Prophage maintenance is determined by environment-dependent selective sweeps rather than mutational availability

Journal Article

Author(s):
Bailey, Zachary M.; Igler, Claudia; Wendling, Carolin C.

Publication date:
2024-04-22

Permanent link:
https://doi.org/10.3929/ethz-b-000669704

Originally published in:
Current Biology 34(8), https://doi.org/10.1016/j.cub.2024.03.025

Funding acknowledgement:
179743 - Evolution and Spread of Antibiotic Resistance Encoded by Prophages (SNF)
19-2 FEL-74 - Dynamics of phage lysis systems and their evolutionary advantages (ETHZ)

Current Biology

CellPress
OPEN ACCESS

Article

Prophage maintenance is determined by environment-dependent selective sweeps rather than mutational availability

Zachary M. Bailey,[1,4,6,*] Claudia Igler,[1,2] and Carolin C. Wendling[1,3,5]
[1]Institute of Integrative Biology, ETH Zürich, 8092 Zürich, Switzerland
[2]Division of Evolution, Infection and Genomics, School of Biological Sciences, University of Manchester, Manchester M13 9PL, UK
[3]Present address: Max von Pettenkofer Institute of Hygiene and Medical Microbiology, Faculty of Medicine, LMU Munich, 80336 Munich, Germany
[4]X (formerly Twitter): @zbailey1991
[5]X (formerly Twitter): @CCWendling
[6]Lead contact
*Correspondence: zachary.bailey@env.ethz.ch
https://doi.org/10.1016/j.cub.2024.03.025

SUMMARY

Prophages, viral sequences integrated into bacterial genomes, can be beneficial and costly. Despite the risk of prophage activation and subsequent bacterial death, active prophages are present in most bacterial genomes. However, our understanding of the selective forces that maintain prophages in bacterial populations is limited. Combining experimental evolution with stochastic modeling, we show that prophage maintenance and loss are primarily determined by environmental conditions that alter the net fitness effect of a prophage on its bacterial host. When prophages are too costly, they are rapidly lost through environment-specific sequences of selective sweeps. Conflicting selection pressures that select against the prophage but for a prophage-encoded accessory gene can maintain prophages. The dynamics of prophage maintenance additionally depend on the sociality of this accessory gene. Prophage-encoded genes that exclusively benefit the lysogen maintain prophages at higher frequencies compared with genes that benefit the entire population. That is because the latter can protect phage-free "cheaters," reducing the benefit of maintaining the prophage. Our simulations suggest that environmental variation plays a larger role than mutation rates in determining prophage maintenance. These findings highlight the complexity of selection pressures that act on mobile genetic elements and challenge our understanding of the role of environmental factors relative to random chance events in shaping the evolutionary trajectory of bacterial populations. By shedding light on the key factors that shape microbial populations in the face of environmental changes, our study significantly advances our understanding of the complex dynamics of microbial evolution and diversification.

INTRODUCTION

Prophages, viruses integrated into bacterial genomes, present a challenging conflict for their hosts. The ability of a prophage to revert to the lytic cycle, leading to the production of new phage particles that subsequently lyse and kill the bacterial cell, puts their hosts at constant risk of death. Yet, their presence in bacterial genomes can also increase bacterial fitness in several ways. First, lysogens, cells carrying one or more prophages, are protected from superinfection of bacteriophages from the same immunity class through superinfection exclusion provided by the integrated prophage.[1] Second, prophages can encode accessory genes that can modify their host's phenotype through a process known as lysogenic conversion.[2] The expression of these genes, whose products often confer a cooperative advantage that benefits all cells in the population irrespective of whether they carry that gene or not,[3] can increase the fitness of lysogens and potentially surrounding bacteria in certain environments. Lastly, prophages can switch back to the lytic cycle, during which they replicate within the bacterial host cell. The subsequently released phage particles can kill phage-susceptible competitors, which enhances the fitness of the remaining lysogen population.[4,5]

While the benefits of prophages are evident, they also come with costs. The expression of viral proteins during lysogenic conversion can impose a metabolic burden on the host, making the prophage costly in certain environments.[6,7] Additionally, induction of the lytic cycle, which can either occur spontaneously or in response to external DNA-damaging stressors,[8–11] results in the death of the bacterial cell. Depending on the rate of induction, i.e., the fraction of the bacterial population that undergoes lysis, such events can drive entire lysogen populations to extinction (Figure S1A). Bacteria can mitigate those costs in several ways, including the accumulation of deleterious mutations in the prophage[12–14] or mutations in the bacterial genome that alter phage induction rate.[15] Alternatively, bacteria can eliminate

prophages from their genomes[12,13,16,17] through incomplete activation or complete deletion.[18,19]

Despite these mechanisms to ameliorate the costs of prophages, the persistence of functional prophages in a majority of bacterial genomes, up to 83%,[20] raises important evolutionary questions about the selective forces that maintain or abolish prophages in nature. Specifically, it remains unclear under which conditions prophages become costly enough to be lost from a bacterial population, and under which conditions they provide enough benefits to be maintained. Here, we explore how prophage maintenance and the evolutionary trajectory of lysogen populations are influenced by the frequency of prophage induction and any environment-dependent costs and benefits associated with both the prophage[7] and the survival of the bacterial host.

To explore the dynamics of prophage maintenance, we coevolved lysogenic *E. coli* populations carrying lambda prophages, under selection pressures that altered the net fitness effect of prophage carriage. We found that high rates of prophage induction selected for fast prophage loss driven by strong directional selection, resulting in environment-dependent selective sweeps. The first was a transient sweep during which bacterial populations acquired chromosomal phage resistance to prevent reinfections by free phages. The second sweep was characterized by an increasing frequency of phage-free cells, with a combined phage-resistant and phage-free genotype rapidly reaching fixation after approximately 500 generations. Using an empirically parameterized model, we found that conflicting selection pressures and prophage-encoded benefits can slow down prophage loss. Our parameter sensitivity analysis showed that prophage loss is largely driven by environmental selection pressures rather than mutation rates. By revealing how bacteria coexist with and adapt to prophages, we uncovered a crucial mechanism that drives microbial interactions and influences the evolutionary path of microbial population dynamics and ultimately diversity.

RESULTS

To study the evolutionary dynamics of prophage maintenance in lysogen populations, we evolved three distinct *E. coli* genotypes: *n_lys* (non-lysogens), *s_lys* (lysogens carrying a Lambda+ prophage), and *r_lys* (lysogens carrying a Lambda+ prophage that encodes a beta lactamase gene providing ampicillin resistance [Amp^R]) in four different selection regimes for 30 serial transfers. These regimes imposed varying selection pressures on prophage carriage: (1) a control regime (nutrient-rich broth), (2) a selection regime that selected against the prophage (broth with mitomycin C [MMC] to increases prophage induction), (3) a selection regime for the prophage-encoded Amp^R (broth with sub-inhibitory concentrations of Amp), and (4) a conflicting selection regime against the prophage but for its encoded Amp-resistance gene (broth with both Amp and MMC).

Production of free phages varies across lysogen genotypes and selection regimes

First, we compared the ecological dynamics of bacteria and free phages between genotypes and treatments. The overall number of bacteria was reduced on transfer 1 (T1) in populations exposed to Amp and/or MMC. However, population sizes recovered after the first week and stabilized at levels comparable to the control (Figures S2A and S2B), suggesting rapid adaptation to the different selection regimes.

In contrast, phage densities varied significantly among selection regimes (significant interaction in generalized least square (gls)-model for treatment, genotype, and time: $F_{1,\ 328} = 26.55$, $p < 0.001$; Figures S2B and S2C). In the absence of MMC, all lysogen populations produced a maximum phage titer of 10^5 plaque-forming units (PFUs)/mL on T1 before declining to 0 PFUs/mL (Figure S2B). This dynamic is not surprising because introduction into nutrient-rich media can initially lead to an increase in spontaneous induction before the number of free phages stabilized at approximately $\sim 10^{-8}$ inductions per generation.[21,22]

In the absence of Amp but the presence of MMC, phage production increased from an average of 1.4×10^6 PFUs/mL at T1 to a maximum of 9.1×10^7 PFUs/mL on T8 (Figure 1) before significantly declining (significant time effect in gls-model: $F_{1,\ 52} = 4.97$, $p = 0.03$; Figure 1), irrespective of phage genotype (non-significant genotype effect in gls-model: $F_{1,\ 52} = 0.64$, $p = 0.43$), over the remaining experiment to an average of 4.3×10^5 PFUs/mL at T29. In extreme cases, i.e., in two *r_lys* populations from T22 and one *s_lys* population from T15 onward, we observed zero phage production.

This overall reduction in phage production seen under MMC selection contrasts sharply with populations exposed to the competing selective pressures, Amp + MMC. While the number of free phages in all *r_lys* populations remained stable over time (no significant time effect in gls-model: $F_{1,\ 28} = 0.44$, $p > 0.05$; Figure 1D), we observed strong fluctuations and no clear trend, across *s_lys* populations (Figure 1B).

At the end of the experiment, we tested the control populations exposed to Luria Broth (LB), for the presence of prophages using an MMC induction assay. The assay confirmed that all *n_lys* populations were indeed phage free, whereas *r_lys* and *s_lys* populations remained inducible by MMC at the end of the experiment.

Decline in phage particles at the population level is driven by individual cells

The consistent decline in free phages under MMC selection could be caused by a reduction in average phage production across the entire population, or by individual cells within a given population that did not produce any phage particles. To differentiate between these two possibilities, we quantified phage particle production of 24 randomly selected clones from two replicate populations subjected to MMC from time points with high (T8) or low (T29) levels of free phages (Figure S3A). Overall, phage particle production followed a bimodal distribution (Figure S3B) with one local maximum at zero, indicating no phage particle production, and a second local maximum at $1.64 \times 10^6 \pm 2.93 \times 10^5$ PFUs/mL, which is close to the average level from the start of the experiment at $9.14 \times 10^5 \pm 3.19 \times 10^5$ PFUs/mL (Figure S3B). Considering the change in free phage abundance between T8 and T29 in the MMC selection regime, we observed, irrespective of the underlying genotype, two different dynamics: either the population size of free phages decreased from its maximum on T8 to zero by T29 or there was no phage production at either time point (Figure S3).

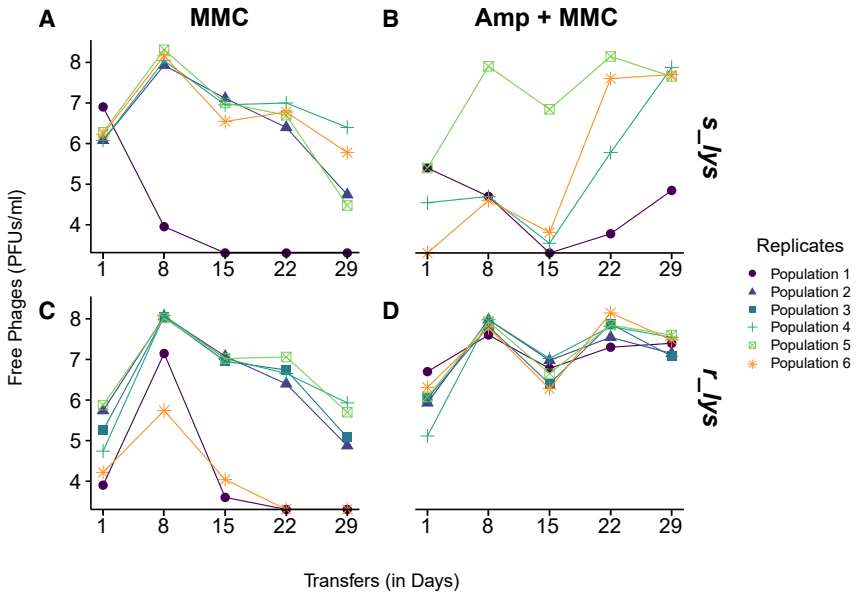

Figure 1. Free phage particle dynamics of evolved lysogens

Ecological dynamics of free phage particles in lysogen populations evolved in the presence of MMC (A and C) or Amp + MMC (B and D). Log values of the number of phages (plaque-forming units [PFUs]/mL) are shown for 29 transfers of selection. Phages either carried no Amp-resistance gene (Amp^R), i.e., susceptible lysogens *s_lys* (A and B), or carried the corresponding Amp^R, i.e., resistant lysogens *r_lys* (C and D). Replicate populations are differentiated by symbol and color.
See also Figures S1 and S2.

Repeating the same analysis for 24 randomly selected clones from two replicate populations subjected to Amp + MMC revealed overall the same bimodal distribution as seen with MMC samples. However, in stark contrast to MMC dynamics, phage production in the Amp + MMC selection regime remained stable within a given population over time (*s_lys*: t = 5.83, p = 1 and *r_lys*: t = 0.30, p = 0.618; Figure S3).

This bimodal distribution of phage particle production suggests that the decline in free phages under MMC selection can be explained by individual cells that have lost the ability to produce free phages potentially due to the loss of Lambda+ within their genome.

Prophage loss is accompanied with phage and antibiotic resistance mutations

To elucidate the molecular mechanisms underlying the absence of phage particle production, we sequenced two clones from two *s_lys* and two *r_lys* populations exposed to MMC or Amp + MMC, and two clones from one *s_lys* and one *r_lys* control population at T8 and T29. From these, we identified three main sets of genetic changes, each comprising one or more genetic loci, which can account for the observed phenotypic changes and ecological dynamics in the selection experiment. Strong parallel signatures observed across populations for each of these three groups indicate adaptive evolution to the respective selection regime. To determine the final frequencies of each mutation within every population, we additionally sequenced whole populations from the last transfer of the experiment (Figure S3C).

The first set of genetic changes involves molecular resistance of bacterial cells to Lambda+ adsorption through surface receptor mutations (SRMs; Figure 2). This set includes *malT*, which is involved in the uptake and utilization of maltose,[23] and confers resistance to Lambda+.[24,25] In lysogens exposed to MMC and Amp + MMC, the frequency of non-synonymous single-nucleotide variants (SNVs), insertions, and non-sense mutations within *malT* increased from 75% of the sampled clones on T8 to 100% on T29 (Figure 2A). To test whether these mutations confer phage resistance through SRMs, i.e., lambda phage receptor inactivation, we assessed the ability of 24 individual clones from two replicate populations from T1–T5, T7, and T29 in *maltose* uptake by using a tetrazolium tetra-chloride (TTC) assay.[26] This assay revealed that the frequency of SRMs via mutations in the maltose pathway rapidly increased to above 75% within the first 48 h and reached a frequency of 100% in all *s_lys* and *r_lys* populations treated with MMC and Amp + MMC by T29 (Figure S3D).

The second genetic change involves variations in the presence or absence of the Lambda+ prophage (Figure 2). In the absence of Amp but presence of MMC, Lambda+ was lost from all sequenced clones. This was supported by our population-level sequencing data, which suggested that Lambda+ was present at less than 1% in each lysogenic population from this treatment at T29 (Figures 2 and S3C). This stands in sharp contrast to the Amp + MMC selection regime, where Lambda+ persisted at nearly 100% in every *s_lys* population or ranged from ~23% to 100% across *r_lys* populations (Figure S3C). In all other *s_lys* and *r_lys* populations which were not treated with MMC, Lambda+ persisted in the bacterial genomes.

The third set of genetic changes includes various genes linked to Amp^R, which did not appear in the absence of Amp. This set includes SNVs, deletions, and insertions in four different loci: the multidrug resistance genes *marR* and *mdtC*, the multidrug efflux pump *acrR*, and the outer membrane gene *ompF* (Figure 2; Tables S1 and S2).[27–30] Mutations in *marR* or *acrR* reached fixation in three out of four *s_lys* populations (Figure S3C; Table S2). *MdtC* mutations are present in one of the eight *s_lys* clones treated with Amp + MMC (Figure 2). We also confirmed the resistance provided by these four mutations through a minimum inhibitory concentration (MIC) assay.[27–30] We found that these mutations confer resistance to the Amp concentration used in the selection experiment (Figure S3E) which explains the significant increase in final population sizes of initially Amp-susceptible populations in the selection experiment over time (effect of Amp on population sizes in gls: F = 120.6; p < 0.01; Figures S1B and S2A).

Amp targets the synthesis of new peptidoglycan for the cell wall of bacteria leading to cell death by cell wall disruption.[31] We observed several loci that could be associated with Amp^R as they are targeted by Amp such as *ftsI, mepS, ddlA*, and *pgaB*. *DdlA*, present in one clone and *ftsI*, present in two clones

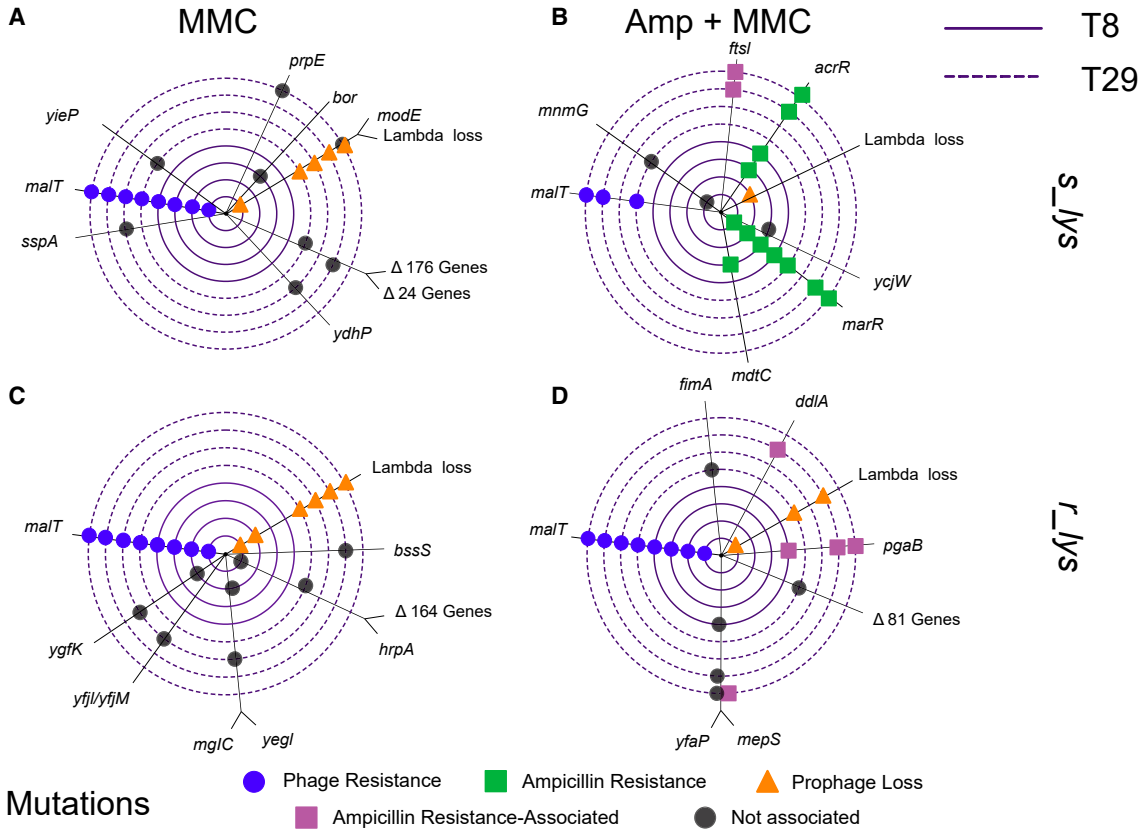

Figure 2. Parallel genomic evolution in *s_lys* and *r_lys* populations

Genetic loci under positive selection as indicated by parallel genomic evolution in *s_lys* (A and B) and *r_lys* (C and D) populations exposed to MMC (A and C) or Amp + MMC (B and D). Each concentric circle corresponds to a single clone; inner circles (solid) correspond to clones from T8, outer circles (dashed) correspond to clones from T29. Each colored point corresponds to one mutation event on the respective clone. Mutations that cause ampicillin resistance (*marR* and *acrR*) are shown in green, mutations that are associated with ampicillin resistance but do not confer resistance alone are in pink, mutations associated with bacteriophage resistance (*MalT*) in blue, and deletion of Lambda+ in orange. Mutations that could not be associated with the different phenotypes are shown in gray. For more detailed information on the underlying mutations see Table S1. For more detailed information on mutations and population frequencies see also Figure S3 and Tables S1 and S2.

are for instance both involved in peptidoglycan synthesis.[32–34] *MepS* and *pgaB* are outer membrane lipoproteins that are associated with the maintenance of the *E. coli* exoskeleton, including peptidoglycan, and are present in one and three clones, respectively.[35,36]

Although not visible at the clonal level, our population-level sequencing revealed that two out of six *r_lys* populations exposed to Amp + MMC acquired mutations associated with Amp^R as well (Figure S3C). These two populations exhibited the highest frequency in prophage loss among all six-replicate *r_lys* populations, i.e., over 75% of the clones lost Lambda+ and its encoded Amp^R. This suggests that the chromosomal Amp^R (c_Amp^R) mutations compensate for the loss of the phage-encoded Amp^R in *r_lys* populations subjected to the Amp + MMC selection regime. Further, our population-level sequencing of *n_lys* and *s_lys* populations treated with Amp alone also showed Amp-resistance mutations (Table S2).

Taken together, the genetic changes observed in this study indicate strong parallel adaptation of *s_lys* and *r_lys* populations to frequent prophage induction—via phage resistance mutations and prophage loss—and Amp selection—via c_Amp^R mutations.

Frequent prophage induction selects for rapid prophage loss characterized by distinct selective sweeps

As the loss of prophages occurred together with phage- and Amp^R mutations, we next inferred the temporal dynamics of these genetic changes within the second replicate population of both *r_lys* and *s_lys* subjected to MMC or Amp + MMC selection. We found that both *r_lys* and *s_lys* populations adapted to MMC through two selective sweeps: during the first sweep, lysogens acquired SRMs that prevented reinfections by free phages. Once SRM lysogens reached a frequency of ~75%, this genotype was replaced by fixation of a phage-free SRM genotype (Figure 3), resulting in the second selective sweep.

To better understand the selective processes underlying prophage loss under MMC selection, we implemented our empirical observations into a stochastic model of lysogen evolution (see STAR Methods and Methods S1). Our simulations demonstrate that these two selective sweeps can be observed if (1) lysogens carry a growth cost in the presence of free phages despite superinfection immunity and (2) MMC selection increases prophage costs due to frequent induction (Figure 3). The proposed growth cost of free phages on lysogens seems surprising but could be

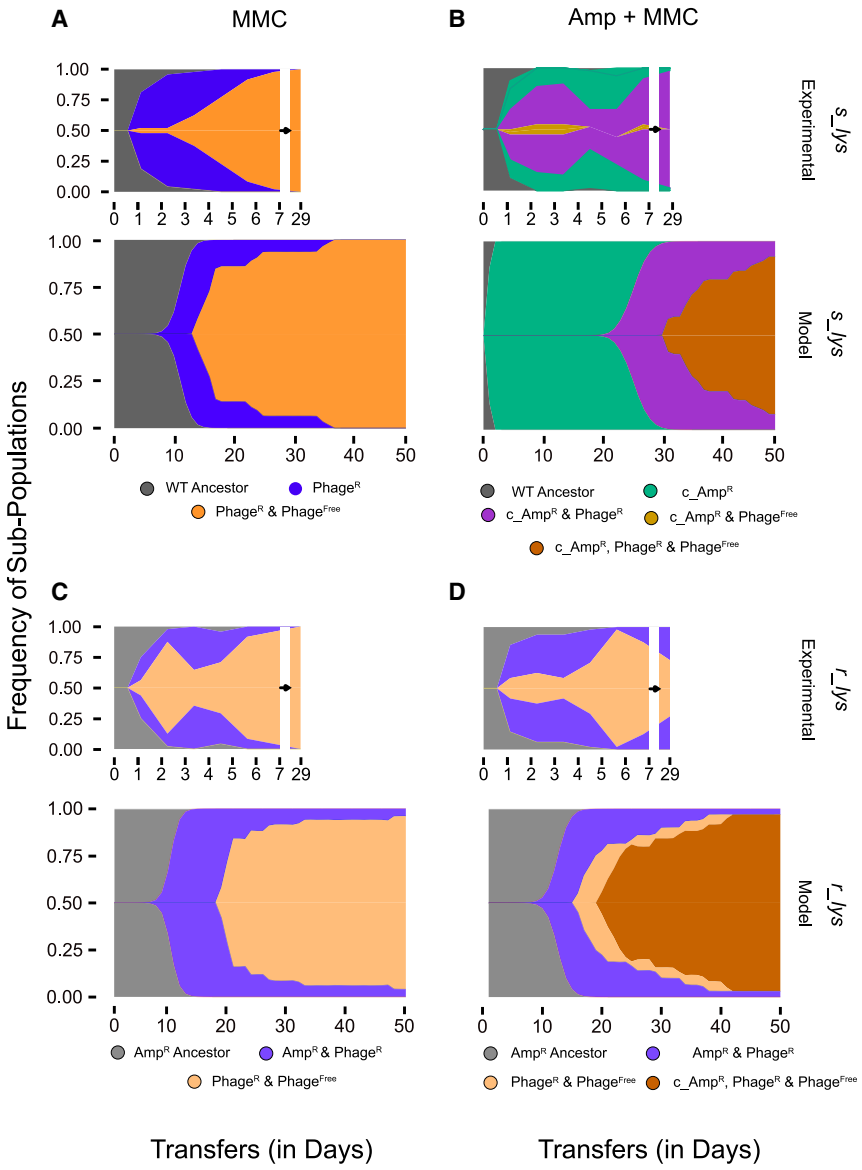

Figure 3. Muller plots showing sub-population frequencies over time

Muller plots showing sub-population frequencies over time of a single replicate population from the selection experiment and our model data averaged over 100 simulations for *s_lys* (A and B) and *r_lys* (C and D) populations exposed to MMC (A and C) or Amp + MMC (B and D). Colored areas are proportional to the frequencies of each sub-population in the population: ancestral WT (gray), phage-resistant lysogens (blue), c_Amp^R from chromosomal mutations (green), c_Amp^R/Amp^R with phage resistance mutations (purple), the c_Amp^R and phage free (gold), and lastly phage resistant and phage free with or without c_Amp^R (orange/brown). Plots of empirical data were generated by combining phenotypic and genotypic assays on sets of 24 individual clones per population for one replicate from each treatment and genotype combination. Plots of model data were generated from an averaging of 100 simulations of that particular treatment and genotype combination. For details of model simulations, parameters used, and a model sketch see also STAR Methods, Figures S3–S5, Table S3, and Methods S1.

confirmed experimentally (Figure S4A) and can be explained by repeated adsorption and disruption of the cell membrane.[37,38] The presence of free phages significantly reduces the growth rate of *s_lys* ($t = -26.7$, $p < 0.001$) and *r_lys* ($t = -8.45$, $p < 0.001$) compared with bacteria that have SRMs preventing attachment of phage Lambda+ to the bacterial cell.

Prophage-encoded benefits and conflicting selective forces prolong prophage maintenance

In contrast to selection in the absence of Amp, the sequenced *s_lys* population subjected to Amp + MMC exhibited distinct evolutionary dynamics. While different types of adaptive mutations were observed early on, the ones that swept through the population first conferred c_Amp^R (Figure 3), followed by the acquisition of SRMs. This new genotype, which is resistant to Amp plus the burden of phage absorption, rapidly spread through the population, reaching a frequency of nearly 100%. Surprisingly, although phage-free c_Amp^R variants did emerge, they did not increase in frequency until the end of the experiment.

Given the limited temporal resolution achievable with one replicate population and the apparent need for three mutational events, we turned to our stochastic model to investigate the evolutionary trajectory of *s_lys* populations under Amp + MMC selection over longer timescales. Running our simulations over 20 days longer than the experiment revealed a distinct cascade of three selective sweeps. As in the experiment, c_Amp^R swept rapidly through the *s_lys* populations, which suggests that in our experiment, the selection pressure imposed by Amp was stronger than the selection pressure imposed by MMC. In the simulations, c_Amp was followed by a sweep of phage resistance via SRM, eventually leading to the loss of prophages in c_Amp- and phage-resistant cells. However, the sweep of prophage-free cells only started after T30, with the mean frequency of prophage carriage decreasing from 99% on T30 to 20% on T50 (Figure 3), which explains why prophage carriage was still very high at the end of our experiment. Our model shows that, even though prophages are costly, they can be maintained longer when facing additional selective pressures. This pressure limits evolvability by necessitating consecutive selective sweeps and reducing the population size, thereby decreasing the mutational supply.

In contrast, *r_lys* populations exposed to Amp + MMC acquired c_Amp^R less frequently and instead acquired the same two mutation types as in the absence of Amp: emergence of phage resistance followed by prophage loss. However, in the presence of both selection pressures, phage-free SRMs did

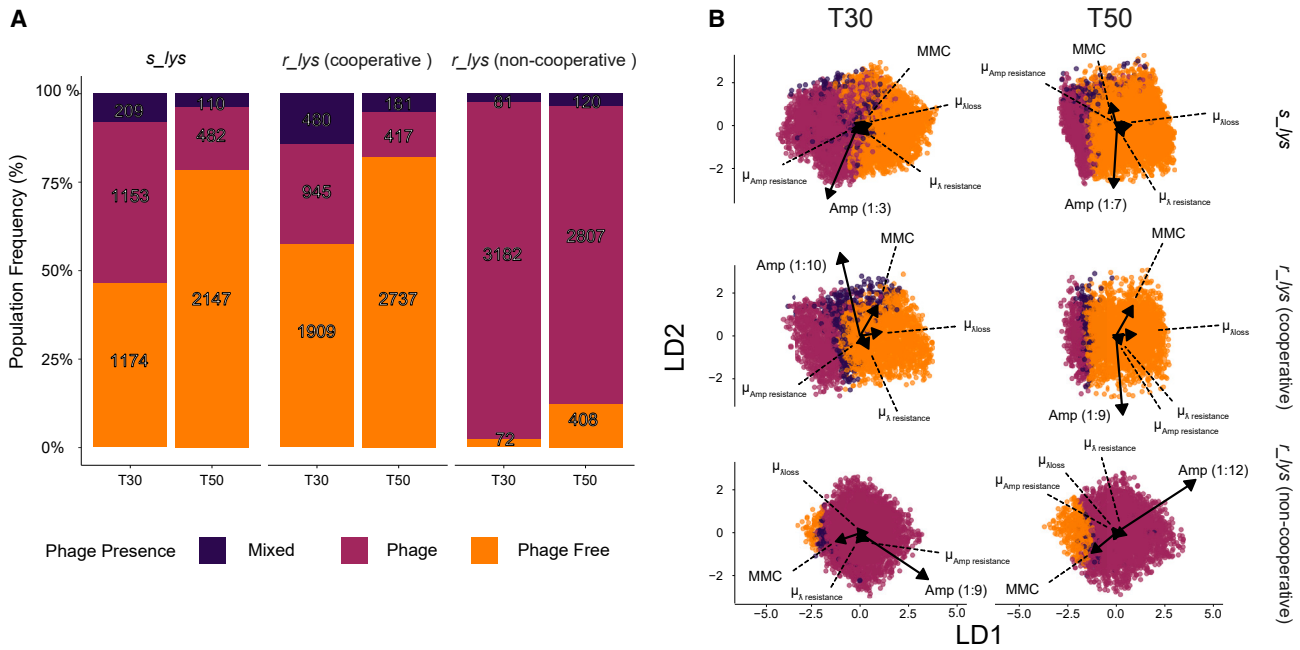

Figure 4. Parameter space exploration for prophage maintenance in the Amp + MMC treatment

Results from model simulations where we perform 100 simulation runs for each set of parameters with the Amp + MMC treatment for either *s_lys* or *r_lys* bacteria (list of parameters is in Table S4). We classified the simulation results based on whether the majority (>51%) of the population is phage free (orange), phage containing (amaranth purple), or the mixed with no majority (violet).

(A) Stacked bar plot of 3,335 sets of parameter values that varied the mutation rates and effect of both Amp and MMC within the simulation, but where parameter values were the same across the different simulation scenarios. We compare these same parameter sets for *s_lys* bacteria, *r_lys* bacteria that carry cooperative antibiotic resistance, and *r_lys* bacteria that carry non-cooperative antibiotic resistance.

(B) Linear discriminant analysis determining the impact of the three mutation rates as well as the two environmental effects—Amp, and MMC—on the frequency of prophage loss. LD1 and LD2 are the principal axes of the LDA. Arrows show the magnitude and direction of a parameter in separating the classes. The effect magnitude of Amp, for example, was generally much larger than all other factors (rescaling of Amp arrow to fit the plot is shown in brackets). For details of model simulations, parameters used, and a model sketch see also Figures S4 and S5, Tables S3 and S4, and Methods S1.

not reach fixation but instead co-existed with SRM *r_lys* lysogens at varying frequencies over time (Figure 3). We hypothesized that this co-existence could be explained by the fact that enzymatic resistance to ampicillin via beta lactamase is a cooperative trait that benefits both Amp-susceptible and Amp-resistant cells. That is because the beta lactamase enzyme secreted by prophage-carrying *r_lys* cells can break down the Amp in the culture, allowing phage-free SRM cheaters to arise and take advantage of their enzyme-producing kin. We confirmed the ability of *r_lys* lysogens to protect phage-free bacteria from an otherwise lethal concentration of Amp using an ampicillin removal assay (Figure S4B). By implementing this cooperative effect of Amp^R in our model, we found a lower selection pressure for chromosomal resistance, while it allowed for faster loss of prophages in *r_lys* compared with *s_lys* populations (Figure 3). This explains why phage-free cells cannot reach fixation until c_Amp^R is acquired. Otherwise, the protective effect of secreted beta lactamase would be lost. Taken together, our results suggest that the selective advantage of the prophage-encoded cooperative Amp^R allows the prophage to persist in environments that would otherwise select for prophage loss.

Simulating the evolution of *r_lys* lysogens encoding a non-cooperative resistance mechanism, such as a drug-efflux pump, predicts even longer maintenance of prophages in lysogen populations (Figure 4A). This can be explained by the non-cooperative benefit, which does not allow phage-free cheaters to emerge as they will not be protected from Amp. Thus, akin to *s_lys* populations, conflicting selection in environments that otherwise select against prophages hinders prophage loss as chromosomal resistance is required for competitive growth in these conditions.

Environmental factors trump mutation rates in determining prophage maintenance

The evolutionary dynamics of prophage loss showed environment-dependent distinct successions of selective sweeps. To determine how robust these succession patterns are with respect to the availability of each genetic change (i.e., the variation in mutation rates of phage and antibiotic resistance and prophage loss) and the strength of individual selection pressures (i.e., the effect sizes of MMC and Amp), we explored the sensitivity of the simulation results to parameter variation, using a linear discriminant analysis (LDA). Our results revealed that mutation rates had little impact on the outcome of our simulations. Instead, irrespective of the lysogen genotype, Amp selection had by far the strongest effect on the evolutionary trajectories, followed by MMC (Figure 4B).

Generally, we find that not only the magnitude but also the direction of the effect caused by the various parameters can change over time. For *s_lys* populations for example, the LDA showed

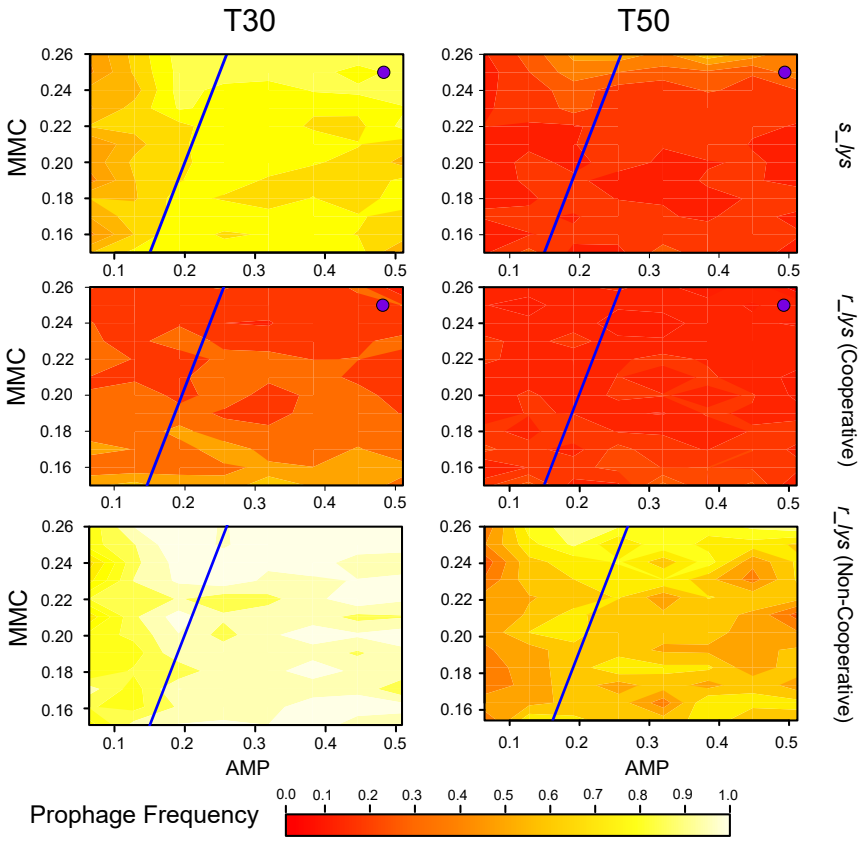

Figure 5. Contour plot demonstrating the final frequencies of prophages within our simulations

Contour plot showing the average frequency of prophage-carrying bacteria in simulated populations for different MMC and Amp effect values. The frequency was averaged over all the parameter sets that share the same Amp and MMC value (but they contain different mutation rates). Here, we use 10,000 parameter sets for the cooperative *r_lys* and *s_lys* simulations contour plots and 3,335 parameter sets for non-cooperative *r_lys* simulations. The effect of Amp and MMC pertains to the killing capacity of the drug (population killed/unit time) times the concentration of that drug (μg/mL). The blue solid line indicates where the selective pressure of MMC is equal to that of Amp. The purple dots are the MMC and Amp equivalent of the amount utilized in the selection experiment. See also Figures S4 and S5, Tables S3 and S4, and Methods S1.

that Amp had the greatest impact on prophage maintenance. However, while it selects weakly for the prophage at T30, it became neutral on T50 (Figure 4B). That is because Amp substantially decreased the population size, reducing mutational supply, and because selection for c_Amp^R was stronger than selection against the prophage. However, once c_Amp^R became more prevalent, population sizes increased again, and the selection pressure regarding prophage maintenance exerted by Amp became neutral. This temporal change in Amp selection effect can explain the differences in phage production that were seen between replicate populations of *s_lys* under Amp + MMC selection (Figure 1B).

For *r_lys* bacteria, the frequency of prophage loss was more strongly determined by the relative effect size of selection pressure for the phage-encoded benefit (Amp) and against the phage itself (MMC; Figures 4 and 5). *r_lys* populations with a cooperative resistance gene lost the phage more quickly and to a greater extent than *s_lys* populations, and this effect increased with increasing selection against prophages (Figures 4A and 5). However, when the resistance gene on the phage was non-cooperative, prophage loss in *r_lys* populations was significantly lower than in any of the other scenarios and most clearly determined by the cost-benefit trade-off between Amp and MMC (Figure 4). Specifically, selection by Amp outweighs MMC selection in non-cooperative *r_lys* populations across a range of sub-MIC values for both Amp and MMC (Figure 5).

In summary, our findings suggest that prophage maintenance is influenced more strongly by conflicting selection patterns resulting from environmental factors than mutational availability.

DISCUSSION

Active prophages are key players in the ecology and evolution of bacterial populations, with important consequences for higher order ecological interactions.[39] While the benefits of prophages to their hosts are well established, they also pose a constant risk of death to their host bacterium. Here, we investigated the role of environmental conditions in determining the maintenance and loss of prophages, considering the fitness benefits and costs associated with active prophages.

We found that prophage maintenance and loss is primarily determined by environmental conditions that alter the fitness benefits or costs of an active prophage. In environments that solely select against the prophage, lysogenic populations can lose the prophage rapidly. However, conflicting environmental selection pressures and prophage-encoded benefits can maintain prophages, even if prophages are costly. For instance, prophage-encoded resistance genes, which offer benefits exclusively to the lysogen that is carrying the prophage i.e., non-cooperative benefits, can substantially prolong prophage maintenance. That is because prophage loss in the absence of chromosomal drug resistance reduces bacterial fitness. In this case, prophage loss can be thought of as crossing a fitness valley, as chromosomal resistance mutations are required first, but they are either selectively neutral or costly[40–42] when the same benefit is already provided by the prophage. However, when the prophage-encoded traits are beneficial to other cells in the populations i.e., cooperative benefits, prophages are maintained, but only in a small part of the population. In this scenario, prophage frequencies are most likely contingent upon the level of protection offered by their gene products, which can protect phage-free cheaters that are likely to emerge if selection against prophages is high. However, even in the absence of prophage-encoded benefits, the presence of an additional, unrelated selection pressure can maintain

prophages by reducing population size and selecting more strongly for other beneficial mutations.

Our combined approach of genomic analysis coupled with mathematical modeling revealed that prophage loss occurs through environment-specific successions of selective sweeps. During the first sweep, lysogenic bacteria acquired SRMs providing protection against phage infections. This enabled the second sweep, during which phage-free SRM rapidly increased in frequency. The rapid spread of SRM in lysogenic populations may seem surprising, as the integrated phage itself already protects the bacterium from superinfection. However, it can be explained by a repeated physiological burden of phage adsorption to the cell surface.[38] Preventing phage adsorption via SRM might thus be a beneficial strategy in the presence of high phage titers[43,44] and allows for subsequent prophage loss, which ameliorates the cost of frequent induction.

Our simulations demonstrated that prophage maintenance within a population is predominantly determined by environmentally derived selection pressures. These environmental pressures can also influence the evolutionary pathways of the populations by reducing population sizes and thus mutational supply. Mutational availability, i.e., the rate of obtaining a specific mutation, on the other hand, was not a major determinant of the selective sweeps in our experiments (Figure 4B). The availability of different mutations could have been expected to play an important role as we observed two types of genetic changes, single-nucleotide mutations and large deletions. Our simulations show that these different mutation rates alone do not lead to significant changes in population structure (Figure 4B). By highlighting the impact of environmental factors relative to mutation rates in determining prophage loss trajectories, we suggest that adaptation of bacterial populations to environmental conditions is less constrained by the availability of a certain mutation but by the strength of selection for different traits.

The maintenance of active prophages is potentially analogous to the plasmid paradox.[45–47] Plasmids should be lost from a population due to their high maintenance costs but are frequently found in bacterial populations because they can carry beneficial genes, such as resistances to antibiotics or heavy metals.[45–47] Similar to plasmids, the net fitness effect of prophages is highly environment dependent, but their ability to kill and lyse the host bacterium imposes an additional conflict for their hosts. Despite this death risk, which can be very costly for a lysogen population, we show that active prophages can be maintained within bacterial populations if they carry beneficial genes that confer a fitness advantage to their host. For example, temperate bacteriophages can contribute to bacterial pathogenesis by helping the bacterium attach to cells, evade the immune system, or produce virulence proteins.[48,49–54] These genes are present on prophages spanning many genera including *Vibrio cholera*, *Salmonella typhimurium*, *E. coli*, and *Shigella*.[49,50,55–57] Such beneficial factors can be thought of as a cooperation between the prophage and the bacterium, which may be further strengthened by the ability to deploy free phages as weapons during bacterial warfare.[5,58]

Our study provides a comprehensive understanding of the intricate interplay between genetic and environmental factors that shape the dynamics of prophage maintenance and loss. By unraveling the complex selection pressures that act upon mobile genetic elements, including the importance of accessory genes encoded thereon, we shed light on the underlying mechanisms that drive the selective forces acting on prophages and their bacterial hosts. Considering the omnipresence of prophages in the bacterial world, our findings hold important implications for our understanding of bacterial evolution and diversification in changing environments. By deepening our understanding of the interplay between genetic and environmental factors, we can improve our ability to predict and manage the evolution of microbial populations, and the genetic elements within them. Such knowledge may contribute to the development of strategies for effectively harnessing the potential of microbial communities and their genetic diversity for various applications in fields such as biotechnology, ecology, and evolutionary medicine.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Study system and lysogen construction
  - Selection experiment
  - Bacterial and phage population dynamics
  - Whole genome sequencing
  - Quantifying candidate mutations over time
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis
  - Stochastic model of prophage maintenance across environments

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.03.025.

ACKNOWLEDGMENTS

We thank the Pathogen Ecology group at ETH Zurich for feedback, especially Ricardo Leon-Sampedro, Ayush Pathak, and Mathilde Boumasmoud. We thank Rahel Bruhlman for assistance in performing PCRs and Christopher Witzany and Daniel Angst for lending cluster computational time. This project was funded by a Swiss National Science Foundation grant (no. PZ00P3_179743) given to C.C.W. and was supported by an ETH Zurich Postdoctoral Fellowship (19-2 FEL-74) received by C.I.

AUTHOR CONTRIBUTIONS

C.C.W. and C.I. obtained funding. C.C.W. and Z.M.B. designed the experiment. Z.M.B. performed the experiment and analyzed the experimental and genomic data. Z.M.B. and C.I. developed the model. Z.M.B. ran the simulations. All authors wrote the manuscript and approved the final version of this manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 6, 2023
Revised: January 19, 2024
Accepted: March 14, 2024
Published: April 9, 2024

REFERENCES

1. Kaiser, A.D., and Jacob, F. (1957). Recombination between related temperate bacteriophages and the genetic control of immunity and prophage localization. Virology *4*, 509–521. https://doi.org/10.1016/0042-6822(57)90083-1.
2. Harrison, E., and Brockhurst, M.A. (2017). Ecological and evolutionary benefits of temperate phage: what does or doesn't kill you makes you stronger. BioEssays *39*, 1700112. https://doi.org/10.1002/bies.201700112.
3. Rankin, D.J., Rocha, E.P.C., and Brown, S.P. (2011). What traits are carried on mobile genetic elements, and why? Heredity *106*, 1–10. https://doi.org/10.1038/hdy.2010.24.
4. Davies, E.V., James, C.E., Kukavica-Ibrulj, I., Levesque, R.C., Brockhurst, M.A., and Winstanley, C. (2016). Temperate phages enhance pathogen fitness in chronic lung infection. ISME J. *10*, 2553–2555. https://doi.org/10.1038/ismej.2016.51.
5. Burns, N., James, C.E., and Harrison, E. (2015). Polylysogeny magnifies competitiveness of a bacterial pathogen *in vivo*. Evol. Appl. *8*, 346–351. https://doi.org/10.1111/eva.12243.
6. Wagner, P.L., and Waldor, M.K. (2005). Bacteriophages in Bacterial Pathogenesis. In The Bacteriophages: Second Edition, R. Calendar, and S.T. Abedon, eds. (Oxford University Press), pp. 710–724. https://academic.oup.com/book/52769.
7. Wendling, C.C., Refardt, D., and Hall, A.R. (2021). Fitness benefits to bacteria of carrying prophages and prophage-encoded antibiotic-resistance genes peak in different environments. Evolution *75*, 515–528. https://doi.org/10.1111/evo.14153.
8. McEntee, K. (1977). Protein X is the product of the recA gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA *74*, 5275–5279. https://doi.org/10.1073/pnas.74.12.5275.
9. Roberts, J.W., and Roberts, C.W. (1975). Proteolytic cleavage of bacteriophage lambda repressor in induction. Proc. Natl. Acad. Sci. USA *72*, 147–151. https://doi.org/10.1073/pnas.72.1.147.
10. Roberts, J.W., Roberts, C.W., and Mount, D.W. (1977). Inactivation and proteolytic cleavage of phage lambda repressor in vitro in an ATP-dependent reaction. Proc. Natl. Acad. Sci. USA *74*, 2283–2287. https://doi.org/10.1073/pnas.74.6.2283.
11. Hendrix, R.W. (1983). Lambda II (Cold Spring Harbor Laboratory Press). https://cshmonographs.org/index.php/monographs/issue/view/087969150.13.
12. Khan, A., and Wahl, L.M. (2020). Quantifying the forces that maintain prophages in bacterial genomes. Theor. Popul. Biol. *133*, 168–179. https://doi.org/10.1016/j.tpb.2019.11.003.
13. Bobay, L.M., Touchon, M., and Rocha, E.P.C. (2014). Pervasive domestication of defective prophages by bacteria. Proc. Natl. Acad. Sci. USA *111*, 12127–12132. https://doi.org/10.1073/pnas.1405336111.
14. Touchon, M., Bobay, L.M., and Rocha, E.P. (2014). The chromosomal accommodation and domestication of mobile genetic elements. Curr. Opin. Microbiol. *22*, 22–29. https://doi.org/10.1016/j.mib.2014.09.010.
15. Frazão, N., Konrad, A., Amicone, M., Seixas, E., Güleresi, D., Lässig, M., and Gordo, I. (2022). Two modes of evolution shape bacterial strain diversity in the mammalian gut for thousands of generations. Nat. Commun. *13*, 5604. https://doi.org/10.1038/s41467-022-33412-8.
16. Bobay, L.M., Rocha, E.P., and Touchon, M. (2013). The adaptation of temperate bacteriophages to their host genomes. Mol. Biol. Evol. *30*, 737–751. https://doi.org/10.1093/molbev/mss279.
17. Lederberg, E.M., and Lederberg, J. (1953). Genetic studies of lysogenicity in *Escherichia coli*. Genetics *38*, 51–64. https://doi.org/10.1093/genetics/38.1.51.
18. Khan, A., Burmeister, A.R., and Wahl, L.M. (2020). Evolution along the parasitism-mutualism continuum determines the genetic repertoire of prophages. PLoS Comput. Biol. *16*, e1008482. https://doi.org/10.1371/journal.pcbi.1008482.
19. Ramisetty, B.C.M., and Sudhakari, P.A. (2019). Bacterial 'grounded' prophages: hotspots for genetic renovation and innovation. Front. Genet. *10*, 65. https://doi.org/10.3389/fgene.2019.00065.
20. Kang, H.S., McNair, K., Cuevas, D.A., Bailey, B.A., Segall, A.M., and Edwards, R.A. (2017 Mar). Prophage genomics reveals patterns in phage genome organization and replication. Preprint at bioRxiv. https://doi.org/10.1101/114819.
21. Chatterjee, A., and Duerkop, B.A. (2019). Sugar and fatty acids Ack-celerate prophage induction. Cell Host Microbe *25*, 175–176. https://doi.org/10.1016/j.chom.2019.01.012.
22. Little, J.W., and Michalowski, C.B. (2010). Stability and instability in the lysogenic state of phage lambda. J. Bacteriol. *192*, 6064–6076. https://doi.org/10.1128/JB.00726-10.
23. Boos, W., and Böhm, A. (2000). Learning new tricks from an old dog: *MalT* of the *Escherichia coli* maltose system is part of a complex regulatory network. Trends Genet. *16*, 404–409. https://doi.org/10.1016/s0168-9525(00)02086-2.
24. Andrews, B., and Fields, S. (2020). Distinct patterns of mutational sensitivity for λ resistance and maltodextrin transport in *Escherichia coli* LamB. Microb. Genomics *6*, e000364. https://doi.org/10.1099/mgen.0.000364.
25. Thirion, J.P., and Hofnung, M. (1972). On some genetic aspects of phage λ resistance in *E. coli* K12. Genetics *71*, 207–216. https://doi.org/10.1093/genetics/71.2.207.
26. Burmeister, A.R., Sullivan, R.M., and Lenski, R.E. (2020). Fitness costs and benefits of resistance to phage lambda in experimentally evolved Escherichia coli. In Evolution in Action: Past, Present and Future: A Festschrift in Honor of Erik D Goodman, W. Banzhaf, B.H.C. Cheng, K. Deb, K.E. Holekamp, R.E. Lenski, and C. Ofria, et al., eds. (Springer International Publishing), pp. 123–143.
27. Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., and Luisi, B.F. (2014). Structure of the *AcrAB-TolC* multidrug efflux pump. Nature *509*, 512–515. https://doi.org/10.1038/nature13205.
28. Kobylka, J., Kuth, M.S., Müller, R.T., Geertsma, E.R., and Pos, K.M. (2020). *AcrB*: a mean, keen, drug efflux machine. Ann. N. Y. Acad. Sci. *1459*, 38–68. https://doi.org/10.1111/nyas.14239.
29. Schuster, S., Vavra, M., Greim, L., and Kern, W.V. (2021). Exploring the contribution of the AcrB homolog MdtF to drug resistance and dye efflux in a multidrug resistant *E. coli* isolate. Antibiotics (Basel) *10*, 503. https://doi.org/10.3390/antibiotics10050503.
30. Sulavik, M.C., Gambino, L.F., and Miller, P.F. (1995). The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. Mol. Med. *1*, 436–446. https://doi.org/10.1007/BF03401581.
31. Raynor, B.D. (1997). Penicillin and ampicillin. Prim. Care Update Ob Gyns *4*, 147–152. https://doi.org/10.1016/S1068-607X(97)00012-7.
32. Zawadzke, L.E., Bugg, T.D., and Walsh, C.T. (1991). Existence of two D-alanine:D-alanine ligases in Escherichia coli: cloning and sequencing of the ddlA gene and purification and characterization of the DdlA and DdlB enzymes. Biochemistry *30*, 1673–1682. https://doi.org/10.1021/bi00220a033.
33. Spratt, B.G. (1975). Distinct penicillin binding proteins involved in the division, elongation, and shape of Escherichia coli K12. Proc. Natl. Acad. Sci. USA *72*, 2999–3003. https://doi.org/10.1073/pnas.72.8.2999.

34. Curtis, N.A., Eisenstadt, R.L., Turner, K.A., and White, A.J. (1985). Inhibition of penicillin-binding protein 3 of *Escherichia coli* K-12. Effects upon growth, viability and outer membrane barrier function. J. Antimicrob. Chemother. *16*, 287–296. https://doi.org/10.1093/jac/16.3.287.

35. Park, S.H., Kim, Y.J., Lee, H.B., Seok, Y.J., and Lee, C.R. (2020). Genetic evidence for distinct functions of peptidoglycan endopeptidases in Escherichia coli. Front. Microbiol. *11*, 565767. https://doi.org/10.3389/fmicb.2020.565767.

36. Itoh, Y., Rice, J.D., Goller, C., Pannuri, A., Taylor, J., Meisner, J., Beveridge, T.J., Preston, J.F., and Romeo, T. (2008). Roles of pgaABCD genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-β-1,6-N-acetyl-d-glucosamine. J. Bacteriol. *190*, 3670–3680. https://doi.org/10.1128/JB.01920-07.

37. Berryhill, B.A., Garcia, R., McCall, I.C., Manuel, J.A., Chaudhry, W., Petit, M.A., and Levin, B.R. (2023). The book of Lambda does not tell us that naturally occurring lysogens of *Escherichia coli* are likely to be resistant as well as immune. Proc. Natl. Acad. Sci. USA *120*, e2212121120. https://doi.org/10.1073/pnas.2212121120.

38. Abedon, S.T. (2011). Lysis from without. Bacteriophage *1*, 46–49. https://doi.org/10.4161/bact.1.1.13980.

39. Wendling, C.C. (2023). Prophage mediated control of higher order interactions – insights from systems approaches. Preprint at EcoEvoRxiv. https://doi.org/10.32942/X2ZK52.

40. Schrag, S.J., Perrot, V., and Levin, B.R. (1997). Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. Proc. Biol. Sci. *264*, 1287–1291. https://doi.org/10.1098/rspb.1997.0178.

41. Hernando-Amado, S., Sanz-García, F., Blanco, P., and Martínez, J.L. (2017). Fitness costs associated with the acquisition of antibiotic resistance. Essays Biochem. *61*, 37–48. https://doi.org/10.1042/EBC20160057.

42. Lenski, R.E. (1998). Bacterial evolution and the cost of antibiotic resistance. Int. Microbiol. *1*, 265–270.

43. Taslem Mourosi, J., Awe, A., Guo, W., Batra, H., Ganesh, H., Wu, X., and Zhu, J. (2022). Understanding bacteriophage tail fiber interaction with host surface receptor: the key "blueprint" for reprogramming phage host range. Int. J. Mol. Sci. *23*, 12146. https://doi.org/10.3390/ijms232012146.

44. Garen, A., and Puck, T.T. (1951). The first two steps of the invasion of host cells by bacterial viruses. II. J. Exp. Med. *94*, 177–189. https://doi.org/10.1084/jem.94.3.177.

45. Millan, A.S., Peña-Miller, R., Toll-Riera, M., Halbert, Z.V., McLean, A.R., Cooper, B.S., and MacLean, R.C. (2014). Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. Nat. Commun. *5*, 5208. https://doi.org/10.1038/ncomms6208.

46. San Millan, A., and MacLean, R.C. (2017). Fitness costs of plasmids: a limit to plasmid transmission. Microbiol. Spectr. *5*, 1–12.

47. Simonsen, L. (1991). The existence conditions for bacterial plasmids: theory and reality. Microb. Ecol. *22*, 187–205. https://doi.org/10.1007/BF02540223.

48. Campbell, A. (2005). General Aspects of Lysogeny. In The Bacteriophages: Second Edition, R. Calendar, and S.T. Abedon, eds. (Oxford University Press), pp. 66–73. https://academic.oup.com/book/52769.

49. Figueroa-Bossi, N., and Bossi, L. (1999). Inducible prophages contribute to Salmonella virulence in mice. Mol. Microbiol. *33*, 167–176. https://doi.org/10.1046/j.1365-2958.1999.01461.x.

50. Lemire, S., Figueroa-Bossi, N., and Bossi, L. (2008). Prophage contribution to salmonella virulence and diversity. In Horizontal Gene Transfer in the Evolution of Pathogenesis, H. Schmidt, and M. Hensel, eds. (Cambridge University Press), pp. 159–192.

51. Vaca, S., Arce, J., Oliver, G., Arenas, D., and Arguello, F. (1989). FIZ15 bacteriophage increases the adhesion of Pseudomonas aeruginosa to human buccal epithelial cells. Rev. Latinoam. Microbiol. *31*, 1–5.

52. Stanley, T.L., Ellermeier, C.D., and Slauch, J.M. (2000). Tissue-specific gene expression identifies a gene in the lysogenic phage Gifsy-1 that affects Salmonella enterica serovar Typhimurium survival in Peyer's patches. J. Bacteriol. *182*, 4406–4413. https://doi.org/10.1128/JB.182.16.4406-4413.2000.

53. Robbins, P.W., and Uchida, T. (1965). Chemical and Macromolecular Structure of O-Antigens from *Salmonella anatum* Strains Carrying Mutants of bacteriophage ε15. J. Biol. Chem. *240*, 375–383. https://doi.org/10.1016/S0021-9258(18)97659-3.

54. Robbins, P.W., and Uchida, T. (1962). Studies on the chemical basis of the phage conversion of O-antigens in the E-group Salmonellae. Biochemistry *1*, 323–335. https://doi.org/10.1021/bi00908a020.

55. Krüger, A., and Lucchesi, P.M.A. (2015). Shiga toxins and *stx* phages: highly diverse entities. Microbiology (Reading) *161*, 451–462. https://doi.org/10.1099/mic.0.000003.

56. Davis, B.M., Moyer, K.E., Boyd, E.F., and Waldor, M.K. (2000). CTX prophages in classical biotype Vibrio cholerae: functional phage genes but dysfunctional phage genomes. J. Bacteriol. *182*, 6992–6998. https://doi.org/10.1128/JB.182.24.6992-6998.2000.

57. Waldor, M.K., and Mekalanos, J.J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. Science *272*, 1910–1914. https://doi.org/10.1126/science.272.5270.1910.

58. Brown, S.P., Le Chat, L., De Paepe, M., and Taddei, F. (2006). Ecology of microbial invasions: amplification allows virus carriers to invade more rapidly when rare. Curr. Biol. *16*, 2048–2052. https://doi.org/10.1016/j.cub.2006.08.089.

59. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008. https://doi.org/10.1038/msb4100050.

60. Powell, B.S., Rivas, M.P., Court, D.L., Nakamura, Y., and Turnbough, C.L. (1994). Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic Acids Res. *22*, 5765–5766. https://doi.org/10.1093/nar/22.25.5765.

61. Deatherage, D.E., and Barrick, J.E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. In Engineering and Analyzing Multicellular Systems: Methods and Protocols, L. Sun, and W. Shou, eds. (Springer), pp. 165–188.

62. Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R.L., Rebelo, A.R., Florensa, A.F., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. J. Antimicrob. Chemother. *75*, 3491–3500. https://doi.org/10.1093/jac/dkaa345.

63. Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., Wingett, S., and Fast, Q.C. (2012). FastQC. Babraham, UK. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

64. Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., and Wingett, S. (2012). Trim galore. Babraham, UK. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

65. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. Current Protocols in Bionformatics *70*, e102. https://doi.org/10.1002/cpbi.102.

66. Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., and Wishart, D.S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. *44*, W16–W21. https://doi.org/10.1093/nar/gkw387.

67. Otsuji, N., Sekiguchi, M., Iijima, T., and Takagi, Y. (1959). Induction of phage formation in the lysogenic *Escherichia coli* K-12 by Mitomycin C. Nature *184*, 1079–1080. https://doi.org/10.1038/1841079b0.

68. Levine, M. (1961). Effect of Mitomycin C on interactions between temperate phages and bacteria. Virology *13*, 493–499. https://doi.org/10.1016/0042-6822(61)90280-x.

69. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

70. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.

71. Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., and Wishart, D.S. (2011). PHAST: A fast phage search tool. Nucleic Acids Res. *39*, W347–W352. https://doi.org/10.1093/nar/gkr485.

72. Kleinheinz, K.A., Joensen, K.G., and Larsen, M.V. (2014). Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and E. coli virulence genes in bacteriophage and prophage nucleotide sequences. Bacteriophage *4*, e27943. https://doi.org/10.4161/bact.27943.

73. Hofnung, M., Jezierska, A., and Braun-Breton, C. (1976). LamB mutations in *E. coli* K12: growth of λ host range mutants and effect of nonsense suppressors. Mol. Gen. Genet. *145*, 207–213. https://doi.org/10.1007/BF00269595.

74. Lwoff, A. (1953). Lysogeny. Bacteriol. Rev. *17*, 269–337. https://doi.org/10.1128/br.17.4.269-337.1953.

75. Regoes, R.R., Wiuff, C., Zappala, R.M., Garner, K.N., Baquero, F., and Levin, B.R. (2004). Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens. Antimicrob. Agents Chemother. *48*, 3670–3676. https://doi.org/10.1128/AAC.48.10.3670-3676.2004.

76. Hill, A. (1910). The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. J. Physiol. *40*. i–vii.

77. Fox, J., and Weisberg, S. (2010). Time-series regression and generalized least squares in R. https://socialsciences.mcmaster.ca/jfox/Books/Companion-2E/appendix/Appendix-Timeseries-Regression.pdf.

78. Cao, Y., Gillespie, D.T., and Petzold, L.R. (2007). Adaptive explicit-implicit tau-leaping method with automatic tau selection. J. Chem. Phys. *126*, 224101. https://doi.org/10.1063/1.2745299.

79. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.

80. Drake, J.W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA *88*, 7160–7164. https://doi.org/10.1073/pnas.88.16.7160.

81. Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998). Rates of spontaneous mutation. Genetics *148*, 1667–1686. https://doi.org/10.1093/genetics/148.4.1667.

82. Fisher, R.A. (1938). The statistical utilization of multiple measurements. Ann. Eugen. *8*, 376–386. https://doi.org/10.1111/j.1469-1809.1938.tb02189.x.

83. Carnell, R. (2022). lhs: Latin Hypercube Samples. https://CRAN.R-project.org/package=lhs.

84. Venables, W.N., and Ripley, B.D. (2002). Modern Applied Statistics with S, Fourth Edition (Springer). https://www.stats.ox.ac.uk/pub/MASS4/.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli K12 MG1655 ΔgalK::yfp-bla, constitutive YFP expression	Fokko Smakman	N/A
Escherichia coli K12 BW25113 ΔmalT	Baba et al.[59]	CGSC#10518
Escherichia coli K12 BW25113 ΔlamB	Baba et al.[59]	CGSC#10877
Escherichia coli phage lambda	Wendling et al.[7]	ENA: ERS13655752
Escherichia coli phage lambda Δbor::cm	Wendling et al.[7]	N/A
Chemicals, peptides, and recombinant proteins		
Ampicillin Sodium Salt	Sigma-Aldrich	Cat#A9518-100G
Mitomycin C from *Streptomyces caespitosus*	Sigma-Aldrich	Cat#M4287-2MG
LB Broth (Lennox)	Sigma-Aldrich	Cat#L3022-1KG
LB Agar (Lennox)	Sigma-Aldrich	Cat#L2897-1KG
Maltose monohydrate	Merck	Cat#1.05910.0500
Phosphate buffered saline	Sigma-Aldrich	Cat#P4417-100TAB
Glycerol	Sigma-Aldrich	Cat#G5516-500ML
Tetrazolium Tetrachloride	Sigma-Aldrich	Cat#T8877-10G
Critical commercial assays		
GenElute™ Bacterial Genomic DNA Kits	Sigma-Aldrich	Cat#NA2110-1KT
DNeasy 96 Blood & Tissue Kit	Qiagen	Cat#69581
Deposited data		
Raw data	This paper	https://doi.org/10.5061/dryad.v15dv420
Oligonucleotides		
Escherichia coli attB Forward: GAGGTACCAGCGCGGTTT GATC	Powell et al.[60]	N/A
Escherichia coli phage Lambda *int* Reverse: ACTCGTCGCGAACCGC TTTC	Powell et al.[60]	N/A
Software and algorithms		
R (v 4.3.1)	R Core Team, 2021	https://cran.r-project.org/
Rstudio	RStudio Team, 2020	https://posit.co/download/rstudio-desktop/
Breseq	Deatherage and Barrick[61]	https://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing
ResFinder	Bortolaia et al.[62]	https://cge.food.dtu.dk/services/ResFinder/
Trim Galore	Andrews et al.[63]	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
FastQC	Andrews et al.[64]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
SPAdes	Prjibelski et al.[65]	https://github.com/ablab/spades
PHASTER	Arndt et al.[66]	https://phaster.ca/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Zachary M. Bailey (zachary.bailey@env.ethz.ch).

Materials availability

All constructed lysogens and bacterial strains used are available upon request.

Data and code availability

- All genome sequences from this study are deposited to ENA; the accession numbers are present in the key resources table. All data files are deposited on Dryad at https://doi.org/10.5061/dryad.v15dv4206.
- This paper does not report original code
- Any additional information needed to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The strains of *Escherichia coli* that we used for our experiments are listed in the key resources table. The phage Lambda+ that we used is also listed in the key resources table. The construction of lysogens with phage Lambda+ is described in method details. *E. coli* was cultured by streaking frozen stocks onto LB agar plates and incubated overnight at 37°C. Individual colonies were picked from this and placed in 4ml of LB broth incubating for ~16 hours at 37° C and 180 RPM. *E. coli* was regularly grown on LB agar plates or with LB broth. Isolates were used to make frozen stocks with 25% glycerol at -80°C.

We used *Escherichia coli* K-12 MG1655, which is cured of phage lambda and contains a constitutively expressed YFP marker to construct two different lysogens. One lysogen contained Lambda+ which encodes a constitutively expressed Ampicillin (Amp) resistance gene (BlaTEM116).[7] The other lysogen contained the same prophage yet lacked the resistance gene. Lysogenization was confirmed by PCR using primers that target the attB site in *E. coli* and the *int* gene of Lambda+.[60] In our work, we refer to lambda free bacteria as non-lysogens (*n_lys*), to lysogenized bacteria without the resistance gene as susceptible lysogens (*s_lys*), and to lysogenized bacteria with the Amp-resistance gene (Amp^R) as resistant lysogens, (*r_lys*).

METHOD DETAILS

Study system and lysogen construction

We used *Escherichia coli* K-12 MG1655, which is cured of phage Lambda and contains a constitutively expressed YFP marker to construct two different lysogens. One lysogen contained Lambda+ which encodes a constitutively expressed Ampicillin (Amp) resistance gene (BlaTEM116).[7] The other lysogen contained the same prophage yet lacked the resistance gene. Naïve *E. coli* was infected with Lambda+ phage at a modality of infection (MOI) >1. Lysogenization was confirmed by PCR using primers that target the attB site in *E. coli* and the *int* gene of Lambda+.[60] Hereafter, we refer to lambda free bacteria as non-lysogens (*n_lys*), to lysogenized bacteria without the resistance gene as susceptible lysogens (*s_lys*), and to lysogenized bacteria with the Amp-resistance gene (Amp^R) as resistant lysogens, (*r_lys*).

Selection experiment

Experimental design

We initiated a selection experiment to observe prophage maintenance over time in four different selection regimes using three bacterial genotypes, i.e., *n_lys, s_lys* and *r_lys* in a full-factorial design. With six biological replicates, each originating from single colonies, our design resulted in a total of 72 populations. We used the following selection regimes: (i) Amp at IC_{50} (i.e., 6.4 μg/ml) selecting for the prophage encoded Amp^R present in the *r_lys* population, (ii) Mitomycin C (MMC) at a concentration of 0.1 ug/ml which induces the lytic cycle of prophages causing massive cell death due to host lysis[67,68] (Figure S1) and increases the costs of prophage carriage for the *s_lys* and *r_lys* populations, (iii) a combination of Amp and MMC at the same concentrations as above, to create opposing selective pressures on the *r_lys* populations and two competing selective pressures on *s_lys* populations, and (iv) LB broth as control.

Experimental setup

We used 10 μl of an overnight culture to inoculate each treatment. All 72 populations were grown in 96 deep-well plates at a total volume of 1 ml. Every day, we transferred 1% of each population to fresh medium for a total of 30 transfers. After the first 24 hours, and subsequently every 7 transfers (T), we quantified bacterial and phage population dynamics. Additionally, we froze 100 μl of each population every transfer in 25% glycerol at -80° C. Concurrently, every seven transfers, we made glycerol cryo stocks of 24 individual clones from each population which we used for follow-up analyses. At the end of the experiment, we detected contamination in two *s_lys* populations and excluded them from later analysis.

Bacterial and phage population dynamics

Bacterial dynamics

We measured bacterial cell counts as cells/ml using a Novocyte Novosampler Pro using 1:1000 dilutions of our cell cultures, following a previously established protocol in Wendling et al.[7]

Phage dynamics

We quantified the number of free phages as plaque forming units (PFU/ml) from each population using a plaque spot assay on a lawn of phage susceptible ancestral *E. coli* K-12 MG1655 as described in Wendling et al.[7] In populations exposed to MMC, we observed a decrease in free phage particles over time (Figure 1). This suggested either (i) a gradual reduction at the population level, resulting from lower burst size and phage productivity, or (ii) the presence of two different phenotypes, of which one stopped producing phages completely while the other produced phages at a rate comparable to the ancestral lysogen. To infer the underlying mechanism, we quantified phage production as PFU/ml, as described above, on ancestral clones and transfers 8, and 29 of 24 randomly selected clones from two *s_lys* and *r_lys* populations originating from either the MMC, or the Amp+MMC treatment. This assay confirmed that the observed decline in free phage particles observed at the population level results from individual clones that did not produce free phages. This suggested that these clones either lost or inactivated their prophages.

Whole genome sequencing

To infer the underlying molecular mechanism(s) behind the observed loss in phage particle production of individual clones that evolved in the presence of MMC, we sequenced both single clones and whole populations from the selection experiment.

Sequencing single clones

We sequenced single clones from the *s_lys* and *r_lys* populations that originated either from (i) LB (ii) the MMC, or the (iii) Amp+MMC treatment. We randomly selected two clones from two replicate populations from T8 and T29 as well as one ancestral clone per genotype.

Sequencing whole populations

We sequenced whole population samples from all 72 populations from the end of the experiment. Additionally, we included two populations from the Amp+MMC treatment from intermediate time-points, i.e., T8, T15, and T22 to track the presence of candidate mutations across time.

We extracted whole genome DNA using the GenElute Bacterial Genomic DNA extraction kit from Sigma-Aldrich. Library preparation and whole genome sequencing was done at the Functional Genomics Center Zurich (FGCZ) using a NovaSeq Illumina sequencer giving us FASTQ files with an average coverage of over 100 ×.

Sequence analysis of individual clones

We performed a quality check and trimmed Illumina adapter sequences using Trim Galore with FastQC on the trimmed reads.[63,64,69] We analyzed these trimmed and filtered reads using breseq (Bacterial Genome Resequencing Tool) (v 0.35.5) using a high-quality genome of the lysogenic *E. coli* K-12 MG1655 as reference (ERA18570755).[61] Next, we filtered the mutation predictions from breseq using the R statistical language[70] as follows: first, we removed mutations resulting from differences between our ancestor and the reference strain. Second, we filtered mutations so that only those present in populations that were treated with Amp, MMC or Amp + MMC remained. In addition, we assembled our reads into contigs using the de novo genome assembly algorithm of SPAdes. We used our genome assemblies with PHASTER to verify the presence/absence of phage Lambda+.[66,71] Lastly, we used ResFinder (v 4.0) to search for antibiotic resistance genes in each sample.[62,72]

Sequence analysis of whole populations

Quality check and trimming was done as described for single clones. We examined the presence of mutations using breseq with the same reference as above (ERA18570755) using the polymorphism option.[61] We exported breseq results into a table and extracted mutations related to phage resistance, chromosomal Amp-resistance (c_Amp^R) and the frequency of Lambda+ carriage using the same method as for individual clones (Table S3).

Quantifying candidate mutations over time

Whole genome sequencing revealed three major types of genetic changes. This includes presence/absence of Lambda+ as well as mutations associated with phage and Amp resistance. Next, we analyzed the frequency and dynamics of these mutations over time. We used the same set of clones, i.e., 24 randomly selected clones from transfers 1-7, and 29 from two *s_lys* and *r_lys* populations that were either exposed to MMC or Amp+MMC, across analyses to generate a clear timeline of when phage loss as well as phage and antibiotic resistances emerged and spread in each population. We refer to this set of clones from here on as "candidate clones".

Phage loss

Whole genome sequencing revealed that clones that no longer produced free phage particles lost Lambda+. To infer the dynamics of this prophage loss we used clonal PCRs on our candidate clones with primers for Lambda+ as described above.

Genomic evidence of phage resistance

First, we inferred the molecular mechanisms underlying phage resistance and their temporal dynamics. Whole genome sequencing revealed two different phage resistance mechanisms: (i) super-infection exclusion (SIE), where the presence of Lambda+ protects the host cell from additional infections, or (ii) surface receptor mutations (SRM) where mutations in *MalT*, a protein involved in cell surface receptor production, prevent Lambda+ from binding to the host cell.[24–26,73,74] To infer the nature of these two phage resistance mechanisms over time, we quantified the presence of Lambda+ and SRMs on our candidate clones.

Surface receptor mutations and phage resistance

Whole genome sequencing of single clones and whole populations revealed SRMs in *MalT*. Mutations in *MalT* come with a pleiotropic effect, i.e., a deficiency in the uptake of maltose. This deficiency can be used to phenotypically infer phage resistance of individual clones using TTC agar plates[26] on which we tested our candidate clones. As positive and negative controls, we used the Δ*MalT* strain from the KEIO collection and each respective ancestral clone.[59]

We combined the two datasets, prophage and SRM presence, to construct the temporal dynamics of prophage loss and phage resistance and the molecular mechanism behind it, i.e., SIE or SRM.

Surface receptor mutations versus superinfection exclusion

Using our genomic and phenotypic data, we selected 24 clones of *s_lys* and *r_lys* from the ancestral population and another 24 that developed SRM. We also selected 24 clones of *n_lys* bacteria as a control. We tested these clones for their resistance to phage infection by performing an RBG (Reduction of Bacterial Growth) assay as described in Wendling et al.[7]. Briefly, we used overnight cultures of each clone and inoculated 2 sets of 96 well LB plates with $\sim 1 \times 10^8$ CFUs/ml; we subsequently added $\sim 8 \times 10^8$ PFUs/ml of phage Lambda+ to each well of one of those sets of 96 well plates. After 24 hours, we plate the same amount of bacteria in the phage containing and phage-free plates. The RBG values are calculated from the difference in growth between these two types of plates.

Ampicillin resistance

Whole genome sequencing revealed two trends regarding phage encoded ampicillin resistance (Amp^R). (1) Under MMC selection, *r_lys* clones that lost their prophage also lost the phage encoded Amp^R and likely became susceptible to Amp (2) under Amp selection, populations previously susceptible to Amp acquired non-synonymous mutations in genes commonly associated with resistance to beta-lactam antibiotics.[27–31] Therefore, we tested the phenotypic resistance to Amp of these two populations.

r_lys populations. To test whether *r_lys* populations that lost Lambda+ and its encoded Amp^R, we performed a Minimum Inhibitory Concentration (MIC) assay using six clones from every evolved *r_lys* population subjected to all four selection regimes. We used Amp concentrations of 2, 4, 8, 16, 32, 64, 128, and 256 μg/ml to estimate the MIC and measured their OD_{600} at T0 and after 24 hours. We then used the difference between the two OD measurements to infer Amp resistance. A difference greater than 0.10 between the final and initial bacterial density indicated appreciable growth and thus resistance to Amp. To detect changes in the IC_{50}, the concentration needed to inhibit 50% of the population, we fitted these growth data to a Hill function.[75,76]

s_lys and n_lys populations. We tested s_lys and n_lys populations subjected to Amp for the presence of chromosomal Amp-resistance (c_Amp^R). We analyzed the Amp MIC of six clones that were used for PCRs and TTC assays using the same protocol as above for *r_lys* bacteria. To get a complete resistance profile of each population, we additionally tested the resistance of all 24 clones at the IC_{90} and IC_{99} concentrations of Amp. We grew each clone overnight at 37° C and re-inoculated them into 96-well plates containing one of two different concentrations of Amp corresponding to the IC_{90} and IC_{99} of the ancestral WT and measured OD_{600} at T0 and after 24 hours.

Ampicillin removal

In the Amp+MMC treatment, we noticed the coexistence between phage-carrying and phage-free clones. This suggested that phages carrying the BLA-TEM resistance gene were degrading Amp allowing clones that were both phage free and antibiotic susceptible to grow even when they lacked chromosomal mutations conferring Amp resistance. To confirm this, we tested the ability of lysogens carrying Amp^R to degrade Amp in liquid culture by performing an antibiotic removal test. To do so, we separately inoculated ancestral *n_lys*, *s_lys*, and *r_lys* bacterial populations containing the maximum concentration of Amp that *r_lys* bacteria can tolerate, i.e., 256 μg/ml and incubated them at 180 rpm overnight at 37° C. Cultures were then centrifuged at 5000 RPM for 10 minutes and the supernatant of each population was removed, and filter sterilized with a pore size of 0.2 μm. We plated the filtered supernatants on LB agar plates to confirm that no viable bacteria were present therein, and we performed a phage spot with the filtered supernatant to confirm the absence of phages. Next, we inoculated the supernatant of the *r_lys* cultures, which we expected to contain beta lactamase and the control supernatants of the *n_lys* and *s_lys* cultures with both the ancestral *n_lys* or ancestral *s_lys* populations, both of which are susceptible to Amp and incubated them overnight at 37° C at 180 rpm. We measured OD_{600} at inoculation and after 24 hours to measure bacterial growth in each of the supernatants. We also tested the ability of Amp to induce phages by performing phage spot assays on the supernatant of our Amp treated populations (Figure S2C) and spotting our MIC assay supernatant as well. We do not see production of phages in these populations outside of the very first day which was the same for those not treated with Amp as well. We therefore concluded that Amp did not influence the induction of phages. This was also tested in previous work, where we did not find evidence of Amp inducing Lambda+.[7]

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

All statistical analysis was done in R (v 4.1.3) using RStudio (v 2021.09.0).[70] If not otherwise stated, we report mean(s) and standard error(s). For all analyses aimed to compare differences in free phages and bacteria we excluded control populations, i.e., *n_lys* populations as well as control treatments, i.e., LB and Amp. To compare temporal dynamics of phage and bacterial population sizes across the remaining genotypes and treatments we used a generalized least squares model (gls) suitable for time series data (package *nlme* v3.1 – 160) with time, treatment, and genotype as well as all interactions as fixed effects we also tested for autocorrelation

within our time series data using the Durbin Watson test (package *car* v3.1 – 2).[77] We used pairwise t-tests (package *stats*) to compare the change in phage population size from T8 to T29 in specific populations.

Stochastic model of prophage maintenance across environments

Model structure

We used a stochastic model of a well-mixed environment to simulate the evolutionary dynamics of Lambda+ lysogens, including all experimentally observed bacterial populations. We used a separate set of ordinary differential equations for each of the four treatment types with either type of bacterial lysogens, N_{s_lys} and N_{r_lys} (see Methods S1). The equations are based on population genetics models where populations grow at population-specific net growth rates (g_x) and competition is implemented through a common carrying capacity (K_{total}). We do not explicitly model the free bacteriophage population because lysogen populations do not get (productively) superinfected. However, we approximate killing due to phages in MMC treatments for phage-free and not phage-resistant bacteria by a death rate β. Bacterial cells also have a reduced growth rate (g_{s_lys}, g_{r_lys}) when they are susceptible to phage attachment (see below).

Depending on the treatment, bacteria die due to the killing effect of Amp (ε) and/or due to the killing caused by prophage induction through MMC (α). The effect of a drug on bacterial growth, i.e., ε or α, is usually related to the concentration of a drug via a sigmoid curve (pharmacodynamic function). As the drug concentrations used in our model and LDA are in a regime far below the MIC (minimal inhibitory concentration), we approximate the drug effect however through a linear effect parameter. Subsequently, the concentrations of Amp and MMC are given by A and M, and are the same as in our experiment, 6.4 μg/ml and 0.1 μg/ml, respectively. The impact of Amp or MMC on a bacterial population is then given by the combination of the drug concentration and its killing effect. A and M are decaying at rates δ_A and δ_M respectively, and Amp is degraded additionally in *r_lys* simulations with cooperative prophage-encoded resistance at a rate γ and proportional to the population size of prophage carrying bacteria.

Here, we describe the genetic changes allowed in the most complex model variant, where antibiotic susceptible lysogens (N_{s_lys}) are subjected to both MMC and Amp (see Methods S1 for details on adjustment of subpopulations and genetic changes in other model variants). We allow for three possible mutations to occur in *s_lys* populations in any order but without back-mutations: (i) chromosomal antibiotic resistance mutations at rate μ_σ that confer complete resistance to the population (N_σ), (ii) chromosomal surface receptor mutations at rate $\mu_{\lambda res}$ conferring resistance to bacteriophage infection ($N_{\lambda res}$),[38] and (iii) loss of the prophage at rate $\mu_{\Delta\lambda}$ ($N_{\Delta\lambda}$). Note, that prophage loss is usually caused by a large deletion, hence we assume this mutation rate to be lower than for phage and Amp resistance, which can be caused by one or few mutations. We assume that mutational changes occur during replication, i.e. in the model equations, mutation terms are dependent on the growth rate of the mutating strain. We allow however for mutational events even when the carrying capacity is reached (strictly speaking, K_{total} describes the bacterial limit without taking mutational changes into account).

We assume that chromosomal receptor mutations increase the growth rate of the mutant cells ($g_{\lambda res}$) as they are no longer negatively affected by repeated phage adsorption (35) (Figure S4A). While cells that carry the lysogen are immune against superinfection, phage adsorption is still possible and costly, presumably due to disruption of the membrane integrity (Figure S4A). In the model, the costs stemming from phage adsorption are captured implicitly in reduced growth rates of cells without surface modifications as they might not die but still suffer a growth cost.

For *r_lys* populations, we did not allow chromosomal resistance to arise as the lysogens were already Amp^R. Emergence of chromosomal resistance was therefore only possible after a cell lost the prophage. We tested this assumption in an additional set of simulations, where we allowed chromosomal resistance to appear in any of the populations, but the pattern of prophage loss did not change substantially as double resistance was not more beneficial than either chromosomal or lysogen resistance alone (Figure S5).

To investigate the effect of cooperativity in the phage-encoded benefit, we performed model simulations where the prophages carried antibiotic resistance that provided purely an individual, non-cooperative benefit and compared it with the simulations where prophages provided cooperative antibiotic resistance. The non-cooperative benefit was achieved by removing the additional Amp degradation by resistant lysogens (i.e., setting γ to 0).

Details of equations and populations used in different model versions were adjusted as appropriate to the experimental setup and are detailed in Methods S1. We simulated the ODE model stochastically in R 3.6.0 using the *adaptivetau* package.[70,78] We simulated this model for 50 transfers, and after each transfer 1% of the population was chosen randomly (R function *sample*) to make up the starting population of the next transfer. We started each transfer again with the chosen initial concentration of Amp and MMC. Based on empirical observations Amp and MMC affect bacterial population size only after a set time-period, which we included in our model as a delayed input time for Amp and MMC. We used the *tidyverse* suite of packages to take the averages of the 100 runs for each version of our model and plot them.[79] Below are the series of equations used for implementing our model for *s_lys* bacteria being treated with Amp+MMC.

$$N_{total} = N_{s\_lys} + N_{s\_lys,\sigma} + N_{s\_lys,\Delta\lambda} + N_{s\_lys,\Delta\lambda,\sigma} + N_{s\_lys,\lambda res} + N_{s\_lys,\lambda res,\sigma} + N_{s_{lys},\lambda res,\Delta\lambda} + N_{s\_lys,\lambda res,\Delta\lambda,\sigma}$$

$$\frac{dN_{s\_lys}}{dt} = g_{s\_lys}N_{s\_lys}\left(1 - \frac{N_{total}}{K_{total}}\right) - g_{s\_lys}\mu_{\Delta\lambda}N_{s\_lys} - g_{s\_lys}\mu_{\lambda res}N_{s\_lys} - g_{s\_lys}\mu_\sigma N_{s\_lys} - \varepsilon A N_{s\_lys} - \alpha M N_{s\_lys}$$

$$\frac{dN_{s\_lys,\sigma}}{dt} = g_{s\_lys,\sigma}N_{s\_lys,\sigma}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys}\mu_{\sigma}N_{s\_lys} - g_{s\_lys,\sigma}\mu_{\Delta\lambda}N_{s\_lys,\sigma} - g_{s\_lys,\sigma}\mu_{\lambda res}N_{s\_lys,\sigma} - \alpha MN_{s\_lys,\sigma}$$

$$\frac{dN_{s\_lys,\Delta\lambda}}{dt} = g_{s\_lys,\Delta\lambda}N_{s\_lys,\Delta\lambda}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys}\mu_{\Delta\lambda}N_{s\_lys} - g_{s\_lys,\Delta\lambda}\mu_{\lambda res}N_{s_{lys},\Delta\lambda} - g_{s\_lys,\Delta\lambda}\mu_{\sigma}N_{s_{lys},\Delta\lambda} - \varepsilon AN_{s_{lys},\Delta\lambda} - \beta N_{s_{lys},\Delta\lambda}$$

$$\frac{dN_{s\_lys,\Delta\lambda,\sigma}}{dt} = g_{s\_lys,\Delta\lambda,\sigma}N_{s\_lys,\Delta\lambda,\sigma}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys,\sigma}\mu_{\Delta\lambda}N_{s\_lys,\sigma} + g_{s\_lys,\Delta\lambda}\mu_{\sigma}N_{s\_lys,\Delta\lambda} - g_{s\_lys,\Delta\lambda,\sigma}\mu_{\lambda res}N_{s_{lys},\Delta\lambda,\sigma} - \beta N_{s_{lys},\Delta\lambda,\sigma}$$

$$\frac{dN_{s\_lys,\lambda res}}{dt} = g_{s\_lys,\lambda res}N_{s\_lys,\lambda res}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys}\mu_{\lambda res}N_{s\_lys} - g_{s\_lys,\lambda res}\mu_{\Delta\lambda}N_{s\_lys,\lambda res} - g_{s\_lys,\lambda res}\mu_{\sigma}N_{s\_lys,\lambda res} - \varepsilon AN_{s\_lys,\lambda res} - \alpha MN_{s\_lys,\lambda res}$$

$$\frac{dN_{s\_lys,\lambda res,\sigma}}{dt} = g_{s\_lys,\lambda res,\sigma}N_{s\_lys,\lambda res,\sigma}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys,\sigma}\mu_{\lambda res}N_{s\_lys,\sigma} + g_{s\_lys,\lambda res}\mu_{\sigma}N_{s\_lys,\lambda res} - g_{s\_lys,\lambda res,\sigma}\mu_{\Delta\lambda}N_{s\_lys,\lambda res,\sigma} - \alpha MN_{s\_lys,\lambda res,\sigma}$$

$$\frac{dN_{s\_lys,\lambda res,\Delta\lambda}}{dt} = g_{s\_lys,\lambda res,\Delta\lambda}N_{s\_lys,\lambda res,\Delta\lambda}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys,\lambda res}\mu_{\Delta\lambda}N_{s\_lys,\lambda res} + g_{s\_lys,\Delta\lambda}\mu_{\lambda res}N_{s\_lys,\Delta\lambda} - g_{s\_lys,\lambda res,\Delta\lambda}\mu_{\sigma}N_{s\_lys,\lambda res,\Delta\lambda} - \varepsilon AN_{s_{lys}\lambda res,\Delta\lambda}$$

$$\frac{dN_{s\_lys,\lambda res,\Delta\lambda,\sigma}}{dt} = g_{s\_lys,\lambda res,\Delta\lambda\sigma}N_{s\_lys,\lambda res,\Delta\lambda\sigma}\left(1 - \frac{N_{total}}{K_{total}}\right) + g_{s\_lys,\lambda res,\sigma}\mu_{\Delta\lambda}N_{s\_lys,\lambda res,\sigma} + g_{s\_lys,\Delta\lambda,\sigma}\mu_{\lambda res}N_{s\_lys,\Delta\lambda,\sigma} + g_{s\_lys,\lambda res,\Delta\lambda}\mu_{\sigma}N_{s\_lys,\lambda res,\Delta\lambda}$$

$$\frac{dA}{dt} = -\delta_A A$$

$$\frac{dM}{dt} = -\delta_M M$$

Model fitting

We fitted individual parameters of the model using a direct search algorithm designed for nonlinear multivariable functions in R (*fminsearch* in the *Pracma* package) based on a deterministic version of our model. We typically fitted two parameters at a time, starting with growth rate and carrying capacity of *s_lys* and *r_lys* populations. We used a growth curve of the ancestral lysogen grown in nutrient-rich media (Figure S1B) for fitting, using as initial values estimates for the maximal growth rate of *s_lys* and *r_lys* as obtained with the package *growthcurver* in R on those growth curves and 1×10^9 as the starting carrying capacity. We used the fitted growth rates and carrying capacities for further parameter fitting. We fitted the effects of the two drugs Amp and MMC (ε and α) together with their decay rates (δ_A and δ_M) to growth curves of *s_lys* treated separately with MMC or Amp at the same concentration as in the selection experiment (Figure S1A). We used starting values ranging from 0.1 to 10 to fit the effects of Amp and MMC and their decay rates based on the two extremes of fast death or decay within an hour to very slow death or decay of much longer than 24h. To fit the mutation rates, we used data on the frequency of each type of mutation estimated from 24 clones per population per transfer for the first seven transfers of the experiment, and we used mutation rates estimated from the literature as the initial value for each rate (Figure 3).[80,81] The mutation rates were fitted one at a time with all the other rates being fixed. The fitted mutation rate approximations are in agreement with published mutation rates of *E. coli* genes.[80]

The package *fminsearch* is sensitive to the initial values given as it can get stuck in local optima. Hence, we tested the robustness of our fitted values by varying our initial values over a wide range: for mutation rates we use initial values spanning 4 orders of magnitude, for growth rates and rates of drug effectiveness we used initial values from 0.5-2x the initial estimate described above. Using the fitted parameter set within our stochastic model gave a qualitatively good fit between the model and our empirical data (Figure 3). Further, we tested the robustness of our overall results with a parameter sensitivity analysis in the mutation rates and

effect sizes of Amp and MMC. For a list of parameter values used in the simulations and the parameter sensitivity analysis (LDA) see Tables S3 and S4.

Parameter sensitivity analysis (linear discriminant analysis)

We tested the sensitivity of our simulation results to variations in the magnitude of drug effects and to variations in mutation rates using Latin hypercube sampling and Linear Discriminant Analysis (LDA) as we expect mutation rate and drug concentration to be influential in the loss of prophages.[82] We created 10,000 sets of parameters using Latin hyper-cube sampling (R package *lhs*)[83] to evenly cover the sampling space for variation in five parameters: mutation rate to phage resistance ($\mu_{\lambda res}$), mutation rate to Amp-resistance (μ_{σ}), mutation rate to prophage loss ($\mu_{\Delta\lambda}$), the effect of MMC (α), and the effect of Amp (ε). We used parameter values for mutations rates that evenly covered either side of our fitted values, but we used Amp and MMC effect values which were centered on a smaller value than our fitted one (but included our fitted one) as higher values led to a high percentage of simulations where all populations die out. We ran 100 replicate simulations of each set of parameters for 50 transfers, at which point the system has generally reached steady state (as tested for a subset of parameters).

The results were categorized into different classes and analyzed using LDA in R (*lda* from the package *MASS*).[84] The classes were based on prophage presence within the population, resulting in a majority phage containing class (> 51%), a majority phage free class (> 51%) and a mixed class where neither phage containing, nor phage free bacteria were dominant. We plotted an overview of the relative amounts of these classes as bar plots (Figures 4A and S5) by utilizing a subset of 3335 parameter sets of the 10,000 parameter sets that had the same parameter values in all three simulation scenarios, *s_lys*, *r_lys* (cooperative), *r_lys* (non-cooperative). (We used 2000 parameter sets for the simulations where chromosomal Amp-resistance mutations were allowed in *r_lys* bacteria.) LDA was then used to determine which parameters separate these classes maximally.[82] The magnitude and direction of the parameter arrows show their significance in separating certain classes, which correlates with the magnitude of their impact on the simulation results (Figure 4B).

The results of the LDA were used to visualize the relative impact of the two selection pressures MMC and Amp on prophage maintenance at T30 and T50 (Figures 4 and 5). For our visualizations, we removed simulations where every sub-population becomes zero. For the LDA plots, we averaged the frequency of prophage-carriers over 100 different simulations using the same parameter set of mutation rates, MMC and Amp effect values. We also visualized our results with contour plots that show the frequency of prophage carrying populations at a particular combination of Amp and MMC effect. For our contour plots, we averaged the frequency of prophage carriers over all simulations done for a parameter combination of MMC and Amp values (meaning that mutation rate parameters were not necessarily the same between simulations used for averaging).