Phage-mediated resolution of genetic conflict alters the evolutionary trajectory of *Pseudomonas aeruginosa* lysogens

Authors: Laura C. Suttenfield[1,2], Zoi Rapti[2,3], Jayadevi H. Chandrashekhar[2], Amelia C. Steinlein[1,2], Juan Cristobal Vera[2], Ted Kim[1,2], Rachel J. Whitaker[1,2*]

1. Department of Microbiology, School of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois
2. Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois
3. Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois

*Correspondence: rwhitakr@illinois.edu

Abstract

The opportunistic human pathogen *Pseudomonas aeruginosa* is naturally infected by a large class of temperate, transposable, Mu-like phages. We examined the genotypic and phenotypic diversity of *P. aeruginosa* PA14 populations as they resolve CRISPR autoimmunity, mediated by an imperfect CRISPR match to the Mu-like DMS3 prophage, and show that lysogen evolution is profoundly impacted by CRISPR autoimmunity and phage transposition around the chromosome. After 12 days of evolution, we measured a decrease in spontaneous induction in both exponential and stationary phase growth. Co-existing variation in spontaneous induction rates in exponential phase corresponded to a difference in the type of CRISPR self-targeting resolution, mediated either by host mutation or phage transposition. Multiple mutational modes to resolve genetic conflict between host and phage resulted in coexistence in evolved populations of single lysogens that maintained CRISPR immunity to other phages and polylysogens that have lost immunity completely. This work highlights a new dimension of the role of lysogenic phages in the evolution of their hosts.

Importance

The chronic opportunistic multi-drug resistant pathogen *Pseudomonas aeruginosa* is persistently infected by temperate phages. We assess the contribution of temperate phage infection to the evolution of the clinically relevant strain UCBPP-PA14. We found that a low level of CRISPR-mediated self-targeting resulted in polylysogeny evolution and large genome rearrangements in lysogens; we also found extensive diversification in CRISPR spacers and *cas* genes. These genomic modifications resulted in decreased spontaneous induction in both exponential and stationary phase growth, increasing lysogen fitness. This work shows the importance of considering latent phage infection in characterizing the evolution of bacterial populations.

Keywords: lysogen; spontaneous induction; evolution; CRISPR; transposable phages

Introduction

Cystic fibrosis (CF) is a genetic disorder which makes patients vulnerable to respiratory infections by commensal and environmental bacterial pathogens like *Pseudomonas aeruginosa*. *Pseudomonas aeruginosa* is a common human pathogen whose increase in multi-drug antibiotic resistance has made it the focus of targeted phage therapy (Kortright et al., 2019). In chronic *Pseudomonas* infections of CF patients, it is common to find that all isolates track their origin to a single ancestral genotype (Feliziani et al., 2014; Folkesson et al., 2012; Jorth et al., 2015; Markussen et al., 2014; Schick & Kassen, 2018; Tai et al., 2017; Workentine et al., 2013). Diversity of *P. aeruginosa* strains diverge from this common ancestor through *de novo* mutations generated by mutation and hypermutation genotypes (Vanderwoude et al., 2023), recombination (Darch et al., 2015; Vanderwoude et al., 2023) and large deletions (Cramer et al., 2011; Jeukens et al., 2014; Rau et al., 2012).

Pseudomonas aeruginosa is commonly infected by latent phages (bacteriophage) when colonizing CF patients (Budzik et al., 2004; Burgener et al., 2019; Folkesson et al., 2012; Ojeniyi et al., 1991; Rossi et al., 2021; Tariq et al., 2019; Vanderwoude et al., 2023; Winstanley et al., 2016). Temperate and chronic phages act as both sources of genetic novelty (Brüssow et al., 2004) and as potential assassins that can be induced to kill their hosts (James et al., 2015; Willner et al., 2012). In the lung environment, the presence of antibiotics and reactive oxygen species (Kettle et al., 2014; Malhotra et al., 2019; McGrath et al., 1999) may act as inducing agents for phage (Bondy-Denomy et al., 2016; Fothergill et al., 2011; James et al., 2015; Nanda et al., 2015; Rolain et al., 2009). Bacterial lysis through phage induction is hypothesized to help control bacterial growth in the lung (James et al., 2015) and may be used in synergy with antibiotics (Al-Anany et al., 2021; Clifton et al., 2019). Additionally, lysogeny has been shown to co-occur with host genome rearrangements in chronically infecting *Staphylococcus aureus* (Goerke et al., 2004, 2006; Golubchik et al., 2013; Guérillot et al., 2019), and *Streptococcus pyogenes* (Nakagawa et al., 2003). However, it remains unclear if and how lysogeny alters the evolution of the host genome.

Mu-like transposable phages are a diverse family of phages which infect an equally diverse range of bacteria (Zhang et al., 2023). Upon infection of the host, Mu-like phages integrate into the host chromosome through a conservative (cut-and-paste) transposition step that occurs with low sequence preference (Chaconas & Harshey, 2007; Vergnaud et al., 2018; Walker et al., 2020). Lytic replication occurs via replicative (copy-and-paste) transposition around the genome (Walker & Harshey, 2020), which occurs approximately 100 times and terminates in headful packaging into the virion. In lysogeny, which is established in approximately 10% of infections, a low-specificity insertion can increase the variation available to natural selection in *P. aeruginosa* populations through knockout mutations (Davies et al., 2016; O'Brien et al., 2019; Rollie et al., 2020). These *P. aeruginosa* lysogens have previously been found to have high viral titers in culture (Bondy-Denomy et al., 2016; James et al., 2012). Due to the nature of the chemistry that governs the transposition reaction, these insertions may also cause structural rearrangements such as deletions (Toussaint & Rice, 2017).

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) is a bacterial and archaeal adaptive immune system which incorporates foreign DNA fragments into an array as a spacer, and subsequently targets any piece of invading DNA (the protospacer) which is complementary to the spacer (Barrangou et al., 2007; Vorontsova et al., 2015). The Mu-like temperate phage are the most commonly targeted phages by CRISPR-Cas in *Pseudomonas aeruginosa* (England et al., 2018). Phage DMS3, a member of this group, was recovered from a *P. aeruginosa* CF isolate and infects the type strain UCBPP-PA14 (Budzik et al., 2004). DMS3 inhibits quorum sensing and pilus formation in PA14 lysogens (Shah et al., 2021). PA14 contains a Type 1-F CRISPR system with a partial spacer match to DMS3 (Zegans et al., 2009). This spacer has 5 mismatches to the phage protospacer, which is not sufficient to mediate immunity to the phage but leads to genetic conflict in DMS3 lysogens. This degenerate protospacer-spacer mismatch between DMS3 and PA14 targets the PA14 chromosome, causing enough DNA damage to stimulate the SOS response, which leads to the expression of pyocin genes, cell death, and limitation of biofilm formation (Heussler et al., 2015). Lysogens arising from PA14 cultures infected with free DMS3 virions evolved to have a lower spontaneous induction in stationary phase and lost their CRISPR systems over a seven-day evolutionary period. This was suggested to resolve genetic conflict caused by CRISPR self-targeting (immunopathology), a phenomenon which is predicted to be common in bacteria with Type 1 CRISPR systems and temperate phages (Rollie et al., 2020). However, how lysogeny alone impacts the bacterial genome during evolution in the absence of new phage infection has not yet been explored.

Here we directly assess the contribution of CRISPR-mediated genetic conflict between host and temperate phage to the evolution of *Pseudomonas* by analyzing evolved lysogen populations. We show that selection to resolve genetic conflict alters the evolutionary landscape of lysogen populations. Experimental work combined with genomic analysis demonstrates that transposable phages are a major source of variation beyond mutation that impacts the evolutionary direction of *P. aeruginosa* lysogens.

Methods

Experimental evolution

To establish the contribution of phages to host genome evolution, we evolved the uninfected laboratory strain UCBPP-PA14 (Rahme et al., 1995) and the established lysogen Lys2 (Zegans et al., 2009) for 12 days by serial transfer. Our strains are listed in Table S1. Lys2 is derived from PA14 and contains DMS3, which mediated a ~20 kb host deletion from 806,169 to 826,108 bp on our PA14 reference chromosome, and a single A<G nonsynonymous point mutation at position 4,755,306 in a hypothetical protein (deleted genes are listed in Table S2). For each strain we randomly selected three colonies and grew up independent overnight cultures in LB (10 g tryptone; 5 g yeast extract; 5 g sodium chloride per liter of water). We subcultured the cells and normalized to an OD_{600} of 0.2 in 25 mL microcosms in three parallel 250 mL flasks. We serially passaged triplicate cultures with daily 1:100 transfers for 12 days, with shaking and at 37°C, for approximately 72 bacterial generations. At the end of the experiment, we colony-purified six colonies isolated from each

replicate, and six colonies from the Lys2 ancestor stock for further analysis. Six isolates from each uninfected population were also sequenced and the data is presented in Supplementary Figure 1.

One-step growth curves

In order to calculate the burst size of DMS3, we performed one-step growth curves (Ellis & Delbrück, 1939). Overnight, stationary phase PA14 cultures were diluted 1:100 and grown in LB until they reached an OD of 0.5. Phage was added to these cultures at a 1:10 volume ratio, for a final MOI of 0.01, mixed well, and incubated at 37°C for 5 minutes. In order to calculate adsorption, a fraction of the sample volume was immediately spun down for 5 mins at 30,000 rpm, and the supernatant was stored with 10% v/v chloroform to quantify the remaining free phage. To begin the one-step growth curve, the remainder of the samples were added to pre-warmed medium at a 1:100 ratio to stop adsorption, and grown on a roller drum at 37°C for the next 2.5 hours. Samples were taken every 10 minutes and mixed with a 10% v/v of chloroform for later quantification of PFUs.

We calculated burst size with the following equation: the total phage produced (the difference between the average PFUs before and after the burst) was divided by number of cells that were infected (estimated by the number of adsorbed phages multiplied by the estimated number of cells which proceeded through lysis). Our measurements of an 8% lysogenization frequency, estimated by spot-on-lawn assay, corresponded to measures in (Cady & O'Toole, 2011) (data not shown).

$$\beta = \frac{p_2 - p_1}{c_i * (1 - l)}$$

Here, β is burst size; p_2 and p_1 are the second and first plateaus, respectively; c_i is the number of infected cells; l is the lysogenization frequency. The burst size of DMS3 is 41.8 +/- 8.4 phages per lysed cell (Fig S2). All PFUs were enumerated by spotting the phage-containing fraction on 0.5% double agar overlay plates.

Spontaneous induction measurements

Because multiple inputs could contribute to a higher raw PFU value in stationary phase, and because stationary phase itself could be an inducing condition for some viruses (including phage Mu) (Ranquet et al., 2005), we chose to measure spontaneous induction in exponential phase, which necessitated the normalization of PFU values with the CFU values. This was also a desirable method for separating whether increased growth rate was responsible for increased PFU values. Due to these factors, we chose to measure spontaneous induction separately in both exponential and stationary phase. Our metric corresponds to other methods used to estimate these rates (Zong et al., 2010).

Growth curves were started from overnight cultures of replicate purified isolates. We washed cultures three times by resuspension in fresh LB media, normalized the OD to 0.2, and diluted them 1000-fold. Time points were taken at 0 and 2-7 hours to capture exponential phase growth, and plated for both CFUs and PFUs. Samples were incubated at 37°C on a roller drum. We measured spontaneous induction (q) by taking the

difference in the number of viral particles released by cells growing in exponential phase, and normalized by the estimated burst size, the average growth of the culture, and the total time the culture grew.

$$q = \frac{\Delta V}{(\beta * \Delta t * C)}$$

Here, q is spontaneous induction, and has units of burst cells per time; ΔV is the total increase in virion particles; β is the burst size; Δt is the change in time; C is the average amount of cells in the culture. To account for exponential growth, all calculations were based on the linear regression of the $\log_{10}$ transformed ΔV and C values.

We also calculated spontaneous induction based on calculations in (Zong et al., 2010). This paper approximated spontaneous induction at each time point with this formula: $q = \frac{V}{\beta C}$; here, V is the number of virion particles at that time; β is the burst size; and C is the average amount of CFUs at that time. These methods produced qualitatively similar results (Fig S3).

To measure induction in stationary phase, we began growth curves in the same way as the exponential phase measurements. Time points were taken at 0, 10, 12, 15, and 18 hours, and plated for both CFUs and PFUs. To calculate the rate of spontaneous induction, we used the same formula as above, except with un-regressed logged values.

Genome sequencing

In order to create a viral reference, DMS3 DNA was extracted from filtered Lys2 overnight supernatant using the Phage DNA Isolation Kit (Norgen Biotek Corp, Cat #: 46800), following the manufacturer's instructions. Libraries were prepared with a Biomek 4000 liquid handler (Beckman-Coulter). We quantitated libraries with a Qubit fluorometer (Life Technologies Corporation, REF #: Q32866). Libraries were pooled and submitted for 2x250 paired-end sequencing by the Roy J. Carver Biotechnology Center at the University of Illinois Urbana-Champaign with an Illumina NovaSeq 6000. We received ~3.8 million reads with about 100X coverage.

We inoculated evolved isolates and ancestral controls in 2 mL deep well plates in LB and grew them overnight. We extracted gDNA with the Beckman-Coulter gDNA extraction kit as above using the Nextera Flex Library Preparation Kit (Illumina). We quantitated libraries with a Qubit fluorometer (Life Technologies Corporation, REF #: Q32866). Libraries were pooled and submitted for 2x250 paired-end sequencing by the Roy J. Carver Biotechnology Center at the University of Illinois Urbana-Champaign with an Illumina NovaSeq 6000. We received an average of about 5 million reads per genome. All raw reads are available on the NCBI database under BioProject number PRJNA1021667.

Genome analysis

We ran a custom QC pipeline on our raw FASTQ reads, available on Github (http://www.github.com/igoh-illinois). Briefly, the Illumina adaptor sequences were trimmed using TrimmomaticPE v0. Read quality was checked with FastQC v0.11.9 (options: --noextract -k 5 -f fastq). Reads were aligned using BWA-MEM (Li, 2013) with default options to a 2-contig reference genome containing both our reference PA14 sequence and our reference DMS3 sequence. To identify chromosomal mutations, we ran Breseq (Deatherage & Barrick, 2014) on the trimmed and quality-controlled reads. SAM files were checked manually in both IGV 2.12.3 (J. T. Robinson et al., 2017) and Tablet (Milne et al., 2013).

To identify insertion sites of transposable phage, we ran a second pipeline, available on Github (http://www.github.com/igoh-illinois). The pipeline identifies insertion sites at nucleotide resolution by identifying reads that map to both the host and viral chromosome ("split" reads). It further identifies insertion sites by finding reads that have been split on either side of a 5-bp window, creating a small overlapping region when mapped back to the host genome. This is the result of a 5-bp duplication, which is characteristic of Mu and Mu-like phage transposition (Allet, 1979; Morgan et al., 2002).

While manually verifying the phage insertion sites in IGV, we found putative large duplicated regions of the host chromosome. To verify these duplicated regions we used the depth command in Samtools (Danecek et al., 2021) to find the number of reads that covered each position in the genome, and graphed this using RStudio (R version 4.3.1) (R Core Team, 2023). In order to assess the protein content of deleted and duplicated regions, FASTA files of sequence the reference PA14 genome was given to the eggNOG-mapper-v2 pipeline (settings: genomic data; default options) (Cantalapiedra et al., 2021).

We used the CRISPR Comparison Tool Kit (CCTK v1.0.0) to identify and compare CRISPR arrays (settings: crisprdiff; default options) (Collins & Whitaker, 2022).

Mitomycin C induction experiments

Single colonies of the strains of interest were inoculated in LB, grown overnight at 37°C on a roller drum, and subcultured until they reached an OD of 0.5. Cultures were normalized, split, and incubated with or without 0.5 µg/µL mitomycin C (MMC) for 3.5 hours, after which CFUs and PFUs were enumerated.

Model information

We use a compartmental model based on a system of ordinary differential equations. There are six lysogen compartments each representing a lysogen characterized by a distinct rate of spontaneous induction, but otherwise being identical. Lysogens are induced at their associated rates and transition into the lytic state which is followed by phage production and bursting. The model is along the lines of the ones presented and analyzed in our previous works (Clifton et al., 2019, 2021; Landa et al., 2021). All lysogens are assumed to grow at the same rate, and the total bacterial population grows logistically. The model equations read:

$$\frac{dL_i}{dt} = rL_i\left(1 - \sum_{j=1}^{6}(L_j + I_j)\right) - q_i L_i$$

$$\frac{dI_i}{dt} = -\delta I_i + q_i L_i$$

$$\frac{dV_i}{dt} = \beta\delta I_i - \mu V_i$$

We have partially non-dimensionalized the model so that 1 time unit in the simulations corresponds to about 50 minutes. The lysogeny growth rate is denoted by r, the spontaneous induction rates by $q_i, i = 1, \ldots, 6$, the rate of phage production by δ, the burst size is by β, and the rate of viral degradation by μ. The first two equations are decoupled from the last one describing the phage (since we do not consider superinfections in this model). Therefore, the dynamics of the bacterial compartments resemble those of generalized Lotka-Volterra competition.

Statistical measures

Data visualization and statistical analyses were performed in R version 4.3.1 (R Core Team, 2023) using the packages tidyverse version 2.0.0 (Wickham et al., 2019), car version 3.1-2 (Fox & Weisberg, 2019), rstatix version 0.7.2, and emmeans version 1.8.6.

Results

Evolution of a self-targeting lysogen results in decreased spontaneous induction in exponential and stationary phase.

After 12 days and ~72 generations of exponential growth in rich media, we found that the spontaneous induction of DMS3 lysogens was significantly reduced compared to ancestral isolates in exponential phase (Fig 1; ANOVA, $F_{3,68} = 16.7$, $P < 1e\text{-}8$). Isolate induction rates within and between experimental replicates ranged from 0.1% to 0.71% of the culture, while the induction rates of ancestral PA14 lysogen (Lys2) isolates ranged from 0.38% to 1.1% of the culture (Fig 1A). Significant differences were robust to the use of other metrics to estimate spontaneous induction (Zong et al., 2010) (see Fig S3). Spontaneous induction was also significantly decreased in stationary phase for all evolved lysogens (Fig S4; ANOVA, $F_{3,109} = 86.98$, $P < 2.2e\text{-}16$), although compared to exponential phase, induction was very reduced in stationary phase overall, indicating that the majority of spontaneous induction takes place in exponential phase in DMS3 lysogens.

Lysogen populations maintain diversity in CRISPR presence and function.

Sequencing of lysogens from each population showed that lysogens evolved the CRISPR locus through a combination of mutation and large deletions and other structural variants (Table 1, Fig 2). These mutations did

not overlap with uninfected evolved populations, which exhibited point mutations in flagellar and quorum sensing loci (Fig S1, Table S3), typical of other laboratory evolution experiments (Schick et al., 2022). 39% (7/18) of isolates mutated or deleted the mismatched spacer in the CRISPR2 array, which contains the mismatched spacer which targets the integrated DMS3 (Fig 2B, Fig S5). Two mutations occurred in parallel between two different replicate populations in the evolved lysogen treatment: an A to G point mutation in the seed region of the self-targeting spacer, and an exact deletion of the self-targeting spacer and its upstream repeat (Fig 2B, Fig S6). 22% (4/18) of isolates had disruptions (three independent frameshift mutations and a small deletion) within the *cas* genes *cas7* and *cas8* which form part of the complex that mediates interference in *P. aeruginosa* (Chowdhury et al., 2017). Of these three frameshift mutations, one led to the predicted loss of the *cas7* RNA-binding domain, and two are predicted to interfere with the cas3-recruitment or interaction domain of cas8 (Fig 2B, Table 1). One strain was recovered with a 3 kb deletion in the *cas* gene region that spans *cas5* (also part of the interference complex) and *cas7*. While each strain with mutations and indels in the CRISPR array likely no longer target the DMS3 prophage, they maintain CRISPR function. We confirmed that the evolved lysogens did not acquire new spacers (Fig S6).

We observed that 44% (8/18) of evolved lysogen isolates carried deletions of varying sizes; the smallest being the 3 kb deletion in the *cas* genes, and the largest deletion, 335,331 bp, comprising about 5% of the PA14 chromosome. Of these eight deletions, seven were centered on the CRISPR2 array (Fig 2C). Therefore, in the evolved lysogen populations, we observed extensive coexistence between combinations of CRISPR spacer mutations and large entire deletions of CRISPR, demonstrating the importance of phage infection in determining distinct evolutionary trajectories in isolates in the same environment.

Evolved lysogens with large deletions are polylysogens.

While confirming that the evolved lysogens had retained the phage at its original integration site (Methods), we found that the boundaries of the deletions of the CRISPR regions were composed of reads that mapped to both the PA14 and DMS3 chromosomes (hereafter referred to as “split” reads), indicating that the large deletions in these seven isolates resulted from a DMS3 transposition event which occurred from within the chromosome (Fig 2C). Therefore, we consider these deletions to be phage-mediated. These phage-mediated deletions ranged from approximately 34-335 kb (mean = 163.3 kb ± 100.3 kb). Deletions occurred in each of the replicates independently, with no shared deletion boundaries even within the same culture (Fig 2C). The regions of the phage chromosome to which the split reads mapped and the orientation of phage reads at the boundaries of the deletions in two samples (1_4 and 2_3) suggests that more than one phage genome may be inserted in the gap (Fig 3D-E, Table 2). Notably, we did not recover any mutations in the phage chromosome.

The large deletions, though centered around the CRISPR locus, had different deletion boundaries. To assess the shared gene content in these regions, we used eggNOG-mapper to query the functional protein content that was lost (Methods). In addition to CRISPR, the deleted regions were enriched with genes from COG category S (genes of unknown function), which included Type 6 secretion system-related genes *tssF* and *tssG*,

and multiple pyoverdine system genes which are often lost during lung colonization (Nguyen et al., 2014; O'Brien et al., 2017; Schick & Kassen, 2018). Several pyoverdine biosynthesis and transport genes (*fpvA, pvdE, pvdH, pvdI, pvdM, pvdM, pvdO*) were deleted in 6 of the 7 isolates with large deletions. *PvdS,* a regulator of pyoverdine biosynthesis genes (Leoni et al., 2000), and *pvdR,* a component of the pyoverdine efflux transporter (Imperi et al., 2009), were also deleted in 5 of the 7 isolates with large deletions. The co-localization of these virulence and defense gene cassettes with CRISPR may contribute to variation in these regions (Makarova et al., 2011; L. A. Robinson et al., 2023).

In this study, 100% (7/7) of isolates with large deletions were polylysogens; however, polylysogeny did not evolve in cells that resolved self-targeting via CRISPR mutations from SNPs or indels (Fig S7). Similarly, a previous study showed that PA14, when challenged with free DMS3 virions and subjected to a short-term evolution experiment, evolved genome deletions encompassing the CRISPR region when the susceptible host did not have self-targeting (e.g. contained a *cas7* deletion) (see: Extended Data Table 1 in Rollie *et al*, 2020). In combination with our data, we infer that self-targeting increases the rate of DMS3 transposition, resulting in polylysogens, which then are maintained in the population only when they are associated with adaptive mutations, such as those which decrease spontaneous induction by resolving genetic conflict (Fig S7; for model, see Fig 6).

Evolved polylysogens contain large duplications.

Three isolates containing phage-mediated large deletions in CRISPR also had large duplications elsewhere in the genome (27, 188, and 244 kb; mean = 153 ± 113 kb). In these cases, these regions were not deleted but were doubled in coverage (Fig 3A-C). The location of these large duplications exactly corresponded to additional insertion sites which were recovered by our pipeline (Methods). Due to the orientation of the PA14/DMS3 split reads at these insertion sites, which faced away from each other rather than toward each other, and due to the fact that the split reads at the boundaries of the duplicated regions only represented about 50% of the total coverage, we interpret these regions to be large duplications with a phage genome in the middle, as opposed to two independent viral insertions (Fig 3D-F). As the duplicated regions were not centered around a shared core, they were almost completely non-overlapping in their gene content. Only one known gene was duplicated in two of the three isolates (Lys2_ev_pop2_3 and Lys2_ev_pop2_6), which was *betT,* a choline transporter known to accumulate mutations in clinical isolates from CF patients (Malek et al., 2011; Marvig et al., 2015; Stribling et al., 2023). Broadly, genes from category H (coenzyme metabolism) were represented in all three isolates, as well as from P (inorganic ion metabolism) and S (genes of unknown function), as in the large deletions. Several genes annotated as part of the major facilitator superfamily, a class of membrane-associated transporter proteins, were also duplicated in two of the three isolates, as well as genes from the *moa* family, which have recently been implicated in biofilm formation (Kaleta & Sauer, 2023).

Two of these duplications (in 1_4 and 2_6) independently evolved a shared boundary six nucleotides apart (at positions 1120309 and 1120303, respectively), in an intergenic region between the 3' end of a hypothetical

protein and the 3' end of an AraC transcriptional regulator. An analysis of all new lysogen insertion sites (including the deletion and duplication boundaries) using the motif-finding software MEME Suite did not return any motifs, either using MEME (searching for a motif in a 15 bp region centered on the insertion site), or MEME-ChIP (searching for centrally enriched motifs 250 bp around the insertion site) (Bailey & Elkan, 1994; Machanick & Bailey, 2011); this suggests parallel duplications may be advantageous in this environment.

Neither the isolates containing large deletions nor the ones containing large duplications differed in their growth from other evolved strains that did not have large structural variation (Fig S8A). This suggests that, under these conditions, the fitness costs to deletions, duplications or carrying additional copies of the phage in the chromosome are smaller than the fitness gains by removing self-targeting. Although these phage-mediated deletions of CRISPR represent the addition of one to two phage genomes to the lysogen chromosome, the spontaneous induction rate of these isolates remains reduced relative to the ancestral PA14 lysogen strain (Lys2) (Fig 1, Fig S8B). Additionally, although viral output does not change with phage genome copy number after challenge with mitomycin C, cell survival is significantly increased with increased phage genomes (Fig S8C), suggesting a possible mechanism of viral interference leading to cell survival which may also contribute to a decreased spontaneous induction measurement.

Spontaneous induction correlates with mutation.

We observed that spontaneous induction was variable within replicate populations (Fig 1). We asked whether this variation might correlate with differences in the type of CRISPR mutation (SNP, deletions, viral transpositions). To address this, we grouped lysogens into 1 of 5 categories based on the type of mutation that occurred in the genome ("cas deletion"; "cas mutation"; "spacer deletion"; "spacer mutation"; and "large deletion polylysogen") and asked whether including mutation type reduced the variation within these groups. We observed that all groups (with the exception of the cas deletion group, which had only one isolate in its group) had significantly lower spontaneous induction than the ancestral strains (Fig 4); and the large deletion polylysogen group was significantly lower than lysogens that had lowered spontaneous induction via SNPs or indels. Another set of experiments which included a lysogenized ΔCRISPR strain showed that evolved lysogens which resolved genetic conflict via genome rearrangements, SNPs, and indels, reduced CRISPR function to the level of a ΔCRISPR mutant (Fig S9). A model incorporating mutation type was a significantly better fit than the model by experimental replicate (Fig 4, ANOVA, $F_{2,66} = 13.777$, $P < 1e\text{-}6$). In view of these results, we find that heterogeneity in genotype correlates to the heterogeneity in phenotype of spontaneous induction in our evolved lysogens.

Given these small but significant variations in spontaneous induction that are maintained within groups yet replaced the ancestral lysogen genotype, we wanted to understand how long the diversity we observed within our experimental replicates could persist in exponentially growing cultures. To do this, we developed a mathematical model to compare six lysogens with six different rates of spontaneous induction (two values from the ancestral group, two values from the host-mutation group, two values from the large deletion polylysogen

group). In a deterministic model of lysogen growth in exponential phase with varied spontaneous induction rates, we founded populations with low densities of high inducers (representing the ancestor strain) and allowed them to grow for 10 hours. At that time, we introduced either strains with low rates of spontaneous induction (representing the large deletions) or medium rates of spontaneous induction (representing host mutations), one at 10 hours and the other at 24 hours, and tested whether they could invade. We found that when low inducers are introduced to a high-inducing population, medium inducers cannot subsequently invade (Fig 5A). When we introduced medium inducers to a high-inducer population after 10 hours of growth, and then low inducers after 24 hours, medium and low inducers outcompeted the high inducers and then coexisted in the absence of the high inducers for about two days (Fig 5B). This coexistence recapitulates the observed recovery of low and medium inducers, but not high inducers, in our experimental data (Fig 4). From these observations we find that the weak selection imposed by these small differences in spontaneous induction, which are caused by different mutational mechanisms, combined with the order in which they were introduced, may allow evolution of diversity in CRISPR self-targeting resolution in *P. aeruginosa* lysogens and preserve coexistence of CRISPR+ and CRISPR- strains in the absence of additional selective variables or environmental change.

Discussion

Although temperate phages are frequently recovered from long-term, evolving *Pseudomonas* infections of the lungs of CF patients, phages are often considered at single time points or outside of relationship to their bacterial hosts (Ambroa et al., 2020; Gabrielaite et al., 2020; Tariq et al., 2015, 2019). Evolution experiments involving transposable phages often include susceptible cells, or are begun with free phages instead of established lysogens (Davies et al., 2016; Marshall et al., 2021; O'Brien et al., 2019; Rollie et al., 2020). Because these transposable phages may insert into the bacterial chromosome in many different locations, this approach can unite selection on the phage with the production of beneficial bacterial mutations that are generated upon infection. Here, we bypass this issue to query the impact of a phage on the evolution of its host genome by studying the context of an established lysogen (vertical transmission), and find that the presence of temperate phage alone profoundly changed the genome of the host through genomic rearrangements mediated by polylysogeny from within. This is especially relevant to understanding *P. aeruginosa* infections which may persist for years in the absence of phage lysis and horizontal transmission.

Our work supports previous studies showing that DMS3 lysogens evolve decreased rates of spontaneous induction over time (Rollie et al., 2020; Zegans et al., 2009); here, we distinguish between exponential and stationary phase, and find that exponential phase induction accounts for the majority of free phages in the medium. Our spontaneous induction estimates in exponential phase are approximately 0.37% and 0.72% of the evolved and ancestral populations, respectively, which are both significantly higher than lambda-like phages (Imamovic et al., 2016; Little et al., 1999; Zong et al., 2010), but are in line with other studies that report high spontaneous induction (Owen et al., 2017). Mu-like phages have been reported previously to naturally produce high titers during lysogen growth (Bondy-Denomy et al., 2016; James et al., 2012; Zegans et

al., 2009), suggesting that this evolutionary pressure is not restricted to DMS3 and Mu-like phages as a family have high spontaneous induction rates. Because Mu-like phages are both highly prevalent and highly targeted in *P. aeruginosa* (England et al., 2018), it is likely that CRISPR self-targeting will inform evolutionary outcomes in these related phages. Here, the resolution of that self-targeting resulted in coexisting variation in lysogen spontaneous induction rates in exponential growth, with polylysogens having the lowest rates perhaps due to viral interference preventing cell death, and lowering virion production (Refardt, 2011). These lysogens could also have lower rates not because of polylysogeny, but because of the large deletion itself.

We find that the genetic conflict between Mu-like phage and host results in a tradeoff between CRISPR immunity and spontaneous induction, which could help explain the maintenance of CRISPR systems in *P. aeruginosa* (Cady et al., 2011; Soliman et al., 2022; Van Belkum et al., 2015). Previously, mutations in *cas7* (Zegans et al., 2009), and deletions of the entire CRISPR region (Rollie et al., 2020) have been found to reduce phage induction in late log phase. We find that these differences in self-targeting resolution reduce phage induction to different degrees and have profoundly different effects on the host genome. Half of the evolved lysogen isolates decreased their spontaneous induction while maintaining either CRISPR function (in mutations or deletions of the self-targeting spacer) or the potential of CRISPR function (frameshift mutations in *cas* genes), while the rest deleted CRISPR and lost any potential immunity, but more substantially decreased their induction. One explanation for the maintenance of these two genotypes is the order in which these mutations were introduced. Modeling simulations show that spontaneous induction rates from strains which resolved self-targeting via SNPs and indels (“medium” inducers) cannot invade established populations of strains with spontaneous induction rates from polylysogens. This indicates that host-mediated SNPs and indels likely arose before polylysogeny and large deletions, as both are maintained together despite spontaneous induction differences. This suggestion of an order (host-mediated before phage-mediated) then suggests that the basal rate of phage transposition is lower than host mutation. A lower rate of formation of higher-fitness polylysogen “low” inducers, which compete with lower-fitness host-mutation “medium” inducers with a faster rate of formation, may work to maintain a pool of diversity that selection may subsequently act on in different ways, given the presence of other phages or other ecological factors (Watson et al., 2023).

In an isolated environment, our model predicts that diversity would likely be resolved by the eventual replacement of CRISPR-immune cells with non-immune polylysogens, thereby increasing the phage copy number in an environment that does not have new susceptible hosts for horizontal transmission (Weitz et al., 2019). However, complex communities with additional phage and bacteria may then select for cells which maintain CRISPR function but which have higher spontaneous induction (Alseth et al., 2019; Gloag et al., 2019). A limitation of this study is that it does not take into account the highly polymicrobial nature of the cystic fibrosis lung and uses only one phage-host pair to explore the evolutionary outcomes of lysogeny. These outcomes may be specific and vary with the evolutionary history of each phage-host pair. How phages evolve increased fitness while maintaining CRISPR immunity and defense against other phages and bacterial competitors in complex environments remains an open question.

The results of genetic conflict in evolved lysogens are not limited to CRISPR deletions and may impact the rate of evolution of bacteria with latent infections. Gene loss and genome reduction have also been shown to occur in *P. aeruginosa* lineages during adaptation to the human lung, although the contribution of phages to this loss is unclear (Gabrielaite et al., 2020; Rau et al., 2012). In this study, pyoverdine and Type 6 secretion system genes were lost in the majority of the polylysogens, and one evolved lysogen isolate (1_1) had a nonsynonymous mutation in the Type 6 secretion system baseplate gene *tssK* (Table 1). These genes, lost under laboratory conditions, are also often lost in chronic CF isolates (Marvig et al., 2015; O'Brien et al., 2017; Perault et al., 2020).

The mechanism of this gene loss is unclear in two isolates. In many evolved polylysogens, a phage genome simply replaced the deleted sequence as in (Rollie et al., 2020). In these isolates, the phage genome appeared in an head-tail or tail-head configuration, which could occur as a result of the replicative transposition reaction itself, or as a result of recombination between two preexisting phages. Recombination between a duplicated sequence is an attractive hypothesis to explain the deletions because it may lead to deletion or duplication of the intervening sequence (Reams & Roth, 2015), and because Cas3 targeting of the phage chromosome may stimulate this recombination. However, two of three isolates with a duplication (1_4 and 2_3) exhibit both non-canonical deletion and duplication structures, where we recover host-phage junctions which suggest two phage genomes facing either head-head (reads recovered at both junctions which map to the 3' end of the genome) or tail-tail (reads recovered at both junctions which map to the 5' end of the phage genome). Additionally, recombination may result in two phage-host junctions on the 3' end of the duplication which lead into different ends of the phage chromosome, whereas we only recover reads which lead into one end of the phage chromosome (Reams et al., 2012). Whether and how phage presence continues to alter the evolution of its host from the uninfected state, and how phage infection influences the rate and mechanism of this evolution, are questions which require future study to explain the ubiquity of phage infection in many clinical environments (Burgener et al., 2019; Holloway et al., 1960; Vaca-Pacheco et al., 1999).

P. aeruginosa evolution in the context of the CF lung can occur via a slow accumulation of SNPs and indels (Marvig et al., 2015) or a more rapid accumulation of SNPs due to the evolution of hypermutator genotypes (Marvig et al., 2015; Oliver et al., 2000). Lysogenization by transposable phages may offer a different mechanism of within-lung diversification which operates in addition to the baseline mutation rate. Due to the nature of short-read sequencing, it is likely that polylysogeny of the same virus, and resulting genome rearrangements, have been under-represented in current datasets of *P. aeruginosa* clinical isolates. Future studies should continue to identify signatures of multiple phage infection in clinical isolates, and look for deletions and duplications that may be associated with phages. The inferred dependence of transposition on the presence of self-targeting, and by extension on low levels of DNA damage, also opens the possibility of other causes of low levels of DNA damage – for example, subinhibitory concentrations of DNA-damaging antibiotics – to be driving evolutionary adaptation and diversity in bacteria lysogenized by transposable

phages. The extent to which phage infection informs differences in evolutionary and functional outcomes in a clinical context is an important subject for future work.

Acknowledgements

We thank Santiago Elena at I2SysBio and the University of Valencia for help conceptualizing and establishing initial experimental evolution studies on lysogen fitness. We are grateful to George O'Toole for discussions as well as strains of bacteria and phages described herein including Lys2 and DMS3. We gratefully acknowledge Dr. Alan Collins for helpful discussions during early stages of this project. We thank Whitaker lab members Jiayue Yang for helpful discussions, and Sierra Bedwell and Izzy Lakis for comments on the manuscript. We thank Alvaro Hernandez and Chris Wright of the Roy J. Carver Biotechnology Institute for sequencing expertise, as well as Jeff Haas of the School of Integrative Biology for crucial help with data storage and server access. This work is funded by grants from the Cystic Fibrosis Foundation and the Allen Distinguished Investigator Award from the Paul G. Allen Foundation to R.J.W. and the National Science Foundation grant DMS-1815764 to Z.R. This research is also a contribution of the GEMS Biology Integration Institute, funded by the National Science Foundation DBI Integration Institutes Program, Award #2022049.

Author Contributions

Conceptualization: L.C.S., R.J.W. Formal analysis: L.C.S. Funding acquisition: L.C.S., R.J.W. Investigation: L.C.S., Z.R., A.C.S. Methodology: L.C.S., Z.R., J.H.C., R.J.W. Resources: J.C.V., R.J.W. Software: J.C.V. Visualization: L.C.S., Z.R. Writing – original draft: L.C.S., R.J.W. Writing – review and editing: L.C.S., R.J.W.

Group	Sample ID	Mutated region	Description	Due to phage	Location on PA14-REF chromosome
ev_pop1	ev_pop1_1	CRISPR2 sp1-8	Deletion	No	2936222-2936676
ev_pop1	ev_pop1_1	*tssK*	Type 6 secretion system baseplate gene; Nonsynonymous	No	94370
ev_pop1	ev_pop1_2	CRISPR2 sp1-4	Deletion	No	2936462-2936675
ev_pop1	ev_pop1_3	Δ 243,737 bp	Includes CRISPR region	Yes	2835111-3078848
ev_pop1	ev_pop1_4	Δ 186,866 bp	Includes CRISPR region	Yes	2839227-3026093
ev_pop1	ev_pop1_4	27,769 bp	Duplicated region	Yes	1120309-1148078
ev_pop1	ev_pop1_4	*cysT*	Sulfate transport protein; $(CAG)_{4\rightarrow3}$	No	329528
ev_pop1	ev_pop1_5	CRISPR2 sp1	A<G mutation in the 2nd nt	No	2936674
ev_pop1	ev_pop1_6	CRISPR2 sp1	GAT<G deletion of 1st and 2nd nt	No	2936673
ev_pop2	ev_pop2_1	Δ 88,123 bp	Includes CRISPR region	Yes	2921860
ev_pop2	ev_pop2_2	*cas8*	Frameshift (predicted loss of C-term helical bundle region)	No	2929846
ev_pop2	ev_pop2_3	Δ 194,512 bp	Includes CRISPR region	Yes	2890191-3084703
ev_pop2	ev_pop2_3	187,901 bp	Duplicated region	Yes	6222744-6410645
ev_pop2	ev_pop2_4	CRISPR2 sp1	Deletion	No	2936643-2936675
ev_pop2	ev_pop2_5	CRISPR2 sp1	A<G mutation in the 2nd nt	No	2936674
ev_pop2	ev_pop2_6	Δ 335,331 bp	Includes CRISPR region	Yes	2811042-3146373
ev_pop2	ev_pop2_6	244,019 bp	Duplicated region	Yes	876284-1120303
ev_pop2	ev_pop2_6	Intergenic region	$(GCCAAC)_{11\rightarrow8}$	No	3792568
ev_pop3	ev_pop3_1	*cas8*	Frameshift (retention of first 57/435 aa)	No	2930725
ev_pop3	ev_pop3_2	CRISPR2 sp1	Deletion	No	2936644-2936675
ev_pop3	ev_pop3_3	*cas7*	Frameshift (predicted loss of RNA-binding domain)	No	2927776
ev_pop3	ev_pop3_4	Δ 60,538 bp	Begins in *cas3* gene and includes CRISPR2 array	Yes	2934007-2994541
ev_pop3	ev_pop3_5	Δ 34,249 bp	Includes CRISPR region	Yes	2903549-2937794
ev_pop3	ev_pop3_6	Cas gene deletion	Deletion of *cas7* and *cas5*, partial deletion of *cas6* and *cas8*	No	2927288-2930093

Table 1. Description of mutations recovered in evolved lysogens.

Sample ID	Group	Upstream deletion boundary Read mate maps to:	Downstream deletion boundary Read mate maps to:
Lys2_ev_pop1_3	ev_pop_1	end of DMS3 genome	start of DMS3 genome
***Lys2_ev_pop1_4**	ev_pop_1	start of DMS3 genome	start of DMS3 genome
Lys2_ev_pop2_1	ev_pop_2	end of DMS3 genome	start of DMS3 genome
***Lys2_ev_pop2_3**	ev_pop_2	start of DMS3 genome	start of DMS3 genome
Lys2_ev_pop2_6	ev_pop_2	end of DMS3 genome	start of DMS3 genome
Lys2_ev_pop3_4	ev_pop_3	end of DMS3 genome	start of DMS3 genome
Lys2_ev_pop3_5	ev_pop_3	end of DMS3 genome	start of DMS3 genome

Table 2. Origin of phage reads at the deletion boundaries. Asterisks denote samples where reads from only one end of the phage genome were recovered.

References

Al-Anany, A. M., Fatima, R., & Hynes, A. P. (2021). Temperate phage-antibiotic synergy eradicates bacteria through depletion of lysogens. *Cell Reports*, *35*(8), 109172. https://doi.org/10.1016/j.celrep.2021.109172

Allet, B. (1979). Mu insertion duplicates a 5 base pair sequence at the host inserted site. *Cell*, *16*(1), 123–129. https://doi.org/10.1016/0092-8674(79)90193-4

Alseth, E. O., Pursey, E., Luján, A. M., McLeod, I., Rollie, C., & Westra, E. R. (2019). Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature*, *574*(7779), 549–552. https://doi.org/10.1038/s41586-019-1662-9

Ambroa, A., Blasco, L., López-Causapé, C., Trastoy, R., Fernandez-García, L., Bleriot, I., Ponce-Alonso, M., Pacios, O., López, M., Cantón, R., Kidd, T. J., Bou, G., Oliver, A., & Tomás, M. (2020). Temperate Bacteriophages (Prophages) in Pseudomonas aeruginosa Isolates Belonging to the International Cystic Fibrosis Clone (CC274). *Frontiers in Microbiology*, *11*, 556706. https://doi.org/10.3389/fmicb.2020.556706

Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, *2*, 28–36.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Bondy-Denomy, J., Qian, J., Westra, E. R., Buckling, A., Guttman, D. S., Davidson, A. R., & Maxwell, K. L. (2016). Prophages mediate defense against phage infection through diverse mechanisms. *The ISME Journal*, *10*(12), 2854–2866. https://doi.org/10.1038/ismej.2016.79

Brüssow, H., Canchaya, C., & Hardt, W.-D. (2004). Phages and the Evolution of Bacterial Pathogens: From Genomic Rearrangements to Lysogenic Conversion. *Microbiology and Molecular Biology Reviews*, *68*(3), 560–602. https://doi.org/10.1128/MMBR.68.3.560-602.2004

Budzik, J. M., Rosche, W. A., Rietsch, A., & O'Toole, G. A. (2004). Isolation and Characterization of a Generalized Transducing Phage for Pseudomonas aeruginosa Strains PAO1 and PA14. *Journal of Bacteriology*, *186*(10), 3270–3273. https://doi.org/10.1128/JB.186.10.3270-3273.2004

Burgener, E. B., Sweere, J. M., Bach, M. S., Secor, P. R., Haddock, N., Jennings, L. K., Marvig, R. L., Johansen, H. K., Rossi, E., Cao, X., Tian, L., Nedelec, L., Molin, S., Bollyky, P. L., & Milla, C. E. (2019). Filamentous bacteriophages are associated with chronic Pseudomonas lung infections and antibiotic resistance in cystic fibrosis. *Science Translational Medicine*, *11*(488), eaau9748. https://doi.org/10.1126/scitranslmed.aau9748

Cady, K. C., & O'Toole, G. A. (2011). Non-Identity-Mediated CRISPR-Bacteriophage Interaction Mediated via the Csy and Cas3 Proteins. *Journal of Bacteriology*, *193*(14), 3433–3445. https://doi.org/10.1128/JB.01411-10

Cady, K. C., White, A. S., Hammond, J. H., Abendroth, M. D., Karthikeyan, R. S. G., Lalitha, P., Zegans, M. E., & O'Toole, G. A. (2011). Prevalence, conservation and functional analysis of Yersinia and Escherichia CRISPR regions in clinical Pseudomonas aeruginosa isolates. *Microbiology*, *157*(2), 430–437. https://doi.org/10.1099/mic.0.045732-0

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). *eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale* [Preprint]. Bioinformatics. https://doi.org/10.1101/2021.06.03.446934

Chaconas, G., & Harshey, R. M. (2007). Transposition of Phage Mu DNA. In N. L. Craig, R. Craigie, M. Gellert, & A. M. Lambowitz (Eds.), *Mobile DNA II* (1st ed., pp. 384–402). Wiley. https://doi.org/10.1128/9781555817954.ch17

Chowdhury, S., Carter, J., Rollins, M. F., Golden, S. M., Jackson, R. N., Hoffmann, C., Nosaka, L., Bondy-Denomy, J., Maxwell, K. L., Davidson, A. R., Fischer, E. R., Lander, G. C., & Wiedenheft, B. (2017). Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. *Cell*, *169*(1), 47-57.e11. https://doi.org/10.1016/j.cell.2017.03.012

Clifton, S. M., Kim, T., Chandrashekhar, J. H., O'Toole, G. A., Rapti, Z., & Whitaker, R. J. (2019). Lying in Wait: Modeling the Control of Bacterial Infections via Antibiotic-Induced Proviruses. *mSystems*, *4*(5). https://doi.org/10.1128/mSystems.00221-19

Clifton, S. M., Whitaker, R. J., & Rapti, Z. (2021). Temperate and chronic virus competition leads to low lysogen frequency. *Journal of Theoretical Biology*, *523*, 110710. https://doi.org/10.1016/j.jtbi.2021.110710

Collins, A. J., & Whitaker, R. J. (2022). *CRISPR comparison toolkit (CCTK): Rapid identification, visualization, and analysis of CRISPR array diversity* [Preprint]. Bioinformatics. https://doi.org/10.1101/2022.07.31.502198

Cramer, N., Klockgether, J., Wrasman, K., Schmidt, M., Davenport, C. F., & Tümmler, B. (2011). Microevolution of the major common Pseudomonas aeruginosa clones C and PA14 in cystic fibrosis lungs: P. aeruginosa microevolution in cystic fibrosis. *Environmental Microbiology*, *13*(7), 1690–1704. https://doi.org/10.1111/j.1462-2920.2011.02483.x

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Darch, S. E., McNally, A., Harrison, F., Corander, J., Barr, H. L., Paszkiewicz, K., Holden, S., Fogarty, A., Crusz, S. A., & Diggle, S. P. (2015). Recombination is a key driver of genomic and phenotypic diversity in a Pseudomonas aeruginosa population during cystic fibrosis infection. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep07649

Davies, E. V., James, C. E., Williams, D., O'Brien, S., Fothergill, J. L., Haldenby, S., Paterson, S., Winstanley, C., & Brockhurst, M. A. (2016). Temperate phages both mediate and drive adaptive evolution in pathogen

biofilms. *Proceedings of the National Academy of Sciences*, *113*(29), 8266–8271. https://doi.org/10.1073/pnas.1520056113

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology (Clifton, N.J.)*, *1151*, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12

Ellis, E. L., & Delbrück, M. (1939). The Growth of Bacteriophage. *The Journal of General Physiology*, *22*(3), 365–384. https://doi.org/10.1085/jgp.22.3.365

England, W. E., Kim, T., & Whitaker, R. J. (2018). Metapopulation Structure of CRISPR-Cas Immunity in Pseudomonas aeruginosa and Its Viruses. *mSystems*, *3*(5), e00075-18. https://doi.org/10.1128/mSystems.00075-18

Feliziani, S., Marvig, R. L., Luján, A. M., Moyano, A. J., Rienzo, J. A. D., Johansen, H. K., Molin, S., & Smania, A. M. (2014). Coexistence and Within-Host Evolution of Diversified Lineages of Hypermutable Pseudomonas aeruginosa in Long-term Cystic Fibrosis Infections. *PLOS Genetics*, *10*(10), e1004651. https://doi.org/10.1371/journal.pgen.1004651

Folkesson, A., Jelsbak, L., Yang, L., Johansen, H. K., Ciofu, O., Høiby, N., & Molin, S. (2012). Adaptation of Pseudomonas aeruginosa to the cystic fibrosis airway: An evolutionary perspective. *Nature Reviews Microbiology*, *10*(12), 841–851. https://doi.org/10.1038/nrmicro2907

Fothergill, J. L., Mowat, E., Walshaw, M. J., Ledson, M. J., James, C. E., & Winstanley, C. (2011). Effect of Antibiotic Treatment on Bacteriophage Production by a Cystic Fibrosis Epidemic Strain of Pseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy*, *55*(1), 426–428. https://doi.org/10.1128/AAC.01257-10

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression*. Sage Publications. https://socialsciences.mcmaster.ca/jfox/Books/Companion

Gabrielaite, M., Johansen, H. K., Molin, S., Nielsen, F. C., & Marvig, R. L. (2020). Gene Loss and Acquisition in Lineages of Pseudomonas aeruginosa Evolving in Cystic Fibrosis Patient Airways. *mBio*, *11*(5), e02359-20. https://doi.org/10.1128/mBio.02359-20

Gloag, E. S., Marshall, C. W., Snyder, D., Lewin, G. R., Harris, J. S., Santos-Lopez, A., Chaney, S. B., Whiteley, M., Cooper, V. S., & Wozniak, D. J. (2019). Pseudomonas aeruginosa Interstrain Dynamics and Selection of Hyperbiofilm Mutants during a Chronic Infection. *mBio*, *10*(4), e01698-19. https://doi.org/10.1128/mBio.01698-19

Goerke, C., Matias y Papenberg, S., Dasbach, S., Dietz, K., Ziebach, R., Kahl, B. C., & Wolz, C. (2004). Increased Frequency of Genomic Alterations in *Staphylococcus aureus* during Chronic Infection Is in Part Due to Phage Mobilization. *The Journal of Infectious Diseases*, *189*(4), 724–734. https://doi.org/10.1086/381502

Goerke, C., Wirtz, C., Fluckiger, U., & Wolz, C. (2006). Extensive phage dynamics in Staphylococcus aureus contributes to adaptation to the human host during infection. *Molecular Microbiology*, *61*(6), 1673–1685. https://doi.org/10.1111/j.1365-2958.2006.05354.x

Golubchik, T., Batty, E. M., Miller, R. R., Farr, H., Young, B. C., Larner-Svensson, H., Fung, R., Godwin, H., Knox, K., Votintseva, A., Everitt, R. G., Street, T., Cule, M., Ip, C. L. C., Didelot, X., Peto, T. E. A., Harding, R. M., Wilson, D. J., Crook, D. W., & Bowden, R. (2013). Within-Host Evolution of Staphylococcus aureus during Asymptomatic Carriage. *PLoS ONE*, *8*(5), e61319. https://doi.org/10.1371/journal.pone.0061319

Guérillot, R., Kostoulias, X., Donovan, L., Li, L., Carter, G. P., Hachani, A., Vandelannoote, K., Giulieri, S., Monk, I. R., Kunimoto, M., Starrs, L., Burgio, G., Seemann, T., Peleg, A. Y., Stinear, T. P., & Howden, B. P. (2019). Unstable chromosome rearrangements in *Staphylococcus aureus* cause phenotype switching associated with persistent infections. *Proceedings of the National Academy of Sciences*, *116*(40), 20135–20140. https://doi.org/10.1073/pnas.1904861116

Heussler, G. E., Cady, K. C., Koeppen, K., Bhuju, S., Stanton, B. A., & O'Toole, G. A. (2015). Clustered Regularly Interspaced Short Palindromic Repeat-Dependent, Biofilm-Specific Death of Pseudomonas aeruginosa

Mediated by Increased Expression of Phage-Related Genes. *mBio*, *6*(3), e00129-15. https://doi.org/10.1128/mBio.00129-15

Holloway, B. W., Egan, J. B., & Monk, M. (1960). Lysogeny in Pseudomonas Aeruginosa. *Australian Journal of Experimental Biology and Medical Science*, *38*(4), 321–330. https://doi.org/10.1038/icb.1960.34

Imamovic, L., Ballesté, E., Martínez-Castillo, A., García-Aljaro, C., & Muniesa, M. (2016). Heterogeneity in phage induction enables the survival of the lysogenic population. *Environmental Microbiology*, *18*(3), 957–969. https://doi.org/10.1111/1462-2920.13151

Imperi, F., Tiburzi, F., & Visca, P. (2009). Molecular basis of pyoverdine siderophore recycling in Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences*, *106*(48), 20440–20445. https://doi.org/10.1073/pnas.0908760106

James, C. E., Davies, E. V., Fothergill, J. L., Walshaw, M. J., Beale, C. M., Brockhurst, M. A., & Winstanley, C. (2015). Lytic activity by temperate phages of Pseudomonas aeruginosa in long-term cystic fibrosis chronic lung infections. *The ISME Journal*, *9*(6), Article 6. https://doi.org/10.1038/ismej.2014.223

James, C. E., Fothergill, J. L., Kalwij, H., Hall, A. J., Cottell, J., Brockhurst, M. A., & Winstanley, C. (2012). Differential infection properties of three inducible prophages from an epidemic strain of Pseudomonas aeruginosa. *BMC Microbiology*, *12*(1), 216. https://doi.org/10.1186/1471-2180-12-216

Jeukens, J., Boyle, B., Kukavica-Ibrulj, I., Ouellet, M. M., Aaron, S. D., Charette, S. J., Fothergill, J. L., Tucker, N. P., Winstanley, C., & Levesque, R. C. (2014). Comparative Genomics of Isolates of a Pseudomonas aeruginosa Epidemic Strain Associated with Chronic Lung Infections of Cystic Fibrosis Patients. *PLoS ONE*, *9*(2), e87611. https://doi.org/10.1371/journal.pone.0087611

Jorth, P., Staudinger, B. J., Wu, X., Hisert, K. B., Hayden, H., Garudathri, J., Harding, C. L., Radey, M. C., Rezayat, A., Bautista, G., Berrington, W. R., Goddard, A. F., Zheng, C., Angermeyer, A., Brittnacher, M. J., Kitzman, J., Shendure, J., Fligner, C. L., Mittler, J., … Singh, P. K. (2015). Regional Isolation Drives Bacterial Diversification within Cystic Fibrosis Lungs. *Cell Host & Microbe*, *18*(3), 307–319. https://doi.org/10.1016/j.chom.2015.07.006

Kaleta, M. F., & Sauer, K. (2023). MoaB1 Homologs Contribute to Biofilm Formation and Motility by Pseudomonas aeruginosa and Escherichia coli. *Journal of Bacteriology*, *205*(5), e00004-23. https://doi.org/10.1128/jb.00004-23

Kettle, A. J., Turner, R., Gangell, C. L., Harwood, D. T., Khalilova, I. S., Chapman, A. L., Winterbourn, C. C., Sly, P. D., & on behalf of AREST CF. (2014). Oxidation contributes to low glutathione in the airways of children with cystic fibrosis. *European Respiratory Journal*, *44*(1), 122–129. https://doi.org/10.1183/09031936.00170213

Kortright, K. E., Chan, B. K., Koff, J. L., & Turner, P. E. (2019). Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria. *Cell Host & Microbe*, *25*(2), 219–232. https://doi.org/10.1016/j.chom.2019.01.014

Landa, K. J., Mossman, L. M., Whitaker, R. J., Rapti, Z., & Clifton, S. M. (2021). *Phage-antibiotic synergy inhibited by temperate and chronic virus competition*. https://arxiv.org/abs/2104.08989v1

Leoni, L., Orsi, N., De Lorenzo, V., & Visca, P. (2000). Functional Analysis of PvdS, an Iron Starvation Sigma Factor of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, *182*(6), 1481–1491. https://doi.org/10.1128/JB.182.6.1481-1491.2000

Little, J. W., Shepley, D. P., & Wert, D. W. (1999). Robustness of a gene regulatory circuit. *The EMBO Journal*, *18*(15), 4299–4307. https://doi.org/10.1093/emboj/18.15.4299

Machanick, P., & Bailey, T. L. (2011). MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics*, *27*(12), 1696–1697. https://doi.org/10.1093/bioinformatics/btr189

Makarova, K. S., Wolf, Y. I., Snir, S., & Koonin, E. V. (2011). Defense Islands in Bacterial and Archaeal Genomes and Prediction of Novel Defense Systems. *Journal of Bacteriology*, *193*(21), 6039–6056. https://doi.org/10.1128/JB.05535-11

Malek, A. A., Chen, C., Wargo, M. J., Beattie, G. A., & Hogan, D. A. (2011). Roles of Three Transporters, CbcXWV, BetT1, and BetT3, in Pseudomonas aeruginosa Choline Uptake for Catabolism. *Journal of Bacteriology*, *193*(12), 3033–3041. https://doi.org/10.1128/JB.00160-11

690 Malhotra, S., Hayes, D., & Wozniak, D. J. (2019). Cystic Fibrosis and Pseudomonas aeruginosa: The Host-
691 Microbe Interface. *Clinical Microbiology Reviews*, *32*(3), e00138-18.
692 https://doi.org/10.1128/CMR.00138-18
693 Markussen, T., Marvig, R. L., Gómez-Lozano, M., Aanæs, K., Burleigh, A. E., Høiby, N., Johansen, H. K., Molin, S.,
694 & Jelsbak, L. (2014). Environmental Heterogeneity Drives Within-Host Diversification and Evolution of
695 Pseudomonas aeruginosa. *mBio*, *5*(5). https://doi.org/10.1128/mBio.01592-14
696 Marshall, C. W., Gloag, E. S., Lim, C., Wozniak, D. J., & Cooper, V. S. (2021). Rampant prophage movement
697 among transient competitors drives rapid adaptation during infection. *Science Advances*, *7*(29),
698 eabh1489. https://doi.org/10.1126/sciadv.abh1489
699 Marvig, R. L., Sommer, L. M., Molin, S., & Johansen, H. K. (2015). Convergent evolution and adaptation of
700 Pseudomonas aeruginosa within patients with cystic fibrosis. *Nature Genetics*, *47*(1), 57–64.
701 https://doi.org/10.1038/ng.3148
702 McGrath, L. T., Mallon, P., Dowey, L., Silke, B., McClean, E., McDonnell, M., Devine, A., Copeland, S., & Elborn,
703 S. (1999). Oxidative stress during acute respiratory exacerbations in cystic fibrosis. *Thorax*, *54*(6), 518–
704 523. https://doi.org/10.1136/thx.54.6.518
705 Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D., & Marshall, D. (2013). Using
706 Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics*, *14*(2),
707 193–202. https://doi.org/10.1093/bib/bbs012
708 Morgan, G. J., Hatfull, G. F., Casjens, S., & Hendrix, R. W. (2002). Bacteriophage Mu genome sequence:
709 Analysis and comparison with Mu-like prophages in Haemophilus, Neisseria and Deinococcus. *Journal*
710 *of Molecular Biology*, *317*(3), 337–359. https://doi.org/10.1006/jmbi.2002.5437
711 Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K.,
712 Shiba, T., Yasunaga, T., Hayashi, H., Hattori, M., & Hamada, S. (2003). Genome Sequence of an M3
713 Strain of *Streptococcus pyogenes* Reveals a Large-Scale Genomic Rearrangement in Invasive Strains and

New Insights into Phage Evolution. *Genome Research*, *13*(6a), 1042–1055. https://doi.org/10.1101/gr.1096703

Nanda, A. M., Thormann, K., & Frunzke, J. (2015). Impact of Spontaneous Prophage Induction on the Fitness of Bacterial Populations and Host-Microbe Interactions. *Journal of Bacteriology*, *197*(3), 410–419. https://doi.org/10.1128/JB.02230-14

Nguyen, A. T., O'Neill, M. J., Watts, A. M., Robson, C. L., Lamont, I. L., Wilks, A., & Oglesby-Sherrouse, A. G. (2014). Adaptation of Iron Homeostasis Pathways by a Pseudomonas aeruginosa Pyoverdine Mutant in the Cystic Fibrosis Lung. *Journal of Bacteriology*, *196*(12), 2265–2276. https://doi.org/10.1128/JB.01491-14

O'Brien, S., Kümmerli, R., Paterson, S., Winstanley, C., & Brockhurst, M. A. (2019). Transposable temperate phages promote the evolution of divergent social strategies in Pseudomonas aeruginosa populations. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1912), 20191794. https://doi.org/10.1098/rspb.2019.1794

O'Brien, S., Williams, D., Fothergill, J. L., Paterson, S., Winstanley, C., & Brockhurst, M. A. (2017). High virulence sub-populations in Pseudomonas aeruginosa long-term cystic fibrosis airway infections. *BMC Microbiology*, *17*(1), 30. https://doi.org/10.1186/s12866-017-0941-6

Ojeniyi, B., Birch-Andersen, A., Mansa, B., Rosdahl, V. T., & HØIby, N. (1991). Morphology Of *Pseudomonas aeruginosa* phages from the sputum of cystic fibrosis patients and from the phage typing set. *APMIS*, *99*(7–12), 925–930. https://doi.org/10.1111/j.1699-0463.1991.tb01280.x

Oliver, A., Cantón, R., Campo, P., Baquero, F., & Blázquez, J. (2000). High Frequency of Hypermutable Pseudomonas aeruginosa in Cystic Fibrosis Lung Infection. *Science*, *288*(5469), 1251–1253. https://doi.org/10.1126/science.288.5469.1251

Owen, S. V., Wenner, N., Canals, R., Makumi, A., Hammarlöf, D. L., Gordon, M. A., Aertsen, A., Feasey, N. A., & Hinton, J. C. D. (2017). Characterization of the Prophage Repertoire of African Salmonella Typhimurium

738 ST313 Reveals High Levels of Spontaneous Induction of Novel Phage BTP1. *Frontiers in Microbiology*, *8*.
739 https://doi.org/10.3389/fmicb.2017.00235
740 Perault, A. I., Chandler, C. E., Rasko, D. A., Ernst, R. K., Wolfgang, M. C., & Cotter, P. A. (2020). Host Adaptation
741 Predisposes Pseudomonas aeruginosa to Type VI Secretion System-Mediated Predation by the
742 Burkholderia cepacia Complex. *Cell Host & Microbe*, *28*(4), 534-547.e3.
743 https://doi.org/10.1016/j.chom.2020.06.019
744 R Core Team. (2023). *R: A Language and Environment for Statistical Computing* (4.3.1) [Computer software]. R
745 Foundation for Statistical Computing. https://www.R-project.org
746 Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G., & Ausubel, F. M. (1995). Common
747 virulence factors for bacterial pathogenicity in plants and animals. *Science*, *268*(5219), 1899–1902.
748 https://doi.org/10.1126/science.7604262
749 Ranquet, C., Toussaint, A., de Jong, H., Maenhaut-Michel, G., & Geiselmann, J. (2005). Control of
750 Bacteriophage Mu Lysogenic Repression. *Journal of Molecular Biology*, *353*(1), 186–195.
751 https://doi.org/10.1016/j.jmb.2005.08.015
752 Rau, M. H., Marvig, R. L., Ehrlich, G. D., Molin, S., & Jelsbak, L. (2012). Deletion and acquisition of genomic
753 content during early stage adaptation of Pseudomonas aeruginosa to a human host environment:
754 Bacterial genome reduction in human host. *Environmental Microbiology*, *14*(8), 2200–2211.
755 https://doi.org/10.1111/j.1462-2920.2012.02795.x
756 Reams, A. B., Kofoid, E., Kugelberg, E., & Roth, J. R. (2012). Multiple Pathways of Duplication Formation with
757 and Without Recombination (RecA) in *Salmonella enterica*. *Genetics*, *192*(2), 397–415.
758 https://doi.org/10.1534/genetics.112.142570
759 Reams, A. B., & Roth, J. R. (2015). Mechanisms of Gene Duplication and Amplification. *Cold Spring Harbor*
760 *Perspectives in Biology*, *7*(2), a016592. https://doi.org/10.1101/cshperspect.a016592
761 Refardt, D. (2011). Within-host competition determines reproductive success of temperate bacteriophages.
762 *The ISME Journal*, *5*(9), 1451–1460. https://doi.org/10.1038/ismej.2011.30

763 Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A., & Mesirov, J. P. (2017). Variant Review with the
764 Integrative Genomics Viewer. *Cancer Research*, *77*(21), e31–e34. https://doi.org/10.1158/0008-
765 5472.CAN-17-0337
766 Robinson, L. A., Collins, A. C. Z., Murphy, R. A., Davies, J. C., & Allsopp, L. P. (2023). Diversity and prevalence of
767 type VI secretion system effectors in clinical Pseudomonas aeruginosa isolates. *Frontiers in*
768 *Microbiology*, *13*, 1042505. https://doi.org/10.3389/fmicb.2022.1042505
769 Rolain, J.-M., Francois, P., Hernandez, D., Bittar, F., Richet, H., Fournous, G., Mattenberger, Y., Bosdure, E.,
770 Stremler, N., Dubus, J.-C., Sarles, J., Reynaud-Gaubert, M., Boniface, S., Schrenzel, J., & Raoult, D.
771 (2009). Genomic analysis of an emerging multiresistant Staphylococcus aureus strain rapidly spreading
772 in cystic fibrosis patients revealed the presence of an antibiotic inducible bacteriophage. *Biology Direct*,
773 *4*(1), 1. https://doi.org/10.1186/1745-6150-4-1
774 Rollie, C., Chevallereau, A., Watson, B. N. J., Chyou, T., Fradet, O., McLeod, I., Fineran, P. C., Brown, C. M.,
775 Gandon, S., & Westra, E. R. (2020). Targeting of temperate phages drives loss of type I CRISPR–Cas
776 systems. *Nature*, *578*(7793), Article 7793. https://doi.org/10.1038/s41586-020-1936-2
777 Rossi, E., La Rosa, R., Bartell, J. A., Marvig, R. L., Haagensen, J. A. J., Sommer, L. M., Molin, S., & Johansen, H. K.
778 (2021). Pseudomonas aeruginosa adaptation and evolution in patients with cystic fibrosis. *Nature*
779 *Reviews Microbiology*, *19*(5), 331–342. https://doi.org/10.1038/s41579-020-00477-5
780 Schick, A., & Kassen, R. (2018). Rapid diversification of Pseudomonas aeruginosa in cystic fibrosis lung-like
781 conditions. *Proceedings of the National Academy of Sciences*, *115*(42), 10714–10719.
782 https://doi.org/10.1073/pnas.1721270115
783 Schick, A., Shewaramani, S., & Kassen, R. (2022). Genomics of Diversification of *Pseudomonas aeruginosa* in
784 Cystic Fibrosis Lung-like Conditions. *Genome Biology and Evolution*, *14*(6), evac074.
785 https://doi.org/10.1093/gbe/evac074
786 Shah, M., Taylor, V. L., Bona, D., Tsao, Y., Stanley, S. Y., Pimentel-Elardo, S. M., McCallum, M., Bondy-Denomy,
787 J., Howell, P. L., Nodwell, J. R., Davidson, A. R., Moraes, T. F., & Maxwell, K. L. (2021). A phage-encoded

anti-activator inhibits quorum sensing in Pseudomonas aeruginosa. *Molecular Cell*, *81*(3), 571-583.e6. https://doi.org/10.1016/j.molcel.2020.12.011

Soliman, M., Said, H. S., El-Mowafy, M., & Barwa, R. (2022). Novel PCR detection of CRISPR/Cas systems in Pseudomonas aeruginosa and its correlation with antibiotic resistance. *Applied Microbiology and Biotechnology*, *106*(21), 7223–7234. https://doi.org/10.1007/s00253-022-12144-1

Stribling, W., Hall, L. R., Powell, A., Harless, C., Martin, M. J., Corey, B. W., Snesrud, E., Ong, A., Maybank, R., Stam, J., Bartlett, K., Jones, B. T., Preston, L. N., Lane, K. F., Thompson, B., Young, L. M., Kwak, Y. I., Barsoumian, A. E., Markelz, A.-E., … Lebreton, F. (2023). *Detecting, mapping, and suppressing the spread of a decade-long* Pseudomonas aeruginosa *nosocomial outbreak with genomics* [Preprint]. Microbiology. https://doi.org/10.1101/2023.07.24.550326

Tai, A. S., Sherrard, L. J., Kidd, T. J., Ramsay, K. A., Buckley, C., Syrmis, M., Grimwood, K., Bell, S. C., & Whiley, D. M. (2017). Antibiotic perturbation of mixed-strain Pseudomonas aeruginosa infection in patients with cystic fibrosis. *BMC Pulmonary Medicine*, *17*. https://doi.org/10.1186/s12890-017-0482-7

Tariq, M. A., Everest, F. L. C., Cowley, L. A., De Soyza, A., Holt, G. S., Bridge, S. H., Perry, A., Perry, J. D., Bourke, S. J., Cummings, S. P., Lanyon, C. V., Barr, J. J., & Smith, D. L. (2015). A metagenomic approach to characterize temperate bacteriophage populations from Cystic Fibrosis and non-Cystic Fibrosis bronchiectasis patients. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00097

Tariq, M. A., Everest, F. L. C., Cowley, L. A., Wright, R., Holt, G. S., Ingram, H., Duignan, L. A. M., Nelson, A., Lanyon, C. V., Perry, A., Perry, J. D., Bourke, S., Brockhurst, M. A., Bridge, S. H., Soyza, A. D., & Smith, D. L. (2019). Temperate Bacteriophages from Chronic Pseudomonas aeruginosa Lung Infections Show Disease-Specific Changes in Host Range and Modulate Antimicrobial Susceptibility. *mSystems*, *4*(4). https://doi.org/10.1128/mSystems.00191-18

Toussaint, A., & Rice, P. A. (2017). Transposable phages, DNA reorganization and transfer. *Current Opinion in Microbiology*, *38*, 88–94. https://doi.org/10.1016/j.mib.2017.04.009

Vaca-Pacheco, S., Paniagua-Contreras, G. L., García-González, O., & de la Garza, M. (1999). The Clinically Isolated FIZ15 Bacteriophage Causes Lysogenic Conversion in Pseudomonas aeruginosa PAO1. *Current Microbiology*, *38*(4), 239–243. https://doi.org/10.1007/PL00006794

Van Belkum, A., Soriaga, L. B., LaFave, M. C., Akella, S., Veyrieras, J.-B., Barbu, E. M., Shortridge, D., Blanc, B., Hannum, G., Zambardi, G., Miller, K., Enright, M. C., Mugnier, N., Brami, D., Schicklin, S., Felderman, M., Schwartz, A. S., Richardson, T. H., Peterson, T. C., … Cady, K. C. (2015). Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant Pseudomonas aeruginosa. *mBio*, *6*(6), e01796-15. https://doi.org/10.1128/mBio.01796-15

Vanderwoude, J., Azimi, S., Read, T., & Diggle, S. P. (2023). *Unraveling Genomic Diversity in Pseudomonas aeruginosa Cystic Fibrosis Lung Infection and Its Impact on Antimicrobial Resistance* [Preprint]. Microbiology. https://doi.org/10.1101/2023.06.14.544983

Vergnaud, G., Midoux, C., Blouin, Y., Bourkaltseva, M., Krylov, V., & Pourcel, C. (2018). Transposition Behavior Revealed by High-Resolution Description of Pseudomonas Aeruginosa Saltovirus Integration Sites. *Viruses*, *10*(5), Article 5. https://doi.org/10.3390/v10050245

Vorontsova, D., Datsenko, K. A., Medvedeva, S., Bondy-Denomy, J., Savitskaya, E. E., Pougach, K., Logacheva, M., Wiedenheft, B., Davidson, A. R., Severinov, K., & Semenova, E. (2015). Foreign DNA acquisition by the I-F CRISPR–Cas system requires all components of the interference machinery. *Nucleic Acids Research*, *43*(22), 10848–10860. https://doi.org/10.1093/nar/gkv1261

Walker, D. M., Freddolino, P. L., & Harshey, R. M. (2020). A Well-Mixed E. coli Genome: Widespread Contacts Revealed by Tracking Mu Transposition. *Cell*, *180*(4), 703-716.e18. https://doi.org/10.1016/j.cell.2020.01.031

Walker, D. M., & Harshey, R. M. (2020). Deep sequencing reveals new roles for MuB in transposition immunity and target-capture, and redefines the insular Ter region of E. coli. *Mobile DNA*, *11*(1), 26. https://doi.org/10.1186/s13100-020-00217-9

Watson, B. N. J., Pursey, E., Gandon, S., & Westra, E. R. (2023). Transient eco-evolutionary dynamics early in a phage epidemic have strong and lasting impact on the long-term evolution of bacterial defences. *PLOS Biology*, *21*(9), e3002122. https://doi.org/10.1371/journal.pbio.3002122

Weitz, J. S., Li, G., Gulbudak, H., Cortez, M. H., & Whitaker, R. J. (2019). Viral invasion fitness across a continuum from lysis to latency. *Virus Evolution*, *5*(1). https://doi.org/10.1093/ve/vez006

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Willner, D., Haynes, M. R., Furlan, M., Hanson, N., Kirby, B., Lim, Y. W., Rainey, P. B., Schmieder, R., Youle, M., Conrad, D., & Rohwer, F. (2012). Case Studies of the Spatial Heterogeneity of DNA Viruses in the Cystic Fibrosis Lung. *American Journal of Respiratory Cell and Molecular Biology*, *46*(2), 127–131. https://doi.org/10.1165/rcmb.2011-0253OC

Winstanley, C., O'Brien, S., & Brockhurst, M. A. (2016). Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends in Microbiology*, *24*(5), 327–337. https://doi.org/10.1016/j.tim.2016.01.008

Workentine, M. L., Sibley, C. D., Glezerson, B., Purighalla, S., Norgaard-Gron, J. C., Parkins, M. D., Rabin, H. R., & Surette, M. G. (2013). Phenotypic Heterogeneity of Pseudomonas aeruginosa Populations in a Cystic Fibrosis Patient. *PLoS ONE*, *8*(4), e60225. https://doi.org/10.1371/journal.pone.0060225

Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., & O'Toole, G. A. (2009). Interaction between Bacteriophage DMS3 and Host CRISPR Region Inhibits Group Behaviors of Pseudomonas aeruginosa. *Journal of Bacteriology*, *191*(1), 210–219. https://doi.org/10.1128/JB.00797-08

Zhang, M., Hao, Y., Yi, Y., Liu, S., Sun, Q., Tan, X., Tang, S., Xiao, X., & Jian, H. (2023). Unexplored diversity and ecological functions of transposable phages. *The ISME Journal*. https://doi.org/10.1038/s41396-023-01414-z

Zong, C., So, L., Sepúlveda, L. A., Skinner, S. O., & Golding, I. (2010). Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene. *Molecular Systems Biology*, *6*(1), 440. https://doi.org/10.1038/msb.2010.96

Supplemental Information

The following parameter values were used in the model equations:

$$\frac{dL_i}{dt} = rL_i\left(1 - \sum_{j=1}^{6}(L_j + I_j)\right) - q_i L_i$$

$$\frac{dI_i}{dt} = -\delta I_i + q_i L_i$$

$$\frac{dV_i}{dt} = \beta\delta I_i - \mu V_i.$$

$$q_1 = 0.1, q_2 = 0.102, q_3 = 0.17, q_4 = 0.175, q_5 = 0.288, q_6 = 0.293$$

and

$$r = 1, \beta = 45, \delta = 4, \mu = 1.$$

Each generation lasts 38 minutes. The initial conditions used were $\frac{1}{6} \times 10^{-4}$ for each L_i, which was added at the prescribed times. I_i and V_i were initially zero for each $i = 1, \ldots, 6$.

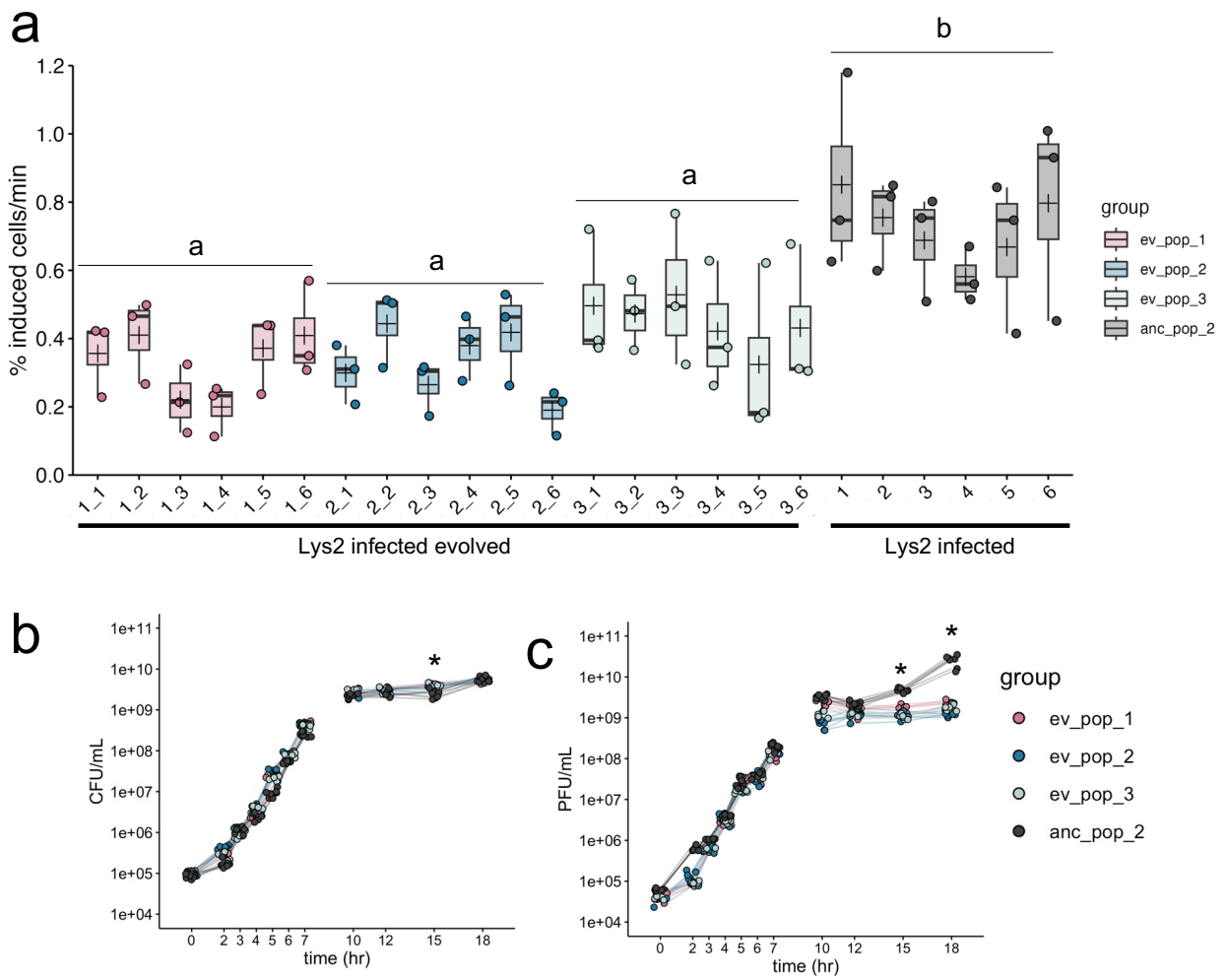

Figure 1. Experimental evolution results in lowered lysogen spontaneous induction. A) Spontaneous induction was measured in exponential phase in six individual isolates from each of three evolved lysogen replicates (ev_pop_1, pink), (ev_pop_2, blue), (ev_pop_3, light blue) and the ancestral strain Lys2 (anc_pop_2, grey). Points are the means of three technical replicates. Bars in the boxplots represent the median; crosses represent the means. Upper and lower bounds of the box are the upper and lower interquartile ranges. Significance was tested with an ANOVA ($F_{3,68}$ = 23.27, p-value = 1.78e-10). Letters indicate significance; groups with different letters have a p-value <0.05; groups with the same or overlapping letters have a p-value of >0.05. B) Bacterial growth curve of all isolates through 18 hours, measured by CFU. C) PFUs sampled through growth curve of all isolates (except the non-lysogenic WT PA14). In B) and C), stationary phase was measured in separate experiments, as indicated by the line breaks between 7 and 10 hours. Points represent the mean of three biological replicates. Asterisks represent a significant difference between the ancestral and evolved populations. Significance was calculated with a Mann-Whitney U test per time point.

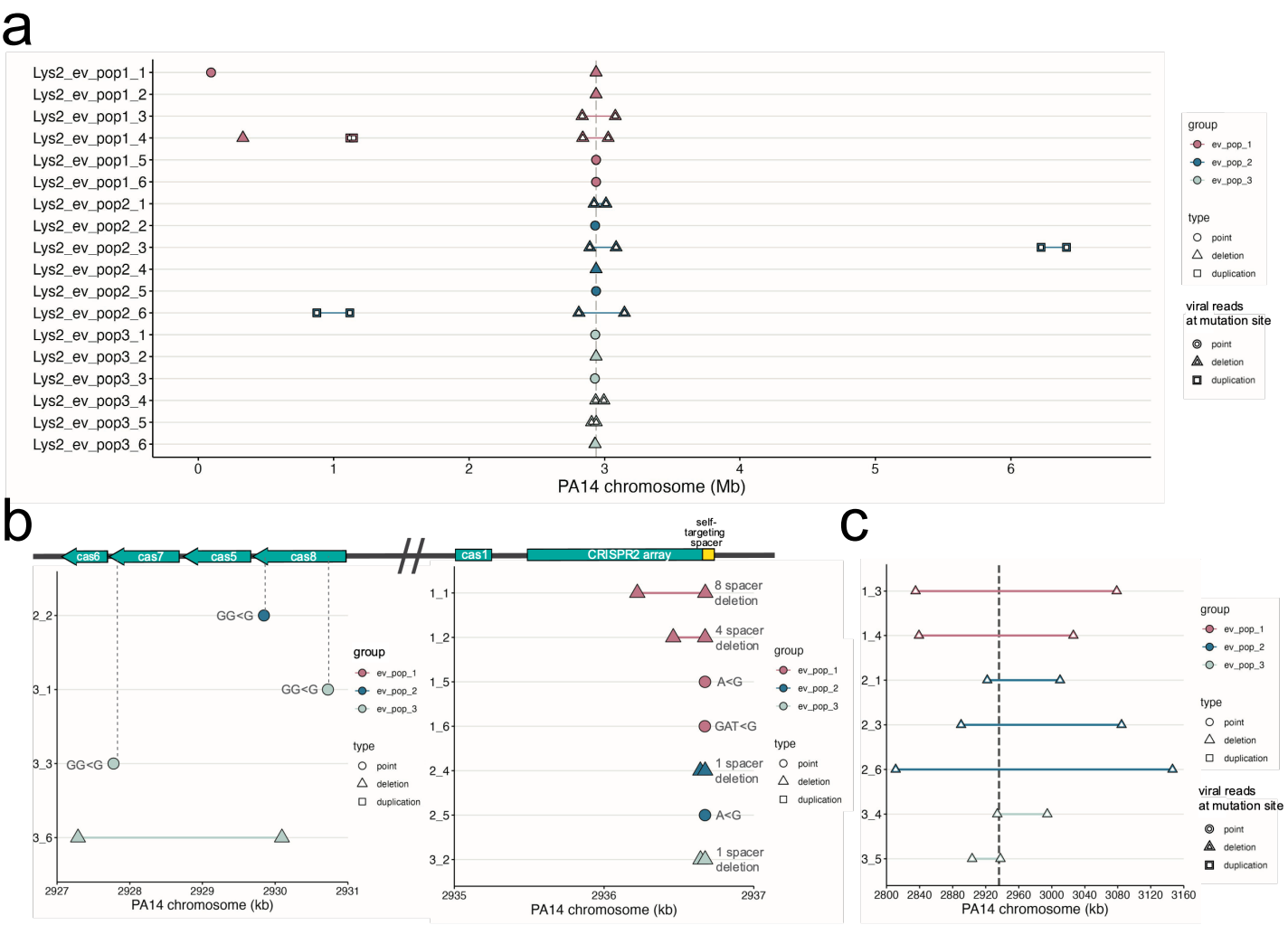

Figure 2. Mutations in infected evolved strains are distinct from uninfected evolved strains. A) Graph of mutations found in all infected evolved samples. B) Close-up of point mutations and small deletions in infected evolved strains. The majority are found in the self-targeting spacer 1 in the 2nd CRISPR array. Gray dashed lines indicate location in the genes. C) Close-up of large deletions in infected evolved strains. Gray dashed double lines indicate the boundaries of the CRISPR-Cas region. In A-C), the y-axis indicates sample ID; x-axis indicates position on the PA14 reference chromosome. Points represent point mutations; triangles spanned by a segment represent deletions of the spanned region; squares spanned by a segment represent duplications of the spanned region. White fill indicates a mutation that was caused by a virus.

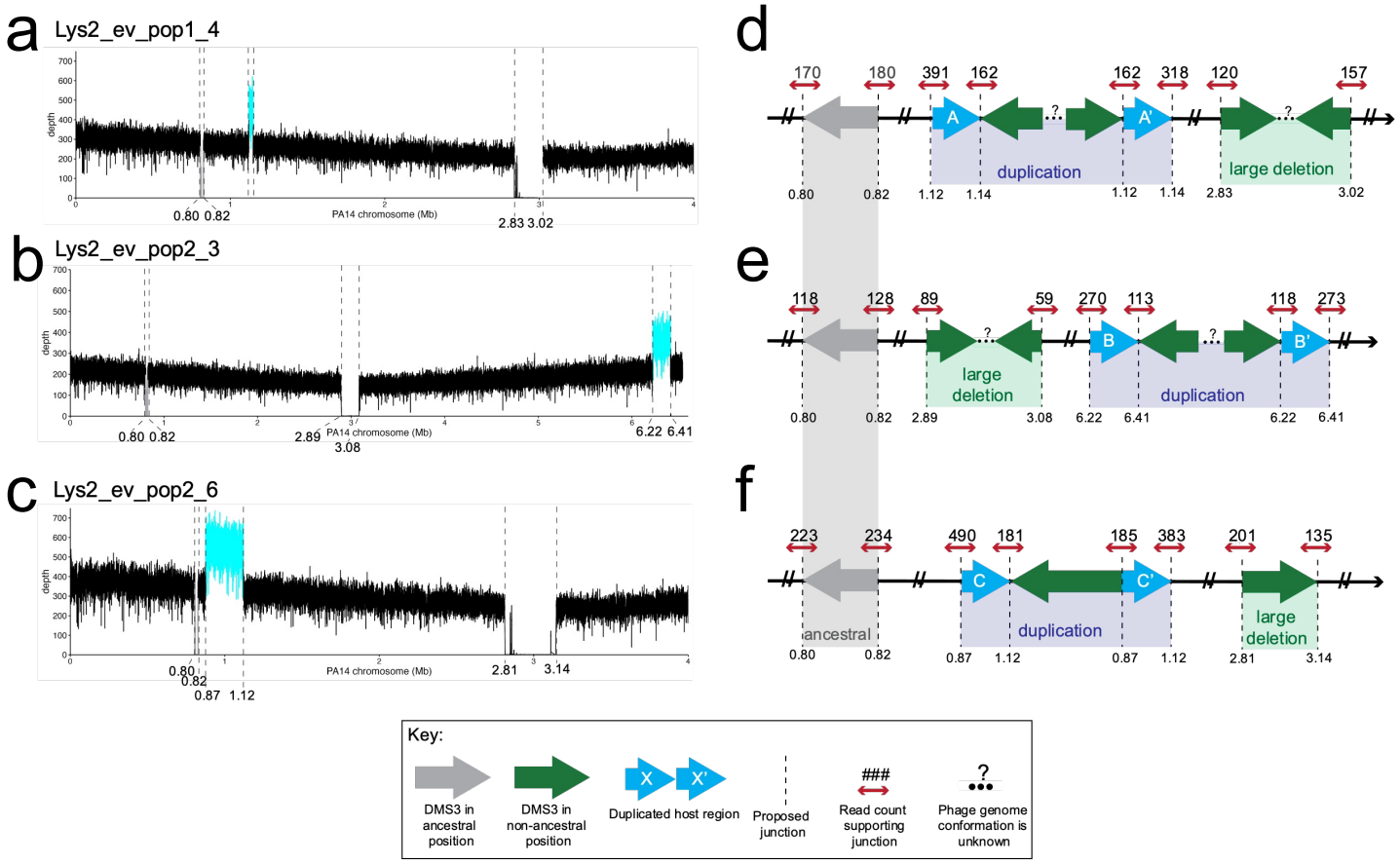

Figure 3. Location and characterization of large duplicated regions in three evolved lysogen isolates. A-C) Coverage plots of the genome. Dashed lines indicate evidence of host-phage boundary. X-axis is position on the PA14 chromosome. Where possible, the x-axis has been truncated to zoom on all possible insertion sites. Y-axis is depth of coverage at that nucleotide. Cyan line represents the putative duplicated region in between two viral insertion sites. Gray line is the region between the ancestral insertion sites. In D-F) Cartoon of resulting genome architecture of the evolved lysogens. Dashed lines indicate a putative host-phage boundary. Gray arrows represent the phage at its ancestral insertion site; green arrows represent phage at new insertion sites. Blue arrows represent duplicate host sequence. The direction of the arrow indicates 5' to 3', with 3' ending at the tip of the arrowhead. Dashed lines represent the junctions between phage and host. Small red arrows indicate host reads; numbers above the arrow indicate read count at that site that supports that junction. Ellipses and question marks between phage genomes represent uncertainty. Read files for each sample and at each location can be found in Supplementary Data. In A and C, regions of coverage in the deletion region in Lys2_ev_pop1_4 and Lys2_ev_pop2_6 are from domains in a TpsA1 and TpsB1 protein (annotated as a filamentous hemagglutinin protein), and in the 3' end of the Lys2_ev_pop2_6 deletion, a domain in OprB (annotated as a porin). Both were identified via BLASTX.

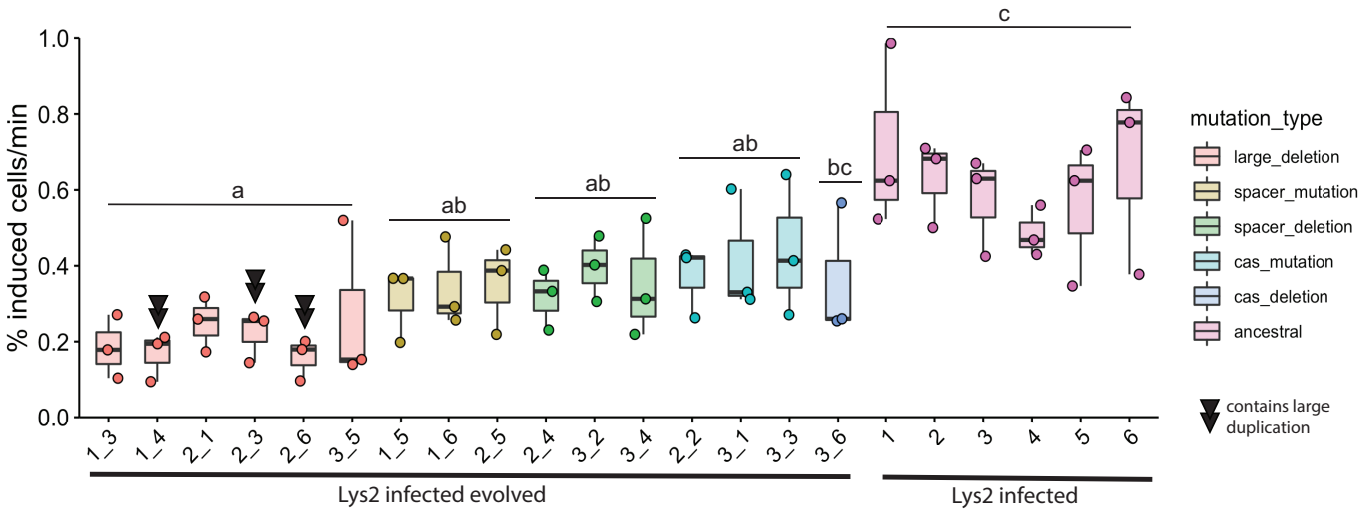

Figure 4. Genomic variation may explain phenotypic variation. Evolved isolates were grouped by type of mutation. Colors of points and boxes represent groups. Points are means of experimental replicates. Box plots are as in Figure 1. Isolates with double black triangles have large duplications which are due to another viral insertion. Significance was determined with an ANOVA with Tukey adjustment using the log-transformed values ($F_{5,66}$ = 19.3, p-value = 8.725e-12). Letters indicate significance; groups with different letters have a p-value <0.05 between them; groups with the same or overlapping letters have a p-value of >0.05 between them.

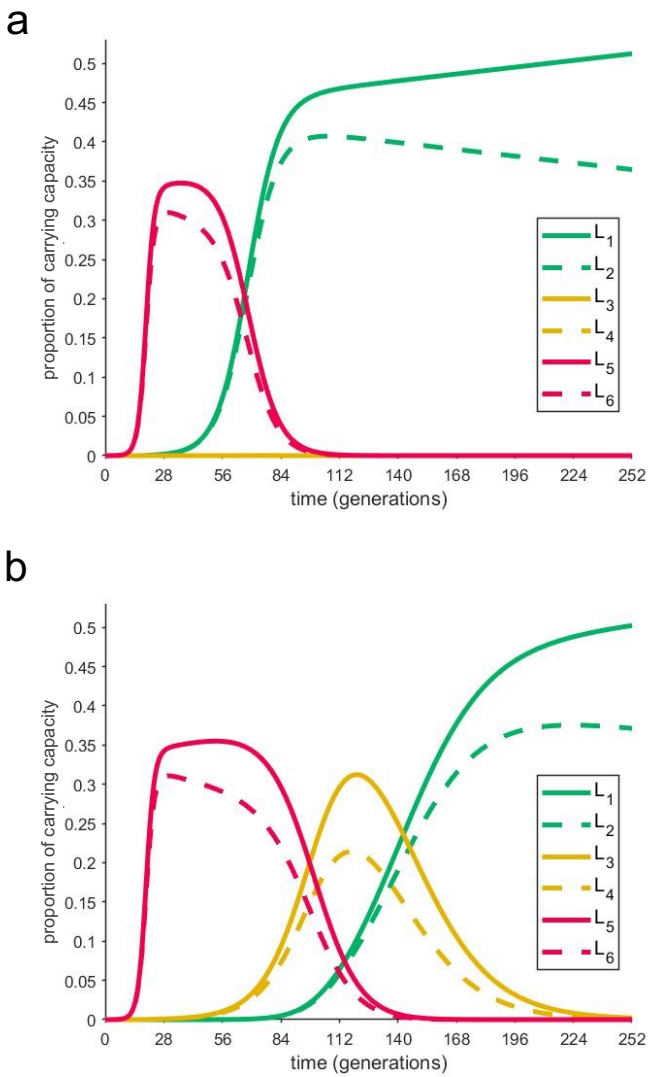

Figure 5. Coexistence of different rates of spontaneous induction depends on the order of their introduction. A mathematical model describes the behavior of several lysogens with different spontaneous induction values. These graphs describe the growth of six lysogens over time in one continuously growing culture. The y-axis represents the proportion of carrying capacity of the medium; the x-axis is time in PA14 generations. Each lysogen was assigned a spontaneous induction value from the experimentally determined range. Each pair (L1/L2; L3/L4; L5/L6) in a certain range are taken from values in the same group, from lowest to highest spontaneous induction values (L1 is the lowest; L6 is the highest). In A) the highest spontaneous inducers are introduced first. At approximately 9 hours, low inducers are introduced; after about 1 day, medium inducers are introduced and do not establish. In B) the highest spontaneous inducers are introduced first. At approximately 9 hours, medium inducers are introduced; after about 1 day, low inducers are introduced and establish, resulting in a period of coexistence between medium and low inducers that is observed experimentally.

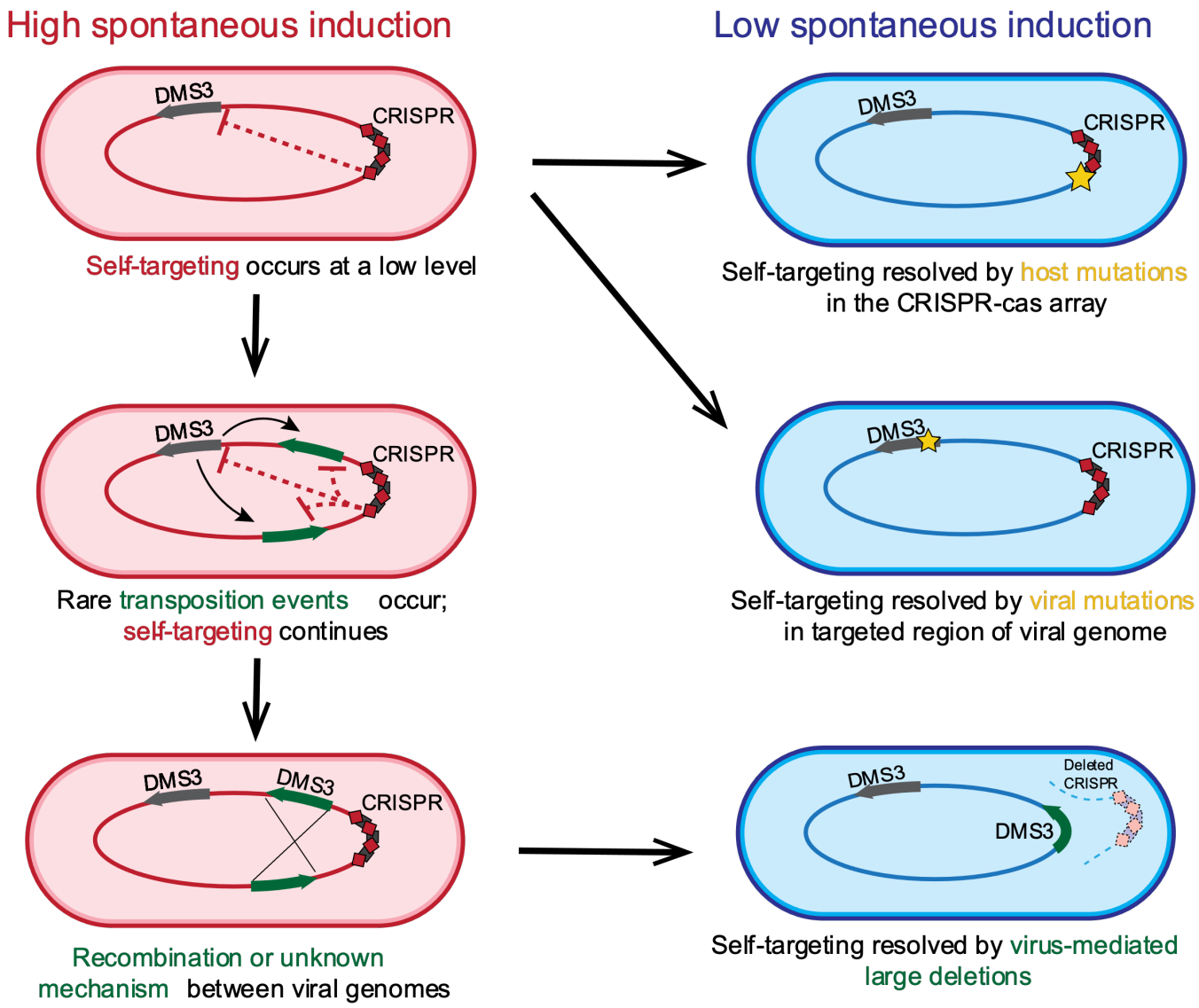

Figure 6. Model of self-targeting leading to DMS3 transposition around the genome. Spontaneous induction caused by CRISPR self-targeting is resolved in two modes. Cells with high spontaneous induction are indicated in the first column, in red. One mode, in the first column, relies on low levels of DMS3 escape from lysogenic repression which result in rare transposition events around the genome. Recombination between multiple DMS3 chromosomes leads to large deletions, which may include the CRISPR area and result in cells with lower spontaneous induction (blue). In the second mode, self-targeting may be resolved by host mutations in the CRISPR-Cas region, or by viral mutations in the targeted region (not recovered in this study).

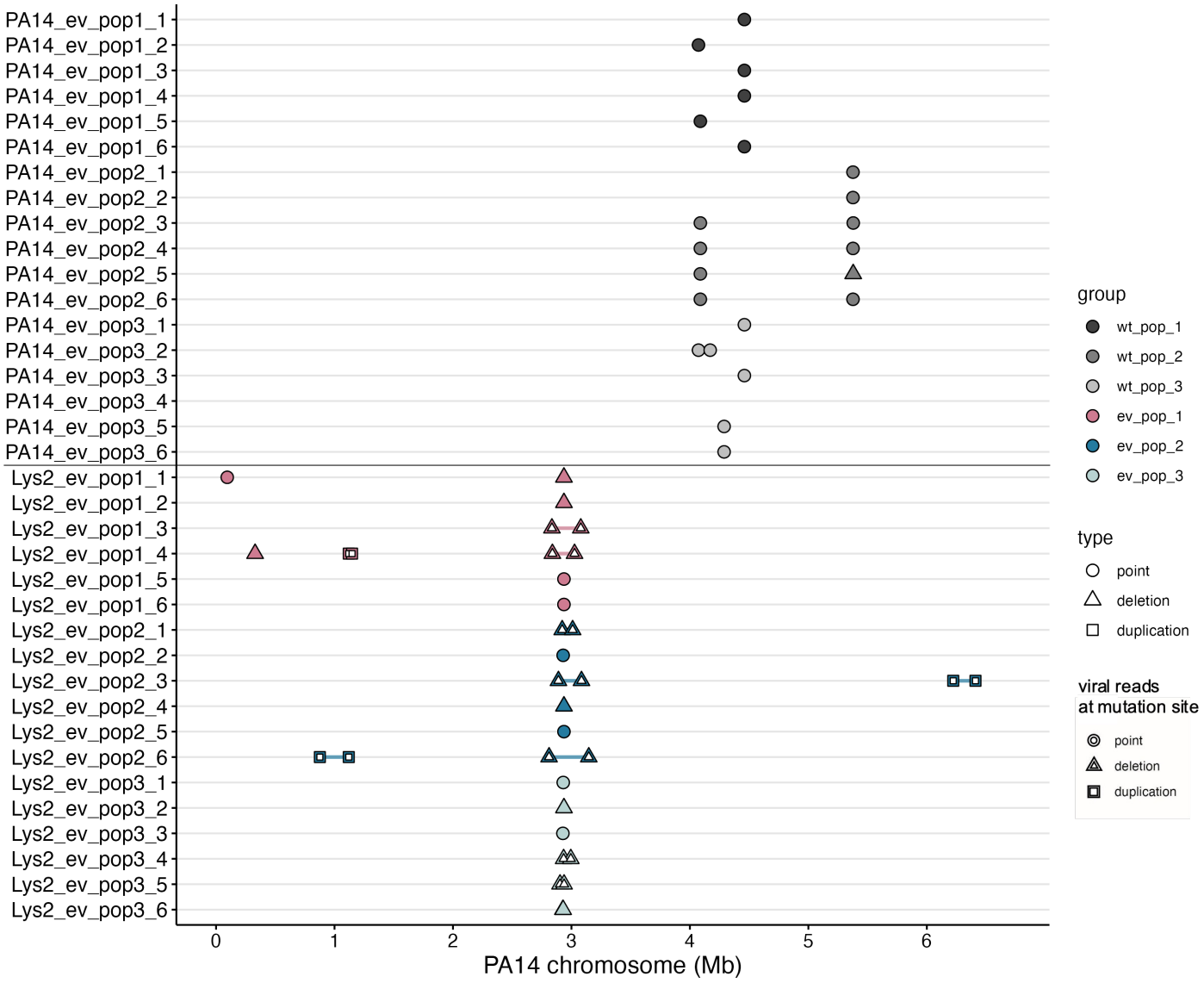

Figure S1. Uninfected populations do not share mutations with lysogen populations. The Y-axis indicates sample ID; X-axis indicates position on the PA14 reference chromosome. Circles represent point mutations; triangles spanned by a segment represent deletions of the spanned region; squares spanned by a segment represent duplications of the spanned region. Inset white shapes indicate a mutation that was caused by a virus. Uninfected population PA14_ev_pop2 contains evidence of phage selection pressure as 100% of isolates have a mutation in the Type 4 pilus. Lysogen data is the same as in Figure 2.

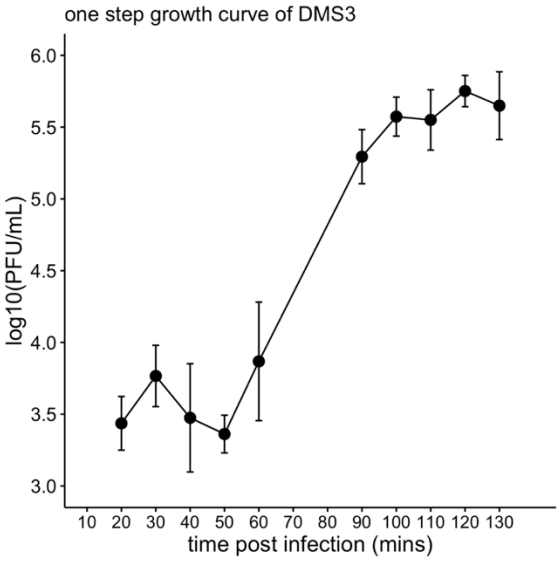

Figure S2. One step growth curve of DMS3. Points represent the mean of experimental replicates (n=6 for all other time points except 20 mins (n=3) and 100 mins (n = 2). Error bars represent standard deviation. The x-axis is time post-infection in minutes; the y-axis is the $\log_{10}$ of PFU/mL. Plateau 1 is the average of all values from 20-50 minutes; Plateau 2 is the average of all values from 100-130 minutes.

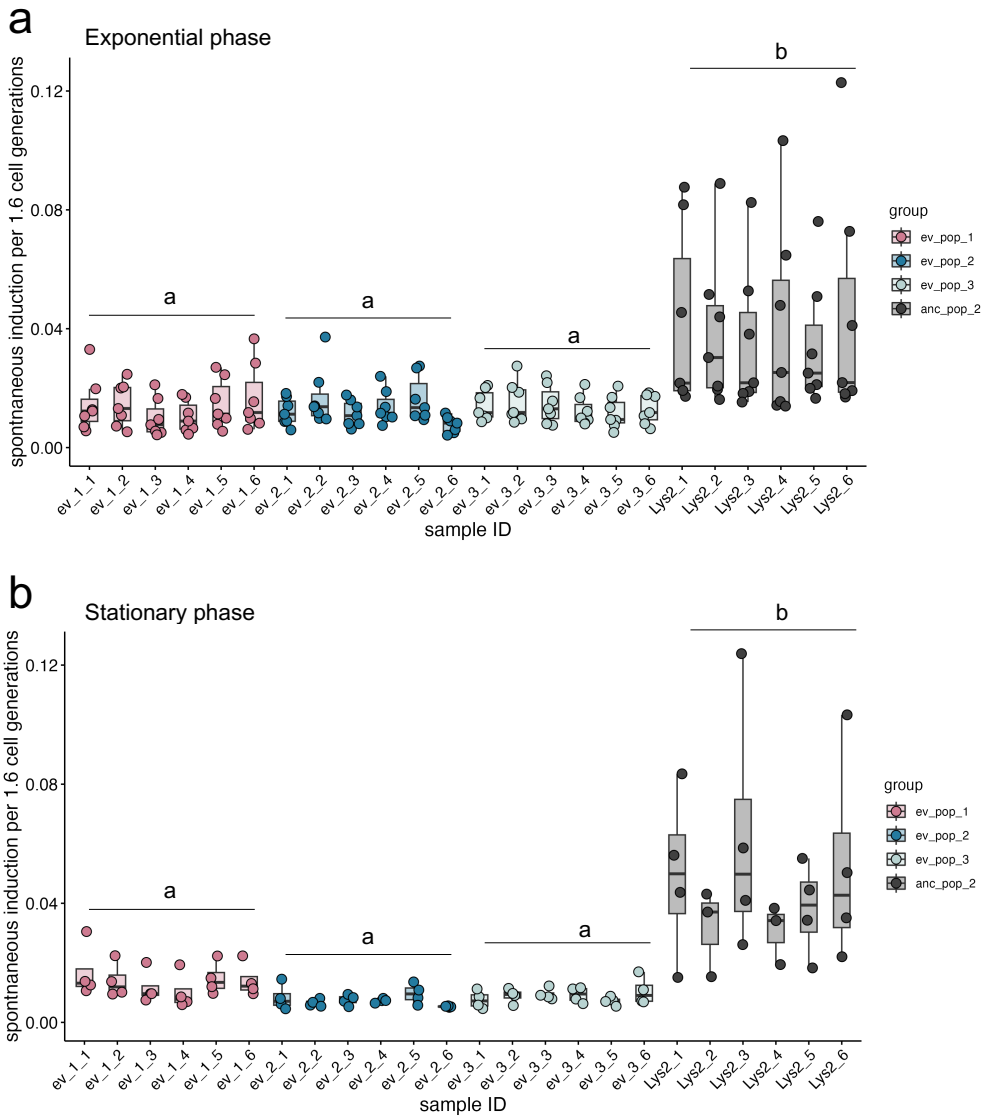

Figure S3. Spontaneous induction calculated with Zong et al formula. A) Spontaneous induction in exponential phase. X-axis is time in hours; y-axis is spontaneous induction per hour, or 1.6 cell generations of PA14. B) Spontaneous induction in stationary phase. X-axis is sample ID; y-axis is spontaneous induction per hour. Significance was calculated as the spontaneous induction value as a function of group with Dunn's Test of Multiple Comparisons and Bonferroni correction. Letters indicate significance; groups with different letters have a p-value <0.05 between them; groups with the same or overlapping letters have a p-value of >0.05 between them.

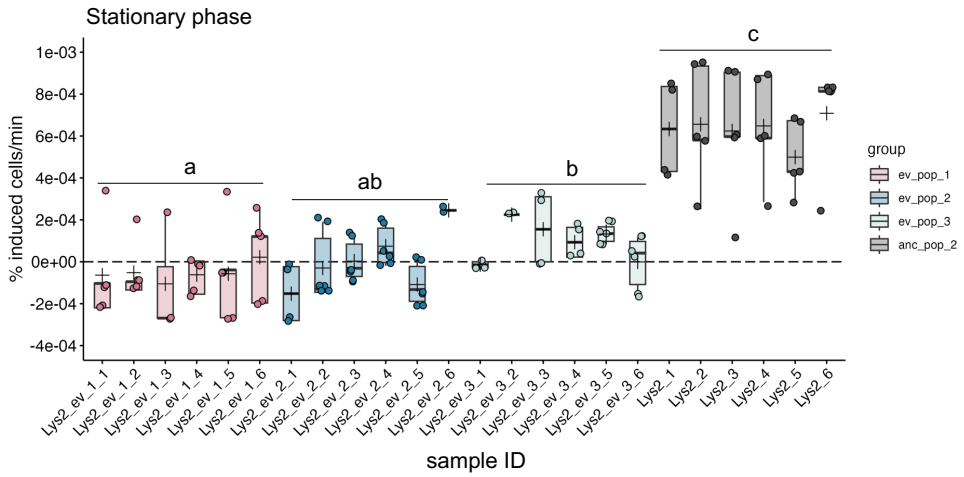

Figure S4. Experimental evolution results in lowered lysogen spontaneous induction in stationary phase. Spontaneous induction was measured in stationary phase. Six individual isolates from each of three lysogen replicates (ev_pop_1, pink), (ev_pop_2, blue), (ev_pop_3, light blue) and the ancestral strain Lys2 (anc_pop_2, grey) were measured. Points are the means of one experimental replicate from three technical replicates. Boxplots are as in Figure 1. Significance was tested with an ANOVA ($F_{3,109}$ = 86.98, P < 2.2e-16). Letters indicate significance; groups with different letters have a p-value <0.05; groups with the same or overlapping letters have a p-value of > 0.05.

Figure S5. DMS3 is partially matched by CRISPR spacers in PA14 and is resolved by host mutations. Map of spacer-protospacer match compared to escape mutations in the host spacer that evolved in two single isolates in parallel cultures. Five mismatched bases on the virus are highlighted in bold red font and lack lines indicating base pairing; escape mutations are highlighted in blue and lack lines indicating base pairing. Hyphens indicate a deletion of those bases.

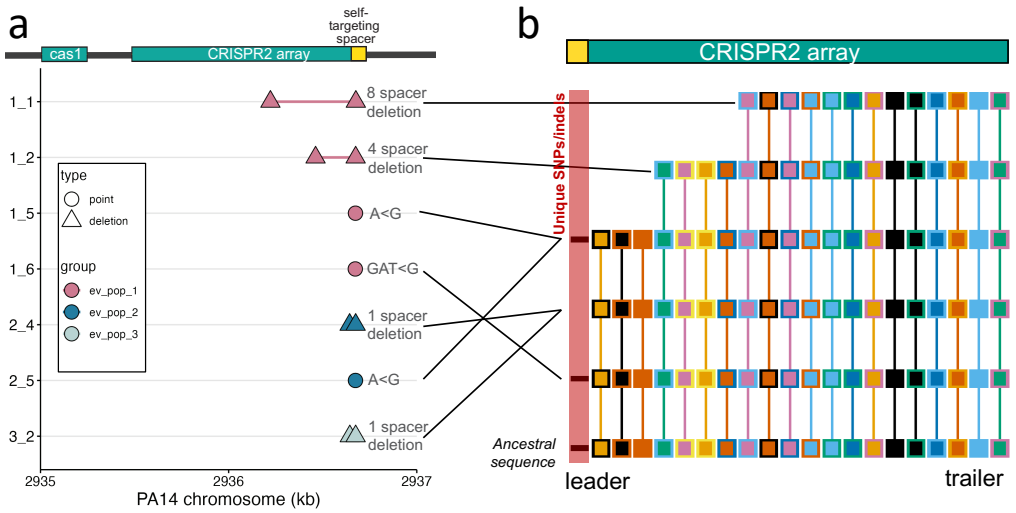

Figure S6. No new spacer acquisition, only spacer evolution in the CRISPR2 array after lysogen evolution. We ran CCTK on the evolved lysogen genomes to look for spacer acquisiton. CCTK recovered all known spacers. A), same figure and data as Figure 2C; B) a CCTK graphical output representing all different arrays from these genomes. Spacers shared between arrays are denoted by boxes with the same colored outline and fill. Spacers that are unique only to that array and which are not shared with any other array are represented as a black line (highlighted with a red box). In our case, these are due not to spacer acquisition but to point mutation which make them non-identical and "unique" with respect to the others. Three cases of unique spacers are due the representation of the ancestral spacer (maintained in all strains which resolve self-targeting via the cas genes) and two unique spacer genotypes.

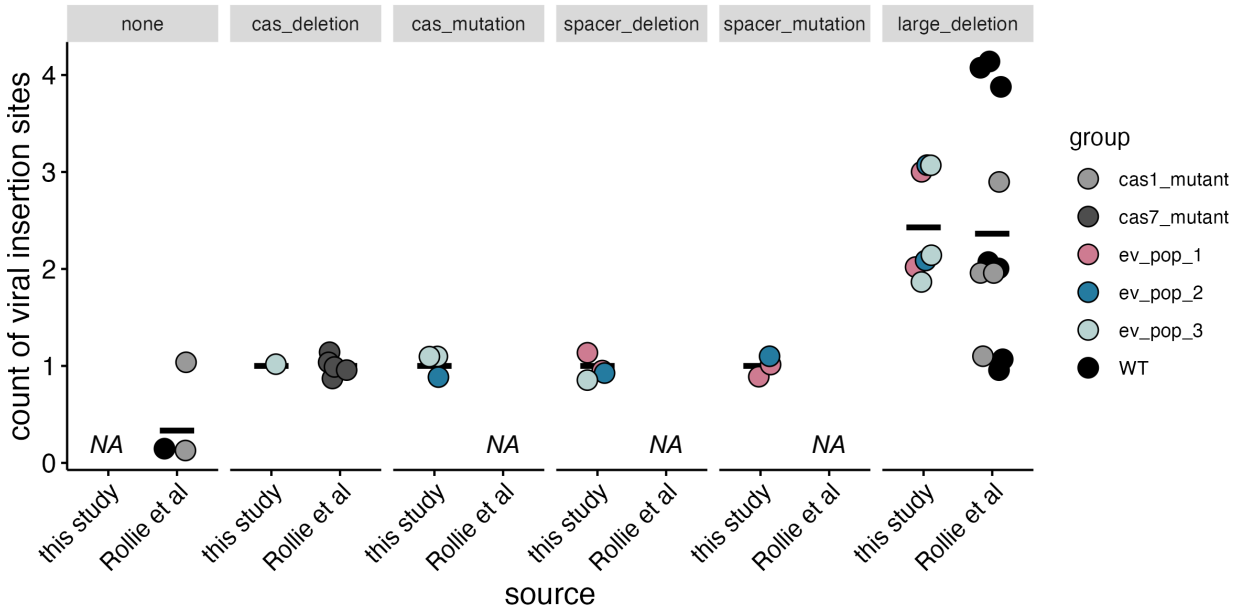

Fig S7. Polylysogeny does not arise in the absence of self-targeting. Y-axis describes the number of viral insertion sites recovered from each isolate; X-axis describes the source of the isolates (either this study or Rollie et al) and is broken up by the type of mutation which the isolate is classified by. Black bars indicate the mean. Points represent one isolate. Point color indicates either what replicate it came from (as in this study) or in what strain background the evolution experiment was started (in Rollie et al). NA: no isolates were recovered in that mutation type category.

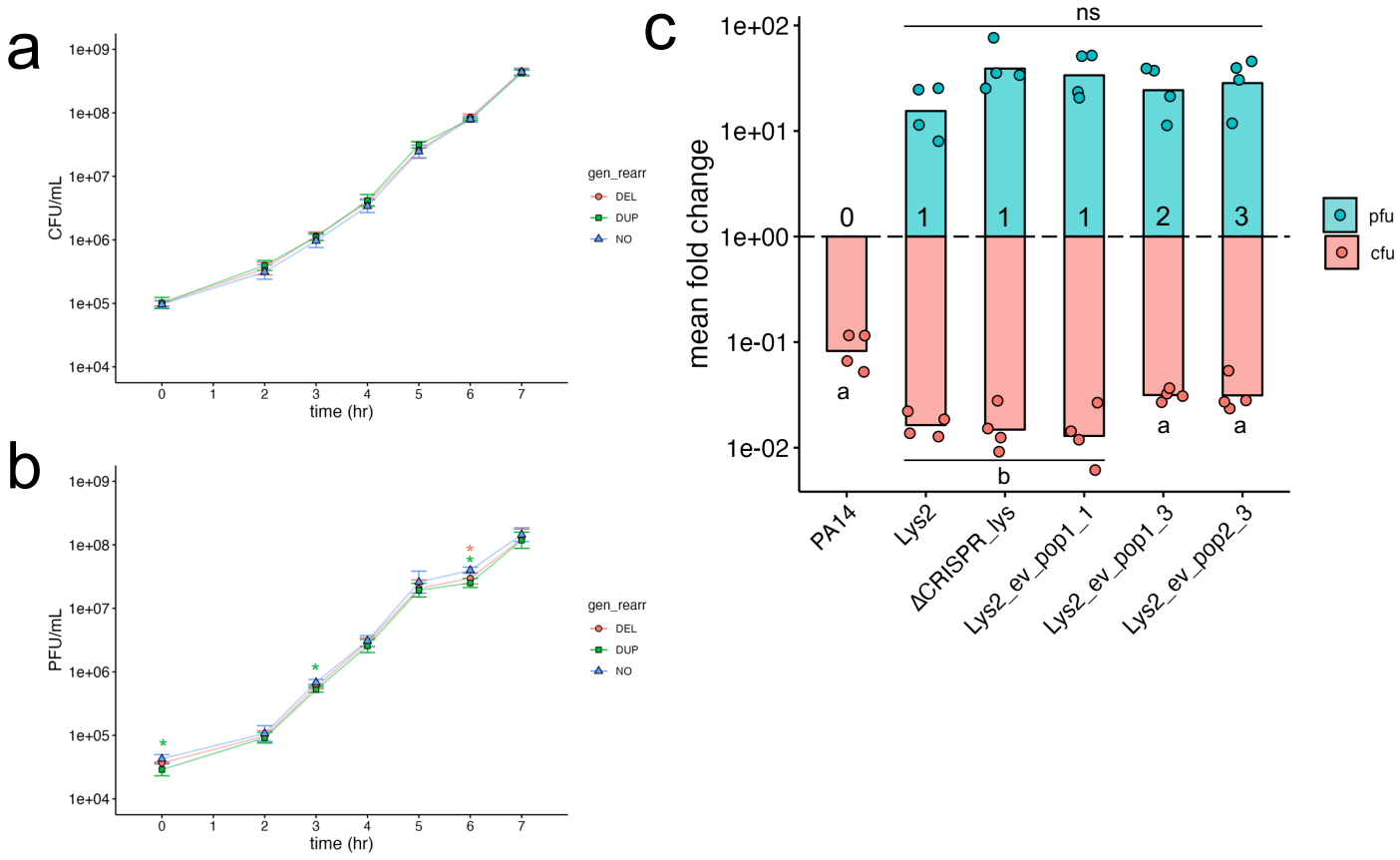

Figure S8. Genome rearrangements are tolerated without growth defects in rich medium in evolved lysogens. A) CFUs and B) PFUs of evolved lysogens in exponential growth. Time is indicated on the x-axis in hours, and CFUs or PFUs are indicated on the y-axis. Asterisks indicate significant differences by Dunn's test between the genomic rearrangement ("DEL": deletion; "DUP": duplication) indicated by the color, and the "NO" genomic rearrangement group (the average of all other lysogens). C) Induction of lysogens with increasing numbers of viral genomes. Numbers in the PFU column represent how many viral genomes are present in the lysogen. Dunn's Test of Multiple Comparisons with a Bonferroni correction was used to test the interaction between the mean fold change and the number of phages present. Points represent the mean of one experimental replicate (three technical replicates each). Bars represent the mean of all experimental replicates. Letters indicate statistical significance between groups; groups with the same letter are not statistically significant. "NS" = not significant.

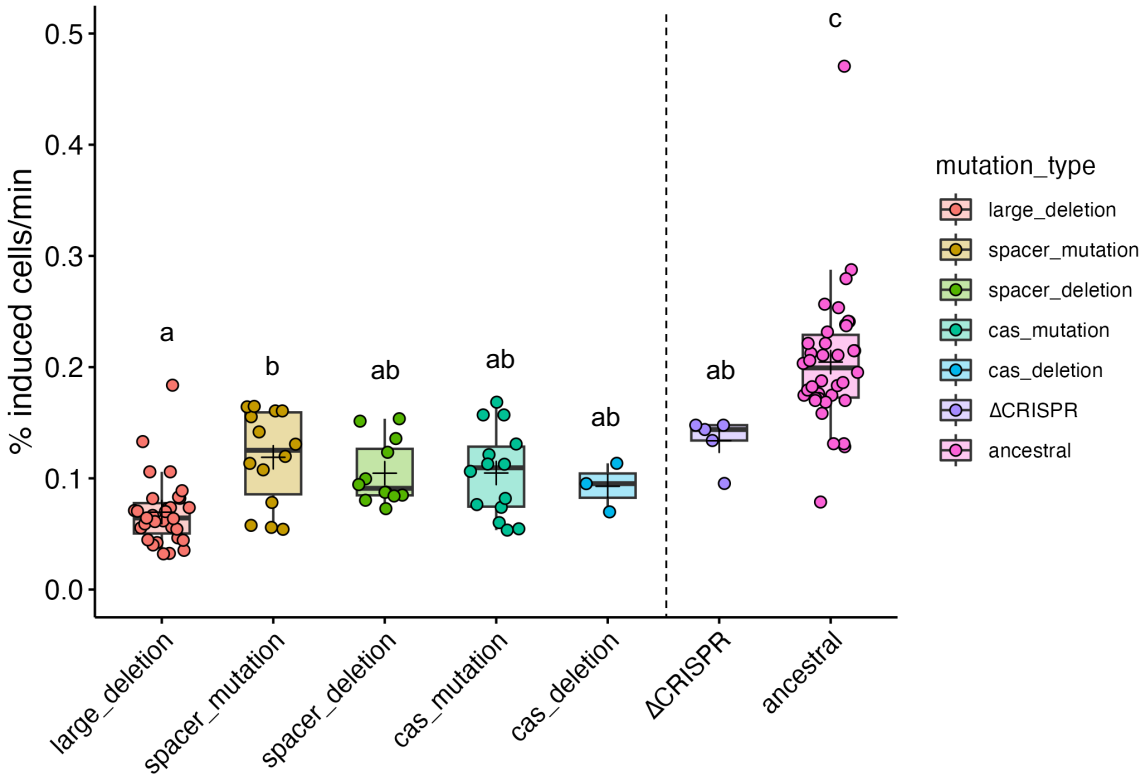

Figure S9. Spontaneous induction differences are maintained between groups and ΔCRISPR lysogen. Spontaneous induction was measured as above, except the growth curve was begun with a 1:100 diluton instead of 1:1000 (OD600 = 0.002). All differences between mutation groups was recovered, except the cas deletion group became significantly different from the ancestral, and the large deletion polylysogen group became significantly different from the spacer mutation group. Significance was tested with an ANOVA ($F_{6,110}$ = 28.35, P < 2.2e-16) with a Tukey adjustment. Letters indicate significance; groups with different letters have a p-value <0.05; groups with the same or overlapping letters have a p-value of > 0.05. Points are the means of one experimental replicate from three technical replicates. A small jitter was added to the horizontal to increase visibility. Bars in the boxplots represent the median; crosses represent the means. Upper and lower bounds of the box are the upper and lower interquartile ranges.

Strain	Description	Source
PA14	WT P. aeruginosa strain	George O'Toole
Lys2	PA14(DMS3) lysogen	Zegans et al, 2009
PA14ΔCRISPR	CRISPR deletion	Cady et al, 2011
PA14ΔCRISPR(DMS3)	CRISPR deletion lysogen	Cady et al, 2011

Table S1. List of strains used in this study.

Lys2 Deletion Boundaries: 806165-826108 on PA14 Reference Genome		
Gene no. (5'-3')	**Deleted gene**	**Description**
1	PhzG	phenazine biosynthesis FMN-dependent oxidase
2	PhzF	phenazine biosynthesis protein; trans-2,3-dihydro-3-hydroxyanthranilate isomerase
3	PhzE	phenazine biosynthesis protein
4	PhzD	phenazine biosynthesis protein
5	PhzC	phenazine biosynthesis protein; phospho-2-dehydro-3-deoxyheptonate aldolase
6	PhzB	phenazine biosynthesis protein
7	PhzA	phenazine biosynthesis protein;
8	PhzM	phenazine-1-carboxylate N-methyltransferase
9	OpmD	multidrug efflux transporter outer membrane subunit; TolC family protein
10	MexI	MexW/MexI family multidrug efflux RND transporter permease subunit
11	MexH	MexH family multidrug efflux RND transporter periplasmic adaptor subunit
12	MexG	DoxX family protein; multidrug efflux RND transporter inhibitory subunit
13	PpgL	Gluconolactonase; 3-carboxymuconate cyclase
14	NmoR	LysR family transcriptional regulator
15	NmoA	nitronate monooxygenase
16	-	D-alanine--D-alanine ligase
17	-	MBL fold metallo-hydrolase

Table S2. Genes contained in Lys2 deletion. All functional predictions were confirmed by BLASTX using the protein RefSeq database.

Group	Sample ID	Mutated region	Description	Location on PA14-REF chromosome
ev_pop1	ev_pop1_1	*fliF*	Nonsense mutation in flagellar M-ring protein $_{G→A}$	4459793
ev_pop1	ev_pop1_2	*fliP*	Nonsynonymous mutation in start codon in flagellar biosynthesis protein $_{C→T}$	4072428
ev_pop1	ev_pop1_3	*fliF*	Nonsense mutation in flagellar M-ring protein $_{G→A}$	4459793
ev_pop1	ev_pop1_4	*fliF*	Nonsense mutation in flagellar M-ring protein $_{G→A}$	4459793
ev_pop1	ev_pop1_5	*lasR*	Nonsynonymous mutation in transcriptional activator $_{C→T}$	4088064
ev_pop1	ev_pop1_6	*fliF*	Nonsense mutation in flagellar M-ring protein $_{G→A}$	4459793
ev_pop1	ev_pop2_1	*pilY1*	Frameshift mutation in T4P biogenesis factor protein $_{GG→G}$	5376474
ev_pop1	ev_pop2_2	*pilY1*	Frameshift mutation in T4P biogenesis factor protein $_{GG→G}$	5376474
ev_pop1	ev_pop2_3	*lasR*	Nonsynonymous mutation in transcriptional activator $_{C→T}$	4088064
ev_pop2	ev_pop2_3	*pilY1*	Nonsynonymous mutation in T4P biogenesis factor protein $_{T→A}$	5378668
ev_pop2	ev_pop2_4	*lasR*	Nonsynonymous mutation in transcriptional activator $_{G→A}$	4088043
ev_pop2	ev_pop2_4	*pilY1*	Frameshift mutation in T4P biogenesis factor protein $_{GG→G}$	5376588
ev_pop2	ev_pop2_5	*lasR*	Nonsynonymous mutation in transcriptional activator $_{C→T}$	4088064
ev_pop2	ev_pop2_5	*pilY1*	1kb deletion in T4P biogenesis factor protein	5377966
ev_pop2	ev_pop2_6	*lasR*	Nonsynonymous mutation in transcriptional activator $_{G→A}$	4088043
ev_pop2	ev_pop2_6	*pilY1*	Frameshift mutation in type 4 pilus biogenesis factor protein $_{GG→G}$	5376588
ev_pop2	ev_pop3_1	*fliF*	Nonsense mutation in flagellar M-ring protein $_{G→A}$	4459793
ev_pop2	ev_pop3_2	*fliP*	Nonsynonymous mutation in start codon in flagellar biosynthesis protein $_{C→T}$	4072428
ev_pop3	ev_pop3_2	MFS transporter	Synonymous mutation $_{G→A}$	4170934
ev_pop3	ev_pop3_3	*fliF*	Nonsense mutation in flagellar M-ring protein	4459793
ev_pop3	ev_pop3_4	*No mutations*	*NA*	*NA*
ev_pop3	ev_pop3_5	C4 RNA	Point mutation in predicted C4 RNA	4288779
ev_pop3	ev_pop3_6	C4 RNA	Point mutation in predicted C4 RNA	4288779

Table S3. Description of mutations recovered in evolved uninfected PA14.