UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Genome-wide identification of Pseudomonas syringae B728a genes required for competitive fitness during colonization of diverse plant hosts, with a focus on toxin efflux transporters

Permalink
https://escholarship.org/uc/item/9m34g17p

Author
Helmann, Tyler Cal

Publication Date
2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Genome-wide identification of *Pseudomonas syringae* B728a genes required for competitive fitness during colonization of diverse plant hosts, with a focus on toxin efflux transporters

By

Tyler Cal Helmann

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Steven E. Lindow, Chair
Professor Shauna C. Somerville
Professor Ellen L. Simms

Summer 2019

Abstract

Genome-wide identification of *Pseudomonas syringae* B728a genes required for competitive fitness during colonization of diverse plant hosts, with a focus on toxin efflux transporters

by

Tyler Cal Helmann

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Steven E. Lindow, Chair

Bacterial colonization of a plant host requires both specialized and general traits. On the leaf surface and in the leaf interior (apoplast), bacteria encounter resources needed to support growth but are also exposed to multiple abiotic and biotic stresses that they must overcome. The foliar plant pathogen *Pseudomonas syringae* can establish large epiphytic populations on leaf surfaces before moving through stomata or other openings such as wounds into the apoplast, where it can multiply to very high numbers before inciting disease symptoms. Strain B728a is pathogenic to common bean (*Phaseolus vulgaris*), lima bean (*P. lunatus*), pepper (*Capsicum annuum*), *Nicotiana benthamiana*, and I show that it is rather promiscuous, capable of growth within other plant species. Strains of *P. syringae* are commonly found as epiphytes on a variety of both host- and non-host plants in agricultural systems, native plants, and in various aquatic environments. *P. syringae* has thus presumably evolved or acquired a diversity of traits needed for success in these varied habitats, but we lack an understanding of the identity and role of these genes and the degree to which they are linked to fitness in a given context.

To examine the context-dependent fitness contributions of genes in *P. syringae* for growth on or in leaves, I created a barcoded transposon insertional mutant library in strain B728a. Randomly-barcoded TnSeq (RB-TnSeq) enables the same heterogeneous mutant population to be used to interrogate gene contribution to bacterial fitness in a variety of distinct hosts or environments. I determined the contributions of 4,296 nonessential genes to fitness on the leaf surface and in the apoplast of common bean. Genes within the functional categories of amino acid and polysaccharide biosynthesis contributed most to fitness both on the leaf surface and in the apoplast, while genes involved in type III secretion and syringomycin synthesis were important primarily in the apoplast. Interestingly, for most genes no relationship was seen between fitness *in planta* and either the magnitude of their expression or degree of induction *in planta* compared in this host plant to *in vitro* conditions measured in other studies. Given the robust bacterial growth of *P. syringae* also in lima bean and pepper, I utilized the barcoded transposon library to identify genes that differentially contributed to

apoplastic colonization in those hosts. Most genes contributed to apoplastic fitness at similar magnitudes in each of the plant hosts. However, 87 genes significantly differed in their fitness contribution in at least one host, and 50 of these genes were highly important for apoplastic colonization. Genes in the functional categories of amino acid metabolism, phytotoxin biosynthesis, alginate biosynthesis, and cofactor metabolism significantly differed in contributions to fitness in the plant species tested. In addition, six genes that encoded hypothetical proteins contributed differentially to growth in these hosts.

Because *P. syringae* colonizes such a large number of plant species and various habitats it presumably is exposed to a wide variety of inhibitory compounds that it must tolerate. To test the hypothesis that plants would differ in the intensity and nature of chemical defenses against bacterial colonists I used a *mexB* mutant of *P. syringae* strain B728a that was blocked in expression of an important efflux pump as an indicator of the chemical defenses of various plants. While MexB was required for full virulence in common bean and pepper, it was dispensable for apoplastic growth in lima bean, *Nicotiana benthamiana*, sunflower (*Helianthus*), and tomato (*Solanum lycopersicum*). Within these plants, B728a grew to varying population sizes, indicating differential susceptibilities among these compatible host plant species. *P. syringae* strain B728a encodes 668 predicted transport proteins. However, the substrates of a majority of these transporters are unknown. Many of these transporters confer resistance to diverse toxicants present in the environment, but due to a high level of functional redundancy, it is difficult to identify those that are of most importance in conferring resistance to specific compounds. To identify transporters in B728a that are important for tolerance of external toxicants but redundant with MexB, I exposed the WT barcoded transposon mutant library as well as a library constructed in the *mexB* deletion strain to diverse toxic compounds *in vitro* to identify mutants with reduced ability to grow in the presence of those compounds. The homologous transporters MexCD and MexEF were partially redundant with MexAB, sharing a subset of substrates tested. I hypothesized that inner membrane transporters contribute to substrate specificity, particularly when functioning cooperatively with outer membrane transporters or porins. In support of this model, the inner membrane transporters Psyr_0228 and Psyr_0541 were highly substrate specific, and contributed to growth independent of *mexB* genotype. This study of isogenic transposon libraries revealed the substrates of efflux transporters that are masked by the dominant activity of MexAB.

These studies provide a more comprehensive insight into the chemical ecology of *P. syringae*. It is clear that plant species differ in their chemical defenses, as seen in the differential fitness of the hyper-susceptible *mexB* efflux pump mutant in various host plants. The fitness of various auxotrophs within a transposon library *in planta* also indicates resources that are often in low amounts in the host, such as amino acids and cofactors. These studies identified genes contributing to competitive fitness in diverse environmental conditions, including tolerance of toxicants that might be encountered by *P. syringae.* The resulting large datasets represent a valuable tool for the identification of genes required for compatibility with a given plant host. Improving our understanding of genes that contribute to bacterial fitness in diverse environments, including the phyllosphere, will improve our understanding of those habitats and the bacterial traits needed to survive and thrive.

Table of Contents

List of Figures ii

List of Tables v

Acknowledgements vii

Chapter 1 1
Introduction

Chapter 2 8
Genome-wide identification of *Pseudomonas syringae* genes required for competitive fitness during colonization of the leaf surface and apoplast

Chapter 3 64
The efflux transporter MexAB-OprM contributes to *Pseudomonas syringae* B728a virulence in a habitat, temporal, and host-dependent manner

Chapter 4 93
Identification of *Pseudomonas syringae* host-specific leaf colonization genes using randomly barcoded TnSeq

Chapter 5 123
Genome-wide transposon screen of a *Pseudomonas syringae* Δ*mexB* mutant reveals the substrates of efflux transporters

Chapter 6 159
Discussion

References 163

List of Figures

Chapter 2

Figure 2-1. The B728a *mariner* transposon library contains 281,417 mapped insertions.

Figure 2-2. Dendrogram of experiments generated using *P. syringae* fitness values.

Figure 2-3. Fitness could not be calculated for 920 genes.

Figure 2-4. Representative growth of B728a in the apoplast.

Figure 2-5. Rank ordered mean gene fitness values for each condition in which *P. syringae* was grown.

Figure 2-6. Genes with significant contributions to competitive fitness.

Figure 2-7. Fitness contributions of genes involved in phytotoxin synthesis and transport, the type III secretion system, and polysaccharide synthesis and regulation.

Figure 2-8. Consistency in identifying genes contributing to epiphytic and apoplastic growth in different replicate experiments.

Figure 2-9. Average apoplastic fitness and absolute average *t*-score for all genes.

Figure 2-10. Growth of deletion mutants in KB and LB.

Figure 2-11. Growth of B728a auxotrophic mutant strains Δ*trpA* and Δ*hisD* in M9.

Figure 2-12. Inhibition of *Geotrichum candidum* Fr260 growth by B728a WT and Δ*syrP.*

Figure 2-13. Apoplastic growth of B728a and deletion strains in bean.

Figure 2-14. The magnitude of fitness contributions of genes in *P. syringae* do not correlate well with their absolute level of expression or fold-change *in planta*.

Chapter 3

Figure 3-1. B728a Δ*mexB* grows as well as WT in King's B (KB) broth.

Figure 3-2. Δ*mexB* establishes lower populations sizes on the leaf surface.

Figure 3-3. Apoplastic populations sizes of B728a WT and Δ*mexB.*

Figure 3-4. Population size of WT B728a in the apoplast is proportional to the inoculum concentration.

Figure 3-5. Frequency distribution of the number of lesions formed on individual leaves.

Figure 3-6. Bacterial population sizes within individual lesions.

Figure 3-7. Frequency distribution of the number of cells recovered from asymptomatic regions of leaves.

Figure 3-8. Number of lesions formed by B728a WT on diverse *P. vulgaris* varieties.

Figure 3-9. Apoplastic population of B728a WT and Δ*mexB* in *Phaseolus* varieties.

Figure 3-10. Apoplastic populations of B728a WT and Δ*mexB* in alternate host plants.

Figure 3-11. Apoplastic populations of B728a WT in *Brassica nigra and Trifolium pratense*.

Figure 3-12. Number of lesions detected on individual leaves over time in common bean and *Nicotiana benthamiana*.

Figure 3-13. Bacterial populations within lesions decline over time in common and lima bean, and not *Nicotiana benthamiana*.

Figure 3-14. Proposed model of a population size threshold for lesion formation and the role of MexB in lesion formation.

Chapter 4

Figure 4-1. Growth of B728a in the apoplast of the susceptible hosts common bean, lima bean, and pepper.

Figure 4-2. Principle component analysis of *P. syringae* B728a grown in KB, common bean, lima bean, or pepper.

Figure 4-3. Rank ordered mean fitness values for each condition and plant species in which *P. syringae* was grown.

Figure 4-4. Distribution of standard deviation of gene fitness values for each gene.

Figure 4-5. The standard deviation for estimates of gene fitness values is highest at both high and low measures of gene fitness contribution.

Figure 4-6. Pairwise comparisons of gene fitness values between plant hosts.

Figure 4-7. Genes with significant contributions to competitive fitness in each host.

Figure 4-8. Heatmap of apoplastic fitness contribution of histidine and isoleucine/valine biosynthetic genes to growth in various plant species.

Figure 4-9. Heatmap of apoplastic fitness values of genes involved in phytotoxin biosynthesis.

Figure 4-10. Heatmap of apoplastic fitness for alginate biosynthetic genes.

Chapter 5

Figure 5-1. Zone of growth inhibition assays to test antibiotic sensitivity of B728a Δ*mexB* complementation variants.

Figure 5-2. Fitness of transposon insertional mutants in the operon *mexAB-oprM*.

Figure 5-3. Fitness of transposon insertional mutants in the operon *mexEF-oprN.*

Figure 5-4. Fitness of transposon insertional mutants in the operon *muxABC-opmB.*

Figure 5-5. Fitness of transposon insertional mutants in the operon *Psyr_2282-3.*

Figure 5-6. Fitness of transposon insertional mutants in the operon *Psyr_0344-6.*

Figure 5-7. Fitness of transposon insertional mutants in the MFS gene *Psyr_0228*.

Figure 5-8. Fitness of transposon insertional mutants in the SMR gene *Psyr_0541.*

Figure 5-9. Fitness of transposon insertional mutants in the MFS gene *Psyr_4519.*

Figure 5-10. Fitness of transposon insertional mutants in the operon *Psyr_0917-8.*

Figure 5-11. Fitness of transposon insertional mutants in the operon *Psyr_1754-9.*

Figure 5-12. Likely substrates of *P. syringae* B728a multidrug resistance RND transporters, the MFS transporter Psyr_0228, and the SMR pump Psyr_0541.

Figure 5-13. Growth inhibition assays to test antibiotic sensitivity of B728a efflux pump deletion mutants Δ*mexB*Δ*mexF*, Δ*mexB*Δ*muxB*, and Δ*mexB*Δ*mexK*.

Figure 5-14. Structures of acridine orange and proflavine / acriflavine.

List of Tables

Chapter 2

Table 2-1. Number of genes within each functional category predicted to be essential or nearly essential in LB, based on TnSeq data.

Table 2-2. B728a genes with 0 TA dinucleotide sites, that are inaccessible by the *mariner* transposon.

Table 2-3. B728a genes predicted to be essential compared to their predicted homologs in *P. aeruginosa* PAO1, *P. simiae* WCS417, and *P. stutzeri* RCH2.

Table 2-4. Genes that when disrupted confer large decreases in fitness in KB.

Table 2-5. Numbers of unique barcodes and median reads per gene in each of three replicate experiments.

Table 2-6. Genes with average fitness < -2 in epiphytic experiments.

Table 2-7. Genes with average fitness < -2 in apoplastic experiments.

Table 2-8. Functional categories that are enriched among genes important for fitness *in planta*.

Table 2-9. Genes with average fitness < -1 but > -2 in KB, epiphytic and apoplastic experiments.

Chapter 3

Table 3-1. *Phaseolus* varieties used in this study.

Chapter 4

Table 4-1. Numbers of unique barcodes and median reads per gene.

Table 4-2. Genes for which their median fitness contribution differed between plant species (average fitness < -0.5).

Table 4-3. Genes for which their median fitness contribution differed between plant species (average fitness > -0.5).

Table 4-4. Unique and shared gene loci among the three hosts tested having average fitness values < -2 or average fitness values < -1.

Chapter 5

Table 5-1. Characteristics of the barcoded *mariner* transposon libraries in B728a WT and Δ*mexB* backgrounds.

Table 5-2. *P. syringae* B728a contains 668 total transporter proteins.

Table 5-3. Antimicrobial compounds used in growth assays.

Table 5-4. RND operons homologous to *mexAB-oprM* that likely contribute to multidrug resistance.

Table 5-5. Shared amino acid identity between *P. syringae* B728a Psyr_2482-5 and *P. aeruginosa* PAO1 MuxABC-OpmB.

Acknowledgements

I would like to start by thanking my graduate advisor Steven Lindow for his tremendous contribution to this thesis and my growth as a scientist. In addition to having a wealth of knowledge, Steve has always been extremely generous with his time and feedback. He would frequently organize trips outside of the lab, to explore the campgrounds and hikes of Lake Tahoe, Yosemite, and other destinations in the Sierras. These lab trips were a welcome opportunity to explore California and relax with lab members and their significant others. My only regret is that we never had the chance to ski together on artificial snow produced with the aid of ice nucleating bacteria.

I would also like to thank the other members of my dissertation committee: Shauna Somerville and Ellen Simms. Their feedback during committee meetings helped to organize my goals and projects, and worked to highlight the importance of sharing my science. I would additionally like to thank the members of my qualifying exam committee for their early feedback on my work: Shauna Somerville, Jennifer Lewis, Pat Zambryski, and Rodrigo Almeida.

Many current and former lab members supported this dissertation along the way, including Monica Hernandez, Tiffany Lowe-Power, Renee Koutsoukis, Despoina Lymperopoulou, Clelia Baccari, Norma Morella, Elena Antonova, Marzia Milettoe, Russell Scott, Adrien Burch, Jules Cho, and Kevin Hockett. Adrien helped to orient me when I first joined the lab. Monica has been very supportive, and always willing to give feedback. Tiffany provided much appreciated advice during the writing process. Renee's work to keep the lab organized and well stocked made performing experiments much more manageable. This work was conducted with the assistance of the undergraduate students with whom I worked closely (in chronological order): Dana King, Kendra Paskvan, Caitlin Ongsarte, Olga Smelik, and Jennifer Lam. I thank each of them for their contributions to these projects throughout my time at Berkeley. Plants used in my greenhouse experiments were maintained with the assistance of the greenhouse staff, led by Tina Wistrom.

The Plant and Microbial Biology Department has been extremely supportive, and there are too many friends and colleagues to thank for their academic support and friendship during my time at Berkeley. I must make special exception for my fellow PMB students and first Berkeley housemates on Woolsey Street: Gabe Rosenfield and Emma Kovak. I would also like to acknowledge a few close friends in science and beyond: Darae Jun, Johan Jaenisch, Justin Zik, and Will Skyrud.

For my RB-TnSeq projects, I thank my collaborators Adam Deutschbauer and Morgan Price. Their assistance throughout the construction and testing of my transposon libraries expanded the scope of my dissertation dramatically, and allowed me to test hypotheses at a truly exciting scale. This dissertation would not be the same without Adam's generosity with his time and resources. I would also like to thank my scientific mentors from my time at Cornell: Alan Collmer, Christine Smart, Suma Chakravarthy, and Ahmed Gaballa.

Creative processes require taking the occasional break from work. To my oldest friends, Dylan Marcus, Alexander Sahn, David Halpern, Zach Montague, and Salmaan Rizvi: let's go skiing. Madeline Zehnder also deserves my thanks for her support and occasional letters throughout graduate school.

Most importantly, I would like to thank my family. To my dad, John, thank you for being my first and most prominent scientific mentor. To my mom, Robin, thank you for teaching me empathy, kindness, and sincerity. To my sister, Sierra, thank you for planning fun excursions for the family and encouraging me out of my comfort zone. Finally, a huge thank you goes to Ariel, my partner and best friend. I could not have done this without you.

Chapter 1

Introduction

The typical plant is heavily colonized by microorganisms. Commensal bacteria and fungi are found both belowground living in association with the roots (the rhizosphere) or aboveground on the leaves, flowers, and other aerial structures (the phyllosphere) (1, 2). This combination of eukaryotic host and associated microorganisms is considered the "holobiont", a relatively recent term used to recognize the genetic and phenotypic contributions of the individual parts to the collective whole (3, 4). Thus, the phenotypic outcome for the plant host is dependent upon the culmination of the positive and negative interactions with associated microorganisms (5). With this in mind, there is considerable motivation to exploit root-associated bacteria and fungi to promote plant health and increase soil nutrient availability. However, we have historically known less about analogous residents of the phyllosphere (2), despite their potential negative and positive effects on the plant. Bacteria are the most abundant residents of the phyllosphere, often found in densities up to 10^8 cells/g leaf (6). Given the large overall terrestrial leaf surface area, it is also clear that bacteria can impact global geochemical processes beyond the individual plant host (6).

Interest in bacteria, fungi, and oomycetes in the phyllosphere often focuses on the subset of these organisms that have the potential to cause disease. Diseases of leaves, fruits, flowers, and other aboveground tissues can be devastating for the individual plant, and decrease food availability for humans and other animals that consume that plant material. However, a majority of phyllosphere microorganisms do not cause disease, and exist as commensal residents of this habitat. In addition, even those that do cause disease do not cause disease on all plant species. Pathogens, especially those that interact with living host tissue, are typically relatively host specific, in that they are limited to a few plant host species. This incompatibility with potential hosts is due to a variety of factors, including physical barriers and conserved immune responses of the plant (7).

How do phyllosphere microorganisms thrive if they are not infecting a host? Commensal phyllosphere bacteria and fungi presumably consume carbon and nitrogen-containing nutrients that have leached to the leaf surface, or potentially been deposited by insects or other external sources such as pollen (8, 9). The distribution of nutritional resources on the leaf, and likely as a result, the distribution of bacterial cells, is spatially heterogeneous on plants (10). The presence of bacteria on leaf surfaces has been shown to prevent establishment of pathogenic or ice-nucleating bacterial species, due to competition for resources or direct inhibition (11, 12). In addition to the biotic stresses found in the phyllosphere, there are numerous abiotic stressors these colonists face. The leaf surface is exposed to fluctuating UV radiation, temperature, wetness, and relative humidity (13, 14). While the apoplast, or leaf intercellular space, provides a degree of protection from these stresses, endophytic colonization represents direct exposure to plant cells and thus likely the induced host immune response.

The biotrophic pathogen Pseudomonas syringae

In this dissertation, I examine host colonization traits of the model bacterial plant pathogen *Pseudomonas syringae*. This species, as a group, is capable of colonizing a wide variety of host plant species, although individual strains are typically thought to be quite host specific (13). Strains of *P. syringae* are commonly found as epiphytes on a variety of both host- and non-host plants, both in agricultural systems as well as native plant communities (15, 16). Many strains are capable of catalyzing ice formation, and because they can be found in rainfall, pristine snow, as well as in water sources around the world, are thought to play an important role in the water cycle by initiating ice formation central to the precipitation process (16). The life cycle of *P. syringae* in the phyllosphere is typically considered in two phases: initial epiphytic growth and subsequent colonization of the apoplast (17). Under favorable environmental conditions, such as high humidity and moderate temperature, bacterial populations can reach high levels, resulting in the formation of visible leaf symptoms (17).

Many traits are necessary to enable the processes of bacterial colonization of the leaf surface, movement into the apoplast, and subsequent multiplication within the leaf. On the leaf surface, cells must survive the various stresses present and achieve a sufficiently high population density to induce quorum sensing-dependent virulence pathways (18). Exopolysaccharide production on the leaf surface shields epiphytic cells from desiccation and reactive oxygen stress (6). The production of biosurfactants such as syringafactin can alter the hydrophobicity of the leaf, to promote water availability (19). Within the apoplast, it is necessary to alter the local environment to suppress host immunity as well as promote water and nutrient availability (20, 21). Foliar lesions follow bacterial-induced water soaking in localized sites on leaves. These lesions subsequently appear necrotic and dry (22). Lesions can then serve as a source of inoculum for subsequent infections.

High-throughput parallel experiments to identify bacterial virulence traits

High-throughput DNA sequencing has dramatically altered the landscape of methodologies to quantify gene expression, measure the abundance of single cells and populations, and calculate taxonomic diversity. Genomic comparisons between strains can reveal differential gene presence, as well as those genes that are under putative positive selection, suggesting that they might be contributing to habitat-specific fitness (23–25). Quantification of gene expression through both direct and indirect methods can reveal those genes, and thus mechanisms, that contribute to survival and growth in a given environment. Most relevant to this dissertation, the use of transposon mutagenesis followed by insertion site sequencing (TnSeq) quantifies the fitness contributions of nonessential genes by measuring the relative changes in insertional mutant abundance after a selective pressure is applied (26). Additionally, when diverse insertional mutant strains are generated on a rich medium, genes lacking recoverable insertions in the mutant population are most likely essential or nearly so in that condition (27).

TnSeq is a relatively laborious and costly method since both the site localization of a transposon in a particular gene and the quantification of the number of such insertions in a population need to be ascertained in every experiment. Random-barcoded TnSeq (RB-TnSeq) was developed to address some of the shortcomings of TnSeq (28). In RB-TnSeq, the strain of interest is transformed using a collection of transposons, each containing a unique 20-nucleotide "barcode" sequence. When transposon insertions are mapped to the genome, these barcodes are thus also mapped to the insertion sites (28). The strength of this method lies in the ability to use amplicon sequencing to simply recover and enumerate the number of occurrences of that barcode in the population of mutants. By this process one can characterize the relative proportion of the barcodes, and therefore the proportional representation of each of the insertional mutant strains within the population. This change to using a previously mapped mutant library decreases the cost and effort required to perform multiple fitness experiments. RB-TnSeq has been used to screen over 32 diverse bacteria - each against dozens of growth conditions, such as the utilization of different carbon and nitrogen sources as well as their tolerance of toxic compounds including antibiotics and metals (29). The change in relative proportion of strains containing transposon insertions in a given gene is used to calculate a fitness score for that gene. Generating genome-wide fitness data at this scale is fundamental to assigning functions to unknown genes, and enables the correction of erroneous genome annotations based on genetic screens rather than merely computational assignments based on gene homology. The use of RB-TnSeq to better understand the genes required for bacterial colonization of plants is a major focus of this dissertation. Prior work *in planta* has utilized transposon mutagenesis, TnSeq, and related methods to examine virulence and other plant colonization genes in diverse pathosystems (30–32). To date, RB-TnSeq has been used only in the rhizosphere, to determine genes required for root colonization (33), with far less known about factors contributing to bacterial colonization of the phyllosphere.

Specialized traits are required for apoplastic colonization

Studies *in planta* using heterogeneous bacterial populations require large overall populations to maintain mutant strain diversity and avoid bottleneck effects that will skew fitness calculations. Insertional mutant populations used for RB-TnSeq can be much larger than those used for TnSeq since such mutant libraries contain multiple insertions per gene. When inoculated into the host, the presence of an excessively large number of bacterial cells can trigger an immune response that would limit the overall growth of the library. This is an important issue since growth of the population is required to generate a measurable change in relative mutant abundance. Conversely, the presence of too few cells in inoculated plants would remove mutant strains from the population prematurely by chance, thus providing a false indication of the importance of that gene in the colonization process. Such an effect would decrease the statistical power of the assay for estimating the fitness contribution of a given gene, and potentially reduce the number of genes for which we could generate fitness values. TnSeq experiments *in planta* have gotten around these challenges in the past by using relatively small mutant libraries, containing many fewer unique mutant strains than used

here. In Chapter 2, I detail our initial studies using the B728a barcoded transposon library inoculated onto the leaf surface and apoplast of common bean (*Phaseolus vulgaris*). An important component of this study was to develop and optimize methods that balanced the need for extensive growth of bacteria in or on plants with the simultaneous need to introduce sufficiently large numbers of mutants disrupting a given gene to avoid bottleneck effects.

Many traits have been implicated in phyllosphere colonization, including type III effector proteins, toxins, siderophores, adhesins, and stress tolerance (34). While many genes contribute generally to stress tolerance and survival in varied environments, canonical virulence factors such as type III effectors and phytotoxins contribute to the suppression of the host immune system and are specifically important in the apoplast (35). While we expected genes such as these to appear in our screen for genes contributing to apoplastic fitness, a screen of this magnitude had never been performed for a foliar plant pathogen. At this scale, we anticipated that RB-TnSeq would prove to be a sensitive assay of those genes contributing to fitness on the leaf surface and/or the apoplast. Furthermore, an assessment of important metabolic traits required for growth of bacteria *in planta* can indicate the nutrients made available to bacterial cells by plants in those locations. Molecules such as amino acids, organic acids, and sugars have been directly measured in the apoplast of common bean (36), but the presence of these nutrients does not guarantee that they are available to *P. syringae* or sufficiently abundant to make a meaningful contribution to its growth. We used RB-TnSeq to identify those genes in *P. syringae* that directly and measurably contribute to fitness, assessed via the relative growth of various mutants over a period of several days.

An indirect method to measure variation in host chemical defenses

Many bacterial phenotypes can be considered virulence traits because they are required for successful colonization of a plant host and/or lead to subsequent disease formation. In the apoplast, colonizing strains must be able to overcome the host immune response and acquire nutrients. The acquisition of nutrients in turn, apparently also requires the presence of water in the intercellular spaces (21). In Chapter 3, I examine the ability of B728a to tolerate chemical stresses, and how this trait varies in importance across different plant hosts. Specifically, we utilize the additional susceptible plant hosts lima bean and pepper (37, 38), to assess the importance of the multidrug resistance efflux transporter *mexAB-oprM*. This transporter has been shown to contribute to virulence of *P. syringae* (39), although efflux transporters have primarily been examined for their role in antibiotic resistance. I focus on multidrug resistance efflux transporters because these systems typically make the largest contribution to the tolerance of external chemical stresses. Because of the importance of antibiotics in medicine, the vast majority of information on the chemical defenses deployed by bacteria has investigated their tolerance of medically important antibiotics. The study, while also addressing the tolerance of some antibiotics by *P. syringae* also explores its tolerance to various other toxic secondary compounds made by plants.

Multidrug resistance efflux transporters have been studied extensively in the context of antibiotic resistance, but they are also important for stress tolerance in the environment. Efflux transporters can remove host-derived toxic substrates, making them important virulence factors (40). When MexAB is disrupted, the mutant strain is hypersusceptible to diverse antibiotics and a wide variety of other toxic chemicals (39). By exploiting the hypersensitivity of a MexAB mutant, we can use it as a bioreporter to examine chemical defenses *in planta*. In parallel growth assays in a variety of plant species that we found to be hosts for *P. syringae*, we could examine not just variation in overall pathogen growth by the wild type strain, but also whether the Δ*mexB* mutant strain is inhibited *in planta* relative to the wild type. Considering the complexity of metabolomics studies, this indirect method indicates those hosts that apparently differ in their chemical defenses.

A screen to identify variation in traits used to colonize alternative hosts

With the knowledge from our studies that strain B728a is capable of colonizing diverse plant hosts, we utilized the barcoded transposon library in the apoplast of additional plant species to determine the genes needed for colonization of the interior of those hosts. We hypothesized that the genes required to colonize these hosts would differ from each other. Building upon the experimental optimization of the use of the mutant library in common bean, Chapter 4 examines the genes that contribute to fitness in the alternative host plants lima bean and pepper. Our current understanding of host-specificity in *P. syringae*, has focused on variations in type III effectors that differ between strains pathogenic on different plant hosts (41). Various strains also produce different phytotoxins, although these can often perform redundant functions and their contribution to the host range of a strain is unclear (42). Studies of host range have been most often based on genomic comparisons between pathogenic and nonpathogenic strains, or those infecting common and unique host plants (25, 43). Unfortunately the function of many of the genes in the genomes of all bacteria, including plant pathogens, remains unknown, so the efficacy of such a strategy is limited. In expanding our *in planta* screen of barcoded transposon mutants to identify those encoding important virulence traits, we aimed to test the same mutagenized population in several diverse plant species to identify host-specific virulence traits.

This RB-TnSeq study can be viewed as a proof of concept – to use this new method to examine virulence traits for multiple host species in parallel. This method has also been optimized to identify those genes that are most important for tolerance of a given stress condition or for growth using various nutrients *in vitro*. However, as discussed previously, the plant environment presents multiple complex abiotic stresses. Biotic stresses such as completion and antagonism within microbial communities in the phyllosphere are not addressed here, but would provide interesting material for follow-up studies using the methods developed here. The RB-TnSeq method allows us to both determine those genes that contribute to fitness and growth in complex environments such as the phyllosphere, as well as to perform follow-up studies *in vitro* to better characterize the function of individual genes. In fact, its use in functional annotation of bacterial genes is among our most powerful assets, and it will be the linkage of gene

function with ecological contribution that should be the ultimate benefit of this powerful method.

Exploring seemingly redundant bacterial efflux transporters and likely substrates

P. syringae is found in diverse habitats that presumably provide highly diverse chemical environments to the bacterium. As such, it must presumably tolerate a range of chemically diverse and potentially harmful molecules that it would encounter. A primary resistance mechanism for diverse bacteria to such toxic compounds is the expression of one or more transporters that remove toxins from the cytoplasm of the cell. Bacterial efflux transporters can function both to achieve stress tolerance as well as contribute to virulence by secreting toxins or effectors of various kinds (40, 44). A given bacterial species commonly possesses a wide variety of efflux transporters, and they differ in the specificity of the toxins that they export. Multidrug resistance (MDR) efflux transporters are structurally diverse, comprising at least five distinct protein families: the major facilitator superfamily (MFS), the small multidrug resistance (SMR) family, the multidrug and toxic compound extrusion (MATE) family, the ATP-binding cassette (ABC) superfamily, and the resistance-nodulation-division (RND) family (45, 46). Many of these transporter classes contain both substrate-specific transporters as well as less specific MDR-type efflux pumps (47). The B728a genome encodes 668 transport proteins (48), but it is unclear to what extent the context in which the majority of these transporters are active and their function in these varying conditions. MexAB-OprM is the best-characterized RND transporter present in *Pseudomonas* species, and is required for full virulence in *P. syringae* strains B728a, DC3000, and 1448A (39). While this protein complex is proposed to secrete the iron-chelating molecule pyoverdine, it also contributes significantly to antibiotic resistance (45). In *P. syringae* strain B728a, genes in this operon are generally expressed at much higher levels than those encoding the other RND transporters, both in culture as well as in cells on the leaf surface and in the apoplast (49).

We hypothesized that the MexAB-OprM efflux system in B728a might be masking the activity of other homologous RND transporters, and that the deletion of this operon would increase its reliance on these potentially redundant transporters. To address this hypothesis, we utilized RB-TnSeq to compare the identity of genes required for toxin tolerance in both the WT and the Δ*mexB* strain. The Δ*mexB* mutant is generally more susceptible to a variety of diverse antibiotics and toxins (39). Antibiotic profiling of a hyper-susceptible mutant strain has been used in *Escherichia coli* to identify redundant homologous efflux transporters (50). We took this same approach in Chapter 5 to identify redundant efflux transporters *in vitro.* These experiments to determine the identity of the substrates of a given transporter and their overlap *in vitro* will generate a starting dataset that should ultimately provide insight as to the potential for efflux pumps to have differential roles in different environmental settings such as in various plant hosts. In addition to multiple efflux transporters sharing the same substrates, Gram-negative cells have two membranes that these pumps must span. As several MDR-type efflux transporters are located in the inner membrane, they must pair with outer membrane pores or efflux system components in order to export toxic

compounds from the cytoplasm to the exterior of the cell (51). There is limited evidence that inner membrane transporters (such as those in the MFS, MATE family, and SMR family) can function cooperatively with outer membrane transporters (52). This has been proposed to synergistically increase overall toxin resistance for the cell, by coupling such presumably specific inner membrane transporters with more promiscuous outer membrane components (53). This would prevent unnecessary or inappropriate efflux of cellular components, attenuating the proton motive force. If this were true, we would expect to observe that inner membrane transporters would each contribute strongly to resistance to only a few toxicants. The largest challenge to any test of this hypothesis is that it requires knowledge of the presumably limited substrate range of a given inner membrane pump. Fortunately, results of our studies have provided some evidence in support of such a model and will be discussed in Chapter 5. In the diverse habitats in which microbes such as *P. syringae* would be found, it seems likely that these abundant and partially redundant efflux transporters contribute differentially to fitness, and are thus maintained in the genome. Using RB-TnSeq, I was able to generate abundant data linking the many efflux transporters in strain B728a to some of their likely substrates. These studies can easily be expanded to other antimicrobial compounds or related bacterial strain to provide a more comprehensive insight into the chemical ecology of *P. syringae*.

The studies in this dissertation provide extensive genome-wide data on the genes used by *P. syringae* strain B728a to tolerate external chemical stresses and colonize various plant tissues and species. Through detection of these important genes required for fitness, it is possible to draw conclusions as to the specific molecular mechanisms required to tolerate the various stresses present for life in the phyllosphere. In addition, these foundational studies are easily expandable using the barcoded transposon libraries created here. For example, while I only examine genes that contribute to fitness in the phyllosphere, studies in other environments such as water and soil are now feasible by using the mapped RB-TnSeq libraries I have produced. Furthermore, testing of additional chemical stresses *in vitro* can be used to clarify the specific roles of hypothetical genes and those currently having only unknown functions. We would be particularly interested in testing additional antimicrobial compounds produced by plants that might be encountered by *P. syringae.* Using mutant strains, such as the hyper-susceptible Δ*mexB* mutant and auxotrophic strains examined here, as reporters provides some indication of the physical and chemical environment experienced by bacteria at the small spatial scales that bacteria typically encounter. For this reason, RB-TnSeq can be used to generate hundreds to thousands of reporter strains that can be tested in any environment in parallel. We believe that the work described in this dissertation sets the stage to dramatically expand our knowledge of molecular plant microbe interactions, and our understanding of genes that directly contribute to fitness in diverse habitats.

Chapter 2

Genome-wide identification of *Pseudomonas syringae* genes required for competitive fitness during colonization of the leaf surface and apoplast

Abstract

The foliar plant pathogen *Pseudomonas syringae* can establish large epiphytic populations on leaf surfaces before apoplastic colonization. However, the bacterial genes that contribute to these lifestyles have not been completely defined. The fitness contributions of 4,296 genes in *P. syringae* pv. *syringae* B728a were determined by genome-wide fitness profiling with a randomly barcoded transposon mutant library that was grown on the leaf surface and in the apoplast of the susceptible plant *Phaseolus vulgaris*. Genes within the functional categories of amino acid and polysaccharide (including alginate) biosynthesis contributed most to fitness both on the leaf surface (epiphytic) and in the leaf interior (apoplast), while genes involved in type III secretion system and syringomycin synthesis were primarily important in the apoplast. Numerous other genes that had not been previously associated with *in planta* growth were also required for maximum epiphytic or apoplastic fitness. 14 hypothetical proteins and uncategorized glycosyltransferases were also required for maximum competitive fitness in and on leaves. For most genes, no relationship was seen between fitness *in planta* and either the magnitude of their expression *in planta* or degree of induction *in planta* compared to *in vitro* conditions measured in other studies. A lack of association of gene expression and fitness has important implications for the interpretation of transcriptional information and our broad understanding of plant-microbe interactions.

Introduction

Many plant pathogenic bacteria are capable of extensive colonization of leaf surfaces before their entry and multiplication within the leaf. As such, epiphytic (leaf surface) populations on asymptomatic plants are considered a reservoir of inoculum, that under the appropriate conditions can lead to infection. The likelihood of disease could be predicted from the epiphytic population size of the pathogen several weeks before disease symptoms were observed (54). The ability to form large epiphytic populations therefore is a measure of success for such a pathogen, and factors that determine its ability to grow on leaves would be considered fitness factors. After entry into the apoplast, bacterial numbers often increase greatly and disease is associated with those sites in which large internal population sizes have been achieved (55). The ability to grow within the apoplast of plants is thus also a measure of its fitness. In addition to plant pathogens, a diversity of other bacteria and fungi typically colonize the surface of aboveground plant parts. Such commensal bacteria and fungi are, however, typically limited to epiphytic growth with only very small numbers of such taxa found within plant tissue as endophytes (17). It is presumed that the growth of epiphytic bacteria is supported by their consumption of a variety of carbon and nitrogen-containing compounds that transit from the interior of the plant to the leaf surface, or are deposited from external sources such as aphid honeydew or pollen (8, 9). A variety of mono- and disaccharides are thought to constitute the majority of the carbon containing compounds on leaf surfaces, with smaller amounts of other sugars, organic acids, and

amino acids also present (2, 56). The absolute amount of such nutrients on leaves is generally low, and the growth of epiphytic bacteria is typically carbon-limited (56). Furthermore, the abundance of such nutrient sources on plants is spatially heterogeneous (4, 6). Because of the apparent chemical complexity, and spatially heterogeneous chemical and physical features of leaves, those traits needed for epiphytic fitness remain largely uncharacterized (6). Only limited descriptions of the chemical and physical environment found within the apoplast of plants have appeared (36). While many of the nutrient resources on the surface are apparently also present in the apoplast, the chemical and physical environment there is largely unknown. Water availability apparently limits intercellular growth (57), and the ability of pathogens to induce plants to release water into the apoplast may be a major feature required for exploitation of this habitat (20, 21). While the apoplast provides bacterial cells protection from environmental stresses on the leaf surface, they are in intimate proximity to living plant cells, and thus subject to inhibitory responses by the plant mediated by the innate immune system (2, 17). Taken together, it is clear that a large repertoire of traits beyond resource acquisition, such as motility, habitat modification, and various interactions with the plant may be needed by a plant pathogenic bacterium to exploit both the leaf surface and the leaf interior. Unfortunately, very few such fitness traits beyond those associated with interactions with the plant immune system have been identified.

Pseudomonas syringae is a plant pathogenic bacterial species that includes strains pathogenic on a wide variety of different plant species (58). Most strains have a prominent epiphytic phase, especially on the plant species for which they can also cause infection. Strains of *P. syringae* are commonly found as epiphytes on a variety of both host-and non-host plants, both in agricultural systems as well as native plant communities (15, 16). Many strains are capable of catalyzing ice formation, and because they can be found in rainfall, pristine snow, as well as in water sources around the world, are thought to play an important role in the water cycle by initiating ice formation central to the precipitation process (16). Such a connection to precipitation may also serve as a vehicle for long-distance dispersal as well as a mechanism for migration to plants after dispersal (59).

The model strain *P. syringae* pv. *syringae* B728a (B728a) is a strong epiphytic colonizer that was originally isolated from common bean (*Phaseolus vulgaris*), and is capable of causing brown spot disease (60). It is the best-studied member within *P. syringae* phylogroup II. This monophyletic clade contains strains that are overrepresented in environmental samples and are generally better epiphytes than members of other clades (61, 62). In addition, phylogroup II contains many strains with broad host ranges (62). Strain B728a is also pathogenic on *Nicotiana benthamiana* (37) and pepper (*Capsicum annuum*) (38). Like other ice nucleation active strains of *P. syringae*, this strain contributes to frost damage in frost sensitive plant species by limiting their ability to supercool and avoid damaging ice formation (63). Strain B728a is also a model for phytotoxin production, and contains a much smaller type III effector repertoire than many other strains such as *P. syringae* pv. *tomato* strain DC3000 (42, 64). As such, it has been hypothesized that this lower type III effector repertoire belies an increased reliance on broad-spectrum toxins as well as ice nucleation ability that contributes to its broad host range and more general environmental distribution (65).

This robust epiphytic colonizer and ubiquitous plant pathogen is thus a useful model to examine traits needed for bacterial success in diverse environments.

The genes that are putatively the most ecologically relevant to the success of bacteria on plants have typically been identified on the basis of their transcriptional induction or expression in a given habitat (66). Measurements of gene expression, directly via microarray or RNAseq, or indirectly through reporter genes or *in vivo* expression technology (IVET), have been used in *Pseudomonas syringae* to identify genes that have host-responsive expression patterns (49, 67–70). Validation of the role of genes identified by this method however usually involves targeted disruption of such genes individually with subsequent assessment of changes in behavior. This is a laborious procedure that cannot be readily applied to the genome as a whole. The high variability of the population size of a given strain after inoculation especially on leaf surfaces makes the comparison of population sizes between mutants and parental strains difficult. Large numbers of replicate samples are required to distinguish differential growth of such strains unless they differ greatly (71). Moreover, transcriptional studies may be of limited use in identifying host-colonization genes due to the lack of correlation between gene expression and contribution of those genes to fitness that is often observed *in vitro* (72). Examples of such a lack of correspondence of gene expression and fitness contribution in some *Pseudomonas* species on hosts have appeared (73, 74) but it is unclear how prevalent such a lack of connection might be.

To ascertain the roles of individual genes in *P. syringae* during epiphytic and apoplastic colonization on a genome-wide scale, we utilized a highly parallel transposon-based genomic screen. A variety of techniques taking advantage of high-throughput sequencing have been used to identify genes contributing to host colonization (75, 76). For example, genomic comparisons can reveal differential gene abundance in strains and genes that are under putative positive selection, suggesting that they might be contributing to host-specific fitness (23–25). However, confirmation of the role of such genes typically requires laborious mutation analysis as discussed above. Random mutagenesis techniques can enable genome-wide, gene-specific fitness contributions to be measured in a given habitat. One such strategy employs transposon sequencing (TnSeq) wherein the relative proportion of a transposon mutant in a given gene within a mixture of such mutants is assessed both before and after the strain mixture experiences a given condition. The number and genome location of the mutants in the mixture is determined by determining the sequences adjacent to the transposon by high throughput sequencing in each experiment (26). Recently developed random-barcoded transposon sequencing (RB-TnSeq) (28) is a modification of TnSeq that enables a transposon library to be used more easily for multiple assays since tagged transposons are used in mutagenesis and need be mapped only once. Eliminating the need to re-map transposon insertions in each experiment dramatically reduces the effort to carry out fitness screens in multiple conditions by using a single library of mutants. In a transposon mutant pool where each transposon is linked to a unique 20-nucleotide barcode, insertion mutant fitness is calculated through amplicon sequencing of the barcode regions to calculate the relative abundance of a given strain. Change in barcode relative abundance over time is used as a proxy for strain fitness within the population. RB-TnSeq was recently used to identify genes required by *P.*

simiae for its invasion of *Arabidopsis thaliana* roots (33). In this study, we used RB-TnSeq to identify genes in *P. syringae* needed for its colonization of both the surface and interior habitats of bean. Since the stimulon for these two habitats had previously been determined (15), we also addressed the extent to which transcriptional changes in gene expression were predictive of the fitness contributions of these same genes.

Results

Adapting RB-TnSeq for an epiphyte and foliar pathogen

In order to screen for genes in *P. syringae* strain B728a contributing to host colonization, we generated a randomly DNA barcoded *mariner* transposon library using the *Escherichia coli* donor library created by Wetmore *et al* (28). The sequenced B728a mutant library consisted of 281,417 strains with insertions that map to the B728a genome, each containing a unique DNA barcode. Computationally removing insertional mutants outside the central 10 – 90% of coding region of a given gene resulted in 169,826 genic strains for analysis, with a median of 21 insertions per gene. The number of usable insertional mutants for each gene was correlated with the number of TA dinucleotide sites within each coding region (Pearson correlation coefficient r = 0.71) (Fig. 1). We analyzed fitness contributions for 4,296 of 5,137 (84%) protein-coding genes that harbored sufficient insertions for analysis.

The rich medium King's B (KB) was used for library recovery prior to plant inoculations, so overnight growth in this condition was used as the control against which growth of the mutants on the leaf surface and in the apoplast (Fig. 2) was compared. All experiments analyzed herein passed quality control metrics that were previously established for *in vitro* studies in (28). The requirements for a successful experiment include ≥ 50 median reads per gene and consistency in the calculated fitness estimate obtained from mutants with insertions in the 3' and 5' half of a gene (28).

For each experiment, an aliquot of the mutant library was grown to mid-log phase in KB (ca. 5 generations) and a sample of the library was taken immediately before inoculating either the surface or interior of plants (time0). After growth in each condition, cells were recovered from either the surface or interior of the plants and prepared for sequencing. Fitness for each strain was calculated as the $\log_2$ of the ratio of barcode abundance following growth in or on plants with that barcode abundance obtained initially at time0. Gene fitness is calculated as the weighted average of the individual strain fitness values (28). Insertions in the majority of genes did not change fitness as measured by relative barcode abundance in the population, and thus the fitness values for most genes were close to 0.

Identification of the essential gene set of B728a

For 920 genes, fitness could not be calculated due to a lack of sufficient insertional mutants. Of those, only 5 do not contain TA dinucleotide sites and thus are not accessible by mutagenesis with the *mariner* transposon we used (Table 2). 512 of these genes did contain at least one mapped insertion, but we were unable to calculate fitness values for them due to the small number of sequencing reads at time0 (Fig. 3), suggesting that they were relatively unfit *in vitro* compared to other mutants, and thus in low relative population size in the library. Based on analysis of the TnSeq data, we predicted at least 392 genes to be essential or nearly essential for B728a growth on LB,

as they contain numerous TA sites but do not contain any mapped insertion strains in our library. The 392 predicted essential genes include many annotated as being involved in translation (including tRNAs), energy generation, and cofactor metabolism (Table 1). Using the Integrated Microbial Genomes database (77), we identified orthologs in *P. aeruginosa* PAO1, *P simiae* WCS417, and *P. stutzeri* RCH2 for 2,349, 2,560, and 1,450 B728a genes respectively. Based on predicted essential genes in these strains (29, 78, 79), we identified the degree of overlap in predicted essential or nearly essential genes for growth on LB (Table 3). 181 genes are predicted to be essential or nearly essential in all 4 strains examined here. Of those genes with predicted orthologs, 48 to 133 were annotated as essential in B728a and nonessential in the comparison strain, or vice versa.

Disruption mutants with fitness defects in rich medium

The pooled library was initially generated in LB but inoculum for the experiments was grown on KB. While these are both rich media, the presence of yeast extract in LB apparently provides resources not found in KB and thus the growth of certain biosynthetic mutants was limited in KB. We identified 20 genes that were required for maximal growth in KB medium, 9 of which are involved in cofactor metabolism (Table 4). Thirteen of these 20 genes contributed to fitness in or on plants. For example, the ability to synthesize the cofactor biotin was required for growth in KB, suggesting its limitation compared to other components such as amino acids and carbon compounds in this medium. However, the presence of biotin auxotrophs in the library indicates that these mutants were able to grow in LB. This suggests that cofactors such as biotin are lacking in KB medium, as well as in the *in planta* habitats.

Maintaining genetic diversity of the transposon library *in planta*

A major challenge for the use of complex mixtures to study the relative fitness of component strains in any experiment, especially studies done *in planta,* is to ensure that all of the mutants in the mixture are well represented after inoculation so as to avoid bottleneck effects. The strength of saturated transposon mutagenesis methods lies in internal replication: the contribution of each gene is assessed by interrogation of the behavior of multiple independent insertional mutant strains. A loss of diversity at the time of inoculation reduces the statistical power for analysis of a given gene. We aimed to maximize the total number of inoculated bacterial cells to maintain population diversity, while achieving a sufficiently low initial inoculum in or on plants so that substantial, competitive growth of the mixture could be assured. Given that the mutant library contains over 200,000 distinctive members, we inoculated a total of approximately 10^7 cells either into or on a collection of plants. So as to achieve a sufficiently low initial population size (ca. 10^5 cells/g leaf) that would enable a maximal number of generations to occur in and on leaves over which growth defects could be estimated, we chose to inoculate a large number of leaves (100 pots of plants). Although we observed slight bottlenecks given the concentration of inoculated cells we used, particularly in apoplastic conditions, these samples provided sufficient reads for most mutants to enable analysis of the fitness contribution of nearly all genes (Table 5); more than 80% and 68% of the unique barcoded mutants were retained in studies of

epiphytic and apoplastic growth, respectively. More than 99% of the unique barcoded mutants in the library were retained during *in vitro* experiments.

During the growth of strain B728a on leaf surfaces for 2 days the total number of cells increased approximately 100-fold to 10^7 CFU/g (49), corresponding to 6 to 7 population doublings. Similarly, during growth in the apoplast for 6 days population size increased about 1,000-fold to over 10^8 CFU/g (Fig. 4), indicating at least 10 cell divisions. In theory, in an experiment on leaves in which most strains exhibited 6 generations of growth, mutants completely incapable of growth should exhibit a fitness of about -6 (28). In practice, insertions in very few genes exhibited such an extreme lack of fitness (Fig. 5). While several genes, such as those conferring auxotrophy upon disruption, had large negative fitness values *in planta*, these values were less than that of the number of generations to which the population as a whole underwent, perhaps due to the consumption of rare nutrients or to residual cellular resources from the period of population recovery. 80 genes were identified as contributing strongly to growth, in which the corresponding mutants exhibited fitness values < -2 (exhibiting only 25% as much growth in the population relative to that of the typical strain in the mutant population) (Fig. 6a). Mutants in an additional 69 genes exhibited fitness values less than -1 but greater than -2 (Fig. 6b) suggesting that these genes contributed somewhat less to fitness (28). We did not normalize fitness values by the average number of generations in a given experiment, as these values are difficult to estimate and likely vary by plant within an experiment.

Despite the differences in the number of generations in epiphytic versus apoplastic growth in plants (49), we observed similar ranges in overall fitness values for individual genes in these two habitats (Fig. 7). Mean fitness values ranged from -4.6 to +1.7 (epiphytic) and -6.0 to +1.8 (apoplastic). For each gene, we averaged fitness values for the 2 replicate growth experiments in KB and the 3 epiphytic and 3 apoplastic experiments performed. We focused our analysis on genes contributing most strongly to fitness - those for which mutants had an average fitness < -2 and for which the *t*-score was < -3 in at least two replicate experiments. Since the plant host constitutes a more variable environment than most *in vitro* experiments, and expecting that many genes would not individually make large contributions to fitness, we also examined genes with fitness values < -1 but with $t < -3$. Special attention was placed on those genes with such values that are operative in a given metabolic pathway or could be placed in the same functional category. Approximately 50% of all genes exhibiting fitness values less than -2 or -1 in either epiphytic or apoplastic habitats were verified in at least two of three replicate experiments (Fig. 8).

Genes required specifically for colonization of the leaf surface

We identified 31 genes that were highly important for fitness on the leaf surface (Table 6), although all but 8 were also important in the apoplast. Among these 8, genes in the predicted operon *Psyr_2461-2* had a particularly strong epiphytic phenotype, with average fitness values of -2.1 and -3.2. *Psyr_2461* is a hypothetical protein containing a domain of unknown function (DUF934) and *Psyr_2462* is homologous to the sulfite reductase *cysI* in *P. aeruginosa.* Glutamate synthase (NADPH) subunit genes *gltB* (*Psyr_0411*) and *gltD* (*Psyr_0412*) also contributed strongly to epiphytic growth, having

average fitness values of -2.0 and -1.1. Disruption of the putative phage-related protein Psyr_4512 also strongly reduced epiphytic fitness (average fitness value = -2.1).

Genes contributing specifically to colonization of the leaf apoplast

Disruption of many genes encoding known virulence factors, including those in the type III secretion system (T3SS) (Fig. 9) and phytotoxin biosynthesis genes greatly reduced the growth of *P. syringae* in the apoplast. Of the 65 genes that were highly important (average fitness < -2) for apoplastic colonization (Table 7), 36 were important in this habitat but not on leaf surfaces. The T3SS genes we observed as essential for successful apoplastic colonization are exclusively involved in the T3SS machinery, as transposon insertions in most individual effector genes generally had no fitness phenotype (Fig. 7). Of the secreted type III effectors, *hopAB1* had the largest negative average fitness value (-0.55). While the fitness contribution of this gene was less than many others, growth of mutants in this gene was decreased in all three experimental replicates (standard deviation = 0.065, $t < -3.5$ for all). As *t*-scores are positively correlated with measures of fitness, the low variance in fitness seen among the 21 insertional mutants for this gene provide confidence in the rather modest fitness estimates for this gene.

Production and secretion of the phytotoxin syringomycin was required for competitive fitness in the apoplast. The syringomycin regulator *syrP* and syringomycin efflux transporter *syrD* both had large apoplastic-specific phenotypes when disrupted. In contrast, syringopeptin and syringolin mutants did not have significant apoplastic fitness defects in our experiments. Polysaccharide synthesis and regulation was highly important for competitive fitness in the apoplast. Mutants in alginate regulation (*algU*) and biosynthesis were dramatically less competitive than the typical mutant in the library. Group 1 glycosyltransferase encoding genes (*Psyr_0920* and *wbpYZ*) also contributed substantially to apoplastic growth (Fig. 7).

The two-component system GacA/GacS was moderately important in the apoplast (average fitness = -0.9 and -1.5), but interestingly, their disruption resulted in increased fitness on the leaf surface (average fitness = 1.3 and 1.7). Conversely, glutathione synthase (*gshB*) was important in KB (average fitness = -1.5) and on the leaf surface (-1.2), but disruption of this gene increased competitive fitness in the apoplast (+1.8). Generally, however, insertional mutations rarely significantly increased fitness in any experiment.

Genes required for the colonization of both the leaf surface and apoplast

Overall, the categories of “amino acid metabolism and transport”, “polysaccharide synthesis and regulation”, and “nucleotide metabolism and transport” were enriched in genes with average fitness less than -2 in both epiphytic and apoplastic habitats relative to that in rich medium (Table 8). We identified 31 genes that were highly important for epiphytic colonization, and 65 genes that contributed to apoplastic growth. Approximately 1/3 of all genes contributing to leaf colonization were also important in the apoplast (Fig. 6).

Genes required for the biosynthesis of several different amino acids were highly important in colonization of both the leaf surface and the apoplastic space. Genes required for biosynthesis of tryptophan, proline, and the shared biosynthetic pathway of

isoleucine/leucine/valine were among those with the largest contributions to fitness in both *in planta* conditions tested. Additionally, genes involved in biosynthesis of methionine were important for epiphytic survival, as seen previously (80), and disruption of these genes caused even greater decreases in apoplastic growth suggesting that these resources are in low abundance in these habitats. For example, average fitness values for *metW* and *metZ* auxotrophs were < -4 in the apoplast, but approximately -0.8 on leaf surfaces. A similar, albeit less dramatic, pattern of proportionally larger requirements for histidine biosynthesis under apoplastic growth was also seen. The production of cofactors such as pantothenate (vitamin B_5) requiring *panC* was important in both *in planta* conditions but contributed more to the growth on the leaf surface. Genes involved in nucleotide biosynthesis (*purFL*) were also highly important for growth both in and on leaves.

The genes *mdoGH* encoding glucan synthesis were required for optimal growth on both the leaf surface and in the apoplast. Hypothetical proteins encoded by *Psyr_0532*, *Psyr_2461*, and *Psyr_4158* (*eftA*) all made significant contributions to fitness both epiphytically and in the apoplast. Psyr_0532 contains a group 1 glycosyltransferase domain.

Validation of fitness estimated in disruption mutant mixtures with targeted deletion strains

To determine whether the growth deficiencies of mutant strains in the pooled assays were predictive of that when grown in isolation, we constructed targeted deletion mutants of several genes that contributed differentially to apoplastic fitness of *P. syringae* (Fig. 10, 11, 12). Amino acid auxotrophs Δ*trpA* (average apoplastic fitness = -2.7) and Δ*hisD* (average fitness = -3.0) were inoculated individually into the apoplast. While Δ*trpA* was almost incapable of growth, the population size of Δ*hisD* was about 10-fold lower than the WT strain 4 days after inoculation. Similarly, while a Δ*hrpL* type III secretion mutant (average fitness = -2.2) achieved an apoplastic population size that was only about 1% that of the WT strain, the population size of a Δ*syrP* mutant deficient in production of syringomycin (average fitness = -2.1) was only slightly lower than that of the WT strain when inoculated separately into plants (Fig. 13).

We also assessed fitness of directed mutants of three poorly understood genes, two of which are predicted group 1 glycosyltransferases, that contributed substantially to competitive apoplastic growth. Deletion mutants of Δ*eftA* (average fitness = -1.4, a hypothetical protein), and Δ*Psyr_0920* (average fitness = -2.4, a group 1 glycosyltransferase) both achieved apoplastic population sizes that were 10-fold lower than that of the WT strain when assessed both 4 and 6 days after inoculation. In contrast, the apoplastic population size of Δ*Psyr_0532* (average fitness = -1.6, a hypothetical protein containing a group 1 glycosyltransferase domain), was only slightly less than that of the WT strain (Fig. 13).

Fitness contributions of genes do not correlate well with their level of transcriptional expression or inducibility in or on plants

We compared previously published global transcriptional patterns for the genes in strain B728a when grown on leaf surfaces and in the apoplast (49), as well as for genes in the related strain DC3000 grown in the apoplast of *Arabidopsis* (67) with that

of the fitness contributions of these genes measured here to determine how predictive gene expression was to the functional role of these genes in growth in and on leaves. Both the absolute levels of gene expression in various *in planta* conditions as well as that of the changes in expression of a given gene *in planta* relative to that in cells grown in a minimal medium (for B728a) or KB (for DC3000) were used as predictors. In general, while many genes exhibited substantial elevated or depressed expression on or in plants compared to that in culture medium, disruption of these same genes often had little or no impact on the competitive fitness of the mutant strain in this study (Fig. 14). For example, while many amino acid auxotrophs were significantly less fit on the leaf surface and leaf interior, expression of biosynthesis genes for tryptophan, histidine, proline, and methionine was not induced, and instead was repressed, in these habitats compared to *in vitro* conditions. Similarly, while genes involved in biosynthesis of the cofactor pantothenate (shared with biosynthesis of the branched amino acids) were required for competitive fitness, their expression was down-regulated *in planta.* Likewise although the expression of genes encoding several hypothetical proteins were strongly increased *in planta*, suggesting that they may play an important role in growth on plants, their disruption had little effect on the competitive fitness of these mutant strains. Exceptions to this lack of association between gene induction and contribution to fitness are the genes (*syrP* and *syrD*) required for the biosynthesis of syringomycin in the apoplast; these genes were highly up-regulated *in planta* and mutants in this gene cluster also were much less fit. Likewise, *scrB*, involved in sucrose metabolism, is strongly up-regulated specifically in the apoplast and was also specifically required for competitive fitness in that environment. Generally, however, examples of genes in which a concordance between absolute or plant-dependent levels of transcription and their fitness contribution *in planta* were rare.

Discussion

Competitive colonization assays are a very sensitive method by which differences in relative fitness can be assessed. In a phenotypically heterogeneous population, changes in relative proportion of a given member provide a direct assessment of relative fitness. A notable exception that would preclude such a process would be one in which there is the production of shared goods (such as siderophores) that can be co-opted by non-producers, as predicted by the Black Queen Hypothesis (81). Random mutagenesis methods such as TnSeq, in which insertional mutant strains are grown in large pooled mixtures, have the advantage of identifying conditionally important genes in a genome-wide manner. By being intrinsically parallel in their structure, the ability to readily distinguish and enumerate each of the individuals in such a mixture by RB-TnSeq provides both a high throughput and highly sensitive means by which relative fitness of the individual strains can be assessed. Furthermore, the creation of multiple independent insertional mutants for each gene provides substantial internal replication, increasing confidence in the fitness phenotype quantified for any given gene. An advantage that RB-TnSeq has over more classical TnSeq is that the association of a given transposon insertion within a gene need be done only once, since a random DNA barcode can then be unambiguously associated with that insertional event thereafter (28). Such a process then allows the use of the same barcoded transposon library for multiple experiments by simply sequencing and enumerating the

DNA barcodes, enabling repeated interrogation of the role of the genes in a species such as *P. syringae* in many different environmental settings. The utility of RB-TnSeq has been demonstrated by its application to a myriad of different bacterial species exposed to hundreds of distinct environmental settings, enabling functions to be assigned to many previously uncharacterized genes (29). Our demonstration of the utility of RB-TnSeq in this study should enable us and others to greatly expand the association of genes in *P. syringae* to the myriad of functions in which it might be expected to participate, in the many chemically and physically different habitats that it colonizes.

P. syringae is a robust colonizer of both leaf surfaces and the apoplastic space of the host plant common bean. In these habitats, this strain exhibited sufficient growth to enable RB-TnSeq to quantify the contribution of individual genes that directly contribute to competitive fitness in a heterogeneous population. It would be expected that the ability of such a method to resolve differences in fitness contributions of these genes would increase with the number of generations of growth that the population of mutants would have undergone during an experiment. Given the large number of genes for which some fitness contribution could be measured, we focused our analysis here on those genes having the largest fitness contribution. Furthermore, there is generally higher statistical support for the validity of fitness estimates for those genes contributing more to fitness (Fig. 9), given that they were large and reproducible across replicate experiments. Genes associated with somewhat lower, but consistent, fitness values (Table 9) are likely also biologically significant, and future studies can explore the roles of these genes during *in planta* growth in more depth. In the current study the high internal replication intrinsic to the barcoded transposon library, and the use of several replicate experiments for each *in planta* condition has provided a compelling list of broadly important genes for further analysis.

Transposon-based approaches have been useful in identifying essential bacterial genes in other taxa, although our knowledge of essential genes in *Pseudomonas* species is limited. Analysis of a 100,000 strain insertional library in *P. aeruginosa* identified 336 of the 5,606 genes to be essential in LB (79). Recent work has identified 473 genes as likely to be essential in *P. simiae,* 430 in *P. stutzeri*, and 325 to 442 in *P. fluorescens* (depending on the strain) (29). We identified 392 genes to be likely essential for B728a growth in LB, comprised of functional categories generally seen to be essential in diverse bacteria. Given that we could calculate fitness contributions for 84% of the protein-coding genes in the environmental conditions tested here, the proportion of genes found to be essential in *P. syringae* appears similar to that *E. coli* and *P. stutzeri* (28). Differences of essential gene predictions between related strains highlight the value of making essential gene predictions using multiple conditions or strains, such as this TnSeq study of *P. aeruginosa* grown on six different media (82).

The identification of genes involved in anabolic processes such as cofactor production and amino acid biosynthesis that contribute to the fitness of *P. syringae* in a given habitat provides some insight into the availability of such resources in that setting. This logic of anabolic mutants as reporters of habitat resources provides insight into the resources on the surface and in the intercellular spaces of plants. The finding of fitness defects for many amino acid biosynthetic genes is a clear example of this concept. The much lower fitness of auxotrophs for several different amino acids suggests an acute

limitation of these essential metabolites both on the leaf surface and in the apoplast. Genes within the biosynthetic pathway of tryptophan had the largest effect on fitness when disrupted in our study, both on the leaf surface and in the apoplast. This is consistent with a previous study showing tryptophan synthesis as critical for both epiphytic and apoplast growth of B728a (83). Likewise, amino acids such as methionine and histidine are present in low levels in the apoplast, imposing a need for their synthesis by bacterial colonists (36, 84). Similar to these observations, biosynthetic genes for tryptophan (and its precursor anthranilate) were identified as important for fitness in a TnSeq study examining *Pantoea stewartii* colonization of maize xylem (31). A TnSeq screen in *Dickeya dadantii* in rotting plant tissue also noted a significant decrease in competitive fitness *in planta* for leucine, cysteine, and lysine auxotrophs that could be negated through the external addition of amino acids (30). Since most amino acids are apparently present at relatively low concentrations in the bean apoplast (36), it could be expected that many auxotrophs are incapable of growth without the ability to synthesize these essential non-substitutable metabolites, a model supported by the observations of this study.

Unlike certain anabolic genes, those involved in catabolism typically had more subtle phenotypes when disrupted. This is likely due to the presence of diverse and substitutable carbon and nitrogen sources such as sugars and organic acids in and on plants (36). It was noteworthy that the fitness of mutants in which sucrose 6-phophate hydrolase encoded by *scrB* was disrupted was lower in the apoplast (fitness value -1.7). Such an observation is consistent with sucrose being the most abundant sugar in intercellular spaces (36). On the other hand, the genes involved in the metabolism of compounds of lesser abundance that that are not essential would be expected to contribute somewhat incrementally to the fitness of *P. syringae* in or on leaves. While carbon availability appears to limit the growth of bacteria such as *P. syringae* on leaves (56) and might also limit the growth of this species in the apoplast, it might be expected that these various compounds represent substitutable resources (36), and that any given compound would contribute relatively little to the overall growth of such a strain if many were present in similar concentrations. In support of this conjecture, while small fitness defects were observed for several mutants harboring disruption of genes essential for catabolism of nutritive compounds, the magnitude of these fitness defects was usually low.

Many genes involved in polysaccharide synthesis and regulation were highly important in leaf colonization. The lipopolysaccharide found in the outer membrane of Gram-negative bacteria is known to induce the innate immune response of plant and animal hosts, yet it is required for bacterial stress tolerance in diverse environments (85). Many of the hypothetical proteins having strong plant phenotypes when disrupted here contain glycosyltransferase domains. We hypothesize that these hypothetical proteins contribute to the biosynthesis of O-antigen that decorates LPS, and thus might be involved in camouflaging the cells so as to not be perceived by plant surveillance systems. O-antigen is an essential virulence factor for *P. aeruginosa* in its colonization of animal tissues (86). While O-antigen is dispensable for growth in culture, it has been recently shown to delay the host immune response during *Xylella fastidiosa* colonization of grape xylem (87). Alternatively, glycosyltransferase activity may contribute to flagellar modifications in order to avoid plant recognition (88–90). However, this is unlikely to be

the major role of these selected genes in *P. syringae* since known flagellar glycosyltransferases in *P. syringae* (*fgt1* and *fgt2*) located adjacent to flagellar biosynthesis genes did not measurably contribute to competitive fitness *in planta* in this study. Nonetheless, we show that 8 genes containing glycosyltransferase domains made large individual contributions to host colonization, suggesting that there may be other important targets for such modification.

Biosynthesis of the exopolysaccharide alginate contributed strongly to growth in the apoplast but not on leaf surfaces. While alginate had been shown to contribute to epiphytic fitness and thus to subsequent disease severity of *P. syringae* (91, 92), the apoplastic colonization of mutants was not distinguished from epiphytic growth in those studies. The apoplast is thought to be a water-limiting environment for endophytic pathogens (20, 49, 57). Furthermore, the transcriptional activation of the key alginate biosynthetic enzyme *algD* is induced by high osmolarity (93) and thus alginate biosynthesis would be expected to contribute to fitness in the apoplast, as observed here. While the AlgU regulon includes genes involved in additional stress-responsive processes such as osmoprotectant transporters (OpuC and Cbc), antioxidant enzymes (KatE, SodC, and CpoF), and the type 6 secretion system (94), we did not observe any fitness decrease for insertional mutants in those genes. In *P. putida,* alginate production is required for biofilm-mediated survival under desiccating conditions (95). We did not see a significant role of alginate on the leaf surface, potentially due either its potential to be a shared good or to the high humidity used for these epiphytic studies that negated any benefit of water absorption. However, alginate biosynthesis was a clear virulence factor in the apoplast.

Our screen highlighted the fitness role of many known virulence factors including the type III secretion system and the phytotoxin syringomycin. It was encouraging to detect type III secretion system mutants, indicating the bacterial apoplastic densities were not high enough to complement these mutants *in trans*, perhaps because of the low spatial density of the inoculated cells within the plant. Individual secreted effector proteins did not generally contribute measurably to apoplastic colonization, while mutations in type III pilus genes significantly decreased fitness. This supports existing dogma, whereby type III effector proteins are individually dispensable and collectively essential (96). HopAB1, a secreted type III effector which we found to have the largest contribution to apoplastic fitness among all secreted effectors, has been shown to have a measurable contribution to B728a growth in the bean apoplast (37). While phytotoxin production in strain B728a has been shown previously to induce symptom formation, there has not been compelling data showing a contribution to bacterial growth in plants (97). It is interesting that the genes involved in syringomycin and syringopeptin have strong negative fitness values in our study, suggesting that mutants in these pathways are impaired in growth relative to wild type. Our results support the model that *P. syringae* strains such as B728a with relatively fewer type III effectors have an increased reliance on phytotoxin production for growth in the apoplast (43, 64). The biosynthetic gene cluster for syringomycin also was distinctive in that it was among the few genes that are both up-regulated *in planta* (15) and contribute to fitness. On the other hand, genes for syringolin biosynthesis while up-regulated in the apoplast (49), did not contribute to apoplastic fitness in our study. Syringolin contributes to virulence through host proteasome inhibition, which has been shown to counteract stomatal innate

immunity (98). Therefore, the role of syringolin is likely limited to the transition from epiphytic to apoplastic growth, a process that was not tested here. Syringopeptin, which is also up-regulated in the apoplast (49), contributed to apoplastic fitness to a much lesser extent than syringomycin. Syringomycin and syringopeptin have the same mechanisms of action, creating membrane pores and causing ion leakage (99). It is unclear why B728a produces two seemingly redundant phytotoxins, although it has been proposed that their differential antimicrobial activities contribute to epiphytic survival (100). Both syringomycin and syringopeptin contribute to virulence on cherry (99) and lysis of tobacco protoplasts (100). Since we observed a much larger contribution to bacterial fitness in the bean apoplast from syringomycin than syringopeptin, it is tempting to speculate that the functions of these phytotoxins in virulence may be somewhat host specific.

Despite the dogma that gene expression is fine-tuned to the metabolic demands of a cell, recent studies of gene expression have shown such regulation to be suboptimal for many bacterial species (72). Despite classic examples of biosynthetic pathways in *E. coli* having adaptive regulation, many genes in diverse bacteria show little correlation between when they are important for fitness and when they are most highly expressed (72). For example, constitutive expression and regulation by growth rate are common indirect gene regulation strategies that occur for genes with diverse functions and yet are often suboptimal in the laboratory and presumably also in natural environments (72). In *P. aeruginosa* wound infections, gene expression was also not well correlated with gene contributions to fitness (74). A proposed explanation for such incongruence was that given that *P. aeruginosa* is considered an opportunistic pathogen that might not have evolved primarily in association with mammalian tissues its patterns of gene expression might have optimized fitness in very different settings (74). Moreover, in persistent, long-lasting infections such as the cystic fibrosis lung, adaptive changes in global patterns of gene expression in *P. aeruginosa* have been observed over time (101), reflecting adaptation to this new habitat.

While *P. syringae* is a model plant pathogen, it is also commonly observed in many other environmental settings (62). The conditions that the cell would experience on the surface of the plant are likely to be quite different from those in the apoplast (17, 58). Here, we see no correlation between gene expression (either absolute or relative change) and contribution to fitness in the host. It was surprising that genes that were highly expressed and/or highly induced in cells in or on leaves did not make large contributions to the fitness of the strain. Studies of gene expression are very common, and tend to focus on those genes that are induced in a given environment or condition. However, calculations of inducibility are inherently relative to the reference condition, often a rich or minimal medium. As the calculations of B728a gene induction *in planta* was performed relative to hrp-inducing minimal medium, we reasoned that some genes would perhaps already be induced. To get around this, we also compared our fitness data to gene expression measurements in the closely related strain DC3000 growth *in planta* relative to the more typically used rich culture medium. While the timing of sampling of RNA from the apoplast for these comparative studies was somewhat earlier in the infection process (Yu *et al*. (49) sampled bacterial cells 2 days post inoculation and Nobori *et al.* (67) sampled after 6 hours, while we sampled after 6 days), we would not have expected temporal changes in gene expression to overwhelm any context-

dependent patterns of gene expression. Despite the different conditions used as reference points, the overall trends were similar between these two strains *in planta* relative to the contributions of those genes to fitness measured here. Likewise, many genes that were either weakly expressed or un-induced on or in plants proved particularly important for fitness in these habitats. This lack of congruence can be explained by the fact that many genes are involved in catabolic processes wherein individual pathways would be expected to contribute only incrementally to the success of a strain. Genes for anabolic pathways, on the other hand, might prove essential irrespective of how highly expressed they are. There remain many genes for which the lack of linkage between expression and contribution to fitness remain unexplained. It is evident that directly measuring the contribution of a gene to fitness in different environments is a necessary complement to global transcriptional profiling to understanding the function and behavior of a cell in a given setting.

Although *P. syringae* is a ubiquitous species, it is most commonly studied in its agriculturally relevant, disease susceptible plant hosts. Random mutagenesis studies typically observe that a majority of genes in the genome are dispensable, as seen in the relatively small number of essential genes across diverse bacteria (102). This is likely due to functional redundancy in addition to many genes contributing to bacterial fitness in untested habitats outside of the laboratory (29). Although previous transposon screens in *P. syringae* have provided information on traits required for epiphytic fitness and virulence, these have either uncovered only those genes with large effects on behavior (103, 104). RB-TnSeq greatly expands our ability to interrogate the ecological determinants of such a cosmopolitan bacterium. Testing *P. syringae* and other bacterial species in a range of conditions, including those of ecological relevance such as on and in additional host and non-host plants, will enable the designation of functions for hypothetical or otherwise uncategorized proteins. Comparisons of these fitness assessments with specific *in vitro* experiments will enable the dissection of how individual genes contribute to a given process and to fitness on a eukaryotic host, a complex habitat with many distinct abiotic and biotic stressors. In such an approach, Cole *et al.* used this method to examine specific nutrient requirements for *P. simiae* colonization of *Arabidopsis* roots (33). Many of the genes found to contribute to fitness had only small effects *in planta*. Expansion of these screens through additional generations of growth will increase the accumulated fitness defects, as seen in a recent study that sequentially passaged a *Caulobacter crescentus* transposon library to identify genes affecting attachment (105). Barcoded transposon libraries were originally developed as a highly scalable tool to identify gene function in diverse *in vitro* conditions such as different growth conditions or abiotic stresses. Here we show that these same libraries can be used to better understand conditionally important genes that contribute to growth on the leaf surface and during colonization of the apoplast, expanding our understanding of the ecological fitness requirements on a genome-wide scale.

Materials and Methods

Bacterial strains and growth media.

P. syringae pv. *syringae* B728a was originally isolated from a bean leaf (*Phaseolus vulgaris*) in Wisconsin (60). The complete genome for B728a is available on NCBI GenBank as accession CP000075.1 (106). B728a and derivative mutant strains

were grown on King's B agar or in broth (107), at 28˚C. *E. coli* strains S17-1, TOP10, and XL1-Blue were grown on LB agar or in LB broth at 37˚C. *Geotrichum candidum* Fr260 was grown at 25˚C on Potato Dextrose Agar (BD Difco). When appropriate, the following antibiotics were used at the indicated concentrations: 100 µg/ml rifampicin, 50 µg/ml kanamycin, 15 µg/ml tetracycline, 40 µg/ml nitrofurantoin, and 21.6 µg/ml natamycin.

Construction of bar-coded transposon library.

The bar-coded transposon insertion library was constructed by transposon mutagenesis using a bar-coded *mariner* transposon library, followed by TnSeq mapping and barcode association as previously described (28). The *E. coli* WM3064 donor library containing the barcoded *mariner* plasmid, pKMW3, was recovered from a glycerol stock in LB kanamycin containing 300 µM diaminopimelic acid (DAP) and conjugated into B728a overnight on 50 different LB plates containing DAP. The many separate conjugations were then resuspended as a single pooled mixture and spread on 200 LB kanamycin plates for selecting mutants. Over 220,000 kanamycin resistant B728a colonies were pooled for the library. All colonies were resuspended in 250 ml LB kanamycin and diluted to a starting OD_{600} 0.2 for outgrowth at 28˚C with shaking to OD_{600} 1.0. Finally, 250 µl 80% glycerol was added to 1 ml aliquots and frozen at -80˚C.

Plant growth conditions.

Common bean (*P. vulgaris* var. Blue Lake Bush 274) seeds (5 to 7 per 10 cm diameter pot) were planted in Super Soil (Scotts Miracle-Gro) and grown in a greenhouse for two weeks before inoculation. Leaves were kept dry to minimize epiphytic bacterial populations. 1000 W metal halide lights were used to provide supplemental lighting for a 16 hour day length. Greenhouse temperatures ranged from 18˚C at night to as high as 40˚C during the day.

Library recovery and growth in KB.

For each inoculation, a 1.25 ml glycerol stock containing the transposon library was inoculated from -80˚C into 25 ml fresh KB with 100 µg/ml kanamycin and grown for approximately 7 hours at 28˚C with shaking until the culture reached mid-log phase, OD_{600} 0.5 - 0.7. Time0 samples were collected at this point during recovery; 1 ml aliquots were pelleted by centrifugation and the pellets were frozen until DNA purification. Cells were then washed twice in 10 mM KPO_4 (pH = 7.0) prior to plant inoculation. To assay library growth in KB, 50 µl log phase cell culture (OD_{600} 0.5) was inoculated into 950 µl KB with kanamycin in a 24-well plate. The plate was incubated at 28˚C with shaking for 15 hours. Cells were collected by centrifugation, and frozen prior to DNA purification.

Plant inoculations of the transposon library.

For epiphytic inoculations, cells were resuspended to a concentration of $2x10^6$ CFU/ml in 10 mM KPO_4 (OD_{600} = 0.001, by dilution from OD_{600} = 0.1), and sprayed onto the leaf surface until runoff. 100 pots were inoculated for each of three replicate experiments. Plants were then placed in a high humidity chamber for two days.

For apoplastic inoculations, cells were resuspended to a concentration of $2x10^5$ CFU/ml in 1 mM KPO_4. The soil was covered with cotton to hold it in place, and the pots were inverted in ~1.5 L inoculum in a glass bell jar. A vacuum was applied for 1.25 minutes and then removed rapidly to force the inoculum into the apoplast. Ca. 100 pots were inoculated for each of three replicate experiments. Plants were allowed to dry overnight and then moved to the greenhouse for six days.

Library isolation from the leaf surface.

Leaves were collected and placed in a glass dish filled with dH_20, which was placed in a sonication water bath (Branson 5510, output frequency 40 kHz), and sonicated for 10 min to remove cells. Approximately 1 L dH_20 + leaves was sonicated at a time. The resulting leaf wash was filtered through a 6 µm filter (Whatman #3), and then cells were collected on 0.2 µm filters. Cells were removed from the filters by vortexing in 30 ml total 10 mM KPO_4, and centrifuged at 17,000 x g for 1 minute to pellet. Cell pellet aliquots were frozen prior to DNA purification.

Library isolation from the apoplast.

Leaves were chopped in a blender to yield fragments of about 2-3 mm in diameter, and placed in a glass dish filled with dH_20, placed in a sonication water bath, and sonicated for 15 min to remove cells. Leaves were not surface sterilized to enable leaf sampling at this scale. Since the number of epiphytic cells on leaves kept at low humidity was several orders of magnitude lower than those in the apoplast the contribution of any epiphytic cells to the recovered apoplastic cells was negligible. Approximately 1 L leaf slurry was sonicated at a time. The resulting slurry was filtered through a coffee filter to minimize plant debris. 10% of the ~5-10 L buffer was taken for additional filtration steps. This sample was filtered through several Whatman filters (20 µm, 10 µm, and 6 µm), and then concentrated by centrifugation at 4696 x g for 10 minutes. The pellet was resuspended in 10 mM KPO_4, and aliquots of cell pellets were frozen prior to DNA purification.

DNA isolation and library preparation.

DNA from frozen pellets was isolated using the Qiagen DNeasy Blood & Tissue Kit according to manufacturer's instructions. Cell lysis was done at 50˚C for 10 minutes as per optional instructions. For apoplastic samples with excess plant material, lysed cells were centrifuged at 1,500 x g for 5 minutes before loading the supernatant onto Qiagen purification columns. Purified genomic DNA was measured on a nanodrop and 200 ng of total DNA was used as a template for DNA barcode amplification and adapter ligation as established previously (28). For each time0 and plant experimental sample, two separately purified DNA samples were sequenced as technical replicates.

Sequencing and fitness data generation.

Barcode sequencing, mapping, and analysis to calculate the relative abundance of barcodes was done using the RB-TnSeq methodology and computation pipeline developed by Wetmore *et al.* (28); code available at bitbucket.org/berkeleylab/feba/. TnSeq was used to map the insertion sites and associate the DNA barcodes to these insertions. Based on the TnSeq data, standard computational methods (29) were used

to predict which genes are likely essential for viability in LB. For these data, the minimum gene length to call a gene essential was 325 bp. For each experiment, fitness values for each gene are calculated as a $\log_2$ ratio of relative barcode abundance following library growth in a given condition divided by relative abundance in the time0 sample. Fitness values are normalized across the genome so the typical gene has a fitness value of 0. All experiments passed previously described quality control metrics (28), with the exception of one technical replicate of apoplast_2, which was removed from our analysis. Fitness values from sequencing replicates were averaged for each experiment. Experimental fitness values are publically available at fit.genomics.lbl.gov. Additional data, including BarSeq counts for all insertional strains, are publically available at github.com/thelmann/P.syringae-leaf-colonization.

Comparison of predicted essential genes in B728a to other *Pseudomonas* species.

We used the Integrated Microbial Genomes (IMG) database (77) to identify homologs for B728a genes in *P. aeruginosa* PAO1, *P. stutzeri* RCH2, and *P. simiae* WCS417 using the genome-gene best homologs function at 70% identity. Turner *et al.* predicted 336 essential genes in *P. aeruginosa* PAO1 using a Monte Carlo statistical analysis (79). Essential genes in *P. stutzeri* RCH2 and *P. simiae* WCS417 were predicted previously (29). A comparison of B728a genes predicted to be essential or nearly essential (N = 392) with their orthologs in these strains identified three categories: predicted essential and nonessential genes, as well as B728a genes with no orthologs identified.

Genomic fitness data analysis.

A dendrogram of experiments was generated from the matrix of fitness values using the dist and hclust functions in R (108) with the default clustering method "Euclidean". To better classify genes based on their genomic annotation, we assigned gene names, gene product descriptions, and broad functional categories based on the previously annotated genomic metadata (49). For each gene, fitness values for experimental replicates were averaged to calculate an average gene fitness value for each treatment. We focused our analysis on genes with average fitness < -2 and t < -3 in at least two experimental replicates. However, we also considered genes for analysis with average fitness < -1 and t < -2.5 in at least two experimental replicates. The t-score is a test statistic used to assess the statistical significance of the gene fitness value (28). For each functional category, we used a the phyper function in R to perform a hypergeometric test to examine category enrichment, using average fitness < -2. Graphs were plotted in R using the ggplot2 package, version 3.1.1 (109). Heatmaps were plotted in R using the function heatmap.2 in the gplots package, version 3.0.1.1 (110).

Construction of targeted deletion mutants.

Deletion strains were constructed using an overlap extension PCR protocol as described previously (111). Briefly, 1kb DNA fragments upstream and downstream of the genes of interest were amplified along with a kanamycin resistance cassette from pKD13 (112). These three fragments were joined by overlap extension PCR. The resulting fragment was blunt-end ligated into the SmaI site of pT*sacB* (113), and

transformed into the *E. coli* subcloning strains TOP10 or XL1-Blue, and then the *E. coli* conjugation donor strain S17-1. This suicide plasmid was conjugated into B728a on KB overnight, and then selected for 3 days on KB containing kanamycin and nitrofurantoin (*E. coli* counter selection). Putative double-crossover colonies that were kanamycin resistant and tetracycline sensitive were selected for screening using external primers and insertion sites further confirmed by PCR and Sanger sequencing.

Bacterial *in vitro* growth measurements.

Strains were grown overnight in KB broth with rifampicin, washed three times in 10 mM KPO_4, and standardized to OD_{600} = 0.3. 20 µl was inoculated per well into a 96-well plate containing 100 µl KB broth, LB broth, or M9 broth supplemented with 0.2% v/v glycerol (3 to 5 replicate wells for each strain). Auxotrophic mutant strains were functionally complemented by the addition of either 20.8 µg/ml tryptophan or 62 µg/ml histidine. The plate was sealed using a semi-permeable membrane (Breathe-Easy Sealing Film, Diversified Biotech). Cells were grown at 28˚C with shaking, taking absorbance measurements at 600 nm every 30 minutes. Absorbance values are reported, with the average absorbance for two replicate media blanks subtracted from the total absorbance for each well.

Bioassay of the syringomycin mutant strain.

Geotrichum candidum was grown at 25˚C for two days on a potato dextrose agar (PDA) plate. Bacterial strains were grown overnight in KB with rifampicin, washed in 10 mM KPO_4, and standardized to OD_{600} = 0.3. 20 µl suspension was spotted onto PDA plates, and grown at 25˚C for two days. After incubation, the plates were lightly oversprayed with a fungal suspension and zones of fungal growth inhibition were measured after 20 hours.

Bacterial apoplastic growth measurements.

Strains were grown overnight on KB with rifampicin, washed in 10 mM KPO_4, and standardized to $2x10^5$ CFU/ml in 1 mM KPO_4. Cells were inoculated into leaves of two-week old plants using a blunt syringe. Leaf samples (3 discs per leaf) were taken using a 5 mm-diameter cork borer into tubes containing 200 µl 10 mM KPO_4 and two 3 mm glass beads, and ground for 30 seconds at 2400 rpm in a Mini-Beadbeater-96 (Biospec Products) before dilution plating on KB with rifampicin and natamycin.

Comparison of gene expression and contribution to fitness.

B728a gene expression was previously determined as a measure of microarray fluorescence from RNA isolated from cells in a hrp-inducing minimal medium (MM), epiphytic growth, or apoplastic growth in *P. vulgaris* (49). These data were originally generated by two laboratories, and within-laboratory technical replicates were averaged before calculating the mean fluorescence for epiphytic and apoplastic treatments. Fold-change was calculated as a $\log_2$ of the ratio of epiphytic or apoplastic fluorescence divided by fluorescence in MM. Gene expression for strain DC3000 in the apoplast of *Arabidopsis thaliana* and in KB was previously measured by RNAseq (67). Fold-change for a given gene was calculated as a $\log_2$ of the ratio of apoplastic reads/million divided by reads/million in KB. Orthologs for 4,209 B728a genes were identified in strain

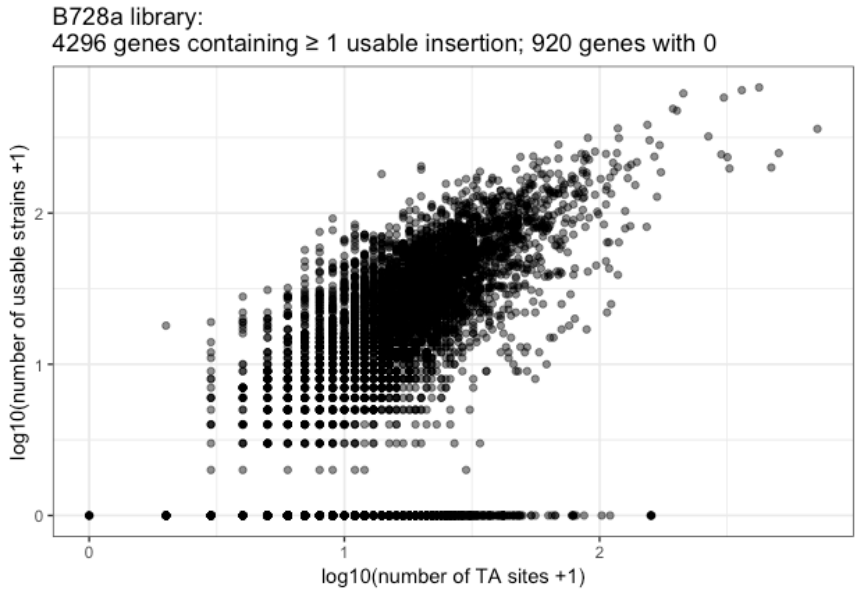
DC3000 using the IMG genome-gene best homologs function at 70% identity (77). Of these, fitness values could be calculated for 1,925 genes in B728a for which RNAseq expression patterns were available in DC3000, and were plotted here.

Acknowledgements

We thank Morgan Price for assistance with RB-TnSeq sequence analysis, and Jayashree Ray for mapping the insertions sites of the B728a transposon library. We thank Russell Scott for the B728a Δ*hrpL* strain and Dennis Gross for the *G. candidum* isolate. We thank Dana King, Caitlin Ongsarte, and Jennifer Lam for assistance with plant inoculations. Funding for Tyler Helmann was partially provided by the Arnon Graduate Fellowship and the William Carroll Smith Fellowship. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

Figures

A.



B.

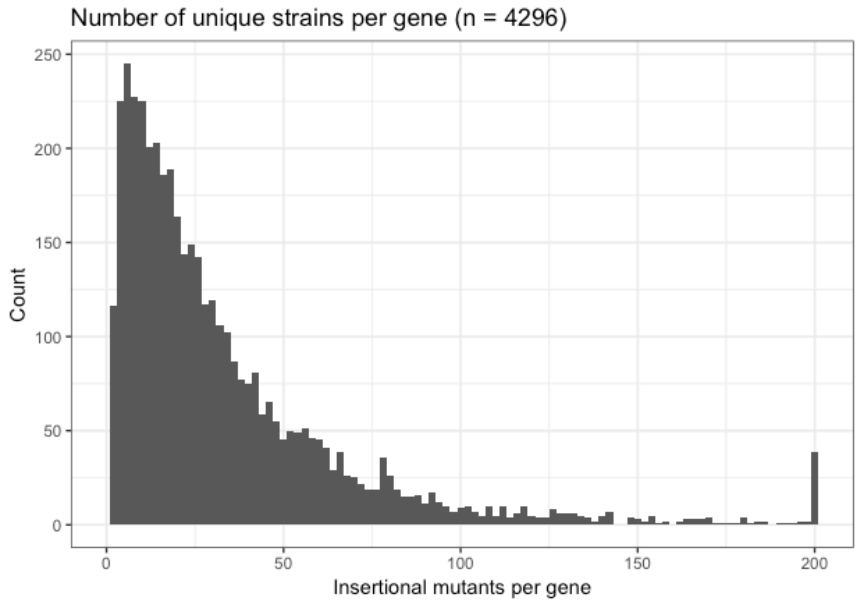


Figure 1. The B728a *mariner* transposon library contains 281,417 mapped insertions, 169,826 of which lie within the central 10 – 90 % of a gene. Only these central insertions are used to calculate fitness. (A) The number of usable insertions for each gene is correlated with the number of TA dinucleotide sites within each coding region. The Pearson correlation coefficient calculated for all genes having at least 1 usable insertion strain: r = 0.71. (B) 4,296 genes contained at least one usable insertion. Genes for which their contribution to fitness could be calculated are represented by a median of 24 insertion strains per gene. 38 genes with > 200 insertions each (range = 203 to 676) are plotted at 200 for clarity.

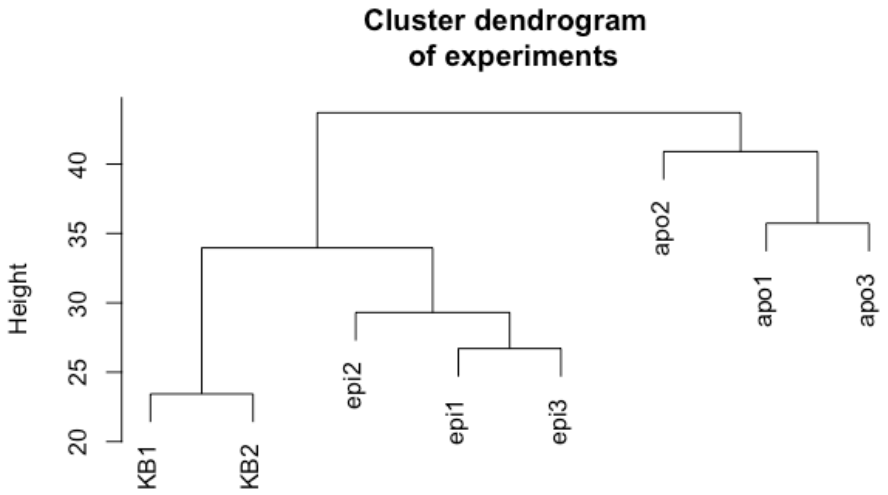


Figure 2. Dendrogram of experiments, generated using *P. syringae* gene fitness values determined in experimental replicates of the three conditions tested. Rich medium King's B (KB) experiments cluster more closely with the epiphytic experiments (epi) than the apoplastic experiments (apo).

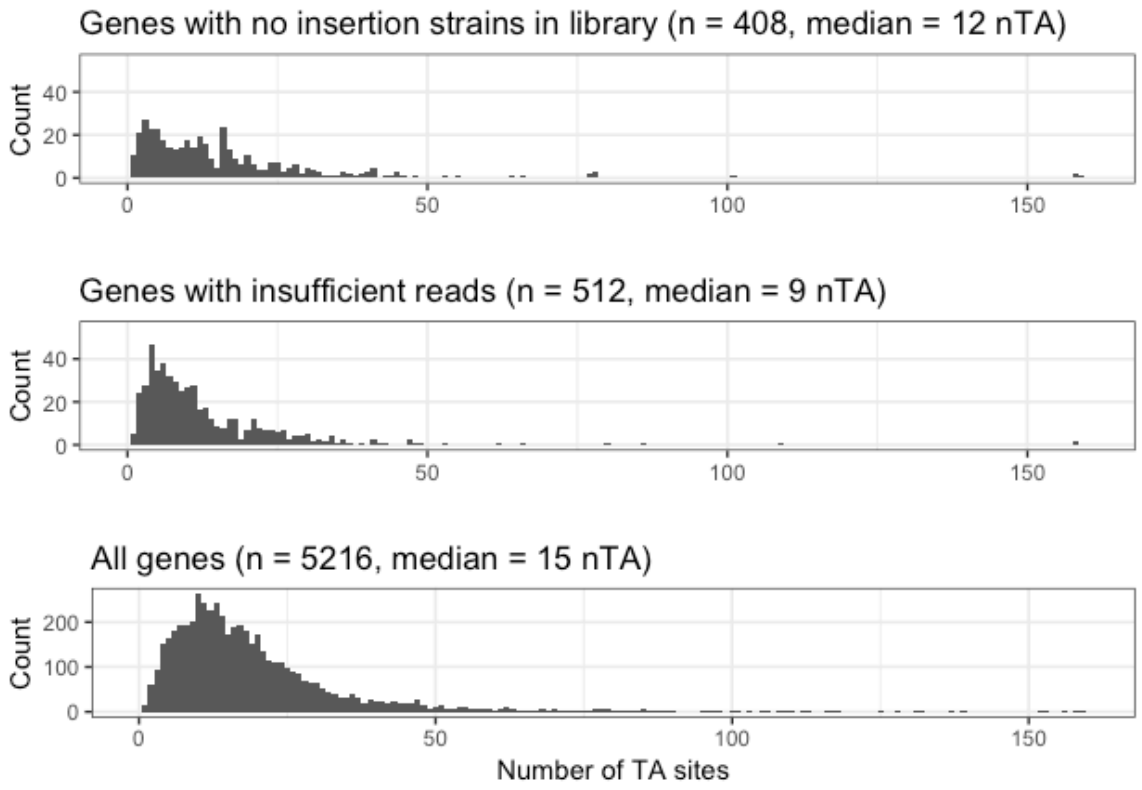


Figure 3. Of the 920 genes for which their fitness contribution could not be calculated, 408 have no insertion strains in the library, while 512 have at least one insertion but an insufficient number of sequencing reads in time0 samples to calculate fitness. These 920 genes range in size from 73 bp to 4199 bp, with a median size of 575 bp. 5 genes have no TA sites. The histogram of all genes excludes 17 genes containing 163 to 713 TA sites.

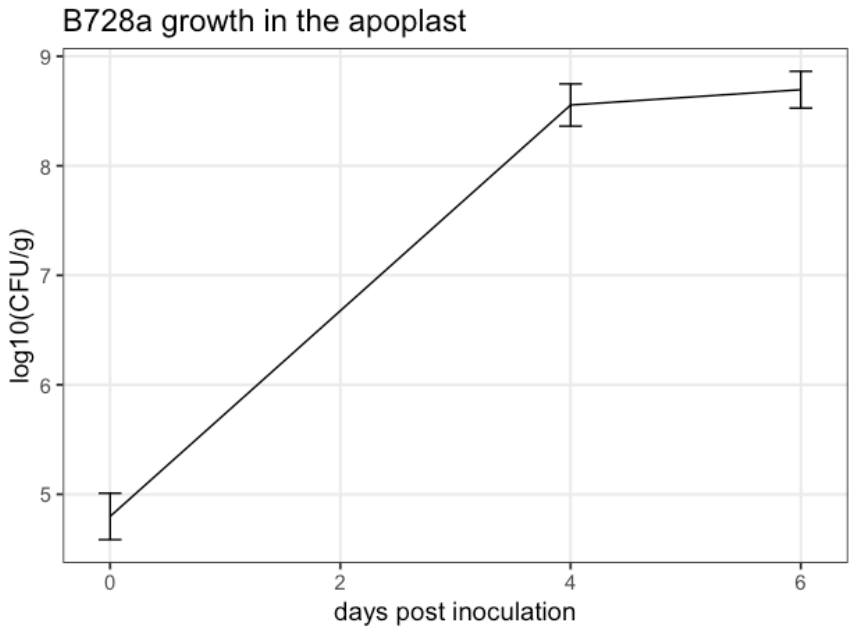


Figure 4. Representative growth of B728a in the apoplast. B728a WT cells were inoculated into the apoplast by submerging the plants in 1.5 L of inoculum containing 10^5 CFU/ml in 1 mM KPO_4 buffer and subjecting the plants to a vacuum for 1.25 minutes. Rapidly restoring atmospheric pressure forced the bacterial suspension into the intercellular spaces of the leaves. The plants were allowed to dry on a laboratory bench for at least five hours, and then moved to the greenhouse. Cells were recovered from 8 to 12 replicate leaves at each sampling time by homogenization of the leaves, dilution plating the homogenate on selective medium, followed by enumeration of colonies. Error bars represent the standard error of the mean of log-transformed population size.

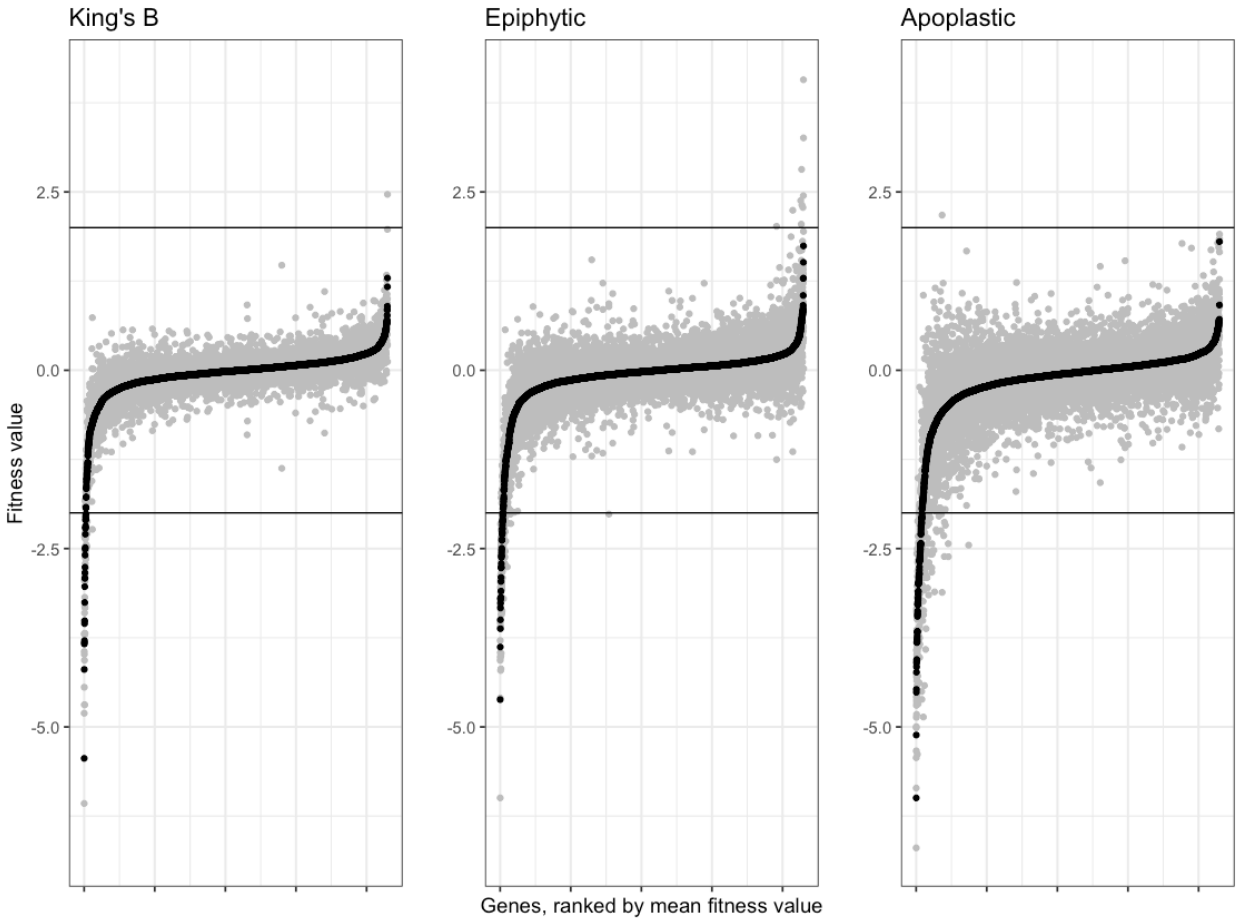


Figure 5. Rank ordered mean gene fitness values for each condition in which *P. syringae* was grown. Fitness values for independent replicate experiments are shown in grey, while mean fitness values are plotted in black. Gene fitness values are calculated as the $\log_2$ ratio of the barcode counts following growth in a given condition compared to the barcode counts from mid-log phase cultures following library outgrowth and before inoculation. Fitness values are normalized across the genome so the typical gene has a fitness value of 0. Black lines at fitness values of -2 and 2 are used to indicate strong phenotypes; for example a value of -2 indicates that mutants in that gene were 25% as fit as the typical strain in the mutant library. In each dataset, fitness values < -2 or > 2 are more than 3 standard deviations from the mean (approximately 0).

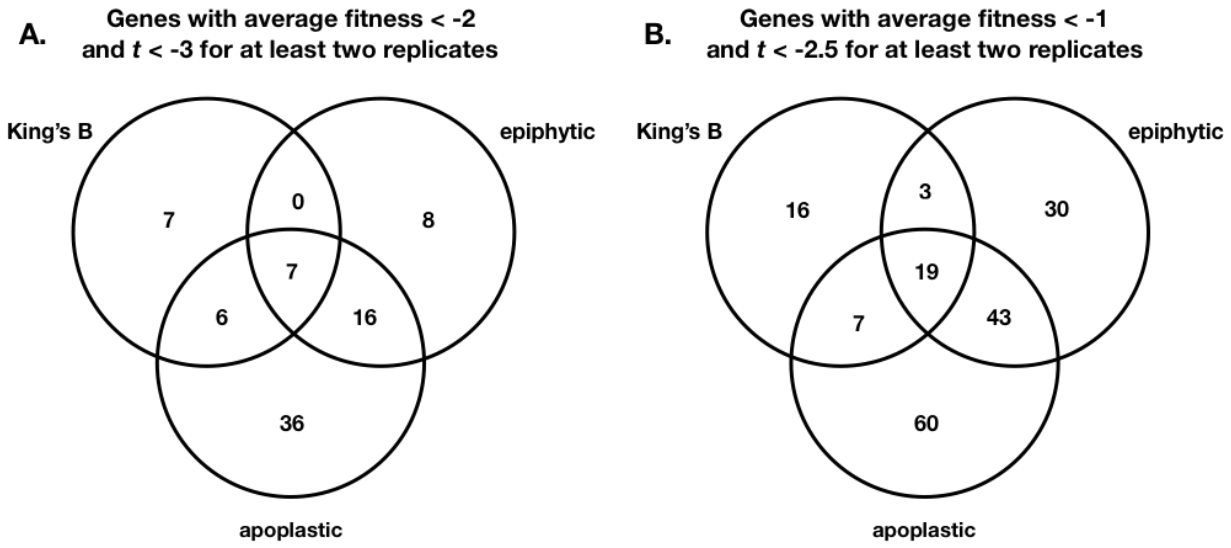


Figure 6. Genes with significant contributions to competitive fitness in the experimental conditions tested. (A) Venn diagram of genes with average fitness values < -2, and *t* < -3 for at least two experimental replicates. (B) Venn diagram of genes with average fitness values < -1, and *t* < -2.5 for at least two experimental replicates.

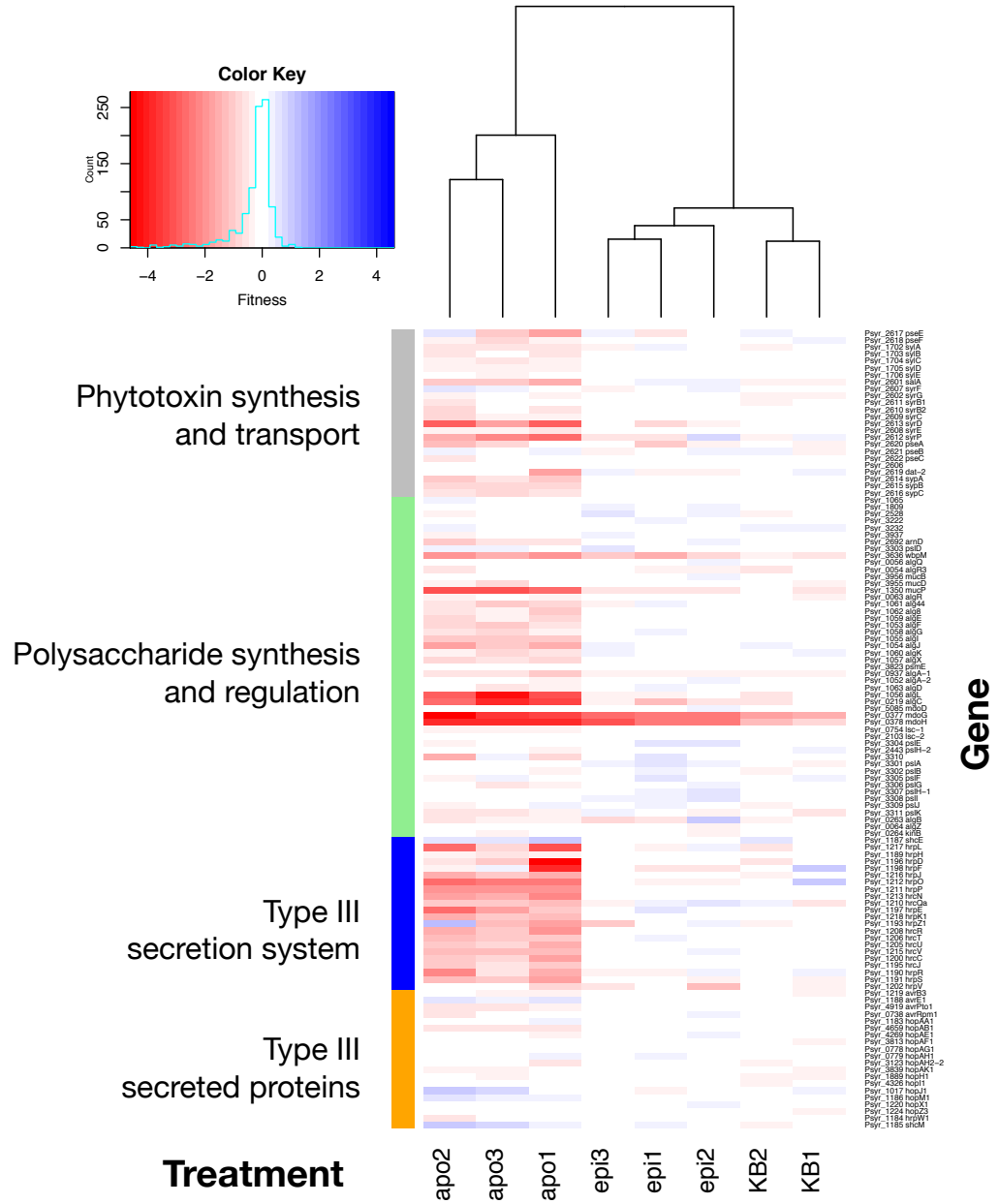
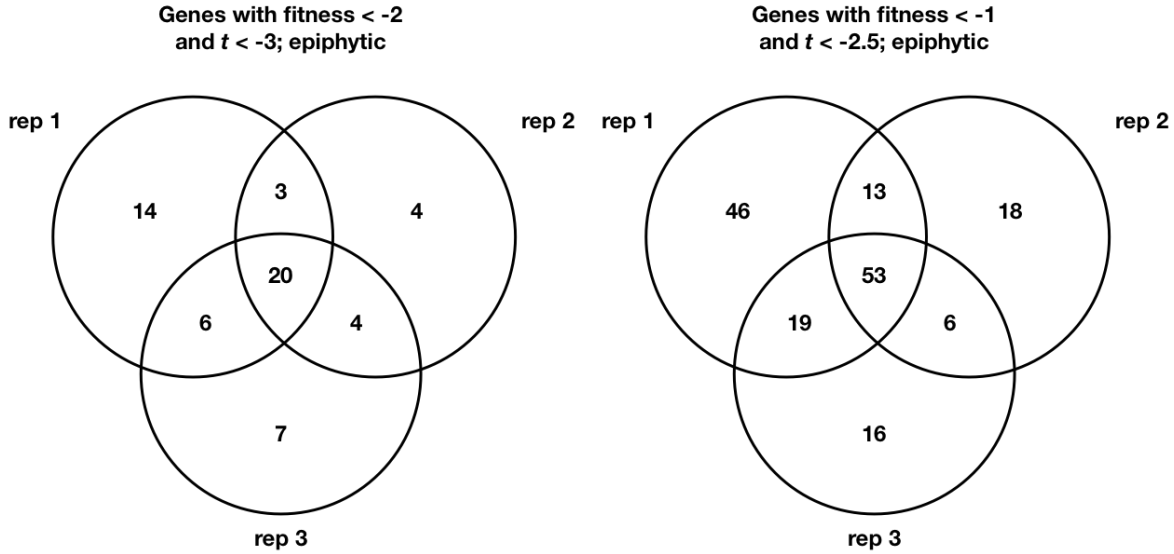


Figure 7. Fitness contributions of genes involved in phytotoxin synthesis and transport, the type III secretion system, and polysaccharide synthesis and regulation are required for apoplastic colonization. Gene fitness values are shown from 3 replicate epiphytic experiments (epi), 3 replicate apoplastic experiments (apo), and 2 replicate experiments in King's B medium (KB). The treatment ordering and corresponding dendrogram are calculated based on these genes only.

A.



B.

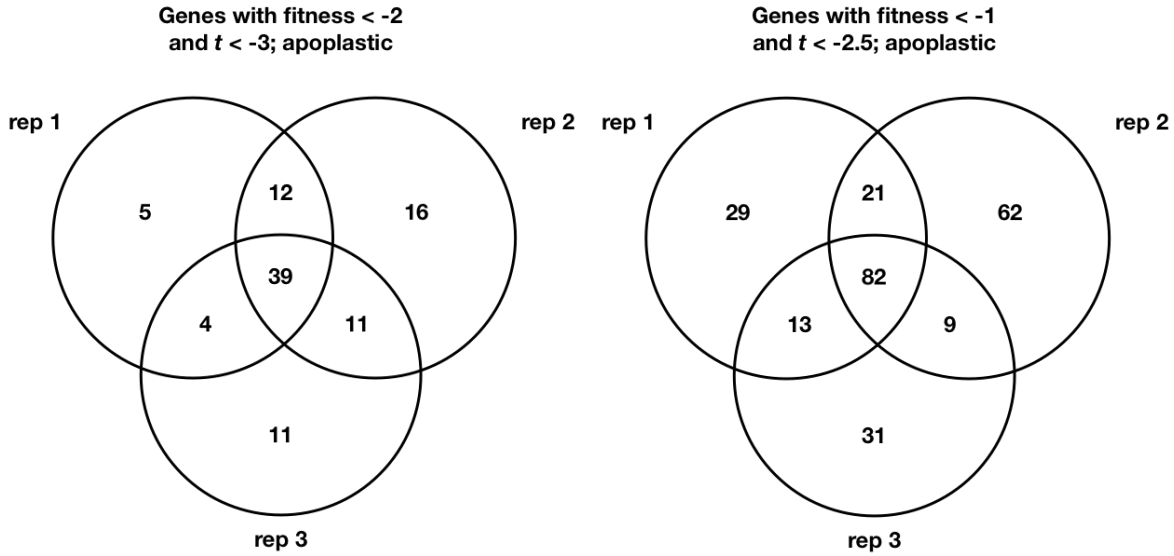


Figure 8. Venn diagrams showing consistency in identifying genes contributing to epiphytic (A) and apoplastic growth (B) in different replicate experiments. Fitness contribution thresholds are fitness < -2, *t* < -3 (left) and fitness < -1, *t* < -2.5 (right).

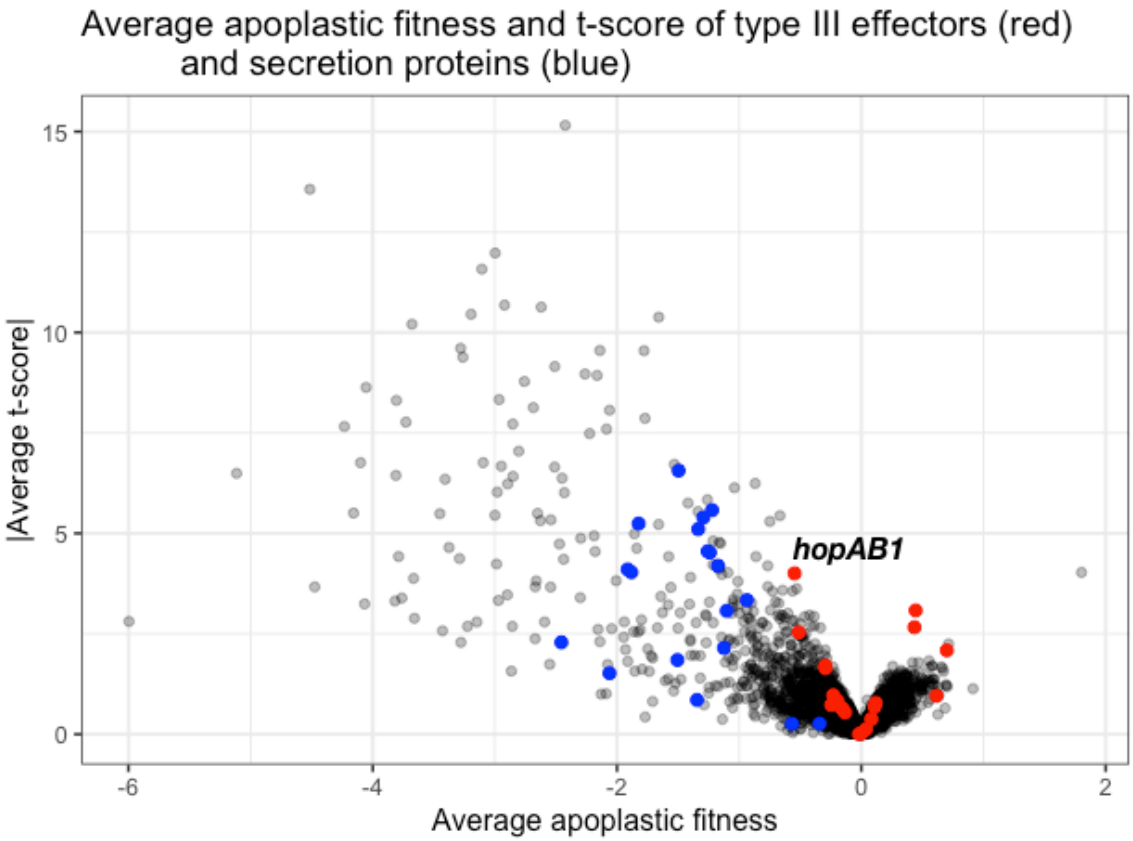
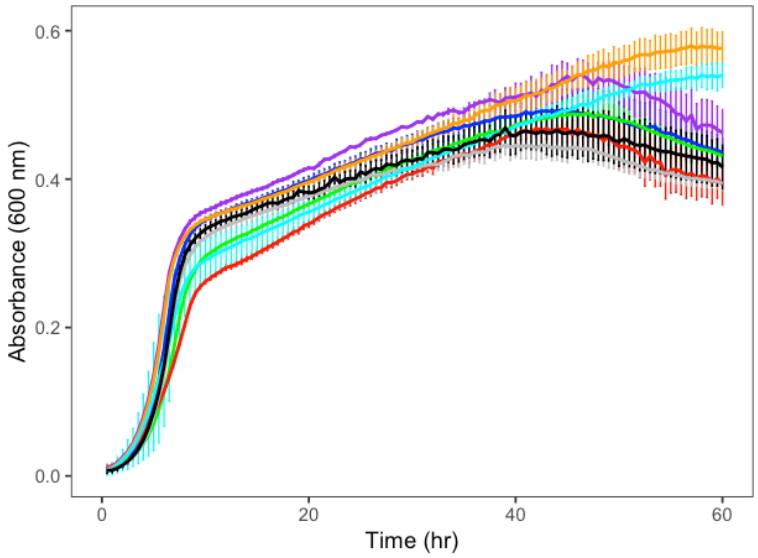


Figure 9. Average apoplastic fitness and absolute average *t*-score for all genes, averaging values from three replicate experiments. Genes encoding type III secreted proteins are shown in red, with those encoding type III secretion system components and regulatory elements shown in blue. Of the secreted effector proteins, HopAB1 has a consistent but relatively small contribution to apoplastic fitness. Disruptions in individual type III secretion system pilus components have large effects. All other genes are shown in grey.

A.



B.

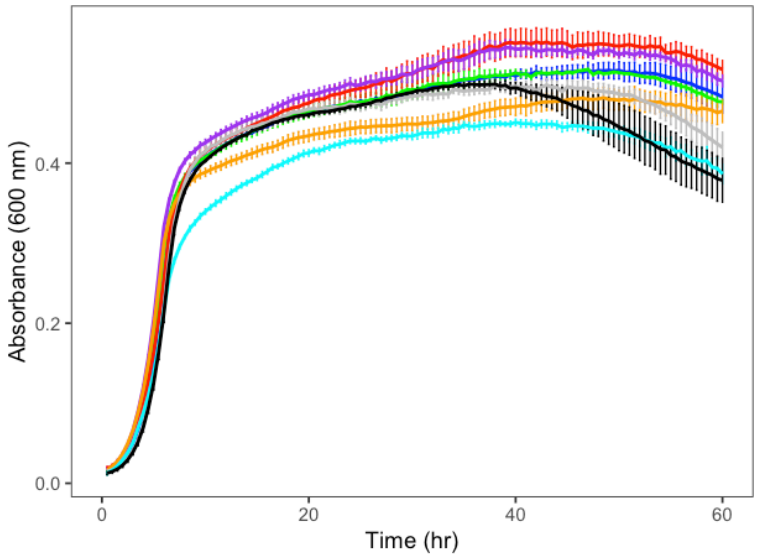
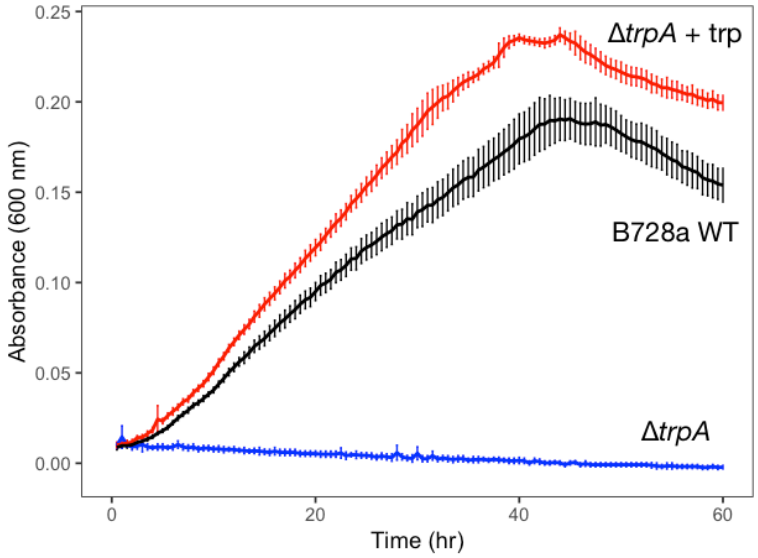


Figure 10. Growth of deletion mutants in KB (A) and LB (B). Absorbance at 600 nm was measured in 4 to 5 replicate wells per strain. Strains shown are B728a WT (black), Δ*syrP* (blue), Δ*hisD* (red), Δ*trpA* (green), Δ*hrpL* (purple), Δ*eftA* (grey), Δ*Psyr_0532* (orange), and Δ*Psyr_0920* (cyan). Mean absorbance of 4 replicate wells per strain is shown. Vertical lines indicate the standard deviation.

A.



B.

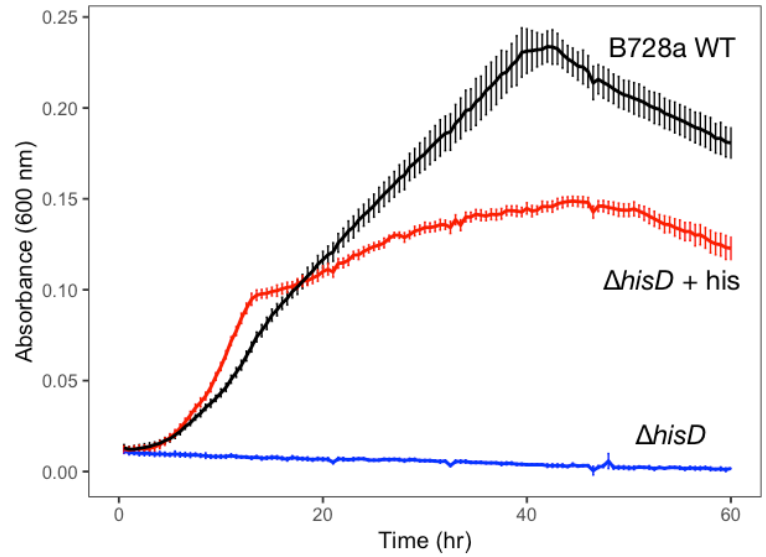


Figure 11. Growth of B728a auxotrophic mutant strains Δ*trpA* (A) and Δ*hisD* (B) in M9 minimal medium supplemented by 0.2% glycerol as a carbon source. Auxotrophic deletion strains were complemented by the addition of 20.8 µg/ml tryptophan or 62 µg/ml histidine. Mean absorbance of 3 replicate wells per strain is shown. Vertical lines indicate the standard deviation.

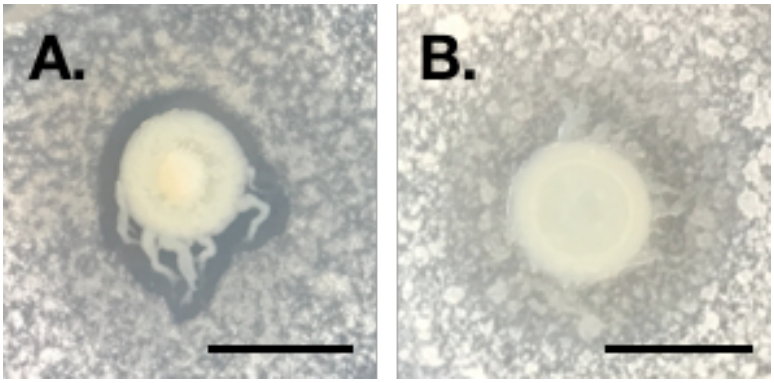


Figure 12. Inhibition of *Geotrichum candidum* Fr260 growth by B728a WT (A) and Δ*syrP* (B). Scale bars are 10 mm in length.

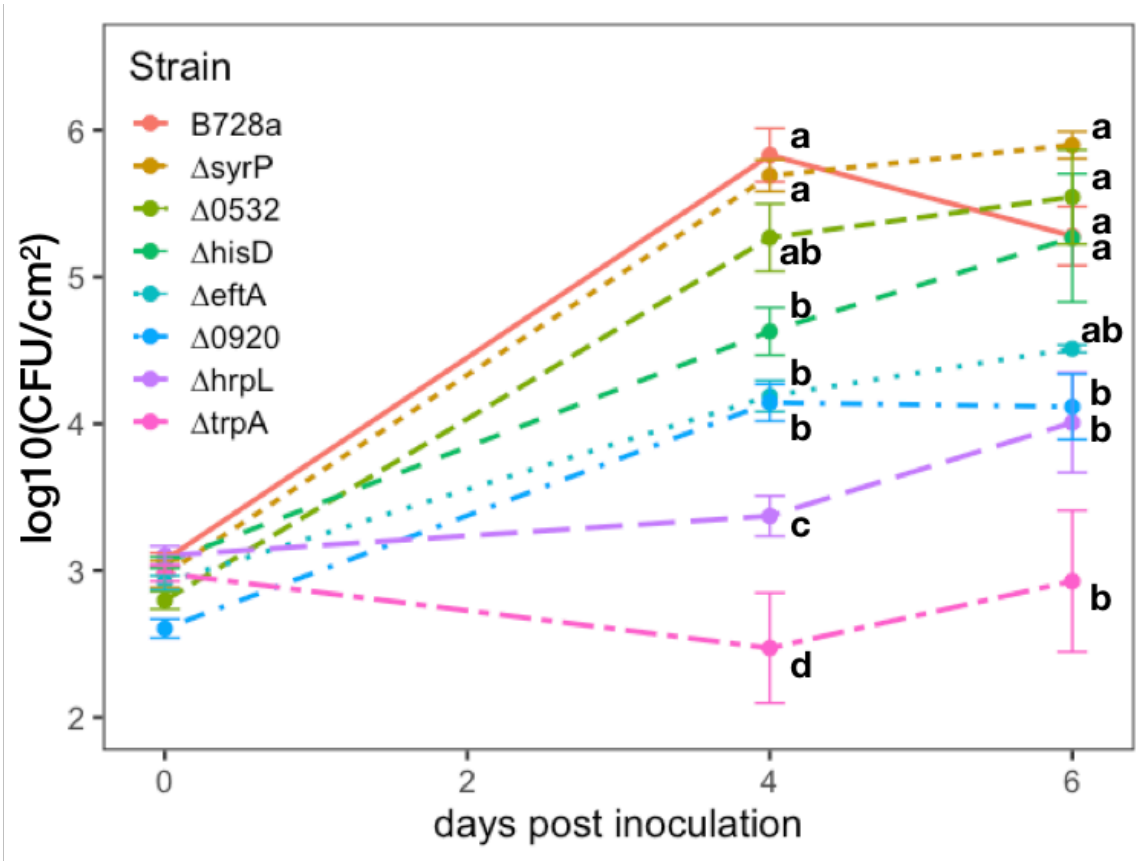


Figure 13. Apoplastic growth of B728a and deletion strains in bean. Growth of the amino acid auxotrophs Δ*trpA* and Δ*hisD*, type III regulatory mutant Δ*hrpL*, and syringomycin mutant Δ*syrP*. Apoplastic fitness of deletion mutants of glycosyltransferase genes Psyr_0532 and Psyr_0920, and hypothetical protein *eftA.* Error bars represent the standard error of the mean. At each time point, strains labeled with the same letter are not significantly different (Tukey's HSD test, $p < 0.05$).

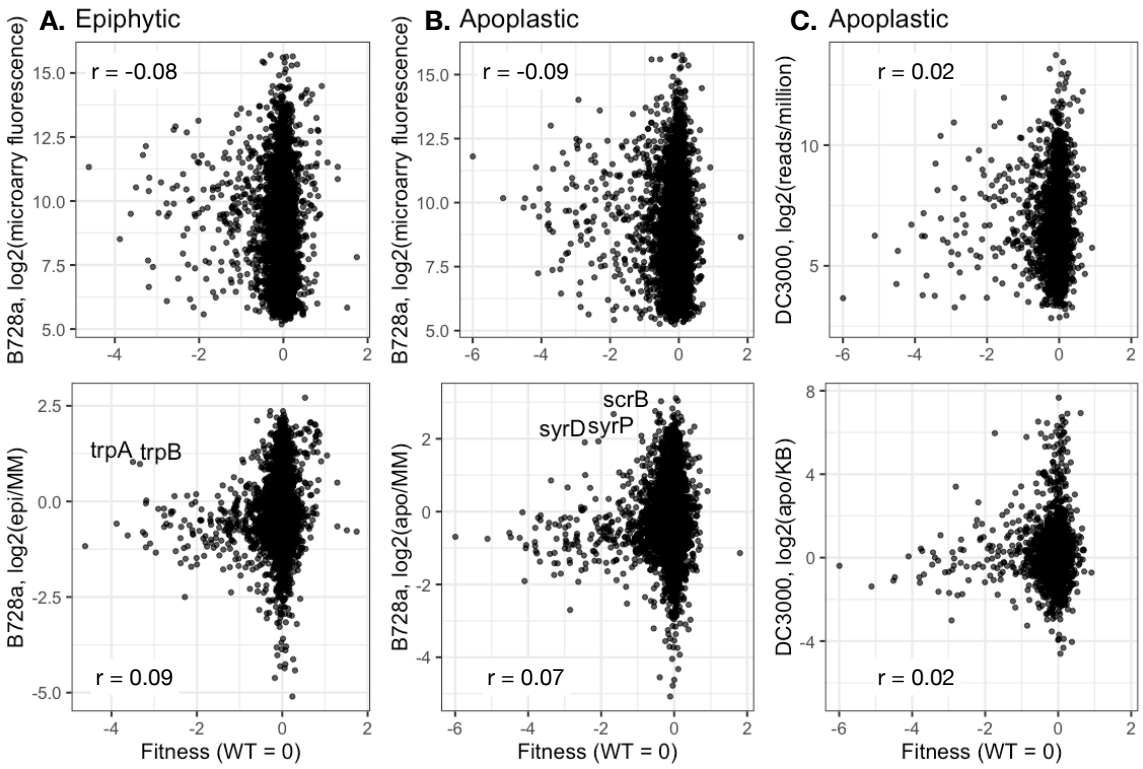


Figure 14. The magnitude of fitness contributions of genes in *P. syringae* do not correlate well with their absolute level of expression (top) or fold-change of these genes *in planta* compared to that in a hrp-inducing minimal medium (MM) or KB (bottom). Epiphytic (A) and apoplastic (B) gene expression in B728a is a measure of fluorescence in microarrays (49). A comparison of B728a apoplastic gene fitness to expression of orthologs in the closely related strain DC3000, measured in *Arabidopsis thaliana*, reveals a similar trend (C). DC3000 gene expression is calculated as the number of reads/million (67). Fold change in gene expression was calculated as a $\log_2$ of the ratio of gene expression estimated from microarray fluorescence (B728a) or RNAseq (DC3000) *in planta* relative to that in MM or in KB (49). The Pearson correlation coefficients are shown.

Tables

Strains and plasmids

Strains		
Organism	Description	Source
E. coli TOP10	For general cloning	Invitrogen
E. coli XL1-Blue	For general cloning	QB3 Macrolab
E. coli S17-1	Conjugation donor strain	(114)
E. coli WM3064	Strain APA752; barcoded *mariner* transposon vector (Kan^R) in *E. coli* conjugation strain	(28)
P. syringae B728a	Wild type strain (Rif^R)	(60)
P. syringae B728a	Whole genome barcoded *mariner* transposon library (Rif^R Kan^R)	This work
P. syringae B728a	Δ*hrpL* (Rif^R)	(94)
P. syringae B728a	Δ*trpA* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*hisD* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*syrP* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*eftA* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*Psyr_0532* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*Psyr_0920* (Rif^R Kan^R)	This work
Geotrichum candidum Fr260	To assay for the production of syringomycin	Dennis Gross

Plasmids			
Plasmid name	Description	Antibiotic	Reference
pKD13	Source of kanamycin resistance	Kan	(112)
pT*sacB*	Suicide plasmid to introduce DNA into *P. syringae*	Tet	(113)
pT:0033-kan	To delete *trpA*, contains *Psyr_0033* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:4133-kan	To delete *hisD*, contains *Psyr_4133* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:2612-kan	To delete *syrP*, contains *Psyr_2612* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:4158-kan	To delete *eftA*, contains *Psyr_4158* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:0532-kan	To delete *Psyr_0532*, contains *Psyr_0532* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:0920-kan	To delete *Psyr_0920*, contains *Psyr_0920* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work

Primers

Bold sequence complements FRT-Kan sequence for splicing by overlap extension protocol. The primers used for TnSeq mapping and BarSeq amplification are described in (28).

Name	Sequence
FRT-KanF	GTGTAGGCTGGAGCTGCTTC
FRT-KanR	ATTCCGGGGATCCGTCGACC
0033 up F	TGATCGGCTGCCCTTATGTG
FRT 0033 up R	**GAAGCAGCTCCAGCCTACAC**AGCAAGACACAAGGGGTTCA
FRT 0033 dwn F	**GGTCGACGGATCCCCGGAAT**TTTGCATGTCTTTGTCGCCG
0033 dwn R	TGGTGTTAGACCTCAACCGC
4133 up F	CTTCCAGCAACGCCTGATGT
FRT 4133 up R	**GAAGCAGCTCCAGCCTACAC**TGAGCAAATTCTGGAGCCCT
FRT 4133 dwn F	**GGTCGACGGATCCCCGGAAT**GGGCCTCAATAATTGGCGGA
4133 dwn R	TGATCGACCGCATCTACCAC
4133 up F	CTTCCAGCAACGCCTGATGT
2612 up F	CACCCCAGATTTCCCAGACC
FRT 2612 up R	**GAAGCAGCTCCAGCCTACAC**GACCTCAGCCCTTCACATCC
FRT 2612 dwn F	**GGTCGACGGATCCCCGGAAT**TCCCGTTATCAAGCCAGGAC
2612 dwn R	TGGAGAATCCGAAATCCGCC
4158 up F	CAGGACTCGGAGATCATGCC
FRT 4158 up R	**GAAGCAGCTCCAGCCTACAC**CGCCTCATGGAGTACAGTGG
FRT 4158 dwn F	**GGTCGACGGATCCCCGGAAT**GGTGCAAAGAGCAGAATCGG
4158 dwn R	GTATCGACTCGCGGGAAACT
0532 up F	ACCTCGTCTCTGGCTGTTTC
FRT 0532 up R	**GAAGCAGCTCCAGCCTACAC**CAGTACTGCGCCTGCTGAAT
FRT 0532 dwn F	**GGTCGACGGATCCCCGGAAT**ATGATGGTATTCAGCGAAAACAG
0532 dwn R	TAATCCCGGCCACGACAAAG
0920 up F	GTACGCTGGAAGAATCGGGT
FRT 0920 up R	**GAAGCAGCTCCAGCCTACAC**TTTCCTTGCGCTCAAAAGCC
FRT 0920 dwn F	**GGTCGACGGATCCCCGGAAT**ACAGCCGATTTGAACCTGGG
0920 dwn R	ACTTCCATGCCAGAAGGTGG

Table 1. Number of genes within each functional category predicted to be essential or nearly essential in LB (N = 392), based on the TnSeq data. Protein coding genes were considered essential or nearly essential for growth in LB if normalized read density (reads/nucleotides across the entire gene) and normalized insertion density (sites/nucleotides in the central 10 – 90% of the gene) were below 0.2 (29). Genes shorter than 325 bp were excluded from this analysis. Functional category annotations for B728a genes are primarily based on COG (77) and KEGG (115) annotations, with manual additions and corrections originally published by Yu *et al.* (49).

Category	Predicted essential genes
Translation	66
None	38
Cofactor metabolism	36
Energy generation	30
Amino acid metabolism and transport	22
Nucleotide metabolism and transport	22
Replication and DNA repair	20
LPS synthesis and transport	19
Hypothetical	15
Secretion/Efflux/Export	15
Peptidoglycan/cell wall polymers	12
Fatty acid metabolism	10
Sulfur metabolism and transport	8
Terpenoid backbone synthesis	8
Carbohydrate metabolism and transport	7
Cell division	7
Siderophore synthesis and transport	7
Transcription	5
Chaperones/Heat shock proteins	4
Nitrogen metabolism	4
Phospholipid metabolism	4
Transcription - Sigma factor	4
Transport (organic compounds)	3
Iron metabolism and transport	2
Iron-sulfur proteins	2
Organic acid metabolism and transport	2
Outer membrane proteins	2
Oxidative stress tolerance	2
Post-translational modification	2
Stress resistance	2
Transcriptional regulation	2
Compatible solute synthesis	1
Osmosensing & regulation	1
Oxidative stress tolerance (Antioxidant enzyme)	1
Phage & IS elements	1
Phosphate metabolism and transport	1
Polysaccharide synthesis and regulation	1
Proteases	1

Signal transduction mechanisms	1
Transport	1
Transport (inorganic ions)	1

Table 2. B728a genes with 0 TA dinucleotide sites, that are inaccessible by the *mariner* transposon. The RegionGC.pl script (https://bitbucket.org/berkeleylab/feba) was used to calculate the number of TA sites per gene in the B728a genome. The script as written considers TA sites in the middle 10 – 80% of a gene, and was altered to calculate the number of TA sites per [entire] gene.

Locus	Description	Size (bp)
Psyr_0794	Protein of unknown function DUF329	209
Psyr_1136	conserved hypothetical protein	224
Psyr_2216	conserved hypothetical protein	125
Psyr_3320	conserved hypothetical protein	149
Psyr_3582	Protein of unknown function DUF1289	191

Table 3. B728a genes predicted to be essential compared to their predicted homologs in *P. aeruginosa* PAO1 (A), *P. simiae* WCS417 (B), and *P. stutzeri* RCH2 (C). Homologs to B728a genes were identified using the IMG database genome-gene best homologs function, at 70% identity (77). Essential vs. nonessential characterization of PAO1 genes was predicted in (79), and predicted for *P. simiae* and *P. stutzeri* in (29). For these studies, all transposon libraries of all strains (including B728a) were generated on LB (29, 78).

A.

	B728a essential	B728a nonessential
P. aeruginosa essential	259	48
P. aeruginosa nonessential	104	1938
No homolog in *P. aeruginosa*	29	2610
Total	392	4596

B.

	B728a essential	B728a nonessential
P. simiae essential	282	133
P. simiae nonessential	75	2070
No homolog in *P. simiae*	35	2393
Total	392	4596

C.

	B728a essential	B728a nonessential
P. stutzeri essential	254	76
P. stutzeri nonessential	63	1057
No homolog in *P. stutzeri*	75	3463
Total	392	4596

Table 4. Genes that when disrupted confer individually large decreases in fitness (average fitness < -2 from two replicate experiments) in King's B medium. All genes represent those in which *t* < -3 in both replicate experiments.

Locus	Name	Description	Fitness in King's B			Classification
			1	2	Average	
Psyr_4581	trpG	anthranilate synthase, component II	-2.47	-1.71	-2.09	Amino acid metabolism and transport
Psyr_0826	pgi	glucose-6-phosphate isomerase	-1.91	-2.21	-2.06	Carbohydrate metabolism and transport
Psyr_1613	minC	septum site-determining protein MinC	-3.40	-3.70	-3.55	Cell division
Psyr_0454	bioA	adenosylmethionine-8-amino-7-oxononanoate aminotransferase apoenzyme	-2.56	-2.62	-2.59	Cofactor metabolism
Psyr_4687	bioB	biotin synthase	-2.87	-2.96	-2.91	Cofactor metabolism
Psyr_4683	bioD	dethiobiotin synthase	-3.30	-2.53	-2.92	Cofactor metabolism
Psyr_4686	bioF	8-amino-7-oxononanoate synthase	-2.89	-3.18	-3.03	Cofactor metabolism
Psyr_0951	moeB	UBA/THIF-type NAD/FAD binding fold:MoeZ/MoeB	-4.81	-6.07	-5.44	Cofactor metabolism
Psyr_2080	pabB	aminodeoxychorismate synthase, subunit I	-1.38	-2.65	-2.01	Cofactor metabolism
Psyr_4341	thiE	thiamine-phosphate diphosphorylase	-2.80	-3.71	-3.26	Cofactor metabolism
Psyr_4740	thiG	thiazole-phosphate synthase	-3.69	-3.98	-3.84	Cofactor metabolism
Psyr_4340		phosphomethylpyrimidine kinase, putative	-3.51	-4.07	-3.79	Cofactor metabolism
Psyr_0565		Protein of unknown function UPF0126	-2.35	-3.33	-2.84	Hypothetical
Psyr_0917	rfbA-2	ABC-2	-1.90	-2.53	-2.21	LPS synthesis and transport
Psyr_0918	rfbB-2	ABC transporter	-2.05	-2.37	-2.21	LPS synthesis and transport
Psyr_0259	envZ	Osmolarity sensor protein envZ	-2.29	-1.91	-2.10	Osmosensing & regulation
Psyr_3008	uppP	Undecaprenyl-diphosphatase	-2.95	-4.69	-3.82	Peptidoglycan/cell wall polymers
Psyr_0849	pssA-1	CDP-diacylglycerol--serine O-phosphatidyltransferase	-2.04	-2.32	-2.18	Phospholipid metabolism
Psyr_4091		8-oxo-dGTPase	-2.54	-2.47	-2.51	Replication and DNA repair
Psyr_1544		SirA-like protein	-3.94	-4.44	-4.19	

Table 5. Numbers of unique barcodes and median reads per gene in each of three replicate experiments and samples obtained from mid-log phase cultures following library outgrowth and before inoculation (time0) and from leaves pooled from 100 pots after growth in a particular habitat (sample). Unique barcodes were calculated as the total that mapped to the genome and had 3 or more reads in a given experiment. The total number of mapped barcodes in the library = 281,417. Technical (sequencing) replicates are listed separately ("a" and "b"), and share the same time0 reference sample. For an experiment to pass quality control, the median reads per gene in the sample must be ≥ 50 (28).

Experiment	Unique barcodes at time0	Unique barcodes in sample	% Recovery	Median reads/gene at time0	Median reads/gene in sample
KB_1	187,482	192,078	>100*	278	150
KB_2	197,089	226,857	>100*	378.5	251
Epiphytic_1a	196,894	191,648	97.3	279	167
Epiphytic_1b	196,894	182,055	92.5	279	148
Epiphytic_2a	216,700	177,679	82.0	381	175.5
Epiphytic_2b	216,700	173,567	80.1	381	164.5
Epiphytic_3a	230,482	201,171	87.3	453.5	248.5
Epiphytic_3b	230,482	196,807	85.4	453.5	211
Apoplast_1a	218,149	149,311	68.4	409.5	163
Apoplast_1b	218,149	151,211	69.3	409.5	155
Apoplast_2	214,346	156,416	73.0	390	160
Apoplast_3a	222,473	185,183	83.2	397	203
Apoplast_3b	222,473	187,289	84.2	397	216

* More unique barcodes sequenced at the end of an experiment indicates that additional unique barcodes were present but were not sequenced at the start of the experiment (time0).

Table 6. Genes with average fitness < -2 in epiphytic experiments. All genes represent those in which *t* < -3 in at least two experiments.

Locus	Name	Description	Epiphytic Fitness				Classification
			1	2	3	Average	
Psyr_4270	glyA	serine hydroxymethyltransferase	-2.35	-2.94	-3.42	-2.90	Amino acid metabolism and transport
Psyr_4369	proA	glutamate-5-semialdehyde dehydrogenase	-3.02	-2.42	-2.84	-2.76	Amino acid metabolism and transport
Psyr_0704	proB	glutamate 5-kinase	-2.91	-1.48	-2.12	-2.17	Amino acid metabolism and transport
Psyr_0557	serB	phosphoserine phosphatase	-2.16	-4.21	-3.27	-3.21	Amino acid metabolism and transport
Psyr_0033	trpA	tryptophan synthase, alpha chain	-3.26	-3.21	-4.03	-3.50	Amino acid metabolism and transport
Psyr_0034	trpB	tryptophan synthase, beta chain	-2.95	-3.04	-4.01	-3.33	Amino acid metabolism and transport
Psyr_4609	trpE	anthranilate synthase, component I	-3.60	-3.31	-3.96	-3.62	Amino acid metabolism and transport
Psyr_1663	trpF	phosphoribosylanthranilate isomerase	-3.54	-3.50	-4.59	-3.88	Amino acid metabolism and transport
Psyr_4852		D-3-phosphoglycerate dehydrogenase	-2.97	-3.51	-3.31	-3.26	Amino acid metabolism and transport
Psyr_2980	galU	UDP-glucose pyrophosphorylase	-2.64	-1.34	-2.07	-2.02	Carbohydrate metabolism and transport
Psyr_1613*	minC	septum site-determining protein MinC	-1.46	-2.27	-2.53	-2.09	Cell division
Psyr_0454*	bioA	adenosylmethionine-8-amino-7-oxononanoate aminotransferase apoenzyme	-2.43	-2.01	-2.35	-2.26	Cofactor metabolism
Psyr_4687*	bioB	biotin synthase	-2.03	-2.76	-2.05	-2.28	Cofactor metabolism
Psyr_4686*	bioF	8-amino-7-oxononanoate synthase	-1.94	-2.28	-2.18	-2.13	Cofactor metabolism
Psyr_3174	cysG	uroporphyrinogen-III C-methyltransferase / precorrin-2 dehydrogenase	-1.33	-2.70	-3.07	-2.37	Cofactor metabolism
Psyr_0846	ilvI	acetolactate synthase, large subunit	-2.34	-2.45	-2.35	-2.38	Cofactor metabolism
Psyr_0827	panC	pantothenate synthetase	-2.60	-4.00	-2.69	-3.10	Cofactor metabolism
Psyr_0532		conserved hypothetical protein	-3.40	-2.35	-1.73	-2.50	Hypothetical
Psyr_2461		Uncharacterized conserved protein UCP030820	-2.13	-1.79	-2.44	-2.12	Hypothetical
Psyr_0917*	rfbA-2	ABC-2	-3.05	-2.13	-2.62	-2.60	LPS synthesis and transport
Psyr_0918*	rfbB-2	ABC transporter	-2.83	-2.00	-2.05	-2.29	LPS synthesis and transport
Psyr_0411	gltB	glutamate synthase (NADPH) large subunit	-2.73	-1.65	-1.64	-2.01	Nitrogen metabolism

Psyr_1668	purF	amidophosphoribosyltransferase	-3.20	-3.21	-3.15	-3.18	Nucleotide metabolism and transport
Psyr_1269	purL	phosphoribosylformylglycinamidine synthase	-3.08	-2.52	-2.54	-2.71	Nucleotide metabolism and transport
Psyr_3008*	uppP	Undecaprenyl-diphosphatase	-1.86	-2.29	-3.01	-2.38	Peptidoglycan/cell wall polymers
Psyr_4512		putative phage-related protein	-2.24	-1.70	-2.38	-2.11	Phage & IS elements
Psyr_0377	mdoG	Periplasmic glucan biosynthesis protein, MdoG	-2.47	-2.48	-2.73	-2.56	Polysaccharide synthesis and regulation
Psyr_0378	mdoH	Glycosyl transferase, family 2	-2.38	-2.37	-3.05	-2.60	Polysaccharide synthesis and regulation
Psyr_1408	ruvC	Holliday junction endonuclease RuvC	-3.43	-2.74	-1.71	-2.63	Replication and DNA repair
Psyr_2462		Nitrite/sulfite reductase, hemoprotein beta-component, ferrodoxin-like:Nitrite and sulphite reductase 4Fe-4S region	-2.60	-3.62	-3.36	-3.19	Sulfur metabolism and transport
Psyr_0529		Glycosyl transferase, group 1	-3.59	-3.11	-2.88	-3.19	

* These genes also have a fitness of < -2 and *t* < -3 for growth in KB relative to the input library.

Table 7. Genes with average fitness < -2 in apoplastic experiments. All genes represent those in which *t* < -3 in at least two experiments.

Locus	Name	Description	Apoplastic Fitness				Classification
			1	2	3	Average	
Psyr_4270	glyA	serine hydroxymethyltransferase	-5.00	-5.34	-5.01	-5.11	Amino acid metabolism and transport
Psyr_4897	hisB	imidazoleglycerol-phosphate dehydratase	-3.76	-3.42	-2.69	-3.29	Amino acid metabolism and transport
Psyr_4132	hisC	histidinol phosphate aminotransferase apoenzyme	-1.90	-3.58	-2.57	-2.69	Amino acid metabolism and transport
Psyr_4133	hisD	histidinol dehydrogenase	-2.55	-3.66	-2.79	-3.00	Amino acid metabolism and transport
Psyr_4134	hisG	ATP phosphoribosyltransferase (homohexameric)	-2.51	-3.37	-2.40	-2.76	Amino acid metabolism and transport
Psyr_4896	hisH	imidazole glycerol phosphate synthase subunit hisH	-2.02	-2.56	-1.96	-2.18	Amino acid metabolism and transport
Psyr_1257	leuA	2-isopropylmalate synthase	-2.16	-2.56	-1.95	-2.22	Amino acid metabolism and transport
Psyr_1985	leuB	3-isopropylmalate dehydrogenase	-4.35	-3.99	-3.82	-4.05	Amino acid metabolism and transport
Psyr_1983	leuC	3-isopropylmalate dehydratase, large subunit	-4.30	-4.82	-3.35	-4.16	Amino acid metabolism and transport
Psyr_0473	metW	Methionine biosynthesis MetW	-4.66	-4.25	-3.38	-4.10	Amino acid metabolism and transport
Psyr_0474	metX	homoserine O-acetyltransferase	-4.04	-3.02	-3.97	-3.68	Amino acid metabolism and transport
Psyr_1669	metZ	O-succinylhomoserine sulfhydrylase	-4.14	-4.63	-3.93	-4.23	Amino acid metabolism and transport
Psyr_4369	proA	glutamate-5-semialdehyde dehydrogenase	-3.72	-2.12	-3.15	-3.00	Amino acid metabolism and transport
Psyr_0704	proB	glutamate 5-kinase	-3.24	-3.19	-1.98	-2.80	Amino acid metabolism and transport
Psyr_0557	serB	phosphoserine phosphatase	-2.88	-2.37	-1.64	-2.30	Amino acid metabolism and transport
Psyr_0033	trpA	tryptophan synthase, alpha chain	-2.89	-3.23	-1.87	-2.66	Amino acid metabolism and transport
Psyr_0034	trpB	tryptophan synthase, beta chain	-4.03	-3.75	-0.24	-2.67	Amino acid metabolism and transport
Psyr_4580	trpD	anthranilate phosphoribosyltransferase	-4.47	-4.70	-4.26	-4.47	Amino acid metabolism and transport
Psyr_4609	trpE	anthranilate synthase,	-4.87	-4.58	-4.10	-4.51	Amino acid

		component I					metabolism and transport
Psyr_1663	trpF	phosphoribosylanthranilate isomerase	-3.43	-3.93	-3.64	-3.67	Amino acid metabolism and transport
Psyr_4852		D-3-phosphoglycerate dehydrogenase	-3.06	-3.30	-2.49	-2.95	Amino acid metabolism and transport
Psyr_4894		1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-1.87	-3.00	-2.76	-2.54	Amino acid metabolism and transport
Psyr_2980	galU	UDP-glucose pyrophosphorylase	-2.10	-2.53	-2.28	-2.30	Carbohydrate metabolism and transport
Psyr_0826*	pgi	glucose-6-phosphate isomerase	-2.87	-2.96	-3.46	-3.10	Carbohydrate metabolism and transport
Psyr_3179	ftsK	DNA translocase FtsK	-2.86	-2.53	-2.23	-2.54	Cell division
Psyr_1613*	minC	septum site-determining protein MinC	-3.94	-3.75	-3.68	-3.79	Cell division
Psyr_4194	dnaJ-1	DnaJ central region:Heat shock protein DnaJ, N-terminal:Chaperone DnaJ, C-terminal	-1.85	-2.30	-1.87	-2.01	Chaperones/Heat shock proteins
Psyr_0454*	bioA	adenosylmethionine-8-amino-7-oxononanoate aminotransferase apoenzyme	-2.32	-2.17	-2.07	-2.19	Cofactor metabolism
Psyr_4687*	bioB	biotin synthase	-2.83	-2.85	-2.88	-2.85	Cofactor metabolism
Psyr_4683*	bioD	dethiobiotin synthase	-2.99	-2.49	-3.21	-2.90	Cofactor metabolism
Psyr_4686*	bioF	8-amino-7-oxononanoate synthase	-2.54	-2.81	-2.53	-2.63	Cofactor metabolism
Psyr_0848	ilvC	ketol-acid reductoisomerase	-2.91	-3.88	-2.11	-2.97	Cofactor metabolism
Psyr_0469	ilvD	dihydroxyacid dehydratase	-3.70	-4.02	-2.11	-3.28	Cofactor metabolism
Psyr_0847	ilvH	acetolactate synthase, small subunit	-2.85	-2.72	-1.72	-2.43	Cofactor metabolism
Psyr_0846	ilvI	acetolactate synthase, large subunit	-3.29	-3.70	-2.33	-3.11	Cofactor metabolism
Psyr_4341*	thiE	thiamine-phosphate diphosphorylase	-1.94	-2.69	-2.91	-2.51	Cofactor metabolism
Psyr_4740*	thiG	thiazole-phosphate synthase	-1.93	-2.18	-3.84	-2.65	Cofactor metabolism
Psyr_4340*		phosphomethylpyrimidine kinase, putative	-2.01	-2.79	-3.76	-2.85	Cofactor metabolism
Psyr_0167		hypothetical protein	-2.86	-3.64	-2.41	-2.97	Hypothetical
Psyr_1614	htrB	lipid A biosynthesis acyltransferase	-4.30	-4.58	-2.57	-3.82	LPS synthesis and transport
Psyr_0917*	rfbA-2	ABC-2	-2.43	-3.24	-3.27	-2.98	LPS synthesis and transport
Psyr_0918*	rfbB-2	ABC transporter	-2.08	-2.40	-3.38	-2.62	LPS synthesis and transport
Psyr_0014		lipid A biosynthesis acyltransferase	-3.14	-2.06	-2.81	-2.67	LPS synthesis and transport
Psyr_1668	purF	amidophosphoribosyltransferase	-3.64	-3.27	-4.53	-3.81	Nucleotide metabolism and transport
Psyr_1269	purL	phosphoribosylformylglycinamidi	-3.19	-4.03	-4.07	-3.76	Nucleotide

		ne synthase					metabolism and transport
Psyr_3008*	uppP	Undecaprenyl-diphosphatase	-3.42	-3.44	-3.36	-3.41	Peptidoglycan/cell wall polymers
Psyr_2613	syrD	Cyclic peptide transporter	-2.77	-2.88	-1.70	-2.45	Phytotoxin synthesis and transport
Psyr_0219	algC	phosphomannomutase	-3.35	-2.67	-3.77	-3.26	Polysaccharide synthesis and regulation
Psyr_1056	algL	Poly(beta-D-mannuronate) lyase	-3.02	-2.82	-4.29	-3.38	Polysaccharide synthesis and regulation
Psyr_0377	mdoG	Periplasmic glucan biosynthesis protein, MdoG	-3.27	-4.42	-3.49	-3.73	Polysaccharide synthesis and regulation
Psyr_0378	mdoH	Glycosyl transferase, family 2	-3.83	-3.70	-3.89	-3.81	Polysaccharide synthesis and regulation
Psyr_1350	mucP	site-2 protease, Metallo peptidase, MEROPS family M50B	-2.71	-3.14	-3.11	-2.99	Polysaccharide synthesis and regulation
Psyr_1748	clpX	ATP-dependent Clp protease ATP-binding subunit ClpX	-2.97	-2.42	-3.29	-2.90	Proteases
Psyr_1410	ruvB	Holliday junction DNA helicase RuvB	-1.88	-2.79	-2.74	-2.47	Replication and DNA repair
Psyr_0919		Chromosome segregation ATPase-like protein	-2.02	-1.91	-2.49	-2.14	Replication and DNA repair
Psyr_3958	algU	RNA polymerase, sigma-24 subunit, RpoE	-2.78	-2.67	-3.31	-2.92	Transcription - Sigma factor
Psyr_4408	retS	Response regulator receiver:ATP-binding region, ATPase-like:Histidine kinase A, N-terminal:Histidine kinase:Histidine kinase	-2.50	-2.32	-2.71	-2.51	Transcriptional regulation
Psyr_3637	wbpL	Glycosyl transferase, family 4	-2.50	-2.04	-1.95	-2.16	
Psyr_0936	wbpY	Glycosyl transferase, group 1	-1.90	-1.68	-2.61	-2.06	
Psyr_0914	wbpZ	Glycosyl transferase, group 1	-2.22	-1.78	-2.26	-2.09	
Psyr_0529		Glycosyl transferase, group 1	-3.78	-2.89	-3.69	-3.45	
Psyr_0915		NAD-dependent epimerase/dehydratase	-3.09	-3.05	-3.44	-3.20	
Psyr_1544*		SirA-like protein	-1.67	-2.64	-3.01	-2.44	
Psyr_0920		Glycosyl transferase, group 1	-2.69	-2.34	-2.24	-2.42	
Psyr_4130		Peptidase S1, chymotrypsin:PDZ/DHR/GLGF	-2.51	-2.20	-2.08	-2.26	

* These genes also fulfill average fitness < -2 and *t* < -3 for growth in KB relative to the input library.

Regarding *wbpYZ*: in the original gene metadata these genes are annotated as "*wpbY*" and "*wpbZ*". This is likely a typo, as *wbp* genes are involved in O-antigen biosynthesis in *P. aeruginosa*, and "*wpb*" genes could not be identified in *P. syringae* or other Pseudomonads.

Table 8. Functional categories that are enriched among genes important for fitness *in planta*, based on a hypergeometric test. For this analysis, we only included genes with average fitness < -2 in the *in planta* habitats. The KB data are shown for comparison. Functional categories were assigned based on the previously annotated genomic metadata (49).

Category	**Category size (total)**	**King's B**	**Epiphytic**	**Apoplastic**
		p-value		
Amino acid metabolism and transport	210	0.329	3.294E-08	5.127E-16
Polysaccharide synthesis and regulation	49	0.241	0.011	3.332E-04
Nucleotide metabolism and transport	48	0.029	7.801E-05	2.244E-03
Type III secretion system	42	0.211	0.333	8.653E-03
Phytotoxin synthesis and transport	24	0.126	0.206	0.01085

Table 9. Genes with average fitness < -1 but > -2 in King's B (A), epiphytic (B) and apoplastic experiments (C). All genes represent those in which *t* < -2.5 in at least two experiments.

A.

Locus	Name	Description	Fitness in King's B			Classification
			1	2	Average	
Psyr_2855	metE	methionine synthase (B12-independent)	-1.70	-1.34	-1.52	Amino acid metabolism and transport
Psyr_2980	galU	UDP-glucose pyrophosphorylase	-1.20	-1.68	-1.44	Carbohydrate metabolism and transport
Psyr_1751	ppiD	PpiC-type peptidyl-prolyl cis-trans isomerase	-1.44	-1.17	-1.31	Chaperones/Heat shock proteins
Psyr_1650	pabC	aminodeoxychorismate lyase apoprotein	-1.66	-2.18	-1.92	Cofactor metabolism
Psyr_0487	gshB	glutathione synthase	-1.81	-1.27	-1.54	Glutathione metabolism
Psyr_0475		Protein of unknown function YGGT	-2.19	-1.37	-1.78	Hypothetical
Psyr_3581		ribosomal large subunit pseudouridine synthase B	-1.31	-1.08	-1.19	Nucleotide metabolism and transport
Psyr_2854		conserved hypothetical protein	-1.80	-1.76	-1.78	Phage & IS elements
Psyr_0377	mdoG	Periplasmic glucan biosynthesis protein, MdoG	-1.39	-1.62	-1.50	Polysaccharide synthesis and regulation
Psyr_0378	mdoH	Glycosyl transferase, family 2	-0.81	-1.27	-1.04	Polysaccharide synthesis and regulation
Psyr_1748	clpX	ATP-dependent Clp protease ATP-binding subunit ClpX	-1.09	-1.55	-1.32	Proteases
Psyr_0574	hflK	protease FtsH subunit HflK	-1.51	-0.81	-1.16	Proteases
Psyr_1749	lon-1	ATP-dependent proteinase, Serine peptidase, MEROPS family S16	-0.97	-1.09	-1.03	Proteases
Psyr_1410	ruvB	Holliday junction DNA helicase RuvB	-1.35	-1.33	-1.34	Replication and DNA repair
Psyr_1408	ruvC	Holliday junction endonuclease RuvC	-1.57	-1.28	-1.42	Replication and DNA repair
Psyr_4408	retS	Response regulator receiver:ATP-binding region, ATPase-like:Histidine kinase A, N-terminal:Histidine kinase:Histidine kinase	-1.28	-0.92	-1.10	Transcriptional regulation
Psyr_5133	trmE	tRNA modification GTPase trmE	-0.92	-1.66	-1.29	Translation
Psyr_1935	hcp-3	Protein of unknown function DUF796	-1.47	-0.60	-1.04	Type VI secretion system
Psyr_3954	lepA	GTP-binding protein LepA	-1.17	-1.56	-1.37	
Psyr_0529		Glycosyl transferase,	-1.57	-1.48	-1.52	

		group 1				
Psyr_4622		Nucleotidyl transferase	-0.98	-1.85	-1.42	
Psyr_4424		Propeptide, PepSY amd peptidase M4:PepSY-associated TM helix	-0.91	-1.20	-1.05	

B.

Locus	Name	Description	Epiphytic Fitness				Classification
			1	2	3	Average	
Psyr_0025	aroE	shikimate dehydrogenase	-1.32	-1.36	-1.06	-1.25	Amino acid metabolism and transport
Psyr_4132	hisC	histidinol phosphate aminotransferase apoenzyme	-1.76	-1.67	-1.17	-1.54	Amino acid metabolism and transport
Psyr_4133	hisD	histidinol dehydrogenase	-1.87	-0.89	-1.03	-1.27	Amino acid metabolism and transport
Psyr_4134	hisG	ATP phosphoribosyltransferase (homohexameric)	-1.63	-1.11	-0.84	-1.19	Amino acid metabolism and transport
Psyr_4896	hisH	imidazole glycerol phosphate synthase subunit hisH	-2.68	-1.93	-1.20	-1.94	Amino acid metabolism and transport
Psyr_1985	leuB	3-isopropylmalate dehydrogenase	-1.58	-1.19	-0.90	-1.23	Amino acid metabolism and transport
Psyr_1983	leuC	3-isopropylmalate dehydratase, large subunit	-1.53	-2.03	-0.41	-1.33	Amino acid metabolism and transport
Psyr_4894		1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-2.42	-1.35	-1.22	-1.66	Amino acid metabolism and transport
Psyr_0916	gmd	GDP-mannose 4,6-dehydratase	-1.64	-0.61	-0.77	-1.01	Carbohydrate metabolism and transport
Psyr_0826	pgi	glucose-6-phosphate isomerase	-2.08	-1.05	-2.55	-1.89	Carbohydrate metabolism and transport
Psyr_4842		Phosphoenolpyruvate-protein phosphotransf	-1.67	-1.49	-1.13	-1.43	Carbohydrate metabolism and transport

		erase					
Psyr_3179	ftsK	DNA translocase FtsK	-1.98	-1.15	-1.90	-1.68	Cell division
Psyr_0848	ilvC	ketol-acid reductoisomerase	-1.11	-1.40	-1.62	-1.38	Cofactor metabolism
Psyr_0469	ilvD	dihydroxyacid dehydratase	-1.62	-2.48	-1.64	-1.91	Cofactor metabolism
Psyr_0847	ilvH	acetolactate synthase, small subunit	-2.09	-2.01	-1.90	-2.00	Cofactor metabolism
Psyr_1120	zwf-1	glucose-6-phosphate 1-dehydrogenase	-1.09	-1.37	-0.60	-1.02	Energy generation
Psyr_0983		Protein of unknown function DUF159	-1.93	-1.43	-1.64	-1.67	Hypothetical
Psyr_0923		hypothetical protein	-1.64	-1.28	-1.43	-1.45	Hypothetical
Psyr_0534		membrane protein, putative	-1.70	-0.76	-1.11	-1.19	Hypothetical
Psyr_3805		hypothetical protein	-2.26	0.57	-2.54	-1.41	Hypothetical
Psyr_3691		conserved hypothetical protein	-1.53	-0.22	-1.32	-1.03	Hypothetical
Psyr_0412	gltD	glutamate synthase (NADPH) small subunit	-1.85	-0.65	-0.78	-1.09	Nitrogen metabolism
Psyr_4822	ntrC	Helix-turn-helix, Fis-type:Nitrogen regulation protein NR(I)	-1.80	-1.74	-0.61	-1.38	Nitrogen metabolism
Psyr_0294	ppx	Exopolyphosphatase	-0.87	-0.78	-1.58	-1.08	Nucleotide metabolism and transport
Psyr_3581		ribosomal large subunit pseudouridine synthase B	-0.61	-1.44	-1.57	-1.21	Nucleotide metabolism and transport
Psyr_0976	mqo	Malate:quinone-oxidoreductase	-1.89	-1.55	-1.62	-1.69	Organic acid metabolism and transport
Psyr_0630	mpl	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase	-1.62	-0.82	-1.23	-1.22	Peptidoglycan/ cell wall polymers
Psyr_3281	nagZ	Glycoside hydrolase, family 3, N-	-2.33	-0.80	-2.37	-1.83	Peptidoglycan/ cell wall polymers

		terminal					
Psyr_0402	ponA	Peptidoglycan glycosyltransferase	-1.20	-0.78	-1.17	-1.05	Peptidoglycan/ cell wall polymers
Psyr_4158	eftA	conserved hypothetical protein	-1.85	-0.52	-1.32	-1.23	Plant-associated proteins
Psyr_3636	wbpM	Polysaccharide biosynthesis protein CapD	-1.56	-0.76	-1.07	-1.13	Polysaccharide synthesis and regulation
Psyr_4887	ctpA	Peptidase S41A, C-terminal protease	-2.42	-1.64	-0.98	-1.68	Post-translational modification
Psyr_1747	clpP	ATP-dependent Clp protease proteolytic subunit ClpP	-0.95	-1.53	-1.16	-1.21	Proteases
Psyr_1748	clpX	ATP-dependent Clp protease ATP-binding subunit ClpX	-1.47	-1.27	-1.06	-1.27	Proteases
Psyr_1378	recA	RecA protein	-2.01	-0.97	-2.41	-1.80	Replication and DNA repair
Psyr_0201	recG	ATP-dependent DNA helicase RecG	-1.50	-0.58	-0.99	-1.02	Replication and DNA repair
Psyr_1410	ruvB	Holliday junction DNA helicase RuvB	-1.65	-1.38	-2.38	-1.80	Replication and DNA repair
Psyr_5065	uvrD	ATP-dependent DNA helicase UvrD	-1.84	-1.41	-2.30	-1.85	Replication and DNA repair
Psyr_0919		Chromosome segregation ATPase-like protein	-1.66	-1.05	-1.16	-1.29	Replication and DNA repair
Psyr_0832	cbrA-1	Two-component sensor kinase CbrA	-1.28	-1.65	-1.35	-1.43	Signal transduction mechanisms
Psyr_0811	terC	Integral membrane protein TerC	-1.11	-0.74	-1.46	-1.10	Stress resistance
Psyr_4128	cysD	sulfate adenylyltransferase subunit 2	-1.28	-1.07	-0.83	-1.06	Sulfur metabolism and transport
Psyr_4126	cysNC	adenylylsulfate kinase / sulfate adenylyltransferase subunit 1	-1.08	-1.15	-1.40	-1.21	Sulfur metabolism and transport
Psyr_4408	retS	Response regulator	-2.05	-1.39	-1.10	-1.51	Transcriptional regulation

		receiver:ATP-binding region, ATPase-like:Histidine kinase A, N-terminal:Histidine kinase:Histidine kinase					
Psyr_5133	trmE	tRNA modification GTPase trmE	-0.86	-1.29	-1.16	-1.10	Translation
Psyr_4567	anmK	Protein of unknown function UPF0075	-1.22	-0.90	-1.17	-1.10	
Psyr_4362	rlpA-2	Rare lipoprotein A	-2.16	-1.93	-1.67	-1.92	
Psyr_3637	wbpL	Glycosyl transferase, family 4	-1.68	-1.09	-1.04	-1.27	
Psyr_0936	wpbY	Glycosyl transferase, group 1	-1.77	-0.52	-0.87	-1.06	
Psyr_0914	wpbZ	Glycosyl transferase, group 1	-1.65	-0.78	-0.83	-1.09	
Psyr_0920		Glycosyl transferase, group 1	-1.78	-0.95	-1.05	-1.26	
Psyr_0947		TPR repeat protein:TPR repeat protein	-1.54	-1.02	-1.08	-1.21	
Psyr_1251		quinoprotein	-1.04	-1.55	-1.03	-1.21	
Psyr_0915		NAD-dependent epimerase/dehydratase	-1.63	-0.97	-0.92	-1.17	
Psyr_4623		Aminoglycoside phosphotransferase	-1.56	-0.96	-0.83	-1.12	
Psyr_4844		HAD-superfamily hydrolase, subfamily IB (PSPase-like):HAD-superfamily subfamily IB hydrolase, hypothetical 2	-1.60	-0.97	-0.66	-1.08	
Psyr_4078		AmpG-related permease	-1.32	-0.74	-1.02	-1.03	
Psyr_4622		Nucleotidyl transferase	-1.65	-1.47	-0.93	-1.35	
Psyr_1544		SirA-like protein	-1.92	-1.54	0.22	-1.08	

C.

Locus	Name	Description	Apoplastic Fitness				Classification
			1	2	3	Average	
Psyr_0385	hisE	phosphoribosyl-ATP pyrophosphatase	-2.30	-2.21	-0.92	-1.81	Amino acid metabolism and transport
Psyr_1109	edd	6-phosphogluconate dehydratase	-0.95	-1.86	-0.69	-1.17	Carbohydrate metabolism and transport
Psyr_0916	gmd	GDP-mannose 4,6-dehydratase	-1.06	-1.38	-1.58	-1.34	Carbohydrate metabolism and transport
Psyr_0758	scrB	beta-fructofuranosidase	-1.44	-2.02	-1.52	-1.66	Carbohydrate metabolism and transport
Psyr_1914	talB	transaldolase	-0.96	-1.82	-0.67	-1.15	Carbohydrate metabolism and transport
Psyr_3676	cobQ	adenosylcobyric acid synthase (glutamine-hydrolysing)	-2.19	-0.79	-1.22	-1.40	Cofactor metabolism
Psyr_3675	cobU	adenosylcobinamide kinase	-3.24	-2.59	0.44	-1.80	Cofactor metabolism
Psyr_3174	cysG	uroporphyrinogen-III C-methyltransferase / precorrin-2 dehydrogenase	-1.89	-1.78	-0.39	-1.35	Cofactor metabolism
Psyr_1542	nadA	quinolinate synthetase	-1.65	-1.15	-0.67	-1.16	Cofactor metabolism
Psyr_5130		chromosome segregation ATPase	-0.91	-1.18	-1.18	-1.09	Cofactor metabolism
Psyr_0923		hypothetical protein	-1.67	-1.56	-2.34	-1.86	Hypothetical
Psyr_5067		conserved hypothetical protein	-2.92	-1.07	-1.53	-1.84	Hypothetical
Psyr_0534		membrane protein, putative	-1.26	-2.07	-2.00	-1.78	Hypothetical
Psyr_0532		conserved hypothetical protein	-1.83	-1.16	-1.68	-1.56	Hypothetical
Psyr_0533		conserved hypothetical protein	-1.46	-1.23	-0.57	-1.09	Hypothetical
Psyr_5053		asparaginase	-1.34	-0.72	-1.12	-1.06	Nitrogen metabolism
Psyr_3690	purN	formyltetrahydrofolate-dependent phosphoribosylglycinamide formyltransferase	-1.84	-1.13	-1.25	-1.41	Nucleotide metabolism and transport
Psyr_4018	purU	Formyltetrahydrofolate deformylase	-1.76	-1.39	-1.82	-1.66	Nucleotide metabolism and transport
Psyr_4044		Phospholipase D/Transphosphatidylase	-1.84	-1.04	-0.25	-1.04	Phospholipid metabolism
Psyr_2601	salA	regulatory protein, LuxR	-1.46	-1.14	-1.05	-1.21	Phytotoxin synthesis and transport
Psyr_4158	eftA	conserved hypothetical protein	-1.99	-0.99	-1.28	-1.42	Plant-associated proteins
Psyr_1055	algI	Membrane bound O-acyl transferase, MBOAT	-1.02	-1.00	-1.10	-1.04	Polysaccharide synthesis and regulation
Psyr_1054	algJ	alginate biosynthesis protein AlgJ	-1.50	-1.76	-0.94	-1.40	Polysaccharide synthesis and regulation
Psyr_3636	wbpM	Polysaccharide biosynthesis protein CapD	-1.88	-1.87	-1.56	-1.77	Polysaccharide synthesis and

							regulation
Psyr_1747	clpP	ATP-dependent Clp protease proteolytic subunit ClpP	-0.99	-2.55	-1.39	-1.64	Proteases
Psyr_0574	hflK	protease FtsH subunit HflK	-0.89	-1.34	-0.81	-1.01	Proteases
Psyr_0201	recG	ATP-dependent DNA helicase RecG	-0.95	-1.36	-0.84	-1.05	Replication and DNA repair
Psyr_3287	topA	DNA topoisomerase I	-1.69	-1.49	-1.82	-1.67	Replication and DNA repair
Psyr_5065	uvrD	ATP-dependent DNA helicase UvrD	-1.52	-0.96	-1.01	-1.16	Replication and DNA repair
Psyr_4091		8-oxo-dGTPase	-1.52	-1.76	-1.46	-1.58	Replication and DNA repair
Psyr_0579	rnr	RNAse R	-1.79	-1.32	-0.67	-1.26	RNA degradation
Psyr_4843		NUDIX hydrolase	-0.87	-1.97	-0.99	-1.28	RNA degradation
Psyr_4882	secB	protein translocase subunit secB	-1.27	-0.69	-1.46	-1.14	Secretion/Efflux/Export
Psyr_4069	colS	ATP-binding region, ATPase-like:Histidine kinase, HAMP region:Histidine kinase A, N-terminal	-1.93	-0.79	-0.93	-1.22	Signal transduction mechanisms
Psyr_2077	cysB	regulatory protein, LysR:LysR, substrate-binding protein	-1.45	-1.17	-1.12	-1.24	Sulfur metabolism and transport
Psyr_3698	gacS	Response regulator receiver:ATP-binding region, ATPase-like:Histidine kinase, HAMP region:Histidine kinase A, N-terminal:Hpt	-1.47	-1.55	-1.43	-1.48	Transcriptional regulation
Psyr_5133	trmE	tRNA modification GTPase trmE	-0.74	-1.77	-1.29	-1.27	Translation
Psyr_1200	hrcC	outer-membrane type III secretion protein HrcC	-1.76	-1.15	-0.76	-1.22	Type III secretion system
Psyr_1213	hrcN	type III secretion cytoplasmic ATPase HrcN	-2.08	-1.78	-1.61	-1.82	Type III secretion system
Psyr_1208	hrcR	type III secretion protein HrcR	-1.93	-1.56	-1.00	-1.50	Type III secretion system
Psyr_1206	hrcT	type III secretion protein HrcT	-1.11	-1.22	-0.98	-1.10	Type III secretion system
Psyr_1205	hrcU	type III secretion protein HrcU	-1.61	-0.99	-0.92	-1.17	Type III secretion system
Psyr_1215	hrcV	Type III secretion protein HrcV	-1.25	-1.24	-1.23	-1.24	Type III secretion system
Psyr_1197	hrpE	type III secretion protein HrpE	-1.11	-2.74	-1.79	-1.88	Type III secretion system
Psyr_1216	hrpJ	type III secretion outer membrane protein PopN	-1.58	-1.48	-0.94	-1.34	Type III secretion system
Psyr_1218	hrpK1	type III helper protein HrpK1	-1.22	-1.50	-1.15	-1.29	Type III secretion system
Psyr_1211	hrpP	type III secretion protein HrpP	-2.00	-1.87	-1.87	-1.91	Type III secretion system
Psyr_1191	hrpS	type III transcriptional regulator HrpS	-1.62	-1.21	-0.94	-1.26	Type III secretion system
Psyr_5132	gidA	Glucose-inhibited division protein A subfamily	-1.25	-1.92	-2.65	-1.94	
Psyr_1588	rdgC	Putative exonuclease, RdgC	-2.81	-1.76	-0.91	-1.83	
Psyr_4362	rlpA-2	Rare lipoprotein A	-2.88	-0.85	-2.10	-1.95	

Psyr_1395		virulence	-1.40	-1.50	-1.84	-1.58	
Psyr_4566		Peptidase M23B	-2.14	-1.11	-1.34	-1.53	
Psyr_1417		TPR repeat protein	-1.06	-1.09	-0.89	-1.02	
Psyr_4844		HAD-superfamily hydrolase, subfamily IB (PSPase-like):HAD-superfamily subfamily IB hydrolase, hypothetical 2	-0.96	-1.16	-0.92	-1.01	

Chapter 3

The efflux transporter MexAB-OprM contributes to *Pseudomonas syringae* B728a virulence in a habitat, temporal, and host-dependent manner

Abstract

Efflux transporters contribute to bacterial resistance to diverse antimicrobial compounds. MexAB-OprM is a dominant toxin efflux transporter in *Pseudomonas* species, and is required for full virulence to plant hosts. Here, we show that MexB contributes to epiphytic growth, apoplastic growth, and lesion formation in common bean (*Phaseolus vulgaris*). However, MexB is not required for colonization of the apoplast of common bean for the first two days following inoculation, likely due to the delayed and gradual accumulation of defensive secondary metabolites in this compatible host. While the *mexB* deletion mutant forms fewer lesions after topical application to common bean, these lesions contain the same number of cells as the WT strain. The internalized population size of *P. syringae* within small segments of asymptomatic leaves varied from undetectably low to as high as 10^5 cells/cm^2. We propose that localized populations of bacteria within the leaf must exceed a threshold size of about 10^5 cells/cm^2 in order for a visible lesion to form. Strain B728a is capable of extensive apoplastic growth in diverse host plants including lima bean (*P. lunatus*), fava bean (*Vicia faba*), pepper (*Capsicum annuum*), *Nicotiana benthamiana*, sunflower (*Helianthus*), and tomato (*Solanum lycopersicum*). While growth continues in common bean for only approximately 48 hours after inoculation, in other hosts such as lima bean and *Nicotiana benthamiana* growth continues for a longer time and higher population sizes are eventually attained compared to those in common bean. MexB is not required for growth in a subset of plant species, particularly those in which growth is apparently not temporally constrained, indicating apparent variation in onset time or magnitude of plant chemical defenses among those hosts. The use of a hypersusceptible efflux pump mutant strain is an informative reporter to explore the diversity of host chemical immune responses.

Introduction

Plants produce diverse specialized antimicrobial metabolites to protect against potential pathogens and pests (116). The identity and abundance of these compounds differ between plant clades (117). Together, they constitute a chemical defense that potential plant pathogens must detoxify or tolerate to successfully colonize a host. Most antimicrobial compounds derived from plants can be placed into one of several chemical classes, including phenolics, terpenoids, alkaloids, and lectins/polypeptides (118). These specialized metabolites may be either pre-formed (“phytoanticipins”) or synthesized *de novo* in response to a pathogen (“phytoalexins”), although this distinction is poorly defined for many compounds (117). Plant-produced antimicrobial metabolites may target many cellular targets including plasma membranes, proteins, or nucleic acids in potential plant pathogens (116).

Bacteria have evolved several mechanisms to tolerate the presence of plant-produced antimicrobial compounds. These traits can therefore be considered necessary compatibility factors enabling the colonization of a given habitat or plant host. For

example, *Pseudomonas syringae* DC3000 requires several genes to overcome the toxicity of aliphatic isothiocyanates produced by *Arabidopsis* and these genes confer resistance to strains that are otherwise incapable of tolerating these molecules (119). These "survival in *Arabidopsis* extracts", or *sax,* genes encode both detoxifying enzymes as well as components of a transporter in the resistance-nodulation-division (RND) efflux protein family (119). Members of the RND family of transporters have been widely studied due to their relatively large contribution to bacterial resistance to antibiotics *in vitro.* Only now is there beginning to be further understanding of their broader role in other environmental habitats (120). By far the best-characterized RND transporter is AcrAB-TolC in *Escherichia coli* (121). Homologous multidrug resistance efflux transporters, such as the MexAB-OprM complex of Pseudomonads, have also been extensively examined. Related transporters have been shown to be required for full virulence of several model plant-colonizing bacterial, including *Ralstonia solanacearum*, *Erwinia amylovora*, and *Xylella fastidiosa* (122–124). A common strategy to identify the targets of efflux transporters involves documentation of increased tolerance of various antimicrobial compounds when the pumps are over-expressed (50, 125). While such strategies are an efficient method for identifying the potential substrates for efflux pumps, they do not always reveal the biological significance of such exporters. The local repressor of MexAB-OprM (PmeR) in *P. syringae* DC3000 responds to plant secondary products such as flavonoids (126), suggesting that these molecules are true substrates rather than artifacts of non-specific substrate binding sometimes seen when overexpressing efflux transporters. Members of the RND protein family are responsible for the removal of a diverse range of toxic substrates from the cytoplasm of Gram-negative bacteria. However, relatively little is known about the breadth of compounds that they are able to transport, particularly for the large number of naturally occurring inhibitory compounds. Because of this, there is a lack of insight as to whether these transporters function in an environmentally-specific manner and thus the extent to which each of the many such efflux pumps participate in any one of the myriad of different habitats that many bacteria may colonize. In this study we use a hyper-susceptible *mexB* mutant strain as a bioreporter to examine differences in chemical defense operative against a plant pathogen capable of colonizing several susceptible host plant species.

P. syringae pv. *syringae* B728a (B728a) is pathogenic of common bean (*Phaseolus vulgaris*) and *Nicotiana benthamiana* (37), and was recently reported to cause disease symptoms on pepper (*Capsicum annuum*) (38). Strain B728a has also been reported to be a poor colonizer of tomato (*Solanum lycopersicum*) (127). Within this strain, MexAB-OprM (*Psyr_4007-9*) is required for full virulence on common bean (39). In strain B728a and the related strains DC3000 and 1448A, this multidrug efflux pump is capable of conferring resistance to a broad range of toxic substrates including antibiotics, plant-derived antimicrobial compounds, detergents, and dyes, presumably by exporting them from the cell (39). MexB is located in the inner membrane, and is responsible for substrate specificity of the protein complex (128), while the outer membrane protein OprM functions also with other RND efflux systems to enable passage of toxicants from the cytoplasm to the exterior of the cell (129). The process of infection of plants by foliar phytopathogens such as *P. syringae* is relatively well characterized. Typically, an epiphytic colonization stage facilitates sufficient

multiplication of immigrant bacteria on leaves to enable invasion of the leaf interior, that can be followed by growth in the apoplast leading to disease symptoms on leaves under favorable environmental conditions and if plants are susceptible to infection (58). Foliar lesions result from the localized multiplication of bacteria inside leaves. The release of water from the plant into intercellular spaces in leaves creates watersoaking that facilitates release and distribution of nutrients from plant cells, is apparently induced by effectors produced by bacteria (20). The assembly of large numbers of bacteria within leaves often causes chlorosis and subsequent necrosis and dehydration in such lesion sites (22). Bacteria within disease lesions can function as the source of inoculum for spread of secondary inoculum via water splash or aerial dispersal (130). While there has been extensive study of bacterial traits needed for growth within plants after their introduction into the plant interior, there has been relatively little study of the early stages of infection, particularly the transition between epiphytic and endophytic populations that lead to disease symptoms. Here we examine these various stages of bacterial growth *in planta* that are ultimately required for infection with special attention to the role of efflux transporters such as MexAB-OprM. Most attention to the process of infection has focused on particular model plants that interact with a given pathogen. Many plant pathogenic bacteria, particularly those in *P. syringae* phylogroup 2, are now recognized to have relatively broad host ranges, and yet it remains unclear to what extent a given virulence factor such as an efflux transporter would contribute to virulence on plants that would be expected to differ substantially, for example, in chemical defenses. In this study, we will address the extent to which a given *P. syringae* strain is a promiscuous colonizer of a variety of plants and the relative importance of MexAB-OprM to its virulence. In this process we will elucidate its role during colonization of diverse plant hosts.

Within the genus *Phaseolus*, 5 species have been domesticated. There is also abundant diversity within common bean (*P. vulgaris*) since it was domesticated from two geographically divergent gene pools (131). The range of susceptibility within *Phaseolus* species to infection by compatible bacterial pathogens indicates a level of quantitative resistance, which might readily be explained by variation in their content of antimicrobial secondary metabolites (132). Based on preliminary studies that revealed that that the growth of a Δ*mexB* mutant strain was inhibited in common bean but not lima bean compared to that of wild type strain, we hypothesized that these susceptible hosts differed in their immune responses to an otherwise compatible pathogen. Furthering this conjecture, we hypothesized that susceptible host plants would vary in their antimicrobial chemical defenses, and that not all antimicrobial metabolites present in a given plant would be substrates of MexAB-OprM, and instead would be the substrates for alternative efflux transporters. By examining the growth of the hyper-susceptible B728a Δ*mexB* strain in a variety of plant species, we addressed the apparent differential role of chemical defenses in these important crop plant species.

Results

MexB contributes to fitness on the leaf surface

Given that B728a Δ*mexB* had been shown to be hyper-susceptible to diverse antimicrobial compounds (39), we reasoned it might be an effective bioreporter strain to detect chemical toxicants present in and on the leaves of various plant species. The

MexAB-OprM efflux pump is apparently completely dispensable for growth in the rich medium King's B (KB) since no difference in growth was seen between it and the wild type strain (Fig. 1). However, when sprayed onto the leaves of the susceptible common bean variety Bush Blue Lake 274, subsequent epiphytic bacterial populations established by Δ*mexB* after 2 days incubation on moist leaves were significantly lower than that of the wild-type (WT) strain (Fig. 2). This suggests that a chemical defense or other inhibitory compound was present on the leaf surface, necessitating the role of the MexAB-OprM efflux transporter for full fitness in this habitat**.**

MexB contributes to late-stage multiplication and survival in common bean

When inoculated directly into the apoplast of common bean, the population size of both WT B728a and Δ*mexB* increased similarly with time for two days (Fig. 3A). While little additional increase in population size of the WT strain occurred after two days, the population size of Δ*mexB* tended to decrease with time and its population in the apoplast at 6 and 8 dpi was significantly lower than that of wild type strain (Fig. 3B). The plateauing of population sizes after two days incubation, particularly for that of the WT strain suggested either that this maximal apoplastic population reflected an intrinsic carrying capacity of the leaf interior for this species or the intervention of an active immune response from the plant that limited further growth after such a time. To discriminate these scenarios we inoculated the interior of leaves with bacterial inoculum of varying concentrations and measured the resultant bacterial population sizes after four days - sufficient time for even low initial numbers of bacteria to reach high internal leaf populations. Importantly, irrespective of the number of cells initially introduced into the leaf interior, bacterial population sizes increased approximately 1000 fold and thus were highly correlated with the original inoculum concentration (Pearson's correlation coefficient = 0.90) (Fig. 4). Such a result is quite inconsistent with the concept of carrying capacity, and instead is highly supportive of a model in which, irrespective of their initial concentration in the plant, bacterial cells are able to grow only for a limited period of time, apparently about two days, before growth is halted. Such a pattern likely reflects a delayed but effective resistance response in the host, and because of hypersensitivity of Δ*mexB* to conditions after this time point, presumably chemically mediated plant response.

Bacterial population size in leaves is driven by the number of visible lesions

When B728a is sprayed onto leaf surfaces, macroscopic lesions typically form only after about 7 days. Some leaves seem to be more susceptible to lesion formation than others, resulting in a rather strongly right hand skewed frequency distribution on the number of lesions is considered over a large number of leaves (Fig. 5). While it is unclear why some leaves appear to be more susceptible to infection than others, younger leaves tended to have more lesions than more fully developed trifoliate leaves. Overall, B728a Δ*mexB* formed fewer lesions per leaf than WT (Fig. 5). Inoculation with the WT strain resulted in a median of 15 and a mean of 28.8 lesions per leaf, while inoculation with the Δ*mexB* strain resulted in a median of 2 and a mean of 5.7 lesions per leaf. It was interesting to note however, that the number of viable cells these two strains found within discrete lesions that were excised from these leaves did not differ statistically (Fig. 6). The mean number of WT cells recovered from a given lesion was

1.3 x 10^6 while 8.7 x 10^5 were recovered from Δ*mexB* lesions. Viable cells of these two strains could also be detected within similarly small segments of healthy leaves that were sampled at the same time and in a similar way. While most leaf segments harvested either 7 or 21 days after inoculation harbored no detectable cells, samples from many asymptomatic areas of the leaf contained up to 10^5 CFU/cm^2 (Fig. 7). Thus while internal colonization of leaves by both strains was relatively common, it is important to note that disease symptoms were observed only when local internal population sizes in excess of about 10^5 cells/cm^2 were present.

Common bean varieties vary in their susceptibility to B728a

Given that some evidence for chemically-mediated resistance to strain B728a was observed in the susceptible *P. vulgaris* variety Bush Blue Lake 274, we explored the variation in apparent disease resistance within the *Phaseolus* genus. Wild type strain B728a was topically applied to the leaves of 17 additional varieties of *P. vulgaris*, 2 varieties of lima bean *(P. lunatus*), and 1 variety of tepary bean (*P. acutifolius*) and the incidence of subsequent lesion formation was enumerated (Table 1). There was a wide range in the number of lesions formed per leaf among the varieties of *P. vulgaris* tested, with no lesions forming on 3 varieties (Pinto San Rafael, SEA5, and Tio Canela). In addition, *P. lunatus* var. Haskell and var. UC92 were more susceptible, and *P. acutifolius* var. G10 less susceptible, to lesion formation than *P. vulgaris* var. Nichols (Fig. 8).

MexB is not required for apoplastic colonization of all plant hosts

Since varieties of common bean differed in their apparent susceptibility to lesion formation by B728a we hypothesized that their relative disease susceptibility was proportional to their ability to support the growth of this strain in the apoplast. We therefore introduced either the WT or Δ*mexB* strain into the leaf interior and measured increases in population size with time. Variation in the number of lesions formed on *Phaseolus* species was predicted by the extent of apoplastic growth on a subset of these hosts that were tested. For example, while no apoplastic growth of B728a was measured in *P. vulgaris* var. Anasazi (Fig. 9A), a variety for which very few leaf lesions had been observed (Fig. 8). In contrast, *P. lunatus* (lima bean) supported both much higher levels of internal bacterial colonization than observed in *P. vulgaris* var. Blue Lake Bush or var. Nichols (Fig. 9B) and also very high numbers of leaf lesions following spray application to leaves (Fig. 8). Interestingly, the population size of both Δ*mexB* and the WT strain were similar and both were much higher than those on common bean varieties Blue Lake Bush and Nichols (Fig. 9) suggesting that unlike in these other bean varieties MexB was not required for high levels of growth in lima bean. In both Blue Lake Bush and Nichols varieties the WT strain reached population levels over 10-fold greater than the Δ*mexB* mutant by 4 and 6 days after inoculation.

To expand beyond the Fabaceae, we also tested additional plant species for their ability to support apoplastic growth of strain B728a and the requirement of MexB for such growth. *P. syringae* strain B728a is capable of apoplastic growth in several additional plant hosts. While very high apoplastic population sizes were seen in both *N. benthamiana* and pepper (*C. annuum*), growth within leaves was much lower (100- to 1,000-fold increase over 6 days) in sunflower and tomato. Surprisingly, *mexB* was not

required for virulence in all of the host plants tested. While *mexB* was required for full growth in pepper (*C. annuum*), the ∆*mexB* mutant strain grew as well as the WT strain in fava bean (*V. faba*), *N. benthamiana*, sunflower (*Helianthus*), and tomato (*S. lycopersicum*) (Fig. 10). Disease symptom development was generally proportional to the extent of apoplastic growth of these strains. While extensive necrotic lesions were incited by both the wild type and ∆*mexB* mutant formed on *N. benthamiana,* only mild discoloration was induced by the wild type strain but not ∆*mexB* on pepper. No symptoms were seen on any of the other plant species that harbored lower apoplastic populations of both strains. In contrast to the several other plant species for which some apoplastic growth but no symptoms were observed, very little growth of the wild type strain was seen in either black mustard (*Brassica nigra*) or red clover (*Trifolium pratense*) (Fig. 11).

Apparent lack of chemical defenses limiting lesion formation in *N. benthamiana*

In addition to the robust apoplastic growth of strain B728a in *N. benthamiana* that continued well beyond the 2 to 4 days seen in common bean varieties, achieving higher population sizes than seen on bean (Figs. 9 and 10), we also observed greater numbers of lesions formed relative to that on common bean. Furthermore, while lesion formation on bean was relatively synchronized, with most lesions that would eventually appear on a given leaf becoming apparent about seven days after topical application of pathogen to the leaf surface (Fig. 12) lesions on *N. benthamiana* leaves often became apparent sooner after inoculation those on common bean, and more importantly, the number of lesions on a given leaf continued to increase with time (Fig. 12). The continuing appearance of new lesions on *N. benthamiana* with increasing time after inoculation contributed to its being a much more susceptible host than common bean. The continual appearance of new lesions with time on *N. benthamiana* apparently is associated with a weak or lacking chemical defense to bacterial colonization as evidenced by the continued increase in apoplastic bacterial population size with time (Fig. 10) and the dispensability of MexB in the colonization of this host by strain B728a.

We hypothesized that since chemical defenses were apparently operative in common bean 2 to 4 days or more after inoculation of strain B728a into leaves, bacterial populations in lesions would decline over time, as plant-produced antimicrobial compounds would accumulate with time. Furthermore, given that the ∆*mexB* mutant strain is apparently more susceptible to plant derived toxicants than the WT strain, its viability would decline at a faster rate than the WT strain. To test this, viable bacterial population sizes were determined in lesions excised from infected leaves at various times after formation. Bacterial population sizes within a given lesion of common bean, lima bean, and *N. benthamiana* generally all ranged from about 10^5 to 10^6 CFU soon after formation. However, the population size of both the WT and ∆*mexB* mutant strains declined over time up to 28 days after inoculation in both common bean and lima bean, but conspicuously, not in *N. benthamiana* (Fig. 13). Furthermore, viable populations of the ∆*mexB* mutant strain tended to decrease more rapidly with time in lesions in lima beans than that of the WT strain (Fig. 13).

Discussion

Not only was the type strain for *P. syringae* originally isolated from diseased lilac, but most research on this species has focused on its ability to colonize diverse plant species and cause disease (58). Strain B728a, isolated from infected common bean, is a member of phylogroup 2, which consists of many pathogenic strains that are also capable of causing ice nucleation, but also includes many strains that are non-pathogenic and can be found in a variety of environmental settings not obviously associated with plants (133). The paradigm of pathogenesis for such foliar pathogens posits that infection follows the multiplication of cells on plant surfaces following their immigration from other plants or airborne or other environmental reservoirs. Only following entry of at least some of these epiphytic cells into the leaf interior by passage through stomata, hydathodes, or other opening such as wounds do cells gain entry into the intercellular spaces where, through their secretion of various effectors and toxins, modify the plant environment to one where extensive growth can occur. Disease symptoms seem normally to only follow the development of large internalized populations. Strains such as B728a are particularly strong epiphytic colonizers (58, 106), and their high epiphytic populations can increase the likelihood of movement through stomata or leaf wounds into the apoplast and the formation of visible symptoms (54). Though stomata close in response to detecting microbes through the perception of pathogen-associated molecular patterns (PAMPS), *P. syringae* can secrete various phytotoxins (dependent on strain) to induce their reopening (17). Foliar lesions usually follow bacterial-induced water soaking, with lesions subsequently appearing necrotic and dry (22).

If bacterial populations above a specific threshold were required for the formation of a visible lesion, we would expect asymptomatic areas of the leaf to host bacterial cells only below that threshold. While this was observed in this study, the frequency at which cells were detected within these asymptomatic areas of the leaf was relatively low. A higher number of samples would have provided a better estimate of the threshold bacterial population size below which disease symptoms are not seen. Based on the sampling time in this study, it would appear that such an apoplastic bacterial threshold is approximately 10^5 cells per site (Fig. 7). The Δ*mexB* mutant of strain B728a formed fewer lesions on common bean in this study, and such mutants in other *P. syringae* strains also have been observed to form fewer lesions than WT strains (39). We therefore hypothesized that while this strain might enter leaves in a similar frequency as that of the WT strain, its subsequent localized apoplastic population sizes would tend to be lower than that of the WT strain because of its apparent hyper-sensitivity to induced host defenses at such a site. As such, it would be expected to be less capable of reaching the necessary “threshold population” for symptom formation. While an attractive model, it proved difficult to test because of the relatively few sites in which any apoplastic growth of these strains occurred. While there were recoverable B728a cells in some surface-sterilized asymptomatic leaf samples, the majority of samples contained undetectably low numbers of viable bacteria. Future work will need to assess many more samples than was feasible in this study. Areas of the leaf without detectable bacterial cells likely represented areas where cells were unable to enter or grow in the apoplast. In a low humidity environment, such as the greenhouse, bacterial populations on the leaf surface decline over time (19), and, coupled with the absence of environmental features such as large raindrops that could facilitate the entry of cells into

the plant, internal colonization and subsequent lesion formation is probably less than what might be expected under certain field conditions. It is noteworthy however that under these greenhouse conditions, when extrapolating across the area of the leaf as a whole, there were many more sites where *P. syringae* was detected within a leaf than a number of sites where lesions formed. It would thus appear that lesion formation might be a relatively uncommon outcome following the introduction of this pathogen into the apoplast.

While bacterial populations within lesions of common bean and lima bean both decreased with lesion age, there were approximately 10^6 CFU per lesion 7 days after epiphytic inoculation, soon after lesions first appear (134), suggesting that such a threshold population size must be achieved at a given site for symptom development. Given that epiphytic bacterial populations presumably must always precede lesion formation, by providing inoculum to enter and subsequently grow within the apoplast, there have been efforts to anticipate disease from knowledge of epiphytic communities. The concept of an "infection threshold" has been used to predict disease incidence at the field level, with a focus on correlating the population size of epiphytic pathogens with disease severity later in the growing season (54, 135). Such predictive models are premised by knowledge of 1) the frequency with which a given leaf would harbor an epiphytic population size of a given size or larger and 2) the likelihood that cells within such a population would enter the leaf, multiply, and cause lesion formation. Mechanistically, feature #2 is driven by the likelihood of lesion formation, given a particular apoplastic bacterial population. If lesion formation requires surpassing a certain internal population size threshold, then the Δ*mexB* strain is hindered in virulence because it appears to be more sensitive to chemical defenses that it elicits during the infection process. Because of that, fewer sites of colonization are likely to reach the threshold for symptom development (Fig. 14). Perhaps more importantly, Δ*mexB* strain suffers from issues related to feature #1. Specifically, the Δ*mexB* strain achieves a lower epiphytic bacterial population than the WT strain apparently because of its sensitivity to inhibitory compounds found on the leaf surface (Fig. 2). As such, fewer leaves (or leaf sites) would harbor the relatively high epiphytic populations that are associated with meaningfully high frequencies of invasion of the leaf (Fig. 14).

Following the observation that bacterial populations within lesions decline over time in common bean and lima bean, we hypothesized that the Δ*mexB* strain would decline at a faster rate. This hypothesis is driven by our conjecture that the accumulation of inhibitory chemicals continues to increase with time after lesion formation, eventually becoming sufficiently high that even the WT strain cannot tolerate them. If the plant were actively reducing the bacterial load on the leaf by producing antimicrobial metabolites, then the Δ*mexB* strain would be more susceptible, and therefore less able to survive in this local environment. While greater mortality of the Δ*mexB* mutant strain in lesions as they aged was not apparent on common bean, its population tended to decline more rapidly than that of the WT strain on lima bean. Importantly, there was a dramatically lower rate of population decline of both the WT and Δ*mexB* mutant between lesions in lesions in common bean and lima bean compared to that in *N. benthamiana* (Fig. 13). Lesions on *N. benthamiana* leaves maintained high bacterial populations for 28 days. We have generally observed that *N. benthamiana* is particularly susceptible to infection by B728a, forming many more

lesions following topical inoculation with a given concentration of cells. The growth within the apoplast is also more extensive and prolonged than that in either common bean or lima bean (Figs. 9 and 10). It is our conjecture that this hyper-susceptibility of *N. benthamiana* is at least partially due to a lack of a robust chemical defense. Conversely, the initial chemical defenses of common bean variety Bush Blue Lake appear to be much more robust than that of lima bean. Both *Phaseolus* species however appear to accumulate antimicrobial compounds in lesions, resulting in the death of cells as lesions age. However, based on our observations, these antimicrobial compounds are not necessarily substrates of MexB. For example, plants generate redox-active molecules as part of the basal immune response (39) and these may differ between plant species. Efflux transporter such as MexB thus appear to play an important epidemiological role in species such as *P. syringae* not only in facilitating initial colonization of plants by contributing to the tolerance of stressful chemical conditions on the leaf surface and early in lesion formation, but perhaps more importantly, by maintaining the viability of cells in lesions as they age. Such lesions provide the secondary inoculum to facilitate epidemics of diseases caused by *P. syringae*, and hence inoculum survival is an important process in the disease cycle.

The efflux transporter MexAB-OprM has been previously shown to contribute to virulence of *P. syringae* B728a on common bean leaves (39). However, this study was not temporally intensive and focused only on long-term differences in bacterial populations assessed in whole infected leaves once a week for three weeks after inoculation. Stoitsova *et al.* noted that lesions formed by Δ*mexB* were less abundant but morphologically indistinguishable from those formed by the WT strain (39). We confirm these findings, and we also show that bacterial populations within these lesions are not significantly different between the two strains. This decrease in the number of lesions, given that all lesions contain approximately the same number of bacterial cells, likely contributes to the decrease in Δ*mexB* mutant cells per leaf that has been observed previously (10, 56). This suggests that the difference in total population sizes previously observed in whole leaves is due primarily to the decrease in total lesions formed by Δ*mexB*. While we observed that bacterial cells were present in asymptomatic areas of the leaf, these cells potentially represent a small fraction of the leaf total when considering the high density of bacterial cells present within a single lesion. Bacterial cell density on the leaf surface is heterogeneous, presumably due to the heterogeneous distribution of nutrients (18). At higher densities, quorum-sensing-dependent traits such as swarming and alginate production promote movement into the leaf and overall virulence (136). The uniformity in bacterial populations within lesions suggests that there may be a bacterial population threshold that must be surpassed for the formation of visible lesions on the leaf surface, and hence the number of lesions on a leaf may be driven by the number of localized sites where epiphytic bacterial populations are sufficiently large that invasion into the apoplast is successful; lesions would form at such sites only if subsequent apoplastic growth was sufficient to exceed a threshold for symptom development. Since *mexB* is required for full growth on the leaf surface and in the apoplast, it would contribute to both the frequency with which bacteria might enter the plant, and also its success in multiplying once in the plant - thereby influencing the number of lesions that would form (Fig. 14).

It is intriguing that topical inoculations of the primary leaves of common bean with *P. syringae* frequently do not yield foliar infections, unlike trifoliate leaves. Such a finding is surprising given that these strains multiply well in primary leaves. *P. syringae* appears to grow equally well in primary and trifoliate leaves of common bean after inoculation into intercellular spaces and it is primarily due to the practicality of inoculating smaller plants with primary leaves that studies of its colonization of this species is typically performed at this stage of plant development. These observations suggest that there is a difference in disease susceptibility between leaf types, or leaf age. Age has been shown to impact disease susceptibility in plants. For example, *Arabidopsis* plants become more resistant to *P. syringae* as they age (137). Early studies in soybean showed infection by *Sclerotinia sclerotiorum* caused symptoms on trifoliate leaves and not primary leaves (138–141). It is unclear whether the preferential formation of lesions on trifoliate leaves compared to primary leaves is due to easier entry into the leaf from epiphytic sources or differences in subsequent apoplastic growth, or even another mechanism. Addition work is needed to confirm this phenotype in common bean, and to identify a mechanism for this variation in disease susceptibility.

Of the plant species tested here, the ability of the host to restrict growth of *P. syringae* by chemical defenses in a MexB-dependent manner was most pronounced in common bean and pepper. Furthermore, even within common bean there appears to be a substantial difference in the contribution of MexB to apoplastic growth and lesion formation (Figs. 8 and 9). This diversity in apparent disease susceptibility within both common bean and lima bean has been noted also for diseases caused by fungal pathogens such as anthracnose (*Colletotrichum lindemuthianum*) and angular leaf spot (*Phaeoisariopsis griseola*) (142). Hagedorn *et al.* screened 383 lines of lima bean for tolerance to *P. syringae* and showed some variation in disease intensity in the field, but similar susceptibilities of all lines in the greenhouse (143). A panel of 215 common bean lines was tested in the field for susceptibility to common bacterial blight (*Xanthomonas axonopodis* pv. *phaseoli*), and these measurements formed a unimodal distribution around an intermediate level of disease susceptibility (144). We hypothesized that the Δ*mexB* mutant strain was limited in growth in a host- or cultivar-specific manner due to the differential presence of antimicrobial metabolites produced by these plants. Lyon and Wood showed that *P. vulgaris* produces several phytoalexins including coumestrol and phaseollin that gradually accumulate in the apoplast within one to five days following infection by *P. phaseolicola* (145–147). These compounds can also form in bean in response to wounding, fungal infection, and viral infection (117). Such phytoalexins, produced after infection, therefore appear to have a primary role in limiting the success (growth) of pathogens after infection has occurred rather than preventing infection, as is the case for effector-induced resistance responses. It thus appears that there is commonly a delay in accumulation of such phytoalexins levels that would be toxic to pathogens. It is tempting to speculate that the strong temporal dynamics of growth of *P. syringae* in plants such as common bean may be explained based on the accumulation of phytoalexins in plants only two days or more following infection. It is noteworthy that strain B728a multiplied rapidly only for about two days, at which point growth largely ceased, independent of the population size that it had achieved by that time (Figs. 3 and 4). Intriguingly, MexB only contributed to growth of strain B728a in plants such as common bean variety Bush Blue Lake or Nichols two days or more after

inoculation. The efflux transporter MexAB-OprM may be required to tolerate the presence of these plant defense compounds as they accumulate to high levels in the apoplast. As different plant species produce different defensive compounds (148), despite its wide substrate range MexB is most likely not capable of excluding all potential antimicrobial molecules. It is tempting to speculate, therefore, that those common bean varieties and other plant species in which strain B728a achieves the lowest population size and for which MexB was required for maximal growth produced inhibitory compounds and that these compounds were also substrates for MexB. Alternatively, one or more alternative efflux transporters found in strain B728a may play a more important role than MexB in tolerating chemicals found in lima bean and *N. benthamiana*.

Gram-negative bacterial species are typically highly resistant to toxins *in vitro* due to the activity of a variety of conserved multidrug resistance efflux pumps. However, the chemical inhibition of these pumps can result in the increased toxicity of plant-produced antimicrobial compounds, including in *P. syringae* (149). It is possible that some of the hosts tested here, such as tomato, in which *mexB* appears dispensable are simply producing one or more synergistically acting efflux pump inhibitors, rendering the transport complex non-functional. This has been seen in barberry (*Berberis* species) plants, which produces the alkaloid berberine as well as the multidrug inhibitor 5'-methoxyhydnocarpin (38). In addition, we observed a substantial variation in the amount of bacterial growth in the apoplast of the plants tested here, with some plant species such as *N. benthamiana* supporting more than 10,000-fold more multiplication than in species such as *V. faba* or *B. nigra* (Figs. 10 and 11). The factors limiting growth in these species remains unknown. Given that the population size of strain B728a increased 100-fold or more after inoculation in all species except *T. pretence* (Figs. 10 and 11), this suggests that they might all be considered hosts for this *P. syringae* strain. Considering the relatively low population sizes that were achieved in several of these species it is not surprising that symptoms were not seen in this study nor have they been typically seen under field conditions. These results however do add to the growing perspective that not all strains of *P. syringae*, and perhaps other plant pathogenic bacteria, are as specific as once thought (38).

B728a is capable of causing disease in diverse host plants (60). While the efflux transporter MexAB-OprM is generally considered a necessary virulence factor in this and other Gram-negative bacteria, here we show that it is contributes to virulence in a temporally- and host-specific manner. The presence of this pump contributes to fitness on the leaf surface, but it appears that it is required in the apoplast only once host chemical defenses have been activated. Further exploration is necessary to better understand the specific metabolic differences between these hosts, the substrate specificities of MexB *in planta*, as well as whether other pumps are complementary to, or replace, the role of MexAB-OprM for tolerating the toxins present in other settings.

Materials and Methods

Bacterial strains and growth media

Pseudomonas syringae pv. *syringae* B728a was originally isolated from a common bean leaf (*Phaseolus vulgaris*) in Wisconsin (107). B728a and the kanamycin-resistant derivative mutant strain Δ*mexB* were grown on King's B (KB) agar or in broth

(111), at 28˚C. *Escherichia coli* strains S17-1, and TOP10 were grown on LB agar or in LB broth at 37˚C. When appropriate, the following antibiotics were used at the indicated concentrations: 100 µg/ml rifampicin, 50 µg/ml kanamycin, 15 µg/ml tetracycline, 40 µg/ml nitrofurantoin, and 21.6 µg/ml natamycin (an anti-fungal).

Construction of the B728a Δ*mexB* mutant strain

The construction of the Δ*mexB* strain followed an overlap extension PCR protocol as described previously (112). DNA fragments upstream (1 kb) and downstream (1.3 kb) of *Psyr_4008* (*mexB*) were amplified along with a kanamycin resistance cassette from pKD13 (113). These three fragments were joined by overlap extension PCR. The resulting fragment was blunt-end ligated into the SmaI site of pT*sacB* (114), and transformed into the *E. coli* subcloning strain TOP10, and then the *E. coli* conjugation donor strain S17-1. This suicide plasmid was conjugated into B728a on KB overnight, and then selected by growing the exconjugants for 3 days on KB containing kanamycin and nitrofurantoin (*E. coli* counter selection). Putative double-crossover colonies that were kanamycin resistant and tetracycline sensitive were selected for screening using external primers and further confirmed by PCR and Sanger sequencing.

Bacterial *in vitro* growth measurements

B728a WT and Δ*mexB* overnight cultures were grown in KB containing rifampicin and standardized to OD_{600} = 0.3. 20 µl was inoculated into a 96-well plate containing 100 µl King's B broth (5 replicate wells each). Cells were grown at 28˚C with shaking, with absorbance measurements made at 600 nm every 30 minutes. Absorbance is reported as the mean and standard deviation of absorbance at 600 nm with the average absorbance of two KB blanks subtracted from the total absorbance for each sample.

Plant growth conditions

All *Phaseolus* varieties with the exception of Bush Blue Lake were acquired as seed from UC Davis. Common bean (*P. vulgaris*), lima bean (*P. lunatus*), tepary bean (*P. acutifolius*), fava bean (*V. faba*), sunflower (*Helianthus*), mustard (*B. nigra*), and clover (*T. pratense*) seeds (5 - 7 per 10 cm diameter pot) were planted in Super Soil (Scotts MiracleGro) and grown in a greenhouse until primary leaves were fully expanded before inoculation (approximately two weeks for bean species). *N. benthamiana,* pepper (*C. annuum* cv. Cal Wonder), and tomato (*S. lycopersicum* cv. Moneymaker) seeds were sown on Sunshine Mix #4 (SunGro Horticulture) to germinate, and then transplanted into 10 cm diameter pots containing Super Soil. *P. vulgaris* var. Blue Lake Bush 274 was used for all experiments in common bean unless otherwise specified. Leaves were kept dry to minimize epiphytic contamination. 1000 W metal halide lights were used to provide supplemental lighting for a 16-hour day length. Greenhouse temperatures ranged from 18˚C at night to approximately 40˚C during the day.

Spray inoculations to measure bacterial epiphytic population size

Strains were grown overnight on KB plates containing rifampicin, washed in 10 mM KPO_4, (pH 7.0) and resuspended in 10 mM KPO_4. Cultures were diluted to a

concentration of 10^6 CFU/ml (OD_{600} = 0.001, by dilution from OD_{600} = 0.1) and sprayed onto the surface of trifoliate leaves until runoff. Plants were placed in a high humidity mist chamber for two days. After this time, leaves were removed and placed into glass tubes containing 25 mL 10 mM KPO_4, where cells were dislodged from the leaves using a water bath sonicator (Branson 5510, output frequency 40 kHz) for 15 min. Bacterial populations were measured by dilution plating onto KB agar containing rifampicin and natamycin.

Bacterial apoplastic growth measurements

Strains were grown overnight on KB plates containing rifampicin, washed in 10 mM KPO_4, and standardized to $2x10^5$ CFU/ml (OD_{600} = 0.0001, by dilution from OD_{600} = 0.1) (*Phaseolus* spp. and *Helianthus*), 10^4 CFU/ml (*C. annuum*, *S. lycopersicum*, and *V. faba*), or 10^3 CFU/ml (*N. benthamiana*, *Brassica*, and *Trifolium*) in 1 mM KPO_4. Cells were inoculated into leaves using a blunt syringe. Leaf samples (3 discs per leaf) were excised using a 5 mm-diameter cork borer into tubes containing 200 µl 10 mM KPO_4 and two 3 mm glass beads, and macerated for 30 seconds at 2400 rpm in a Mini-Beadbeater-96 (Biospec Products) before dilution plating on KB with rifampicin and natamycin.

Spray inoculations to measure lesion formation and bacterial populations within symptomatic/asymptomatic areas of leaf tissue

Bacterial strains were grown overnight on KB plates containing rifampicin, washed in 10 mM KPO_4, and resuspended in 10 mM KPO_4. Cultures were diluted to a concentration of 10^7 CFU/ml in 10 mM KPO_4 (OD_{600} = 0.01, by dilution from OD_{600} = 0.1) and sprayed onto the surface of trifoliate leaves of Bush Blue Lake bean until runoff. Due to its higher susceptibility, *N. benthamiana* plants were sprayed with inoculum diluted to 10^5 CFU/ml. Plants were placed in a high humidity mist chamber for two days, and then moved to the greenhouse. Individual lesions and asymptomatic areas of leaves were sampled at least 7 days after spraying, the specific timing noted for each experiment. Leaf samples (1 disc = 1 lesion) were excised with a 5 mm-diameter cork borer and placed in tubes containing 200 µl 10 mM KPO_4 and two 3 mm glass beads, and ground for 30 seconds at 2400 rpm in a Mini-Beadbeater-96 before dilution plating on KB with rifampicin and natamycin.

Acknowledgements

We thank Dr. Paul Gepts at UC Davis for providing all seeds for all *Phaseolus* species and varieties, with the exception of *P. vulgaris* var. Blue Lake Bush. Plants were grown and maintained with the assistance of the UC Berkeley greenhouse staff, led by Tina Wistrom. We thank Kendra Paskvan and Caitlin Ongsarte for assistance with plant inoculations and sampling. Funding for Tyler Helmann was partially provided by the Arnon Graduate Fellowship and the William Carroll Smith Fellowship.

Figures

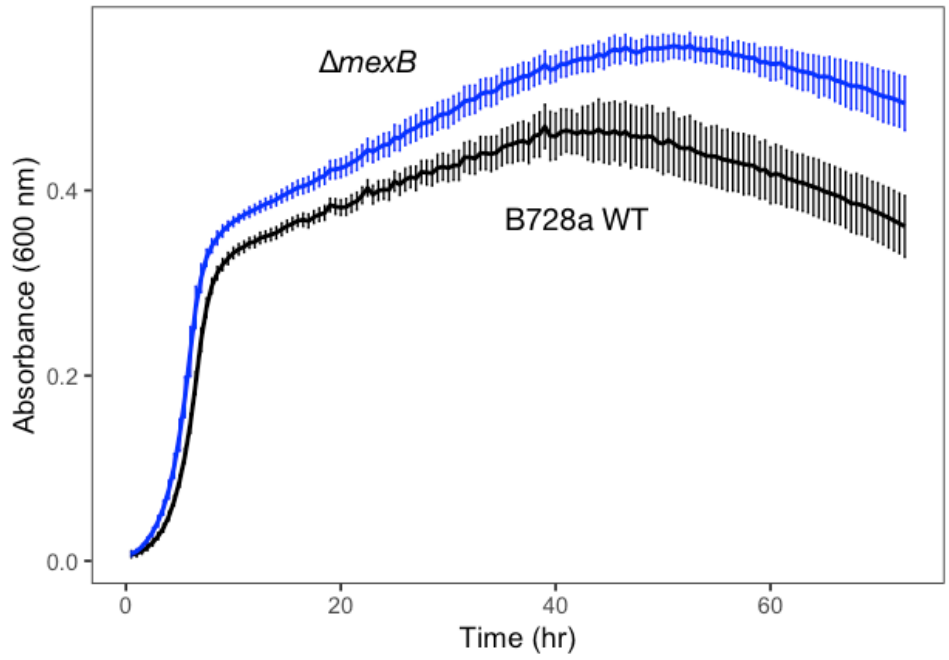


Figure 1. B728a Δ*mexB* (blue) grows as well as WT (black) in King's B (KB) broth. Average absorbance at 600 nm is shown here for 5 replicate samples each, with the average absorbance of two KB blanks subtracted from the total absorbance for each sample. Vertical lines indicate the standard deviation.

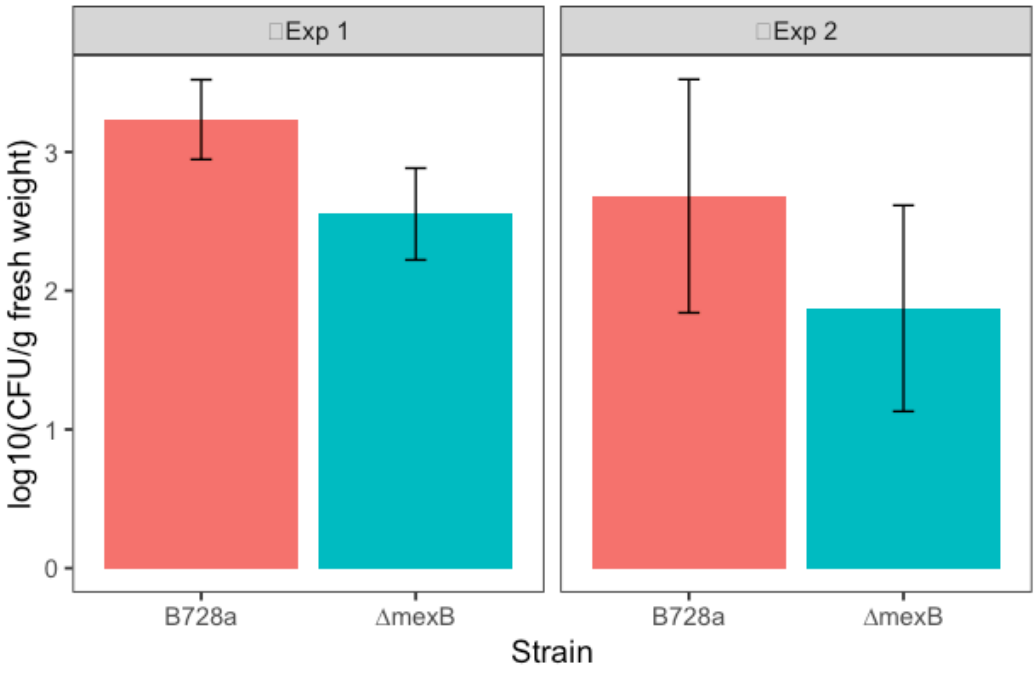
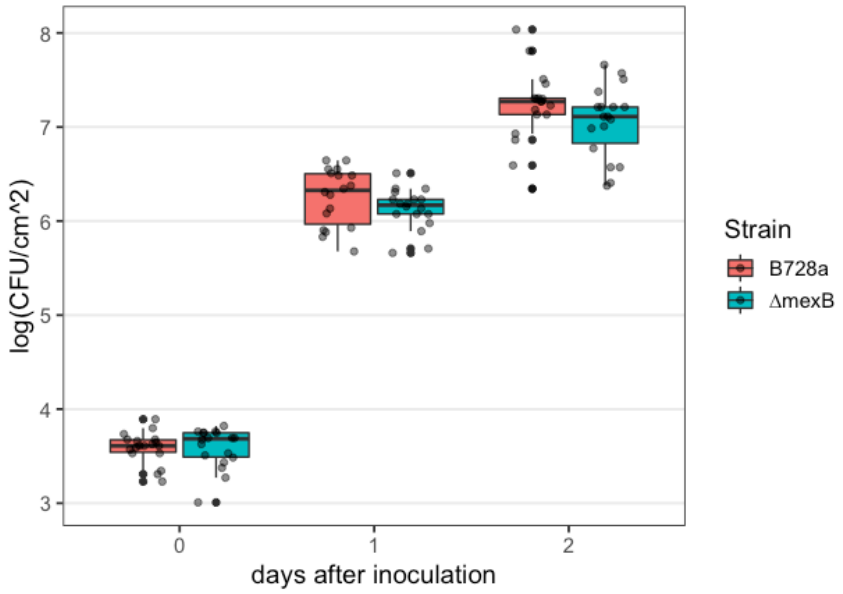


Figure 2. B728a Δ*mexB* establishes lower population sizes on the surface of common bean leaves than the WT strain. In each experiment, epiphytic populations at high humidity were significantly different 2 days following a spray inoculation (Welch Two Sample t-test, $p = 6*10^{-4}$ and $p = 0.03$). Mean CFU per g leaf tissue are shown, with error bars indicating the standard deviation. N = 8 (exp 1) and 10 (exp 2) per strain.

A.



B.

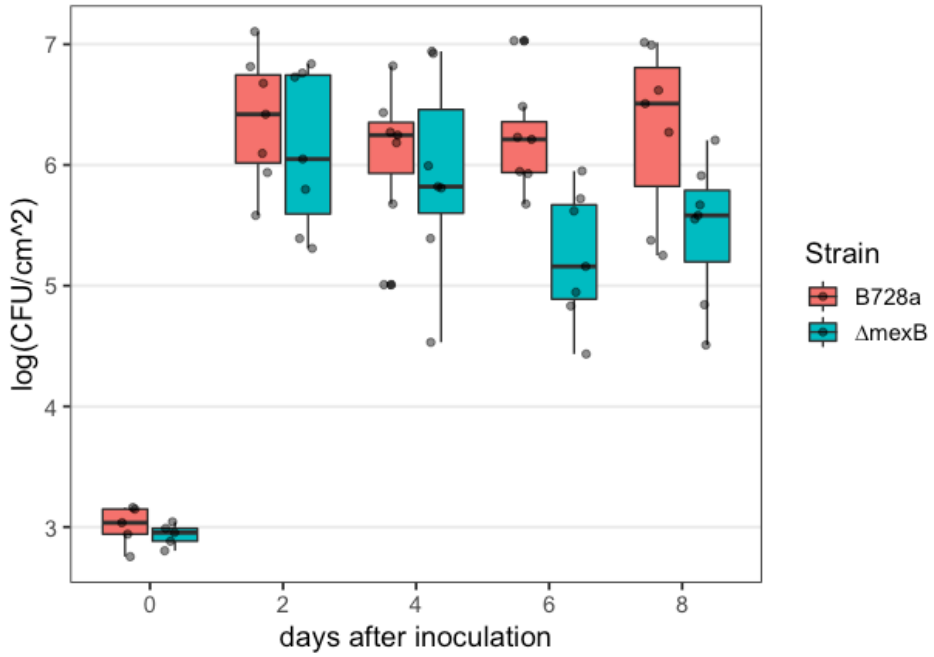


Figure 3. Apoplastic population sizes of *P. syringae* WT (red) and Δ*mexB* mutant (blue) at various times after inoculation when inoculated at 10^5 CFU/ml (A) or 10^4 CFU/ml (B). The MexAB-OprM efflux pump is not required during early stages of apoplast colonization (0 to 2 days post inoculation), and becomes more important later (~ 6 to 8 dpi). A) Bacterial population sizes at each time point (0, 1, and 2 dpi) are not significantly different between strains (Student's t-test). B) Bacterial population sizes at 6 and 8 dpi are significantly different (Student's t-test, $p < 0.05$).

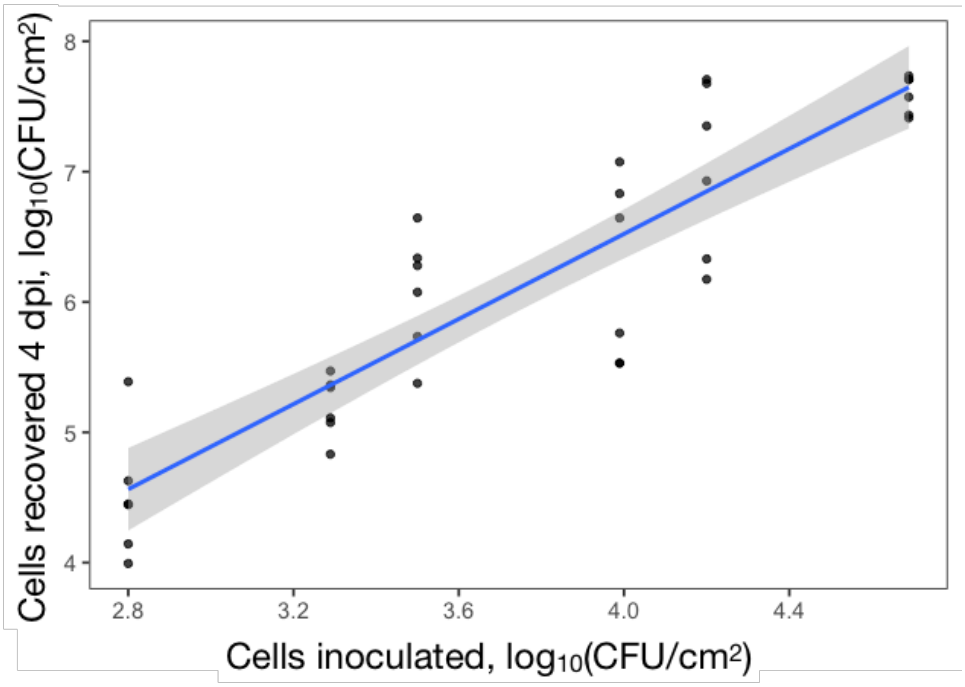


Figure 4. Population size of WT *P. syringae* strain B728a achieved after growth in common bean variety Bush blue Lake for 4 days after introduction into the interior of common bean (*P. vulgaris*) leaves at various concentrations is directly proportional with the concentration of inoculum introduced into the plant. A linear model was used to create the regression line. Pearson's correlation coefficient = 0.90, p = 1.2e-13.

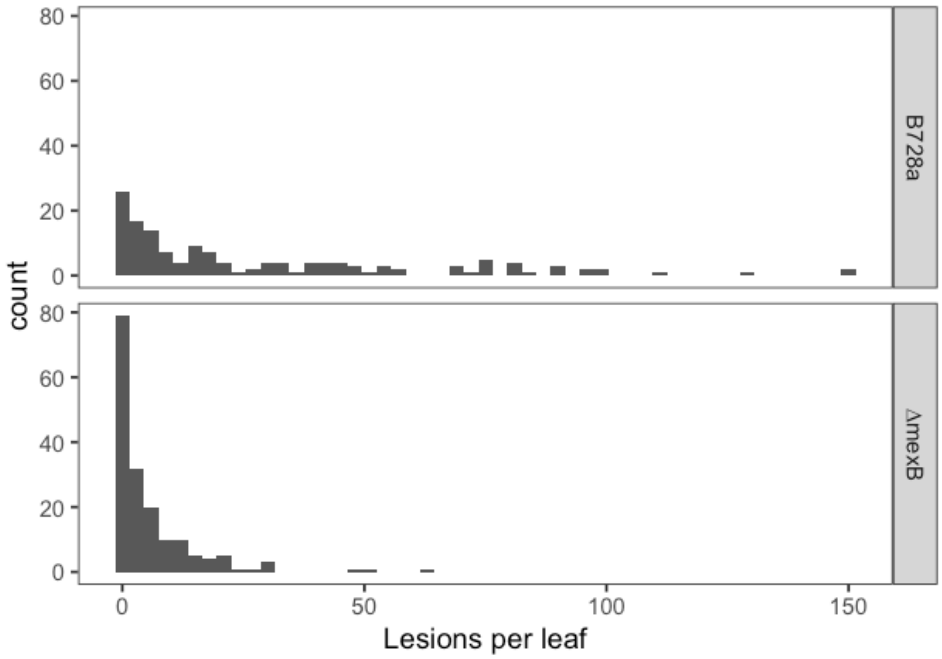


Figure 5. Frequency distribution of the number of lesions formed on individual leaves of common bean (*P. vulgaris*) variety Bush Blue Lake after topical application of *P. syringae* WT strain B728a (top panel) or a Δ*mexB* mutant (bottom panel). B728a Δ*mexB* forms fewer lesions than WT on common bean. B728a WT and Δ*mexB* lesion counts are significantly different (Wilcoxon rank sum test with continuity correction, $p = 7*10^{-15}$). N = 146 leaves (WT) and 173 leaves (Δ*mexB*), measured 10 days post inoculation.

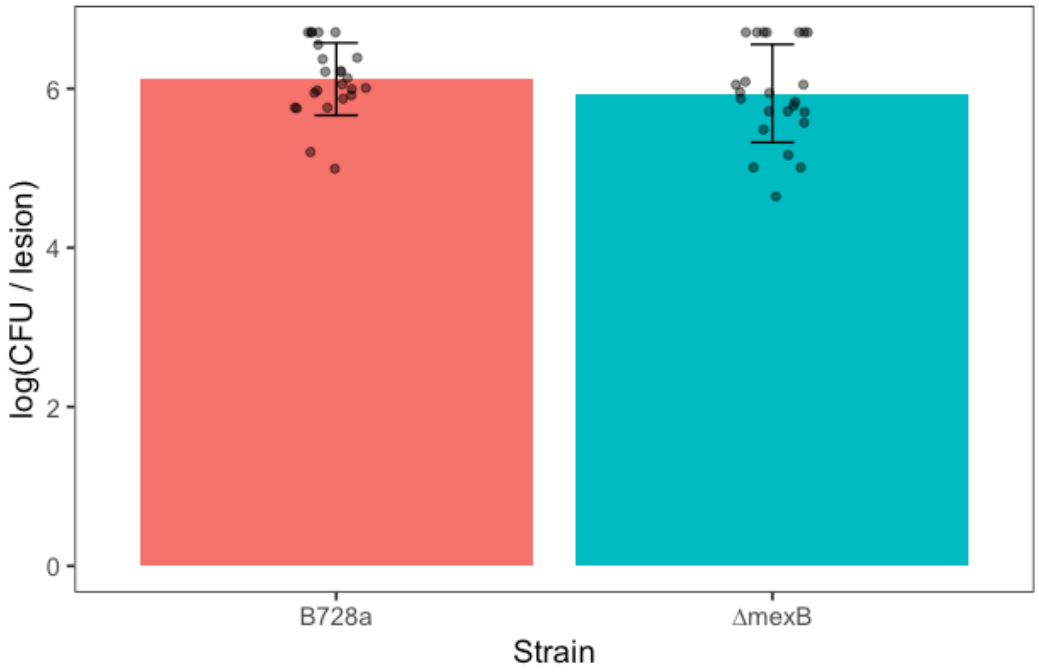


Figure 6. Population size of cells found within disease lesions formed after topical application of *P. syringae* wild type strain B728a (red) or a Δ*mexB* mutant (blue) that were excised from leaves of common bean variety Bush Blue Lake. Bacterial populations within newly formed lesions are not statistically different between B728a WT and Δ*mexB* (Welch Two Sample t-test, p = 0.26). Individual lesions in common bean leaves were sampled 7 days post inoculation to estimate bacterial population sizes. Black dots represent the population size in a given lesion.

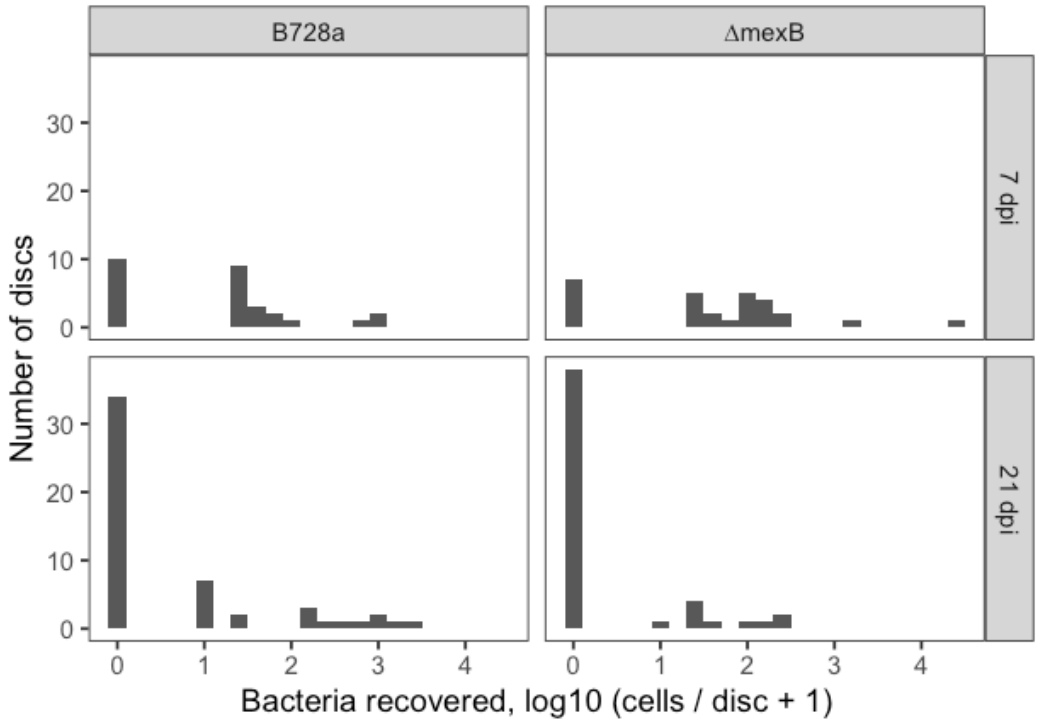


Figure 7. Frequency distribution of the number of bacterial cells recovered from small discs (19.6 mm^2) excised from asymptomatic regions of common bean (*P. vulgaris*) variety Bush Blue Lake leaves that also harbored disease lesions when sampled 7 days (top panel) or 21 days (bottom panel) after topical application of *P. syringae* wild type strain B728a (left panel) or a Δ*mexB* mutant (right panel). B728a WT and Δ*mexB* cells are present within asymptomatic areas of the common bean leaf, 7 and 21 days post inoculation (dpi). N = 28 (each WT and ΔmexB) 7 dpi, 53 (WT) and 48 (Δ*mexB*) 21 dpi.

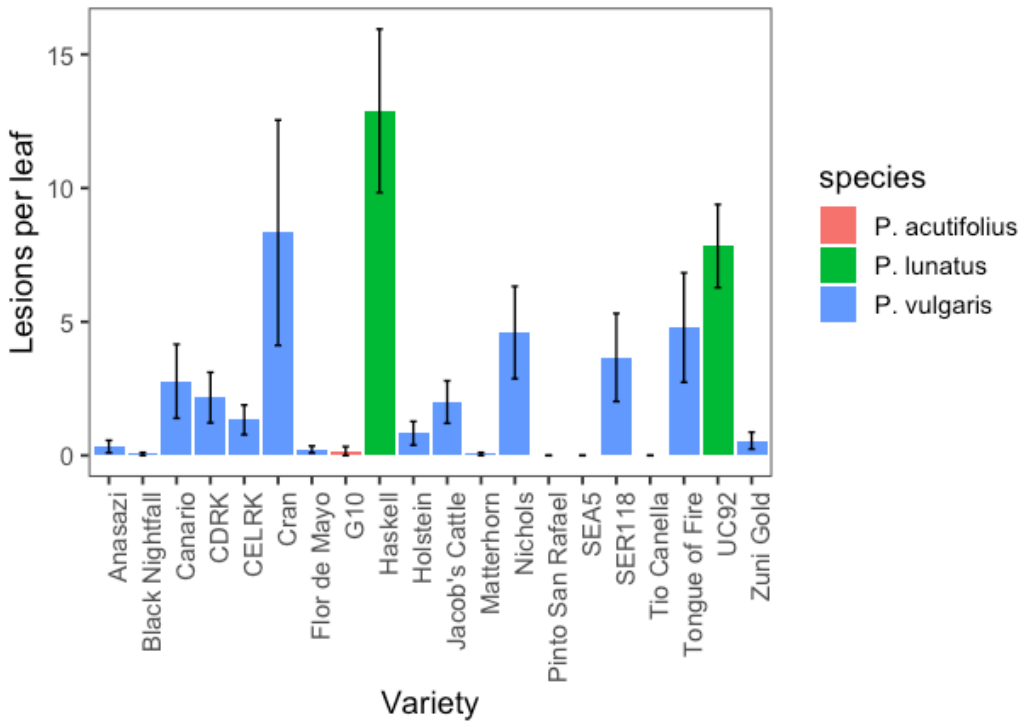
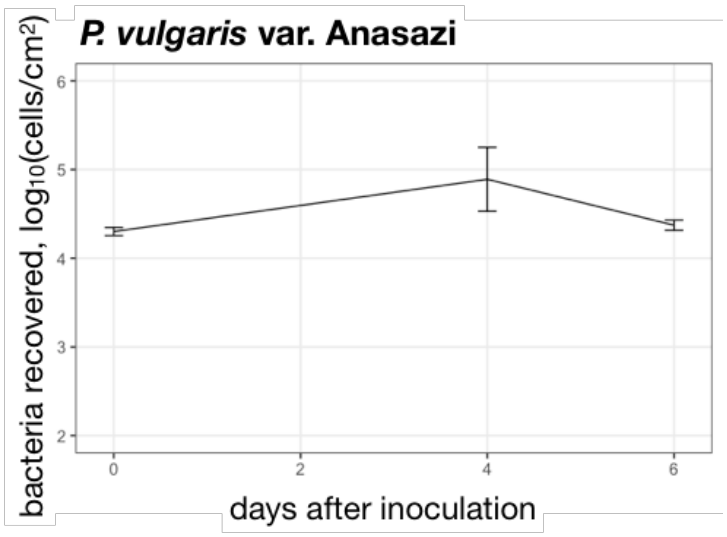


Figure 8. Number of lesions formed by *P. syringae* wild type strain B728a on leaves of varieties of three *Phaseolus* species, measured 10 days following a spray inoculation. *P. lunatus* is more susceptible, and *P. acutifolius* is less susceptible relative to other varieties tested here, to disease symptom formation. The number of leaves assessed ranged from 6 to 18 per variety. Error bars indicate standard error of the mean.

A.



B.

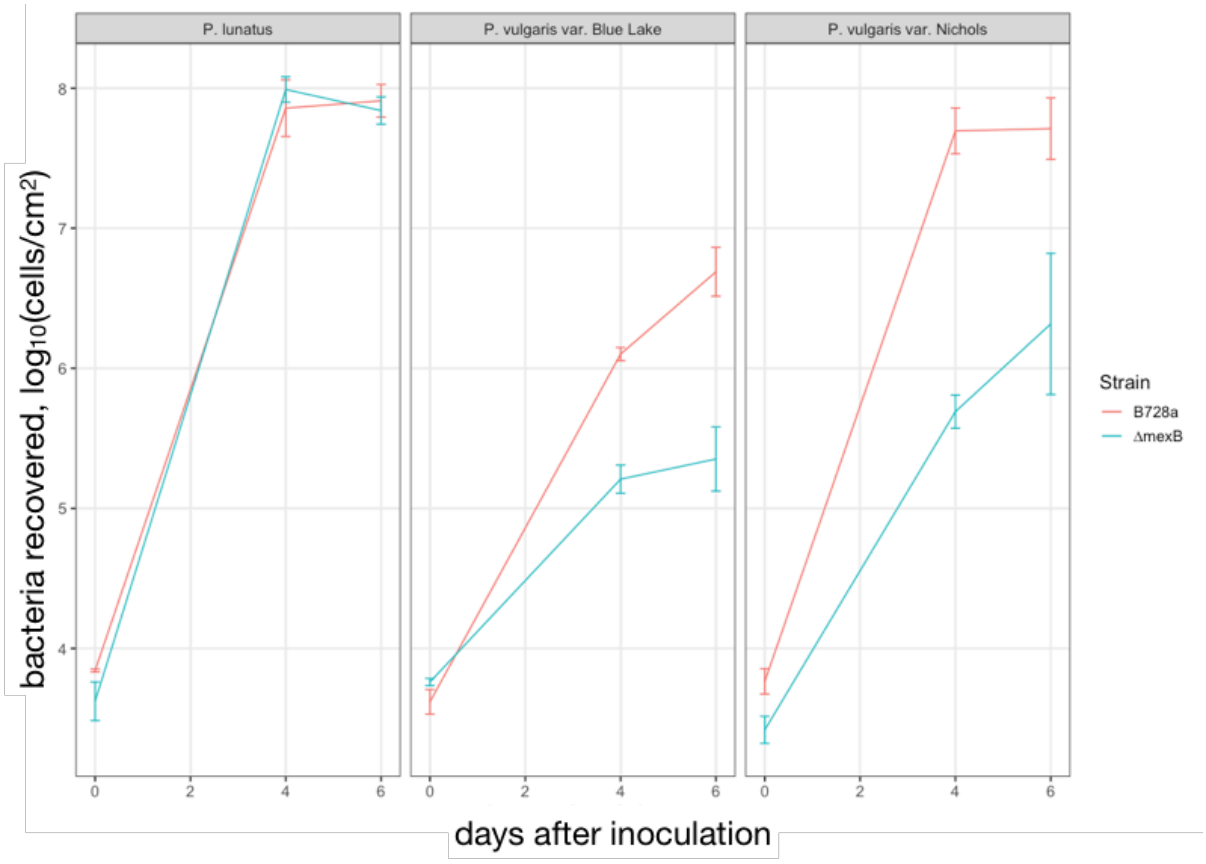


Figure 9. Apoplastic population size of *P. syringae* WT strain B728a (black) in *P. vulgaris* var. Anasazi (A), and B728a WT (red) and Δ*mexB* (blue) in *P. vulgaris* var. Blue Lake Bush, *P. vulgaris* var. Nichols, and *P. lunatus* var. Haskell (B). The vertical bars represent the standard error of the mean of log-transformed populations per cm^2.

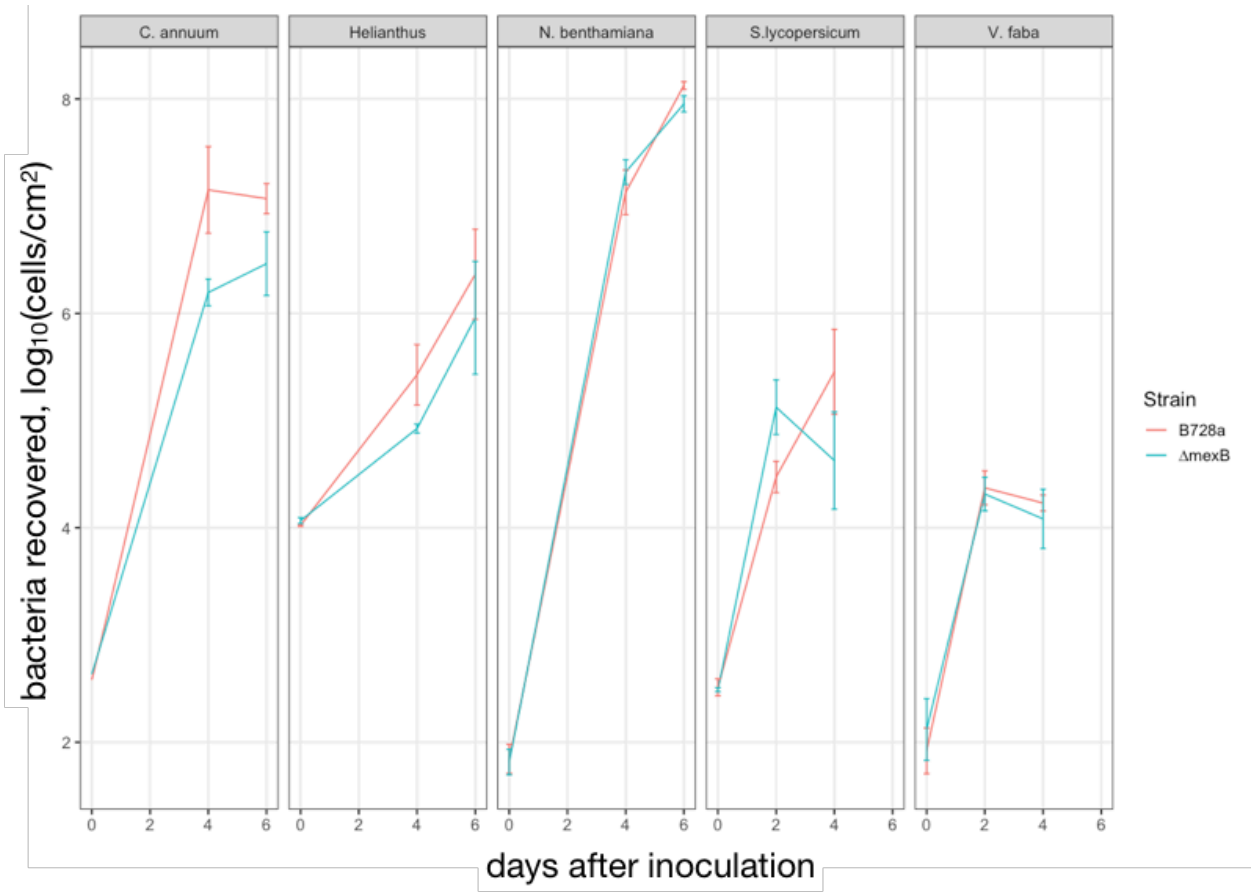


Figure 10. Apoplastic population size of *P. syringae* wild type strain B728a (red) and Δ*mexB* (blue) recovered from a variety of plant species at various times after inoculation. MexB is required for full apoplastic growth in pepper (*C. annuum*), but not fava bean (*V. faba*), *N. benthamiana*, sunflower (*Helianthus*), or tomato (*S. lycopersicum*). Total bacterial growth is relatively low (100- to 1,000-fold increase over 6 days) in sunflower and tomato. The vertical bars represent the standard error of the mean of log-transformed populations per cm^2.

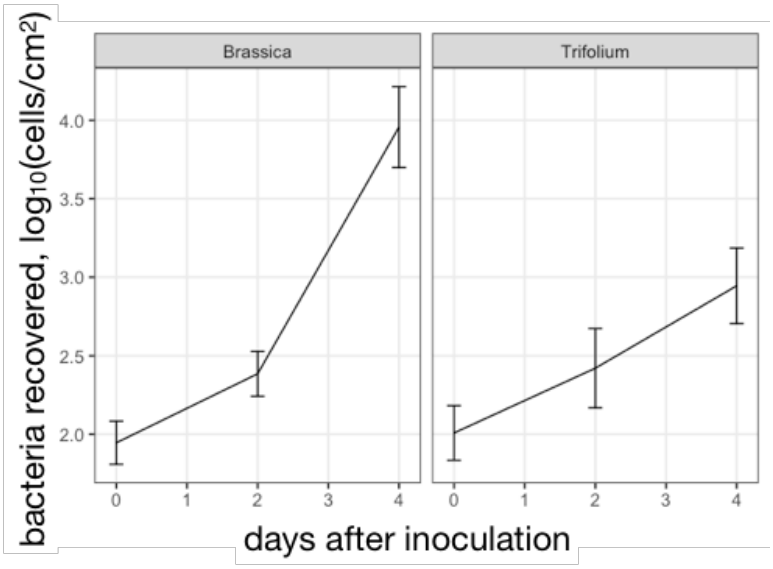


Figure 11. Apoplastic population size of *P. syringae* wild type strain B728a recovered from *Brassica nigra* (left panel) and *Trifolium pratense* at various times after inoculation. Strain B728a is marginally able to colonize the apoplast of mustard (*Brassica*) and clover (*Trifolium*). The vertical bars represent the standard error of the mean of log-transformed populations per cm^2. For each time point, N = 3 (*Brassica*) and 6 (*Trifolium*) replicate samples were assessed.

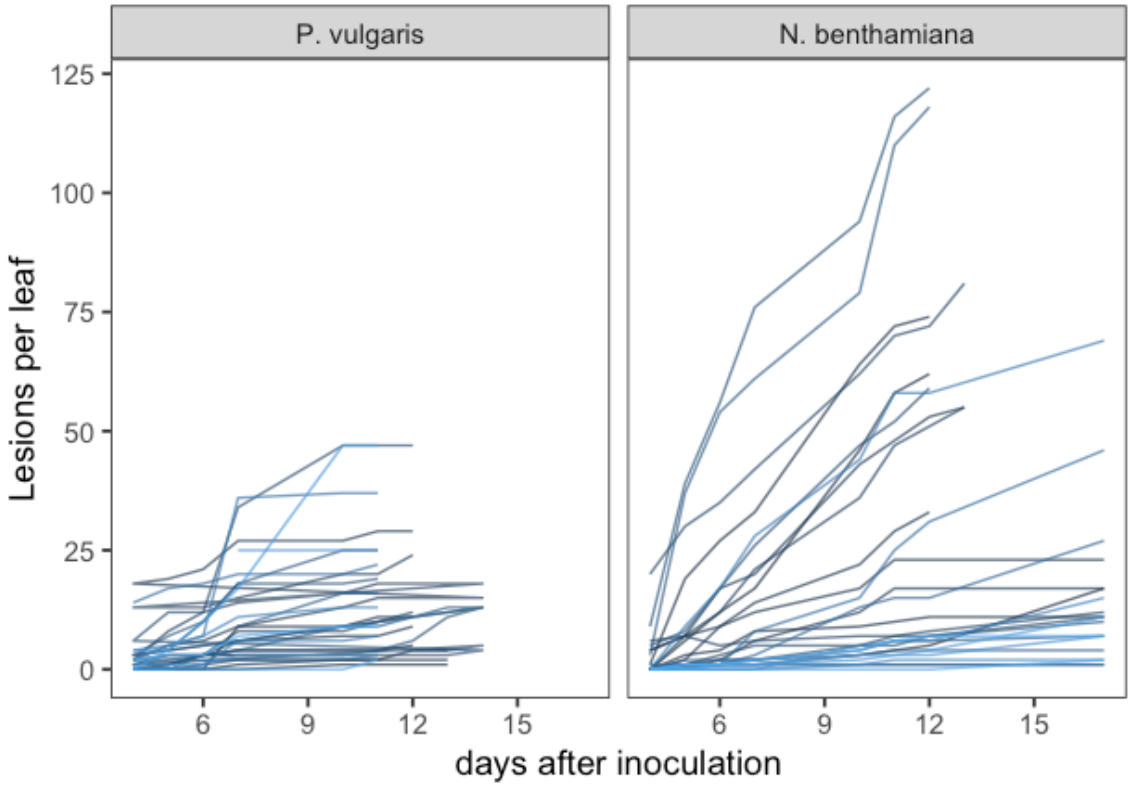


Figure 12. Number of lesions detected on individual leaves in common bean (*P. vulgaris*) and *N. benthamiana* when assessed at various times after inoculation. Each line represents the number of lesions on a given leaf assessed at the times after inoculation shown on the abscissa.

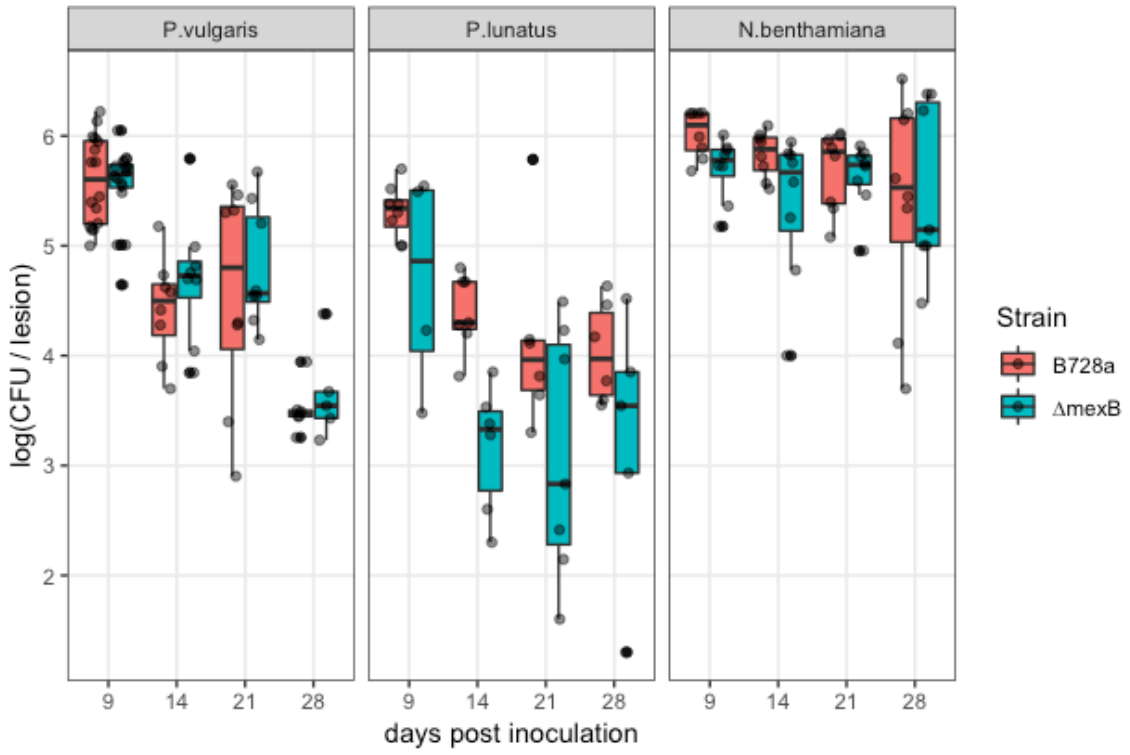


Figure 13. Population size of cells found within disease lesions formed after topical application of *P. syringae* wild type strain B728a (red) or a Δ*mexB* mutant (blue) that were excised from leaves of common bean variety Bush Blue Lake (*P. vulgaris*) (left panel), lima bean (*P. lunatus*) (center panel) and *N. benthamiana* (right panel) at the various times after inoculation shown on the abscissa. Bacterial populations within individual lesions decline over time in common bean and lima bean but not in *N. benthamiana*. Differences between populations in lesions of B728a WT and Δ*mexB* sampled at a given time were not statistically significant at any sample time, with the exception of lima bean measured at 14 days post inoculation (Welch Two Sample t-test, p = 0.002).

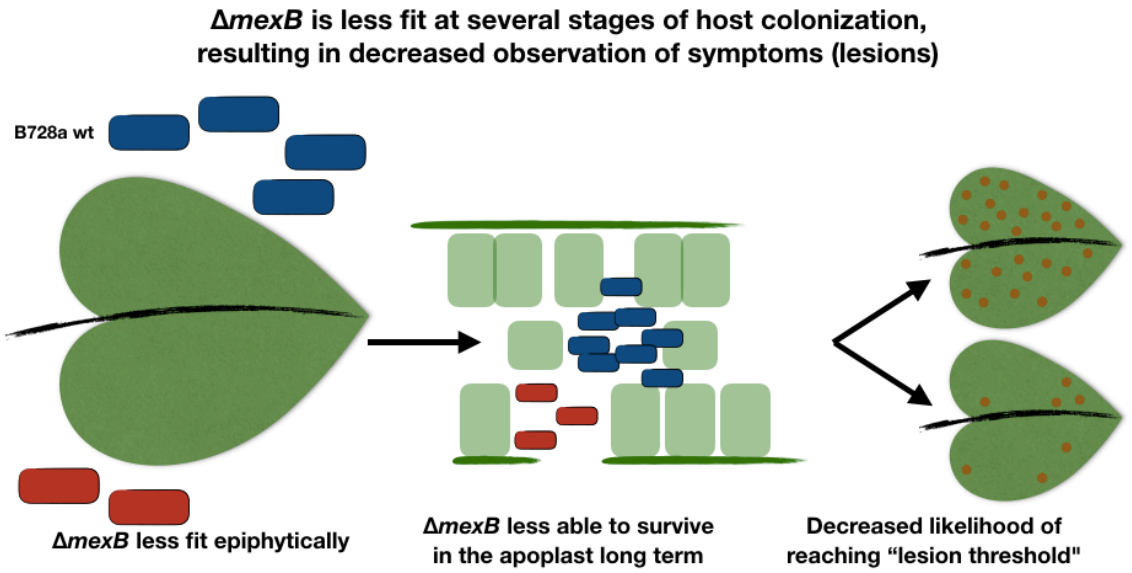


Figure 14. Proposed model of a population size threshold for lesion formation and the role of MexB in lesion formation. B728a Δ*mexB* is less fit both on the leaf surfaces well as in apoplastic growth, resulting in decreased incidence of lesion formation. A threshold population in the apoplast is required for visible symptoms to form on the leaf. B728a Δ*mexB* is less fit epiphytically, as well as in the apoplast (within the time frame required for lesions to form: approximately 1 week). These factors likely function synergistically, resulting in a decreased frequency of Δ*mexB* both entering the leaf and its subsequent growth in the apoplast to reach the required “lesion threshold”. These combined deficits decrease the likelihood of its epiphytic community to cause disease symptoms .

Tables

Strains and plasmids

Strains		
Organism	Description	Source
E. coli TOP10	For general cloning	Invitrogen
E. coli S17-1	Conjugation donor strain	(60)
P. syringae B728a	Wild type strain (Rif^R)	(112)
P. syringae B728a	Δ*mexB* (Rif^R Kan^R)	This work

Plasmids			
Plasmid name	Description	Antibiotic	Source
pKD13	Source of kanamycin resistance	Kan	(113)
pT*sacB*	Suicide plasmid to introduce DNA into *P. syringae*	Tet	(113)
pT:4008-kan	To delete *Psyr_4008*, contains *Psyr_4008* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work

Primers
Bold sequence complements FRT-Kan sequence for SOE protocol.

Name	Sequence
FRT-KanF	GTGTAGGCTGGAGCTGCTTC
FRT-KanR	ATTCCGGGGATCCGTCGACC
4008 up F	AGCACAGCCCTATACCCTGA
FRT 4008 up R	**GAAGCAGCTCCAGCCTACAC**TTACTCCCCTTTGCTGCCTG
FRT 4008 down F	**GGTCGACGGATCCCCGGAAT**TGCAGTTACCGCGTTCATTC
4008 down R	AGGAAGGTCAGGTTGCTGTC
4008 check F	CAGTGGCTGTAGCAAGAAGGA
4008 check R	TTTGTTGCGCACTGAACAGC

Table 1. *Phaseolus* varieties used in this study.

Species	Name	Source
P. vulgaris	Blue Lake Bush 274	*NA*
P. vulgaris	Anasazi	COB_HEIR_15
P. vulgaris	Black Nightfall	COB_HEIR_15
P. vulgaris	UCD 707 Canario	UCD_RL_2013-02 CR1, 124-A
P. vulgaris	California Early Light Red Kidney (CELRK)	CDBN_2013 174-3 CR1, 120-8
P. vulgaris	California Dark Red Kidney (CDRK)	2010-EXP MAIN53 CR1, 118-B
P. vulgaris	UCD 0801 Cran	CDT_SUBS_15 236-18
P. vulgaris	Flor de Mayo de Euglina	CDT_WXD_15 196-6
P. vulgaris	UCD Holstein	COB_HEIR_15
P. vulgaris	UCD Jacob's Cattle 0908	CDBN_2015 CR1, 116-8
P. vulgaris	Matterhorn	CDT_SUBS_15 204-2
P. vulgaris	UC Nichols	28(10)77 RIII E31 CR1, 125-8
P. vulgaris	Pinto San Rafael	CDT_SUBS_15 248-15
P. vulgaris	SEA 5	CBS_PINK_15 280-18
P. vulgaris	SER 118	CDT_SUBS_15 188-18
P. vulgaris	Tio Canela 75	CDT_WXD_15 180-4
P. vulgaris	Tongue of Fire	COB_HEIR_15
P. vulgaris	Zuni Gold	COB_HEIR_14
P. lunatus	UC Haskell	2010 CR1, 103-B
P. lunatus	UC 92	2010-EXP 19-E1 19(10)E/Bulk CR1, 101A
P. acutifolius	G40010	TDT_FSI_15 248-4

Chapter 4

Differential contribution of genes in *Pseudomonas syringae* to its growth in diverse host plant species

Abstract

Diverse traits are necessary for bacterial colonization of the interior of plant hosts, ranging from classic virulence traits to those more generally required for bacterial growth in both the apoplast and leaf surfaces. High-throughput methods such as transposon sequencing (TnSeq) are powerful tools used to identify genes required for host colonization, especially for bacterial pathogens. However, there is little information as to the distinct genetic requirements for bacteria to colonize different hosts, particularly among plant pathogens. Here, we utilize randomly barcoded TnSeq (RB-TnSeq) to identify the genes that contribute to the ability of *Pseudomonas syringae* strain B728a to grow within three different plant species spanning two families. Most genes contributed at similar magnitudes to apoplastic fitness in each of the plant hosts. However, 50 genes significantly differed in their fitness contributions to growth in various species. These genes encoded proteins in various functional categories including polysaccharide synthesis and transport, amino acid metabolism and transport, cofactor metabolism, and phytotoxin synthesis and transport. Six genes that encoded unannotated, hypothetical proteins also contributed differentially to growth in these hosts. Using this easily scalable method, this same heterogeneous mutant population can be used to interrogate gene contribution to success in a variety of distinct hosts or environments. This dataset represents a valuable tool for the identification of host-specific virulence genes.

Introduction

Pseudomonas syringae is a ubiquitous species complex containing strains commonly found in association with a variety of both healthy and diseased plants. More than 60 pathovars have been identified in this species, where taxonomic placement is primarily associated with their host range on different groups of host plants (58). The life cycle of *P. syringae* is typically considered to consist of two overlapping phases: an initial period of epiphytic growth on the surface of plants and subsequent invasion and then growth in the intracellular spaces in plants (apoplast) (17). Under favorable environmental conditions, such as high humidity or surface moisture and moderate temperature, epiphytic bacterial populations can reach high levels, apparently increasing the likelihood that at least some cells will invade the interior of the leaf. After invasion of the leaf through stomata or openings caused by wounds rapid bacterial multiplication can occur in plants that are susceptible to infection, resulting in the formation of visible leaf symptoms and eventually localized host cell death (17). Most attention to traits involved in plant colonization has focused on those contributing to apoplastic growth. Several such traits including the type III secretion system and certain effector proteins, phytotoxins, siderophores, adhesins, and genes contributing to stress tolerance have been identified (34). While many genes contribute to growth in varied environments, canonical virulence factors such as type III effectors and phytotoxins that

contribute to the suppression of the host immune system and are specifically important to growth in the apoplast (35).

In *P. syringae* and other bacterial pathogens, type III effectors are likely to be important determinants of host specificity since they are usually involved in suppression of plant defenses against bacteria that are recognized by the plant. The linkage between type III effector repertoires and host range however remains unclear (150). In addition, individual type III effector genes can be disrupted with little to no effect on virulence, commonly because some of the many effectors produced by a given bacterium have redundant functions (151). On the other hand, toxins produced by *P. syringae* show no host specificity and typically do not contribute to bacterial multiplication in plants, instead are required for inciting characteristic symptoms that follow infection (152). While some other traits required for colonization of a given host plant, such as those described above, have been investigated, in nearly all cases studies have been limited to a given pathosystem. There has been almost no work to identify to what extent such traits contribute to bacterial growth in a plant host-dependent manner.

While the identification of individual virulence factors of plant pathogenic bacteria by classical gene by gene disruption studies is a very laborious process, transposon sequencing (TnSeq), a process by which the behavior of a large mixture of disruption mutants can be studied in parallel, provides comprehensive information on the contribution of nonessential genes required for growth in and or on a given plant host. For example, about 100 genes were identified that significantly increased or decreased fitness in *Dickeya dadantii* in chicory when disrupted using this method (30). The importance of the biosynthetic pathways for leucine, cysteine, and lysine for bacterial growth within the plant were highlighted in this study. A similar study in *Pantoea stewartii* ssp. *stewartii* identified genes that were important for its growth within corn (*Zea mays*) xylem. Typical of most studies, these reports focused on a single pathosystem. It is unknown to what extent the genes that contributed to their virulence in most plant species investigated would also be required for virulence to other potential host plants or in environmental habitats. That is, given that there is an appreciation that many plant pathogenic bacteria such as *P. syringae* are relatively promiscuous, growing both on and in a variety of often unrelated plant species (38), such pathogens might have evolved and retained virulence genes that are selectively important only on a subset of the plants with which they interact. This important question in the ecology and epidemiology of such species will only be answered by determining the extent to which traits needed for successful colonization are distinctive to a given host plant. The global assessment of the contribution of genes of a strain to its success in a myriad of different habitats, however, until recently was untenable.

Random barcoded TnSeq (RB-TnSeq) is a modification of TnSeq in which 20 nucleotide barcodes are associated with distinct individual transposon insertion sites (28) enabling a single such mapped transposon library to be easily used in multiple experiments. This method greatly facilitates the use of a large transposon-mutagenized population to assess the fitness contribution of genes in a strain to large numbers of conditions, or in the case of the study, to various host plants. Because the location of the transposons in the mutant population only needs to be mapped a single time using TnSeq, the composition of the individual members of the mutagenized community can

be easily quantified in each of many subsequent experiments by sequencing of only the amplicon barcode regions (BarSeq) (28). This technique has been used to identify *P. simiae* genes that enable colonization of *Arabidopsis thaliana* roots (33). Combined with *in vitro* profiling that aids assignment of genes to various functions, this high-throughput transposon mutant study provided considerable insight into the various traits required for rhizosphere colonization.

We previously (Chapter 2) used RB-TnSeq to identify genes that contributed to competitive growth of *P. syringae* B728a on both the leaf surface and in the apoplast of common bean (*Phaseolus vulgaris*). This study revealed the importance of a variety of non-essential genes for growth on leaf surfaces including those encoding proteins involved in processes as varied as cofactor biosynthesis, amino acid biosynthesis, nucleotide metabolism, and polysaccharide synthesis. Genes involved in phytotoxin biosynthesis, the type III secretion system, and alginate biosynthesis were important in growth in the apoplast, but apparently not on leaf surfaces. In addition to common bean, strain B728a is capable of both multiplication and elicitation of disease symptoms in additional plant hosts including lima bean (*P. lunatus*) and pepper (*Capsicum annuum*) (38). Since apoplastic colonization requires bacterial growth while in intimate contact with host mesophyll cells, we hypothesized that the genes required to colonize these phylogenetically distinct plant species would differ. Because of the distinct metabolites expected within these plants and differential chemical or physical responses to bacterial infection, we hypothesized that a non-overlapping set of genes in *P. syringae* would contribute to its success in these various hosts. Furthermore, we hypothesized that there would be a greater overlap among those genes that contribute most to growth in common bean and lima bean than to those in pepper due to the close phylogenetic relationship of these two *Phaseolus* species in the Fabaceae family compared to that of *Capsicum*, a member of the Solanaceae family, Here, we examine non-essential genes in strain B728a for their differential contribution to apoplastic growth in three plant species.

Results

P. syringae B728a growth is supported in diverse plant hosts

As has been discussed in more detail previously (Chapter 2), the use of TnSeq requires sufficient growth of the mutagenized population to sensitively measure the relative changes in abundance of individual mutant strains within the total population. We determined that in addition to common bean, lima bean and pepper would support substantial growth of B728a (Fig. 1). While population sizes increased approximately 1,000 fold during growth in common bean, populations increased 10,000 fold or more in both lima bean and pepper within six days after inoculation. As such, on average each inoculated cell had undergone 10 generations of growth in common bean and 13 or more generations of growth in both lima bean and pepper. While lima bean appeared to be relatively more susceptible to infection than common bean, enabling more growth of bacterial cells within their tissues as well as with visible symptoms forming more quickly than in common bean, the same starting concentration of bacterial inoculum could successfully be used in all three species. A concentration of 10^5 cells/ml was found to be sufficiently high to avoid bottleneck effects when the transposon library was

inoculated by infiltration of intercellular spaces and yet low enough that substantial growth of the inoculated bacterial cells could occur.

Inoculation of a barcoded transposon library in multiple plant hosts

The randomly barcoded transposon library used in other studies (initially described in Chapter 2) was inoculated by vacuum infiltration into the leaves of the three species to identify genes contributing to host colonization. This mutant library contains 281,417 total strains with insertions that map to the B728a genome. A subset of these 169,826 genic strains in which transposon insertions were in the middle 10 – 90% portion of a given gene, were used for analysis. We analyzed fitness contributions for 4,296 of 5,137 (84%) protein-coding genes that harbored sufficient numbers of transposon insertions to accurately assess changes in proportional representation of a given mutant in the population of mutants. The mutant library was cultured in the rich medium King's B (KB) from a frozen common stock immediately prior to plant inoculations. For this reason, any changes in the proportional representation of transposon mutants in a given gene during overnight growth in this condition was used as the control against which growth of the mutants *in planta* was compared. Fitness values were calculated and the genes that were essential for growth in this condition were performed as described in Chapter 2. All experiments discussed here passed quality control metrics that were previously established in *in vitro* studies in (28). The requirements for a successful experiment included the detection of more than 50 median reads per gene and a consistency in the calculated fitness value estimated from mutants with insertions in the 5' and 3' distal half of a given gene (28).

Samples of the transposon library recovered from each plant host contained at least 68% the total unique barcodes that were inoculated into leaves at time0 (Table 1). Unique barcoded mutants were identified as those that both mapped to the genome and for which 3 or more reads were obtained during sequencing of the amplified barcodes in a given experiment. Such a result indicated that a sufficiently large proportion of the transposon mutant population was successfully introduced into plants during inoculation into the apoplast. When comparing these experiments, replicates of gene fitness values for a given plant host cluster together, and separately from the samples in KB (Fig. 2). It is noteworthy that the patterns of gene fitness values obtained in pepper are more disparate from those seen in both lima bean and common bean. Despite the differential overall growth in these hosts, fitness values ranged from approximately -5.5 to +1 in all experiments (Fig. 3). For each gene, fitness values for the 2 replicate growth experiments in KB and the 3 replicate apoplastic growth experiments performed in each plant host were averaged. While there was variation in the calculated standard deviation of fitness values of a given gene in a given host, with that of a few genes being relatively high, the mean standard deviation did not differ between the three plant species tested here (Fig. 4). It was noteworthy that the standard deviation of genes having both very high and low fitness values were somewhat higher than those of genes that contributed little to fitness in a given host (Fig. 5), suggesting that the fitness contribution of the traits encoded by these genes was perhaps somewhat context–dependent and thus influenced by variation between particular experiments.

Most plant colonization genes contribute to growth in all hosts tested

Genes that contributed specifically to apoplastic colonization in common bean are detailed in Chapter 2. To determine the extent to which these genes were also required for maximum fitness in both lima bean and pepper, as well as to determine genes that were more important in pepper and lima bean than common bean, we compared the gene fitness values after inoculating the mutant library into the apoplast of common bean, lima bean, and pepper. Fitness contributions of 87 genes were significantly different in at least one plant host from that in the other plant hosts (Kruskal-Wallis, $p < 0.05$). Fifty of those genes had average fitness values of less than -0.5 or greater than 0.5 in at least one host (Table 2). These genes that had a differential impact on fitness in different hosts included 8 that were annotated as being involved in polysaccharide synthesis and regulation, 6 in cofactor metabolism, 5 in phytotoxin synthesis and transport, and 4 in amino acid metabolism and transport. We also identified 6 genes that encoded hypothetical proteins that were differentially important in these hosts. The identity of the other 37 genes that did not contribute substantially to fitness on one or more host are described in Table 3.

Most commonly, genes that contributed to fitness in a given host contributed similarly to the fitness of *P. syringae* in the other two plant species tested. Pairwise comparisons of gene fitness values between plant hosts demonstrated the high correlation between the fitness contribution of genes in the apoplast of common bean, lima bean, and pepper (Fig. 6). Pearson correlation coefficients, comparing average fitness values from all genes, ranged from 0.725 (lima bean - pepper) to 0.796 (common – lima bean). With the exceptions noted above, it is clear that most genes involved in colonization of the host apoplast were broadly important regardless of the plant species. At average fitness value thresholds of both < -2 and < -1, approximately 1/3 of all genes that contributed to competitive fitness did so in all three host plants tested (Fig. 7). Genes with average fitness values below these two fitness thresholds are listed in Table 4.

Variations in amino acid availability between plant species revealed through assessment of the fitness of auxotrophs

Genes involved in the biosynthesis of several amino acids were generally required for competitive fitness in the apoplast of all three plant species tested. However, auxotrophs in histidine and isoleucine/valine biosynthesis were less fit in common bean and pepper than in lima bean (Fig. 8) suggesting that these amino acids were more abundant in lima bean than in the other two species. Within these pathways, average fitness values for *hisD*, *hisH*, and *ilvHI* differed significantly in at least one plant species (Kruskal-Wallis, $p < 0.05$).

Phytotoxin production appears to be a host-specific virulence trait

We previously showed that the production of phytotoxins such as syringomycin contributes to competitive fitness in the apoplast of common bean (Chapter 2). It is noteworthy that such phytotoxin biosynthesis genes contributed less to fitness in lima bean than in both common bean and pepper (Fig. 9) suggesting that this host was perhaps less susceptible to the damage caused by phytotoxins such as syringomycin. The average fitness values of genes encoding the regulatory proteins SalA, SylA, and SyrP, as well as encoding the secretion protein PseB and the amino acid adenylation

protein SypA differed significantly among these three plant species (Kruskal-Wallis $p < 0.05$).

Alginate production varies in importance among host plants

Polysaccharide biosynthesis and regulation is a broadly important plant colonization trait. While the genes involved in the biosynthesis of alginate contributed to fitness in all three plant species examined, these genes were more important in lima bean than in pepper (Fig. 10). Specifically, the alginate biosynthesis genes *alg44*, *alg8*, *algA-1*, *algF*, *algI*, and *algK* contributed differentially to fitness on lima bean (Kruskal-Wallis, $p < 0.05$).

Discussion

RB-TnSeq is a powerful method that has been used in a variety of bacterial species to identify those genes that contribute to fitness in diverse *in vitro* conditions (29). However, until now there has been little application of this novel technology to identify bacterial fitness traits in bacteria that live in complex symbiotic relationships with other organisms, such as plant-associated bacteria. The use of populations of isogenic bacteria harboring individual transposons to disrupt gene function is a powerful strategy wherein even relatively small changes in the relative proportion of mutants in that complex mutagenized population can be assessed. Studies of insertional mutant populations have usually focused on identifying genes that are important in a single context, such as a particular plant host or environmental condition. This somewhat limited use of this powerful technology is largely due to the cost and effort required to map genomic insertions in the mutagenized population at the start and end of each such condition to which the mutagenized population is exposed. In contrast, after initially associating a variety of uniquely tagged transposons with a given gene in RB-TnSeq the relative proportions of the different insertional mutants in large mutant populations can be readily assessed by quantifying the amplified tags. This enables the same mutant library to be used repeatedly to interrogate gene contributions in many different settings. In this study, we exploited such a well-characterized mutant population to interrogate the role of genes in *P. syringae* in growth in several different host plants, in many replicate experiments. This method provided confident estimates of the differential contribution of each of a myriad of genes to fitness in these different hosts. This proved to be a very powerful tool to generate hypotheses regarding host-specific virulence traits. In this study we identified many common and host-specific genes that enable growth in the intercellular spaces of plants for further analysis.

We had hypothesized that at least a subset of the genes in *P. syringae* required for colonization of the apoplast of a given host plant would not be required for colonization of other plants, with the assumption that distinctive features of the niche provided by the interior of these plants would differ in resources and chemical defenses. We had previously observed that a larger number of genes were required for colonization of the apoplast of common bean compared to the leaf surface (Chapter 2) suggesting that the leaf interior provides a chemically or physically more complex habitat than the leaf surface. Through the fitness profiling of amino acid auxotrophs, the previous study demonstrated variation in the importance of various amino acid biosynthetic pathways for growth both on the leaf surface and in the apoplast of

common bean. This suggested that some amino acids are less abundant than others in both of these environments, supporting direct measurements made by others (36). Based on fitness profiling of transposon mutants in the apoplast of additional plants, we find that the relative abundance of available amino acids seems to vary between plant species. Specifically, the apoplast of lima bean appears to contain more accessible histidine, isoleucine, and valine than in both common bean and pepper. On the other hand, similar relative amounts of other amino acids are apparently present in the three plant species tested since the genes in other amino acid biosynthetic pathways did not contribute to fitness differently between the host plants examined here.

While they are not essential for growth *in planta*, *P. syringae* produces various phytotoxins that contribute substantially to symptom development in plants (42, 97). Strain B728a produces the lipodesinonapeptides syringomycin and syringopeptin, as well as the peptide derivative syringolin A (106, 153). Syringolin A has been shown to counteract stomatal immunity through proteasome inhibition (154), and genes in this biosynthetic pathway did not contribute to apoplastic growth in any of the three plant studied here. Syringomycin and syringopeptin function as pore-forming cytotoxins, contributing to ion leakage (100). It was interesting to note that the biosynthetic genes for these non-specific toxins contributed more to fitness in common bean and pepper than in lima bean. Curiously, while strain B728a grew to higher population sizes in pepper than in common bean (Fig. 1), symptom development is much less apparent on pepper than in either common or lima bean, consisting only of leaf yellowing without the appearance of any necrosis. While these three plant species were all susceptible to disease caused by strain B728a, lima bean appeared much more susceptible than common bean because of the substantially higher apoplastic populations that were attained as well as an earlier onset and more extensive necrotic lesion formation that it supported. Thus while it appears that syringomycin is a general virulence trait required for maximum growth within plants, and not merely a factor contributing to necrotic lesion development, the mechanism by which it contributes to growth in plants remains unclear. It is possible that differences in the immune responses between these plant species that might be modulated by syringomycin may contribute to the differential contribution of this trait to aplastic growth.

Given that most of the genes that contribute most to virulence in any of the three host plant studied here contributed similarly to virulence and the other two hosts as well, it remains unclear why *P. syringae* strain B728a multiplied to substantially different population sizes in the three plant species. Low water availability on both the leaf surface and in the apoplast limits bacterial growth (57). In response, *P. syringae* apparently upregulates the production of the exopolysaccharide alginate, to help maintain available moisture for bacterial growth (58). Alginate production is an important factor for both virulence and epiphytic fitness of *P. syringae* (58, 91). While alginate biosynthetic mutants were less fit relative to WT in the apoplast of all three hosts tested in this study, alginate appears to be a more important virulence trait in lima bean relative to that in pepper. In *P. syringae* pv. *tomato* DC3000, alginate biosynthetic genes are expressed in both host and non-host plants (155). While alginate production is a generally important virulence trait, differences in the relative importance of alginate biosynthesis may indicate variation in the relative moisture content between the apoplastic space of lima bean and pepper. It is possible if the anatomical differences

between these plants could contribute to differences in water availability. For example, pepper leaves are somewhat thicker than that of either of the two bean species, and thus water loss during the infection process may be somewhat diminished in pepper, thus enabling leaves to better retain water that may be released by the plant to the apoplast during the infection process (20).

Using RB-TnSeq, we have determined those genes that are most important for growth inside various plant species. Given that strain B728a is capable of apoplastic growth in several additional plant species beyond those tested here (described in Chapter 3), it should prove informative to interrogate the contribution of both the common and unique virulence genes identified here in a more disparate collection of plant species that presumably provide yet more distinctive plant habitats that this strain can colonize. Further work on the plant species examined here should also enable us to identify additional genes that play a quantitatively less important role in the colonization of these plants. While the experimental design used here was somewhat ambitious given the effort needed to inoculate and recover bacteria from the large numbers of plants needed to perform RB-TnSeq, the three experimental replicates performed for each host somewhat limited our ability to identify genes that did not contribute strongly to fitness. The fitness values observed for many genes reflected 50% or less growth of disruption mutants compared to wild type strains. Such a large effect of these genes on growth over the relatively few (10-13 generations) in a given experiment suggests that there is strong selection for these genes in plants, especially when considering the many generations that the species would undergo in a given year. Because of the limited number replications used in the study we did not use a multiple testing correction for the Kruskal-Wallis test because the resulting p-value (1e-05) was deemed to be too conservative and no genes would have been identified as significant contributors to fitness among these hosts at such a corrected p-value. This risks the identification of a number of false positive claims of genes annotated as “host-specific”. However, it is important to note that all of the genes that contributed differentially to colonization of these three plant species (Table 2) were found to be strongly required for fitness in at least one host plant species, having an average fitness value less than -0.5 in at least one host. There is somewhat less support for the 37 other genes for which differential contribution fitness was indicated (Table 3) and not analyzed further were statistically significant ($p < 0.05$), but were less convincingly biologically relevant. Low magnitude fitness values are more likely to be generated by chance during barcode amplification and sequencing, especially when observed across all three hosts. Our less conservative approach to nominating fitness genes here provides a list of genes for which hypotheses as to what traits are most likely host-specific can be further tested.

P. syringae is a commonly observed phyllosphere resident that is also found in the global water cycle (65). These habitats presumably require very different genes to enable it to exploit potentially very different habitats. For example, those genes required for growth on leaf surfaces are probably very different than those required for growth and survival in agricultural irrigation waters, and likely very different than those required for apoplastic growth. Bacterial biosensors have been used to characterize conditions experienced by bacteria such as nutrient availability or salt stress, and have revealed that these factors can often differ substantially in different settings (56, 156). There is therefore much we still need to learn about habitat specific contributions of the many

genes in a bacterium such as *P. syringae.* Transposon mutagenesis and subsequent characterization of relative mutant fitness provides genome-wide data on many traits simultaneously. This high throughput analysis of mixtures of insertional mutants not only indicates important traits for a given environment, but also can be extrapolated to suggest limiting factors in those habitats. The scalability of a method such as RB-TnSeq now provides the potential to examine the function of the many genes in the genome of species such as *P. syringae*, and the myriad of different settings in which it is often found. As such, further exploitation of the methods developed here brings functional genomics to such a species even in diverse natural habitats. While we considered here the contribution of genes to fitness in several of the many plants on which *P. syringae* might find itself in, there are many other settings in which a subset of its genes may become particularly important in fitness. Clearly, *P. syringae* has maintained genes preferentially important in some hosts, and many other genes may play similar habitat-specific roles.

Materials and Methods

Bacterial strains and growth media.

P. syringae pv. *syringae* B728a was originally isolated from a bean leaf (*Phaseolus vulgaris*) in Wisconsin (60). The complete genome for B728a is available on NCBI GenBank as accession CP000075.1 (106). B728a was grown on King's B (KB) agar or in KB broth (107), at 28˚C. When appropriate, the following antibiotics were used at the indicated concentrations: 100 µg/ml rifampicin, 100 µg/ml kanamycin, and 21.6 µg/ml natamycin (an anti-fungal).

Plant growth conditions.

Common bean (*P. vulgaris* var. Blue Lake Bush 274) and lima bean (*P. lunatus* var. Haskell) (5 - 7 plants per 10 cm diameter pot) were grown in Super Soil (Scotts Miracle-Gro) in a greenhouse for two weeks before inoculation. Pepper (*C. annuum*) seeds were sown on Sunshine Mix #4 (SunGro Horticulture) to a height of about 10 cm, and then transplanted into 10 cm diameter pots (3 to 5 plants per pot) containing Super Soil and grown in a greenhouse to 4 to 5 weeks old. Leaves were kept dry to minimize epiphytic bacterial populations. 1000 W metal halide lights were used to provide supplemental lighting for a 16-hour day length.

Bacterial apoplastic growth measurements.

Wild type strain B728a was grown overnight on KB with rifampicin, washed in 10 mM KPO_4 (pH 7.0), and adjusted to a cell concentration of $2x10^5$ CFU/ml in 1 mM KPO_4. Cells were inoculated into leaves using a blunt syringe. Leaf samples (3 discs per leaf) were excised using a 5 mm-diameter cork borer and placed into microfuge tubes containing 200 µl 10 mM KPO_4 and two 3 mm glass beads, and macerated by shaking for 30 seconds at 2400 rpm in a Mini-Beadbeater-96 (Biospec Products) before dilution plating of appropriate serial dilutions on KB containing rifampicin and natamycin.

Library recovery and growth in KB.

For each inoculation, a 1.25 ml aliquot of a glycerol stock containing the transposon library that had been stored at -80˚C was placed in 25 ml fresh KB with

kanamycin and grown for approximately 7 hours at 28˚C with shaking until the culture reached mid-log phase (OD_{600} 0.5 - 0.7). Time0 samples enumerating the relative abundance of each mutant were collected at this point during recovery; 1 ml aliquots were pelleted by centrifugation and the pellets frozen until DNA purification. The remaining cells were then washed twice in 10 mM KPO_4 prior to plant inoculation.

To determine the growth of the mutant library in KB, a 50 µl log phase cell culture (OD_{600} 0.5) was inoculated into 950 µl KB with kanamycin in a 24-well plate. The plate was incubated at 28˚C with shaking for 15 hours. Cells were collected by centrifugation, and frozen prior to DNA purification. Calculation of the fitness values from the KB control treatment is discussed in Chapter 2.

Inoculations of the transposon library into plants.

Cells were re-suspended to a concentration of $2x10^5$ CFU/ml in 1 mM KPO_4. The soil of potted plants was covered with cotton to hold the soil in place, and the pots were inverted in open containers containing ca. 1.5 L of bacterial cell suspension in an open glass bell jar. A vacuum was then applied to the plants sealed within the bell jar for 1.25 minutes and then removed rapidly to force the inoculum into the evacuated apoplast. Approximately 100 pots of plants of a given species were inoculated for a given replicate experiment. Plants were allowed to dry overnight and then moved to the greenhouse for six days.

Bacterial isolation from the apoplast.

Leaves of each plant species were excised from the plants and then chopped in a blender to yield fragments with an average diameter of about 1 to 4 mm. The slurry of plant fragments was then placed in a water-filled glass dish and placed in a sonication water bath (Branson 5510, output frequency 40 kHz) for 15 min to remove cells. The resulting slurry containing cells that were dislodged from the leaf interior was filtered through a coffee filter to minimize plant debris. 10% of the ca. 5-10 L of buffer containing the largely plant-free cell suspension was then subjected to sequential additional filtration steps of Whatman filters (20 µm, 10 µm, and 6 µm) to remove much of the remaining plant debris. Bacterial cells were then removed from suspension by centrifugation at 4696 x g for 10 minutes. The pellet was resuspended in water, and aliquots of cell pellets were frozen prior to DNA purification.

DNA isolation and library preparation.

DNA from frozen pellets was isolated using the Qiagen DNeasy Blood & Tissue Kit according to manufacturer's instructions. Cell lysis was done at 50˚C for 10 minutes as per optional instructions. For those apoplastic samples having excess residual plant material, lysed cells were centrifuged at 1,500 x g for 5 minutes before loading the supernatant onto purification columns. Purified genomic DNA was measured on a nanodrop device and 200 ng of total DNA was used as a template for DNA barcode amplification and adapter ligation as established previously (28). For each time0 and plant experimental sample, two separately purified DNA samples were sequenced as technical replicates.

Sequencing and fitness value estimation.

Barcode sequencing, mapping, and analysis to calculate the relative abundance of barcodes was done using the RB-TnSeq methodology and computation pipeline developed by Wetmore *et al.* (28); code available at bitbucket.org/berkeleylab/feba/. TnSeq was used to map the insertion sites and associate the DNA barcodes to these insertions, as described in Chapter 2. For each experiment, fitness values for each gene were calculated as the $\log_2$ of the ratio of relative barcode abundance following library growth in a given condition divided by relative abundance in the time0 sample. Fitness values were normalized across the genome so the typical gene had a fitness value of 0. Fitness values from sequencing replicates were averaged for each experiment. All experiments passed previously described quality control metrics (28), with the exception of one technical replicate of *P.vulgaris*_2, which was removed from analysis. Experimental fitness values are publically available at fit.genomics.lbl.gov.

Genomic fitness data analysis.

All analysis of gene fitness values and gene metadata was done in R (108). A PCA plot of experiments was generated from the matrix of gene fitness values using the function prcomp. To better classify genes based on their genomic annotation, we assigned gene names, gene product descriptions, and broad functional categories based on the previously annotated genomic metadata (49). For each gene, fitness values for experimental replicates were averaged to calculate an average gene fitness value for each plant species. A Kruskal-Wallis test was used to identify those genes in which the median fitness in at least one plant host was significantly different than that in other plant species ($p < 0.05$). Fitness values in KB were not included in this analysis. From these results, genes were removed to a separate list if the absolute average fitness was less than 0.5 in all conditions. Graphs were plotted in R using the ggplot2 package, version 3.1.1 (109). Heatmaps were plotted in R using the gplots package, version 3.0.1.1 (110).

Acknowledgements

We thank Morgan Price for assistance with RB-TnSeq sequence analysis. Plants were grown and maintained with the assistance of the UC Berkeley greenhouse staff, led by Tina Wistrom. Funding for Tyler Helmann was partially provided by the Arnon Graduate Fellowship and the William Carroll Smith Fellowship. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

Figures

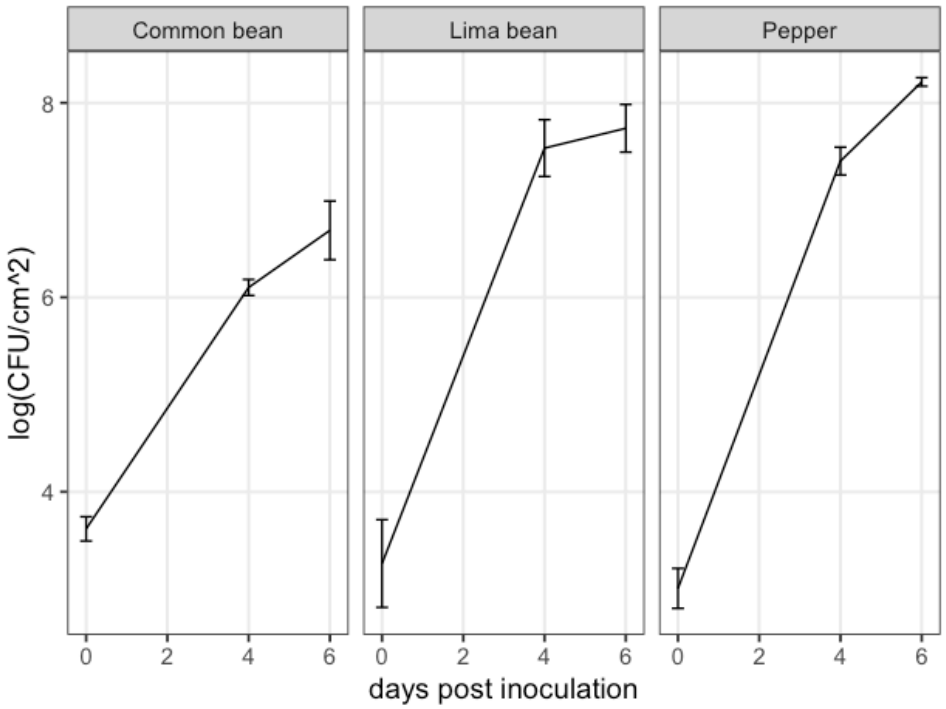


Figure 1. Growth of *P. syringae* B728a in the apoplast of the susceptible hosts common bean (*Phaseolus vulgaris*), lima bean (*P. lunatus*), and pepper (*Capsicum annuum*). Error bars represent the standard deviation.

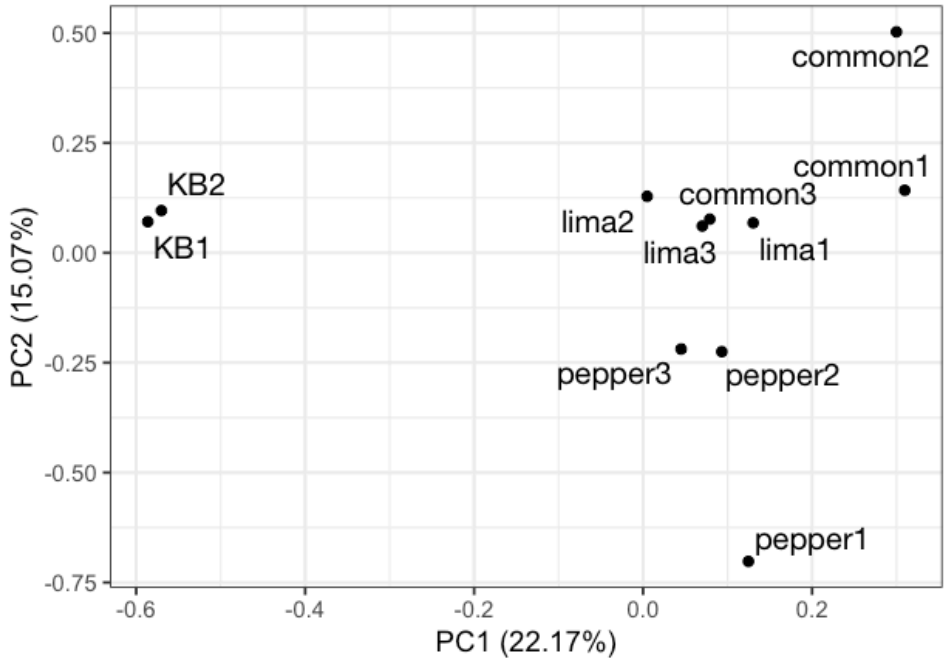


Figure 2. Principle component analysis of *P. syringae* B728a grown in King's B (KB), common bean ("common"), lima bean ("lima"), or pepper, based on gene fitness values estimated for 4,296 genes. Replicate experiments are noted as the numbers 1 to 3.

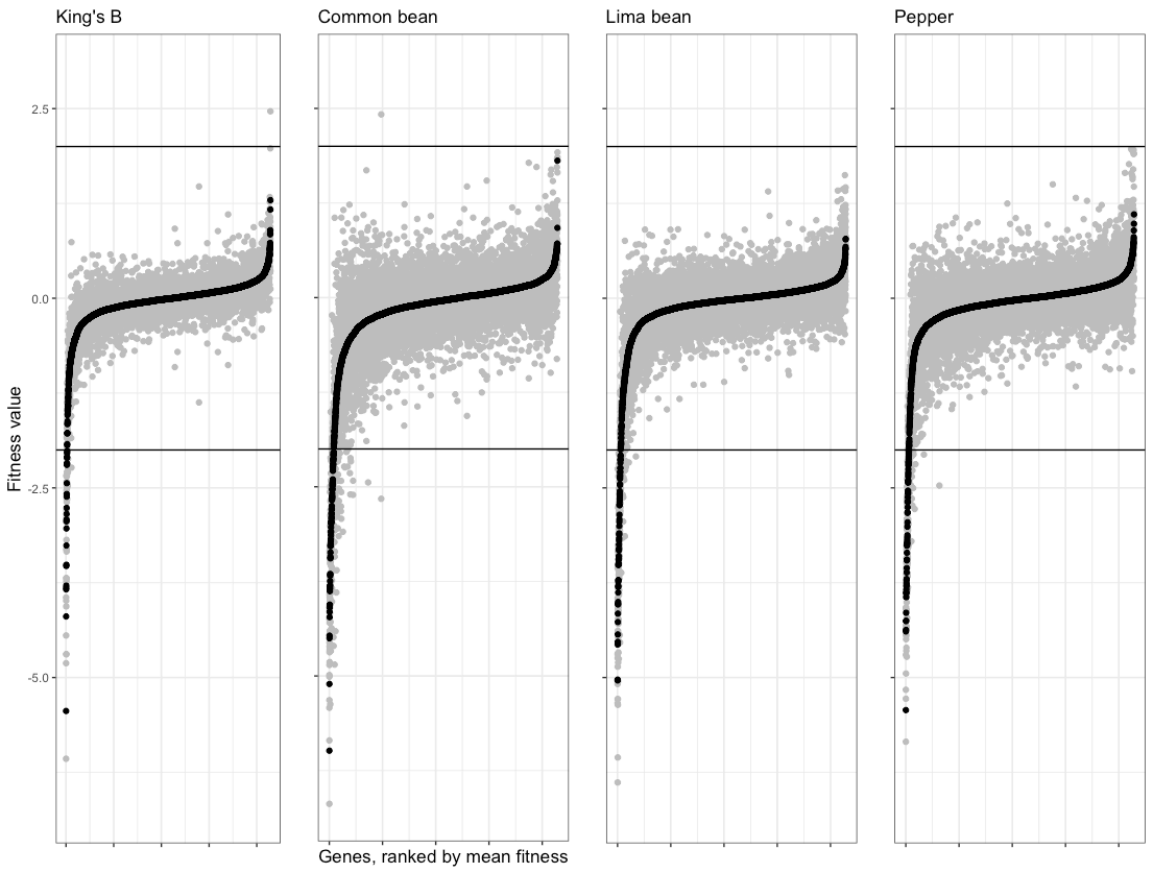


Figure 3. Rank ordered mean fitness values for each condition and plant species in which *P. syringae* was grown. Fitness values for independent replicate experiments are shown in grey, while mean fitness values are shown in black. Gene fitness value is calculated as the $\log_2$ of the ratio of the barcode counts following growth in a given condition compared to the barcode counts before inoculation. Black lines at fitness values of -2 and 2 are used to indicate strong phenotypes; for example a value of -2 indicates that mutants are 25% as fit as the typical strain in the mutant library. In each dataset, fitness values < -2 or > 2 are more than 3 standard deviations from the mean (approximately 0).

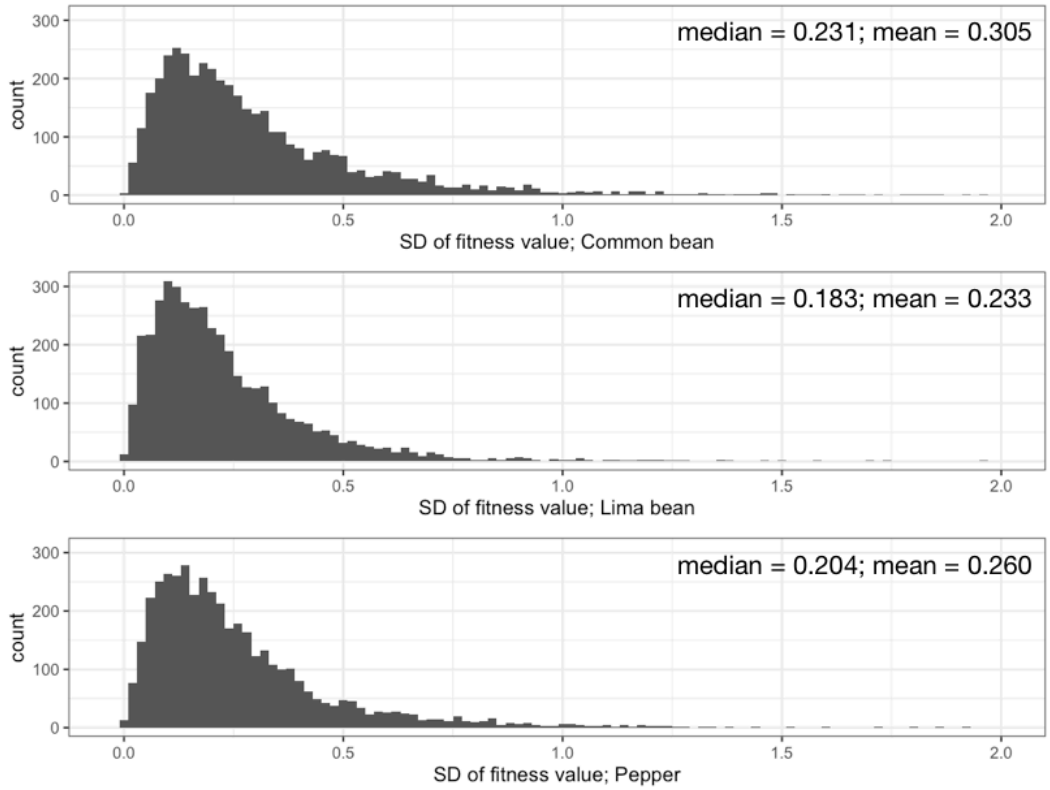


Figure 4. Distribution of standard deviation of gene fitness values for each gene calculated from 3 replicate *in planta* experiments. These distributions are similar for the 3 plant species.

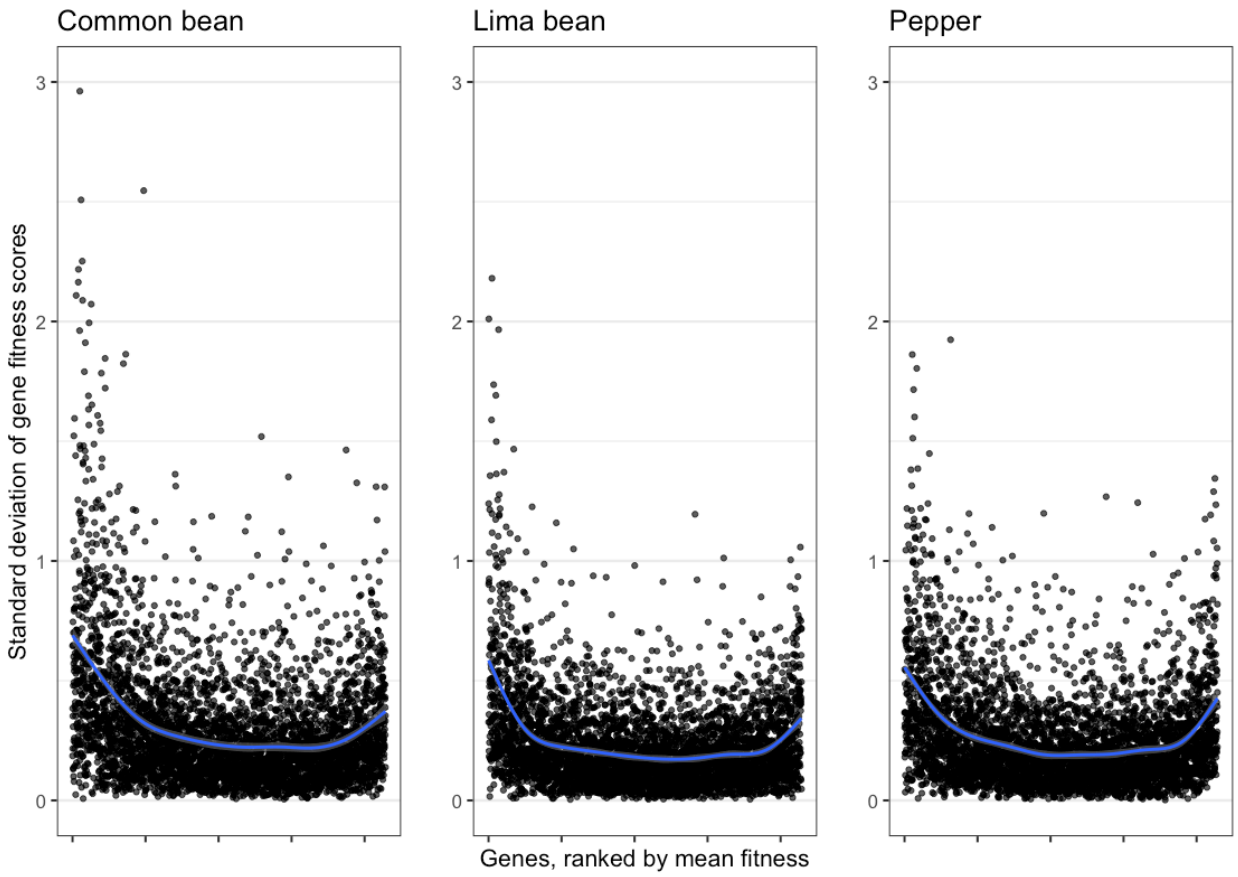


Figure 5. The standard deviation for estimates of gene fitness values is highest at both high and low measures of gene fitness contribution. For each host plant, the standard deviations of gene fitness values are ranked by average gene fitness, and are highest on average for genes with very low or high fitness value. A generalized additive model (GAM) was used to fit the regression lines.

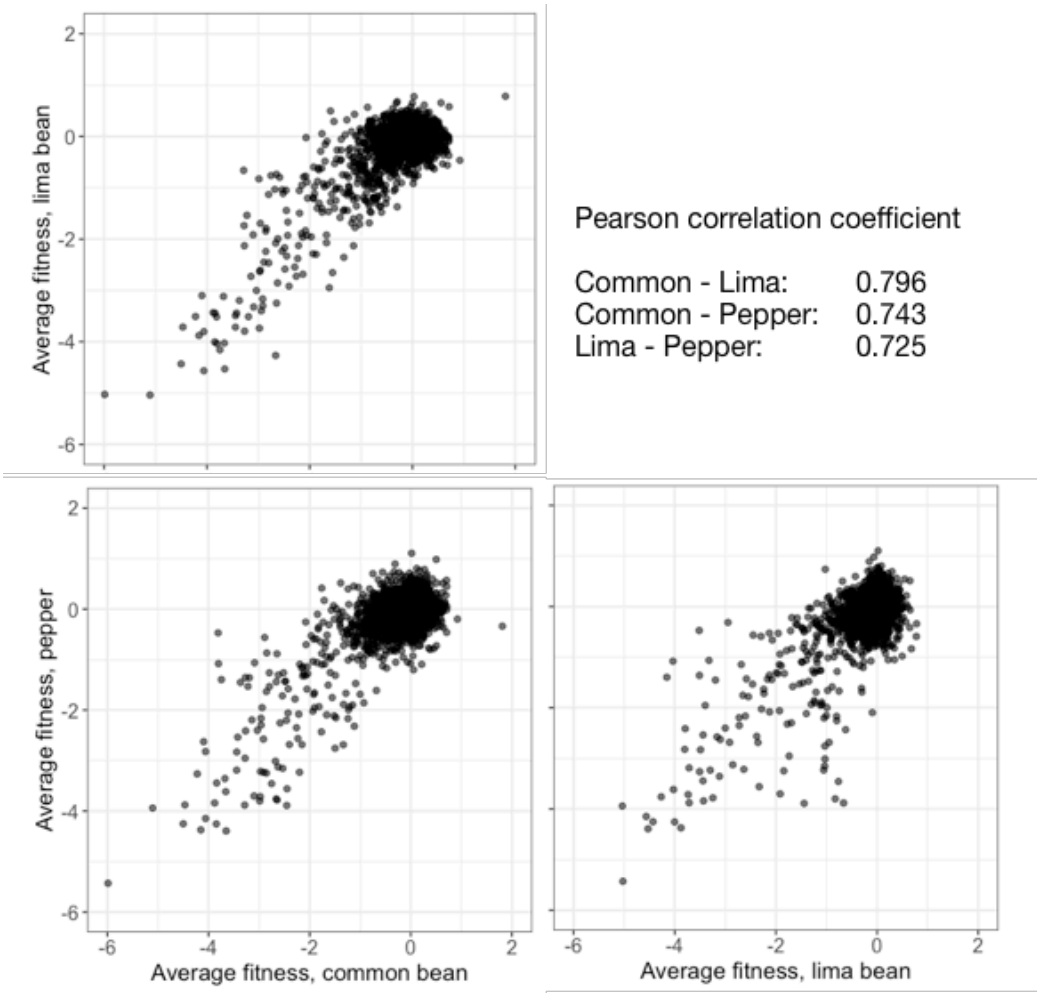
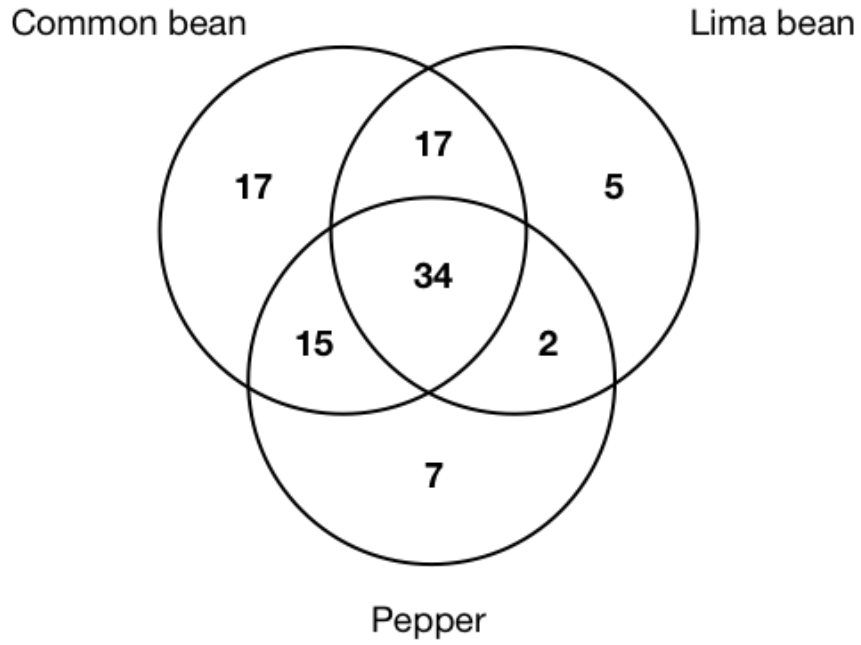


Figure 6. Pairwise comparisons of gene fitness values between plant hosts demonstrate the high correlation between gene fitness contributions to growth in the apoplast of common bean, lima bean, and pepper.

A.



B.

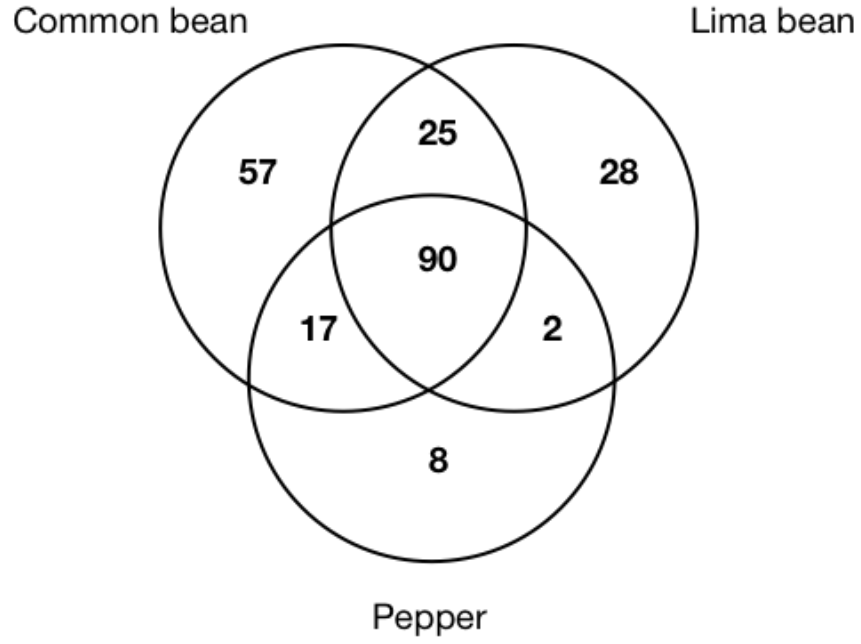


Figure 7. Genes with significant contributions to competitive fitness in each host. Venn diagram of the number of genes with average fitness values < -2 (A) or the number of genes with average fitness values < -1 (B).

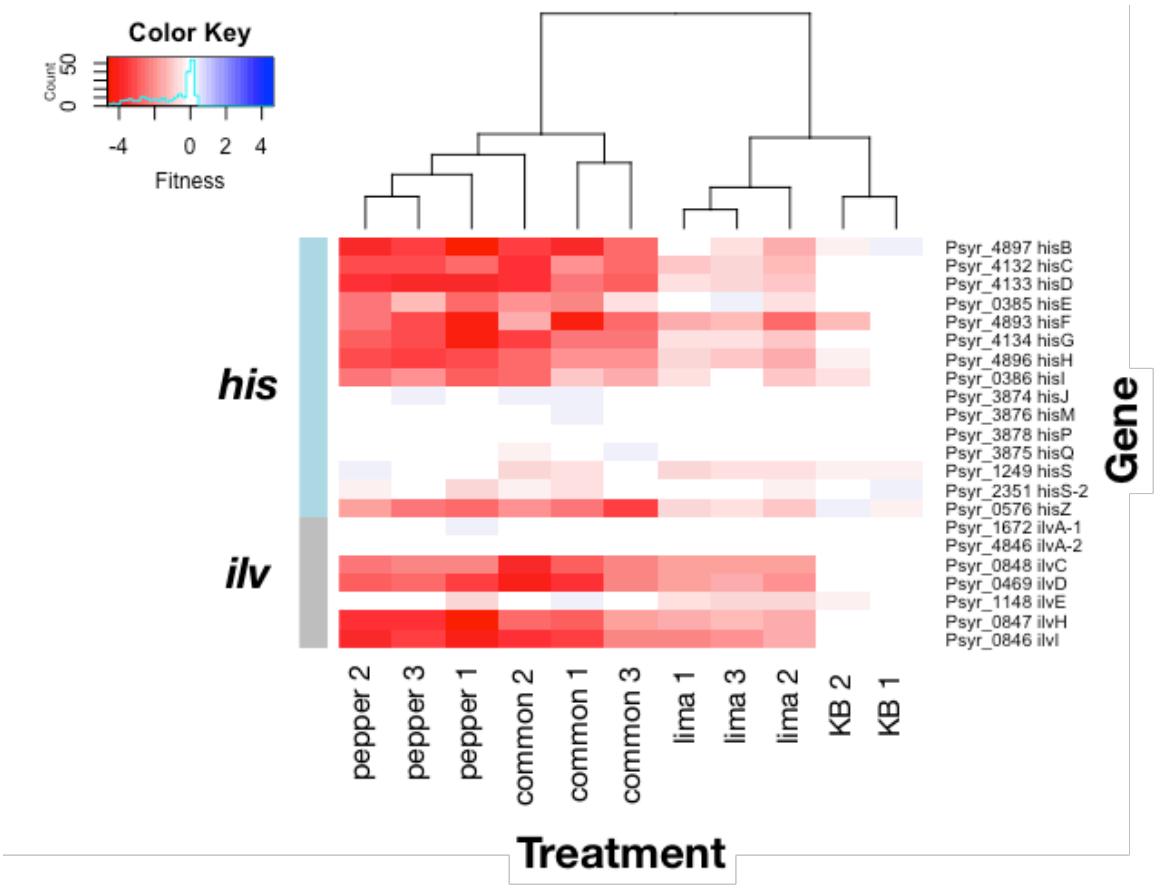


Figure 8. Heatmap of apoplastic fitness contribution of histidine and isoleucine/valine biosynthetic genes to growth in various plant species. Genes *hisD*, *hisH*, and *ilvHI* had significantly different fitness values in different plant species by a Kruskal-Wallis test ($p < 0.05$).

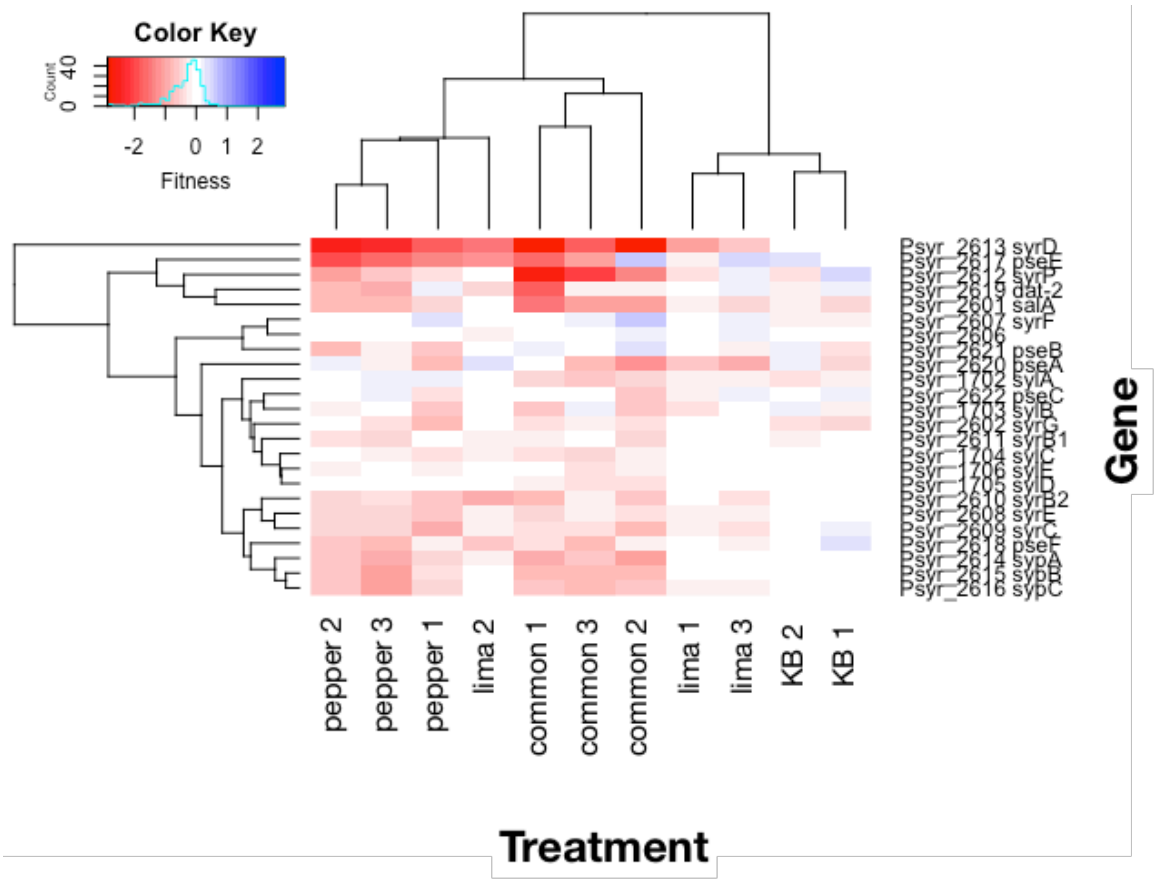


Figure 9. Heatmap of apoplastic fitness values of genes involved in phytotoxin biosynthesis. Genes *sylA*, *salA*, *syrP*, *sypA*, and *pseB* had significantly different fitness values in these 3 plant species by a Kruskal-Wallis test ($p < 0.05$).

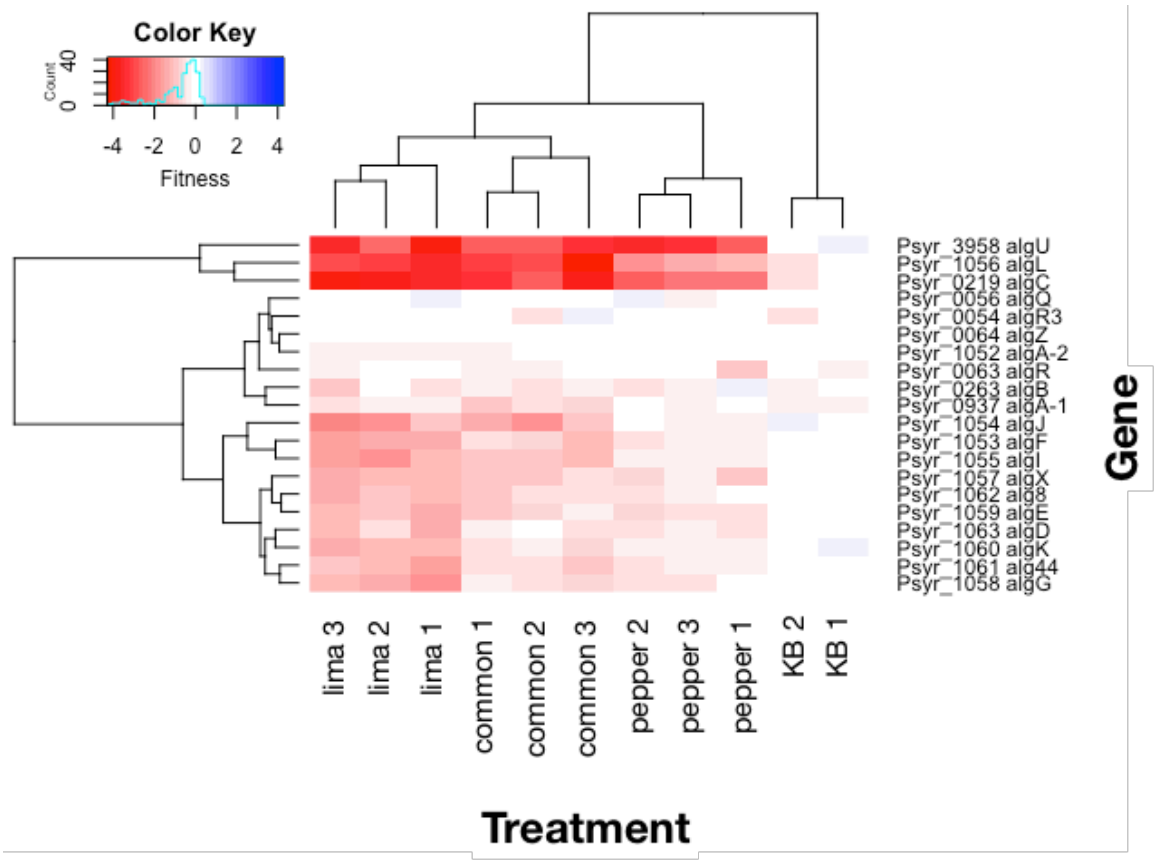


Figure 10. Heatmap of apoplastic fitness for alginate biosynthetic genes. Genes *alg44*, *alg8*, *algA-1*, *algF*, *algI*, and *algK* had significantly different fitness values in the 3 plant species by a Kruskal-Wallis test ($p < 0.05$).

Tables

Strains

Strains	Genotype	Reference
P. syringae B728a	Wild type strain (Rif^R)	(60)
P. syringae B728a	Whole genome barcoded *mariner* transposon library (Rif^R Kan^R)	Chapter 2

Primers

Primers used for BarSeq amplification are described in (28).

Table 1. Numbers of unique barcodes and median reads per gene obtained from mid-log phase cultures following library outgrowth and before inoculation (time0) and from leaves pooled after growth in 100 pots of plants of a particular host (sample). Unique barcodes were calculated as the total that mapped to the genome and for which 3 or more reads were obtained in a given experiment. The total number of genome mapped barcodes in the library was 281,417. Technical (sequencing) replicates are listed separately ("a" and "b"), and share the same time0 reference sample. For an experiment to pass quality control, the median reads per gene in the sample must be ≥ 50 (28).

Experiment	Unique barcodes at time 0	Unique barcodes in sample	Recovery (%)	Median reads/gene at time 0	Median reads/gene in sample
KB_1	187,482	192,078	>100*	278	150
KB_2	197,089	226,857	>100*	378.5	251
*P.vulgaris*_1a	218,149	149,311	68.4	409.5	163
*P.vulgaris*_1b	218,149	151,211	69.3	409.5	155
*P.vulgaris*_2	214,346	156,416	73.0	390	160
*P.vulgaris*_3a	222,473	185,183	83.2	397	203
*P.vulgaris*_3b	222,473	187,289	84.2	397	216
*P.lunatus*_1a	238,593	197,334	82.7	511.5	245
*P.lunatus*_1b	238,593	198,378	83.1	511.5	252
*P.lunatus*_2a	211,661	182,104	86.0	356	175
*P.lunatus*_2b	211,661	181,831	85.9	356	166
*P.lunatus*_3a	208,924	199,475	95.5	368	214
*P.lunatus*_3b	208,924	196,817	94.2	368	203
*C.annuum*_1a	220,024	163,572	74.3	398	198
*C.annuum*_1b	220,024	166,370	75.6	398	206
*C.annuum*_2a	208,449	194,666	93.4	383.5	212
*C.annuum*_2b	208,449	201,173	96.5	383.5	220
*C.annuum*_3a	220,860	198,541	89.9	406	219
*C.annuum*_3b	220,860	204,372	92.5	406	239.5

* More unique barcodes sequenced at the end of an experiment indicates that additional unique barcodes were present but were not sequenced at the start of the experiment (time0).

Table 2. Genes for which their median fitness contribution differed between plant species by a Kruskal-Wallis rank sum test ($p < 0.05$). Fitness values in KB are included for comparison. This table does not include 37 genes where average fitness values were greater than -0.5 and less than 0.5 in all host plant species.

Locus	Name	Description	Classification	Average fitness			
				KB	Common bean	Lima bean	Pepper
Psyr_0025	aroE	shikimate dehydrogenase	Amino acid metabolism and transport	-0.25	-1.62	-2.95	-0.31
Psyr_4133	hisD	histidinol dehydrogenase	Amino acid metabolism and transport	0.12	-2.99	-0.83	-3.80
Psyr_4896	hisH	imidazole glycerol phosphate synthase subunit hisH	Amino acid metabolism and transport	-0.10	-2.20	-1.06	-3.23
Psyr_4852		D-3-phosphoglycerate dehydrogenase	Amino acid metabolism and transport	-0.13	-2.94	-3.40	-1.95
Psyr_0758	scrB	beta-fructofuranosidase	Carbohydrate metabolism and transport	0.05	-1.66	-0.94	-0.52
Psyr_2188		Histidine kinase, HAMP region:Bacterial chemotaxis sensory transducer:CHASE3	Chemosensing & chemotaxis	-0.23	-0.88	0.18	-0.15
Psyr_4687	bioB	biotin synthase	Cofactor metabolism	-2.94	-2.86	-2.24	-1.63
Psyr_4684	bioC	biotin synthesis protein BioC	Cofactor metabolism	-2.76	-2.85	-1.79	-0.87
Psyr_4686	bioF	8-amino-7-oxononanoate synthase	Cofactor metabolism	-3.03	-2.62	-1.99	-1.30
Psyr_2098	cobA	uroporphyrinogen-III C-methyltransferase	Cofactor metabolism	0.27	-0.56	0.08	-0.23
Psyr_0847	ilvH	acetolactate synthase, small subunit	Cofactor metabolism	0.08	-2.45	-1.44	-3.89
Psyr_0846	ilvI	acetolactate synthase, large subunit	Cofactor metabolism	-0.18	-3.10	-1.92	-3.70
Psyr_3487	flgA	flagellar protein FlgA	Flagellar synthesis and motility	0.14	0.27	0.02	0.78
Psyr_0487	gshB	glutathione synthase	Glutathione metabolism	-1.54	1.81	0.78	-0.34
Psyr_1716		conserved hypothetical protein	Hypothetical	0.26	-0.87	-0.40	0.00
Psyr_2543		conserved hypothetical protein	Hypothetical	0.57	-0.07	0.34	-1.05
Psyr_3107		conserved hypothetical protein	Hypothetical	-0.18	-0.54	0.20	0.44
Psyr_3172		Glycosyl transferase, family 3	Hypothetical	-0.04	0.52	-0.30	0.11
Psyr_3611		Protein of unknown function DUF815	Hypothetical	0.22	0.68	-0.56	0.10
Psyr_3889		conserved hypothetical protein	Hypothetical	-0.01	-1.39	-0.06	0.52
Psyr_0014		lipid A biosynthesis acyltransferase	LPS synthesis and transport	-0.53	-2.66	-2.08	-1.44
Psyr_3369		Twin-arginine translocation pathway signal:Tat-	Oxidative stress tolerance	-0.10	-0.55	0.22	-0.25

		translocated enzyme:Dyp-type peroxidase	(Antioxidant enzyme)				
Psyr_2621	pseB	Secretion protein HlyD	Phytotoxin synthesis and transport	-0.06	0.25	-0.03	-0.53
Psyr_2601	salA	regulatory protein, LuxR	Phytotoxin synthesis and transport	-0.37	-1.21	-0.23	-0.68
Psyr_1702	sylA	regulatory protein, LuxR	Phytotoxin synthesis and transport	-0.26	-0.54	-0.10	0.18
Psyr_2614	sypA	Amino acid adenylation	Phytotoxin synthesis and transport	0.03	-0.85	-0.15	-0.69
Psyr_2612	syrP	syrP protein, putative	Phytotoxin synthesis and transport	0.04	-2.07	-0.02	-0.70
Psyr_4158	eftA	conserved hypothetical protein	Plant-associated proteins	-0.39	-1.42	-1.01	-0.09
Psyr_1061	alg44	alginate biosynthesis protein Alg44	Polysaccharide synthesis and regulation	0.02	-0.72	-1.28	-0.34
Psyr_1062	alg8	alginate biosynthesis protein Alg8	Polysaccharide synthesis and regulation	-0.04	-0.70	-1.22	-0.33
Psyr_0937	algA-1	mannose-6-phosphate isomerase, type 2 / mannose-1-phosphate guanylyltransferase (GDP)	Polysaccharide synthesis and regulation	-0.30	-0.70	-0.47	-0.16
Psyr_1053	algF	alginate biosynthesis protein AlgF	Polysaccharide synthesis and regulation	-0.04	-0.81	-1.48	-0.43
Psyr_1055	algI	Membrane bound O-acyl transferase, MBOAT	Polysaccharide synthesis and regulation	0.00	-1.03	-1.57	-0.36
Psyr_1060	algK	Sel1-like repeat protein	Polysaccharide synthesis and regulation	0.17	-0.51	-1.21	-0.28
Psyr_0378	mdoH	Glycosyl transferase, family 2	Polysaccharide synthesis and regulation	-1.04	-3.80	-4.04	-1.08
Psyr_3636	wbpM	Polysaccharide biosynthesis protein CapD	Polysaccharide synthesis and regulation	-0.45	-1.76	-1.48	-1.10
Psyr_3161	aprD	Type I secretion system ATPase, PrtD	Secretion/Efflux/ Export	0.39	0.11	0.38	-0.61
Psyr_4009	oprM	RND efflux system, outer membrane lipoprotein, NodT	Secretion/Efflux/ Export	0.02	-0.52	-0.95	0.20
Psyr_0831	cbrB-1	Two-component response regulator CbrB	Signal transduction mechanisms	-1.45	-1.71	-0.98	0.17
Psyr_4069	colS	ATP-binding region, ATPase-like:Histidine kinase, HAMP region:Histidine kinase A, N-terminal	Signal transduction mechanisms	-0.08	-1.21	-0.57	-0.38
Psyr_4138		Toluene tolerance	Stress resistance	-0.22	-0.20	-0.59	0.28
Psyr_3698	gacS	Response regulator receiver:ATP-binding region, ATPase-like:Histidine kinase,	Transcriptional regulation	0.07	-1.47	-0.44	-0.85

		HAMP region:Histidine kinase A, N-terminal:Hpt					
Psyr_4239	dppB	Binding-protein-dependent transport systems inner membrane component	Transport (peptides)	-0.09	-0.58	-0.33	-0.07
Psyr_4240	dppC	Binding-protein-dependent transport systems inner membrane component	Transport (peptides)	0.14	-0.57	-0.34	0.01
Psyr_1218	hrpK1	type III helper protein HrpK1	Type III secretion system	0.08	-1.28	-0.75	-1.67
Psyr_0914	wpbZ	Glycosyl transferase, group 1		-0.19	-2.08	-1.83	-1.32
Psyr_0915		NAD-dependent epimerase/dehydratase		-0.21	-3.19	-3.51	-1.36
Psyr_1419		preQ(0) biosynthesis protein QueC		-0.65	-1.85	-0.87	-0.10
Psyr_4844		HAD-superfamily hydrolase, subfamily IB (PSPase-like):HAD-superfamily subfamily IB hydrolase, hypothetical 2		-0.10	-1.00	-0.30	-1.69
Psyr_4886		Peptidase M23B		-0.05	-0.53	-0.30	-0.07

Table 3. Genes for which their median fitness contribution differed between plant species by a Kruskal-Wallis rank sum test ($p < 0.05$). Fitness values in KB are included for comparison. These genes are less likely biologically relevant, as their average fitness values were not less than -0.5 or greater than 0.5 in all host plant species examined.

Locus	Name	Description	Classification	Average fitness			
				KB	Common bean	Lima bean	Pepper
Psyr_1074	aapM	amino acid ABC transporter membrane protein 2, PAAT family	Amino acid metabolism and transport	0.17	0.10	0.28	-0.06
Psyr_3876	hisM	amino acid ABC transporter membrane protein 2, PAAT family	Amino acid metabolism and transport	-0.04	0.23	0.08	-0.09
Psyr_2715		Major facilitator superfamily	Carbohydrate metabolism and transport	0.02	0.30	-0.16	0.12
Psyr_3406	aer-2	PAS	Chemosensing & chemotaxis	-0.06	-0.29	0.05	-0.13
Psyr_2995	treY	maltooligosyl trehalose synthase	Compatible solute synthesis	0.02	-0.45	-0.18	0.01
Psyr_1481	ppa-2	Inorganic diphosphatase	Energy generation	-0.09	-0.06	0.03	-0.20
Psyr_1770		Enoyl-CoA hydratase/isomerase	Fatty acid metabolism	-0.07	-0.06	-0.27	0.30
Psyr_3456	fliG	Flagellar motor switch protein FliG	Flagellar synthesis and motility	0.09	-0.24	0.04	0.38
Psyr_3613		Glutathione peroxidase	Glutathione metabolism	0.12	0.39	-0.19	0.08
Psyr_0332		hypothetical protein	Hypothetical	-0.06	0.21	-0.10	0.05
Psyr_1137		Protein of unknown function UPF0153	Hypothetical	0.00	-0.12	0.26	0.12
Psyr_1407		Protein of unknown function DUF28	Hypothetical	-0.07	0.41	0.07	-0.31
Psyr_1533		hypothetical protein	Hypothetical	0.16	-0.39	-0.11	0.15
Psyr_2339		hypothetical protein	Hypothetical	0.09	0.46	-0.37	0.18
Psyr_2947		hypothetical protein	Hypothetical	-0.22	0.47	-0.01	-0.39
Psyr_3006		Protein of unknown function DUF419	Hypothetical	-0.10	0.40	-0.10	0.06
Psyr_3066		conserved hypothetical protein	Hypothetical	-0.22	-0.10	-0.39	0.27
Psyr_3740		Protein of unknown function DUF454	Hypothetical	0.03	-0.33	0.01	0.13
Psyr_3798		conserved domain protein	Hypothetical	0.01	0.16	0.01	-0.09
Psyr_4248		hypothetical protein	Hypothetical	0.01	0.08	0.14	-0.01
Psyr_5112		conserved hypothetical protein	Hypothetical	-0.13	0.03	-0.06	0.16
Psyr_1318	ppc	Phosphoenolpyruvate carboxylase	Organic acid metabolism and transport	0.15	-0.18	0.22	0.09
Psyr_1705	sylD	Amino acid adenylation	Phytotoxin synthesis and transport	0.03	-0.32	-0.02	-0.06
Psyr_4615		Spermidine/putrescine	Polyamine	-0.09	-0.15	0.10	0.20

		ABC transporter ATP-binding subunit	metabolism and transport				
Psyr_3956	mucB	sigma E regulatory protein, MucB/RseB	Polysaccharide synthesis and regulation	-0.04	0.06	0.01	0.33
Psyr_3236	dhcR	transcriptional regulator, LysR family	QAC metabolism and transport	0.32	-0.12	0.06	0.29
Psyr_2576	syfA	Amino acid adenylation	Secondary metabolism	0.10	-0.03	0.10	0.31
Psyr_2864		RND efflux system, outer membrane lipoprotein, NodT	Secretion/Efflux/ Export	0.14	-0.11	-0.01	-0.28
Psyr_3989	xaxA	hypothetical protein	Special	-0.02	0.15	0.08	-0.02
Psyr_4273	cstA	Carbon starvation protein CstA	Stress resistance	0.05	-0.05	0.04	0.15
Psyr_0986	rsmC	16S rRNA m(2)G 1207 methyltransferase		-0.19	-0.37	0.00	0.37
Psyr_0929		Glycosyl transferase, family 2		0.11	-0.20	-0.07	-0.02
Psyr_0989		Lysine exporter protein (LYSE/YGGA)		0.06	-0.25	-0.03	0.32
Psyr_3859		Purine nucleoside permease		-0.01	-0.23	-0.08	0.04
Psyr_4631		PrkA serine kinase		0.07	-0.09	-0.03	0.07
Psyr_5082		Band 7 protein		-0.02	-0.01	-0.24	0.11
Psyr_5111		dTDP-glucose 4,6-dehydratase		-0.06	0.08	-0.03	-0.13

Table 4. Unique and shared gene loci among the three hosts tested having average fitness values less than -2 (A) or average fitness values less than -1 (B). These totals are shown as Venn diagrams in Figure 5.

A.

Host[s]	Total	Genes
P. vulgaris *P. lunatus* *C. annuum*	34	Psyr_0033 Psyr_0034 Psyr_0167 Psyr_0219 Psyr_0473 Psyr_0474 Psyr_0529 Psyr_0531 Psyr_0917 Psyr_0918 Psyr_1212 Psyr_1257 Psyr_1269 Psyr_1350 Psyr_1373 Psyr_1613 Psyr_1663 Psyr_1668 Psyr_1669 Psyr_1748 Psyr_1983 Psyr_1984 Psyr_1985 Psyr_2980 Psyr_3008 Psyr_3958 Psyr_4130 Psyr_4270 Psyr_4369 Psyr_4407 Psyr_4580 Psyr_4581 Psyr_4609 Psyr_4991
P. vulgaris *P. lunatus*	17	Psyr_0014 Psyr_0377 Psyr_0378 Psyr_0557 Psyr_0704 Psyr_0826 Psyr_0915 Psyr_1056 Psyr_1408 Psyr_1410 Psyr_1614 Psyr_3179 Psyr_3883 Psyr_4408 Psyr_4683 Psyr_4687 Psyr_4852
P. vulgaris *C. annuum*	15	Psyr_0469 Psyr_0576 Psyr_0846 Psyr_0847 Psyr_0848 Psyr_1217 Psyr_1544 Psyr_2613 Psyr_4132 Psyr_4133 Psyr_4134 Psyr_4893 Psyr_4894 Psyr_4896 Psyr_4897
P. lunatus *C. annuum*	2	Psyr_1198 Psyr_4362
P. vulgaris	17	Psyr_0454 Psyr_0528 Psyr_0914 Psyr_0919 Psyr_0920 Psyr_0936 Psyr_0951 Psyr_1196 Psyr_2612 Psyr_3637 Psyr_4116 Psyr_4144 Psyr_4340 Psyr_4341 Psyr_4684 Psyr_4686 Psyr_4740
P. lunatus	5	Psyr_0025 Psyr_0532 Psyr_0923 Psyr_1914 Psyr_4018
C. annuum	7	Psyr_0385 Psyr_0386 Psyr_1190 Psyr_1191 Psyr_1208 Psyr_1210 Psyr_4566

B.

Host[s]	Total	Genes
P. vulgaris *P. lunatus* *C. annuum*	90	Psyr_0014 Psyr_0033 Psyr_0034 Psyr_0167 Psyr_0219 Psyr_0377 Psyr_0378 Psyr_0454 Psyr_0469 Psyr_0473 Psyr_0474 Psyr_0528 Psyr_0529 Psyr_0531 Psyr_0532 Psyr_0557 Psyr_0704 Psyr_0826 Psyr_0827 Psyr_0846 Psyr_0847 Psyr_0848 Psyr_0914 Psyr_0915 Psyr_0917 Psyr_0918 Psyr_0919 Psyr_0920 Psyr_0936 Psyr_0951 Psyr_1056 Psyr_1190 Psyr_1197 Psyr_1198 Psyr_1200 Psyr_1205 Psyr_1206 Psyr_1208 Psyr_1210 Psyr_1211 Psyr_1212 Psyr_1213 Psyr_1216 Psyr_1257 Psyr_1269 Psyr_1350 Psyr_1373 Psyr_1408 Psyr_1410 Psyr_1544 Psyr_1613 Psyr_1663 Psyr_1668 Psyr_1669 Psyr_1747 Psyr_1748 Psyr_1983 Psyr_1984 Psyr_1985 Psyr_2613 Psyr_2980 Psyr_3008 Psyr_3179 Psyr_3636 Psyr_3637 Psyr_3883 Psyr_3958 Psyr_4091 Psyr_4130 Psyr_4132 Psyr_4270 Psyr_4340 Psyr_4341 Psyr_4361 Psyr_4362 Psyr_4369 Psyr_4407 Psyr_4408 Psyr_4566 Psyr_4580 Psyr_4581 Psyr_4609 Psyr_4686 Psyr_4687 Psyr_4852 Psyr_4893 Psyr_4894 Psyr_4896 Psyr_4991 Psyr_5132
P. vulgaris *P. lunatus*	25	Psyr_0025 Psyr_0534 Psyr_0579 Psyr_0916 Psyr_0923 Psyr_1054 Psyr_1055 Psyr_1121 Psyr_1401 Psyr_1614 Psyr_1914 Psyr_2077 Psyr_2461 Psyr_3174 Psyr_3199 Psyr_3287 Psyr_4018 Psyr_4158 Psyr_4194 Psyr_4683 Psyr_4684 Psyr_4843 Psyr_5065 Psyr_5130 Psyr_5133

P. vulgaris *C. annuum*	17	Psyr_0385 Psyr_0386 Psyr_0576 Psyr_1191 Psyr_1196 Psyr_1215 Psyr_1217 Psyr_1218 Psyr_3193 Psyr_4133 Psyr_4134 Psyr_4144 Psyr_4512 Psyr_4844 Psyr_4897 Psyr_4940 Psyr_5072
P. lunatus *C. annuum*	2	Psyr_0555 Psyr_1195
P. vulgaris	57	Psyr_0103 Psyr_0201 Psyr_0268 Psyr_0475 Psyr_0533 Psyr_0550 Psyr_0758 Psyr_0831 Psyr_1097 Psyr_1109 Psyr_1247 Psyr_1395 Psyr_1417 Psyr_1419 Psyr_1487 Psyr_1542 Psyr_1588 Psyr_1733 Psyr_1751 Psyr_1907 Psyr_1998 Psyr_2221 Psyr_2264 Psyr_2396 Psyr_2474 Psyr_2501 Psyr_2557 Psyr_2601 Psyr_2612 Psyr_3028 Psyr_3427 Psyr_3552 Psyr_3597 Psyr_3667 Psyr_3675 Psyr_3676 Psyr_3678 Psyr_3690 Psyr_3698 Psyr_3791 Psyr_3889 Psyr_4015 Psyr_4044 Psyr_4069 Psyr_4116 Psyr_4125 Psyr_4136 Psyr_4143 Psyr_4224 Psyr_4740 Psyr_4754 Psyr_4774 Psyr_4882 Psyr_4895 Psyr_5053 Psyr_5067 Psyr_5129
P. lunatus	28	Psyr_0202 Psyr_0259 Psyr_0435 Psyr_0478 Psyr_0524 Psyr_0796 Psyr_0822 Psyr_1053 Psyr_1057 Psyr_1058 Psyr_1059 Psyr_1060 Psyr_1061 Psyr_1062 Psyr_1063 Psyr_1140 Psyr_1667 Psyr_1749 Psyr_2462 Psyr_3146 Psyr_3684 Psyr_3691 Psyr_4008 Psyr_4019 Psyr_4100 Psyr_4627 Psyr_4842 Psyr_4898
C. annuum	8	Psyr_1111 Psyr_2080 Psyr_2245 Psyr_2543 Psyr_2617 Psyr_3459 Psyr_3835 Psyr_5135

Chapter 5

Genome-wide transposon screen of a *Pseudomonas syringae mexB* mutant reveals the substrates of efflux transporters

Abstract

Bacteria express numerous efflux transporters that confer resistance to diverse toxicants present in their environment. Due to a high level of functional redundancy of these transporters, it is difficult to identify those that are of most importance in conferring resistance to specific compounds. The resistance-nodulation-division (RND) protein family is one such example of redundant transporters that are widespread among Gram-negative bacteria. Within this family, the MexAB-OprM protein complex is highly-expressed and conserved among *Pseudomonas* species. We exposed barcoded transposon mutant libraries in isogenic wild-type and Δ*mexB* backgrounds in *P. syringae* B728a to diverse toxic compounds *in vitro* to identify mutants with increased susceptibility to these compounds. Mutants in genes encoding both known and novel redundant transporters, but with partially overlapping substrate specificities were observed in a Δ*mexB* background. Psyr_0228, an uncharacterized member of the Major Facilitator Superfamily of transporters, preferentially contributes to tolerance of acridine orange and acriflavine. Another transporter located in the inner membrane, Psyr_0541, contributes to tolerance to acriflavine and berberine. The presence of multiple redundant, genomically encoded, efflux transporters appears to enable bacterial strains to tolerate a diversity of environmental toxins. This genome-wide screen in a hyper-susceptible mutant strain revealed numerous transporters that would otherwise be dispensable in these conditions. Bacterial strains such as *P. syringae* that likely encounter diverse toxins in their environment such as in association with many different plant species, probably benefit from possessing multiple redundant transporters that enable versatility to tolerate novel toxicants.

Introduction

Bacteria, like all living organisms, must tolerate a variety of potentially harmful, chemically diverse molecules present in the environment. While many of these compounds can be degraded to prevent their accumulation to harmful levels within cells, a common means of tolerance of toxins is to export them from the cell. Bacterial efflux transporters can function both to achieve stress tolerance as well as to contribute to virulence by secreting toxins or effectors of various kinds (40, 44). A given bacterial species commonly possesses a wide variety of efflux transporters. It is presumed that they differ in the specificity of toxins that they export. Broad-specificity transporters actively remove toxins from the cell, and their substrates can include heavy metals, solvents, dyes, detergents, antibiotics, as well as certain host-derived products (157–160). An example of such promiscuous transporters are multidrug resistant (MDR) efflux transporters that remove a wide range of structurally diverse chemical compounds from the cell interior (161). While genes encoding these exporters can be found on plasmids, both pathogenic and nonpathogenic bacteria have comparable numbers of chromosomally encoded MDR systems (47). For those bacteria that colonize eukaryotic hosts, MDR efflux pumps not only export antibiotics and other toxic compounds found in

their environment, but also host-derived antimicrobial compounds (40). Given that genes encoding efflux transporters are highly conserved among diverse bacteria, their physiological role is likely not involved in resistance to clinically relevant antibiotics (162). A high level of redundancy of transporters exists within a given strain, where they may share many of the same substrates (52, 53). These observations prompt the question of why so many transporters are present in a given bacterium, given the broad substrate ranges of these MDR pumps.

MDR efflux transporters are structurally diverse, being found in at least five distinct protein families: the major facilitator superfamily (MFS), the small multidrug resistance (SMR) family, the multidrug and toxic compound extrusion (MATE) family, the ATP-binding cassette (ABC) superfamily, and the resistance-nodulation-division (RND) family (45, 46). Many of these transporter classes contain both substrate-specific transporters as well as less specific MDR-type efflux pumps (47). For example, characterized MFS transporters in *Escherichia coli* range from highly specific sugar:proton symporters such as LacY or XylE to multidrug efflux transporters like EmrD (163). Our understanding of efflux-mediated resistance to toxins and antibiotics would benefit from both identifying the MDR transporters from within the larger collection of genomic transporters for a given organism, and also identifying the specific substrates for those transporters.

Multiple transporters, in both the same and different protein families, can share the same substrates (40, 52). This redundancy poses a challenge for their study, as these alternative transporters can mask mutations that disrupt the function of a given pump under investigation (52). The large number of potential MDR transporters in a given strain also makes their investigation by analysis of targeted gene deletion strains laborious. An example of one of the few such intensive studies is that of Sulavik *et al.* (50), who tested the susceptibility of *E. coli* strains with null mutations in 7 known and 9 predicted efflux genes against 35 toxicants. This study, and others (164, 165), identified the RND transporter AcrAB-TolC as the major determinant of intrinsic toxicant resistance in *E. coli*.

Homologs of AcrAB-TolC are common in Gram-negative bacteria, and have been shown to contribute to virulence of plant pathogens as diverse as *Erwinia amylovora*, *Pseudomonas syringae*, *Ralstonia solanacearum*, and *Xylella fastidiosa* (39, 122–124). Interestingly, inhibition of efflux pumps by known chemical inhibitors can increase the antimicrobial activity of compounds produced by plants (148, 149, 166) and might be considered a plant disease resistance trait. *P. syringae* pv. *syringae* is a plant pathogen that is commonly found both in association with plants as well as in the water cycle (16). These are habitats in which it might be expected to encounter a variety of toxic compounds. MexAB-OprM is the best-characterized AcrAB-TolC homolog present in *Pseudomonas* species. While this protein complex is proposed to secrete the iron-chelating molecule pyoverdine (167), it also contributes significantly to antibiotic resistance (45). In *P. syringae* pv. *syringae* strain B728a, genes in this operon are generally expressed at much higher levels than other RND transporters, both in culture as well as in cells on the leaf surface and in the apoplast (49). However, in *P. aeruginosa*, expression of *mexAB-oprM* is inversely correlated with expression of the related RND transporters *mexEF-oprN* and *mexCD-oprJ* (168). This suggests that regulation of the overall repertoire of efflux transporters is tightly regulated, and that

alterations in one or more can lead to compensatory changes in the others. To examine the role of MexAB-OprM homologs and other potential MDR transporters present in strain B728a, we interrogated the fitness of a large library of randomly-barcoded *mariner* transposon mutants (RB-TnSeq) in culture media containing diverse antimicrobial compounds. RB-TnSeq is a modification of TnSeq where each transposon insertion is tagged with a unique 20-nucleotide barcode (28). As transposon insertions are only mapped once for a given library, this reduces the effort to use that library for fitness contributions of genes in multiple conditions. Changes in relative barcode abundance over time are used as a proxy for the relative fitness contribution of a given gene in a given condition. This method can be used to associate bacterial genes with their importance to fitness in different growth conditions, and has been used to improve genome annotations for diverse bacteria (29). Of particular importance for this current study is the ease and scale with which this method could be used to associate genes encoding transporter proteins with their likely substrates. Here, we used RB-TnSeq to identify likely substrates for B728a efflux transporters, with a particular focus on complementary RND homologs of MexAB-OprM.

Results

Creation of a barcoded transposon library in B728a Δ*mexB*

To test the role of redundant RND efflux proteins *in vitro*, we created a barcoded *mariner* transposon library in a Δ*mexB* strain. Comparisons of gene fitness in the mutant library with that of the barcoded transposon library in wild-type (WT) strain B728a (Chapter 2), allowed us to directly test the fitness contributions of B728a transporters in both genetic backgrounds, enabling the complementarity of other transporters with MexAB-OprM to be quantified. The Δ*mexB* transposon library contains 237,285 unique insertion strains, and is a similar size and insertion density to the WT library (Table 1). We hypothesized that insertional mutants in efflux genes in the Δ*mexB* genotype would be less fit than mutants in the WT genotype, particularly when exposed to toxic substrates of MexAB-OprM. The B728a genome encodes 668 predicted transport proteins (48); here we primarily focus on the RND transporters but also investigate members of the other protein families that typically encode MDR efflux transporters (Table 2).

The *mexB* deletion strain completely eliminates the activity of MexAB-OprM, including inactivating OprM. To determine if the deletion of *mexB* resulted in a polar mutation, we expressed *mexAB* and *mexAB-oprM* under the native promoter in the Δ*mexB* mutant strain. While the Δ*mexB* strain containing the plasmid to express the entire operon was able to tolerate acriflavine and berberine at WT levels, the Δ*mexB* strain expressing only *mexAB* was not (Fig. 1). This indicates that the Δ*mexB* transposon library does not produce a functional OprM.

Testing the fitness of mutants in the WT and Δ*mexB* transposon libraries *in vitro*

Both mutant libraries were grown in the rich culture medium King's B (KB) with 28 different antimicrobial compounds. For compounds where the minimum inhibitory concentration (MIC) was known for B728a WT and Δ*mexB* (39), we used ¼ the MIC. However, 14 compounds for the WT library and 15 compounds for the Δ*mexB* library did not pass previously established metrics (28) for quality control of the RB-TnSeq

experiments. The most common reason a given experiment did not pass quality control was the high correlation between GC content and fitness or normalization of fitness estimates (high correlation between fitness of adjacent genes on opposite strands). Failure of some experiments on this basis is expected, particularly for higher GC content bacteria (28), and we proceeded with experiments that passed quality control for both libraries (16 unique antimicrobial compounds). B728a has a moderately high GC content of 59% (106). In experiments that did pass quality controls, the antimicrobial compounds and their final concentration are described in Table 3.

For each gene, fitness is calculated as the $\log_2$ ratio of barcode abundance following growth in a given condition relative to barcode abundance measured initially at time0. As expected, insertions in the majority of genes did not contribute to fitness as measured by relative barcode abundance in the population, and thus the fitness values for most genes were close to 0. A mutant with a fitness value of -1 is approximately 50% less abundant relative to the typical strain in the library under that experimental condition.

Identification of MexAB-OprM substrates

For many substrates, the dominant activity of MexAB-OprM seen in other species was expected to mask any requirement for complementary transporters. We therefore hypothesized that the Δ*mexB* deletion library would unmask the contributions of such alternative transporters to tolerance of various toxins. The majority of antimicrobial compounds tested here are substrates of the MexAB-OprM transporter, and have known MICs (39). Other previously non-investigated compounds include capsaicin, flavone, rifampicin, and rotenone.

For the majority of the antimicrobial compounds tested, disruption of any gene in the MexAB-OprM operon resulted in decreased fitness relative to insertions in the average gene in the WT strain (Fig. 2). No fitness change for insertional strains in this operon was observed when cells were exposed to capsaicin, rifampicin, or rotenone. There was also no fitness change when exposed to ampicillin, a previously-reported substrate (39). For erythromycin, only transposon insertions in *oprM* decreased mutant fitness. Flavone is a likely a substrate of this complex due to the large decreases in fitness comparable to that seen upon exposure to known substrates that accompanied disruptions across the *mexAB-oprM* operon. Using this method, we would expect insertional mutations in *mexA* and *oprM* to be neutral in the Δ*mexB* genetic background. All strains in the Δ*mexB* transposon library indeed were unable to assemble a complete MexAB-OprM complex, since insertions elsewhere in the *mexAB-oprM* operon did not alter fitness. For all antimicrobial compounds tested, *mexA* and *oprM* are dispensable in the Δ*mexB* library, with fitness values close to 0 (Fig. 2).

Transporters homologous to MexAB contribute to resistance to diverse toxicants

To identify efflux transporters that had the largest contribution to tolerance of various toxicants, we calculated the number of compounds in which genes encoding transporters or transporter components had a fitness value less than -1. RND transporters generally had a larger contribution to toxin tolerance in the Δ*mexB* library than in the WT, but some were important in both backgrounds (Table 4). B728a contains 16 RND transporters, including MexAB-OprM (48). Due to our experimental

design that focused on known MexAB-OprM substrates, *mexA* and *mexB* were required for competitive fitness in the presence of 10 compounds. The outer membrane transporter OprM is likely shared with additional efflux pumps, and so is required for tolerance of an additional 3 compounds.

In addition to MexAB, four homologous RND transporters contributed substantially to tolerance of several toxicants (Table 4). The *mexEF-oprN* operon however, was only required for tolerance of these compounds in the absence of MexB (Fig. 3). MexEF-OprN is likely redundant with MexAB-OprM for the substrates acridine orange, acriflavine, berberine, chloramphenicol, flavone, nalidixic acid, nitrofurantoin, and phloretin. The lack of a phenotype for mutants in the WT background is consistent with the likely subsidiary role of this transporter for these shared substrates, and therefore masking by the more highly expressed MexAB-OprM under these conditions.

Psyr_2483-5 is likely redundant with MexAB-OprM for the substrates acriflavine, berberine, carbenicillin, chloramphenicol, erythromycin, and phloretin, with decreased fitness for disruptions in the Δ*mexB* genetic background (Fig. 4). Psyr_2483-5 independently contributes to resistance to rifampicin. This apparent operon encodes an unnamed RND transporter, and is most homologous to the MuxABC-OpmB complex in other Pseudomonads, such as *P. aeruginosa* PAO1 (Table 5).

MexCD (*Psyr_2282-3*) was required for full competitive fitness in the presence of acriflavine, berberine, erythromycin, and nalidixic acid (Fig. 5). Importantly, these phenotypes could be seen in both the WT and Δ*mexB* backgrounds. Negative fitness values for insertional mutants in this operon were, however, greater in the WT background, although this may have been at least partially due to the higher toxicant concentrations used in testing the mutants in the WT background than that of the Δ*mexB* library. This operon does not encode an outer membrane protein, and a gene for such a required component is likely located elsewhere in the genome.

Of the compounds tested here, Psyr_0344-6 only contributed to tolerance of phloretin in the Δ*mexB* background (Fig. 6). Interestingly, the disruption of any gene in this operon resulted in a mutant strain that was more fit than mutants of other genes when exposed to acriflavine, berberine, or nalidixic acid, but only if *mexB* was also absent. This operon likely requires an unknown outer membrane protein located elsewhere in the genome.

Inner membrane transporters appear more substrate specific

The same computational analysis was used to interrogate the MFS and other inner membrane transporters to identify those with fitness contributions in the presence of various toxic compounds in either or both genetic backgrounds. Among 68 MFS transporters examined, only Psyr_0228 contributed to tolerance of any compounds. Psyr_0228 contributed to competitive fitness in both acriflavine and acridine orange (Fig. 7). Interestingly, this gene contributed to acriflavine resistance even in a WT background, indicating that its role was independent of *mexB*. In contrast, this gene contributed to acridine orange resistance only in the Δ*mexB* genotype. Disruption of the gene encoding the SMR transporter Psyr_0541 strongly decreased fitness in berberine and mildly decreased fitness in acriflavine, independent of the *mexB* genotype (Fig. 8). In the Δ*mexB* strain, disruption of *Psyr_0541* resulted in a mild susceptibility to carbenicillin (Fig. 8).

Expanding this analysis to include compounds from experiments that failed the quality control cutoffs revealed an additional apparently substrate specific MFS transporter. Psyr_4519 contributed strongly and specifically to competitive fitness in fusaric acid, in both the WT and Δ*mexB* strains (Fig. 9). For the Δ*mexB* transposon library, fitness calculations in fusaric acid failed quality control due to a high correlation between gene fitness and GC content (Pearson correlation coefficient r = 0.33, where $|r| \leq 0.2$ is required to pass quality control).

Several ABC transporters contribute to growth in rich media

Most genes annotated as encoding ABC transporter subunits were putative amino acid or carbohydrate transporters. In the conditions tested, several ABC transporters were required for competitive fitness in the rich media controls. For example, insertions in *Psyr_0917-8*, encoding the polysaccharide permease ABC transporter RfbAB-2 strongly decreased fitness in all conditions, including the controls containing only KB medium, in both the WT and Δ*mexB* genotypes (Fig. 10). RfbAB-2 appears nearly essential for growth in KB. Genes encoding a putative peptide ABC transporter, *Psyr_1754-9*, were required for competitive fitness under all conditions, including the KB controls, but only in a Δ*mexB* mutant background (Fig. 11).

Overlapping substrate specificities between transporters

The MDR efflux transporters tested here displayed a range of apparent substrates, with varying degrees of overlap with MexAB-OprM (Fig. 12). Compounds such as acriflavine, berberine, and phloretin are probable substrates of multiple RND transporters. MexCD was required for tolerance of fewer compounds tested than MexAB, MexEF, or MuxABC. As hypothesized, the inner membrane transporters Psyr_0228, Psyr_4519, and Psyr_0541 contributed to resistance to only a few toxicants. However, for the compounds examined here, no other inner membrane transporters contributed to competitive fitness.

Zone of inhibition assays to test predicted mutant phenotypes

The fitness contributions of the various transporters were all conducted at a given concentration of a particular toxicant. The lower fitness seen for various mutants suggested a hypersensitivity to that compound. To demonstrate however that disruption of these transporters reduced the concentration of the toxicant at which any growth could occur we constructed targeted deletion mutants of the substrate binding proteins MexF, MuxB, and MexK (Psyr_0346). We used zone of inhibition (ZOI) assays to provide semi-quantitative estimates of the MIC of the toxicants. The Δ*mexB*Δ*mexF* double mutant was significantly more susceptible to berberine, phloretin, and rifampicin than the WT and either single deletion strain (Fig. 13a). The Δ*mexB*Δ*muxB* double mutant was significantly more sensitive to acriflavine and rifampicin than the WT and either single deletion strain (Fig. 13b). As suggested by positive fitness values observed in mixture studies, the Δ*mexBmexK* double mutant exhibited significantly decreased susceptibility to acriflavine and berberine than the Δ*mexB* mutant (Fig. 13c).

Discussion

Bacteria encounter and must tolerate diverse and potentially toxic molecules in the various habitats they occupy. While efflux transporters are generally thought to play a major role in the tolerance of toxins, it is still unclear why bacterial genomes encode such a large number of MDR transporters. With high expression and wide substrate ranges, RND efflux transporters such as AcrAB-TolC and MexAB-OprM are essential for environmental survival and colonization of eukaryotic hosts (169). However, for these same reasons, these transporters are thought to mask the contributions to resistance of homologous RND transporters (50). Prompting our investigation in *P. syringae*, studies have shown that expression of multiple alternative transporters is increased when these primary RND pumps are disrupted (168, 170). Our studies of these alternative and redundant pumps, has benefited from a method that has previously been successfully employed that relies on the use of hypersusceptible pump mutants, to unmask redundant transporters (171–173).

B728a contains a diverse array of transporters, many with likely roles in multidrug resistance (48). The MexAB-OprM complex has been shown to be important for tolerance of several *P. syringae* strains to a variety of antibiotics and other toxins, as well as virulence during plant colonization (39). However, it has remained a challenge to associate particular transporters with their substrates, especially for host-produced compounds. One strategy to overcome the masking effect of such dominant transporters is either to genetically disrupt genes encoding MDR efflux transporters, as discussed previously, or chemically inhibit them (148, 166, 174). Here, we were able to use a transposon library constructed in a Δ*mexB* strain to reveal the role of several partially redundant RND efflux pumps. While several of these pumps appeared completely masked by the role of MexAB-OprM under the in vitro conditions used here, some were complementary to MexAB-OprM, contributing incrementally to tolerance of various toxicants.

Among those transporters with an incremental role in tolerance were *mexEF-oprN* and *muxABC-opmB.* These transporters however, were only important for a subset of the MexAB-OprM substrates tested. We expected MexEF-OprN to be important in the Δ*mexB* genotype for two main reasons. First, previous transcriptional measurements in *P. syringae* strain B728a show *mexEF-oprN* is generally the second-highest expressed RND transporter after *mexAB-oprM* under a variety of conditions in culture and *in planta* (49). Second, over-expression of *mexEF-oprN* in *P. aeruginosa* strains lacking *mexAB-oprM* increased resistance to several antibiotics (53). MuxABC-OpmB is an unusual RND operon because it is expected to have two inner membrane components. This transporter was only recently characterized in *P. aeruginosa* (175), and is homologous to the *E. coli* complex MdtABCD (also annotated as YegMNOB) (171, 172). Rifampicin was not previously known to be a substrate of any Mex pumps (53), but full resistance appears to require the MuxABC-OpmB complex.

Of the eight RND operons likely involved in multidrug resistance, four do not contain genes encoding outer membrane proteins. Transporters such as these likely require the presence of outer membrane proteins encoded by genes located elsewhere in the genome. The archetypal outer membrane protein TolC and its homologs (such as OprM and OprN in *Pseudomonas* species) can couple to many different transporters, and apparently do not play a role in substrate specificity (176). Since the construction of the unmarked Δ*mexB* strain also disrupted the *oprM* gene, this may have reduced the

functionality of transporters MexCD, Psyr_0344-6, Psyr_2193-4, and Psyr_3130-2 in that strain if they require OprM for maximum functionality. Combined with the minor roles of these other transporters relative to that of MexAB-OprM, this would potentially explain the lack of fitness cost upon their disruption in both the WT and Δ*mexB* genotypes. We should emphasize however that nearly all of the compounds tested here were known MexAB-OprM substrates. In such a scenario, it might be expected that MexAB-OprM would play a dominant role in tolerance to such compounds. These other transporters, however, probably are not masked by the role of MexAB-OprM for many other compounds and conditions in which the cells might encounter them. Very few studies have addressed the breath of the substrate range for MexAB-OprM beyond that of clinically important antibiotics, and given the wide variety of other toxic natural products that bacteria such as *P. syringae* would encounter in its myriad of habitats, it seems likely that many would be found for which these other efflux transporters might play a dominant role. Indeed, evidence for such a role for Psyr_0228 in the tolerance of acriflavin was seen here. It should prove fruitful to test this hypothesis further by the testing of these transposon libraries for their tolerance of diverse additional compounds. Not only will this provide further evidence of the constraints of MexAB-OprM in the chemical ecology of *P. syringae*, but also provide contextual evidence for the retention of the myriad of other efflux pumps in this species that presumably remain under selection, suggesting the importance in at least some settings encountered by this cosmopolitan bacterium. While we observed redundancy in efflux transporter function, seen in the overlapping substrate range of several RND transporters, it may simply benefit the cell to have multiple mechanisms of resistance for a given toxicant.

In Gram-negative bacteria, inner membrane transporters such as those in the MFS can function cooperatively with outer membrane transporters (like RND transporters) (52). This has been proposed to synergistically increase overall toxin resistance for the cell (53). Together with the high level of substrate promiscuity for most MDR RND transporters, we reasoned that inner membrane transporters might contribute to substrate specificity. This would prevent unnecessary efflux, attenuating the proton motive force. If this were true, we would expect to observe that inner membrane transporters would each contribute strongly to resistance of only a few toxicants. For Psyr_0228, this appeared to be the case, which contributed only to the tolerance of acridine orange and acriflavine, two molecules that are structurally similar (Fig. 14). (Technically three molecules, as acriflavine is typically sold as a mixture of the related acriflavine + proflavine). The SMR protein Psyr_0541 is homologous to the quaternary ammonium compound-resistance protein QacE in *P. fluorescens* SBW25 and QacH in *P. aeruginosa* PA14 (177). Both acriflavine and berberine are quaternary ammonium compounds, and it is noteworthy that *Psyr_0541* was found to be necessary for tolerance to both of these compounds. Insertional mutations at this locus were particularly susceptible to berberine. The inner membrane transporters Psyr_0228 and Psyr_0541 appear to share some substrates with MexAB-OprM. More information is required to determine to what extent these transporters function cooperatively if at all. In addition, it is possible that these transporters interact with molecules that are not substrates of MexAB-OprM, potentially with the assistance of outer membrane transporters other than OprM.

Efflux transporters are typically easily detected and annotated computationally due to their sequence homology and transmembrane domains. Determining their substrates is another matter. Here we show that RB-TnSeq can be used to characterize the substrates of MDR efflux transporters, using a range of structurally diverse antimicrobial compounds. This method is sensitive because it measures competitive fitness, rather than simple inhibition of growth, a phenotype important to bacteria in complex environments where they face competition with other microbes. RB-TnSeq is also very cost effective once transposon mutants are initially mapped, since amplification of the barcodes in the same library enables up to 96 experiments run on a single Illumina sequencing lane (28). This high throughput makes it practical to readily test a large number of conditions and compounds. Characterizing MDR transporter redundancy is essential to our understanding of not just clinical antibiotic resistance, but also the role of these abundant proteins for bacterial survival in diverse environments.

Using RB-TnSeq in a hyper-susceptible mutant strain allowed us to examine the role of alternative transporters that might not otherwise be active or discernible. Furthermore, plant-produced compounds are often not as toxic to bacteria as common antibiotics produced by bacteria or fungi, and this has been hypothesized to be due to the activity of efflux transporters (148). There is some evidence that plants can produce MDR pump inhibitors in addition to antimicrobial compounds (178), and this reinforces the importance of multi-"drug" resistance transporters in tolerating diverse plant-produced and other naturally occurring antimicrobial molecules. Multiple-component, outer membrane transporters are apparently highly redundant, while inner membrane transporters can be more substrate specific. This may help explain the abundance of partially redundant MDR efflux transporters found in many bacteria, especially those strains typically found in environments with exposure to diverse toxic chemicals.

Materials and Methods

Bacterial strains and growth media

P. syringae pv. *syringae* B728a was originally isolated from a bean leaf (*Phaseolus vulgaris*) in Wisconsin (60). The complete genome sequence for B728a is available on NCBI GenBank as accession CP000075.1 (106). B728a and derivative mutant strains were grown on King's B (KB) agar or in KB broth (107), at 28˚C. *E. coli* strains S17-1, TOP10, and XL1-Blue were grown on LB agar or in LB broth at 37˚C. When appropriate, the following antibiotics were used at the indicated concentrations: 100 µg/ml rifampicin, 50 µg/ml kanamycin, 15 µg/ml tetracycline, 100 µg/ml spectinomycin, and 40 µg/ml nitrofurantoin.

Complementation of the *mexB* deletion strain

Expression constructs of *mexAB* and *mexAB-OprM* were constructed through PCR amplification of this partial or complete operon, as well as 1161 bp of upstream sequence that would include potential promoter regions. These PCR products were ligated into the XbaI and EcoRI restriction enzyme sites of the plasmid p519*ngfp* (179). The ligation mixture was subsequently transformed into chemically competent *E. coli* XL1-Blue. Plasmids were confirmed to contain the correct insertion sequences by Sanger sequencing, and then electroporated (1.8 kV) into the *E. coli* donor strain S17-1. Plasmids were introduced into B728a WT and B728a Δ*mexB* by conjugation on KB

overnight, and then selected for 3 days on KB containing kanamycin and nitrofurantoin (*E. coli* counter selection).

Construction of bar-coded transposon libraries

Construction of the B728a WT barcoded transposon insertion library was described in Chapter 2. The Δ*mexB* library was constructed in a similar fashion. For the Δ*mexB* library, we first removed the kanamycin resistance cassette from the B728a Δ*mexB* deletion strain. The pFLP2Ω plasmid containing the Flp recombinase and a spectinomycin resistance cassette (180) was introduced into the Δ*mexB* strain by conjugation. Exconjugants were selected on spectinomycin. Colonies were screened for the loss of kanamycin resistance by plating, and the loss of the kanamycin cassette in sensitive colonies was confirmed by PCR. The pFLP2Ω plasmid was cured from the deletion strain by several overnight passages in KB containing rifampicin only, and spectinomycin sensitivity was confirmed by plating. The rifampicin resistant Δ*mexB* deletion strain was used as the recipient for conjugation with the barcoded *mariner* transposon library using the same protocol.

In vitro growth of the library

Aliquots of the transposon libraries that were stored at -80˚C were removed from cold storage, thawed, and inoculated into 25 ml fresh KB with 100 µg/ml kanamycin and grown for approximately 7 hours at 28˚C with shaking until the culture reached mid-log phase, OD_{600} 0.5 - 0.7. For time0 samples, 1 ml aliquots were pelleted by centrifugation and the pellets were frozen at -20˚C until DNA purification. All *in vitro* experiments were performed in 24-well plates, containing 1 ml total volume per well. Each compound used was added to KB containing 100 µg/ml kanamycin to a final volume of 950 µl. Cell suspensions (50 µl) were added to each well last. Libraries were grown overnight (15 hours) at 28˚C with shaking. The cells from each well were then pelleted and frozen until DNA purification.

DNA isolation and library preparation

DNA from frozen pellets was isolated using the Qiagen DNeasy Blood & Tissue Kit according to manufacturer's instructions. Cell lysis was performed at 50˚C for 10 minutes as per manufacturer instructions. Purified genomic DNA was measured on a nanodrop device and 200 ng of total DNA was used as a template for DNA barcode amplification and adapter ligation as established previously (28). For each time0 sample, two separately purified DNA samples were sequenced as technical replicates.

Sequencing and fitness data generation

Barcode sequencing, mapping, and analysis to calculate the relative abundance of barcodes was done using the RB-TnSeq methodology and computation pipeline developed by Wetmore *et al.* (28); code available at bitbucket.org/berkeleylab/feba/. Briefly, TnSeq was used to map the insertion sites and associate the DNA barcodes to these insertions. Fitness values for each gene are calculated as the $\log_2$ ratio of relative barcode abundance following library growth in a given condition divided by the relative abundance in the time0 sample. Barcode counts were summed between replicate time0 samples. For analysis, genes must have adequate coverage in the time0 sample: at

least 3 reads per strain and 30 reads per gene (28). Fitness values are normalized across the genome so the median gene has a fitness value of 0. All experiments described herein passed previously described quality control metrics (28). Experimental fitness values are publically available at fit.genomics.lbl.gov.

Identification of efflux pumps and assessment of fitness

We focused our initial analysis on genes that are homologous to *mexAB-oprM*. In the Transport Database (membranetransport.org) (48), B728a is annotated as containing 16 RND transporters. We manually curated this list of genes to focus on those operons potentially involved in "multidrug resistance": namely known *mex* genes or operons containing an "acriflavin resistance" or "hydrophobe/amphiphile efflux-1" and HlyD secretion protein. We filtered out 2 pseudogenes and 6 genes with known or likely unrelated function: protein export proteins SecD and SecF (Psyr_1230-1), syringomycin efflux (Psyr_2622), and heavy metal efflux pump CzcA (Psyr_4803). *Psyr_1128* is adjacent to a heavy metal (Cu/Ag) two-component system. *Psyr_1701* is a hypothetical protein in the syringolin A biosynthesis gene cluster.

For the remaining 8 RND operons, we compared fitness values for all genes within each operon across all experiments that passed quality control in both libraries. We focused on pumps where multiple encoding genes in the same operon contributed strongly to competitive fitness (fitness value < -1). As a negative control, the 8 genes with expected functions unrelated to drug resistance were confirmed to not be required in any conditions tested (fitness > -0.5 in both libraries). We performed the same analysis using 68 predicted MFS genes (a mixture of potential drug resistance transporters as well as diverse sugar transporters), 210 ABC transporter genes, the SMR gene *Psyr_0541*, and *norA* (*Psyr_0073*). Heatmaps were plotted in R (108) using the gplots package, version 3.0.1.1 (110).

Construction of targeted deletion mutants

Deletion strains were constructed using an overlap extension PCR protocol as described previously (111). Briefly, 1 kb DNA fragments upstream and downstream of the genes of interest were amplified along with a kanamycin resistance cassette from pKD13 (112). These three fragments were joined by overlap extension PCR. The resulting fragment was blunt-end ligated into the SmaI site of pT*sacB* (113), and transformed into the *E. coli* subcloning strains TOP10 or XL1-Blue, and then the *E. coli* conjugation donor strain S17-1. This suicide plasmid was conjugated into B728 on KB overnight, and then selected for 3 days on KB containing kanamycin and nitrofurantoin. Putative double-crossover colonies that were kanamycin resistant and tetracycline sensitive were selected for screening using external primers and further confirmed by PCR and Sanger sequencing.

Zone of inhibition assays

Zone of inhibition assays were used to test deletion mutants for altered susceptibility to selected toxicants *in vitro*. Strains were grown on KB agar containing rifampicin overnight, resuspended in 10 mM KPO_4, and diluted to an OD_{600} of 0.1. 200 µl of the cell suspension was spread on each plate. 6 mm diameter paper discs were used to absorb antibiotics at the following concentrations: acriflavin, 20 µl of 3 mg/ml

H_2O; berberine, 30 µl of 50 mg/ml DMSO; flavone, 15µl of 20 mg/ml DMSO; and phloretin, 30 µl of 30 mg/ml DMSO. 3 replicate discs were used per plate. Plates were incubated at 28˚C overnight, and the diameter of the clearance zone was measured. Graphs were plotted in R using the ggplot2 package version 3.1.1 (109).

Acknowledgements

We thank Morgan Price for assistance with RB-TnSeq sequence analysis, and Jayashree Ray for mapping the insertions sites of the B728a WT and Δ*mexB* transposon libraries. We thank Olga Smelik for experimental assistance. Capsaicin, flavone, and rotenone were kindly provided by Cat Adams and Dr. Thomas Bruns at UC Berkeley. Funding for Tyler Helmann was partially provided by the Arnon Graduate Fellowship and the William Carroll Smith Fellowship. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

Figures

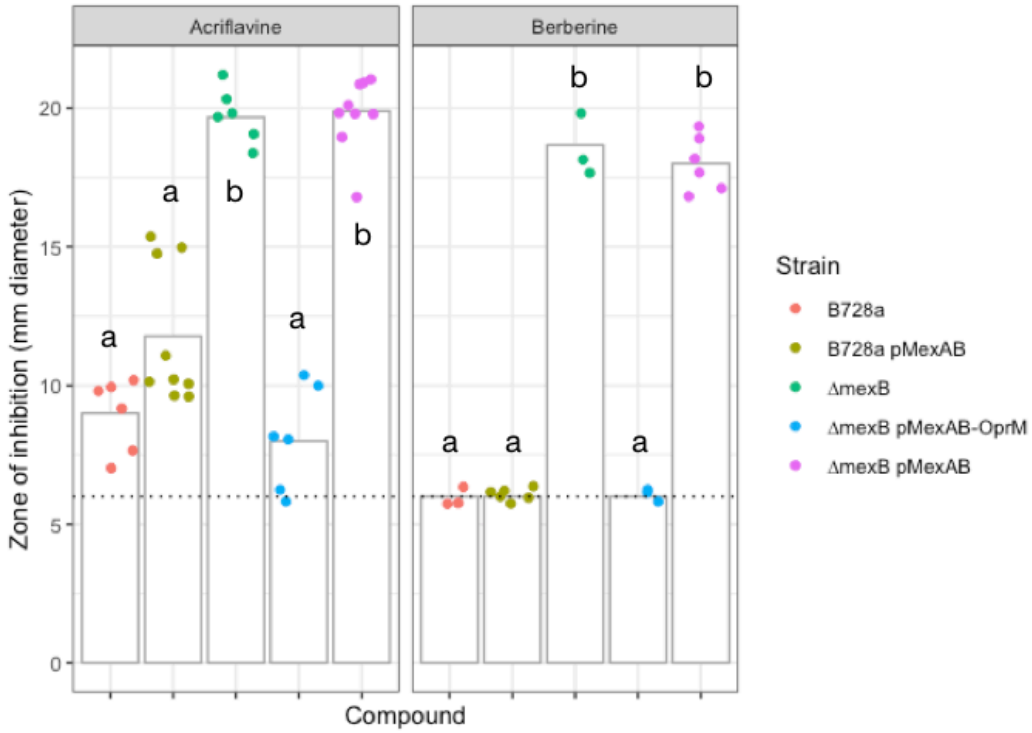


Figure 1. Zone of growth inhibition assays to test antibiotic sensitivity of B728a Δ*mexB* complementation variants. Mean values displayed as column height. For each compound tested, means marked with the same letter do not differ at p = 0.01 using Tukey's HSD test. The lower limit of measurement is 6 mm (dotted line).

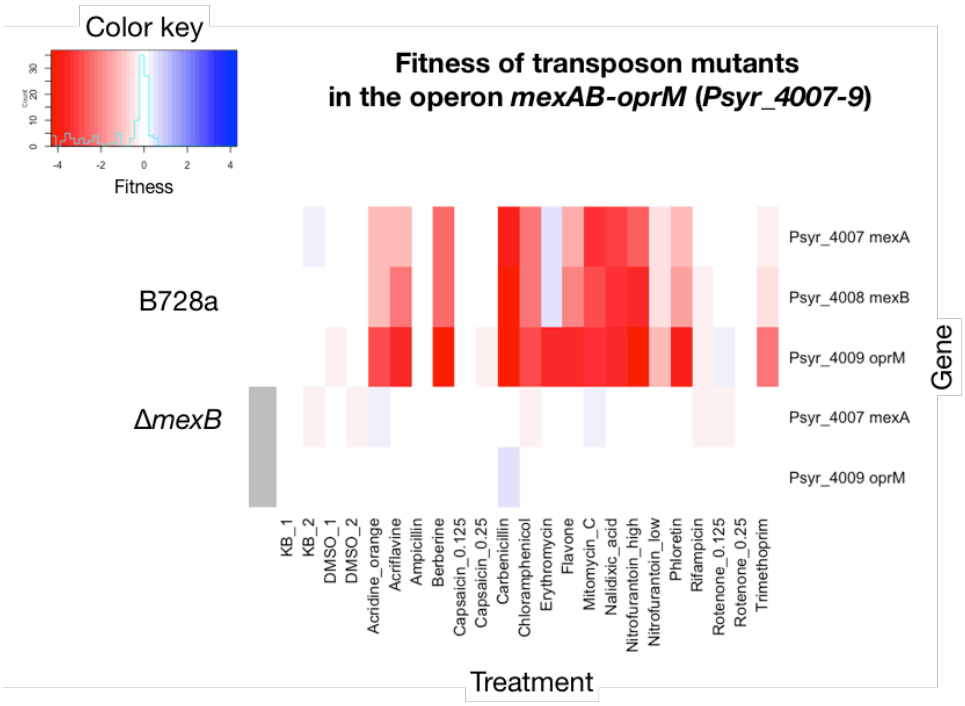


Figure 2. Fitness of transposon insertional mutants in the operon *mexAB-oprM* (*Psyr_4007-9*) in both a WT background and in cells in which *mexAB-oprM* has been disrupted. KB and DMSO are the media and media + solvent control respectively. Fitness is calculated as $\log_2$ change in relative insertion strain barcode abundance for a given gene.

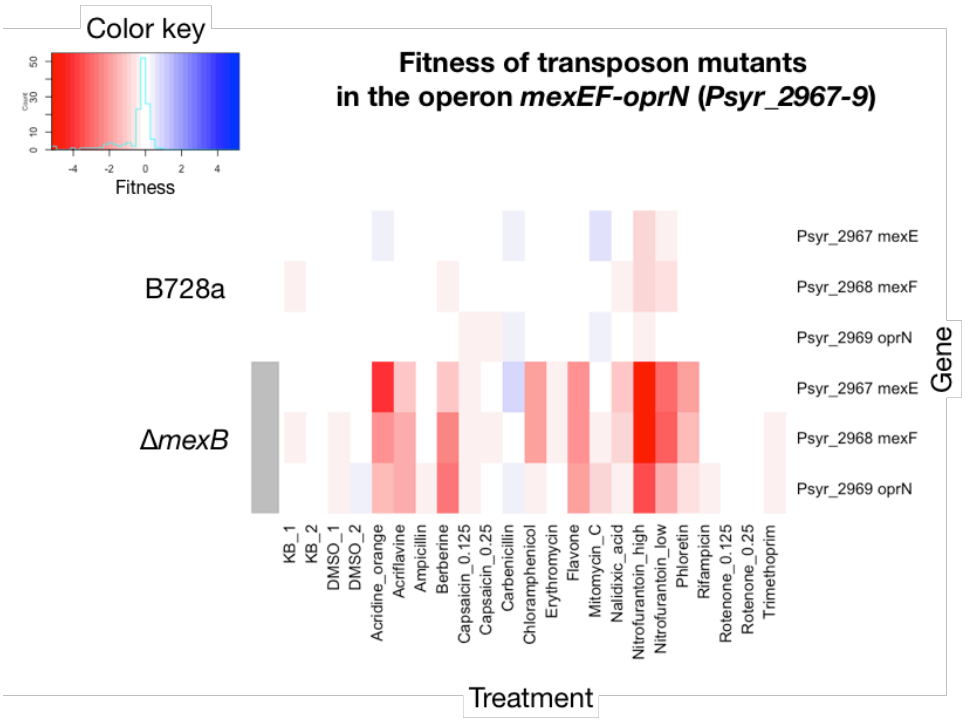


Figure 3. Fitness of transposon insertional mutants in the operon *mexEF-oprN* (*Psyr_2967-9*) in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as $\log_2$ change in relative insertion strain barcode abundance for a given gene.

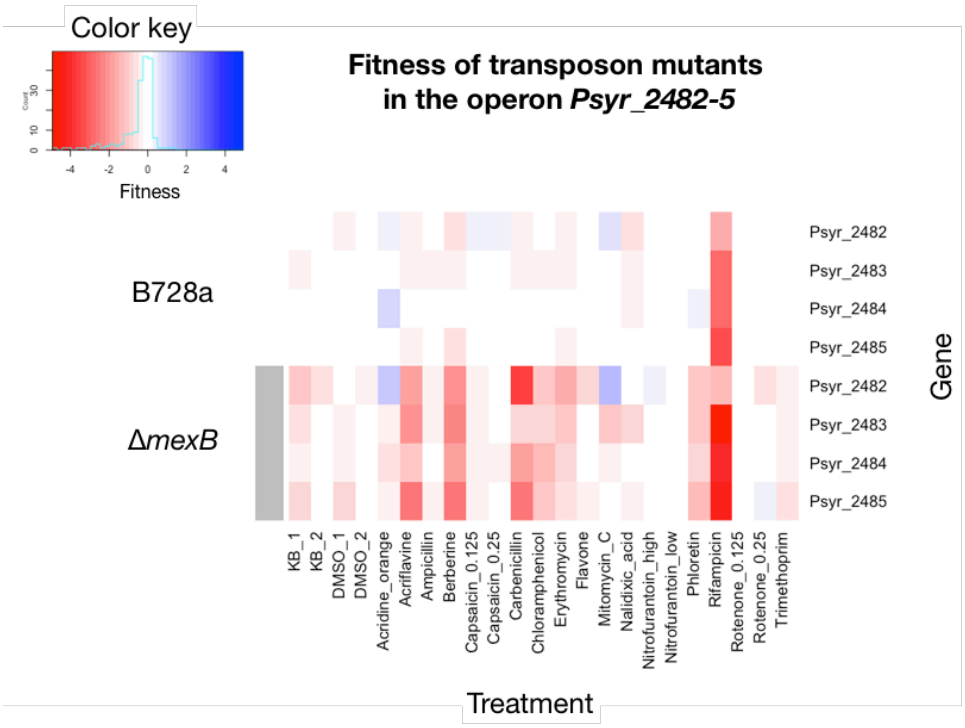


Figure 4. Fitness of transposon insertional mutants in the operon *muxABC-opmB* (*Psyr_2482-5*) in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as $\log_2$ change in relative insertion strain barcode abundance for a given gene.

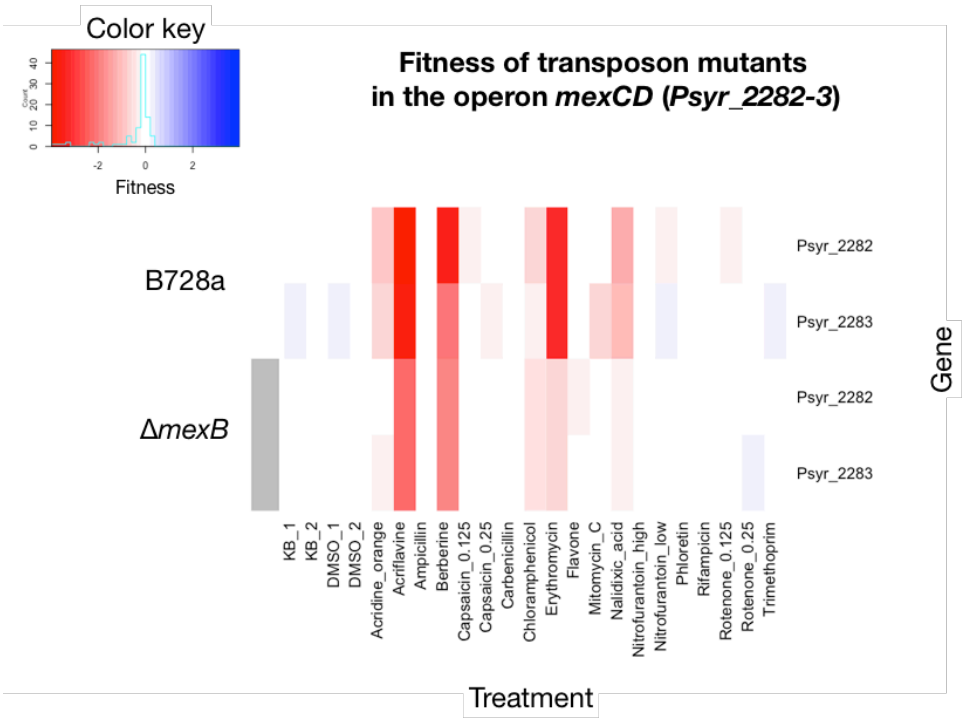


Figure 5. Fitness of transposon insertional mutants in the operon *Psyr_2282-3* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

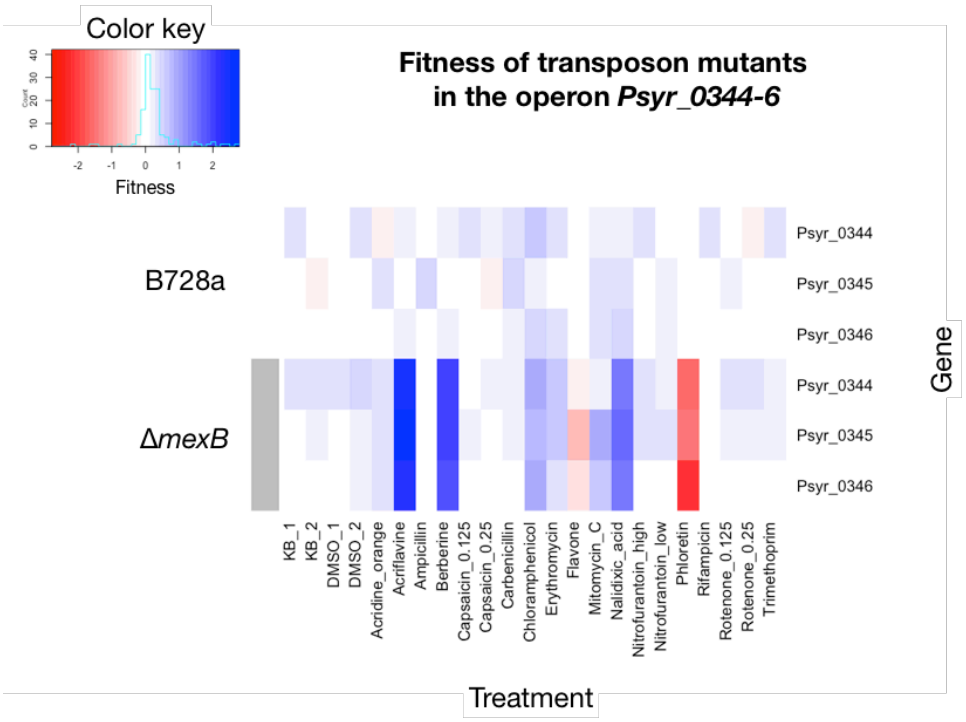


Figure 6. Fitness of transposon insertional mutants in the operon *Psyr_0344-6* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as $\log_2$ change in relative insertion strain barcode abundance for a given gene.

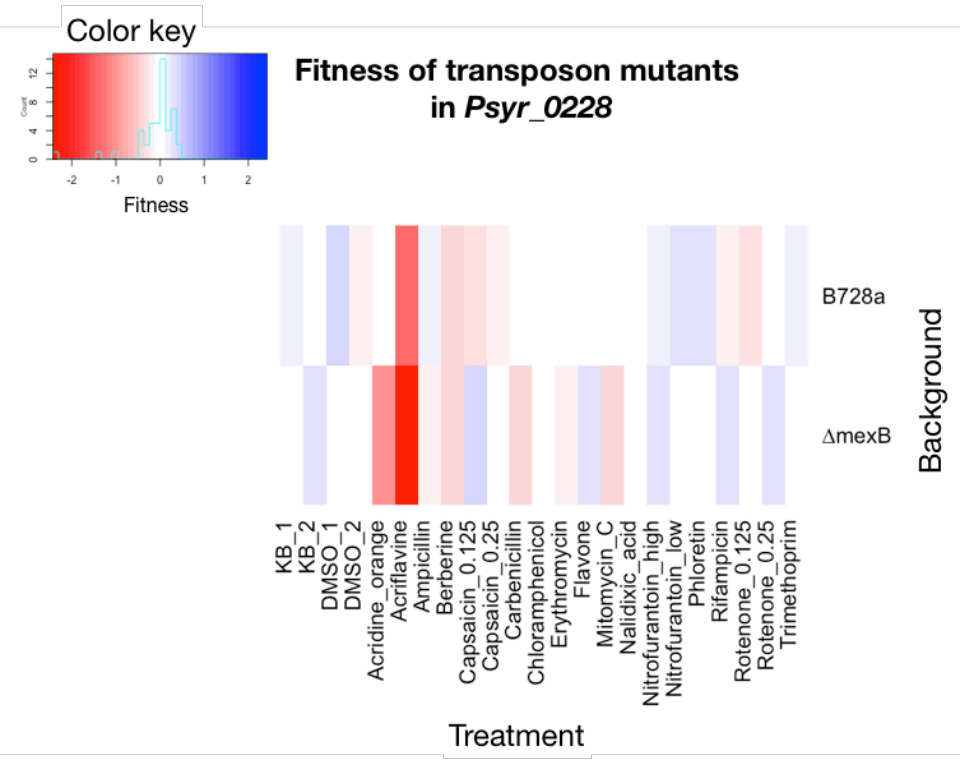


Figure 7. Fitness of transposon insertional mutants in the Major Facilitator Superfamily transporter gene *Psyr_0228* in both a WT background ad in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

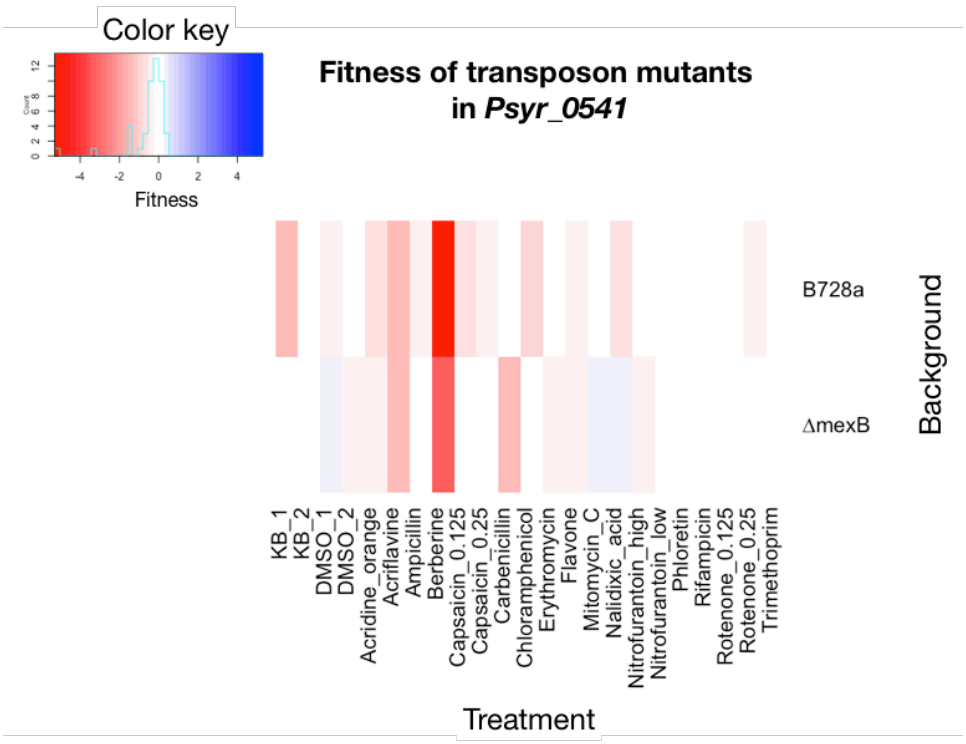


Figure 8. Fitness of transposon insertional mutants in the SMR gene *Psyr_0541* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

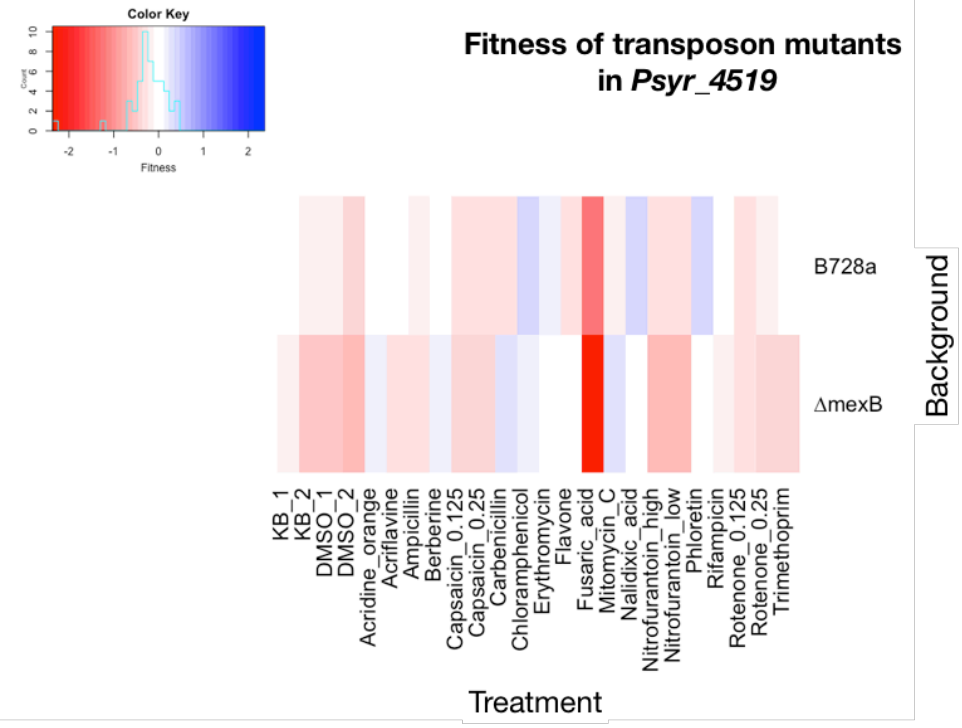


Figure 9. Fitness of transposon insertional mutants in the Major Facilitator Superfamily transporter gene *Psyr_4519* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

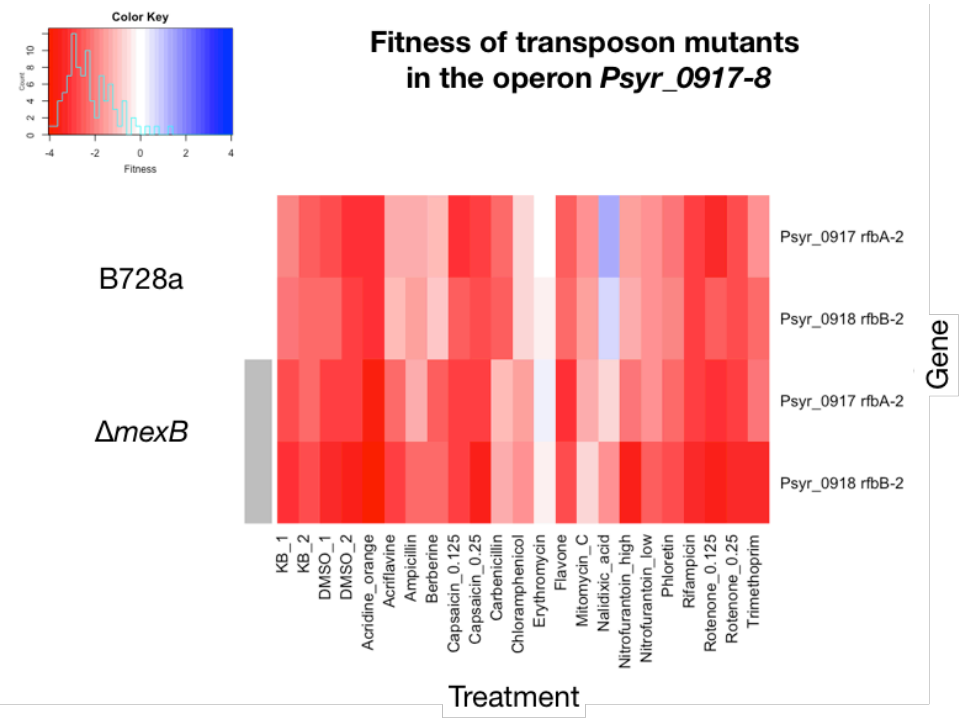


Figure 10. Fitness of transposon insertional mutants in the operon *Psyr_0917-8* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

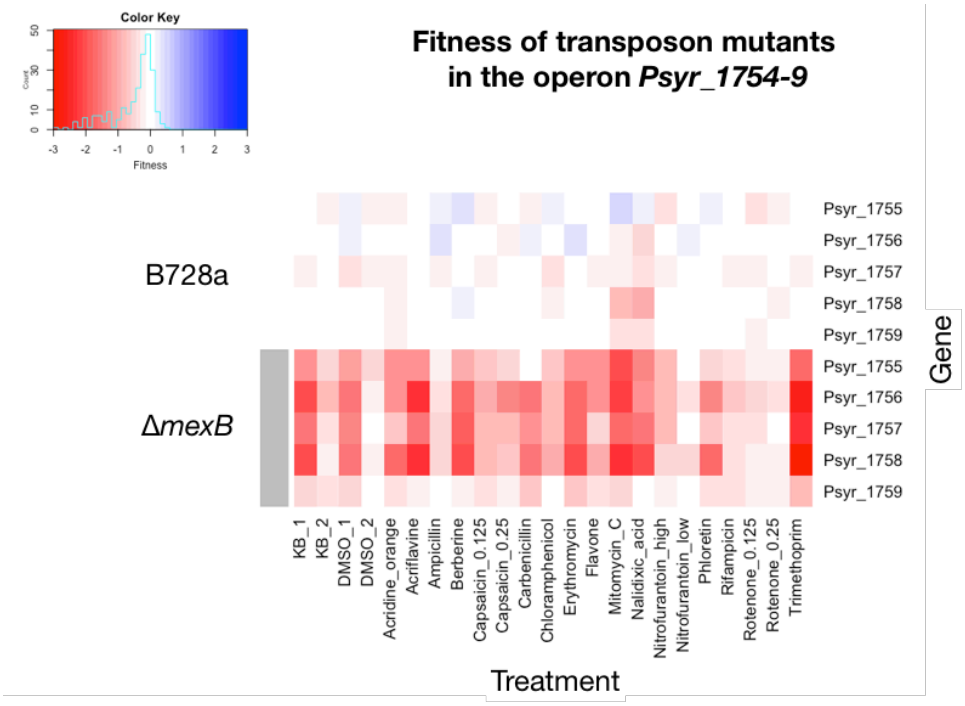


Figure 11. Fitness of transposon insertional mutants in the operon *Psyr_1754-9* in both a WT background and in cells in which *mexAB-oprM* has been disrupted. Fitness is calculated as the $\log_2$ change in relative insertion strain barcode abundance for a given gene.

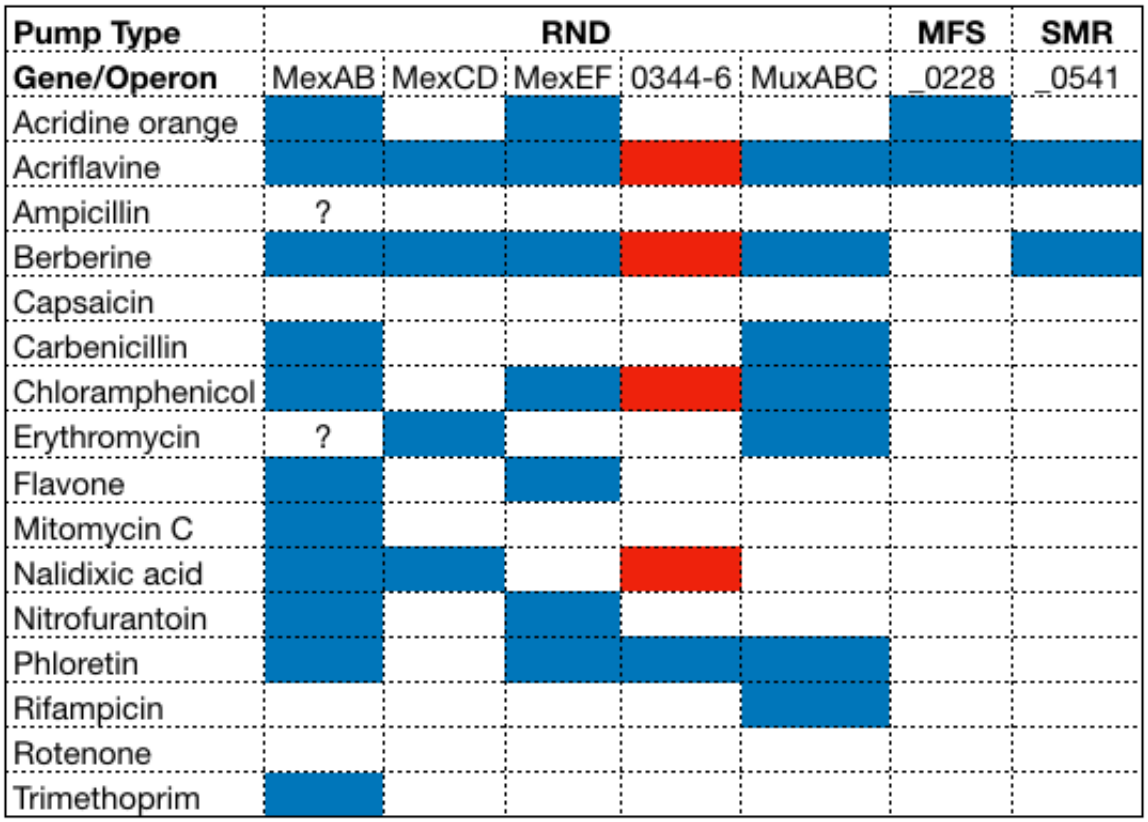
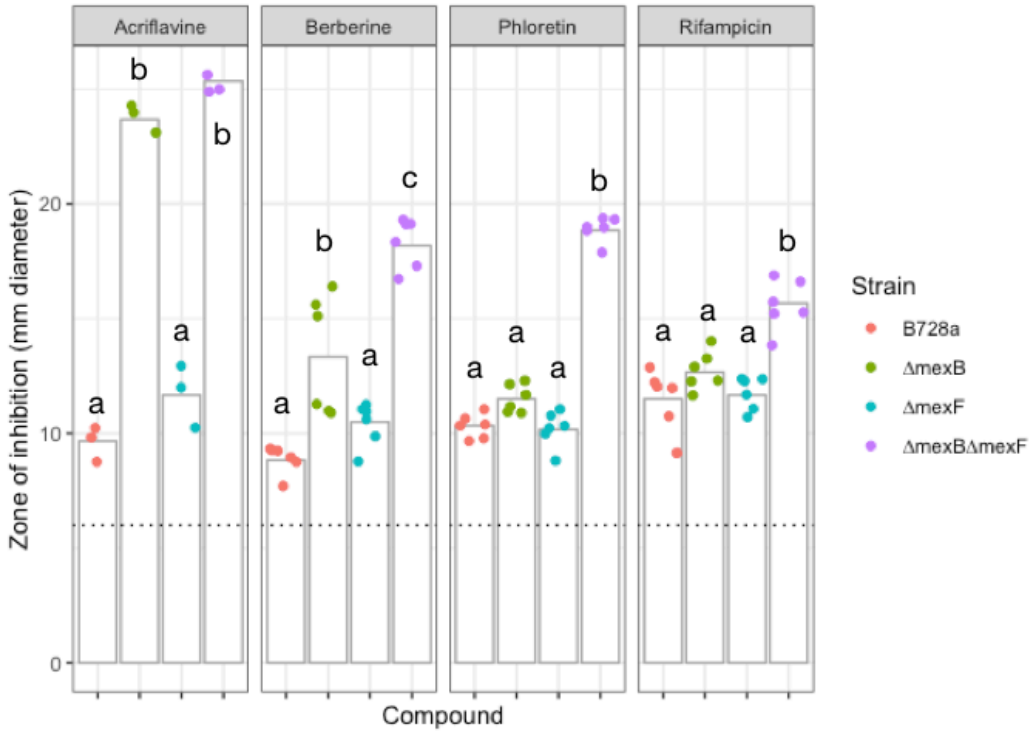
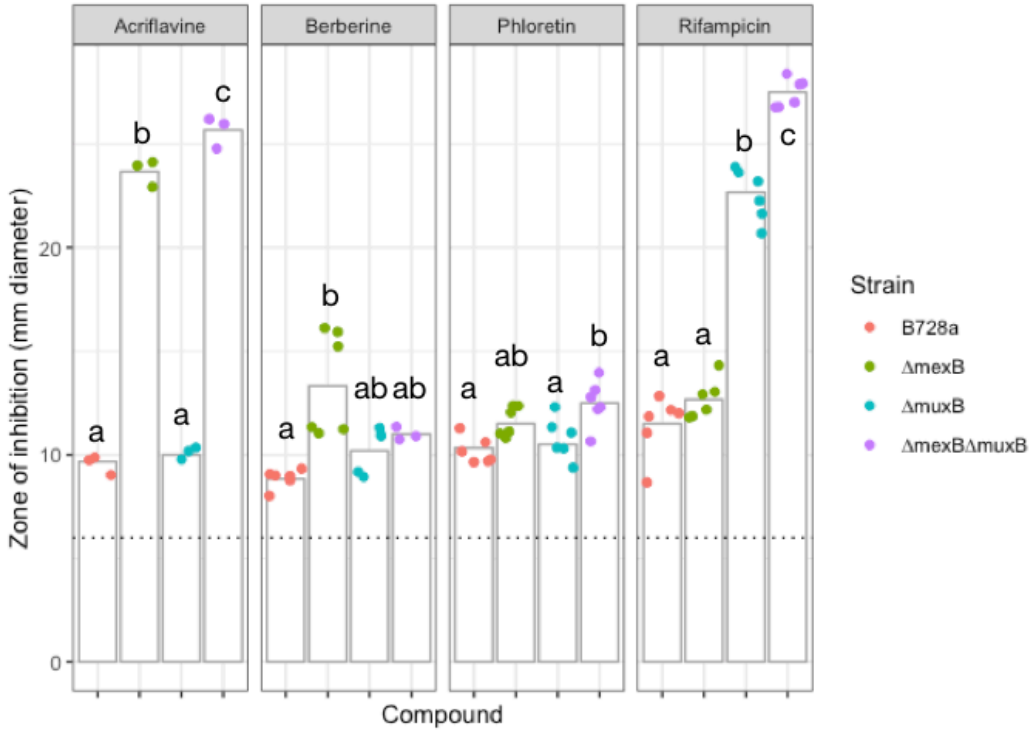


Figure 12. Likely substrates of *P. syringae* B728a multidrug resistance RND transporters, the MFS transporter Psyr_0228, and the SMR pump Psyr_0541. Blue cells indicate the compound is an apparent substrate of that transporter, with insertional mutants of a given gene having decreased fitness (fitness value less than -1) in the WT and/or the Δ*mexB* genetic backgrounds. Red cells indicate an increased fitness of insertional mutants in this operon.

A.


Acriflavine
Berberine
Phloretin
Rifampicin
Zone of inhibition (mm diameter)
0
10
20
Compound
Strain
B728a
ΔmexB
ΔmexF
ΔmexBΔmexF

B.


Acriflavine
Berberine
Phloretin
Rifampicin
Zone of inhibition (mm diameter)
0
10
20
Compound
Strain
B728a
ΔmexB
ΔmuxB
ΔmexBΔmuxB

C.

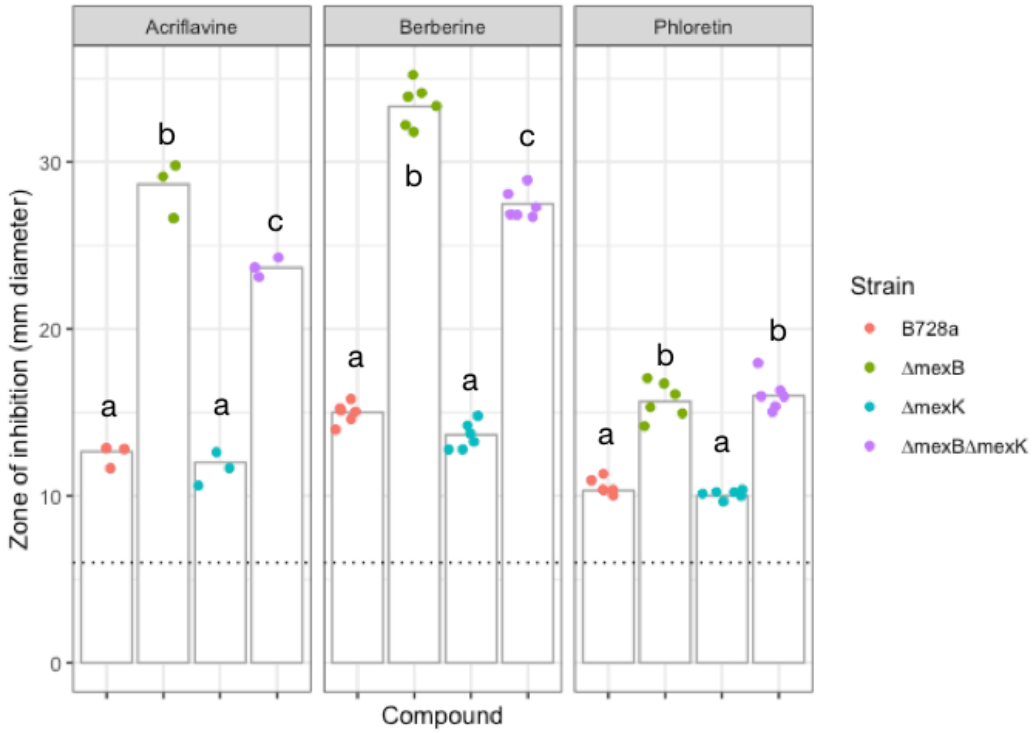
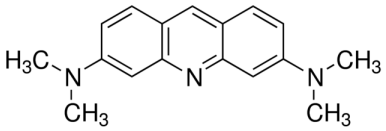


Figure 13. Zone of growth inhibition assays to test antibiotic sensitivity of B728a efflux pump deletion mutants (A) Δ*mexB*Δ*mexF*, (B) Δ*mexB*Δ*muxB*, and (C) Δ*mexB*Δ*mexK*. Mean values are displayed as column height. For each compound tested, means marked with the same letter do not differ at p = 0.01 using Tukey's HSD test. The lower limit of measurement is 6 mm (dotted line).

A.



B.

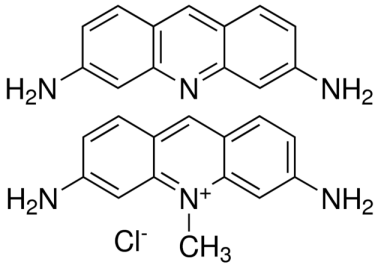


Figure 14. Structures of (A) acridine orange and (B) proflavine (top) / acriflavine (bottom). Acriflavine is typically sold as a mixture of 3,6-Diamino-10-methylacridinium chloride (acriflavine) with 3,6-diaminoacridine (proflavine) (www.sigmaaldrich.com).

Tables

Strains and plasmids

Strains		
Organism	Description	Source
E. coli TOP10	For general cloning	Invitrogen
E. coli XL1-Blue	For general cloning	QB3 Macrolab
E. coli S17-1	Conjugation donor strain	(114)
E. coli WM3064	Strain APA752; barcoded *mariner* transposon vector (Kan^R) in *E. coli* conjugation strain	(28)
P. syringae B728a	Wild type strain (Rif^R)	(60)
P. syringae B728a	Whole genome barcoded *mariner* transposon library (Rif^R Kan^R)	This work
P. syringae B728a	Whole genome barcoded *mariner* transposon library (Rif^R Kan^R), in Δ*mexB* genotype	This work
P. syringae B728a	Δ*mexB* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* (Rif^R) – cured of pFLP2Ω	This work
P. syringae B728a	p519-MexAB (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* p519-MexAB (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* p519-MexAB-OprM (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexF* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* Δ*mexF* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*muxB* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* Δ*muxB* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexK* (Rif^R Kan^R)	This work
P. syringae B728a	Δ*mexB* Δ*mexK* (Rif^R Kan^R)	This work

Plasmids			
Plasmid name	Description	Antibiotic	Reference
pKD13	Source of kanamycin resistance	Kan	(112)
pT*sacB*	Suicide plasmid to introduce DNA into *P. syringae*	Tet	(113)
pFLP2Ω	Contains Flp recombinase	Spec	(180)
p519*ngfp*	*lac* promoter and *pnpt2* in front of *gfp*	Kan	(179)
p519-MexAB	*mexAB* and upstream 1161 bp, replacing *gfp* in p519*ngfp*	Kan	This work
p519-MexAB-OprM	*mexAB-oprM* and upstream 1161 bp, replacing *gfp* in p519*ngfp*	Kan	This work
pT:2968-kan	To delete *mexF*, contains *Psyr_2968* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:2483-kan	To delete *muxB*, contains *Psyr_2483* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work
pT:0346-kan	To delete *mexK*, contains *Psyr_0346* flanking regions bordering kan^R cassette, inserted into SmaI site of pT*sacB*	Tet Kan	This work

Primers.

Primers used to construct MexAB (*Psyr_4007-8*) and MexAB-OprM (*Psyr_4007-9*) expression plasmids. Restriction cut sites are lowercase.

Name	Sequence
XbaI 4007 F	TAAtctagaTTTTCCTGCGGGGCAATACT
EcoRI 4008 R	ATgaattcGAACGCGGTAACTGCAAGTG
EcoRI 4009 R	ATgaattcCACTGATCGCAGGTGTGTTT

Bold sequence complements FRT-Kan sequence for splicing by overlap extension protocol. The primers used for TnSeq mapping and BarSeq amplification are described (28).

Name	Sequence
FRT-KanF	GTGTAGGCTGGAGCTGCTTC
FRT-KanR	ATTCCGGGGATCCGTCGACC
4008 up F	AGCACAGCCCTATACCCTGA
FRT 4008 up R	**GAAGCAGCTCCAGCCTACAC**TTACTCCCCTTTGCTGCCTG
FRT 4008 down F	**GGTCGACGGATCCCCGGAAT**TGCAGTTACCGCGTTCATTC
4008 down R	AGGAAGGTCAGGTTGCTGTC
2968 up F	GGGATGAATTCACCGGTCGT
FRT 2968 up R	**GAAGCAGCTCCAGCCTACAC**CGGACGAGTCCCTTAGCCGC
FRT 2968 down F	**GGTCGACGGATCCCCGGAAT**GAGTGCGTACAAAGTCTTCATCCCG
2968 down R	TTGTCGTAGTCGCTGAACGC
2483 up F	GCAGCTATCAAGTGGCGTTG
FRT 2483 up R	**GAAGCAGCTCCAGCCTACAC**GGTTTTTCCACCTTGCCGAG
FRT 2483 down F	**GGTCGACGGATCCCCGGAAT**GTGAGCTTTCGCTTGTTGCC
2483 down R	GAGCGGTCCATCGCGATATT
0346 up F	CAAAGTGCGCGAAGTAACCC
FRT 0346 up R	**GAAGCAGCTCCAGCCTACAC**TGACTGGCGACGGTGAATAC
FRT 0346 down F	**GGTCGACGGATCCCCGGAAT**GGCTGCTGACGCGGTAATC
0346 down R	CTGCGAAGTGGTGCAGCAAA

Primers used to confirm deletion strains.

Name	Sequence
4008 check F	CAGTGGCTGTAGCAAGAAGGA
4008 check R	TTTGTTGCGCACTGAACAGC
2968 check F	GTACTGGAGCAGCCAGTCAA
2968 check R	CCAGCAGGACCAGAAAGTCC
2483 check F	TTCCAGGAAGGGCAGATGGT
2483 check R	AGGTGCATGATCGCAAAGGT
0346 check F	ACGTTGCTCGATAACCCGAA
0346 check R	TCCGAACGTTTCGCAGAGAT

Table 1. Characteristics of the barcoded *mariner* transposon libraries in B728a WT and Δ*mexB* backgrounds. Genic strains contain insertions within the central 10 – 90% of the gene; only these insertions are used to calculate fitness. The total number of B728a genes = 5,216.

Library	**B728a WT**	**B728a Δ*mexB***
Mapped barcodes	281,419	237,285
Genic strains	169,826	143,498
Median insertions / gene	21	16
Genes with enough reads	4,296	4,255

Table 2. *P. syringae* B728a contains 668 total transporter proteins (48).

ATP-Dependent	**382 (57%)**
The ATP-binding Cassette (ABC) Superfamily	344
The Bacterial Competence-related DNA Transformation Transporter (DNA-T) Family	2
The H+- or Na+-translocating F-type, V-type and A-type ATPase (F-ATPase) Superfamily	9
The Type III (Virulence-related) Secretory Pathway (IIISP) Family	11
The Type IV (Conjugal DNA-Protein Transfer or VirB) Secretory Pathway (IVSP) Family	4
The Outer Membrane Protein Secreting Main Terminal Branch (MTB)	5
The P-type ATPase (P-ATPase) Superfamily	7

Ion Channels	**34 (5%)**
The Ammonia Transporter Channel (Amt) Family	2
The Anion Channel-forming Bestrophin (Bestrophin) Family	2
The Major Intrinsic Protein (MIP) Family	1
The CorA Metal Ion Transporter (MIT) Family	4
The H+- or Na+-translocating Bacterial Flagellar Motor 1ExbBD Outer Membrane Transport Energizer	13
The Large Conductance Mechanosensitive Ion Channel (MscL) Family	1
The Small Conductance Mechanosensitive Ion Channel (MscS) Family	9
The Homotrimeric Cation Channel (TRIC) Family	1
The Voltage-gated Ion Channel (VIC) Superfamily	1

Phosphotransferase System (PTS)	**5 (1%)**
General PTS	3
Sugar Specific PTS	2

Secondary Transporter	**231 (35%)**
The Arsenical Resistance-3 (ACR3) Family	1
The Auxin Efflux Carrier (AEC) Family	1
The Alanine or Glycine:Cation Symporter (AGCS) Family	1
The Autoinducer-2 Exporter (AI-2E) Family (Formerly the PerM Family, TC #9.B.22)	5
The Amino Acid-Polyamine-Organocation (APC) Family	6
The Aromatic Acid Exporter (ArAE) Family	4
The Arsenite-Antimonite (ArsB) Efflux Family	2
The Short Chain Fatty Acid Uptake (AtoE) Family	1
The Bile Acid:Na+ Symporter (BASS) Family	1
The Betaine/Carnitine/Choline Transporter (BCCT) Family	1
The Benzoate:H+ Symporter (BenE) Family	1
The Ca2+:Cation Antiporter (CaCA) Family	2
The Cation Diffusion Facilitator (CDF) Family	3
The Citrate-Mg2+:H+ (CitM) Citrate-Ca2+:H+ (CitH) Symporter (CitMHS) Family	2
The Chloride Carrier/Channel (ClC) Family	1
The Monovalent Cation:Proton Antiporter-1 (CPA1) Family	3
The Monovalent Cation:Proton Antiporter-2 (CPA2) Family	2
The Monovalent Cation (K+ or Na+):Proton Antiporter-3 (CPA3) Family	7
The Dicarboxylate/Amino Acid:Cation (Na+ or H+) Symporter (DAACS) Family	2
The Drug/Metabolite Transporter (DMT) Superfamily	17
The Formate-Nitrite Transporter (FNT) Family	2
The Gluconate:H+ Symporter (GntP) Family	4

The Glycerol Uptake (GUP) Family	1
The Hydroxy/Aromatic Amino Acid Permease (HAAAP) Family	1
The K+ Uptake Permease (KUP) Family	1
The Branched Chain Amino Acid:Cation Symporter (LIVCS) Family	1
The L-Lysine Exporter (LysE) Family	1
The Major Facilitator Superfamily (MFS)	72
The Multidrug/Oligosaccharidyl-lipid/Polysaccharide (MOP) Flippase Superfamily	6
The Malonate:Na+ Symporter (MSS) Family	2
The Nucleobase:Cation Symporter-1 (NCS1) Family	3
The Nucleobase:Cation Symporter-2 (NCS2) Family	3
The NhaA Na+:H+ Antiporter (NhaA) Family	2
The Ni2+-Co2+ Transporter (NiCoT) Family	1
The Metal Ion (Mn2+-iron) Transporter (Nramp) Family	3
The Neurotransmitter:Sodium Symporter (NSS) Family	1
The Oligopeptide Transporter (OPT) Family	1
The Cytochrome Oxidase Biogenesis (Oxa1) Family	1
The Inorganic Phosphate Transporter (PiT) Family	2
The Phosphate:Na+ Symporter (PNaS) Family	1
The Reduced Folate Carrier (RFC) Family	1
The Resistance to Homoserine/Threonine (RhtB) Family	10
The Resistance-Nodulation-Cell Division (RND) Superfamily	16
The Solute:Sodium Symporter (SSS) Family	14
The Sulfate Permease (SulP) Family	4
The Twin Arginine Targeting (Tat) Family	3
The Telurite-resistance/Dicarboxylate Transporter (TDT) Family	2
The Threonine/Serine Exporter (ThrE) Family	1
The Tripartite ATP-independent Periplasmic Transporter (TRAP-T) Family	3
The K+ Transporter (Trk) Family	1
The Tricarboxylate Transporter (TTT) Family	3
The Zinc (Zn2+)-Iron (Fe2+) Permease (ZIP) Family	1

Unclassified	**12 (2%)**
The HlyC/CorC (HCC) Family	2
The Iron/Lead Transporter (ILT) Superfamily	4
The Mg2+ Transporter-E (MgtE) Family	1
The Tellurium Ion Resistance (TerC) Family	3
The Unknown BART Superfamily-1 (UBS1) Family	1
The YggT or Fanciful K+ Uptake-B (FkuB; YggT) Family	1

Table 3. Antimicrobial compounds used in growth assays of the barcoded *mariner* transposon libraries. For compounds with known (and different) MIC values for the WT and Δ*mexB* genotypes, ¼ MIC values were used. Otherwise, the same concentration was used for both libraries. DMSO controls were used at concentrations representative of the highest solvent concentration used.

Compound	Solvent	µg/ml for B728a	µg/ml for Δ*mexB*
DMSO_1		1%	0.42%
DMSO_2		1.25%	1.25%
Acridine orange	H_20	62.5	3.125
Acriflavine	H_20	7.75	1.25
Ampicillin	H_20	15.6	0.775
Berberine	DMSO	250	15.6
Capsaicin	DMSO	250	250
Capsaicin	DMSO	125	125
Carbenicillin	H_20	250	3.1
Chloramphenicol	EtOH	12.5	1.55
Erythromycin	EtOH	25	6.25
Flavone	DMSO	125	125
Fusaric acid	DMSO	250	15.6
Mitomycin C	H_20	0.625	0.05
Nalidixic acid	H_20	12.5	0.775
Nitrofurantoin	DMSO	250	62.5
Nitrofurantoin	DMSO	125	31.25
Phloretin	DMSO	250	125
Rifampicin	MeOH	100	100
Rotenone	DMSO	250	250
Rotenone	DMSO	125	125
Trimethoprim	H_20	125	15.6

Table 4. RND operons homologous to *mexAB-oprM* that likely contribute to multidrug resistance. The number of experimental conditions where they have a significant contribution to fitness (fitness value less than -1) is shown for mutants in a WT and Δ*mexB* background. A total of 23 treatments were examined, including 2 KB controls, 2 DMSO controls, and 16 unique compounds (3 at two concentrations).

Locus	**Gene**	**Description**	**B728a**	**Δ*mexB***
Psyr_4007	mexA	Secretion protein HlyD	**10**	0
Psyr_4008	mexB	Hydrophobe/amphiphile efflux-1 HAE1	**10**	*NA*
Psyr_4009	oprM	RND efflux system, outer membrane lipoprotein, NodT	**13**	*NA*
Psyr_2967	mexE	Secretion protein HlyD	**1**	**9**
Psyr_2968	mexF	Hydrophobe/amphiphile efflux-1 HAE1	**1**	**8**
Psyr_2969	oprN	RND efflux system, outer membrane lipoprotein, NodT	0	**6**
Psyr_2482	muxA	Secretion protein HlyD	**1**	**8**
Psyr_2483	muxB	Acriflavin resistance protein	**1**	**5**
Psyr_2484	muxC	Acriflavin resistance protein	**1**	**5**
Psyr_2485	opmB	RND efflux system, outer membrane lipoprotein, NodT	**1**	**6**
Psyr_2283	mexC	Secretion protein HlyD	**4**	**2**
Psyr_2282	mexD	Hydrophobe/amphiphile efflux-1 HAE1	**4**	**2**
		(No outer membrane protein in operon.)		
Psyr_0344		Secretion protein HlyD	0	**1**
Psyr_0345		Secretion protein HlyD	0	**1**
Psyr_0346	mexK	Acriflavin resistance protein	0	**1**
		(No outer membrane protein in operon.)		
Psyr_2193		Secretion protein HlyD	0	0
Psyr_2194		Acriflavin resistance protein	0	0
		(No outer membrane protein in operon.)		
Psyr_2864		RND efflux system, outer membrane lipoprotein, NodT	0	0
Psyr_2865		Hydrophobe/amphiphile efflux-1 HAE1	0	0
Psyr_2866		Secretion protein HlyD	**1**	0
Psyr_3130		Secretion protein HlyD	0	0
Psyr_3131		Secretion protein HlyD	0	0
Psyr_3132		Acriflavin resistance protein	0	0
		(No outer membrane protein in operon.)		

Table 5. Shared amino acid identity between *P. syringae* B728a Psyr_2482-5 and *P. aeruginosa* PAO1 MuxABC-OpmB. Amino acid sequence homology was calculated using NCBI blastp. “Positives” includes amino acids having similar chemical properties.

P. syringae B728a	*P. aeruginosa* PAO1	Gene	A.A. identity	Positives
Psyr_2482	PA2528	*muxA*	219/420 (52%)	279/420 (66%)
Psyr_2483	PA2527	*muxB*	731/1022 (72%)	875/1022 (85%)
Psyr_2484	PA2526	*muxC*	611/1039 (59%)	777/1039 (74%)
Psyr_2485	PA2525	*opmB*	288/459 (63%)	350/459 (76%)

Chapter 6

Discussion

Pseudomonas syringae is a valuable model organism for many reasons, including its economic importance to agriculture, decades of existing research, culturability, genetic tractability, and natural diversity. Its ubiquity in various natural habitats lends significance to these studies – this is a bacterial species where findings can have substantial real world impact. In this dissertation, I examine those *P. syringae* traits that are necessary for its colonization of the phyllosphere. Bacterial growth on or in a plant host reflects a complex multifaceted relationship, where one must consider factors such as the biology of both the plant and potential pathogen, local environmental conditions, and temporal issues in the interaction – leading from colonization to infection. This concept is typically described as the "disease triangle" (58). The majority of the work described herein focuses on traits that contribute to bacterial fitness in one or more plant-associated habitats (the leaf surface and apoplast) or host species. In addition, I used a hyper-susceptible efflux pump mutant to characterize variation between susceptible hosts in plant content of preformed phytoanticipins as well as induced antimicrobial compounds known as phytoalexins. This mutant also enabled me to functionally annotate a variety of other genes contributing to tolerance of toxic compounds *in vitro*. While this particular transporter, when disrupted, had been seen to have a major role in bacterial virulence, we did not find this to be the case in all of the host plant species tested we tested. This indicated that there is wide variation in plant chemical defenses, and therefore a range of needed pathogen compatibility with the toxins associated with these susceptible host plants. I used a genome-wide *in vitro* screen to assess the importance of alternative and redundant efflux pumps. The conclusions drawn from bacterial tolerance of highly specific chemical stresses *in vitro* provide the groundwork to expand *in planta* studies to better understand the diverse chemical stresses encountered in these environments. As *P. syringae* strain B728a is capable of colonizing diverse host plants in addition to growth in other habitats, these genome-wide studies should provide abundant data with which to design and perform additional experiments.

Throughout this dissertation, I spent a considerable amount of effort describing the development and application of randomly-barcoded transposon sequencing (RB-TnSeq). While new techniques and methodologies can be considered a means to an end – a way to ask and address new scientific questions – it is hard not to be excited about the ability to delve into the unknown in new ways. In this case, while transposon-based studies have been performed for years in numerous biological systems, RB-TnSeq allowed me to expand this dissertation beyond what I could have imagined. We were able to generate genome-wide fitness data across multiple conditions, as well as replicate these experiments to control for biological and environmental variability. The most important strength of RB-TnSeq relevant here is its scalability – the ability to replicate studies and compare them across related growth conditions provides tremendous statistical and analytical power to support conclusions of the role of these genes in a given process (28).

P. syringae strain B728a is capable of growth in diverse crop and native plant species (38). I measured the growth of this strain in known and unknown host plant species to guide the genome-wide screens performed for this dissertation. Prior to the massive RB-TnSeq inoculations, I identified the plant species that would most likely lead to success for these experiments. For example, we determined that within even a single species such as common bean (*Phaseolus vulgaris*), different varieties sustained a wide range of bacterial colonization. These studies are detailed in Chapter 3, but really provided the foundation upon which we designed all subsequent *in planta* RB-TnSeq studies. Additionally, we observed that closely related plant species such as common bean and lima bean (*P. lunatus*) differed greatly in their overall susceptibility to infection and support of *P. syringae* populations, as well as in their chemical defenses, determined through the differential growth of a hyper-susceptible bacterial mutant strain. This observation motived an untargeted metabolomics study that was not included in this dissertation. The primary challenge encountered with untargeted metabolomics was the extremely limited ability to identify unknown metabolites using existing reference databases, particularly in complex mixtures such as the leaf. As there were numerous differentially abundant metabolites seen in different plant species, and many changes upon infection, it simply was not feasible to associate particular metabolites responsible for this range in host susceptibility. Individual compounds having characteristic masses were not reliably identified. Even the hypothesis that plant species that seem to have a more robust chemical defense than others would harbor a larger array of different chemicals could not be easily supported given the very large number of compounds we encountered and the large variation in their apparent abundance. As available databases of diverse molecules expand (such as has been seen with databases of DNA sequences), the analytical power of metabolomics studies will greatly improve and it might eventually be possible to associate particular chemicals with the resistance to *P. syringae* colonization. Given the current intractability of any of the host plants that we investigated here to genetic manipulation, the unambiguous association of any compounds with resistant to colonization may prove elusive. Relying on the hyper-susceptible mutant strain however, I could identify a subset of host plants where chemical defenses appeared to contribute significantly to the limitation of bacterial growth. I also clarified the timing during infection where bacterial efflux becomes most important, expanding upon previous work generally highlighting the necessity of these transporter proteins *in planta*. It was heartening to find that the time of onset of apparent mechanisms to suppress bacterial growth was consistent with the known timing of phytoalexin induction in common bean. Future studies could more specifically characterize metabolic differences between host plants, which may indicate why the primary efflux transporter was dispensable in several hosts. It might also be possible to exploit our findings of temporal onsets of apparent resistance to better interpret the complex metabolome in these plants during infection. It would be interesting if alternative transporters were preferentially used to tolerate plant-produced antimicrobial compounds; such knowledge might have important translational applications.

Based on the robust growth of *P. syringae* strain B728a in the hosts common bean, lima bean, and pepper (*Capsicum annuum*), we determined that these plant

species would be informative hosts in which to test the RB-TnSeq library (Chapters 2 and 4). Considering two species within the same genus (*Phaseolus*), and another in a different plant family (Solanaceae), we expected to find that a subset of virulence or plant colonization genes would be host specific. The topic of phytopathogen host-specificity is an interesting one, which is not specifically addressed in this work. Much of our current understanding on this topic has come from a focus of a tremendous amount of work that has examined the taxonomic distribution of type III effector proteins, and their function and subsequent recognition by the host. A relatively simple example was the discovery that a mutant of *P. syringae* strain DC3000 lacking the single effector HopQ1-1 is able to cause disease on the otherwise non-host *Nicotiana benthamiana* (181). As expected, the type III secretion system significantly contributed to bacterial growth in all three plant species tested here. This suggested that the heterogeneous mutant population was inoculated into the apoplast at a low enough concentration to avoid complementation by nearby cells containing the wild type allele at a given locus. However, using RB-TnSeq, I also identified a wide variety of "compatibility genes" – genes contributing to fitness in the plant habitat. These important traits included the production of exopolysaccharide such as alginate, the metabolic capability to grow in the nutrient and cofactor-limiting host, and the production of other putative virulence factors. Most interestingly, some of these very general compatibility traits contributed to bacterial fitness differentially between these host plants, and this deserves to be examined further. That is, the fundamental chemical makeup of different plant species differentially imposes catabolic and anabolic processes on the pathogen to maximize its fitness in the particular host. The direct measurements of bacterial fitness mediated by a given gene that were measured here nicely complement previous studies and provide an expanded view of important bacterial traits for life in the phyllosphere.

We foresee that this richness of publically available information on the fitness contribution of genes will be a powerful tool to expand upon existing knowledge in several ways. First, fitness values generated for non-essential genes in these complex host environments can be compared to *in vitro* work to identify and better annotate genes that are important in those habitats. Second, measurements of gene expression via transcript abundance (RNAseq) are common in many fields of biology, yet have several glaring issues (72). Most relevant to this dissertation, and which is discussed in some detail, is the observation that global gene expression is not correlated with fitness. Specifically, genes with large contributions to fitness are often not highly expressed or induced in a given habitat. Third, in this dissertation I identified numerous hypothetical or otherwise uncharacterized genes that contributed significantly to fitness in the environments examined. However, only a few were validated and I was not able to determine molecular mechanisms for function. Follow-up studies should be able to validate these findings and identify novel genes for environmental growth. Indeed, RB-TnSeq is particularly powerful in assigning functions to genes - especially if *in vitro* assays can be employed. I see a great benefit in the further functional annotation of environmentally important bacteria such as *P. syringae*.

I only began to explore the possibilities of testing the two transposon libraries created here. Comparative studies between insertional mutants in the wild type strain

and the mutant lacking a functional MexAB-OprM efflux pump system revealed important functions for some partially redundant homologous transporters. The relatively limited number of antimicrobial compounds tested to date somewhat limited the analysis from these experiments. However, the power of these assays (Chapter 5) derived from the direct comparisons that could be made between these isogenic transposon libraries. I would expect that interrogation of the fitness contribution of these genes to a large collection of naturally occurring antimicrobial compounds would provide real insight into the linkage of genes in *P. syringae* to the chemical ecology of the habitats in which it exists. As the technology to perform TnSeq screens improves and becomes cheaper, there will be much to learn from mutagenesis studies in non-wild type backgrounds.

To expand upon the studies described herein, we also hope to assay gene contribution to fitness in additional parts of the *P. syringae* life cycle. For example, aphids commonly found on the leaf surface can function as both a host and a vector for *P. syringae* (182). This infection results in rapid aphid mortality, but enhances pathogen dispersal throughout the phyllosphere. We are interested to see how the bacterial traits necessary for this interaction compare to those related to epiphytic and endophytic growth. The bacterial consumption of aphid excretions (honeydew) represents a potentially important exogenous source of nutrition on the leaf surface (183), that may require distinct catabolic genes than those that we observed in our greenhouse studies that were done in the absence of such aphid-provided resources.

As with any good genome-wide screen, these experiments generated more questions than answers. It would be valuable to expand the examination of the breadth of environmentally representative and stress-specific conditions experienced by *P. syringae* by exploiting these transposon libraries. It would also be valuable to gain more depth in our understanding of its lifestyle by more focused study of the individual bacterial genes identified as significant here. Improving our understanding of genes that contribute to bacterial fitness in diverse environments, including the phyllosphere, will also improve our understanding of those habitats, at the scale at which bacterial cells interact with the plant. Furthermore, these broad findings on bacterial colonization of plants can be applied to counter diseases of crop plants through targeted plant breeding, field management, or chemical or biological control.

References

1. Berendsen RL, Pieterse CMJ, Bakker PAHM. 2012. The rhizosphere microbiome and plant health. Trends Plant Sci 17:478–486.
2. Vorholt JA. 2012. Microbial life in the phyllosphere. Nat Rev Microbiol 10:828–840.
3. Zilber-Rosenberg I, Rosenberg E. 2008. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. FEMS Microbiol Rev 32:723–735.
4. Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A. 2015. The importance of the microbiome of the plant holobiont. New Phytol 206:1196–1206.
5. Partida-Martinez L, Heil M. 2011. The microbe-free plant: fact or artifact? Front Plant Sci 2:1–16.
6. Lindow SE, Brandl MT. 2003. Microbiology of the phyllosphere. Appl Environ Microbiol 69:1875–1883.
7. Mysore KS, Ryu CM. 2004. Nonhost resistance: How much do we know? Trends Plant Sci 9:97–104.
8. Leveau JHJ, Lindow SE. 2001. Appetite of an epiphyte: Quantitative monitoring of bacterial sugar consumption in the phyllosphere. Proc Natl Acad Sci 98:3446–3453.
9. Van Der Wal A, Leveau JHJ. 2011. Modelling sugar diffusion across plant leaf cuticles: the effect of free water on substrate availability to phyllosphere bacteria. Environ Microbiol 13:792–797.
10. Monier JM, Lindow SE. 2004. Frequency, size, and localization of bacterial aggregates on bean leaf surfaces. Appl Environ Microbiol 70:346–355.
11. Mercier J, Lindow SE. 2001. Field performance of antagonistic bacteria identified in a novel laboratory assay for biological control of fire blight of pear. Biol Control 22:66–71.
12. Wilson M, Lindow SE. 1994. Ecological similarity and coexistance of epiphytic ice-nucleating (Ice+) *Pseudomonas syringae* strains and a non-ice-nucleating (Ice-) biological control agent. Appl Enviromental Microbiol 60:3128–3137.
13. Hirano SS, Upper CD. 2000. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte. Microbiol Mol Biol Rev 64:624–653.
14. Lindow SE, Leveau JHJ. 2002. Phyllosphere microbiology. Curr Opin Biotechnol 13:238–243.
15. Lindemann J, Arny DC, Upper CD. 1984. Epiphytic populations of *Pseudomonas syringae* pv. *syringae* on snap bean and nonhost plants and the incidence of bacterial brown spot disease in relation to cropping patterns. Ecol Epidemiol 74:1329–33.
16. Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, Yan S, Dominguez H, Thompson BM. 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J 2:321–334.
17. Melotto M, Underwood W, He SY. 2008. Role of stomata in plant innate immunity and foliar bacterial diseases. Annu Rev Phytopathol 46:101–122.
18. Quiñones B, Dulla G, Lindow SE. 2005. Quorum sensing regulates

exopolysaccharide production, motility, and virulence in *Pseudomonas syringae*. Mol Plant-Microbe Interact 18:682–693.
19. Burch AY, Zeisler V, Yokota K, Schreiber L, Lindow SE. 2014. The hygroscopic biosurfactant syringafactin produced by *Pseudomonas syringae* enhances fitness on leaf surfaces during fluctuating humidity. Environ Microbiol 16:2086–2098.
20. Xin XF, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, Chang JH, Zipfel C, He SY. 2016. Bacteria establish an aqueous living space in plants crucial for virulence. Nature 539:524–529.
21. Aung K, Jiang Y, He SY. 2018. The role of water in plant–microbe interactions. Plant J 93:771–780.
22. Leben C. 1981. How plant-pathogenic bacteria survive. Plant Dis 65:633–637.
23. Van Sluys MA, Monteiro-Vitorello CB, Camargo LEA, Menck CFM, da Silva ACR, Ferro JA, Oliveira MC, Setubal JC, Kitajima JP, Simpson AJ. 2002. Comparative genomic analysis of plant-associated bacteria. Annu Rev Phytopathol 40:169–189.
24. Levy A, Salas Gonzalez I, Mittelviefhaus M, Clingenpeel S, Herrera Paredes S, Miao J, Wang K, Devescovi G, Stillman K, Monteiro F, Rangel Alvarez B, Lundberg DS, Lu TY, Lebeis S, Jin Z, McDonald M, Klein AP, Feltcher ME, Rio TG, Grant SR, Doty SL, Ley RE, Zhao B, Venturi V, Pelletier DA, Vorholt JA, Tringe SG, Woyke T, Dangl JL. 2018. Genomic features of bacterial adaptation to plants. Nat Genet 50:138–150.
25. Melnyk RA, Hossain SS, Haney CH. 2019. Convergent gain and loss of genomic islands drives lifestyle changes in plant-associated *Pseudomonas*. ISME J 13:1575–1588.
26. van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 6:767–772.
27. Judson N, Mekalanos JJ. 2000. Transposon-based approaches to identify essential bacterial genes. Trends Microbiol 8:521–6.
28. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, Blow MJ, Bristow J, Butland G, Arkin AP. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. MBio 6:1–15.
29. Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, Kuehl J V., Melnyk RA, Lamson JS, Suh Y, Carlson HK, Esquivel Z, Sadeeshkumar H, Chakraborty R, Zane GM, Rubin BE, Wall JD, Visel A, Bristow J, Blow MJ, Arkin AP, Deutschbauer AM. 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 557:503–509.
30. Royet K, Parisot N, Rodrigue A, Gueguen E, Condemine G. 2019. Identification by Tn-seq of *Dickeya dadantii* genes required for survival in chicory plants. Mol Plant Pathol 20:287–306.
31. Duong DA, Jensen R V., Stevens AM. 2018. Discovery of *Pantoea stewartii* ssp. *stewartii* genes important for survival in corn xylem through a Tn-Seq analysis. Mol Plant Pathol 19:1929–1941.
32. Liu Z, Beskrovnaya P, Melnyk RA, Hossain SS, Khorasani S, O'Sullivan LR, Wiesmann CL, Bush J, Richard JD, Haney CH. 2018. A genome-wide screen identifies genes in rhizosphere-associated *Pseudomonas* required to evade plant

defenses. MBio 9:e00433-18.
33. Cole BJ, Feltcher ME, Waters RJ, Wetmore KM, Mucyn TS, Ryan EM, Wang G, Ul-Hasan S, McDonald M, Yoshikuni Y, Malmstrom RR, Deutschbauer AM, Dangl JL, Visel A. 2017. Genome-wide identification of bacterial plant colonization genes. PLoS Biol 15:1–24.
34. Lindeberg M, Myers CR, Collmer A, Schneider DJ. 2008. Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. Mol Plant-Microbe Interact 21:685–700.
35. Nomura K, Melotto M, He SY. 2005. Suppression of host defense in compatible plant-*Pseudomonas syringae* interactions. Curr Opin Plant Biol 8:361–368.
36. O'Leary BM, Neale HC, Geilfus CM, Jackson RW, Arnold DL, Preston GM. 2016. Early changes in apoplast composition associated with defence and disease in interactions between *Phaseolus vulgaris* and the halo blight pathogen *Pseudomonas syringae* pv. *phaseolicola*. Plant Cell Environ 39:2172–2184.
37. Vinatzer BA, Teitzel GM, Lee MW, Jelenska J, Hotton S, Fairfax K, Jenrette J, Greenberg JT. 2006. The type III effector repertoire of *Pseudomonas syringae* pv. *syringae* B728a and its role in survival and disease on host and non-host plants. Mol Microbiol 62:26–44.
38. Morris CE, Lamichhane JR, Nikolić I, Stanković S, Moury B. 2019. The overlapping continuum of host range among strains in the *Pseudomonas syringae* complex. Phytopathol Res 1:1–16.
39. Stoitsova SO, Braun Y, Ullrich MS, Weingart H. 2008. Characterization of the RND-type multidrug efflux pump MexAB-OprM of the plant pathogen *Pseudomonas syringae*. Appl Environ Microbiol 74:3387–3393.
40. Piddock LJ V. 2006. Multidrug-resistance efflux pumps - not just for resistance. Nat Rev Microbiol 4:629–636.
41. Lindeberg M, Cunnac S, Collmer A. 2009. The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. Mol Plant Pathol 10:767–775.
42. Bender CL, Alarcon-Chaidez F, Gross DC. 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol Rev 63:266–292.
43. Hulin MT, Armitage AD, Vicente JG, Holub EB, Baxter L, Bates HJ, Mansfield JW, Jackson RW, Harrison RJ. 2018. Comparative genomics of *Pseudomonas syringae* reveals convergent gene gain and loss associated with specialization onto cherry (*Prunus avium*). New Phytol 219:672–696.
44. Blanco P, Hernando-Amado S, Reales-Calderon J, Corona F, Lira F, Alcalde-Rico M, Bernardini A, Sanchez M, Martinez J. 2016. Bacterial multidrug efflux pumps: Much more than antibiotic resistance determinants. Microorganisms 4:14.
45. Paulsen IT, Brown MH, Skurray RA. 1996. Proton-dependent multidrug efflux systems. Microbiol Rev 60:575–608.
46. Paulsen IT. 2003. Multidrug efflux pumps and resistance: regulation and evolution. Curr Opin Microbiol 6:446–451.
47. Saier MH, Paulsen IT, Sliwinski MK, Pao SS, Skurray RA, Nikaido H. 1998. Evolutionary origins of multidrug and drug-specific efflux pumps in bacteria. FASEB J 12:265–274.

48. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: A database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Res 45:D320–D324.
49. Yu X, Lund SP, Scott RA, Greenwald JW, Records AH, Nettleton D, Lindow SE, Gross DC, Beattie GA. 2013. Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. Proc Natl Acad Sci 110:E425–E434.
50. Sulavik MC, Houseweart C, Cramer C, Jiwani N, Murgolo N, Greene J, DiDomenico B, Shaw KJ, Miller G, Hare R, Shimer G. 2001. Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes. Antimicrob Agents Chemother 45:1126–1136.
51. Zgurskaya HI, Nikaido H. 2000. Multidrug resistance mechanisms: Drug efflux across two membranes. Mol Microbiol 37:219–225.
52. Tal N, Schuldiner S. 2009. A coordinated network of transporters with overlapping specificities provides a robust survival strategy. Proc Natl Acad Sci 106:9051–9056.
53. Lee A, Mao W, Warren MS, Mistry A, Hoshino K, Okumura R, Ishida H, Lomovskaya O. 2000. Interplay between efflux pumps may provide either additive or multiplicative effects on drug resistance. J Bacteriol 182:3142–3150.
54. Lindemann J, Arny DC, Upper CD. 1984. Use of an apparent infection threshold population of *Pseudomonas syringae* to predict incidence and severity of brown spot of bean. Phytopathology 74:1334–1339.
55. Hirano SS, Upper CD. 1983. Ecology and epidemiology of foliar bacterial plant pathogens. Annu Rev Phytopathol 21:243–269.
56. Mercier J, Lindow SE. 2000. Role of leaf surface sugars in colonization of plants by bacterial epiphytes. Appl Environ Microbiol 66:369–374.
57. Beattie GA. 2011. Water relations in the interaction of foliar bacterial pathogens with plants. Annu Rev Phytopathol 49:533–555.
58. Xin XF, Kvitko BH, He SY. 2018. *Pseudomonas syringae*: what it takes to be a pathogen. Nat Rev Microbiol 16:316–328.
59. Monteil CL, Bardin M, Morris CE. 2014. Features of air masses associated with the deposition of *Pseudomonas syringae* and *Botrytis cinerea* by rain and snowfall. ISME J 8:2290–2304.
60. Loper JE, Lindow SE. 1987. Lack of evidence for *in situ* fluorescent pigment production by *Pseudomonas syringae* pv. *syringae* on bean leaf surfaces. Phytopathology 77:1449–1454.
61. Baltrus DA, Mccann HC, Guttman DS. 2017. Evolution, genomics and epidemiology of *Pseudomonas syringae*. Mol Plant Pathol 18:152–168.
62. Morris CE, Sands DC, Vanneste JL, Montarray J, Oakley B, Guilbaud C, Glaux C. 2010. Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. MBio 1:1–11.
63. Lindow SE. 1983. The role of bacterial ice nucleation in frost injury to plants. Ann Rev Phytopathol 21:363–384.
64. Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL. 2011. Dynamic evolution of

pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog 7:e1002132.
65. Morris CE, Monteil CL, Berge O. 2013. The life history of *Pseudomonas syringae*: Linking agriculture to Earth system processes. Annu Rev Phytopathol 51:85–104.
66. Gal M, Preston GM, Massey RC, Spiers AJ, Rainey PB. 2003. Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. Mol Ecol 12:3109–3121.
67. Nobori T, Velásquez AC, Wu J, Kvitko BH, Kremer JM, Wang Y, He SY, Tsuda K. 2018. Transcriptome landscape of a bacterial pathogen under plant immunity. Proc Natl Acad Sci 115:E3055–E3064.
68. Lindgren PB, Frederick R, Govindarajan A, Panopoulos N, Staskawicz BJ, Lindow SE. 1989. An ice nucleation reporter gene system: identification of inducible pathogenicity genes in *Pseudomonas syringae* pv. *phaseolicola*. EMBO J 8:1291–1301.
69. Boch J, Joardar V, Gao L, Robertson TL, Lim M, Kunkel BN. 2002. Identification of *Pseudomonas syringae* pv. *tomato* genes induced during infection of *Arabidopsis thaliana*. Mol Microbiol 44:73–88.
70. Marco ML, Legac J, Lindow SE. 2005. *Pseudomonas syringae* genes induced during colonization of leaf surfaces. Environ Microbiol 7:1379–1391.
71. Kinkel LL, Wilson M, Lindow SE. 1995. Effect of sampling scale on the assessment of epiphytic bacterial populations. Microb Ecol 29:283–297.
72. Price MN, Deutschbauer AM, Skerker JM, Wetmore KM, Ruths T, Mar JS, Kuehl J V., Shao W, Arkin AP. 2013. Indirect and suboptimal control of gene expression is widespread in bacteria. Mol Syst Biol 9:1–18.
73. Zhang XX, George A, Bailey MJ, Rainey PB. 2006. The histidine utilization (*hut*) genes of *Pseudomonas fluorescens* SBW25 are active on plant surfaces, but are not required for competitive colonization of sugar beet seedlings. Microbiology 152:1867–1875.
74. Turner KH, Everett J, Trivedi U, Rumbaugh KP, Whiteley M. 2014. Requirements for *Pseudomonas aeruginosa* acute burn and chronic surgical wound infection. PLoS Genet 10:e1004518.
75. McAdam PR, Richardson EJ, Fitzgerald RR. 2014. High-throughput sequencing for the study of bacterial pathogen biology. Curr Opin Microbiol 19:106–113.
76. Rastogi G, Coaker GL, Leveau JHJ. 2013. New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. FEMS Microbiol Lett 348:1–10.
77. Chen IMA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. 2019. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res 47:D666–D677.
78. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson M V., Manoil C. 2003. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. Proc Natl Acad Sci 100:14339–14344.
79. Turner KH, Wessel AK, Palmer GC, Murray JL, Whiteley M. 2015. Essential

genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. Proc Natl Acad Sci 112:4110–4115.
80. Andersen GL, Beattie GA, Lindow SE. 1998. Molecular characterization and sequence of a methionine biosynthetic locus from *Pseudomonas syringae*. J Bacteriol 180:4497–4507.
81. Morris J, Lenski RE, Zinser E. 2012. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. MBio 3:e00036-12.
82. Lee SA, Gallagher LA, Thongdee M, Staudinger BJ, Lippman S, Singh PK, Manoil C. 2015. General and condition-specific essential functions of *Pseudomonas aeruginosa* . Proc Natl Acad Sci 112:5189–5194.
83. Beattie GA, Lindow SE. 1994. Survival, growth, and localization of epiphytic fitness mutants of *Pseudomonas syringae* on leaves. Appl Environ Microbiol 60:3790–3798.
84. Rico A, Preston GM. 2008. *Pseudomonas syringae pv. tomato* DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. Mol Plant-Microbe Interact 21:269–282.
85. Erbs G, Newman MA. 2012. The role of lipopolysaccharide and peptidoglycan, two glycosylated bacterial microbe-associated molecular patterns (MAMPs), in plant innate immunity. Mol Plant Pathol 13:95–104.
86. Rocchetta HL, Burrows LL, Lam JS. 1999. Genetics of O-antigen biosynthesis in *Pseudomonas aeruginosa*. Microbiol Mol Biol Rev 63:523–553.
87. Rapicavoli JN, Blanco-Ulate B, Muszyński A, Figueroa-Balderas R, Morales-Cruz A, Azadi P, Dobruchowska JM, Castro C, Cantu D, Roper MC. 2018. Lipopolysaccharide O-antigen delays plant innate immune recognition of *Xylella fastidiosa*. Nat Commun 9:390.
88. Takeuchi K, Taguchi F, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y. 2003. Flagellin glycosylation island in *Pseudomonas syringae* pv. *glycinea* and its role in host specificity. J Bacteriol 185:6658–6665.
89. Taguchi F, Takeuchi K, Katoh E, Murata K, Suzuki T, Marutani M, Kawasaki T, Eguchi M, Katoh S, Kaku H, Yasuda C, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y. 2006. Identification of glycosylation genes and glycosylated amino acids of flagellin in *Pseudomonas syringae* pv. *tabaci*. Cell Microbiol 8:923–938.
90. Buscaill P, Chandrasekar B, Sanguankiattichai N, Kourelis J, Kaschani F, Thomas EL, Morimoto K, Kaiser M, Preston GM, Ichinose Y, Van Der Hoorn RAL. 2019. Glycosidase and glycan polymorphism control hydrolytic release of immunogenic flagellin peptides. Science (80-) 364.
91. Yu J, Penaloza-Vázquez A, Chakrabarty AM, Bender CL. 1999. Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. Mol Microbiol 33:712–720.
92. Markel E, Stodghill P V., Bao Z, Myers CR, Swingle B. 2016. AlgU controls expression of virulence genes in *Pseudomonas syringae* pv. *tomato* DC3000. J Bacteriol 198:2330–2344.
93. Peñaloza-Vázquez A, Kidambi SP, Chakrabarty AM, Bender CL. 1997. Characterization of the alginate biosynthetic gene cluster in *Pseudomonas syringae* pv. *syringae*. J Bacteriol 179:4464–4472.

94. Yu X, Lund SP, Greenwald JW, Records AH, Scott RA, Nettleton D, Lindow SE, Gross DC, Beattie GA. 2014. Transcriptional analysis of the global regulatory networks active in *Pseudomonas syringae* during leaf colonization. MBio 5:e01683-14.
95. Chang WS, Van De Mortel M, Nielsen L, De Guzman GN, Li X, Halverson LJ. 2007. Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. J Bacteriol 189:8290–8299.
96. Kvitko BH, Park DH, Velásquez AC, Wei CF, Russell AB, Martin GB, Schneider DJ, Collmer A. 2009. Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors. PLoS Pathog 5:e1000388.
97. Xu GW, Gross DC. 1988. Evaluation of the role of syringomycin in plant pathogenesis by using Tn5 mutants of *Pseudomonas syringae* pv. *syringae* defective in syringomycin production. Appl Environ Microbiol 54:1345–1353.
98. Schellenberg B, Ramel C, Dudler R. 2010. *Pseudomonas syringae* virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. Mol Plant-Microbe Interact 23:1287–93.
99. Scholz-Schroeder BK, Hutchison ML, Grgurina I, Gross DC. 2001. The contribution of syringopeptin and syringomycin to virulence of *Pseudomonas syringae* pv. *syringae* strain B301D on the basis of *sypA* and *syrB1* biosynthesis mutant analysis. Mol Plant-Microbe Interact 14:336–348.
100. Hutchison ML, Gross DC. 1997. Lipopeptide phytotoxins produced by *Pseudomonas syringae* pv. *syringae*: comparison of the biosurfactant and ion channel-forming activities of syringopeptin and syringomycin. Mol Plant-Microbe Interact 10:347–354.
101. Huse HK, Kwon T, Zlosnik JEA, Speert DP, Marcotte EM, Whiteley M. 2010. Parallel evolution in *Pseudomonas aeruginosa* over 39,000 generations *in vivo*. MBio 1:e00199-10.
102. Luo H, Lin Y, Gao F, Zhang CT, Zhang R. 2014. DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic Acids Res 42:D574–D580.
103. Manoharan B, Neale HC, Hancock JT, Jackson RW, Arnold DL. 2015. The identification of genes important in *Pseudomonas syringae* pv. *phaseolicola* plant colonisation using *in vitro* screening of transposon libraries. PLoS One 10:e0137355.
104. Lindow SE, Andersen GL, Beattie GA. 1993. Characteristics of insertional mutants of *Pseudomonas syringae* with reduced epiphytic fitness. Appl Environ Microbiol 59:1593–1601.
105. Hershey DM, Fiebig A, Crosson S. 2019. A genome-wide analysis of adhesion in *Caulobacter crescentus* identifies new regulatory and biosynthetic components for holdfast assembly. MBio 10:e02273-18.
106. Feil H, Feil WS, Chain P, Larimer F, DiBartolo G, Copeland A, Lykidis A, Trong S, Nolan M, Goltsman E, Thiel J, Malfatti S, Loper JE, Lapidus A, Detter JC, Land M, Richardson PM, Kyrpides NC, Ivanova NN, Lindow SE. 2005. Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and

pv. *tomato* DC3000. Proc Natl Acad Sci USA 102:11064–11069.
107. King EO, Ward MK, Raney DE. 1954. Two simple media for the demonstration of pyocyanin and fluorescin. J Lab Clin Med 44:301–307.
108. R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
109. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
110. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. 2019. gplots: Various R Programming Tools for Plotting Data.
111. Hockett KL, Burch AY, Lindow SE. 2013. Thermo-regulation of genes mediating motility and plant interactions in *Pseudomonas syringae*. PLoS One 8:e59850.
112. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci 97:6640–6645.
113. Chen C, Malek AA, Wargo MJ, Hogan DA, Beattie GA. 2010. The ATP-binding cassette transporter Cbc (choline/betaine/carnitine) recruits multiple substrate-binding proteins with strong specificity for distinct quaternary ammonium compounds. Mol Microbiol 75:29–45.
114. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Nat Biotechnol 1:784–791.
115. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30.
116. Piasecka A, Jedrzejczak-Rey N, Bednarek P. 2015. Secondary metabolites in plant innate immunity: Conserved function of divergent chemicals. New Phytol 206:948–964.
117. Dixon RA. 2001. Natural products and plant disease resistance. Nature 411:843–847.
118. Cowan MM. 1997. Plant products as antimicrobial agents. Clin Microbiol Rev 12:564–582.
119. Fan J, Crooks C, Creissen G, Hill L, Fairhurst S, Doerner P, Lamb C. 2011. *Pseudomonas sax* genes overcome aliphatic isothiocyanate-mediated non-host resistance in *Arabidopsis*. Science (80-) 331:1185–1188.
120. Alvarez-Ortega C, Olivares J, Martínez JL. 2013. RND multidrug efflux pumps: What are they good for? Front Microbiol 4:1–11.
121. Nikaido H. 1996. Multidrug efflux pumps of Gram-negative bacteria. J Bacteriol 178:5853–5859.
122. Brown DG, Swanson JK, Allen C. 2007. Two host-induced *Ralstonia solanacearum* genes, *acrA* and *dinF*, encode multidrug efflux pumps and contribute to bacterial wilt virulence. Appl Environ Microbiol 73:2777–2786.
123. Burse A, Weingart H, Ullrich MS. 2004. The phytoalexin-inducible multidrug efflux pump AcrAB contributes to virulence in the fire blight pathogen, *Erwinia amylovora*. Mol Plant-Microbe Interact 17:43–54.
124. Reddy JD, Reddy SL, Hopkins DL, Gabriel DW. 2007. TolC is required for pathogenicity of *Xylella fastidiosa* in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:403–410.

125. Lomovskaya O, Lee A, Hoshino K, Ishida H, Mistry A, Warren MS, Boyer E, Chamberland S, Lee VJ. 1999. Use of a genetic approach to evaluate the consequences of inhibition of efflux pumps in *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 43:1340–1346.
126. Vargas P, Felipe A, Michán C, Gallegos M-T. 2011. Induction of *Pseudomonas syringae* pv. *tomato* DC3000 MexAB-OprM multidrug efflux pump by flavonoids Is mediated by the repressor PmeR. Mol Plant-Microbe Interact 24:1207–1219.
127. Lin N-C, Martin GB. 2007. Pto- and Prf-mediated recognition of AvrPto and AvrPtoB restricts the ability of diverse *Pseudomonas syringae* pathovars to infect tomato. Mol Plant-Microbe Interact 20:806–815.
128. Tikhonova EB, Wang Q, Zgurskaya HI. 2002. Chimeric analysis of the multicomponent multidrug efflux transporters from Gram-negative bacteria. J Bacteriol 184:6499–6507.
129. Schweizer HP. 2003. Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions. Genet Mol Res 2:48–26.
130. Arnold DL, Lovell HC, Jackson RW, Mansfield JW. 2011. *Pseudomonas syringae* pv. *phaseolicola*: From “has bean” to supermodel. Mol Plant Pathol 12:617–627.
131. Kwak M, Gepts P. 2009. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theor Appl Genet 118:979–992.
132. Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. 2009. Shades of gray: the world of quantitative disease resistance. Trends Plant Sci 14:21–29.
133. Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, Sands DC, Morris CE. 2014. A user’s guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. PLoS One 9:e105547.
134. Cuppels DA, Elmhirst J. 1999. Disease development and changes in the natural *Pseudomonas syringae* pv. *tomato* populations on field tomato plants. Plant Dis 83:759-764.
135. Stromberg KD, Kinkel LL, Leonard KJ. 1999. Relationship between phyllosphere population sizes of *Xanthomonas translucens* pv. *translucens* and bacterial leaf streak severity on wheat seedlings. Epidemiology 89:131–135.
136. Kus J V, Zaton K, Sarkar R, Cameron RK. 2002. Age-related resistance in *Arabidopsis* is a developmentally regulated defense response to *Pseudomonas syringae*. Plant Cell 14:479–490.
137. Sutton DC, Deverall BJ. 1983. Studies on infection of bean (*Phaseolus vulgaris*) and soybean (*Glycine max*) by ascospores of *Sclerotinia sclerotiarum*. Plant Pathol 32:251–261.
138. Zuiderveen GH, Padder BA, Kamfwa K, Song Q, Kelly JD. 2016. Genome-wide association study of anthracnose resistance in Andean beans (*Phaseolus vulgaris*). PLoS One 11:1–17.
139. Balardin RS, Kelly JD. 1998. Interaction between *Colletotrichum lindemuthianum* races and gene pool diversity in *Phaseolus vulgaris*. J Am Soc Hortic Sci 123:1038–1047.
140. Young RA, Kelly JD. 1996. Characterization of the genetic resistance to

Colletotrichum lindemuthianum in common bean differential cultivars. Plant Dis 80:650–654.
141. Guzmán P, Gilbertson RL, Nodari R, Johnson WC, Temple SR, Mandala D, Mkandawire ABC, Gepts P. 1995. Characterization of variability in the fungus *Phaeoisariopsis griseola* suggests coevolution with the common bean (*Phaseolus vulgaris*). Genetics 85:600–607.
142. Hagedorn DJ, Rand RE, Saad SM. 1971. Tolerance of *Phaseolus lunatus* to bacterial brown spot. Phytopathology 61:1406–1407.
143. Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, Grusak MA, Kamfwa K, Katuuramu DN, McClean P, Mndolwa E, Nchimbi-Msolla S, Pastor-Corrales MA, Miklas PN. 2015. A *Phaseolus vulgaris* diversity panel for Andean bean improvement. Crop Sci 55:2149–2160.
144. Lyon FM, Wood RKS. 1975. Production of phaseollin, coumestrol and related compounds in bean leaves inoculated with *Pseudomonas* spp. Physiol Plant Pathol 6:117–124.
145. Bailey JA, Burden RS. 1973. Biochemical changes and phytoalexin accumulation in *Phaseolus vulgaris* following cellular browning caused by tobacco necrosis virus. Physiol Plant Pathol 3:171–177.
146. Gnanamanickam SS, Patil SS. 1977. Accumulation of antibacterial isoflavonoids in hypersensitively responding bean leaf tissues inoculated with *Pseudomonas phaseolicola*. Physiol Plant Pathol 10:159–168.
147. Woodward MD. 1980. Phaseollin formation and metabolism in *Phaseolus vulgaris*. Phytochemistry 19:921–927.
148. Tegos GP, Stermitz FR, Lomovskaya O, Lewis K. 2002. Multidrug pump inhibitors uncover remarkable activity of plant antimicrobials. Antimicrob Agents Chemother 46:3133–3141.
149. Lewis K, Ausubel FM. 2006. Prospects for plant-derived antibacterials. Nat Biotechnol 24:1504–1507.
150. Alfano JR, Collmer A. 2004. Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Annu Rev Phytopathol 42:385–414.
151. Cunnac S, Lindeberg M, Collmer A. 2009. *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. Curr Opin Microbiol 12:53–60.
152. Alfano JR. 1996. Bacterial Pathogens in Plants: Life up against the Wall. Plant Cell Online 8:1683–1698.
153. Grgurina I, Mariotti F, Fogliano V, Gallo M, Scaloni A, Iacobellis NS, Lo Cantore P, Mannina L, Van Axel Castelli V, Greco ML, Graniti A. 2002. A new syringopeptin produced by bean strains of *Pseudomonas syringae* pv. *syringae*. Biochim Biophys Acta - Protein Struct Mol Enzymol 1597:81–89.
154. Schellenberg B, Ramel C, Dudler R. 2010. *Pseudomonas syringae* virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. Mol Plant Microbe Interact 23:1287–1293.
155. Keith RC, Keith LMW, Hernández-Guzmán G, Uppalapati SR, Bender CL. 2003. Alginate gene expression by *Pseudomonas syringae* pv. *tomato* DC3000 in host and non-host plants. Microbiology 149:1127–1138.

156. Burch AY, Finkel OM, Cho JK, Belkin S, Lindow SE. 2013. Diverse microhabitats experienced by *Halomonas variabilis* on salt-secreting leaves. Appl Environ Microbiol 79:845–852.
157. Nies DH. 2003. Efflux-mediated heavy metal resistance in prokaryotes. FEMS Microbiol Rev 27:313–339.
158. Fernandes P, Sommer Ferreira B, Sampaio Cabral JM. 2003. Solvent tolerance in bacteria: role of efflux pumps and cross-resistance with antibiotics. Int J Antimicrob Agents 22:211–216.
159. Nikaido H. 1996. Multidrug efflux pumps of gram-negative bacteria. J Bacteriol 178:5853–5859.
160. Stermitz FR, Lorenz P, Tawara JN, Zenewicz LA, Lewis K. 2000. Synergy in a medicinal plant: Antimicrobial action of berberine potentiated by 5'-methoxyhydnocarpin, a multidrug pump inhibitor. Proc Natl Acad Sci 97:1433–1437.
161. Blair JM, Piddock LJ V. 2009. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. Curr Opin Microbiol 12:512–519.
162. Martinez JL, Sánchez MB, Martínez-Solano L, Hernandez A, Garmendia L, Fajardo A, Alvarez-Ortega C. 2009. Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. FEMS Microbiol Rev 33:430–449.
163. Yan N. 2013. Structural advances for the major facilitator superfamily (MFS) transporters. Trends Biochem Sci 38:151–159.
164. Nishino K, Yamaguchi A. 2001. Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. J Bacteriol 183:5803–5812.
165. Okusu H, Ma D, Nikaido H. 1996. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. J Bacteriol 178:306–308.
166. Rempe CS, Burris KP, Lenaghan SC, Stewart CN. 2017. The potential of systems biology to discover antibacterial mechanisms of plant phenolics. Front Microbiol 8:422.
167. Poole K, Heinrichs DE, Neshat S. 1993. Cloning and sequence analysis of an EnvCD homologue in *Pseudomonas aeruginosa*: regulation by iron and possible involvement in the secretion of the siderophore pyoverdine. Mol Microbiol 10:529–544.
168. Li X-Z, Barre N, Poole K. 2000. Influence of the MexA-MexB-OprM multidrug efflux system on expression of the MexC-MexD-OprJ and MexE-MexF-OprN multidrug efflux systems in *Pseudomonas aeruginosa*. J Antimicrob Chemother 46:885–893.
169. Li XZ, Plésiat P, Nikaido H. 2015. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. Clin Microbiol Rev 28:337–418.
170. Blair JM, Smith HE, Ricci V, Lawler AJ, Thompson LJ, Piddock LJ V. 2015. Expression of homologous RND efflux pump genes is dependent upon AcrB expression: Implications for efflux and virulence inhibitor design. J Antimicrob Chemother 70:424–431.
171. Baranova N, Nikaido H. 2002. The BaeSR two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. J

Bacteriol 15:4168–4176.
172. Nagakubo S, Nishino K, Hirata T, Yamaguchi A. 2002. The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. J Bacteriol 15:4161–4167.
173. Pletzer D, Weingart H. 2014. Characterization and regulation of the Resistance-Nodulation-Cell Division-type multidrug efflux pumps MdtABC and MdtUVW from the fire blight pathogen *Erwinia amylovora*. BMC Microbiol 14:1–15.
174. Dreier J, Ruggerone P. 2015. Interaction of antibacterial compounds with RND efflux pumps in *Pseudomonas aeruginosa*. Front Microbiol 6:1–21.
175. Mima T, Kohira N, Li Y, Sekiya H, Ogawa W, Kuroda T, Tsuchiya T. 2009. Gene cloning and characteristics of the RND-type multidrug efflux pump MuxABC-OpmB possessing two RND components in *Pseudomonas aeruginosa*. Microbiology 155:3509–3517.
176. Higgins CF. 2007. Multiple molecular mechanisms for multidrug resistance transporters. Nature 446:749–757.
177. Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FSL. 2016. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the Pseudomonas genome database. Nucleic Acids Res 44:D646–D653.
178. Belofsky G, Percivill D, Lewis K, Tegos GP, Ekart J. 2004. Phenolic metabolites of *Dalea versicolor* that enhance antibiotic activity against model pathogenic bacteria. J Nat Prod 67:481–484.
179. Matthysse AG, Stretton S, Dandie C, McClure NC, Goodman AE. 1996. Construction of GFP vectors for use in Gram-negative bacteria other than *Escherichia coli*. FEMS Microbiol Lett 145:87–94.
180. Hoang T, Karkhoff-Schweizer R, Kutchma A, Schweizer HP. 1998. A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene 212:77–86.
181. Wei CF, Kvitko BH, Shimizu R, Crabill E, Alfano JR, Lin NC, Martin GB, Huang HC, Collmer A. 2007. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J 51:32–46.
182. Stavrinides J, McCloskey JK, Ochman H. 2009. Pea aphid as both host and vector for the phytopathogenic bacterium *Pseudomonas syringae*. Appl Environ Microbiol 75:2230–2235.
183. Wilkinson TL, Ashford DA, Pritchard J, Douglas AE. 1997. Honeydew sugars and osmoregulation in the pea aphid *Acyrthosiphon pisum*. J Exp Biol 200:2137–2143.