bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Immune Loss as a Driver of Coexistence During Host-Phage Coevolution

Jake L Weissman[1], Rayshawn Holmes[4], Rodolphe Barrangou[2], Sylvain Moineau[3], William F Fagan[1], Bruce Levin[4], Philip LF Johnson[1*],

**1** Department of Biology, University of Maryland College Park, College Park, MD, USA
**2** Department of Food, Bioprocessing, & Nutrition Sciences, North Carolina State University, Raleigh, NC, USA
**3** Department of Biochemistry, Microbiology and Bioinformatics, Université Laval, Québec, Québec, Canada
**4** Department of Biology, Emory University, Atlanta, GA, USA

* plfj@umd.edu

## Abstract

Bacteria and their viral pathogens face a constant pressure for augmented immune and infective capabilities respectively. Under this reciprocally imposed selective regime we expect to see a runaway evolutionary arms race that ultimately leads to the extinction of bacteriophages, either due to immunity making their host inaccessible or the loss of their host due to exhaustion of susceptible subpopulations. Despite this prediction, in many systems host and phage coexist with minimal coevolution. Previous work explained this puzzling phenomenon by invoking spatial structure or fitness tradeoffs, which can drive coexistence and the diminishment of an arms race dynamic. However, these explanations do not apply to all systems. Here we propose a new hypothesis, that the regular loss of immunity by the bacterial host can also produce robust host-phage coexistence. We pair a general model of immunity with an experimental and theoretical case study of the CRISPR-Cas immune system to characterize and contrast the behavior of tradeoff and loss mechanisms in well-mixed systems. We find that, while both a cost of immunity and the loss of immunity can lead to stable coexistence, only a loss mechanism can do so robustly within a realistic region of parameter space.

## Author Summary

Bacteria in natural environments must deal with the constant threat of infection by bacteriophages (i.e., viruses that infect bacteria). This creates a situation in which bacteria that are able to develop defense against these "phages", generally via novel mutations, have a selective advantage in the population and may rise to dominance. This rise of resistant mutants in turn puts pressure on phages so that mutations allowing phages to maintain infectivity will spread through the phage population. In this coevolutionary arms race bacteria and phages each trade mutations in an escalating battle to expand their range of resistance and infectivity respectively. Theory predicts that such arms races are unsustainable in the long term. Thus it is unclear what environmental conditions and characteristics of bacteria and phages can lead to the long-term coexistence of both parties. Here we draw into question the generality of past explanations that applied a cost in terms of growth rate to resistant bacterial strains.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We propose additionally that if bacterial defense fails or is lost at a sufficiently high rate this can produce long-term coexistence. Furthermore, this loss-driven coexistence is not sensitive to perturbations of the system.

# 1 Introduction

While the abundance of bacteria observed globally is impressive [1–3], any apparent microbial dominance is rivaled by the ubiquity, diversity, and abundance of predatory bacteriophages, which target these microbes [4–8]. As one might expect, "phages" are powerful modulators of microbial population and evolutionary dynamics, and thus of global nutrient cycles [1, 4, 6, 7, 9–14]. Despite these indications of ecological importance, the diversity and distribution of phages is poorly characterized, leading some to call them the "dark matter" of biology [15]. Mirroring our lack of knowledge about the general characteristics of phage populations, we lack a comprehensive understanding of the the dynamical behavior of these populations. More specifically, it is an open question what processes sustain phage populations in the long term across habitats.

Phages persist in a variety of natural environments (e.g. [16, 17]) and artificial laboratory systems (e.g. [18–25]) despite the fact that their host can evade them using both passive forms of resistance (e.g. receptor loss, modification, and masking) and active immune systems that lead to the degradation of phages (e.g. restriction-modification systems, CRISPR-Cas). The processes that maintain this coexistence of host and phage are still being elucidated, and while several hypotheses have been set forth, the known drivers of coexistence are insufficient to explain the many observed cases where phages and host persist.

Coexistence poses an issue because in the classical view host and phage are under constant selective pressure for augmented defense and infective capabilities respectively. This reciprocal conflict can produce an escalating arms race dynamic in which host and pathogen each drive the evolution of their respective partner [26, 27]. In the context of such an arms race, each party must continually keep up with the evolutionary pace of the other, or the race must end [28–30]. Despite basic theory that predicts that the dynamics of well-mixed host-phage systems will generally be unstable and sensitive to initial conditions [22], there are many examples in which both parties coexist stably over long periods of time in natural and laboratory systems (e.g [16, 18, 20–25, 31]).

These examples of apparently stable coexistence have motivated a search for mechanisms that might explain the deescalation and eventual cessation of a coevolutionary arms race dynamic, despite fitness gains each player might achieve by continuing to acquire novel phenotypes. Previous authors have identified (1) fluctuating selection, (2) spatial heterogeneity, and (3) costs of defense as potential drivers of coexistence. Here we propose (4) the loss of immunity or resistance as an additional mechanism. We differentiate between intracellular *immunity* (e.g., CRISPR-Cas) in which immune host act as a sink for the phage population versus extracellular *resistance* (e.g., receptor modifications) that prevent phages from adsorbing to the host outright. Here, we focus on immunity rather than resistance because phages will generally have a more difficult time persisting in the face of an active immune system than passive resistance, and thus this case poses more of a puzzle when explaining long-term coexistence. Intracellularly acting forms of resistance also exist, but should behave qualitatively similarly to intracellular immunity since both create sinks for the phage population.

Host and phage almost certainly cannot escalate an arms race indefinitely via novel mutations [21, 32, 33]. The reemergence of or reversion to old, rare genotypes provides an alternative to the generation of new mutations, and can lead to a fluctuating selection dynamic in which alleles cycle in frequency over time [28, 29, 34, 35]. There is

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

empirical evidence that escalating arms races give way to this type of fluctuating selection dynamic over time in some host-phage systems [33]. In contrast to this result, when novel immune or infective phenotypes correspond to increased generalism we do not expect past phenotypes to recur [34, 35] since they will no longer be adaptive. The expansion of generalism during coevolution has been shown to be typical of many experimental microbial systems [36]. Thus we might expect a clear "winner" of the arms race in these situations, leading to a breakdown in coexistence. Therefore, while a fluctuating selection dynamic may explain long term coexistence in some systems, it does not seem to be sufficient to explain the majority of host-phage coexistence observed in the laboratory.

Both empirical and theoretical work has concluded that spatial structure can support coexistence [22, 37, 38]. If spatial refuges serve to partially protect a susceptible host population then phages may persist even in cases where we would expect them to be cleared from the system [22]. Additionally, if local adaptation keeps host subpopulations in distinct spatial regions from becoming globally defended against phages then a "geographic mosaic" may be produced which promotes phage persistence [36, 39]. Spatial effects cannot account for the behavior of all systems though, because coexistence has been observed in serial transfer experiments where populations are well-mixed (e.g., [25]).

In well-mixed systems it has been shown that costs associated with immunity can produce stable coexistence. When host and phage face tradeoffs between growth and immunity or host range respectively, coexistence can occur [20, 31, 40, 41]. If phage evolution is constrained either due to architectural constraints or pleiotropic costs then the bacterial host faces no pressure to evolve heightened defense beyond a certain threshold. In this case a tradeoff between increased immunity and growth rate in the host can lead to the maintenance of a susceptible host population on which phages can persist [19–21, 40, 42, 43]. Tradeoff-based mechanisms may drive coexistence in some systems, but they require a high cost of immunity that does not apply to all systems (e.g., [22]). In the case of host immunity, phages experience adsorption-dependent death such that their growth rate depends on the ratio of the densities of susceptible to immune host. In cases of host resistance the growth rate of phages solely depends on the density of susceptible hosts. Thus an even higher cost of defense will be required in cases of host immunity than host resistance in order to maintain a phage population. Here we show that in the case of intracellular immunity a tradeoff mechanism requires an unreasonably high cost of immunity to produce a coexistence regime that is robust to the initial conditions of the system. In general, we conclude that cost-based mechanisms do not produce as robust coexistence regimes as alternative mechanisms, namely immune loss.

In large host populations, which are typical of bacterial cultures, even a low rate of loss of immunity/resistance may produce a substantial susceptible host population. Thus loss could produce long-term host-phage coexistence by maintaining enough susceptible host for phage to reproduce. Delbrück [44] initially describes this hypothesis of loss of defense via back-mutation in order to challenge the evidence for lysogeny, differentiating between "true" and "apparent" lysogenesis. Lenski [45] reiterates this hypothesis in terms of phenotypic plasticity and notes that conditioning the production of a susceptible host population on a resistant one could lead to very robust, host-dominated coexistence.

Schragg and Mittler [22] look for loss of resistance in an *E. coli*-phage system and see evidence of loss in certain cases, but only observe coexistence over limited time spans in serial transfer. Nevertheless, they conclusively rule out a tradeoff mechanism in their system since the cost of acquiring resistance is simply not high enough to sustain a susceptible host population. Thus, even in the absence of spatial heterogeneity and costs

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

on immunity, some mechanism is stabilizing their system in part, quite possibly the loss of resistance. More recently, Meyer *et al.* [46] present an empirical example of a system in which stochastic phenotypic loss of resistance leads to persistence of a coevolving phage population even given an almost entirely resistant host population. In cases where the host encodes a dedicated immune system such as a restriction modification system or CRISPR, corruption or wholesale loss of that system via mutation will lead to genetic, permanent immune loss. In the case of receptor modification or down-regulation of receptors, failure of resistance may be entirely phenotypic without any genetic component determining loss. While permanent and temporary loss mechanisms should behave qualitatively similarly, here we focus on the permanent loss case.

We hypothesize that coexistence equilibria will be more robust under an immune loss mechanism than in the case of costly immunity because a loss mechanism conditions the production of susceptible host on an extremely robust, resource-limited immune host population. We build a general mathematical model to demonstrate this point and then use a combination of experimental evidence and simulation-based modeling to apply this result to the coevolution of *Streptococcus thermophilus* and its lytic phage 2972 in the context of CRISPR immunity.

## 2 General Immune Loss Model

Here we modify a previous model of rotifer-algal cryptic predator-prey dynamics [43] to fit typical host-phage dynamics [31, 47] such that resistant host face autoimmunity and immune system loss. "Defended" bacteria have a parameterized degree of immunity from phages ($\phi_d$). In reality, it is unlikely that coevolution would maintain a constant fraction of the host population in a susceptible state; however, our parameterization of immunity accounts for the net effect of phage coevolution, which is not otherwise incorporated into this model. This simplified model allows us to compare the steady states and general behavior of various parameter combinations analytically. Coevolutionary dynamics will be included explicitly in our examination of the CRISPR immune system below.

We choose to have autoimmunity be the primary cost of immunity in our model as it allows for us to readily apply our results to the CRISPR system examined later. Additionally, since autoimmunity and immune loss can both be implemented identically into the equation for defended host they are more easily compared. Autoimmunity can decrease the growth rate of individual hosts [48], or be lethal; we focus on the latter as it will have a stronger effect on the population dynamics [48–50]. We implement autoimmunity as an additional death term in the defended host population ($\alpha$). We find similar general results when applying a penalty to the resource affinity or maximum growth rate of the defended host (S1 Text, S1 Fig-S8 Fig). We also add flow from the defended to "undefended" host population representing loss of immunity at rate $\mu$. We refer to defended and undefended host populations here rather than "immune" and "susceptible" populations because this deterministic model allows for partially effective immunity due to coevolution, which reduces the mean immunity across the defended population, and additional secondary innate resistance/immunity in the undefended host that provides some level of protection. Thus defended host may not be completely immune nor undefended hosts completely susceptible.

We examine the chemostat system with resources:

$$\dot{R} = w(A - R) - \frac{evR}{z + R}(D + U) \tag{1}$$

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

defended host:

$$\dot{D} = D\left(\frac{vR}{z+R} - \delta\phi_d P - \alpha - \mu - w\right), \tag{2}$$

undefended host:

$$\dot{U} = U\left(\frac{vR}{z+R} - \delta\phi_u P - w\right) + \mu D, \tag{3}$$

and phage:

$$\dot{P} = P\left(\delta U(\phi_u\beta - 1) + \delta D(\phi_d\beta - 1) - w\right), \tag{4}$$

where $0 < \phi_u, \phi_d \leq 1$ are the degrees to which phages are able to infect the undefended and defended populations respectively. The former ($\phi_u$) allows for secondary forms of defense, such as a restriction-modification system [51, 52]. The latter ($\phi_d$) allows for coevolution of phages that can predate even the defended host population. Parameter definitions and values for all analytical models can be found in Table 1 and rationale/references for parameter values in S2 Text.

**Table 1. Definitions and oft used values/initial values of variables, functions, and parameters for the general mathematical model**

| Symbol | Definition | Value |
|---|---|---|
| $R$ | Resources | $R(0) = 350$ $\mu$g/mL |
| $D$ | Defended Host | $D(0) = 10^6$ cells/mL |
| $U$ | Undefended Host | $U(0) = 10^2$ cells/mL |
| $P$ | Phage | $P(0) = 10^6$ particles/mL |
| $e$ | Resource consumption rate of growing bacteria | $5 \times 10^{-7}$ $\mu$g/cell |
| $v$ | Maximum bacterial growth rate | 1.4 divisions/hr |
| $z$ | Resource concentration for half-maximal growth | 1 $\mu$g/mL |
| $A$ | Resource pool concentration | 350 $\mu$g/mL |
| $w$ | Flow rate | 0.3 mL/hr |
| $\delta$ | Adsorption rate | $10^{-8}$ mL per cell per phage per hr |
| $\beta$ | Burst Size | 80 particles per infected cell |
| $\phi_u$ | Degree of susceptibility of undefended host | 1 |
| $\phi_d$ | Degree of susceptibility of defended host | 0 |
| $\alpha$ | Autoimmunity rate | $2.5 \times 10^{-5}$ deaths per individual per hr |
| $\mu$ | Rate of immune inactivation/loss | $5 \times 10^{-4}$ losses per individual per hr |

Assuming no phage coevolution ($\phi_d = 0$), this model has a single equilibrium in which all population coexist (S1 Table). We analyze this equilibrium analytically and also use a numerical solution to verify that the equilibrium is reachable from plausible (e.g., experimental) starting values (S3 Text). Our question becomes under what parameter regimes will phage persist but defended host dominate at a stable equilibrium such that $D > P$ at equilibrium?

In Fig 1 we explore model behavior under varying rates of autoimmunity ($\alpha$) and immune loss ($\mu$). Clearly there is a boundary near where autoimmunity and loss rates surpass unity above which defended host go extinct in the face of excessive immune loss and autoimmune targeting. At the opposite end of the parameter spectrum, we see coexistence disappear from the low end of the parameter ranges as phage populations collapse. This leads to a band of parameter space where coexistence is possible, stable, robust, and host-dominated. In this band autoimmunity and/or immune loss occur at a high enough rates to ensure maintenance of coexistence, but not so high as to place an excessive cost on immunity. We note that this band is much more constrained in the $\alpha$-dimension, which essentially restricts the rates of autoimmunity that can cause stable coexistence to an implausibly high and narrow region of parameter space. Under a

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

loss-mechanism, excess undefended host are produced and thus phage populations increase, with the culture even becoming phage-dominated at high $\mu$ (S2 Fig). This allows for a greater robustness of coexistence under a loss mechanism even at low loss rates (Fig 1, S2 Fig-S9 Fig).

We can also determine the effects of alternative forms of defense. In the case of phage coevolution ($\phi_d > 0$), the equilibria still have closed forms, but are not easily representable as simple equations (S3 Text). When $\phi_d > \frac{1}{\beta}$, defended host begin to contribute positively to phage growth, which leads to an eventual shift in the coexistence equilibrium from host to phage dominance (S10 Fig). Adding innate immunity to the system by decreasing $\phi_u$, new regions of the parameter space become phage dominated, particularly at high $\alpha$ (S11 Fig). This is in line with counterintuitive results that suggest that higher immunity may increase phage density by allowing the host population to increase in size [53]. We note though, that the model requires innate immunity with over a 50% effectiveness in combating phage infection to see even a small shift in behavior, suggesting that secondary defense in the undefended strain (e.g. innate envelope resistance) has minimal effects unless it provides near-complete protection.

# 3 A Case Study: CRISPR-Phage Coevolution

The CRISPR (Clustered Regularly Inter-spaced Short Palindromic Repeats) prokaryotic adaptive immune system incorporates novel and highly specific immune "memory" and then uses this memory to clear infections by selfish genetic elements such as bacteriophages and plasmids [54–57]. These features place a high degree of selective pressure on phages to modify targeted sequences in order to escape immune action, and can lead to accelerated phage and bacterial diversification [25, 58, 59]. The effects of this rapid coevolution on host-phage population dynamics remain largely unexplored.

A CRISPR immune system is composed of two genomic regions: a set of short ($\sim$ 30bp) "spacer" segments that match the targeted foreign sequences ("protospacers") interspersed by conserved "repeat" segments, and a set of *cas* (CRISPR-associated) genes that provide the machinery for the acquisition of novel immunity and the immune function itself [60]. We refer to a "CRISPR locus" as the combined set of *cas* genes and a CRISPR repeat-spacer array, which typically, but not always, are located near each other in the genome [61, 62].

In type I and II CRISPR systems, protospacers are flanked by a short motif called a protospacer adjacent motif (PAM) that is essential for the adaptation and targeting steps of CRISPR immunity and prevents targeting of content within the CRISPR locus itself [58, 63–65]. Any mutation in this motif completely negates the effect of an associated spacer [66], although as we move away from the PAM along the protospacer sequence more substitutions are tolerated by the matching machinery [66]. Thus a single base pair substitution in or near the PAM can nullify the effects of CRISPR immunity, providing phages with a straightforward evolutionary escape route [58, 63, 66, 67], and indeed there is a bias in observed phage SNPs towards these regions [25]. This can create an escalating arms race in which immunity in bacteria selects for escape mutants in phages, which in turn select for novel spacer acquisitions in bacteria, and so on [25, 68, 69].

While coevolutionary arms races occur across diverse taxa (e.g., phytophagous insects [70], garter snakes and newts [71], and bacteria and bacteriophages [72]), predicting the outcome of an arms race in a CRISPR-phage system poses a particularly difficult problem as evolutionary dynamics occur on the same timescale as population dynamics and clonal interference can play a significant role [25, 68, 73, 74]. Because CRISPR essentially allows bacteria to pay a "lump-sum" cost for the acquisition of the system but does not "charge" for the acquisition of novel immunity in the form of

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

spacers [75], it provides a unique coevolutionary structure in which bacteria can continue to escalate the arms race for longer than otherwise possible via alternative "innate" immune strategies (e.g., receptor modifications [20, 21, 72]). How phages can persist in the face of this extremely adaptable immune system remains unclear. Previous theoretical and limited experimental work has explained persistence by invoking spatial heterogeneity [38, 76], negative frequency-dependent selection [68], and host switching to a constitutive defense strategy such as surface receptor modification [77, 78].

While these hypotheses may play a role in persistence, they do not explain simple co-evolution experiments with *Streptococcus thermophilus* (type II-A CRISPR system) and its lytic phage 2972 which resulted in either quasi-stable long-term coexistence or rapid phage loss [25]. These experiments produced extended periods of seemingly stable host-phage coexistence wherein host appear to be resource rather than phage limited and phages remain at low density. This resource-limitation implies that the bacterial host has "won" the arms race, and that CRISPR immunity is providing complete or near complete defense against phages infection, yet phages are able to persist on a susceptible sub-population of host. We rule out a Red Queen dynamic based on the apparent stability of the host population density and based on the sequencing results we present below. Additionally, since experiments are carried out in liquid culture with daily serial dilutions we do not expect spatial heterogeneity to play a role. This implies that either (1) costs associated with CRISPR immunity or (2) the loss of CRISPR immunity is playing a role in maintaining susceptible host populations on which phages can persist.

There are two primary costs that accompany a functional CRISPR system: autoimmunity via the acquisition self-targeting spacers and the inhibition of horizontal gene transfer. Autoimmunity may play an important role in the evolution of CRISPR defense. It is unclear how or if bacteria distinguish self from non-self during the adaptation step of CRISPR immunity [79–83]. Experimental evidence suggests that there is no mechanism of self vs. non-self recognition in *S. thermophilus* so that self-targeting spacers are acquired frequently and may have a significant impact on fitness [79]. Previous work on cryptic population dynamics has shown that, when a resistant prey phenotype appears that is less fit, an equilibrium condition or stable limit cycle can result in which the resistant phenotype dominates the prey population but the cost of resistance allows a small susceptible subpopulation to persist [20, 21, 42, 43]. Horizontal gene transfer is unlikely to play a role in determining fitness during the experiments we consider.

While an immunity-growth tradeoff could provide the competitive boost needed to maintain a persisting susceptible population, the outright loss of CRISPR immunity at a high enough rate could also lead to the accumulation of susceptible host. The bacterial host *Staphylococcus epidermidis* loses functionality in its CRISPR system, either due to wholesale loss of the system or mutation of essential sequences (i.e. the leader sequence or *cas* genes), at a rate of $10^{-4}$-$10^{-3}$ inactivation/loss events per individual per generation [84], and CRISPR loss has been observed in other systems as well [85, 86]. We consider any event that leads to a non-functional CRISPR system as a "loss" event, regardless of whether or not any of the locus remains in the genome.

The observed CRISPR loss rate is exceptionally high, and may maintain a sufficient susceptible host density for phages to persist even in the absence of any cost on the CRISPR system. Our general model above hints that regular loss of the CRISPR locus may lead to a more robust coexistence equilibrium than the imposition of a CRISPR-growth rate tradeoff, since the production of the phage food source is linked explicitly to the presence of resistant host. Similarly, the loss of individual spacers at a high rate could account for a stable pool of susceptible host. Spacer loss can likely be attributed to homologous recombination [58] and thus can occur anywhere in the

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CRISPR locus. Therefore this process should leave an observable signature in CRISPR locus sequencing data if it occurs frequently enough to have a significant effect on population dynamics.

Below we replicate the serial-transfer coevolution experiments performed by Paez-Espino et al. [25, 50] and develop a simulation-based model to explain the phenomenon of coexistence. Our empirical results effectively rule out the existence of a Red Queen dynamic and the widespread loss of spacers. Our simulations indicate that the cost imposed by autoimmunity cannot produce robust coexistence in realistic parameter ranges, as implied by our more general model above, but that the functional loss of the CRISPR-Cas system can. Therefore, we propose that the stochastic loss of the CRISPR system can lead to host-phage coexistence in the long term following the observed host-dominated pattern and that the effects of autoimmunity are insufficient to explain the persistence of a susceptible host population.

## 3.1 Experiments

We performed long-term daily serial transfer experiments with *Streptococcus thermophilus* and its lytic phage 2972 in milk, a model system for studying CRISPR evolution (see S4 Text for detailed methods). We measured bacteria and phage densities on a daily basis. Further, on selected days we PCR-amplified and sequenced the CRISPR1 locus, which is known to be the most active in the *S. thermophilus* host used here [56].

From the perspective of density, phages transiently dominated the system early on, but the bacteria quickly took over and by day five appeared to be resource-limited rather than phage-limited (Fig 2). This switch to host-dominance corresponded to a drop in phage populations to a density two to three orders of magnitude below that of the bacteria. Once arriving at this host-dominated state, the system either maintained quasi-stable coexistence on an extended timescale (over a month and a half), or phages continued to decline and went extinct relatively quickly (Fig 2). We performed six additional replicate experiments which confirmed this dichotomy between either extended co-existence (4 lines quasi-stable for $> 2$ weeks) or quick phage extinction (2 lines $< 1$ week) (S12 Fig). In addition, a previous long-term coevolution experiment in the same system observed host-dominated coexistence for the better part of a year and was robust to starting conditions [25, 50]. In these experiments, phage chimerism due to a secondary contaminating phage may have extended the period of coexistence, but even in the case where there was no contamination they observe host-dominated coexistence for a period of approximately one month.

Sequencing of the CRISPR1 locus reveals the rapid gain of a single spacer (albeit different spacers are found in different sequenced clones) followed by a very slow increase in spacer counts with time (S13 Fig) that is not consistent with a rapid red queen / arms race dynamic. We did not observe a signature of frequent spacer loss in the CRISPR1 array.

## 3.2 CRISPR-phage Coevolutionary Model

While our simplified model yielded closed form expressions for equilibria, it lacked the realistic coevolutionary dynamics of the CRISPR-phage system wherein bacteria acquire immunity via the addition of novel spacers and phages escape this immune response via protospacer mutations. We next built a hybrid deterministic/stochastic lineage-based model similar to an earlier model by Childs et al. [68, 87]. In this model coevolution is incorporated explicitly, and we are able to examine the effects that early steps in the host-phage arms race have on the overall outcome of the system. Our simulations also replicate the resource dynamics of a serial dilution experiment rather than a chemostat.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

This final modification is essential as the two resource environments are not always comparable (e.g., [22]). Population dynamics are dictated by a set of ordinary differential equations and new equations are added to the system stochastically to simulate spacer acquisition and phage mutation. Simulation procedures are detailed in S3 Text.

We consider only mutations in the PAM region in our model, which is the dominant form of CRISPR escape by phages [25]. This approach differs from other models which consider mutations in the protospacer region that negate matching but allow the spacer to be reacquired (e.g., [53, 68, 88]). A PAM mutation completely removes a protospacer from further consideration rather than simply modifying its sequence. We expect that the acquisition of novel spacers from alternative locations on the phage genome is more relevant in a coevolutionary context than the re-acquisition of "obsolete" spacers. Since the probability of re-acquisition will be quite low if there are many protospacers, and since an acquisition elsewhere in the genome that has not undergone selection for mutation in the phage population provides an opportunity for much more broad-based immunity, we hold that re-acquisition is the less important phenomenon. Therefore we assume that PAM mutations will have a far more important effect on the overall coevolutionary dynamics. This difference in relevance is further compounded by the fact that as we move away from the PAM along the protospacer sequence more substitutions are tolerated by the CRISPR matching machinery [66], meaning that mutations farther away from the PAM will be less able to combat immunity.

Population dynamics are modeled by a set of differential equations for resources:

$$\dot{R} = \frac{-evR}{z+R}\left(U + \sum_i D_i\right) \tag{5}$$

CRISPR-enabled bacteria with spacer set $X_i$:

$$\dot{D}_i = D_i\left(\frac{vR}{z+R} - \delta\left(\sum_j (1 - M(X_i, Y_j))P_j\right) - \alpha - \mu_L\right) \tag{6}$$

a pool of undefended bacteria with a missing or defective CRISPR system:

$$\dot{U} = U\left(\frac{vR}{z+R} - \delta\sum_i P_i\right) + \mu_L \sum_i D_i \tag{7}$$

and phages with protospacer set $Y_i$ :

$$\dot{P}_i = \delta P_i\left(U(\beta_i - 1) + \sum_j D_j(\beta_i(1 - M(X_j, Y_i)) - 1)\right), \tag{8}$$

and stochastic events occur according to a Poisson process with rate $\lambda$:

$$\lambda = \sum_i \lambda_{B_i} + \sum_i \lambda_{P_i} + \sum_i \lambda_{K_i} \tag{9}$$

which is a sum of the total per-strain spacer-acquisition rates:

$$\lambda_{B_i} = \mu_b \delta D_i \sum_j P_j \tag{10}$$

total per-strain PAM mutation rates:

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$\lambda_{P_i} = \mu_p \beta_i \delta P_i \left( U + \sum_j (1 - M(X_i, Y_j)) D_i \right) \tag{11}$$

and total per-strain PAM back mutation rates:

$$\lambda_{Q_i} = \mu_q \beta_i \delta P_i \left( U + \sum_j (1 - M(X_i, Y_j)) D_i \right). \tag{12}$$

The function $M(X_i, Y_j)$ is a binary matching function between (proto)spacer content of bacterial and phage genomes that determines the presence or absence of immunity. We refer the the "order" of a host or phage strain, which is the number of evolutionary events that strain has undergone, $|X_i|$ or $n_s - |Y_i|$ respectively. The PAM back mutation rate describes the rate at which we expect a mutated PAM to revert to it's original sequence (assuming the mutation is a substitution), where the back mutation rate parameter $\mu_q$ should be considerably smaller than the forward rate $\mu_p$. Because there is no reason for a PAM with more than a single mutation to be selected for since a single substitution can completely inhibit immune action [66], we assume reversion to the original sequence requires a single specific substitution at the swapped nucleotide. While back mutation is not required to generate stable host-dominated coexistence, it greatly expands the relevant region of parameter space because it allows phages to avoid the cost we will impose on PAM mutations, discussed below, when those immune escape mutations are no longer beneficial (Fig 4). A similar effect is seen by adding viral recombination to the model, which provides another route to an un-mutated or less mutated genome (results not shown).

We assume that the number of PAM mutations in a single phage genome is constrained by a tradeoff with phage fitness, as this is necessary to prevent the total clearance of protospacers from a single strain at high mutation rates. There is empirical evidence that increases in host breadth generally come at a cost for viruses due to pleiotropic effects [89], although not all individual mutations necessarily come with a cost [90]. These studies, though, generally examine host range expansions to entirely novel host species rather than much more closely related CRISPR variants. That said, mutations tend to be deleterious on average (e.g [91]) and so it is reasonable to expect PAM mutations to, on average, come with a cost. Additionally, assuming the phage in question has been associated with its host over a long evolutionary history, and that the host has possessed a CRISPR system during this time, it is reasonable to speculate that the phage has been under pressure to lose any active PAMs on its genome, and thus that the persisting PAMs may have been preserved because their loss is associated with a fitness cost. In our model, burst size presents itself as an attractive option for the incorporation of a cost, since a decrease in burst size can signify a decrease in either fecundity or viability.

The function

$$\beta_i = -\frac{c\beta_{\text{base}}}{n_s}|Y_i| + \beta_{\text{base}} \tag{13}$$

incorporates a linear cost of mutation into the phage burst size. PAM back mutation allows the phage population to recover in fitness by escaping these costs in the case where phages with relatively few mutations have already gone extinct. See Table 2 for further definitions of variables, functions, and parameters in Equations 5-13. Simulation procedures and rationale for parameter values, including phage genome size, are detailed in S3 Text.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Table 2. Definitions and oft used values/initial values of variables, functions, and parameters for the simulation model**

| Symbol | Definition | Value |
|---|---|---|
| $R$ | Resource concentration | 350 $\mu$g/mL |
| $B_i$ | Population size of CRISPR$^+$ bacterial strain $i$ | $10^6$ |
| $P_i$ | Population size of phage strain $i$ | $10^6$ |
| $B_u$ | Population size of CRISPR$^-$ bacteria | $10^2$ |
| $\lambda_{B_i}$ | Mutation rate of bacterial strain $i$ | n/a |
| $\lambda_{P_i}$ | PAM mutation rate of phage strain $i$ | n/a |
| $\lambda_{Q_i}$ | PAM back mutation rate of phage strain $i$ | n/a |
| $\lambda$ | Total rate of mutation events occurring in model | n/a |
| $M(X_i, Y_j)$ | Matching function between spacer set of bacterial strain $i$ and protospacer set of phage strain $j$ | no matches initially |
| $\beta(\lvert Y_i \rvert)$ | Burst size as a function of the order of phage strain $i$ | $\beta(0) = 80$ |
| $\lvert X_i \rvert$ | Order of bacterial strain $i$ | 0 |
| $\lvert Y_i \rvert$ | Order of phage strain $i$ | 0 |
| $e$ | Resource consumption rate of growing bacteria | $5 \times 10^{-7}$ $\mu$g |
| $v$ | Maximum bacterial growth rate | 1.4/hr |
| $z$ | Resource concentration for half-maximal growth | 1 $\mu$g |
| $\delta$ | Adsorption rate | $10^{-8}$ mL per cell per phage per hr |
| $\beta_{\text{base}}$ | Maximum burst size | 80 particles per infected cell |
| $n_s$ | Size of phage genome | 10 protospacers |
| $c$ | Cost weight per PAM mutation | 3 |
| $\mu_L$ | Per individual per generation CRISPR inactivation/loss rate | $5 \times 10^{-4}$ |
| $\alpha$ | Rate of autoimmunity | $50\mu_b$ deaths per individual per hr |
| $\mu_b$ | Spacer acquisition rate | $5 \times 10^{-7}$ acquisitions per individual per hr |
| $\mu_p$ | Per-protospacer PAM mutation rate | $5 \times 10^{-8}$ mutations per spacer per individual per hr |
| $\mu_q$ | PAM back mutation rate | $5 \times 10^{-9}$ mutations per spacer per individual per hr |

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

#### 3.2.1 Stable Host-Dominated Coexistence

The results of our simulations accounting for host-phage coevolution confirm our qualitative analytic results that stable long-term coexistence can occur as the byproduct of a CRISPR-lacking host subpopulation (e.g., Figs 2, 5). It is important to note that whereas our analysis of the deterministic model largely focused on equilibrium states of the system, our coevolutionary simulations serve to investigate the early, transient stages of immune development and thus capture if and how a system arrives at a state of stable-coexistence rather than whether that state exists in the first place.

Simulations with immune loss reliably produce extended coexistence within a realistic region of the parameter space (Fig 3). This region does not overlap with the parameter region shown to produce robust stable coexistence with the simpler deterministic model (Fig 1). We observed no simulations in which autoimmunity alone produced stable coexistence. This agrees with our earlier numerical results where unrealistically high rates of autoimmunity were required to produce coexistence.

As seen in our experimental results, for a single set of parameters the host-phage system can stochastically fall into either stable coexistence or into a phage-free state (Fig 3). The frequency with which we see each outcome, as well as the distribution of times that phages are able to persist, is dependent on the specific set of parameters chosen. As seen in Fig 4, the rate of PAM back mutation in phages is important for producing coexistence, although even in the absence of back mutation the system can reach a point of stable coexistence, albeit rarely. This dependence on back mutation is caused by the combined effects of the cumulative cost we impose on PAM mutations and the inability of phages to keep up with host in a continuing arms race. In the early stages of the arms race it is optimal for phages to continue undergoing PAM mutations as the most abundant available hosts are high-order CRISPR variants, whereas once hosts are able to pull sufficiently ahead of phage in the arms race it becomes optimal for phages to feed on the lower-density but consistently available CRISPR-lacking host population (S14 Fig).

The adsorption rate, on a coarse scale, has an important effect on how the model behaves (S15 Fig). At high values of $\delta$ where we would expect phages to cause host extinction in the absence of CRISPR immunity ($\delta = 10^{-7}$) we see that long-term coexistence of phages and hosts does not occur frequently, and is negatively associated with the phage back mutation rate. In this case phages will rapidly consume the susceptible host population and crash to extinction unless they have undergone PAM mutations that lower their growth rate. The appearance of a lower order phage strain can cause a rapid decline in the susceptible host population and precipitate phage extinction. This causes a reversal in the previous trend seen with back mutation where the ability of phages to escape the costs of PAM mutation was essential to their persistence. A decrease in the adsorption rate to a very low value ($\delta = 10^{-9}$) leads to most simulations persisting in host-dominated coexistence until the 80 day cutoff. Because both evolutionary and demographic dynamics occur much more slowly in this case, long term persistence does not necessarily imply actual stability, as suggested by our and the previous Paez-Espino experimental results in which coexistence eventually ends [25]. In general, lower adsorption rates lead to longer periods of host-dominated coexistence and reduce the chance of phage extinction.

The failure of autoimmunity to produce coexistence warrants further investigation. Upon closer examination, it is clear that in early stages of the arms race where CRISPR-enabled bacteria have not yet obtained spacers or been selected for in the host population, phages are able to proliferate to extremely high levels. During this period the CRISPR-lacking host are overwhelmed by phages and go extinct. Because autoimmunity as a mechanism of coexistence relies on the continued presence of immune-lacking host, it may not be able to function in the face of this early phage

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

burst, which is consistently seen across all simulations where CRISPR-enabled bacteria are initiated with naive CRISPR arrays. This would further act to rule out autoimmunity as a mechanism capable of producing coexistence. There is a possibility that very low locus loss rates that reintroduce CRISPR-lacking bacteria but do not appreciably contribute to their density combined with high rates of autoimmunity could maintain high enough density susceptible host populations to sustain phage. To investigate this possibility we imposed a floor of $U > 1$ and ran another round of simulations. Even with very high rates of autoimmunity based on an upper limit of likely spacer acquisition rates ($\alpha = 50\mu_b$, $\mu_b = 10^{-5}$) the susceptible host population does not grow quickly enough to sufficiently high levels to sustain phage (S16 Fig). Thus it is not early dynamics that rule out autoimmunity but the insufficiency of the mechanism itself for maintaining large enough susceptible host populations.

### 3.2.2 Transient Coexistence with Low Density Phage

While we do not observe stable coexistence in any case where there is not loss of the CRISPR immune system, we did observe prolonged phage persistence in some cases where $\mu_L = \alpha = 0$ (Fig 3) and in cases with autoimmunity only ($\mu_L = 0$). Phages were able to persist at very low density ($\sim 10 - 100$ particles/mL) for as long as two months in a host-dominated setting without the presence of a CRISPR-lacking host sub-population (Fig 3, S17 Fig). It appears that in these cases phages are at sufficiently low density as to have a minimal effect on their host population and thus that host strain is selected against very slowly. Because the phages have undergone many PAM mutations at this point they are unable to proliferate rapidly enough between dilution events to have an easily measurable impact on the host population. Essentially, phages delay their collapse by consuming their host extremely slowly (S17 Fig). However, with an active locus loss mechanism (i.e., $\mu_L > 0$), we did not see this sustained but unstable coexistence occur, likely because the undefended phages would have driven the phage population to higher levels and increased selection on the susceptible CRISPR variants.

## 4 Discussion

We paired a general model of immunity with a case study of the CRISPR immune system to characterize and contrast the potential drivers of long-term host-phage coexistence in well-mixed systems. We found that, while both a cost of immunity and the loss of immunity can lead to stable coexistence, only a loss mechanism can do so robustly within a realistic region of parameter space. We were able to reject a cost imposed on immunity (i.e., a tradeoff) as a driver of coexistence based on the lack of robustness of the equilibria produced by this mechanism. This rejection calls into question the generality of tradeoff-based explanations of host-phage coexistence, which have often been presented as the primary cause of coexistence in well-mixed systems (e.g. [92]). The specific features of a given type of defense (e.g., intracellular or extracellular action) determine which mechanisms can plausibly lead to robust coexistence.

We showed that the loss of immunity can play an important role in determining coevolutionary trajectories and global population dynamics in host-parasite systems, even when only a small segment of the host population experiences these losses. Furthermore, this result, though derived from a simple, analytically tractable model, is robust to the initial conditions of the system and the addition of realistic coevolutionary dynamics. In fact, the addition of coevolutionary dynamics allowed us to reproduce additional patterns observed in our experiments, such as an early peak in the phage population and dip in the bacterial population before a transition to host dominance, as

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

well as stochastic switching between the possible outcomes of long term coexistence and rapid phage clearance. Our simulations reliably demonstrated a transition from an escalating arms race dynamic to a fluctuating selection dynamic, and finally to stable predator-prey oscillations without further coevolution. The first of these transitions has also been observed in previous experimental work [33].

Our experiments in the *S. thermophilus* system reject a sustained arms race dynamic, since spacers did not continue to accumulate rapidly over the long term. We also reject a fluctuating selection dynamic based on simulations that show this type of dynamic to be unstable in the long term. Sequencing of the CRISPR locus did not reveal pervasive spacer loss events. This supports our hypothesis that immune loss is at the system rather than spacer level so that susceptible host do not posses a functional CRISPR locus. While our experiments do not speak to the relative importance of locus loss versus autoimmunity in the maintenance of this susceptible populations, our theoretical results reject autoimmunity as a realistic mechanism of coexistence. Our experimental setup was in serial dilution, which effectively subjects the culture to large daily perturbations in population size, ruling out any mechanism that does not produce an extremely robust coexistence regime. While autoimmunity does not lead to robust host-phage coexistence, it is a common phenomenon across organisms possessing CRISPR immune systems and may have an important effect on the evolution of host immunity in the absence of phages [48–50, 79].

We note that while we explored the possibility of alternate immune tradeoff regimes (S1 Text), all of our results were derived in the context of an intracellular immune system (e.g., restriction-modification systems, CRISPR). In such systems phages are free to adsorb to defended host, and thus immunity causes phage death rather than simply preventing phage growth as in the case of extracellular resistance. Therefore, for us to observe coexistence, phages require a susceptible host population of high enough density to offset the death rate caused by immune hosts. The resulting threshold density is higher than would otherwise be needed to simply sustain phages in the absence of the adsorption-death term, which in turn increases the requisite rate of autoimmunity needed to maintain the required susceptible host population. This contributes to the finding that autoimmunity cannot account for coexistence in the systems we examine. In systems where resistance prevents phage adsorption, a resistance-growth tradeoff in the host will likely produce a more robust coexistence regime than observed here with tradeoffs. Regardless, in either type of system a loss mechanism should reliably produce host-phage coexistence.

It is not immediately clear why bacteria would lose the functionality of their immune systems at such a high rate. Perhaps in the case of CRISPR there is some inherent instability of the locus, leading to higher rates of horizontally transfer [85, 86, 93]. Jiang et al. [84] propose that CRISPR loss is a bet-hedging strategy that allows horizontal gene transfer to occur in stressful environments (e.g., under selection for antibiotic resistance). Similarly, loss could allow escape from the cost of immunity when phages are not present, although this seems unlikely in the case of CRISPR which is up-regulated when phages are detected and relatively dormant otherwise [94–96], and because such a strategy would lead to poor performance in environments of frequent but transient infections. We note that a high rate of CRISPR loss and inactivation could produce a pressure for bacteria to frequently acquire new CRISPR systems through horizontal gene transfer, perhaps explaining why strains with multiple redundant CRISPR systems are frequently observed [97, 98].

In the case of costly and constitutive forms of defense a loss mechanism may be the only route to increased fitness in the absence of phages, perhaps driving a high loss rate in some forms of innate defense. If resistance mechanisms have a certain error rate, as seen in some cases of receptor masking or modification, then stochastic failure can also

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

produce a susceptible host population on which phage can persist [46].

Our immune loss hypothesis predicts that the CRISPR immune system is lost at a nontrivial rate in *S. thermophilus* in addition to *S. epidermidis* [84], and possibly a range of other species. Competition experiments with phages that quantify the costs of PAM mutations, and whether or not they are cumulative, are needed to better understand the coevolutionary structure of CRISPR-phage systems in general. Other paths to sustained coexistence between CRISPR-enabled host and phages may also exist, notably anti-CRISPR phage proteins that allow phages to escape immune action similar to the methods of inactivation considered here [99, 100], though it is possible the presence of such proteins would simply lead to another level of host-phage arms race.

In order to verify that the loss of defense is generally relevant to host-phage systems and not a CRISPR-specific phenomenon we must observe cases of the loss of both immunity and resistance across diverse taxa. Some examples already exist for both cases, as mentioned earlier [46, 84], but in any example of host-phage coexistence it is important to consider loss of defense as a potential driver. Additionally, our work highlights that in order to invoke a tradeoff mechanism to explain the maintenance of coexistence it is necessary not only to show that a growth-immunity tradeoff exists, but that it is also sufficiently severe.

Finally, we would like to point out that both tradeoff and loss mechanisms essentially condition the dynamics of the host-phage system on the absence of immune system functionality in some segment of the population. Our results show that the regular loss of immunity can sustain a viable phage population, leading to the maintenance of selective pressure and thus keeping immunity prevalent in the population overall. Counterintuitively, this leads us to suggest that the periodic loss of immunity drives the maintenance of a high population immune prevalence. Thus conversations about host-phage coevolution, specifically those concerning CRISPR, cannot neglect the potential susceptible individuals in a population.

**Fig 1. Model behavior under variations in the rates of autoimmunity ($\alpha$) and CRISPR locus loss ($\mu$)** Equilibria (S1 Table) derived from Equations 1-4 are shown in (a) where orange indicates a stable equilibrium with all populations coexisting and defended host dominating phage populations, green indicates that all populations coexist but phages dominate, and blue indicates that defended bacteria have gone extinct but phages and undefended bacteria coexist. In (b) we find numerical solutions to the model at 80 days using realistic initial conditions more specific to the experimental setup ($R(0) = 350$, $D(0) = 10^6$, $U(0) = 100$, $P(0) = 10^6$). In this case orange indicates coexistence at 80 days with defended host at higher density than phages, green indicates a phage-dominated coexistence at 80 days, and blue indicates that coexistence did not occur. Numerical error is apparent as noise near the orange-blue boundary. We neglect coevolution and innate immunity in this analysis ($\phi_u = 1$, $\phi_d = 0$).

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Fig 2. Serial transfer experiments carried out with *S. thermophilus* and lytic phage 2972** Bacteria are resource-limited rather than phage-limited by day five and phages can either (a) persist at relatively low density in the system on long timescales (greater than 1 month) or (b) collapse relatively quickly. These results agree with those of Paez-Espino [25] where coexistence was observed in *S. thermophilus* and phage 2972 serially transferred culture for as long as a year. Experiments were initiated with identical starting populations and carried out following the same procedure. In (c-e) we show that our simulations replicate the qualitative patterns seen in the data, with an early phage peak, followed by host-dominated coexistence that can either be (c) stable, (d) sustained but unstable, or (e) short-lived. Each plot is a single representative simulation and simulations were ended when phages went extinct. Note that experimental data has a resolution of one time point per day, preventing conclusions about the underlying population dynamics (e.g., cycling), whereas simulations are continuous in time.

**Fig 3. Distribution of phage extinction times in bacterial-dominated cultures with different possible combinations of coexistence mechanisms** The peak at $\geq 75$ corresponds to what we call stable coexistence (simulations ran for a maximum of 80 days). There is no significant difference between the top two panels in the number of simulations reaching the 80 day mark ($\chi^2 = 2.8904$, $df = 1$, $p - \text{value} = 0.08911$). Back mutation was set at $\mu_q = 5 \times 10^{-9}$.

**Fig 4. Distribution of phage extinction times in bacterial-dominated cultures with different rates of PAM back mutation in phages ($\mu_q$)** The peak at 80 corresponds to what we call stable coexistence (simulations ran for a maximum of 80 days). These results are shown for a locus-loss mechanism only ($\mu_L = 5 \times 10^{-4}$, $\alpha = 0$). The histogram for $\mu_q = 5 \times 10^{-8}$ is omitted as it is nearly identical to that for $\mu_q = 5 \times 10^{-9}$, indicating that the height of the coexistence peak saturates at high back mutation.

**Fig 5. Representative example of a simulation demonstrating stable coexistence under a loss mechanism ($\mu_L = 5 \times 10^{-4}$, $\alpha = 0$, $\mu_q = 5 \times 10^{-9}$)** In (a) we show the archetypal shift from phage-dominance during an initial arms race to unstable host-dominated coexistence where fluctuating selection dynamics are observed to stable host-dominated predator-prey cycling of phages and CRISPR-lacking hosts as seen in (d) where evolution ceases to occur. In (b) we see that a drop in mean phage order leads to stable cycling and in (c) that this corresponds to a single phage strain becoming dominant after previous cycling of strains. This corresponds to a shift away from fluctuating selection dynamics. In (c) the colors specify different phage strains.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Supporting Information

**S1 Text. Alternative immune costs.**

**S2 Text. Parameter Values.**

**S3 Text. Analysis and Simulation Procedures.**

**S4 Text. Detailed experimental methods.**

**S1 Table. Equilibrium of analytical model without coevolution.** Equilibria for the autoimmunity/locus-loss model where $\mu > 0$ and $\alpha > 0$. Not all substitutions have been made here in the interest of readability, but equilibrium values can be easily computed using the above expressions. The stability of these equilibria can be assessed by linearizing our system around them (i.e., taking the Jacobian) as described in S3 Text.

**S1 Fig Equilibria with alternative costs of immunity.** Model behavior under variations in the immune system loss rate and (a) resource affinity coefficient or (b) growth rate penalty. Equilibria derived from our equations in S1 Text are shown where orange indicates a stable equilibrium with all populations coexisting and defended host dominating phage populations, green indicates that all populations coexist but phages dominate, light blue indicates that defended bacteria have gone extinct but phages and undefended bacteria coexist, and dark blue indicates that there is no stable equilibrium. We neglect coevolution and innate immunity in this analysis ($\phi_u = 1$, $\phi_d = 0$) and do not consider the effects of autoimmunity ($\alpha = 0$).

**S2 Fig Equilibria with each coexistence mechanism in isolation.** Behavior of coexistence equilibrium when (a) there is only CRISPR loss without autoimmunity, (b) there is only autoimmunity without CRISPR loss, (c) there is only a cost applied to resource affinity (S1 Text), and (d) there is only a cost applied to maximum growth rate (S1 Text). Notice that immune loss and autoimmune mechanisms essentially act in the same manner, except that the loss mechanism produces a larger phage population by flushing extra susceptible bacteria into the system. This is consistent with theoretical results showing that increasing resource availability in a host-phage system increases phage rather than host populations [101]. The upper bound of the $x$-axis in (a-d) represents the upper limit of the cost of immunity, above which coexistence will not occur because immune host cannot survive.

**S3 Fig Numerical solutions to model at 80 days with realistic initial conditions.** Numerical solutions to the alternative cost model (S1 Text) at 80 days using realistic initial conditions more specific to the experimental setup ($R(0) = 350$, $D(0) = 10^6$, $U(0) = 100$, $P(0) = 10^6$). Results only shown for cases in which all three populations remained extant.

**S4 Fig Simulations of perturbed starting conditions (small perturbations).** We find numerical solutions to the alternative cost model (S1 Text) at 80 days with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We ran 50 simulations for each condition. We let $\gamma = 0.1$.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**S5 Fig Simulations of perturbed starting conditions (intermediate perturbations).** We find numerical solutions to the alternative cost model (S1 Text) at 80 days with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We ran 50 simulations for each condition. We let $\gamma = 1$.

**S6 Fig Simulations of perturbed starting conditions (large perturbations).** We find numerical solutions to the alternative cost model (S1 Text) at 80 days with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We ran 50 simulations for each condition. We let $\gamma = 10$.

**S7 Fig Simulations of perturbed starting conditions (very large perturbations).** We find numerical solutions to the alternative cost model (S1 Text) at 80 days with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We ran 50 simulations for each condition. We let $\gamma = 100$.

**S8 Fig Mean population size with perturbed starting conditions (intermediate perturbations).** We find numerical solutions to the alternative cost model (S1 Text) at 80 days with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We ran 50 simulations for each condition. We let $\gamma = 10$. Mean population across all simulations (including cases of phage or host extinction) shown by bold line and two standard deviations away from the mean are represented by the thin lines.

**S9 Fig Phase diagram with perturbed starting conditions.** Numerical simulations with starting conditions ($X(0) = [R(0), D(0), U(0), P(0)]$) perturbed by a proportion of the equilibrium condition $X(0) = \tilde{X}(1 + \gamma Y)$ where $Y \sim U[0, 1]$ and $\tilde{X}$ signifies an equilibrium value to explore how robust the equilibria are to starting conditions. We neglect coevolution and innate immunity in this analysis ($\phi_u = 1$, $\phi_d = 0$). A single simulation was run for each parameter combination.

**S10 Fig Phase diagram of analytical model with phage coevolution.** Phase diagrams of simple coevolutionary model behavior under variations in the rates of autoimmunity ($\alpha$) and CRISPR locus loss ($\mu$) over various coevolutionary scenarios ($\phi_d$). Values of $\phi_d$ were chosen so as to demonstrate the rapid shift that occurs from host to phage dominated equilibrium as the infected fraction of defended host increases. Orange indicates a stable equilibrium with all populations coexisting and defended host dominating phage populations, green indicates that all populations coexist but phages dominate, and blue indicates that defended bacteria have gone extinct but phages and undefended bacteria coexist.

**S11 Fig Phase diagram of analytical model with innate immunity.** Phase diagram of model behavior under variations in the rates of autoimmunity ($\alpha$) and CRISPR locus loss ($\mu$) for different values of ($\phi_u$). Orange indicates a stable equilibrium with all populations coexisting and defended host dominating phage populations, green

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

indicates that all populations coexist but phages dominate, and blue indicates that defended bacteria have gone extinct but phages and undefended bacteria coexist.

**S12 Fig Replicate serial transfer experiments.** Densities of (a) phage and (b) bacteria measured daily at serial transfer. All replicate epxeriments start with the same conditions and strains as in the main text.

**S13 Fig Mean sequenced order of host over time in serial transfer experiments 1 and 2.**

**S14 Fig Optimal host order for phages to infect over time.** The optimal host strain that is either currently infected by a phage strain or one PAM mutation away from being infected. Optimality is defined in terms of population size times the burst size of the phage strain that does or could infect that strain, so that the balance between abundant host and mutation cost is taken into account. In (a) we track the order of the "best" available host strain at any given point in a single example simulation (Fig 5), and in (b) look at the timing of the peak optimal order across 100 simulations (Fig 4, $\mu_q = 5 \times 10^{-9}$). Note that after the initial arms race dynamic the best available host strain is the CRISPR-lacking host in all simulations. The order of the best host strain peaks early on in all simulations and then drops to zero (CRISPR-lacking), signifying an early end to the arms race between host and phage.

**S15 Fig Effect on simulations of varied phage adsorption rates.** Adsorption rate has a profound effect on the outcome of host-phage interactions. At a high adsorption rate ($\delta = 10^{-7}$) either phages or bacteria tend to go extinct early on in the simulation, and phages are able to drive their host extinct approximately 50% of the time (49/100 simulations for all back-mutation rates). We see a reversed relationship of time to extinction with $\mu_q$ from our base adsorption rate ($\delta = 10^{-8}$, Fig 4), although in general at a very high adsorption rate few simulations demonstrate long-term coexistence as phage consume all host early on. This suggests that the costs associated with PAM mutations are required to keep phage growth rate low enough to prevent overconsumption of host, and indeed upon closer examination of individual simulations it is clear that back mutations to lower phage orders precipitate phage collapse. In the lowest panel we demonstrate that coexistence in the long term at high $\delta$ is associated with a high mean phage order over the course of a simulation, while the opposite is true of our typical intermediate $\delta$. At a low adsorption rate ($\delta = 10^{-9}$) we see populations coexisting until the 80 day mark (max simulation length) in almost all simulations.

**S16 Fig Representative simulation with a floor on the susceptible host population and high autoimmunity .** We let $B_s > 1 \, \forall t$ and $\alpha = 5 \times 10^{-4}$.

**S17 Fig Transient phage survival at low density** Example of low-level phage persistence due to slow evolutionary dynamics. Here we see that (a) in the absence of the constant production of susceptible bacteria by CRISPR-enabled strains (i.e., $\mu = 0$) phages are still able to paradoxically persist despite a clear advantage to bacteria in the arms race and an absence of other sustaining mechanisms. In (b) a small fraction of the CRISPR enabled bacterial population is maintained that lacks spacers towards the infecting phages and in (c) we zoom in to show that this population is declining due to this infection, but extremely slowly, implying this coexistence is not stable in the long term. The number of bacterial strains that can be infected by phages over time is shown in (d), and (e) shows how the richness of phage and bacterial strains changes over time. Note that although the number of bacterial strains increases asymptotically, after an

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

initial spike the number of strains that can be infected by phages drops dramatically to the single digits. This keeps the overall phage population growth constrained (balanced by adsorption to immune bacteria). In fact, over time phage act to suppress their own growth by negatively infecting the competitiveness of their host (although this effect is so small that phages can persist for an extended period of time seemingly stably). What looks flat is actually monotonically decreasing. All parameters as in Table **??** and $\alpha = 0$.

**S18 Fig Effect of changes in PAM mutation cost ($c$).** Distribution of phage extinction times in bacterial-dominated cultures with different costs on PAM mutation in phage ($c$). The peak at 80 corresponds to stable coexistence (simulations ran for a maximum of 80 days). These results of for a locus-loss mechanism only ($\mu_L = 5 \times 10^{-4}$, $\alpha = 0$).

**S19 Fig Phase diagram of analytical model with phage coevolution.** Phase diagrams of model behavior without coevolution or other forms of immunity ($\phi_d = 0$, $\phi_u = 1$) under variations in the rates of autoimmunity ($\alpha$) and CRISPR locus loss ($\mu$) over various adsorption rates ($\delta$). Orange indicates a stable equilibrium with all populations coexisting and defended host dominating phage populations, green indicates that all populations coexist but phages dominate, and blue indicates that defended bacteria have gone extinct but phages and undefended bacteria coexist. There is an apparent increase in the area of the coexistence region in which host dominate as adsorption rate increases.

**S20 Fig Equilibrium phage population during coexistence.** Equilibrium (analytical) population of phages when there is full coexistence over a range of $\alpha$ and $\mu_L$ values for our simple mathematical model without coevolution ($\phi_u = 1$, $\phi_d = 0$).

**S21 Fig Distribution of phage extinction times in bacterial-dominated cultures with an MOI of 10**. The peak at 80 corresponds to what we call stable coexistence (simulations ran for a maximum of 80 days).

**S22 Fig Distributions of phage extinction times in bacterial-dominated cultures with various burst sizes**. The peak at 80 corresponds to what we call stable coexistence (simulations ran for a maximum of 80 days).

## Acknowledgments

This work was supported by funding from the University of Maryland and the U.S. Department of Education GAANN program. This material is based upon work supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number #W911NF-14-1-0490. SM acknowledges funding from the Natural Sciences and Engineering Research Council of Canada (Discovery program). SM holds a T1 Canada Research Chair in Bacteriophages. PLFJ was supported in part by NIH R00 GM104158.

## References

1. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. Proceedings of the National Academy of Sciences. 1998;95(12):6578–6583.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the tree of life. Nature Microbiology. 2016;1:16048. doi:10.1038/nmicrobiol.2016.48.

3. Schloss PD, Girard RA, Martin T, Edwards J, Thrash JC. Status of the Archaeal and Bacterial Census: an Update. mBio. 2016;7(3):e00201–16. doi:10.1128/mBio.00201-16.

4. Wilhelm SW, Suttle CA. Viruses and Nutrient Cycles in the SeaViruses play critical roles in the structure and function of aquatic food webs. BioScience. 1999;49(10):781–788. doi:10.2307/1313569.

5. Wommack KE, Colwell RR. Virioplankton: Viruses in Aquatic Ecosystems. Microbiology and Molecular Biology Reviews. 2000;64(1):69–114.

6. Suttle CA. Viruses in the sea. Nature. 2005;437(7057):356–361. doi:10.1038/nature04160.

7. Weitz JS, Wilhelm SW. Ocean viruses and their effects on microbial communities and biogeochemical cycles. F1000 Biol Rep. 2012;4. doi:10.3410/B4-17.

8. Wigington CH, Sonderegger D, Brussaard CPD, Buchan A, Finke JF, Fuhrman JA, et al. Re-examination of the relationship between marine virus and microbial cell abundances. Nature Microbiology. 2016;1:15024. doi:10.1038/nmicrobiol.2015.24.

9. Bergh Ø, BØrsheim KY, Bratbak G, Heldal M. High abundance of viruses found in aquatic environments. Nature. 1989;340(6233):467–468. doi:10.1038/340467a0.

10. Sieburth JM, Johnson PW, Hargraves PE. Ultrastructure and Ecology of Aureococcus Anophageferens Gen. Et Sp. Nov. (chrysophyceae): The Dominant Picoplankter During a Bloom in Narragansett Bay, Rhode Island, Summer 19851. Journal of Phycology. 1988;24(3):416–425. doi:10.1111/j.1529-8817.1988.tb04485.x.

11. Proctor LM, Fuhrman JA. Viral mortality of marine bacteria and cyanobacteria. Nature. 1990;343(6253):60–62. doi:10.1038/343060a0.

12. Bratbak G, Heldal M, Norland S, Thingstad TF. Viruses as Partners in Spring Bloom Microbial Trophodynamics. Appl Environ Microbiol. 1990;56(5):1400–1405.

13. Bratbak G, Thingstad F, Heldal M. Viruses and the Microbial Loop. Microbial Ecology. 1994;28(2):209–221.

14. Weinbauer MG, Rassoulzadegan F. Are viruses driving microbial diversification and diversity? Environmental Microbiology. 2004;6(1):1–11. doi:10.1046/j.1462-2920.2003.00539.x.

15. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, et al. Origins of Highly Mosaic Mycobacteriophage Genomes. Cell. 2003;113(2):171–182. doi:10.1016/S0092-8674(03)00233-2.

16. Waterbury JB, Valois FW. Resistance to Co-Occurring Phages Enables Marine Synechococcus Communities To Coexist with Cyanophages Abundant in Seawater. Appl Environ Microbiol. 1993;59(10):3393–3399.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

17. Gómez P, Buckling A. Bacteria-Phage Antagonistic Coevolution in Soil. Science. 2011;332(6025):106–109. doi:10.1126/science.1198767.

18. Horne MT. Coevolution of Escherichia coli and Bacteriophages in Chemostat Culture. Science. 1970;168(3934):992–993.

19. Levin SA, Udovic JD. A Mathematical Model of Coevolving Populations. The American Naturalist. 1977;111(980):657–675.

20. Chao L, Levin BR, Stewart FM. A Complex Community in a Simple Habitat: An Experimental Study with Bacteria and Phage. Ecology. 1977;58(2):369–378. doi:10.2307/1935611.

21. Lenski RE, Levin BR. Constraints on the Coevolution of Bacteria and Virulent Phage: A Model, Some Experiments, and Predictions for Natural Communities. The American Naturalist. 1985;125(4):585–602.

22. Schrag SJ, Mittler JE. Host-Parasite Coexistence: The Role of Spatial Refuges in Stabilizing Bacteria-Phage Interactions. The American Naturalist. 1996;148(2):348–377.

23. Bohannan BJM, Lenski RE, Holt AERD. Effect of Prey Heterogeneity on the Response of a Model Food Chain to Resource Enrichment. The American Naturalist. 1999;153(1):73–82. doi:10.1086/303151.

24. Wei Y, Kirby A, Levin BR, Rohani AEP, Bronstein EJL. The Population and Evolutionary Dynamics of Vibrio cholerae and Its Bacteriophage: Conditions for Maintaining Phage-Limited Communities. The American Naturalist. 2011;178(6):715–725. doi:10.1086/662677.

25. Paez-Espino D, Sharon I, Morovic W, Stahl B, Thomas BC, Barrangou R, et al. CRISPR Immunity Drives Rapid Phage Genome Evolution in Streptococcus thermophilus. mBio. 2015;6(2):e00262–15. doi:10.1128/mBio.00262-15.

26. Rodin SN, Ratner VA. Some theoretical aspects of protein coevolution in the ecosystem “phage-bacteria” I. The problem. Journal of Theoretical Biology. 1983;100(2):185–195. doi:10.1016/0022-5193(83)90346-6.

27. Rodin SN, Ratner VA. Some theoretical aspects of protein coevolution in the ecosystem “phage-bacteria” II. The deterministic model of microevolution. Journal of Theoretical Biology. 1983;100(2):197–210. doi:10.1016/0022-5193(83)90347-8.

28. Van Valen L. A new evolutionary law. Evolutionary Theory. 1973;1:1–30.

29. Van Valen L. Molecular evolution as predicted by natural selection. J Mol Evol. 1974;3(2):89–101. doi:10.1007/BF01796554.

30. Dawkins R, Krebs JR. Arms Races between and within Species. Proceedings of the Royal Society of London Series B, Biological Sciences. 1979;205(1161):489–511.

31. Levin BR, Stewart FM, Chao L. Resource-Limited Growth, Competition, and Predation: A Model and Experimental Studies with Bacteria and Bacteriophage. The American Naturalist. 1977;111(977):3–24.

32. Lenski RE. Coevolution of bacteria and phage: Are there endless cycles of bacterial defenses and phage counterdefenses? Journal of Theoretical Biology. 1984;108(3):319–325. doi:10.1016/S0022-5193(84)80035-1.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.


33. Hall AR, Scanlan PD, Morgan AD, Buckling A. Host–parasite coevolutionary arms races give way to fluctuating selection. Ecology Letters. 2011;14(7):635–642. doi:10.1111/j.1461-0248.2011.01624.x.

34. Agrawal* A, Lively CM. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. Evol Ecol Res. 2002;4(1):91–107.

35. Gandon S, Buckling A, Decaestecker E, Day T. Host–parasite coevolution and patterns of adaptation across time and space. Journal of Evolutionary Biology. 2008;21(6):1861–1866. doi:10.1111/j.1420-9101.2008.01598.x.

36. Buckling A, Rainey PB. Antagonistic coevolution between a bacterium and a bacteriophage. Proceedings of the Royal Society of London B: Biological Sciences. 2002;269(1494):931–936. doi:10.1098/rspb.2001.1945.

37. Hassell MP, Comins HN, Mayt RM. Spatial structure and chaos in insect population dynamics. Nature. 1991;353(6341):255–258. doi:10.1038/353255a0.

38. Haerter JO, Sneppen K. Spatial Structure and Lamarckian Adaptation Explain Extreme Genetic Diversity at CRISPR Locus. mBio. 2012;3(4):e00126–12. doi:10.1128/mBio.00126-12.

39. Thompson JN. Coevolution: The Geographic Mosaic of Coevolutionary Arms Races. Current Biology. 2005;15(24):R992–R994. doi:10.1016/j.cub.2005.11.046.

40. Jover LF, Cortez MH, Weitz JS. Mechanisms of multi-strain coexistence in host–phage systems with nested infection networks. Journal of Theoretical Biology. 2013;332:65–77. doi:10.1016/j.jtbi.2013.04.011.

41. Meyer JR, Dobias DT, Medina SJ, Servilio L, Gupta A, Lenski RE. Ecological speciation of bacteriophage lambda in allopatry and sympatry. Science. 2016;354(6317):1301–1304. doi:10.1126/science.aai8446.

42. Yoshida T, Ellner SP, Jones LE, Bohannan BJM, Lenski RE, Jr NGH. Cryptic Population Dynamics: Rapid Evolution Masks Trophic Interactions. PLOS Biol. 2007;5(9):e235. doi:10.1371/journal.pbio.0050235.

43. Jones LE, Ellner SP. Effects of rapid prey evolution on predator–prey cycles. J Math Biol. 2007;55(4):541–573. doi:10.1007/s00285-007-0094-6.

44. Delbrück M. Bacterial Viruses or Bacteriophages. Biological Reviews. 1946;21(1):30–40. doi:10.1111/j.1469-185X.1946.tb00451.x.

45. Lenski RE. Dynamics of Interactions between Bacteria and Virulent Bacteriophage. In: Marshall KC, editor. Advances in Microbial Ecology. No. 10 in Advances in Microbial Ecology. Springer US; 1988. p. 1–44. Available from: `http://link.springer.com/chapter/10.1007/978-1-4684-5409-3_1`.

46. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science. 2012;335(6067):428–432. doi:10.1126/science.1214449.

47. Weitz JS. Quantitative Viral Ecology: Dynamics of Viruses and Their Microbial Hosts. Princeton University Press; 2016.

48. Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, Clulow JS, et al. Cytotoxic Chromosomal Targeting by CRISPR/Cas Systems Can Reshape Bacterial Genomes and Expel or Remodel Pathogenicity Islands. PLOS Genet. 2013;9(4):e1003454. doi:10.1371/journal.pgen.1003454.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

49. Dy RL, Pitman AR, Fineran PC. Chromosomal targeting by CRISPR-Cas systems can contribute to genome plasticity in bacteria. Mobile Genetic Elements. 2013;3(5):e26831. doi:10.4161/mge.26831.

50. Paez-Espino D, Morovic W, Sun CL, Thomas BC, Ueda Ki, Stahl B, et al. Strong bias in the bacterial CRISPR elements that confer immunity to phage. Nat Commun. 2013;4:1430. doi:10.1038/ncomms2440.

51. Luria SE, Human ML. A NONHEREDITARY, HOST-INDUCED VARIATION OF BACTERIAL VIRUSES1. J Bacteriol. 1952;64(4):557–569.

52. Dussoix D, Arber W. Host specificity of DNA produced by Escherichia coli. Journal of Molecular Biology. 1962;5(1):37–49. doi:10.1016/S0022-2836(62)80059-X.

53. Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV. Evolutionary Dynamics of the Prokaryotic Adaptive Immunity System CRISPR-Cas in an Explicit Ecological Context. J Bacteriol. 2013;195(17):3834–3844. doi:10.1128/JB.00412-13.

54. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J Mol Evol. 2005;60(2):174–182. doi:10.1007/s00239-004-0046-3.

55. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151(8):2551–2561. doi:10.1099/mic.0.28048-0.

56. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science. 2007;315(5819):1709–1712. doi:10.1126/science.1138140.

57. Garneau JE, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67–71. doi:10.1038/nature09523.

58. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage Response to CRISPR-Encoded Resistance in Streptococcus thermophilus. J Bacteriol. 2008;190(4):1390–1400. doi:10.1128/JB.01412-07.

59. Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, et al. Diversity, Activity, and Evolution of CRISPR Loci in Streptococcus thermophilus. J Bacteriol. 2008;190(4):1401–1412. doi:10.1128/JB.01415-07.

60. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biology Direct. 2006;1:7. doi:10.1186/1745-6150-1-7.

61. Haft DH, Selengut J, Mongodin EF, Nelson KE. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. PLoS Comput Biol. 2005;1(6):e60. doi:10.1371/journal.pcbi.0010060.

62. Zhang Q, Ye Y. Not all predicted CRISPR–Cas systems are equal: isolated cas genes and classes of CRISPR like elements. BMC Bioinformatics. 2017;18:92. doi:10.1186/s12859-017-1512-4.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

63. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155(3):733–740. doi:10.1099/mic.0.023960-0.

64. Shah SA, Erdmann S, Mojica FJM, Garrett RA. Protospacer recognition motifs. RNA Biology. 2013;10(5):891–899. doi:10.4161/rna.23764.

65. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature. 2010;463(7280):568–571. doi:10.1038/nature08703.

66. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. PNAS. 2011;108(25):10098–10103. doi:10.1073/pnas.1104144108.

67. Andersson AF, Banfield JF. Virus Population Dynamics and Acquired Virus Resistance in Natural Microbial Communities. Science. 2008;320(5879):1047–1050. doi:10.1126/science.1157358.

68. Childs LM, Held NL, Young MJ, Whitaker RJ, Weitz JS. Multiscale Model of Crispr-Induced Coevolutionary Dynamics: Diversification at the Interface of Lamarck and Darwin. Evolution. 2012;66(7):2015–2029. doi:10.1111/j.1558-5646.2012.01595.x.

69. Levin BR, Moineau S, Bushman M, Barrangou R. The Population and Evolutionary Dynamics of Phage and Bacteria with CRISPR–Mediated Immunity. PLoS Genet. 2013;9(3):e1003312. doi:10.1371/journal.pgen.1003312.

70. Ehrlich PR, Raven PH. Butterflies and Plants: A Study in Coevolution. Evolution. 1964;18(4):586–608. doi:10.2307/2406212.

71. Hanifin CT, Jr EDB, Iii EDB. Phenotypic Mismatches Reveal Escape from Arms-Race Coevolution. PLOS Biol. 2008;6(3):e60. doi:10.1371/journal.pbio.0060060.

72. Luria SE. Mutations of Bacterial Viruses Affecting Their Host Range. Genetics. 1945;30(1):84–99.

73. Gerrish PJ, Lenski RE. The fate of competing beneficial mutations in an asexual population. Genetica. 1998;102-103(0):127. doi:10.1023/A:1017067816551.

74. Desai MM, Fisher DS, Murray AW. The Speed of Evolution and Maintenance of Variation in Asexual Populations. Current Biology. 2007;17(5):385–394. doi:10.1016/j.cub.2007.01.072.

75. Vale PF, Lafforgue G, Gatchitch F, Gardan R, Moineau S, Gandon S. Costs of CRISPR-Cas-mediated resistance in Streptococcus thermophilus. Proc R Soc B. 2015;282(1812):20151270. doi:10.1098/rspb.2015.1270.

76. Haerter JO, Trusina A, Sneppen K. Targeted Bacterial Immunity Buffers Phage Diversity. J Virol. 2011;85(20):10554–10560. doi:10.1128/JVI.05222-11.

77. Westra ER, van Houte S, Oyesiku-Blakemore S, Makin B, Broniewski JM, Best A, et al. Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. Current Biology. 2015;25(8):1043–1049. doi:10.1016/j.cub.2015.01.065.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

78. Chabas H, van Houte S, Høyland-Kroghsbo NM, Buckling A, Westra ER. Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. Proc Biol Sci. 2016;283(1837). doi:10.1098/rspb.2016.0721.

79. Wei Y, Terns RM, Terns MP. Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. Genes Dev. 2015;29(4):356–361. doi:10.1101/gad.257550.114.

80. Kumar MS, Plotkin JB, Hannenhalli S. Regulated CRISPR Modules Exploit a Dual Defense Strategy of Restriction and Abortive Infection in a Model of Prokaryote-Phage Coevolution. PLoS Comput Biol. 2015;11(11):e1004603. doi:10.1371/journal.pcbi.1004603.

81. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucl Acids Res. 2012; p. gks216. doi:10.1093/nar/gks216.

82. Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature. 2015;520(7548):505–510. doi:10.1038/nature14302.

83. Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. Self-targeting by CRISPR: gene regulation or autoimmunity? Trends in Genetics. 2010;26(8):335–340. doi:10.1016/j.tig.2010.05.008.

84. Jiang W, Maniv I, Arain F, Wang Y, Levin BR, Marraffini LA. Dealing with the Evolutionary Downside of CRISPR Immunity: Bacteria and Beneficial Plasmids. PLoS Genet. 2013;9(9):e1003844. doi:10.1371/journal.pgen.1003844.

85. Garrett RA, Shah SA, Vestergaard G, Deng L, Gudbergsdottir S, Kenchappa CS, et al. CRISPR-based immune systems of the Sulfolobales: complexity and diversity. Biochemical Society Transactions. 2011;39(1):51–57. doi:10.1042/BST0390051.

86. Palmer KL, Gilmore MS. Multidrug-Resistant Enterococci Lack CRISPR-cas. mBio. 2010;1(4):e00227–10. doi:10.1128/mBio.00227-10.

87. Childs LM, England WE, Young MJ, Weitz JS, Whitaker RJ. CRISPR-Induced Distributed Immunity in Microbial Populations. PLoS ONE. 2014;9(7):e101710. doi:10.1371/journal.pone.0101710.

88. Weinberger AD, Wolf YI, Lobkovsky AE, Gilmore MS, Koonin EV. Viral Diversity Threshold for Adaptive Immunity in Prokaryotes. mBio. 2012;3(6):e00456–12. doi:10.1128/mBio.00456-12.

89. Ferris MT, Joyce P, Burch CL. High Frequency of Mutations That Expand the Host Range of an RNA Virus. Genetics. 2007;176(2):1013–1022. doi:10.1534/genetics.106.064634.

90. Duffy S, Turner PE, Burch CL. Pleiotropic Costs of Niche Expansion in the RNA Bacteriophage Φ6. Genetics. 2006;172(2):751–757. doi:10.1534/genetics.105.051136.

91. Chao L. Fitness of RNA virus decreased by Muller's ratchet. Nature. 1990;348(6300):454–455. doi:10.1038/348454a0.

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

92. Bohannan BJM, Lenski RE. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecology Letters. 2000;3(4):362–377. doi:10.1046/j.1461-0248.2000.00161.x.

93. Shah SA, Garrett RA. CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems. Research in Microbiology. 2011;162(1):27–38. doi:10.1016/j.resmic.2010.09.001.

94. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. Transcription Profile of Thermus thermophilus CRISPR Systems after Phage Infection. Journal of Molecular Biology. 2010;395(2):270–281. doi:10.1016/j.jmb.2009.10.057.

95. Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, Shah M, et al. Phage-Induced Expression of CRISPR-Associated Proteins Is Revealed by Shotgun Proteomics in Streptococcus thermophilus. PLoS ONE. 2012;7(5):e38077. doi:10.1371/journal.pone.0038077.

96. Quax TEF, Prangishvili D, Voet M, Sismeiro O, Dillies MA, Jagla B, et al. Massive activation of archaeal defense genes during viral infection. Journal of Virology. 2013;87(15):8419–8428. doi:10.1128/JVI.01020-13.

97. Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol. 2009;131(1):62–70. doi:10.1016/j.ijfoodmicro.2008.05.030.

98. Cai F, Axen SD, Kerfeld CA. Evidence for the widespread distribution of CRISPR-Cas system in the Phylum Cyanobacteria. RNA Biology. 2013;10(5):687–693. doi:10.4161/rna.24571.

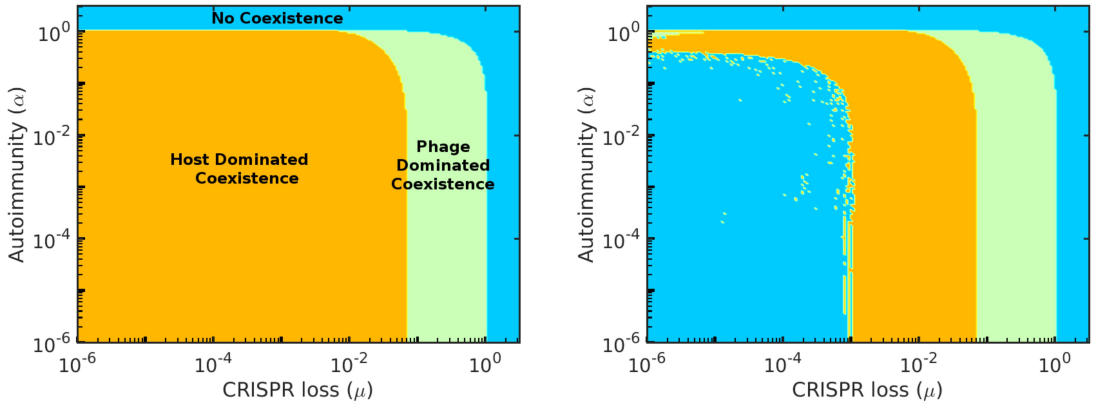
99. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 2013;493(7432):429–432. doi:10.1038/nature11723.

100. Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, et al. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature. 2015;526(7571):136–139. doi:10.1038/nature15254.

101. Bohannan BJM, Lenski RE. Effect of Resource Enrichment on a Chemostat Community of Bacteria and Bacteriophage. Ecology. 1997;78(8):2303–2315. doi:10.1890/0012-9658(1997)078[2303:EOREOA]2.0.CO;2.

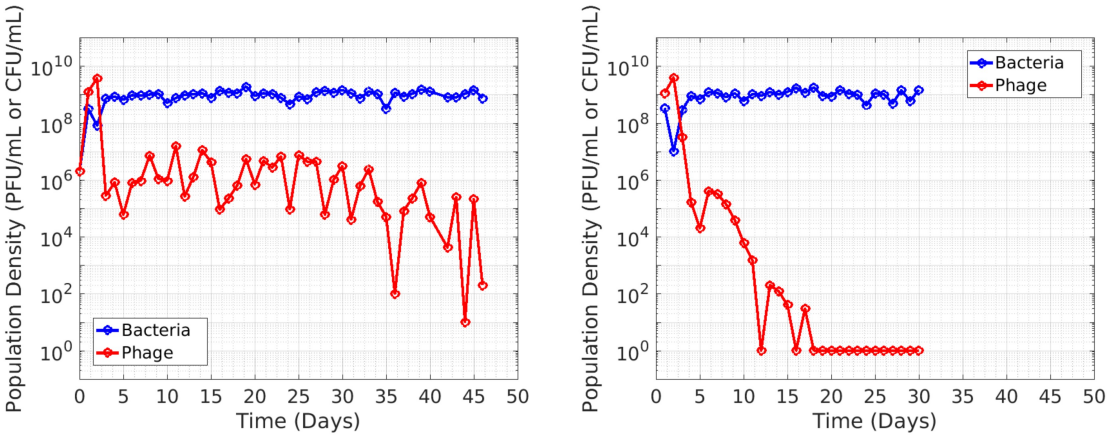
bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig 1.

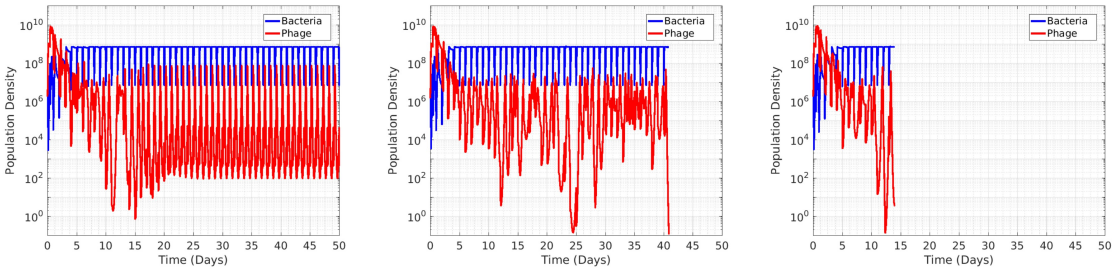



(a) (b)

Fig 2.




(a) (b)




(c) (d) (e)

bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig 3.

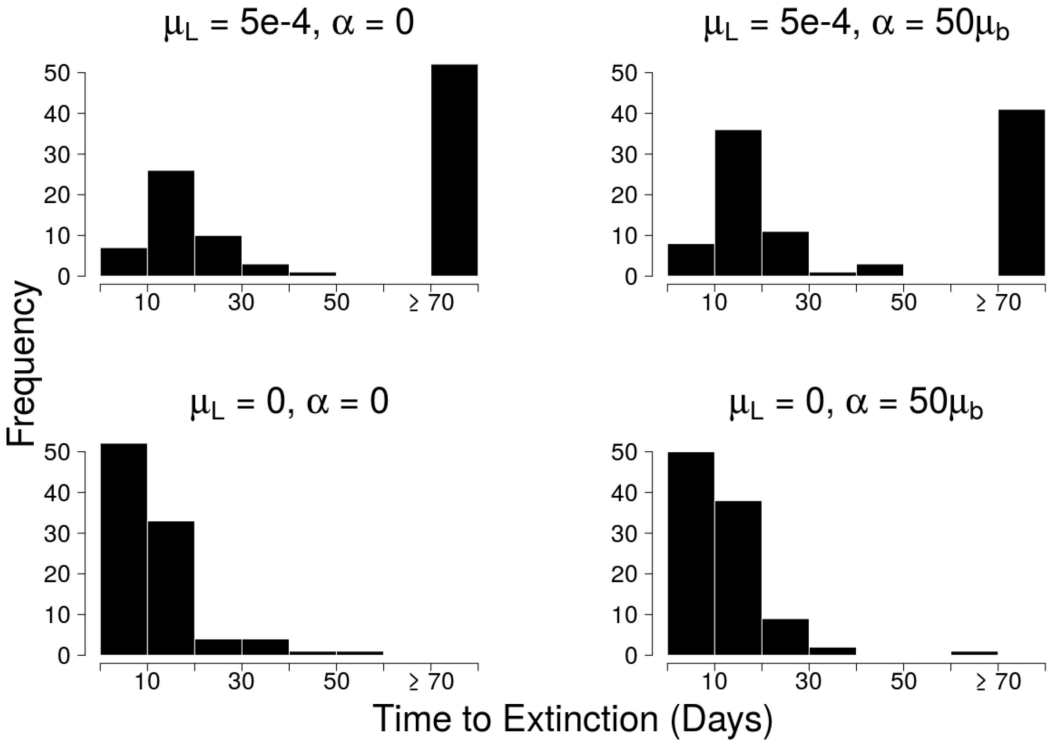
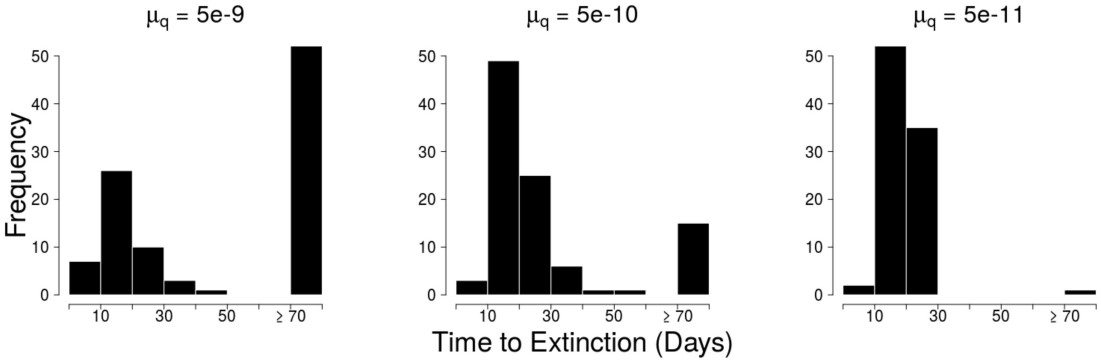



Fig 4.



bioRxiv preprint doi: https://doi.org/10.1101/105908; this version posted February 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

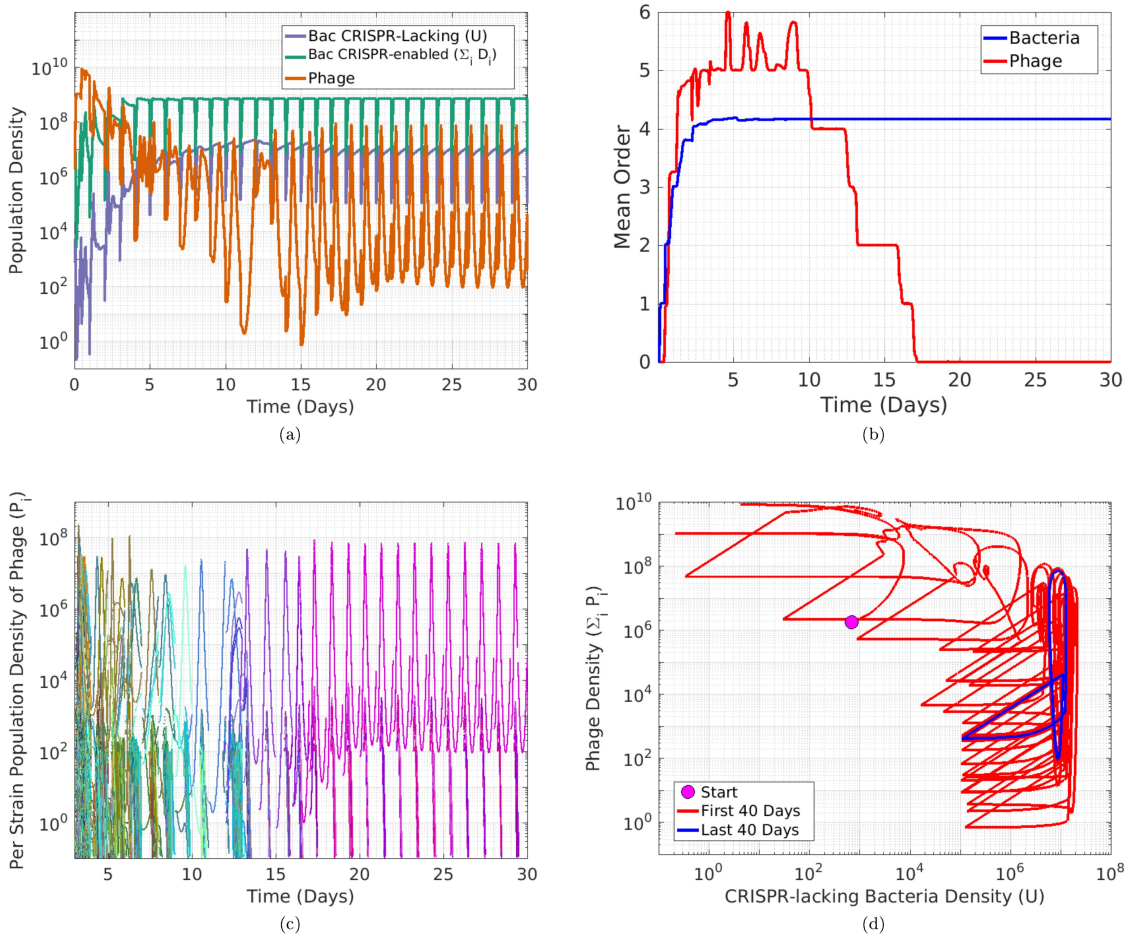



Fig 5.