol), *X. o.* pv. *oryzae* (Xoo), *X. o.* pv. *oryzicola* (Xoc), and US *X. oryzae* (Xo), do not induce spreading, watersoaked lesions on barley after infiltration.

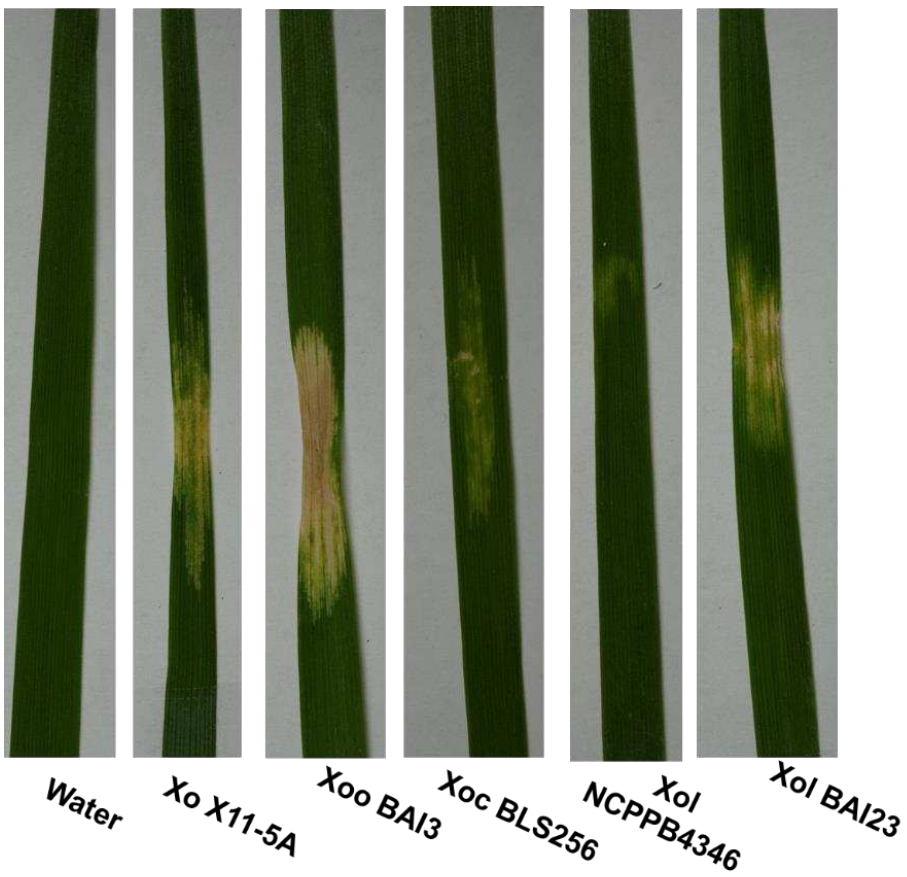

Figure A.2.4. *X. oryzae* strains, including *X. oryzae* pv. *leersiae* (Xol), *X. o.* pv. *oryzae* (Xoo), *X. o.* pv. *oryzicola* (Xoc), and US *X. oryzae* (Xo), do not induce spreading, watersoaked lesions on wheat after infiltration.

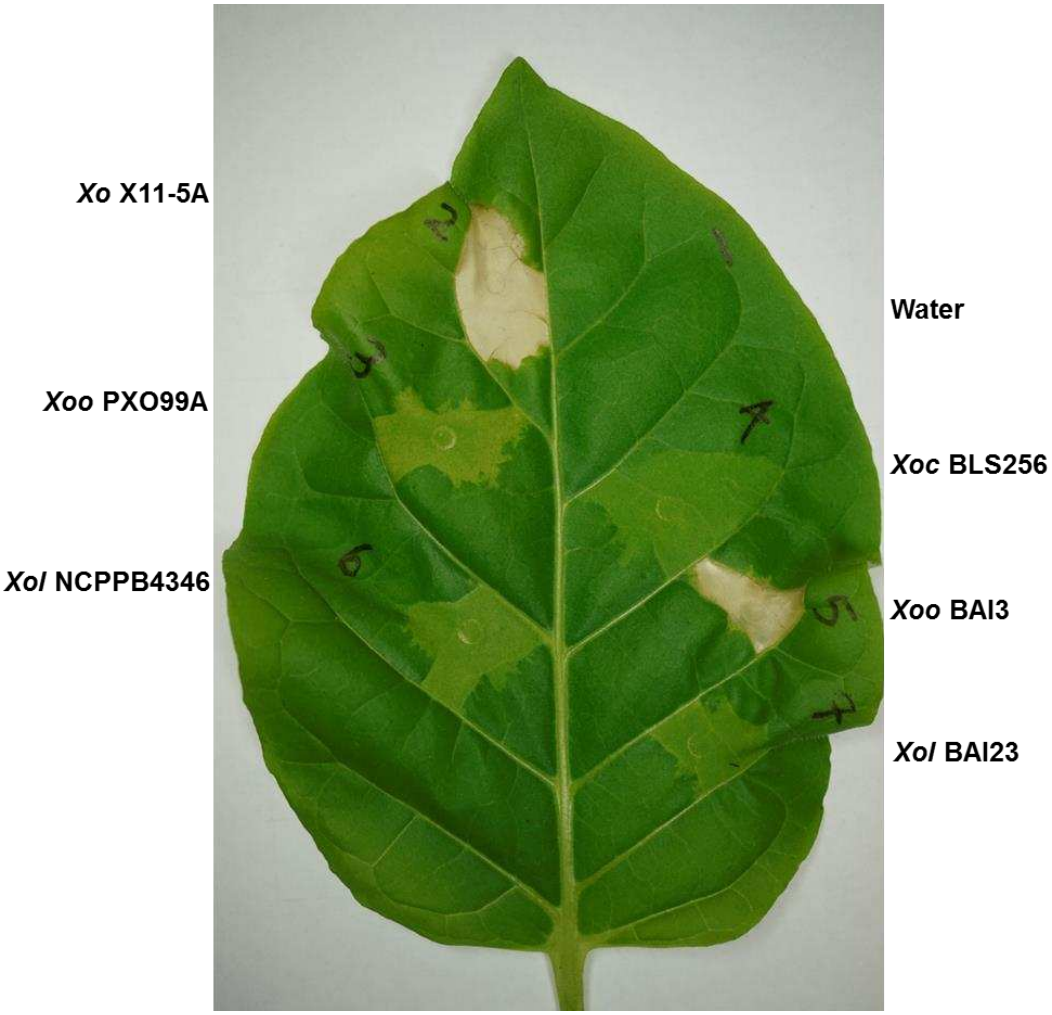

Figure A.2.5. *X. oryzae* pv. *leersiae* can evade host defense responses in *Nicotiana tabacum*. *N. tabacum* leaves were infiltrated with water (1), *X. oryzae* X11-5A (2), X. *oryzae* pv. *oryzae,* BAI3 (5), PXO99A (3), *X. oryzae* pv. *oryzicola* BLS256 (4) and *X. oryzae* pv. *leersiae* NCPPB4346 (6) and BAI23 (7). Plants were observed at 4 dpi.

A.3 CHAPTER 4 – SUPPLEMENTARY MATERIALS

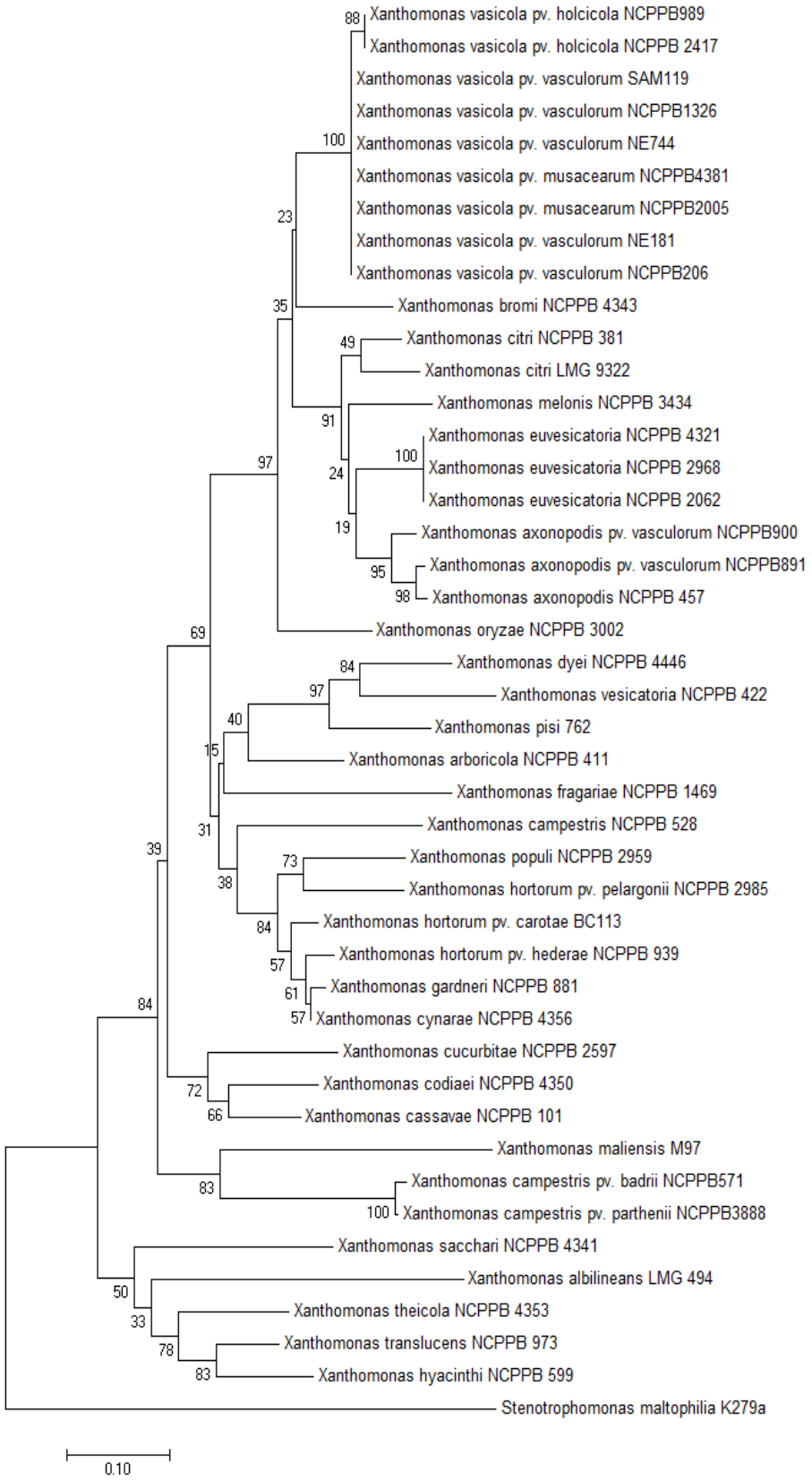

Figure A.3.1. Phylogeny of what *Xanthomonas* species based on partial *gyrB* sequence alignment. Bootstrap values are shown at nodes generated from 1000 replicates. Branch lengths measured in the number of substitutions per site. A total of 540 bp from the *gyrB* open reading frame were used in the final dataset.

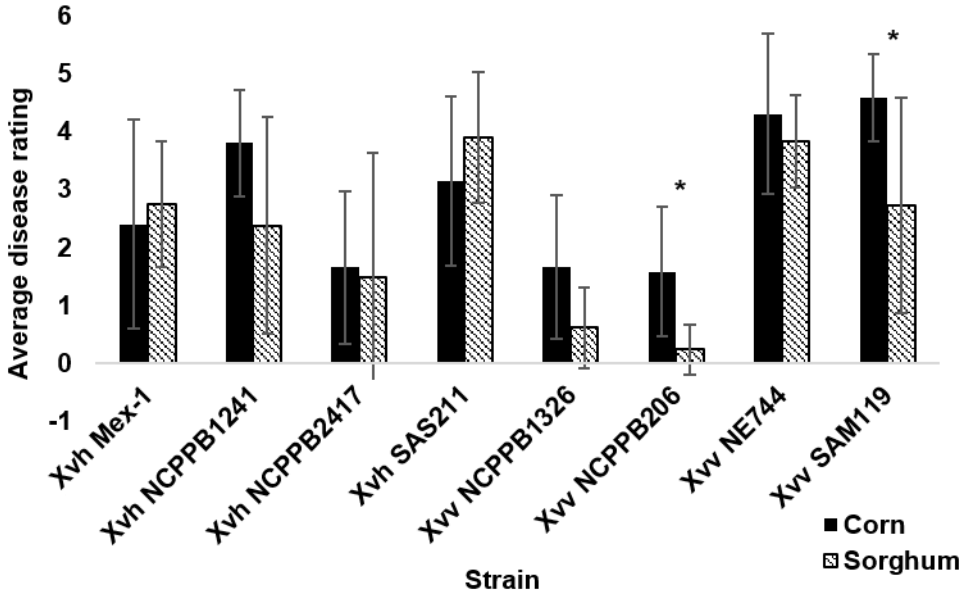

Figure A.3.2. Disease caused by *X. vasicola* pv. *vasculorum* (Xvv) and *X. v.* pv. *holcicola* (Xvh) on corn (cv. hybrid DKC 61-88) and sorghum (cv. Mycogen IG588). Stems of 4-week-old plants were injected with 108 CFU ml-1 of each strain, and disease was assessed at 7 days post inoculation (dpi). Disease was rated according a scale of 1 to 5, with 1 demonstrating minimal disease and 5 meaning extensive lesions (Coutinho, 1988). The entire experiment was replicated four times and combined data from all replications is shown here. Bars represent standard error and asterisks designate significance at $P < 0.0021$ and $P < 0.0040$ for Xvv NCPPB206 and SAM119, respectively.

Figure A.3.3. Qualitative disease phenotyping of *X. v.* pv. *vasculorum* (Xvv) and *X. v.* pv. *holcicola* (Xvh) on sorghum (cv.Mycogen IG588). Four-week-old plants were stabbed and infiltrated with 10^8 CFU ml^{-1} of each strain. Reactions were recorded 7 dpi.

Figure. A.3.4. *X. v.* pv. *vasculorum* NE744 bacterial exudate on corn (cv. hybrid DKC 61-88) 7 dpi.

Supplementary Table A.3.1. Quantitative disease response of sugarcane to infection with diverse strains

Strain	Host[a]	Mean Lesion Length (cm)[bc]
Xanthomonas vasicola* pv. *vasculorum		
NCPPB1326	Sugarcane	1.21 ± 0.67 c
ZCP611	Sugarcane	5.26 ± 5.01 ab
UVZ411	Sugarcane	4.47 ± 4.87 abc
CO-5	Corn	4.54 ± 4.07 c
NE744	Corn	3.91 ± 2.98 bc
SAM119	Corn	6.18 ± 4.82 ab
X. v.* pv. *holcicola		
Mex-1	Sorghum	3.32 ± 2.63 bc
NCPPB2417	Sorghum	6.01 ± 4.57 ab
NCPPB1241	Sorghum	7.65 ± 5.45 a
SAS211	Sorghum	6.34 ± 5.40 ab
X. campestris* pv. *musacearum		
NCPPB4381	Banana	0 ± 0 d

[a]Original host bacterial strain was isolated from

[b]Leaves were infiltrated with 10^8 CFU/ml of bacteria suspended in sterile water in two spots per leaf, on either side of the main vein.

[c]Mean lesion length ± standard deviation. Lesion lengths averaged over two independent replications. Means followed by the same letter are not significantly different according to Tukey's honestly significant differences test ($P < 0.05$).

APPENDIX B

OUTREACH

B.1 INTRODUCTION

Scientific communication is integral to advancing research and impact. Outreach components have become a regular accessory to federal and private funding opportunities. As a graduate student, I had the opportunity to organize, train and collaborate with international partners. Locally, I created partnerships within our school district to educate kindergarteners up to high school students from directly entering classrooms to hosting large summer school science, technology and math (STEM) programs on campus. In one case, I worked with two local high school science teachers to develop a proposal: 'Adding to the Power of Your Laptop: Using Probeware Technology to Enhance the Life Sciences,' that earned $25,000 to purchase upgraded scientific equipment and software for their classrooms.

Mentorship of undergraduate and graduate students became a central part of my time as a graduate student. I have and continue to advise several students at all levels of engagement in our lab and greenhouse spaces in addition to two international MS interns. Incorporating these diverse scopes of engaged communication and teaching was incredibly rewarding and has taught me a diverse set of skills in time management, interpersonal relationships, conflict resolution, counseling and project management.

B.2 INTERNATIONAL WORKSHOPS

B.2.1 'Harmonizing Detection of Xanthomonas oryzae Pathovars,' International Rice Research Institute (IRRI), Los Baños, Philippines 2012

The vision of this workshop was to engage international partners in the development of a universal approach to accurately detecting *X. o.* pv. *oryzae* and *X. o.* pv. *oryzicola.* At the time of this workshop, advances had been made in genome-based molecular diagnostic tools. However, these had not been widely accepted internationally, in part due to disparities in equipment, supplies and facilities for molecular biology. We also sought to educate international partners from industry, academia and

governments in how molecular diagnostic assays could be designed, developed, and then validated for use with other pathogens. This workshop integrated support from industry partners Bayer Crop Science (https://www.cropscience.bayer.com/en) and OptiGene (http://www.optigene.co.uk/). OptiGene donated reagents and equipment for LAMP training while Bayer Crop Science provided funding for participant travel and accommodations to reflect their investment in seed pathology and advanced diagnostics.

Several integral collaborations developed from this workshop. With input from participants and continued optimization of existing protocols, I collaborated with other workshop trainers to develop a harmonized detection strategy for *X. oryzae* in rice seeds; this was distributed to workshop participants. It was later expanded and incorporated into a chapter of a peer-reviewed manual published by the American Pathological Society (1). Furthermore, Dr. Gavin Ash (University Southern Queensland, Australia), an attendee with similar interests to our research group, arranged to join our research group on a short-term sabbatical in 2014. We were able to use our combined knowledge and resources to analyze draft genome sequence for the rice brown sheath rot pathogen, *Pseudomonas fuscovaginae*, then design and screen molecular diagnostic assays. This eight-week research collaboration resulted in a peer-reviewed publication, of which I am a co-author (2).

B.2.2 'Molecular Diagnostics Workshop', Biosciences in eastern and central Africa-International Livestock Research Institute (BecA-ILRI), Nairobi, Kenya 2014

This capacity building activity was an outreach component of a proposal titled 'Adapting existing *Xanthomonas oryzae* diagnostics from genomic information' that was funded by both BecA and Sweden; I was a coPI on this proposal, along with Drs. Jan Leach (CSU), Valerie Verdier (IRD), Timothy Holton (ILRI) and Appolinaire Djeking (ILRI). Workshop attendees traveled from Uganda, Tanzania, Burkina Faso, Burundi, Ethiopia and within Kenya to learn basic principles of molecular diagnostics and gain hands on experience in performing sensitive detection assays with real agricultural samples such as cassava and sheep blood infected with viruses or bacteria. They applied their knowledge in hands on lab exercises in extracting DNA and RNA from these samples, preparing assays and analyzing results. Attendees of this workshop were employed in both animal and plant regulatory agencies, universities,

extension offices and industry. This collaboration also supported a research component that funded the publication of chapter 2 of this dissertation (3). I particularly enjoyed this workshop because I was able to work with people from countries that I had never traveled to, nor had prior interaction. It was warm and fruitful and expanded my own knowledge on international agriculture.

B.2.3 'Resistance, Genomics and Molecular Diagnostics of Bacterial Plant Pathogens' (IRRI), Los Baños, Philippines, August 2016

To build on our prior training exercises, in October 2016, I co-coordinated a workshop on molecular diagnostic assay design. We recruited 16 local Philippine Agricultural Extension Workers (AEWs) from Region XII in Mindanao. Most were recipients of (Agri-Pinoy) Rice Achievers' Award of a Philippine program in agriculture and as AEWs, they require training on symptomatology, disease diagnosis and pathogen detection. The remaining diverse set of participants represented academia, industry and federal governments from Burundi, Cambodia, Myanmar, Pakistan and Thailand. Resource persons (including IRRI staff and visiting scientists from IRD, Montpellier, France and Cornell University, USA) provided lectures, hands-on lab experience as well as field diagnosis experience for participants.

B.3 PRIMARY AND SECONDARY EDUCATION

On a local level, I've truly enjoyed extending learning opportunities into our community's Poudre School District. Throughout my PhD I have engaged with several elementary schools, connecting with individual class rooms to share with young children basic principles of plant biology and agriculture. Most importantly, I have tailored lessons for children about where food comes from through hands-on materials, such as live rice plants and seed for them to have a tactile, connective experience. I have taught in high school biotechnology classes on innovative advances in genome editing and applications to agriculture with exercises that engage students in defining genetically modified organisms and most recently, having them design their own CRISPR-Cas9 constructs. I introduced them to accessing genome sequence and defined how this impressive amount of information is publicly available, even to young adults.

For the past three summers I have facilitated local high school interns to join our research group and gain experience in advanced research. One of these students asked to return for a second summer because he enjoyed the experience and wanted to gain more hours with our group. Both interns worked over 50 hours to obtain high school credit under my co-supervision and their achievement was significant. They were registered through a Science, Technology, Engineering and Mathematics (STEM) internship program in cooperation with CSU.

Annually, I host over 50 middle school students enrolled in STEM summer school programs from both local and regional Colorado schools. During their daily field trips, we tour our lab and greenhouse facilities and I engage them in research fields of agriculture, plant biotechnology and pathology. It's thrilling to show them sophisticated equipment and crops they've eaten, yet never seen or touched in person and guide them to apply principles and theories they're learning in the classroom to real world applications that directly impact their own lives.

B.4 MENTORSHIP

Perhaps the most rewarding form of education and outreach for me has been individual mentorship. The opportunity to supervise and advise undergraduate and graduate research students throughout my PhD work has been unparalleled. Research in Dr. Leach's lab is open, dynamic, diverse and challenging. We have employed so many excellent individuals. I personally have overseen five independent credit-earning research projects by undergraduate students encompassing plant physiology, hydroponic plant production, bioenergy, molecular diagnostics, emerging plant diseases and bacterial genomics. These students have gone on to graduate school, prominent internships, government employment and even law school. One-on-one instruction allows for trusting relationships, challenging both our skills and knowledge and most importantly, rewarding excellence in research. In Chapter 2 of this dissertation, the second author was an undergraduate student, Paul Langlois, who I advised through this project and he went to on to do his own MS degree in molecular diagnostics and bacterial genomics, which was recently published (4). Similarly, in Chapter 4, second co-author Elysa DuCharme was

instrumental in completing this project and I was fortunate to advise her as an undergraduate then as a research associate in our lab after she graduated.

B.5 BIOSAFETY AND BIOSECURITY TRAINING COURSE

Starting in 2011, I have presented annually in the Biosafety and Biosecurity Training Course coordinated by Dr. Robert Ellis, on agricultural biosecurity using select agents as a model for compliance. This course provides intensive, participatory training on most aspects of biosafety and biosecurity including exercises, tours, audit and inspection discussions relevant to both animal and plant based research (http://www.bbtcfortcollins.com/). International participants include regulatory officials from academia, government and industry. I have used our lab's experience in adapting to a BSL2 and BSL2P+ regulatory status to work with *Xanthomonas oryzae* as select agents and how we maintain compliance. I also describe applications of molecular diagnostics in plant pathology in the scope of agricultural biosecurity. Over the years, I transitioned to making this a more participatory and discussion based training by encouraging questions and gaining input in how our experience can help and guide their work at their own institutions.

REFERENCES

1. **Vera Cruz CM**, **Cottyn B**, **Nguyen MH**, **Lang J**, **Verdier V**, **Mew TW**, **Leach JE**. 2017. Detection of plant-pathogenic bacteria in seed and other planting material. 2nd Edition. American Phytopathological Society. Minneapolis, MN, USA.
2. **Ash GJ**, **Lang JM**, **Triplett LR**, **Stodart BJ**, **Verdier V**, **Cruz CV**, **Rott P**, **Leach JE**. 2014. Development of a genomics-based LAMP (Loop-Mediated Isothermal Amplification) assay for detection of *Pseudomonas fuscovaginae* from rice. Plant Dis **98**:909–915.
3. **Lang JM**, **Langlois P**, **Nguyen MHR**, **Triplett LR**, **Purdie L**, **Holton TA**, **Djikeng A**, **Vera Cruz CM**, **Verdier V**, **Leach JE**. 2014. Sensitive detection of *Xanthomonas oryzae* pathovars oryzae and oryzicola by loop-mediated isothermal amplification. Appl Environ Microbiol **80**:4519–30.
4. **Langlois PA**, **Snelling J**, **Hamilton JP**, **Bragard C**, **Koebnik R**, **Verdier V**, **Triplett LR**, **Blom J**, **Tisserat NA**, **Leach JE**. 2017. Characterization of the *Xanthomonas translucens* Complex Using Draft Genomes, Comparative Genomics, Phylogenetic Analysis, and Diagnostic LAMP Assays. Phytopathology **107**:519–527.