Genomic Correlates of Tailocin Sensitivity in *Pseudomonas syringae*

David A. Baltrus[1,2]
Savannah Weaver[3]
Laura Krings[1]
Anh Evy Nguyen[1]

[1]School of Plant Sciences, University of Arizona, Tucson AZ , USA
[2]School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson AZ , USA
[3]Department of Molecular and Cellular Biology, University of Arizona, Tucson AZ, USA

Abstract

Phage derived bacteriocins, also referred to as tailocins, are structures encoded by bacterial genomes and deployed into the extracellular environment to target and kill sensitive cells. Tailocins display great potential as agricultural antimicrobials due to their durability, efficiency, and specificity of killing with prophylactic application of these molecules having been shown to prevent infection by multiple phytopathogens. Although previous reports have strongly suggested that tailocins of *Pseudomonas syringae* bind sugar moieties in the lipopolysaccharide of target cells, the molecular mechanisms and binding interactions that enable tailocins from *P. syringae* to kill sensitive targets remain unclear. We therefore carried out a genome-wide association study investigating tailocin sensitivity across a diverse set of *P. syringae* genomes. Our results demonstrate that genes strongly correlated with tailocin sensitivity are localized to one contiguous region on the *P. syringae* chromosome encoding LPS structures similar to the Common Polysaccharide Antigen of *P. aeruginosa*. We further find that enzymes involved in the biosynthesis and transport of D-rhamnose and L-rhamnose are associated with tailocin sensitivity classes A and B, respectively, with large-scale recombination of the O-antigen biosynthesis region likely underlying rapid and fundamental changes in LPS structure between strains. Lastly, we identify the *rfbD* gene as an additional genomic indicator to predict tailocin sensitivity and use this information to test tailocin interactions with previously unscreened strains from across phylogroups.

Introduction

Phage derived bacteriocins, hereafter referred to as tailocins, are proteinaceous structures encoded by bacterial genomes thought to be used by strains in an extracellular manner to target and kill competing strains under natural conditions[1–3]. The efficacy of killing, durability, and strain specificity of tailocins have made them exceptional candidates for development and deployment as antimicrobials in clinical and agricultural environments[4, 5]. To best maximize the efficiency of use and develop methods to optimize application strategies, as well as to predict sensitivity of untested strains, we must better understand the genetic basis of resistance to tailocins across diverse suites of strains. Here, we present a genome wide association analysis of loci that affect tailocin based killing within *Pseudomonas syringae* strains across a set of phylogenetically diverse strains previously screened for tailocin sensitivity and use this information to predict tailocin sensitivity for additional strains from their genome sequences.

Tailocins are part of a broad class of structures, also including type VI secretion systems and extracellular contractile injection systems (eCIS), coopted by bacterial genomes from phages and structurally resembling phage tails[2, 3]. They are produced by bacterial cells and released through cell lysis, but unlike either the type VI systems and eCIS, are not currently known to deliver effector proteins to target cells. Tailocins have evolved at least thrice independently from different phage: with the R-type pyocins and R-type syringacins evolved independently from myoviridae phage resembling P2 and Mu respectively, and the F-type pyocins having been derived from a phage in the siphoviridae family [1, 6]. Although a variety of bacteria besides Pseudomonads have been shown to produce tailocins, much of our current knowledge concerning tailocin killing and resistance comes from studies investigating R-type tailocins (also known as R-type pyocins) of *P. aeruginosa* [2, 3, 7–10]. R-type pyocins have been shown to efficiently kill sensitive cells through direct lysis by mechanical disruption of target cell membranes, with binding of tailocins determined largely by receptor binding proteins (Rbps) and their associated chaperones[8, 9, 11]. These two loci together appear to determine the targeting spectrum for both types of Pseudomonad R-type tailocins (pyocins and syringocins), and the targeting spectrum for R-type syringacins found in *Pseudomonas syringae* has been shown to rapidly evolve through localized recombination of regions containing the Rbps [12].

The outer membranes of Gram negative bacteria are decorated with a variety of proteins and sugars, with one of the most well-known structures being the lipopolysaccharide (LPS) layer [13]. The LPS is composed of covalently bound chains of various sugar moieties, the compositions of which vary within and between species, anchored into the outer membrane

[14]. Since the LPS is a major component determining interactions of bacteria with the immune systems of hosts, is often used for attachment of phage, and is likely involved in a variety of other processes important for survival across bacterial populations and communities, composition of the LPS is thought to be under strong diversifying selection and can be notoriously variable even between closely related strains [15]. Earlier studies on the R-pyocins of *P. aeruginosa* implicated the LPS as the binding sites for these tailocins and specifically by the Rbps, and pointed towards either rhamnose or glucose as the interacting sugar moieties [9]. Additionally, *P. aeruginosa* strains are able to produce multiple types of O-antigen that are potentially linked to the LPS core: one type of chain is referred to as the Common Polysaccharide Antigen (CPA) or A band and is composed of repetitive chains of D-rhamnose linked together, while the other chain can be composed of more diverse sugars and is referred to as the B band [14, 16]. The presence, composition, and length of these different O-antigen bands has been previously implicated as playing a role in shielding or enabling attachment of tailocins to LPS-associated sugars [9, 17]. Further, RB-TnSeq across a handful of *P. fluorescens* strains implicated both the core and O-antigen bands of the LPS as interaction sites for these molecules, with disruption of genes from all of these regions providing resistance against tailocin-based killing [18]. Conversely, multiple papers have implicated the Common Polysaccharide Antigen or A band of the LPS as a main site of interactions of tailocins from *P. syringae* as genetic data suggests that *P. syringae* strains may only produce variations on the CPA and deletions or modification of genes within the CPA production pathway differentially affect tailocin targeting to otherwise sensitive strains [19, 20]. This difference in tailocin targeting between R-pyocins and R-syringacins is perhaps not completely surprising as it echoes from our current understanding that tailocins of *P. syringae* are derived from a different progenitor phage than the R-type pyocins of *P. aeruginosa* and many of the R-type tailocins of *P. fluorescens* [6].

Here, we report a genome-wide association study to map the genetic basis of tailocin sensitivity across a variety of closely related *P. syringae* strains and in order to identify genomic regions associated with differing sensitivity to R-type syringacins. We then extrapolate from these results to map and predict tailocin sensitivity across a diverse range of isolates from this species and uncover genomic trends that appear to enable prediction of tailocin sensitivity for these strains. Overall, our results highlight the variability of pathways determining LPS structures throughout *P. syringae*, suggest that large-scale recombination of one region on the chromosome is responsible for fundamental shifts in LPS composition in *P. syringae*, and provide information that could enable prediction of tailocin sensitivity across this phytopathogenic species.

Methods

Genome Sequences Used for Genome-Association Studies and Tailocin Prediction Studies

All genomes referenced in this manuscript are listed in Table 1, which also lists accessions for these sequences at Genbank and abbreviations used for each strain in the manuscript.

Genome-wise Association with Tailocin Sensitivity

We downloaded .fasta sequences for each of the genomes in Table 1 that had been previously screened for tailocin sensitivity in [12], and used Prokka to reannotate each in a consistent manner [21]. We then carried out a clustering analysis of the .gff files from these genomes using Roary and with default parameters [22]. Last, we used Scoary to associate tailocin sensitivity with clusters of orthologous genes found by Roary [23]. Tailocin sensitivity classes for each gene as used in the Scoary analysis are listed in Table 1, and we note that sensitivity classes for strains CC1630, *Pto*T1, and *Pla*106 differ from what was published in [12] due to labelling errors in the previous manuscript. All input and output files for Prokka, Roary, and Scoary analyses can be found as supplemental files at 10.6084/m9.figshare.22688020.

Phylogenetic Analysis

We obtained protein sequences for RfbD from each of the genomes listed in Table 1, and aligned these using Clustal Omega with default parameters [24]. We used Modeltest to find the best evolutionary model describing this alignment [25], which was JTT+G4. We then inferred phylogenetic relationships from this alignment using RaxML-ng [26]. We performed bootstrapping analysis in RaxML-ng, with phylogenetic models converging after 850. We then calculated bootstrap values for nodes on the best tree according to RaxML-ng using the transfer bootstrap extrapolation (-tbe) function. All input and output files for phylogenetic analyses can be found as supplemental files at 10.6084/m9.figshare.22688020.

Soft-agar Overlay of Previously Unscreened Strains for Tailocin Sensitivity

Soft-agar overlays were carried out as previously described [27]. Briefly, overnight cultures of strains CC440, USA011, USA007, and CC1416 were diluted back 1:100 in King's B media and grown for 4 hours at 27°C while shaking at 220rpm. At this point, mitomycin C was added to these cultures to a final concentration of 0.75 μg/mL and the cultures were incubated under the same conditions overnight. The next day, tailocins from these strains were purified from the supernatant using PEG precipitation.

For overlays, strains of interest were grown overnight in KB media at 27°C while shaking at 220rpm. The next morning, cultures were diluted 1:100 in fresh KB media and incubated an additional 4 hours under the same environmental conditions. At this point, 100μL of bacterial cultures was mixed with 3mL of 0.4% molten agar and poured onto the top of a KB agar plate. Top agar was allowed to solidify for ~10 minutes, at which point a 10μL sample of each tailocin was added. Killing zones for each strain were observed the next day. Overlay assays were carried out at least three independent times, with representative results reported. Unedited pictures of the overlay results can be found in supplemental files at 10.6084/m9.figshare.22688020.

Results

Genomic Association for Tailocin Sensitivity Across P. syringae

There have been numerous switches in the phenotype of tailocin sensitivity across *P. syringae* strains. We therefore hypothesized that genomic investigation of strains could highlight strong and independent signals pointing to the identity of genes and pathways contributing to differential tailocin sensitivity. The results of an analysis using Roary and Scoary to correlate genomic changes with tailocin sensitivity are shown in Table 2, and we find that 17 different loci are significantly ($p<0.05$) associated with tailocin sensitivity by these analyses and after Benjamini-Hochberg correction for multiple tests. A subset of 11 of these remain significantly associated with tailocin sensitivity after Bonferroni correction for multiple tests ($p<0.05$). Looking at the larger set more specifically, there are 9 loci whose presence is strongly correlated with tailocin sensitivity group A (as per [12]) and 8 whose presence is strongly correlated with tailocin sensitivity group B. Each of the 17 loci associated with each sensitivity group are colocated in syntenic genomic locations across strains, a location which corresponds to gene clusters implicated in production of the Common Polysaccharide Antigen (CPA) chain of the LPS as per [19].

To gain deeper insight into this genomic cluster and these results and to highlight diversification of this region between closely related strains, we focused on investigating gene presence and synteny within the entirety of the CPA locus across a subset of phylogroup 2 strains (Figure 1). We specifically found that this region was bracketed by conserved loci in nearly every strain, with the *ildD* coding for L-lactate dehydrogenase on one end and the gene for *ychF* coding for an ATPase on the other end. Genes found in this region in tailocin sensitivity group A include those involved in the production and modification of rhamnose moieties in the CPA (*gmd, rmd, wzm, algA,* and numerous glycosyl transferases and sugar modification enzymes). This region also encodes the TET operon, a type I secretion system to transport CPA

sugar chains from the cytoplasm, but as highlighted in Jayaraman et al. and elaborated on here there is extensive sequence variation between these operons across sensitivity groups. Genes found in this region in tailocin sensitivity group B also include those involved in the production and modification of rhamnose moieties in the CPA (*rfbABC,* AKA *rmlABC*) as well as other glycosyl transferases and sugar modification enzymes and two potential secretion systems). Despite diversity within this *ildD-ychF* region across these strains, it is clear that the genes implicated as contributing to tailocin sensitivity by the Scoary analysis are colocalized and found in syntenic locations across sensitivity classes. However, more pointedly, organization and sequence of this region completely changes with tailocin sensitivity and this change is consistent across all independent switches in sensitivity. Such changes appear to be much broader than that highlighted in [19], as a much greater portion of this region is swapped out through what appears to be a recombination event.

rbfD As an Indicator Gene for Tailocin Sensitivity

A previous report identified gene sequences within the TET operon as fairly accurate indicators of tailocin sensitivity classes [19]. Upon inspection of the CPA locus across tailocin sensitivity classes, we observed that *rfbD* gene, encoding dTDP-4-dehydrorhamnose reductase, appears to be the only gene located in the *ildD-ychF* region across all strains. However, unlike the TET operon throughout *P. syringae*, sequence similarity across alleles of *rfbD* is extensive enough to be identified and analyzed across strains by sequence searches alone. We therefore pulled sequences for alleles of this locus from all strains from the main *P. syringae* phylogroups as described in [28] (phylogroups 1-6 and 10) and analyzed in [12], as well as 5 strains for which tailocin sensitivity has yet to be screened and inferred a maximum likelihood phylogeny from these sequences. As one can see in Figure 2A, alleles of RfbD form two distinct clades, with clade membership showing a complete match to tailocin sensitivity in all previously screened strains (with the caveat that tailocin sensitivity of three strains is misidentified in [12]; *Pto*T1, *Pla*106, CC1630 are labeled as sensitivity class A in this prior publication but should clearly be class B upon further comparison). From this phylogeny, we predict that strains *Pto*2170, Cit7, *Por* 1_36, *Phel* LMG 5067, and *Ptg* ICMP 4091 would be characterized as sensitivity class B,B,B,A and A respectively based on *rfbD* sequences.

We performed overlay assays using these five strains, which were not included in our previous tailocin sensitivity matrix and which represent multiple phylogroups of *P. syringae.* As a means to categorize tailocin sensitivity classes for these strains, we, used tailocins produced by four different strains from two phylogroups that delineate sensitivity groups A and B as per Baltrus et al. 2019: CC440 (phylogroup 2, killing class 1), USA011 (phylogroup 2, killing class 2), USA007 (phylogroup 3, killing class 1), and CC1416 (phylogroup 3, killing class 2). Results from these overlays are largely in line with predictions from *rfbD* phylogenies for strains Cit7, *Por*

1_36, and *Pto*2170, and show that these strains are classified as sensitivity group B (Figure 3B). Likewise, phylogenetic relationships accurately predict that *Ptg* ICMP 4091 would be sensitivity class A. *Phel* LMB 5067 appears to be resistant to all four tailocins, a result which is rare but not unprecedented.

Discussion

Composition of the lipopolysaccharide layer of the outer membrane has been shown to impact multiple phenotypes crucial for survival of bacteria that are widely distributed throughout the environment [13]. Likewise, since the LPS often forms the interface between bacteria and hosts, changes in the LPS can have widespread effects during infection of plants by phytopathogenic bacteria [29, 30]. Our findings support previously published results implicating the LPS, and specifically a structure analogous to the CPA chain of *P. aerguinosa* formed by rhamnose polymers, as the target of *P. syringae* tailocins. Previous results suggested that changes in LPS composition across strains of pathovar *actinidae* were due to localized recombination of allelic variation of the TET operon [19], which is part of this CPA locus and encodes ABC transporter subunits, a bifunctional glycosyltransferase and an o-methyltransferase. The TET operon is responsible for the translocation, ligation, and modification of D-rhamnose moieties as they move from construction in the cytoplasm to the outer membrane. Jayaraman et al. were also able to extrapolate to strains outside of pathovar *actinidiae* and showed that allelic diversity in genes resembling the TET operon in context and sequence could predict tailocin sensitivity broadly across *P. syringae and P. viridiflava* strains.

While our results echo those of Jayaraman et al., the patterns reported here skew more towards presence/absence of genes and broad structural differences in the arrangement of the operons rather than *just* redistribution of allelic variation. Specifically, we show that operons encoding pathways for biosynthesis of precursors of the rhamnose moieties, multiple type I secretion systems including the TET operon, and various glycosyl transferases and sugar modification enzymes are syntenic and colocalized across many strains but also appear to undergo repeated wholesale swaps in gene content when a broader set of *P. syringae* genomes is analyzed for correlations with tailocin sensitivity. We further evaluate each of these operons in the context of LPS production below:

gmd, rmd, algA

Alleles of *gmd* and *rmd* are differentially present within the CPA locus of strains classified as sensitivity class A (represented by strain *Pseudomonas syringae* pv. *syringae* B728a, *Psy*B728a, in Fig. 1) and form orthologous clusters according to the Roary pipeline which are

significantly associated with tailocin sensitivity according to Scoary. Together with *wbpZ,* these genes are predicted to form an operon involved in the biosynthesis and linkage of D-rhamnose in the O-antigen of the LPS [19]. While spacing and orientation of these genes and the lack of factor independent transcriptional terminator suggests that all three could coregulated as an operon within *Psy*B728a, it is unclear whether regions upstream of *wbpZ* can independently control transcription of this gene.

D-rhamnose in the LPS is ultimately formed through the production of GDP-D-rhamnose from GDP-D-mannose. Gmd is predicted to encode GDP-mannose 4,6-dehydratase, an enzyme that carries out the first step in this pathway and converts GDP-D-mannose to GDP-4-dehydro-6-deoxy-D-mannose (aka GDP-6-deoxy-d-lyxo-hexos-4-ulose) [31]. Rmd is predicted to encode GDP-6-deoxy-d-lyxo-hexos-4-ulose-4-reductase, which produces GDP-D-rhamnose from the substrate produced by Gmd [31]. *wbpZ* is a glycosyltransferase that can attach D-rhamnose to the growing LPS chain from GDP-D-rhamnose [32]. Alleles of *algA* are encoded by genes within the CPA locus in strains from *P. syringae* tailocin sensitivity class A, also form an orthologous cluster according to the Roary pipeline, and are significantly associated with tailocin sensitivity class A strains according to Scoary. AlgA is predicted to encode a bifunctional protein with phosphomannose isomerase (PMI) and GDP-mannose pyrophosphorylase (GMP) activities, and which is therefore predicted to catalyze the conversion of fructose-6-phosphate to mannose-6-phosphate and mannose-1-phosphate to GDP-mannose [33]. It is likely that this enzyme catalyzes the substrate (GDP-mannose) that is acted upon by other genes in the CPA locus significantly associated with tailocin sensitivity in class A strains and which produce GDP-D-rhamnose, because GDP-mannose is the substrate for the Gmd protein.

Many strains from *P. syringae* sensitivity class B can possess relatively distant orthologues of both *gmd* (~80% protein identity) and *rmd* (~30% protein identity), and while they are positioned in close proximity within these strains, they are located outside of the CPA locus and do not appear to be closely associated with an additional glycosyltransferase (data not shown). Disruption of this operon in sensitivity class A strain *P. syringae* pv. *actinidiae* ICMP 18884, and specifically within *gmd* (mutant B12), and *wbpZ* (mutant A1) has been shown to provide resistance to tailocin killing [20]. It has also been shown that a point mutation in *rmd* can provide tailocin resistance in the class A strain *P. syringae* pv. *phaseolicola Pph*1448a [34]. The significant association of D-rhamnose biosynthesis genes with tailocin sensitivity class A strains is strong evidence suggesting D-rhamnose as a potential binding target of tailocins that can target these strains.

rfbA, rfbB, wecE

Alleles of *rfbA* and *rfbB* from sensitivity class B form an orthologous cluster according to the Roary pipeline which are significantly associated with tailocin sensitivity according to Scoary. Additionally, an orthologue of *wecE* appears to be found within a separate operon within the CPA locus, but is also significantly correlated with tailocin sensitivity class B strains.

RfbA (also known as RmlA) is a glucose-1-phosphate thymidylyltransferase, which catalyzes the formation of dTDP-glucose from dTTP and glucose 1-phosphate [35]. It should be noted that there are two distinct enzymes (RfbA and RffH) identified in *E. coli* that can carry out this same reaction, and that these enzymes function in different pathways and are present in different operons in *E. coli* [36, 37]. RfbB is a dTDP-glucose 4,6-dehydratase, which catalyzes the dehydration of dTDP-D-glucose to form dTDP-6-deoxy-D-xylo-4-hexulose [38]. WecE is a dTDP-4-amino-4,6-dideoxygalactose transaminase [39]. L-rhamnose in the LPS is ultimately formed through the production of dTDP-L-rhamnose from GDP-L-mannose, and each of these enzymes is therefore predicted to be involved in the biosynthesis or modification of L-rhamnose for the O-antigen of tailocin sensitivity class B strains. As with the D-rhamnose biosynthesis gene associations with tailocin sensitivity class A genomes, the significant association of L-rhamnose biosynthesis genes with tailocin sensitivity class B strains strongly suggests L-rhamnose as a potential binding target of tailocins that can target these strains.

Lastly, we note that an independent orthologous cluster of *rfbB* alleles is also significantly associated with tailocin sensitivity class A strains. While this set of genes likely encodes an enzyme with similar functionality as those that are found in sensitivity class A strains, there is significant divergence between alleles of *rfbB* from each sensitivity class.

Potential Type I Secretion and Transporter Loci (*wzm,* ATPase, *prsD, prsE,* TolC-like, *rfbC,* Permease)

For class A genomes, the presence of an apparent orthologue of *wzm* as well an ATPase are significantly correlated with tailocin sensitivity. In *P. aeruginosa,* rhamnose precursors of the CPA are transported from the cytoplasm of the cell by an ABC transporter (*wzt/wzm*), with energy provided by an ATPase, and with a fourth glycosyltransferase gene present in the operon to attach these sugars to the growing chain [19, 40]. This particular operon in *P. syringae* is the same as highlighted by Jayaraman et al. for its patterns of recombination and it is likely that this operon in tailocin sensitivity class A strains is carrying out the same activities as the *wzt/wzm* ABC transporter in *P. aeruginosa*. Also worth noting is that there is a fourth gene encoding a glycosyltransferase (potentially referred to as *wbdD*) which appears to be part of this operon in sensitivity class A genomes. However, significant divergence for this gene within sensitivity classes due to recombination likely leads to a lack of overall significant associations overall for this enzyme with class A strains, as pointed out by Jayaraman et al. Disruption of this operon, and specifically within *wzm* (mutant A4), *wzt* (mutant B4), and *wbdD* (mutants A6

and B6) has been shown to provide resistance to tailocin killing of the sensitivity class A strain *P. syringae* pv. *actinidiae* ICMP 18884 [19, 20]. Given the predicted enzymatic activities of these genes, it is likely that they are responsible for attaching D-rhamnose to the O-antigen of *P. syringae* tailocin sensitivity class A strains. This provides an additional line of evidence suggesting D-rhamnose as a potential binding target of tailocins that can target sensitivity class A strains.

Tailocin sensitivity class B genomes potentially encode multiple transporters that each contain subsets of genes significantly associated with tailocin sensitivity. One of these transporter operons appears to encode four genes that likely encode an ABC transporter used by Jayaraman in phylogenetic analyses as a predictor of tailocin sensitivity [19](and which is analogous to the potential *wzt/wzm* ABC transporter system mentioned above but divergent both in sequence and in annotation). Three of these genes in this operon are significantly associated with tailocin sensitivity for class B: a permease, *rfbC* (dTDP-4-dehydrorhamnose 3,5-epimerase), and an ATPase. As with the *wzt/wzm* operon mentioned above, there is often a fourth glycosyltransferase gene downstream of these three but which is divergent enough across strains so that significant associations with tailocin sensitivity are likely obscured. The presence of an *rfbC* orthologue as part of this pathway further hints that this may be the ABC transporter that enables building L-rhamnose chains as part of the *P. syringae* O-antigen, and provides additional support for the idea that L-rhamnose is a potential binding target of tailocins that can target sensitivity class B strains.

One additional type I secretion operon is significantly associated with tailocin sensitivity in many class B genomes, and contains genes similar to *prsD, prsE,* as well as a locus potentially encoding a TolC-like protein. Unlike the two transporter operons mentioned above, there does not appear to be a glycosyltransferase as part of this operon across strains. PrsD and PrsE are involved in exporting acidic polysaccharide in *Rhizobium* as part of biofilm formation, with PrsD acting as an ATP-binding protein and PrsE encoding a membrane fusion protein [41], but their functions within these *P. syringae* strains is unknown. TolC is an outer membrane channel protein for type I secretion systems [42]. By annotation, it is likely that this operon encodes proteins responsible for transporting sugars as part of the LPS of *P. syringae*, but at present it is difficult to identify the precise substrates. TolC outer membrane proteins are predicted to form "tunnels" through the outer membrane for transport of substrates for a variety of different type I secretion pathways.

Glycosyl-transferases

Sugar chains of the LPS are built through the action of glycosyltransferases, which take sugars linked to a variety of substrates and transfer them to the growing LPS chains [43, 44]. One glycosyltransferase, potentially found in an operon with *wecE* mentioned above, is

significantly correlated with tailocin sensitivity group B strains and therefore is potentially involved in the transfer of L-rhamnose. One glycosyltransferase is significantly associated with tailocin sensitivity class A strains, and appears to be present as part of a larger operon across these strains, but is the only gene from this operon that shows a significant association. Outside of context clues, such as other enzymes encoded in operons with these loci, it is difficult to pinpoint the functions of these glycosyl transferases without further experiments.

dapH

dapH putatively encodes 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase, which catalyzes the transfer of an acetyl group from acetyl-CoA to tetrahydrodipicolinate [45]. There are multiple copies of similar enzymes throughout the *P. syringae* genome, but this particular copy is correlated with tailocin sensitivity group B genomes and potentially points towards the role of additional modifications of sugars in LPS production and tailocin sensitivity.

Previous manuscripts demonstrate that allelic variation within the TET operon could be effectively used for prediction of tailocin sensitivity [19]. Evaluation of genetic diversity in this pathway likely acts as a good metric for tailocin sensitivity when whole genome sequences are in hand, however, the TET operon shares little sequence similarity across strains within different tailocin sensitivity classes and thus there may not be an easy way to design probes to identify tailocin sensitivity classes without whole genomes. Our analyses further identify *rfbD* as potential additional candidate genes for extrapolating and predicting tailocin sensitivity classes from genome sequence alone for most *P. syringae* strains. Upon examination of the CPA locus of numerous *P. syringae* genomes, and unlike many of the genes significantly associated with tailocin sensitivity listed in Table 2, we observed that alleles of *rfbD* were always present within this island even though these alleles did not rise to the level of significance in the genome wide association study. Phylogenetic inference of RfbD clearly showed that sequences of each sensitivity group cluster with each other, and to the exclusion of sequences from the alternative sensitivity group, with strong support. Further, using this placement allowed for relatively high precision for assignment of previously unscreened strains to sensitivity classes. Strain *P. syringae* pv. *helianthii* LMG 5067 is predicted to be part of sensitivity class A by *rfbD* allele classification, but the lack of killing of this strain by either tailocin class likely points to biological nuances in LPS modification. Although RfbD appears critical for the creation of L-rhamnose [46, 47], a main component of the CPA chain of the LPS, it is unclear why there would be two different allelic clades for these genes across *P. syringae* classes and whether enzymes encoded by both clades have similar functionalities for the cell or whether each allele

is optimized to work in concert with additional loci that are associated with each tailocin sensitivity class to encode LPS chains.

An abundance of data pulled together from different sources strongly suggests that the R type syringacins interact with rhamnose moieties that form the core chain of the O-antigen across *P. syringae* isolates. First, the R-type pyocins of *P. aeruginosa* have been shown to interact with rhamnose and glucose moieties in the LPS of target strains [9]. Although tailocins from *P. syringae* have been independently coopted from different phage than the R-pyocins, that these sugars have already been shown to be targeted by some tailoicns sets a prior precedent for evolutionary comparisons and predictions. More directly, studies characterizing the LPS of *P. syringae* demonstrate that the LPS of these strains fundamentally differs from what is known from *P. aeruginosa* and that strains can be classified into two LPS groups determined by whether the O-antigen is composed of repeating units of either D-rhamnose or L-rhamnose [48]. We report here that tailocin sensitivity differentially correlates with enzymes implicated in the synthesis of either D or L-rhamnose, with D-rhamnose chains implicated in sensitivity class A and L-rhamnose chains implicated for sensitivity class B. Evidence from multiple strains also highlights that disruption of the production of D-rhamnose provides resistance to tailocins. Taken together, our results create a compelling narrative that the chirality of rhamnose in the O-antigen of *P. syringae* at least partially determines classification for tailocin sensitivity: Chemical analysis suggests that *P. syringae* create either an L-rhamnose or D-rhamnose based O-antigen, genetic evidence shows that loss of O-antigen leads to tailocin resistance for most *P. syringae* tailocins where evaluated, and that two large scale tailocin sensitivity classes have been identified across *P. syringae* strains [12] and strongly correlate with presence/absence of operons involved in biosynthesis and transport of D-rhamnose and L-rhamnose. A direct link remains unproven at the moment though because, to our knowledge, strains that have been fully characterized for LPS have not had their genomes sequenced or been screened for tailocin sensitivity and strains that have been characterized for tailocin sensitivity have not had their LPS fully characterized. However, assuming that this link does exist, it is worth noting that *P. syringae* pv. *glycinea* R4 is categorized as tailocin sensitivity class A and possesses an O-antigen that is dominated by D-rhamnose moieties [49]. Future work will explore the possibility that the two distinct rhamnose biosynthesis and transport pathways present within *P. syringae* determine whether the LPS is dominated by L or D-rhamnose and will determine whether changing the chirality of the core CPA sugar is sufficient for changing tailocin sensitivity class.

It is no surprise that composition of the LPS is extremely variable between bacterial isolates, as past studies have focused on using LPS characterization (in part) for classification of different lineages of *P. syringae* [48], but relatively few studies have evaluated the genomic basis of LPS diversity across *Pseudomonas syringae* (sensu lato). Our results here suggest that

presence/absence diversity driven by relatively large-scale (~20Kb) recombination events can dramatically and repeatedly alter LPS composition between closely related isolates and that such changes likely drive differences in sensitivity to different classes of tailocins. While the genetic basis of LPS configuration in *P. syringae* differs from canonical pathways established in *E. coli, P. aeruginosa* and other systems, in that the O-antigen appears to be dominated by enantiomers of rhamnose rather than a mix of variable sugar chains in addition to rhamnose, it is unknown how such variation impacts recognition of these strains by the plant immune system. It also remains unknown what correlated phenotypic changes, in the context of behaviors like biofilm formation or in responses to LPS interacting molecules like cationic antimicrobials [50], might occur across *P. syringae* strains with fundamentally different LPS conformations. It is also highly likely that phage attachment and infection also greatly contributes to, and may indeed itself drive, modification of the LPS under natural conditions [40, 51]. Our results highlight the potential for tailocins, or other LPS targeting bacteriocins [52], to be used as tools for uncovering and exploring genetic diversity in pathways and genes that underlie LPS production in these and other strains. Moreover, our study demonstrates that it is possible to uncover genetic and evolutionary patterns in LPS biosynthesis pathways by pairing large-scale phenotypic screens for tailocin sensitivity with genome sequencing, and such protocols could be used broadly to uncover interesting correlations in any system where strains display differential tailocin sensitivity.

There are multiple reports showing that independent recombination events occur at different locations in the chromosome in concert [12, 19], affecting LPS targeting ability and sensitivity across *P. syringae* strains. While there is currently no general mechanism identified to explain to how these independent recombination events co-occur, the coincident nature of these independent events hints that one of if not the main force driving fundamental changes in LPS composition within this phytopathogen over relevant evolutionary time scales is selection to prevent mismatches between tailocin killing capabilities and sensitivity. That these changes happen relatively frequently throughout the *P. syringae* phylogeny also suggests that the ability to diversity tailocin targeting capabilities, whether because of direct benefits during interstrain competition or in order to remedy phenotypic mismatches that may occur if alternative selection pressures shift LPS conformation prior to tailocin targeting, is an important trait for survival of natural *P. syringae* populations.

In sum, our results shed light on the potential genetic basis of fundamental switches in LPS conformation in the *P. syringae,* which appear to be strongly correlated with shifts in sensitivity to phage derived bacteriocions. They suggest that differences in binding of two previously identified tailocin killing types could in part be due to differential binding of rhamnose enantiomers that form the basis of the O-antigen of the LPS of this phytopathogen

and point towards additional genomic loci that can be used to predict LPS and tailocin phenotypes from genomic sequences alone. Lastly, our study provides a clear proof of concept for combining large-scale phenotypic screens of tailocin sensitivity with genome wide association studies to identify the basis of bacteriocin interacting proteins and pathways.

Table 1. Strains Used in this Study

Strain	Abbreviation	Genome Accession	Tailocin Sensitivity Group
Pseudomonas syringae pv. *glycinea* A29-2	*Pgy* R4	ADWY00000000.1	A
Pseudomonas syringae pv. *phaseolicola* 1644	*Pph* 1644	AGAS00000000.1	A
Pseudomonas syringae pv. phaseolicola 1448a	*Pph* 1448a	CP000058.1	A
Pseudomonas syringae pv. *aesculi* 0893_23	*Pae*	AEAD00000000.1	A
Pseudomonas savastanoi	*Psv*	CP008742.1	B
Pseudomonas syringae pv. *mori* MAFF301020	*Pmo*	AEAG00000000.1	B
Pseudomonas syringae pv. *lachrymans* MAFF370057	*Pla* YM7902	LGLI00000000.1	B
Pseudomonas syringae TLP2	TLP2	CP053029.1	B
Pseudomonas syringae CC1458	CC1458	AVEN00000000.2	A
Pseudomonas syringae CC440	CC440	CP093321.1	A
Pseudomonas syringae pv. *japonica* MAFF301072	*Pja*	AEAH00000000.1	A
Pseudomonas syringae pv. *atrofaciens* DSM50255	*Paf*	AWUI00000000.1	A
Pseudomonas syringae CC457	CC457	AVEB00000000.2	B
Pseudomonas syringae pv. *aptata* DSM50252	*Ptt*	AEAN00000000.1	A
Pseudomonas syringae CC1543	CC1543	AVEJ00000000.2	A
Pseudomonas syringae pv. *pisi* 1704B	*Ppi*	AEAI00000000.1	A
Pseudomonas syringae pv. *syringae* B48	B48	TBD	A
Pseudomonas syringae CC94	CC94	AVEA00000000.2	B
Pseudomonas syringae USA011	USA011	CP045799.1	B
Pseudomonas syringae pv. *syringae* B301D	B301D	CP005969.1	A
Pseudomonas syringae UB303	UB303	CP047267.1	B
Pseudomonas syringae pv. *syringae* B728a	*Psy* B728a	CP000075.1	A
Pseudomonas syringae CC1629	CC1629	AVEE00000000.2	B
Pseudomonas syringae CC1513	CC1513	AVEL00000000.2	B
Pseudomonas syringae pv. *oryzae* 1_6	*Por* 1_6	NZ_CP046035.1	B
Pseudomonas syringae pv. *tomato* DC3000	*Pto* DC3000	AE016853.1	B
Pseudomonas syringae pv. *lachrymans* MAFF302278	*Pla* 106	AEAM00000000.1	B
Pseudomonas syringae pv. *tomato* T1	*Pto* T1	ABSM00000000.1	B
Pseudomonas syringae CC1630	CC1630	AVED00000000.2	B
Pseudomonas syringae CC1544	CC1544	AVEI00000000.2	B
Pseudomonas syringae USA007	USA007	AVDY00000000.2	A
Pseudomonas syringae CC1416	CC1416	AVEP00000000.2	B
Pseudomonas syringae pv. *morsprunorum* MAFF302280	*Pmp*	AEAE00000000.1	A
Pseudomonas syringae pv. *actinidiae* ICMP 18884	*Pan* ICMP 18884	NZ_CP011972.2	A
Pseudomonas syringae CC1559	CC1559	AVEG00000000.2	A
Pseudomonas syringae CC1466	CC1466	AVEM00000000.2	B
Pseudomonas syringae CC1557	CC1557	CP007014.1	B
Pseudomonas cannabina pv. *maculicola* ES4326	*Pma* ES4326	NZ_CP047260.1	B
Pseudomonas syringae pv. *helianthii* LMG 5067	*Phel* LMG 5096	NZ_CP092918.1	Unscreened
Pseudomonas syringae pv. *tagetis* IMCP 4091	*Ptg* ICMP 4091	NZ_CP092923.1	Unscreened
Pseudomonas syringae pv. *tomato* 2170	*Pto* 2170	JAPTOE000000000.1	Unscreened
Pseudomonas syringae Cit7	Cit7	CP073636.1	Unscreened
Pseudomonas syringae pv *oryzae* 36_1	*Por* 36_1	NZ_LGKZ00000000.1	Unscreened

Table 2. Results from Genome-Wide Association Analysis

Gene	Annotation in *Psy* B728a	Annotation in USA011	Annotation	# Class A Present In (of 17)	# Class B Present In (of 20)	Bonferroni Corrected p value	Benjamini-Hochberg Corrected p value
rmd	Psyr_0915	N/A	GDP-6-deoxy-D-talose 4-dehydrogenase	13	0	0.027422528	0.002492957
gmd	Psyr_0916	N/A	GDP-mannose 4,6-dehydratase	16	0	5.42E-05	1.35E-05
wzm	Psyr_0917	N/A	Sugar transport protein	15	0	0.000596142	8.52E-05
rfbB	Psyr_0926	N/A	dTDP-glucose 4,6-dehydratase	11	0	0.594154778	0.034950281
Glycosyltransferase	Psyr_0930	N/A	putative glycosyltransferase	12	0	0.137112641	0.00856954
prsD	Psyr_0933	N/A	Type I secretion system ATP-binding protein	13	0	0.027422528	0.002492957
prsE	Psyr_0934	N/A	Type I secretion system membrane fusion protein	13	0	0.027422528	0.002492957
TolC-like	Psyr_0935	N/A	Outer membrane protein	14	0	0.004570421	0.000571303
algA	Psyr_0937	N/A	Alginate biosynthesis protein AlgA	16	0	5.42E-05	1.35E-05
dapH	N/A	N028_RS04635	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase	0	15	0.128170512	0.008544701
Glycosyltransferase	N/A	N028_RS04640	putative glycosyltransferase	0	15	0.128170512	0.008544701
wecE	N/A	N028_RS04645	dTDP-3-amino-3,6-dideoxy-alpha-D-galactopyranose transaminase	0	15	0.128170512	0.008544701
ATPase	N/A	N028_RS04660	Type I secretion system ATPase	0	18	0.000487752	8.13E-05
Permease	N/A	N028_RS04665	ABC transporter permease	0	18	0.000487752	8.13E-05
rfbC	N/A	N028_RS04670	dTDP-4-dehydrorhamnose 3,5-epimerase	0	15	0.128170512	0.008544701
rfbA	N/A	N028_RS04680	Glucose-1-phosphate thymidylyltransferase	0	19	4.65E-05	1.35E-05
rfbB	N/A	N028_RS04690	dTDP-glucose 4,6-dehydratase	0	20	2.58E-06	2.58E-06

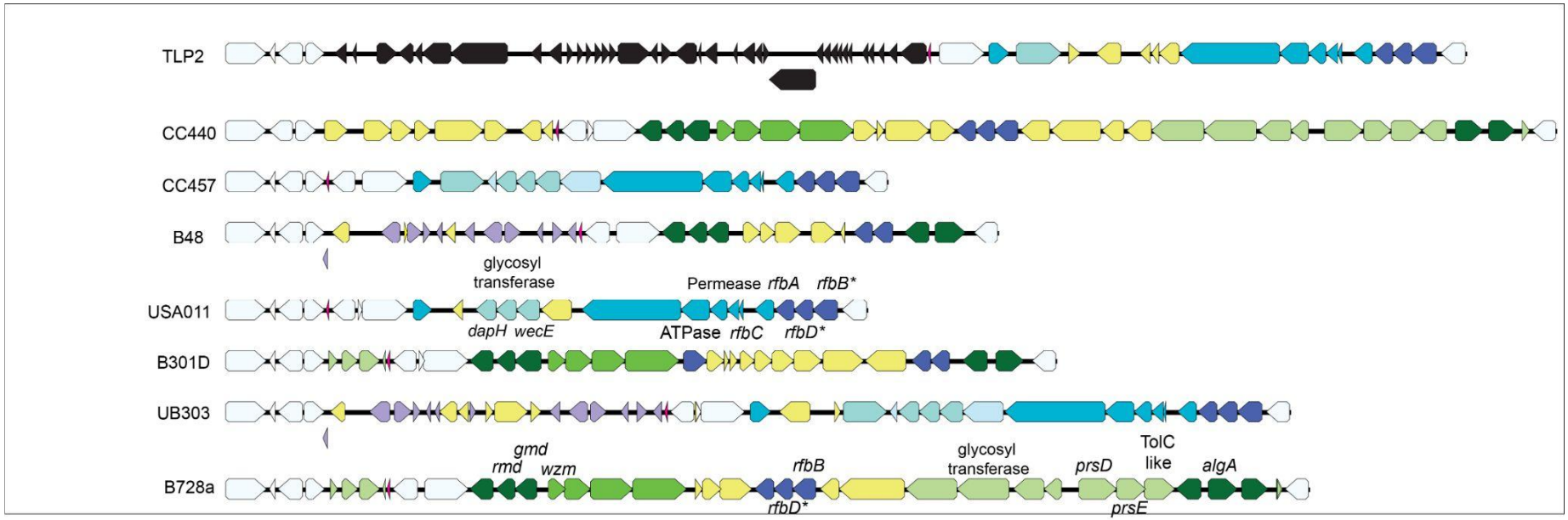
TLP2
CC440
CC457
B48
glycosyl
transferase
Permease
rfbA
rfbB*
USA011
dapH
wecE
ATPase
rfbC
rfbD*
B301D
UB303
gmd
rmd
wzm
rfbB
glycosyl
transferase
prsD
TolC
like
algA
B728a
rfbD*
prsE

Figure 1. O-antigen / Common Polysaccharide Antigen Locus of *P. syringae* Phylogroup 2 Strains. The genomic island found between *ildD-ychF* is shown for 8 phylogroup 2 strains. Strains are arranged vertically in phylogenetic order according to Baltrus et al. 2019, and were picked for the figure because they represent situations where closely related strains have switched tailocin sensitivity class phenotypes. Common loci not thought to be involved in LPS construction are shown in white, and loci that are part of a possible prophage in strain TLP2 are shown in black. Loci that are possibly involved in LPS construction but which are found across tailocin sensitivity class strains are colored purple, with stronger coloration indicating presence in a larger number of strains. Yellow loci are only found in single strains among the eight shown. Loci that are possibly involved in LPS construction but which are found only within tailocin sensitivity class A and B strains are shown in green and blue, respectively, with stronger coloration indicating presence in a larger number of strains. Loci that were significantly associated with tailocin sensitivity class according to our analyses are labelled in strains *Psy*B728a (sensitivity class A) and USA011 (sensitivity class B), and we additionally label *rfbD* since we highlight its use as a predictor of tailocin sensitivity.

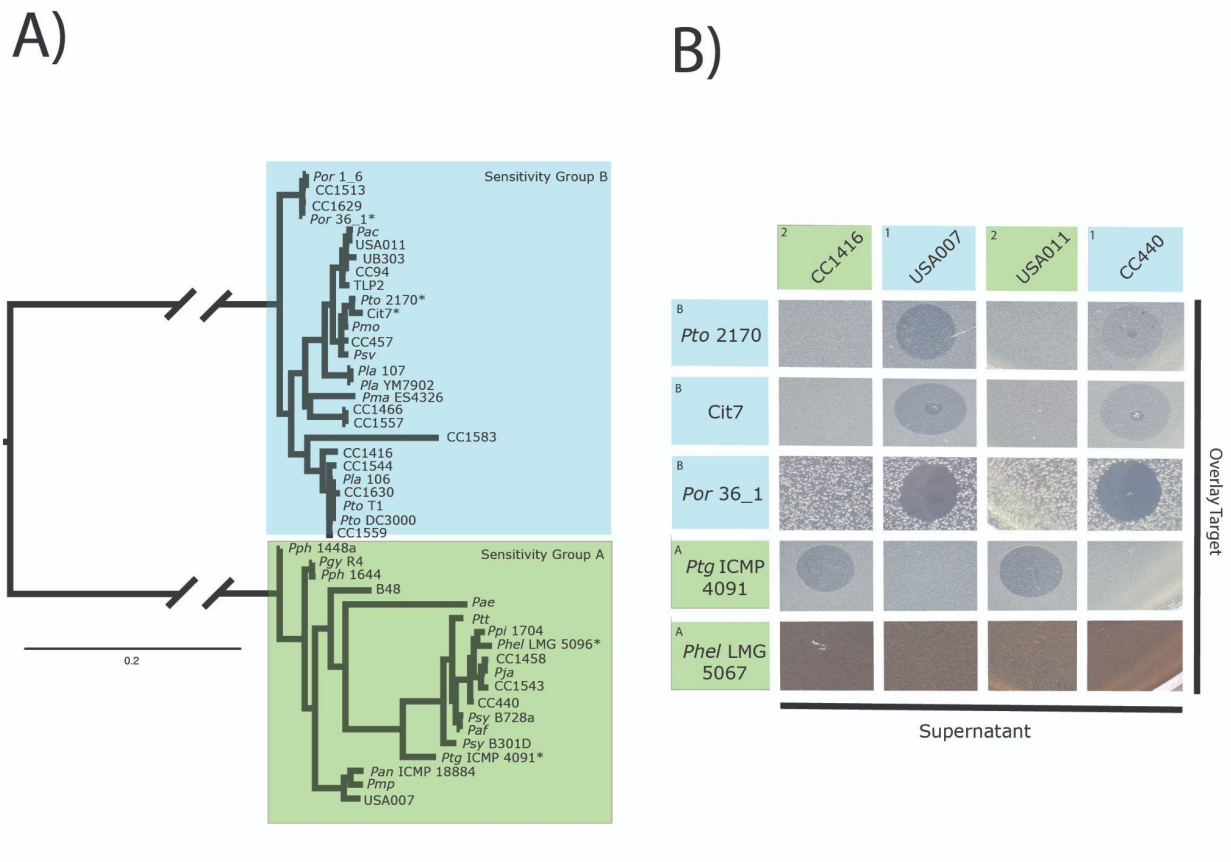

Figure 2. RfbD as a Predictor of Tailocin Sensitivity Classes in *P. syringae*. A) We inferred a maximum likelihood phylogeny from protein sequences of RfbD from across *P. syringae* strains tested for tailocin sensitivity in Baltrus et al. 2019 and from five additional strains which have had their genomes sequenced but which have not been previously screened. Strain names are shown matching abbreviations in Table 1. Clades containing strains previously typed as tailocin sensitivity class A and B are colored green and blue, respectively. We further include alleles of RfbD from previously unscreened strains in the phylogeny, designate these strains with *. All files for phylogenetic analysis can be found in the supplemental data at 10.6084/m9.figshare.22688020. **B)** We performed soft agar overlay assays where tailocins from four different strains and two different phylogenetic groups of *P. syringae* (USA007 and CC440, tailocin killing class 1, blue boxes; CC1416 and USA011, tailocin killing class 2, green boxes) were tested against 5 strains for which tailocin sensitivity has not been previously reported (strains listed in the Y axis, with strain names colored for predicted tailocin sensitivity class as per phylogeny in Figure 2A).

1. **Ghequire MGK, De Mot R**. The Tailocin Tale: Peeling off Phage Tails. *Trends Microbiol* 2015;23:587–590.

2. **Patz S, Becker Y, Richert-Pöggeler KR, Berger B, Ruppel S, *et al.*** Phage tail-like particles are versatile bacterial nanomachines - A mini-review. *J Advert Res* 2019;19:75–84.

3. **Scholl D**. Phage Tail-Like Bacteriocins. *Annu Rev Virol* 2017;4:453–467.

4. **Príncipe A, Fernandez M, Torasso M, Godino A, Fischer S**. Effectiveness of tailocins produced by Pseudomonas fluorescens SF4c in controlling the bacterial-spot disease in tomatoes caused by Xanthomonas vesicatoria. *Microbiol Res* 2018;212–213:94–102.

5. **Baltrus DA, Clark M, Hockett KL, Mollico M, Smith C, *et al.*** Prophylactic Application of Tailocins Prevents Infection by Pseudomonas syringae. *Phytopathology* 2022;112:561–566.

6. **Hockett KL, Renner T, Baltrus DA**. Independent Co-Option of a Tailed Bacteriophage into a Killing Complex in Pseudomonas. *MBio* 2015;6:e00452.

7. **Ge P, Scholl D, Prokhorov NS, Avaylon J, Shneider MM, *et al.*** Action of a minimal contractile bactericidal nanomachine. *Nature* 2020;580:658–662.

8. **Williams SR, Gebhart D, Martin DW, Scholl D**. Retargeting R-type pyocins to generate novel bactericidal protein complexes. *Appl Environ Microbiol* 2008;74:3868–3876.

9. **Köhler T, Donner V, van Delden C**. Lipopolysaccharide as shield and receptor for R-pyocin-mediated killing in Pseudomonas aeruginosa. *J Bacteriol* 2010;192:1921–1928.

10. **Scholl D, Martin DW Jr**. Antibacterial efficacy of R-type pyocins towards Pseudomonas aeruginosa in a murine peritonitis model. *Antimicrob Agents Chemother* 2008;52:1647–1652.

11. **Salazar AJ, Sherekar M, Tsai J, Sacchettini JC**. R pyocin tail fiber structure reveals a receptor-binding domain with a lectin fold. *PLoS One* 2019;14:e0211432.

12. **Baltrus DA, Clark M, Smith C, Hockett KL**. Localized recombination drives diversification of killing spectra for phage-derived syringacins. *ISME J*. Epub ahead of print August 31, 2018. DOI: 10.1038/s41396-018-0261-3.

13. **Bertani B, Ruiz N**. Function and biogenesis of lipopolysaccharides. *EcoSal Plus*;8. Epub ahead of print August 2018. DOI: 10.1128/ecosalplus.ESP-0001-2018.

14. **King JD, Kocíncová D, Westman EL, Lam JS**. Review: Lipopolysaccharide biosynthesis in Pseudomonas aeruginosa. *Innate Immun* 2009;15:261–312.

15. **Lam JS, Taylor VL, Islam ST, Hao Y, Kocíncová D**. Genetic and Functional Diversity of Pseudomonas aeruginosa Lipopolysaccharide. *Front Microbiol* 2011;2:118.

16. **Pier GB**. Pseudomonas aeruginosa lipopolysaccharide: a major virulence factor, initiator of inflammation and target for effective immunity. *Int J Med Microbiol* 2007;297:277–295.

17. **Heiman CM, Maurhofer M, Calderon S, Dupasquier M, Marquis J, *et al.*** Pivotal role of O-antigenic polysaccharide display in the sensitivity against phage tail-like particles in environmental Pseudomonas kin competition. *ISME J* 2022;16:1683–1693.

18. **Carim S, Azadeh AL, Kazakov AE, Price MN, Walian PJ, *et al.*** Systematic discovery of pseudomonad genetic factors involved in sensitivity to tailocins. *ISME J* 2021;15:2289–2305.

19. **Jayaraman J, Jones WT, Harvey D, Hemara LM, McCann HC, *et al.*** Variation at the common polysaccharide antigen locus drives lipopolysaccharide diversity within the Pseudomonas syringae species complex. *Environ Microbiol* 2020;22:5356–5372.

20. **Hockett KL, Clark M, Scott S, Baltrus DA**. Conditionally Redundant Bacteriocin Targeting by Pseudomonas syringae. *bioRxiv* 2017;167593.

21. **Seemann T**. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.

22. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, *et al.*** Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.

23. **Brynildsrud O, Bohlin J, Scheffer L, Eldholm V**. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* 2016;17:238.

24. **Sievers F, Higgins DG**. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 2018;27:135–145.

25. **Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, *et al.*** ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 2020;37:291–294.

26. **Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A**. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019;35:4453–4455.

27. **Hockett KL, Baltrus DA**. Use of the Soft-agar Overlay Technique to Screen for Bacterially Produced Inhibitory Compounds. *J Vis Exp*. Epub ahead of print January 14, 2017. DOI: 10.3791/55064.

28. **Dillon MM, Thakur S, Almeida RND, Wang PW, Weir BS, *et al.*** Recombination of ecologically and evolutionarily significant loci maintains genetic cohesion in the Pseudomonas syringae species complex. *Genome Biol* 2019;20:3.

29. **Meyer A, Pühler A, Niehaus K**. The lipopolysaccharides of the phytopathogen Xanthomonas campestris pv. campestris induce an oxidative burst reaction in cell cultures of Nicotiana tabacum. *Planta* 2001;213:214–222.

30. **Newman MA, von Roepenack E, Daniels M, Dow M**. Lipopolysaccharides and plant responses to phytopathogenic bacteria. *Mol Plant Pathol* 2000;1:25–31.

31. **King JD, Poon KKH, Webb NA, Anderson EM, McNally DJ, *et al.*** The structural basis for catalytic function of GMD and RMD, two closely related enzymes from the GDP-D-rhamnose biosynthesis pathway. *FEBS J* 2009;276:2686–2700.

32. **Wang S, Hao Y, Lam JS, Vlahakis JZ, Szarek WA, *et al.*** Biosynthesis of the common polysaccharide antigen of Pseudomonas aeruginosa PAO1: Characterization and role of GDP-D-rhamnose:GlcNAc/GalNAc-diphosphate-lipid α1,3-D-rhamnosyltransferase WbpZ. *J Bacteriol* 2015;197:2012–2019.

33. **Rocchetta HL, Pacan JC, Lam JS**. Synthesis of the A-band polysaccharide sugar D-rhamnose requires Rmd and WbpW: identification of multiple AlgA homologues, WbpW and ORF488, in Pseudomonas aeruginosa. *Mol Microbiol* 1998;29:1419–1434.

34. **Kandel PP, Baltrus DA, Hockett KL**. Pseudomonas Can Survive Tailocin Killing via Persistence-Like and Heterogenous Resistance Mechanisms. *J Bacteriol*;202. Epub ahead of print June 9, 2020. DOI: 10.1128/JB.00142-20.

35. **Rocchetta HL, Burrows LL, Pacan JC, Lam JS**. Three rhamnosyltransferases responsible for assembly of the A-band D-rhamnan polysaccharide in Pseudomonas aeruginosa: a fourth transferase, WbpL, is required for the initiation of both A-band and B-band lipopolysaccharide synthesis. *Mol Microbiol* 1998;28:1103–1119.

36. **Marolda CL, Valvano MA**. Genetic analysis of the dTDP-rhamnose biosynthesis region of the Escherichia coli VW187 (O7:K1) rfb gene cluster: identification of functional homologs of rfbB and rfbA in the rff cluster and correct location of the rffE gene. *J Bacteriol* 1995;177:5539–5546.

37. **Sivaraman J, Sauvé V, Matte A, Cygler M**. Crystal structure of Escherichia coli glucose-1-phosphate thymidylyltransferase (RffH) complexed with dTTP and Mg2+. *J Biol Chem* 2002;277:44214–44219.

38. **Rocchetta HL, Burrows LL, Lam JS**. Genetics of O-antigen biosynthesis in Pseudomonas aeruginosa. *Microbiol Mol Biol Rev* 1999;63:523–553.

39. **Hwang B-Y, Lee H-J, Yang Y-H, Joo H-S, Kim B-G**. Characterization and investigation of substrate specificity of the sugar aminotransferase WecE from E. coli K12. *Chem Biol* 2004;11:915–925.

40. **Warring SL, Malone LM, Jayaraman J, Easingwood RA, Rigano LA, *et al.*** A lipopolysaccharide-dependent phage infects a pseudomonad phytopathogen and can evolve to evade phage resistance. *Environ Microbiol* 2022;24:4834–4852.

41. **Russo DM, Williams A, Edwards A, Posadas DM, Finnie C, *et al.*** Proteins exported via the PrsD-PrsE type I secretion system and the acidic exopolysaccharide are involved in biofilm formation by Rhizobium leguminosarum. *J Bacteriol* 2006;188:4474–4486.

42. **Vakharia H, German GJ, Misra R**. Isolation and characterization of Escherichia coli tolC mutants defective in secreting enzymatically active alpha-hemolysin. *J Bacteriol* 2001;183:6908–6916.

43. **Cote JM, Taylor EA**. The glycosyltransferases of LPS core: A review of four heptosyltransferase enzymes in context. *Int J Mol Sci*;18. Epub ahead of print October 27, 2017. DOI: 10.3390/ijms18112256.

44. **Whitfield C, Williams DM, Kelly SD**. Lipopolysaccharide O-antigens-bacterial glycans made to measure. *J Biol Chem* 2020;295:10593–10609.

45. **Beaman TW, Blanchard JS, Roderick SL**. The conformational change and active site structure of tetrahydrodipicolinate N-succinyltransferase. *Biochemistry* 1998;37:10363–10369.

46. **Lerouge I, Vanderleyden J**. O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. *FEMS Microbiol Rev* 2002;26:17–47.

47. **Nian H, Zhang J, Song F, Fan L, Huang D**. Isolation of transposon mutants and characterization of genes involved in biofilm formation by Pseudomonas fluorescens TC222. *Arch Microbiol* 2007;188:205–213.

48. **Zdorovenko GM, Zdorovenko EL**. Pseudomonas syringae lipopolysaccharides: Immunochemical characteristics and structure as a basis for strain classification. *Microbiology* 2010;79:47–57.

49. **Chiku K, Tsunemi K, Yamamoto M, Ohnishi-Kameyama M, Yoshida M, *et al.*** Defects in D-rhamnosyl residue biosynthetic genes affect lipopolysaccharide structure, motility, and cell-surface hydrophobicity in Pseudomonas syringae pathovar glycinea race 4. *Biosci Biotechnol Biochem* 2013;77:505–510.

50. **Huszczynski SM, Lam JS, Khursigara CM**. The role of Pseudomonas aeruginosa lipopolysaccharide in bacterial pathogenesis and physiology. *Pathogens* 2019;9:6.

51. **Meaden S, Paszkiewicz K, Koskella B**. The cost of phage resistance in a plant pathogenic bacterium is context-dependent. *Evolution* 2015;69:1321–1328.

52. **McCaughey LC, Grinter R, Josts I, Roszak AW, Waløen KI, *et al.*** Lectin-like bacteriocins from Pseudomonas spp. utilise D-rhamnose containing lipopolysaccharide as a cellular receptor. *PLoS Pathog* 2014;10:e1003898.