University of Southern Queensland

HOST-INDUCED ADAPTATION OF *Xanthomonas oryzae* pv. *oryzae*

A Thesis submitted by

Marian Hanna R. Nguyen, M. Sc.

For the award of

Doctor of Philosophy

November 2021

ABSTRACT

This dissertation reports research on the adaptation mechanisms of the Bacterial Blight (BB) pathogen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) in response to its growth in rice, as well as developing a tool for monitoring this pathogen intended to facilitate the specific deployment of resistance genes for breeding. BB is one of the major diseases of rice. The negative impact of virulent forms of this pathogen on rice production has been widely reported. It has a range of adaptations which contribute to its virulence. One of the less studied aspects of the rice-*Xoo* interaction is the biofilm forming capability of the pathogen in the rice xylem. In this research, the genetic diversity of biofilm-related proteins in *Xoo* was investigated. The results show both conservation of majority of the genes, but variability in a smaller subset. This variation could be part of the adaption of the pathogen to new rice varieties or geographic locations. To determine the incremental steps in *Xoo* adaptation to rice, a serial passage experiment was conducted using the Philippine *Xoo* strain PXO61, and subsequent experiments were performed comparing the wild-type strain (designated in this study as PXO61-0) and its derivative, PXO61-4. Characterization of both PXO61-0 and PXO61-4 revealed that PXO61-4 exhibited an increase in fitness during *in vitro* and *in planta* growth. A SNP was detected by genomic comparison of both variants, which was confirmed in RNA-sequencing reads. Finally, a prototype of a single nucleotide polymorphism (SNP)-based monitoring system was developed and tested on simulated populations of *Xoo*. The detection pipeline successfully detected target populations in samples prepared or collected from artificial and field conditions.

CERTIFICATION OF THESIS

I, **Marian Hanna R. Nguyen**, declare that the PhD Thesis entitled "Host-Induced adaptation of *Xanthomonas oryzae* pv. *oryzae"* is not more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes. The thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signed: Date:

Endorsed by:

Professor **Gavin Ash**
Principal Supervisor

Dr. Dante Adorada
Associate Supervisor

Dr. Ricardo Oliva
Associate Supervisor

Student and supervisors' signatures of endorsement are held at the University.

ACKNOWLEDGEMENT

I would like to acknowledge the generous support from the University of Southern Queensland International Fees Scholarship. Thank you very much for this huge opportunity. I also wish to acknowledge the support from the office of the Graduate Research School, USQ. Sincere thanks to the International Rice Research Institute (IRRI), where the experiments described in this thesis were conducted.

To my mentors at the University of Southern Queensland (USQ), Prof. Dr. Gavin Ash and Dr. Dante Adorada, thank you very much for allowing me to pursue this program under your supervision. I am very grateful for all your mentorship and guidance. With your expert guidance and flexibility in communication media, the remote learning program was not as difficult as I have thought it would be. Thus, I have survived up to this point and I have learned so much. To my mentors at IRRI, Dr. Casiana "Nollie" Vera Cruz and Dr. Ricardo Oliva, I am very grateful for all the teachings and opportunities that you have provided to me and to other young researchers as well. Thank you very for instilling an encouraging environment at Host Plant Resistance (HPR), where new ideas are welcomed and reinforced.

I would like to express my sincerest thanks to present and past colleagues of HPR. To Epifania "Fanny" Garcia, Crispulo "Boyet" Cura, Elenita "Ellen" Silab, and Ruby Burgos, who generously imparted techniques gained through many years of work experience in studying rice diseases in the laboratory and field, techniques are not found in conventional manuals or procedurals. Special thanks to Tita Fanny, who was always eager to share the details of the techniques and observations gained through more than 30 years of working on *Xoo*, the pathogen focused on in this research. To Jonas Padilla, thank you very much for all the rich discussions and for being a sounding board to my experimental plans throughout the years. Thank you for those times that you too stayed late in the laboratory to help me with work that I have underestimated. Thanks to Ismael Mamiit, Maritess Carillaga, Ronald Javier, and Carlo Figuerra for the technical assistance in all the greenhouse experiments. Thanks to Dr. Ana Cope who encouraged me to apply for this PhD program at USQ. To Genelou Atienza-Grande, Eula Oreiro, Thea Coronejo, Pauline Capistrano, Paolo Balahadia, Dr. Gilda Jonson, Dr. Jeanie Yanoria, thank you very much for all the technical assistance and professional advice. To Ian Quibod, thank you very much for

all your patience in helping me out with my questions, for all the brilliant discussions, and for all the words of encouragement.

Sincere gratitude to all the scientists and researchers who have contributed to the experiments described in this thesis: (Chapter 4: Exploring the adaptation potential of *Xanthomonas oryzae* pv. *oryzae* in rice through serial passage in bacterial blight resistant rice line IRBB4) Jonas Padilla, Ian Quibod, Dr. Sara Carpenter, Dr. Veronica Roman-Reyna, Dr. Sylvestre Dossa, Dr. Casiana Vera Cruz, Dr. Ramil Mauleon, (Chapter 5: Monitoring *Xanthomonas oryzae* pv. *oryzae* populations with single nucleotide polymorphism-based primers) Genelou Grande, Thea Coronejo, Dr. Nafisah Nafisah, Dr. Celvia Rosa, Dr. Joko Prasetiyono, Dr. Fatimah Fatimah, Joanne Caguiat, Frodie Waing, Jelyn Manangkil, Dr. Si Si Myint, Dr. Sumera Yasmin, Dr. Muhammad Arif, Dr. Muhammad Asif, Dr. Nguyen Duy Phuong, Ha Anh Tran, Nguyen Thi Xuan Mai, Dr. Hafiz Imran Arshad, Dr. Gouri Sankar Laha, Dr. Raman Meenakashi Sundaram, Dr. Socheath Ong, Dr. Wichai Kositrana, and Dr. Sujin Patarapuwadol.

Finally, I wish to thank my friends and family. To Dannah Torres and Joh Gomez, sincere thanks for all the encouragement. Thanks to my brother William, who readily lent me his computer which I used for some of the data analysis. Thank you very much to my sisters, Faye, and Nova, for the rich discussions on statistical analysis, among many other topics. Thank you to my brother Gerald, for the sacrifices you have made to keep everyone safe. Special thanks to my dear nephew, Ali, who brings joy to the entire family.

I wish to thank my parents, Miriam, and Hoanh, for everything. I dedicate this thesis to you. Thank you, Papa, for bringing Science to our home – from magnets to soil kits and for teaching me how to use a computer during the time of MS-DOS. The command line tools I used were less scary because of this. I will always remember the engineer's field notebook that you gave me when I was a kid, where I pretended to be a scientist while writing on it. Coincidentally, the research notebooks at IRRI also have that grid printed on its pages. Thank you, Ma. You are my inspiration. The image of you managing to finish your own dissertation with four young children and a full-time job has pushed me to just keep moving forward. Most importantly, thank you both for all your sacrifices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ALE	Adaptive laboratory evolution
AUDPC	Area under disease progress curve
BAPS	Bayesian analysis of population structure
BB	Bacterial blight
BLAST	Basic Local Alignment Search Tool
CFA	Cyclopropane fatty acids
DNA	Deoxyribonucleic acid
EBE	Effector binding site
EPS	Extracellular polymeric substances
ETI	Effector triggered immunity
FAOSTAT	A database for food and agriculture data provided by the Food and Agriculture Organization of the United Nations. It can be accessed online through this link: https://www.fao.org/faostat/en/#home
IRRI	International Rice Research Institute
KASP	Kompetitive allele specific PCR, a genotyping assay for detection of bi-allelic single nucleotide polymorphisms. A description of this assay is available through this link: https://www.biosearchtech.com/support/education/kasp-genotyping-reagents/how-does-kasp-work.
LPS	Lipopolysaccharides
MDS	Multi-dimensional scaling
ML	Maximum likelihood
MSA	Multiple sequence alignment
NIL	Near-isogenic line

NJ	Neighbor-Joining
NPPO	National Plant Protection Oranization
PathoTracer	A decision support system to reduce the risk of rice diseases. It can be accessed online through this link: https://sites.google.com/irri.org/pathotracer.
PAMP	Pathogen-associated molecular patterns
popID	putative population
PTI	PAMP-triggered immunity
RFLP	Restriction fragment length polymorphism
R gene	Resistance gene
SNP	Single nucleotide polymorphism
SMRT	Single Molecule, Real-Time
SWEET	Sugars will eventually be exported transporter
T2SS	Type II secretion system
T3SS	Type III secretion system
T6SS	Type VI secretion system
Xoo	*Xanthomonas oryzae* pv. *oryzae*

CHAPTER 1: INTRODUCTION

1.1 Background of the study

Rice is an important agricultural crop, with an average production value of about 7.0 Mt from 1994 to 2019 (Figure 1.1A). It is a staple food in Asia, where nine out of the top ten rice-producing countries are from Asia (Figure 1.1B). Rice (*Oryza* L.) has two cultivated species, *O. sativa* L. and *O. glaberrima* Steud (Khush, 2000). Improvements in agricultural practices during the Green Revolution has been attributed to the steady increase of rice yield since the 1960s (Pinstrup-Andersen & Hazell, 1985) and initiated by the introduction of semi-dwarf high yielding Indica cultivars (Hedden, 2003).

Rice production is constantly threatened by pests and pathogens. One of the major pathogens of rice, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is the causal organism of bacterial blight (BB), one of the most serious threats to rice production in Asia (Swings et al., 1990), and one of the top ten most important pathogenic bacteria in crops (Mansfield et al., 2012). Since the early 1970s, the International Rice Research Institute (IRRI) has been curating strains of *Xoo* isolated from leaves showing symptoms of BB, collected throughout the Philippines. This unique resource provided the basis for the research undertaken in this dissertation.

1.2 Statement of the problem

The genus *Xanthomonas* is known to cause diseases in a wide range of plant hosts. In fact, three out of the ten most important plant pathogenic bacteria are *Xanthomonas* spp with *Xoo* ranked fourth (Mansfield et al., 2012). *Xanthomonas oryzae* pv. *oryzae* was selected as the organism of focus in this study due to the intriguing diversity of pathogenicity, physiology, and genomic characteristics that it possesses. Despite it being a model plant pathogen (Niño-Liu, Ronald, & Bogdanove, 2006), the understanding of its virulence mechanisms is incomplete (White & Yang, 2009).

Numerous studies have demonstrated that this bacterium is efficient in responding to and adapting to its host (Gonzalez et al., 2007; Khaeruni & Wijayanto, 2013; Mew, Vera Cruz, & Medalla, 1992; Midha et al., 2017; Noda et al., 2001; Quibod et al., 2016), with the maintenance of variability as a key mechanism for its adaptation. Pathogenic profiling of *Xoo* strains collected by IRRI scientists from

various locations in the Philippines from 1972 to 2015 show a shift in the pathogen populations, from strains that cannot infect rice carrying the BB resistance gene Xa4, to those that can infect hosts carrying that gene (Quibod et al., 2020).

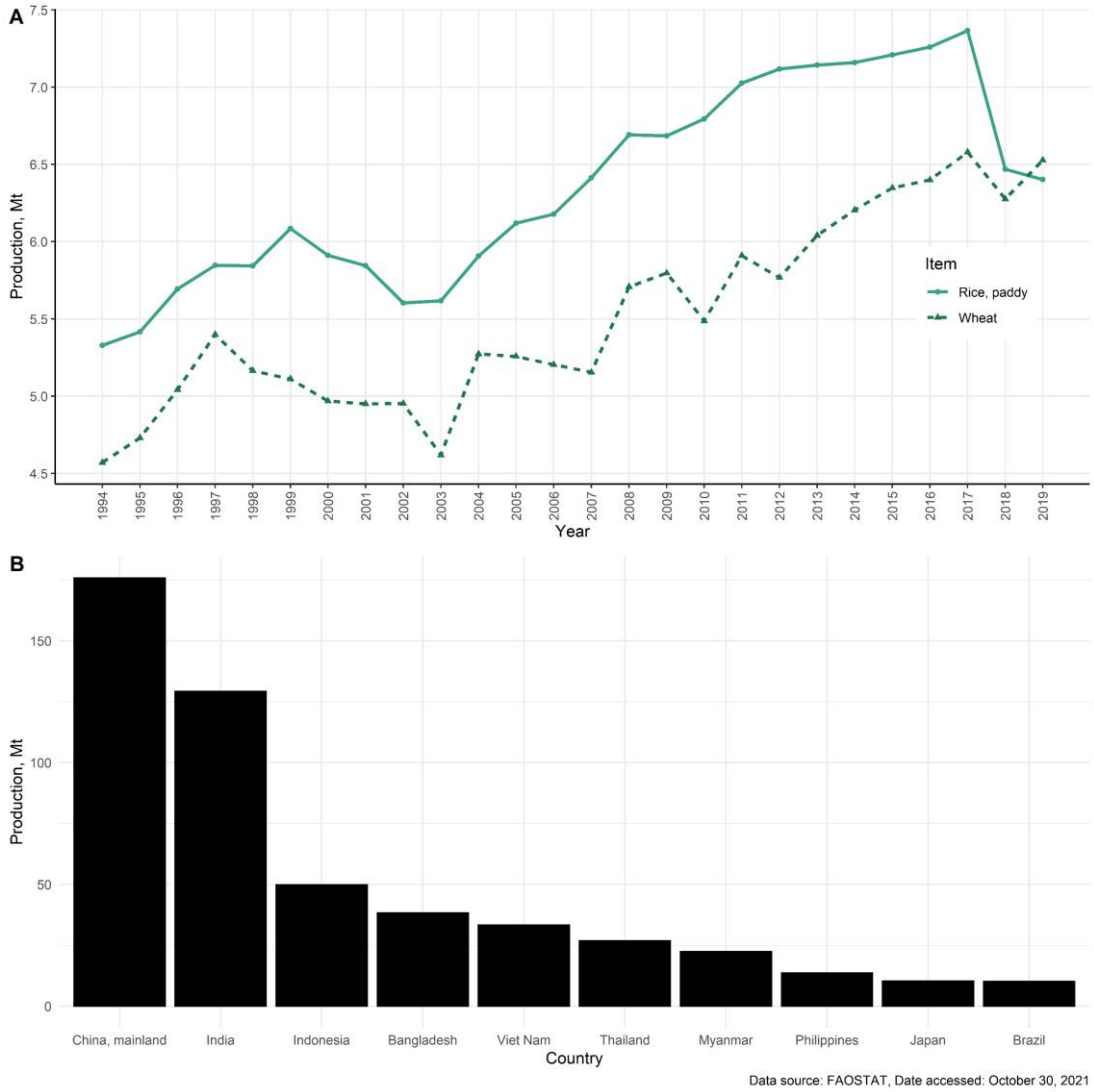


Figure 1.1. Status of global rice production. A. The annual mean production value of rice in million metric tonnes from 1994 to 2019. Production value of wheat is added for additional reference. B. The top ten countries that produce rice. Data was downloaded from the FAOSTAT online database (FAOSTAT, 2021).

It is common knowledge that bacteria exist as sessile (attached) or planktonic (free-living) entities. The ability to form biofilms plays an essential role in the survival of bacteria in different environments, this includes plant pathogens and their survival on the host surface and within host tissues (Donlan & Costerton, 2002; Muhammad et al., 2020). Its potential for a means to manage BB has been recognized, and there have

been reports of niclosamide (Sahu, Zheng, & Yao, 2018) and thyme oil (Singh, Gupta, Tandon, & Pandey, 2017) as an effective means to inhibit biofilm-formation of Xoo which has been linked to reduced virulence in rice.

However, the incremental changes that lead to population shifts in response to its host environment have not been studied. Understanding the molecular mechanisms of plant-microbe coevolution would initiate the development of new measures for disease control (Dodds, 2010). Moreover, a highly accessible, universal, single nucleotide polymorphism (SNP)-based system of monitoring populations of this pathogen at the genetic level has not yet been established. The collection of experiments in this project aims to explore the adaptation of *Xoo* in artificial and natural systems, as well as developing a SNP-based *Xoo* monitoring and surveillance system.

1.3 Research questions

This dissertation focuses on the following research questions:

1. Several studies have described the genetic variability of *Xoo*. Is there any evidence of this variability in the biofilm-related genes of *Xoo*?
2. Can *Xoo* adaptation be induced by repeated cycles of inoculation on and re-isolation of the pathogen from rice? If so, what are the incremental changes in *Xoo* populations in response to growth in rice, that cause a shift in virulence?
3. Can SNP-based markers be used to monitor population shifts of *Xoo,* and serve as a useful tool in disease surveillance?

1.4 Research aims and objectives

The research reported in this dissertation aims to investigate the adaptation mechanisms of *Xoo* in response to its growth on rice, and to develop a monitoring tool for this pathogen that would be useful in the strategic deployment of specific resistance genes to areas based on the pathogen population data. The objectives are grouped into the following experimental chapters:

1. Investigation of the biofilm-related genes of Xanthomonads in *Xoo* using a bioinformatics approach.

2. Characterization of *Xoo* after serial passage in rice using a combination of disease phenotyping, metabolic profiling, transcriptome analysis, and genomic analysis.
3. Development of a monitoring tool for *Xoo* populations using SNP-derived markers.

1.5 Significance of the study

The relatively short life cycles of microbes in comparison to their plant hosts confers an inherent advantage in gaining counter-adaptations. *Xoo* and other plant pathogens are well-adapted to their hosts and are continuing to adapt as the hosts develop or gain counter-adaptations to these pathogens. This is classically known as the host-pathogen arms race. *Xoo* populations have been shown to shift in response to deployment of rice carrying genes for resistance to BB. Determining the mechanisms of adaptations that lead to these shifts in populations could help scientists further understand the disease etiology, predict BB outbreaks, and reduce yield loss by controlling the disease.

CHAPTER 2: LITERATURE REVIEW

2.1 The impact of bacterial blight on rice production

Bacterial blight is a vascular disease of rice, which presents as straw white lesions on the leaves of infected rice plants (Figure 2.1), or in more severe cases of infection, it causes wilting or *kresek* (T. W. Mew, A. M. Alvarez, J. Leach, & J. Swings, 1993), causing yield losses of up to 50% (Ou, 1985). Quantification of yield loss in controlled conditions by inoculation of *Xoo* into a susceptible variety found up to 24.5% loss in grains per panicle, and up to a 36.8% decrease in 1,000-grain weight (Adhikari, Shrestha, Basnyat, & Mew, 1999). This disease is commonly observed in areas within Asia and Africa. A recent survey in Pakistan reported BB incidence values ranging from 37% up to 80% in provinces where BB was observed (Ahsan et al., 2021). Khan et al (Khan, Rafi, Abbas, Ali, & Hassan, 2015) surveyed distinct areas in Pakistan and reported similar values, with disease incidence reaching up to 75%, and yield loss (1,000-grain weight) of up to 15.59%. In Senegal, BB incidence was reported to be ranging from 30 to 50% (Tall et al., 2020). BB has also been reported in Bangladesh (Rashid et al., 2021), India (Lore et al., 2011), Korea (Jeung, Heu, Shin, Vera Cruz, & Jena, 2006), Malaysia (Saad & Habibuddin, 2010), and the Philippines (Mew et al., 1992).

Because of its potential threat to agriculture, *Xoo* is recognized as a quarantine pest. It is considered as a select agent by the US Department of Agriculture (USDA, https://www.selectagents.gov/sat/list.htm) and as a regulated pathogen by the European and Mediterranean Plant Protection Organization (EPPO, https://www.eppo.int/ACTIVITIES/plant_quarantine/A1_list). It also ranks fourth in the in the list of bacterial plant pathogens with scientific and economic importance (Mansfield et al., 2012).

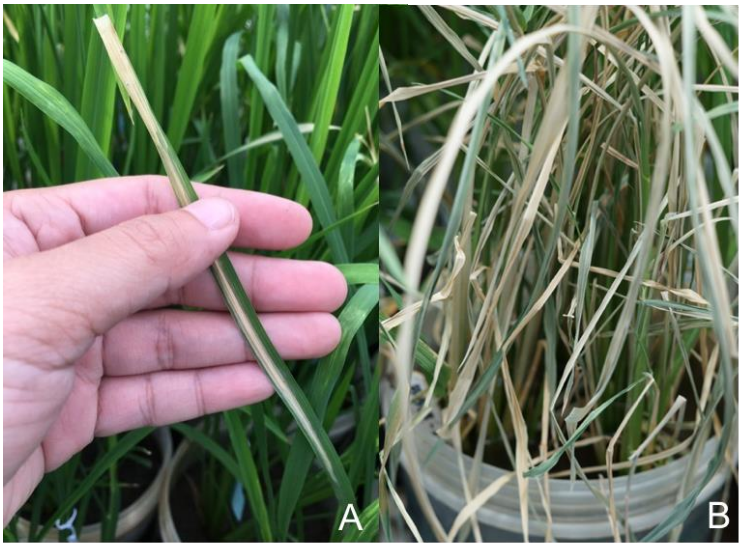


Figure 2.1. Typical symptoms of bacterial blight observed on inoculated rice plants. A. Lesions on artificially inoculated leaves of a susceptible variety 14 days post-inoculation. B. Wilt observed during severe infection by *Xanthomonas oryzae* pv. *oryzae*.

2.2 *Xanthomonas oryzae* pv. *oryzae, causal organism of Bacterial blight in rice*

The discovery of *Xoo* as the causal organism of bacterial blight is summarized in the review paper by Niño-Liu and co-workers (Niño-Liu et al., 2006), where it is described that early studies on BB by Takaishi in 1909 allowed the discovery of bacteria present in dew drops from infected leaves, and disproved the initial hypothesis that BB was a physiological disorder of rice caused by acidified soils (Ou, 1985). After numerous instances of renaming, the bacterium is now known as *Xanthomonas oryzae* pv. *oryzae* (Swings et al., 1990).

Xanthomonas oryzae pv. *oryzae* is a gram-negative member of the Gammaproteobacteria, order Xanthomonadales (Saddler & Bradbury, 2005). The order Xanthomonadales comprises *Xanthomonas* and *Xylella* species and is reviewed extensively by H. S. Naushad and R. S. Gupta (Naushad & Gupta, 2013), both economically important plant pathogens (Mansfield et al., 2012).

There are numerous sources of *Xoo* inoculum in nature, which includes: infected seeds, irrigation water, weeds, contaminated materials from the previous season, and farmers' tools for planting (Ou, 1985). This bacterium colonizes the rice xylem (Hilaire et al., 2001; Noda & Kaku, 1999) by gaining entry through natural

openings or wounds of the plant, where the interaction of the host, and environment, allows the pathogen to infect the host and produce observable lesions on leaves of susceptible rice plants.

2.3 Plant-pathogen interaction and host-driven adaptation of *Xoo* populations

The nature of the interaction between plant hosts and pathogens is reliant on how the plant perceives or detects the invading pathogen, and the subsequent defense response rendered by the plant through innate immune responses. These immune responses are now known as pathogen-associated molecular pattern- (PAMP) triggered immunity and effector triggered immunity (ETI), previously known as horizontal disease resistance and vertical disease resistance (Boller & Felix, 2009).

Within the context of microbial invasion of plant tissues, PTI is dependent on the recognition of PAMPs (or microbe-associated molecular patterns, MAMPs) by cell surface-localized receptors known as pattern recognition receptors (PRR) (Macho & Zipfel, 2014) which belong to families of receptor-like kinases (RLK) and receptor-like proteins (RLP) (Jones & Dangle, 2006). One example of PAMP is chitin (Akamatsu et al., 2013), which is implicated in rice immunity against the blast fungus, *Magnaporthe oryzae*. Other well-known examples of PAMPs are flagellin and elongation factor EF-Tu (Boller & Felix, 2009). It is notable that the *Xoo* flagellin is not detected by the known receptor kinase cognate in rice, suggesting that *Xoo* has evolved to escape detection via this pathway thereby avoiding mitogen-activated protein kinase (MAPK)-mediated or reactive oxygen species (ROS) production which are involved in defense responses against invading pathogens (S. Wang et al., 2015).

While the PTI is known to detect conserved PAMPs, the ETI response is dependent on the specific recognition of pathogen avirulence (Avr) genes by the host resistance (R) genes, which confers resistance to the disease (Flor, 1971) by triggering rapid cell death at the point of pathogen invasion or the hypersensitive response (Balint-Kurti, 2019). This specificity between host and pathogen genes has been demonstrated in classical investigations of the *Xoo*-rice interaction where a set of rice lines that contain single BB resistance genes were used to group *Xoo* strains into races (Nelson et al., 1994), or, strains that share a unique virulence pattern when inoculated into hosts with specific resistance genes (Caten, 1987). The populations of *Xoo* in the Philippines is composed of ten major races (Table A.1) (Ardales et al., 1996; Mew &

Vera Cruz, 1979; Mew et al., 1992; Nelson et al., 1994; Vera Cruz et al., 2000) represented by 14 strains (Table A.2).

The Avr gene products have been recently categorized under the term effectors, which are proteins and other small molecules secreted by plant pathogens to trigger changes in the host. This change in nomenclature addresses the dual functionality of effectors not only as triggers of resistance but also dictating susceptibility of the host (Hogenhout, Van der Hoorn, Terauchi, & Kamoun, 2009). For instance, the R genes xa13, xa25, and xa41 which encode SWEET (Sugar Will Eventually be Exported Transporter) proteins have corresponding dominant alleles that can be induced by specific *Xoo* strains (Jiang et al., 2020). Strains that produce the secreted effector protein PthXo1 induce susceptibility of rice carrying the dominant Xa13, also known as *OsSWEET11* (L.-Q. Chen et al., 2010). Similarly, strains that secrete the effectors PthXo3, AvrXa7 (Antony et al., 2010), Tal5, and TalC (Streubel et al., 2013) enhance the susceptibility of rice carrying the dominant allele of Xa41, also known as OsSWEET14.

Table A.1. Pathogenic specialization of Philippine *Xanthomonas oryzae* pv. *oryzae* races derived from strains inoculated into rice lines introgressed with single bacterial blight resistance genes using susceptible variety IR24 as background.[α]

Near-isogenic lines	*Xa*-gene	Race 1	Race 2	Race 3b	Race 3c	Race 4	Race 5	Race 6	Race 7	Race 8	Race 9a	Race 9b	Race 9c	Race 9d	Race 10
IR24		S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB4	Xa4	R	S	S	S	MR-MS	R	S	R	R	S	S	S	S	R
IRBB5	xa5	R	R	R	R	S	R	S	R	R	R	R	R	R	R
IRBB7	Xa7	MS	R	R	R	S	R	S	R	R	S	MR-MS	MS-S	R	R
IRBB10	Xa10	S	R	S	S	S	R	S	R	S	S	S	S	S	S
IRBB13	xa13	S	S	S	S	S	S	R	S	S	S	S	S	S	S
IRBB14	Xa14	S	S	S	S	S	R	S	S	R	S	S	S	S	S
IRBB21	Xa21	R	R	R	MR	R	R	MR	MR	MR	MR	MR	MR	MR	S

[α] Resistance or susceptibility of rice plants to *Xoo* is expressed in lesion lengths measured 14 days after inoculation of *Xoo* strains into rice leaves 40-45 days after sowing. Resistant (R): <5 cm; moderately resistant (MR): 5-10 cm; moderately susceptible (MS): 10-15 cm; susceptible (S): >15 cm

Table A.2. Strains of *Xanthomonas oryzae* pv. *oryzae* representative of race groupings in the Philippines.

	Strain	Race	Collection site
1	PXO61	1	IRRI experimental plot
2	PXO86	2	IRRI experimental plot
3	PXO79	3b	Davao
4	PXO340	3c	Calauan, Laguna
5	PXO71	4	Palawan
6	PXO112	5	Banaue
7	PXO99	6	IRRI experimental plot
8	PXO145	7	Bontoc
9	PXO280	8	Namulditan, Ifugao
10	PXO339	9a	Calauan, Laguna
11	PXO349	9b	Calauan, Laguna
12	PXO347	9c	Mabitac, Laguna
13	PXO363	9d	Calauan, Laguna
14	PXO341	10	Calauan, Laguna

Host-driven adaptation of *Xoo* populations has been extensively reported in the Philippines (Mew et al., 1992; Nelson et al., 1994; Quibod et al., 2016; Vera Cruz et al., 2000) and in India (Midha et al., 2017). Pathotype profiling of 1822 Xoo strains in the Philippines collected between 1970 and 2015 show that the deployment of the R gene Xa4 has favoured populations which are able to overcome the gene (Quibod et al., 2020). Similarly, a shift in *Xoo* populations favouring more virulent pathotypes have been reported in Indonesia (Khaeruni & Wijayanto, 2013); in China, this pathogen has been reported to migrate from the northern to southern part of Yunnan through irrigation, acquiring new pathogenic characteristics in the process (Noda et al., 2001). Gonzalez et al. (Gonzalez et al., 2007) reported that West African strains were only recently exposed to selective pressure due to recent rice cultivation practices in that region.

2.4 Management of bacterial blight

The search for novel sources of BB resistance genes and subsequent breeding efforts to produce resistant varieties is the most successful method to manage BB (Niño-Liu et al., 2006). This began with the development of near-isogenic rice lines (NILs) at IRRI as a necessary prerequisite in the establishment of breeding for resistance to BB (Ogawa, Yamamoto, Khush, & Mew, 1991). Each of these lines were bred to contain single *Xa* (BB resistance) genes from previously identified resistant varieties and were designated with the prefix IRBB. The most frequently used IRBB NILs in experiments related to *Xoo* are lines with *Xa* genes in the background of the susceptible variety IR24. A list of these lines can be found in Table A.1. To date, about 40 genes for resistance to BB have been identified, with the most recently discovered gene designated as *Xa43*(t) (S.-M. Kim & Reinke, 2019).

Natural variations in the EBEs of the SWEET genes *OsSWEET13 (xa25)* and *OsSWEET14* (xa41) have been recently discovered (Zaka et al., 2018), which led to the development of BB-resistant rice lines with genome-edited EBEs (Eom et al., 2019), a modern approach to developing resistant materials via CRISPR/Cas9 editing.

Monitoring pathogen populations is an important component of disease management. The effective deployment of resistant varieties strongly depends on pre-existing knowledge of the predominant and emerging pathogen populations (Dossa, Sparks, Vera Cruz, & Oliva, 2015). When the composition of the pathogen population in a targeted area is known, the specific R gene(s) suitable for deployment in that area can be selected with more certainty. However, this requires knowledge of the genetic pool and virulence factors associated with the populations (Quibod et al., 2016).

2.5 General adaptation mechanisms of bacterial plant pathogens

Bacteria constantly adapt to changing environmental conditions, because of selection pressure (M. Kang et al., 2019), a concept in biology known as evolutionary adaptation (Waddington, 1959). Environmental factors such as temperature (Beales, 2004), pH (Fernandez, Mercader, Planas-Fèlix, & Torrents, 2014), and extreme environmental conditions (S.-J. Li et al., 2014) are known to affect bacterial adaptation and evolution. Numerous studies have been conducted that report on host-induced adaptation of plant pathogens and plant-associated bacteria (Darling, Miklós, &

Ragan, 2008; C.-L. Huang et al., 2015; R. A. Melnyk, Hossain, & Haney, 2019; Quibod et al., 2016).

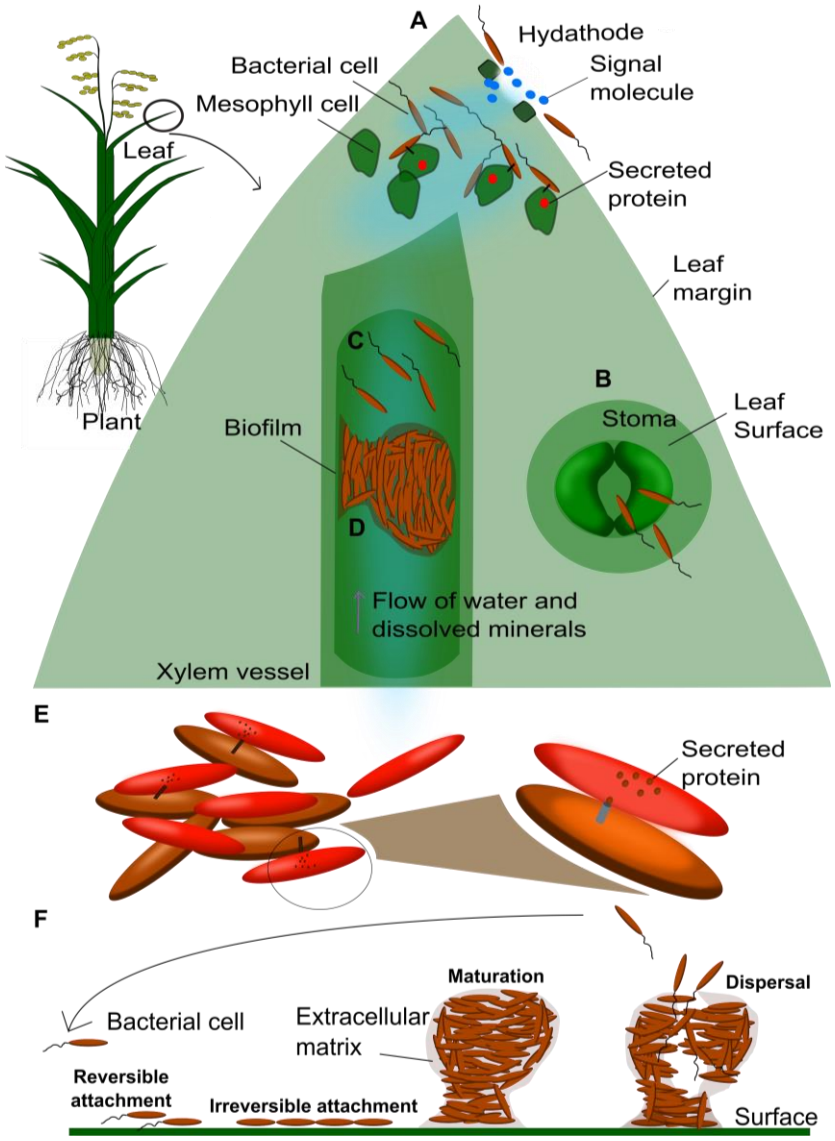


Figure 2.2. General adaptation mechanisms of bacterial plant pathogens. Depiction of bacteria entering water pores or hydathodes (**A**) via chemotactic movement (Kumar Verma, Samal, & Chatterjee, 2018) as attracted by signal molecules from xylem sap. Upon entry, pathogens inject effectors directly into plant cells via the type III secretion system (T3SS) (Coburn, Sekirov, & Finlay, 2007). Stoma (**B**) are natural openings that also facilitate pathogen entry into plants (Romantschuk & Bamford, 1986). Finally, bacteria enter the host tissue and acquire nutrients by secretion of host tissue-degrading enzymes (Cianciotto & White, 2017) (**C**), where some pathogens such as *Xylella fastidiosa* (Nascimento et al., 2016) are known to form biofilms (**D**). Plant pathogens are also known to outcompete other bacteria (**E**), by secreting T6SS

effectors directly into the cytoplasm of bacterial competitors (Hood et al., 2010). In response to extreme environmental conditions, bacteria form biofilms, which is a complex process (**F**) that involves attachment, maturation, and dispersal (Sauer, Camper, Ehrlich, Costerton, & Davies, 2002). The image was created using Inkscape 1.1 (https://inkscape.org/).

2.5.1 Adaptation and the concept of fitness

Perhaps one of the best-known examples of microbial adaptation is the development of antibiotic resistance. To put it simply, when a bacterial population with both resistant and susceptible phenotypes is exposed to a bacteriocidal antibiotic at its effective concentration, this results in the death of the susceptible variant. The resistant variant survives and subsequently multiplies, becoming the dominant variant of the population. Maintenance of resistance characteristics is highly dependent on the fitness cost of these mutations such that, the lower the fitness cost, the more likely are the mutations to be maintained in a population (A. H. Melnyk, Wong, & Kassen, 2015).

In the example above, the application of the antibiotic is the selection pressure, which favors the survival of the resistant phenotype. Within this context, the fitness of an organism is defined as the quantification of the total reproductive output of a specific genotype (Gregory, 2009). At the time of exposure to the antibiotic, the resistant variant is relatively more fit, hence adapted well to the environmental change. The resistant allele, which confers the resistant phenotype, will dominate the population in one generation following the equation $\Delta_p = \frac{pqs}{1-ps}$ (Orr, 2009), where p is the frequency of the susceptible allele, q is the frequency of the resistant allele, and s denotes the selection coefficient.

To understand fitness and adaptation, it is best to think at the level of populations, and proportions of variants within populations (Gregory, 2009) which arise primarily through mutations, or changes in the genetic code that may occur spontaneously in the processes of DNA replication (Schaaper & Dunn, 1987), transcription, and translation (Ninio, 1991). Homologous recombination (Vos & Didelot, 2009) and horizontal gene transfer (Woods et al., 2020) are also a source of variants. In plant pathogens, mutations are relatively frequent in genomic regions that contain genes encoding effectors (Karasov, Horton, & Bergelson, 2014), suggesting a

definitive role of these genes in increasing pathogen fitness against selection pressure from resistance genes of the host. Specific examples of adaptation mechanisms and fitness components are described in the following sections.

2.5.2 Adaptations to the host environment

Fatima and Senthil-Kumar (Fatima & Senthil-Kumar, 2015) describes six tissues or "nutrient niches" in plants that are colonized by pathogenic bacteria – phyllosphere, rhizosphere, apoplast, xylem, phloem, and cell organelles. This section focuses on the plant vascular system, specifically the xylem tissue.

2.5.3 Pathogen mobility in plants and entry into the vascular system

In general, bacterial pathogens enter the host tissues via natural openings facilitated by motility. For instance, *Pseudomonas syringae* pv. *phaseolicola,* the causal organism of bean halo blight, was demonstrated to enter the leaf stomata via pili-mediated motility (Romantschuk & Bamford, 1986). *Ralstonia solanacearum*, which causes bacterial wilt in a wide host range, was found to gain entry in the natural openings of roots of tomato host via chemotactic movement (Yao & Allen, 2006). Entry of *Xoo* into natural openings of the leaf has also been associated with chemotaxis (Kumar Verma et al., 2018). *X. campestris* pv. *campestris,* which causes black rot disease in crucifers, also enters the host system via hydathodes (Akimoto-Tomiyama, Furutani, & Ochiai, 2014; Meyer, Lauber, Roby, Arlat, & Kroj, 2005), however the requirement for motility is unsupported (Kamoun & Kado, 1990).

2.5.4 Life in the xylem

The anatomy and history of the plant xylem has been reviewed elsewhere (Růžička, Ursache, Hejátko, & Helariutta, 2015). It is part of the plant vascular system and is known to function in the transport of water and minerals to the leaves. The xylem contains a limited amount of nutrients such as amino acids, minerals, organic acids, and sugars compared to the phloem, which is nutrient-rich (Fatima & Senthil-Kumar, 2015).

Adaptations of bacterial pathogens to life in the nutrient-poor xylem environment have been described recently in several pathogens. *In planta* transcriptome analysis of *R. solanacearum* in tomato stems highlighted the role of sucrose uptake and catabolism, specifically the gene *scrA*, during infection (Jacobs et al., 2012). The movement of *Erwinia tracheiphila* (cause of bacterial wilt of cucurbits)

inside the xylem has been attributed to an expansin (Rocha, Shapiro, & Kolter, 2020), a protein that loosens plant cell walls non-enzymatically (Georgelis, Nikolaidis, & Cosgrove, 2015). In contrast, cellobiohydrolase CbsA, an enzyme that attacks the 1,4-β-D-glycosidic bonds of cellulose (Gilbert & Hazlewood, 1993), was determined to be associated with Xanthomonadaceae genera that inhabit the vascular system of host plants (Gluck-Thaler et al., 2020). The utilization of host machinery for nutrient acquisition through the induction of sugar efflux from plant cells mediated by effectors from the pathogen has also been speculated in the *Xoo pthXo1* and rice *OsSWEET11* interaction (L.-Q. Chen et al., 2010).

Aside from the nutrient-poor environment in the xylem, bacterial invaders must overcome the plant defense responses. These responses are categorized as physical and chemical, which prevents pathogen spread in the xylem tissue or kills the pathogen (Yadeta & J Thomma, 2013). An example of a physical defense response is the production of pectins and tyloses by grapevines (*Vitis vinifera*) during infection by *Xylella fastidiosa* (Fry & Milholland, 1990).

The escape of pathogens from plant defense mechanisms can be facilitated by delaying its initiation. For instance, *X. fastidiosa* has been observed to have a long chain O-antigen that allows it to evade host recognition (Rapicavoli et al., 2018).

2.5.4.1 Secretion systems as adaptation mechanisms to the host environment

Bacteria employ an arsenal of protein secretion systems which have specialized functions. Several types have been characterized, depending on the substrate and mechanisms of delivery of the secreted proteins (Green, Mecsas, & Kudva, 2016). In host-plant pathogen interactions, the type II and type III secretion systems are highly attributed to the successful colonization of plant tissues by bacterial pathogens (Alvarez-Martinez et al., 2021).

2.5.4.2 Type II secretion system

The type II secretion system (T2SS) or general secretion pathway was first described in *Klebsiella oxytoca* (d'Enfert, Ryter, & Pugsley, 1987), which secreted a starch debranching enzyme pullalanase under regulation of the maltose regulon. This pathway is most prevalent in *Proteobacteria*, and is related to secretion of carbohydrate-degrading enzymes in plant pathogens (Cianciotto & White, 2017). The mechanism of T2SS has been reviewed by Jha and co-workers (Jha, Rajeshwari, &

Sonti, 2005), where it is discussed that secretion of proteins via the T2SS is implemented in two steps, translocation across the cytoplasmic membrane, and secretion from the periplasmic space to the external environment, passing through the outer membrane. In general, plant pathogens are equipped with T2SS that secrete enzymes for the degradation of host tissues (Cianciotto & White, 2017).

2.5.4.3 Type III secretion system

One of the most characterized pathogen adaptation mechanism is the type III secretion system (T3SS) (McCann & Guttman, 2008). In contrast to the T2SS, the T3SS provides a continuous path for effectors to be injected into the host cytoplasm (Coburn et al., 2007), thus enabling plant pathogens to inject effectors directly into host plant cells. Components of this system are encoded by around 20 *hrp* (hypersensitivity response and pathogenicity) genes (Büttner & Bonas, 2002). In Xanthomonas, T3SS is highly linked to virulence (White, Potnis, Jones, & Koebnik, 2009; White & Yang, 2009), and escape from host defense mechanisms (Sonnewald et al., 2012).

2.5.5 Coexistence, competition, and biofilm formation

The preference of some pathogens to colonize the xylem has been attributed to the low interbacterial competition due to limiting nutrient conditions (Yadeta & J Thomma, 2013). Naturally, plant pathogens spend a portion of their life cycles in relatively nutrient-rich environments, arguably a very different environment than the xylem. Also, epiphytic growth of plant pathogens on the leaf surfaces of non-host plants have been reported (Beattie & Lindow, 1994; Zarei, Taghavi, Hamzehzarghani, Osdaghi, & Lamichhane, 2018), where biofilm formation has been found to be advantageous for some pathogens such as *X. axonopodis* pv. *citri* (Rigano et al., 2007). General mechanisms of intra- and interstrain survival are described in the following sections.

2.5.5.1 Inter- and intrastrain coexistence and competition

Bacteria exist as mixtures of populations composed of different species in natural environments (L. Liu, Zhu, Wurzburger, & Zhang, 2020; Wagg, Schlaeppi, Banerjee, Kuramae, & van der Heijden, 2019). Although there are several examples of adaptations for bacterial competitiveness (Hibbing, Fuqua, Parsek, & Peterson,

2010), this section aims to describe two general mechanisms – to coexist or to compete.

Bacterial motility is one of the main mechanisms to outcompete other bacterial species by providing a means to access nutrients. This characteristic allows bacterial populations to coexist, simply by colonizing delineated localities in a niche, based on differences in motility (Gude et al., 2020). Other reports have shown that motility is a key component of fitness. Competition experiments have demonstrated that a slower growing population is fit to compete with a faster growing population as long as it possesses superior chemotactic motility (Kelly, Dapsis, & Lauffenburger, 1988). In *R. solanacearum*, inoculation of a 1:1 mixture of a wild-type strain and a non-chemotactic mutant resulted in the domination of the tactic wild-type strain (Yao & Allen, 2006).

Another adaptation mechanism employed by bacteria is to outcompete rather than to coexist. To outcompete other bacteria, gram-negative bacteria are known to produce effectors which kill the competition through the type VI secretion system (T6SS) (Pukatzki et al., 2006). A detailed review of the T6SS (Coulthurst, 2019) describes that the VgrG-PAAR proteins as well as the Hcp facilitate the delivery of anti-bacterial effectors directly into the "enemy" cytoplasm. Like host R genes and the cognate pathogen *Avr* genes, effectivity of the T6SS operates in a similar way via effector-immunity pairs (X. Yang, Long, & Shen, 2018). This mechanism has been reported to be utilized in eliminating other species (Speare et al., 2018), however intra-strain variation of T6SS components have been reported (Shyntum, Venter, Moleleki, Toth, & Coutinho, 2014), implying that the T6SS may be involved in intra-strain competition as well. Phylogenetic analysis of the T6SS genes of Xanthomonads revealed three clades, where variability within *Xoo* is apparent (Bayer-Santos, Ceseti, Farah, & Alvarez-Martinez, 2019). In contrast to the known assumption that T6SS gene products function primarily for bacterial competition, recent work by Choi and co-workers have shown that the T6SS gene *hcp2* is essential for virulence of *Xoo* in rice and not for interbacterial competition against the tester strain *Escherichia coli* (Choi et al., 2020).

2.5.5.2 Dual lifestyles of bacteria and the formation of biofilms

In natural settings bacterial populations exist as biofilms, which are aggregates of bacteria attached to a surface (Costerton et al., 1987; O'Toole & Kolter, 1998). Biofilm formation is a switch from a free-living (planktonic) to an attached (sessile) lifestyle, a key adaptation mechanism in response to environmental conditions (de la Fuente-Núñez, Reffuveille, Fernández, & Hancock, 2013). The processes involving biofilm formation involves reversible attachment, irreversible attachment to a surface, maturation (maturation-1 and maturation-2), and dispersal (Sauer et al., 2002).

The initiation of biofilm formation is dependent on environmental signals, which designate the movement or taxis of bacteria. *R. solanacearum* is capable of moving towards a location based on potentially maximum intracellular energy levels, or aerotaxis, and form biofilms on those surfaces (Yao & Allen, 2007). In *Xanthomonas axonopodis* pv. *citri (Malamud et al., 2011)*, the ability to form biofilms depends on flagellar movement and the structure itself. High osmolarity has been shown to inhibit the formation of biofilms (O'Toole & Kolter, 1998), suggesting that biofilm formation is supported only within specific nutrient concentrations. The importance of factors specific to the surface to be attached, such as charge, stiffness, and roughness, have been identified, and discussed thoroughly in a recent review article (Zheng et al., 2021). Attachment to surfaces is facilitated by production of extracellular polysaccharides (EPS) (Killiny, Martinez, Dumenyo, Cooksey, & Almeida, 2013).

Biofilm maturation involves the clustering and aggregation of bacterial cells, and thickening of the biofilm structure (Sauer et al., 2002), which is enveloped by an extracellular polymeric matrix (Stanley & Lazazzera, 2004). The biofilm ultrastructure is regulated by quorum sensing mechanisms (J. Zhu & Mekalanos, 2003), and could also be regulated by type III secretion systems and the three component system *sadARS* as discovered in *Pseudomonas aeruginosa* (Kuchma, Connolly, & O'Toole, 2005).

As the biofilm matures, some of the cells detach from the matrix. This dispersal stage is further classified into three stages: detachment from the colony, translocation of cells, and attachment of cells onto a new surface (Kaplan, 2010). In *Vibrio cholerae*, spermidine has been implicated in the regulation of biofilm dispersal (Bridges & Bassler, 2021). At this stage, bacteria are detached from the colony, exhibiting a

planktonic lifestyle as the cells spread to another location for colonization. Signal molecules involved in cell-to-cell communication or quorum sensing have been implicated in the induction of biofilm dispersal (Solano, Echeverz, & Lasa, 2014).

The switch from a planktonic to a sessile life phase during biofilm formation is facilitated by complex gene regulatory processes that include transcriptional changes in genes involved in metabolic pathways (Guan, Luo, Fang, & Zhou, 2018) to genes involved in motility and cell wall biogenesis (Chin et al., 2015). These global changes in gene expression are attributed to a form of centralized control, such as the second messenger molecule cyclic-di-GMP, which regulates the expression of biofilm formation-related genes at the post-transcriptional level (Vadyvaloo & Martínez, 2014). In *Xanthomonas* sp., the link between virulence, motility, nutrient acquisition, and biofilm formation is apparent due to the identification of multifunctional proteins involved in these processes (H. Park et al., 2020; Yaryura et al., 2015). However, studies on biofilm-related genes of *Xoo* at the population level has not been done.

2.6 Investigating adaptation and fitness

Inducing microbial adaptation in the laboratory or in field conditions allows the researcher to study adaptation in a step-wise manner, a strategy known as experimental evolution (Kawecki et al., 2012) or adaptive laboratory evolution (ALE), where a microbe is grown under defined and controlled conditions until selection of improved phenotypes is possible (Dragosits & Mattanovich, 2013). Using this approach to understand adaptation, coupled with whole genome sequencing, allows the observation of both phenotypic variations that occur in microbial populations, as well exploring the genes that may contribute to these changes (Dettman et al., 2012). In ALE, microbes are grown for a period under selective pressure or stress, and mutations that contribute to the emergence of an improved strain are identified, as well as their improved phenotypes (Choe et al., 2019; Dettman et al., 2012; Dragosits & Mattanovich, 2013; Horinouchi et al., 2015; M. Kang et al., 2019). Improved phenotypes show an increase in fitness, which can include increased doubling time (M. Kang et al., 2019), increased growth rates (Oxman, Alon, & Dekel, 2008), modified regulation of stress responses (Dettman et al., 2012), and diversification of nutrient metabolism (Dragosits & Mattanovich, 2013).

In studying the experimental evolution of bacterial pathogens, the serial passage experiment is employed, which involves subsequent passages or inoculations and re-isolations of the pathogen from its host (Chapuis, Pagès, Emelianoff, Givaudan, & Ferdy, 2011; Guidot et al., 2014; Meaden & Koskella, 2017; W. M. Silva et al., 2017). The derived strains are then subject to rigorous characterization of phenotypic characteristics, such as: proteomic profiling (W. M. Silva et al., 2017), within-host multiplication (Meaden & Koskella, 2017), virulence (Chapuis et al., 2011), and fitness (Guidot et al., 2014); as well as explorative analyses of their genomes (Guidot et al., 2014). Understanding pathogen adaptation mechanisms to growth in the host may reveal data important in developing strategies to manage the disease caused by the pathogen.

Virulence and fitness are components of pathogen aggressiveness (Pariaud et al., 2009). *Xoo* fitness in planta could thus be inferred by aggressiveness in terms of observed lesion lengths on leaves inoculated with the pathogen, physiological characteristics of the pathogen, and its genome plasticity. Growth characteristics such as growth rate, lag phase duration (Oxman et al., 2008), and doubling time (M. Kang et al., 2019) are also used to measure fitness.

2.7 Inferring phylogenetic relationships

The work of Charles Darwin (Darwin, 1859) pioneered the visualization of ancestor and descendant relationships using a tree-like diagram. Through rigorous investigations of the similarities of organisms at the molecular level, Carl Woese introduced the system of classification that is accepted today, consisting of three domains – Archaea, Bacteria, and Eucarya (Woese, Kandler, & Wheelis, 1990). A phylogenetic tree is a visualization (Figure 2.3), with a root (for rooted trees), nodes, branches, and leaves.

Building a tree based on DNA or protein data can be done in four steps – identify homologous DNA or protein sequences, align the sequences, generate a best-fit tree based on the alignment, and present the tree in a clear manner (Hall, 2013). BLAST or basic local alignment search tool (Altschul, Gish, Miller, Myers, & Lipman, 1990) is probably the most widely used algorithm to identify homologous sequences, or similar sequences that share a common origin (Koonin & Galperin, 2003).

Various algorithms have been developed for multiple sequence alignment (MSA), which is the process of assembling alignments considering their biological structure by estimating a binary tree (Notredame, 2007). CLUSTAL W (Thompson, Higgins, & Gibson, 1994) and MUSCLE (Edgar, 2004) are examples of these algorithms.

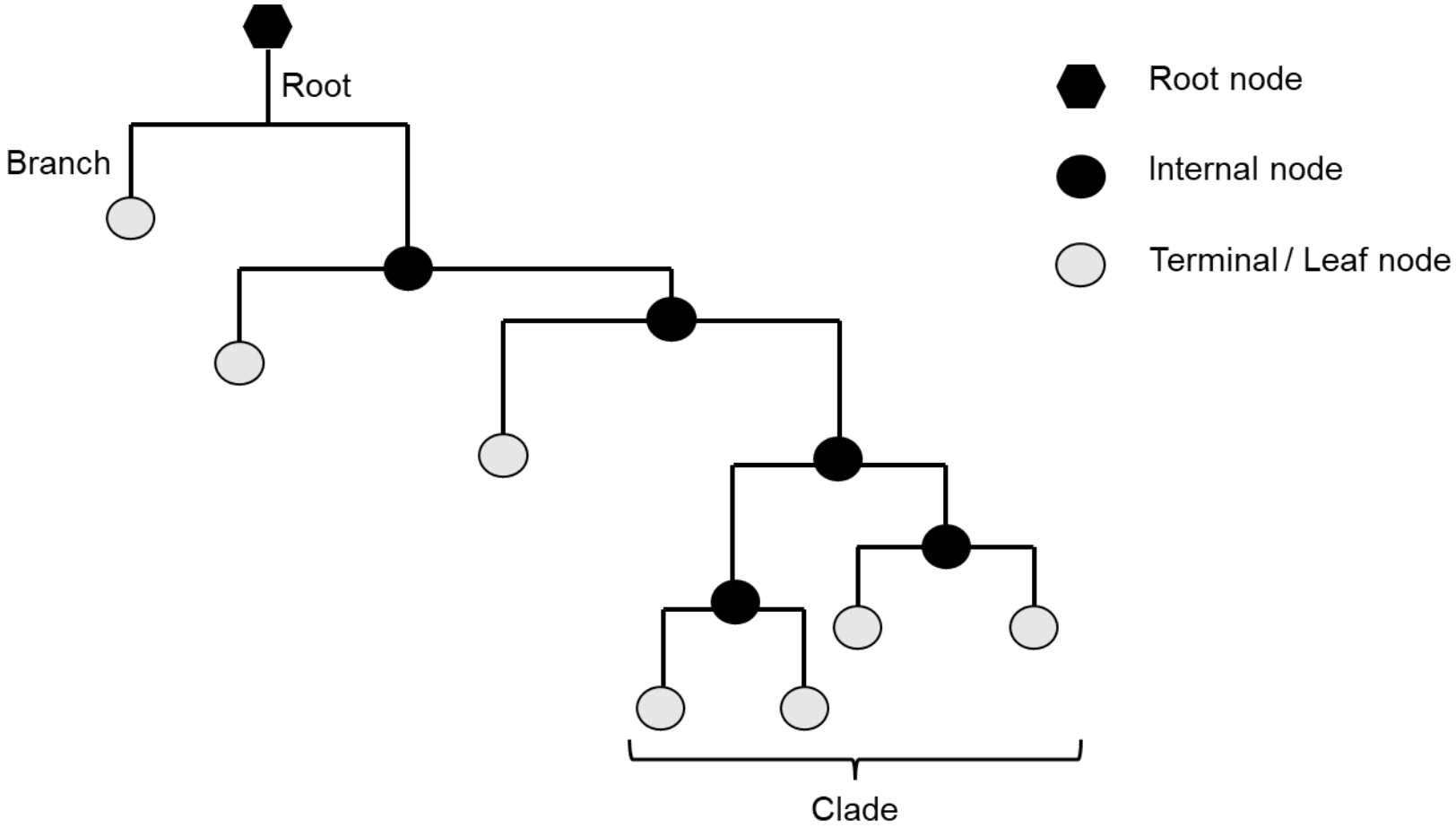


Figure 2.3. Parts of a phylogenetic tree.

The phylogenetic tree represents the relationship of species or genes with respect to a common ancestor (Baum, 2008). Several algorithms have been developed for this purpose. One of the early methods for tree generation is the Neighbor-Joining method (NJ), which calculates the tree topology and branch lengths by successive clustering of neighbor pairs based on differences, initially with the assumption that there is no clustering (Saitou & Nei, 1987). For genomic data, clustering is based on substitutions. Algorithms based on Bayesian analysis have also been developed, which makes it possible to analyze mixed data, i.e. phenotypic and genotypic datasets (Ronquist & Huelsenbeck, 2003). In estimating phylogenetic trees based on genomic sequences, the maximum likelihood method (Felsenstein, 1981) is often considered, as the best tree is estimated by exhaustively determining the tree with the highest likelihood to represent the data (Hollich, Milchert, Arvestad, & Sonnhammer, 2005).

The reduction in costs of whole genome sequencing, and the development of advanced tools and computer programs have enabled biologists and bioinformaticians to perform rather complex analyses of phylogeny based on genomic variation of organisms. This has led to a vast increase in research related to the phylogenetic inferencing based on whole genome sequences or genes of interest. For instance, Morris and co-workers (Morris et al., 2010) investigated the evolution of *P. syringae* by inferring phylogeny based on four housekeeping genes. By attaching habitat data to the phylogenetic tree, the research uncovered "genetic ecotypes" of *P. syringae* were discovered. With enough computational resources, it is even possible to combine trees. In one study, 730 trees were combined to build supertrees using both NJ and ML approaches, providing additional evidence to support that the three domains of life indeed originate from one ancestor (Daubin, Gouy, & Perrière, 2002), as previously reported (Woese et al., 1990).

2.8 Investigating host-induced adaptation mechanisms of *Xoo*

This dissertation reports three studies that aimed to investigate host-induced adaptation mechanisms of *Xoo*. Chapter 3 describes research on biofilm-related genes of Xanthomonadales in *Xoo*. The study was designed to consolidate some of the characterized biofilm-related genes in literature and examine the variation of these genes among well-characterized *Xoo* strains from the Philippines, while considering the population structure of these strains. Chapter 4 reports research on the close

examination of the adaptation potential of *Xoo* as induced by serial passage into a resistant rice line, and subsequent characterization of the wild-type strain and the variant which was passaged into the host. The aim of the study was to gain some insight on small yet important variations that potentially lead to the population shifts that are extensively reported in literature. Chapter 5 describes the development of a monitoring tool for *Xoo* populations, and pilot testing-through an extensive inter-institute collaboration.

CHAPTER 3: VARIATION IN THE BIOFILM - RELATED PROTEINS OF XANTHOMONADS IN *Xanthomonas oryzae* pv. *oryzae*

3.1 Introduction

Although the extent of genetic diversity in *Xoo* has been established in previous work based on restriction fragment length polymorphism (RFLP) analysis with avrXa10 and insertion sequence IS1112 as probes (Adhikari et al., 1995; Islam et al., 2016; Mishra et al., 2013), and whole genome sequencing (Quibod et al., 2020; Quibod et al., 2016; Salzberg et al., 2008), multi-gene genetic analysis of *Xoo* based on biofilm-related genes has not been conducted.

Recent research on the rice-*Xoo* interaction has been focused on the hijacking of the OsSWEET genes by *Xoo* (Antony et al., 2010; L.-Q. Chen et al., 2010; Streubel et al., 2013; Zaka et al., 2018), giving rise to new approaches in developing resistant varieties to BB through gene-editing (Eom et al., 2019). Since bacteria exist as biofilms or surface-attached microbial communities in nature (Costerton et al., 1987; O'Toole & Kolter, 1998), and biofilm formation is known to affect virulence of plant pathogens (Conforte et al., 2019; Gouran et al., 2016; M. S. Silva, De Souza, Takita, Labate, & Machado, 2011; Yaryura et al., 2015), further research on biofilm formation of *Xoo* could help in the development of new methods to manage the disease. For instance, the chemical control of BB by niclosamide (S.-I. Kim, Song, Jeong, & Seo, 2016) has been linked to its inhibitory action against *Xoo* biofilm formation (Sahu et al., 2018). Moreover, biofilm formation of plant pathogens is associated with disease resistance in some pathosystems (Castiblanco & Sundin, 2016).

Biofilm formation is a complex process, which is initiated by movement based on environmental conditions such as osmolarity (O'Toole & Kolter, 1998) that trigger bacterial taxis (Malamud et al., 2011; Yao & Allen, 2007) and subsequent attachment to surfaces with conducive physical attributes such as roughness or charge (Zheng et al., 2021) facilitated by production of EPS (Killiny et al., 2013). Thus, bacterial attachment to surfaces is indeed a non-random process (Grinberg, Orevi, & Kashtan, 2019). The maturation stage of biofilms is mediated by quorum sensing mechanisms (J. Zhu & Mekalanos, 2003) characterized by the aggregation of bacterial cells to form complex ultrastructure (Sauer et al., 2002). Finally, the dispersal stage occurs where

cells detach from the colony and translocate, and then attaches onto a new surface (Kaplan, 2010). *In Xoo*, biofilm formation is regulated by a diffusible signal factor (DSF) which negatively regulates motility and positively regulates attachment and biofilm formation (Rai, Ranjan, Pradhan, & Chatterjee, 2012). Several genes have been identified to be involved in biofilm formation in *Xoo*, as well as other processes such as carbon source utilization, EPS production, and motility (Bae, Park, Park, Kim, & Han, 2018; Cho, Yoon, Lee, Noh, & Cha, 2013; X.-Y. Wang, Zhou, Yang, Ji, & He, 2016).

The ability to form biofilms in plant tissues is considered as a virulence factor (Marques, Ceri, Manfio, Reid, & Olson, 2002; Yao & Allen, 2007), increasing pathogen fitness *in planta*. For instance, impairment of biofilm formation by the citrus canker pathogen *X. citri* subsp. *citri* (formerly *X. axonopodis* pv. *citri*) resulted in reduction of growth and epiphytic survival (Rigano et al., 2007). In order of significance : *Pseudomonas syringae* (Aragón, Pérez-Mendoza, Gallegos, & Ramos, 2015), *R. solanacearum* (Mori et al., 2016), *Agrobacterium tumefaciens* (Merritt, Danhorn, & Fuqua, 2007), and *E. amylovora* (Koczan, McGrath, Zhao, & Sundin, 2009) are examples of plant pathogenic species that are known to produce biofilms. Biofilm formation of clinical pathogens such as *Escherichia coli* and *Salmonella enterica* on leaf surfaces contribute to their lifestyles outside human and animal hosts and to disease outbreaks related to fresh produce (Yaron & Römling, 2014).

This study aims to aid subsequent research on biofilm formation of *Xoo* strains by investigating the occurrence and variation of previously identified biofilm-related genes of Xanthomonads in the genomes of Philippine *Xoo* strains and identifying genes that are potential drivers of *Xoo* evolution.

3.2 Methodology

3.2.1 Genomes used in this study

The genomes of 80 *Xoo* strains and the corresponding race classifications were downloaded from previous research of Quibod et al (Quibod et al., 2020), where the effect of the R gene *Xa4* as a selection pressure driving the evolution of *Xoo* populations in the Philippines was demonstrated through comparative analysis of *Xoo* whole genome sequences and association with the deployment of rice varieties containing *Xa4*. DNA sequences from the Genbank formatted files were translated

into amino acid sequences in Biophyton Seq.IO and were stored in the fasta format for subsequent analysis.

3.2.2 Biofilm-related genes of Xanthomonads

A preliminary search was performed in Pubmed using the keywords "biofilm AND xanthomonas". The search results were downloaded in RIS format and input into R (R Core Team, 2021) using the *import_results* function of the package ***litsearchr*** (E. Grames, Stillman, Tingley, & Elphick, 2020; E. M. Grames, Stillman, Tingley, & Elphick, 2019). The package vignette (https://elizagrames.github.io/litsearchr/litsearchr_vignette.html) provides a detailed protocol for the generation of Boolean search terms to increase literature coverage. Briefly, possible key words were extracted from the titles and abstracts of the search results using the *extract_terms* function and were grouped manually into biofilm-related terms or *Xanthomonas*-related terms. A list of the extracted terms is shown in Table A.1. The Boolean search was generated using the *write_search* function of the same package. Search results were sorted by "best match" and downloaded as a citation file, which was uploaded into the software ASReview Lab v. 0.17 for semi-automated screening of articles (van de Schoot et al., 2021). Prior to screening, a rule was set to discontinue the screening process if 20 consecutive irrelevant articles in succession was predicted by the software. In creating the gene list, only genes that were identified through mutation or gene expression experiments and where biofilm-formation was directly assayed were considered. All genes found within genera belonging to the order Xanthomonadales were included. The gene sequences were downloaded from the NCBI website and from the UNIPROT database. Also downloaded were the 13 reference genomes from which the genes were derived from. These include *Stenotrophomonas maltophilia* K279a (AM743169.1), *Xanthomonas axonopodis* pv. *citri* str. 306 (AE008923.1), *Xanthomonas axonopodis* Xac29-1 (CP004399.1), *Xanthomonas campestris* pv. *campestris* str. 8004 (CP000050.1), *Xanthomonas campestris* pv. *campestris* str. ATCC 33913 (AE008922.1), *Xanthomonas campestris* pv. *campestris* strain 17 (NZ_CP011946.1), *Xanthomonas campestris* pv. *vesicatoria* (AM039952.1), *Xanthomonas citri* subsp. *citri* A306 (CP006857.1), *Xanthomonas oryzae* pv. *oryzae* KACC 10331 (AE013598.1), *Xanthomonas oryzae* pv. *oryzicola* BLS256 (NC_017267.2), *Xanthomonas oryzae* pv.

oryzicola strain RS105 (CP011961.1), *Xylella fastidiosa* Temecula1 (AE009442.1), *Xanthomonas oryzae* pv. *oryzae* PXO99A (CP000967.2).

3.2.3 Identification of homologous regions in Philippine Xanthomonas oryzae pv. oryzae genomes

Command-line BLAST v. 2.9.0-2 was used to find homologous regions of the identified biofilm-related genes of Xanthomonads in the Philippine *Xoo* genomes. First, a local database of the translated genomes was created using the makeblastdb function. Then, the genes were used as query sequences against the local blast database using the tblastn function. Results were filtered according to percent identity and E-value of less than 0.00001.

The 13 reference genomes were included in the local blast database to serve as controls during the runs.

3.2.4 Multiple sequence alignment and phylogenetic inference

Multiple sequence alignment (MSA) was performed using MUSCLE (Edgar, 2004) program v3.8.1551 using the options *-diags1 -sv -distance1 kbit20_3* as recommended in the user manual. The aligned sequences were concatenated, and the best evolutionary models were inferred using the software Partition Finder (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) v2. To estimate the maximum likelihood tree, a partitioned analysis was performed in IQ-Tree (Nguyen, Schmidt, von Haeseler, & Minh, 2015) v2 where branch supports were analyzed at with the built-in SH-like approximate likelihood ratio test (Guindon et al., 2010) and ultrafast boostrap functions at 1000 replications. Tree visualization was performed in the R package ggtree (G. Yu, Smith, Zhu, Guan, & Lam, 2017). The 12,904 single nucleotide polymorphisms (SNPs) from the core genome of 91 Philippine *Xoo* strains (Quibod et al., 2020) was used to infer phylogenetic structure at the genome level by applying the Bayesian Analysis of Population Structure (BAPS) (L. Cheng, Connor, Sirén, Aanensen, & Corander, 2013) in the R package Rhierbaps (Tonkin-Hill, Lees, Bentley, Frost, & Corander, 2018).

3.2.5 Multidimensional scaling

The R (R Core Team, 2021) package bios2mds v.1.2.3 (Pelé, Bécu, Abdi, & Chabbert, 2012) was used to generate multidimensional scaling (MDS) plots. First, a dissimilarity matrix was calculated with the function *mat.dis* using PAM250 amino

acid substitution matrix. Then, *mmds.2D.plot* and *mmds.3D.plot* functions were used to visualize the multidimensional data in 2D and 3D, respectively.

3.3 Results

A list of 56 biofilm-related genes (Table A.1) of Xanthomonadales was populated by searching publicly available data. Literature search from the Pubmed database using the expanded search garnered 1,010,338 hits in the Pubmed database derived from extracted terms. This facilitated the creation of a database with 10000 publications. The top 10000 hits were selected by sorting the list according to relevance. A total of 247 articles were reviewed in ASReview Lab (van de Schoot et al., 2021), where 117 were scored as relevant to the study. The final list included genes that were identified or characterized in *Stenotrophomonas maltophilia, X. axonopodis* pv. *glycines, X. campestris* pv. *campestris, X. citri* subsp. *citri* (formerly *X. axonopodis* pv. *citri*), *Xoo, X. oryzae* pv. *oryzicola,* and in *Xylella fastidiosa*. Only genes reported in studies where the biofilm-formation was quantified were included in the list of genes, except for *hshC, hscA, rpoN2, rpfB* and *rpfF* which were included based on their links to cellular processes that are known to be highly related to biofilm formation.

Table A.1. Biofilm-related genes of Xanthomonadales extracted from published literature and used in this study to search for homologs in *Xanthomonas oryzae* pv. *oryzae* strains from the Philippines.

Accession no.	Gene name	Locus tag	Function	Reference
Stenotrophomonas maltophilia, nosocomial infections				
CAQ45732.1	*rpfF*	Smlt2235	Biofilm formation, swimming, and twitching motility	(Alcaraz, García, Friedman, & de Rossi, 2019)
CAQ45793.1	*fsnR*	Smlt2299	Regulation of flagellar assembly, motility, and biofilm formation	(X.-M. Kang et al., 2015)
X. axonopodis pv. *glycines*, bacterial pustule in soybean				
AAM35374.1	*clp*	XAC0483	Negative regulation of carbohydrate utilization, biofilm formation, EPS production	(Guo et al., 2019)
X. campestris pv. *campestris*, black rot disease in crucifers				
AAM41224.1	*rpoN2*	XCC1935	Biofilm formation, EPS production, swimming ability, flagellar synthesis	(K. Li et al., 2020)
AAY47792.1	*prc*	XC_0714	Positive regulator of extracellular protease production, attachment, biofilm formation	(C.-T. Liao et al., 2016)
AAY49266.1	*lolA*	XC_2211	Attachment, biofilm formation, extracellular enzymes	(C.-T. Liao, Chiang, & Hsiao, 2019)
AAY49386.1	*rpfC*	XC_2333	Positive regulation of biofilm formation.	(Cai et al., 2017)
AAY49556.1	*rsmA*	XC_2506	Negative regulator of biofilm formation	(Lu et al., 2012)
AAY49748.1	*mip*	XC_2699	positive regulator of EPS production	(Meng et al., 2011)
WP_011037861.1	XCC2731	XCC2731	Aggregation, negative regulator of motility, negative regulator of EPS production	(Hsiao, Liu, Fang, & Song, 2011)
CAJ25451.1	*wxcB*	XCV3720	Negative regulator of biofilm formation. Positive regulator of motility	(H.-J. Park, Jung, & Han, 2014)
X. citri subsp. *citri* (formerly *X. axonopodis* pv. *citri*), citrus canker				
AAM35959.1	*hupB*	XAC1081	Flagella synthesis (Flagella is required in biofilm formation)	(Conforte et al., 2019)
AAM36155.1	*bfdR*	XAC1284	Biofilm formation, regulates *rpfF*	(T.-P. Huang, Lu, & Chen, 2013)
AAM37406.1	*fixL*	XAC2555	Positive relationship with biofilm formation, adhesion; Negative relationship with EPS production and swarming motility	(Kraiselburd et al., 2012)

Table 3.1 Continued

Accession no.	Gene name	Locus tag	Function	Reference
AAM37955.1	*gpsX*	XAC3110	EPS and LPS biosynthesis, motility, biofilm formation	(J. Li & Wang, 2012)
AAM38443.1	*wzt*	XAC3600	Biosynthesis of lipopolysaccharides (LPS) and consequently, biofilm formation	(Petrocelli, Tondo, Daurelio, & Orellano, 2012)
AAM38844.1	*ecnA*	XAC4008	Positive relationship with biofilm formation and EPS production	(Granato et al., 2019)
AAM38858.1	*phoP*	XAC4023	Positive relationship with biofilm formation motility, and virulence	(Wei et al., 2019)
AGH77232.1	*hfq*	XAC29_08740	Positive regulator of motility and biofilm formation	(X. Liu, Yan, Wu, Zhou, & Wang, 2019)
AJD70189.1	*nlxA*		EPS production, biofilm formation,	(Yan, Hu, & Wang, 2012)
X. oryzae pv. *oryzae*, bacterial blight of rice				
AAW76139.1	*rpfE*	XOO2885	Regulation of swarming motility and utilization of carbon sources; not associated with DSF production	(Cho et al., 2013)
AAW77242.1	*dgcA*	XOO3988	Positive regulator of biofilm formation	(J. Su et al., 2016)
ACD56999.1	*edpX1*	PXO_03877	Positive association with biofilm formation	(Fenghuan Yang et al., 2019)
ACD58005.1	*edpX2*	PXO_04753	Negative regulation of biofilm formation	(Shahbaz et al., 2020)
ACD58006.1	*cheW1*	PXO_04752	Negative regulation of biofilm formation	(Shahbaz et al., 2020)
ACD58121.1	*pilCX*	PXO_04887	Positive regulation of biofilm formation	(C. Yu et al., 2020)
ACD58122.1	*pilAX*	PXO_04886	Positive regulation of biofilm formation	(C. Yu et al., 2020)
ACD58395.1	*rpfB*$^{\alpha}$	PXO_00067	Negative regulation of DSF family signals. Positive regulation of EPS production, extracellular amylase activity and pathogenicity. May be involved in production of xanthomonadin.	(X.-Y. Wang et al., 2016)
ACD58842.1	*gdpX1*	PXO_00649	biofilm formation, xylanase activity, repress flagella synthesis	(F. Yang et al., 2016)
ACD59167.1	*rpoN2*$^{\alpha}$	PXO_00995	Positive regulation of flagellar motility, contributes to virulence	(Tian et al., 2015)
ACD59169.1	*fleQ*	PXO_00993	Biofilm formation, twitching and swarming motility	(Bae, Park, Park, Kim, Do, et al., 2018)
ACD59175.1	PXO_00987	PXO_00987	Positive association with motility, EPS	(Haiyun Li, Yu, Chen, Tian, & He, 2015)

Table 3.1 Continued

Accession no.	Gene name	Locus tag	Function	Reference
			production, biofilm formation, and virulence	
ACD59367.1	*gigX6*	PXO_00987	positive association with EPS production, biofilm formation	(C. Yu et al., 2018)
ACD60174.1	*tdrX2*	PXO_01883	Negative regulation of biofilm formation	(Shahbaz et al., 2020)
ACD60218.1	*thiG*	PXO_01841	Negative regulator of biofilm formation related genes *rpfC* and *rpfG*	(X. Yu et al., 2015)
ACD60919.1	*rr35*	PXO_02637	Positive relationship with biofilm formation and motility	(Antar, Lee, Yoo, Cho, & Lee, 2020)
ACD61037.1	PXO_03177	PXO_03177	Positive association with biofilm formation	(B. Wang, Wu, Zhang, Qian, & Liu, 2018)
AEQ96721.1	*zwf*	XOC_2607	Regulation of *rpfF*, *rpfG*, *clp*; motility; EPS production. Indirect relationship with biofilm formation.	(Guo, Zou, Cai, & Chen, 2015)
AWN56701.1	*oprBXo*		Negative regulator of biofilm formation, twitching, and swarming motility	(Bae, Park, Park, Kim, & Han, 2018)
X. oryzae pv. *oryzae* / *X. oryzae* pv. *oryzicola*, Bacterial blight of rice / Bacterial leaf streak of rice				
AEQ98395.1	*ankB*	XOC_4324	Positive association with EPS production, swimming motility, biofilm formation, catalase activity	(Pan et al., 2018)
X. oryzae pv. *oryzicola*, Bacterial leaf streak of rice				
AKO16449.1	*gumD*	ACU12_12720	Associated with biofilm formation, EPS production, and virulence	(Y. Zhang et al., 2013)
WP_014503217.1	*rpfG*	XOC_2264	Positive regulator of biofilm formation.	(Y. Zhang et al., 2013)
WP_014504558.1	*hshC*$^{\alpha}$	XOC_3768	Important in epiphytic survival of Xoc	(Song et al., 2015)
WP_082350419.1	*hshA*$^{\alpha}$	XOC_3767	Important in epiphytic survival of Xoc	(Song et al., 2015)
AEQ97862.1	*hshB*	XOC_3772	Motility, biofilm formation	(Song, Zhao, Qian, Odhiambo, & Liu, 2017)
AEQ97874.1	*xagB*	XOC_3784	Biofilm formation	(Y. Zhang et al., 2013)
AEQ97875.1	*xagA*	XOC_3785	Biofilm formation	(Y. Zhang et al., 2013)
Xylella fastidiosa, Pierce's disease of grapevines; multiple hosts				
AAO28287.1	*rpfF*$^{\alpha}$	PD_0407	Global regulator, Diffusible signal factor (DSF) synthase	(Ionescu et al., 2013)
AAO28716.1	*pilL*	PD_0848	Pil-Chp operon has a positive relationship	(Cursino et al., 2011)

Table 3.1 Continued

Accession no.	Gene name	Locus tag	Function	Reference
			with biofilm formation and motility	
AAO29233.1	*xhpT*	PD_1386	Involved in cell-cell aggregation and attachment	(Voegel et al., 2013)
AAO29511.1		PD_1671	Negative regulation of gum genes (involved in EPS synthesis) and biofilm formation	(Cursino et al., 2015)
AAO29541.1	*lesA*	PD_1703	Negative regulator of biofilm formation	(Nascimento et al., 2016)
AAO29547.1	*mopB*	PD_1709	Biofilm formation, cell aggregation, attachment	(H. Chen, Kandel, Cruz, Cobine, & De La Fuente, 2017)
AAO29754.1	*pilA1*	PD_1924	Type IV pili	(Kandel, Chen, Fuente, & Stabb, 2018)
AAO29756.1	*pilA2*	PD_1926	Twitching motility, biofilm formation	(Kandel et al., 2018)
AAO28818.1	*prtA*	PD_0956	Positive regulation of genes involved in the biofilm formation process	(Gouran et al., 2016)

[α]The effect of deletion of these genes on biofilm-formation was not experimentally determined by the reporting study indicated in the last column.

3.3.1 Biofilm-related proteins of Xanthomonads are conserved in Philippine Xoo strains

The importance of biofilms in plant-associated lifestyles of bacteria (Bogino, Oliva, Sorroche, & Giordano, 2013) and as a virulence factor (Marques et al., 2002; Yao & Allen, 2007) leads to the hypothesis that biofilm-related genes are conserved in a vascular pathogen such as *Xoo*. To test this hypothesis, a homology search was conducted using the BLAST algorithm (Altschul et al., 1990) to query a local database of 80 *Xoo* genomes with the amino acid translations of the 56 genes as query sequences. Out of the 56 genes, 50 were found positive for homologous regions in the 80 *Xoo* genomes included in this study. The results of the BLAST analysis are shown in Figure 3.1. Percent identity ranged from 29 to 100%, with most of the query proteins showing high sequence similarity with the BLAST hits in the *Xoo* genomes. Indeed, majority of the proteins (48 out of the 50) were conserved in the 80 Philippine *Xoo* genomes included in this study and did not differ among the strains.

3.3.2 A minor set of genes showed variability in Philippine Xoo strains

Out of the 56 test proteins, six proteins were not found in any the Philippine *Xoo* genomes used in this study (AAO28818.1, AEQ97862.1, AEQ97874.1, AEQ97875.1, AJD70189.1, CAJ25451.1). AAO28818.1 is a secreted protein encoded by the gene *prtA* in *X. fastidiosa* and was found to be involved in growth and pathogenicity as well as biofilm formation (Gouran et al., 2016). AEQ97862.1 was characterized in *X. oryzae* pv. *oryzicola*, encoded by the gene *hshB* and is important in motility and biofilm formation (Song et al., 2017). AEQ97874.1 and AEQ97875.1 are encoded by the genes *xagB* and *xagA*, respectively. These were also characterized in *X. oryzae* pv. *oryzicola* (Y. Zhang et al., 2013). AJD70189.1, encoded by the gene *nlxA*, was characterized in *X. citri* subsp. *citri* to be involved in EPS production as well as biofilm formation (Yan et al., 2012). CAJ25451.1 is encoded by the gene *wxcB* and was characterized in *X. campestris* pv. *vesicatoria* to be involved in motility and biofilm formation (H.-J. Park et al., 2014).

Inferring the sequence variability of each protein based on the BLAST percent identity (Figure 3.1 and Figure 3.2), five proteins were observed to vary across the *Xoo* strains - AAM38443.1, AAO28716.1, AAM41224.1, AAY49386.1, AAO29541.1. AAM38443.1 is encoded by the gene *wzt*, and is involved in the biosynthesis of lipopolysaccharides (Petrocelli et al., 2012). Experiments done by

Petrocelli and co-workers on *X. axonopodis* pv. *citri* also showed its involvement in biofilm formation. AAO28716.1 is encoded by the gene *pilL* and was characterized in *X. fastidiosa* (Cursino et al., 2011). AAM41224.1 is encoded by *rpoN2* and was determined to be involved in biofilm formation, production of EPS, swimming motility, and flagellar synthesis in *X. campestris* pv. *campestris* (K. Li et al., 2020). AAY49386.1 is encoded by *rpfC* and was studied also in *X. campestris* pv. *campestris* (Cai et al., 2017). AAO29541.1 is encoded by the gene *lesA* in *X. fastidiosa* and negatively regulates biofilm formation (Nascimento et al., 2016).

Out of the 56 proteins, three proteins were observed to vary in terms of presence or absence in the 80 Philippine *Xoo* genomes (Figure 3.2). WP_014504558.1 (present in 38 out of 80 genomes) and WP_082350419.1 (present in 43 out of 80 genomes) are encoded by the genes *hshC* and *hshA*, respectively, in *X. oryzae* pv. *oryzicola* and were found to be associated with *in planta* growth and epiphytic survival of this pathogen (Song et al., 2015). ACD61037.1 (present in 78 out of 80 genomes) is a hypothetical protein (locus tag PXO_03177) characterized in *Xoo* strain PXO99A as a required protein for virulence as well as biofilm formation and tolerance to sodium dodecyl sulfate (B. Wang et al., 2018).

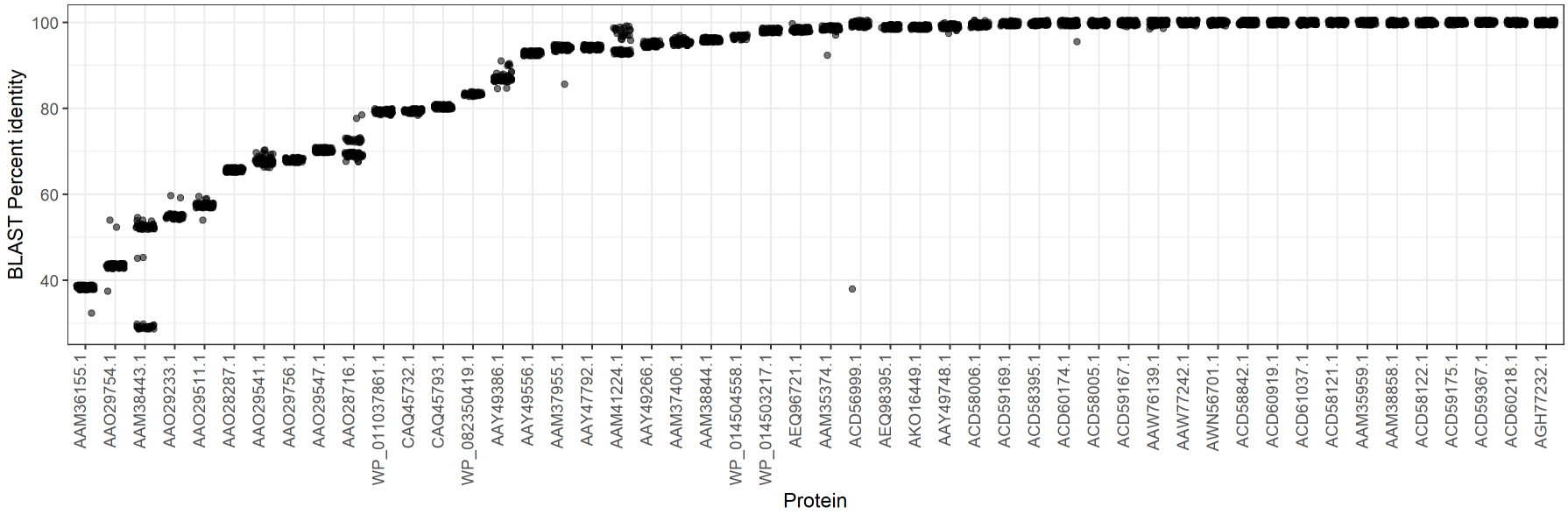


Figure 3.1. BLAST percent identity scores of the fifty biofilm-related proteins previously characterized in *Stenotrophomonas maltophilia*, *Xanthomonas axonopodis* pv. *glycines*, *X. campestris* pv. *campestris*, *X. citri* subsp. *citri* (formerly *X. axonopodis* pv. *citri*), *X. oryzae* pv. *oryzae*, *X. oryzae* pv. *oryzicola*, and in *Xylella fastidiosa* against 80 genomes of Philippine *X. oryzae* pv. *oryzae* strains.

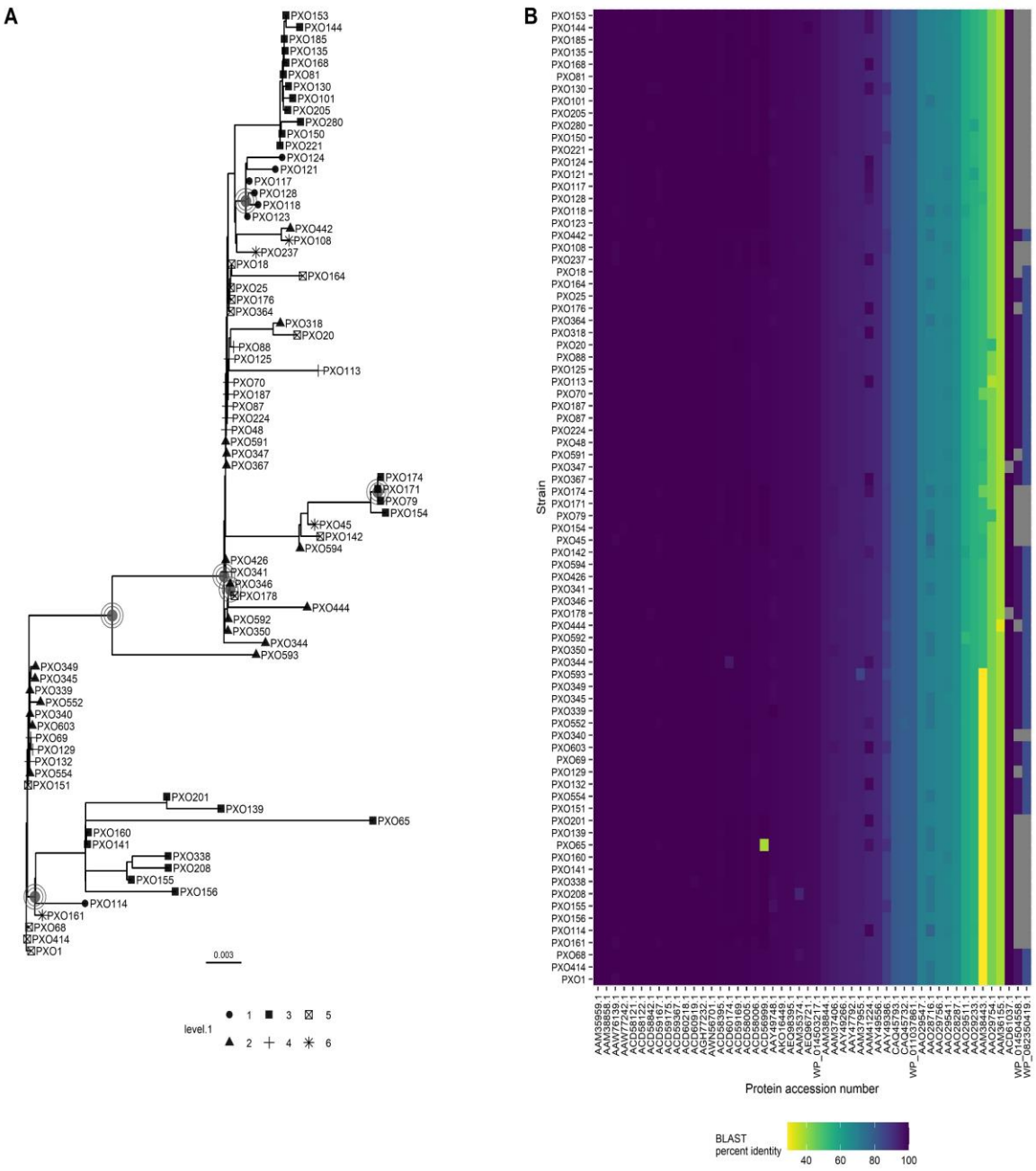


Figure 3.2. Phylogeny of Philippine *Xanthomonas oryzae* pv. *oryzae* based on biofilm-related proteins of Xanthomonads. A. Maximum likelihood tree estimated in IQ-Tree (Lanfear et al., 2016) from multiple sequence alignment of concatenated amino acid sequences from 50 biofilm-related genes previously reported in Xanthomonads. Nodes with concentric circles represent nodes with support values of Shimodaira-Hasegawa approximate likelihood ratio test, SH-aLRT (Guindon et al., 2010) >= 80 and ultrafast bootstrap (Hoang, Chernomor, von Haeseler, Minh, & Vinh, 2017) >= 95. Clustering was estimated based on the 12,904 core SNPs of Philippine *Xoo* strains (Quibod et al., 2020) by

applying the Bayesian analysis of population structure algorithm (L. Cheng et al., 2013) implemented in Rhierbaps (Pelé et al., 2012; Tonkin-Hill et al., 2018). B. Percent identity of homologous proteins found in *Xanthomonas oryzae* pv. *oryzae* genomes using the *blastp* algorithm. Gray boxes represent absence of the gene in the genomes of the strains. Visualizations generated in ggtree (G. Yu et al., 2017).

3.3.3 Biofilm-related proteins of Xanthomonadales are variable across inferred population structure of Xoo

To infer the phylogeny of Philippine *Xoo* strains based on the 50 biofilm-related proteins, the MSA-based maximum-likelihood (ML) tree was estimated in IQ-Tree (Nguyen et al., 2015) and is shown in Figure 3.2. The results show that the 80 strains are divided into two main clades based on K-means clustering of amino acid dissimilarity scores of the 50 genes (Figure A.2).

In applying these BAPS clusters to the ML tree (Figure 3.2), it can be observed that strains predicted to be in the same BAPS cluster tend to cluster together in sub-groups along the ML tree based on MSA of 50 biofilm-related proteins.

3.3.4 The fifty biofilm-related proteins of Xanthomonads found in Xanthomonas oryzae pv. oryzae can be grouped into 20 subsets according to best-fit evolutionary models

An expanded literature search of biofilm-related genes in Xanthomonadales yielded genes that are involved in diverse functions such as motility, production of extracellular polysaccharides (EPS), and carbon source utilization. With this diversity, these genes could be evolving differently or are affected by independent environmental factors. To test this hypothesis, the resultant amino acid sequences from the *Xoo* genomes were analyzed to predict the best-fit evolutionary models and partition the data set according to these models. The concatenated alignments were analyzed with Partition Finder 2 (Lanfear et al., 2016), which identified 20 subsets or partitions (Table A.1).

Table A.1. Best-fit partitions and substitution models of concatenated amino acid alignments of homologous biofilm-related proteins of Xanthomonadales in *Xanthomonas oryzae* pv. *oryzae* as predicted by the Partition Finder 2 (Lanfear et al., 2016) software.

Subset	Best model[α]	Number of sites	Protein names
1	WAG+G+F	1246	AAM35374, AAO29511, AAO29233
2	WAG+G	90	AAM35959
3	WAG+G+F	1385	AAM38443, ACD56999, AAM36155
4	WAG+G+F	502	AAM37406
5	WAG+G+F	2192	ACD61037, AAM37955, AAY49386
6	LG+G+F	332	AAM38844, AAO29756, ACD58122
7	WAG+G+F	1770	AAM38858, AEQ98395, AEQ96721, WP_014503217, CAQ45793, ACD60218
8	WAG+G+F	1599	ACD59167, AAM41224, ACD58006, ACD59169
9	WAG+G+F	1625	AAO28287, CAQ45732, ACD58005, ACD58842, AAW77242
10	WAG+G+F	2308	AAO28716
11	LG+G+F	1905	ACD60174, AAO29541, AWN56701
12	WAG+G+F	474	AAO29547, AGH77232
13	WAG	160	AAO29754
14	WAG	836	AAW76139
15	WAG+G+F	950	AAY49266, AAY47792
16	LG+G+F	572	AAY49556, WP_011037861
17	WAG+G	245	AAY49748
18	LG+G+F	1107	ACD60919, ACD58121, ACD58395
19	LG+G+F	432	ACD59175, ACD59367
20	LG+G+F	704	WP_014504558, AKO16449, WP_082350419

[α]The substitution, rate heterogeneity, and amino acid frequency models are explained in the IQ-Tree software manual. In this column, the models for each component are separated by a plus (+) sign. WAG = Whelan and Goldman (Whelan & Goldman, 2001), LG = Le and Gascuel (Le & Gascuel, 2008) ;+G = Free rate model (Z. Yang, 1994); +F = amino acid frequencies are calculated empirically. Best-fit models were determined by the software based on the Akaike information criterion (AIC).

3.3.5 Variability of biofilm-associated proteins found in the 80 Philippine Xoo genomes are associated with BAPS clustering based 12,904 core genome SNPs of Philippine Xoo strains.

To validate the observations in combining the phylogenetic analysis with population structure analysis, a multidimensional scaling (MDS) analysis was performed which was based on dissimilarity scores calculated from amino acid substitutions. Like the results obtained from the phylogenetic approach described above, simplifying the data into two axes by MDS analysis showed that the strains separate along PC1, and form sub-groups along PC2 (Figure 3.3A). However, the separation of BAPS Cluster 3 along PC2 was evident. The similarity of the other BAPS clusters can be observed when PC3 is considered in the visualization (Figure 3.3B). Overlaying the BAPS clusters in the visualization reveal that most of the strains that grouped together based on MDS also belonged to the same BAPS cluster.

In section 3.3.2, it was presented that there were five proteins that showed variability in terms of sequence, and three in terms of presence or absence in the 80 Philippine *Xoo* genomes. Figure 3.4 shows a visualization of these proteins as well as the BAPS clusters. It was determined that the occurrence of the two proteins WP_014504558.1 and WP_082350419.1 in Philippine *Xoo* genomes are strongly associated with the BAPS clusters (Cramer's V = 0.8295 and 0.9489, respectively).

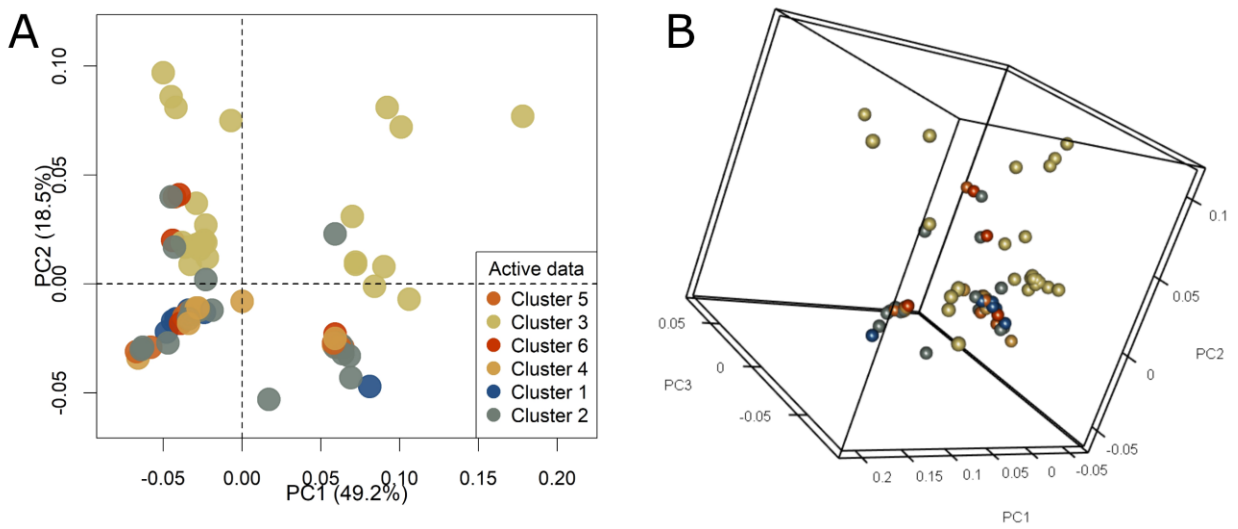


Figure 3.3. The relationship of biofilm-related proteins and population structure of 80 *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) strains based on multidimensional scaling (MDS) of dissimilarity scores calculated from amino acid substitutions. Colors are based on clusters obtained from Bayesian analysis of population structure (BAPS) inferred from 12,904 core genome SNPs of Philippine *Xoo* strains (Quibod et al., 2020). MDS analysis was performed on dissimilarity scores based on the amino acid substitutions found in previously reported 50 biofilm-related proteins of Xanthomonadales using the R package in bios2mds (Pelé et al., 2012). BAPS (L. Cheng et al., 2013) clustering was conducted using Rhierbaps (Tonkin-Hill et al., 2018). A. MDS plot showing two principal components, PC1 and PC2. B. A 3D rendering of the MDS analysis showing PC1, PC2, and PC3.

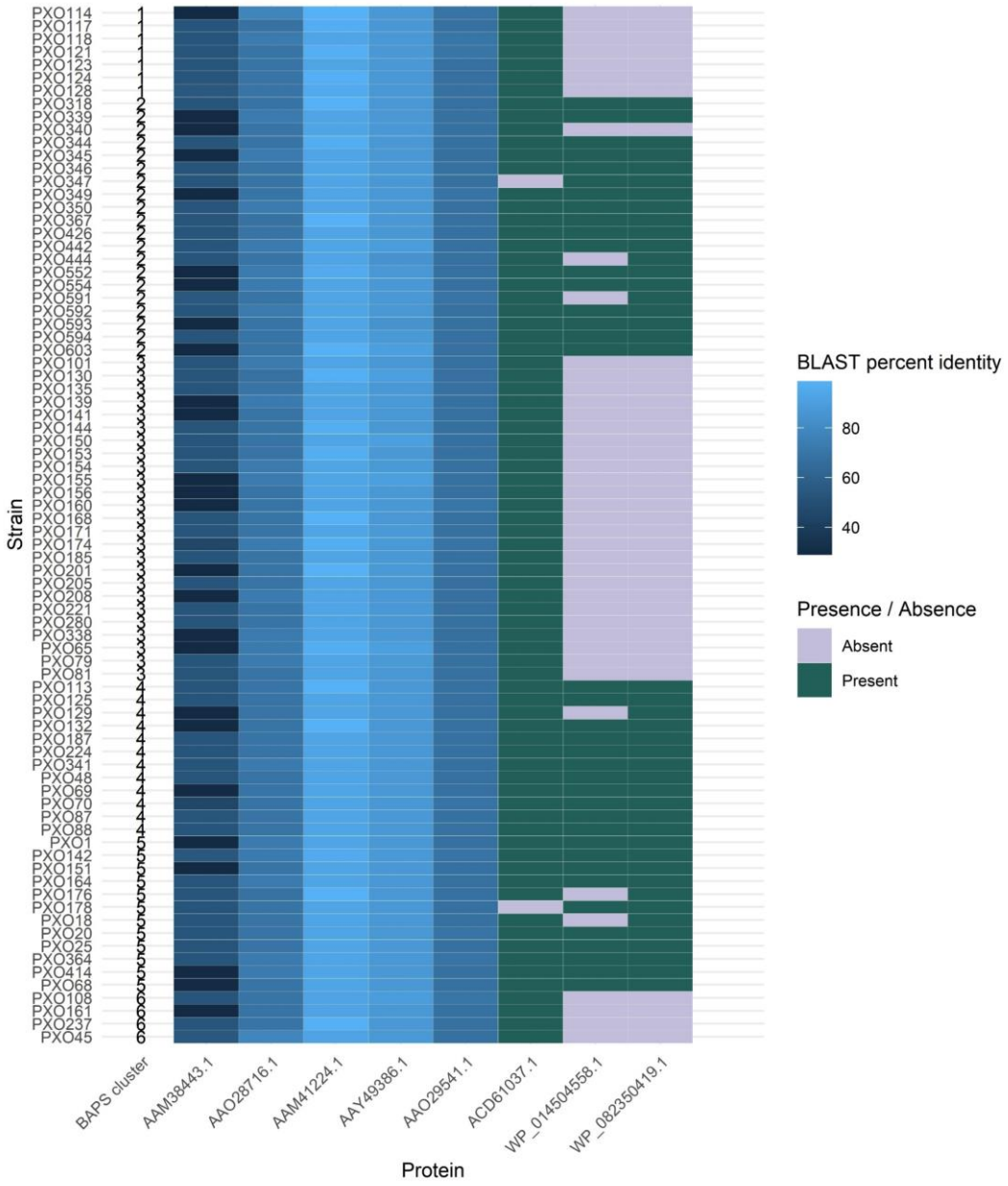


Figure 3.4. Heatmap visualization of Xanthomonadales biofilm-related proteins that showed variability in terms of sequence and presence or absence in the genomes of 80 Philippine *Xanthomonas oryzae* pv. *oryzae* strains. The column after the strain labels indicate the clusters obtained from Bayesian analysis of population structure (BAPS) inferred from 12,904 core genome SNPs of Philippine *Xanthomonas oryzae* pv. *oryzae* strains (Quibod et al., 2020). The first five proteins are those that varied in BLAST % identity, and the last three columns show the three proteins that varied in occurrence.

3.3.6 Multi-dimensional scaling (MDS) analysis from the alignment of 50 biofilm-related genes of Xanthomonads in 80 strains of Xanthomonas oryzae pv. oryzae may be related with Xoo virulence in IRBB4, IRBB5, IRBB7, and IRBB10

As biofilm formation is a virulence factor, R genes may be associated with the variability of biofilm-related proteins in Xoo. To test this hypothesis, the MDS plot was combined with pathotype data of each strain (Quibod et al., 2020; Quibod et al., 2016) and was used to associate the reaction pattern of BB R genes to the variation in biofilm-related genes (Figure 3.5).

In Figure 3.5, it can be observed that strains which clustered together based on based on amino acid dissimilarity scores tend to elicit the same reaction from near-isogenic lines carrying individual R genes. This was observed in near-isogenic lines IRBB4 (*Xa4*), IRBB5 (*xa5*), IRBB7 (*Xa7*), and IRBB10 (*Xa10*), but was not observed in IRBB13 (*xa13*), IRBB14 (*Xa14*), and IRBB21 (*Xa21*). If the quadrants are labelled as in a cartesian plane, starting from the top right corner, and proceeding in a counterclockwise direction, one can denote the quadrants as I, II, III, IV. In IRBB4, IRBB5, IRBB7, and IRBB10, most of the virulent strains are in quadrants III and IV. In contrast, the avirulent strains in IRBB13 are in quadrant III. IRBB21 exhibits resistant to almost all the strains used in this study.

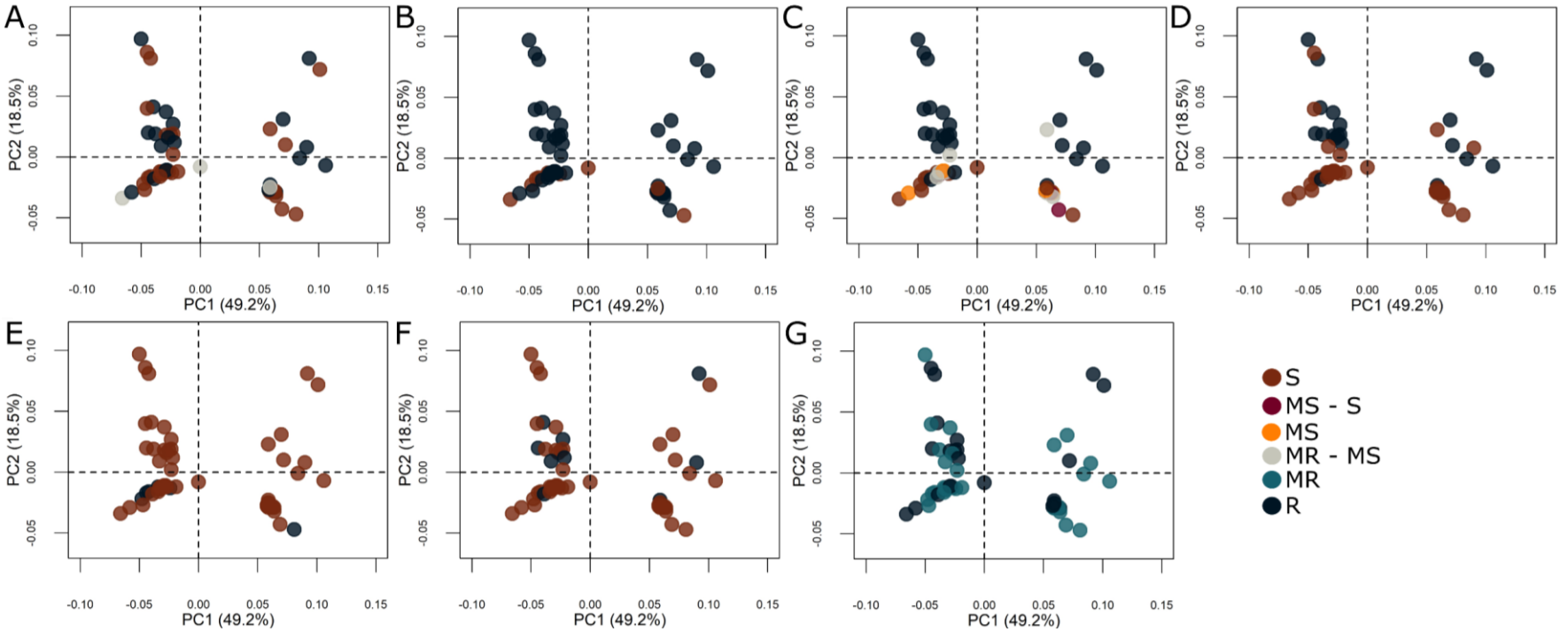


Figure 3.5. Multi-dimensional scaling (MDS) analysis from the alignment of 50 biofilm-related genes of Xanthomonads in 80 strains of *Xanthomonas oryzae* pv. *oryzae*. MDS analysis was performed in bios2mds (Pelé et al., 2012) with the function *mat.dis* where a dissimilarity matrix was calculated using PAM250 amino acid substitution matrix and visualized using the function *mmds.2D.plot.* The first panel shows the output of the analysis. Subsequent panels show the MDS plot with colours that represent the elicited reaction in rice lines introgressed with

bacterial blight resistance genes. A: IRBB4 (*Xa4*), B: IRBB5 (*xa5*), C: IRBB7 (*Xa7*), D: IRBB10 (*Xa10*), E: IRBB13 (*xa13*), F: IRBB14 (*Xa14*), G: IRBB21 (*Xa21*). The reactions S (susceptible), R (resistant), MS (moderately susceptible), MR (moderately resistant) are based on race classification of the strains (Quibod et al., 2016).

3.4 Discussion

This study aimed to obtain data on the occurrence and variation of previously identified biofilm-related genes of Xanthomonads in the genomes of Philippine *Xoo* strains and identifying genes that are potential drivers of *Xoo* evolution. The panel of genes obtained from published literature represents a diverse set of pathogenic Xanthomonads that include *Stenotrophomonas maltophilia, X. axonopodis* pv. *glycines, X. campestris* pv. *campestris*, *X. citri* subsp. *citri* (formerly *X. axonopodis* pv. *citri*), *Xoo, X. oryzae* pv. *oryzicola*, and from *Xylella fastidiosa*. These genes were first transcribed into proteins and then were used to query a dataset of 80 strains collected from different areas in the Philippines between 1972 to 2010 (Table A.2) and have been previously characterized (Quibod et al., 2020).

The results indicate that biofilm-related proteins are conserved in a diverse set of Philippine *Xoo* strains. Out of the 50 homologs found in 80 *Xoo* strains from the Philippines, 90% were conserved across the 80 strains. This suggests that biofilm-formation is indeed an essential process in *Xoo*. The six proteins (AAO28818.1, AEQ97862.1, AEQ97874.1, AEQ97875.1, AJD70189.1, CAJ25451.1) that were not found in any of the Philippine *Xoo* genomes used in this study were characterized in *Xanthomonas* sp. apart from *Xoo*: *X. oryzae* pv. *oryzicola, X. fastidiosa,* and *X. citri* subsp. *citri*. Three of these were characterized in *X. oryzae* pv. *oryzicola* and may be specific to this pathovar. This is corroborated by a previous report (Y. Zhao et al., 2012) that AEQ97862.1 (*hshB,* from *X. oryzae* pv. *oryzicola*) was not found in *Xoo* strain PXO99A reference genome. Similarly, AEQ97874.1 (*xagA,* from *X. oryzae* pv. *oryzicola*) and AEQ97875.1 (*xagB,* from *X. oryzae* pv. *oryzicola*) were also previously not found in *Xoo* (Y. Zhang et al., 2013). The remaining two proteins may be specific to *X. fastidiosa,* and *X. citri* subsp. *citri*, respectively.

However, sequence variability and presence/absence variability were observed in a minor set of proteins. Aside from biofilm formation, each of the five proteins with observed sequence variability based on BLAST (AAM38443.1, AAO28716.1, AAM41224.1, AAY49386.1, AAO29541.1) have all been linked to virulence mechanisms. All four of these were also linked to motility except for AAY49386.1 (*rpfC, X. campestris* pv. *campestris*). Interestingly, AAM38443.1 (*wzt, X. axonopodis* pv. *citri*) is an O-antigen ABC transporter and has been proposed to be a PAMP

involved in triggering the innate immune response of citrus during invasion by *X. axonopodis* pv. *citri* (Casabuono, Petrocelli, Ottado, Orellano, & Couto, 2011). Of the three proteins with variability in terms of presence absence in the 80 Philippine *Xoo* genomes, two were characterized from *X. oryzae* pv. *oryzicola* (Song et al., 2015). Interestingly, the third protein was characterized in *Xoo* strain PXO99A (B. Wang, Wu, Zhang, Qian, & Liu, 2018). WP_082350419.1 (*hshA*, from *X. oryzae* pv. *oryzicola*) and WP_014504558.1 (*hshC*, from *X. oryzae* pv. *oryzicola*) were not found in the reference genome PXO99A, similar to previous reports (Song et al., 2015; Y. Zhao et al., 2012), but were found in about half of the Philippine *Xoo* genomes used in this study as well as in *Xoo* reference genome KACC10331, with a percent identity of 83%.

The observed variations of these proteins may be linked to *Xoo* adaptation mechanisms to the host, while the results of the partition analysis could hint that these mechanisms may be independent. As these proteins are related to a diverse set of functions and involved in the biofilm formation process (Table A.1), it is interesting to explore the relationship of the protein diversity of biofilm-related proteins to the *Xoo* population structure. The population structure of Philippine *Xoo* strains has been estimated in previous work by Quibod et al (Quibod et al., 2020), using BAPS clustering (L. Cheng et al., 2013) of 12904 core genome SNPs from Philippine *Xoo* strains. Based on this, there are six *Xoo* populations in the Philippines, which was linked to *Xoo* adaptation caused by extensive deployment of the R gene Xa4. A repeat of the BAPS clustering in this study using the R package Rhierbaps (Tonkin-Hill et al., 2018) also yield 6 clusters. Multivariate analysis is often employed to simplify the visualization of the relationships of variables in complex datasets (Ivosev, Burton, & Bonner, 2008). In this study, when the coordinates obtained from the MDS analysis of concatenated amino acid alignments of the 50 biofilm-related proteins in Philippine *Xoo* genomes was overlayed with the phylogenetic data from the BAPS clusters revealed that most of the strains that grouped together based on MDS also belonged to the same BAPS cluster. Moreover, two proteins WP_082350419.1 (*hshA*, from *X. oryzae* pv. *oryzicola*) and WP_014504558.1 (*hshC*, from *X. oryzae* pv. *oryzicola*) which varied in terms of their occurrence in the 80 Philippine *Xoo* genomes were strongly associated with the BAPS clusters. These observations suggest that there is indeed a link between variation of biofilm-related proteins and *Xoo* adaptation to rice.

Because the resistance or susceptibility of the IRBB near-isogenic lines is controlled by the specificity of R-Avr gene interactions, the observed clustering based on MDS analysis of strains which elicit similar reactions from IRBB4, IRBB5, IRBB7, and IRBB10 was not expected and must be investigated further in future studies. Virulence and persistence characteristics gained through biofilm formation could play a role in successful colonization of hosts with specific R genes.

Biofilm formation has been associated with virulence. In the enteric bacterium and human pathogen *Proteus mirabilis*, biofilm formation was determined to be significantly associated with higher expression levels of virulence genes (Sun et al., 2020). The same has been reported in plant pathogens. For instance, in *X. axonopodis* pv. *citri*, the BfdR protein (biofilm formation defective regulator) has been demonstrated to be important in leaf colonization and development of disease symptoms (T.-P. Huang et al., 2013). Coordinated expression of biofilm genes and virulence genes is enhanced by global regulators which control a myriad of inter-related phenotypes, such as the protein RpfG (Y. Zhang et al., 2013), which was shown to regulate T3SS expression, biofilm formation, and EPS production in *X. oryzae* pv. *oryzicola*. In clinical pathogens, the formation of biofilms is linked to persistence (Soares, Alexandre, & Etienne, 2020), a survival mechanism wherein bacteria enter a dormant state when exposed to unfavorable environmental conditions (Eisenreich, Rudel, Heesemann, & Goebel, 2021). Evidence of this relationship has also been reported in some plant pathogens such as *Xylella fastidiosa* (Martins, Merfa, Takita, & De Souza, 2018). In *Xoo*, inhibition of biofilm formation has been reported as a potential means to control BB (S.-I. Kim et al., 2016; Sahu et al., 2018).

This study employed a simple bioinformatic approach to studying the genetic diversity of biofilm-related genes in *Xoo* by first compiling biofilm-related proteins in published research that were identified by mutagenesis and phenotyping. The results show that while most of the proteins are conserved in *Xoo*, there is evidence of adaptation as observed in some of the proteins with variability across Philippine *Xoo* populations and may be linked to *Xoo* adaptation to growth in rice. This diversity must be explored further in future studies. Moreover, the potential role of biofilm formation in the rice-*Xoo* interaction should be pursued further as a possible source of alternative mechanisms for BB management and control.

CHAPTER 4: EXPLORING THE ADAPTATION POTENTIAL OF *Xanthomonas oryzae* pv. *oryzae* IN RICE THROUGH SERIAL PASSAGE IN BACTERIAL BLIGHT RESISTANT RICE LINE IRBB4 (*XA4*).

4.1 Introduction

Selection pressure drives the constant adaptation of bacteria (M. Kang et al., 2019), a concept in biology known as evolutionary adaptation (Waddington, 1959). Numerous studies report on host-induced adaptation of plant pathogens and plant-associated bacteria (C.-L. Huang et al., 2015; R. A. Melnyk et al., 2019; Quibod et al., 2020; Quibod et al., 2016), where variation is acquired primarily through mutations (Stukenbrock & Bataillon, 2012). Mutations that incur a lower fitness cost are more likely to be maintained by a population (A. H. Melnyk et al., 2015).

Adaptive laboratory evolution (ALE) experiments are often employed to study microbial adaptation in controlled conditions which allows the observation of adaptation in a step-wise manner, also known as experimental evolution (Kawecki et al., 2012) experiments. In ALE experiments, variants with improved fitness can be characterized, exhibiting increased growth rates, modified stress responses, or specialization of nutrient metabolism as well as modifications in membrane proteins (Choe et al., 2019; Dettman et al., 2012; Dragosits & Mattanovich, 2013; Horinouchi et al., 2015; M. Kang et al., 2019). Whole genome sequencing of variants with phenotype data enables an in-depth analysis of the genes that are possibly involved in the variations in phenotype (Dettman et al., 2012).

As *Xoo* is considered an important pathogen of rice due to the potential for immense yield losses caused by BB, various aspects of the pathogen biology and its interaction with rice has been studied. Population studies of *Xoo* have shown that this pathogen has developed numerous adaptation mechanisms in response to rice (Mew et al., 1992; Quibod et al., 2020; F. Zhang et al., 2021). These adaptations are attributed to the selection pressure caused by extensive adoption of resistant rice varieties since the Green Revolution (Mew et al., 1992; Quibod et al., 2020; F. Zhang et al., 2021). The interaction between rice and *Xoo* is known to follow the gene-for-gene (Flor, 1971) model, where the induction of susceptibility (Antony et al., 2010; L.-Q. Chen

et al., 2010; Streubel et al., 2013) or resistance (Hopkins, White, Choi, Guo, & Leach, 1992) in rice is caused by the specific interaction of genes or gene products from both *Xoo* and rice.

In this study, Philippine *Xoo* strain PXO61 was passaged in resistant rice line IRBB4, which contains the BB resistance gene Xa4 (Ogawa et al., 1991). Xa4 was selected in this study because it has been linked to the emergence of more virulent *Xoo* populations in the Philippines following widespread deployment of Xa4, a consequence of adaptation to this resistance gene (Quibod et al., 2020). Varieties that carry Xa4 are known to be resistant to PXO61. This study aimed to characterize *Xoo* after serial passage in rice using a combination of disease phenotyping, metabolic profiling, transcriptome analysis, and genomic analysis. These procedures were done to test if *Xoo* adaptation can be induced by serial passage in rice, and to observe for any small, incremental changes in *Xoo* as a response to growth in rice.

4.2 Methodology

An overview of the methods used in this study is shown in Figure 4.1.

4.2.1 Strain growth and storage

Philippine *Xoo* strain PXO61 was revived from the original lyophilized stock stored in an ampoule. The strain was grown on modified Wakimoto's agar (WF-P) (Karganilla, Natural, & Ou, 1973) and allowed to recover from any injury it may have incurred during storage by aerobic incubation at 30 ˚C for 4 to 5 days. After appearance of typical the typical yellow, mucoidal, *Xoo* colonies, these were sub-cultured onto WF-P plates and incubated for 24 to 48 h at 30 ˚C.

Working stocks for subsequent experiments were prepared from 24 h-old cultures subcultured from skim milk stocks and incubated at 30ºC. Single, isolated colonies were then selected and streaked onto WF-P agar for multiplication.

4.2.2 Inoculations into plants

For all inoculations, 24 to 48 h old *Xoo* cultures grown on WF-P agar were suspended in peptone broth and homogenized by vortex-mixing at maximum setting. Cell suspensions were adjusted to OD_{600} = 0.9 and used for clip-inoculations (Kauffman, Reddy, Hsieh, & Merca, 1973) or infiltration with a needleless syringe.

Clip-inoculations were performed using sterilized stainless-steel scissors dipped in inoculum suspension contained in a 50-ml centrifuge tube. After dipping in

inoculum, about 1 to 2 cm from the tips of the youngest fully expanded leaves or pre-identified leaf positions were cut, introducing inoculum into the wound. The scissors were dipped in inoculum each time a new leaf was cut. A fresh pair of scissors were used per treatment.

4.2.3 Serial passage experiment

Inoculation was done using the clipping method (Kauffman et al., 1973) on 21 day old rice lines carrying BB R genes Xa4, xa5, and Xa7 (IRBB4, IRBB5, and IRBB7, respectively) grown in a growth chamber with 29/21 ˚C day/night temperature setting. Only the second leaf position, counting from the top leaf, was inoculated per plant. Twenty-four plants were inoculated in each batch. Each batch of inoculation will hereafter be referred to as a cycle. Lesion lengths were recorded 11-days post-inoculation.

Leaves with lesion lengths closest to the mean lesion length were selected, and processed for re-isolation following the standard isolation procedure at IRRI: after surface disinfection of the leaves by dipping for a few seconds in 70% ethanol and rinsing in sterile distilled water, the leaves were cut into small pieces using a sterile pair of scissors, and soaked in 1X phosphate buffered saline solution (PBS, NaCl 8 g L^{-1}, KCl 0.2 g L^{-1}, Na_2HPO_4 1.44 g L^{-1}, KH_2PO_4 0.24 g L^{-1}). A loopful of the suspension was streaked onto pre-plated WF-P agar and incubated at 30 ˚C. Isolated colonies were randomly picked using a sterilized wire loop and transferred into a fresh plate of WF-P agar. Subcultures of the purified isolate were stored in skim milk (-20 ˚C), with working stocks prepared for the subsequent round of inoculation. Only one purified isolate per round of inoculation was stored and included in the subsequent inoculation for practical reasons. All strains isolated from leaves were preserved in skim milk as described previously (Vera Cruz et al., 2017).

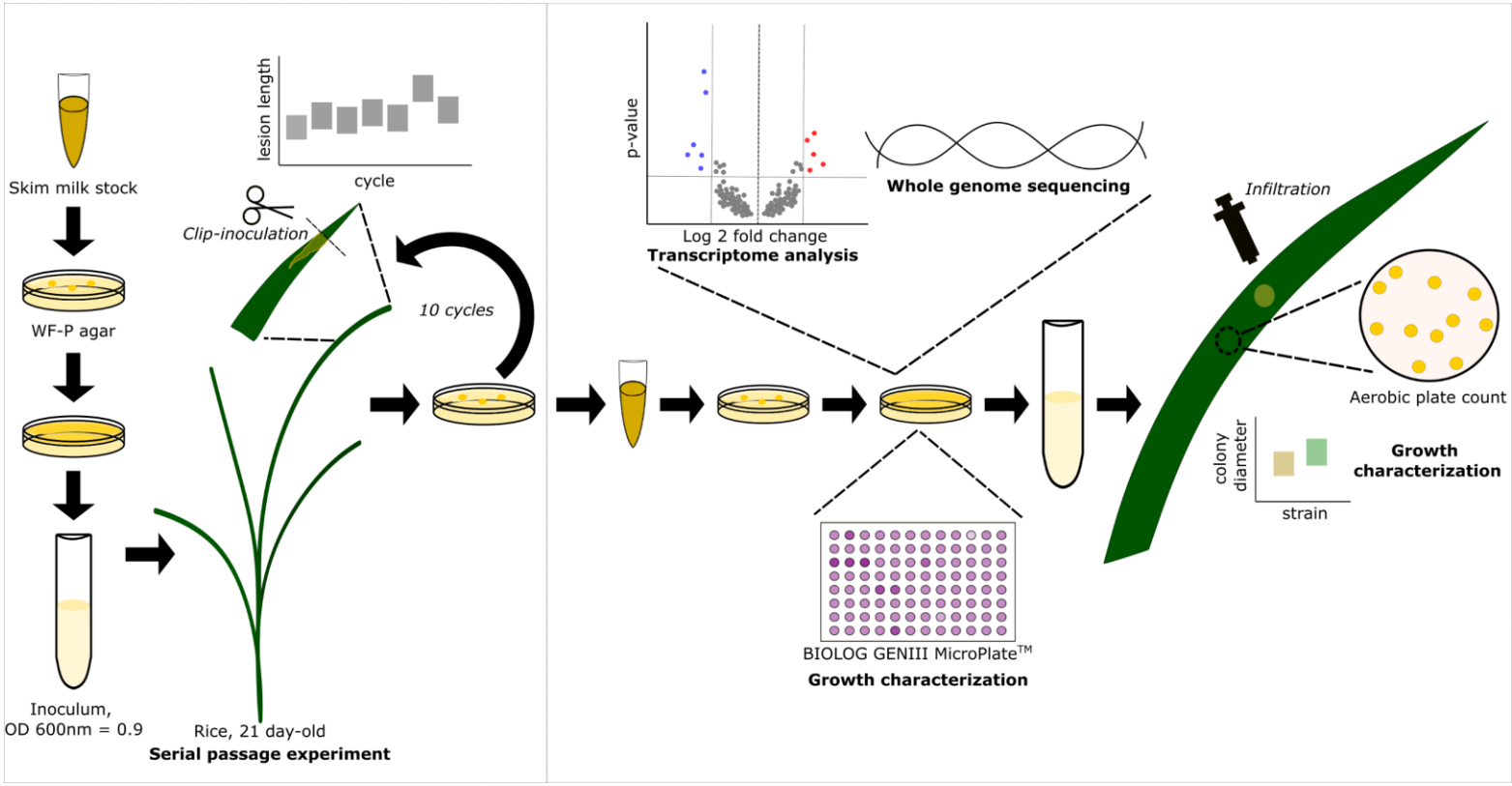


Figure 4.1. Schematic overview of the methods used in studying the adaptation potential of *Xanthomonas oryzae* pv. *oryzae* (*Xoo*). The left panel illustrates the serial passage of *Xoo* strain PXO61 in IRBB4. A total of ten passages or cyclic inoculations were performed. The characterization experiments which include transcriptome analysis, whole genome sequencing, and growth characterizations conducted on both the wild-type strain (designated as PXO61-0) and the variant from the sixth cycle are shown in the right. The image was created using Inkscape 1.1 (https://inkscape.org/).

Clip-inoculations in the greenhouse were done in a similar manner. Rice variety IR24 and near-isogenic line IRBB4 were inoculated with PXO 61-0 and PXO 61-4 in a split plot-RCBD with variety/line as main plot and isolates as subplot with replications (culture batch) as blocking factor. Five replications of each treatment were set up, with each replicate composed of three hills. The lesions formed on each inoculated leaf were be recorded at 5-, 7-, 9-, and 11-days post-inoculation. Measurement units were expressed as mean lesion length. Comparison of aggressiveness in terms of mean lesion lengths were conducted using a linear mixed model with the area under disease progression curve (AUDPC) as the response variable (Schandry, 2017). Wilt data (Table B.17) was collected by recording all tillers with wilted leaves.

4.2.4 High-throughput metabolic profiling in microplates

The inoculum suspension in IF-A (BIOLOG, Cat. No. 72401) was prepared according to manufacturer's specifications. Each well of the GENIII microplate™ was inoculated with 100 µl of the cell suspension and incubated at 30°C. Hourly absorbance measurements at 600nm were collected using a multi-mode plate reader (FLUOstar OPTIMA, BMG LABTECH) for 36 h. The *SummarizeGrowthByPlate* function of the R package growthcurver (Sprouffske & Wagner, 2016) was used to process the data in plate format. This function estimates growth parameters for each microplate well using the collected absorbance data. Comparison of PXO61-0 and PXO61-4 growth characteristics was done by comparing the areas under the logistic growth curve generated by the function.

4.2.5 Within host multiplication and colony diameter of PXO61-0 and PXO61-4

Disc samples (1 mm) from inoculated leaves infiltrated with a needleless syringe were collected using a sterile single-hole puncher at 24- and 48-hours post-inoculation (hpi). These were obtained 2-cm below the visible lesion and kept in a sterile microtube on ice for transport to the laboratory. Samples were then suspended in PBS buffer. Selected ten-fold dilutions were plated on WF-P agar for aerobic plate count. Plates were incubated at 30°C for 72h. Estimated plate counts were calculated using the formula (Maturin & Peeler, 2001):

$$N = \frac{\sum colonies}{[(1 \times n1) + (0.1 \times n2) \times d]}$$

where N is the CFU mL^{-1}, n1 and n2 correspond to the number of plates in each dilution that has 25 to 250 colonies, and "d" is the dilution.

For colony diameter measurements, randomly selected colonies were viewed 3 to 4 d after incubation with an Olympus SZX16 Stereo Microscope and photographed with a QImaging Go-21 camera. Colony diameter was measured using the QCapture Pro software.

4.2.6 Pathogen transcriptome on agar

Strains were streaked onto WF-P agar and incubated at 30°C for 24 h. The growth was then scraped and fixed onto GE Whatman® FTA® WB120365 PlantSaver Cards. Sections of the cards were cut using sterilized scissors and forceps and placed onto sterilized microtubes. An aliquot of 60 – 100 µl of RNA Rapid Extraction Solution (AM9775, Ambion, USA) was added to each sample. The samples were then purified using Quick-RNA Plant Miniprep Kit (Zymo Research, USA) and sent to BGI (China) for transcriptome sequencing.

The quality of the RNA-seq reads was checked using FastQC (Andrews, 2010), and any adapter sequences trimmed with Trimmomatic (Bolger, Lohse, & Usadel, 2014). Reads were then aligned to the PXO99A genome (Salzberg et al., 2008) using the R package Rsubread (Y. Liao, Smyth, & Shi, 2019). Differential expression analysis was performed using DESeq2 (Love, Huber, & Anders, 2014).

4.2.7 Whole genome sequencing and variant calling

Strains were grown in Nutrient Broth (BD Difco™) overnight at 28°C with shaking (100 rpm). DNA extraction was performed using Easy-DNA™ gDNA purification kit (Invitrogen) following the manufacturer's specifications for extraction of DNA from bacterial cells. Purified DNA samples were then submitted for PacBio Single Molecule, Real-Time (SMRT) whole genome sequencing (Icahn School of Medicine, Mount Sinai). The raw sequence reads were mapped to PXO99A (Salzberg et al., 2008) using minimap2 (Heng Li, 2018), and variant calling was done using snippy (Seemann, 2015). Putative variant sites were inspected using the *pileLettersAt* of the R package GenomicAlignments (Lawrence et al., 2013).

4.2.8 Data analysis

All statistical calculations and comparisons were performed in R (R Core Team, 2021). Mean comparisons of continuous variables were performed using the Anova function of the car (Fox & Weisberg, 2019) package. In cases where the assumptions of ANOVA were not fulfilled, the Wilcoxon test was used to compare treatments. For this, the compare_means function of the R package ggpubr was used (Kassambara, 2020). All the tests are summarized in Table B.1..

4.3 Results

4.3.1 Serial passage of Xoo strain PXO61 in the resistant lines showed increased aggressiveness

To investigate how *Xoo* persists on hostile conditions within the resistant plants, an ALE setup involving the cyclic inoculation, re-isolation, and re-inoculation of Philippine *Xoo* strain PXO61 in 21-day old IRBB4 rice lines was performed. Evaluation of the lesion lengths obtained from inoculated leaves showed progressive increase in aggressiveness of the pathogen from the sixth inoculation cycle up to the tenth cycle (Figure 4.2A). Hence, the re-isolated *Xoo* strain from the 6^{th} cycle was selected for subsequent experiments. Figure 4.1 shows an overview of the serial passage experiment as well as the subsequent comparative analysis of the variants. In this setup, aggressiveness was inferred by measuring the lesion lengths caused by the pathogen 11 days after inoculation of the leaves. Lesion lengths obtained from this setup ranged from 0.3 cm to 8.1 cm. The overall mean lesion length obtained from the ten inoculation cycles is 3.5 ± 1.7 cm. Treatment means and summary statistics of all the cycles are listed in Table B.2. Cycles 1, 2, and 4 were determined to have treatment means lower than the overall mean, and cycles 6 – 10 significantly higher than the overall mean. Cycles 3 and 5 did not vary significantly from the overall mean. The results of the mean comparison are shown in Table B.3. However, serial passage of PXO61 in the susceptible background IR24 (Figure 4.3A) showed random shifts in mean lesion lengths. Similar to passage in IRBB4, a pattern of increased aggressiveness was observed in resistant rice lines IRBB5 (Figure 4.3B), and IRBB7 (Figure 4.3C), suggesting that the increase in aggressiveness is linked to growth in resistant rice genotypes. The summarized data and mean comparisons can be found in Table B.4 to Table B.9. From this point forward in this paper, the wild-type strain used

in the first inoculation event is denoted as PXO61-0, and the selected variant strain from the 6th cycle as PXO61-4 (PXO61 passed through IRBB4).

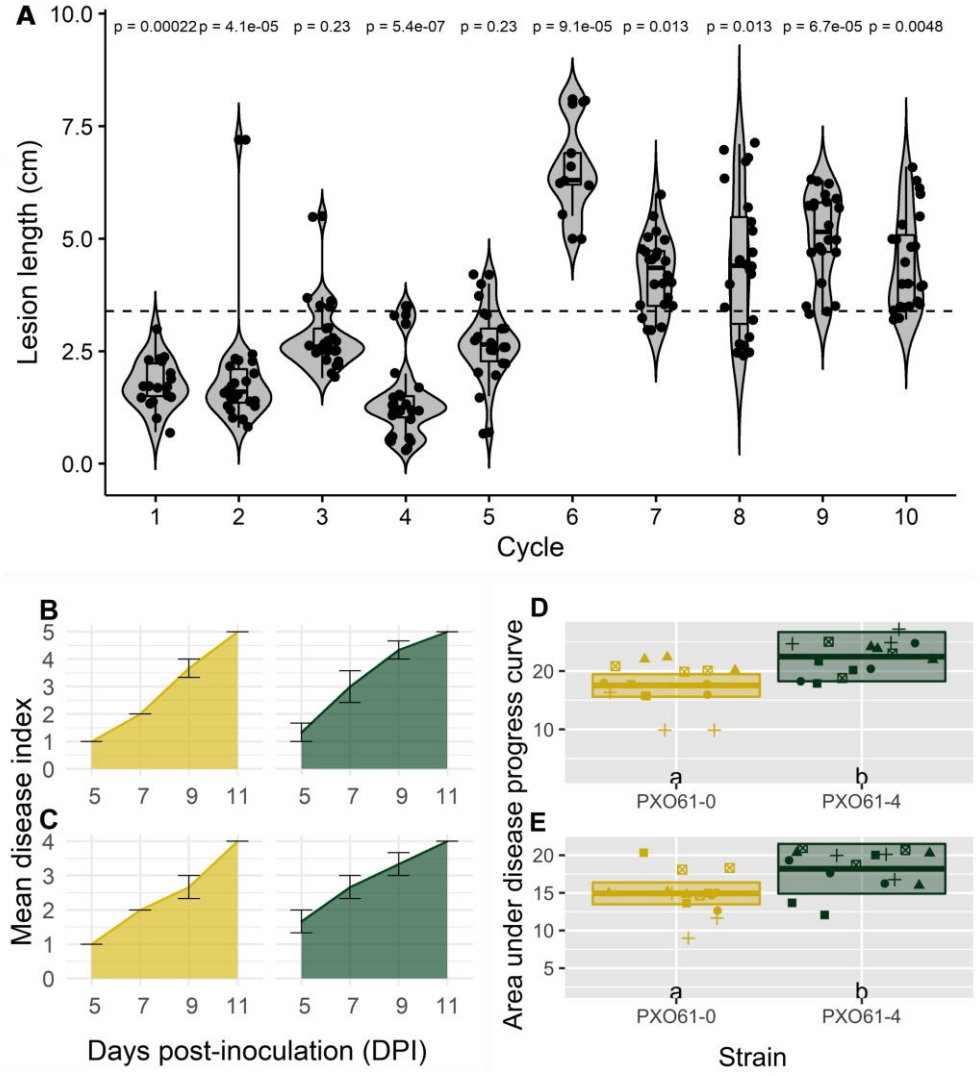


Figure 4.2. *In planta* fitness of *Xanthomonas oryzae* pv. *oryzae* PXO61. A. Lesion length data obtained from the inoculation of PXO61 into resistant rice variety IRBB4 in a serial passage experiment. Subsequent experiments used the strain from the 6th cycle, designated as PXO61-4. For clarity, the wild-type strain is designated as PXO61-0. B – E. Comparisons of the area under disease progress curve (AUDPC) of PXO61-0 and PXO61-4 in IR24 and in IRBB4. Lesion lengths were measured at 5-, 7-, 9-, and 11-days post-inoculation (DPI) and were used to estimate the AUDPC. B and C show the estimated disease progress over time for one replication in IR24 and IRBB4, respectively. Yellow: PXO61-0, Green: PXO61-4. D and E show the comparisons of the AUDPC for all the replications in IR24 and IRBB4, respectively.

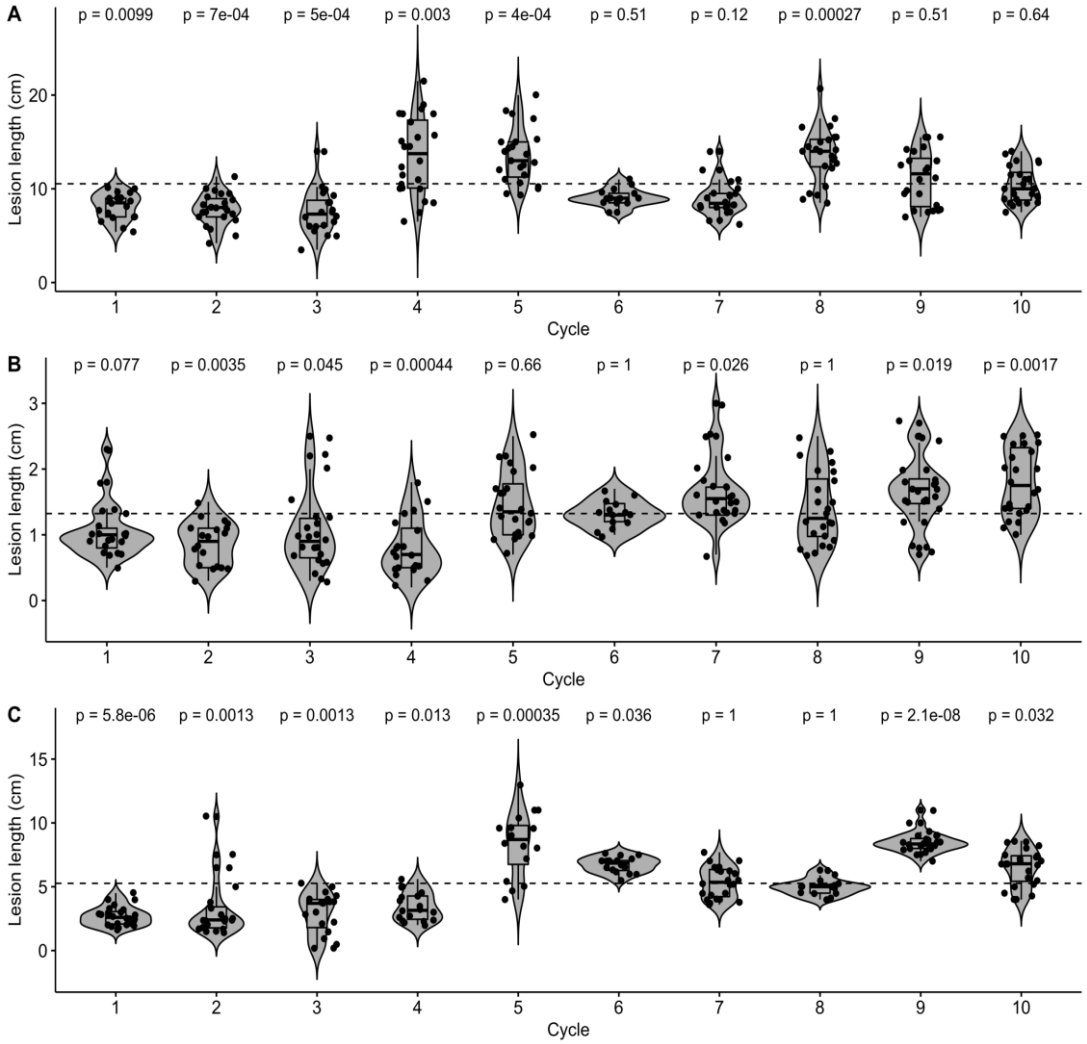


Figure 4.3. Fitness of PXO61 *in planta* as estimated by lesion lengths obtained from the serial passage of PXO61 in 21-day old plants of susceptible variety IR24 (A), and resistant lines IRBB5 (B), and IRBB7 (C). Shown are data from ten cycles of inoculation, re-isolation, and re-inoculation. Cyclic inoculations of each variety/line were performed independently with a minimum of four replications under controlled conditions in the growth chamber (29/21 ℃ day/night). The overall mean of all cycles for each variety/ line is shown by the dotted horizontal line. P-values were determined by comparing the treatment means to the overall mean by the student's t-test.

4.3.1 PXO61-4 shows increased fitness in planta

To validate the difference in aggressiveness observed in the cycling experiment, PXO61-0 and PXO61-4 were inoculated under greenhouse conditions and the area under disease progress curve (AUDPC) (Madden, Hughes, & Van Den Bosch, 2007) was calculated. In this controlled setup, the increased aggressiveness of PXO61-4 variant in vascular rice tissues was confirmed (Figure 4.2B - E). Variants were inoculated in resistance rice line IRBB4 and the susceptible background IR24 under greenhouse conditions, and lesion length data was collected at 5,7,9, and 11 dpi. The AUDPC values for variety IR24 (n = 30) ranged from 10 to 27, with a mean of 20 ± 4.1 across all the treatments. The disease progress caused by PXO61-4 (M = 22.5, 95% CI [26.7, 18.2]) is significantly higher (ANOVA, p =.0000172, Cohen's effect size = 1.2) than the disease progress caused by PXO61-0 (M = 17.5, 95% CI [19.5, 15.6]). Inoculations into IRBB4 (n = 30) yield an overall mean AUDPC of 16.6 ± 9.4, with values ranging from 9 to 21. The disease progress caused by PXO61-4 inoculation (M = 18.2, 95% CI [21.5, 14.9]) is significantly higher (ANOVA, p=.000421, Cohen's effect size = 1.06) than the disease progress caused by PXO61-0 inoculation (M = 14.9, 95% CI [16.4, 13.5]) (Figure 4.2D - E). In addition, Table B.10 and Table B.11 contains the summarized disease progression data. Table B.12 contains the comparison of mean AUDPC for all the treatments.

Severe infection with *Xoo* can cause wilting or *kresek* (T. Mew, A. Alvarez, J. Leach, & J. Swings, 1993). In the same greenhouse experiment, wilt was quantified at 17 dpi in susceptible IR24 plants inoculated with PXO61-0 and PXO61-4. About 16% of the tillers inoculated with PXO61-0 showed wilt (47 were observed with wilt out of 290 tillers). On the other hand, 53.8% of plants inoculated with PXO61-4 showed wilting (132 tillers were wilted out of 245 tillers). Wilting was found to be associated with IR24 plants inoculated with PXO61-4 (Figure 4.4) and the inoculation with PXO61-4 increased the odds of observing wilt by 6.01 compared to inoculation with PXO61-0 (Fisher's exact test for count data, $p = 1.78 \times 10^{-20}$, 95% CI = [3.97, 9.22]). This data is presented in Table B.17. Wilting was not observed in a parallel setup where PXO61-0 and PXO61-4 were inoculated into IRBB4.

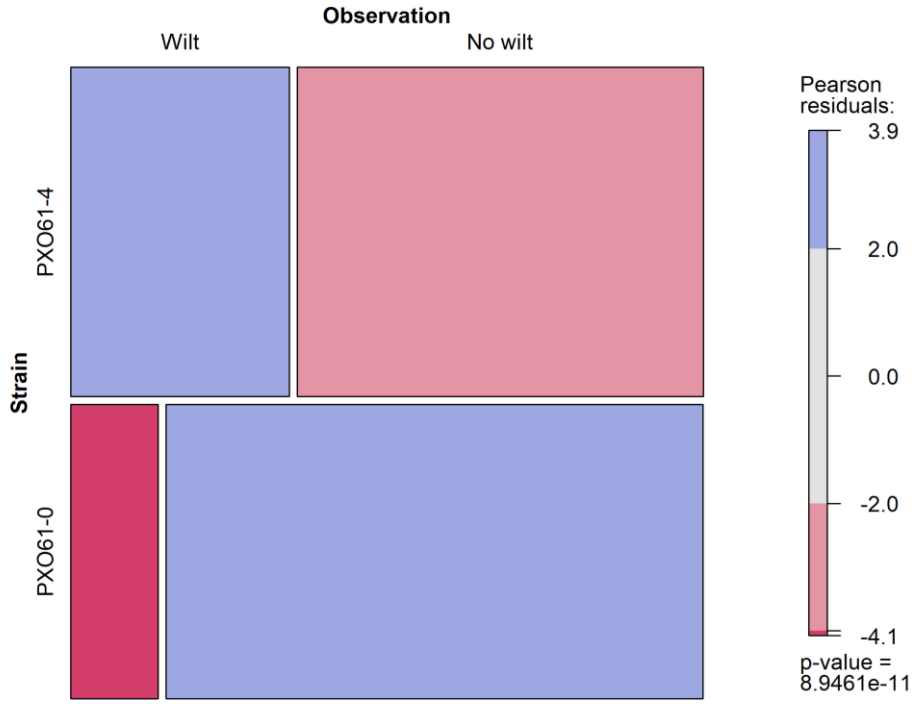


Figure 4.4. Comparison of observed wilt in tillers of IR24 plants inoculated with wildtype strain PXO61-0 and variant PXO61-4 which was passaged in IRBB4 for six cycles of inoculation, re-isolation, and re-inoculation.

4.3.2 PXO61-4 shows increased fitness in vitro

ALE experiments have been reported to produce mutants or variants that have increased growth rates (Ackermann, Schauerte, Stearns, & Jenal, 2007; Bachmann, Starrenburg, Molenaar, Kleerebezem, & van Hylckama Vlieg, 2012; Knöppel et al., 2018) hence increased fitness *in vitro*. To investigate variations of growth *in vitro*, the phenotype profiles (Figure 4.5A) of PXO61-0 and PXO61-4 were used in a microplate with 94 substrates (GEN III MicroPlate^{TM}, BIOLOG, USA). These phenotyping plates provided a way to grow the strains in different substrates, using the same inoculum. PXO61-4 yielded a higher area under the logistic curve than PXO61-0 (Wilcoxon rank sum test, $p = 6.79 \times 10^{-6}$), which is shown in Figure 4.5B. Table B.18contains the output of the SummarizeGrowthByPlate function of the R package growthcurver.

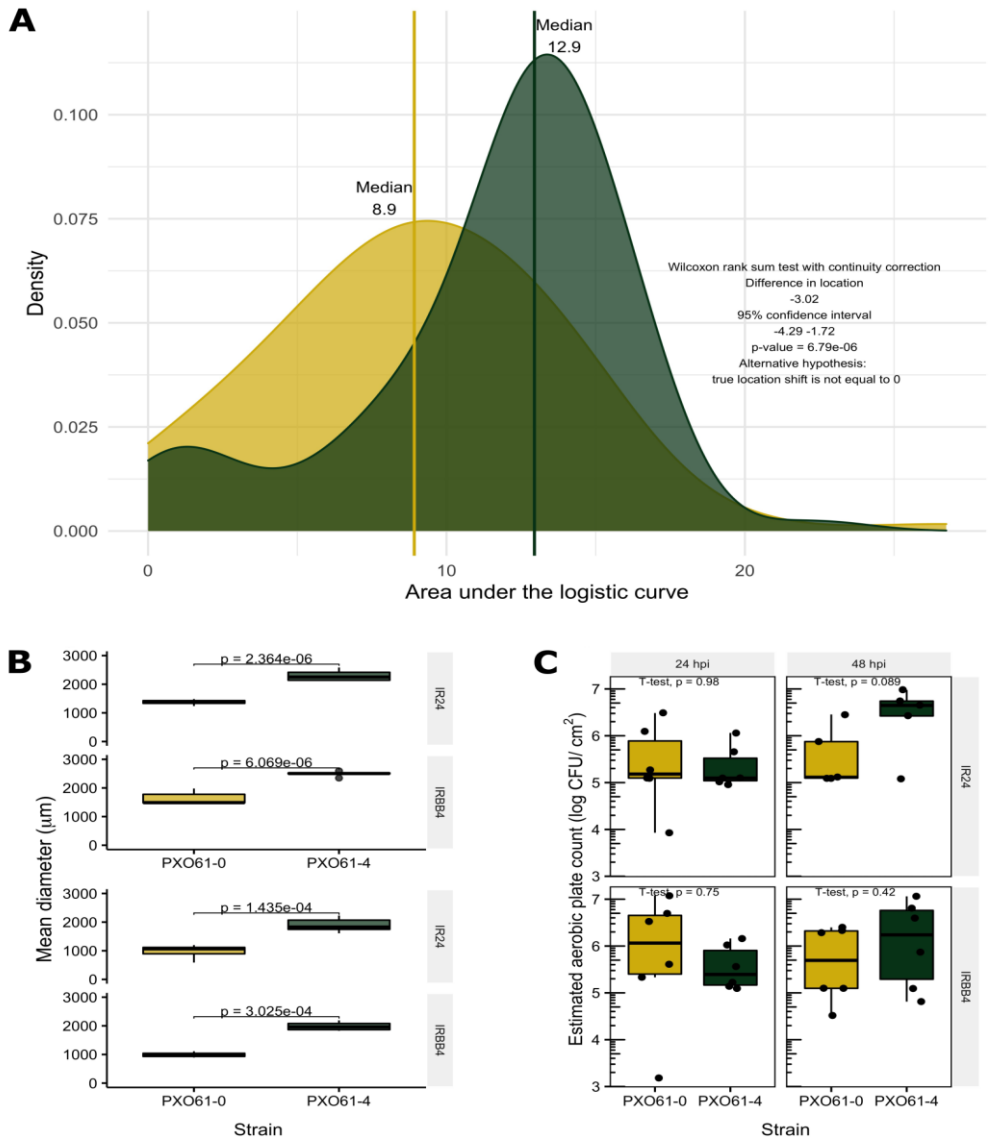


Figure 4.5. *In vitro* fitness of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4). A. Comparison of growth characteristics by fitting the OD data from hourly measurements of optical density at 450 nm for 36 h and fitting into a logarithmic growth model and calculating the area under the logistic curve using the R package growthcurveR (Sprouffske & Wagner, 2016). B. Colony diameter of strains re-isolated from IR24 and IRBB4 at 24 and 48 h after infiltration. Leaf discs (1 mm) were taken 2 cm below the margin of visible lesions that originate from the point of inoculation. Serial tenfold dilutions were plated on WF-P agar for observation under the dissecting microscope after 3 days of incubation at 30 °C. Colonies for diameter measurement were randomly selected by dividing the plate into a grid of 1 cm^2 squares, and

randomly selecting the square for microscopic observation. If more than one colony was observed within the square, the largest colony was selected. If there were no colonies within the square, a colony on the left, top, right, or bottom boundary was selected. C. Estimated aerobic plate count of PXO61-0 and PXO61-4 re-isolated from infiltrated leaves of 21-day old plants at 24 and 48 hpi. Leaf discs (1 mm) were taken 2 cm below the margin of visible lesions that originate from the point of inoculation and resuspended in sterile distilled water. Serial, tenfold dilutions were plated on WF-P agar. Colonies were counted after 3 days of incubation at 30 ℃.

In order to check if the variation reverted after passage through rice, PXO61-0 and PXO61-4 were inoculated into IRBB4 and IR24, re-isolated, and then plated on modified Wakimoto's agar [0.5 g of Ca(NO3)2·4H2, 0.82 g of Na2HPO4, 5 g of peptone, 20 g of sucrose, 300 g of potato 15 g agar per liter of water] (Karganilla et al., 1973). The variation in colony diameter of *Xoo* re-isolated from IR24 and IRBB4 is presented in Figure 4.5C and D. See also Figure B.1. In this experiment, the mean diameter of PXO61-4 (2303.3 ± 107.6 µm) colonies was observed to be higher than PXO61-0 colonies (1375.6 ± 30.7µm) that were re-isolated from IR24 at 24 hpi (*ANOVA*, $p = 2.364 \times 10^{-06}$, $\eta^2 = 0.92$, 95% CI [0.76, 0.96]). This variation was also observed in IRBB4, where re-isolated colonies of PXO61-4 (2494.4 ± 33.3 µm) showed a higher mean diameter than PXO61-0 colonies (1631.1 ± 94.4 µm) recovered at 24 hpi (*ANOVA*, $p = 6.069 \times 10^{-06}$, $\eta^2 = 0.88$, 95% CI [0.66, 0.94]). A similar trend was observed in colonies re-isolated at 48 hpi. Table B.20 to Table B.23 contains the summary data of all the treatments and the results of the statistical analysis. Estimation of growth *in planta* was done by plating sub-samples from inoculated leaves. No significant difference was observed in the estimated bacterial counts of PXO61-0 and PXO61-4 inoculated into IR24 and IRBB4 recovered at 24 and 48 hpi (Figure 4.5E). The summarized estimated plate count data and statistical analyses can be found in Table B.25 to Table B.28.

4.3.3 Whole genome sequencing of PXO61-0 and PXO61-4 suggests an adaptation mechanism involving maintenance of variability

To investigate sequence variation between PXO61-0 and PXO61-4, whole genome sequencing (WGS) reads were mapped to the PXO99A genome, and two genomic coordinates of potential interest were identified. Figure 4.6 outlines the

pipeline we applied to the WGS data and the result that is produced for each step. Initially, we identified 694 and 931 unique coordinates for PXO61-0 replicates 1 and 2. Variant calls for PXO61-4 reads of technical replicates 1 and 2 were 713 and 182, respectively. Each replicate was analyzed independently until the visualization step, where the final two coordinates of interest were identified, 789258 bp and 3108141 bp. Figure 4.7A-B shows the visualization of each coordinate in PXO61-0 and PXO61-4, as well as in the WGS reads of other variants from the passage of PXO61-0 into IRBB5 (PXO61-5) and IRBB7 (PXO61-7). Further investigation of these positions via inspection of aligned RNA-seq reads confirms the T to G mutation in position 3108141 bp but not the mutation in 789258. Interestingly, a small proportion (2%) of reads in position 3108141 of PXO61-0 are called as G and a proportion (10%) of reads in that same position are called as T in PXO61-4 (Table A.1). This suggests that the variants co-exist and a shift in the dominant variant is induced by certain environmental stimuli. The SNP is located within PXO_01777 (cyclopropane-fatty-acyl-phospholipid synthase), which produces a non-synonymous mutation from isoleucine to serine in the non-binding region of the predicted protein, which influenced a change in the predicted structure (Figure 4.7).

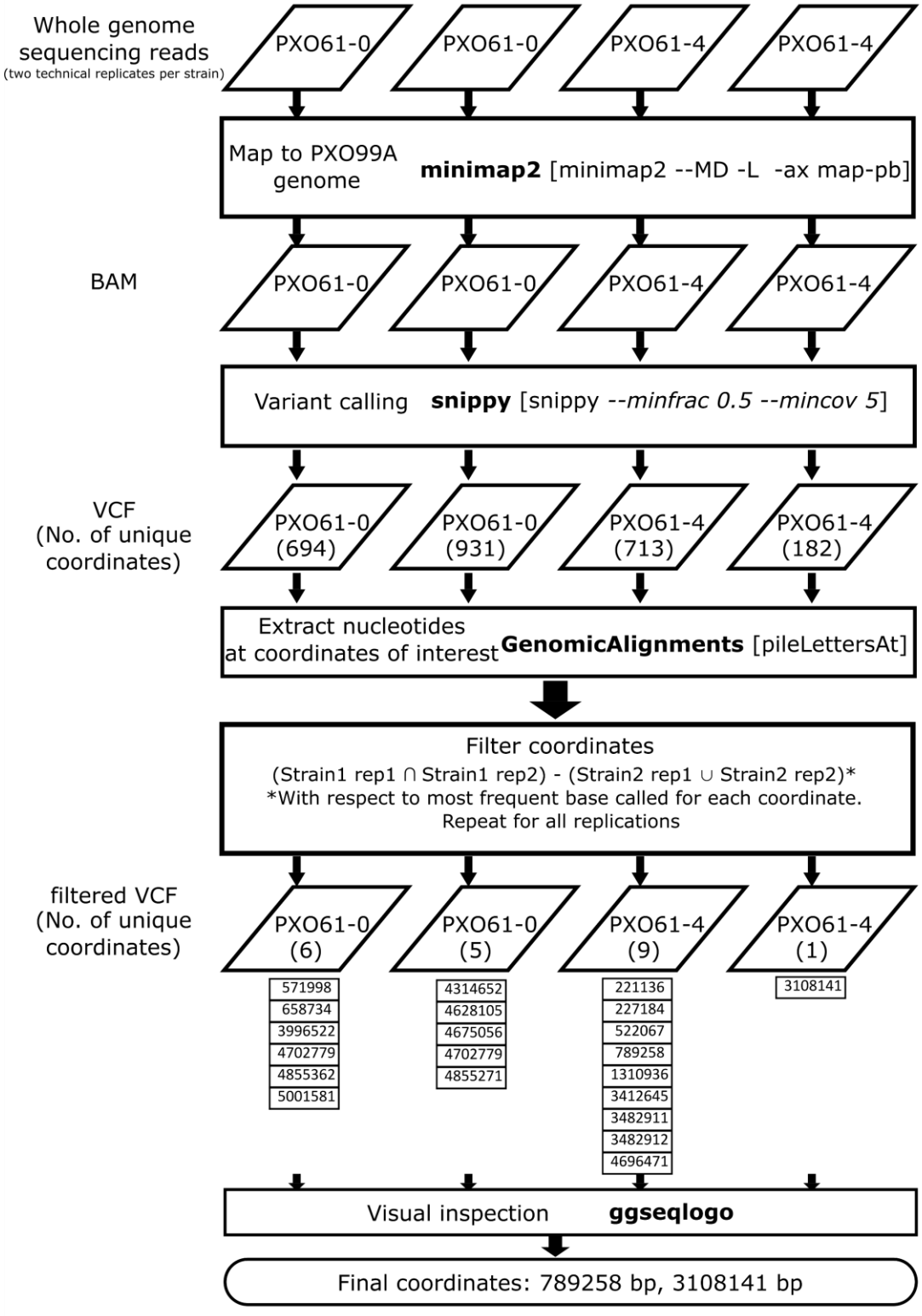


Figure 4.6. Overview of the pipeline used to identify genome sequence variation between *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4).

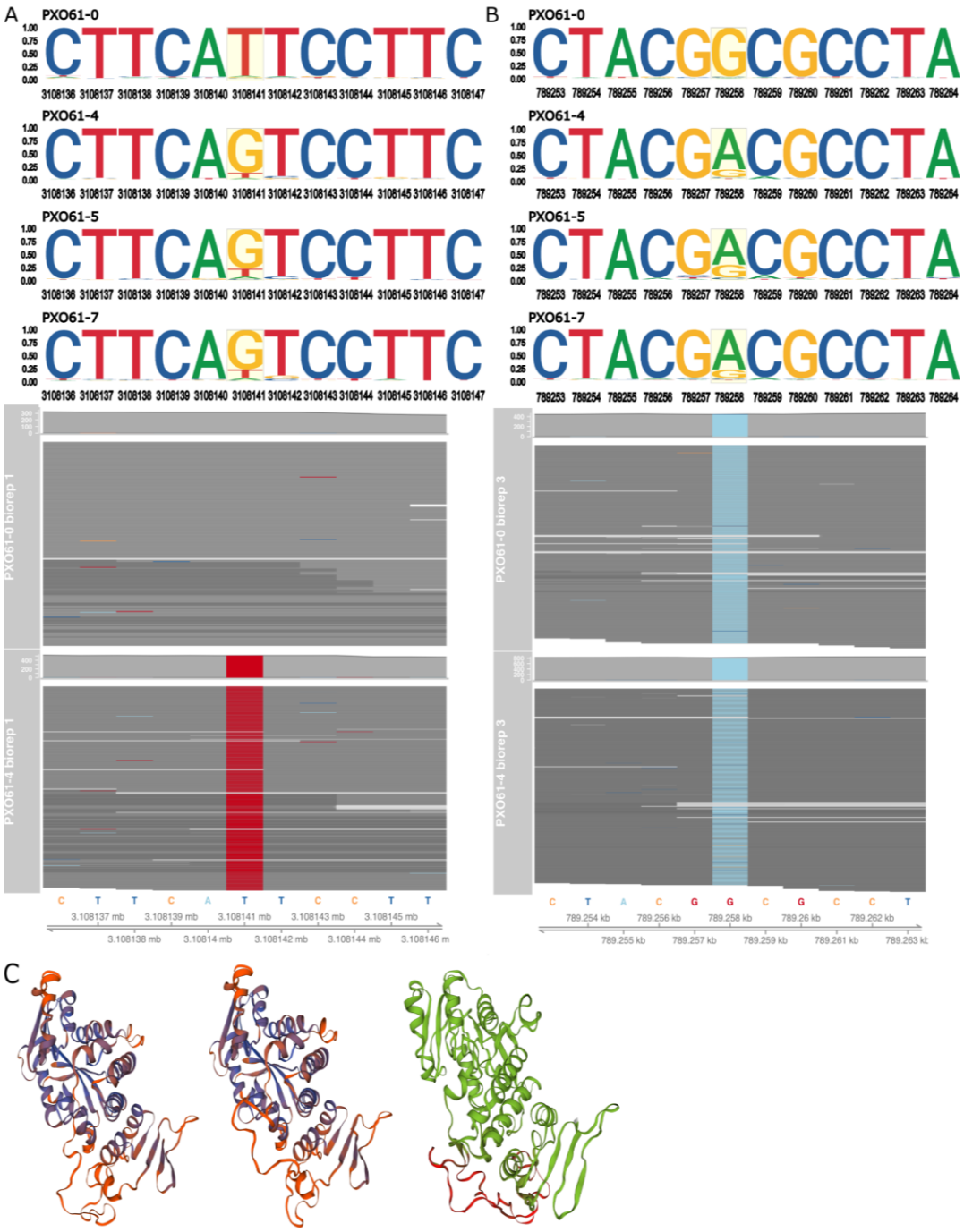


Figure 4.7. Genomic comparison of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4). A (top to bottom). Visualization of the mapped reads at position 3108141, showing sequence variation between the wildtype strain and strains passaged in resistant rice lines IRBB4, IRBB5, and IRBB7. Below, RNA-seq read visualization at the same position confirms the variation between PXO61-0 and PXO61-4. B (top to bottom). Visualization of the mapped reads at position 789258, showing sequence variation

between the wildtype strain and strains grown in IRBB4, IRBB5, and IRBB7. RNA-seq read visualization on the bottom of the panel shows that both strains have the same adenine base at that position. C (left to right). Predicted structures of the Cyclopropane-fatty-acyl-phospholipid synthase with a T to G point mutation causing a substitution from Isoleucine to Serine at the 331 amino acid (aa) position. The first model is the predicted structure of the wild-type protein, the next model shows the predicted structural variation when the T at position 3108141 is replaced with a G. A super-imposition of both models is shown in the right, with structural variation highlighted in red. The red-blue ribbon diagrams from left to right show the native conformation and the altered conformation produced by the substitution, respectively. The green-red ribbon diagram is a superimposition of both conformations. In the right-most superimposed diagram, residues that show inconsistent structures between the two conformations are shown in red. Models were generated and visualized with the SWISS-MODEL Workspace (https://swissmodel.expasy.org/).

Table A.1. Base calls of mapped reads from whole genome sequences of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) to position 3108141 of PXO99A (Salzberg et al., 2008). Extraction of the reads at this coordinate was performed using GenomicAlignments (Lawrence et al., 2013) function pileLettersAt, the variant position was identified using snippy (Seeman, 2015), and read mapping was performed using minimap2 (Li, 2018).

Variant [α]	Nucleotide				Maximum probability	Base call	Nucleotide piles
	A	C	G	T			
PXO61-0 rep 1	0.02	0.00	0.01	0.97	0.97	T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTG TTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTT TTTTTTTTTTTTTTTTTTGTTTTTTTTTTATTTT
PXO61-0 rep 2	0.02	0.01	0.02	0.95	0.95	T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTCTTCTTTTTTTTTGTTTTTTTTTTTTTT GTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTT TTATTTTTTTTTTGTTTTTTTTATTTTTTTTTTTTTT TTTTTTTTTTTATTTTT
PXO61-4 rep 1	0.03	0.00	0.87	0.10	0.87	G	GGGGGGGGGTTGGGGGTGGGGGGGGTGGGGGGG TGGGGGGAGGGGGGTTGGGGGGGGGGGTGGGGG GAGGGGGGGGG
PXO61-4 rep 2	0.02	0.02	0.79	0.17	0.79	G	TGGGGGTTGGTTGGGGTGGGGGGGTGGGGGGGG GGGGGACGGGGTTTGGGGGGGGGGGGGGGGGGG GGGGTTGGGGGGGGGGCCGGGGTTGTTTGGGGTT GGGGGGGGAAGGGGTTTTGGGTGGGGGGGGGGG GGTGGGGGGGGGGTGGG

[α]Rep = technical replicates

4.3.4 Comparative transcriptomic analysis of PXO61-0 and PXO61-4 show differential expression of classical virulence genes as well as genes involved in metabolic pathways

Because variation in growth curves was observed between the two variants, the total RNA PXO61-0 and PXO61-4 after growing for 24 h on modified Wakimoto's agar (Karganilla et al., 1973)was sequenced. Comparative transcriptomic analysis of PXO61-0 and PXO61-4 revealed 163 differentially expressed genes (*p* corrected < 0.05) (Figure 4.8). In this study, the alignment of PXO61-0 and PXO61-4 RNA-seq reads to PXO99 reference genome (Salzberg et al., 2008) showed a total of 4147 transcripts. The gene PXO_RS24615 annotated as *ssrA* had the highest abundance of transcripts, followed by PXO_02523, and PXO_RS24085 *rnpB*. In this study, differentially expressed genes were inferred by filtering the transcripts with p-adjusted values of less than 0.05. Using PXO61-0 expression as the point of reference, 51 genes were upregulated and 112 were downregulated in PXO61-4 (Table B.29). GO Term analysis shows that most of the differentially expressed genes were associated with the GO Term Cell Membrane (*p. adj* = 5.81 x 10^{-05}). Out of 25 hypothetical proteins, 24 were downregulated and 1 was upregulated in PXO61-4.

The most downregulated genes were PXO_02250 (log2FC -3.9, $p = 5.79x10^{-43}$), PXO_03644 (log2FC -3.1, $p = 1.39x10^{-70}$), and PXO_02464 (log2FC -2.1, $p = 5.79 \times 10^{-43}$). PXO_2250 codes for a DNA transport competence protein, while both PXO_03644 and PXO_02464 both encode for *vgrG*, a component of the bacterial type VI secretion system (T6SS). Putative effectors associated with these proteins (Bayer-Santos et al., 2019) - PXO_2459 and PXO_2463, which are predicted to be associated with PXO_2464; as well as PXO_3639 and PXO_3643 which are predicted to be associated with PXO_3644, were also down-regulated in PXO61-4.

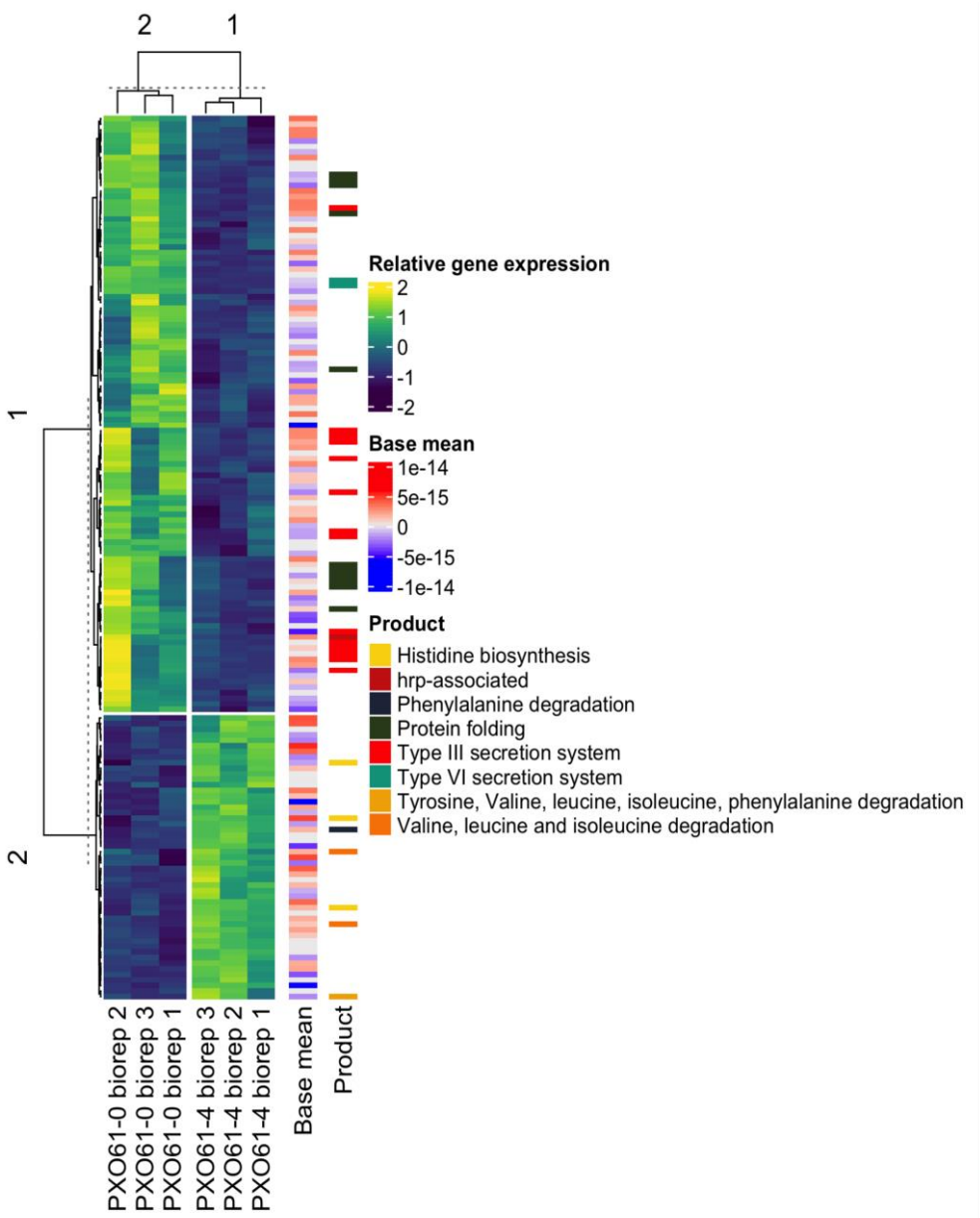


Figure 4.8. Transcriptomic comparison of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4). The heatmap shows relative gene expression of PXO61-0 and PXO61-4 grown in agar media, determined using the R package DEseq2. Total RNA was extracted from three biological replicates derived from three colonies of each strain grown independently for 24 hours on modified Wakimoto's agar. In this comparison, 112 genes are down-regulated, and 51 genes are up-regulated in PXO61-4 with respect to PXO61-0 (p-adjusted values < 0.05). All genomic positions correspond to the PXO99A genome which was used as the reference genome for genomic and transcriptomic analysis. biorep = biological replicate.

The expression of genes associated with the T3SS is less in PXO61-4. These include PXO_03702 *xopX*, PXO_03414 *hpa3*, PXO_03411 *hrpE*, PXO_03392 *hpa1*, PXO_02760 XopN effector, PXO_03417 *hrpF*, PXO_02475 *lon*, PXO_03401 T3S protein HrpB1/HrpK, PXO_00505 *tal3b*, PXO_03409 HrpD5, PXO_03400 T3S protein HrpB2, PXO_03413 xopF1*Xoo*, PXO_03399 EscT/YscT/HrcT family T3SS export apparatus protein, PXO_03403 *hrcV*, PXO_01951 *hrpG*. Components of the bacterial chaperone network were also expressed less in PXO61-4. These include PXO_03704 *groS*, PXO_03703 *groL*, PXO_1186 *dnaJ*, PXO_1185 *dnaK*, PXO_02158 *clpB*, PXO_01184 *grpE*, PXO_00404 *htpX*, PXO_00313 *htpG*, PXO_04280 *hslV*, and PXO_04279 *hslU*.

The most upregulated genes in PXO61-4 were PXO_01706 (Enoyl-CoA hydratase), PXO_1707 *mmsB*, PXO_1704 (Acyl-CoA dehydrogenase family member 8), PXO_01703 *mmsA*, and PXO_1705 (enoyl-CoA). One of the upregulated genes, PXO_RS24800, is annotated as non-coding RNA sX9 sRNA. Three genes that are part of the histidine biosynthetic pathway (PXO_00836 *hisH*, PXO_00835 *hisB*, PXO_00838 *hisF*) was significantly more expressed in PXO61-4.

4.4 Discussion

A serial passage experiment was conducted to investigate the incremental steps in the adaptation of the bacterial blight pathogen *Xoo*, using the strain PXO61, a strain isolated in the Philippines. Since the 1970s, it has been commonly used as a control strain in experiments conducted at IRRI (Mew & Vera Cruz, 1979). This experiment uncovered the potential co-occurrence of two variants of this pathogen, designated as PXO61-0, the strain used in the first cycle of inoculation, and PXO61-4, the selected variant from the sixth cycle of the passage in resistant rice line IRBB4 (*Xa4*). Further comparative analysis of PXO61-0 and PXO61-4 show that the variant PXO61-4 appears to be more fit in terms of *in planta* aggressiveness (Figure 4.2) and *in vitro* metabolism **(**Figure 4.5**)**.

Comparison of WGS reads reveal the presence of one mutation in PXO_01777 (cyclopropane-fatty-acyl-phospholipid synthase) that can differentiate PXO61-0 from the variant PXO61-4, as well as other variants *passaged* in resistant lines IRBB5 (PXO61-5) and IRBB7 (PXO61-7) (Figure 4.7), but there is not enough evidence to correlate its effect on the phenotypic variance we have observed in this study. However, this observation is notable since it hints on the maintenance of variability as a key adaptation mechanism for *Xoo*. The observation of the PXO61-4 allele at a low frequency in the PXO61-0 WGS reads (Table A.1) mapped to that position implies that these strains most likely co-exist as a mixture. In addition, the PXO61-0 allele was also observed in PXO61-4 WGS reads. The variations observed in this study mapped to the PXO99 reference genome is shown in Figure 4.9.

An early review (Law, 1971) of the chemical reactions involved in the biosynthesis of cyclopropanes highlights the activity of cyclopropane synthase as a key enzyme in this process. This enzyme is involved in the modification of existing unsaturated fatty acids in bacterial membrane phospholipids into cyclopropane fatty acids (CFA) and has been reviewed elsewhere (Grogan & Cronan, 1997). CFAs have been associated with adaptations to pH (Brown, Ross, McMeekin, & Nichols, 1997) and proton homeostasis (Shabala & Ross, 2008) in *E. coli* and with the virulence of *Mycobacterium tuberculosis* (Glickman, Cox, & Jacobs, 2000).

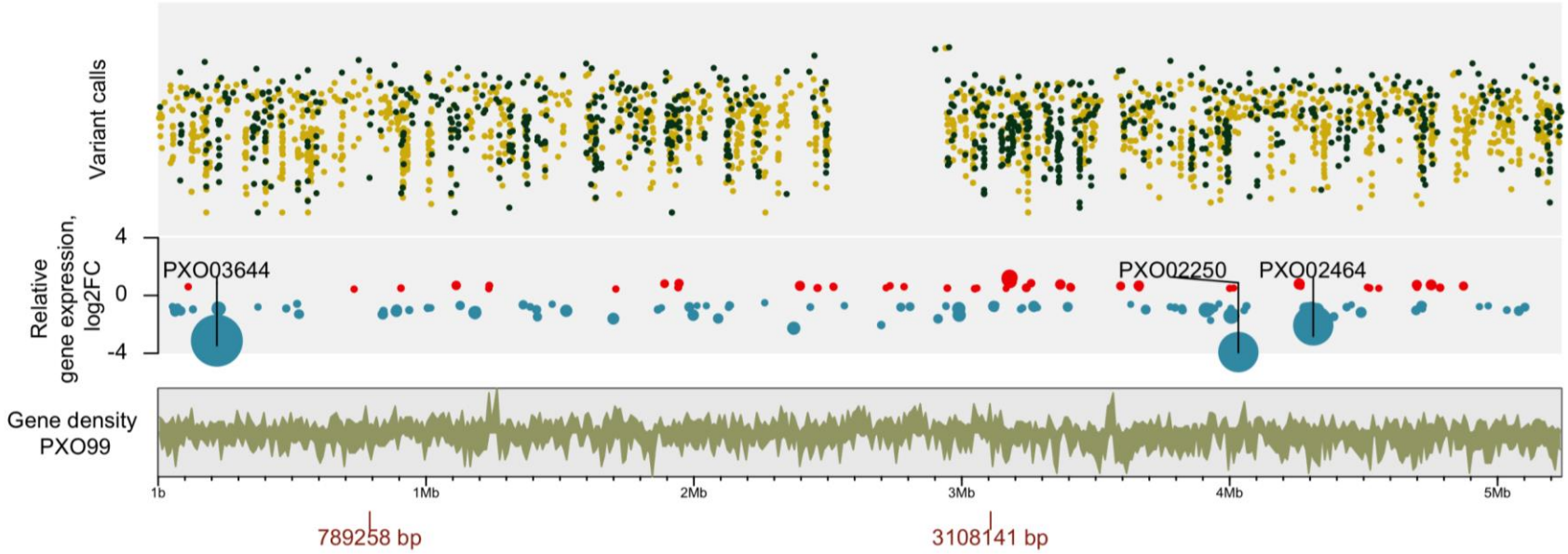


Figure 4.9 . A visualization of the transcriptomic and genomic variation observed between *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4). Shown in the first row is the variant calls from mapping whole-genome sequencing reads to the reference genome, PXO99, with yellow dots referring to PXO61-0 and green dots denoting PXO61-4. The next row is a visualization of the differentially expressed genes plotted along the coordinates of PXO99, the circle diameter is relative to the adjusted p-values. Red dots denote genes that are more expressed in PXO61-4 relative to PXO61-0. Blue dots denote genes that are less expressed in PXO61-4 relative to PXO61-0. The third row is a visualization of the gene annotation in PXO99, red text shows the position of the single nucleotide polymorphisms (SNPs) found after filtering of the variant calls in the first row. Further analysis of the two positions resulted in the validation of position 3,108,141 bp as a SNP, which is found in PXO_01777 (cyclopropane-fatty-acyl-phospholipid synthase).

A notable observation in the comparative transcriptome analysis (Figure 4.8) of PXO61-0 and PXO61-4 is the relatively less expression levels of *hrp* genes in PXO61-4. First characterized in the 1980s (Lindgren, Peet, & Panopoulos, 1986), *hrp* genes or hypersensitive response and pathogenicity genes are components of a type III secretion system (T3SS) which have been found to be drivers of the interaction between plant pathogens and their host plants (for reviews, see (Alfano & Collmer, 1997); (Büttner & He, 2009)). T3SS function in the secretion of pathogenicity proteins into host cells (Hueck, 1998). In *Xoo*, *hrpF* mutants have been found to be less virulent than the wild type strain upon inoculation into susceptible rice variety IR24 (Sugio, Yang, & White, 2005). Also notable is that PXO_02464 and PXO_03644, which are components of the bacterial type VI secretion system (T6SS), both annotated as *vgrG*, coding for an Rhs element Vgr protein, are less expressed in PXO61-4. In *Xanthomonas oryzae* pv. *oryzicola*, the T6SS has been implicated in antagonism against other bacteria by competition assays against *E. coli* and *Ochrobactrum* oryzae (P. C. Zhu et al., 2020). Downregulation of PXO_02250 (DNA transport competence protein) in PXO61-4 may be related to the observed downregulation of T6SS components in the same strain. The T6SS has been reported to be a component of a natural competence system in *Vibrio cholerae* (Borgeaud, Metzger, Scrignari, & Blokesch, 2015), where killing of adjacent cells results in the release of exogenous DNA that can be targeted for uptake by the bacterium. Interestingly, PXO_01511 (Vi polysaccharide biosynthesis protein VipA/tviB) is overexpressed in PXO61-4. This gene is predicted to code for a UDP-glucose/GDP-mannose dehydrogenase family protein, which are proteins that have been reported to be involved in exopolysaccharide biosynthesis (Cole et al., 2012; da Silva, Vettore, Kemper, Leite, & Arruda, 2001). It has been reported that the exopolysaccharide produced by *Vibrio cholerae* functions as a barrier against T6SS-mediated attacks (Toska, Ho, & Mekalanos, 2018). The *udgH* gene (UDP-glucose dehydrogenase) has been reported to be required for virulence in *X. campestris* pv. *campestris* (K.-W. Chang, Weng, & Tseng, 2001).

In bacteria, fatty acid synthesis and fatty acid degradation pathways are regulated in a coordinated manner (Fujita, Matsuoka, & Hirooka, 2007). The downregulation of PXO_01663 *fadE*, which is involved in fatty acid beta-oxidation (Campbell & Cronan, 2002), and the upregulation of PXO_01118 *fabZ*, a key enzyme

in fatty acid biosynthesis (Dodge et al., 2019), suggest a differential activity of PXO61-0 and PXO61-4 with respect to fatty acid metabolism. The upregulation of PXO_01706, and PXO_01705, as well as the upregulation of PXO_01704 which are predicted to code for enoyl-CoA hydratases and an acyl-CoA dehydrogenase, respectively, support this observation. Enoyl-CoA hydratase is an enzyme in the fatty acid beta-oxidation pathway that catalyzes the reversible reaction in the second step of that pathway (Agnihotri & Liu, 2003), while acyl-CoA dehydrogenases are components of both fatty acid beta-oxidation and amino acid catabolism (Swigonová, Mohsen, & Vockley, 2009).

One can postulate that the upregulation of genes involved in catabolic pathways in PXO61-4 could be viewed as a shift from an offensive/defensive lifestyle to one that is dedicated primarily to the expenditure of energy for production of enzymes related to metabolism. In addition to beta-oxidation genes described above, *mmsA* and *mmsB* are related to valine catabolism in *Pseudomonas* (M. I. Steele, Lorenz, Hatter, Park, & Sokatch, 1992) and have been identified to be involved in the monosaccharide utilization of *Agrobacterium* (J. Zhao & Binns, 2014).

On the other hand, biosynthesis of histidine could hint at the employment of virulence mechanisms apart from T3SS. This speculation could be verified in future studies by investigating the transcriptome of PXO61-4 and PXO61-0 *in planta*. Histidine biosynthesis has been identified to play an important role in the pathogenicity of *Xoo* and its growth *in planta* (T. Li et al., 2019). In *X. oryzae* pv. *oryzicola*, a close relative of *Xoo* that also infects rice, the gene *hisB*, which is part of the histidine biosynthesis operon, is required for growth and virulence (P. Su et al., 2018).

Differential expression of genes that are part of the bacterial chaperone network and related proteases could be caused by a variety of reasons. While best known for their roles in the ATP-dependent stress response (Hartl & Hayer-Hartl, 2002), these proteins have been implicated in the survival mechanisms of intracellular clinical pathogens during infection of a host (Stewart & Young, 2004)) HtpG has been reported to be a regulator of DnaK and DnaJ activity in *E. coli* cells that are not exposed to stress (Fauvet et al., 2021). Moreover, the expression of *clpB*, *dnaJ*, *dnaK*, *grpE*, *hslU*, *hslV*, *hslO*, *groS*, *groL*, and *htpG*, has been reported to be least expressed when *E. coli* cells are exposed to environments with high pH (Maurer, Yohannes,

Bondurant, Radmacher, & Slonczewski, 2005), and are more expressed at pH 5.0 and pH 7.0.

The Philippine *Xoo* strain PXO61 appears to possess adaptive mechanisms in place that allows it to change its lifestyle. This duality of bacterial lifestyle is a common theme in several studies, including the ones that investigate biofilm formation, antibiotic tolerance, and persistence. A recent review describes these mechanisms in clinical bacterial pathogens (Desai & Kenney, 2019) – how these microbes are able to switch between planktonic and biofilm conformations, for example, in response to environmental signals. Another example of lifestyle switching in bacteria is the existence of persister (Bigger, 1944) cells, which are associated with tolerance to antibiotics (for a review article, consult (Van den Bergh, Fauvart, & Michiels, 2017)). The plant pathogen *X. citri* subsp. *citri* (Martins, Wood, & de Souza, 2021) as well as *Xylella fastidiosa* (Muranaka, Takita, Olivato, Kishi, & de Souza, 2012) has been demonstrated form persister cells in response to environmental triggers. These phenotypic variation mechanisms are more attributed to isogenic populations, and is reviewed in detail in Ref (Schröter & Dersch, 2019). However, two SNPs were confirmed when PXO61-0 and PXO61-4 WGS and RNA-seq data were analyzed.

The variant PXO61-4 most likely is the variant that is favored once the pathogen successfully enters the host. In this study, the methodology allowed the pathogen to grow without exposure to environmental stresses outside its natural host. Thus, exposure to non-PXO61 competitive bacteria during experimentation was highly unlikely. Re-isolated colonies from the inoculated leaves resulted in pure cultures, except in one plate in which some contamination which appeared to be fungal was observed. Further, the pathogen was directly introduced into the host vascular system via clip-inoculation or infiltration. The only stress it was exposed to is presumed to be the host defense mechanisms in relation to the resistance genes. However, transcriptome analysis of IRBB4 inoculated with PXO61-0 and PXO61-4 (10^{th} cycle) showed that there is down-regulation of defense-related genes at 12 hpi in plants inoculated with PXO61-4 (10^{th} cycle) (Zaka, 2019). This implies that the variant passaged in IRBB4 is undetectable by the host during infection at that time point or can inhibit defense-related genes.

Hence, we have uncovered two variants of *Xoo* strain PXO61, both with the ability to infect rice. But the observations show that PXO61-4 is more fit in terms of aggressiveness *in planta* and *in vitro*, and that the PXO61-0 wild-type appears to spend additional energy to express genes related to type III and type VI secretion systems as well as molecular chaperones, implying fitness to a lifestyle of microbial competition and canonical mechanisms of pathogenesis.

CHAPTER 5: MONITORING *Xanthomonas oryzae* pv. *oryzae* POPULATIONS WITH SINGLE NUCLEOTIDE POLYMORPHISM-BASED PRIMERS

5.1 Introduction

Xanthomonas oryzae pv. *oryzae* has been extensively reported to adapt and shift its population structure in response to the deployment of resistant rice cultivars (Mew et al., 1992; Midha et al., 2017; Nelson et al., 1994; Quibod et al., 2016; Vera Cruz et al., 2000), which may lead to the existence of more virulent pathotypes (Khaeruni & Wijayanto, 2013) or to the acquisition of new pathogenic characteristics (Noda et al., 2001). Due to its distribution and economic importance, *Xoo* has been placed in the list of top ten most important pathogenic bacteria in crop plants (Mansfield et al., 2012). Therefore, monitoring the changes in pathogen population structure is essential to detect the emergence of a new, more virulent strain. Further, regular and frequent sampling is essential for early detection of an invading pathogen and subsequent management of the disease (Parnell, Gottwald, Cunniffe, Alonso Chavez, & van Den Bosch, 2015). However, the standard method for *Xoo* characterization is slow and cumbersome to perform routinely.

Monitoring the local distribution of *Xoo* populations is conventionally done by first isolating *Xoo* from diseased leaf samples collected from rice fields, then conducting a pathogenicity assay under controlled conditions by inoculating the strains into rice varieties or lines with known resistance genes. The routine conduct of this procedure is very tedious to perform and is a major constraint in ensuring that current pathogen population data is readily available. Ideally, pathogen population data should be available in real-time to enable timely rice breeding intervention or resistant variety deployment decisions (Dossa et al., 2015).

Globally, the impact of rice diseases result to estimated yield losses of up to 30% (Savary et al., 2019). Because of the potential damage caused by plant pathogens, the development of tools to monitor these pathogens is a continuously expanding field. Recent advances in genomics facilitate the development of DNA-based detection protocols. The availability of whole genome sequences has enabled the design of highly specific primers suitable for diagnostic assays. For instance, In *Xoo* diagnostics,

the availability of Xanthomonad and plant-associated bacterial genomes led to the development of pathovar-specific primers for the detection of *Xoo* via multiplex PCR (Lang et al., 2010), and via loop-mediated isothermal amplification (LAMP) (Lang et al., 2014), a method to amplify DNA that does not require thermal cycling (Notomi et al., 2000). Genomics-based methods for the monitoring of pathogens have been developed in a variety of fields – food safety (Alegbeleye & Sant'Ana, 2020; W. Liu et al., 2020; Szarvas et al., 2020; Uelze et al., 2020), public health (Cremers et al., 2020; Gardy & Loman, 2018; Roe et al., 2016; Varghese et al., 2020), and in agriculture (Radhakrishnan et al., 2019). With the decline in the cost of implementing high throughput sequencing methods, using molecular methods to routinely analyze a high number of samples from the field has become more feasible.

Previously, Quibod et al (Quibod et al., 2020) demonstrated the use of core SNPs obtained from the alignment of 91 Philippine *Xoo* genomes to predict *Xoo* phylogenetic structure which generated three groups, confirming the results of a previous experiment using a smaller set of genomes (Quibod et al., 2016). Using both phylogenetic and phenotype data, these studies demonstrated that *Xoo* adapted to the deployment of the R gene Xa4, and that SNP-based populations may be associated with the *Xoo* phenotype. The results described in Chapter 3 and Chapter 4 also showed that adaptation in *Xoo* could potentially be tracked with signatures in their genomes.

The KASP (Kompetitive allele-specific PCR) assay is a method that is gaining traction in applications that require detection of genetic variants. In plant breeding applications, KASP assays have been developed to detect the presence of specific loci that are associated with a targeted trait (Fu, Mason, Zhang, & Yu, 2021; Ur Rehman et al., 2021). This assay permits the simultaneous detection of two SNP alleles by requiring that two forward primers, one for SNP allele 1 and another for SNP allele 2, are included in the reaction which does include a common reverse primer (Suo, Shi, Xu, Li, & Lin, 2020). Because the forward primers are labelled with a fluorescent tag, it is possible to detect which allele is present in the sample based on the fluorescence signals.

In this study, SNPs that are linked to putative populations of *Xoo* inferred through Bayesian analysis were transformed into KASP markers and validated against various sample types and sample sources. By developing a monitoring tool for *Xoo* populations using SNP-derived markers, this study aimed to answer the research

question, can SNP-based markers be used to monitor population shifts of *Xoo* and serve as a useful tool in disease monitoring?

5.2 Methodology

5.2.1 Selection of putative population-specific SNPs for genotyping

The development of the SNP markers is described in Figure 5.1. *Xoo* genomes (Table S1) were downloaded from publicly available datasets (Chien, Chou, Chen, & Shih, 2019; Midha et al., 2017; Quibod et al., 2020). Extraction of the core SNPs was applied thru Harvest 1.1.2 (Treangen, Ondov, Koren, & Phillippy, 2014) using parsnp. A total of 5,000 SNPs were finally annotated from PXO99A reference genome (Salzberg et al., 2008) using SNPeff (Cingolani et al., 2012). Putative population structure was estimated using heirBAPS (L. Cheng et al., 2013). Recombinant SNPs were filtered by clonalframeML 1.25 (Didelot & Wilson, 2015). Phylogenetic reconstruction was implemented using a maximum likelihood (ML) approach in RaxML 8.2.9 (Stamatakis, 2014). The parameters for the ML tree are 1000 bootstrap runs employing general time reversal model of nucleotide substitution with the Gamma model of rate heterogeneity. The visualization of the ML tree was performed in the R package GGTREE (G. Yu et al., 2017).

After extracting the SNPs in the core genome alignment and inferring the population structure, the SNPs for building the markers were selected. Putative population-specific SNPs were selected based on the following criteria: a) the SNP is within a conserved coding region, (b) it is outside regions with mobile genetic elements, (c) rooted in the phylogenetic tree, (d) with no close signatures of recombination, and (e) with flanking sequences of less than 97% similarity to other microbes or rice genomes. Selected SNPs were transformed into Kompetitive allele-specific PCR (KASP^{TM}, LGC Biosearch Technologies) markers at Intertek (Hyderabad, India). A total of 14 SNPs were included in the pilot test.

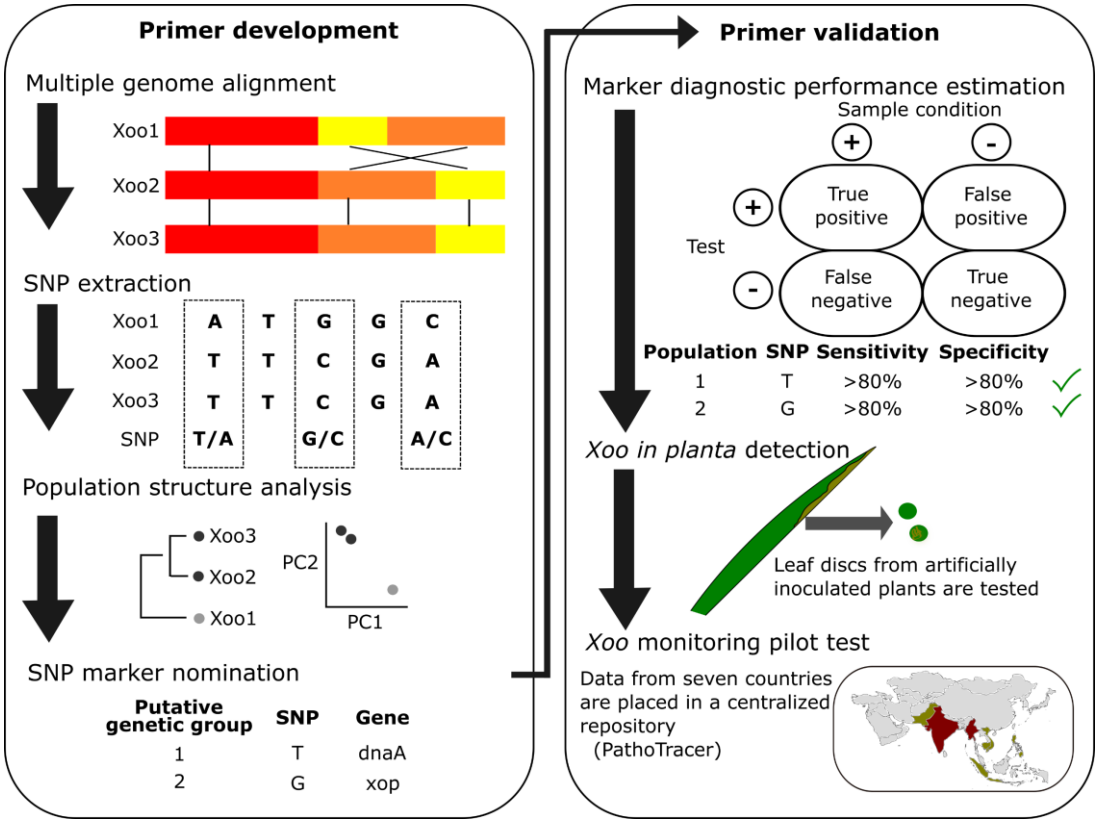


Figure 5.1. The complete *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) marker development and validation workflow, from the bioinformatics analysis, to *in vitro* and *in planta* testing of the markers and to the pilot test of the real-time tracking of the pathogen. The first box describes the workflow used to develop genetic group-specific markers, which starts with whole genome alignment and ends with single nucleotide polymorphism positions which are unique to a genetic cluster. The second box depicts how the markers are validated and finally monitored. The selected markers are tested using trial *Xoo* strains in each genetic group. The markers are then tested for *in planta* detection using four *Xoo* strains. Lastly, the markers are then used to track simulated *Xoo* populations across Asia. The image was created using Inkscape 1.1 (https://inkscape.org/).

5.2.2 Sample preparation and SNP assay validation

Three sets of samples were tested as described in Figure 5.1. To evaluate the assay performance, DNA samples from standard *Xoo* strains were used to estimate the diagnostic specificity, sensitivity, and accuracy of the primers. Selected isolates were incubated overnight on nutrient broth (BBL™, USA) at 30 °C and genomic DNA was extracted using the Easy-DNA kit (Invitrogen, USA) following the manufacturer's protocol. The concentration of each DNA sample was measured using a NanoDrop Spectrophotometer (ThermoFisher Scientific, USA) and was adjusted to 2 µg before being aliquoted in 96-well plates. The list of samples included in the marker validation can be found in Table C.2.

5.2.3 In planta detection of Xoo

Leaf discs were collected from artificially inoculated plants to determine the detection potential *in planta*. Inoculation of tester strains PXO71, PXO99, PXO112, and PXO349 was performed using the leaf-clip method (Kauffman et al., 1973). Leaf samples were collected 15- and 30-days post- inoculation (dpi). These were surface disinfected by dipping in 70% ethanol for 1-2 seconds, rinsed three times in sterile distilled water, and were dried using sterile paper towels. A single-hole punch was used to collect 2-mm diameter discs from the base of lesions and about 5 cm below the lesions. In the older leaves (30 dpi), the samples were obtained by tracing the lesion margin and punching holes within the lesion. The single-hole punch was soaked in 95% ethanol and passed through a Bunsen burner flame in between samples. Two leaf discs for each sample were placed into 96-well microplates and run against all the 14 primers.

5.2.4 Pilot test

Field samples collected from rice-growing areas, as well as purified DNA from *Xoo* collections in Cambodia, India, Indonesia, Pakistan, Philippines, and Vietnam were used in the pilot test. The sample types from each country are summarized in Appendix A**.** Samples from Cambodia were processed using a quick extraction method to isolate plant and *Xoo* DNA from Whatman® FTA® PlantSaver Cards. The leaves were rubbed onto the cards following the manufacturer's protocol. Then, 0.5 cm x 0.5 cm sections were cut from the card using sterilized forceps and scissors and transferred into a 1.5 ml microtube. The tube was washed with approximately 100 µl of 80%

ethanol and vortexed at maximum setting for 10 s. Then the ethanol was removed with a pipette and 100 µl of 1X Tris-EDTA (TE) buffer (pH 8.0) (Thermo Fisher Scientific, USA) was added to the tube. After gentle rotation and inversion of the tube several times, the tube was spun down briefly for 5 s and the 1X TE was removed using a pipette. The washing steps were done twice. Finally, 100 µl of 1X TE was added to the tube and stored for at least 1 h prior to dispensing into a 96-well plate for the analysis. DNA samples and leaf discs were prepared as described above.

5.2.5 Data analysis and visualization

Estimates of primer specificity and sensitivity were calculated using the R package caret (Kuhn, 2008) version 6.0-88. Genotype data was re-coded into simulated population codes (popID), representing genetic group-specific diagnostic primers using reference (PXO99A allele) or alternative forms. Samples with null or with a positive signal for both reference and alternative alleles were verified manually. Predominant populations were organized by country and sub-region using R package dplyr (Wickham, Francois, Henry, & Müller, 2015). Shapefiles were downloaded from GADM.org (https://gadm.org) and imported into Rstudio (R Core Team, 2021) using rgdal (Bivand et al., 2018). These were merged with pathogen data using the geo_join function of tigris (Walker, 2018). Visualization of the combined data was done using leaflet (J. Cheng, Karambelkar, & Xie, 2018) and shiny (W. Chang, Cheng, Allaire, Xie, & McPherson, 2017), and is deployed at https://mhn1.shinyapps.io/PathoTracer/. Version tracking was done by saving all scripts in the github repository https://github.com/marianhannanguyen/PathoTracer The final version of the code used to visualize the data was uploaded to the Zenodo repository https://doi.org/10.5281/zenodo.5883068. The interactive map was then uploaded into the PathoTracer platform, which was built for centralized access to rice pathogen data (https://sites.google.com/irri.org/pathotracer),) as a decision tool for customized deployment of resistant rice varieties (Dossa et al., 2015). To simplify the visualization, the simulated population data was summarized using unique colors at the administrative boundary level 1, or the first subdivision of each country, by determining the simulated population mode and displaying it as the dominant population in the map. Areas with two modes were categorized as having “no dominant population”. The category “uncharacterized” represents a sample that tested

positive for more than one group-specific SNP and warrants further characterization of the strains. Samples that test negative for any of the group-specific alleles were categorized as “others”. Clicking on each province reveals the simulated population composition in detail and the number of samples analyzed for that area.

5.3 Results

5.3.1 Phylogenetic analysis and primer selection

To define putative populations of the pathogen, a Bayesian algorithm (L. Cheng et al., 2013) was applied to analyze 197 Asian *Xoo* genomes (Table C.1) and resulted in the identification of at least 9 clusters. This method was selected due to its flexibility to integrate new *Xoo* genomes in the future. A total of 5,000 SNPs were annotated from the PXO99A reference genome (Salzberg et al., 2008) which was filtered based on the following criteria: a) the SNP is within a conserved coding region, (b) it is outside regions with mobile genetic elements, (c) rooted in the phylogenetic tree, (d) with no close signatures of recombination, and, (e) with flanking sequences of less than 97% similarity to other microbes or rice genomes. A total of 14 SNPs rooted in the phylogenetic tree (Figure C.1) was nominated based on the inferred genetic clusters. Details of the selected SNPs, genomic locations, gene annotation, and putative population codes (popID) can be found in Table A.1.

Table A.1. List of the genetic group-specific single nucleotide polymorphisms inferred through genetic similarity structure analysis of 197 Asian *Xanthomonas oryzae* pv. *oryzae* genomes selected for the pilot test of monitoring simulated *Xanthomonas oryzae* pv. *oryzae* populations.

Intertek SNP ID	Gene name	Old tag	Reference allele	Alternative alelle	Putative population*
snpOS0345	acetylglutamate kinase	PXO_00349	G	A	popID 1
snpOS0256	pectate lyase	PXO_02650	G	T	popID 2
snpOS0346	moaA/ molybdenum cofactor biosynthesis protein A	PXO_02449	C	A	popID 3
snpOS0344	topA/ DNA topoisomerase I	PXO_04049	G	T	popID 4
snpOS0338	lysyl-tRNA synthetase	PXO_00071	C	G	popID 5
snpOS0343	trmB/tRNA (guanine-N(7)-)-methyltransferase	PXO_01840	C	A	popID 6
snpOS0342	TonB-dependent outer membrane receptor	PXO_03612	T	C	popID 7
snpOS0339	membrane protein, putative	PXO_02484	C	T	popID 8
snpOS0260	recG	PXO_02362	G	T	popID 9
snpOS0341	mltB/lytic murein transglycosylase B	PXO_04282	C	T	popID 10
snpOS0258	endonuclease	PXO_04291	C	A	popID 11
snpOS0263	ribosome-binding A	PXO_01304	G	A	popID 12
snpOS0262	TonB- receptor	PXO_02512	G	T	popID 13
snpOS0340	mgtE/ magnesium transporter	PXO_02891	C	G	popID 14

*Arbitrary naming system used in this study

5.3.2 Diagnostic sensitivity and specificity

Evaluating diagnostic performance by obtaining estimates of diagnostic accuracy is essential in assay validation (Cardwell et al., 2018). To estimate the diagnostic sensitivity and specificity of the assay, 131 purified DNA controls (Table C.2), composed of 51 target (*Xoo*) and 80 non-target controls (*Xoc*) were genotyped using the 14 primers. A confusion matrix of the results was generated and is shown in C.10. The estimated specificity values ranged from 0.99 to 1.0, and the sensitivity values obtained are all 1.0 (Table A.1). Overall, the accuracy of the assay was calculated to be 0.9847 95% CI [0.9459, 0.9815] (*p-value* [Acc > NIR] = 3.103647 x 10^{-25}). Though most of the strains were tested repeatedly, these samples were prepared independently. Each result was considered as a data point for the estimation of both parameters.

Table A.1. Diagnostic specificity and sensitivity of selected genetic group-specific primers for pathogen monitoring obtained by analyzing purified genomic DNA of *Xanthomonas oryzae* pv. *oryzae* and *Xanthomonas oryzae* pv. *oryzicola* strains. Included in the categories is the no ID class which corresponds to samples which are expected test negative for any of the group-associated SNPs. The table is an output of the confusionMatrix function of the R (R Core Team, 2021) package caret (Kuhn, 2008). Also shown are other parameters such as Positive predictive value, Negative predictive value, Precision, Recall, F1, Prevalence, Detection Rate, Detection prevalence, and Balanced accuracy. [α]

SNP primer	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Precision	Recall	F1	Prevalence	Detection Rate	Detection Prevalence	Balanced Accuracy
no ID	0.975	1.000	1.000	0.962	1.000	0.975	0.987	0.611	0.595	0.595	0.988
snpOS0256	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.031	0.031	0.031	1.000
snpOS0258	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.031	0.031	0.031	1.000
snpOS0260	1.000	0.992	0.875	1.000	0.875	1.000	0.933	0.053	0.053	0.061	0.996
snpOS0262	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.084	0.084	0.084	1.000
snpOS0263	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.023	0.023	0.023	1.000
snpOS0338	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.023	0.023	0.023	1.000
snpOS0339	1.000	0.991	0.938	1.000	0.938	1.000	0.968	0.115	0.115	0.122	0.996
snpOS0340	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.031	0.031	0.031	1.000

[α]The formulas used to calculate these values are shown in detail in the package vignette https://cran.r-project.org/web/packages/caret/caret.pdf

5.3.3 In planta detection

To evaluate the detection accuracy of *Xoo* DNA in rice leaves, a susceptible rice variety was inoculated with Philippine *Xoo* strains PXO71, PXO99, PXO112, and PXO349, representing four different simulated populations and tested these against the 14 primers. Diagnostic accuracy of 100% was obtained from *Xoo* population detection in artificially inoculated leaf discs (Figure 5.2). Random testing of 11 samples collected from asymptomatic plants in the greenhouse, where disease experiments are routinely conducted, resulted in the detection of popID 8 (8 out of 11) and popID 3 (1 out of 11).

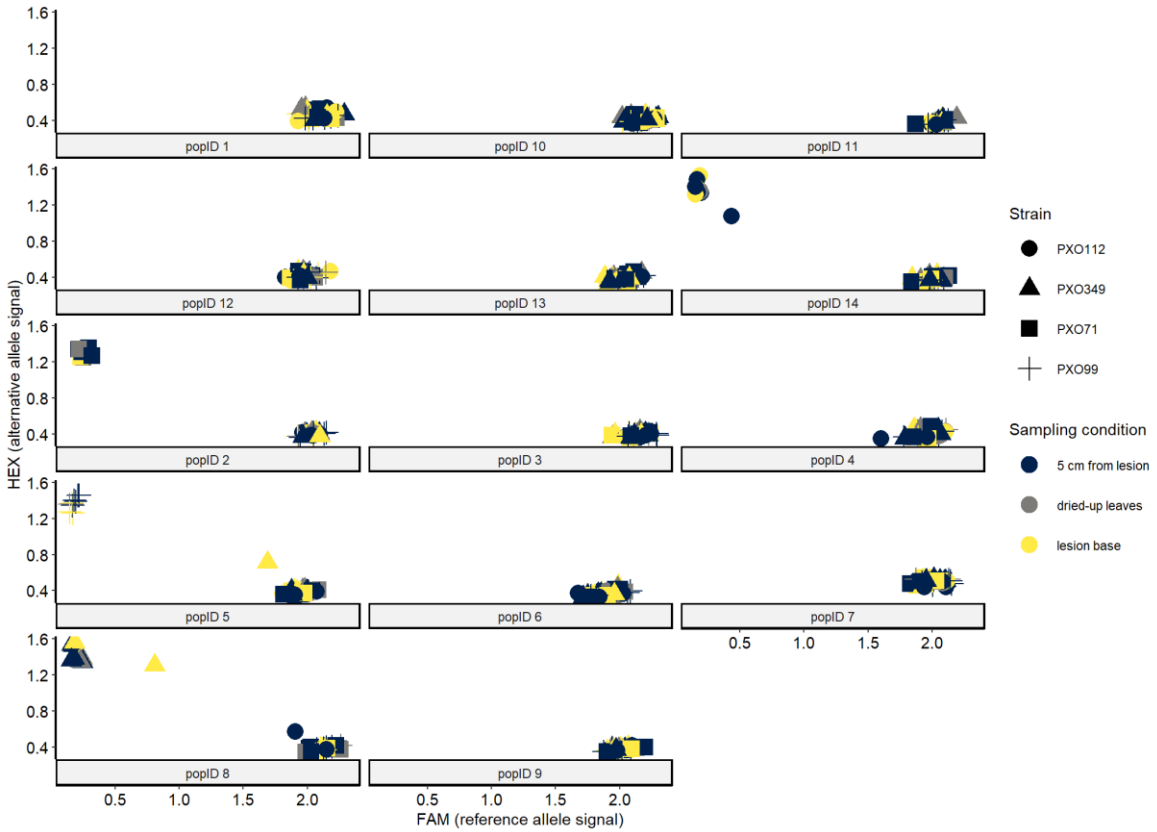


Figure 5.2. *In planta* detection of *Xanthomonas oryzae* pv. *oryzae* in the rice leaf matrix of artificially inoculated plants by genetic group-specific SNP primers. Two discs of 1 mm in diameter were collected from leaves of rice variety IR24, inoculated with Philippine *Xanthomonas oryzae* pv. *oryzae* strains PXO71, PXO99, PXO112, and PXO349. Sampling was conducted at least 20 days after inoculation. Discs were collected 5 cm below the lesion, directly at the lesion base, or from dried-up leaves.

5.3.4 Pilot test and data visualization in PathoTracer

An extensive pilot test of the monitoring pipeline was conducted. This test included samples from rice-growing areas in Cambodia, India, Indonesia, Myanmar, Pakistan, Philippines, and Vietnam. The samples were derived from leaf samples from the field, purified DNA, and DNA stored on Whatman® FTA® PlantSaver cards (Appendix A). Out of the 1067 samples included in the pilot test, 914 (85.7%) tested positive for one putative *Xoo* population-specific SNP, 15 (1.4%) tested positive for more than one SNP, and 138 (12.9%) samples tested negative for the 14 putative *Xoo* population-specific SNPs included in this test. A visualization of the pilot test data is shown in Figure 5.3 and is available online as an interactive map, which is accessible through the PathoTracer platform (https://sites.google.com/irri.org/pathotracer). Figure 5.4 is a consolidated rendering of the pilot test genotyping results.

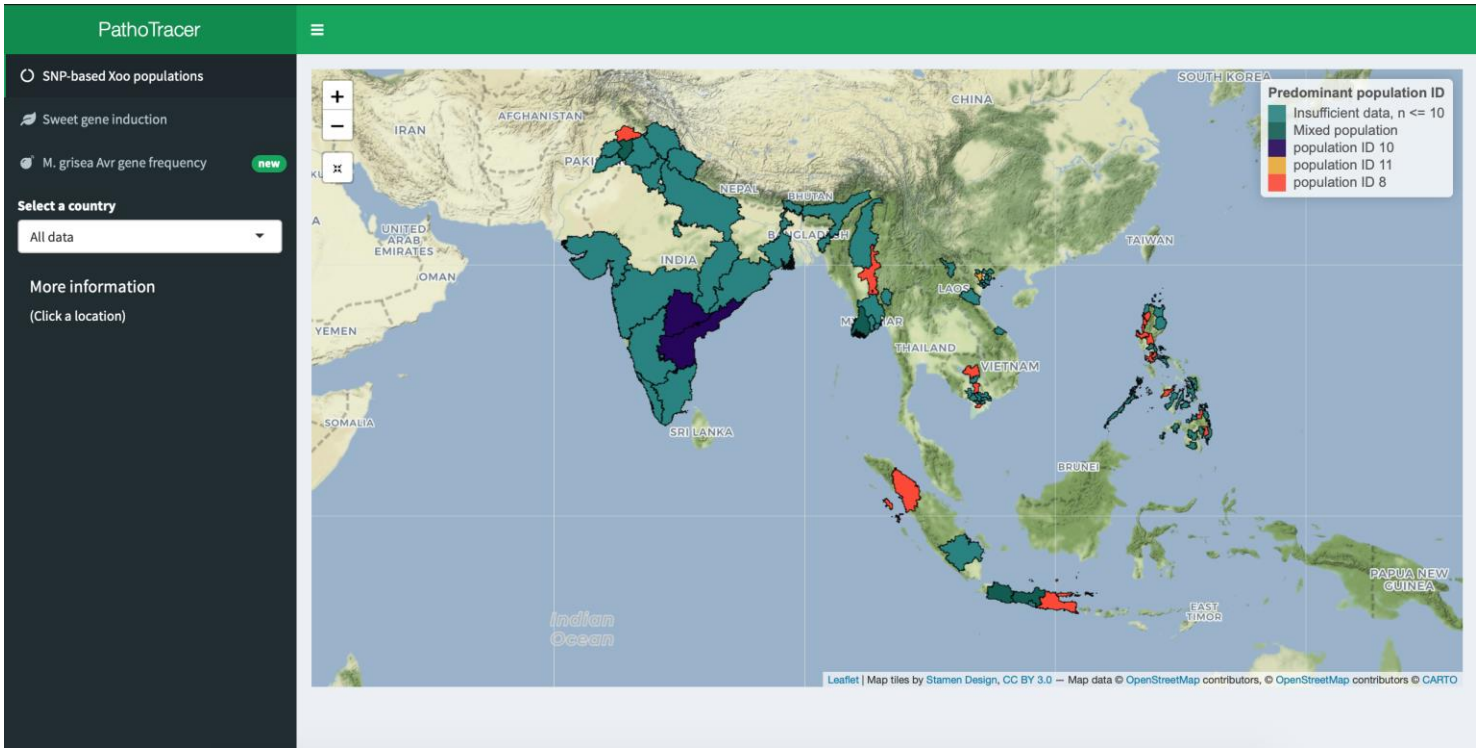


Figure 5.3. An interactive visualization of was constructed for real-time monitoring and detection of simulated *Xanthomonas oryzae* pv. *oryzae* populations across Asia. The interactive visualization was uploaded into PathoTracer (https://sites.google.com/irri.org/pathotracer), a decision tool (Dossa et al. 2015) for rice breeders and a centralized repository created for rice pathogen population information. The color reflects the simulated populations identified by putative group-specific primers in sites which we have acquired samples. Samples include rice leaf discs and DNA extracted from bacterial cultures or leaf samples by DNA extraction kits or Whatman® FTA® PlantSaver Cards.

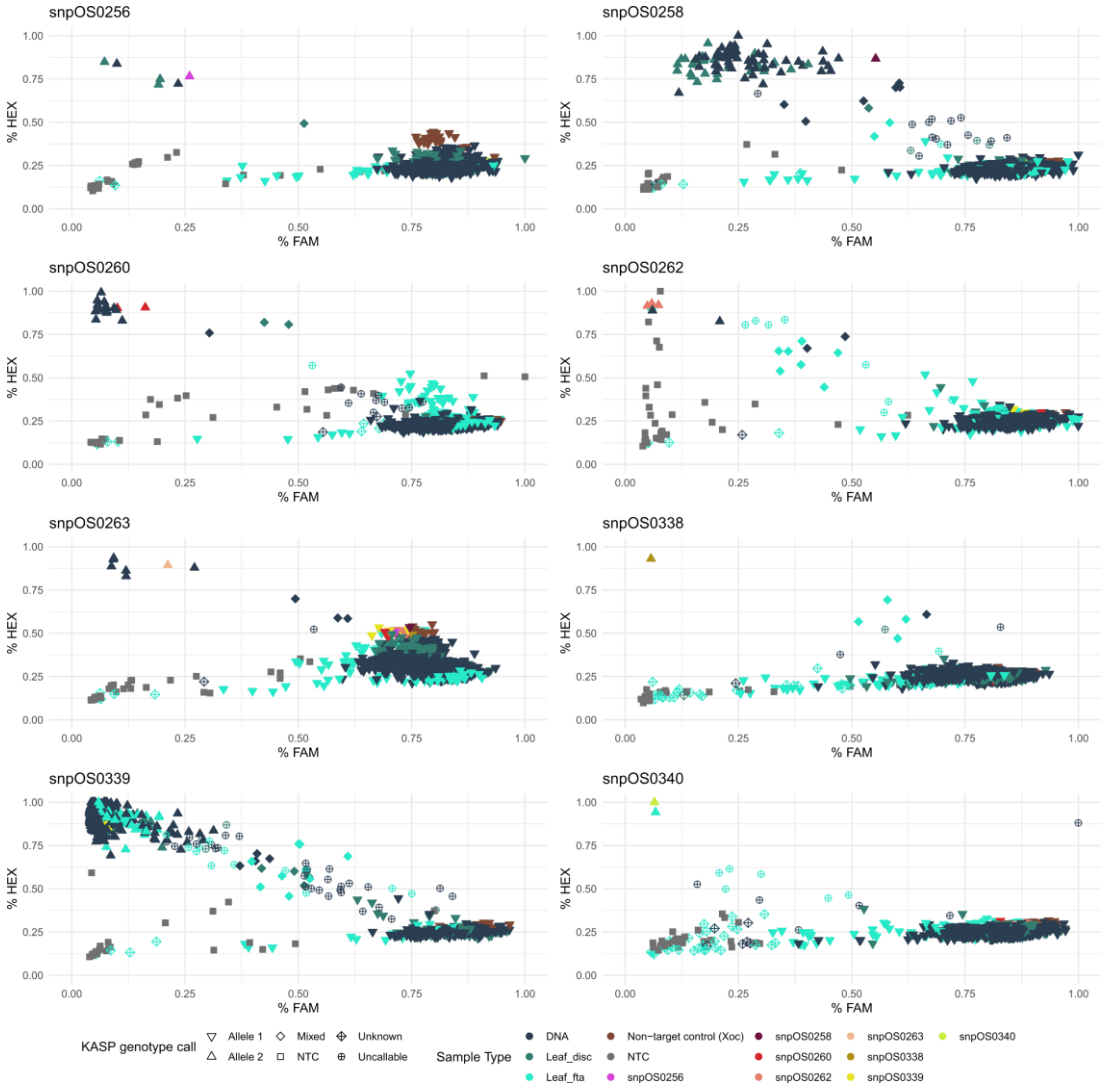
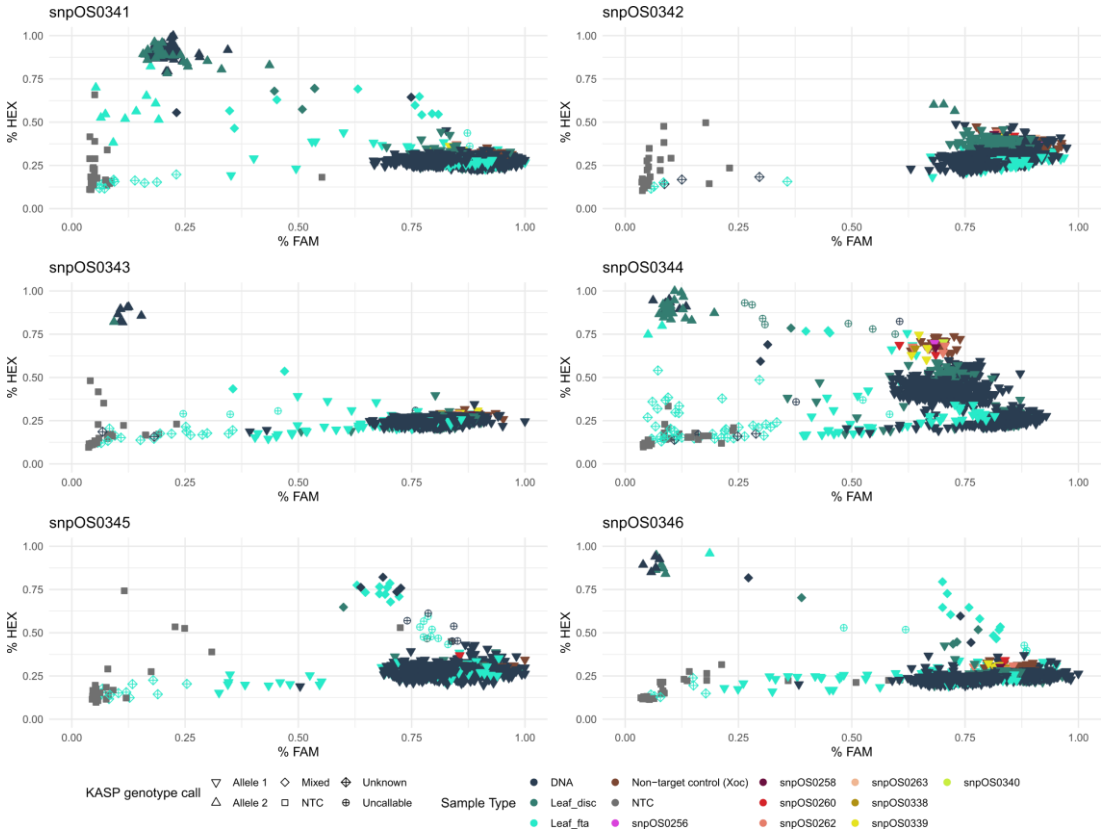


Figure 5.4. Pilot testing the SNP-based detection of *Xanthomonas oryzae* pv. *oryzae*. Consolidated results are visualized in each graph per KASP- (Kompetitive Allele-Specific PCR, LGC Biosciences) SNP primer, where the hex and fam dye signals associated with the putative population-specific and *Xoo* reference alleles, respectively, are expressed as percentages to minimize variation between plates. Shapes represent the genotype calls assigned by the LGC Klustercaller (LGC Biosciences) software, and colors indicate sample types. Whenever available, positive DNA controls are included in the visualization and differentiated by colors.

Figure 5.4 Continued



5.4 Discussion

In this study, genomic data from the genomes of 197 Asia *Xoo* strains were used to develop and validate a field monitoring pipeline for *Xoo* populations. Because the dataset included genomes of *Xoo* strains from India, Japan, Philippines, South Korea, Taiwan, and Thailand, population structure analysis resulted in the prediction of at least nine BAPS clusters. This is a relatively higher number compared to previous genome-based phylogenetic analysis (Adhikari et al., 1995; Quibod et al., 2020; Quibod et al., 2016), which suggests that the 197 Asia *Xoo* strains represents a more diverse collection of Xoo. Still, these genetic groupings were considered as putative because these might only represent a subset of the *Xoo* diversity in Asia.

By selecting SNPs obtained from core genome alignments that were rooted in the phylogenetic tree obtained via BAPS, a total of 14 SNPs were selected for the subsequent steps. These SNPs were transformed into KASP markers and were used to investigate the suitability of SNP-derived primers to monitor *Xoo* populations. The KASP genotyping method has been proven to be useful in breeding programs (K. A. Steele et al., 2018), and also in differentiating variants of the human-infecting virus

SARS-CoV2 (Harper et al., 2021). This method is highly compatible with SNP-derived primers and offers a potentially cost-effective solution for high-throughput monitoring of plant pathogen populations.

In testing the primers against a panel of DNA controls, the estimated values of diagnostic sensitivity were 1.0 for all the primers tested. And the estimated diagnostic specificity obtained ranged from 0.99 to 1.0. Sensitivity, which is the probability of obtaining a positive result in samples with the target DNA, and specificity, which is the probability of obtaining a negative result in samples without the target DNA, are measures of diagnostic accuracy (Šimundić, 2009). Diagnostic assays typically require a minimum of 0.8 for each. For reference, the World Health Organization recommends 0.8 sensitivity and 0.97 specificity for the rapid antigen tests used to detect SARS-CoV2, the human-infecting virus that causes COVID-19 (Peeling, Olliaro, Boeras, & Fongwen, 2021). Sensitivity and specificity values were estimated for all primers with available positive control DNA (snpOS0256, snpOS0258, snpOS0260, snpOS0262, snpOS0263, snpOS0338, snpOS0339, snpOS0340). It is noteworthy that the six remaining primers without available positive DNA controls did not result in a false positive detection when tested against all the 131 DNA controls.

In the validation runs which included artificially inoculated leaves of as samples, it was confirmed that *in planta* detection of *Xoo* populations is possible using the assay. Only a minimum of two leaf discs (2 mm diameter) were required for sensitive detection of these populations in leaves. It is ideal for an assay intended for real-time monitoring of a vascular pathogen to be able to detect the pathogen within the rice leaf matrix. In the same test, random samples taken from healthy plants in the same greenhouse were included in the run, which resulted in the detection of popID 8 and popID 3 putative populations. Interestingly, popID 8 is detected in most of the samples from the Philippines, while popID 3 is more associated with samples from India. Thus, we have shown that leaf discs can be potentially used to directly sample for *Xoo* populations, which could hasten the intended routine process of monitoring *Xoo* groups. Leaf samples can be collected from the rice fields, processed in the laboratory, and sent to the external service provider for analysis. This workflow would bypass the need to routinely isolate, purify, and then extract the DNA of the pathogen, thereby providing real-time results. Subsamples of the leaves can be lyophilized or stored in the -20 ºC freezer if further characterization of the samples is needed.

The KASP genotyping analysis enabled genotype assignments per sample which was obtained automatically through the LGC Klustercaller (LGC Biosciences). Using the 14 primers in this study, the genotypes of each sample was determined based on the 14 SNPs or 14 alleles of the genes where the SNPs were identified in. Based on this genotype, a popID was assigned each sample. It is also possible to view the results of the analysis on a cartesian plane as shown in Figure 5.4 and assess the validity of the genotype calls manually by visual inspection or computationally via cluster analysis. Because only 14 primers have been developed, assigning a popID is relatively simple and straightforward using the KASP genotypes. As more markers are included in the assay, it is highly likely that popID predictions are done via machine learning algorithms, which this type of data is compatible with.

Most importantly, to enhance the intended standardized comparison and monitoring of *Xoo* populations, the aggregated data is available online. A central repository for pathogen population data increases the power of pathogen monitoring systems (Aarestrup et al., 2012; Carvajal-Yepes et al., 2019). The interactive map generated in this study facilitates the visualization of uniformly genotyped *Xoo* population data. This, coupled with a standard method of molecular typing, such as the one described here, enables comparison of these populations at a global scale. Moreover, the variability associated with handling samples and reagents in the laboratory is reduced since the molecular typing is facilitated by an external service provider.

The results of the study confirm that a SNP-based approach is a suitable method for monitoring *Xoo* populations. Genomic information of the vascular rice pathogen *Xoo* was utilized to develop and validate a high-throughput method for the *in planta* detection and field monitoring of putative *Xoo* populations. Artificial genetic groups of *Xoo* inferred by population structure analysis of 197 *Xoo* genomes were successfully detected and differentiated from each other in this pilot test using *Xoo* DNA prepared in the laboratory, leaf samples collected from plants artificially inoculated with *Xoo*, and from leaf samples collected from naturally infected plants in rice fields. The use of these markers and a centralized repository of data to monitor *Xoo* populations could play an important role in the management of BB. However, the importance of the traditional methods of isolation and pathotyping to characterize new strains or populations is also recognized. The data obtained from strain

characterization, coupled with whole genome sequencing will ensure that new populations are added to the database and corresponding population-specific primers are added to the existing marker panel. As more whole genome sequence data of strains become available, inferencing the *Xoo* population structure can be performed at a much higher resolution. This would also allow the discovery of more SNPs associated with genetic clusters. Thus, more than one marker could potentially be used to detect one population. In addition, a hierarchical approach can be employed in the nomination and deployment of suitable primers. A study is underway to establish the *Xoo* populations in Asia using an exhaustive set of available genomes, which is a potential source of these additional markers. In conclusion, the experiments in this study confirmed that SNP-derived primers are indeed useful for monitoring populations of *Xoo*.

CHAPTER 6: CONCLUSION

The research reported in this dissertation focused on *Xoo*, the causal pathogen of BB in rice. Because of the known impact of this disease on yield, rice breeding programs commonly include BB resistance in breeding populations, even if the targeted trait is not BB resistance. Much of the research on BB has been focused on the specificity of the rice-*Xoo* interaction. This project aimed to investigate the adaptation mechanisms of *Xoo* in response to growth in rice, focusing on more general adaptations that increase the fitness of this pathogen *in planta*,– the small incremental changes that could eventually lead to population shifts and to the emergence of a more virulent strain. It is important to study these subtle adaptation mechanisms to find new ways of managing the disease in the future. The experiments related to this aim are presented in Chapter 3 and in Chapter 4. In developing area-specific management strategies for BB, knowledge of existing populations is an advantage. In this regard, Chapter 5 describes the development of a high-throughput and cost-effective pipeline to monitor *Xoo*.

Chapter 3 describes an investigation of the biofilm-related proteins of Xanthomonads in *Xoo* using a bioinformatics approach. In this study, the aim was to gather evidence of variability in biofilm-related proteins in *Xoo* by querying 80 Philippine *Xoo* strains with biofilm-related proteins characterized in Xanthomonadales. This research was initiated by *standing on the shoulders of giants*, where a database of well-characterized biofilm-related proteins was established based on previously published research. Using this database, the presence or absence of proteins in Philippine *Xoo* genomes was confirmed and detailed analyses of multiple sequence alignments were performed. It was found that most of the proteins were conserved in all the tester strains. However, sequence variability and presence/absence variability were also observed in a small set of the query protein sequences. The observed variation in two of the proteins were found to be associated with genetic groups that were inferred based on multiple sequence alignments of the genomes. In a previous report (Quibod et al., 2020), these groups were linked to the widespread adoption of rice varieties that carry the BB resistance gene Xa4. Interestingly, strains which clustered together based on MDS analysis of amino acid dissimilarity scores tend to elicit the same reaction from near-isogenic lines carrying individual R genes.

These observations suggest that there is indeed a link between variation of biofilm-related proteins and *Xoo* adaptation to rice. However, this link requires further validation through rigorous phenotyping experiments.

While the experiments conducted in this study satisfied the aim and answered the research question, the data obtained from this study could also prove to be important for future research in *Xoo* biofilm-formation. This study enabled the identification of a set of biofilm-related proteins which may be linked to host-driven adaptation of *Xoo*. In addition, the genetic clusters of these strains obtained from comparing alignments of the concatenated amino acid sequences could be used as a guide in selecting representative strains for subsequent experiments such as phenotypic characterization of biofilm-formation.

The results of the first study raised a need to first seek out the incremental changes in *Xoo* populations in response to growth in rice. This was the premise for conducting a serial passage experiment involving one *Xoo* strain, PXO61 (Chapter 4). By characterization of *Xoo* after serial passage in rice using a combination of disease phenotyping, metabolic profiling, transcriptome analysis, and genomic analysis. This study aimed to answer two questions: Can *Xoo* adaptation be induced by repeated cycles of inoculation on and re-isolation of the pathogen from rice? If so, what are the incremental changes in *Xoo* populations in response to growth in rice, that cause a shift in virulence?

When *Xoo* was passaged into the resistant line IRBB4 (containing *Xa4*), a variant that exhibited increased fitness *in planta* and *in vitro* was discovered through several experiments that quantified the *Xoo* phenotype *in planta* and *in vitro* (Chapter 4). Analysis of WGS reads suggests that the wild-type sub-population and variant sub-population are maintained by the strain used in the study, PXO61. This maintenance of sub-populations potentially allows the pathogen to increase its fitness in response to changing environments. Further, transcriptomic comparisons of PXO61-0 (wild type) and PXO61-4 (passaged in IRBB4) suggests differences in substrate preferences of these two strains. In this study, one may consider that PXO61 maintains at least two populations, the PXO61-0 sub-population and the PXO61-4 sub-population, and that the ratio of PXO61-0 and PXO61-4 in this population fluctuates and changes in response to environmental changes or signals.

The study described in Chapter 4 involved experiments using one *Xoo* strain. Future studies of serial passage in resistant lines involving more strains are recommended. Moreover, *in planta* transcriptome analysis of *Xoo* variants passaged in R lines could complement the transcriptome analysis described in Chapter 4. There is also a potential question to address, if the PXO61-4 sub-population has increased fitness *in planta*, what environment is the PXO61-0 sub-population more adapted to? Since the experiments were conducted with pure cultures of *Xoo,* the characterization of population shifts when *Xoo* is grown in mixed cultures could be a potential source of more information. It is also possible that the PXO61-0 sub-population is more adapted to stringent conditions outside the host, or in the soil. Since new isolates are routinely collected from rice fields when the symptoms are readily observed, that is, when the infection is several weeks old, it is also quite possible that PXO61-0 sub-population is favored just as infected leaves dry up or when the infection has progressed for several weeks.

An important observation in the serial passage study (Chapter 4) is the detection of one SNP that is specific to the sub-populations that exhibit different fitness attributes. The SNP was also detected in variants passaged in other resistant rice lines (IRBB5 and IRBB7), which supports the inference that PXO61 maintains at least two populations, the PXO61-0 sub-population and the PXO61-4 sub-population. This also implies that the sub-populations are maintained, and the dominant variant is selected in response to the environment or conditions within the host. This led to the hypothesis that SNPs could be a useful tool in tagging *Xoo* populations and serve as tool for monitoring these populations. Chapter 5 describes how a SNP-based monitoring tool was developed based on putative population-linked SNPs. The tool was successfully validated in the laboratory and pilot-tested using field samples from an intensive inter-Institute collection scheme across Asia. This is an important contribution to the field of *Xoo* monitoring, it potentially enables rice breeders to quickly assess which R genes are suitable for breeding and deployment in specific regions based on real-time data.

The success of monitoring *Xoo* populations strongly depends on the continuous characterization of new and emerging populations *Xoo* for the primers to be kept up to date. This requires frequent sampling and early detection of emerging populations, followed by whole genome sequencing of selected strains in order develop new

primers for inclusion in the detection pipeline, after primer validation. These activities are only possible when there is a strong collaborative effort among research entities and National Plant Protection Organizations (NPPOs) to engage in rigorous sample collection from farmers' fields and in the isolation and characterization of nominated strains that are suspected to be part of an emerging population that may pose a new threat to rice production.

In conclusion, the three studies presented in this thesis show a key feature in the adaptation of *Xoo* – the maintenance of variability as an adaptive mechanism of *Xoo* during its growth in rice. This variability was observed in a specialized set of proteins related to biofilm formation and was uncovered in a strain that was passaged in resistant lines (Chapter 3). This variability appears to be maintained by *Xoo*, as shown in Chapter 4, such that it can be exploited to develop monitoring tools for this *Xoo*, as demonstrated in Chapter 5. This observed variability also implies that *Xoo* is still evolving and further research on *Xoo* is required to develop new ways of managing the disease and to maintain the food supply of almost half of the world's population who depend on rice.

REFERENCES

Aarestrup, F. M., Brown, E. W., Detter, C., Gerner-Smidt, P., Gilmour, M. W., Harmsen, D., . . . Johansson, K. (2012). Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response. *Emerging infectious diseases, 18*(11), e1.

Ackermann, M., Schauerte, A., Stearns, S. C., & Jenal, U. (2007). Experimental evolution of aging in a bacterium. *BMC evolutionary biology, 7*(1), 126. doi:10.1186/1471-2148-7-126

Adhikari, T. B., Shrestha, A., Basnyat, R. C., & Mew, T. W. (1999). Use of Partial Host Resistance in the Management of Bacterial Blight of Rice. *Plant disease, 83*(10), 896-901. doi:10.1094/pdis.1999.83.10.896

Adhikari, T. B., Vera Cruz, C., Zhang, Q., Nelson, R. J., Skinner, D. Z., Mew, T. W., & Leach, J. E. (1995). Genetic Diversity of *Xanthomonas oryzae* pv. *oryzae* in Asia. *Applied and Environmental Microbiology, 61*(3), 966-971. doi:doi:10.1128/aem.61.3.966-971.1995

Agnihotri, G., & Liu, H.-w. (2003). Enoyl-CoA hydratase: Reaction, mechanism, and inhibition. *Bioorganic & Medicinal Chemistry, 11*(1), 9-20. doi:https://doi.org/10.1016/S0968-0896(02)00333-4

Ahsan, R., Ullah, S., Yaseen, I., Fateh, F. S., Fayyaz, M., Asad, S., . . . Zakria, M. (2021). Assessment of bacterial leaf blight incidence and severity in rice growing areas of Pakistan. *Pakistan Journal of Agricultural Research, 34*(4), 693-699.

Akamatsu, A., Wong, Hann L., Fujiwara, M., Okuda, J., Nishide, K., Uno, K., . . . Shimamoto, K. (2013). An OsCEBiP/OsCERK1-OsRacGEF1-OsRac1 Module Is an Essential Early Component of Chitin-Induced Rice Immunity. *Cell Host & Microbe, 13*(4), 465-476. doi:https://doi.org/10.1016/j.chom.2013.03.007

Akimoto-Tomiyama, C., Furutani, A., & Ochiai, H. (2014). Real Time Live Imaging of Phytopathogenic Bacteria *Xanthomonas campestris* pv. *campestris* MAFF106712 in 'Plant Sweet Home'. *PLoS One, 9*(4), e94386. doi:10.1371/journal.pone.0094386

Alcaraz, E., García, C., Friedman, L., & de Rossi, B. P. (2019). The rpf/DSF signalling system of *Stenotrophomonas maltophilia* positively regulates biofilm formation, production of virulence-associated factors and β-lactamase induction. *FEMS Microbiology Letters, 366*(6). doi:10.1093/femsle/fnz069

Alegbeleye, O. O., & Sant'Ana, A. S. (2020). Pathogen subtyping tools for risk assessment and management of produce-borne outbreaks. *Current Opinion in Food Science*.

Alfano, J. R., & Collmer, A. (1997). The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *Journal of bacteriology, 179*(18), 5655-5662.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410. doi:https://doi.org/10.1016/S0022-2836(05)80360-2

Alvarez-Martinez, C. E., Sgro, G. G., Araujo, G. G., Paiva, M. R. N., Matsuyama, B. Y., Guzzo, C. R., . . . Farah, C. S. (2021). Secrete or perish: The role of secretion systems in *Xanthomonas* biology. *Computational and Structural Biotechnology Journal, 19*, 279-302. doi:https://doi.org/10.1016/j.csbj.2020.12.020

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Antar, A., Lee, M.-A., Yoo, Y., Cho, M.-H., & Lee, S.-W. (2020). PXO_RS20535, Encoding a Novel Response Regulator, Is Required for Chemotactic Motility, Biofilm Formation, and Tolerance to Oxidative Stress in *Xanthomonas oryzae* pv. *oryzae*. *Pathogens, 9*(11), 956. Retrieved from https://www.mdpi.com/2076-0817/9/11/956

Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., & Yang, B. (2010). Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3. *The Plant Cell, 22*(11), 3864-3876.

Aragón, I. M., Pérez-Mendoza, D., Gallegos, M.-T., & Ramos, C. (2015). The c-di-GMP phosphodiesterase BifA is involved in the virulence of bacteria from the *Pseudomonas syringae* complex. *Molecular Plant Pathology, 16*(6), 604-615. doi:https://doi.org/10.1111/mpp.12218

Ardales, E. Y., Leung, H., Vera Cruz, C. M., Mew, T. W., Leach, J. E., & Nelson, R. J. (1996). Hierarchical analysis of spatial variation of the rice bacterial blight pathogen across diverse agroecosystems in the Philippines. *Phytopathology, 86*, 241-252.

Bachmann, H., Starrenburg, M. J., Molenaar, D., Kleerebezem, M., & van Hylckama Vlieg, J. E. (2012). Microbial domestication signatures of Lactococcus lactis can be reproduced by experimental evolution. *Genome research, 22*(1), 115-124.

Bae, N., Park, H.-J., Park, H., Kim, M., Do, E., & Han, S.-W. (2018). Elucidating Functions of FleQ in *Xanthomonas oryzae* pv. *oryzae* by Comparative Proteomic and Phenotypic Analyses. *International Journal of Molecular Sciences, 19*(10), 3038. Retrieved from https://www.mdpi.com/1422-0067/19/10/3038

Bae, N., Park, H.-J., Park, H., Kim, M., & Han, S.-W. (2018). Deciphering the functions of the outer membrane porin OprBXo involved in virulence, motility, exopolysaccharide production, biofilm formation and stress tolerance in *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology, 19*(12), 2527-2542. doi:https://doi.org/10.1111/mpp.12727

Balint-Kurti, P. (2019). The plant hypersensitive response: concepts, control and consequences. *Molecular Plant Pathology, 20*(8), 1163-1178. doi:https://doi.org/10.1111/mpp.12821

Baum, D. (2008). Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups. *1*(1), 190.

Bayer-Santos, E., Ceseti, L. d. M., Farah, C. S., & Alvarez-Martinez, C. E. (2019). Distribution, Function and Regulation of Type 6 Secretion Systems of Xanthomonadales. *Frontiers in Microbiology, 10*(1635). doi:10.3389/fmicb.2019.01635

Beales, N. (2004). Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: a review. *Comprehensive reviews in food science and food safety, 3*(1), 1-20.

Beattie, G. A., & Lindow, S. E. (1994). Survival, Growth, and Localization of Epiphytic Fitness Mutants of *Pseudomonas syringae* on Leaves. *Applied and Environmental Microbiology, 60*(10), 3790-3798. doi:doi:10.1128/aem.60.10.3790-3798.1994

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software, 5*(56), 2815.

Bigger, J. (1944). Treatment of Staphyloeoeeal Infections with Penicillin by Intermittent Sterilisation. *Lancet*, 497-500.

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., & Hijmans, R. (2018). rgdal: Bindings for the Geospatial Data Abstraction Library 2017. *R package version 0.8-13*.

Bogino, P. C., Oliva, M. d. l. M., Sorroche, F. G., & Giordano, W. (2013). The role of bacterial biofilms and surface components in plant-bacterial associations.

International Journal of Molecular Sciences, 14(8), 15838-15859. doi:10.3390/ijms140815838

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England), 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Boller, T., & Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu Rev Plant Biol, 60*, 379-406. doi:10.1146/annurev.arplant.57.032905.105346

Borgeaud, S., Metzger, L. C., Scrignari, T., & Blokesch, M. (2015). The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science, 347*(6217), 63-67. doi:10.1126/science.1260064

Bridges, A. A., & Bassler, B. L. (2021). Inverse regulation of Vibrio cholerae biofilm dispersal by polyamine signals. *eLife, 10*, e65487. doi:10.7554/eLife.65487

Brown, J. L., Ross, T., McMeekin, T. A., & Nichols, P. D. (1997). Acid habituation of Escherichia coli and the potential role of cyclopropane fatty acids in low pH tolerance. *International journal of food microbiology, 37*(2-3), 163-173.

Büttner, D., & Bonas, U. (2002). Getting across—bacterial type III effector proteins on their way to the plant cell. *The EMBO Journal, 21*(20), 5313-5322. doi:https://doi.org/10.1093/emboj/cdf536

Büttner, D., & He, S. Y. (2009). Type III Protein Secretion in Plant Pathogenic Bacteria. *Plant Physiology, 150*(4), 1656-1664. doi:10.1104/pp.109.139089

Cai, Z., Yuan, Z.-H., Zhang, H., Pan, Y., Wu, Y., Tian, X.-Q., . . . Qian, W. (2017). Fatty acid DSF binds and allosterically activates histidine kinase RpfC of phytopathogenic bacterium *Xanthomonas campestris* pv. *campestris* to regulate quorum-sensing and virulence. *PLOS Pathogens, 13*(4), e1006304. doi:10.1371/journal.ppat.1006304

Campbell, J. W., & Cronan, J. E. (2002). The Enigmatic *Escherichia coli fadE* Gene Is *yafH*. *Journal of bacteriology, 184*(13), 3759-3764. doi:doi:10.1128/JB.184.13.3759-3764.2002

Cardwell, K., Dennis, G., Flannery, A. R., Fletcher, J., Luster, D., Nakhla, M., . . . Walsh, C. (2018). Principles of Diagnostic Assay Validation for Plant Pathogens: A Basic Review of Concepts. *Plant Health Progress, 19*(4), 272-278.

Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K. A., Giovani, B., Saunders, D. G. O., . . . Tohme, J. (2019). A global surveillance system for crop diseases. *Science, 364*(6447), 1237. doi:10.1126/science.aaw1572

Casabuono, A., Petrocelli, S., Ottado, J., Orellano, E. G., & Couto, A. S. (2011). Structural analysis and involvement in plant innate immunity of Xanthomonas axonopodis pv. citri lipopolysaccharide. *Journal of Biological Chemistry, 286*(29), 25628-25643.

Castiblanco, L. F., & Sundin, G. W. (2016). New insights on molecular regulation of biofilm formation in plant-associated bacteria. *Journal of Integrative Plant Biology, 58*(4), 362-372. doi:https://doi.org/10.1111/jipb.12428

Caten, C. E. (1987). The concept of race in plant pathology. In M. S. Wolfe & C. E. Caten (Eds.), *Populations of plant pathogens: their dynamics and genetics*. Oxford: Blackwell Scientific Publications.

Chang, K.-W., Weng, S.-F., & Tseng, Y.-H. (2001). UDP-Glucose Dehydrogenase Gene of *Xanthomonas campestris* Is Required for Virulence. *Biochemical and Biophysical Research Communications, 287*(2), 550-555. doi:https://doi.org/10.1006/bbrc.2001.5591

Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2017). Shiny: web application framework for R. *R package version, 1*(5).

Chapuis, E., Pagès, S., Emelianoff, V., Givaudan, A., & Ferdy, J.-B. (2011). Virulence and pathogen multiplication: a serial passage experiment in the hypervirulent bacterial insect-pathogen *Xenorhabdus nematophila*. *PLoS One, 6*(1), e15872.

Chen, H., Kandel, P. P., Cruz, L. F., Cobine, P. A., & De La Fuente, L. (2017). The Major Outer Membrane Protein MopB Is Required for Twitching Movement and Affects Biofilm Formation and Virulence in Two *Xylella fastidiosa* strains. *Molecular Plant-Microbe Interactions®, 30*(11), 896-905. doi:10.1094/mpmi-07-17-0161-r

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., . . . Frommer, W. B. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature, 468*(7323), 527-532. doi:10.1038/nature09606

Cheng, J., Karambelkar, B., & Xie, Y. (2018). Leaflet: Create interactive web maps with the javascript'leaflet'library. *R package version, 2*(2).

Cheng, L., Connor, T. R., Sirén, J., Aanensen, D. M., & Corander, J. (2013). Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Molecular biology and evolution, 30*(5), 1224-1228.

Chien, C.-C., Chou, M.-Y., Chen, C.-Y., & Shih, M.-C. (2019). Analysis of genetic diversity of Xanthomonas oryzae pv. oryzae populations in Taiwan. *Scientific reports, 9*(1), 1-15.

Chin, C.-Y., Hara, Y., Ghazali, A.-K., Yap, S.-J., Kong, C., Wong, Y.-C., . . . Nathan, S. (2015). Global transcriptional analysis of *Burkholderia pseudomallei* high and low biofilm producers reveals insights into biofilm production and virulence. *BMC Genomics, 16*(1), 471. doi:10.1186/s12864-015-1692-0

Cho, J.-H., Yoon, J.-M., Lee, S.-W., Noh, Y.-H., & Cha, J.-S. (2013). *Xanthomonas oryzae* pv. *oryzae* RpfE Regulates Virulence and Carbon Source Utilization without Change of the DSF Production. *Plant Pathol J, 29*(4), 364-373. doi:10.5423/PPJ.OA.06.2013.0057

Choe, D., Lee, J. H., Yoo, M., Hwang, S., Sung, B. H., Cho, S., . . . Cho, B.-K. (2019). Adaptive laboratory evolution of a genome-reduced Escherichia coli. *Nature communications, 10*(1), 1-14.

Choi, Y., Kim, N., Mannaa, M., Kim, H., Park, J., Jung, H., . . . Seo, Y.-S. (2020). Characterization of Type VI Secretion System in *Xanthomonas oryzae* pv. *oryzae* and Its Role in Virulence to Rice. *The plant pathology journal, 36*(3), 289-296. doi:10.5423/PPJ.NT.02.2020.0026

Cianciotto, N. P., & White, R. C. (2017). Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. *Infection and Immunity, 85*(5), e00014-00017. doi:10.1128/IAI.00014-17

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin), 6*(2), 80-92. doi:10.4161/fly.19695

Coburn, B., Sekirov, I., & Finlay, B. B. (2007). Type III secretion systems and disease. *Clinical microbiology reviews, 20*(4), 535-549. doi:10.1128/CMR.00013-07

Cole, J. N., Aziz, R. K., Kuipers, K., Timmer, A. M., Nizet, V., & Sorge, N. M. v. (2012). A Conserved UDP-Glucose Dehydrogenase Encoded outside the *hasABC* Operon Contributes to Capsule Biogenesis in Group A

Streptococcus. *Journal of bacteriology, 194*(22), 6154-6161. doi:doi:10.1128/JB.01317-12

Conforte, V. P., Malamud, F., Yaryura, P. M., Toum Terrones, L., Torres, P. S., De Pino, V., . . . Vojnov, A. A. (2019). The histone-like protein HupB influences biofilm formation and virulence in *Xanthomonas citri* ssp. *citri* through the regulation of flagellar biosynthesis. *Molecular Plant Pathology, 20*(4), 589-598. doi:https://doi.org/10.1111/mpp.12777

Costerton, J. W., Cheng, K., Geesey, G. G., Ladd, T. I., Nickel, J. C., Dasgupta, M., & Marrie, T. J. (1987). Bacterial biofilms in nature and disease. *Annual Reviews in Microbiology, 41*(1), 435-464.

Coulthurst, S. (2019). The Type VI secretion system: a versatile bacterial weapon. *Microbiology, 165*(5), 503-515. doi:https://doi.org/10.1099/mic.0.000789

Cremers, A., Coolen, J., Bleeker-Rovers, C., van der Geest-Blankert, A., Haverkate, D., Hendriks, H., . . . Liem, D. (2020). Surveillance-embedded genomic outbreak resolution of methicillin-susceptible Staphylococcus aureus in a neonatal intensive care unit. *Scientific reports, 10*(1), 1-10.

Cursino, L., Athinuwat, D., Patel, K. R., Galvani, C. D., Zaini, P. A., Li, Y., . . . Mowery, P. (2015). Characterization of the *Xylella fastidiosa* PD1671 Gene Encoding Degenerate c-di-GMP GGDEF/EAL Domains, and Its Role in the Development of Pierce's Disease. *PLoS One, 10*(3), e0121851. doi:10.1371/journal.pone.0121851

Cursino, L., Galvani, C. D., Athinuwat, D., Zaini, P. A., Li, Y., De La Fuente, L., . . . Mowery, P. (2011). Identification of an Operon, Pil-Chp, That Controls Twitching Motility and Virulence in *Xylella fastidiosa*. *Molecular Plant-Microbe Interactions®, 24*(10), 1198-1206. doi:10.1094/mpmi-10-10-0252

d'Enfert, C., Ryter, A., & Pugsley, A. P. (1987). Cloning and expression in *Escherichia coli* of the *Klebsiella pneumoniae* genes for production, surface localization and secretion of the lipoprotein pullulanase. *Embo j, 6*(11), 3531-3538.

da Silva, F. R., Vettore, A. L., Kemper, E. L., Leite, A., & Arruda, P. (2001). Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *FEMS Microbiology Letters, 203*(2), 165-171.

Darling, A. E., Miklós, I., & Ragan, M. A. (2008). Dynamics of genome rearrangement in bacterial populations. *PLoS Genet, 4*(7), e1000128.

Darwin, C. (1859). *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life* (6th ed.). London: J. Murray.

Daubin, V., Gouy, M., & Perrière, G. (2002). A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome research, 12*(7), 1080-1090. doi:10.1101/gr.187002

de la Fuente-Núñez, C., Reffuveille, F., Fernández, L., & Hancock, R. E. (2013). Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. *Curr Opin Microbiol, 16*(5), 580-589. doi:10.1016/j.mib.2013.06.013

Desai, S. K., & Kenney, L. J. (2019). Switching Lifestyles Is an *in vivo* Adaptive Strategy of Bacterial Pathogens. *Frontiers in Cellular and Infection Microbiology, 9*(421). doi:10.3389/fcimb.2019.00421

Dettman, J. R., Rodrigue, N., Melnyk, A. H., Wong, A., Bailey, S. F., & Kassen, R. (2012). Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular ecology, 21*(9), 2058-2077.

Didelot, X., & Wilson, D. J. (2015). ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol, 11*(2), e1004041.

Dodds, P. N. (2010). Genome evolution in plant pathogens. *Science, 330*, 1486-1487.

Dodge, G. J., Patel, A., Jaremko, K. L., McCammon, J. A., Smith, J. L., & Burkart, M. D. (2019). Structural and dynamical rationale for fatty acid unsaturation in *Escherichia coli*. *Proceedings of the National Academy of Sciences, 116*(14), 6775-6783. doi:10.1073/pnas.1818686116

Donlan, R. M., & Costerton, J. W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clinical microbiology reviews, 15*(2), 167-193. doi:10.1128/CMR.15.2.167-193.2002

Dossa, G. S., Sparks, A., Vera Cruz, C., & Oliva, R. (2015). Decision tools for bacterial blight resistance gene deployment in rice-based agricultural ecosystems. *Frontiers in Plant Science, 6*(305). doi:10.3389/fpls.2015.00305

Dragosits, M., & Mattanovich, D. (2013). Adaptive laboratory evolution–principles and applications for biotechnology. *Microbial cell factories, 12*(1), 1-17.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics, 5*(1), 113. doi:10.1186/1471-2105-5-113

Eisenreich, W., Rudel, T., Heesemann, J., & Goebel, W. (2021). Persistence of Intracellular Bacterial Pathogens—With a Focus on the Metabolic Perspective. *Frontiers in Cellular and Infection Microbiology, 10*(847). doi:10.3389/fcimb.2020.615450

Eom, J.-S., Luo, D., Atienza-Grande, G., Yang, J., Ji, C., Thi Luu, V., . . . Frommer, W. B. (2019). Diagnostic kit for rice blight resistance. *Nature Biotechnology, 37*(11), 1372-1379. doi:10.1038/s41587-019-0268-y

FAOSTAT. (2021). Production/Yield quantities of Rice, paddy in World. Retrieved October 30, 2021, from Food and Agriculture Organization of the United Nations

Fatima, U., & Senthil-Kumar, M. (2015). Plant and pathogen nutrient acquisition strategies. *Frontiers in Plant Science, 6*(750). doi:10.3389/fpls.2015.00750

Fauvet, B., Finka, A., Castanié-Cornet, M.-P., Cirinesi, A.-M., Genevaux, P., Quadroni, M., & Goloubinoff, P. (2021). Bacterial Hsp90 Facilitates the Degradation of Aggregation-Prone Hsp70–Hsp40 Substrates. *Frontiers in Molecular Biosciences, 8*(225). doi:10.3389/fmolb.2021.653073

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution, 17*(6), 368-376. doi:10.1007/BF01734359

Fernandez, L., Mercader, J. M., Planas-Fèlix, M., & Torrents, D. (2014). Adaptation to environmental factors shapes the organization of regulatory regions in microbial communities. *BMC Genomics, 15*(1), 1-12.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual review of phytopathology, 9*(1), 275-296.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third edition ed.). Thousand Oaks CA: Sage.

Fry, S., & Milholland, R. (1990). Response of resistant, tolerant, and susceptible grapevine tissues to invasion by the Pierce's disease bacterium, *Xylella fastidiosa*. *Phytopathology, 80*(1), 66-69.

Fu, Y., Mason, A. S., Zhang, Y., & Yu, H. (2021). Identification and Development of KASP Markers for Novel Mutant BnFAD2 Alleles Associated With

Elevated Oleic Acid in *Brassica napus*. *Frontiers in Plant Science, 12*. doi:10.3389/fpls.2021.715633

Fujita, Y., Matsuoka, H., & Hirooka, K. (2007). Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology, 66*(4), 829-839. doi:https://doi.org/10.1111/j.1365-2958.2007.05947.x

Gardy, J. L., & Loman, N. J. (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics, 19*(1), 9.

Georgelis, N., Nikolaidis, N., & Cosgrove, D. J. (2015). Bacterial expansins and related proteins from the world of microbes. *Appl Microbiol Biotechnol, 99*(9), 3807-3823. doi:10.1007/s00253-015-6534-0

Gilbert, H. J., & Hazlewood, G. P. (1993). Bacterial cellulases and xylanases. *Microbiology, 139*(2), 187-194.

Glickman, M. S., Cox, J. S., & Jacobs, W. R. (2000). A Novel Mycolic Acid Cyclopropane Synthetase Is Required for Cording, Persistence, and Virulence of *Mycobacterium tuberculosis*. *Molecular Cell, 5*(4), 717-727. doi:https://doi.org/10.1016/S1097-2765(00)80250-6

Gluck-Thaler, E., Cerutti, A., Perez-Quintero, A. L., Butchacas, J., Roman-Reyna, V., Madhavan, V. N., . . . Jacobs, J. M. (2020). Repeated gain and loss of a single gene modulates the evolution of vascular plant pathogen lifestyles. *Science Advances, 6*(46), eabc4516. doi:doi:10.1126/sciadv.abc4516

Gonzalez, C., Szurek, B., Manceau, C., Mathieu, T., Séré, Y., & Verdier, V. (2007). Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. *Molecular plant-microbe interactions, 20*(5), 534-546.

Gouran, H., Gillespie, H., Nascimento, R., Chakraborty, S., Zaini, P. A., Jacobson, A., . . . Dandekar, A. M. (2016). The Secreted Protease PrtA Controls Cell Growth, Biofilm Formation and Pathogenicity in *Xylella fastidiosa*. *Scientific Reports, 6*(1), 31098. doi:10.1038/srep31098

Grames, E., Stillman, A., Tingley, M., & Elphick, C. (2020). litsearchr: Automated Search Term Selection and Search Strategy for Systematic Reviews (Version R package version 1.0.0). Retrieved from https://github.com/elizagrames/litsearchr

Grames, E. M., Stillman, A. N., Tingley, M. W., & Elphick, C. S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution, 10*(10), 1645-1654. doi:https://doi.org/10.1111/2041-210X.13268

Granato, L. M., Picchi, S. C., Andrade, M. d. O., Martins, P. M. M., Takita, M. A., Machado, M. A., . . . Becker, A. (2019). The *ecnA* Antitoxin Is Important Not Only for Human Pathogens: Evidence of Its Role in the Plant Pathogen *Xanthomonas citri* subsp. *citri*. *Journal of bacteriology, 201*(20), e00796-00718. doi:doi:10.1128/JB.00796-18

Green, E. R., Mecsas, J., & Kudva, I. T. (2016). Bacterial Secretion Systems: An Overview. *Microbiology Spectrum, 4*(1), 4.1.13. doi:doi:10.1128/microbiolspec.VMBF-0012-2015

Gregory, T. R. (2009). Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach, 2*(2), 156-175. doi:10.1007/s12052-009-0128-1

Grinberg, M., Orevi, T., & Kashtan, N. (2019). Bacterial surface colonization, preferential attachment and fitness under periodic stress. *PLOS Computational Biology, 15*(3), e1006815. doi:10.1371/journal.pcbi.1006815

Grogan, D. W., & Cronan, J. E. (1997). Cyclopropane ring formation in membrane lipids of bacteria. *Microbiology and molecular biology reviews, 61*(4), 429-441. doi:doi:10.1128/mmbr.61.4.429-441.1997

Guan, C.-P., Luo, H.-X., Fang, H. E., & Zhou, X.-Z. (2018). Global Transcriptome Changes of Biofilm-Forming *Staphylococcus epidermidis* Responding to Total Alkaloids of *Sophorea alopecuroides*. *Polish journal of microbiology, 67*(2), 223-226. doi:10.21307/pjm-2018-024

Gude, S., Pinçe, E., Taute, K. M., Seinen, A.-B., Shimizu, T. S., & Tans, S. J. (2020). Bacterial coexistence driven by motility and spatial competition. *Nature, 578*(7796), 588-592. doi:10.1038/s41586-020-2033-2

Guidot, A., Jiang, W., Ferdy, J. B., Thébaud, C., Barberis, P., Gouzy, J., & Genin, S. (2014). Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Mol Biol Evol, 31*(11), 2913-2928. doi:10.1093/molbev/msu229

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology, 59*(3), 307-321.

Guo, W., Gao, J., Chen, Q., Ma, B., Fang, Y., Liu, X., . . . Liu, J.-Z. (2019). Crp-Like Protein Plays Both Positive and Negative Roles in Regulating the Pathogenicity of Bacterial Pustule Pathogen *Xanthomonas axonopodis* pv. *glycines*. *Phytopathology®, 109*(7), 1171-1183. doi:10.1094/phyto-07-18-0225-r

Guo, W., Zou, L.-F., Cai, L.-L., & Chen, G.-Y. (2015). Glucose-6-phosphate dehydrogenase is required for extracellular polysaccharide production, cell motility and the full virulence of *Xanthomonas oryzae* pv. *oryzicola*. *Microbial Pathogenesis, 78*, 87-94. doi:https://doi.org/10.1016/j.micpath.2014.11.007

Hall, B. G. (2013). Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution, 30*(5), 1229-1235. doi:10.1093/molbev/mst012

Harper, H., Burridge, A., Winfield, M., Finn, A., Davidson, A., Matthews, D., . . . Barker, G. (2021). Detecting SARS-CoV-2 variants with SNP genotyping. *PLoS One, 16*(2), e0243185. doi:10.1371/journal.pone.0243185

Hartl, F. U., & Hayer-Hartl, M. (2002). Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein. *Science, 295*(5561), 1852-1858. doi:10.1126/science.1068408

Hedden, P. (2003). The genes of the Green Revolution. *TRENDS in Genetics, 19*(1), 5-9.

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology, 8*(1), 15-25. doi:10.1038/nrmicro2259

Hilaire, E., Young, S. A., Willard, L. H., McGee, J. D., Sweat, T., Chittoor, J., . . . Leach, J. E. (2001). Vascular defense responses in rice: peroxidase accumulation in xylem parenchyma cells and xylem wall thickening. *Molecular plant-microbe interactions, 14*(12), 1411-1419.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular biology and evolution, 35*(2), 518-522. doi:10.1093/molbev/msx281

Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R., & Kamoun, S. (2009). Emerging Concepts in Effector Biology of Plant-Associated Organisms. *Molecular Plant-Microbe Interactions®, 22*(2), 115-122. doi:10.1094/mpmi-22-2-0115

Hollich, V., Milchert, L., Arvestad, L., & Sonnhammer, E. L. L. (2005). Assessment of Protein Distance Measures and Tree-Building Methods for Phylogenetic Tree Reconstruction. *Molecular Biology and Evolution, 22*(11), 2257-2264. doi:10.1093/molbev/msi224

Hood, R. D., Singh, P., Hsu, F., Güvener, T., Carl, M. A., Trinidad, R. R. S., . . . Mougous, J. D. (2010). A Type VI Secretion System of *Pseudomonas*

aeruginosa Targets a Toxin to Bacteria. *Cell Host & Microbe, 7*(1), 25-37. doi:https://doi.org/10.1016/j.chom.2009.12.007

Hopkins, C. M., White, F., Choi, S., Guo, A., & Leach, J. (1992). Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *Mol Plant Microbe Interact, 5*(6), 451-459.

Horinouchi, T., Suzuki, S., Hirasawa, T., Ono, N., Yomo, T., Shimizu, H., & Furusawa, C. (2015). Phenotypic convergence in bacterial adaptive evolution to ethanol stress. *BMC evolutionary biology, 15*(1), 1-14.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom J, 50*(3), 346-363. doi:10.1002/bimj.200810425

Hsiao, Y.-M., Liu, Y.-F., Fang, M.-C., & Song, W.-L. (2011). XCC2731, a GGDEF domain protein in *Xanthomonas campestris*, is involved in bacterial attachment and is positively regulated by Clp. *Microbiological Research, 166*(7), 548-565. doi:https://doi.org/10.1016/j.micres.2010.11.003

Huang, C.-L., Pu, P.-H., Huang, H.-J., Sung, H.-M., Liaw, H.-J., Chen, Y.-M., . . . Gojobori, T. (2015). Ecological genomics in *Xanthomonas*: the nature of genetic adaptation with homologous recombination and host shifts. *BMC Genomics, 16*(1), 1-13.

Huang, T.-P., Lu, K.-M., & Chen, Y.-H. (2013). A Novel Two-Component Response Regulator Links *rpf* with Biofilm Formation and Virulence of *Xanthomonas axonopodis* pv. *citri*. *PLoS One, 8*(4), e62824. doi:10.1371/journal.pone.0062824

Hueck, C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and molecular biology reviews, 62*(2), 379-433.

Ionescu, M., Baccari, C., Da Silva, A. M., Garcia, A., Yokota, K., & Lindow, S. E. (2013). Diffusible signal factor (DSF) synthase RpfF of *Xylella fastidiosa* is a multifunction protein also required for response to DSF. *Journal of bacteriology, 195*(23), 5273-5284. doi:10.1128/JB.00713-13

Islam, M. R., Alam, M. S., Khan, A. I., Hossain, I., Adam, L. R., & Daayf, F. (2016). Analyses of genetic diversity of bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae* using IS1112 in Bangladesh. *Comptes Rendus Biologies, 339*(9), 399-407. doi:https://doi.org/10.1016/j.crvi.2016.06.002

Ivosev, G., Burton, L., & Bonner, R. (2008). Dimensionality reduction and visualization in principal component analysis. *Analytical chemistry, 80*(13), 4933-4944.

Jacobs, J. M., Babujee, L., Meng, F., Milling, A., Allen, C., & Handelsman, J. (2012). The *In Planta* Transcriptome of *Ralstonia solanacearum*: Conserved Physiological and Virulence Strategies during Bacterial Wilt of Tomato. *mBio, 3*(4), e00114-00112. doi:doi:10.1128/mBio.00114-12

Jeung, J. U., Heu, S. G., Shin, M. S., Vera Cruz, C. M., & Jena, K. K. (2006). Dynamics of *Xanthomonas oryzae* pv. *oryzae* Populations in Korea and Their Relationship to Known Bacterial Blight Resistance Genes. *Phytopathology®, 96*(8), 867-875. doi:10.1094/phyto-96-0867

Jha, G., Rajeshwari, R., & Sonti, R. V. (2005). Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. *Molecular Plant-Microbe Interactions, 18*(9), 891-898.

Jiang, N., Yan, J., Liang, Y., Shi, Y., He, Z., Wu, Y., . . . Peng, J. (2020). Resistance Genes and their Interactions with Bacterial Blight/Leaf Streak Pathogens (*Xanthomonas oryzae*) in Rice (*Oryza sativa* L.)—an Updated Review. *Rice, 13*(1), 3. doi:10.1186/s12284-019-0358-y

Jones, J. D., & Dangle, J. L. (2006). The plant immune system. *Nature, 444*, 323-329.

Kamoun, S., & Kado, C. I. (1990). Phenotypic Switching Affecting Chemotaxis, Xanthan Production, and Virulence in *Xanthomonas campestris*. *Applied and Environmental Microbiology, 56*(12), 3855-3860. doi:doi:10.1128/aem.56.12.3855-3860.1990

Kandel, P. P., Chen, H., Fuente, L. D. L., & Stabb, E. V. (2018). A Short Protocol for Gene Knockout and Complementation in *Xylella fastidiosa* Shows that One of the Type IV Pilin Paralogs (PD1926) Is Needed for Twitching while Another (PD1924) Affects Pilus Number and Location. *Applied and Environmental Microbiology, 84*(18), e01167-01118. doi:doi:10.1128/AEM.01167-18

Kang, M., Kim, K., Choe, D., Cho, S., Kim, S. C., Palsson, B., & Cho, B.-K. (2019). Inactivation of a mismatch-repair system diversifies genotypic landscape of Escherichia coli during adaptive laboratory evolution. *Frontiers in microbiology, 10*, 1845.

Kang, X.-M., Wang, F.-F., Zhang, H., Zhang, Q., Qian, W., & Kivisaar, M. (2015). Genome-Wide Identification of Genes Necessary for Biofilm Formation by Nosocomial Pathogen *Stenotrophomonas maltophilia* Reveals that Orphan Response Regulator FsnR Is a Critical Modulator. *Applied and Environmental Microbiology, 81*(4), 1200-1209. doi:doi:10.1128/AEM.03408-14

Kaplan, J. B. (2010). Biofilm dispersal: mechanisms, clinical implications, and potential therapeutic uses. *Journal of dental research, 89*(3), 205-218. doi:10.1177/0022034509359403

Karasov, T. L., Horton, M. W., & Bergelson, J. (2014). Genomic variability as a driver of plant-pathogen coevolution? *Current opinion in plant biology, 18*, 24-30. doi:10.1016/j.pbi.2013.12.003

Karganilla, A., Natural, M. P., & Ou, S. H. (1973). A comparative study of culture media for *Xanthomonas oryzae* [bacterial leaf blight of rice]. *Philippine Agriculturist, 57*(PH19750009708).

Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. (Version R package version 0.4.0.). Retrieved from https://CRAN.R-project.org/package=ggpubr

Kauffman, H., Reddy, A., Hsieh, S., & Merca, S. (1973). An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. *Plant Dis. Rep, 57*, 537-541.

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., & Whitlock, M. C. (2012). Experimental evolution. *Trends in ecology & evolution, 27*(10), 547-560.

Kelly, F. X., Dapsis, K. J., & Lauffenburger, D. A. (1988). Effect of bacterial chemotaxis on dynamics of microbial competition. *Microbial Ecology, 16*(2), 115-131. doi:10.1007/BF02018908

Khaeruni, A., & Wijayanto, T. (2013). Pathotype grouping of *Xanthomonas oryzae* pv. *oryzae* isolates from South Sulawesi and Southeast Sulawesi. *AGRIVITA, Journal of Agricultural Science, 35*(2), 138-144.

Khan, M., Rafi, A., Abbas, A., Ali, T., & Hassan, A. (2015). Assessment of yield lossess caused by bacterial blight of rice in upper dir, khyber Pakhtunkhwa province. *Asian J. Agric. Biol, 3*(2), 74-78.

Khush, G. (2000). Taxonomy and origin of rice. *Aromatic rices*, 5-13.

Killiny, N., Martinez, R. H., Dumenyo, C. K., Cooksey, D. A., & Almeida, R. P. (2013). The exopolysaccharide of *Xylella fastidiosa* is essential for biofilm formation, plant virulence, and vector transmission. *Mol Plant Microbe Interact, 26*(9), 1044-1053. doi:10.1094/mpmi-09-12-0211-r

Kim, S.-I., Song, J. T., Jeong, J.-Y., & Seo, H. S. (2016). Niclosamide inhibits leaf blight caused by *Xanthomonas oryzae* in rice. *Scientific Reports, 6*(1), 21209. doi:10.1038/srep21209

Kim, S.-M., & Reinke, R. F. (2019). A novel resistance gene for bacterial blight in rice, Xa43(t) identified by GWAS, confirmed by QTL mapping using a bi-parental population. *PLOS ONE, 14*(2), e0211775. doi:10.1371/journal.pone.0211775

Knöppel, A., Knopp, M., Albrecht, L. M., Lundin, E., Lustig, U., Näsvall, J., & Andersson, D. I. (2018). Genetic Adaptation to Growth Under Laboratory Conditions in *Escherichia coli* and *Salmonella enterica*. *Frontiers in Microbiology, 9*(756). doi:10.3389/fmicb.2018.00756

Koczan, J. M., McGrath, M. J., Zhao, Y., & Sundin, G. W. (2009). Contribution of *Erwinia amylovora* exopolysaccharides amylovoran and levan to biofilm formation: implications in pathogenicity. *Phytopathology, 99*(11), 1237-1244. doi:10.1094/phyto-99-11-1237

Koonin, E. V., & Galperin, M. Y. (2003). Evolutionary Concept in Genetics and Genomics. In *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics*. Boston: Kluwer Academic.

Kraiselburd, I., Alet, A. I., Tondo, M. L., Petrocelli, S., Daurelio, L. D., Monzón, J., . . . Orellano, E. G. (2012). A LOV Protein Modulates the Physiological Attributes of *Xanthomonas axonopodis* pv. *citri* Relevant for Host Plant Colonization. *PLoS One, 7*(6), e38226. doi:10.1371/journal.pone.0038226

Kuchma, S. L., Connolly, J. P., & O'Toole, G. A. (2005). A Three-Component Regulatory System Regulates Biofilm Maturation and Type III Secretion in *Pseudomonas aeruginosa*. *Journal of bacteriology, 187*(4), 1441-1454. doi:doi:10.1128/JB.187.4.1441-1454.2005

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of statistical software, 28*(5), 26. doi:10.18637/jss.v028.i05

Kumar Verma, R., Samal, B., & Chatterjee, S. (2018). *Xanthomonas oryzae* pv. *oryzae* chemotaxis components and chemoreceptor Mcp2 are involved in the sensing of constituents of xylem sap and contribute to the regulation of virulence-associated functions and entry into rice. *Molecular Plant Pathology, 19*(11), 2397-2415. doi:10.1111/mpp.12718

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses.

Molecular biology and evolution, 34(3), 772-773. doi:10.1093/molbev/msw260

Lang, J. M., Hamilton, J. P., Diaz, M. G. Q., Van Sluys, M. A., Burgos, M. R. G., Vera Cruz, C. M., . . . Leach, J. E. (2010). Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. *Plant disease, 94*(3), 311-319.

Lang, J. M., Langlois, P., Nguyen, M. H. R., Triplett, L. R., Purdie, L., Holton, T. A., . . . Leach, J. E. (2014). Sensitive detection of *Xanthomonas oryzae* Pathovars *oryzae* and *oryzicola* by loop-mediated isothermal amplification. *Applied and Environmental Microbiology, 80*(15), 4519-4530. doi:10.1128/AEM.00274-14

Law, J. H. (1971). Biosynthesis of cyclopropane rings. *Accounts of Chemical Research, 4*(6), 199-203.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., . . . Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology, 9*(8), e1003118. doi:10.1371/journal.pcbi.1003118

Le, S. Q., & Gascuel, O. (2008). An improved general amino acid replacement matrix. *Molecular Biology and Evolution, 25*(7), 1307-1320.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics, 34*(18), 3094-3100. doi:10.1093/bioinformatics/bty191

Li, H., Yu, C., Chen, H., Tian, F., & He, C. (2015). PXO_00987, a putative acetyltransferase, is required for flagellin glycosylation, and regulates flagellar motility, exopolysaccharide production, and biofilm formation in *Xanthomonas oryzae* pv. *oryzae*. *Microbial Pathogenesis, 85*, 50-57. doi:https://doi.org/10.1016/j.micpath.2015.06.001

Li, J., & Wang, N. (2012). The gpsX gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in *Xanthomonas citri* subsp. *citri*. *BMC Microbiology, 12*(1), 31. doi:10.1186/1471-2180-12-31

Li, K., Wu, G., Liao, Y., Zeng, Q., Wang, H., & Liu, F. (2020). RpoN1 and RpoN2 play different regulatory roles in virulence traits, flagellar biosynthesis, and basal metabolism in *Xanthomonas campestris*. *Molecular Plant Pathology, 21*(7), 907-922. doi:https://doi.org/10.1111/mpp.12938

Li, S.-J., Hua, Z.-S., Huang, L.-N., Li, J., Shi, S.-H., Chen, L.-X., . . . Shu, W.-S. (2014). Microbial communities evolve faster in extreme environments. *Scientific Reports, 4*(1), 1-9.

Li, T., Zhan, Z., Lin, Y., Lin, M., Xie, Q., Chen, Y., . . . Li, C. (2019). Biosynthesis of Amino Acids in *Xanthomonas oryzae* pv. *oryzae* Is Essential to Its Pathogenicity. *Microorganisms, 7*(12), 693.

Liao, C.-T., Chiang, Y.-C., & Hsiao, Y.-M. (2019). Functional characterization and proteomic analysis of lolA in *Xanthomonas campestris* pv. *campestris*. *BMC Microbiology, 19*(1), 20. doi:10.1186/s12866-019-1387-9

Liao, C.-T., Liu, Y.-F., Chiang, Y.-C., Lo, H.-H., Du, S.-C., Hsu, P.-C., & Hsiao, Y.-M. (2016). Functional characterization and transcriptome analysis reveal multiple roles for *prc* in the pathogenicity of the black rot pathogen *Xanthomonas campestris* pv. *campestris*. *Research in microbiology, 167*(4), 299-312. doi:https://doi.org/10.1016/j.resmic.2016.01.002

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research, 47*(8), e47-e47. doi:10.1093/nar/gkz114

Lindgren, P. B., Peet, R. C., & Panopoulos, N. J. (1986). Gene cluster of *Pseudomonas syringae* pv. "phaseolicola" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. *Journal of bacteriology, 168*(2), 512-522. doi:10.1128/jb.168.2.512-522.1986

Liu, L., Zhu, K., Wurzburger, N., & Zhang, J. (2020). Relationships between plant diversity and soil microbial diversity vary across taxonomic groups and spatial scales. *Ecosphere, 11*(1), e02999. doi:https://doi.org/10.1002/ecs2.2999

Liu, W., Zhao, H., Qiu, Z., Jin, M., Yang, D., Xu, Q., . . . Shen, Z. (2020). Identifying geographic origins of the *Escherichia coli* isolates from food by a method based on single-nucleotide polymorphisms. *Journal of Microbiological Methods, 168*, 105807.

Liu, X., Yan, Y., Wu, H., Zhou, C., & Wang, X. (2019). Biological and transcriptomic studies reveal *hfq* is required for swimming, biofilm formation and stress response in *Xanthomonas axonpodis* pv. *citri*. *BMC Microbiology, 19*(1), 103. doi:10.1186/s12866-019-1476-9

Lore, J. S., Vikal, Y., Hunjan, M. S., Goel, R. K., Bharaj, T. S., & Raina, G. L. (2011). Genotypic and Pathotypic Diversity of *Xanthomonas oryzae* pv. *oryzae*, the Cause of Bacterial Blight of Rice in Punjab State of India.

Journal of Phytopathology, 159(7-8), 479-487. doi:https://doi.org/10.1111/j.1439-0434.2011.01789.x

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Lu, X.-H., An, S.-Q., Tang, D.-J., McCarthy, Y., Tang, J.-L., Dow, J. M., & Ryan, R. P. (2012). RsmA Regulates Biofilm Formation in *Xanthomonas campestris* through a Regulatory Network Involving Cyclic di-GMP and the Clp Transcription Factor. *PLoS One, 7*(12), e52646. doi:10.1371/journal.pone.0052646

Macho, Alberto P., & Zipfel, C. (2014). Plant PRRs and the Activation of Innate Immune Signaling. *Molecular Cell, 54*(2), 263-272. doi:https://doi.org/10.1016/j.molcel.2014.03.028

Madden, L. V., Hughes, G., & Van Den Bosch, F. (2007). The study of plant disease epidemics.

Malamud, F., Torres, P. S., Roeschlin, R., Rigano, L. A., Enrique, R., Bonomi, H. R., . . . Vojnov, A. A. (2011). The *Xanthomonas axonopodis* pv. *citri* flagellum is required for mature biofilm and canker development. *Microbiology, 157*(3), 819-829. doi:10.1099/mic.0.044255-0

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., . . . Machado, M. A. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology, 13*(6), 614-629.

Marques, L. L. R., Ceri, H., Manfio, G. P., Reid, D. M., & Olson, M. E. (2002). Characterization of Biofilm Formation by *Xylella fastidiosa* In Vitro. *Plant disease, 86*(6), 633-638. doi:10.1094/pdis.2002.86.6.633

Martins, P. M. M., Merfa, M. V., Takita, M. A., & De Souza, A. A. (2018). Persistence in Phytopathogenic Bacteria: Do We Know Enough? *Frontiers in Microbiology, 9*(1099). doi:10.3389/fmicb.2018.01099

Martins, P. M. M., Wood, T. K., & de Souza, A. A. (2021). Persister Cells Form in the Plant Pathogen *Xanthomonas citri* subsp. *citri* under Different Stress Conditions. *Microorganisms, 9*(2), 384. doi:10.3390/microorganisms9020384

Maturin, L., & Peeler, J. (2001). BAM: Aerobic plate count. *Bacteriological Analytical Manual US Food and Drug Administration, New Hampshire Avenue Silver Spring USA*.

Maurer, L. M., Yohannes, E., Bondurant, S. S., Radmacher, M., & Slonczewski, J. L. (2005). pH Regulates Genes for Flagellar Motility, Catabolism, and Oxidative Stress in *Escherichia coli* K-12. *Journal of bacteriology, 187*(1), 304-319. doi:doi:10.1128/JB.187.1.304-319.2005

McCann, H. C., & Guttman, D. S. (2008). Evolution of the type III secretion system and its effectors in plant–microbe interactions. *New Phytologist, 177*(1), 33-47. doi:https://doi.org/10.1111/j.1469-8137.2007.02293.x

Meaden, S., & Koskella, B. (2017). Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. *Mol Ecol, 26*(7), 1790-1801. doi:10.1111/mec.14060

Melnyk, A. H., Wong, A., & Kassen, R. (2015). The fitness costs of antibiotic resistance mutations. *Evolutionary applications, 8*(3), 273-283. doi:10.1111/eva.12196

Melnyk, R. A., Hossain, S. S., & Haney, C. H. (2019). Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated *Pseudomonas*. *The ISME journal, 13*(6), 1575-1588.

Meng, Q.-L., Tang, D.-J., Fan, Y.-Y., Li, Z.-J., Zhang, H., He, Y.-Q., . . . Tang, J.-L. (2011). Effect of interactions between Mip and PrtA on the full extracellular protease activity of *Xanthomonas campestris* pathovar *campestris*. *FEMS Microbiology Letters, 323*(2), 180-187. doi:10.1111/j.1574-6968.2011.02377.x

Merritt, P. M., Danhorn, T., & Fuqua, C. (2007). Motility and Chemotaxis in *Agrobacterium tumefaciens* Surface Attachment and Biofilm Formation. *Journal of bacteriology, 189*(22), 8005-8014. doi:doi:10.1128/JB.00566-07

Mew, T., Alvarez, A., Leach, J., & Swings, J. (1993). Focus on bacterial blight of rice. *Plant disease, 77*(1), 5-12.

Mew, T., & Vera Cruz, C. (1979). Variability of *Xanthomonas oryzae*: specificity in infection of rice differentials. *Phytopathology, 69*, 152-155.

Mew, T., Vera Cruz, C., & Medalla, E. (1992). Changes in race frequency of *Xanthomonas oryzae* pv. *oryzae* in response to rice cultivars planted in the Philippines. *Plant disease, 76*(10), 1029-1032.

Mew, T. W., Alvarez, A. M., Leach, J., & Swings, J. (1993). Focus on bacterial blight of rice. *Plant disease, 77*(1), 5-12.

Meyer, D., Lauber, E., Roby, D., Arlat, M., & Kroj, T. (2005). Optimization of pathogenicity assays to study the *Arabidopsis thaliana–Xanthomonas campestris* pv. *campestris* pathosystem. *Molecular Plant Pathology, 6*(3), 327-333. doi:https://doi.org/10.1111/j.1364-3703.2005.00287.x

Midha, S., Bansal, K., Kumar, S., Girija, A. M., Mishra, D., Brahma, K., . . . Patil, P. B. (2017). Population genomic insights into variation and evolution of *Xanthomonas oryzae* pv. *oryzae*. *Scientific reports, 7*(1), 1-13.

Mishra, D., Vishnupriya, M. R., Anil, M. G., Konda, K., Raj, Y., & Sonti, R. V. (2013). Pathotype and Genetic Diversity amongst Indian Isolates of *Xanthomonas oryzae* pv. *oryzae*. *PLoS One, 8*(11), e81996. doi:10.1371/journal.pone.0081996

Mori, Y., Inoue, K., Ikeda, K., Nakayashiki, H., Higashimoto, C., Ohnishi, K., . . . Hikichi, Y. (2016). The vascular plant-pathogenic bacterium *Ralstonia solanacearum* produces biofilms required for its virulence on the surfaces of tomato cells adjacent to intercellular spaces. *Mol Plant Pathol, 17*(6), 890-902. doi:10.1111/mpp.12335

Morris, C. E., Sands, D. C., Vanneste, J. L., Montarry, J., Oakley, B., Guilbaud, C., . . . Lindow, S. (2010). Inferring the Evolutionary History of the Plant Pathogen *Pseudomonas syringae* from Its Biogeography in Headwaters of Rivers in North America, Europe, and New Zealand. *mBio, 1*(3), e00107-00110. doi:doi:10.1128/mBio.00107-10

Muhammad, M. H., Idris, A. L., Fan, X., Guo, Y., Yu, Y., Jin, X., . . . Huang, T. (2020). Beyond Risk: Bacterial Biofilms and Their Regulating Approaches. *Frontiers in Microbiology, 11*. doi:10.3389/fmicb.2020.00928

Muranaka, L. S., Takita, M. A., Olivato, J. C., Kishi, L. T., & de Souza, A. A. (2012). Global expression profile of biofilm resistance to antimicrobial compounds in the plant-pathogenic bacterium *Xylella fastidiosa* reveals evidence of persister cells. *Journal of bacteriology, 194*(17), 4561-4569. doi:10.1128/JB.00436-12

Nascimento, R., Gouran, H., Chakraborty, S., Gillespie, H. W., Almeida-Souza, H. O., Tu, A., . . . Dandekar, A. M. (2016). The Type II Secreted Lipase/Esterase LesA is a Key Virulence Factor Required for *Xylella fastidiosa* Pathogenesis in Grapevines. *Scientific Reports, 6*(1), 18598. doi:10.1038/srep18598

Naushad, H. S., & Gupta, R. S. (2013). Phylogenomics and Molecular Signatures for Species from the Plant Pathogen-Containing Order Xanthomonadales. *PLOS ONE, 8*(2), e55216. doi:10.1371/journal.pone.0055216

Nelson, R. J., Baraoidan, M. R., Cruz, C. M. V., Yap, I. V., Leach, J. E., Mew, T. W., & Leung, H. (1994). Relationship between phylogeny and pathotype for the bacterial blight pathogen of rice. *Applied and Environmental Microbiology, 60*(9), 3275-3283.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution, 32*(1), 268-274. doi:10.1093/molbev/msu300

Ninio, J. (1991). Transient mutators: a semiquantitative analysis of the influence of translation and transcription errors on mutation rates. *Genetics, 129*(3), 957-962. Retrieved from https://www.genetics.org/content/genetics/129/3/957.full.pdf

Niño-Liu, D. O., Ronald, P. C., & Bogdanove, A. J. (2006). *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Molecular Plant Pathology, 7*(5), 303-324.

Noda, T., & Kaku, H. (1999). Growth of *Xanthomonas oryzae* pv. *oryzae* in planta and in guttation fluid of rice. *Japanese Journal of Phytopathology, 65*(1), 9-14.

Noda, T., Li, C., Li, J., Ochiai, H., Ise, K., & Kaku, H. (2001). Pathogenic diversity of *Xanthomonas oryzae* pv. *oryzae* strains from Yunnan province, China. *Japan Agricultural Research Quarterly: JARQ, 35*(2), 97-103.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., & Hase, T. (2000). Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research, 28*(12), e63-e63.

Notredame, C. (2007). Recent Evolutions of Multiple Sequence Alignment Algorithms. *PLOS Computational Biology, 3*(8), e123. doi:10.1371/journal.pcbi.0030123

O'Toole, G. A., & Kolter, R. (1998). Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Molecular Microbiology, 28*(3), 449-461. doi:https://doi.org/10.1046/j.1365-2958.1998.00797.x

Ogawa, T., Yamamoto, T., Khush, G. S., & Mew, T.-W. (1991). Breeding of near-isogenic lines of rice with single genes for resrstance to bacterial blight pathogen (*Xanthomonas campestris* pv. *oryzae*). *Japanese Journal of Breeding, 41*(3), 523-529.

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature reviews. Genetics, 10*(8), 531-539. doi:10.1038/nrg2603

Ou, S. H. (1985). *Rice diseases*: IRRI.

Oxman, E., Alon, U., & Dekel, E. (2008). Defined order of evolutionary adaptations: experimental evidence. *Evolution: International Journal of Organic Evolution, 62*(7), 1547-1554.

Pan, X., Xu, S., Wu, J., Duan, Y., Zheng, Z., Wang, J., . . . Schottel, J. L. (2018). Ankyrin-Like Protein AnkB Interacts with CatB, Affects Catalase Activity, and Enhances Resistance of *Xanthomonas oryzae* pv. *oryzae* and *Xanthomonas oryzae* pv. *oryzicola* to Phenazine-1-Carboxylic Acid. *Applied and Environmental Microbiology, 84*(4), e02145-02117. doi:doi:10.1128/AEM.02145-17

Pariaud, B., Ravigné, V., Halkett, F., Goyeau, H., Carlier, J., & Lannou, C. (2009). Aggressiveness and its role in the adaptation of plant pathogens. *Plant Pathology, 58*(3), 409-424. doi:https://doi.org/10.1111/j.1365-3059.2009.02039.x

Park, H.-J., Jung, H. W., & Han, S.-W. (2014). Functional and proteomic analyses reveal that wxcB is involved in virulence, motility, detergent tolerance, and biofilm formation in *Xanthomonas campestris* pv. *vesicatoria*. *Biochemical and Biophysical Research Communications, 452*(3), 389-394. doi:https://doi.org/10.1016/j.bbrc.2014.08.076

Park, H., Do, E., Kim, M., Park, H. J., Lee, J., & Han, S. W. (2020). A LysR-Type Transcriptional Regulator LcrX Is Involved in Virulence, Biofilm Formation, Swimming Motility, Siderophore Secretion, and Growth in Sugar Sources in *Xanthomonas axonopodis* pv. *glycines*. *Frontiers in Plant Science, 10*. doi:10.3389/fpls.2019.01657

Parnell, S., Gottwald, T., Cunniffe, N., Alonso Chavez, V., & van Den Bosch, F. (2015). Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. *Proceedings of the Royal Society B: Biological Sciences, 282*(1814), 20151478.

Peeling, R. W., Olliaro, P. L., Boeras, D. I., & Fongwen, N. (2021). Scaling up COVID-19 rapid antigen tests: promises and challenges. *The Lancet infectious diseases, 21*(9), e290-e295.

Pelé, J., Bécu, J.-M., Abdi, H., & Chabbert, M. (2012). Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *BMC Bioinformatics, 13*(1), 133. doi:10.1186/1471-2105-13-133

Petrocelli, S., Tondo, M. L., Daurelio, L. D., & Orellano, E. G. (2012). Modifications of *Xanthomonas axonopodis* pv. *citri* Lipopolysaccharide Affect the Basal Response and the Virulence Process during Citrus Canker. *PLoS One, 7*(7), e40051. doi:10.1371/journal.pone.0040051

Pinstrup-Andersen, P., & Hazell, P. B. R. (1985). The impact of the green revolution and prospects for the future. *Food Reviews International, 1*(1), 1-25. doi:10.1080/87559128509540765

Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., . . . Mekalanos, J. J. (2006). Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proceedings of the National Academy of Sciences, 103*(5), 1528. doi:10.1073/pnas.0510322103

Quibod, I. L., Atieza-Grande, G., Oreiro, E. G., Palmos, D., Nguyen, M. H., Coronejo, S. T., . . . Burgos, M. R. (2020). The green revolution shaped the population structure of the rice pathogen *Xanthomonas oryzae* pv. *oryzae*. *The ISME journal, 14*(2), 492-505.

Quibod, I. L., Perez-Quintero, A., Booher, N. J., Dossa, G. S., Grande, G., Szurek, B., . . . Oliva, R. (2016). Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. *Scientific Reports, 6*(1), 1-11.

R Core Team. (2021). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Radhakrishnan, G. V., Cook, N. M., Bueno-Sancho, V., Lewis, C. M., Persoons, A., Mitiku, A. D., . . . Alemayehu, Y. (2019). MARPLE, a point-of-care, strain-level disease diagnostics and surveillance tool for complex fungal pathogens. *BMC biology, 17*(1), 1-17.

Rai, R., Ranjan, M., Pradhan, B. B., & Chatterjee, S. (2012). Atypical regulation of virulence-associated functions by a diffusible signal factor in *Xanthomonas oryzae* pv. *oryzae*. *Molecular plant-microbe interactions, 25*(6), 789-801.

Rapicavoli, J. N., Blanco-Ulate, B., Muszyński, A., Figueroa-Balderas, R., Morales-Cruz, A., Azadi, P., . . . Roper, M. C. (2018). Lipopolysaccharide O-antigen delays plant innate immune recognition of *Xylella fastidiosa*. *Nature communications, 9*(1), 390. doi:10.1038/s41467-018-02861-5

Rashid, M. M., Nihad, S. A. I., Khan, M. A. I., Haque, A., Ara, A., Ferdous, T., . . . Latif, M. A. (2021). Pathotype profiling, distribution and virulence analysis of *Xanthomonas oryzae* pv. *oryzae* causing bacterial blight disease of rice in

Bangladesh. *Journal of Phytopathology*, *169*(7-8), 438-446. doi:https://doi.org/10.1111/jph.13000

Rigano, L. A., Siciliano, F., Enrique, R., Sendín, L., Filippone, P., Torres, P. S., . . . Marano, M. R. (2007). Biofilm Formation, Epiphytic Fitness, and Canker Development in *Xanthomonas axonopodis* pv. *citri*. *Molecular Plant-Microbe Interactions®, 20*(10), 1222-1230. doi:10.1094/mpmi-20-10-1222

Rocha, J., Shapiro, L. R., & Kolter, R. (2020). A horizontally acquired expansin gene increases virulence of the emerging plant pathogen *Erwinia tracheiphila*. *Scientific Reports, 10*(1), 21743. doi:10.1038/s41598-020-78157-w

Roe, C. C., Horn, K. S., Driebe, E. M., Bowers, J., Terriquez, J. A., Keim, P., & Engelthaler, D. M. (2016). Whole genome SNP typing to investigate methicillin-resistant *Staphylococcus aureus* carriage in a health-care provider as the source of multiple surgical site infections. *Hereditas, 153*(1), 11.

Romantschuk, M., & Bamford, D. H. (1986). The causal agent of halo blight in bean, *Pseudomonas syringae* pv. *phaseolicola*, attaches to stomata via its pili. *Microbial Pathogenesis, 1*(2), 139-148. doi:https://doi.org/10.1016/0882-4010(86)90016-1

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*(12), 1572-1574. doi:10.1093/bioinformatics/btg180

Růžička, K., Ursache, R., Hejátko, J., & Helariutta, Y. (2015). Xylem development – from the cradle to the grave. *New Phytologist, 207*(3), 519-535. doi:https://doi.org/10.1111/nph.13383

Saad, A., & Habibuddin, H. (2010). Pathotypes and virulence of *Xanthomonas oryzae* causing bacterial blight disease of rice in Peninsular Malaysia. *J. Trop. Agric. and Fd. Sc, 38*(2), 257-266.

Saddler, G. S., & Bradbury, J. F. (2005). Xanthomonadalesord. nov. In *Bergey's manual® of systematic bacteriology* (pp. 63-122): Springer.

Sahu, S. K., Zheng, P., & Yao, N. (2018). Niclosamide blocks rice leaf blight by inhibiting biofilm formation of *Xanthomonas oryzae*. *Frontiers in Plant Science, 9*. doi:10.3389/fpls.2018.00408

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution, 4*(4), 406-425. doi:10.1093/oxfordjournals.molbev.a040454

Salzberg, S. L., Sommer, D. D., Schatz, M. C., Phillippy, A. M., Rabinowicz, P. D., Tsuge, S., . . . Bogdanove, A. J. (2008). Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. *BMC Genomics, 9*(1), 204. doi:10.1186/1471-2164-9-204

Sauer, K., Camper, A. K., Ehrlich, G. D., Costerton, J. W., & Davies, D. G. (2002). *Pseudomonas aeruginosa* Displays Multiple Phenotypes during Development as a Biofilm. *Journal of bacteriology, 184*(4), 1140-1154. doi:doi:10.1128/jb.184.4.1140-1154.2002

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution, 3*(3), 430-439. doi:10.1038/s41559-018-0793-y

Schaaper, R. M., & Dunn, R. L. (1987). Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of in vivo DNA replication errors. *Proceedings of the National Academy of Sciences, 84*(17), 6220. doi:10.1073/pnas.84.17.6220

Schandry, N. (2017). A practical guild to visualization and statistical analysis of *R. solanacearum* infection data using R. *Frontiers in Plant Science, 8*, 623.

Schröter, L., & Dersch, P. (2019). Phenotypic Diversification of Microbial Pathogens—Cooperating and Preparing for the Future. *Journal of Molecular Biology, 431*(23), 4645-4655. doi:https://doi.org/10.1016/j.jmb.2019.06.024

Seemann, T. (2015). snippy: fast bacterial variant calling from NGS reads. Retrieved from https://github.com/tseemann/snippy

Shabala, L., & Ross, T. (2008). Cyclopropane fatty acids improve *Escherichia coli* survival in acidified minimal media by reducing membrane permeability to H+ and enhanced ability to extrude H+. *Research in microbiology, 159*(6), 458-461.

Shahbaz, M. U., Qian, S., Yun, F., Zhang, J., Yu, C., Tian, F., . . . Chen, H. (2020). Identification of the Regulatory Components Mediated by the Cyclic di-GMP Receptor Filp and Its Interactor PilZX3 and Functioning in Virulence of *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant-Microbe Interactions®, 33*(10), 1196-1208. doi:10.1094/mpmi-04-20-0088-r

Shyntum, D. Y., Venter, S. N., Moleleki, L. N., Toth, I., & Coutinho, T. A. (2014). Comparative genomics of type VI secretion systems in strains of *Pantoea ananatis* from different environments. *BMC Genomics, 15*(1), 163. doi:10.1186/1471-2164-15-163

Silva, M. S., De Souza, A. A., Takita, M. A., Labate, C. A., & Machado, M. A. (2011). Analysis of the biofilm proteome of *Xylella fastidiosa*. *Proteome Science*, *9*(1), 58. doi:10.1186/1477-5956-9-58

Silva, W. M., Dorella, F. A., Soares, S. C., Souza, G. H. M. F., Castro, T. L. P., Seyffert, N., . . . Azevedo, V. (2017). A shift in the virulence potential of *Corynebacterium pseudotuberculosis* biovar ovis after passage in a murine host demonstrated through comparative proteomics. *BMC Microbiology*, *17*(1), 55. doi:10.1186/s12866-017-0925-6

Šimundić, A.-M. (2009). Measures of diagnostic accuracy: basic definitions. *Ejifcc*, *19*(4), 203.

Singh, A., Gupta, R., Tandon, S., & Pandey, R. (2017). Thyme Oil reduces biofilm formation and impairs virulence of *Xanthomonas oryzae*. *Frontiers in Microbiology*, *8*(JUN). doi:10.3389/fmicb.2017.01074

Soares, A., Alexandre, K., & Etienne, M. (2020). Tolerance and Persistence of *Pseudomonas aeruginosa* in Biofilms Exposed to Antibiotics: Molecular Mechanisms, Antibiotic Strategies and Therapeutic Perspectives. *Frontiers in Microbiology*, *11*(2057). doi:10.3389/fmicb.2020.02057

Solano, C., Echeverz, M., & Lasa, I. (2014). Biofilm dispersion and quorum sensing. *Current Opinion in Microbiology*, *18*, 96-104. doi:https://doi.org/10.1016/j.mib.2014.02.008

Song, Z., Zhao, Y., Qian, G., Odhiambo, B. O., & Liu, F. (2017). Novel insights into the regulatory roles of gene *hshB* in *Xanthomonas oryzae* pv. *oryzicola*. *Research in microbiology*, *168*(2), 165-173. doi:https://doi.org/10.1016/j.resmic.2016.10.007

Song, Z., Zhao, Y., Zhou, X., Wu, G., Zhang, Y., Qian, G., & Liu, F. (2015). Identification and Characterization of Two Novel DSF-Controlled Virulence-Associated Genes Within the *nodB-rhgB* Locus of *Xanthomonas oryzae* pv. *oryzicola* Rs105. *Phytopathology®*, *105*(5), 588-596. doi:10.1094/phyto-07-14-0190-r

Sonnewald, S., Priller, J. P. R., Schuster, J., Glickmann, E., Hajirezaei, M.-R., Siebig, S., . . . Sonnewald, U. (2012). Regulation of Cell Wall-Bound Invertase in Pepper Leaves by *Xanthomonas campestris* pv. *vesicatoria* Type Three Effectors. *PLoS One*, *7*(12), e51763. doi:10.1371/journal.pone.0051763

Speare, L., Cecere, A. G., Guckes, K. R., Smith, S., Wollenberg, M. S., Mandel, M. J., . . . Septer, A. N. (2018). Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *Proceedings of the*

National Academy of Sciences, 115(36), E8528. doi:10.1073/pnas.1808302115

Sprouffske, K., & Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics, 17*(1), 1-4.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312-1313.

Stanley, N. R., & Lazazzera, B. A. (2004). Environmental signals and regulatory pathways that influence biofilm formation. *Molecular Microbiology, 52*(4), 917-924. doi:https://doi.org/10.1111/j.1365-2958.2004.04036.x

Steele, K. A., Quinton-Tulloch, M. J., Amgai, R. B., Dhakal, R., Khatiwada, S. P., Vyas, D., . . . Witcombe, J. R. (2018). Accelerating public sector rice breeding with high-density KASP markers derived from whole genome sequencing of indica rice. *Molecular breeding : new strategies in plant improvement, 38*(4), 38-38. doi:10.1007/s11032-018-0777-2

Steele, M. I., Lorenz, D., Hatter, K., Park, A., & Sokatch, J. R. (1992). Characterization of the *mmsAB* operon of *Pseudomonas aeruginosa* PAO encoding methylmalonate-semialdehyde dehydrogenase and 3-hydroxyisobutyrate dehydrogenase. *J Biol Chem, 267*(19), 13585-13592.

Stewart, G. R., & Young, D. B. (2004). Heat-shock proteins and the host–pathogen interaction during bacterial infection. *Current Opinion in Immunology, 16*(4), 506-510. doi:https://doi.org/10.1016/j.coi.2004.05.007

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

Stukenbrock, E. H., & Bataillon, T. (2012). A Population Genomics Perspective on the Emergence and Adaptation of New Plant Pathogens in Agro-Ecosystems. *PLOS Pathogens, 8*(9), e1002893. doi:10.1371/journal.ppat.1002893

Su, J., Zou, X., Huang, L., Bai, T., Liu, S., Yuan, M., . . . He, J. (2016). DgcA, a diguanylate cyclase from *Xanthomonas oryzae* pv. *oryzae* regulates bacterial pathogenicity on rice. *Scientific Reports, 6*(1), 25978. doi:10.1038/srep25978

Su, P., Song, Z., Wu, G., Zhao, Y., Zhang, Y., Wang, B., . . . Liu, F. (2018). Insights into the roles of two genes of the histidine biosynthesis operon in pathogenicity of *Xanthomonas oryzae* pv. *oryzicola*. *Phytopathology, 108*(5), 542-551.

Sugio, A., Yang, B., & White, F. F. (2005). Characterization of the hrpF pathogenicity peninsula of *Xanthomonas oryzae* pv. *oryzae*. *Molecular plant-microbe interactions, 18*(6), 546-554.

Sun, Y., Wen, S., Zhao, L., Xia, Q., Pan, Y., Liu, H., . . . Wang, H. (2020). Association among biofilm formation, virulence gene expression, and antibiotic resistance in Proteus mirabilis isolates from diarrhetic animals in Northeast China. *BMC Veterinary Research, 16*(1), 176. doi:10.1186/s12917-020-02372-w

Suo, W., Shi, X., Xu, S., Li, X., & Lin, Y. (2020). Towards low cost, multiplex clinical genotyping: 4-fluorescent Kompetitive Allele-Specific PCR and its application on pharmacogenetics. *PLoS One, 15*(3), e0230445. doi:10.1371/journal.pone.0230445

Swigonová, Z., Mohsen, A.-W., & Vockley, J. (2009). Acyl-CoA dehydrogenases: Dynamic history of protein family evolution. *Journal of molecular evolution, 69*(2), 176-193. doi:10.1007/s00239-009-9263-0

Swings, J., Van den Mooter, M., Vauterin, L., Hoste, B., Gillis, M., Mew, T., & Kersters, K. (1990). Reclassification of the Causal Agents of Bacterial Blight (*Xanthomonas campestris* pv. *oryzae*) and Bacterial Leaf Streak (*Xanthomonas campestris* pv. *oryzicola*) of Rice as Pathovars of *Xanthomonas oryzae* (ex Ishiyama 1922) sp. nov., nom. rev. *International Journal of Systematic and Evolutionary Microbiology, 40*(3), 309-311.

Szarvas, J., Ahrenfeldt, J., Cisneros, J. L. B., Thomsen, M. C. F., Aarestrup, F. M., & Lund, O. (2020). Large scale automated phylogenomic analysis of bacterial isolates and the Evergreen Online platform. *Communications biology, 3*(1), 1-10.

Tall, H., Tékété, C., Noba, K., Koita, O., Cunnac, S., Hutin, M., . . . Verdier, V. (2020). Confirmation Report of Bacterial Blight Caused by Xanthomonas oryzae pv. oryzae on Rice in Senegal. *Plant disease, 104*(3), 968-968. doi:10.1094/pdis-07-19-1464-pdn

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research, 22*(22), 4673-4680. doi:10.1093/nar/22.22.4673

Tian, F., Yu, C., Li, H., Wu, X., Li, B., Chen, H., . . . He, C. (2015). Alternative sigma factor RpoN2 is required for flagellar motility and full virulence of *Xanthomonas oryzae* pv. *oryzae*. *Microbiological Research, 170*, 177-183. doi:https://doi.org/10.1016/j.micres.2014.07.002

Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D., & Corander, J. (2018). RhierBAPS: An R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Research, 3*.

Toska, J., Ho, B. T., & Mekalanos, J. J. (2018). Exopolysaccharide protects *Vibrio cholerae* from exogenous attacks by the type 6 secretion system. *Proceedings of the National Academy of Sciences of the United States of America, 115*(31), 7997-8002. doi:10.1073/pnas.1808469115

Treangen, T. J., Ondov, B. D., Koren, S., & Phillippy, A. M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biology, 15*(11), 524.

Uelze, L., Grützke, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., . . . Malorny, B. (2020). Typing methods based on whole genome sequencing data. *One Health Outlook, 2*(1), 1-19.

Ur Rehman, S., Ali Sher, M., Saddique, M. A. B., Ali, Z., Khan, M. A., Mao, X., . . . Jing, R. (2021). Development and Exploitation of KASP Assays for Genes Underpinning Drought Tolerance Among Wheat Cultivars From Pakistan. *Frontiers in Genetics, 12*. doi:10.3389/fgene.2021.684702

Vadyvaloo, V., & Martínez, L. (2014). Mechanisms of post-transcriptional gene regulation in bacterial biofilms. *Frontiers in Cellular and Infection Microbiology, 4*. doi:10.3389/fcimb.2014.00038

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., . . . Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence, 3*(2), 125-133. doi:10.1038/s42256-020-00287-7

Van den Bergh, B., Fauvart, M., & Michiels, J. (2017). Formation, physiology, ecology, evolution and clinical importance of bacterial persisters. *FEMS Microbiology Reviews, 41*(3), 219-251. doi:10.1093/femsre/fux001

Varghese, J., Chochua, S., Tran, T., Walker, H., Li, Z., Vagnone, P. S., . . . Metcalf, B. (2020). Multistate population and whole genome sequence-based strain surveillance of invasive pneumococci recovered in the USA during 2017. *Clinical Microbiology and Infection, 26*(4), 512. e511-512. e510.

Vera Cruz, C., Bai, J., Oña, I., Leung, H., Nelson, R. J., Mew, T.-W., & Leach, J. E. (2000). Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proceedings of the National Academy of Sciences, 97*(25), 13500-13505.

Vera Cruz, C., Cottyn, B., Nguyen, M., Lang, J., Verdier, V., Mew, T. W., & Leach, J. E. (2017). Detection of *Xanthomonas oryzae* pv. *oryzae*, and *X*. *oryzae* pv. *oryzicola* in rice seeds (Chapter 8). *APS Manual on Detection of Plant Pathogenic Bacteria in Seed and Other Planting Material. APS Press, Minneapolis, MN. ISBN*, 978-970.

Voegel, T. M., Doddapaneni, H., Cheng, D. W., Lin, H., Stenger, D. C., Kirkpatrick, B. C., & Roper, M. C. (2013). Identification of a response regulator involved in surface attachment, cell–cell aggregation, exopolysaccharide production and virulence in the plant pathogen *Xylella fastidiosa*. *Molecular Plant Pathology*, *14*(3), 256-264.

Vos, M., & Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *The ISME journal*, *3*(2), 199-208. doi:10.1038/ismej.2008.93

Waddington, C. H. (1959). Evolutionary adaptation. *Perspectives in biology and medicine*, *2*(4), 379-401.

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E. E., & van der Heijden, M. G. A. (2019). Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature Communications*, *10*(1), 4841. doi:10.1038/s41467-019-12798-y

Walker, K. (2018). Tigris: load census TIGER/line shapefiles. *See https://CRAN. R-project. org/package= tigris*.

Wang, B., Wu, G., Zhang, Y., Qian, G., & Liu, F. (2018). Dissecting the virulence-related functionality and cellular transcription mechanism of a conserved hypothetical protein in *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology*, *19*(8), 1859-1872. doi:https://doi.org/10.1111/mpp.12664

Wang, S., Sun, Z., Wang, H., Liu, L., Lu, F., Yang, J., . . . Bent, A. F. (2015). Rice OsFLS2-mediated perception of bacterial flagellins is evaded by Xanthomonas oryzae pvs. oryzae and oryzicola. *Molecular plant*, *8*(7), 1024-1037.

Wang, X.-Y., Zhou, L., Yang, J., Ji, G.-H., & He, Y.-W. (2016). The RpfB-Dependent Quorum Sensing Signal Turnover System Is Required for Adaptation and Virulence in Rice Bacterial Blight Pathogen *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant-Microbe Interactions®*, *29*(3), 220-230. doi:10.1094/mpmi-09-15-0206-r

Wei, C., Ding, T., Chang, C., Yu, C., Li, X., & Liu, Q. (2019). Global Regulator PhoP is Necessary for Motility, Biofilm Formation, Exoenzyme Production,

and Virulence of *Xanthomonas citri* subsp. *citri* on Citrus Plants. *Genes, 10*(5), 340. Retrieved from https://www.mdpi.com/2073-4425/10/5/340

Whelan, S., & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution, 18*(5), 691-699.

White, F., Potnis, N., Jones, J. B., & Koebnik, R. (2009). The type III effectors of *Xanthomonas*. *Molecular Plant Pathology, 10*(6), 749-766. doi:10.1111/j.1364-3703.2009.00590.x

White, F., & Yang, B. (2009). Host and pathogen factors controlling the rice-*Xanthomonas oryzae* interaction. *Plant Physiology, 150*(4), 1677-1686. doi:10.1104/pp.109.139360

Wickham, H., Francois, R., Henry, L., & Müller, K. (2015). dplyr: A grammar of data manipulation. *R package version 0.4, 3*.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences, 87*(12), 4576. doi:10.1073/pnas.87.12.4576

Woods, L. C., Gorrell, R. J., Taylor, F., Connallon, T., Kwok, T., & McDonald, M. J. (2020). Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proceedings of the National Academy of Sciences, 117*(43), 26868-26875. doi:doi:10.1073/pnas.2005331117

Yadeta, K. A., & J Thomma, B. P. H. (2013). The xylem as battleground for plant hosts and vascular wilt pathogens. *Frontiers in Plant Science, 4*, 97-97. doi:10.3389/fpls.2013.00097

Yan, Q., Hu, X., & Wang, N. (2012). The novel virulence-related gene *nlxA* in the lipopolysaccharide cluster of *Xanthomonas citri* ssp. *citri* is involved in the production of lipopolysaccharide and extracellular polysaccharide, motility, biofilm formation and stress resistance. *Molecular Plant Pathology, 13*(8), 923-934. doi:https://doi.org/10.1111/j.1364-3703.2012.00800.x

Yang, F., Qian, S., Tian, F., Chen, H., Hutchins, W., Yang, C.-H., & He, C. (2016). The GGDEF-domain protein GdpX1 attenuates motility, exopolysaccharide production and virulence in *Xanthomonas oryzae* pv. *oryzae*. *Journal of Applied Microbiology, 120*(6), 1646-1657. doi:https://doi.org/10.1111/jam.13115

Yang, F., Xue, D., Tian, F., Hutchins, W., Yang, C.-H., & He, C. (2019). Identification of c-di-GMP Signaling Components in *Xanthomonas oryzae* and Their Orthologs in Xanthomonads Involved in Regulation of Bacterial Virulence Expression. *Frontiers in Microbiology, 10*(1402). doi:10.3389/fmicb.2019.01402

Yang, X., Long, M., & Shen, X. (2018). Effector–Immunity Pairs Provide the T6SS Nanomachine its Offensive and Defensive Capabilities. *Molecules (Basel, Switzerland), 23*(5), 1009. doi:10.3390/molecules23051009

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution, 39*(3), 306-314.

Yao, J., & Allen, C. (2006). Chemotaxis is required for virulence and competitive fitness of the bacterial wilt pathogen *Ralstonia solanacearum*. *Journal of bacteriology, 188*(10), 3697-3708.

Yao, J., & Allen, C. (2007). The Plant Pathogen *Ralstonia solanacearum* Needs Aerotaxis for Normal Biofilm Formation and Interactions with Its Tomato Host. *Journal of bacteriology, 189*(17), 6415-6424. doi:doi:10.1128/JB.00398-07

Yaron, S., & Römling, U. (2014). Biofilm formation by enteric pathogens and its role in plant colonization and persistence. *Microbial Biotechnology, 7*(6), 496-516. doi:10.1111/1751-7915.12186

Yaryura, P. M., Conforte, V. P., Malamud, F., Roeschlin, R., de Pino, V., Castagnaro, A. P., . . . Vojnov, A. A. (2015). XbmR, a new transcription factor involved in the regulation of chemotaxis, biofilm formation and virulence in *Xanthomonas citri* subsp. *citri*. *Environmental Microbiology, 17*(11), 4164-4176. doi:10.1111/1462-2920.12684

Yu, C., Chen, H., Tian, F., Yang, F., Yuan, X., Yang, C.-H., & He, C. (2018). A ten gene-containing genomic island determines flagellin glycosylation: implication for its regulatory role in motility and virulence of *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology, 19*(3), 579-592. doi:https://doi.org/10.1111/mpp.12543

Yu, C., Nguyen, D.-P., Ren, Z., Liu, J., Yang, F., Tian, F., . . . Chen, H. (2020). The RpoN2-PilRX regulatory system governs type IV pilus gene transcription and is required for bacterial motility and virulence in *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology, 21*(5), 652-666. doi:https://doi.org/10.1111/mpp.12920

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution, 8*(1), 28-36.

Yu, X., Liang, X., Liu, K., Dong, W., Wang, J., & Zhou, M.-g. (2015). The thiG Gene Is Required for Full Virulence of *Xanthomonas oryzae* pv. *oryzae* by Preventing Cell Aggregation. *PLoS One, 10*(7), e0134237. doi:10.1371/journal.pone.0134237

Zaka, A. (2019). *Molecular characterization of BLB resistant rice germplasm.* (Doctoral dissertation). Retrieved from http://prr.hec.gov.pk/jspui/handle/123456789/12295 Pakistan Research Repository database.

Zaka, A., Grande, G., Coronejo, T., Quibod, I. L., Chen, C.-W., Chang, S.-J., . . . Oliva, R. (2018). Natural variations in the promoter of OsSWEET13 and OsSWEET14 expand the range of resistance against *Xanthomonas oryzae* pv. *oryzae*. *PLoS One, 13*(9), e0203711. doi:10.1371/journal.pone.0203711

Zarei, S., Taghavi, S. M., Hamzehzarghani, H., Osdaghi, E., & Lamichhane, J. R. (2018). Epiphytic growth of Xanthomonas arboricola and Xanthomonas citri on non-host plants. *Plant Pathology, 67*(3), 660-670. doi:https://doi.org/10.1111/ppa.12769

Zhang, F., Hu, Z., Wu, Z., Lu, J., Shi, Y., Xu, J., . . . Li, Z. (2021). Reciprocal adaptation of rice and *Xanthomonas oryzae* pv. *oryzae*: cross-species 2D GWAS reveals the underlying genetics. *The Plant Cell*. doi:10.1093/plcell/koab146

Zhang, Y., Wei, C., Jiang, W., Wang, L., Li, C., Wang, Y., . . . Sun, W. (2013). The HD-GYP Domain Protein RpfG of *Xanthomonas oryzae* pv. *oryzicola* Regulates Synthesis of Extracellular Polysaccharides that Contribute to Biofilm Formation and Virulence on Rice. *PLoS One, 8*(3), e59428. doi:10.1371/journal.pone.0059428

Zhao, J., & Binns, A. N. (2014). GxySBA ABC transporter of *Agrobacterium tumefaciens* and its role in sugar utilization and *vir* gene expression. *Journal of bacteriology, 196*(17), 3150-3159.

Zhao, Y., Qian, G., Fan, J., Yin, F., Zhou, Y., Liu, C., . . . Liu, F. (2012). Identification and Characterization of a Novel Gene, hshB, in *Xanthomonas oryzae* pv. *oryzicola* Co-Regulated by Quorum Sensing and clp. *Phytopathology®, 102*(3), 252-259. doi:10.1094/phyto-06-11-0169

Zheng, S., Bawazir, M., Dhall, A., Kim, H.-E., He, L., Heo, J., & Hwang, G. (2021). Implication of Surface Properties, Bacterial Motility, and Hydrodynamic Conditions on Bacterial Surface Sensing and Their Initial Adhesion. *Frontiers in Bioengineering and Biotechnology*, *9*(82). doi:10.3389/fbioe.2021.643722

Zhu, J., & Mekalanos, J. J. (2003). Quorum sensing-dependent biofilms enhance colonization in Vibrio cholerae. *Developmental cell*, *5*(4), 647-656.

Zhu, P. C., Li, Y. M., Yang, X., Zou, H. F., Zhu, X. L., Niu, X. N., . . . He, Y. Q. (2020). Type VI secretion system is not required for virulence on rice but for inter-bacterial competition in *Xanthomonas oryzae* pv. *oryzicola*. *Res Microbiol*, *171*(2), 64-73. doi:10.1016/j.resmic.2019.10.004

Appendix A

Table A.1. Terms used to conduct an expanded search of publications that report findings on the biofilm-related genes specific to Xanthomonadales. The list was generated by extracting terms from the titles and abstracts of search results using the keywords "Xanthomonas AND biofilm" using the *extract_terms* function of the R package ***litsearchr*** (E. Grames et al., 2020; E. M. Grames et al., 2019) and were grouped manually into biofilm-related terms or *Xanthomonas*-related terms.

Term	Group
amyloliquefaciens mep218	biofilm
analyses reveal	biofilm, *Xanthomonas*
antibacterial activity	biofilm
Bacillus amyloliquefaciens	biofilm
Bacillus amyloliquefaciens mep218	biofilm
bacterial pathogen	biofilm, *Xanthomonas*
bacterial pathogens	biofilm, *Xanthomonas*
bacterial virulence	biofilm, *Xanthomonas*
bacterium *Xanthomonas*	*Xanthomonas*
bacterium *Xanthomonas citri*	*Xanthomonas*
biofilm formation	biofilm, *Xanthomonas*
canker disease	*Xanthomonas*
citri strains	*Xanthomonas*
citri subsp	*Xanthomonas*
citrus canker	*Xanthomonas*
citrus canker disease	*Xanthomonas*
cyclic di-gmp	biofilm, *Xanthomonas*
diffusible signal	biofilm, *Xanthomonas*
diffusible signal factor	biofilm, *Xanthomonas*
diguanylate cyclase	biofilm, *Xanthomonas*
domain protein	biofilm, *Xanthomonas*
exopolysaccharide production	biofilm, *Xanthomonas*
functional analysis	biofilm, *Xanthomonas*
functional characterization	biofilm, *Xanthomonas*
gram-negative bacterium	xanthomonas
green synthesis	biofilm, *Xanthomonas*
growth promotion	biofilm, *Xanthomonas*
hd-gyp domain	biofilm, *Xanthomonas*
influences biofilm	biofilm, *Xanthomonas*
influences biofilm formation	biofilm, *Xanthomonas*
membrane porin	biofilm, *Xanthomonas*
multiple roles	biofilm, *Xanthomonas*
outer membrane	biofilm, *Xanthomonas*

Table A.1. Continued

Term	Group
outer membrane porin	biofilm, *Xanthomonas*
oxide nanoparticles	biofilm, *Xanthomonas*
pathogen *Xanthomonas*	*Xanthomonas*
pathogen *Xanthomonas campestris*	*Xanthomonas*
pathogen *Xanthomonas citri*	*Xanthomonas*
pathogen *Xanthomonas citri* subsp	*Xanthomonas*
pathogenic bacteria	*Xanthomonas*
pathogens xanthomonas	*Xanthomonas*
phytopathogen *Xanthomonas*	*Xanthomonas*
phytopathogenic bacterium	*Xanthomonas*
phytopathogenic bacterium *Xanthomonas*	*Xanthomonas*
phytopathogenic bacterium *Xanthomonas citri*	*Xanthomonas*
plant growth	*Xanthomonas*
plant growth promotion	*Xanthomonas*
plant pathogen	*Xanthomonas*
plant pathogen *Xanthomonas*	*Xanthomonas*
plant pathogen *Xanthomonas citri*	*Xanthomonas*
plant pathogen *Xanthomonas citri* subsp	*Xanthomonas*
polysaccharide production	biofilm, *Xanthomonas*
protein-mediated quorum	biofilm, *Xanthomonas*
protein-mediated quorum sensing	biofilm, *Xanthomonas*
proteomic analyses	biofilm, *Xanthomonas*
proteomic analyses reveal	biofilm, *Xanthomonas*
Pseudomonas aeruginosa	biofilm, *Xanthomonas*
quorum sensing	biofilm, *Xanthomonas*
secretion system	biofilm, *Xanthomonas*
signal factor	biofilm, *Xanthomonas*
small protein-mediated	biofilm, *Xanthomonas*
small protein-mediated quorum	biofilm, *Xanthomonas*
small protein-mediated quorum sensing	biofilm, *Xanthomonas*
Stenotrophomonas maltophilia	biofilm, *Xanthomonas*
stress tolerance	biofilm, *Xanthomonas*
two-component system	biofilm, *Xanthomonas*
virulence factor	biofilm, *Xanthomonas*
virulence factors	biofilm, *Xanthomonas*
Xanthomonas axonopodis	*Xanthomonas*
Xanthomonas campestris	*Xanthomonas*
Xanthomonas citri	*Xanthomonas*
Xanthomonas citri subsp	*Xanthomonas*
Xanthomonas oryzae	*Xanthomonas*
Xylella fastidiosa	*Xanthomonas*

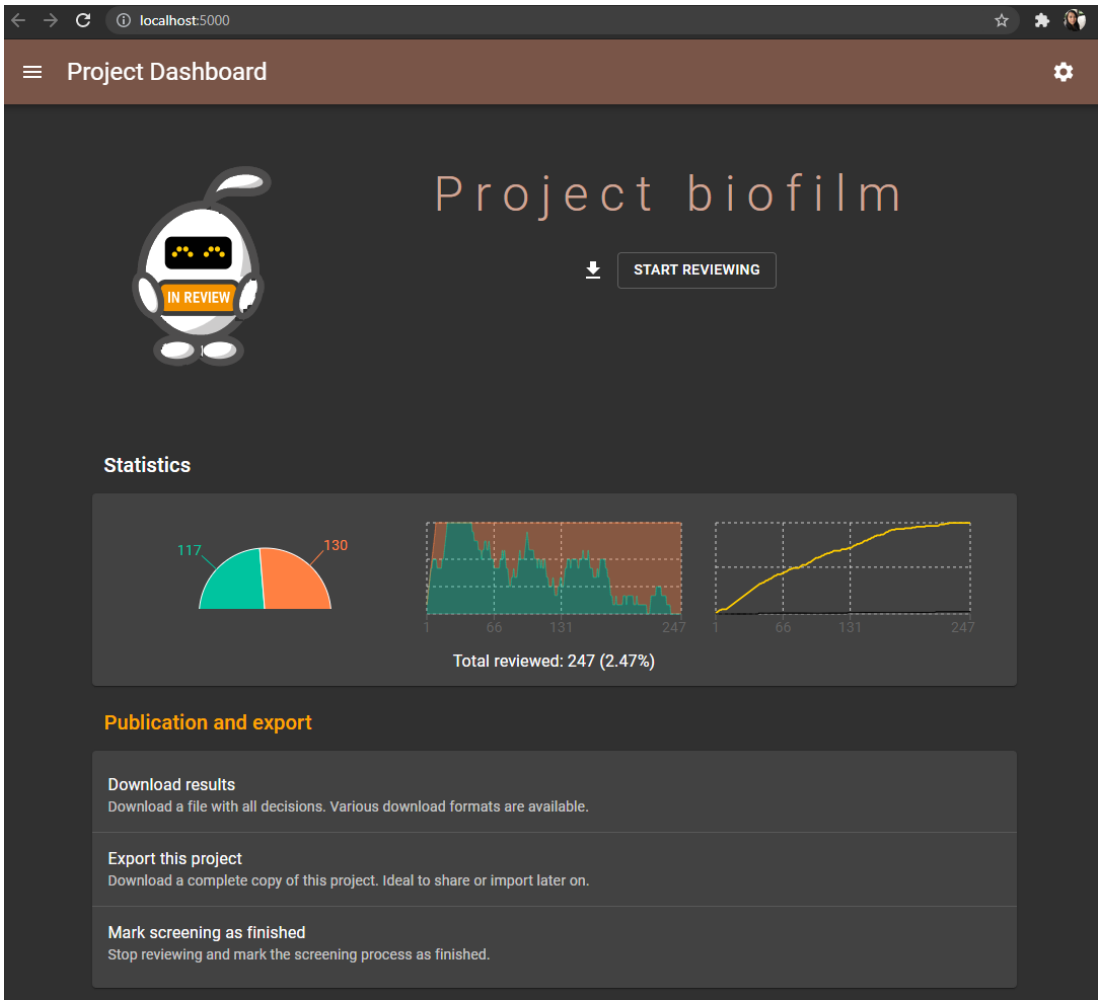


Figure A.1. The user interface of ASReview (van de Schoot et al., 2021) which was used for semi-automated screening for published research on the biofilm-related genes of Xanthomonadales. This software applies a Bayesian algorithm to predict the next best-fit article for review based on previous decisions by the user. Prior to screening, a rule was set to stop the process when 20 consecutive irrelevant articles are suggested by the software. Charts in the user-interface can be observed where teal color indicates publications categorized by the user as "relevant", and the orange color denotes publications categorized as "irrelevant" to the research topic.

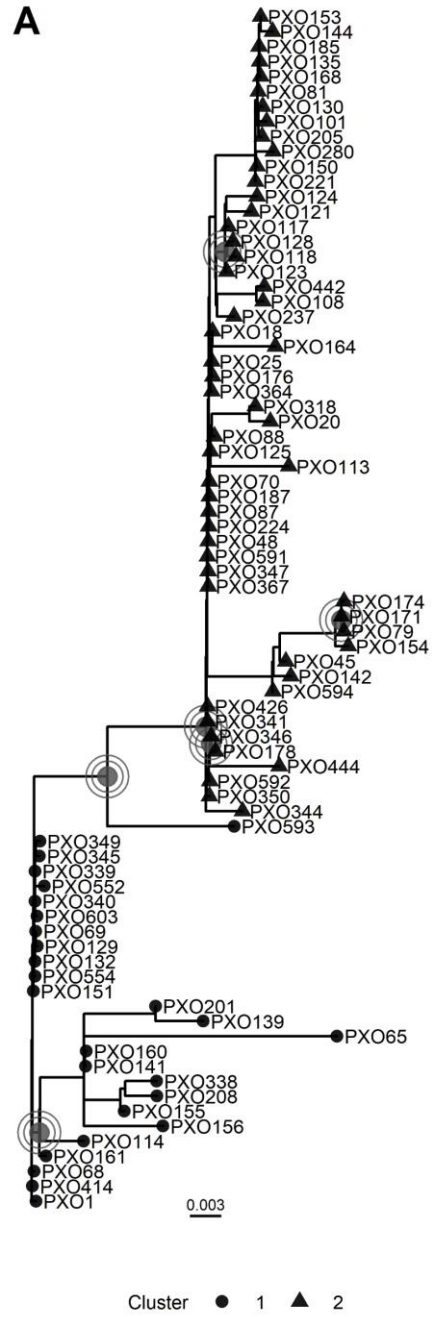


Figure A.2. Maximum likelihood tree of Philippine *Xanthomonas oryzae* pv. *oryzae* estimated in IQ-Tree (Lanfear et al., 2016) from multiple sequence alignment of concatenated amino acid sequences from 50 biofilm-related proteins previously reported in Xanthomonads. Nodes with concentric circles represent nodes with support values of Shimodaira-Hasegawa approximate likelihood ratio test, SH-aLRT (Guindon et al., 2010) >= 80 and ultrafast bootstrap (Hoang et al., 2017) >= 95. Shapes denote K-means clusters calculated in the package bios2mds v.1.2.3 (Pelé et al., 2012)

from a dissimilarity matrix generated from amino acid substitutions gathered from multiple sequence alignment of the 50 biofilm-related proteins.

Table A.2. Variability of selected biofilm-related proteins in *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) based on Basic Local Alignment Search Tool (BLAST) percent identity scores of selected biofilm-related proteins reported previously in Xanthomonads queried against a database of 80 Philippine *Xoo* genomes.

Strain	Sequence variability					Presence / absence variability			BAPS cluster [a]	Collection site [b]	Collection date [b]
	AAM38443.1	AAO28716.1	AAM41224.1	AAY49386.1	AAO29541.1	ACD61037.1	WP_014504558.1	WP_082350419.1			
PXO117	52.5	68.0	97.5	86.8	67.6	100.0	Absent	Absent	1	IRRI	1979
PXO114	29.3	77.5	98.8	86.8	68.9	100.0	Absent	Absent	1	IRRI	1979
PXO118	52.5	72.6	93.1	86.8	70.0	100.0	Absent	Absent	1	IRRI	1979
PXO121	52.5	69.2	97.8	86.8	70.0	100.0	Absent	Absent	1	IRRI	1979
PXO123	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	1	IRRI	1980
PXO124	52.5	69.2	98.3	86.8	67.6	100.0	Absent	Absent	1	IRRI	1980
PXO128	54.3	69.2	95.6	86.8	67.6	100.0	Absent	Absent	1	IRRI	1982
PXO318	52.5	68.6	98.8	86.8	67.6	100.0	96.6	83.3	2	Central Luzon	1990
PXO444	52.5	69.2	93.1	84.6	67.6	100.0	Absent	83.3	2	Laguna	1993
PXO426	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	1993
PXO367	52.5	68.0	98.8	86.8	67.6	100.0	96.6	83.3	2	Laguna	1993
PXO442	52.5	72.6	93.1	89.0	67.6	100.0	96.6	83.3	2	Laguna	1993
PXO345	29.3	72.6	93.1	86.8	66.8	100.0	96.6	83.3	2	Laguna	1994
PXO340	29.3	69.2	93.1	86.8	67.6	100.0	Absent	Absent	2	Laguna	1994
PXO346	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994
PXO347	52.5	69.2	93.1	86.8	67.6	Absent	96.6	83.3	2	Laguna	1994
PXO339	29.3	72.6	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994
PXO554	29.3	72.6	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994

Table A.2. Continued.

Strain	Sequence variability					Presence / absence variability			BAPS cluster [a]	Collection site [b]	Collection date [b]
	AAM38443.1	AAO28716.1	AAM41224.1	AAY49386.1	AAO29541.1	ACD61037.1	WP_014504558.1	WP_082350419.1			
PXO350	52.5	72.6	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994
PXO552	29.3	72.6	95.6	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994
PXO344	52.5	69.2	97.2	86.8	67.6	100.0	96.6	83.3	2	Laguna	1994
PXO349	29.3	69.2	93.1	86.8	68.9	100.0	96.6	83.3	2	Laguna	1994
PXO593	29.0	69.2	93.1	84.1	67.6	100.0	96.6	83.3	2	Laguna	2006
PXO592	52.5	72.6	93.1	86.8	67.6	100.0	96.6	83.3	2	Laguna	2006
PXO594	52.5	69.2	93.1	87.8	67.6	100.0	96.6	83.3	2	Laguna	2006
PXO591	54.3	69.2	93.1	86.8	69.0	100.0	Absent	83.3	2	Laguna	2006
PXO603	29.3	69.2	98.8	90.5	67.6	100.0	96.6	83.3	2	Laguna	2010
PXO65	29.3	72.6	97.2	90.5	67.6	100.0	Absent	Absent	3	Isabela	1973
PXO79	52.5	72.6	93.1	87.8	67.6	100.0	Absent	Absent	3	Davao	1975
PXO81	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Davao	1976
PXO141	29.3	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Bicol	1980
PXO101	52.5	72.6	93.1	86.8	67.6	100.0	Absent	Absent	3	Laguna	1980
PXO156	29.3	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Ifugao	1981
PXO135	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Ilocos Sur	1981
PXO154	52.5	72.6	93.1	87.8	67.6	100.0	Absent	Absent	3	Ifugao	1981
PXO155	29.3	72.6	93.1	90.5	67.6	100.0	Absent	Absent	3	Ifugao	1981
PXO130	52.5	68.0	97.8	90.5	67.6	100.0	Absent	Absent	3	Ifugao	1981
PXO144	52.5	69.2	97.2	86.8	66.8	100.0	Absent	Absent	3	Ifugao	1982
PXO139	29.3	72.6	93.1	87.8	67.6	100.0	Absent	Absent	3	Butuan City	1982
PXO168	52.5	69.2	98.8	86.8	67.6	100.0	Absent	Absent	3	Bulacan	1984
PXO171	52.5	69.2	93.1	87.8	67.6	100.0	Absent	Absent	3	Negros Occidental	1984

Table A.2. Continued.

Strain	Sequence variability					Presence / absence variability			BAPS cluster [a]	Collection site [b]	Collection date [b]
	AAM38443.1	AAO28716.1	AAM41224.1	AAY49386.1	AAO29541.1	ACD61037.1	WP_014504558.1	WP_082350419.1			
PXO201	29.3	69.2	98.8	87.8	67.6	100.0	Absent	Absent	3	Ifugao	1984
PXO150	52.5	69.2	93.1	90.5	67.6	100.0	Absent	Absent	3	Ifugao	1984
PXO153	52.5	69.2	97.2	86.8	67.6	100.0	Absent	Absent	3	Camarines Sur	1985
PXO174	45.1	72.6	97.2	87.8	69.0	100.0	Absent	Absent	3	Albay	1985
PXO160	29.3	69.2	93.1	86.8	70.0	100.0	Absent	Absent	3	Leyte	1985
PXO185	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Tarlac	1986
PXO205	52.5	69.2	93.1	86.8	66.8	100.0	Absent	Absent	3	Ifugao	1989
PXO221	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Ifugao	1989
PXO208	29.3	72.6	93.1	86.8	67.6	100.0	Absent	Absent	3	Ifugao	1989
PXO280	52.5	69.2	93.1	86.8	67.6	100.0	Absent	Absent	3	Ifugao	1990
PXO338	29.3	72.6	93.1	86.8	66.8	100.0	Absent	Absent	3	Ifugao	1991
PXO48	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	Pampanga	1973
PXO69	29.3	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	Palawan	1974
PXO88	52.5	69.2	93.1	86.8	67.8	99.7	96.6	83.3	4	Isabela	1974
PXO70	45.1	69.2	93.1	86.8	69.0	100.0	96.6	83.3	4	Palawan	1974
PXO87	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	Isabela	1978
PXO113	52.5	69.2	98.8	86.8	67.6	100.0	96.6	83.3	4	IRRI	1978
PXO132	29.3	69.2	98.8	86.8	67.6	100.0	96.6	83.3	4	Cagayan	1979
PXO125	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	IRRI	1981
PXO129	29.3	69.2	93.1	86.8	67.6	100.0	Absent	83.3	4	IRRI	1982
PXO187	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	Ilocos Norte	1986
PXO224	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	4	Ifugao	1989
PXO341	52.5	72.6	93.1	86.8	67.6	100.0	96.6	83.3	4	Laguna	1998

Table A.2. Continued.

Strain	Sequence variability					Presence / absence variability			BAPS cluster [a]	Collection site [b]	Collection date [b]
	AAM38443.1	AAO28716.1	AAM41224.1	AAY49386.1	AAO29541.1	ACD61037.1	WP_014504558.1	WP_082350419.1			
PXO18	52.5	69.2	93.1	86.8	67.6	100.0	Absent	83.3	5	IRRI	1972
PXO25	52.5	69.2	93.1	86.8	67.6	100.0	96.6	83.3	5	Laguna	1972
PXO1	29.3	72.6	93.1	86.8	67.6	100.0	96.6	83.3	5	IRRI	1972
PXO20	52.5	69.2	93.1	86.8	68.9	100.0	96.6	83.3	5	Laguna	1972
PXO68	29.3	69.2	93.1	86.8	67.6	100.0	96.6	83.3	5	Palawan	1974
PXO164	52.5	72.6	93.1	86.8	67.6	100.0	96.6	83.3	5	Albay	1979
PXO142	54.3	72.6	95.6	87.8	68.9	100.0	96.6	83.3	5	Davao	1981
PXO151	29.3	69.2	93.1	86.8	67.6	100.0	96.6	83.3	5	Palawan	1985
PXO178	52.5	69.2	93.1	86.8	67.6	Absent	96.6	83.3	5	Albay	1986
PXO176	52.5	68.0	98.8	86.8	67.6	100.0	Absent	83.3	5	Sorsogon	1986
PXO414	29.3	72.6	93.1	86.8	67.6	100.0	96.6	83.3	5	Laguna	1993
PXO364	52.5	72.6	93.1	86.8	69.0	100.0	96.6	83.3	5	Laguna	1993
PXO45	54.3	77.9	92.7	87.8	67.6	100.0	Absent	Absent	6	Ifugao	1973
PXO108	52.5	69.2	92.7	89.0	70.2	100.0	Absent	Absent	6	Ifugao	1976
PXO161	29.3	69.2	92.7	86.8	67.6	100.0	Absent	Absent	6	Ifugao	1985
PXO237	52.5	69.2	98.1	86.8	67.6	100.0	Absent	Absent	6	Ifugao	1989

[a] Bayesian analysis of population structure (L. Cheng et al., 2013) estimated using the R package RhierBAPS (Tonkin-Hill et al., 2018).

[b] (Quibod et al., 2020)

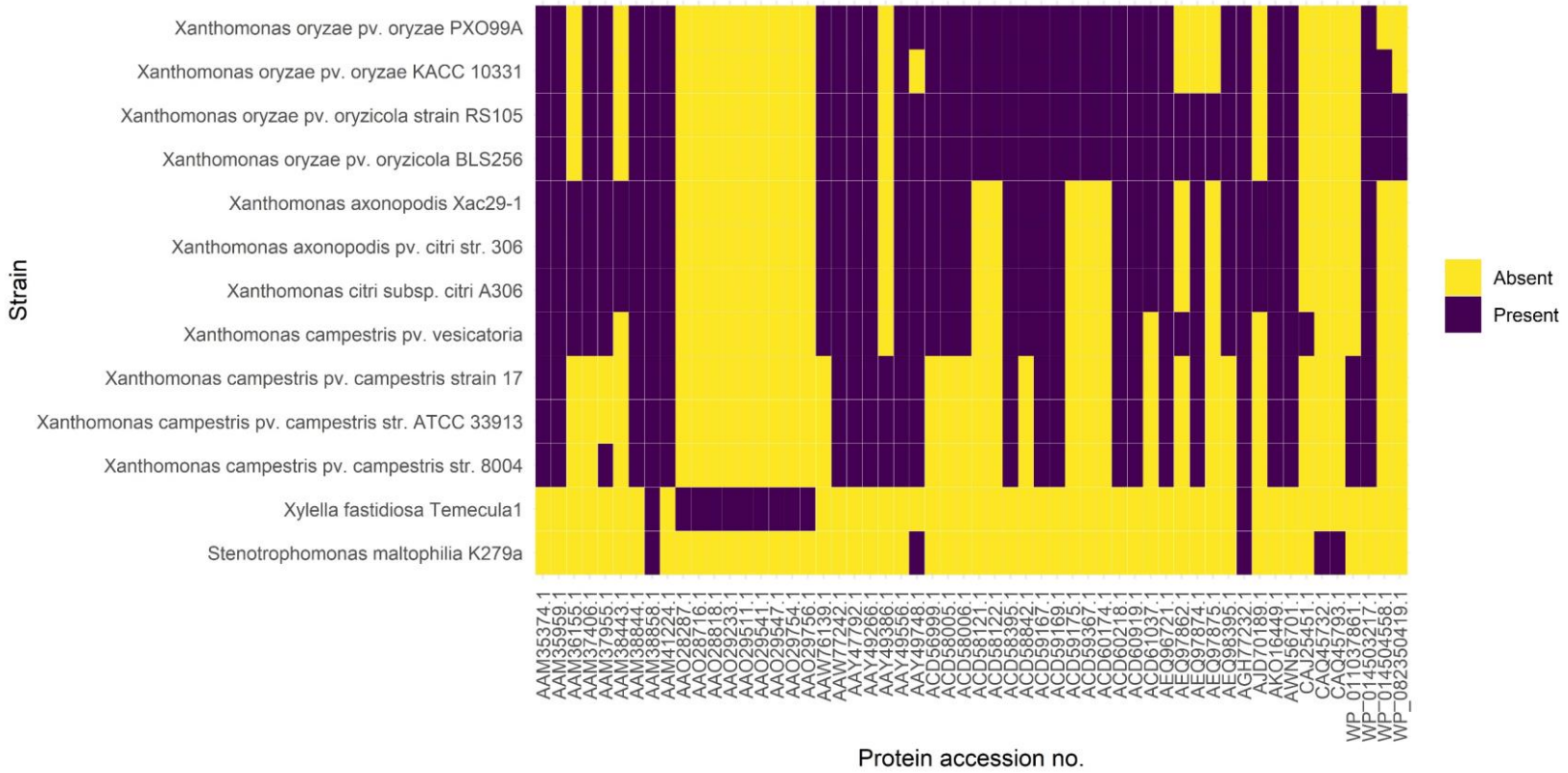


Figure A.3. The 56 biofilm-related proteins downloaded from previously published literature shown here as a presence/absence (90% BLAST percent identity threshold) plot in representative members of the order Xanthomonadales.

Appendix B

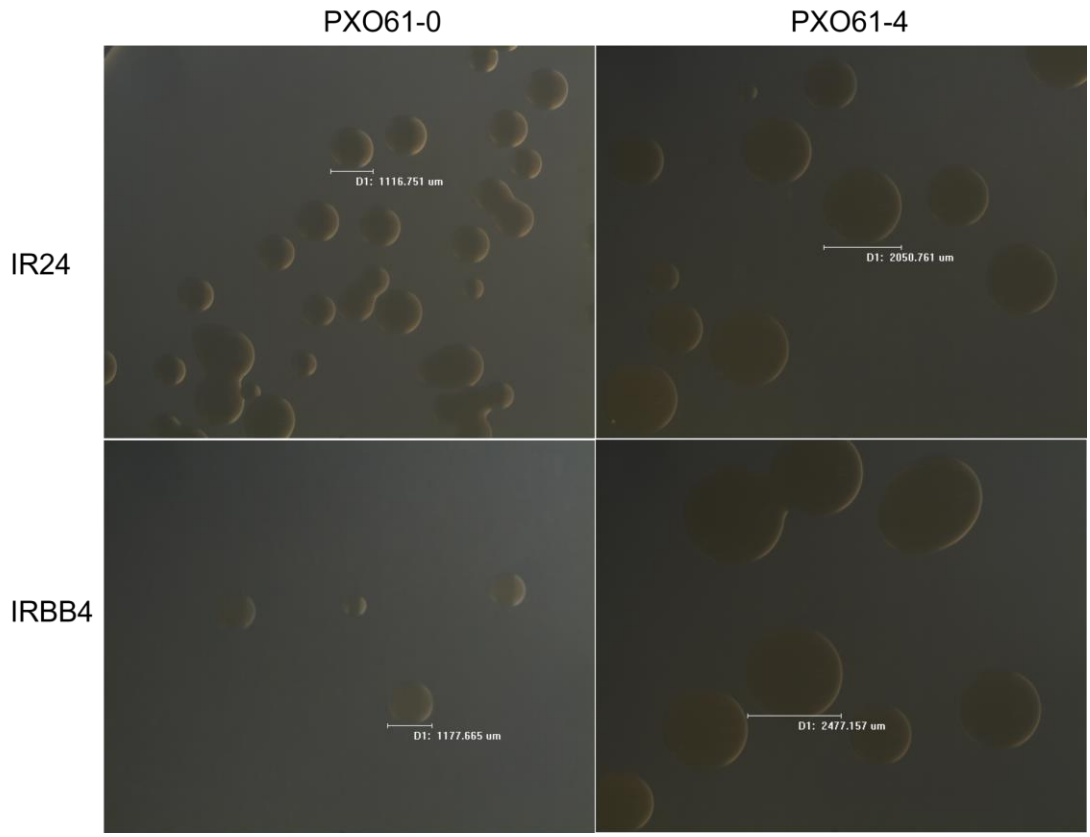


Figure B.1. Representative photos of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) taken 4 d after incubation at 30 °C on modified Wakimoto's agar. Colonies were re-isolated 24 hours after inoculation by infiltration in IRBB4 leaves and photographed with a QImaging Go-21 camera attached to an Olympus SZX16 Stereo Microscope. Colony diameter was measured using the QCapture Pro software.

Table B.1 Software and tests used for statistical analysis in Chapter 4 Exploring the adaptation potential of *Xanthomonas oryzae* pv. *oryzae* in rice through serial passage in resistant rice line IRBB4 (*Xa4*).

Data	Statistical test	Function	Software	Effect size	Software
Cycling	Wilcoxon	compare_means	ggpubr (Kassambara, 2020)	rank biserial correlation	Effectsize (Ben-Shachar, Lüdecke, & Makowski, 2020)
AUDPC	General linear hypothesis	glht	multcomp (Hothorn, Bretz, & Westfall, 2008)	Cohen's D	Effectsize (Ben-Shachar et al., 2020)
Wilt	Fisher's exact test	fisher.test	stats (R Core Team, 2021)	odds ratio	stats
Colony diameter	ANOVA	Anova	car (Fox & Weisberg, 2019)	eta_squared	Effectsize (Ben-Shachar et al., 2020)
Log CFU	ANOVA	Anova	car (Fox & Weisberg, 2019)		

Table B.2. Summarized leaf lesion length (cm) data measured from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB4 (*Xa4*). Data collection was performed 11 days post-inoculation.

Cycle	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
1	18	0.7	3.0	1.7	1.8	0.7	0.4	0.6	0.1	0.3	2.3	0.3	0.3
2	23	0.8	7.2	1.6	1.9	0.8	0.4	1.2	0.3	0.5	6.4	0.7	1.6
3	24	1.9	5.5	2.6	2.8	0.5	0.4	0.8	0.2	0.3	3.6	0.3	0.6
4	22	0.3	3.5	1.2	1.4	0.5	0.4	0.9	0.2	0.4	3.2	0.6	0.8
5	22	0.7	4.2	2.6	2.7	0.7	0.6	0.8	0.2	0.3	3.5	0.3	0.6
6	9	5.0	8.1	6.3	6.5	0.7	0.9	1.0	0.3	0.8	3.1	0.2	1.1
7	24	3.0	6.0	4.3	4.2	1.2	1.0	0.8	0.2	0.3	3.0	0.2	0.7
8	24	2.4	7.1	4.4	4.5	2.4	1.9	1.5	0.3	0.6	4.7	0.3	2.4
9	22	3.3	6.3	5.2	5.1	1.1	1.0	1.0	0.2	0.4	3.0	0.2	1.0
10	24	3.2	6.6	4.0	4.4	1.6	1.0	1.1	0.2	0.5	3.4	0.2	1.1

Table B.3. Multiple pairwise comparisons of mean lesion length data from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB4 (*Xa4*) against the base mean. Shown is the raw output from the R package ggpubr function compare_means. P-adj column indicates Holm adjusted p-values.

.y.	*group1*	*group2*	*p*	*p.adj*	*p.format*	*p.signif*	*method*
Lesion.Length	.all.	1	0.000	0.000	3.7e-05	****	Wilcoxon
Lesion.Length	.all.	2	0.000	0.000	4.6e-06	****	Wilcoxon
Lesion.Length	.all.	3	0.208	0.230	0.20790	ns	Wilcoxon
Lesion.Length	.all.	4	0.000	0.000	5.4e-08	****	Wilcoxon
Lesion.Length	.all.	5	0.113	0.230	0.11280	ns	Wilcoxon
Lesion.Length	.all.	6	0.000	0.000	1.3e-05	****	Wilcoxon
Lesion.Length	.all.	7	0.004	0.013	0.00392	**	Wilcoxon
Lesion.Length	.all.	8	0.003	0.013	0.00325	**	Wilcoxon
Lesion.Length	.all.	9	0.000	0.000	8.4e-06	****	Wilcoxon
Lesion.Length	.all.	10	0.001	0.005	0.00096	***	Wilcoxon

Table B.4. Summarized leaf lesion length (cm) data measured from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in susceptible rice variety IR24. Data collection was performed 11 days post-inoculation.

Cycle	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
1	19	5.4	10.1	8.5	8.1	2.0	1.6	1.4	0.3	0.7	4.7	0.2	2.0
2	23	4.2	11.3	8.0	7.8	1.9	1.5	1.7	0.3	0.7	7.1	0.2	2.8
3	20	3.5	14.0	7.3	7.5	2.8	1.9	2.4	0.5	1.1	10.5	0.3	5.6
4	24	6.5	21.5	13.8	13.5	7.3	5.5	4.2	0.8	1.8	15.0	0.3	17.3
5	23	9.3	20.0	13.0	13.5	3.8	3.0	3.0	0.6	1.3	10.7	0.2	8.7
6	16	7.5	11.0	9.0	9.0	1.0	0.7	0.9	0.2	0.5	3.5	0.1	0.9
7	22	6.2	14.0	8.4	8.9	1.5	1.4	1.8	0.4	0.8	7.8	0.2	3.4
8	24	8.5	20.7	14.0	13.6	2.9	2.2	2.9	0.6	1.2	12.2	0.2	8.6
9	24	7.0	15.5	11.6	11.2	5.2	3.7	2.9	0.6	1.2	8.5	0.3	8.4
10	24	7.5	14.0	10.0	10.4	2.9	2.2	1.9	0.4	0.8	6.5	0.2	3.7

Table B.5. Summarized leaf lesion length (cm) data measured from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB5 (*xa5*). Data collection was performed 11 days post-inoculation.

Cycle	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
1	21	0.5	2.3	1.0	1.0	0.3	0.3	0.4	0.1	0.2	1.8	0.4	0.2
2	20	0.3	1.5	0.9	0.9	0.6	0.4	0.3	0.1	0.2	1.2	0.4	0.1
3	23	0.3	2.5	0.9	1.0	0.6	0.4	0.6	0.1	0.2	2.2	0.6	0.3
4	21	0.2	1.8	0.7	0.8	0.6	0.3	0.4	0.1	0.2	1.6	0.5	0.2
5	24	0.7	2.5	1.4	1.5	0.8	0.5	0.5	0.1	0.2	1.8	0.3	0.2
6	14	1.0	1.7	1.3	1.3	0.3	0.2	0.2	0.1	0.1	0.7	0.2	0.0
7	24	0.7	3.0	1.6	1.6	0.4	0.4	0.5	0.1	0.2	2.3	0.3	0.2
8	24	0.7	2.5	1.2	1.4	0.9	0.7	0.6	0.1	0.2	1.8	0.4	0.3
9	24	0.7	2.7	1.7	1.7	0.4	0.4	0.5	0.1	0.2	2.0	0.3	0.3
10	24	1.0	2.5	1.8	1.8	0.9	0.7	0.5	0.1	0.2	1.5	0.3	0.3

Table B.6. Summarized leaf lesion length (cm) data measured from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB7 (*Xa7*). Data collection was performed 11 days post-inoculation.

Cycle	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
1	24	1.6	4.5	2.6	2.7	1.0	0.8	0.8	0.2	0.3	2.9	0.3	0.6
2	20	1.4	10.5	2.4	3.3	1.6	1.0	2.4	0.5	1.1	9.1	0.7	5.6
3	19	0.2	5.3	3.7	2.9	2.2	1.3	1.6	0.4	0.8	5.1	0.6	2.7
4	18	2.0	5.6	3.2	3.4	1.8	1.3	1.1	0.3	0.6	3.6	0.3	1.2
5	16	4.0	13.0	8.7	8.4	3.0	2.4	2.6	0.6	1.4	9.0	0.3	6.6
6	16	5.5	7.6	6.8	6.7	0.8	0.7	0.6	0.2	0.3	2.1	0.1	0.4
7	24	3.6	7.7	5.3	5.4	2.1	1.6	1.2	0.2	0.5	4.1	0.2	1.5
8	14	4.0	6.3	5.0	5.0	0.7	0.7	0.6	0.2	0.4	2.3	0.1	0.4
9	24	7.0	11.0	8.4	8.5	0.8	0.5	0.8	0.2	0.4	4.0	0.1	0.7
10	24	4.0	8.6	6.8	6.5	2.0	1.7	1.4	0.3	0.6	4.6	0.2	2.1

Table B.7. Multiple pairwise comparisons of mean lesion length data from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in susceptible rice variety IR24 against the base mean. Shown is the raw output from the R package ggpubr function compare_means. P-adj column indicates Holm adjusted p-values.

.y.	*group1*	*group2*	*p*	*p.adj*	*p.format*	*p.signif*	*method*
Lesion.Length	.all.	1	0.002	0.010	0.0020	**	Wilcoxon
Lesion.Length	.all.	2	0.000	0.001	9.9e-05	****	Wilcoxon
Lesion.Length	.all.	3	0.000	0.000	6.2e-05	****	Wilcoxon
Lesion.Length	.all.	4	0.000	0.003	0.0005	***	Wilcoxon
Lesion.Length	.all.	5	0.000	0.000	4.4e-05	****	Wilcoxon
Lesion.Length	.all.	6	0.171	0.510	0.1706	ns	Wilcoxon
Lesion.Length	.all.	7	0.031	0.120	0.0312	*	Wilcoxon
Lesion.Length	.all.	8	0.000	0.000	2.7e-05	****	Wilcoxon
Lesion.Length	.all.	9	0.241	0.510	0.2412	ns	Wilcoxon
Lesion.Length	.all.	10	0.642	0.640	0.6418	ns	Wilcoxon

Output from R package ggpubr function compare_means. P adj column indicates Holm adjusted p-values

Table B.8. Multiple pairwise comparisons of mean lesion length data from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB5 (*xa5*) against the base mean. Shown is the raw output from the R package ggpubr function compare_means. P-adj column indicates Holm adjusted p-values.

.y.	*group1*	*group2*	*p*	*p.adj*	*p.format*	*p.signif*	*method*
Lesion.Length	.all.	1	0.019	0.077	0.01932	*	Wilcoxon
Lesion.Length	.all.	2	0.000	0.004	0.00044	***	Wilcoxon
Lesion.Length	.all.	3	0.009	0.045	0.00910	**	Wilcoxon
Lesion.Length	.all.	4	0.000	0.000	4.4e-05	****	Wilcoxon
Lesion.Length	.all.	5	0.221	0.660	0.22136	ns	Wilcoxon
Lesion.Length	.all.	6	0.699	1.000	0.69855	ns	Wilcoxon
Lesion.Length	.all.	7	0.004	0.026	0.00435	**	Wilcoxon
Lesion.Length	.all.	8	0.549	1.000	0.54906	ns	Wilcoxon
Lesion.Length	.all.	9	0.003	0.019	0.00264	**	Wilcoxon
Lesion.Length	.all.	10	0.000	0.002	0.00019	***	Wilcoxon

Table B.9. Multiple pairwise comparisons of mean lesion length data from the serial passage of *Xanthomonas oryzae* pv. *oryzae* strain PXO61 by leaf clip-inoculation in resistant rice line IRBB7 (*Xa7*) against the base mean. Shown is the raw output from the R package ggpubr function compare_means. P-adj column indicates Holm adjusted p-values.

.y.	*group1*	*group2*	*p*	*p.adj*	*p.format*	*p.signif*	*method*
Lesion.Length	.all.	1	0.000	0.000	6.5e-07	****	Wilcoxon
Lesion.Length	.all.	2	0.000	0.001	0.00020	***	Wilcoxon
Lesion.Length	.all.	3	0.000	0.001	0.00018	***	Wilcoxon
Lesion.Length	.all.	4	0.003	0.013	0.00253	**	Wilcoxon
Lesion.Length	.all.	5	0.000	0.000	4.4e-05	****	Wilcoxon
Lesion.Length	.all.	6	0.012	0.036	0.01210	*	Wilcoxon
Lesion.Length	.all.	7	0.675	1.000	0.67535	ns	Wilcoxon
Lesion.Length	.all.	8	0.865	1.000	0.86458	ns	Wilcoxon
Lesion.Length	.all.	9	0.000	0.000	2.1e-09	****	Wilcoxon
Lesion.Length	.all.	10	0.008	0.032	0.00809	**	Wilcoxon

Output from R package ggpubr function compare_means. P adj column indicates Holm adjusted p-values

Table B.10. Summarized area under the disease progress curve (AUDPC) values calculated from lesion lengths measured at 5-, 7-, 9-, and 11-days post-inoculation from rice variety IR24 inoculated with *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4).

Strain	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
PXO61-0	15	10	22	18	17.5	4	3.0	3.7	1.0	2.1	12	0.2	13.8
PXO61-4	15	18	27	23	22.5	5	3.0	2.9	0.7	1.6	9	0.1	8.3

Table B.11. Summarized area under the disease progress curve (AUDPC) values calculated from lesion lengths measured at 5-, 7-, 9-, and 11-days post-inoculation from resistant rice line IRBB4 (*Xa4*) inoculated with *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4).

Strain	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
PXO61-0	15	9	20	15	14.9	0.5	0.0	2.6	0.7	1.4	11	0.2	6.6
PXO61-4	15	12	21	19	18.2	3.5	1.5	2.7	0.7	1.5	9	0.1	7.2

Table B.12. Comparison of mean AUDPC values calculated from lesion lengths measured at 5-, 7-, 9-, and 11-days post-inoculation measured from susceptible rice variety IR24 and resistant rice line IRBB4 (*Xa4*) inoculated with *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) in the greenhouse setting.

	AUDPC, IR24			**AUDPC, IRBB4**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	17.53	15.64 – 19.43	**<0.001**	14.93	13.50 – 16.37	**<0.001**
StrainPXO61-4	4.93	2.68 – 7.18	**<0.001**	3.27	1.45 – 5.08	**0.0004**
Random Effects						
σ^2	9.88			6.44		
τ_{00}	1.37 $_{Rep}$			0.55 $_{Rep}$		
ICC	0.12			0.08		
N	5 $_{Rep}$			5 $_{Rep}$		
Observations	30			30		
Marginal R2 / Conditional R2		0.359 / 0.437		0.283 / 0.339		

Table B.13. Comparison of estimated aerobic plate counts obtained from leaf discs sampled at 24 hours post-inoculation of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) by leaf clipping in susceptible rice variety IR24. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df F value    Pr(>F)
(Intercept) 170.364  1 341.865 4.618e-09 ***
Strain        0.001  1   0.001    0.9749
Residuals     4.983 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.14. Comparison of estimated aerobic plate counts obtained from leaf discs sampled at 48 hours post-inoculation of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) by leaf clipping in susceptible rice variety IR24. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df  F value    Pr(>F)
(Intercept) 152.768  1 323.6982 9.343e-08 ***
Strain        1.790  1   3.7935    0.0873 .
Residuals     3.776  8
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.15. Comparison of estimated aerobic plate counts obtained from leaf discs sampled at 24 hours post-inoculation of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) by leaf clipping in resistant rice line IRBB4 (Xa4). Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df  F value    Pr(>F)
(Intercept) 197.372  1 177.1523 1.097e-07 ***
Strain        0.120  1   0.1078    0.7495
Residuals    11.141 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.16. Comparison of estimated aerobic plate counts obtained from leaf discs sampled at 48 hours post-inoculation of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) by leaf clipping in resistant rice line IRBB4 (Xa4). Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df  F value    Pr(>F)
(Intercept) 189.575  1 245.7453 2.287e-08 ***
Strain        0.535  1   0.6939    0.4243
Residuals     7.714 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.17. Wilt in susceptible rice variety IR24 observed 17 days after inoculation with *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4).

Strain	*Observation*		***Total***
	No wilt	Wilt	
PXO61-0	243	47	290
PXO61-4	113	132	245
Total	356	179	535

χ^2=82.965 · df=1 · φ=0.398 · p=0.000

Table B.18. Growth parameters calculated after fitting hourly measurements of optical density (405 nm) obtained from *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) inoculated into BIOLOG GENIII plates into a logarithmic growth model generated in the growthcurveR (Sprouffske & Wagner, 2016) package. Wells with a "questionable fit" or "cannot fit data" in the note column were excluded in subsequent analysis.

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	Negative control	0.2634	0.0325	0.3974	4.9333	1.7443	8.0970	8.0720	0.0231	
PXO61-0	Dextrin	0.1079	0.0187	0.2594	6.0277	2.6724	3.1537	3.1520	0.0219	
PXO61-0	D-Maltose	0.5591	0.0210	0.3386	9.5825	2.0472	14.7068	14.6595	0.0205	
PXO61-0	D-Trehalose	0.1930	0.0163	0.3788	6.2914	1.8296	5.6892	5.6740	0.0137	
PXO61-0	D-Cellobiose	0.0939	0.0000	9.1616	2.7676	0.0757	3.1208	3.1505	0.0308	
PXO61-0	Gentiobiose	0.4979	0.0494	0.2284	9.6621	3.0349	12.8908	12.8325	0.0251	
PXO61-0	Sucrose	0.0338	0.0000	0.9615	33.5109	0.7209	0.0873	0.1325	0.0040	
PXO61-0	D-Turanose	0.4895	0.0626	0.2564	7.4870	2.7035	13.6977	13.6965	0.0209	
PXO61-0	Stachyose	0.3088	0.0048	1.1789	3.5252	0.5880	10.0227	10.0010	0.0105	
PXO61-0	Positive control	0.2549	0.0425	0.2249	7.1535	3.0827	7.1481	7.1410	0.0262	
PXO61-0	pH6	0.5701	0.0303	0.2836	10.1617	2.4444	14.6232	14.5820	0.0230	
PXO61-0	pH5	0.1367	0.0079	0.3977	7.0303	1.7428	3.9394	3.9315	0.0167	
PXO61-0	D-Raffinose	0.3234	0.0150	0.3814	7.9237	1.8175	9.0402	9.0070	0.0166	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	a-D-Lactose	0.3309	0.0308	0.4271	5.3335	1.6228	10.0731	10.0420	0.0182	
PXO61-0	D-Mellibiose	0.5096	0.0116	0.2793	13.4728	2.4820	11.4421	11.3495	0.0163	
PXO61-0	ß-Methyl-D-Glucoside	0.1898	0.0123	0.3112	8.5913	2.2276	5.1617	5.1230	0.0140	
PXO61-0	D-Salicin	0.3285	0.0495	0.1749	9.8954	3.9641	8.2875	8.2690	0.0308	
PXO61-0	N-Acetyl-D-Glucosamine	0.3046	0.0361	0.2547	7.8752	2.7213	8.4177	8.3905	0.0257	
PXO61-0	N-Acetyl-ß-D-Mannosamine	0.5764	0.0503	0.2561	9.1637	2.7065	15.2639	15.2070	0.0285	
PXO61-0	N-Acetyl-D-Galactosamine	0.3565	0.0139	0.5103	6.2870	1.3584	10.5657	10.5375	0.0214	
PXO61-0	N-Acetyl Neuraminic Acid	0.3829	0.0277	0.1980	12.8885	3.5001	8.7248	8.6480	0.0207	
PXO61-0	1% NaCl	0.2009	0.0016	0.5696	8.4693	1.2168	5.5275	5.5115	0.0111	
PXO61-0	4% NaCl	0.2467	0.0133	0.3497	8.1884	1.9824	6.8222	6.7910	0.0161	
PXO61-0	8% NaCl	0.2048	0.0007	0.2481	23.0779	2.7937	2.6764	2.6030	0.0087	
PXO61-0	a-D-Glucose	0.2075	0.0000	3.7146	2.6718	0.1866	6.9155	6.8900	0.0281	
PXO61-0	D-Mannose	0.3398	0.0151	0.2759	11.1203	2.5122	8.3998	8.3660	0.0175	
PXO61-0	D-Fructose	0.4519	0.0082	0.4138	9.6579	1.6753	11.8833	11.8560	0.0151	
PXO61-0	D-Galactose	0.3288	0.0174	0.3601	8.0182	1.9247	9.1513	9.1155	0.0197	
PXO61-0	3-Methyl Glucose	0.3846	0.0200	0.3373	8.6054	2.0548	10.4741	10.3760	0.0237	
PXO61-0	D-Fucose	0.3190	0.0231	0.3393	7.5099	2.0428	9.0167	8.9515	0.0194	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	L-Fucose	0.5130	0.0149	0.3607	9.7274	1.9215	13.4358	13.3395	0.0222	
PXO61-0	L-Rhamnose	0.1989	0.0011	0.4223	12.3177	1.6412	4.7079	4.6535	0.0117	
PXO61-0	Inosine	0.6810	0.0640	0.1645	13.7716	4.2132	14.8338	14.7835	0.0309	
PXO61-0	1% Sodium Lactate	0.4851	0.0161	0.2434	13.8608	2.8474	10.6810	10.6275	0.0138	
PXO61-0	Fusidic acid	0.1907	0.0121	0.3471	7.7542	1.9970	5.3491	5.3250	0.0171	
PXO61-0	D-Serine INH	0.2777	0.0227	0.3060	7.9085	2.2650	7.7250	7.7045	0.0195	
PXO61-0	D-Sorbitol	0.4788	0.0211	0.2954	10.4178	2.3462	12.1766	12.0895	0.0200	
PXO61-0	D-Mannitol	0.2995	0.0166	0.4424	6.4147	1.5666	8.8224	8.7680	0.0196	
PXO61-0	D-Arabitol	0.4826	0.0129	0.4330	8.2972	1.6008	13.3391	13.2530	0.0255	
PXO61-0	myo-Inositol	0.0007	0.0270	11.1991	-0.0025	0.0619	0.0272	0.0270	0.0065	questionable fit (k < n0)
PXO61-0	Glycerol	0.3460	0.0678	0.2367	5.9620	2.9282	10.0755	10.0345	0.0344	
PXO61-0	D-Glucose-6-PO4	0.4834	0.0085	0.3875	10.3877	1.7890	12.3587	12.3065	0.0189	
PXO61-0	D-Fructose-6-PO4	0.2092	0.0068	0.2685	12.6494	2.5817	4.8612	4.8400	0.0132	
PXO61-0	D-Aspartic Acid	0.3074	0.0236	0.3174	7.8329	2.1838	8.5805	8.5295	0.0192	
PXO61-0	D-Serine	0.4721	0.0182	0.3564	9.0286	1.9449	12.6808	12.6265	0.0152	
PXO61-0	Troleandomycin	0.2280	0.0166	0.4484	5.6700	1.5459	6.8753	6.8405	0.0172	
PXO61-0	Rifamycin SV	0.4689	0.0227	0.3020	9.8667	2.2955	12.1771	12.0725	0.0205	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	Minocycline	0.4580	0.0056	1.3129	3.3495	0.5280	14.9509	14.9305	0.0180	
PXO61-0	Gelatin	0.5950	0.0643	0.2019	10.4528	3.4330	14.8802	14.8240	0.0350	
PXO61-0	Glycyl-L-Proline	0.7029	0.1315	0.2047	7.1783	3.3865	19.5564	19.5605	0.0672	
PXO61-0	L-Alanine	0.5628	0.0474	0.2890	8.2545	2.3980	15.4437	15.3985	0.0302	
PXO61-0	L-Arginine	0.4471	0.0360	0.3503	6.9559	1.9789	12.8785	12.8335	0.0240	
PXO61-0	L-Aspartic Acid	0.5436	0.0476	0.1058	22.1570	6.5542	8.1230	8.1050	0.0209	
PXO61-0	L-Glutamic Acid	0.3209	0.0186	0.3751	7.4315	1.8477	9.1166	9.0665	0.0220	
PXO61-0	L-Histidine	0.5361	0.0313	0.3159	8.8068	2.1941	14.4770	14.3870	0.0214	
PXO61-0	L-Pyroglutamic Acid	0.3540	0.0105	0.3591	9.7103	1.9304	9.2765	9.1905	0.0159	
PXO61-0	L-Serine	0.2428	0.0326	0.2426	7.6776	2.8568	6.7340	6.6845	0.0229	
PXO61-0	Lincomycin	0.2536	0.0088	0.3270	10.1559	2.1197	6.5277	6.4770	0.0177	
PXO61-0	Guanidine HCl	0.1730	0.0080	0.3718	8.1551	1.8642	4.7939	4.7820	0.0143	
PXO61-0	Niaproof 4	0.2767	0.0035	0.4218	10.3283	1.6431	7.0938	7.0500	0.0150	
PXO61-0	Pectin	0.2746	0.0081	0.3741	9.3270	1.8529	7.3023	7.2495	0.0148	
PXO61-0	D-Galacturonic Acid	0.2502	0.0010	0.6626	8.2669	1.0460	6.9379	6.9015	0.0180	
PXO61-0	L-Galactonic Acid Lactone	0.5788	0.0235	0.3892	8.1280	1.7810	16.0712	15.9445	0.0337	
PXO61-0	D-Gluconic Acid	0.5364	0.0284	0.4388	6.5755	1.5795	15.7159	15.5975	0.0284	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	D-Glucuronic Acid	0.2620	0.0339	0.2255	8.4551	3.0732	7.0577	7.0210	0.0216	
PXO61-0	Glucuronamide	0.3111	0.0330	0.2322	9.1763	2.9850	8.1963	8.1875	0.0194	
PXO61-0	Muric Acid	0.2118	0.0208	0.2526	8.7806	2.7440	5.6788	5.6650	0.0178	
PXO61-0	Quinic Acid	0.1733	0.0168	0.2294	9.7248	3.0216	4.4782	4.4645	0.0167	
PXO61-0	D-Saccharic Acid	0.4771	0.0166	0.3341	9.9457	2.0747	12.3809	12.2910	0.0196	
PXO61-0	Vancomycin	0.3325	0.0187	0.3776	7.4681	1.8358	9.4359	9.3925	0.0162	
PXO61-0	Tetrazolium Violet	0.4746	0.0099	0.3129	12.3047	2.2152	11.2138	11.1240	0.0183	
PXO61-0	Tetrazolium Blue	0.1941	0.0009	0.2671	20.1971	2.5950	3.0754	2.9890	0.0125	
PXO61-0	p-Hydroxy-Phenylacetic Acid	0.6721	0.0495	0.2110	11.9985	3.2856	15.9079	15.7435	0.0378	
PXO61-0	Methyl Pyruvate	0.5193	0.0372	0.1949	13.1445	3.5561	11.7013	11.4995	0.0492	
PXO61-0	D-Lactic Acid Methyl Ester	0.4398	0.0154	0.3115	10.6352	2.2250	11.1061	10.9490	0.0344	
PXO61-0	L-Lactic Acid	0.3391	0.0257	0.2469	10.1236	2.8069	8.6686	8.6045	0.0209	
PXO61-0	Citric Acid	0.4046	0.0144	0.3220	10.2432	2.1528	10.3766	10.2715	0.0225	
PXO61-0	a-Keto-Glutaric Acid	0.3086	0.0029	0.5674	8.1970	1.2217	8.5747	8.5270	0.0158	
PXO61-0	D-Malic Acid	0.4151	0.0166	0.2945	10.8017	2.3539	10.4030	10.2950	0.0200	
PXO61-0	L-Malic Acid	0.4019	0.0262	0.3868	6.8902	1.7921	11.6301	11.5870	0.0150	
PXO61-0	Bromo-Succinic Acid	0.5746	0.0197	0.2673	12.4948	2.5933	13.4341	13.3125	0.0223	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-0	Nalidixic Acid	0.2397	0.0002	0.3466	20.2990	1.9999	3.7658	3.7155	0.0156	
PXO61-0	Lithium Chloride	0.0120	0.0000	0.9701	25.6227	0.7145	0.1245	0.1265	0.0054	
PXO61-0	Potassium Tellurite	0.0793	0.0005	0.4745	10.6309	1.4606	2.0115	2.0105	0.0129	
PXO61-0	Tween 40	34523.1394	0.0003	0.1224	151.1297	5.6626	0.2111	0.2035	0.0074	
PXO61-0	-Amino-Butryric Acid	0.0937	0.0019	3.6256	1.0716	0.1912	3.2720	3.2400	0.0184	
PXO61-0	a-Hydroxy-Butyric Acid	0.5098	0.0210	0.3021	10.4163	2.2943	12.9713	12.8795	0.0260	
PXO61-0	a-Hydroxy-D,L-Butyric Acid	0.5143	0.0316	0.2222	12.2688	3.1199	12.0702	12.0305	0.0214	
PXO61-0	a-Keto-Butyric Acid	0.0586	0.0005	0.2070	23.0318	3.3492	0.7764	0.7455	0.0114	
PXO61-0	Acetoacetic acid	0.0869	0.0000	1.1763	36.8766	0.5892	0.0225	0.0200	0.0017	
PXO61-0	Propionic Acid	0.4572	0.0147	0.2930	11.6321	2.3660	11.0903	11.0045	0.0213	
PXO61-0	Acetic Acid	0.1622	0.0019	0.3696	12.0238	1.8755	3.8830	3.8565	0.0119	
PXO61-0	Formic Acid	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	cannot fit data
PXO61-0	Aztreonam	0.8946	0.0330	0.5407	6.0340	1.2820	26.7456	26.7395	0.0259	
PXO61-0	Sodium butyrate	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	cannot fit data
PXO61-0	Sodium bromate	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	cannot fit data
PXO61-4	Negative control	0.3719	0.0318	0.1756	13.4968	3.9478	8.2198	8.1955	0.0215	
PXO61-4	Dextrin	0.2744	0.0416	0.1944	8.8518	3.5654	7.2245	7.2200	0.0177	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	D-Maltose	0.5705	0.0425	0.2662	9.4657	2.6036	14.9751	14.9180	0.0295	
PXO61-4	D-Trehalose	0.3299	0.0318	0.1888	11.8537	3.6714	7.8071	7.7720	0.0217	
PXO61-4	D-Cellobiose	0.2453	0.0326	0.2808	6.6840	2.4682	7.0667	7.0745	0.0160	
PXO61-4	Gentiobiose	0.5437	0.0587	0.1559	13.5466	4.4455	11.9124	11.8825	0.0256	
PXO61-4	Sucrose	0.1131	0.0044	0.1836	17.4092	3.7754	2.0974	2.0580	0.0124	
PXO61-4	D-Turanose	0.6246	0.0715	0.1876	10.8992	3.6939	15.3023	15.3225	0.0254	
PXO61-4	Stachyose	0.5238	0.0739	0.3218	5.6146	2.1537	15.6681	15.5920	0.0417	
PXO61-4	Positive control	0.4674	0.0741	0.1911	8.7310	3.6263	12.3364	12.3080	0.0295	
PXO61-4	pH6	0.5868	0.0565	0.2544	8.8027	2.7245	15.7290	15.6885	0.0337	
PXO61-4	pH5	0.3749	0.0408	0.2264	9.2884	3.0620	9.8261	9.7790	0.0231	
PXO61-4	D-Raffinose	0.4720	0.0357	0.2339	10.6943	2.9632	11.7896	11.7610	0.0170	
PXO61-4	a-D-Lactose	0.5039	0.0466	0.2791	8.1796	2.4832	13.8435	13.7565	0.0294	
PXO61-4	D-Mellibiose	0.6147	0.0295	0.2434	12.2710	2.8483	14.4705	14.3760	0.0237	
PXO61-4	ß-Methyl-D-Glucoside	0.3549	0.0154	0.2682	11.5209	2.5847	8.6297	8.5740	0.0179	
PXO61-4	D-Salicin	0.5166	0.0932	0.1772	8.5389	3.9108	13.6282	13.5990	0.0402	
PXO61-4	N-Acetyl-D-Glucosamine	0.4484	0.0400	0.2561	9.0747	2.7070	11.9104	11.8700	0.0247	
PXO61-4	N-Acetyl-ß-D-Mannosamine	0.6029	0.0485	0.2730	8.9250	2.5390	16.1400	16.0560	0.0317	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	N-Acetyl-D-Galactosamine	0.5043	0.0466	0.2439	9.3710	2.8424	13.2322	13.1930	0.0240	
PXO61-4	N-Acetyl Neuraminic Acid	0.5260	0.0354	0.1915	13.7219	3.6188	11.5652	11.5045	0.0242	
PXO61-4	1% NaCl	0.5128	0.0479	0.2788	8.1480	2.4863	14.1020	14.0575	0.0261	
PXO61-4	4% NaCl	0.4498	0.0308	0.2448	10.6647	2.8319	11.2701	11.2380	0.0203	
PXO61-4	8% NaCl	0.3856	0.0328	0.1250	19.0093	5.5456	6.6256	6.6245	0.0155	
PXO61-4	a-D-Glucose	0.4094	0.0766	0.1892	7.7682	3.6645	11.1206	11.0955	0.0336	
PXO61-4	D-Mannose	0.5313	0.0308	0.2447	11.3931	2.8326	12.9493	12.8895	0.0234	
PXO61-4	D-Fructose	0.5128	0.0239	0.3179	9.5000	2.1802	13.5117	13.4580	0.0205	
PXO61-4	D-Galactose	0.5350	0.0455	0.2380	9.9785	2.9120	13.7268	13.6510	0.0264	
PXO61-4	3-Methyl Glucose	0.5272	0.0356	0.2379	11.0365	2.9137	13.0111	12.9390	0.0237	
PXO61-4	D-Fucose	0.5398	0.0418	0.2490	9.9460	2.7838	13.8911	13.8320	0.0231	
PXO61-4	L-Fucose	0.5537	0.0340	0.3032	8.9907	2.2865	14.8388	14.7770	0.0284	
PXO61-4	L-Rhamnose	0.4476	0.0457	0.2314	9.3990	2.9953	11.7033	11.6750	0.0253	
PXO61-4	Inosine	0.7061	0.0753	0.1722	12.3374	4.0247	16.3138	16.2615	0.0304	
PXO61-4	1% Sodium Lactate	0.5879	0.0381	0.2218	12.0396	3.1254	13.9208	13.8260	0.0276	
PXO61-4	Fusidic acid	0.2203	0.0263	0.1967	10.1531	3.5241	5.5592	5.5215	0.0187	
PXO61-4	D-Serine INH	0.4771	0.0340	0.2270	11.3155	3.0536	11.6285	11.5995	0.0198	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	D-Sorbitol	0.5475	0.0390	0.2601	9.8686	2.6645	14.1549	14.0995	0.0264	
PXO61-4	D-Mannitol	0.4760	0.0305	0.2840	9.4438	2.4411	12.5308	12.4540	0.0271	
PXO61-4	D-Arabitol	0.5253	0.0170	0.4083	8.3184	1.6977	14.4975	14.4230	0.0277	
PXO61-4	myo-Inositol	0.1243	0.0004	0.2392	24.3840	2.8982	1.4732	1.4365	0.0107	
PXO61-4	Glycerol	0.5356	0.0514	0.2364	9.4911	2.9327	13.9735	13.9095	0.0281	
PXO61-4	D-Glucose-6-PO4	0.5993	0.0452	0.3154	7.9478	2.1975	16.6626	16.5945	0.0291	
PXO61-4	D-Fructose-6-PO4	0.3820	0.0294	0.2231	11.1379	3.1065	9.3658	9.2895	0.0251	
PXO61-4	D-Aspartic Acid	0.5711	0.0477	0.2571	9.3199	2.6957	15.0449	14.9970	0.0226	
PXO61-4	D-Serine	0.6090	0.0570	0.2656	8.5512	2.6098	16.4926	16.4240	0.0263	
PXO61-4	Troleandomycin	0.4182	0.0486	0.2742	7.4028	2.5279	11.7717	11.7275	0.0245	
PXO61-4	Rifamycin SV	0.5219	0.0192	0.3556	9.1836	1.9493	13.9395	13.8385	0.0300	
PXO61-4	Minocycline	0.5125	0.0207	0.9177	3.4512	0.7553	16.6568	16.6005	0.0232	
PXO61-4	Gelatin	0.6349	0.0857	0.2040	9.1077	3.3986	16.6362	16.5555	0.0439	
PXO61-4	Glycyl-L-Proline	0.6041	0.0481	0.2242	10.9167	3.0921	14.9390	14.8795	0.0268	
PXO61-4	L-Alanine	0.5813	0.0447	0.3307	7.5160	2.0960	16.4159	16.3660	0.0288	
PXO61-4	L-Arginine	0.4869	0.0215	0.3278	9.3845	2.1144	12.8923	12.8395	0.0226	
PXO61-4	L-Aspartic Acid	0.5183	0.0454	0.1504	15.5773	4.6081	10.4252	10.3730	0.0222	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	L-Glutamic Acid	0.5386	0.0338	0.2414	11.2005	2.8713	13.2184	13.1715	0.0194	
PXO61-4	L-Histidine	0.6242	0.0478	0.2300	10.8284	3.0136	15.5033	15.4030	0.0303	
PXO61-4	L-Pyroglutamic Acid	0.5445	0.0374	0.5052	5.1586	1.3720	16.7178	16.6555	0.0290	
PXO61-4	L-Serine	0.5072	0.0644	0.2134	9.0347	3.2480	13.3621	13.3375	0.0238	
PXO61-4	Lincomycin	0.4266	0.0176	0.2835	11.0874	2.4451	10.5655	10.4615	0.0243	
PXO61-4	Guanidine HCl	0.3216	0.0308	0.2728	8.2279	2.5409	8.8131	8.7395	0.0241	
PXO61-4	Niaproof 4	0.4618	0.0330	0.1926	13.3096	3.5991	10.3299	10.2880	0.0192	
PXO61-4	Pectin	0.4913	0.0402	0.2173	11.1213	3.1892	12.0409	12.0150	0.0218	
PXO61-4	D-Galacturonic Acid	0.5454	0.0591	0.2259	9.3271	3.0689	14.2748	14.2600	0.0239	
PXO61-4	L-Galactonic Acid Lactone	0.6107	0.0384	0.3179	8.4966	2.1801	16.6718	16.6085	0.0233	
PXO61-4	D-Gluconic Acid	0.5217	0.0155	1.1889	2.9338	0.5830	17.2376	17.2295	0.0246	
PXO61-4	D-Glucuronic Acid	0.5149	0.0517	0.2323	9.4417	2.9841	13.4448	13.4290	0.0236	
PXO61-4	Glucuronamide	0.5155	0.0344	0.2554	10.3261	2.7142	13.0993	13.0080	0.0264	
PXO61-4	Muric Acid	0.4350	0.0566	0.1736	10.9439	3.9928	10.5829	10.5560	0.0257	
PXO61-4	Quinic Acid	0.3541	0.0604	0.1548	10.2080	4.4767	8.7454	8.7395	0.0224	
PXO61-4	D-Saccharic Acid	0.6085	0.0682	0.2154	9.6122	3.2181	15.7302	15.6825	0.0303	
PXO61-4	Vancomycin	0.5321	0.0412	0.2423	10.2300	2.8611	13.5390	13.5195	0.0249	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	Tetrazolium Violet	0.5756	0.0184	0.2589	13.1761	2.6772	13.0721	12.9380	0.0273	
PXO61-4	Tetrazolium Blue	0.4749	0.0409	0.2819	8.3773	2.4589	12.9676	12.8940	0.0253	
PXO61-4	p-Hydroxy-Phenylacetic Acid	0.5861	0.0500	0.2223	10.6678	3.1183	14.6204	14.5065	0.0435	
PXO61-4	Methyl Pyruvate	0.3621	0.0054	0.5026	8.3222	1.3792	10.0106	9.9820	0.0354	
PXO61-4	D-Lactic Acid Methyl Ester	0.3532	0.0083	0.3740	9.9676	1.8532	9.1732	9.0565	0.0310	
PXO61-4	L-Lactic Acid	0.4083	0.0523	0.2862	6.6999	2.4219	11.7686	11.7425	0.0351	
PXO61-4	Citric Acid	0.4654	0.0292	0.2884	9.3757	2.4037	12.2881	12.1960	0.0334	
PXO61-4	a-Keto-Glutaric Acid	0.4019	0.0175	0.3692	8.3701	1.8773	11.0567	10.9845	0.0374	
PXO61-4	D-Malic Acid	0.5509	0.0337	0.2305	11.8473	3.0075	13.1640	13.0800	0.0337	
PXO61-4	L-Malic Acid	0.4894	0.0418	0.3025	7.8344	2.2910	13.6391	13.5855	0.0335	
PXO61-4	Bromo-Succinic Acid	0.5882	0.0220	0.2270	14.3138	3.0534	12.6755	12.6210	0.0372	
PXO61-4	Nalidixic Acid	0.4324	0.0104	0.1882	19.6778	3.6829	7.1055	7.2235	0.0313	
PXO61-4	Lithium Chloride	0.0000	0.0010	12.1706	-0.0023	0.0570	0.0010	0.0010	0.0002	questionable fit (k < n0)
PXO61-4	Potassium Tellurite	0.1799	0.0017	0.3055	15.1648	2.2691	3.7443	3.7140	0.0230	
PXO61-4	Tween 40	0.1005	0.0012	0.1437	30.7335	4.8220	0.7901	0.7905	0.0104	
PXO61-4	-Amino-Butryric Acid	0.3694	0.0079	4.5748	0.8371	0.1515	12.9873	12.9660	0.0843	
PXO61-4	a-Hydroxy-Butyric Acid	0.4987	0.0319	0.4293	6.2529	1.6147	14.7577	14.6710	0.0385	

Table B.18. Continued

Strain	*Sample*	*k*	*n0*	*r*	*t_mid*	*t_gen*	*auc_l*	*auc_e*	*sigma*	*note*
PXO61-4	a-Hydroxy-D,L-Butyric Acid	0.4465	0.0202	0.4416	6.9049	1.5696	12.9442	12.8610	0.0341	
PXO61-4	a-Keto-Butyric Acid	0.1273	0.0155	0.1099	17.9607	6.3078	2.2960	2.2735	0.0205	
PXO61-4	Acetoacetic acid	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	cannot fit data
PXO61-4	Propionic Acid	0.5257	0.0349	0.2154	12.2692	3.2174	12.3213	12.2450	0.0338	
PXO61-4	Acetic Acid	0.2167	0.0100	0.2262	13.3746	3.0637	4.8628	4.7880	0.0231	
PXO61-4	Formic Acid	0.0786	0.0116	0.0301	58.2296	23.0583	0.6632	0.6665	0.0131	
PXO61-4	Aztreonam	0.6819	0.2764	0.2861	1.3395	2.4223	22.3953	22.3710	0.1191	
PXO61-4	Sodium butyrate	56663.8282	0.0004	0.1291	145.0728	5.3695	0.3336	0.3145	0.0097	
PXO61-4	Sodium bromate	0.0587	0.0018	0.6228	5.5554	1.1129	1.7845	1.7590	0.0256	

Table B.19. Summarized data of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) colony diameter after re-isolation from inoculation in susceptible rice variety IR24 and resistant line IRBB4 (*Xa4*) at 24 and 48 hours post-inoculation[α]

VarietyLine	*Time.point*	*Strain*	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
IR24	24	PXO61-0	7	1238.6	1482.2	1385.8	1375.6	94.8	72.8	81.2	30.7	75.1	243.7	0.1	6585.9
IR24	24	PXO61-4	4	2126.9	2583.8	2251.3	2303.3	277.9	173.1	215.3	107.6	342.5	456.9	0.1	46335.9
IR24	48	PXO61-0	5	588.8	1198.0	1066.0	972.6	223.4	195.7	241.8	108.1	300.2	609.1	0.2	58461.1
IR24	48	PXO61-4	6	1609.1	2228.4	1827.4	1893.9	321.8	237.1	242.3	98.9	254.3	619.3	0.1	58702.5
IRBB4	24	PXO61-0	6	1477.2	1979.7	1487.3	1631.1	295.7	11.3	231.2	94.4	242.7	502.5	0.1	53468.7
IRBB4	24	PXO61-4	6	2345.9	2593.9	2503.9	2494.4	28.6	27.6	81.6	33.3	85.6	248.0	0.0	6657.0
IRBB4	48	PXO61-0	3	893.4	1111.7	969.5	991.5	109.1	112.9	110.8	64.0	275.2	218.3	0.1	12273.8
IRBB4	48	PXO61-4	4	1827.4	2192.9	1961.9	1986.0	217.0	161.8	166.5	83.2	264.9	365.5	0.1	27714.8

[α]Measurements expressed in micrometers. Colony diameter was measured 3 d and 4 d after incubation of samples collected at the 48 hours post-inoculation (hpi) and 24 hpi time points, respectively.

Table B.20. Comparison of mean colony diameter of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from susceptible variety IR24 at 24 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: mean diameter
           Sum Sq Df F value    Pr(>F)
Strain    2190522  1  110.43 2.364e-06 ***
Residuals  178523  9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.21. Comparison of mean colony diameter of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from re-isolated from the resistant line IRBB4 (*Xa4*) at 24 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: mean diameter
           Sum Sq Df F value    Pr(>F)
Strain    2235438  1  74.359 6.069e-06 ***
Residuals  300628 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.22. Comparison of mean colony diameter of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from susceptible variety IR24 at 48 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: mean diameter
           Sum Sq Df F value    Pr(>F)
Strain    2314728  1  39.504 0.0001435 ***
Residuals  527357  9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.23. Comparison of mean colony diameter of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from re-isolated from the resistant line IRBB4 (*Xa4*) at 48 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: mean diameter
           Sum Sq Df F value    Pr(>F)
Strain    1695483  1  78.719 0.0003025 ***
Residuals  107692  5
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.24. Summarized data of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) estimated aerobic plate count (log CFU per cm^2) after re-isolation from inoculation in susceptible rice variety IR24 and resistant line IRBB4 (*Xa4*) at 24 and 48 hours post-inoculation[α].

VarietyLine	*Time.point*	*Strain*	*length*	*min*	*max*	*median*	*mean*	*iqr*	*mad*	*sd*	*se*	*ci*	*range*	*cv*	*var*
IR24	24	PXO61-0	6	3.9	6.5	5.2	5.3	0.8	0.7	0.9	0.4	0.9	2.6	0.2	0.8
IR24	24	PXO61-4	6	5.0	6.1	5.1	5.3	0.5	0.2	0.4	0.2	0.5	1.1	0.1	0.2
IR24	48	PXO61-0	5	5.1	6.5	5.1	5.5	0.8	0.0	0.6	0.3	0.8	1.4	0.1	0.4
IR24	48	PXO61-4	5	5.1	7.0	6.6	6.4	0.3	0.3	0.8	0.3	0.9	1.9	0.1	0.6
IRBB4	24	PXO61-0	6	3.2	7.1	6.1	5.7	1.3	1.0	1.4	0.6	1.5	3.9	0.2	2.0
IRBB4	24	PXO61-4	6	5.1	6.2	5.4	5.5	0.7	0.4	0.5	0.2	0.5	1.1	0.1	0.2
IRBB4	48	PXO61-0	6	4.5	6.4	5.7	5.6	1.2	0.9	0.8	0.3	0.9	1.9	0.1	0.7
IRBB4	48	PXO61-4	6	4.8	7.1	6.2	6.0	1.5	1.0	0.9	0.4	1.0	2.2	0.2	0.9

[α]Colonies were counted 3 d and 4 d after incubation of samples collected at the 48 hours post-inoculation (hpi) and 24 hpi time points, respectively.

Table B.25. Comparison of estimated aerobic plate counts of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from susceptible variety IR24 at 24 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df F value    Pr(>F)
(Intercept) 170.364  1 341.865 4.618e-09 ***
Strain        0.001  1   0.001    0.9749
Residuals     4.983 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.26. Comparison of estimated aerobic plate counts of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from susceptible variety IR24 at 48 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

Response: logCFU

```
             Sum Sq Df  F value    Pr(>F)
(Intercept) 152.768  1 323.6982 9.343e-08 ***
Strain        1.790  1   3.7935    0.0873 .
Residuals     3.776  8
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.27. Comparison of estimated aerobic plate counts of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from resistant line IRBB4 (*Xa4*) at 24 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

Response: logCFU

```
             Sum Sq Df  F value    Pr(>F)
(Intercept) 197.372  1 177.1523 1.097e-07 ***
Strain        0.120  1   0.1078    0.7495
Residuals    11.141 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table B.28. Comparison of estimated aerobic plate counts of *Xanthomonas oryzae* pv. *oryzae* PXO61 wild type (designated as PXO61-0) and PXO61-4 (passaged in resistant line IRBB4) re-isolated from resistant line IRBB4 (*Xa4*) at 48 hours post-inoculation. Shown is the raw output of the ANOVA (Type III tests) performed in R using the package car (Fox & Weisberg, 2019).

```
Response: logCFU
             Sum Sq Df  F value    Pr(>F)
(Intercept) 189.575  1 245.7453 2.287e-08 ***
Strain        0.535  1   0.6939    0.4243
Residuals     7.714 10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

Table B.29. Differentially expressed genes (p. adj < 0.05) from the comparative transcriptomic analysis of PXO61 variants PXO61-0 and PXO61-4 (passed through IRBB4) performed using the R (R Core Team, 2021) package DESeq2 (Love et al., 2014).

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene3860	PXO_02250	42090.68695	-3.92095	0.275007	-14.2576	4.02E-46	5.79E-43		DNA transport competence protein
gene209	PXO_03644	6408.294498	-3.12707	0.171595	-18.2236	3.36E-74	1.39E-70	*vgrG*	Rhs element Vgr protein
gene2274		102.1720289	-2.2665	0.426412	-5.31528	1.06E-07	2.94E-05		hypothetical protein
gene4138	PXO_02464	2596.861521	-2.06685	0.144993	-14.2548	4.19E-46	5.79E-43	*vgrG*	Rhs element Vgr protein
gene2583	PXO_01075	82.21851506	-2.04492	0.557508	-3.66796	0.000244	0.012868		energy transducer TonB
gene3760		67.68823846	-1.72556	0.53735	-3.21125	0.001322	0.037539		sX9 sRNA
gene2779	PXO_06244	1063.272502	-1.61589	0.393314	-4.1084	3.98E-05	0.003524		energy transducer TonB
gene1635	PXO_00175	202.629646	-1.60336	0.326478	-4.9111	9.06E-07	0.000198		Uncharacterized protein
gene2002	PXO_00313	8748.461237	-1.59224	0.363138	-4.38468	1.16E-05	0.001376	*htpG*	Chaperone protein HtpG (Heat shock protein HtpG) (High temperature protein G)
gene1384	PXO_04806	30446.98826	-1.47854	0.373329	-3.96042	7.48E-05	0.005642		Outer membrane protein
gene4219	PXO_02539	4209.255428	-1.46467	0.377926	-3.87555	0.000106	0.007146		Cell Wall Hydrolase superfamily
gene3831	PXO_02158	6318.204923	-1.40138	0.22005	-6.36845	1.91E-10	1.32E-07	*clpB*	Chaperone protein ClpB
gene1911	PXO_00404	3892.34686	-1.36179	0.284946	-4.77912	1.76E-06	0.000348	*htpX*	Protease HtpX (EC 3.4.24.-) (Heat shock protein HtpX)
gene2858	PXO_01185	35561.4851	-1.36154	0.252555	-5.39106	7.00E-08	2.23E-05	*dnaK*	Chaperone protein DnaK (HSP70) (Heat shock 70 kDa protein) (Heat shock protein 70)
gene799	PXO_04172	11036.0871	-1.30555	0.295753	-4.41433	1.01E-05	0.001313		Uncharacterized protein
gene487	PXO_03704	43974.46972	-1.30304	0.335932	-3.87889	0.000105	0.007146	*groS*	10 kDa chaperonin (GroES protein) (Protein Cpn10)
gene489	PXO_03702	48486.02895	-1.29724	0.323688	-4.0077	6.13E-05	0.00489	*xopX*	XopX effector protein
gene488	PXO_03703	132604.9634	-1.28576	0.324013	-3.96823	7.24E-05	0.005642	*groL*	60 kDa chaperonin (GroEL protein) (Protein Cpn60)
gene1141		349.2602227	-1.283	0.286238	-4.48227	7.39E-06	0.000988		hypothetical protein
gene4137	PXO_02463	2360.338048	-1.25717	0.163594	-7.68469	1.53E-14	1.59E-11		Uncharacterized protein

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene2859	PXO_01184	1813.11522	-1.24986	0.311802	-4.00849	6.11E-05	0.00489	*grpE*	Protein GrpE (HSP-70 cofactor)
gene1140	PXO_04598	2938.753244	-1.1781	0.222746	-5.28899	1.23E-07	3.19E-05	*msrA*	Peptide methionine sulfoxide reductase MsrA (Protein-methionine-S-oxide reductase) (EC 1.8.4.11) (Peptide-methionine (S)-S-oxide reductase) (Peptide Met(O) reductase)
gene4313	PXO_02640	10623.1399	-1.16733	0.254061	-4.59469	4.33E-06	0.000666		Low molecular weight heat shock protein
gene68	PXO_03414	1115.888103	-1.12047	0.266678	-4.20157	2.65E-05	0.002599	*hpa3*	Hpa3
gene847	PXO_04280	2690.121202	-1.08829	0.228324	-4.7664	1.88E-06	0.000354	*hslV*	ATP-dependent protease HslV
gene4893	PXO_03200	714.2371222	-1.0682	0.254466	-4.19782	2.69E-05	0.002599	*mutM*	Formamidopyrimidine-DNA glycosylase (Fapy-DNA glycosylase) (EC 3.2.2.23) (DNA-(apurinic or apyrimidinic site) lyase MutM) (AP lyase MutM) (EC 4.2.99.18)
gene92		1262.311879	-1.06733	0.259829	-4.10782	3.99E-05	0.003524		lytic transglycosylase
gene802	PXO_04169	927.1005984	-1.06418	0.295403	-3.60245	0.000315	0.015201		Uncharacterized protein
gene72	PXO_03411	5857.831464	-1.06384	0.275843	-3.85671	0.000115	0.007392	*hrpE*	HrpE type III pilin
gene1481	PXO_04858	1374.123663	-1.05747	0.210211	-5.03049	4.89E-07	0.000113		EF hand domain protein
gene1147	PXO_04592	3045.558563	-1.0532	0.269368	-3.90989	9.23E-05	0.006602		Regulator of nucleoside diphosphate kinase
gene3759	PXO_01983	652.1633741	-1.04422	0.286395	-3.64609	0.000266	0.013802		Intracellular protease I
gene91	PXO_03392	22350.72266	-1.03499	0.289237	-3.57834	0.000346	0.016295	*hpa1*	Hpa1
gene3885	PXO_02226	3942.791067	-1.03397	0.307687	-3.36047	0.000778	0.028811	*yfiA*	Ribosomal subunit interface protein
gene4510	PXO_03097	30130.58897	-1.03382	0.251663	-4.10794	3.99E-05	0.003524		Outer membrane protein
gene3666	PXO_02078	617.1288229	-1.01587	0.271612	-3.74016	0.000184	0.010592		Serine/threonine kinase
gene891	PXO_04299	2640.280768	-1.01287	0.282022	-3.59145	0.000329	0.015675		Surface antigen protein
gene4852	PXO_02760	12723.81643	-1.01019	0.291204	-3.469	0.000522	0.021883		XopN effector
gene3738	PXO_02005	10182.52863	-0.99769	0.167786	-5.94617	2.74E-09	1.26E-06		Surface antigen protein
gene4134	PXO_02459	1958.896794	-0.99467	0.154489	-6.43845	1.21E-10	1.00E-07		Uncharacterized protein
gene848	PXO_04279	6902.33246	-0.99365	0.220079	-4.51497	6.33E-06	0.000875	*hslU*	ATP-dependent protease ATPase subunit HslU (Unfoldase HslU)

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene66	PXO_03417	21692.88729	-0.9915	0.26662	-3.71876	0.0002	0.011373	*hrpF*	HrpF/NolX/HrpK
gene4156	PXO_02475	17796.47659	-0.98932	0.179211	-5.52043	3.38E-08	1.17E-05	*lon*	Lon protease (EC 3.4.21.53) (ATP-dependent protease La)
gene1379	PXO_04811	7466.250988	-0.97821	0.247267	-3.95609	7.62E-05	0.005642		Peptidoglycan-associated outer membrane lipoprotein
gene82	PXO_03401	7516.63849	-0.97763	0.271033	-3.60706	0.00031	0.015201		Type III secretion protein HrpB1/HrpK
gene3526	PXO_01885	2132.896885	-0.97673	0.233425	-4.18432	2.86E-05	0.002696		Outer membrane protein
gene1799	PXO_00505	14127.18475	-0.97489	0.269134	-3.62231	0.000292	0.014767	*tal3b*	Tal3b, TAL effector AvrBs3/PthA family
gene132	PXO_03355	1903.57349	-0.95996	0.272498	-3.52283	0.000427	0.018444		Uncharacterized protein
gene74	PXO_03409	8564.538725	-0.9527	0.224639	-4.24105	2.22E-05	0.002366		HrpD5
gene83	PXO_03400	7531.345364	-0.95138	0.275796	-3.44956	0.000561	0.023055		Type III secretion protein HrpB2
gene3078	PXO_01668	1509.26199	-0.94723	0.278073	-3.40641	0.000658	0.025511	*hslO*	Chaperonin HslO
gene69	PXO_03413	3953.341321	-0.94427	0.290144	-3.25449	0.001136	0.036237	*xopF1Xoo*	XopF1 effector
gene4538	PXO_03070	1059.594302	-0.92416	0.286206	-3.22902	0.001242	0.036722	*msrB*	Peptide methionine sulfoxide reductase MsrB (EC 1.8.4.12) (Peptide-methionine (R)-S-oxide reductase)
gene213	PXO_03639	2648.543256	-0.91976	0.161276	-5.70303	1.18E-08	4.88E-06		Uncharacterized protein
gene3758	PXO_01984	2636.043318	-0.91178	0.197381	-4.61936	3.85E-06	0.000614		Uncharacterized protein
gene79		2616.831912	-0.91125	0.274047	-3.32516	0.000884	0.031322		type III secretion protein HpaP
gene2857	PXO_01186	5085.321679	-0.90398	0.16836	-5.36929	7.90E-08	2.34E-05	*dnaJ*	Chaperone protein DnaJ
gene443	PXO_03755	18550.0483	-0.90003	0.249813	-3.60284	0.000315	0.015201		5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase
gene84	PXO_03399	4443.864903	-0.89816	0.280983	-3.1965	0.001391	0.038395	*yscJ*	Lipoprotein
gene3830	PXO_02159	1446.419555	-0.89245	0.276198	-3.23118	0.001233	0.036722		Uncharacterized protein
gene3777	PXO_01964	506.2772951	-0.89116	0.271452	-3.28295	0.001027	0.034081	*metY*	O-acetylhomoserine sulfhydrylase
gene4106	PXO_02433	3876.459257	-0.88088	0.183944	-4.78883	1.68E-06	0.000348	*def*	Peptide deformylase (PDF) (EC 3.5.1.88) (Polypeptide deformylase)
gene3664		2508.51566	-0.87947	0.247005	-3.56055	0.00037	0.016865		hypothetical protein

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene3083	PXO_01663	28680.01261	-0.86686	0.271314	-3.19505	0.001398	0.038395		Acyl-coenzyme A dehydrogenase (EC 1.3.8.7) (EC 1.3.8.8)
gene954	PXO_04425	1437.03467	-0.86463	0.242544	-3.56485	0.000364	0.016775		Glutamate--cysteine ligase (EC 6.3.2.2)
gene967	PXO_04410	346.6564669	-0.85925	0.266073	-3.22939	0.001241	0.036722		Lipoprotein, putative
gene1806		155.2154485	-0.85704	0.263069	-3.25785	0.001123	0.036237		hypothetical protein
gene3082	PXO_01664	3528.009706	-0.84635	0.267597	-3.1628	0.001563	0.041808		Hydrolase, alpha/beta fold family, putative
gene80	PXO_03403	2489.508251	-0.84329	0.22737	-3.70888	0.000208	0.011666	*hrcV*	HrcV
gene3636	PXO_02107	782.7103147	-0.83903	0.236489	-3.54784	0.000388	0.01732		Uncharacterized protein
gene3129	PXO_01620	1334.411385	-0.8383	0.266032	-3.15113	0.001626	0.043235		Leucin rich protein
gene2985	PXO_01763	13780.01549	-0.83765	0.21852	-3.83329	0.000126	0.007711		GTN Reductase
gene2660	PXO_06125	4279.114639	-0.80983	0.207645	-3.90006	9.62E-05	0.006759		hypothetical protein
gene4915	PXO_03298	8789.391532	-0.80925	0.210766	-3.83955	0.000123	0.007629		Endoglucanase
gene4267	PXO_02592	1787.682428	-0.80598	0.218667	-3.68588	0.000228	0.012436		Uncharacterized protein
gene2044	PXO_00536	1265.675527	-0.80534	0.234529	-3.43387	0.000595	0.023957	*phoU*	Phosphate-specific transport system accessory protein PhoU
gene2335	PXO_00794	432.9551732	-0.80413	0.23576	-3.41079	0.000648	0.025511		Uncharacterized protein
gene343	PXO_03859	13917.78855	-0.79842	0.243368	-3.28069	0.001036	0.034082		Uncharacterized protein
gene210	PXO_03643	1412.60425	-0.79667	0.173847	-4.58256	4.59E-06	0.00068		Uncharacterized protein
gene1898	PXO_00416	3807.90971	-0.79567	0.188085	-4.23037	2.33E-05	0.002369		Virulence protein
gene3222	PXO_01519	3622.702402	-0.7915	0.186609	-4.24152	2.22E-05	0.002366		Uncharacterized protein
gene1351	PXO_04739	872.8031351	-0.79029	0.231009	-3.42105	0.000624	0.024874		Cysteine protease
gene3621	PXO_02120	5159.340376	-0.78836	0.244207	-3.22824	0.001246	0.036722		O-acetyl-ADP-ribose deacetylase
gene4105		945.6153058	-0.78427	0.221493	-3.54084	0.000399	0.017596		sX9 sRNA
gene4113	PXO_02439	15656.32354	-0.77318	0.200016	-3.86561	0.000111	0.007294		Outer membrane protein Slp
gene4740		1818.822897	-0.77274	0.240971	-3.20676	0.001342	0.03787		amidophosphoribosyltransferase
gene2689	PXO_06156	2406.711828	-0.77052	0.200332	-3.84619	0.00012	0.007538		flagellar hook protein

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene60	PXO_03424	7733.813616	-0.7668	0.239948	-3.1957	0.001395	0.038395		Periplasmic serine endoprotease DegP-like (EC 3.4.21.107)
gene2812	PXO_01233	3272.717335	-0.7598	0.235768	-3.22264	0.00127	0.036722	*metH*	Methionine synthase
gene3117	PXO_01629	794.2260377	-0.75384	0.160137	-4.70747	2.51E-06	0.000452		Membrane-fusion protein
gene2981	PXO_01766	31677.70238	-0.74834	0.160153	-4.67264	2.97E-06	0.000514		DUF2383 domain-containing protein
gene3895	PXO_02215	4986.80032	-0.73785	0.20656	-3.5721	0.000354	0.016501	*cydB*	Cytochrome D ubiquinol oxidase, subunit II
gene3896	PXO_02214	35563.06666	-0.72307	0.225841	-3.20171	0.001366	0.03828		Cytochrome D ubiquinol oxidase subunit I
gene4778	PXO_02836	8652.431351	-0.71505	0.221686	-3.22549	0.001258	0.036722	*raxR2*	Two-component system regulatory protein (Response regulator) required for AvrXa21 activity R2 (RaxR2)
gene2811	PXO_01234	2835.714757	-0.71265	0.221636	-3.21539	0.001303	0.037256		methyltransferase domain-containing protein
gene1960	PXO_00353	4368.736998	-0.71211	0.22452	-3.17169	0.001516	0.040811	*argF'*	N-acetylornithine carbamoyltransferase (EC 2.1.3.9) (N-acetyl-L-ornithine transcarbamylase) (AOTCase) (Acetylornithine transcarbamylase)
gene4533	PXO_03075	3841.854663	-0.70783	0.166055	-4.2626	2.02E-05	0.002265		Uncharacterized protein
gene2464	PXO_01041	1916.500794	-0.70735	0.212241	-3.33276	0.00086	0.030759		hypothetical protein
gene1926	PXO_00390	1046.281119	-0.70385	0.227369	-3.09563	0.001964	0.049967		Ribonuclease
gene1073	PXO_04661	827.1490543	-0.69596	0.172998	-4.02295	5.75E-05	0.004864		ABC transporter ATP-binding protein
gene2050	PXO_00788	1786.997104	-0.68681	0.206933	-3.31898	0.000903	0.031488		Polyphosphate-selective porin O
gene4539	PXO_03069	1358.642812	-0.68355	0.192478	-3.55129	0.000383	0.01728		Putative secreted protein
gene2047	PXO_00542	1424.127987	-0.65789	0.194514	-3.38223	0.000719	0.027316	*pstC*	Phosphate transport system permease protein
gene2991	PXO_01757	1364.110879	-0.64661	0.206947	-3.12453	0.001781	0.04704	*slyd*	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)
gene1326	PXO_04762	1602.345099	-0.64439	0.160447	-4.01619	5.91E-05	0.00489	*nfi*	Endonuclease V (EC 3.1.21.7) (Deoxyinosine 3'endonuclease) (Deoxyribonuclease V) (DNase V)
gene4279	PXO_02605	3408.141511	-0.64281	0.207367	-3.09985	0.001936	0.049564	*ilvC*	Ketol-acid reductoisomerase (NADP(+)) (KARI) (EC 1.1.1.86) (Acetohydroxy-acid isomeroreductase) (AHIR) (Alpha-keto-beta-hydroxylacyl reductoisomerase) (Ketol-acid reductoisomerase type 1) (Ketol-acid reductoisomerase type I)
gene2986		608.1240955	-0.63931	0.194018	-3.29508	0.000984	0.033446		hypothetical protein

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene3468	PXO_01268	10379.2635	-0.61564	0.197805	-3.11235	0.001856	0.048106	*trpF*	N-(5'-phosphoribosyl)anthranilate isomerase (PRAI) (EC 5.3.1.24)
gene1434	PXO_04844	582.2256257	-0.60364	0.187137	-3.22563	0.001257	0.036722		IS630 family transposase
gene3790	PXO_01951	3724.770996	-0.57788	0.185871	-3.10905	0.001877	0.048345	*hrpG*	HrpG
gene481	PXO_03711	15720.2943	-0.57658	0.163895	-3.51796	0.000435	0.018592	*rpoH*	RNA polymerase sigma factor RpoH (RNA polymerase sigma-32 factor)
gene2165	PXO_00663	6776.494718	-0.50253	0.154106	-3.26092	0.001111	0.036237		Uncharacterized protein
gene689	PXO_04088	13926.68731	0.438694	0.135722	3.232294	0.001228	0.036722		TonB-dependent outer membrane Receptor
gene1646	PXO_00165	9782.174254	0.44922	0.138592	3.241315	0.00119	0.036722		5'-nucleotidase SurE (EC 3.1.3.5) (Nucleoside 5'-monophosphate phosphohydrolase)
gene1198	PXO_04540	5012.682161	0.472579	0.145358	3.251139	0.001149	0.036387	*pth*	Peptidyl-tRNA hydrolase (PTH) (EC 3.1.1.29)
gene2913	PXO_01126	4152.475903	0.472772	0.138769	3.406895	0.000657	0.025511	*cdsA*	Phosphatidate cytidylyltransferase (EC 2.7.7.41)
gene3032	PXO_01718	2467.725716	0.480738	0.154362	3.114348	0.001844	0.048082		GTPase Era
gene3825	PXO_02166	4979.848724	0.493848	0.153128	3.225062	0.001259	0.036722	*btuB*	TonB-dependent outer membrane Receptor
gene4375	PXO_02685	2425.585319	0.498159	0.156454	3.184058	0.001452	0.039622		General secretion pathway protein N
gene863	PXO_04261	2957.440509	0.505923	0.149746	3.378532	0.000729	0.027316		Cell shape protein MreC (Cell shape-determining protein MreC)
gene2814	PXO_01231	10677.46164	0.507792	0.150125	3.382454	0.000718	0.027316		N-acylglucosamine 2-epimerase
gene2365	PXO_00907	2671.220875	0.508825	0.153417	3.316613	0.000911	0.031488		UPF0753 protein PXO_00907
gene2920	PXO_01118	1490.915568	0.508992	0.157997	3.221521	0.001275	0.036722	*fabZ*	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ (EC 4.2.1.59) ((3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase) ((3R)-hydroxymyristoyl-ACP dehydrase) (Beta-hydroxyacyl-ACP dehydratase)
gene4343	PXO_02712	3761.188806	0.51486	0.152432	3.377628	0.000731	0.027316	*pabB*	Para-aminobenzoate synthase component I
gene3840	PXO_02150	14309.53624	0.515239	0.165407	3.114978	0.00184	0.048082	*rplS*	50S ribosomal protein L19
gene2366	PXO_00906	11371.66358	0.520911	0.147661	3.52774	0.000419	0.018296	*guaA*	GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (GMP synthetase) (Glutamine amidotransferase)
gene3505	PXO_01909	2690.175988	0.524733	0.161579	3.247529	0.001164	0.036573		Transport protein

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene2602	PXO_06063	1313.436432	0.532156	0.167694	3.173372	0.001507	0.040811		heme exporter protein CcmB
gene3097	PXO_01649	4406.009156	0.534432	0.140069	3.815481	0.000136	0.008169	*cysQ*	3'(2'),5'-bisphosphate nucleotidase CysQ (EC 3.1.3.7) (3'(2'),5-bisphosphonucleoside 3'(2')-phosphohydrolase) (3'-phosphoadenosine 5'-phosphate phosphatase) (PAP phosphatase)
gene4612	PXO_02996	6081.372967	0.53956	0.146231	3.689766	0.000224	0.012411		Putative fis-like DNA-binding protein
gene1849	PXO_00459	5228.531152	0.547414	0.163793	3.342112	0.000831	0.030245	*oprN*	Outer membrane protein OprN
gene1848	PXO_00460	4692.007264	0.547535	0.1653	3.31238	0.000925	0.031704		Transport protein
gene3232	PXO_01511	2710.674523	0.548063	0.149171	3.674051	0.000239	0.012858		VI polysaccharide biosynthesis protein VipA/tviB
gene4335	PXO_02718	1569.438973	0.569906	0.176366	3.231385	0.001232	0.036722	*prpB*	2-methylisocitrate lyase (2-MIC) (MICL) (EC 4.1.3.30) ((2R,3S)-2-methylisocitrate lyase)
gene4514	PXO_03093	2409.226963	0.581343	0.16774	3.465747	0.000529	0.021928	*metN*	Methionine import ATP-binding protein MetN (EC 7.4.2.11)
gene3230	PXO_01512	1090.729131	0.589007	0.162329	3.62847	0.000285	0.014597		Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)
gene2421	PXO_00849	1293.687495	0.589387	0.168528	3.497269	0.00047	0.019891		amino acid permease
gene112	PXO_03370	1758.158001	0.605153	0.183983	3.289183	0.001005	0.033877		NAD-dependent dehydratase
gene2669	PXO_06135	3140.484179	0.607384	0.182256	3.332594	0.00086	0.030759		translation initiation factor IF-1
gene2293	PXO_00838	1991.59593	0.610565	0.187493	3.256477	0.001128	0.036237	*hisF*	IGP synthase cyclase subunit (EC 4.3.2.10)
gene2420	PXO_00850	2343.351376	0.631072	0.190238	3.317278	0.000909	0.031488		amino acid permease
gene4088	PXO_02417	3582.200546	0.640036	0.194789	3.285788	0.001017	0.034011		ThiJ/PfpI family protein
gene4685	PXO_02928	949.7120679	0.653789	0.166304	3.931285	8.45E-05	0.006147		Mitomycin resistance protein
gene3429	PXO_01305	1776.036079	0.655086	0.169944	3.854709	0.000116	0.007392	*truB*	tRNA pseudouridine synthase B (EC 5.4.99.25) (tRNA pseudouridine(55) synthase) (Psi55 synthase) (tRNA pseudouridylate synthase) (tRNA-uridine isomerase)
gene2296	PXO_00835	4049.823119	0.66646	0.155381	4.289191	1.79E-05	0.002066	*hisB*	Histidine biosynthesis bifunctional protein HisB [Includes: Histidinol-phosphatase (EC 3.1.3.15); Imidazoleglycerol-phosphate dehydratase (IGPD) (EC 4.2.1.19)]

Table B.29. Continued

ID	Locus tag	Base mean	log2Fold Change	lfcSE	stat	pvalue	padj	Gene names	Protein names
gene2617	PXO_06078	4878.358215	0.667236	0.199082	3.351565	0.000804	0.02949		amino acid permease
gene2295	PXO_00836	1723.567684	0.668831	0.176022	3.799709	0.000145	0.008582	*hisH*	Imidazole glycerol phosphate synthase subunit HisH (EC 4.3.2.10) (IGP synthase glutaminase subunit) (EC 3.5.1.2) (IGP synthase subunit HisH) (IGP synthase subunit HisH) (IGPS subunit HisH)
gene3504	PXO_01910	2574.628466	0.679861	0.154409	4.402992	1.07E-05	0.001342		D-amino acid dehydrogenase subunit
gene1200	PXO_rna4	1384.263739	0.6824	0.198113	3.444494	0.000572	0.023261		tRNA-Tyr
gene1057	PXO_04679	1852.674486	0.697849	0.172861	4.037056	5.41E-05	0.004676		Nucleoside transporter
gene4570	PXO_03036	10893.42066	0.733877	0.161318	4.54927	5.38E-06	0.00077		Amino acid transporter
gene4513	PXO_03094	1696.108458	0.738016	0.174492	4.229503	2.34E-05	0.002369	*metI*	ABC transporter permease
gene3198	PXO_01544	6319.729395	0.761635	0.173336	4.393972	1.11E-05	0.001358		Membrane protein, putative
gene4085	PXO_02413	9626.197647	0.798637	0.17145	4.658131	3.19E-06	0.000529		Sugar transporter
gene1815	PXO_00488	5581.443057	0.816113	0.216698	3.766134	0.000166	0.009684		Phosphofructokinase
gene1851		418.6585577	0.839689	0.212064	3.959609	7.51E-05	0.005642		hypothetical protein
gene3108	PXO_01638	3869.939651	0.848903	0.231479	3.667294	0.000245	0.012868		Glucan 1,4-beta-glucosidase
gene4084		399.1660373	0.880225	0.227184	3.87451	0.000107	0.007146		sX9 sRNA
gene3044	PXO_01705	2645.405578	0.994835	0.176137	5.648058	1.62E-08	6.12E-06		Enoyl-CoA hydratase
gene3046	PXO_01703	3701.12629	1.103905	0.305632	3.611876	0.000304	0.015188	*mmsA*	Methylmalonate-semialdehyde dehydrogenase
gene3045	PXO_01704	3198.012738	1.107584	0.210179	5.26971	1.37E-07	3.33E-05		Acyl-CoA dehydrogenase family member 8
gene3042	PXO_01707	2005.156013	1.128623	0.186238	6.060112	1.36E-09	7.05E-07	*mmsB*	3-hydroxyisobutyrate dehydrogenase (HIBADH) (EC 1.1.1.31)
gene3043	PXO_01706	4476.564712	1.236415	0.196729	6.284855	3.28E-10	1.94E-07		Enoyl-CoA hydratase

Appendix C

Table C.1. List of *Xanthomonas oryzae* pv. *oryzae* genomes included in the population structure analysis and selection of single nucleotide polymorphisms linked to putative *Xanthomonas oryzae* pv. *oryzae* populations for primer development.

Source	Strain	Putative population
Philippines	PXO318	popID 8
Philippines	PXO339	popID 8
Philippines	PXO340	popID 8
Philippines	PXO344	popID 8
Philippines	PXO345	popID 8
Philippines	PXO346	popID 8
Philippines	PXO347	popID 8
Philippines	PXO349	popID 8
Philippines	PXO350	popID 8
Philippines	PXO367	popID 8
Philippines	PXO426	popID 8
Philippines	PXO442	popID 8
Philippines	PXO444	popID 8
Philippines	PXO524	popID 8
Philippines	PXO552	popID 8
Philippines	PXO554	popID 8
Philippines	PXO591	popID 8
Philippines	PXO592	popID 8
Philippines	PXO593	popID 8
Philippines	PXO594	popID 8
Philippines	PXO603	popID 8
Indonesia	S216	popID 8
Indonesia	TRIV	popID 8
Indonesia	TRVIII	popID 8
Indonesia	Xoo12-116-2	popID 8
Indonesia	Xoo12-190	popID 8
India	DXO-331	popID 8
Philippines	PXO282	popID 8
Philippines	PXO602	popID 8
Indonesia	TRIII	popID 8
Indonesia	Xoo93299	popID 8
Philippines	PXO113	popID 2
Philippines	PXO125	popID 2
Philippines	PXO129	popID 2
Philippines	PXO69	popID 2
Philippines	PXO70	popID 2
Philippines	PXO71	popID 2
Philippines	PXO87	popID 11
Philippines	PXO88	popID 11
Indonesia	Xoo11-021	popID 11
Indonesia	Xoo12-183	popID 11
India	DXO-206	popID 9
India	IXO159	popID 9
India	IXO603	popID 9

Table C.1. Continued

Source	Strain	Putative population
India	IXO608	popID 9
India	IXO90	popID 9
India	IXO93	popID 9
India	IXO97	popID 9
Japan	MAFF_311018	popID 9
Philippines	PXO1	popID 9
Philippines	PXO142	popID 9
Philippines	PXO151	popID 9
Philippines	PXO164	popID 9
Philippines	PXO176	popID 9
Philippines	PXO178	popID 9
Philippines	PXO18	popID 9
Philippines	PXO20	popID 9
Philippines	PXO224	popID 9
Philippines	PXO25	popID 9
Philippines	PXO404	popID 9
Philippines	PXO414	popID 9
Philippines	PXO68	popID 9
Taiwan	XF89b	popID 9
India	DXO-216	popID 3
India	IXO597	popID 3
India	IXO599	popID 3
Indonesia	Xoo12-011	popID 3
Indonesia	Xoo12-195	popID 3
Korea	KACC_10331	popID 3
India	DXO-015	popID 6
India	DXO-200	popID 6
India	DXO-242	popID 6
India	DXO-246	popID 6
India	IXO278	popID 6
India	IXO411	popID 6
India	IXO675	popID 6
India	BXO6	popID 4
India	BXO8	popID 4
India	DXO-174	popID 4
India	DXO-233	popID 4
India	IXO1221	popID 4
India	IXO651	popID 4
India	IXO685	popID 4
India	IXO725	popID 4
India	IXO884	popID 4
India	IXO89	popID 4
India	IXO1088	popID 1
India	IXO1104	popID 1
Philippines	PXO114	popID 5
Philippines	PXO117	popID 5
Philippines	PXO118	popID 5
Philippines	PXO121	popID 5
Philippines	PXO123	popID 5
Philippines	PXO124	popID 5
Philippines	PXO128	popID 5

Table C.1. Continued

Source	Strain	Putative population
Philippines	XooPXO99A	popID 5
Philippines	PXO108	popID 14
Philippines	PXO161	popID 14
Philippines	PXO236	popID 14
Philippines	PXO237	popID 14
Philippines	PXO45	popID 14
Philippines	PXO101	popID 13
Philippines	PXO130	popID 13
Philippines	PXO135	popID 13
Philippines	PXO144	popID 13
Philippines	PXO145	popID 13
Philippines	PXO150	popID 13
Philippines	PXO154	popID 13
Philippines	PXO155	popID 13
Philippines	PXO156	popID 13
Philippines	PXO201	popID 13
Philippines	PXO205	popID 13
Philippines	PXO280	popID 13
Philippines	PXO325	popID 13
Philippines	PXO338	popID 13
Philippines	PXO65	popID 13
Philippines	PXO86	popID 13
Philippines	PXO139	popID 12
Philippines	PXO141	popID 12
Philippines	PXO153	popID 12
Philippines	PXO168	popID 12
Philippines	PXO171	popID 12
Philippines	PXO174	popID 12
Philippines	PXO185	popID 12
Philippines	PXO79	popID 12
Philippines	PXO81	popID 12
Philippines	PXO83	popID 12
India	IXO141	popID 10
India	IXO222	popID 10
India	IXO365	popID 10
India	IXO390	popID 10
India	IXO493	popID 10
India	IXO620	popID 10
India	IXO621	popID 10
India	IXO630	popID 10
India	IXO644	popID 10
India	IXO645	popID 10
India	IXO704	popID 10
India	IXO792	popID 10
India	IXO842	popID 10
India	IXO141-1	popID 10
India	BXO1	popID 7
India	BXO2	popID 7
India	BXO25	popID 7
India	BXO33	popID 7
India	BXO34	popID 7

Table C.1. Continued

Source	Strain	Putative population
India	BXO407	popID 7
India	BXO416	popID 7
India	BXO432	popID 7
India	BXO439	popID 7
India	BXO447	popID 7
India	BXO454	popID 7
India	BXO471	popID 7
India	BXO512	popID 7
India	BXO554	popID 7
India	BXO557	popID 7
India	BXO558	popID 7
India	BXO559	popID 7
India	BXO571	popID 7
India	BXO582	popID 7
India	BXO589	popID 7
India	BXO590	popID 7
India	DXO-012	popID 7
India	DXO-027	popID 7
India	DXO-044	popID 7
India	DXO-050	popID 7
India	DXO-052	popID 7
India	DXO-089	popID 7
India	DXO-091	popID 7
India	DXO-116	popID 7
India	DXO-122	popID 7
India	DXO-129	popID 7
India	DXO-131	popID 7
India	DXO-133	popID 7
India	DXO-150	popID 7
India	DXO-165	popID 7
India	DXO-170	popID 7
India	DXO-203	popID 7
India	DXO-248	popID 7
India	DXO-369	popID 7
India	DXO-397	popID 7
India	IXO134	popID 7
India	IXO151	popID 7
India	IXO189	popID 7
India	IXO220	popID 7
India	IXO221	popID 7
India	IXO35	popID 7
India	IXO367	popID 7
India	IXO414	popID 7
India	IXO627	popID 7
India	IXO639	popID 7
India	IXO74	popID 7
India	IXO812	popID 7
India	IXO92	popID 7
India	IXO98	popID 7
India	IXO99	popID 7
Thailand	SK2	popID 7

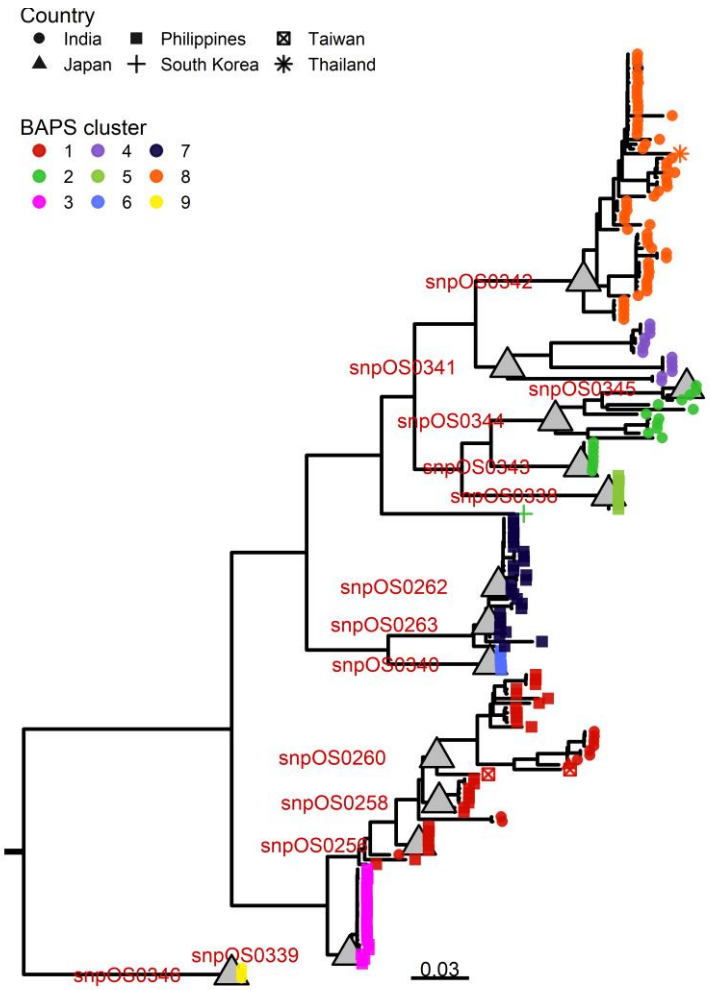


Figure C.1 Phylogenetic relationship and genetic structure of 197 *Xanthomonas oryzae* pv. *oryzae* strains collected across Asian countries. A maximum likelihood tree was build using recombination free core genome alignment from 1000 bootstrap re-sampling. Colored tips designate the country of origin. The right-side panel shows the inferred population demographic of the *Xanthomonas oryzae* pv. *oryzae* strains utilized in heirBAPS (Cheng et al. 2013).

Table C.2. List of samples used to determine the diagnostic specificity and sensitivity of selected *Xanthomonas oryzae* pv. *oryzae* putative genetic group-specific primers included in the pilot test of a field monitoring strategy for *Xanthomonas oryzae* pv. *oryzae* using simulated populations (popID).

Strain	Sample matrix	Organism	Expected group	Expected SNP	Detected group	Detected SNP
BLS256	DNA	*Xoc*	NA	negative	no ID	no ID
BLS500	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2376	DNA	*Xoc*	NA	negative	no ID	no ID
BLS524	DNA	*Xoc*	NA	negative	popID 8	snpOS0339
PXO364	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO61	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO86	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO71	DNA	*Xoo*	popID 2	snpOS0256	popID 2	snpOS0256
PXO112	DNA	*Xoo*	popID 14	snpOS0340	popID 14	snpOS0340
PXO79	DNA	*Xoo*	popID 12	snpOS0263	popID 12	snpOS0263
PXO99	DNA	*Xoo*	popID 5	snpOS0338	popID 5	snpOS0338
PXO145	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO349	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO340	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO339	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO280	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO363	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO347	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO341	DNA	*Xoo*	popID 11	snpOS0258	popID 11	snpOS0258
BLS1873	DNA	*Xoc*	NA	negative	no ID	snpOS0260
BLS1894	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1897	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1913	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1991	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2136	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2052	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2145	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2300	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2389	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2427	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2459	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2468	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2471	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1811	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1847	DNA	*Xoc*	NA	negative	no ID	no ID

Table C.2. Continued

Strain	Sample matrix	Organism	Expected group	Expected SNP	Detected group	Detected SNP
BLS1850	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1854	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2082	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2027	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1890	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1885	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1511	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1656	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1500	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1619	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1468	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1880	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1750	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2312	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1867	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1603	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1616	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1618	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1864	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1503	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1477	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1478	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1749	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1620	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1725	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1866	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1720	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1688	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1765	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1851	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1755	DNA	*Xoc*	NA	negative	no ID	no ID
PXO349	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO347	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO363	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO341	DNA	*Xoo*	popID 11	snpOS0258	popID 11	snpOS0258
PXO71	DNA	*Xoo*	popID 2	snpOS0256	popID 2	snpOS0256
PXO112	DNA	*Xoo*	popID 14	snpOS0340	popID 14	snpOS0340
PXO99	DNA	*Xoo*	popID 5	snpOS0338	popID 5	snpOS0338
PXO145	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO280	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO339	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339

Table C.2. Continued

Strain	Sample matrix	Organism	Expected group	Expected SNP	Detected group	Detected SNP
PXO61	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO86	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO79	DNA	*Xoo*	popID 12	snpOS0263	popID 12	snpOS0263
PXO340	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO349	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO347	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO341	DNA	*Xoo*	popID 11	snpOS0258	popID 11	snpOS0258
PXO71	DNA	*Xoo*	popID 2	snpOS0256	popID 2	snpOS0256
PXO112	DNA	*Xoo*	popID 14	snpOS0340	popID 14	snpOS0340
PXO145	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO280	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO339	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
BLS1873	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1894	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1897	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1913	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1991	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2136	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2052	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2145	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2300	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2389	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2427	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2459	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2468	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2471	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1811	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1847	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1850	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1854	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2082	DNA	*Xoc*	NA	negative	no ID	no ID
BLS2027	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1890	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1885	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1511	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1656	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1500	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1619	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1468	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1880	DNA	*Xoc*	NA	negative	no ID	no ID
BLS1750	DNA	*Xoc*	NA	negative	no ID	no ID

Table C.2. Continued

Strain	Sample matrix	Organism	Expected group	Expected SNP	Detected group	Detected SNP
PXO61	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO86	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO71	DNA	*Xoo*	popID 2	snpOS0256	popID 2	snpOS0256
PXO112	DNA	*Xoo*	popID 14	snpOS0340	popID 14	snpOS0340
PXO79	DNA	*Xoo*	popID 12	snpOS0263	popID 12	snpOS0263
PXO99	DNA	*Xoo*	popID 5	snpOS0338	popID 5	snpOS0338
PXO145	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO349	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO340	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO339	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO280	DNA	*Xoo*	popID 13	snpOS0262	popID 13	snpOS0262
PXO363	DNA	*Xoo*	popID 9	snpOS0260	popID 9	snpOS0260
PXO347	DNA	*Xoo*	popID 8	snpOS0339	popID 8	snpOS0339
PXO341	DNA	*Xoo*	popID 11	snpOS0258	popID 11	snpOS0258

Table C.3 Confusion matrix generated using the R (R Core Team, 2021) package caret (Kuhn, 2008) using data from the analysis of 131 DNA controls by eight primers derived from single nucleotide polymorphisms (SNPs) delineating putative populations of *Xanthomonas oryzae* pv. *oryzae*. Columns designate the expected SNPs to be detected, and rows designate the actual detected SNPs. The category "no ID" is included which signifies samples which are expected to test negative for any of the genetic group-specific alleles.

Detected	Expected								
	no ID	snpOS0256	snpOS0258	snpOS0260	snpOS0262	snpOS0263	snpOS0338	snpOS0339	snpOS0340
no ID	78	0	0	0	0	0	0	0	0
snpOS0256	0	4	0	0	0	0	0	0	0
snpOS0258	0	0	4	0	0	0	0	0	0
snpOS0260	1	0	0	7	0	0	0	0	0
snpOS0262	0	0	0	0	11	0	0	0	0
snpOS0263	0	0	0	0	0	3	0	0	0
snpOS0338	0	0	0	0	0	0	3	0	0
snpOS0339	1	0	0	0	0	0	0	15	0
snpOS0340	0	0	0	0	0	0	0	0	4

Table C.4. Summary of the three sample types included in the pilot test of the field monitoring strategy developed for *Xanthomonas oryzae* pv. *oryzae.*

Sample type	Cambodia	India	Indonesia	Myanmar	Pakistan	Philippines	Vietnam	Grand Total
DNA		94	96		45	520	94	849
DNA[α]	80						44	124
Leaf				94				94
Grand Total	80	94	96	94	45	520	138	**1067**

[α]Extracted from leaf samples stored on Whatman® FTA® PlantSaver Cards using an in-house quick extraction procedure developed in this study.