The adaptation of the plant pathogen *Pseudomonas syringae* onto the novel host *Arabidopsis thaliana* through experimental evolution

by

Andrew Jamnik

A thesis submitted in conformity with the requirements for the degree of Master of Science

Department of Cell and Systems Biology
University of Toronto

©Copyright by Andrew Jamnik 2017
The adaptation of the plant pathogen *Pseudomonas syringae* onto the novel host *Arabidopsis thaliana* through experimental evolution

Andrew Jamnik
Master of Science
Department of Cell and Systems Biology
University of Toronto
2017

Abstract
The bacterial plant pathogen *Pseudomonas syringae* is a highly diverse species complex that can infect and cause disease on a wide range of hosts. Despite this broad host range, individual strains of *P. syringae* are highly host specific. In this thesis, we used *in planta* serial passaging to experimentally evolve twelve independent lineages of *P. syringae* pathovar phaseolicola (*Pph*) 1448A, a bean pathogen, on the novel host *Arabidopsis thaliana* (*Arabidopsis*), with the goal of characterizing the early adaptive steps towards host adaptation. Twelve hypermutating lineages of *Pph* 1448A were constructed for this experiment to increase the evolvability of lineages, and to efficiently study mutations associated with adaptation towards *Arabidopsis*. After eighty days of passaging on the novel host, lineage 7 has shown a significant gain in fitness on *Arabidopsis*. Population sequencing of this lineage may provide novel insights into the adaptive processes underlying host specificity in *P. syringae*. 
Acknowledgements

I would like to thank my supervisor Dr. David S. Guttman for giving me the opportunity to work in his great laboratory for the past two years, and for entrusting me with this wonderful project. I would also like to thank Dr. Darrell Desveaux and Dr. Alex Ensminger for taking the time to being on my supervisory committee, and for their invaluable feedback on this project. I also appreciate Dr. Nicole Mideo for taking the time to be an external examiner for my thesis defense, and for her help in the final steps to submitting my thesis.

Current and previous members of both the Guttman and Desveaux lab were always welcoming and made coming into work an enjoyable experience. I would especially like to thank Dr. Marcus Dillon for his collaboration in this project, and for all his support and guidance throughout. Special thanks to Dr. Pauline Wang, Jessica Schembri and Dr. Patrick Bastedo for training me in the lab, and for sharing their vast knowledge on many laboratory techniques. I would also like to acknowledge the support and guidance from the senior members of the lab including Julio Diaz Caballero, Dr. Shawn Clark, Dr. Mehdi Layeghifard, Julia Copeland and Andre Severino who helped keep me motivated especially during the lows of graduate school.

I would also like to acknowledge my family, especially my parents who have been their me from day one, and for whom I love dearly. Lastly, I would like to thank Hayley Nelles for her care and support throughout my life as both a graduate student and as a person.
# Table of contents

Acknowledgements ........................................................................................................................................ iii

Table of contents ........................................................................................................................................ iv

List of tables ............................................................................................................................................... vii

List of figures ............................................................................................................................................... viii

List of abbreviations ................................................................................................................................ xi

## Chapter 1: Introduction

  1.1 Pseudomonas syringae ....................................................................................................................... 1
  1.2 Plant innate immunity .......................................................................................................................... 2
  1.3 P. syringae and host specificity .......................................................................................................... 3
  1.4 P. syringae pv. phaseolicola and non-host resistance ....................................................................... 4
  1.5 Microbial experimental evolution ..................................................................................................... 6
  1.6 In planta experimental evolution ...................................................................................................... 7
    1.6.2 In planta experimental evolution with P. syringae ................................................................. 9
  1.7 Hypermutators within adapting pathogen populations .................................................................. 11
    1.7.2 Hypermutators in plant pathogens ......................................................................................... 13
  1.8 Beneficial mutations in novel environments ................................................................................... 13
  1.9 Experimental evolution of Pph 1448A on Arabidopsis ................................................................... 14

## Chapter 2: Materials and Methods

  2.1 Bacterial growth conditions ............................................................................................................. 16
  2.2 Cloning ............................................................................................................................................. 16
  2.3 E. coli transformations ..................................................................................................................... 17
  2.4 Genomic DNA isolation .................................................................................................................. 17
  2.5 Sanger sequencing ........................................................................................................................... 18
  2.6 Construction of freezer stocks ......................................................................................................... 18
  2.7 Electroporation of P. syringae ........................................................................................................ 19
  2.8 SacB counter selection ...................................................................................................................... 20
  2.9 Generation of uniquely barcoded pMKA suicide vectors .............................................................. 20
  2.10 Generation of Δmuts Pph 1448A strains ....................................................................................... 24
  2.11 Fluctuation assays ......................................................................................................................... 31
  2.12 In vitro growth assays ................................................................................................................... 33
  2.13 Plant growth conditions ................................................................................................................ 33
2.14 In planta growth assay 34
2.15 In planta competition assay 37
2.16 Preliminary Pph 1448A serial passaging experiments on Arabidopsis 38
2.17 Optimization of recovery step for the LTEE 39
2.18 The LTEE of twelve barcoded hypermutator lineages of Pph 1448A on Arabidopsis 40
  2.18.2 LTEE freezer stock protocol ......................................................... 41
  2.18.3 LTEE lineage recovery from population crashes or extinctions .............. 41
  2.18.4 Estimation of the initial bacterial loads and generations of lineages during the LTEE 42
2.18.5 Cross-contamination checks of the LTEE lineages .................................. 42

Chapter 3: Development of the Long-Term Evolution Experiment 44
3.1 Fluctuation assays of the Δmuts Pph 1448A strains 44
3.2 In vitro growth assays of the twelve ancestral Δmuts Pph 1448A lineages 46
3.3 Optimization of Pph 1448A growth on Arabidopsis 50
3.4 First preliminary serial passaging experiment of Pph 1448A on Arabidopsis 51
3.5 The effects of plant material on Pph 1448A growth 52
3.6 Full leaf preliminary serial passaging experiment of Pph 1448A on Arabidopsis 53
3.7 Growth of Pph 1448A on Arabidopsis grown in Jiffy pots versus a flat 54
3.8 Optimization of recovery step for the LTEE 55
3.9 Discussion 56

Chapter 4: Long-Term Evolution of Pph 1448A on Arabidopsis 59
4.1 The LTEE of twelve barcoded hypermutator lineages of Pph 1448A on Arabidopsis 59
4.2 Estimation of the initial bacterial loads of the LTEE lineages 61
4.3 Growth dynamics of lineages over the eighty days of LTEE 63
4.4 Lineages crashed after eighty days during the LTEE 67
4.5 Changes in growth room conditions after eighty days of the LTEE 68
4.6 Cross-contamination checks of the LTEE lineages 71
4.7 Discussion 73

Chapter 5: Phenotypic Characterization of the LTEE Lineages 75
5.1 Growth assays of the twelve LTEE lineages after eighty days of passaging 75
5.2 In vitro growth assays of D80 populations compared to ancestors 78
5.3 Growth assay comparing population streak to decontaminated isolates 79
5.4 In planta competition assays with L7 82
5.6 Discussion 84
Chapter 6: Discussion and Future Directions

6.1 Overview
6.2 Prospective adaptive strategies of L7 on Arabidopsis
6.3 Future WGS of the LTEE lineages
6.4 Continuation of the LTEE
6.5 Future modifications to the LTEE
6.6 Conclusion

References
List of tables

Table 1. Primers used in this work ............................................................................................................. 24

Table 2. Plasmids used and barcoded pMKA suicide vectors generated in this work. .................. 25

Table 3. Bacterial Strains and barcoded ΔmutS Pph 1448A mutant collection generated in this work.
......................................................................................................................................................... 31

Table 4. Conditions of each recovery method. ...................................................................................... 40

Table 5. Estimated generations and mutations per clone of the twelve LTEE lineages over eighty days on Arabidopsis
............................................................................................................................................................. 93
List of figures

Figure 1. A schematic of the steps taken to develop barcoded strains of $\Delta muts$ Pph 1448A ............ 23

Figure 2. Culture tube fluctuation assay comparing the mutation rate of $\Delta muts$ Pph 1448A to wild-type ........................................................................................................................................................................... 45

Figure 3. 96-well plate fluctuation assay comparing the mutation rates of nine strains of $\Delta muts$ Pph 1448A to wild-type. ........................................................................................................................................................................... 45

Figure 4. In vitro growth curves of the ancestors of the twelve $\Delta muts$ Pph 1448A lineages used in the LTEE ........................................................................................................................................................................... 48

Figure 5. Boxplots of area under the growth curves of the ancestors of the twelve $\Delta muts$ Pph 1448A lineages used in the LTEE. ........................................................................................................................................................................... 49

Figure 6. Time series growth assay of Pph 1448A on Arabidopsis ........................................................................................................................................................................... 50

Figure 7. Total in planta generations of Pph 1448A on Arabidopsis over the time series growth assay ........................................................................................................................................................................... 51

Figure 8. First preliminary serial passaging experiment of Pph 1448A on Arabidopsis ......................... 52

Figure 9. Growth assay of Pph 1448A testing the effects of leaf material within the inoculum ........... 53

Figure 10. Full leaf preliminary serial passaging experiment of Pph 1448A on Arabidopsis .......... 54

Figure 11. The growth of Pph 1448A on Arabidopsis plants grown in Jiffy pots or in a flat .......... 55

Figure 12. The testing of four recovery methods for the LTEE of Pph 1448A on Arabidopsis ........ 56

Figure 13. An overview of the experimental evolution procedure of Pph 1448A on Arabidopsis ...... 60

Figure 14. The final bacterial load of the twelve independent $\Delta muts$ Pph 1448A lineages at the end of each passage on Arabidopsis during the first eighty days of the LTEE ......................................................... 62
Figure 15. The concentrations of the twelve LTEE lineages before and after processing the undiluted homogenate into transfer inoculum ................................................................. 62

Figure 16. Comparison of expected initial bacterial loads to the observed initial bacterial loads ...... 63

Figure 17. The population growth of each lineage over eighty days of experimental evolution on Arabidopsis .................................................................................................................. 65

Figure 18. The total number of in planta generations each lineage had undergone on Arabidopsis over eighty days of passaging ........................................................................................................... 66

Figure 19. Boxplots of the number of in planta generations undergone by all lineages at each passage ...................................................................................................................... 66

Figure 20. The final bacterial load of the twelve independent ΔmutS Pph 1448A lineages after eighty days of passaging .................................................................................................................. 67

Figure 21. A sliding window of the mean number of generations per passage of L1-L7 ............... 68

Figure 22. Temperature and humidity of the growth room throughout the LTEE ....................... 70

Figure 23. The total mass of infected leaves per lineage before and after eighty days of passaging . 71

Figure 24. The mean Phred quality scores per base pair of each lineage’s eight nucleotide barcode over the LTEE .................................................................................................................. 72

Figure 25. Growth assay of the D80 populations of the twelve LTEE lineages on Arabidopsis ......... 76

Figure 26. Growth assay of D80 populations in parallel with ancestral populations of the twelve LTEE lineages on Arabidopsis ............................................................................................................. 77

Figure 27. Boxplots of the normalized area under the growth curves of the D80 and ancestor populations of the twelve LTEE lineages ............................................................................................................. 79
Figure 28. Plates of bacterial streaks of both ancestral and D80 populations from the twelve LTEE lineages.

Figure 29. Growth assay testing if fungal contamination affected D80 population growth on Arabidopsis.

Figure 30. Time series competition assay of L7D80 over six days on Arabidopsis.

Figure 31. Competition assay of L7D80 in parallel with L7D0 on Arabidopsis.

Figure 32. Summary of the growth assay results of the D80 populations from the twelve LTEE lineages on Arabidopsis.
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA</td>
<td>Abscisic acid</td>
</tr>
<tr>
<td>Avr</td>
<td>Avirulence</td>
</tr>
<tr>
<td>CAGEF</td>
<td>Centre for the Analysis of Genome Evolution and Function</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony forming unit</td>
</tr>
<tr>
<td>Cm</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>Col-0</td>
<td>Colombia-0</td>
</tr>
<tr>
<td>CW</td>
<td>Canadian Wonder</td>
</tr>
<tr>
<td>Cyc</td>
<td>Cylohexamide</td>
</tr>
<tr>
<td>DAMP</td>
<td>Damage associated molecular pattern</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>Elongation factor-Tu</td>
</tr>
<tr>
<td>EPS</td>
<td>Extracellular polysaccharides</td>
</tr>
<tr>
<td>ETI</td>
<td>Effector triggered immunity</td>
</tr>
<tr>
<td>FLS2</td>
<td>Flagellin-sensing 2</td>
</tr>
<tr>
<td>Gent</td>
<td>Gentamycin</td>
</tr>
<tr>
<td>GI</td>
<td>Genomic island</td>
</tr>
<tr>
<td>HGT</td>
<td>Horizontal gene transfer</td>
</tr>
<tr>
<td>Hop</td>
<td>Hrp outer protein</td>
</tr>
<tr>
<td>HR</td>
<td>Hypersensitive response</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-Beta-D-thiogalactoside</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>KB</td>
<td>King’s broth</td>
</tr>
<tr>
<td>Km</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny broth</td>
</tr>
<tr>
<td>LRT</td>
<td>Likelihood ratio test</td>
</tr>
<tr>
<td>LTEE</td>
<td>Long term evolution experiment</td>
</tr>
<tr>
<td>MAMP</td>
<td>Microbial associated molecular pattern</td>
</tr>
<tr>
<td>MA</td>
<td>Mutation accumulation</td>
</tr>
<tr>
<td>MDR</td>
<td>Mutagenic DNA repair</td>
</tr>
<tr>
<td>MMR</td>
<td>Methyl-directed mismatch repair</td>
</tr>
<tr>
<td>MSS</td>
<td>Ma-Sandri-Sarkar</td>
</tr>
<tr>
<td>NHR</td>
<td>Non-host resistance</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>Optical density 600 nm</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen associated molecular pattern</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reactions</td>
</tr>
<tr>
<td>Pph</td>
<td><em>Pseudomonas syringae pv. tomato</em></td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern recognition receptor</td>
</tr>
<tr>
<td>Psy</td>
<td><em>Pseudomonas syringae pv. syringae</em></td>
</tr>
<tr>
<td>PTI</td>
<td>PAMP triggered immunity</td>
</tr>
<tr>
<td>Pto</td>
<td><em>Pseudomonas syringae pv. tomato</em></td>
</tr>
<tr>
<td>pv</td>
<td>Pathovar</td>
</tr>
<tr>
<td>r</td>
<td>Selection rate constant</td>
</tr>
<tr>
<td>RA</td>
<td>Representative ancestor</td>
</tr>
<tr>
<td>Rif</td>
<td>Rifampicin</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>Rpm</td>
<td>Rounds per minute</td>
</tr>
<tr>
<td>SA</td>
<td>Salicylic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>Strp</td>
<td>Streptomycin</td>
</tr>
<tr>
<td>T3SE</td>
<td>Type III secreted effector</td>
</tr>
<tr>
<td>T3SS</td>
<td>Type III secretion system</td>
</tr>
<tr>
<td>TEV</td>
<td><em>Tobacco etch virus</em></td>
</tr>
<tr>
<td>TG</td>
<td>Tendergreen</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole genome sequencing</td>
</tr>
<tr>
<td>Ws-3</td>
<td>Wassilewskija-3</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-Beta-D-galactopyranoside</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 *Pseudomonas syringae*

Pathogens encounter a wide variety of obstacles when infecting a host. For a pathogen to be successful, it must be able to acquire essential nutrients from its host (Fatima and Senthil-Kumar, 2015), and effectively suppress the host’s immune response (Jones and Dangl, 2006). The Gram negative phytopathogenic bacterium *Pseudomonas syringae* is a highly diverse species complex with the ability to cause disease on a wide variety of plant hosts, including many economically important crops such as: beans, tomatoes, cereals, kiwi fruits, tobacco, maize and coffee (Höfte and Vos, 2006; Baltrus *et al.*, 2017). This *P. syringae* species complex can be organized into thirteen distinct phylogenetic groups comprising over 60 pathogenic variants (pathovars) categorized by plant host pathogenicity characteristics, and attests to its broad host range (Baltrus *et al.*, 2017). In addition to being a highly successful pathogen of agricultural plants, populations of *P. syringae* have been isolated in a variety of natural reservoirs such as: streams, snow, rain, leaf litter, and the phyllosphere of wild plants indicative of a complex life history outside of the agricultural environment (Morris *et al.*, 2013; Morris *et al.*, 2008).

When *P. syringae* causes disease, it does so by entering and colonizing the intercellular space of host tissue. *Pseudomonas syringae* generally utilizes two main virulence strategies to disrupt and suppress host immunity. By far the most diverse strategy is the injection of type III secreted effector proteins (T3SEs); virulence-associated biomolecules that enter into the host cytosol via the type III secretion system (T3SS) (Lindeberg *et al.*, 2012). These T3SEs have the shared function of producing a favourable intercellular host environment to colonize, and perform an array of functions in the host cell such as: disrupting early immune signaling, inhibiting antimicrobial vesicle trafficking, suppressing the detection of other effectors, and increasing nutrient availability (Lindeberg *et al.*, 2012; Chen *et al.*, 2010; Nomura, 2007). Currently, over 60 effector families have been identified among *P. syringae* complex members and its hypothesized that this large diversity in effectors has allowed this species
complex to successfully adapt to its broad host range (Baltrus *et al.*, 2017). In conjunction with effectors, another virulence strategy utilized by *P. syringae* is the secretion of toxic secondary metabolites, termed phytotoxins. These toxins can disrupt the host metabolism, induce host cell damage and can even mimic specific plant hormones to suppress immune signaling (Bender *et al.*, 1999; Geng *et al.*, 2014).

### 1.2 Plant innate immunity

Plants have evolved two main layers of innate immunity to withstand challenge from phytopathogens. The first layer is initiated when pattern recognition receptors (PRRs) which consist of receptor kinases or receptor like kinases bind with molecular epitopes termed pathogen associated molecular patterns (PAMPs), also known as microbial associated molecular patterns (MAMPS) (Boller and Felix, 2009; Boutrot and Zipfel, 2017). PAMPs, in most cases, are highly conserved molecules that are shared among a broad spectrum of pathogens but not by the host itself. Some examples of bacterial PAMPs include: the flagellin, Elongation factor-Tu (EF-Tu), siderophores, extracellular polysaccharides (EPS) and lipopolysaccharides (Boller and Felix, 2009; Boutrot and Zipfel, 2017). Once these molecular signatures are recognized by PRRs, PAMP triggered immunity (PTI) is initiated, which results in a variety of responses including: the closure of stomata (Melotto *et al.*, 2006), the rapid production of reactive oxygen species (ROS) within the apoplast (Mitchell *et al.*, 2015; Torres *et al.*, 2006), the deposition of callose polysaccharides within the cell wall (Mitchell *et al.*, 2015), the production of antimicrobials phytoalxins (Ahuja *et al.*, 2012), the expression of multiple immune related genes and the induction of systemic acquired resistance (Mishina and Zeier, 2007b). This first layer of immunity gives plants the ability to maintain surveillance over a broad spectrum of pathogens.

The second layer of innate plant immunity, effector triggered immunity (ETI), is more specific than PTI, and is initiated upon detecting the presence of specific effector proteins (T3SEs or other) within the host cytosol (Jones and Dangl, 2006). Generally, ETI is induced when an intracellular nucleotide binding leucine rich repeat receptor detects the presence of a specific effector either through
the modification of the effector’s intended target protein (guardee) or of a non-intended bait (decoy) protein (Cui et al., 2015; Jones and Dangl, 2006). This layer of immunity generally causes an elevated immune response relative to PTI and in most cases leads to local cell death known as the hypersensitive response (HR) which limits pathogen growth (Jones and Dangl, 2006; Cui et al., 2015).

1.3 P. syringae and host specificity

Although P. syringae as a species complex can infect a diverse range of hosts, individual strains within this species complex have relatively narrow host ranges (Baltrus et al., 2011; Cai et al., 2011). Comparative genomic studies exploring the functional T3SE repertoires among diverse strains of P. syringae, found that overall repertoires differed greatly between strains, with very few T3SEs shared universally (Baltrus et al., 2011; O’Brien et al., 2011). A large amount of diversity among T3SE repertoires was also observed between strains which infect the same host (Almeida et al., 2009; O’Brien et al., 2012) indicating that there are potentially many different combinations of virulence factors within the P. syringae species complex which can lead to pathogenicity onto a specific host.

Despite the importance of T3SEs to successfully colonize plant hosts, individual T3SEs within a repertoire have been shown to be dispensable by the pathogen. This redundancy was exposed in a reverse genetic study using the strain P. syringae pv. tomato (Pto) DC3000, as observable reductions in growth on Nicotiana benthamiana only occurred when a combination of redundant effectors were knocked out (Kvitko et al., 2009). Additionally, in a follow up study they identified that out of the twenty-eight functional effectors within Pto DC3000, only eight were required to restore nearly full fitness on N. benthamiana (Cunnac et al., 2011). Interestingly, disruption/truncation mutations within T3SEs are rarely conserved among diverse strains, indicating that these mutations have occurred evolutionarily recently and probably the result of evading host detection (Baltrus et al., 2011). Both the redundancy of T3SE functions and the coevolutionary arms race to avoid host immune detection have contributed to the highly dynamic evolution of T3SE repertoires among strains of P. syringae.
Also contributing to the dynamic evolution of many T3SEs is horizontal gene transfer (HGT), as many T3SEs reside on mobile genetic elements like plasmids and/or genomic islands (GIs) making their evolution highly dynamic (Lovell et al., 2009; Baltrus et al., 2011; O’Brien et al., 2011). Consequently, highly diverged strains of *P. syringae* in some cases can share nearly identical T3SEs (Baltrus et al., 2011; O’Brien et al., 2011). Overall, both the highly dynamic evolution of T3SEs and the functional redundancy of T3SEs within strain repertoires has made it difficult to link patterns in virulence factor repertoires to certain hosts from a comparative genomics approach. This has contributed to our limited understanding of the important adaptive steps a strain of *P. syringae* takes to cause disease onto novel or non-hosts.

1.4 *P. syringae* pv. *phaseolicola* and non-host resistance

*Pph* is the causal agent of halo blight disease on *Phaseolus vulgaris* (common bean), and continues to contribute to crop loss within developing nations (Arnold et al., 2011; Taylor et al., 1996). *Pph* can survive both on the leaf surface as an epiphyte and within the leaf intercellular space (apoplast) as a pathogen (Arnold et al., 2011; Hirano and Upper, 2000). *Pph* enters the leaf apoplast either through wounds or natural openings (stomata), where it then secretes T3SEs into the host cytosol to successfully colonize host tissue (Baltrus et al., 2012; Arnold et al., 2011). The strain of *Pph* 1448A, was originally isolated in 1985 from diseased kidney beans in Ethiopia and is pathogenic on all tested cultivated variants (cultivars) of the common bean (Arnold et al., 2011; Taylor et al., 1996; Joardar et al., 2005). *Pph* 1448A is one of the three *P. syringae* strains to have a fully sequenced genome (Joardar et al., 2005) making it a popular strain of study.

Although strains within the *Pph* group are able to cause disease among a variety of legume hosts, strains within this group are unable to cause disease on plants outside of this host range including the model plant *Arabidopsis* (Forsyth et al., 2010; Mishina and Zeier, 2007a; Ham et al., 2007). The inability for a pathogen to cause disease on all varieties of a plant species is known as non-host resistance (NHR), and this type of resistance is generally multi-layered and robust (Senthil-Kumar and
Both PTI and ETI can potentially contribute to the NHR a pathogen experiences on plants outside of its host range (Senthil-Kumar and Mysore, 2013). When strains of Pph are inoculated onto wild-type Arabidopsis, they undergo restricted growth and leaves remain symptomless (Forsyth et al., 2010; Mishina and Zeier, 2007a; Ham et al., 2007; Wang et al., 2008). To identify the necessary immune signaling pathways involved in the NHR of Pph on Arabidopsis, one study explored how the growth of Pph NPS3121 was affected on plants compromised in known immune signaling pathways (Ham et al., 2007). They found that disrupting individual immune pathways did not compromise NHR against Pph NPS3121, but only when multiple pathways were disrupted by both mutations and heterologous T3SEs, NHR became compromised (Ham et al., 2007). This signified that the NHR of Arabidopsis against Pph consists of multiple layers of immunity including: salicylic acid (SA) signaling, R-protein function (weak ETI) and cell wall immunity (Ham et al., 2007; Mishina and Zeier, 2007b).

Both tiers of innate plant immunity have been suggested to play an important role in the NHR of Arabidopsis against Pph 1448A. One study took advantage of the natural variation in NHR among cultivars of Arabidopsis against Pph 1448A to map genes important for this resistance (Forsyth et al., 2010). Through quantitative trait loci mapping, they identified two loci which significantly linked with resistance, one of which was the PRR Flagellin-sensing 2 (FLS2) which detects the bacterial flagellin PAMP. The second locus mapped on chromosome 1 termed the basal resistance to Pph 1 (BRF1), is hypothesized to be involved with FLS2 detection, as the presence of the resistant BRF1 allele had no effect on Pph 1448A resistance in an fls2 background (Forsyth et al., 2010). In total, this indicated that a major component responsible for the NHR against Pph 1448A is derived from the PTI that is triggered when the flagellin PAMP is detected by the FLS2 receptor (Forsyth et al., 2010). Additionally, the induction of PTI in Arabidopsis triggers the production of the glucosinolate derivative sulforaphane, a toxic secondary metabolite which has shown to limit the growth of non-host P. syringae strains including
Moreover, all *P. syringae* strains which are virulent on *Arabidopsis* have multiple *sax* genes which enable these strains to be resistant to sulforaphane (Fan *et al.*, 2011). Although *Pph* 1448A does not cause HR on *Arabidopsis*, it appears that some T3SEs within the *Pph* 1448A repertoire are contributing to its NHR. When the T3SEs HopAS1 and AvrB4-1 from *Pph* 1448A were expressed in the virulent strain *Pto* DC3000, it resulted in these strains showing a significant decrease in growth on *Arabidopsis* indicative of an ETI response being caused by these effectors and contributing to NHR against *Pph* 1448A (Sohn *et al.*, 2012; Zumaquero *et al.*, 2010). Together, this evidence suggests that the NHR of *Arabidopsis* against *Pph* 1448A is multi-layered and would require *Pph* 1448A to overcome these layers of immunity before it could cause disease on the non-host. An experimental evolution approach has the potential to uncover the adaptive steps a strain such as *Pph* 1448A would need to take to overcome this multi-layered NHR of *Arabidopsis*.

1.5 Microbial experimental evolution

Experimental evolution is the process where replicate populations are propagated in controlled laboratory conditions over many generations with the aim of understanding the process which these populations adapt to a new environment (Buckling *et al.*, 2009). Microbial experimental evolution has been applied to better understand a variety of processes including: the development of antibiotic resistance (Jansen *et al.*, 2013), the adaptation to various environmental stresses (Dhar *et al.*, 2011; Voordekers *et al.*, 2015), the utilization of specific nutrients (Chu *et al.*, 2017), and the basic processes of adaptation (Lenski *et al.*, 1991). The use of microbes in evolution experiments has many benefits including: ease of handling, large population sizes, and quick generation times (Buckling *et al.*, 2009). Throughout an evolution experiment, samples of the adapting populations are frozen down at specific time intervals to maintain a “fossil record” of the adapting populations, which can then be revived and characterized relative the ancestor with fitness assays (Buckling *et al.*, 2009; Dettman *et al.*, 2012). When evolution experiments are coupled with whole genome sequencing (WGS), they allow for us to
directly identify genomic differences between evolved and ancestral populations, which can help link genetic variants directly with differences in phenotypic fitness (Dettman et al., 2012). The hallmark experimental evolution initiated by Richard Lenski’s group in 1988 consists of twelve independent initially isogenic populations of E. coli being serially transferred in glucose limited media (Lenski et al., 1991). This long term evolution experiment (LTEE) is still currently running and has recently surpassed 66,000 generations to date (Lenski, 2017). This simple experiment has enlightened our understanding on many evolutionary questions such as: how the rate of adaptation changes over time (Wiser et al., 2013; Lenski et al., 1991), how mutation rate dynamics affect the rate of adaptation (Sniegowski et al., 1997), and how epistasis affects the selection of beneficial mutations (Wiser et al., 2013). Additionally, the WGS of over 200 clones from these E.coli lineages after 50,000 generations of evolution has uncovered strong signs of genome-wide parallelism among independent populations (Tenaillon et al., 2016), a characteristic which is commonly observed in evolution experiments and represents a strong sign of selection (Dettman et al., 2012).

1.6 In planta experimental evolution

In addition to helping answer basic evolutionary questions, experimental evolution has also been utilized in a multitude of studies to further understand the dynamic interactions between microbes and plant hosts. Experimental evolution was used to characterize the adaptive process a pathogen takes to become an intracellular infecting nitrogen fixing symbiont (Marchetti et al., 2010; Guan et al., 2013). A modified strain of the vascular infecting pathogen Rastonia solanacearum, which had gained the ability to form nodules and produce extracellular infections was serially passaged on the legume host Mimosa pudica. After 300 generations of being passaged on the roots of the legume host, all three independent populations gained the ability to cause intracellular infections within root nodules, because of mutations which suppressed the expression of virulence genes involved in the T3SS (Guan et al., 2013).
**In planta** experimental evolution has also been applied to understand how differences in the host immune response affects the evolution of infecting pathogen populations. In one study, three replicate populations of *Xanthomonas citri*, the causal agent of citrus canker were serially passaged on kumquat a host on which it elicits HR, while another three replicate populations were passaged on the susceptible grapefruit host (Trivedi and Wang, 2014). After fifty-five serial passages, two out of three populations passaged on kumquat lost the ability to cause HR on the resistant host. Conversely, all three of the populations passaged on the susceptible grapefruit host continued to elicit HR on kumquat (Trivedi and Wang, 2014), demonstrating that the immune response of kumquat drove the pathogen to evade HR. Additionally signs of parallel evolution existed between the two populations which evaded HR, as they shared five genes with fixed non-synonymous mutations (Trivedi and Wang, 2014).

Multiple **in planta** experimental evolution studies have focused on characterizing changes important for helping pathogens adapt to new hosts. One evolution experiment using *R. solanacearum* the causal agent of bacterial wilt, serially passaged independent lineages on three original hosts (tomato, eggplant and pelargonium) and two distant hosts (bean and cabbage) (Guidot et al., 2014). After over 300 generations of passaging, populations adapted onto the distant hosts experienced overall larger gains in fitness compared to populations adapted on the original hosts. Additionally, among the four sequenced populations adapted onto beans, three of them had a non-synonymous mutation within the previously uncharacterized transcription regulator protein termed *efpR* (Guidot et al., 2014). Interestingly, *efpR* was further characterized to being a virulence regulator and a catabolic repressor (Perrier et al., 2016), with the non-synonymous mutations which occurred during the experimental evolution reducing its activity (Guidot et al., 2014).

Likewise, another study using the *Tobacco etch virus* (TEV) adapted independent lineages of this pathogen on the non-host pepper and the original host tobacco (Agudelo-Romero et al., 2008). All four populations adapted to pepper displayed significant increases in growth on the new host, while
populations passaged on tobacco showed no significant change in growth after fifteen serial passages. Three of the four pepper populations shared the same mutation within the gene $Hc$-$Pro$, which codes for a protein essential for replication, transmission and movement systemically within the host (Agudelo-Romero et al., 2008). Furthermore, populations adapted to pepper displayed antagonistic pleiotropy as the gain in fitness on the new host came to a cost in fitness on the tobacco host, indicating a host shift (Agudelo-Romero et al., 2008).

Similarly, antagonistic pleiotropy was observed in lineages of the plum pox virus serially passaged on a novel pea host, as these lineages experienced reduced infection rates on the original peach host (Wallis et al., 2007). All four populations passaged on peas gained the same non-synonymous mutation in the $Nlb$ gene, which codes for an RNA dependent RNA polymerase, an essential protein for replication of the viral genome (Wallis et al., 2007). Overall, these in planta experimental evolutions of pathogens onto novel hosts have repeatedly shown: the in planta environment of novel hosts is able to effectively select for beneficial variants, pathogens experience larger fitness gains on non-adapted hosts, and parallel evolution is common when adapting to the same host.

1.6.2 In planta experimental evolution with P. syringae

Few studies to date have used in planta experimental evolution with the phytopathogen $P. syringae$ to further characterize how it adapts to overcome host immunity. One study uncovered the mechanism to which the bean strain $Pph$ 1302A evaded HR in $P. vulgaris$ cultivar Tendergreen (TG) through the use of serial passaging (Pitman et al., 2005). $Pph$ 1302A causes HR on TG through R3 resistance gene recognition of the T3SE hopARI, which exists on a 106 KB GI termed PPHGI-1 (Jackson et al., 2000; Pitman et al., 2005). This type of resistance imposed by TG on $Pph$ 1302A is an example of host resistance (Gill et al., 2015), as the strain $Pph$ 1302A is pathogenic on other cultivars of the species $P. vulgaris$ such as Canadian Wonder (CW) (Pitman et al., 2005). After five passages of $Pph$ 1302A on resistant TG, 98% of isolates within the population lost the ability to cause HR and were fully virulent on
the once resistant TG host. This gain of virulence was the result of the excision PPHG1-1 from the genome as this removed the HR causing effector hopAR1. Additionally, it was the HR itself selecting for the high prevalence of disease causing isolates in the TG passaged population as conversely, when a population of Pph 1302A was passaged on the susceptible cultivar CW, only 2% of isolates lost the ability to cause HR on TG (Pitman et al., 2005). This represents a clear example of how P. syringae can quickly adapt to evade host resistance through the loss or inactivation the T3SE being detected. This experiment exemplifies the evolutionary arms race occurring between hosts and pathogens, but does not address the early evolutionary steps with which a strain like Pph 1302A undergoes to specialize onto a non-host.

To address this question, an evolution experiment was performed with the bean strain Pph 1448A on the non-host tomato (McCann, 2013). In this evolution experiment, the contribution of both HGT and mutations were tested, as one set of lineages were co-inoculated with sheared DNA from Pto DC3000, a strain which causes disease on tomato (McCann, 2013). After twelve serial passages, one lineage which was co-inoculated with Pto DC3000 sheared DNA showed a significant improvement of growth on tomato. Additionally, this lineage experienced a growth deficit on the original bean host, a sign of antagonistic pleiotropy. WGS of this lineage revealed that the improved growth was not the result of HGT, but most likely the result of two nonsynonymous single nucleotide polymorphisms (SNPs). One of these SNPs was within the flagellar biosynthesis protein (FlhA) potentially affecting the structure of the flagellar apparatus. The other SNP was a reversion mutation in levansucrase, an important enzyme for the production of EPS (McCann, 2013). These findings suggest that mutational changes possibly affecting: motility, PAMP detection, and EPS production are important steps towards novel host adaptation in P. syringae. Further characterization and functional validation of these mutations are still required to understand the mechanism to which these mutations contribute to the adaptation on tomato.
1.7 Hypermutators within adapting pathogen populations

Mutations are the ultimate source of genetic variation with which natural selections acts upon to direct adaptive evolution. Theory predicts with the understanding that most mutations with phenotypic effects are deleterious (Eyre-Walker and Keightley, 2007), that asexual populations will adapt to have mutation rates as low as possible (Drake, 1991) to reduce the accumulation of deleterious mutations within the population (Muller, 1964; McDonald et al., 2012). Alternatively, in cases when bacterial populations are adapting to novel environments, the evolution of mutation rates can be highly dynamic dependent on multiple factors including: the magnitude of the mutation rate, the population size, and the rates of beneficial and deleterious mutations (Taddei et al., 1997; Good and Desai, 2016). In both empirical and in silico experiments where asexual populations were adapting within novel environments, the selection and sometimes fixation of hypermutators has repeatedly occurred (Giraud, 2001; Notley-McRobb et al., 2002; Pal et al., 2007; Shaver et al., 2002; Sniegowski et al., 1997; Voordeckers et al., 2015; Taddei et al., 1997). All current evidence suggests that the selection for a hypermutator allele occurs as a result of the allele hitchhiking along with an adaptive (driver) mutation, which given the elevation in mutation rate is more likely to appear in a hypermutator background (Raynes and Sniegowski, 2014; Notley-McRobb et al., 2002; Shaver et al., 2002). Interestingly, this phenomenon appears to be highly prevalent among human pathogens given the overrepresentation of hypermutators found among pathogenic populations of: *E.coli*, *Salmonella enteritidis* (Leclerc et al., 1996), *Staphylococcus aureus* (Prunier et al., 2003), *Pseudomonas aeruginosa* (Oliver et al., 2000) and *Haemophilus influenzae* (Román et al., 2004) relative to natural populations. This overrepresentation of hypermutators among human infections is not surprising given their superior abilities to gain mutations important for adapting to the highly dynamic host environment (Oliver and Mena, 2010). Many of these adaptations include: gaining antibiotic resistance (Oliver et al., 2000; Leclerc et al., 1996; López-Causapé et al., 2013), biofilm adaptation (Luján et al., 2011), resistance against oxidative stresses (Torres-Barceló
et al., 2013) and enhanced persistence within chronic infections (Smith et al., 2006; Mena et al., 2007; Labat et al., 2005).

Many of the hypermutators characterized among pathogenic populations arise due to defects within the methyl-directed mismatch repair (MMR) system, a post replicative repair system (Leclerc et al., 1996; Oliver et al., 2002). Within prokaryotes, the MMR system is responsible for recognizing and repairing single base pair mismatches along with 1-4 base pair insertion/deletion (indel) mutations (Guarné, 2012; Jun et al., 2006). This process of DNA repair is initiated when the MutS homodimer recognizes and binds to a kink within the DNA caused by a mismatch or indel mutation. The MutS homodimer binds to the kink and initiates the MMR system by attracting the necessary machinery needed to excise the newly synthesized errant daughter DNA strand, and to replace it with a correctly synthesized strand. This includes attracting the MutL homodimer which has the role of attracting and activating the endonuclease MutH which nicks the non-methylated daughter strand at the nearest GATC site (Guarné, 2012; Jun et al., 2006). MutL also loads a helicase II called UvrD into the nick generated by MutH to unwind the cleaved DNA. Many bacteria such as P. syringae lack a MutH homolog and therefore distinguish and cut the daughter strand through a currently uncharacterized method (Guarné, 2012). The errant daughter strand is digested by specific exonucleases (5’ to 3’ or 3’ to 5’ digesting) proceeded by being synthesized by DNA polymerase III and stitched together by DNA ligase. Overall, when any one of these proteins within the MMR system are non-functional, the bacterium is essentially MMR deficient and experiences a constitutive elevation in mutation rate anywhere between 100 to 1000-fold (Oliver and Mena, 2010; Dettman et al., 2016). Mutation accumulation (MA) experiments followed by WGS of MMR deficient hypermutators have characterized the mutation rate spectrum of these mutants and have shown that these elevations in mutation rate display specific biases towards transitions and small insertion mutations (Dettman et al., 2016; Long et al., 2014). These biases can be accounted for when exploring the genomic variation of adapting MMR deficient populations.
1.7.2 Hypermutators in plant pathogens

Although constitutive hypermutators are highly prevalent among human pathogen populations, their presence among plant pathogen populations have yet to be revealed. To date, only one study has explored the variation in mutation rates among bacterial plant pathogens, and they looked specifically at three strains of the phytopathogen *Pseudomonas viridiflava* (Bartoli et al., 2015). They found that transparent colony variants within each of these strains had mutation rates between 2-6-fold larger than wild-type, and that these elevations in mutation rate were not the result of MMR defects (Bartoli et al., 2015). Another study using the bean pathogen *P. syringae* pv. syringae (*Psy*) B86-17, showed that under increased UV-B stress, *Psy* B86-17 significantly elevated its mutation rate due to the induction of the *rulAB* operon (Kim and Sundin, 2000). This operon is homologous to the *umuDC* operon a well characterized component of the SOS response in *E. coli* (Smith and Walker, 1998) resulting in mutagenic DNA repair (MDR) (Kim and Sundin, 2000). This elevation in mutation rate was accompanied with an increased survival rate under UV-B stress, as *rulAB* knockouts under UV-B stress displayed a significant reduction in survival (Kim and Sundin, 2000). Additionally, the *rulAB* operon is widely distributed among many strains of *P. syringae* indicative of the importance of UV-B resistance and the prevalence of MDR among *P. syringae* (Sundin and Murillo, 1999; Gutiérrez-Barranquero et al., 2017). Overall, although constitutive hypermutators of plant pathogens such as *P. syringae* have yet to be isolated, the transiently induced mutation rates under UV-B stress may play an important role in accelerating the adaption of *P. syringae* when living in the phyllosphere.

1.8 Beneficial mutations in novel environments

An important factor which determines the advantage of having an elevated mutation rate is the fraction of mutations which are potentially beneficial within the current environment (Good and Desai, 2016). To understand the effects of individual random mutations in specific environments, one study measured the fitness of twenty-six genotypes of *E. coli*, which differed by only a single mutation over
four different environments that varied between temperature (28°C or 37°C) and carbon source (glucose or maltose) (Remold and Lenski, 2001). They identified that a surprisingly large proportion of mutations (approximately 12%) were beneficial in maltose, a carbon source to which the genotypes were not adapted to (Remold and Lenski, 2001). This indicated that within an environment that differed in only a single factor, a large proportion of mutations can be beneficial. Another study looked at the distribution of mutational fitness effects by measuring the fitness of twenty genotypes of the RNA plant virus TEV which differed by a single random substitution mutation, over five natural Solanaceae hosts and three partially susceptible non-hosts (Lalić et al., 2011). Most of the mutations had either deleterious or lethal effects on the fitness of infecting TEV when fitness of these genotypes were measured over the five natural hosts. Conversely, when fitness effects of these same mutations were measured on the three non-hosts, a significantly larger proportion of mutations were beneficial (Lalić et al., 2011). This suggests that when a pathogen is adapting to a novel host, it is likely that a larger portion of all mutations would be beneficial within this new environment, and is supported by evolution experiments where pathogens show larger increases in fitness when adapted onto novel hosts (Guidot et al., 2014; Agudelo-Romero et al., 2008). Therefore, an elevated mutation rate is likely to aide a pathogen in gaining beneficial mutations and to adapting on a new host.

1.9 Experimental evolution of Pph 1448A on Arabidopsis

In this thesis, an experimental evolution approach was used with the goal of characterizing the early adaptive steps a strain of P. syringae takes to overcome NHR, and to become pathogenic on a new host. To do so, twelve independent lineages of the bean pathogen Pph 1448A were serially passaged on the non-host Arabidopsis. Despite the lack of hypermutators present in natural plant pathogen populations, we hypothesized given the ability of hypermutators to rapidly adapt in human pathogen populations (Oliver et al., 2000; Leclerc et al., 1996; López-Causapé et al., 2013; Luján et al., 2011; Torres-Barceló et al., 2013; Smith et al., 2006; Mena et al., 2007; Labat et al., 2005), and the expected
increase in the fraction of mutations which are beneficial in the non-host environment (Lalić et al., 2011), that using hypermutators in this study would rapidly increase the rate at which beneficial mutations arise and are selected for within the passaging Pph 1448A lineages. Therefore, the first objective was to engineer barcoded MMR deficient hypermutating strains of Pph 1448A to use for the LTEE on Arabidopsis. Lineages were uniquely barcoded to allow us to track for cross-contamination between lineages throughout the experiment. The second objective was to optimize and implement an evolution experiment where twelve independent hypermutating lineages of Pph 1448A were serially passaged on Arabidopsis. Lastly, evolved lineages were phenotypically characterized for fitness on the non-host Arabidopsis by using both growth assays and in planta competition assays against the ancestral strain. In this LTEE, we hypothesized that lineages of Pph 1448A passaged on Arabidopsis, would adapt improved growth on the non-host relative to ancestral Pph 1448A. After eighty days of serial passaging, one lineage has displayed a significant increase in growth on Arabidopsis. WGS of this lineage has the possibility of uncovering the early adaptations necessary for a strain P. syringae to infect a new host.
Chapter 2: Materials and Methods

2.1 Bacterial growth conditions

All *P. syringae* strains used were grown at 30°C on King’s broth (KB) media while all *Escherichia coli* strains were grown in Lysogeny broth (LB) media at 37°C. When necessary, liquid cultures or media plates with agar for either *P. syringae* or *E. coli* were supplemented with antibiotics, reagents and sugars at the following concentrations: chloramphenicol (Cm) 25 μg/ml, gentamycin (Gent) 15 μg/ml, kanamycin (Km) 50 μg/ml, rifampicin (Rif) 50 μg/ml, streptomycin (Strp) 100 μg/ml, cycloheximide (Cyc) 50 μg/ml, isopropy-beta-D-thiogalactoside (IPTG) 100 mg/ml, 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal) 20 mg/ml and sucrose at 5%. Media growing wild-type strains of *P. syringae* were supplemented with Rif while ΔmutS strains were supplemented with Rif and Km.

2.2 Cloning

All plasmids were isolated from *E. coli* DH5α using the Thermo Scientific™ Molecular biology GeneJET plasmid Miniprep kit following the product’s manual protocol for high copy plasmids. All polymerase chain reactions (PCRs) were performed using either Thermo Scientific™ Taq DNA polymerase or NEB Q5® High-Fidelity DNA polymerase following the product manual protocols from each product. All primers used in this work are listed in Table 1. The following restriction enzymes were used along with the specified buffer and incubation temperature: SmaI with CutSmart® at 25°C, EcoRV-HF® with CutSmart® at 37°C, BamHI with Tango at 37°C, SsPl-HF® with CutSmart® at 37°C, Pfol with Tango at 37°C, Xhol with CutSmart® at 37°C, and BstAPI with CutSmart® at 60°C. Restriction digests were kept at the incubation temperature for at least 1.5 hours. PCR and restriction digest products which were used for cloning were either gel purified or cleaned using Macherey-Nagel Nucleospin® Gel and PCR Clean-up kit following the product’s manual protocol. All ligations were performed using Thermo Scientific™ T4 DNA ligase following a modified product manual protocol. A 10 μl reaction would be
made consisting of 0.8 µl of T4 DNA ligase, 1 µl of 10X Ligation Buffer, a 1:3 or 1:5 molar ratio of insert to cut plasmid and a top up of sterile ddH_2O. Ligation reactions were left at 16°C overnight, before being transformed.

2.3 E. coli transformations

A ligation to be transformed was first incubated at 70°C for 20 minutes to heat inactivate the T4 DNA ligase. Once heat inactivated, the ligation reaction would be placed on ice along with a frozen tube consisting of approximately 150 µl of concentrated chemically competent E. coli DH5α cells. Once the cells thawed, the entire ligation reaction (10 µl) was added and mixed to the competent cells. Afterwards, competent cells were then kept on ice for 20 minutes, then heat shocked for 50 seconds at 42°C and then placed back on ice for 2 minutes. Cells were then inoculated into 1 ml of Super Optimal broth with Catabolite repression media and left shaking at 37°C for 1 hour. Cells were then plated on LB plates supplemented with the appropriate antibiotics.

2.4 Genomic DNA isolation

Genomic DNA isolation from P. syringae was performed using the Gentra® Puregene® DNA isolation kit following a modified product manual protocol for Gram-Negative bacterium. A 1 ml sample from an overnight culture was aliquoted into a 1.5 ml microfuge tube and was centrifuged for 5 minutes at 4600 rounds per minute (rpm). The supernatant was removed and 600 µl of cell lysis solution was added to re-suspend the bacterial pellet. This suspension was incubated at 80°C for 5 minutes to further lyse cells. Once at room temperature, 1.2 µl of RNase A [10 µg/ml] was added, mixed by inverting the tube several times and then incubated at 37°C for 15-60 minutes. Once at room temperature, 200 µl of protein precipitation solution was added, followed by being vortexed for 20 seconds and then centrifuged at 13,000 rpm for 8 minutes. The DNA containing supernatant was then transferred to a fresh 1.5 ml microfuge tube. To isolate the DNA from solution, 600 µl of 100% 2-propanol was added followed by repeatedly inverting the tube approximately 50 times. The DNA was either centrifuged at
13,000 rpm for 8 minutes followed by the removal of the supernatant and re-suspended in 600 µl of 70% ethanol or was plucked from the 2-propanol solution using a wide bore sterile p200 tip and transferred into 600 µl of 70% ethanol. Once the DNA was transferred to 70% ethanol, the tube was inverted several times and centrifuged at 13,000 rpm for 8 minutes to pellet the DNA. The tube was drained and air dried until all 70% ethanol had evaporated. The dried DNA pellet was rehydrated in 100 µl of Tris EDTA buffer (pH 8) by either leaving DNA suspended at room temperature overnight or by heating to 65°C for 1 hour.

2.5 Sanger sequencing

DNA was prepared for sequencing by first PCR amplifying the locus of interest from genomic or cloning vector template, unless otherwise stated. The PCR template was then cleaned either by the Macherey-Nagel Nucleospin® Gel and PCR Clean-up kit or with NEB Exonuclease I and NEB Calf Intestinal Phosphatase by adding 0.2 µl of each enzyme into a 25 µl PCR reaction. Afterwards, a PCR sequencing reaction was performed with each reaction consisting of: 2 µl of the cleaned PCR template, 0.5 µl of BDTV 3.1 mix, 0.5 µl of a single primer [10 µM], 1.75 µl of 5X sequencing buffer, and 5.25 µl of sterile ddH₂O. If the PCR sequencing reaction was done using vector template, the template was heated to 86°C for five minutes prior to setting up the reaction. Samples were then sequenced by the Centre for the Analysis of Genome Evolution and Function (CAGEF).

2.6 Construction of freezer stocks

Freezer stocks of all bacterial strains used in the work were stocked as such. Liquid cultures were initiated the day before with the appropriate antibiotics and shaken at 225 rpm till reaching saturation. Stocks were stored in a final concentration of 10% glycerol by mixing 900 µl of saturated culture with 900 µl of 20% glycerol solution in a 2ml cryogenic tube. Stocks were stored at -80°C.
2.7 Electroporation of *P. syringae*

To introduce cloning vectors into *Pph* 1448A, cells were made competent through electroporation by the following protocol. A saturated 5 ml overnight culture of *Pph* 1448A was used to inoculate 300 ml of KB Rif and left shaking at 30°C at 225 rpm. Once the culture reached mid-late log phase (OD₆₀₀ 0.2-2), 36 ml aliquots were placed into eight separate 50 ml conical tubes. These tubes were centrifuged at 2400 rpm for 20 minutes at 4°C in the Eppendorf centrifuge 5810R. The supernatant was decanted and each tube was re-suspended with 25 ml of ice cold sterile ddH₂O. The re-suspended tubes were then centrifuged at 2400 rpm for 20 minutes at 4°C in the Eppendorf centrifuge 5810R. The previous wash and centrifugation steps were repeated twice. The final pellets in each conical tube were re-suspended with 600 μl of ice cold sterile ddH₂O along with 150 μl of ice cold sterile 50% glycerol solution forming a final glycerol concentration of 10%. Cell suspensions were then divided into 125 μl aliquots into fresh 1.5 microfuge tubes. Microfuge tubes containing concentrated cell suspensions were flash frozen in liquid nitrogen and stored at -80°C for future use.

When competent cells were ready for use, tubes were placed and thawed on ice. Once thawed, 5 μl of vector miniprep was added. Once mixed, cells in the microfuge tube were aliquoted to individual prechilled Sigma-Aldrich® 0.2 mm electroporation cuvettes. Cells were electrically pulsed using the Bio-Rad Gene Pulser® II at the following setting: 2.5 kV, 400 ohms, and 25 μf. After being pulsed, 1ml of KB was added and gently mixed. Cells were then recovered by pouring cultures into 1ml culture tubes and placed in the 30 °C shaker at 225 rpm for 1-2 hours. Once cells were recovered, they were moved into 1.5 ml microfuge tubes and centrifuged at 4600 rpm for 5 minutes. Afterwards, 800 μl of the supernatant was removed, the pellets were re-suspended in the remaining supernatant and plated onto pre-warmed KB Rif Km plates and placed in the 30 °C incubator for two to three nights.
2.8 SacB counter selection

The expression of the sacB gene from Bacillus subtilis in Gram-negative bacteria is lethal in the presence of sucrose (Gay et al., 1985), making this gene a commonly used counter-selective marker when engineering complete gene deletions. Counter selection of the sacB marker was performed by streaking a recombinant colony onto KB Rif Km plates supplemented with 5% sucrose. A single sucrose resistant colony from each streak was recovered 1-2 days after and was re-streaked onto KB Rif Km plates. These streaks were then used to inoculate KB Rif Km liquid cultures for freezer stocks.

2.9 Generation of uniquely barcoded pMKA suicide vectors

The steps taken to engineer the hypermutating strains of Pph 1448A used in the LTEE are summarized in Figure 1. Suicide vectors cannot replicate in the host cell, therefore colonies expressing the selective marker on the vector are a consequence of the vector integrating into the host’s chromosome through recombination. First, a library of barcoded suicide vectors which would both target the mutS gene through homologous recombination, and replace it with a uniquely barcoded antibiotic marker were made. The cloning vector pEX18Gm was digested using Smal which cut at the multiple cloning site located in the lacZ reporter gene. A nearly complete mutS gene was PCR amplified from Pph 1448A genomic prep. This PCR product was gel purified to remove the smaller secondary PCR product. The gel purified product was blunt end ligated to the Smal cut pEX18Gm. The ligated vector was transformed into chemically competent DH5α E. coli cells and plated on LB Cm plates supplemented with Xgal and IPTG. A blue white screen was used to screen for clones with an insert, as blue colonies represented vector with an intact lacZ reporter gene (re-ligated vector), and white colonies represented vector with an insert disrupting the lacZ reporter gene. A colony PCR on all white colonies was performed amplifying the insert. Products of the appropriate size were Sanger sequenced to confirm that candidate clones had the mutS insert forming pEX18GmMutS (Figure 1A).
We replaced the Gent resistance (Gent<sup>R</sup>) cassette on pEX18GmMutS with a Cm resistance (Cm<sup>R</sup>) cassette as the former was non-functional in <i>P. syringae</i>. First, the Cm<sup>R</sup> cassette was PCR amplified from the cloning vector pSM1695. pEX18GmMutS was then digested with EcoRV which cut within the Gent<sup>R</sup> cassette. The cut vector was blunt end ligated to the PCR amplified Cm<sup>R</sup> cassette forming pEX18CmMutS (Figure 1B). The insertion of the Cm<sup>R</sup> cassette was confirmed through colony PCR amplification of the target region. In addition, a diagnostic restriction digest with SspI of the pEX18CmMutS confirmed the insertion.

The original <i>sacB</i> construct on pEX18CmMutS needed to be substituted because previous experiments had shown that it was non-functional in <i>P. syringae</i>. Alternatively, the <i>sacB</i> construct from the cloning vector pMTN1907 has worked effectively in <i>P. syringae</i> because it contains a constitutive <i>Salmonella</i> trp promoter (Baltrus et al., 2012). Therefore, we removed the original <i>sacB</i> construct from pEX18CmMutS, and replaced it with the <i>sacB</i> construct from pMTN1907. First, pEX18CmMutS was double digested with the restriction enzymes PfoI and BamH1. The larger band consisting of the vector backbone without the <i>sacB</i> construct was extracted through gel purification. Along with removing the original <i>sacB</i> construct, this digest also removed the origin of transfer from the vector meaning that the final suicide vector could not be introduced into <i>P. syringae</i> via conjugation. The <i>sacB</i> gene along with the trp promoter was PCR amplified from pMTN1907 with PfoI and BamH1 restriction sites. This PCR product was double digested with PfoI and BamH1, and then ligated to the double digested pEX18CmMutS forming pEX18CmMutStrpSacB (Figure 1C). The insertion of the pMTN1907 <i>sacB</i> construct was confirmed by both PCR amplification of trp promoter and by Sanger sequencing the trp promoter from the pEX18CmMutStrpSacB vector.

The final step to designing the appropriate library of suicide vectors was to insert a pool of uniquely barcoded kanamycin resistance (Km<sup>R</sup>) cassettes within the middle of the <i>mutS</i> gene in pEX18CmMutStrpSacB. First, pEX18CmMutStrpSacB was double digested with BstAP1 and Xho1. This
removed ~1.5 KB middle portion of the mutS gene, leaving the vector with two flanking ends of mutS with lengths of approximately 500 base pairs each (Figure 1D). We elected not to gel purify the cut vector as this reduced the efficiency of ligations. The forward primer npt+XhoI (Table 1) used to amplify the Km\(^{R}\) cassette from the vector pUCP20TK contained eight degenerate bases upstream of the XhoI restriction site and downstream of the annealing region which resulted in the pool of amplified Km\(^{R}\) cassettes having many unique barcodes. The reverse primer npt-BstAP1 (Table 1) used to amplify the Km\(^{R}\) cassette had a BstAP1 restriction site on the 5’ end resulting in the amplified Km\(^{R}\) cassettes having appropriate restriction sites for a sticky end ligation. The pool of barcoded Km\(^{R}\) cassettes were double digested by BstAP1 and XhoI and then sticky end ligated to the double digested pEX18CmMutStrpSacB. This resulted in the production of a library of Km\(^{R}\) suicide vectors termed the mutS Knockout Allele (pMKA), each with unique barcodes (Figure 1F). This library of vectors was transformed into chemically competent DH5 \(\alpha\) E. coli and selected on LB Km Cm plates. Barcodes of each isolate were identified by first colony PCR amplifying the mutS region with primers Laczp+ and SacB-143 (Table 1). These PCR products were used as template in a sequencing PCR reaction which used the primer MutS-2120 to sequence the barcode. In total, fifty-five uniquely barcoded pMKAs were isolated and confirmed through Sanger sequencing (Table 2).
Figure 1. A schematic of the steps taken to develop barcoded strains of ΔmutS Pph 1448A. A. pEX18Gm was digested with Smal at the lacZ reporter gene and ligated to PCR amplified Pph 1448A mutS gene. B. pEX18GmMutS was digested with EcoRV and ligated to the PCR amplified CmR cassette. C. pEX18CmMutS was double digested with BamH1 and Pfo1 and ligated to the PCR amplified SacB construct from pMTN1907 with the trp promoter and appropriate restriction digest ends. D. pEX18CmMutStrpSacB was double digested with BstAP1 and Xho1. E. KmR cassettes were amplified from pUCP20TK with one primer having eight degenerate bases adding random barcodes. Barcoded KmR cassettes were double digested by BstAP1 and Xho1 and ligated to the cut pEX18CmMutStrpSacB. F. Suicide vectors pMKA* were electroporated into Pph 1448A and recombinants were selected for on KB Km Rif plates. G. Plated single recombinants onto KB Kan Rif 5% sucrose plates for SacB counter, selecting for double recombinants (complete knockouts).
2.10 Generation of Δmuts Pph 1448A strains

Once the appropriate library of suicide pMKA vectors were engineered, they were introduced into component Pph 1448A cells via electroporation. Afterwards, cells were subjected to Km selection which isolated clones with the pMKA vector integrated via homologous recombination at the mutS gene (Figure 1F). Recovered clones were then subjected to sucrose counter selection while still maintaining Km selection, as this selected for the removal of sacB, but the conservation of the barcoded Km\(^\text{R}\) cassette, resulting in the replacement of the mutS gene with the barcoded Km\(^\text{R}\) cassette (Figure 1G). Complete knockouts were confirmed by first PCR amplifying the mutS locus from genomic preps using the primer pair MutS-ext-F and MutS-ext-R (Table 1). These primers amplified just upstream and downstream of the mutS locus. PCR products were then used as template in a sequencing PCR reaction which used the primer MutS-2120 to sequence the barcode. Fourteen uniquely barcoded Δmuts Pph 1448A strains were confirmed by Sanger sequencing, twelve of which were used in the LTEE (Table 3).

Table 1. Primers used in this work.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5'-3')</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutS+17</td>
<td>CGACCTCTCCTCACATACCCCCCAT</td>
<td>None</td>
</tr>
<tr>
<td>MutS-2488</td>
<td>GGTGCGGCGAGACTTGGCGAATATCT</td>
<td>None</td>
</tr>
<tr>
<td>MutS-2120</td>
<td>CCGTGAAGGTACTGGTGCCAC</td>
<td>None</td>
</tr>
<tr>
<td>MutS-ext-F</td>
<td>TTATCTCGTCCACGGTTTCTCGTTC</td>
<td>None</td>
</tr>
<tr>
<td>MutS-ext-R</td>
<td>CAGACTTCTACGCAGTGGTGTAAC</td>
<td>None</td>
</tr>
<tr>
<td>Laczp+</td>
<td>CGTATGTTGTGGAATTTGTGAGCG</td>
<td>None</td>
</tr>
<tr>
<td>Lacz-90</td>
<td>TGCTGCAAGGCGATTAGTTGGG</td>
<td>None</td>
</tr>
<tr>
<td>aacC1+</td>
<td>CATAAGCCCTGTTGCGTTTCGTAATGG</td>
<td>None</td>
</tr>
<tr>
<td>aacC1-</td>
<td>GGCTCTGACGTGCACTGGGAACGG</td>
<td>None</td>
</tr>
<tr>
<td>SmTrp+</td>
<td>CGATCGGCTAGCGGCAAATATACTG</td>
<td>None</td>
</tr>
<tr>
<td>SacB-1389</td>
<td>TTATTTGTTAATCTGTTAATTTGTCTGTCAAGG</td>
<td>None</td>
</tr>
<tr>
<td>SmTrp+BamHI</td>
<td>CGCGGATCCGCGAGTGGCTAGGCGGCAAATATACTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>SacB-1389+BfoI</td>
<td>GCAGCTCCCAGGAGCTTTGTCTGGTTATGCATG</td>
<td>PfoI</td>
</tr>
<tr>
<td>npt+BstAPI</td>
<td>CTGCGACAAGCAAGGTGTGATTTGCTGCATTGCCCT</td>
<td>BstAPI</td>
</tr>
<tr>
<td>SacB-143</td>
<td>TTCGGGATTTCGAGCATACGATG</td>
<td>None</td>
</tr>
</tbody>
</table>
Table 2. Plasmids used and barcoded pMKA suicide vectors generated in this work.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Antibiotic Resistance</th>
<th>Barcode</th>
</tr>
</thead>
<tbody>
<tr>
<td>pEX18Gm</td>
<td>Cloning vector</td>
<td>Gm</td>
<td>None</td>
</tr>
<tr>
<td>pMTN1907</td>
<td>Cloning vector</td>
<td>Tet Km</td>
<td>None</td>
</tr>
<tr>
<td>pUCP20TK</td>
<td>Cloning vector</td>
<td>Km</td>
<td>None</td>
</tr>
<tr>
<td>pEX18GmMutS</td>
<td>pEX18Gm derivative carrying the 1448A MutS gene</td>
<td>Gm</td>
<td>None</td>
</tr>
<tr>
<td>pEX18CmMutS</td>
<td>pEX18GmMutS derivative containing Cm resistance gene</td>
<td>Cm</td>
<td>None</td>
</tr>
<tr>
<td>pEX18CmMutStrpSacB</td>
<td>pEX18CmMutS derivative containing the SacB gene with a trp promoter from pMTN1907</td>
<td>Cm</td>
<td>None</td>
</tr>
<tr>
<td>pMKA1</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGAAGGGG</td>
</tr>
<tr>
<td>pMKA2</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TCCGTCTC</td>
</tr>
<tr>
<td>pMKA3</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GTCGATGG</td>
</tr>
<tr>
<td>pMKA4</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GTCATGGG</td>
</tr>
<tr>
<td>pMKA5</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TCGTGTC</td>
</tr>
<tr>
<td>pMKA6</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GAATGTCA</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>pMKA7</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGTAAAAA</td>
</tr>
<tr>
<td>pMKA8</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGGTGTT</td>
</tr>
<tr>
<td>pMKA9</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGCCTGTG</td>
</tr>
<tr>
<td>pMKA10</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGGATGCT</td>
</tr>
<tr>
<td>pMKA11</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GATGTAGG</td>
</tr>
<tr>
<td>pMKA12</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CTGGGGGA</td>
</tr>
<tr>
<td>pMKA13</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCGAGGCT</td>
</tr>
<tr>
<td>pMKA14</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGTCAGGC</td>
</tr>
<tr>
<td>pMKA15</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CGGCAGGA</td>
</tr>
<tr>
<td>pMKA16</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>ATGTCACG</td>
</tr>
<tr>
<td>pMKA17</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCTGGCTG</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>pMKA18</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGATCCGG</td>
</tr>
<tr>
<td>pMKA19</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGGCGACA</td>
</tr>
<tr>
<td>pMKA20</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CCCATAGG</td>
</tr>
<tr>
<td>pMKA21</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TGGTAGGT</td>
</tr>
<tr>
<td>pMKA22</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CCCTGTTC</td>
</tr>
<tr>
<td>pMKA23</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>ATGACTTT</td>
</tr>
<tr>
<td>pMKA24</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGAGAGGG</td>
</tr>
<tr>
<td>pMKA25</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TCGAGTTA</td>
</tr>
<tr>
<td>pMKA26</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TGCAGTAC</td>
</tr>
<tr>
<td>pMKA27</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TTGTGCTT</td>
</tr>
<tr>
<td>pMKA28</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CAGCCTGG</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>pMKA29</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGGGTAGT</td>
</tr>
<tr>
<td>pMKA30</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AATTGTGC</td>
</tr>
<tr>
<td>pMKA31</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GAAGCGGT</td>
</tr>
<tr>
<td>pMKA32</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GTGCAATG</td>
</tr>
<tr>
<td>pMKA33</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>ATGGGATA</td>
</tr>
<tr>
<td>pMKA34</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGTGGGGT</td>
</tr>
<tr>
<td>pMKA35</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>ATAGGCGG</td>
</tr>
<tr>
<td>pMKA36</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGTTTGTG</td>
</tr>
<tr>
<td>pMKA37</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGCCGTAT</td>
</tr>
<tr>
<td>pMKA38</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CTTATTAG</td>
</tr>
<tr>
<td>pMKA39</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGGTTTGT</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>pMKA40</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GATGTGAG</td>
</tr>
<tr>
<td>pMKA41</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CGCGGCGT</td>
</tr>
<tr>
<td>pMKA42</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>ACCCGTGG</td>
</tr>
<tr>
<td>pMKA43</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGTGCAAA</td>
</tr>
<tr>
<td>pMKA44</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GATACGTC</td>
</tr>
<tr>
<td>pMKA45</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCAAATGA</td>
</tr>
<tr>
<td>pMKA46</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GACGATAAC</td>
</tr>
<tr>
<td>pMKA47</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GGGTGCTT</td>
</tr>
<tr>
<td>pMKA48</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>TGAGGTGA</td>
</tr>
<tr>
<td>pMKA49</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCTGGTAC</td>
</tr>
<tr>
<td>pMKA50</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCCATCGC</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>pMKA51</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GACGTTGA</td>
</tr>
<tr>
<td>pMKA52</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>CTAGAGGA</td>
</tr>
<tr>
<td>pMKA53</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>AGTTGGAG</td>
</tr>
<tr>
<td>pMKA54</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GTGGACGA</td>
</tr>
<tr>
<td>pMKA55</td>
<td>pEX18CmMutStrpSacB derivative carrying the ΔmutS knockout allele with unique barcode</td>
<td>Km Cm</td>
<td>GCTTACAC</td>
</tr>
</tbody>
</table>
Table 3. Bacterial Strains and barcoded ΔmutS Pph 1448A mutant collection generated in this work.

<table>
<thead>
<tr>
<th>Bacterial Strain</th>
<th>Description</th>
<th>Genomic Barcode</th>
<th>Antibiotic Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> DH5α</td>
<td>NA</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td><em>P. syringae</em> pv. phaseolicola 1448A</td>
<td>Race 6; wild-type strain</td>
<td>None</td>
<td>Rif</td>
</tr>
<tr>
<td>L1</td>
<td>ΔmutS AGGATGCT</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L2</td>
<td>ΔmutS GGGTGGA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L3</td>
<td>ΔmutS TCCGTCTC</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L4</td>
<td>ΔmutS CCCATAGG</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L5</td>
<td>ΔmutS CGATTGTA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L6</td>
<td>ΔmutS CAGCCTGG</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L7</td>
<td>ΔmutS GATACGTC</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L8</td>
<td>ΔmutS TGGTACGG</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L9</td>
<td>ΔmutS GGGCGACA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L10</td>
<td>ΔmutS GAATGCTA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L11</td>
<td>ΔmutS ATGGGAGA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>L12</td>
<td>ΔmutS CTGGGAGA</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>RA</td>
<td>ΔmutS AGTTTGTC</td>
<td></td>
<td>Rif Km</td>
</tr>
<tr>
<td>RA2</td>
<td>ΔmutS AGAGAGGG</td>
<td></td>
<td>Rif Km</td>
</tr>
</tbody>
</table>

2.11 Fluctuation assays

Strains used were prepared by first streaking a single colony onto KB plates with the appropriate antibiotics and grown over night. Liquid cultures for strains were made by inoculating 5 ml KB liquid cultures from the fresh streaks. The overnight cultures for each strain were diluted into 80 ml of KB at a concentration of 5000 CFU/ml (OD_{600} of 0.00001). Fifteen replicate liquid cultures of 5 ml were aliquoted for each strain and left shaking at 30 °C at 225 rpm over three days to allow cultures to reach saturation. Afterwards, twelve out of the fifteen cultures from each treatment were plated onto KB Strp plates to quantify the number of spontaneous Strp resistant (Strp^{R}) mutants within each saturated culture. This was done by first dividing each culture into four separate 1.25 ml aliquots in 1.5 ml microfuge tubes. Microfuge tubes were centrifuged at 4600 rpm for 5 minutes to pellet cultures. After centrifugation, 1 ml of supernatant was removed from each tube and pellets were re-suspended in the remaining 250 μl of culture. The four tubes from each original culture were plated over two KB Strp plates each. The
remaining three cultures from each treatment were serially diluted, and 200 µl from the dilution factor of 10⁶ was spread plated onto plain KB plates to determine the mean final cell population (N_f) for each treatment. Plates were incubated at 30°C for two days and then counted. All fluctuation assays performed had a wild-type Pph 1448A control. The web tool FALCOR was used to calculate the average number of mutations per culture (m) and 95% confidence intervals using the Ma-Sandri-Sarkar (MSS) maximum-likelihood method (Hall et al., 2009). This was then converted to mutation rate per cell per generation by dividing m with N_f (Hall et al., 2009). To compare if the mutation rate of the Δmuts strain differed significantly from wild-type, the R-package rSalvador was used to calculate the likelihood ratio test (LRT) incorporating the difference in N_f between treatments (Zheng, 2015).

Alternatively, when the mutation rates of multiple ΔmutS Pph 1448A strains were measured concurrently, the assay was performed in a 96-well plate. Strains were prepared as previously stated except the final dilution was made in 13 ml of KB for each strain to a final concentration of 5000 CFU/ml (OD₆₀₀ of 0.00001). Each strain had twelve replicate cultures of 1 ml distributed over two sterile deep well 96-well plates. Edge wells were filled with 1 ml of sterile ddH₂O to reduce evaporation effects. Wells which had culture contained a sterile small glass bead to ensure of adequate mixing. Plates were sealed with Micropore™ tape and left shaking at 30°C at 225 rpm over two days. Afterwards, nine out of the twelve cultures from each strain were plated onto KB Strp plates to quantify the number of spontaneous Strp⁺ mutants within each saturated culture. This was done by first transferring cultures into two sterile deep well 96-well plates without beads and followed by centrifuging the plates at 4000 rpm for 10 minutes at 4°C to pellet cultures. Once the cultures were pelleted, 720 µl of supernatant was removed from each well and pellets were re-suspended in remaining culture and plated onto KB Strep plates. The remaining three cultures from each strain were serially diluted and 5 µl spots of each dilution were plated onto KB plates with two technical replicates each to determine the mean N_f. All plates were incubated at 30°C. Plates for estimating N_f were counted one day after being plating. Plates for
quantifying Strep\textsuperscript{6} colonies were first counted after two days post plating and then recounted five days post plating. Waiting an additional three days improved the ability to count small colony variants that were not visible two days post plating. The data was analyzed using the same tools as previously stated. P-values were adjusted using the Benjamini-Hochberg method to account for multiple testing.

2.12 In vitro growth assays

Each strain used was prepared by first inoculating a 5 ml KB Rif or KB Rif Cyc liquid culture from a random population sample. A 500 µl aliquot of overnight culture from each strain was transferred into a fresh 1.5 microfuge tube and was centrifuged at 4600 rpm for 5 minutes. The supernatant was removed and pellets were re-suspended in 1ml of KB Rif. These suspensions were diluted into a 1.5 ml solution of KB Rif to a concentration of 5x10\textsuperscript{7} CFU/ml (OD\textsubscript{600} 0.1). Aliquots of 150 µl of culture were transferred into each well of a COSTAR 96-well microtiter plate with each strain having six replicates randomly arranged on the plate including blank media wells. Edge wells were filled with 200 µl of sterile ddH\textsubscript{2}O to reduce evaporation effects. Once setup, the 96-well plate was placed in the Tecan NanoQuant Infinite M200 PRO without a lid and set to a 23-hour growth cycle. In this growth cycle, the OD\textsubscript{600} readings from each well were taken every 30 minutes. A single plate held a total of six treatments.

The R package Growthcurver was used to determine the area under the growth curves of each treatment by fitting a logistical function to each replicate and calculating the area of that function (Sprouffske and Wagner, 2016). Since individual plates would have slightly different time points at which the carrying capacity was reached, the time point at which the entire plate average reached maximum density was used as the endpoint at which the area under the curve for each replicate was calculated. Additionally, all curves were adjusted by subtracting the average blank media OD\textsubscript{600}.

2.13 Plant growth conditions

All plants were grown in a Conair growth room set at 21°C under a 12-hour photoperiod (100-150 µmol/m\textsuperscript{2}s) at a humidity which ranged between 55-70%. All experiments were done using four to
five-week-old non-bolting A. thaliana Colombia-0 (Col-0) plants unless otherwise indicated. Arabidopsis seeds were surface sterilized prior to plating by first mixing within a 10% bleach and 0.01% Triton X-100 solution. Seeds were either inverted or shaken for 15 minutes and rinsed with sterile ddH₂O at least six times. In general, sterilized seeds were sowed out onto Sunshine mix soil moisturized by water mixed with fertilizer at a concentration of 1g/L. Alternatively, plants were grown in peat pellet Jiffy® pots (42X42) using the following protocol. Pellets were expanded by adding lukewarm water with fertilizer at a concentration of 1g/L. Once the pellets were fully expanded, the mesh was gently opened and the naturally forming hole for a seedling was filled in. Sterilized Arabidopsis seeds were then sowed out onto the surface of the pot.

2.14 In planta growth assay

The following standard growth assay protocol was implemented to measure P. syringae growth in planta on Arabidopsis. Bacterial treatments were streaked from freezer stocks at least three days before the assay but never longer than two weeks before the assay. The day before the growth assay, either a single colony or a random population streak from this plate was re-streaked onto KB plates with the appropriate antibiotics and incubated at 30°C overnight. For each strain, a loop of bacteria was picked from the re-streaked plate using a sterile p200 tip and the bacteria was suspended in 1 ml of 10 mM MgSO₄. Each treatment was then diluted into 20-30ml of 10 mM MgSO₄ usually to a concentration of either 10⁵ CFU/ml (OD₆₀₀ 0.0002) or 10⁶ CFU/ml (OD₆₀₀ 0.002). Inoculum was then pressure infiltrated on the abaxial side of the leaf into the apoplastic using a 1ml needless syringe. The surface of infiltrated leaves were dried with Kim wipes. Four leaves were inoculated for each biological replicate (single plant) unless otherwise indicated. Sampling and quantification of the bacterial loads were usually taken immediately after inoculation (day zero) and then three or four days post inoculation. Once all treatments were completed, plants were returned to the growth room until further sampling. Sampling was done by first removing the four infiltrated leaves with sterile tweezers. A 5mm diameter disc (core)
was taken from each leaf using a #3 core borer (four cores per replicate). Cores were surface sterilized in 70% ethanol for 5-10 seconds and dried on a Kim wipe. Each set of four cores from the same biological replicate were put into a sterile 1.5 ml microfuge tube filled with 200 µl of 10 mM MgSO₄ and a small sterile glass bead. Once all replicates were complete, the microfuge tubes were placed in the Mini-Beadbeater™ and ground for 2 minutes to homogenize the leaf tissue. The homogenate was serially diluted, and 5 µl spots of each dilution were plated on KB plates with the appropriate antibiotics. Plates were incubated for two days at 30 °C and then quantified at the lowest countable dilution. To test for significant differences in growth between treatments, bacterial loads were compared using Student’s t-tests, with Benjamini-Hochberg adjusted P-values to account for multiple testing when necessary.

When optimizing the growth of *Pph* 1448A on *Arabidopsis, Pph* 1448A was inoculated at the following concentrations: $10^2$ CFU/ml (OD₆₀₀ 0.0000002), $10^3$ CFU/ml (OD₆₀₀ 0.000002), $10^4$ CFU/ml (OD₆₀₀ 0.00002) and $10^5$ CFU/ml (OD₆₀₀ 0.0002) following the previously stated methodology. For each inoculum concentration, four biological replicates were sampled each day over five days including day zero. Generations were estimated using the following formula:

**Equation 1:** \[ \text{Generations} = \frac{\log_{10} \left( \frac{N_f}{N_i} \right)}{\log_{10}(2)} \]

where $N_i$ was the mean population size from day zero and $N_f$ is the final cell density of a single replicate.

When testing the effects of leaf material on *Pph* 1448A growth, *Pph* 1448A was inoculated onto *Arabidopsis* with or without leaf material at a concentration of $10^5$ CFU/ml (OD₆₀₀ 0.0002). Inoculum for each replicate with leaf material was prepared as such. Six 5mm cores were sampled from the leaf of a naïve *Arabidopsis* plant and placed into a 1.5 microfuge tube containing 200 µl of 10 mM MgSO₄ and a small glass bead. This was repeated for six leaves (thirty-six cores in total). Tubes were homogenized by the Mini-Beadbeater™ and grinded for 2 minutes, and were merged into a fresh 1.5 ml microfuge tube.
generating ≈ 1200 µl of homogenate. This tube was centrifuged at 300xg for 1 minute and 500 µl of supernatant was extracted and transferred into a fresh 1.5 ml microfuge tube. The extracted supernatant was centrifuged at 7500xg for 1 minute and the supernatant was removed leaving only the leaf pellet in the tube. This pellet was re-suspended in 750 µl of Pph 1448A suspension at a concentration of 10^5 CFU/ml (OD_{600} 0.0002). Each treatment had six replicates which were sampled four days post inoculation. Plants used in this experiment were grown in Jiffy pots.

When testing if growth of Arabidopsis plants in Jiffy pots was negatively affecting the growth of Pph 1448A, a growth assay comparing the growth of Pph 1448A on plants grown in Jiffy pots and in a flat of soil at a starting concentration of 10^5 CFU/ml (OD_{600} 0.0002) was performed. One modification imposed in this assay was that six full leaves were inoculated and the entire leaves were sampled. For each treatment three replicates were sampled immediately after inoculation and eleven were sampled four days post inoculation.

The growth of the twelve LTEE lineages’ day 80 (D80) populations were assayed on Arabidopsis as stated previously with inoculums for each treatment generated using a random population samples at a starting concentration of 10^6 CFU/ml (OD_{600} 0.002). Treatments were sampled four days post inoculation with each treatment having six replicates.

When testing if fungal contamination within the D80 freezer stocks of the LTEE lineages was negatively affecting their growth, an assay comparing the growth of D80 decontaminated isolates was done in parallel with normally prepared D80 populations. Decontaminated D80 treatments were prepared by combining and re-streaking five randomly selected isolates. Both D80 treatments were done using L7 and L12 along with their D0 populations. All treatments were inoculated at 10^6 CFU/ml (OD_{600} 0.002) and were sampled zero and four days post inoculation. For each treatment, three replicates were sampled on day zero and seven were sampled on day four.
2.15 in planta competition assay

The relative *in planta* fitness of the LTEE lineages were measured using co-inoculation competition assays. These were employed by co-inoculating any one of the LTEE lineages with wild-type *Pph 1448A* in *Arabidopsis*. Since all LTEE lineages had an integrated Km$^R$ marker at the *mutS* locus, this made them distinguishable from wild-type *Pph 1448A* through Km selection. The preparation of bacterial treatments was the same as in an “*in planta* growth assay” where population samples were always taken and re-streaked onto KB plates with the appropriate antibiotics the day before inoculation. Afterwards, a loop of bacteria was taken from the fresh streak for each treatment using a sterile p200 tip and suspended in 1 ml of 10 mM MgSO$_4$. The LTEE lineage was first diluted into 40-50 ml of 10 mM MgSO$_4$ to a final concentration of 5X10$^5$ CFU/ml (OD$_{600}$ 0.001). Next, wild-type *Pph 1448A* suspension was diluted into the same inoculum to a final concentration of 5X10$^5$ CFU/ml (OD$_{600}$ 0.001), giving the inoculum a total bacterial concentration of 10$^6$ CFU/ml (OD$_{600}$ 0.002). Treatments were inoculated as in an “*in planta* growth assay” and returned to the growth room once completed. Sampling was done by removing infected leaves with a sterile tweezer, and by taking two 5 mm diameter cores from each leaf (eight cores per plant). Cores from each replicate were placed into a 1.5 ml microfuge tube containing 200 µl of 10 mM MgSO$_4$. Tubes were homogenized as in an “*in planta* growth assay”, serially diluted and plated. For each replicate, 100 µl of the appropriate dilution was spread plated onto a KB Rif Km plate to quantify the LTEE lineage, and another 100 µl on KB Rif plate to quantify the LTEE lineage plus wild-type. The dilution factor plated onto KB Rif plates was twice as large the dilution factor plated onto KB Rif Km plates. Plates were incubated at 30 °C for two days and then quantified. To measure the input proportion of each strain within a co-inoculum, at least three biological replicates were quantified immediately post inoculation. Selection rate constant (r) which measures the natural logarithm difference in growth between competitors was calculated as indicated in Lenski *et al.* 1991. If the ancestral D0 population was not included in an assay, one sample Student’s t-tests were used to test if
the r for the LTEE lineage was significantly different to the expected null of zero, with Benjamini-Hochberg adjusted P-values to account for multiple testing when necessary. If the ancestral D0 strain was included, Student’s t-tests of the r between D0 and D80 were performed.

2.16 Preliminary Pph 1448A serial passaging experiments on Arabidopsis

Starting inoculums were generated using a freshly streaked colony of Pph 1448A and made at a starting concentration of 10^5 CFU/ml (OD_{600} 0.0002). Four leaves per replicate were pressure infiltrated with a 1mL needless syringe as in an “in planta growth assay” with a total of seven independent replicates being inoculated. Once all replicates were inoculated, plants were returned to the growth room. Four days post-inoculation each replicate was extracted by: harvesting twenty-four 5mm leaf cores per replicate (six per leaf), surface sterilizing the cores in 70% ethanol for 5-10 seconds and placing four cores each into a single 1.5 ml microfuge tube 200 µl of 10 mM MgSO4 (six tubes per replicate). Tubes were homogenized as in an “in planta growth assay” followed by merging the homogenate from the six tubes for each replicate into a single fresh 1.5 ml microfuge tube. This resulted in each replicate tube having a total of ≈1200 µl of undiluted homogenate. 100 µl of the undiluted homogenate was sampled, serially diluted and plated as in an “in planta growth assay”. Afterwards, the remaining 1100 µl of undiluted homogenate was centrifuged at 300xg for 1 minute to pull down the large chunks of leaf material, but to leave most of the bacteria in the supernatant. Afterwards, 800 µl of the supernatant was extracted, put in a fresh 1.5 ml microfuge tube and centrifuged at 7500xg for 2 minutes. The supernatant was removed and pellets were re-suspended in 800 µl of 10 mM MgSO4 forming the transfer inoculum. The transfer inoculum was pressure infiltrated into a fresh four-week-old Arabidopsis plant as in an “in planta growth assay”. The passaging was terminated after the second passage (eight days in planta).

In the succeeding trial experiment, a few modifications were introduced to the previous protocol to maximize the amount of Pph 1448A recovered in each passage. This included infiltrating and
sampling six full leaves per replicate. Instead of taking six cores per leaf, an entire leaf was rolled up and placed into 1.5 microfuge tube containing 200 μl of 10 mM MgSO₄ and a sterile glass bead. Another modification was that infected leaf material was not surface sterilized before being homogenized. Additionally, 500 μl of supernatant was extracted instead of 800 μl after the first centrifugation step to eliminate any large chunks of leaf material from being extracted. Lastly, second centrifugation step was eliminated and instead 100 μl of 10 mM MgSO₄ was added to the extracted supernatant. This resulted in each replicate having a total of 600 μl of transfer inoculum. This serially passaging experiment had six independent replicates. The passaging was terminated after the second passage (eight days *in planta*). Plants used in both trial experiments were grown in Jiffy pots. To test for significant differences of population sizes between passages, bacterial loads were compared using a paired Student’s t-tests.

### 2.17 Optimization of recovery step for the LTEE

Four different methods (Table 4) of recovering declining passaging populations of *Pph* 1448A in *Arabidopsis* were tested to identify a method which would effectively recover the most *Pph* 1448A CFUs with the least number of contaminants. For all methods, six leaves per plant were inoculated with *Pph* 1448A as in an “*in planta* growth assay” at a starting concentration of $10^3$ CFU/ml (OD₆₀₀ 0.000002). After two days of *in planta* growth, leaves for methods 1 and 4 were surface sterilized by dipping in 70% ethanol, while treatments 2 and 3 were not surface sterilized. Leaves were homogenized and merged as in the full leaf serial passaging trial assay. Tubes from all treatments were then centrifuged at 300×g for 1 minute to pull down large pieces of leaf material. For methods 1 and 2, the entire supernatant was extracted and transferred to a fresh 1.5 ml microfuge tube. For methods 3 and 4, 500 μl of supernatant was extracted and transferred into a fresh 1.5 ml microfuge tube. All methods centrifuged the supernatant at 7500×g for 2 minutes, discarded supernatants and re-suspended the pellet in 100 μl of 10 mM MgSO₄. Suspensions of methods 1 and 3 were spot plated while methods 2 and 4 were spread
plated. All methods were plated on KB Rif Cyc plates. Plants used in this experiment were grown in Jiffy pots.

Table 4. Conditions of each recovery method.

<table>
<thead>
<tr>
<th>Recovery method</th>
<th>Surface sterilization (Yes/No)</th>
<th>Amount of supernatant extracted</th>
<th>Plating method (spot/spread)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>Entire</td>
<td>Spot</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Entire</td>
<td>Spread</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>500 µl</td>
<td>Spot</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>500 µl</td>
<td>Spread</td>
</tr>
</tbody>
</table>

2.18 The LTEE of twelve barcoded hypermutator lineages of *Pph 1448A* on *Arabidopsis*

In this LTEE, twelve uniquely barcoded Δmuts *Pph 1448A* lineages were independently passaged on *Arabidopsis*. For each lineage, the appropriate strain was isolation streaked from freezer stock onto a plate of KB Rif Km and incubated for two days at 30°C. For each lineage, a single colony was picked and re-streaked onto a plate of KB Rif Km a day prior to the beginning of the serial passaging. To initiate the serial passaging, each lineage was scooped from the streak using a sterile P200 tip and was suspended into 1 ml of 10 mM MgSO₄. Bacterial suspensions were diluted to make inoculums at a concentration of 10⁶ CFU/ml (OD₆₀₀ 0.002). Each lineage was syringe inoculated as in an “*in planta* growth assay” into six leaves of an individual 4-week old *Arabidopsis* plant. Plants used in the passaging were grown in flats with fertilized soil. Immediately after inoculating a lineage, leaves were dried with a Kim wipe and fingers were sterilized by dipping in 70% ethanol to avoid cross contamination. After all lineages were infiltrated, plants were placed back into the growth room. Additionally, the remaining bacteria from the re-streak of each lineage were used to inoculate individual 5 ml KB Rif Km liquid cultures. These liquid cultures were used the following day to make ancestral (D0) freezer stocks for each lineage.

After four days of *in planta* growth each lineage was extracted and transferred to a fresh 4-week old plant as such. The six infected leaves from the infected plant was extracted with sterilized tweezers. Each of the six leaves were rolled and placed into six separate 1.5 ml microfuge tubes containing 200 µl of 10 mM MgSO₄ and a small glass bead. The total mass of the six infected leaves for each lineage was
weighed after blanking the scale with six 1.5 ml microfuge tubes containing 200 µl of 10 mM MgSO₄ and a small bead. Tubes were then homogenized as in an “in planta growth assay”. Afterwards, tubes were given a quick spin to pull down all the homogenate. All six tubes were merged into a fresh 1.5 ml microfuge tube to generate ≈1200 µl of undiluted homogenate per lineage. From this combined undiluted homogenate, 100 µl was sampled, serially diluted and plated as in an “in planta growth assay” onto KB Rif Km Cyc plates. Plates were incubated at 30°C for two days and then quantified at the lowest countable dilution. The remaining homogenate was centrifuged at 300xg for 1 minute. 500 µl of the supernatant was extracted and transferred into a fresh 1.5 microfuge tube. An additional 100 µl was added to the extracted supernatant to generate a total of 600 µl of transfer inoculum. The transfer inoculum was then infiltrated into six leaves of a fresh 4-week old plant. This was followed by drying the leaves with a Kim wipe and sterilizing fingers in 70% ethanol between replicates to avoid cross-contamination. Freshly inoculated plants were then placed back into the growth room until the next passage four days later.

2.18.2 LTEE freezer stock protocol

Freezer stocks for all lineages were prepared after every four passages (sixteen days). Each stock was prepared by first inoculating a 5 ml KB Rif Km Cyc liquid culture with 50 µl of the undiluted homogenate. Cultures were incubated at 30 °C shaking at 225 rpm to saturation over 40-48 hours. Once saturated, each lineage was stocked by taking 900 µl of culture and combining it with 900 µl of 20% glycerol in a 2-ml cryogenic vial and then stored at -80°C.

2.18.3 LTEE lineage recovery from population crashes or extinctions

When a lineage fell below a total population size of 5000 CFUs, the most recent freezer stock of that lineage was streaked on KB Rif Km Cyc plate two days before the subsequent passage. On the day of passaging, the transfer inoculum for the recovering lineage would be made from a random sample of the streaked population and diluted to a final concentration of 10⁶ CFU/ml (OD₆₀₀ 0.002).
2.18.4 Estimation of the initial bacterial loads and generations of lineages during the LTEE 

To estimate the number of *in planta* generations, an estimate of the initial bacterial load immediately after inoculation was needed for each passage. The initial bacterial load was estimated based on the concentration of the undiluted homogenate. The dilution factor of preparing the transfer inoculum was empirically determined by serially diluting and plating a sample of both the prepared transfer inoculum and the undiluted homogenate in passage 6. The average volume of inoculum infiltrated into a full leaf of *Arabidopsis* was empirically determined from day zero densities of a growth assay (data not shown) with the following equation:

**Equation 2:** \[ \text{Infiltration volume} = \frac{D_0 \text{ density}}{\text{Inoculum concentration}} \]

where D0 density is the bacterial load measured from one fully inoculated leaf. The following equation was then used to calculate the expected initial bacterial load:

**Equation 3:** \[ D_0 \text{ Bacterial load} = [\text{Undiluted homogenate}] \times \text{Transfer inoculum dilution factor} \times \text{Infiltration volume} \]

*In planta* generations for each lineage at every passage were then calculated using Eq 1.

The accuracy of our estimations was tested by taking empirical measurements sporadically throughout the serial passaging experiment. When possible, extra transfer inoculum after passaging a lineage was used to inoculate two leaves of an extra 4-week old *Arabidopsis* plant. These leaves where then sampled and quantified identically to the method employed in the LTEE protocol.

2.18.5 Cross-contamination checks of the LTEE lineages 

Throughout the LTEE, cross contamination checks were done by Sanger sequencing the eight nucleotide barcodes of each lineage throughout the passaging experiment. First genomic preps of lineages were taken either from the remaining liquid culture after making freezer stocks, or from overnight cultures inoculated from freezer stocks. The primer pair MutS-ext-F and MutS-ext-R (Table 1)
was used to amplify the entire *mutS* locus from the genomic preps. The PCR products were Sanger sequenced using the primer MutS-2488 or MutS-2120 to sequence the barcode.
Chapter 3: Development of the Long-Term Evolution Experiment

3.1 Fluctuation assays of the Δmuts Pph 1448A strains

Prior to optimizing a protocol for the LTEE of Pph 1448A on Arabidopsis, characterization of the barcoded Δmut Pph 1448A strains was necessary to determine: if the strains were hypermutators, and if the process of cloning had effects on the fitness of the strains. We used fluctuation assays to determine if the Δmut Pph 1448A mutants possessed an elevated mutation rate relative to wild-type. Generally, fluctuation assays are performed by growing multiple parallel cultures of a strain to saturation, followed by plating the cultures on selective media to quantify spontaneous mutants (Rosche and Foster, 2000). Mutation rates are then determined based on the distribution of spontaneous mutants which occur over parallel cultures (Rosche and Foster, 2000). Spontaneous Rif<sup>R</sup> mutants are commonly quantified in studies which measure mutation rate/frequencies since these mutants occur at a relatively high frequency and are well characterized (Oliver <i>et al.</i>, 2002; Leclerc <i>et al.</i>, 1996; Baltz, 2014; Garibyan <i>et al.</i>, 2003). Unfortunately, the ΔmutS Pph 1448A strains developed in this work were already Rif<sup>R</sup>, therefore we elected to select for spontaneous Strp<sup>R</sup> mutants in all fluctuation assays. The first assay performed using 5 ml parallel cultures quantified the relative mutation rates of Δmuts Pph 1448A lineage 1 (L1) in parallel with wild-type Pph 1448A. L1 had a significantly higher mutation rate compared to wild-type (LRT; P<10<sup>-11</sup>) with an approximate 10-fold increase (Figure 2). A second fluctuation was performed with nine randomly chosen Δmuts Pph 1448A stains in a 96-well setup. In this second assay, all tested Δmuts Pph 1448A strains had a significantly higher mutation rate compared to wild-type (LRT; 0.0001<P<0.01). Mutants showed anywhere between a 6-fold to 13-fold increase in mutation rate (Figure 3).
Figure 2. Culture tube fluctuation assay comparing the mutation rate of Δmuts Pph 1448A to wild-type. The Δmuts Pph 1448A (L1) has a significantly higher mutation rate compared to the wt (LRT; **** P<0.0001). Relative mutations rates were quantified using a fluctuation assay (n=12) selecting for Strp^R and calculated using the MSS maximum likelihood method. Error bars represent 95% confidence intervals.

Figure 3. 96-well plate fluctuation assay comparing the mutation rates of nine strains of Δmuts Pph 1448A to wild-type. All Δmuts Pph 1448A mutants tested display a significantly higher mutation rate compared to the wild-type after using the Benjamini-Hochberg adjustment (LRT; * P<0.05, ** P<0.01 and *** P<0.001). Relative mutations rates were quantified using a fluctuation assay (n=9) selecting for Strp^R and calculated using the MSS maximum likelihood method. Error bars represent 95% confidence intervals.
3.2 In vitro growth assays of the twelve ancestral Δmuts Pph 1448A lineages

*In vitro* growth assays in KB were performed on the ancestral strains of the twelve LTEE lineages to determine if the cloning procedure affected the fitness of any lineage prior to the LTEE on *Arabidopsis*. Each 96-well plate grew a total of four lineages in parallel with wild-type *Pph* 1448A, and a control ΔmutS *Pph* 1448A strain which was not used in the LTEE on *Arabidopsis* known as a representative ancestor (RA). The RA was also used as a surrogate common ancestor for all evolved lineages in future *in planta* growth assays. There were no distinguishable differences in the growth dynamics among the Δmuts strains compared to wild-type on any plate, although there was a large amount of variation between replicates as seen by the large error bars (Figure 4). To quantitatively assess the growth phenotype of each lineage, the holistic metric of area under the growth curve was used. This unit integrates important growth metrics including: starting density, growth rate and carrying capacity into one value (Sprouffske and Wagner, 2016). Within each plate, every strain was compared to each other using an ANOVA. The area under the curves of all treatments on the first plate showed that there was a significant difference in mean areas between strains (ANOVA; P=0.04), with post hoc analysis identifying that L4 had a significantly larger area than L2 (Figure 5A) (Tukey test; P<0.05). The mean areas of strains on the second plate did not differ significantly among each other (ANOVA; P=0.93) (Figure 5B). Strains on the final plate did display a significant difference in mean areas (ANOVA; P=0.0035) with post hoc analysis identifying that both L11 (Tukey test; P<0.01) and L12 (Tukey test; P<0.01) mean areas being significantly lower than wild-type (Figure 5C) signifying a potential reduction in fitness of these lineages because of the cloning process. To test for plate effects, ANOVAs of wild-type and RA replicates between plates were performed. Both wild-type (ANOVA; P<10⁻⁵) and RA (ANOVA; P<10⁻⁵) displayed significant differences in mean areas between plates, with post hoc analysis identifying plate two and three having significantly larger areas than plate one for both wild-type (Tukey test;
P<0.01 & P<0.01) and RA (Tukey test; P<0.01 & P<0.01). This indicated that plate effects were significantly enhancing the growth of strains on plates two and three relative to plate one.
Figure 4. *In vitro* growth curves of the ancestors of the twelve Δmuts *Pph* 1448A lineages used in the LTEE. Lineages were grown in KB media in parallel with the RA and wild-type controls (n=6). Each letter (A, B and C) represents a single plate on which the listed treatments were grown on. Error bars represent SEM.
Figure 5. Boxplots of area under the growth curves of the ancestors of the twelve Δmuts Pph 1448A lineages used in the LTEE. Areas were derived from the *in vitro* growth assays performed in KB. Mean areas of strains on each plate were compared against each other using ANOVAs followed with post hoc tests if significantly different (Tukey test; * P<0.05 & ** P<0.01). Each letter (A, B and C) represents a single plate on which the listed treatments were grown on.
3.3 Optimization of *Pph 1448A* growth on *Arabidopsis*

Before a protocol for performing the LTEE of *Pph 1448A* on *Arabidopsis* was initiated, a series of preliminary assays were done to identify the conditions at which *Pph 1448A* optimally grows on *Arabidopsis*. In this first assay, we wanted to determine at which inoculum concentration and time of extraction *Pph 1448A* underwent the highest number of generations on *Arabidopsis*. A standard growth assay was done where *Pph 1448A* was inoculated at the following concentrations: $10^2$ CFU/ml, $10^3$ CFU/ml, $10^4$ CFU/ml and $10^5$ CFU/ml. Replicates inoculated at each of these concentrations were sampled every day over five days (including day zero). Replicates inoculated at $10^2$ CFU/ml and $10^3$ CFU/ml were highly variable and in many cases, did not recover any bacteria so these treatments were excluded from future analyses. Over five days, at starting inoculums of $10^4$ CFU/ml and $10^5$ CFU/ml, *Pph 1448A* showed a steady increase in population size over four days followed by a slight drop after five days *in planta* (Figure 6). This similar trend was also observed when estimating the total number of *in planta* generations based on initial and final populations sizes (Eq1) (Figure 7). After four days of *in planta* growth at a starting inoculum of $10^5$ CFU/ml, *Pph 1448A* underwent the highest number of *in planta* generations at approximately 4.5 (Figure 7).

![Figure 6. Time series growth assay of *Pph 1448A* on *Arabidopsis*.](image)

The growth of *Pph 1448A* on *Arabidopsis* was measured over five days at starting inoculums of $10^4$ CFU/ml and $10^5$ CFU/ml. Four replicates at each concentration were sampled per day (n=4) over five days. Error bars represent SEM.
3.4 First preliminary serial passaging experiment of \textit{Pph 1448A} on \textit{Arabidopsis}

Once the optimum day of extraction and inoculum concentration were determined, a preliminary serial passaging experiment of \textit{Pph 1448A} on \textit{Arabidopsis} was tested. The goal of this assay was to maintain direct plant to plant passaging of seven independent \textit{Pph 1448A} populations. The methods used in this assay were determined based on the population densities obtained from the previous growth assay. Therefore, each lineage was inoculated at a starting concentration of $10^5$ CFU/ml and lineages were extracted and transferred to fresh plants after four days of \textit{in planta} growth. Over the course of the first trial assay which was carried over eight days (two passages), all seven lineages decreased significantly on day eight (Paired Student’s t-test; $P< 10^{-4}$) by about 1 log compared to day four population sizes (Figure 8). One hypothesis for the decrease in population size over each progressive passage was that the pieces of plant material within the transfer inoculum contained endogenous immune elicitors or Damage Associated Molecular Patterns (DAMPs), which can further contribute to the induction of the plant immune response and decrease bacterial growth (Boller and Felix, 2009; Benedetti \textit{et al.}, 2015). Although the large pieces of plant material were removed from the
transfer inoculum in the centrifugation step, refined pieces of leaf material were still present in the transfer inoculum and it is possible that this material possessed DAMPs.

3.5 The effects of plant material on Pph 1448A growth

To test exclusively if the presence of leaf material in the transfer inoculum was influencing the growth of infiltrated Pph 1448A, a standard growth assay was performed on Arabidopsis using Pph 1448A inoculums with or without leaf material present. Leaf material was generated as it was within a passaging experiment, to replicate the amount of material present in the transfer inoculum. Leaf material was then mixed in a suspension of Pph 1448A at a concentration of $10^5$ CFU/ml therefore controlling the starting inoculums for both treatments. The presence of plant material in the inoculum did not significantly change the growth of Pph 1448A compared to replicates infiltrated with normal Pph 1448A (Figure 9).

![Graph](image_url)

**Figure 8. First preliminary serial passaging experiment of Pph 1448A on Arabidopsis.** Seven independent lineages (n=7) of Pph 1448A were passaged on Arabidopsis every four days from an initial inoculum of $10^5$ CFU/ml. All lineages showed a significant decline in population size by about 1 log after one passage (Paired t-test; P<10^{-4}).
Another hypothesis to explain the decline seen in the passaging populations of Pph 1448A was that the preparation of the transfer inoculum itself diluted the inoculum below a concentration of $10^5$ CFU/ml. As a result, the next set of plants were inoculated with a lower concentration of Pph 1448A and this resulted in lower final population sizes as evident in the optimization growth assay (Figure 6).

Therefore, the modified passaging protocol attempted to maximize the concentration of Pph 1448A within the transfer inoculum by no longer surface sterilizing infected leaf material, and by extracting Pph 1448A from six full leaves. Even after these modification to the passaging protocol, all lineages significantly decreased in population size by about 1.6 logs after eight days (Paired Student’s t-test; $P=0.022$) (Figure 10), with one going extinct following the second passage (Figure 10). Despite maximizing the amount of Pph 1448A being extracted per passage, the populations of Pph 1448A continued to decline over each passage indicating that the transfer inoculums were still too dilute.
The decline of passaging populations signified that the growth of Pph 1448A on Arabidopsis over the four days between passages was insufficient (Figure 8 & 10) and lower compared to the original optimization growth assay (Figure 6). In both serial passaging assays, the plants used were grown in Jiffy pots while plants used in the optimization assay which experienced higher bacterial growth were grown in flats of soil. Jiffy pot grown plants were originally planned to be used in the LTEE because they offered versatility when infiltrating and allowed for the appropriate spacing between lineages to avoid cross contamination. Despite these benefits, it appeared that the growth of Pph 1448A from plants grown in Jiffy pots was less compared to growth assays from plants grown in flats. Furthermore, we tested if the deficient in planta growth of Pph 1448A on Arabidopsis observed during both trial serial passaging experiments was because the plants used in these assays were grown in Jiffy pots. This was tested by performing a growth assay where Pph 1448A was inoculated to a set of 4-week old Arabidopsis plants grown in Jiffy pots and in a soil flat in parallel. Although Pph 1448A did on average reach a higher density on plants grown in flats (Figure 11), this difference was not significant (Student’s t-test; P=0.13).

### Figure 10. Full leaf preliminary serial passaging experiment of Pph 1448A on Arabidopsis

Six independent lineages (n=6) of Pph 1448A were passaged on Arabidopsis every four days from an initial inoculum of $10^5$ CFU/ml. CFU/cm² were predicted by assuming a full leaf was equivalent to six 5mm diameter cores. All lineages showed a significant decline in population size by an average of 1.6 logs after one passage with one lineage going extinct by day eight (Paired t-test; P=0.022).
Regardless, we decided to carry out the LTEE in plants grown in flats to gain any potential enhancement in growth.

\[ \text{Figure 11. The growth of Pph 1448A on Arabidopsis plants grown in Jiffy pots or in a flat.} \]

Both plants grown in a flat and Jiffy pots were infiltrated at a starting inoculum of $10^5$ CFU/ml. For each treatment, three replicates (n=3) were measured on day zero while eleven replicates (n=11) were measured on day four. Pph 1448A did grow to a slightly higher in flats compared to Jiffy pots but not significantly so (Student’s t-test; $P=0.13$). CFU/cm$^2$ were predicted by assuming a full leaf was equivalent to six 5mm diameter cores. Error bars represent SEM.

### 3.8 Optimization of recovery step for the LTEE

Due to the inability to sustain continuous plant to plant transfers of Pph 1448A on Arabidopsis, we designed and optimized a recovery step where declining populations of Pph 1448A could be recovered by plating the them onto KB. The populations could then be extracted from plates and be infiltrated back into Arabidopsis for the next passage. Since the developed method of passaging was through direct plant to plant transfers, the concentration of the transfer inoculum was not known until two days afterwards once the quantification plates were countable. Therefore, when a lineage fell below a critical population size, the lineage would be extracted that same day from the plant (two days before the next passage) and plated onto KB Rif Cyc for two days. Four different methods of recovering Pph 1448A from Arabidopsis were tested (Table 4). Replicates for all methods tested were inoculated at a low concentration ($10^3$ CFU/ml), to reproduce the condition as if the populations were crashing. The
populations were extracted two days post inoculation as would be in a recovery scenario. These four recovery methods varied: in surface sterilization, the amount of supernatant extracted from the centrifuged homogenate, and the method of plating (Table 4). Overall, method 2 recovered the most Pph 1448A with relatively few bacterial contaminants present (Figure 12).

![Figure 12. The testing of four recovery methods for the LTEE of Pph 1448A on Arabidopsis. The average number of CFUs of Pph 1448A and contaminant colonies extracted from each recovery treatment (n=4) after two days of in planta after being inoculated at 10^3 CFU/ml. Error bars represent SEM.](image)

3.9 Discussion

The objectives prior to initiating the LTEE, were to characterize the ΔmutS Pph 1448A strains and to optimize a protocol for passaging Pph 1448A on Arabidopsis. Fluctuation assays were used to quantify the mutation rates of the ΔmutS Pph 1448A relative to wild-type. Although in both assays all ΔmutS strains tested showed a significant elevation in mutation rate (Figure 2 & 3), this elevation ranged between 6-13 fold, which was smaller than the 100-1000 fold increases reported in other MMR deficient bacteria (Oliver and Mena, 2010). One potential reason for this discrepancy is that the fluctuation assays performed in this work only measured the relative mutation rates of loci responsible for streptomycin
resistance and not the average genomic rate. The most accurate method to measuring mutation rates are through MA experiments coupled with WGS. MA experiments entail propagating multiple lineages through repeated single cell bottlenecks which greatly reduces the power of selection and allows for the accumulation of most mutations except for lethal ones. MA studies using ΔmutS strains of Pseudomonas fluorescens and P. aeruginosa have reported genome wide mutation rates of 2.51x10^-8 and 2.44x10^-8 (/bp/generation) respectively (Long et al., 2014; Dettman et al., 2016). Compared to wild-type P. aeruginosa, this mutation rate was about 267 fold higher, and this magnitude was similar to what has been measured from MA experiments with E.coli (Lee et al., 2012). Therefore, it is likely that the ΔmutS Pph 1448A strains developed in this work have a similar elevation in mutation rate.

After fluctuation assays confirmed the hypermutator phenotype of the ΔmutS Pph 1448A strains, in vitro growth assays were used to determine if the cloning process resulted in any of these strains having reduced fitness. From these assays, it appeared that L4 grew significantly better than L2 (Figure 5A) and that wild-type grew significantly better than L11 and L12 (Figure 5C). Although these results indicate a possible decrease in fitness of L11 and L12 relative to wild-type, we hypothesize that this observed difference is more likely the result of random plate effects caused by evaporation. A large amount of variation existed between replicates on the same plate indicating that certain parts of the plate allowed for better growth than others. Wild-type showed the largest difference in growth relative to RA on the third plate (Figure 5) indicating the growth of wild-type on this plate might have been enhanced by the position of wild-type wells. Although treatments were randomly assigned wells on each plate, it is possible that by chance wild-type replicates were arranged into more wells which permitted for better growth on the third plate. In the future, repeating these experiments with a transparent cover on the plate may reduce the evaporation effects of edge wells and reduce the amount of noise between replicates. Alternatively, the use of co-inoculation competition assays would better characterize
differences in fitness as it reduces the noise introduced from plate effects since both competitors share the same wells.

Despite designing a plant to plant passaging protocol which maximized the amount of *Pph* 1448A extracted per passage, populations continued to decline after each passage (Figure 10). Neither the presences of leaf material (Figure 9) nor the medium which the plants were grown in significantly affected the growth of *Pph* 1448A (Figure 11) signifying that the low growth of *Pph* 1448A on *Arabidopsis*, coupled with the dilution caused by preparing the transfer inoculum was the cause for the decline in populations. Therefore, we optimized a plating step where declining populations could be recovered on KB midway through a four-day passage, and re-infiltrated back into a fresh plant for the subsequent passage. We planned to implement this step when the mean concentration of all lineages fell below $10^4$ CFU/ml. Increasing the number of direct plant to plant passages was preferable as this would increase the number of generations lineages experience *in planta* relative to on plate. Therefore, in the LTEE we decided to use a starting inoculum of $10^6$ CFU/ml instead of $10^5$ CFU/ml which was used in both preliminary experiments, because we hypothesized that this would increase the number of sustainable direct plant to plant transfers done before populations reached the threshold at which they needed to be recovered on KB plates.
Chapter 4: Long-Term Evolution of Pph 1448A on Arabidopsis

4.1 The LTEE of twelve barcoded hypermutator lineages of Pph 1448A on Arabidopsis

Twelve independently and uniquely barcoded lineages of Δmuts Pph 1448A were serially passaged on the non-host Arabidopsis successfully for eighty days. Each independent lineage possessed a unique barcode and was passaged on its own plant (Figure 13). The LTEE was initiated by inoculating each lineage into six full leaves of a 4-week old Arabidopsis plant at a starting concentration of $10^6$ CFU/ml. After four days of *in planta* growth, the bacterial populations from each lineage was extracted from the previously infected leaf tissue, and re-infiltrated onto a fresh set of 4-week old Arabidopsis plants (Figure 13). After every four passages, 50 μl of the undiluted homogenate from each lineage was used to inoculate a liquid culture to generate freezer stocks (Figure 13). The bacterial loads of each lineage were sampled and quantified during each transfer by serially diluting and plating the undiluted homogenate (Figure 14). Despite the decreases in population size of all lineages throughout the two trial passaging experiments (Figure 8 & 10), this universal decline among lineages over each passage did not occur within the eighty days of experimental evolution (Figure 14). As a result, the optimized plate recovery step which was to be implemented when the mean concentration of the undiluted homogenate among all lineages fell below a concentration of $10^4$ CFU/ml was never utilized. Alternatively, when individual lineages declined bellow a critical size of 5000 CFUs, they were re-initiated from the most recent freezer stocks. Throughout the eighty days of passaging, five lineages experienced crashes or extinction events, and were recovered from freezer stocks. L8-L11 were lost on passage 16 (day 64), and re-established from the day 48 freezer stocks on passage 17 (day 68). As a result, these lineages lost five passages in comparison with lineages that did not crash. L12 was lost on passage 18 (day 72) and re-established from day 64 freezer stocks on passage 19 (day 76), and resulted in the loss of three passages.
Figure 13. An overview of the experimental evolution procedure of *Pph* 1448A on *Arabidopsis*. In summary, twelve uniquely barcoded Δmuts *Pph* 1448A lineages were independently serially passaged on the non-host *Arabidopsis* over eighty days (twenty transfers). Passages were done after four days of *in planta* growth, where infected plant tissue was first collected and homogenized. This was followed by a slow centrifugation step which pulled down large chunks of plant material but left *P. syringae* in the supernatant. The supernatant was extracted and then infiltrated into a fresh 4-week old *Arabidopsis* plant.
4.2 Estimation of the initial bacterial loads of the LTEE lineages

Throughout the LTEE, only the final bacterial loads at the end of each passage were sampled and quantified (Figure 14). To estimate the amount of in planta growth experienced through the duration of each passage, an estimate of the initial population size at the beginning of each transfer was required. Since all lineages experienced the same method of extraction, centrifugation and inoculation at each passage, every passage should theoretically have the same dilution factor between the undiluted homogenate to the initial bacterial load. Therefore, the initial bacterial loads of lineages from each passage were estimated by multiplying the final bacterial densities by the dilution factor of preparing the transfer inoculum, and with the volume of transfer inoculum infiltrated into six full Arabidopsis leaves. The dilution factor of forming the transfer inoculum was empirically measured by plating a sample of transfer inoculum along with the undiluted homogenate of each lineage during passage 6 (Figure 15). The average dilution factor between the undiluted homogenate and transfer inoculum was ≈0.77 (Figure 15). An estimate of the volume of inoculum infiltrated into six full leaves of a 4-week old Arabidopsis plant was obtained from day zero values of a growth assay using Eq 2 (data not shown). The expected initial bacterial densities were then determined using Eq 3.

The accuracy of our initial bacterial load estimates was tested by sporadically measuring initial bacterial loads throughout the LTEE. When possible, extra transfer inoculum after passaging a lineage was used to inoculate and sample two leaves on an extra 4-week old plant. This was done for L1-L3 on passage 9, L1, L3 & L4 on passage 12, and L2 & L3 on passage 21. The observed initial bacterial loads did not significantly differ from the expected values calculated from the empirically determined dilution factor (Paired Student’s t-test; P=0.80) (Figure 16 A). Additionally, when comparing individually measured initial bacterial loads to the expected values, exactly half of the expected values were over estimates while the other half were under estimates (Figure 16 B) demonstrating that our estimations were not biased in one direction.
Figure 14. The final bacterial load of the twelve independent ΔmutS Pph 1448A lineages at the end of each passage on Arabidopsis during the first eighty days of the LTEE. Each lineage was initially inoculated at a starting concentration of $10^6$ CFU/ml.

Figure 15. The concentrations of the twelve LTEE lineages before and after processing the undiluted homogenate into transfer inoculum. Both the undiluted homogenate and transfer inoculum of all lineages (n=12) were quantified at passage 6. The calculated dilution factor between the undiluted homogenate and transfer inoculum was $\approx 0.77$. Error bars represent SEM.
Figure 16. Comparison of expected initial bacterial loads to the observed initial bacterial loads. 

A) The mean initial \textit{Pph} 1448A loads measured empirically (observed) and compared to the expected values using the empirically determined dilution factor (n=8). The observed bacterial loads did not significantly differ from the estimates (Paired Student’s t-test; P=0.80). Error bars represent SEM. 

B) Comparisons of each empirically determined initial load to the expected loads.

4.3 Growth dynamics of lineages over the eighty days of LTEE

With accurate estimations of the initial bacterial loads, the growth of lineages throughout the LTEE could be more accurately depicted (Figure 17). These estimations were used to calculate the total number of \textit{in planta} generations of each lineage using Eq 1 (Figure 18). Lineages which did not experience population crashes through the eighty days of LTEE (L1-L7) had each undergone just over 100 generations on \textit{Arabidopsis} (Figure 18). Lineages which crashed and were recovered lost between three to five passages within the eighty days of passaging and experienced between 75-82 generations (Figure
Throughout the eighty days of the LTEE, lineages averaged \( \approx 5 \) generations per passage (\( \bar{x} = 5.13; \ SEM = 0.122 \)). Passages 8, 12 and 15 represented significantly low outliers as the mean number of generations in these passages were greater than 1.5 interquartile ranges (IQRs) below the first quartile of all mean generations per passage (Figure 19). This universal decline in growth among lineages is likely the result of some flat effect. One possibility is that plants on these flats exhibited a stronger immune response which reduced pathogen growth. Additionally, multiple outliers of growth occurred among lineages within passages throughout the LTEE (Figure 19), potentially indicating varying strengths of immunity of plants within a flat. Overall, the large amount of variation present in both the number of generations undergone between lineages and between passages made tracking actual adaptive increases in growth throughout the LTEE difficult.
Figure 17. The population growth of each lineage over eighty days of experimental evolution on *Arabidopsis*. The initial population sizes of each passage were estimated using the final bacterial load of the previous passage and the empirically determined dilution factor of the passaging process. Black lines represent passages which were captured in the freezer stocks of the evolving lineages. Red lines represent passages lost due to population crashes. L8-L11 crashed on D64, and were re-initiated from the most recent freezer stock which was D48. L12 crashed on day 76 and was re-initiated from D64 freezer stocks. Red circle indicates a severe bottleneck L12 underwent on passage 16.
Figure 18. The total number of \textit{in planta} generations each lineage had undergone on \textit{Arabidopsis} over eighty days of passaging. L8-L12 have experienced fewer \textit{in planta} generations over the eighty days of passaging because of the lineage crash events.

Figure 19. Boxplots of the number of \textit{in planta} generations undergone by all lineages at each passage. Asterisk (*) indicate outlier passages (1.5 IQRs outside of the first or third quartile). Passages 8, 12 and 15 were significantly low outliers. Open circles represent outlier lineages within a passage (1.5 IQRs outside of the first or third quartile).
4.4 Lineages crashed after eighty days during the LTEE

Although lineages maintained stable population sizes on *Arabidopsis* throughout the first eighty days of the LTEE (Figure 14), all lineages declined in population size from day eighty to the end of the experiment at day ninety-six. Further, two lineages (L2 and L6) went extinct by day ninety-six (Figure 20). In a sliding window of the mean number of generations per passage, all significantly low outliers (greater than 1.5 IQRs below the first quartile) were of windows which included passages after day eighty (Figure 21). This universal decline in population sizes among lineages was akin to what occurred in both trial passaging assays (Figure 8 & 10), and resulted in the termination of the experiment at day ninety-six. Interestingly, this decline in *Pph* 1448A growth occurred when the LTEE was being performed in mid-April, which coincided with an outdoor climate that was transitioning from winter to spring. We hypothesized that the universal decline of all lineages after eighty days of LTEE was the result of the changing environmental conditions.

![Figure 20](image_url)

**Figure 20.** The final bacterial load of the twelve independent ΔmutS Pph 1448A lineages after eighty days of passage. All lineages display a downward trend in population size over time.
Figure 21. A sliding window of the mean number of generations per passage of L1-L7. Each window consisting of four passages. Only L1-L7 were included because these lineages were not interrupted by population crashes and population recoveries. The dotted red line indicates the point at which the windows to the right of it contain passages which surpass day eighty of the LTEE. Asterisk (*) indicate outlier windows (1.5 IQRs outside of the first or third quartile). All windows which contained passages after eighty days of LTEE were significantly low outliers. Error bars above represent SEM.

4.5 Changes in growth room conditions after eighty days of the LTEE

Throughout the LTEE, both temperature and humidity were sporadically measured in the growth room to keep track of shifts in environmental conditions. The mean of temperatures measured after day eighty were significantly higher (Student’s t-test; P<0.001) compared to the mean of temperatures measured during the first eighty days (Figure 22A). Additionally, the mean of humidity measurements taken after eighty days of LTEE were significantly lower (Student’s t-test; P<0.0001) compared to the mean of measurements taken during the first eighty days (Figure 22B). Overall, both temperature and humidity of the growth room had significantly changed after eighty days of the LTEE.

These changes in growth conditions after eighty days of LTEE appeared to increase the overall size of plants used. The mass of the six infected leaves per lineage were measured prior to being homogenized during each passage starting at passage 6 (day twenty-four) of the LTEE. The average mass of infected leaves per lineage of each passage was significantly larger (Student’s t-test; P<0.001) after
day eighty compared to the average mass during the first eighty days (Figure 23). We hypothesize that the changes in the environmental conditions which occurred after eighty days of LTEE resulted in the growth of healthier plants which imposed a stronger immune response and negatively affected the growth of the passaging lineages.
Figure 22. Temperature and humidity of the growth room throughout the LTEE A) Comparison of the mean temperature of the growth room measured for the first eighty days of the LTEE (n=14) compared to mean temperature measured after eighty days (n=5). The temperature of the growth room was significantly higher after eighty days of the LTEE (Student’s t-test; *** P<0.001). B) Comparison of the mean humidity of the growth room measured for the first eighty days of the LTEE (n=14) compared to mean humidity measured after eighty days (n=5). The temperature of the growth room was significantly higher after eighty days of the LTEE (Student’s t-test; **** P<0.0001 & *** P<0.001). Error bars represent SEM.
Figure 23. The total mass of infected leaves per lineage before and after eighty days of passaging. The average infected leave mass per lineage of each passage during the LTEE up to day eighty (n=15) compared to average mass after day eighty (n=4). The mass of infected leaves per lineage from plants after day eighty was significantly higher (Student’s t-test; *** P<0.001) compared to plants during the first eighty days of the LTEE. Error bars above represent SEM.

4.6 Cross-contamination checks of the LTEE lineages

To check for cross contamination, the unique eight nucleotide barcodes of each lineage were Sanger sequenced from genomic preps all lineages after: sixteen, thirty-two and eighty days of passaging. The average Phred (quality) scores of each nucleotide within the barcode was calculated and compared to the average Phred scores of the conserved sequence flanking the barcode (four nucleotides upstream and four nucleotides downstream of the barcode) within each sequence trace (Figure 24). If lineages had experienced cross-contamination, we would expect low Phred scores relative to the flanking conserved region. Only L1 from D32 (L1D32) had a barcode quality that was significantly lower than the conserved region (Student’s t-test; P=0.0354) potentially signifying cross-contamination. Conversely, L3D16 (Student’s t-test; P<0.001), L9D16 (Student’s t-test; P=0.0116) and L9D80 (Student’s t-test; P=0.0285) had barcodes with mean quality scores that were significantly higher than the mean quality scores of the conserved regions. In all sequenced samples, the correct nucleotides were called in both the conserved and barcode regions.
Figure 24. The mean Phred quality scores per base pair of each lineage’s eight nucleotide barcode over the LTEE. Barcodes of each lineage was Sanger sequenced from isolates taken after: sixteen (D16), thirty-two (D32) and eighty days (D80) of passaging. Average Phred scores of the barcodes were compared to the average Phred scores of the conserved sequence flanking the barcodes (four nucleotides upstream and four nucleotides downstream of the barcode) (Student’s t-test; *** P<0.001, ** P<0.01 & *P<0.05). The mean quality score of the barcode L1D32 was significantly less than the conserved region. Barcodes of L3D16, L9D16 and L9D80 had significantly higher quality scores compared to the conserved region. Error bars above represent SEM.
4.7 Discussion

In this evolution experiment, twelve independent lineages of Δmuts Pph 1448A were successfully serially passaged on the non-host Arabidopsis over eighty days. Contrary to the results of the preliminary serial passaging experiments (Figure 8 & 10), lineages in the LTEE maintained stable population sizes (Figure 14) and did not require the use of a plating step to recover populations. This was surprising since the only changes between the second preliminary passaging experiment and the LTEE were: a starting inoculum concentration of $10^6$ CFU/ml instead of $10^5$ CFU/ml, plants used were grown in flats instead of Jiffy pots, and that the Pph 1448A strains used were hypermutators. It is possible that a combination of these changes enhanced the growth of lineages on Arabidopsis and allowed for direct plant to plant transfers to be maintained. Alternatively, it is possible that differences in external conditions throughout the LTEE may have enhanced the growth of Pph 1448A which will be further discussed.

Although most lineages maintained stable population sizes over the course of the first eighty days of passaging, L8-L12 each experienced population crashes during this time. These crashes were probably initially driven by experimental variation resulting in these lineages having lower population sizes relative to the other seven lineages. Interestingly, L8-L12 crashed on either passage 16 or 17, which was just before a significantly low growing passage (Figure 19). Prior to this low growing passage (passage 15), L8-L12 had the lowest population sizes, suggesting that the further decline of these populations caused by this passage resulted in these lineages to eventually crash. Unfortunately, it remains unclear as to what factors caused specific passages to have significantly less growth relative to most passages during the first eighty days of the LTEE.

After the eighty days of passaging, all lineages began to decline over each subsequent passage (Figure 20), and this resulted in the termination of the LTEE after ninety-six days. This general decline in growth coincided with significant changes in both temperature and humidity within the growth room.
These changes in growth room conditions were likely the reason plants after day eighty of the LTEE grew significantly larger (Figure 23). Therefore, plants throughout the first eighty days of the LTEE were smaller, indicating their growth was stunted potentially due to abiotic stresses (Tuteja, 2007). The production of the phytohormone Abscisic acid (ABA) plays a vital role in helping plants combat abiotic stresses including osmotic stress (Tuteja, 2007). Consequently, the increased production of ABA has shown to dampen SA immune signaling and enhance the growth of infecting hemi-biotrophic pathogens including *P. syringae* (Fan *et al.*, 2009). Therefore, it is possible that a dampened immune response due higher levels of osmotic stress induced in the first eighty days of the LTEE enhanced the growth of lineages, and permitted for twenty rounds of direct plant to plant transfers. When conditions for the plants became more optimal in the spring, plants induced a stronger immune response reducing the growth of lineages, and causing lineages to crash as they did in both trial assays.

Cross contamination checks were done throughout the LTEE by Sanger sequencing the unique barcodes of each lineage within the *mutS* locus. The mean Phred scores of the barcodes were compared with the mean Phred scores of the conserved flanking regions with the expectation that if cross-contamination had occurred, the quality scores of the barcode would be significantly lower relative to the conserved regions. Only L1D32 displayed this signature (Figure 24), but it seems unlikely that this was the result of cross-contamination. First, the mean Phred score of the barcode from L1D32 was 54 which represents a high confident base call accuracy of 99.9996% (Ewing and Green, 1998). Additionally, if this was the result of cross-contamination we would expect to see a similar signature in the D80 time point, which was not the case (Figure 24). Overall, among all sequenced samples, there were no clear signatures of cross-contamination between lineages indicating that the methods employed in the LTEE avoided this issue.
Chapter 5: Phenotypic Characterization of the LTEE Lineages

5.1 Growth assays of the twelve LTEE lineages after eighty days of passaging

After serially passaging the twelve lineages of \textit{Pph} 1448A for eighty days on \textit{Arabidopsis}, we used both growth assays and \textit{in planta} competition assays to determine if any of the lineages had significantly adapted to grow better on the non-host. It is important to note that because of population crashes, L8-L12 did not actually experience eighty days of passaging during the LTEE. L8-L11 lost five passages and therefore the latest freezer stocks represented sixty days of passaging (D60) on \textit{Arabidopsis}. L12 lost three passages and therefore its latest freezer stock represented sixty-eight days of passaging (D68) on \textit{Arabidopsis}. For simplicity, the latest freezer stocks for all evolved populations will be referred to as D80 for the remainder of the thesis. A growth assay of the D80 populations from the twelve LTEE lineages were performed on \textit{Arabidopsis} to characterize potential increases in growth compared to both wild-type \textit{Pph} 1448A and the surrogate ancestral \textit{ΔmutS Pph} 1448A strain RA. After four days of \textit{in planta} growth L3 (Student’s t-test; \textit{P}=0.019), L6 (Student’s t-test; \textit{P}=0.019) L7 (Student’s t-test; \textit{P}=0.017) and L10 (Student’s t-test; \textit{P}=0.025) all grew significantly higher than RA (Figure 25). This significant increase in growth relative to the surrogate ancestor signified potential adaptation of these lineages on \textit{Arabidopsis}, with L7 showing the largest growth advantage of \textit{≈}0.8 logs (Figure 25). Contrary to this, none of the lineages grew significantly more than wild-type \textit{Pph} 1448A (Figure 25). The Benjamini Hochberg adjustment was applied to \textit{P}-values for comparisons to both RA and wild-type to account for multiple testing.

To test if these improvements in growth relative to RA seen in the first growth assay were the result of adaptations occurring during the LTEE, a second four-day growth assay was done where the D80 populations from all lineages were grown in parallel with their respective true ancestor (D0) and with the RA. This growth assay was split into two blocks for feasibility. None of the ancestors from the twelve lineages grew significantly different among each other (ANOVA; \textit{P}=0.52) indicating that there...
were no inherent differences in fitness between lineages prior to the LTEE on *Arabidopsis* (Figure 26 A & B). None of the lineages’ D80 populations grew significantly better than the RA or to their respective ancestors (Figure 26 A & B). Conversely, the D80 populations of L11 (Student’s t-test; P=0.012) and L12 (Student’s t-test; P=0.012) grew significantly less than RA after the Benjamini-Hochberg adjustment. Additionally, the D80 populations of: L3 (Student’s t-test; P=0.039), L8 (Student’s t-test; P=0.03), L11 (Student’s t-test; P=0.015), and L12 (Student’s t-test; P=0.0004) grew significantly less than their respective ancestors. From the following growth assay, it appeared that none of the lineages had adapted to grow better after eighty days of passaging and that some lineages had reduced fitness on *Arabidopsis*.

![Figure 25](image)

**Figure 25. Growth assay of the D80 populations of the twelve LTEE lineages on *Arabidopsis*.** The bacterial loads after four days of *in planta* growth of the D80 populations from each lineage (n=6) from a starting inoculum of 10^6 CFU/ml. Four lineages showed a significant increase in growth compared to RA after the Benjamini-Hochberg adjustment (Student’s t-test; * P<0.05). None of the lineages showed a significant difference in growth compared to wild-type. Error bars above represent SEM.
Figure 26. Growth assay of D80 populations in parallel with ancestral populations of the twelve LTEE lineages on Arabidopsis. The bacterial loads after four days of in planta growth of the D80 populations from each lineage (n=6) in parallel with ancestral populations (n=6) in two separate blocks (A & B). All treatments were inoculated from a starting inoculum of 10^6 CFU/ml. Four lineages of the D80 populations showed a significant decrease in growth compared to their true ancestral populations (Student’s t-test; * P <0.05, ** P<0.01, *** P<0.001). Error bars above represent SEM.
5.2 *In vitro* growth assays of D80 populations compared to ancestors

The decrease in growth on *Arabidopsis* of some lineages after eighty days of passaging was unexpected given that these lineages had experienced between 78 to 109 *in planta* generations (Figure 18) on *Arabidopsis*. A risk in using hypermutators in evolution experiments, is that hypermutators have a greater chance of accumulating deleterious mutations and a higher genetic load (Zeyl et al., 2001; Funchain et al., 2000). We originally postulated that this genetic load caused the reduction in general fitness, and resulted in some of the lineages to grow worse on *Arabidopsis* relative to their ancestors. Therefore, *in vitro* growth assays of the D80 populations of the twelve lineages were grown in parallel with their ancestors to test for general decreases in fitness. Treatments were normalized for overall plate effects by dividing each replicate with the average area under the curve of the entire plate. None of the lineages had D80 populations which grew significantly worse than the ancestral strains in KB (Figure 27). Conversely, the L11D80 grew significantly better than its ancestral population (Student’s *t*-test; *P*=0.03). These results suggest that the reduced fitness of some lineages on *Arabidopsis* was not the result of a high genetic load decreasing general growth.
5.3 Growth assay comparing population streak to decontaminated isolates

The presence of fungal contamination within the freezer stocks of all D80 populations (Figure 28) may have negatively influenced the growth of lineages on Arabidopsis. It is likely that this fungal contamination originated from the undiluted homogenate used to inoculate the liquid cultures to construct the freezer stocks (Figure 13). Although liquid cultures were supplemented with Cycloheximide to kill fungal contaminants, some still managed to persist in the freezer stocks (Figure 28). When the previous growth assays of the D80 populations were performed, streaked plates did not have visible contamination. Only after one to two weeks after being streaked did the D80 populations have visible fungal contamination. Despite this, the potential presence of fungal contamination in the D80 inoculums could be negatively affecting the growth of infecting P. syringae, as fungi possess additional PAMPs like chitin which could further contribute to PTI (Boller and Felix, 2009; Boutrot and Zipfel, 2017). Therefore, we tested if the potential presence of fungal contamination was negatively affecting the growth of the D80 populations. A growth assay was done growing decontaminated D80
populations in parallel with regularly prepared D80 population for L7 and L12. The decontaminated D80 populations were prepared by combining and re-streaking five random isolates for use in a standard growth assay. There was no significant difference in growth on *Arabidopsis* between D80 treatments for either L7 (Student’s t-test; P=0.94) or L12 (Student’s t-test; P=0.28) (Figure 29A & B). For L12, both the D80 population (P=0.0093) and the five random isolates from the D80 population (P=0.0011) grew significantly worse than the ancestor (Figure 29B) after the Benjamini-Hochberg adjustment. This decrease in growth of L12D80 was consistent with the previous growth assay (Figure 26B) and indicative of maladaptation in L12 on *Arabidopsis*. Overall, this assay demonstrated that the potential fungal contamination within the D80 populations in the previous growth assays did not significantly affect growth on *Arabidopsis*.

**Figure 28. Plates of bacterial streaks of both ancestral and D80 populations from the twelve LTEE lineages** Population streaks from freezer stocks from D80 (A, B & C) and D0 (D, E & F) after two weeks of growth on plates. Strains were streaked onto KB Rif Km Cyc plates from freezer stocks. All D80 populations had some fungal contamination while none existed in the D0 populations.
Figure 29. Growth assay testing if fungal contamination affected D80 population growth on Arabidopsis. The bacterial loads of L7 and L12 treatments were measured after zero (n=3) and four days (n=7) in planta. The regularly prepared D80 population was grow in parallel with five isolates from the D80 population (decontaminated) and the true ancestral populations for L7 (A) and L12 (B). D80 populations did not grow significantly different compared to the decontaminated treatment over four days for either L7 (Student’s t-test; P=0.94) or L12 (Student’s t-test; P=0.28). Both D80 treatments grew significantly less than the ancestral population over four days for L12 (Student’s t-test; **P<0.01). Error bars above represent SEM.
5.4 In planta competition assays with L7

The former growth assays of the evolved LTEE lineages exhibited no evidence to suggest that any lineage had adapted on Arabidopsis. One possibility is that the overall power of the growth assays implemented were too weak to detect the finite increases in fitness which some lineages had experienced. Therefore, more sensitive in planta competition assays of the evolved LTEE lineages were implemented, as these assays reduce the noise generated from experimental variation because both competitors share the same biological replicates (Macho et al., 2016). Initially, we performed a time series in planta competition assay with L7D80 competing with wild-type Pph 1448A over six days. L7 was selected for this assay as it showed the most promise of adaptation from the previous growth assays (Figure 25, 26 & 29). The goals of this assay were to identify: if L7D80 was more fit than wild-type Pph 1448A, and if so at which time did L7D80 have the largest difference in fitness. Since only one dilution was spread plated for each replicate, if a replicate grew to a density outside of the countable range, it was excluded from any analysis. One replicate measured on day five was uncountable and therefore only seven replicates were quantified that day. Over the course of the experiment, L7D80 grew significantly more than wild-type after one (Student’s t-test; P=0.045), four (Student’s t-test; P=0.045) and six (Student’s t-test; P=0.045) days post inoculation with the largest difference occurring on day six (Figure 30). P-values were adjusted using the Benjamini-Hochberg method to account for multiple testing. The results of this assay indicated that L7 had adapted to grow significantly better on Arabidopsis relative to wild-type Pph 1448A, and that the largest growth advantage occurred at the latest time point measured.
Figure 30. Time series competition assay of L7D80 over six days on *Arabidopsis*. The mean fitness of L7D80 compared to wild-type *Pph 1448A* over a series of six days (n=8) *in planta* on *Arabidopsis*. Fitness was measured using the selective rate constant which measures the difference in the natural logarithms of growth between competitors. L7D80 grew significantly better than wild-type: one, four, and six days *in planta* after the Benjamini-Hochberg adjustment (Student’s t-test; *P<0.05). Error bars above represent SEM.

To test if the growth advantage of L7D80 on *Arabidopsis* relative to wild-type was the result of genetic adaptions gained through the LTEE, a subsequent competition assay was done where both the L7D80 and L7D0 were competing with wild-type *Pph 1448A*. Eight replicates were quantified each day except for L7D80 on day four, L7D0 on day six, and L7D80 on days six which each lost one replicate because plates were uncountable. L7D80 grew significantly better than L7D0 relative to wild-type both four (Student’s t-test; P=0.0495) and six (Student’s t-test; P=0.0023) days post inoculation with the largest difference seen after six days *in planta* (Figure 31). Overall, the results of both competition assays indicate that L7 had significantly adapted on *Arabidopsis* throughout the eighty days of passaging, and that the largest fitness increase occurred at the latest measured time post inoculation.
**Figure 31. Competition assay of L7D80 in parallel with L7D0 on Arabidopsis.** The mean fitness of L7D80 and L7D0 compared to wild-type Pph 1448A after four (n=8) and six (n=8) days in planta on Arabidopsis. Fitness was measured using the selection rate constant which measures the difference in the natural logarithms of growth between competitors. L7D80 had a significantly higher r compared to the ancestral population on after both four and six days in planta (Student’s t-test; * P<0.05, ** P<0.01). Error bars above represent SEM.

5.6 Discussion

Since the starting inoculums were not controlled throughout the LTEE, this generated a large amount of variation in the final bacterial densities among lineages over the course of the experiment. Additionally, a large amount of variation in growth of all lineages between passages existed likely due to overall flat effects. This noise made detecting adaptive increases in growth from the LTEE data difficult. Therefore, growth assays with controlled inoculums and multiple replicates were used to determine if any of the evolved lineages in the LTEE, adapted to grow better on Arabidopsis. A summary of all growth assay results is provided in Figure 31. In the first growth assay, five lineages showed significant increases in growth to the non-host relative to the surrogate ancestral strain RA (Figure 32). Unfortunately, these results were not repeatable as none of these lineages grew significantly better than RA in the proceeding growth assay, while L11 and L12 grew significantly worse. The conflicting results between the two growth assays (Figure 25 & 26) appeared to be driven by RA growing about 0.5 logs lower in the
first growth assay compared to the two blocks in the second growth assay. The final bacterial loads of RA in both blocks of the second growth assay were significantly higher than the final bacterial load in the first assay (Student’s t-test; P=0.009 & 0.001) after the Benjamini-Hochberg adjustment. Therefore, the significant increases in growth seen in the first the growth assay were most likely the result of RA being a low outlier in that particulate assay, and not because of an adapted growth advantage in the evolved lineages. Additionally, none of the D80 populations grew better than their respective ancestors, while four lineages grew significantly worse suggesting maladaptation on *Arabidopsis*.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>First GA</th>
<th>Second GA</th>
<th>Third GA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth relative to RA</td>
<td>Growth relative to RA</td>
<td>Growth relative to D0</td>
</tr>
<tr>
<td>L1</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L2</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L3</td>
<td>*</td>
<td>ns</td>
<td>*</td>
</tr>
<tr>
<td>L4</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L5</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L6</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L7</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L8</td>
<td>ns</td>
<td>ns</td>
<td>*</td>
</tr>
<tr>
<td>L9</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L10</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>L11</td>
<td>ns</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>L12</td>
<td>ns</td>
<td>*</td>
<td>***</td>
</tr>
</tbody>
</table>

**Figure 32. Summary of the growth assay results of the D80 populations from the twelve LTEE lineages on *Arabidopsis***. Green indicates significant increases in growth while red indicates significant decrease in growth (Student’s t-test; * P <0.05, ** P<0.01, *** P<0.001). GA (growth assay), ns (non-significant), & NA (non-applicable).

From these results, we originally postulated that because the LTEE lineages were hypermutators, they had accumulated a high genetic load causing a general decrease in fitness. The in
vitro growth assays done in KB showed no signs of growth deprivation among any of the evolved lineages, with L11D80 showing a significant increase in general growth relative to its ancestor. Although, this result indicated that lineages have not accumulated mutations which affect general growth, it does not exclude the possibility that mutations specifically affecting in planta fitness have not accumulated. Repeating these in vitro growth assays in hrp inducing media (which induces the expression of the T3SS) or in Arabidopsis apoplastic fluid might better characterize changes in fitness which affect in planta growth. Alternatively, using phenotypic assays which characterize traits important for virulence such as motility, and biofilm formation may better uncover potential in planta maladaptation (Manoharan et al., 2015).

Furthermore, we hypothesized that the fungal contamination within the D80 freezer stocks negatively affected the growth of the evolved lineages in the former growth assays. The growth of decontaminated D80 isolates of L7 and L12 were compared to the growth of the normally prepared populations, with neither treatment growing significantly different indicating that any possible fungal contamination in the D80 populations was not significantly impacting growth. Additionally, in this growth assay, L12D80 grew significantly worse than L12D0 as in the second growth assay (Figure 32), further supporting L12 has become less fit on Arabidopsis. One hypothesis is that the predicted bottleneck in passage 16 of ≈171 CFUs (Figure 17) may have resulted in the fixation of deleterious mutations within L12. Performing growth assays measuring the fitness of L12 before and after passage 16 could uncover if this bottleneck resulted in the reduced in planta fitness. Also, growth assays on the original bean host could test if this reduction in fitness is general to colonizing the in planta environment, or specific to Arabidopsis.

Although some evolved lineages showed significant decreases in growth on Arabidopsis, most showed no difference in growth relative to RA or to their ancestors (Figure 32). It is conceivable that the intrinsic noise of standard growth assays was too large to effectively detect the finite increases in fitness.
which these lineages might have gained throughout the LTEE. Therefore, we implemented more sensitive *in planta* competition assays with L7, as this lineage showed the most promise of adaptation in the first growth assay. L7D80 consistently showed a significant growth advantage in both *in planta* completion assays (Figure 30 & 31) indicating that this lineage has adapted on the non-host. In the future, this assay should be implemented with the remaining eleven lineages as it appears to better characterize finite differences in growth compared to growth assays.
Chapter 6: Discussion and Future Directions

6.1 Overview

The main objective of this thesis was to adapt the bean pathogen Pph 1448A onto the non-host Arabidopsis through experimental evolution, with the goal of characterizing the early adaptive steps a pathogen takes towards novel host adaptation. Twelve independent lineages of Pph 1448A were serially passaged onto Arabidopsis over eighty days, which equated to just over 100 generations of growth for most lineages on the non-host (Table 5). Prior to the passaging experiment, lineages were engineered into MMR deficient hypermutators by knocking out mutS, and replacing the gene with a uniquely barcoded Km\(^r\) cassette (Figure 1). This study was the first to engineer MMR deficient hypermutating strains of P. syringae, and to use hypermutators to study the adaptation of a plant pathogen to a non-host. Furthermore, each lineage possessed a unique eight nucleotide barcode integrated within the genome, with the purpose of allowing us to track for cross contamination throughout the experiment. The genomic integration of the Km\(^r\) cassette was used as a selectable marker allowing us to perform sensitive in planta competition assays of the evolved lineages against wild-type Pph 1448A. From these competition assays, we have shown that at least L7 has adapted to significantly grow better on Arabidopsis after eighty days of passaging (Figure 30-31).

6.2 Prospective adaptive strategies of L7 on Arabidopsis

The further characterization of L7 has the potential to uncover the strategy or strategies to which this lineage has adapted to grow better on the non-host. The detection of PAMPs which triggers PTI represents a major barrier which pathogens must overcome to successfully cause disease (Boller and Felix, 2009; Senthil-Kumar and Mysore, 2013). Although PAMPs are generally categorized as highly conserved epitopes, a remarkable amount of variation exists both among alleles of these epitopes and in their perception by PRRs (Clarke et al., 2013; Ingle et al., 2006; Sun et al., 2006). PAMPs show specific signatures of positive selection within residues that are being detected by the host, a sign of selection for immune
evasion (McCann et al., 2012). Interestingly, these patterns of positive selection have been utilized to detect candidate PAMPs within conserved genes (core genomes) of the plant pathogens *Xanthomonas campetris* and *P. syringae*, in which multiple candidates were validated as PTI elicitors (McCann et al., 2012). Overall, this evidence suggests that one strategy employed by pathogens to overcome host immunity is through PTI evasion by PAMP diversification. The detection of the flagellin PAMP has been implicated as a major component of the NHR of *Arabidopsis* against *Pph 1448A* (Forsyth et al., 2010), and one could hypothesize that mutations within the flagellin epitope which reduce host detection could be responsible for L7’s adaptation. Interestingly, the complete loss of the flagellin has shown no ill effect on *P. syringae* growth when syringe infiltrated, but a significantly negative effect on growth when surface inoculated (Schreiber and Desveaux, 2011). Therefore, L7 may have completely lost the flagellin to evade detection as lineages were syringe inoculated throughout the LTEE, making the flagellin potentially dispensable. If so, swim plate motility assays could be used to detect for changes in motility and the potential loss of the flagellin. Additionally, mutations within other PAMPs which result in the evasion of host detection may also have contributed to L7’s adaptation. Future experiments which measure the output signals of PTI may find that L7D80 elicits a weaker output relative to its ancestor if it has adapted to evade PTI detection. Assays measuring ROS activity (Chakravarthy et al., 2010; Clarke et al., 2013) (Chakravarthy et al 2010, Clarke paper), apoplastic peroxidase activity (Mott et al., 2016) and callose deposition would be good candidate experiments to measuring PTI outputs.

Immune suppression is another strategy which L7 may have utilized to grow better on *Arabidopsis*. *P. syringae* generally suppresses immunity by injecting T3SEs into the host cell (Lindeberg et al., 2012). A candidate mutation which has the possibility to enhance *Pph 1448A*’s suppression of the *Arabidopsis* immunity, is a reversion mutation within the truncated hopM1 effector. HopM1 is responsible for inhibiting vesicle trafficking and its presence is essential for producing a favourable aqueous environment within the apoplast (Nomura, 2007; Xin et al., 2016). This effector is shared among many
diverse strains of *P. syringae* and is essential for virulence on *Arabidopsis* (Baltrus *et al.*, 2011; Xin *et al.*, 2016; Nomura, 2007). The hopM1 allele in *Pph* 1448A is non-functional as it possess a frameshift mutation at codon 545, resulting in a premature stop codon (Joardar *et al.*, 2005). It is likely that this truncation was selected for to evade host immunity as *Pph* 1448A cloned with an intact hopM1 allele experiences a significant growth deficit on its original bean host (Baltrus *et al.*, 2012). The selective pressures of the *Arabidopsis* environment throughout the LTEE may have selected for hopM1’s restoration. Therefore, it is possible that L7’s improved growth on *Arabidopsis* is the result of enhanced immune suppression. Sanger sequencing of hopM1 could determine if this reversion occurred in L7D80.

Alternatively, L7 could have adapted to better tolerate the immune response of *Arabidopsis*. One major component of the *Arabidopsis* immune response which inhibits the growth of many non-host pathogens including *Pph* 1448A, is the production of the secondary metabolite sulforaphane (Fan *et al.*, 2011). All *P. syringae* strains with the ability to infect *Arabidopsis*, possess sax genes which provide these strains resistance against sulphoraphane and other glucosinolate derivatives. *Pto* DC3000 mutants which lack all sax genes experience larger growth deficits on *Arabidopsis* the longer it remains in planta (Fan *et al.*, 2011). This is precisely the opposite trend seen in L7D80’s fitness on *Arabidopsis*, as it possesses the greatest fitness advantage the longer it is in planta (Figure 30 &31). Therefore, L7D80 may have adapted to become more tolerant of sulphoraphane and/or of other antimicrobials produced by *Arabidopsis*. Experiments measuring the *in vitro* growth of L7D80 in *Arabidopsis* extracts containing antimicrobials such as sulforaphane would indicate if this lineage has adapted to be more tolerant of this environment.

### 6.3 Future WGS of the LTEE lineages

WGS of L7D80 and its ancestor (L7D0) could potentially uncover beneficial variants in the adapted population, and enhance our ability formulate and test hypotheses for the mechanism(s) behind the improved fitness on *Arabidopsis*. The one caveat of using hypermutators in this LTEE, is that it will increase the difficulty of identifying beneficial mutations from sequencing data, as most variants detected will likely
be neutral hitchhiker mutations (Tenaillon et al., 2016). To overcome this issue, the detection of parallel evolution among lineages which is a strong signature of adaptation, would be the most effective method in identifying truly beneficial variants responsible for the adaptation on Arabidopsis (Tenaillon et al., 2016; Elena and Lenski, 2003). Therefore, competition assays on Arabidopsis of the remaining lineages should be done to identify if other lineages have significantly adapted onto the non-host. If so, WGS of these lineages along with L7 can be done to identify loci under parallel evolution. The presence of barcodes within each lineage will also allow us to effectively distinguish between parallel evolution and cross-contamination in WGS data.

6.4 Continuation of the LTEE

Although one lineage has shown significant adaptation to grow on Arabidopsis, most lineages have not shown any sign of adaptation over multiple growth assays (Figure 32). As previously stated, it could be that other lineages have significantly adapted to grow on Arabidopsis, but that the increases were too finite to be detected by growth assays. Competition assays with the remaining lineages should elucidate if other lineages have significantly adapted to grow on the non-host.

Another reason why most lineages may not show significant signs of adaptation to Arabidopsis, could be that not enough evolution of the lineages on the non-host has taken place. An estimate of the expected number of mutations per clone within a lineage was determined using the mutation rates of Δmuts strains of fluorescens and P. aeruginosa determined from MA experiments (Long et al., 2014; Dettman et al., 2016) and the estimated number in planta generations of each LTEE lineage (Table 5). Clones within each lineage were estimated to have experienced 11-17 mutations (Table 5) assuming neutrality, which equates to a rate of 15 mutations/100 generations. This rate is about 10-fold higher than the average rate seen over multiple E. coli evolution experiments (Dettman et al., 2012) a result of the strains used in this experiment being hypermutators. Although many beneficial variants may have arisen in multiple LTEE lineages, 100s of generations are required for mutations even with large fitness
advantages to initially stabilize and to rise within a population (Barrick et al., 2009). In comparison with other in planta evolution experiments, the LTEE performed in this work is still within its early stages. One evolution experiment serially passaged the plant pathogen R. solanacearum on five different plant hosts for over 300 generations, and showed that this was a sufficient number of generations for multiple lineages to significantly adapt to their respective hosts (Guidot et al., 2014). Another evolution experiment using a modified R. solanacearum strain with the ability to form root nodules, took approximately 300 generations of serially passing on the roots of the legume host M. pudica for all three independent populations to gain the ability to cause intracellular infections within root nodules (Guan et al., 2013). The evolution experiment which passaged the pathogen of X. citri on the resistant kumquat host, found that after approximately 800 in planta generations, two out of three lineages adapted to evade the elicitation of HR (Trivedi and Wang, 2014). These evolution studies suggest that more generations of in planta growth on Arabidopsis are required for the continual selection and fixation of beneficial variants in the LTEE lineages.
Table 5. Estimated generations and mutations per clone of the twelve LTEE lineages over eighty days on *Arabidopsis*

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Days of passaging</th>
<th>Total generations</th>
<th>Estimated mutations per clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>80</td>
<td>109</td>
<td>16</td>
</tr>
<tr>
<td>L2</td>
<td>80</td>
<td>105</td>
<td>16</td>
</tr>
<tr>
<td>L3</td>
<td>80</td>
<td>109</td>
<td>17</td>
</tr>
<tr>
<td>L4</td>
<td>80</td>
<td>105</td>
<td>16</td>
</tr>
<tr>
<td>L5</td>
<td>80</td>
<td>106</td>
<td>16</td>
</tr>
<tr>
<td>L6</td>
<td>80</td>
<td>103</td>
<td>16</td>
</tr>
<tr>
<td>L7</td>
<td>80</td>
<td>109</td>
<td>16</td>
</tr>
<tr>
<td>L8</td>
<td>60</td>
<td>78</td>
<td>12</td>
</tr>
<tr>
<td>L9</td>
<td>60</td>
<td>80</td>
<td>12</td>
</tr>
<tr>
<td>L10</td>
<td>60</td>
<td>77</td>
<td>12</td>
</tr>
<tr>
<td>L11</td>
<td>60</td>
<td>75</td>
<td>11</td>
</tr>
<tr>
<td>L12</td>
<td>68</td>
<td>82</td>
<td>12</td>
</tr>
</tbody>
</table>

Unfortunately, all twelve lineages experienced population crashes after eighty days of passaging. It appeared that the outdoor climate had subtle but significant effects on both the temperature and humidity within the growth room. Passaging throughout the first eighty days occurred during the winter months when plant mass was significantly less compared to plants used in the spring. Plants grown within the winter were probably under more abiotic stress, which increased ABA production and stunted growth (Tuteja, 2007). Although ABA signaling is an important early response to pathogen defense as it induces the closure of the stomata (Melotto *et al.*, 2006), constitutive production reduces SA defense signaling and enhances *P. syringae* growth (Fan *et al.*, 2009). During the spring and summer, plants are healthier and probably induce a strong immune response causing lineages to crash. This difference in plant health might also explain why evolved lineages did not show significant adaptation in growth assays. These assays were performed in the spring, on plants which did not accurately represent the environment in which the lineages were passaged on throughout the LTEE. Mutations selected for within the LTEE
environment may not be adaptive on healthier plants, and resulted in the appearance of no adaptation or maladaptation in these assays (Figure 32).

If the greater production of ABA in plants grown during the winter months resulted in the enhanced growth of *Pph* 1448A lineages, one possible solution to continuing the passaging during the spring and summer months, could be to reduce the amount of water provided to plants for passaging during these seasons. This would increase ABA production and to enhance the growth of lineages (Fan *et al.*, 2009; Tuteja, 2007). Alternatively, using younger plants at a similar developmental stage as 4-week old winter plants may enhance the growth of lineages during the spring and summer seasons. Lastly, integrating the optimized plating step to recover crashing lineages to the passaging protocol could be done in the spring and summer months to allow for the LTEE to continue. Overall, it is essential that these lineages continue to be passaged on *Arabidopsis* for further adaptation to occur.

### 6.5 Future modifications to the LTEE

The inconsistent growth of *Pph* 1448A lineages on the non-host *Arabidopsis* throughout the LTEE made it difficult to perform direct plant to plant transfers and eventually resulted in the temporary stoppage of the experiment. One modification which could result in a more successful experimental design is the use of a non-host in which *Pph* 1448A can undergo more robust growth. This could be achieved by using the *Arabidopsis* ecotype Wassilewskija-3 (Ws-3), as *Pph* 1448A grows significantly better on Ws-3 relative to Col-0 because of a disruption mutation within the FLS2 gene, resulting in the Ws-3 ecotype being insensitive to the flagellin PAMP (Zipfel *et al.*, 2004; Forsyth *et al.*, 2010). Alternatively, using a non-host species which is phylogenetically more closely related to the common bean, such as soybean or mung bean would also permit better growth of *Pph* 1448A (Baltrus *et al.*, 2012), relative to *Arabidopsis* Col-0. By using a non-host which permits better growth, lineages would undergo more generations per passage and increase the rate at which adaptive mutations are introduced and fixed in the population. Therefore, performing the LTEE on any of these suggested non-
hosts would reduce the difficulty in optimizing a passaging protocol, and could accelerate the rate at which *Pph* 1448A significantly adapts to the non-host.

Another modification which could make the experimental design more effective in selecting for variants important for novel host adaptation, is the use of a surface inoculation. Syringe inoculation was used in this LTEE because it permitted the use of lower inoculum concentrations relative to surface inoculation, thus making direct plant to plant passaging for *Pph* 1448A possible on *Arabidopsis*. By using this method, lineages bypassed the selective pressures which naturally infecting *P. syringae* experience in the epiphytic and initial invasion stages (Hirano and Upper, 2000). In doing so, only mutations which are beneficial for *in planta* growth were selected for, despite the potential costs they may have in pre-infection environment reducing their biological relevance. This was the case for lineages of the bacterial plant pathogen *R. solanaceaeum* which were experimentally evolved onto the novel bean host via direct injection into the stem (Guidot *et al.*, 2014). Mutations which suppressed the activity of the catabolic and virulence suppressor *efpR*, were selected for in multiple lineages and significantly enhanced the *in planta* growth of *R. solanaceaeum* on the bean host. Conversely, clones with the adapted alleles of *efpR* were significantly less fit when inoculated via root soak, a more natural mode of inoculation (Perrier *et al.*, 2016) indicating that these mutations have a cost in the pre-infection environment. This result also explains why these mutations were not found within natural isolates (Perrier *et al.*, 2016). Therefore, implementing a surface inoculation method such as spray or dip inoculation in future evolution experiments should more effectively select for variants important for the overall pathogen fitness on a novel host, and not just within the *in planta* environment.

6.6 Conclusion

In this work, we showed that *in planta* experimental evolution is a viable approach to characterizing the evolution of host-specificity in *P. syringae*. The continuation of this LTEE can enhance our general understanding of the adaptive strategies *P. syringae* uses to overcome novel host immunity
within the early stages of host-specialization. In doing so, we may be able to better understand the general processes of how strains of *P. syringae* adapt to cause disease on novel hosts.
References


Genetics, 204, 1249–1266.


Kim, J.J. and Sundin, G.W. (2000) Regulation of the rulAB mutagenic DNA repair operon of Pseudomonas syringae by UV-B (290 to 320 nanometers) radiation and analysis of rulAB-mediated


