How important is CRISPR-Cas for protecting natural populations of bacteria against infections with badass DNAs?

Edze Westra[1], Bruce Levin[2]

Authors contributed equally

[1]Environment and Sustainability Institute, Biosciences, University of Exeter,

Cornwall Campus, Penryn, Cornwall TR10 9FE, United Kingdom

[2] Department of Biology, Emory University, Atlanta, GA 30307, USA.

Abstract

Articles on CRISPR commonly open with some variant of the phrase 'these short-palindromic repeats and their associated endonucleases (Cas) are an adaptive immune system that exists to protect bacteria and archaea from viruses and infections with other deleterious ("badass") DNAs'. There is an abundance of genomic data consistent with the hypothesis that CRISPR plays this role in natural populations of bacteria and archaea, and experimental demonstrations with a few species of bacteria and their phage and plasmids show that CRISPR-Cas systems can play this role *in vitro*. Not at all clear are the ubiquity, magnitude and nature of the contribution of CRISPR-Cas systems to the ecology and evolution of natural populations of microbes, and the strength of selection mediated by different types of phage and plasmids to the evolution and maintenance of CRISPR-Cas systems. In this perspective, with the aid of heuristic mathematical-computer simulation models, we explore the *a priori* conditions under which exposure to lytic and temperate phage and conjugative plasmids will select for and maintain CRISPR-Cas systems in populations of bacteria and archaea. We review the existing literature addressing these ecological and evolutionary questions and highlight the experimental and other evidence needed to fully understand the conditions responsible for the evolution and maintenance of CRISPR-Cas systems and the contribution of these systems to the ecology and evolution of bacteria, archaea and the mobile genetic elements that infect them.

Significance

There is no question about the importance and utility of CRISPR-Cas for editing and modifying genomes. On the other hand, the mechanisms responsible for the evolution and maintenance of these systems and the magnitude of their importance to the ecology and evolution of bacteria, archaea and their infectious DNAs, are not at all clear. With the aid of heuristic mathematical/simulation models and reviews of the existing literature, we raise questions that have to be answered to elucidate the contribution of selection – mediated by phage and plasmids – to the evolution and maintenance of this adaptive immune system and its consequences for the ecology and evolution of prokaryotes and their viruses and plasmids.

Introduction

In 1987, a study by Ishino and colleagues aimed at analyzing the nucleotide sequence of the *iap* gene (isozyme-converting alkaline phosphatase) in *Escherichia coli* serendipitously led to the first-ever description of a CRISPR array. Specifically, they identified 14 repetitive sequences of 29 base pairs each at the 3' end of the *iap* gene, that were interspersed by 32–33 bp variable sequences (1, 2). Over the next years, identification of CRISPR arrays in other Gram-negative bacteria, Gram-positives and in Archaea (3-7) triggered a quest to identify their biological function (reviewed in (8)). Meanwhile, as more whole-genome sequences became available and CRISPR detection algorithms were developed, it became clear that these arrays of repeating sequences are very common in prokaryotes, with estimated frequencies of approximately 30-40% in Bacteria and 90% in the Archaea (9-13). An important step in understanding the function of CRISPRs was the identification of so-called *cas* genes (for CRISPR-associated) that are often found in the neighborhood of CRISPR arrays (14). Bioinformatics analyses expanded the known repertoire of *cas* genes (11, 15-18), shed light on their evolutionary origins (19, 20) and led to a comprehensive classification of *cas* gene combinations into two classes, and an increasing number of types and subtypes of CRISPR-Cas systems that differ in key mechanistic aspects (12).

The original idea that CRISPR-Cas is an adaptive immune system came from observations that sequences in CRISPR arrays on the host chromosome match those of mobile genetic elements (10, 21, 22). *Cas* genes were – based on their domains and predicted catalytic activities – suggested to encode the protein machinery that carries out the various steps of the immune response (11). The first experimental evidence in support of this hypothesis came from a joint effort by industrial and academic partners, who showed that the lactic acid bacterium *Streptococcus thermophilus* acquired post-infection resistance (hereafter immunity) against viral infection by inserting sequences from the viral genome into CRISPR loci on the bacterial genome (23). Viruses in turn were found to overcome CRISPR-based immunity by mutation of the target sequence on their genome (24). Together, these findings fueled models of how bacteria with CRISPR systems and their viruses might coevolve (25-27). In parallel with this experimental work, genomic data suggested that CRISPR loci evolve rapidly in natural populations of acidophilic bacteria (28), and that the DNA sequences between these palindromic repeats, spacers, were homologous to that of phage, plasmids and transposons (29).

Since these pioneering studies, spacer uptake from phages and other mobile genetic elements in bacteria and archaea from natural and human-associated environments has been inferred from variation in spacer sequences within and between populations of the same species and from their homology to mobile genetic elements (MGE) genomes (30-39). Experimental observations of spacer uptake in the lab in response to plasmid and phage infection have been made amongst others in engineered *E. coli* strains (40-42) and *Staphylococcus aureus* (43-46) and in wild type *Pectobacterium atrosepticum* (47), *Pseudomonas aeruginosa* (48, 49), *Roseburia intestinalis* (50), *Sulfolobus solfataricus* (51), *Streptococcus mutans* (36) and other species (reviewed in (52)). Consistent with the hypothesis that CRISPR-Cas protects bacteria from infections mobile genetic elements that encode so-called anti-CRISPR genes (reviewed in (53)).

How important is CRISPR to the ecology and evolution of bacteria and archaea?

While the preceding evidence and arguments demonstrate that CRISPR-Cas can protect bacteria and archaea from infectious DNAs, it is not at all clear commonly CRISPR plays this role in natural populations of these microbes. If CRISPR commonly protects bacteria from infectious DNAs, one might expect a strong negative correlation between the presence of a CRISPR system and the signatures of horizontal gene transfer in the same genomes. For Restriction Modification systems – the most prevalent innate immune system of bacteria and archaea – such correlations can be readily detected (54). Yet, in the case of CRISPR-Cas the evidence is ambiguous, with some studies suggesting that CRISPR does form a barrier for the movement of mobile genes between microbial species (55, 56), whereas other studies arrive at the opposite conclusion (57-59). Furthermore, most spacers from sequenced isolates have no homology to viral or plasmid sequences in databases (39) and the same holds for spacers extracted from metagenomes (38).

If indeed CRISPR-Cas is commonly important for protecting populations of microbes from phage and preventing the acquisition of mobile genetic elements, there should be no trouble detecting (in the lab or in nature) CRISPR-Cas-encoding bacteria and archaea that acquire spacers from novel sources of infectious DNAs to become immune to those infections. However, other than the bacterial species listed above, we are unaware of wildtype bacteria or archaea for which spacer acquisition from phage or plasmids has been demonstrated to occur at high frequencies. Moreover, even for species that have been reported to acquire spacers, it is not clear whether they do so in response to only few or many phages and plasmids. Because in most bacteria and archaea spacer

acquisition is rare, a range of elegant tricks have been developed to detect these events (43, 60-65), and while this has propelled our understanding of the mechanisms of spacer acquisition (52), and CRISPR-mediated immunity (66), it raises questions concerning the ecological importance of CRISPR-Cas immune systems. Perhaps this is because many of the domesticated bacteria and archaea that we use for research simply lost their ability to rapidly acquire spacers? Could it be that the quest to find culturable bacteria, archaea, phage and plasmids with these properties has not be adequately extensive? Or could it be that the results of these quests are commonly negative and therefore not reported? Sequence data analysis also provides a mixed picture: CRISPR loci in some species appear to evolve rapidly, e.g. refs. (28, 29, 31) and reviewed in (67), whereas others are relatively static over long periods (59, 68).

Open questions

We hope the preceding has convinced the reader that at this juncture that it is not at all clear how commonly CRISPR protects populations of bacteria and archaea against mobile genetic elements, and therefore to what extent CRISPR-Cas genes are maintained by selection mediated by these infectious genetic elements. In this perspective we discuss the following four questions:

1- Under what conditions will immunity be selected for in populations with functional CRISPR-Cas systems ($CRISPR^+$ bacteria and archaea) confronted with phage and plasmids?

2- Under what conditions will bacteria and archaea with functional CRISPR-Cas immune systems invade from rare and become established in populations that lack CRISPR-Cas systems ($CRISPR^-$ bacteria and archaea)?

3- What is the contribution of CRISPR immunity to the population dynamics, ecology and evolution of prokaryotes and their mobile genetic elements?

4- What are the characteristics of mobile genetic elements that lead to spacer acquisition by bacteria during an infection?

Here we discuss the first three of these questions with the aid of simple, heuristic, mathematical-computer simulation models and a review of the experimental and other empirical studies that have explored these questions. We separately consider the three main sources of infectious DNAs: lytic phage, temperate phage, and conjugative plasmids. The models and results are presented in detail in the Supplemental Material, as are the caveats and concerns about the limitations of these models

and our analyses of their properties. The main conclusions of our analysis are summarized in Figure 1.

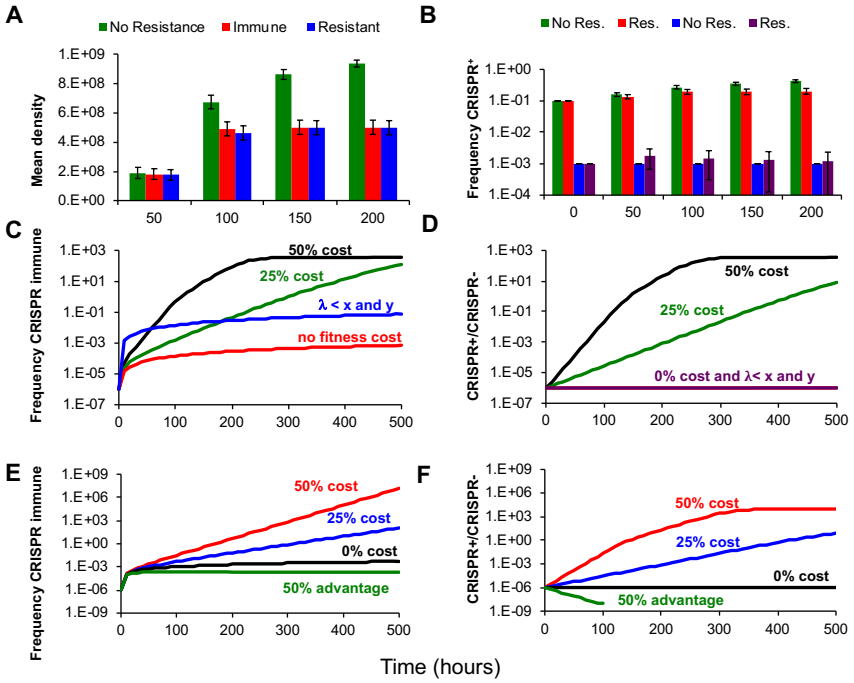

Figure 1: Phage and Plasmid-mediated Selection for CRISPR-Cas immunity. (A) Monte Carlo simulations of selection for CRISPR-Cas immunity and surface-based resistance in a CRISPR$^+$ population initially at equilibrium with phage. The mean and standard error of the changes in density of CRISPR$^+$ immune and resistant bacteria across 100 runs. Green bars are for a population of CRISPR$^+$ bacteria that cannot generate resistant mutants (No Resistance). Red and blue bars are for populations that can evolve both CRISPR-Cas immunity and resistance with an equal probability. **(B)** Invasion of CRISPR$^+$ into a population of CRISPR$^-$ sensitive bacteria at equilibrium with the phage for two different initial frequencies of CRISPR$^+$ phage sensitive cells, 0.1 (red and green bars) and 0.001 (blue and purple bars). Mean and standard error of the frequency of CRISPR$^+$ bacteria over time across 100 runs in populations where resistance cannot be generated (green and blue bars) or where resistance can be generated (red and purple bars). **(C)** The establishment of immune CRISPR$^+$ bacteria in a population at equilibrium with temperate phage. Changes in the ratio CRISPR$^+$ immune cells to the total number of CRISPR$^+$ cells for different costs of the prophage, as indicated. **(D)** The invasion of CRISPR$^+$ bacteria in a CRISPR$^-$ population at equilibrium with a temperate phage. Changes in the ratio of CRISPR$^+$/CRISPR$^-$ are depicted for different costs of the prophage, as indicated. **(E)** The establishment of CRISPR$^+$ immunity in a CRISPR$^+$ population at equilibrium with a conjugative plasmid. Changes in the ratio of immune to non-immune CRISPR$^+$ bacteria are depicted for different costs and benefits of the plasmid, as indicated. **(F)** The establishment of a CRISPR$^+$ population in a CRISPR$^-$ population at equilibrium with a conjugative plasmid. Changes in the ratio of CRISPR$^+$/CRISPR$^-$ are depicted for different costs of the plasmid, as indicated.

I - Lytic phage.

What the theory tells us: To explore the conditions under which lytic phage will select for CRISPR-mediated immunity, we use Monte Carlo simulations where (i) CRISPR immune and surface resistant cells evolve in populations of CRISPR$^+$ cells (i.e. bacteria that are phage sensitive but carry a functional CRISPR-Cas immune system), and (ii) when CRISPR$^+$ populations become established in populations of sensitive CRISPR$^-$ cells (i.e. bacteria lacking a functional CRISPR-Cas immune system). In both cases, we assume at the start of the simulations that the bacterial community being invaded is comprised of a population of sensitive bacteria at equilibrium with a population of lytic phage. In the absence of resistance or immunity, the densities of bacteria in these populations are limited by the phage. As the frequency of immune or resistant bacteria increases, the density of the bacterial population increases to point where it is limited by resources, rather than the phage.

If CRISPR-Cas–mediated immunity is the only way populations of bacteria and archaea can survive infections with lytic phage, there are broad conditions under which selection mediated by these viruses will favor the evolution of this adaptive immune system. This can be seen in Figure 1A from the increases in the mean densities of the bacterial populations in 100 simulations. At the start of these simulations, time 0, there are no resistant or immune bacteria. As time proceeds, resistant and immune cells are generated, by mutation and the acquisition of spacers, respectively. If the population is unable to generate resistant mutants, the green bars, immunity ascends. If the population can generate both immune and resistant bacteria, both are equally favored in these simulations since with the parameters used resistance is as likely to be generated as immunity.

The conditions for CRISPR-Cas to become established in a CRISPR$^-$ population of sensitive microbes at equilibrium with a lytic phage is critically dependent on the initial frequency of the invading population of CRISPR$^+$ cells (Figure 1B). The lower the initial density of the invading CRISPR$^+$ population relative to the CRISPR$^-$ population, the less likely it is that CRISPR$^+$ bacteria will invade, since the probability of generating CRISPR immunity declines. Under conditions where bacteria can evolve both immunity and resistance, the larger population of CRISPR$^-$ bacteria is far more likely to generate resistant bacteria, which further limits the ability of CRISPR$^+$ cells to invade the population. This is particularly critical when the genes coding for this immune system are acquired by horizontal gene transfer, since the number of bacteria acquiring those genes will generally be low.

The invasion conditions theory considered above addresses only one element of the role of phage in the evolution of CRISPR-Cas. The other element is the length of time selection mediated by lytic phage will favor CRISPR-Cas immune systems. This will depend on (i) the capacity of the phage to generate protospacer mutants, (ii) the capacity of the bacteria to acquire novel spacers to counter protospacer mutations in the phage, (iii) the rates at which these novel spacers are acquired, (iv) the rate of mutation to resistance, and (v) the fitness costs associated with the carriage of CRISPR-Cas, surface resistance, and protospacer mutations in the phage. Three theoretical studies that have partly addressed this issue (26, 27, 69), but the models employed do not consider all five of the above-listed factors contributing to the length of time selection for CRISPR-Cas will be maintained.

What the experiments and genomic data say: While we are unaware of experimental studies of the invasion of CRISPR$^+$ into established CRISPR$^-$ populations of bacteria or archaea, there have been several experimental studies of the population and evolutionary dynamics of lytic phage and bacteria with CRISPR-Cas systems. As anticipated by the models, CRISPR-Cas immunity readily evolves in *S. thermophilus* strains exposed to virulent phages (23, 70). In this system, bacteria with envelope resistance are normally not detected and an extended spacer – protospacer arms race can ensue when these bacteria and phage are together in serial transfer culture (23, 71-73). While the phage will eventually be lost, the duration of the arms race and the diversity of spacers and phage protospacer mutants that evolves during this process can be substantial. In these experimental populations, the densities of bacteria remain at levels similar to that of phage-free populations. Stated another way, the bacterial populations are limited by resources, rather than the phage.

In these experiments, it is clear that the co-evolutionary dynamics observed for *Streptococcus thermophilus* and its phage can be attributed to CRISPR-Cas-mediated immunity to the phage. Resistant bacteria only evolve if the CRISPR-Cas system is inactivated by either anti-sense RNA expression (74) or an antiCRISPR (*acr*) gene encoded by the phage (75), with resistance due to mutations in either the receptor or intracellular host genes required for completing the phage life cycle. This system, therefore, corresponds well with the theoretical scenario in Figure S2, which therefore may explain why the evolution of CRISPR immunity is so commonly observed in this model organism.

Also consistent with the theoretical predictions are the results of experiments with bacteria that can evolve both CRISPR immunity and resistance by mutation of the phage receptor. For example, *Pseudomonas aeruginosa* strain PA14 either evolves resistance against phage DMS3*vir* (a

temperate phage locked in the lytic cycle) by mutation of the Type IV pilus or immunity by the acquisition of spacers into its CRISPR arrays (48, 49). Experimental manipulation of the bacterial mutation rate shows that which of these two defense mechanisms prevails strongly depends on the rate at which receptor mutants are generated in the population (76). Similarly, the rates of spacer acquisition matters: bacteria with "primed" CRISPR-Cas systems acquire spacers at a higher rate compared to bacteria with naïve CRISPR-Cas systems; this translates in a strong increase in the proportion of bacteria that evolved CRISPR-Cas immunity following phage exposure (40, 41, 47, 60, 77-80). This phenomenon was observed first for many Type I and, more recently, also for Type II systems CRISPR-Cas systems (46, 77), and relies on an imperfect match between a pre-existing spacer and the infectious genome (see ref. (52) for mechanistic details). When bacteria are exposed to defective phage, or when bacteria carry both a restriction-modification and a CRISPR-Cas system, the rates of spacer acquisition are also elevated, and again this leads to higher levels of evolved CRISPR immunity (81, 82). The typically low frequencies of CRISPR resistant bacteria evolve in the lab may therefore be at least in part explained by the high mutation rates and large population sizes relative to the rates of spacer acquisition in many model systems, although the fitness costs and benefits of CRISPR-based immunity and surface-based resistance will also be important (49, 83), which has been reviewed in (84).

II - Temperate phage.

What the theory tells us: The conditions under which exposure to temperate phage will select for CRISPR-Cas immunity depends on whether the immune cells are entering a $CRISPR^+$ population or a $CRISPR^+$ population is entering a $CRISPR^-$ population. In the former case, if the population is initially at equilibrium with the temperate phage, a CRISPR immune population can be weakly favored if the probability of forming lysogens is less than that of generating CRISPR immune cells (note that the model assumes that lysogens are immune to superinfection). For selection mediated by temperate phage to strongly favor CRISPR-Cas immunity, the carriage of the prophage has to engender a fitness cost (Figure 1C). Furthermore, the only condition under which $CRISPR^+$ bacteria can invade a $CRISPR^-$ population at equilibrium with temperate phage is when lysogeny engenders a fitness cost (Figure 1D). If the carriage of the prophage augments the fitness of the bacteria, as may be the case (85, 86), CRISPR immune bacteria will be selected against in virtually all cases.

What the experiments and data say: There is evidence that CRISPR-Cas systems target temperate phage in nature. For example, spacers encoded by *P. aeruginosa* isolates with Type I CRISPR-Cas systems from cystic fibrosis lungs were found to frequently target related groups of temperate phages (including DMS3). Surprisingly, however, in these patients, no spacer acquisition was observed over time (59). By contrast, in an experimental study where a wound model was infected with a mix of 6 *P. aeruginosa* strains, CRISPR immunity was found to evolve in *P. aeruginosa* strain PA14 against a prophage carried by one of the strains, known as strain B23-2 (87). Another recent study with *P. aeruginosa* PA14 and its phage DMS3 showed that carrying a primed CRISPR-Cas immune system is, in fact, maladaptive, due to immunopathology in $CRISPR^+$ lysogens since the partial matching spacer triggers cleavage of the prophage. The associated fitness costs caused a rapid invasion of spontaneous mutants that had lost their CRISPR-Cas immune system (88). Acquisition of perfectly matching spacers in lysogens amplified these fitness costs, since this further programs the immune system to attack the prophage inside the bacterial genome. Such self-targeting by CRISPR-Cas is well known to be highly toxic (89-96), even for Type III CRISPR-Cas systems that target only transcriptionally active DNA (97, 98). Finally, a recent and exciting study showed that *R. intestinalis* in the mouse gut can evolve high levels of CRISPR-based immunity when one of its active prophages evolves to become hypervirulent (i.e. virulent phage mutants that can infect the lysogen) (50). This, however, brings us to what is anticipated for lytic phage. We are unaware of empirical studies that have explored the contribution of temperate phage to the establishment of CRISPR-Cas in population in $CRISPR^-$ populations.

III – Conjugative Plasmids.

What the theory tells us: The only condition under which conjugative plasmid can select for CRISPR-Cas immunity in our model is when the plasmids engender fitness costs. This can be seen for the establishment of CRISPR-immune cells in a $CRISPR^+$ population initially at equilibrium with a conjugative plasmid (Figure 1E) and the invasion of $CRISPR^+$ bacteria into a $CRISPR^-$ population initially at equilibrium with a conjugative plasmid (Figure 1F). If the plasmid increases the fitness of their host bacteria, CRISPR-Cas would be selected against (also see (99)).

What the experiments and data say: Experimental studies demonstrate that bacteria can evolve CRISPR-based immunity against plasmids, and in the case of Type I CRISPR-Cas systems, spacer acquisition is accelerated if the CRISPR immune system is primed. Most commonly this priming is

accomplished by engineering the plasmid in a way that it contains a sequence with a partial match to a pre-existing spacer on the genome of a CRISPR$^+$ host (40, 47, 60, 100). As anticipated by the model, CRISPR immunity will be selected against if the plasmid provides a net benefit to the host, for example when it confers resistance to an antibiotic that is present in the environment. Rare mutants that lack an intact CRISPR-Cas immune system will quickly replace the dominant CRISPR$^+$ population (101). The model predicts that if the carriage of the plasmid engenders a fitness cost, CRISPR-mediated immunity to that plasmid will be favored. To our knowledge, this has not been demonstrated experimentally. Indeed, the evidence we are aware of is inconsistent with this hypothesis. A study with engineered *E. coli* strains showed that even under these conditions CRISPR immunity can be maladaptive, because the time between infection and clearance of the plasmid may allow for the expression of toxin/anti-toxin cassettes. This, in turn, triggers a significant cost of plasmid removal because (short-lived) antitoxin is no longer produced to neutralize the long-lived toxin molecules (102). How common this is, is not at all clear. We anticipate that CRISPR-immunity will be favored when the plasmid engenders a fitness cost, provided that any costs of plasmid removal do not outweigh the benefits of being plasmid-free. We are unaware of any experimental or other empirical studies that have addressed the question of the conditions under which plasmid-mediated selection will favor the establishment and maintenance of CRISPR-Cas in CRISPR$^-$ populations.

Other reasons why virulent phage may not select for CRISPR immunity: In the heuristic model considered here and the experiments described in the preceding section, the acquisition of spacers confers immunity to the mobile genetic element from whence the spacers are derived. This may not always be the case. More and more mobile genetic elements are found to encode anti-CRISPR (*acr*) genes that suppress CRISPR-Cas immune systems through a range of different mechanisms and with often high specificity for a single subtype (103-112). The ability of *acrs* to bypass immunity of bacteria that are already CRISPR immune varies (113, 114), but even the weakest *acr* genes characterized to date effectively block the evolution of CRISPR immunity (115).

Other viruses can by-pass CRISPR immunity without a need for *acr* genes. So-called "jumbophage", which are a class of phages with genome sizes that exceed 200kb, have been reported to form nucleus-like structure during infection (116-118). These structures contain the phage genomes, which shields it from DNA-targeting CRISPR-Cas systems, but not from systems

that have RNA-targeting activity, such as Type III-A and VI-A CRISPR-Cas (119, 120). This variation in the level of protection explains why in nature spacer acquisition from nucleus-forming jumbophages is detected more frequently for bacteria that carry Type III systems compared to those that carry Type I-E or I-F systems (119). A lack of protection by CRISPR immunity is not limited to jumbophages: *E. coli* strains carrying Type I-E CRISPR-Cas that were engineered to carry a single targeting spacer against different phages revealed a lack of protection against phages R1-37 (a giant phage) and T4 (121). The ability of phage T4 to by-pass Type I-E CRISPR immunity is at least in part attributable to their genome containing glucosyl-5-hydroxymethylcytosine instead of cytosine (122), and this cytosine modification also confers infectivity to the phage when bacteria have Type II-A CRISPR-based immunity (122, 123), but not when they have Type V-A CRISPR-based immunity (122). Type I-E CRISPR-Cas offers protection against phage T7, but only under low phage densities; at high MOIs the cultures were lysed as efficiently as uninduced controls. Efficient protection against T5 was only observed if the CRISPR spacer targeted a pre-early gene (121). Furthermore, a recent study demonstrated that a Type I-F CRISPR-Cas system of *Pectobacterium atrosepticum* reduced the efficiency of plaquing of two virulent phages ϕTE and ϕM1 when the immune system was engineered to carry spacers targeting these phages. Interestingly, CRISPR-Cas was unable to rescue the host from cell death, hence resulting in an Abi-like phenotype that blocks phage propagation (124). It remains to be determined if and when CRISPR immunity would evolve in bacterial populations exposed to these phages.

These observations are in stark contrast with the high levels of protection against virulent mutants of temperate phage. Examples include high levels of CRISPR-based immunity observed in *P. aeruginosa* strain PA14 against its phage DMS3vir, which we discussed above. Mild or strong overexpression of the Type I-E CRISPR-Cas system of *E. coli* targeting the non-lysogenic mutant λvir provides full immunity, with efficiency of plaquing around 10^{-6} (121, 125). Similarly, the Type II-A CRISPR-Cas immune system of *Streptococcus pyogenes* SF370 provides high levels of immunity when expressed in *Staphylococcus aureus* RN4220 against the staphylococcal phage ϕNM4γ4, a lytic mutant of ϕNM4 (43-45) and the Type III-A from *S. epidermidis* RP62a provides high levels of immunity when expressed in *S. aureus* RN4220 against phage ϕNM1γ6, a lytic mutant of the temperate phage ϕNM1 (98). Mapping the variability in the levels of protection conferred by CRISPR-Cas immunity using a wider range of CRISPR immune systems and phages will be critical to understand when and where these systems matter.

Consequences of CRISPR-Cas immunity to the population dynamics, ecology and (co)evolution of bacteria and lytic phage

In experiments where bacteria evolve high levels of CRISPR-based immunity, three possible outcomes have been observed: the phage (i) are be eliminated in short order, (ii) persist for an extended time in spacer-protospacer arms race but are eventually to be lost, (iii) persist without co-evolution. The first outcome is observed when *P. aeruginosa* PA14 is infected with DMS3*vir* (126, 127). This is because in this empirical system, the bacteria acquire many different spacers at the population level, which increases the degree of protection since it reduces the probability that phage evolve to overcome host immunity (126, 128), which they can do by point mutation (24) or deletion of the target regions on the phage genome (129). The second scenario is observed when the spacer diversity at the population-level is reduced. In this case, the phage can evolve to overcome host immunity, and hosts, in turn, need to acquire novel spacers to regain immunity. Such a spacer-protospacer arms race is observed in *S. thermophilus*, where bacteria and phage coexist and coevolve for an extended time eventually for the phage to be lost (71, 73, 130). Phage extinction is in this system due to the arms-race being asymmetrical: acquisition of novel spacers is cost-free for the host (131), whereas accumulating point mutations reduces the fitness of the phage (132). Moreover, the host population gradually increases the diversity of spacers, which makes it harder for the phage to keep up with the host (73). Finally, the phage may continue to be maintained without co-evolution, when bacteria with CRISPR immunity in the population continuously acquire mutations in their CRISPR-Cas immune systems that cause phenotypic reversion to sensitivity (72), or when there is a continuous influx of sensitive bacteria due to immigration (133). While important progress has been made in understanding the consequences of the evolution of CRISPR immunity, most studies have been carried out in highly simplified environments with a single host species infected with a single phage in well-mixed and nutrient-rich broth. Future studies that examine these interactions under more ecologically relevant conditions are desperately needed to understand how CRISPR-Cas systems shape microbial population and evolutionary dynamics in nature.

Conclusion and future directions

We do not question the validity of the hypothesis that CRISPR-Cas systems can protect bacteria and archaea from infections with deleterious ("badass") DNAs. What we question is the magnitude of

the contribution of these systems to the ecology and evolution of populations of bacteria and archaea and their phage and other mobile genetic elements, collectively MGEs. As outlined above, many key questions remain. First, more experimental and observational studies are needed to understand not only how frequently, but also when, where and why CRISPR-Cas systems play a role in defense against MGEs. Second, it remains unclear how commonly selection mediated by MGEs is responsible for the existence and maintenance of CRISPR-Cas systems in populations of bacteria and archaea, and how this is determined by the type of MGE. Third, while CRISPR-Cas systems clearly spread by HGT, it remains unclear how these genes are able to invade a population from rare, especially if the bacteria or archaea can evolve envelope resistance as well. Finally, our understanding of the ecological and evolutionary consequences of CRISPR-Cas immune responses are limited to in vitro experiments that lack much of the biotic and abiotic complexity of natural environments. Could it be that the biotic and abiotic complexity of the real world, where communities of microbes include multiple species and strains as well as diversity in phage and plasmids, are spatially structured and exist in fluctuating environments are critical to the evolution and maintenance of CRISPR (49, 83, 134)? Filling these gaps in our current understanding of CRISPR ecology and evolution requires interdisciplinary approaches that combine observational studies, mathematical and computer simulation models, as well as population and evolutionary dynamics experiments. The question is how can we do experiments in a way that they would also provide a test of the generality of the hypotheses that are being tested? For that, we would need a diverse array of culturable bacteria and archaea with functional CRISPR-Cas systems and a diverse set of phage and plasmids, which inevitably requires a many different research teams to examine these questions. This brings us back to our concern about the dearth of bacteria and archaea phage and plasmid systems amenable for these experimental studies, and the "fishing expedition" dilemma that a quest to find new systems engenders. However, it is always more difficult publishing negative evidence, no matter how informative that evidence would be. We argue that there is a pressing need to publish any negative results of spacer acquisition in response to MGEs, since knowing which culturable bacteria and archaea with functional CRISPR-Cas do and do not acquire spacers and how this depends on the type of infectious DNA will be critical to fully understand the evolutionary ecology of CRISPR-Cas.

Acknowledgments

ERW received funding from the Natural Environment Research Council (http://www.nerc.ac.uk) (NE/M018350/1), the Biotechnology and Biological Sciences Research Council (BB/N017412/1), and the European Research Council (https://erc.europa.eu) (ERC-STG-2016-714478 - EVOIMMECH). BL was funded by grants from the U.S. National Institutes of General Medical Science GM098175-17 (BRL) and the generosity of Emory University.

References

1. Ishino Y, Shinagawa H, Makino K, Amemura M, & Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol* 169(12):5429-5433.
2. Nakata A, Amemura M, & Makino K (1989) Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. *J Bacteriol* 171(6):3553-3556.
3. Hermans PW*, et al.* (1991) Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in Mycobacterium tuberculosis complex strains. *Infect Immun* 59(8):2695-2705.
4. Mojica FJ, Ferrer C, Juez G, & Rodriguez-Valera F (1995) Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. *Mol Microbiol* 17(1):85-93.
5. Bult CJ*, et al.* (1996) Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science* 273(5278):1058-1073.
6. Nelson KE*, et al.* (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima. *Nature* 399(6734):323-329.
7. Mojica FJ, Diez-Villasenor C, Soria E, & Juez G (2000) Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* 36(1):244-246.
8. Mojica FJ & Rodriguez-Valera F (2016) The discovery of CRISPR in archaea and bacteria. *FEBS J* 283(17):3162-3169.
9. Sorek R, Kunin V, & Hugenholtz P (2008) CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* 6(3):181-186.
10. Pourcel C, Salvignol G, & Vergnaud G (2005) CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151(Pt 3):653-663.
11. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, & Koonin EV (2006) 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. *Biol Direct* 1:7.
12. Makarova KS*, et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 13(11):722-736.
13. Koonin EV, Makarova KS, & Wolf YI (2017) Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annu Rev Microbiol* 71:233-261.
14. Jansen R, Embden JD, Gaastra W, & Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43(6):1565-1575.

15. Makarova KS, Aravind L, Grishin NV, Rogozin IB, & Koonin EV (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res* 30(2):482-496.
16. Haft DH, Selengut J, Mongodin EF, & Nelson KE (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* 1(6):e60.
17. Makarova KS & Koonin EV (2015) Annotation and Classification of CRISPR-Cas Systems. *Methods Mol Biol* 1311:47-75.
18. Shmakov SA*, et al.* (2019) Systematic prediction of functionally linked genes in bacterial and archaeal genomes. *Nat Protoc* 14(10):3013-3031.
19. Shmakov S*, et al.* (2017) Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol* 15(3):169-182.
20. Koonin EV & Makarova KS (2019) Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180087.
21. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, & Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 60(2):174-182.
22. Bolotin A, Quinquis B, Sorokin A, & Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(Pt 8):2551-2561.
23. Barrangou R*, et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709-1712.
24. Deveau H*, et al.* (2008) Phage response to CRISPR-Encoded resistance in Streptococcus thermophilus. *Journal of Bacteriology* 190(4):1390-1400.
25. Iranzo J, Lobkovsky AE, Wolf YI, & Koonin EV (2013) Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *J Bacteriol* 195(17):3834-3844.
26. Childs LM, Held NL, Young MJ, Whitaker RJ, & Weitz JS (2012) Multiscale model of CRISPR-induced coevolutionary dynamics: diversification at the interface of Lamarck and Darwin. *Evolution* 66(7):2015-2029.
27. Childs LM, England WE, Young MJ, Weitz JS, & Whitaker RJ (2014) CRISPR-induced distributed immunity in microbial populations. *PLoS One* 9(7):e101710.
28. Tyson GW & Banfield JF (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ Microbiol* 10(1):200-207.
29. Andersson AF & Banfield JF (2008) Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320(5879):1047-1050.
30. Pride DT*, et al.* (2011) Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. *Genome Res* 21(1):126-136.
31. Laanto E, Hoikkala V, Ravantti J, & Sundberg L-R (2017) Long-term genomic coevolution of host-parasite interaction in the natural environment. *Nature communications* 8(1):111.
32. Pauly MD, Bautista MA, Black JA, & Whitaker RJ (2019) Diversified local CRISPR-Cas immunity to viruses of Sulfolobus islandicus. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180093.
33. Bautista MA, Black JA, Youngblut ND, & Whitaker RJ (2017) Differentiation and Structure in Sulfolobus islandicus Rod-Shaped Virus Populations. *Viruses* 9(5).
34. Held NL & Whitaker RJ (2009) Viral biogeography revealed by signatures in Sulfolobus islandicus genomes. *Environ Microbiol* 11(2):457-466.

35. Held NL, Herrera A, Cadillo-Quiroz H, & Whitaker RJ (2010) CRISPR associated diversity within a population of Sulfolobus islandicus. *PLoS One* 5(9).
36. van der Ploeg JR (2009) Analysis of CRISPR in Streptococcus mutans suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology* 155(Pt 6):1966-1976.
37. Heidelberg JF, Nelson WC, Schoenfeld T, & Bhaya D (2009) Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS One* 4(1):e4169.
38. Skennerton CT, Imelfort M, & Tyson GW (2013) Crass: identification and reconstruction of CRISPR from unassembled metagenomic data. *Nucleic Acids Res* 41(10):e105.
39. Shmakov SA*, et al.* (2017) The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *MBio* 8(5).
40. Swarts DC, Mosterd C, van Passel MW, & Brouns SJ (2012) CRISPR interference directs strand specific spacer acquisition. *PLoS One* 7(4):e35888.
41. Datsenko KA*, et al.* (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* 3:945.
42. Yosef I, Goren MG, & Qimron U (2012) Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. *Nucleic Acids Res* 40(12):5569-5576.
43. Heler R*, et al.* (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519(7542):199-202.
44. Modell JW, Jiang WY, & Marraffini LA (2017) CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature* 544(7648):101-+.
45. Nussenzweig PM, McGinn J, & Marraffini LA (2019) Cas9 Cleavage of Viral Genomes Primes the Acquisition of New Immunological Memories. *Cell Host Microbe* 26(4):515-+.
46. Heler R, Wright AV, Vucelja M, Doudna JA, & Marraffini LA (2019) Spacer Acquisition Rates Determine the Immunological Diversity of the Type II CRISPR-Cas Immune Response. *Cell Host Microbe* 25(2):242-249 e243.
47. Richter C*, et al.* (2014) Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res* 42(13):8516-8526.
48. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, & O'Toole GA (2012) The CRISPR/Cas adaptive immune system of Pseudomonas aeruginosa mediates resistance to naturally occurring and engineered phages. *J Bacteriol* 194(21):5728-5738.
49. Westra ER*, et al.* (2015) Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr Biol* 25(8):1043-1049.
50. Cornuault JK*, et al.* (2019) The enemy from within: a prophage of Roseburia intestinalis systematically turns lytic in the mouse gut, driving bacterial adaptation by CRISPR spacer acquisition. *ISME J.*
51. Erdmann S & Garrett RA (2012) Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol Microbiol* 85(6):1044-1056.
52. Jackson SA*, et al.* (2017) CRISPR-Cas: Adapting to change. *Science* 356(6333).
53. Pawluk A, Davidson AR, & Maxwell KL (2018) Anti-CRISPR: discovery, mechanism and function. *Nat Rev Microbiol* 16(1):12-17.
54. Oliveira PH, Touchon M, & Rocha EP (2014) The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res* 42(16):10618-10631.

55. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *MBio* 1(4).
56. O'Meara D & Nunney L (2019) A phylogenetic test of the role of CRISPR-Cas in limiting plasmid acquisition and prophage integration in bacteria. *Plasmid* 104:102418.
57. Gophna U*, et al.* (2015) No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *ISME J* 9(9):2021-2027.
58. Touchon M*, et al.* (2012) Antibiotic resistance plasmids spread among natural isolates of Escherichia coli in spite of CRISPR elements. *Microbiology* 158(12):2997-3004.
59. England WE, Kim T, & Whitaker RJ (2018) Metapopulation Structure of CRISPR-Cas Immunity in Pseudomonas aeruginosa and Its Viruses. *mSystems* 3(5).
60. Staals RH*, et al.* (2016) Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR-Cas system. *Nat Commun* 7:12853.
61. Amlinger L, Hoekzema M, Wagner EGH, Koskiniemi S, & Lundgren M (2017) Fluorescent CRISPR Adaptation Reporter for rapid quantification of spacer acquisition. *Sci Rep* 7(1):10392.
62. Diez-Villasenor C, Guzman NM, Almendros C, Garcia-Martinez J, & Mojica FJ (2013) CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of Escherichia coli. *RNA Biol* 10(5):792-802.
63. Kieper SN*, et al.* (2018) Cas4 Facilitates PAM-Compatible Spacer Selection during CRISPR Adaptation. *Cell Rep* 22(13):3377-3384.
64. McKenzie RE, Almendros C, Vink JNA, & Brouns SJJ (2019) Using CAPTURE to detect spacer acquisition in native CRISPR arrays. *Nat Protoc* 14(3):976-990.
65. Schmidt F, Cherepkova MY, & Platt RJ (2018) Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature* 562(7727):380-385.
66. Mohanraju P*, et al.* (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353(6299):aad5147.
67. Westra ER, Dowling AJ, Broniewski JM, & van Houte S (2016) Evolution and Ecology of CRISPR. *Annu Rev Ecol Evol S* 47:307-331.
68. Touchon M & Rocha EP (2010) The small, slow and specialized CRISPR and anti-CRISPR of Escherichia and Salmonella. *PLoS One* 5(6):e11126.
69. Gurney J, Pleska M, & Levin BR (2019) Why put up with immunity when there is resistance: an excursion into the population and evolutionary dynamics of restriction-modification and CRISPR-Cas. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180096.
70. Horvath P*, et al.* (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J Bacteriol* 190(4):1401-1412.
71. Paez-Espino D*, et al.* (2015) CRISPR immunity drives rapid phage genome evolution in Streptococcus thermophilus. *MBio* 6(2).
72. Weissman J*, et al.* (2018) Immune loss as a driver of coexistence during host-phage coevolution. *ISME*:1-13.
73. Common J, Morley D, Westra ER, & van Houte S (2019) CRISPR-Cas immunity leads to a coevolutionary arms race between Streptococcus thermophilus and lytic phage. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180098.
74. McDonnell B, Mahony J, Hanemaaijer L, Kouwen T, & van Sinderen D (2018) Generation of Bacteriophage-Insensitive Mutants of Streptococcus thermophilus via an Antisense RNA CRISPR-Cas Silencing Approach. *Appl Environ Microbiol* 84(4).
75. Labrie SJ*, et al.* (2019) A mutation in the methionine aminopeptidase gene provides phage resistance in Streptococcus thermophilus. *Sci Rep* 9(1):13816.

76. Chevallereau A, Meaden S, van Houte S, Westra ER, & Rollie C (2019) The effect of bacterial mutation rate on the evolution of CRISPR-Cas adaptive immunity. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180094.
77. Nicholson TJ*, et al.* (2019) Bioinformatic evidence of widespread priming in type I and II CRISPR-Cas systems. *RNA Biol* 16(4):566-576.
78. Musharova O*, et al.* (2019) Systematic analysis of Type I-E Escherichia coli CRISPR-Cas PAM sequences ability to promote interference and primed adaptation. *Mol Microbiol* 111(6):1558-1570.
79. Xue C*, et al.* (2015) CRISPR interference and priming varies with individual spacer sequences. *Nucleic Acids Res* 43(22):10831-10847.
80. Musharova O*, et al.* (2018) Avoidance of Trinucleotide Corresponding to Consensus Protospacer Adjacent Motif Controls the Efficiency of Prespacer Selection during Primed Adaptation. *MBio* 9(6).
81. Hynes AP, Villion M, & Moineau S (2014) Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat Commun* 5:4399.
82. Dupuis ME, Villion M, Magadan AH, & Moineau S (2013) CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat Commun* 4:2087.
83. Alseth EO*, et al.* (2019) Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature* 574(7779):549-552.
84. van Houte S, Buckling A, & Westra ER (2016) Evolutionary Ecology of Prokaryotic Immune Mechanisms. *Microbiol Mol Biol Rev* 80(3):745-763.
85. Taylor VL, Fitzpatrick AD, Islam Z, & Maxwell KL (2019) The Diverse Impacts of Phage Morons on Bacterial Fitness and Virulence. *Adv Virus Res* 103:1-31.
86. Harrison E & Brockhurst MA (2017) Ecological and Evolutionary Benefits of Temperate Phage: What Does or Doesn't Kill You Makes You Stronger. *Bioessays* 39(12).
87. Gloag ES*, et al.* (2019) Pseudomonas aeruginosa Interstrain Dynamics and Selection of Hyperbiofilm Mutants during a Chronic Infection. *MBio* 10(4).
88. Rollie C*, et al.* (2020) Targeting of temperate phages drives loss of type I CRISPR-Cas systems. *Nature* in press.
89. Vercoe RB*, et al.* (2013) Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet* 9(4):e1003454.
90. Edgar R & Qimron U (2010) The Escherichia coli CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J Bacteriol* 192(23):6291-6294.
91. Citorik RJ, Mimee M, & Lu TK (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat Biotechnol* 32(11):1141-1145.
92. Ramachandran G & Bikard D (2019) Editing the microbiome the CRISPR way. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180103.
93. Bikard D & Barrangou R (2017) Using CRISPR-Cas systems as antimicrobials. *Curr Opin Microbiol* 37:155-160.
94. Cui L & Bikard D (2016) Consequences of Cas9 cleavage in the chromosome of Escherichia coli. *Nucleic Acids Res* 44(9):4243-4251.
95. Pursey E, Sunderhauf D, Gaze WH, Westra ER, & van Houte S (2018) CRISPR-Cas antimicrobials: Challenges and future prospects. *PLoS Pathog* 14(6):e1006990.
96. Gomaa AA*, et al.* (2014) Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *MBio* 5(1):e00928-00913.
97. Goldberg GW*, et al.* (2018) Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces the fitness of lysogenic hosts. *Nat Commun* 9(1):61.

98. Goldberg GW, Jiang W, Bikard D, & Marraffini LA (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* 514(7524):633-637.
99. Gandon S & Vale PF (2014) The evolution of resistance against good and bad infections. *J Evolution Biol* 27(2):303-312.
100. Fineran PC*, et al.* (2014) Degenerate target sites mediate rapid primed CRISPR adaptation. *P Natl Acad Sci USA* 111(16):E1629-E1638.
101. Jiang W*, et al.* (2013) Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet* 9(9):e1003844.
102. van Sluijs L*, et al.* (2019) Addiction systems antagonize bacterial adaptive immunity. *FEMS Microbiol Lett*.
103. Bondy-Denomy J, Pawluk A, Maxwell KL, & Davidson AR (2013) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493(7432):429-432.
104. Hynes AP*, et al.* (2018) Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. *Nature Communications* 9.
105. Hynes AP*, et al.* (2017) An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. *Nature Microbiology* 2(10):1374-1380.
106. Athukoralage JS*, et al.* (2020) An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. *Nature* 577(7791):572-575.
107. Bhoobalan-Chitty Y, Johansen TB, Di Cianni N, & Peng X (2019) Inhibition of Type III CRISPR-Cas Immunity by an Archaeal Virus-Encoded Anti-CRISPR Protein. *Cell* 179(2):448-458 e411.
108. He F*, et al.* (2018) Anti-CRISPR proteins encoded by archaeal lytic viruses inhibit subtype I-D immunity. *Nat Microbiol* 3(4):461-469.
109. Watters KE, Fellmann C, Bai HB, Ren SM, & Doudna JA (2018) Systematic discovery of natural CRISPR-Cas12a inhibitors. *Science* 362(6411):236-239.
110. Marino ND*, et al.* (2018) Discovery of widespread type I and type V CRISPR-Cas inhibitors. *Science* 362(6411):240-242.
111. Pawluk A*, et al.* (2016) Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* 167(7):1829-1838 e1829.
112. Pawluk A, Bondy-Denomy J, Cheung VH, Maxwell KL, & Davidson AR (2014) A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of Pseudomonas aeruginosa. *MBio* 5(2):e00896.
113. Landsberger M*, et al.* (2018) Anti-CRISPR Phages Cooperate to Overcome CRISPR-Cas Immunity. *Cell*.
114. Borges AL*, et al.* (2018) Bacteriophage Cooperation Suppresses CRISPR-Cas3 and Cas9 Immunity. *Cell* 174(4):917-+.
115. Chevallereau A*, et al.* (2020) Exploitation of the cooperative behaviours of anti-CRISPR phages. Cell Host Microbe, in press. *Cell Host Microbe* in press.
116. Chaikeeratisak V*, et al.* (2019) Viral Capsid Trafficking along Treadmilling Tubulin Filaments in Bacteria. *Cell* 177(7):1771-+.
117. Chaikeeratisak V*, et al.* (2017) Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355(6321):194-197.
118. Chaikeeratisak V*, et al.* (2017) The Phage Nucleus and Tubulin Spindle Are Conserved among Large Pseudomonas Phages. *Cell Reports* 20(7):1563-1571.
119. Malone LM*, et al.* (2019) A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat Microbiol*.

120. Mendoza SD, *et al.* (2019) A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature*.
121. Strotskaya A, *et al.* (2017) The action of Escherichia coli CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Research* 45(4):1946-1957.
122. Vlot M, *et al.* (2018) Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR-Cas effector complexes. *Nucleic Acids Research* 46(2):873-885.
123. Bryson AL, *et al.* (2015) Covalent Modification of Bacteriophage T4 DNA Inhibits CRISPR-Cas9. *Mbio* 6(3).
124. Watson BNJ, *et al.* (2019) Type I-F CRISPR-Cas resistance against virulent phages results in abortive infection and provides population-level immunity. *Nat Commun* 10(1):5526.
125. Brouns SJJ, *et al.* (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321(5891):960-964.
126. van Houte S, *et al.* (2016) The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* 532(7599):385-388.
127. Morley D, Broniewski JM, Westra ER, Buckling A, & van Houte S (2017) Host diversity limits the evolution of parasite local adaptation. *Mol Ecol* 26(7):1756-1763.
128. Chabas H, *et al.* (2018) Evolutionary emergence of infectious diseases in heterogeneous host populations. *Plos Biol* 16(9).
129. Watson BNJ, *et al.* (2019) Different genetic and morphological outcomes for phages targeted by single or multiple CRISPR-Cas spacers. *Philos T R Soc B* 374(1772).
130. Paez-Espino D, *et al.* (2013) Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat Commun* 4:1430.
131. Vale PF, *et al.* (2015) Costs of CRISPR-Cas-mediated resistance in Streptococcus thermophilus. *Proc Biol Sci* 282(1812):20151270.
132. Chabas H, *et al.* (2019) Variability in the durability of CRISPR-Cas immunity. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180097.
133. Chabas H, van Houte S, Hoyland-Kroghsbo NM, Buckling A, & Westra ER (2016) Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. *Proc Biol Sci* 283(1837).
134. Broniewski JM, Meaden S, Paterson S, Buckling A, & Westra ER (2020) The effect of phage genetic diversity on bacterial resistance evolution. *ISME J,* doi: 10.1038/s41396-019-0577-7.

How important is CRISPR-Cas for protecting natural populations of bacteria against infections with badass DNAs?

Edze Westra and Bruce Levin

Supplemental Material. A theoretical consideration of the population and evolutionary dynamics of CRISPR-Cas mediated immunity

As considered in the body of this perspective, CRISPR-Cas can provide bacteria protection against deleterious (badass) DNAs from three major sources, lytic (virulent) phage, temperate phage, and self-transmissible plasmids. In this supplemental material, we use mathematical and computer simulation models to elucidate the *a priori* conditions under which these three types infectious DNAs will select for (i) CRISPR-Cas mediated immunity in populations with functional CRISPR-Cas systems, $CRISPR^+$, and (ii) the ascent of $CRISPR^+$ populations without functional CRISPR-Cas systems, $CRISPR^-$.

For all these simulations, we assume the populations are maintained in continuous (chemostat) culture. A limiting resource, r µg/ml from a reservoir where it is maintained at a concentration RRµg/ml enters a habitat of unit volume at a rate w per hour, which is the same rate at which excess resources, bacteria, and free phage are removed. The rate of growth of the bacteria of type i is directly proportional to its maximum growth v_i per cell per hour and hyperbolic function of the concentration of the resource (1).

$$Z_i(r) = v_i \bullet \frac{r}{(r+k)}$$

The parameter, *k,* the Monod constant, is the concentration of the limiting resource, where the rate of growth of the population is half its maximum value. As in (2), we assume the limiting resource is consumed at a rate proportional to the growth rate of those bacteria, $Z_i(r)$, and the amount of resource required to produce a new cell, e µg/ml.

The properties of these models were analyzed numerically with Berkeley Madonna. For copies of these programs and instructions for their use, write to blevin@emory.edu. The models used for these numerical analyses are generic and chosen to illustrate the conditions under which selection mediated by phage and plasmids will favor the evolution of CRISPR-

Cas immunity. However, the growth rates, phage infection and lysogeny rates, and plasmid transfer rates are in the ranges estimated in the cited experimental studies.

1. The population dynamics of lytic phage and bacteria with and without CRISPR-Cas immunity

The Model: In this model, depicted in Figure S1, there is a single population of lytic phage, V, and two types of bacteria: those that are CRISPR$^+$ (i.e. carry a functional CRISPR-Cas system) and those that are CRISPR$^-$ (i.e. lack a functional CRISPR-Cas system). CRISPR$^-$ bacteria can exist in two states; they can either be sensitive to the phage or they can have surface resistance, respectively S and SR. CRISPR$^+$ populations can exist in three states: one sensitive to the phage, C, one with surface resistance, CR, and one that is CRISPR immune, CI. The resistant cells, SR and CR are refractory to the phage, while the CRISPR immune population can be infected by the phage, but the infecting phage are lost. The variables V, S, SR, C, CR, and CI are both the designations and densities of these populations, particles, and cells per ml.

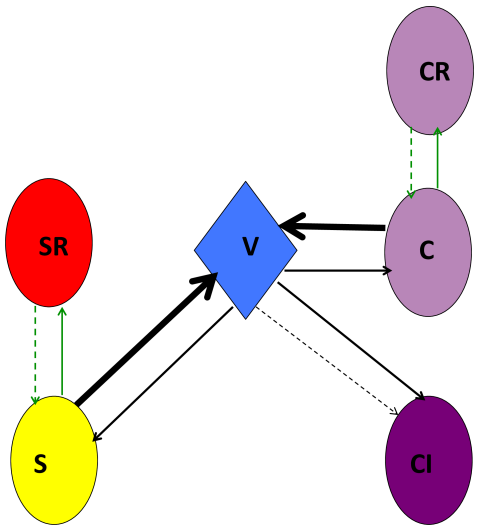

Figure: S1 Lytic phage model: V- free phage, S - CRISPR$^-$ sensitive bacteria, SR - CRISPR$^-$ resistant (refractory) bacteria, C - CRISPR$^+$ phage sensitive bacteria, CR - CRISPR$^+$ resistant (refractory) bacteria, CI - CRISPR$^+$ immune bacteria. The thin black arrows indicate infection and thick black arrow lytic production of phage. The thin green solid and dotted lines designate transitions between states due to mutation or phenotypic processes, respectively, the generation and loss of resistance.

The bacteria grow at maximum rates, v_S, v_{SR}, v_C, v_R and v_{CI} per cell per hour. The phage adsorb to the bacteria at a rate proportional to the product of their densities and a rate parameter, δ cell x ml/hour (3). If the hosts are sensitive to the phage, S or C, upon infection β phage particles per cell, the burst size, are produced and the infected host cells are removed

from the population. Phage infecting CRISPR$^+$ immune cells, CI, are lost and removed from the population. A fraction, x ($0 \leq x \leq 1$) of the adsorptions of phage V to sensitive CRISPR$^+$ cells, C, produce immune cells CI. (4). By mutation at a rate μ_{SR} per cell per hour, sensitive cells generate resistant mutants, S→SR, and C→CR. As in (5), we assume that either by mutation or phenotypic processes, at a rate μ_{RS} per cell per hour resistant cells become susceptible, SR→S, and CR→C. With these definitions, the rate of change in the densities of the bacteria and phage and concentration of resources is given by the following set of coupled differential equations.

$$\frac{dr}{dt} = w \bullet (RR - r) - e \bullet y(r) \bullet (v_S \bullet S + v_{SR} \bullet SR + v_C \bullet C + v_{CR} \bullet CR + v_{CI} \bullet CI)$$

$$\frac{dS}{dt} = v_S \bullet S \bullet y(r) - d \bullet V \bullet S + \mu_{RS} \bullet SR - m_{SR} \bullet S - w \bullet S$$

$$\frac{dSR}{dt} = v_{SR} \bullet SR \bullet y(r) + m_{SR} \bullet S - m_{RS} \bullet SR - w \bullet SR$$

$$\frac{dC}{dt} = v_C \bullet C \bullet y(r) - d \bullet V \bullet C - m_{SR} \bullet C + m_{RS} \bullet CR - w \bullet C$$

$$\frac{dCR}{dt} = v_{CR} \bullet CR \bullet y(r) + m_{SR} \bullet C - m_{RS} \bullet CR - w \bullet CR$$

$$\frac{dCI}{dt} = v_{CI} \bullet CI \bullet y(r) + x \bullet d \bullet V \bullet C - w \bullet CI$$

$$\frac{dV}{dt} = d \bullet V \bullet (S + C) \bullet (b - 1) - d \bullet V \bullet CI - w \bullet V$$

$$where \;\; y(r) = \frac{r}{(r + k)}$$

<u>I. The evolution of CRISPR immunity and surface resistance in CRISPR$^+$ populations</u>

When sensitive populations of CRISPR$^-$ bacteria, S, are confronted with phage, resistant mutants, SR, will ascend and become the dominant bacterial population. In CRISPR$^+$ populations this dominant population will be either resistant bacteria, CR, or immune cells, CI, which are produced by the acquisition of a spacer from the phage. Whether resistant mutants or CRISPR-Cas immune cells will dominate depends on the rate of mutation to resistance, S→SR and C→CR, the rate of spacer acquisition, C→CI. If we assume all the bacterial populations are of equal fitness (have the same maximum growth rates), whether resistance or immunity will prevail will depend on whether resistant or immune bacteria appear first.

To illustrate this, we consider a continuous culture community of sensitive bacteria of density N cells per ml at equilibrium with lytic phage in a habitat where resources are sufficient for the rates of growth to be at their maximum. Under these conditions, the densities of the phage and bacteria populations would be,

$$V^* = \frac{(v-w)}{\delta} \; and \; N^* = \frac{w}{\delta \bullet \beta}$$

Where V^* is the equilibrium density of the phage and N^* the equilibrium density of the bacteria. For example, if the maximum growth rate of the bacteria, v =0.7 per hour, the adsorption rate constant, $\delta = 10^{-7}$, the burst size β=50, and the flow rate w= 0.1 per hour, the equilibrium densities of bacteria and phage would be, respectively $N^* = 2.0 \times 10^4$ and $V^* = 6.0 \times 10^6$.

For a population of CRISPR$^+$ sensitive bacteria at equilibrium with phage, C* and V*, whether resistant, CR or immune CI cells will appear first will depend on the rate of mutation to resistance, μ_{SR}, the likelihood of the bacteria picking up a spacer, the parameters δ and x and the total number of bacteria, C*•Vol. If μ_{SR}•C*•Vol > x•δ•V*•C*•Vol, resistance will more likely appear before CRISPR immunity. In general, for CRISPR-Cas immune cells to be generated before resistant ones,

$$x \bullet \delta \bullet V^* \bullet C^* > \mu_{SR} \bullet C^* \;\; or \; x > \frac{\mu_{SR}}{\delta \bullet V^*} \quad .$$

For example, if the mutation rate to resistance is $\mu_{SR}=10^{-8}$, with $\delta=10^{-7}$ and $V^*=6\times10^6$, for CI to appear before CR, the probability of acquiring a spacer upon infection, x has to exceed 1.67×10^{-8}.

<u>(i) Establishment of immunity and resistance in a CRISPR$^+$ population</u>

In Figure S2, we consider the establishment of immune and resistant bacteria, respectively CI and CR, in a CRISPR$^+$, phage sensitive C, population initially at equilibrium with the phage. With the parameters used, the likelihood of generating resistance, CR, and immunity CI are equal, $\mu_{SR} = 10^{-8}$, and $x=1.67\times10^{-8}$. For these simulations, we use a semi-stochastic version of this model where the generation of mutants or the acquisition of spacers are determined by a Monte Carlo process, with the rest of the transitions between states being deterministic.

Consequently, although all of the runs were started with the same conditions, the dynamics differed because mutation and the acquisition of spacers are stochastic processes. In Figure S2 A, B, and C we present runs with different outcomes, A where the immune cells become dominant population, B- where the resistant cells become the dominant population, and C where the resistant cells become the dominant population but a high density of immune cells are maintained. For a more comprehensive perspective of when these different outcomes are anticipated and the likelihood of their occurrence, we use a Monte Carlo simulation to determine the average density of immune and resistant cells at different times for 100 simulations. (Figure S2D). In the absence of resistance, $\mu_{SR}=\mu_{RS}=0$, within short order immune cells, CI, evolve and dominate the CRISPR^{+} population. When resistance and immune cells are can be generated, both can evolve to dominate the population. The reason that in this simulation, immune and resistant cells are equally likely to ascend to dominate the bacterial community is that with the parameters employed, they are equally likely to be generated.

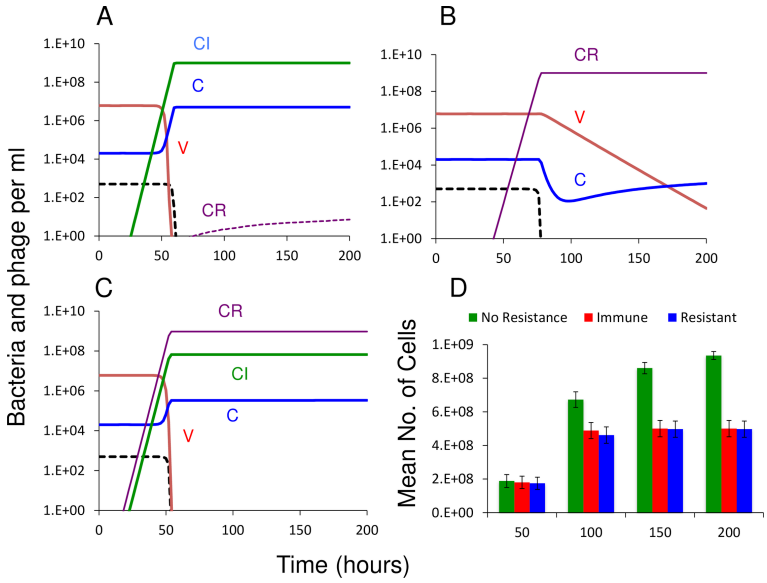

Figure S2: Establishment of immune, CI, and resistant CR cells in a sensitive population of CRISPR$^+$ bacteria C. Standard parameters $v_C=v_{CI}=v_{CR}=0.7$, $\delta=10^{-7}$, $\beta=50$, $e=5x10^{-7}$, k=1, RR=500, w=0.1, r(0) =500, $\mu_{SR}=10^{-8}$ $\mu_{RS}=10^{-8}$, $x=1.667x10^{-8}$, and the total volume of the vessel is Vol=100 ml. The initial densities of bacteria and phage in these simulations are those at the equilibrium for a phage-limited population, respectively $C=2x10^4$, $V=6x10^6$. (A, B and C) Changes in the densities of bacteria and phage, A - A run where, immunity, CI, arises and ascends to dominance before resistance, CR. B- A run where resistance, CR, emerges and ascends to dominance before immune cells are generated. C- A run where resistance emerges and ascends shortly before immune. D- The outcome of 100 independent simulations, mean and standard estimate of the density of CI and CR at different times. The green bars are for runs where resistance cannot be generated, $\mu_{SR}=0$, but immunity can be.

(ii). Establishment of CRISPR$^+$ bacteria in a CRISPR$^-$ populations

In Figure S3, A, B, and C, we consider the dynamics of the changes in the densities of the different populations of bacteria, S, SR, C, CI and CR, the phage V, and the concentration of the resource, r, for situations where resistance can be generated, $\mu_{SR} = \mu_{RS} = 10^{-8}$. In these simulations, initially, the sensitive populations are at equilibrium with the phage, $V= 6x10^6$, and the density of CRISPR$^+$ is 10% of the total population, respectively $C=2x10^3$ and $S=1.8x10^4$. CRISPR$^+$ resistant, CR, or immune, CI, cells can emerge and ascend to dominate the population. To elucidate the relationship between the initial frequencies of CRISPR$^+$ and CRISPR$^-$ on the conditions for CRISPR$^+$ cells to become established in CRISPR$^-$ population, we use Monte Carlo simulations. We follow the changes in the frequency of runs dominated by CRISPR$^+$ cells different at times with different initial frequencies of CRISPR+ cells, Figure

S3E and S3F, respectively, for the situations where resistant mutants cannot be generated, and where they can. In both cases, the initial CRISPR$^+$ cells were neither resistant nor immune, C.

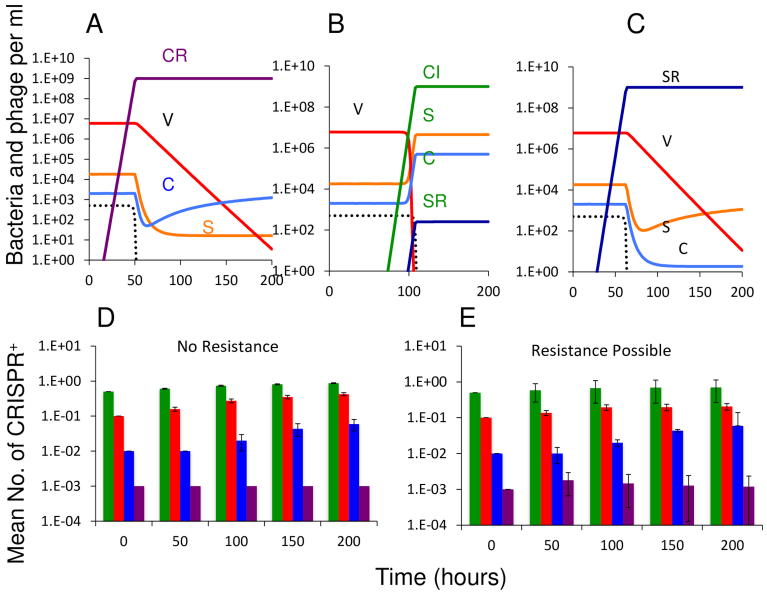

Figure S3 Establishment of a CRISPR$^+$ population in a CRISPR$^-$ phage sensitive population at equilibrium with a lytic phage: The parameter values in these simulations are the same as those in Figure S2. In this simulation, however, the equilibrium population of sensitive bacteria includes both CRISPR$^-$ and CRISPR$^+$ cells. A, B and C - Changes in the densities bacteria and phage in populations initiated with 10% CRISPR$^+$, C, and 90% CRISPR$^-$ S. A – CRISPR$^+$ resistant cells, CR, evolved to dominate. B- CRISPR$^+$ immune cells, CI, evolved to dominate. C - CRISPR$^-$ resistant cells, CR, evolved to dominate. D and Eb - Monte Carlo simulations. Mean and standard errors in the densities of CRISPR$^+$ bacteria in 100 simulated populations of CRISPR$^-$ initiated with different initial frequencies of CRISPR$^+$ C, and CRISPR$^-$ S at equilibrium with the phage. The Green, Orange, Blue, Purple, are respectively sets of runs where the initial frequencies of CRISPR$^+$, C cells, are 0.5, 0.1, 0.01, and 0.001. D- No resistance possible. E- resistance possible.

As our criterion for the establishment of CRISPR$^+$ in a CRISPR$^-$ population, we consider the ratio of the initial mean frequency of CRISPR$^+$ at time 0 to that at 200 hours. By this criterion, save for the runs where resistance cannot be generated, Figure S3 D, CRISPR$^+$ cell invaded at all initial frequencies. However, this criterion need not reflect successful invasions of CRISPR$^+$. Since all of the populations of bacteria equally fit, by chance alone there could be a small increase in the frequency of CRISPR$^+$. For example, in 200 runs with the initial frequency of CRISPR$^+$ of 10^{-4}, the mean the frequency of CRISPR$^+$ at 200 hours was 4.96 ±1.39 x10^{-6}. Moreover, in 100 runs with an initial frequency of CRISPR$^+$ of 10^{-3} with resistance possible, neither CRISPR$^+$ resistant nor immune cells, CR or CI were generated. On the other hand, with resistance possible and an initial frequency of CRISPR$^+$, C=0.01, CRISPR$^+$ resistant cells, CR, dominated the community by 200 hours, but immune, CI bacteria never dominated. With this initial frequency CRISPR$^+$ in this situation where resistance cannot be generated, by 200 hours, CRISPR$^+$ immune cells, CI, dominated three out of 100 runs. Overall, we interpret these simulation results to suggest that if the frequency of CRISPR$^+$ cells that are

neither resistant to immune to the phage is less than 10^{-3}, the likelihood of CRISPR successfully becoming established in a CRISPR$^-$ population at equilibrium with the phage is negligible. This is particularly important when considering the establishment of CRISPR$^+$ cells in CRISPR$^-$ populations by horizontal gene transfer when the initial frequency of CRISPR$^+$ cells is going to be low.

(ii) Caveats

In choosing the parameters for these simulations of lytic phage selecting for CRISPR-Cas mediated immunity, we are making two assumptions that can profoundly affect the predictions. One is that all of the populations are equally fit, have the same maximum growth rates. For example, if CRISPR-Cas engenders a cost relative to sensitive cells, the conditions for the establishment of immunity in a CRISPR$^-$ or CRISPR$^+$ population are going to be less than that predicted by this model. If CRISPR-Cas is less costly than envelope resistance, $v_{CI} > v_{CR}$, immunity is more likely to evolve than resistance. Two is the assumption that the likelihood of acquiring a spacer is equal to that of generating a resistant mutant. Clearly, if the likelihood of acquiring a spacer is greater than that of generating a resistant mutant, immunity will be more likely to evolve than resistance, and the reverse is true if the likelihood of generating a resistant mutant is greater than that of acquiring a spacer. Finally, and most importantly, for the conditions under which CRISPR$^+$ cells will become established in a CRISPR$^-$ population is the assumption that the populations of sensitive cells, S is at equilibrium with and with a density limited by the phage. If indeed, these CRISPR$^-$ populations were confronted by phage, they may already be dominated by resistant cells.

II. Population dynamics of temperate phage and bacteria with and without a CRISPR-Cas system.

The Model: In this model, diagrammed in Figure S4, there is a single population of temperate phage, P, and two types of bacteria: those that are CRISPR$^+$ (i.e., carry a functional CRISPR-Cas system) and those that are CRISPR$^-$ (i.e. lack a functional CRISPR-Cas system). The CRISPR$^-$ bacteria can exist in two states; they can either be sensitive non-lysogens, or they can be lysogens (carry the prophage), respectively S and L. CRISPR$^+$ populations can be present in three states: sensitive non-lysogens, C, lysogens CL, and CRISPR immune, CI. The phage can exist in three states, as free phage, P, as prophage in CRISPR$^-$ lysogens, L, or as CRISPR$^+$ lysogens, CL. The bacteria grow at maximum rates, v_S, v_L, v_C, v_{CL}, and v_{CI} per

cell per hour/ The phage adsorb to the bacteria at a rate proportional to the product of their density, that of the bacteria, and a rate parameter, δ cell x ml/hour (6). The S and C populations support the lytic replication of the phage. As in (7) with a probability λ ($0 \leq \lambda \leq 1$) upon infection with lytic phage, the S and C populations become lysogens, respectively L and CL. These lysogens are immune to super-infection with the temperate phage, as are the CRISPR$^+$ immune cells, CI; phage that infect these immune cells are removed from the population. CRISPR$^+$ immune cells, CI, are generated in two ways, from existing lysogens, at rate y per cell per hour, and by infection with P, with a probability x. In addition to being produced by lytic infections, with a burst size β, free temperate phage are generated by induction of the lysogens, at a rate i per cell per hour. When they lose the prophage, CRISPR$^-$ lysogens revert to sensitivity, S, and when they lose the prophage, CRISPR$^+$ lysogens revert to sensitivity, C. In this model, we assume CRISPR-Cas is not lost.

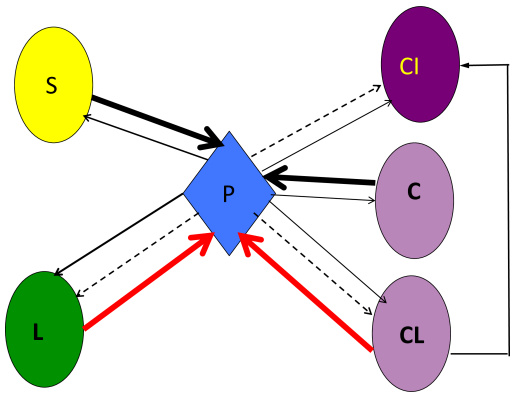

Figure S4. Model of the population dynamics of temperate phage and bacteria with and without CRISPR-Cas systems. P is the density of free phage, S, and L are respectively CRISPR$^-$ sensitive and lysogenic, C, CL, and CI are, respectively the designations and densities of CRISPR$^+$ bacteria that are sensitive to the free phage, lysogens, and immune cells. Solid red lines denote the production of free phage by induction of the lysogens. The thick solid black lines denote the production of phage by lytic infection. The thin broken lines denote the loss of the phage due to the adsorption to immune cells, L, CL, and CI. The spacers responsible for the immunity of non-lysogens, C, are picked up by infection with the phage and from CRISPR$^+$ lysogen.

With these definitions and assumptions, the rates of change in the densities of the bacterial populations, free phage, and the concentration of the limiting resource are given by the following set of coupled differential equations.

$$\frac{dr}{dt} = w \bullet (RR - r) - e \bullet \psi(r) \bullet (v_S \bullet S + v_L \bullet L + v_C \bullet C + v_{CL} \bullet CL + v_{CI} \bullet CI)$$

$$\frac{dS}{dt} = v_S \bullet S \bullet \psi(r) - \delta \bullet P \bullet S + i \bullet L - w \bullet S$$

$$\frac{dL}{dt} = v_L \bullet L \bullet \psi(r) + \delta \bullet P \bullet S \bullet \lambda - i \bullet L - w \bullet L$$

$$\frac{dC}{dt} = v_C \bullet C \bullet \psi(r) - \delta \bullet P \bullet C - w \bullet C$$

$$\frac{dCL}{dt} = v_{CL} \bullet CL \bullet \psi(r) + \delta \bullet P \bullet C \bullet \lambda - y \bullet CL - i \bullet CL - w \bullet CL$$

$$\frac{dCI}{dt} = v_{CI} \bullet CI \bullet \psi(r) + x \bullet \delta \bullet P \bullet C + y \bullet CL - w \bullet CI$$

$$\frac{dP}{dt} = \delta \bullet P \bullet (S + C) \bullet (1 - \lambda) \bullet \beta + i \bullet (L + CL) \bullet \beta - \delta \bullet P \bullet (CI + L + CL) - w \bullet P$$

$$where \quad \psi(r) = \frac{r}{(r + k)}$$

Temperate phage-mediated selection for CRISPR-Cas:

We open our analysis with an exploration of the conditions under which CRISPR-Cas immune cells will have an advantage in CRISPR$^+$ populations. When sensitive bacteria encounter temperate phage, most infections are initially lytic and CRISPR-Cas - mediated immunity would be favored because of the protection provided for sensitive cells. This can be seen in Figure S5. In all cases, free phage, P, select for the ascent of the immune cells, CI, which evolve to dominate the population at which time the phage are lost. In the case where there is no cost to the carriage of the prophage, free temperate phage, P, are maintained at a low density (Figure S5A). If there is a fitness cost associated with the carriage of the prophage, the lysogens decline in frequency (Figure S5B). If there is no fitness cost due to the carriage of the prophage, but the probability of lysogeny by infection of C is less than that of acquiring a spacer, $x < \lambda$ the lysogens continue to maintain their population but at a density less than when the probability of lysogeny exceeds that of the acquisition of spacers (Figure S5C). With selection against the lysogens, the ratio of immune to non-immune CRISPR$^+$ cells continues to increase at a rate proportional to the fitness cost of resistance (i.e., the greater the cost, the higher the rate). Even if there is no fitness cost due to the carriage of the prophage

when the density of free phage is high, CRISPR-mediated immunity is favored, but the rate of ascent is greater when the probability of acquiring spacers exceeds that of lysogeny (Figure S5D)

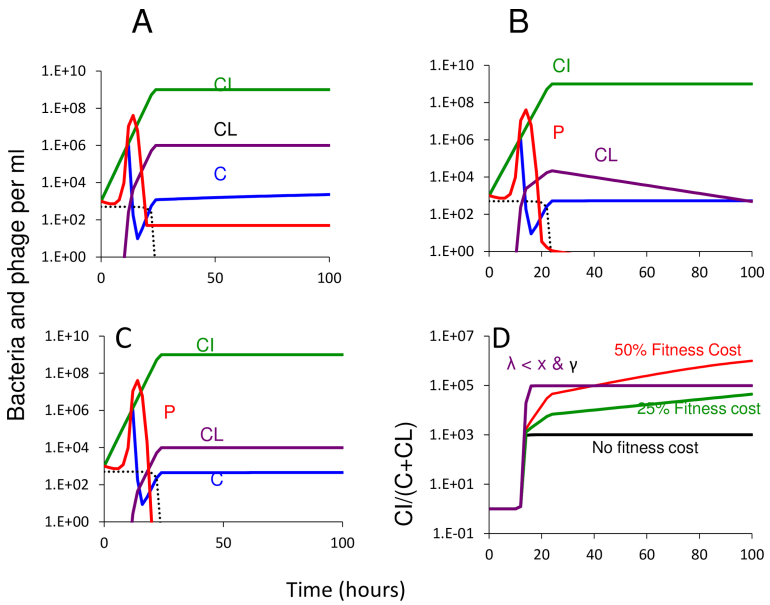

Figure S5 Temperate phage–mediated selection for CRISPR-Cas immunity in a CRISPR$^+$. Dynamics in establishing populations. Common parameters RR=500, w=0.1, k=1.0, e=5x10^{-7}, δ=10$^{-7,}$ β=50, i=10^{-4}. A- C -Changes in the densities of bacteria and phage, A- No cost associated with the carriage of the prophage, v_S=v_C=v_L=0.7, λ=10^{-3} x=y=10$^{-5.}$ B- 50% fitness costs for lysogens v_S=v_C=0.7, vl =0.35, λ=10^{-3} x=y=10^{-5}. (C) No cost associated with the carriage of the prophage, but the probaility of lysogeny is less than that of acquring spacers, v_S=v_C=v_L=0.7, λ=10^{-5} x=γ=10$^{-3.}$. D - Changes in the ratio of immune to non-immune CRISPR$^+$, CI/(C+CP) for different conditions.

In Figure S6A and B, we consider the ascent of immune cells, CI, in a CRISPR$^+$ population at equilibrium with a temperate phage, and in Figure S6C and D, the ascent of the CI in a CRISPR$^-$ population at equilibrium with the temperate phage. The dynamics of the ascent of immune cells in a CRISPR$^+$ population is similar to that in CRISPR$^-$, but the rate of ascent of the immune cells is faster when the dominant population is CRISPR$^+$ than when it is CRISPR$^-$ (compare Figure 6SA and Figure 6SC). The most significant difference between the establishment of CRISPR immune cells in a CRISPR$^+$ population and the establishment of

CRISPR$^+$ bacteria in a CRISPR$^+$ population at equilibrium with the temperate phage can be seen in Figures S6B and S6D. Under all conditions, the frequency of immune, CI, cells increases. The reason for this is that lysogens are continually converted into immune cells, CP →CI, albeit at a low rate, γ per cell per hour. The most significant rate of ascent of CRISPR-Cas immune cells in a CRISPR$^+$ population is observed when lysogeny is costly; the greater the cost, the greater the rate of ascent of immune cells (Figure S6B). On the other hand, the only condition for CRISPR$^+$ to invade a population of CRISPR$^-$ bacteria is when the carriage of the prophage engenders a fitness cost (Figure 6SD).

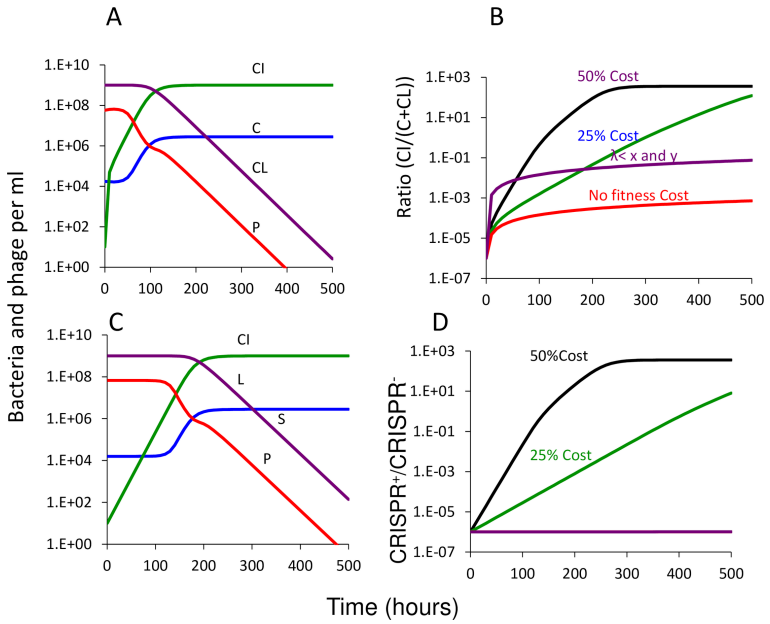

Figure S6 Dynamics of invasion of CRISPR-Cas mediated immunity into populations at equilibrium with temperate phage. Common parameters RR=500, w=0.1, k=1.0, e=5x10^{-7}, $\delta=10^{-7}$, β=50, i=10^{-4}. A- Dynamics of the establishment of CRISPR immune cells in a CRISPR$^+$ population at equilibrium with temperate phage with a 50% cost of lysogeny, $v_S=v_C=0.7$, $v_L=0.35$, $\lambda=10^{-3}$ $x=\gamma=10^{-5}$. B- Changes in the ratio of immune to non-immune CRISPR$^+$, CI/(C+CP) for different conditions. C- Dynamics of the establishment of CRISPR immune cells in a CRISPR$^-$ population at equilibrium with temperate phage with a 50% cost of lysogeny, $v_S=v_C=0.7$, $v_L=0.35$, $\lambda=10^{-3}$ $x=\gamma=10^{-5}$. D- Changes in the ratio of CRISPR$^+$/CRISPR$^-$ for different conditions.

III. Population dynamics of conjugative plasmids and bacteria with and without CRISPR-Cas mediated immunity:

The Model: There are two populations of CRISPR$^-$ cells, one that carries the plasmid and one that does not, DP, and S, respectively, and three populations of CRISPR$^+$ cells, plasmid-free, plasmid-bearing, and immune, respectively C, CP, and CI. These populations grow at maximum rates, v_S, v_P, v_C, v_{CP}, v_{CI} per cell per respectively. The plasmids are transferred at rates proportional to the product of the densities of plasmid-bearing and plasmid-free cells and a rate parameter, γ (2). CRISPR$^+$ cells, C, acquire spacers and become immune to infection with the plasmid, CI, at a rate proportional to the product of their densities, the rate constant of plasmid transfer, γ, and the probability of picking up a spacer x ($0 \leq x \leq 1$) upon conjugation. Immune CRISPR$^+$ cells can also be generated from plasmid-bearing CP at a rate y per cell per hour. Plasmids are lost by vegetative segregation at a rate, τ per cell per hour, resulting in DP cells reverting back to S and CP cells reverting back to C.

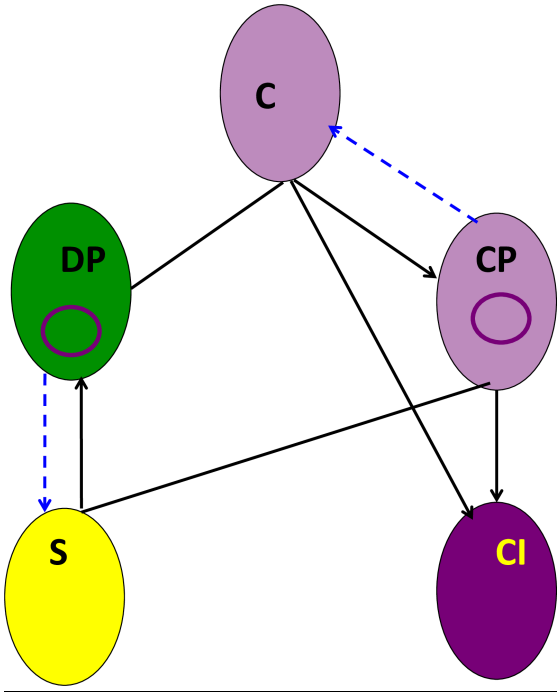

Figure S7 Model of the population dynamics conjugative plasmids and bacteria with and without CRISPR. There are two populations of CRISPR$^-$ cells, plasmid-free, and plasmid bearing, S, and DP. There are three populations of CRISPR$^+$ bacteria, those that are plasmid-free, those that carry the plasmid and those that are immune to the plasmid, respectively, C, CP, and CI. Plasmids are transferred to plasmid-free cells at a rate proportional to the product of their densities and a rate constant, γ, DP X S → DP X C→ CP, CP X S→DP and CP XC → CP. Immune, cells, CI are produced by infection of C by mating with a plasmid bearing cells or from a transition from CP to CI.

With these definitions and assumptions, a chemostat habitat and resource-limited growth, the rates of change in the densities of the different populations and concentration of the resource are given by the following array of coupled differential equations.

$$\frac{dr}{dt} = w \bullet (RR - r) - e \bullet y(r) \bullet (v_S \bullet S + v_{DP} \bullet DP + v_C \bullet C + v_{CP} \bullet CP + v_{CI} \bullet CI)$$

$$\frac{dS}{dt} = v_S \bullet y(r) \bullet S - g \bullet S \bullet (DP + CP) + t \bullet DP - w \bullet S$$

$$\frac{dDP}{dt} = v_{DP} \bullet y(r) \bullet DP + g \bullet S \bullet (DP + CP) - t \bullet DP - w \bullet DP$$

$$\frac{dC}{dt} = v_C \bullet y(r) \bullet C - g \bullet C \bullet (DP + CP) + t \bullet CP - g \bullet C \bullet (DP + CP) - w \bullet C$$

$$\frac{dCP}{dt} = v_{CP} \bullet y(r) \bullet CP + g \bullet C \bullet (DP + CP) - t \bullet CP - y \bullet CP - w \bullet CP$$

$$\frac{dCI}{dt} = v_{CI} \bullet y(r) \bullet CI + x \bullet g \bullet C \bullet (DP + CP) + y \bullet CP - w \bullet CI$$

$$where \;\; y(r) = \frac{r}{(r + k)}$$

When microbes carrying conjugative plasmids are introduced into receptive populations of plasmid-free cells in continuous culture, a stable equilibrium can obtain between plasmid bearing and plasmid-free cells (2). Whether that equilibrium will be obtained or the plasmid will be lost depends on the rate constant of transmission, γ, the segregation rate, τ, the rate of flow through the habitat, w, the fitness cost of the plasmid, and the density of the plasmid-free cells. With the parameters considered in the simulations, see the legend to Figure 8, there will be stable equilibria with plasmid-bearing and plasmid-free cells. With these parameters, the equilibrium points are almost independent of the relative fitness of the plasmid-bearing and plasmid-free cells. For example, with no cost, the equilibrium densities of plasmid-bearing and plasmid-free cells, are, respectively 1.00×10^9, and 1.00×10^5; With a 50% cost to the carriage of the plasmid, they are respectively, 9.99×10^8 and $1.1. \times 10^5$, and with a 50% advantage, these equilibrium densities are 1.00×10^9 and 9.52×10^4. On the other hand, these fitness costs have a marked effect on the conditions under which conjugative plasmids will select for immunity. This can be seen in Figure 8, where we explore the conditions under which (i) CRISPR immune cells will become established in CRISPR$^+$ populations at equilibrium with a conjugative plasmid, and (ii) CRISPR$^+$ cells will become established in CRISPR$^-$ populations at equilibrium with conjugative plasmids.

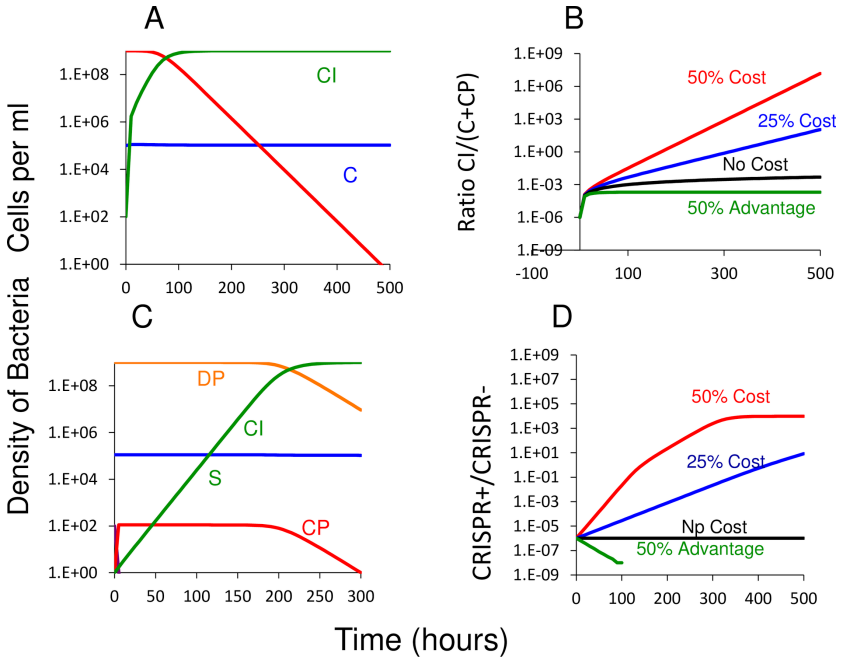

Figure S8 Plasmid-mediated selection for CRISPR-Cas. Common parameters, RR=500, k=1, $\gamma=10^{-9}$, w=0.1, e=5x10^{-7}, x=10^{-4}, y=10^{-4}, $\tau=10^{-4}$. All populations are initially at equilibrium with the plasmid, A- Establishment of immunity in a CRISPR$^+$ population. The densities of the different states of CRISPR$^+$ bacteria when there is a 50% fitness cost associated with the carriage of the plasmid, $v_C=v_{CI}=0.7$, $v_{CP}=0.35$. B-Changes in the ratio of CRISPR immune and non-immune cells for different costs due to the carriage of the plasmid 50% - $v_C=v_{CI}=0.7$, $v_P=0.35$, 25%- $v_C=v_I=0.7$, $v_P=0.525$, no cost, $v_C=v_I,=v_P=0.7$, plasmid advantage, $v_C=v_{CI}=0.35$, $v_P=0.7$. C- Establishment of immunity in a CRISPR$^+$ population in a CRISPR$^-$ population. The densities of the different states of bacteria when there is a 50% fitness cost associated with the carriage of the plasmid, $v_C=v_S=v_{CI}=0.7$, $v_{CP}=v_P=0.35$. D- Changes in the ratio of CRISPR$^+$ immune cells and CRISPR$^-$ cells for different costs due to the carriage of the plasmid 50% - $v_S=v_{CI}=v_C=0.7$, $v_P=v_{CP}=0.35$, 25%- $v_S=v_C=v_{CI}=0.7$, $v_P=v_{CP}=0.525$, no cost, $v_S=v_P=v_C=v_{CP}=v_{CI}=0.7$, plasmid advantage, $v_S=v_C=v_{CI}=0.35$, $v_P=v_{CP}=0.7$.

If the plasmid engenders a fitness cost, CRISPR immune cells, CI will rapidly increase in density in a CRISPR$^+$ population at equilibrium with the plasmid (Figure S8A). Whether the plasmids engender a cost or not, the immune cells, CI, will increase in frequency when they enter a CRISPR$^+$ population at equilibrium with the plasmid. This will not be the case if the plasmid is beneficial since plasmid bearing cells are at an advantage relative to plasmid-free ones. The rate of increase in the ratio of CI/(C+CP) is proportional to the fitness cost associated with the carriage of the plasmid (Figure S8B). The reason for the increase in the frequency of CI in the absence of a cost of plasmid carriage is because plasmid-bearing CRISPR$^+$ cells CP continually produce immune cells CP → CI, albeit at a low rate, y per cell per hour.

If CRISPR$^+$ immune cells, CI, enter a population of CRISPR$^-$ cells at equilibrium with the plasmid and there is a cost associated with the carriage of the plasmid, the immune cells will

increase in density (Figure S8C). The rate of increase in the density of CI is low compared to the corresponding situation of CI entering a CRISPR$^+$ population at equilibrium with the plasmid (compare Figure S8C with S8A). Moreover, the only condition under which CRISPR$^+$ microbes will invade a CRISPR-population at equilibrium with a conjugative plasmid is when the plasmid engenders a fitness cost (Figure S8D).

The Utility of the Models

Arguably, surely to us, the most important role of these heuristic models is to identify the factors, parameters, that govern the conditions under which CRISPR-Cas immunity will evolve and the relative contributions of these parameters to this evolution. The parameters of these models can be independently estimated in experimental populations of bacteria and archaea with lytic and temperate phage or conjugative plasmids and the hypotheses generated from their analysis tested in experimental populations of bacteria and archaea with phage and plasmids,

REFERENCES

1. Monod J (1949) The growth of bacterial cultures. *Annual Review of Microbiology* 3:371-394.
2. Stewart FM & Levin BR (1977) The population biology of bacterial plasmids: a priori conditions for the existence of conjugationally transmitted factors. *Genetics* 87:209-228.
3. Levin BR & Stewart FM (1977) Probability of establishing chimeric plasmids in natural populations of bacteria. *Science* 196(4286):218-220.
4. Gurney J, Pleska M, & Levin BR (2019) Why put up with immunity when there is resistance: an excursion into the population and evolutionary dynamics of restriction-modification and CRISPR-Cas. *Philos Trans R Soc Lond B Biol Sci* 374(1772):20180096.
5. Chaudhry WN*, et al.* (2018) Leaky Resistance and the conditions for the existience of lytic bacteriophage. *PLoS Biology* (IN PRESS).
6. Levin BR, Stewart FM, & Chao L (1977) Resource - limited growth, competition , and predation: a model and experimental studies with bacteria and bacteriophage. *American Naturalist* 977:3-24.
7. Stewart FM & Levin BR (1984) The population biology of bacterial viruses: why be temperate. *Theor Popul Biol* 26(1):93-117.