Check for updates

Game-Theoretical Modeling of Interviral Conflicts Mediated by Mini-CRISPR Arrays

Jaime Iranzo[1,2,3], Guilhem Faure[4], Yuri I. Wolf[3] and Eugene V. Koonin[3]*

[1] Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM) – Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain, [2] Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza, Spain, [3] National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, United States, [4] Broad Institute of MIT and Harvard, Cambridge, MA, United States

All cellular organisms coevolve with multiple viruses, so that both virus-host and intervirus conflicts are major factors of evolution. Accordingly, hosts evolve multiple, elaborate defense systems and viruses respond by evolving means of antidefense. Although less thoroughly characterized, several dedicated mechanisms of intervirus competition have been described as well. Recently, the genomes of some bacterial and archaeal viruses have been shown to harbor CRISPR mini-arrays that typically contain a single spacer targeting a closely related virus. The involvement of mini-arrays in an intervirus conflict has been experimentally demonstrated for a pair of archaeal viruses. We model the evolution of virus-encoded CRISPR mini-arrays using a game theoretical approach. Analysis of the model reveals multiple equilibria that include mutual targeting, unidirectional targeting, no targeting, cyclic polymorphism, and bistability. The choice between these evolutionary regimes depends on the model parameters including the coinfection frequency, differential productivity of the conflicting viruses, and the fitness cost of mini-arrays. At high coinfection frequencies, the model becomes a version of the Prisoner's dilemma in which defection, i.e., mutual targeting between the competing viruses, is the winning strategy.

Keywords: CRISPR-Cas, mini-CRISPR arrays, intervirus conflicts, superinfection exclusion, game theory

INTRODUCTION

Most microbial communities are associated with highly diverse and abundant viral populations (Perez Sepulveda et al., 2016; Duerkop, 2018; Pratama and van Elsas, 2018; Shkoporov and Hill, 2019). Although the ratio of virus particle counts to the counts of microbial cells varies within a broad range, in many habitats viruses outnumber cells by one to two orders of magnitude (Suttle, 2007; Rohwer et al., 2009; Wigington et al., 2016). A major consequence of the imbalance in viral and host abundances is that viruses of bacteria and archaea often have to compete for a limited number of host cells. To minimize the loss of productivity that occurs when two or more viruses have to share the limited resources of the same host cell, prokaryotic viruses have evolved mechanisms to prevent superinfection and cope with coinfecting viruses (Delbruck, 1945; Wagner, 1960; Nowak and May, 1994; Refardt, 2011). Among such mechanisms, those leading to superinfection exclusion in lysogenic viruses are the best studied ones, both experimentally

and theoretically, from a population dynamics perspective (Lu and Henning, 1994; Sun et al., 2006). Other strategies involve increasing virulence (Eshelman et al., 2010; Refardt, 2011), sequestering diffusible intracellular products (Turner and Chao, 1998, 2003), and inducing host defense systems, such as CRISPR-Cas, against the competing virus (Erdmann et al., 2014).

Analysis of viral genomes from databases and environmental samples has revealed that viruses can co-opt parts of the CRISPR-Cas antiviral defense systems from the host and use them as weapons against competing viruses. Specifically, some viruses and prophages contain mini-CRISPR arrays with 1 or 2 spacers that target sequences from related viruses, typically infecting the same host and sometimes isolated from the same environment (Faure et al., 2019). Viruses carrying mini-arrays lack *cas* genes. However, repeats in mini-CRISPR arrays are identical to those found in the host's CRISPR-Cas locus, suggesting compatibility of the viral mini-array with CRISPR-Cas components of the host. In some cases, pairs of viruses infecting the same host target each other through their mini-CRISPR arrays. Such mutual targeting supports the role of mini-CRISPR arrays as weapons in interviral conflicts which, combined with the *cas* genes from the host, target and destroy other viruses that could coinfect the same host. In the case of proviruses and other non-lytic viruses, mini-CRISPR arrays would constitute heterotypic superinfection exclusion mechanisms that promote long-term survival of the carrier by protecting the host from potentially lethal infections. Recently, intervirus competition among archaeal viruses carrying mini-arrays with cross-targeting spacers has been validated experimentally (Medvedeva et al., 2019).

We were interested in investigating the role of mini-arrays in interviral conflicts from an evolutionary cost vs. benefit perspective. To this end, here we develop a game-theoretical model of interviral competition mediated by mini-CRISPR arrays. The main goal of this work is to identify the conditions under which such virus-against-virus targeting becomes a successful strategy. Our results shed light on the ecological and evolutionary trade-offs that lead to the engagement of bacterial and archaeal viruses in CRISPR-mediated arms races. We provide testable predictions on the differential prevalence of mini-CRISPR arrays among viruses with different lifestyles.

RESULTS

Modeling CRISPR-Mediated Competition Among Bacterial and Archaeal Viruses

We present a minimal model of interviral conflict mediated by mini-CRISPR arrays. Because interviral conflicts are driven by mixed coinfection of hosts, the coinfection probability is the main ingredient of our model. Below, we introduce a parameterization of the coinfection probability based on Bayes' theorem and discuss it in detail. Then, we describe the mathematical model, which combines elements from game theory and ordinary differential equations, and obtain its steady state solutions for a general case. We apply the model to three scenarios that involve, respectively, two lytic viruses; a lytic virus and a temperate virus (or a provirus); and two temperate viruses or proviruses. Finally,

we briefly discuss some extensions of the model to account for the loss of CRISPR spacers and the competition among more than two viruses.

Coinfection Probability

Let us consider two viruses, A and B, both infecting the same host. Mixed coinfections, in which viruses A and B infect the same cell, are a central ingredient of the model. From the point of view of virus A, the coinfection probability $p_{B|A}$ is the probability that virus A encounters virus B during the infection process. Similarly, we define the coinfection probability for virus B, $p_{A|B}$, as the probability that virus B encounters A in the course of a single infection. From the perspective of the host population, the *virus-centric* coinfection probability $p_{B|A}$ can be understood as the conditional probability of finding a host infected by virus B given that the host is also infected by virus A. Applying Bayes' theorem, we get

$$p_{B|A} = \frac{p_B}{p_A}\, p_{A|B}$$

where p_A and p_B are the probabilities of finding hosts infected by virus A and virus B, respectively. It is clear from this equation that the coinfection probabilities $p_{B|A}$ and $p_{A|B}$ can take different values. Such differences are caused by differences in abundance, infectivity, spatial distribution, host range, and any other factor that introduces an imbalance in the prevalence of viral infections. All other factors being equal, the coinfection probability will be greater for the virus with the lower abundance, the more restricted distribution, or the narrower host range.

In the case of viruses that target each other via mini-CRISPR arrays, the outcome of a coinfection event likely depends on which virus infected the host first. To account for that, we split each coinfection probability into two components, that is, $p_{B|A} = p_{AB|A} + p_{BA|A}$ and $p_{A|B} = p_{AB|B} + p_{BA|B}$, where the subindices AB and BA indicate the sequence in which infections occurred. We introduce the parameters q_A and q_B to denote the fractions of the coinfection events in which viruses A and B were, respectively, the first to arrive. From these definitions, it follows that $p_{AB|A} = q_A \times p_{B|A}$ and $p_{BA|B} = q_B \times p_{A|B}$.

Mathematical Model

We model the competition among two groups of viruses (species or strains), each of which can possess or lack mini-CRISPR arrays against the other. For simplicity, and without loss of generality, we normalize the productivity of an infection by a single virus that lacks mini-CRISPR arrays to make it equal to 1 (this is an arbitrary choice because multiplying all productivities by the same factor does not affect the relative composition of the population). When two viruses coinfect the same host, the offspring of each virus depends on whether they possess or lack mini-CRISPR arrays targeting each other:

(a) In the absence of such arrays, the average productivity of each virus in a mixed coinfection is given by the parameter w (w_A for virus A, w_B for virus B). If coinfection does not affect the total yield and both viruses are equally competitive, we can simplify $w_A = w_B = 1/2$. By using different values of w_A and w_B, the model

can account for unequal competitive abilities ($w_A \neq w_B$), interference ($w_A + w_B < 1$) and facilitation ($w_A + w_B > 1$) between viruses.

(b) If only one virus has a mini-CRISPR array with a spacer against the other virus, the productivities are $1 - c$ for the virus that carries the array and 0 for the targeted virus. Parameter c (denoted as c_A for virus A and c_B for virus B) is the cost of having a mini-CRISPR array with a spacer that matches a sequence that is often closely similar to a sequence in the genome of the mini-array-carrying virus itself (although not identical because, in such a case, self-targeting would inactivate the virus' own genome; see Discussion).

(c) If both viruses target each other, we assume that "the first that arrives wins." This is biologically plausible because, whereas transcription and processing of mini-CRISPR arrays require some time, scanning and cleavage of the targeted viral DNA is very fast (Singh et al., 2016; Shibata et al., 2017; Raper et al., 2018; Vink et al., 2019). Therefore, unless both viruses enter the cell shortly after each other, the second virus will be incapacitated before its mini-CRISPR array can be transcribed. Mathematically, the average productivity of coinfecting viruses that target each other is q $(1 - c)$. In general, q and c are different for viruses A (q_A, c_A) and B (q_B, c_B). As explained above, q_A can be interpreted as the probability that virus A arrives first and virus B arrives later (and the reverse for q_B). Note, however, that the "first who arrives wins" assumption can be easily relaxed to allow for some degree of mutual destruction upon coinfection, in which case $q_A + q_B < 1$.

The parameters of the model are summarized in **Table 1**. Competition between viruses A and B can be modeled as an evolutionary game, with the strategies being having or not having a spacer against the other virus. The fitness values of each virus are given in **Table 2**, under the assumption that populations of viruses A and B are homogeneous with respect to their strategies (i.e., all members either possess or lack a spacer against the other virus).

TABLE 1 | Variables and parameters of the model.

Symbol	Description
$p_{B\|A}$	Coinfection probability for virus A
$p_{A\|B}$	Coinfection probability for virus B
c, c_A, c_B	Cost of the mini-CRISPR array
w, w_A, w_B	Viral productivity in a mixed coinfection
q, q_A, q_B	Probability that the virus survives when there is mutual targeting
x_A, x_B	Fraction of the viral population that harbors a mini-CRISPR array
P^*, P^{**}	Critical coinfection probabilities for the maintenance of a mini-CRISPR array

The cost of the mini-CRISPR array and the viral productivity in a mixed coinfection are defined in relation with the basal productivity in a pure infection.

TABLE 2 | Mean productivity of a virus (A), which encounters a second virus (B) with probability $p_{B|A}$, depending on the presence or absence of mutually targeting CRISPR arrays.

	Anti-A spacer in virus B	
	+	−
Anti-B spacer in virus A: +	$(1 - c_A)(1 - p_{B\|A} + q_A\, p_{B\|A})$	$1 - c_A$
−	$1 - p_{B\|A}$	$1 - (1 - w_A)\, p_{B\|A}$

Before proceeding with the formal analysis of the system, some immediate insight can be gained from inspecting the fitness matrix of the evolutionary game. If we consider a scenario in which virus B lacks anti-A spacers, the fitness of virus A is given by the second column of **Table 2**. Accordingly, virus A benefits from anti-B spacers if $p_{B|A} > c_A / (1 - w_A)$. Conversely, if the entire population of virus B carries anti-A spacers (first column of **Table 1**), virus A will benefit from anti-B spacers if $p_{B|A} > c_A/(q_A - q_A c_A + c_A)$. Thus, the outcome of the evolutionary process is determined by a trade-off between the cost of the CRISPR mini-array and the potential benefit conferred by the spacer against the competing virus during coinfections.

To analyze the general case of mixed populations, let us introduce new variables x_A and x_B, both with values between 0 and 1, which represent the fraction of viruses of each strain that harbor CRISPR spacers against the other strain. Conversely, $1 - x_A$ and $1 - x_B$ are the relative sizes of the spacer-free subpopulations. Focusing on strain A, the fitness of viruses with and without anti-B spacers (denoted as $f_{A,1}$ and $f_{A,2}$, respectively) results from multiplying the fitness matrix in **Table 2** by the fractions of viruses in strain B that harbor and lack anti-A spacers:

$$f_{A,1}(x_B) = (1 - c_A)(1 - p_{B|A} + q_A\, p_{B|A})\, x_B +$$
$$(1 - c_A)(1 - x_B) \qquad (1)$$

$$f_{A,2}(x_B) = (1 - p_{B|A})\, x_B + (1 - (1 - w_A)\, p_{B|A})(1 - x_B)$$

Because of the complexity of virus-host interaction networks in natural microbial communities, it is unlikely that the host ranges or geographical distributions of two viruses completely overlap (Flores et al., 2013; Munson-McGee et al., 2018). The existence of strain-exclusive reservoirs partially decouples the population dynamics of different viruses, such that the coinfection probabilities cannot be easily calculated by simply considering both viruses and their shared host. Moreover, even in simple models with one virus and one host, the addition of realistic features, such as heterogeneous spatial structure or antiviral defense systems, dramatically alters the infection dynamics, stabilizing virus and hosts populations toward an equilibrium state (Heilmann et al., 2012; Iranzo et al., 2015). Because of these reasons, we do not attempt to model the coinfection dynamics and, instead, treat the coinfection probabilities as independent parameters, under the assumption that these are governed, to a large extent, by the overlap between the ecological niches of the competing viruses. By adopting the framework of evolutionary game theory, we can study the fraction of viruses with and without spacers for

each strain without explicitly modeling the host population, the absolute sizes of viral populations, and the possible differences in the infection rates of each virus. Thus, the dynamics of the mixed population can be represented by the following replicator equations:

$$\begin{cases} \dot{x}_A = x_A \left(f_{A,1}\left(x_B\right) - \phi_A\left(x_A, x_B\right) \right) \\ \dot{x}_B = x_B \left(f_{B,1}\left(x_A\right) - \phi_B\left(x_A, x_B\right) \right) \end{cases} \tag{2}$$

where ϕ_A is the mean fitness of strain A, calculated as

$$\phi_A = x_A f_{A,1}\left(x_B\right) + (1 - x_A) f_{A,2}\left(x_B\right) \tag{3}$$

The mean fitness (ϕ_B) and class–specific fitness values ($f_{B,1}$ and $f_{B,2}$) for strain B are obtained by exchanging subindices A and B in Eqs 3 and 1, respectively. After some manipulation, the equations that describe the population dynamics can be rewritten in a more informative form as

$$\begin{cases} \dot{x}_A = x_A\,(1 - x_A)\,(a + b\,x_B) \\ \dot{x}_B = x_B\,(1 - x_B)\,(a' + b'\,x_A) \end{cases} \tag{4}$$

where

$$a = (1 - w_A)\,p_{B|A} - c_A \tag{5}$$

$$b = p_{B|A}\,\left(q_A - q_A c_A + c_A + w_A - 1\right) \tag{6}$$

$$a' = (1 - w_B)\,p_{A|B} - c_B \tag{7}$$

$$b' = p_{A|B}\,\left(q_B - q_B c_B + c_B + w_B - 1\right) \tag{8}$$

General Solution of the Model

A comprehensive analysis of the solutions of Eq. 4 indicates that the outcome of the evolutionary arms race between the two viruses is given by the relative values of the coinfection probabilities $p_{B|A}$ and $p_{A|B}$ with respect to two pairs of critical thresholds, P_A^* and P_A^{**} for virus A, and P_B^* and P_B^{**} for virus B. The expressions for these critical thresholds are

$$P^* = \frac{c}{1 - w} \tag{9}$$

$$P^{**} = \frac{c}{q - qc + c} \tag{10}$$

(note that we omit the subindices A and B for simplicity). An intuitive interpretation of these critical thresholds is that they represent the coinfection probabilities for which the benefit of targeting the competitor virus outweighs the cost of the CRIPSR array, either when the competitor lacks a cross-targeting spacer and hence does not engage in the arms race (P^*) or when it does (P^{**}). Depending on the coinfection rate, the evolutionary arms race can lead to 5 qualitatively different regimes (**Figure 1**):

(a) No targeting if $p_{B|A} < P_A^*$ and $p_{A|B} < P_B^*$
(b) Unidirectional targeting, with A targeting B if $p_{B|A} > P_A^*$ and $p_{A|B} < P_B^{**}$

(c) Unidirectional targeting, with B targeting A if $p_{B|A} < P_A^{**}$ and $p_{A|B} > P_B^*$
(d) Mutual targeting if $p_{B|A} > P_A^{**}$ and $p_{A|B} > P_B^{**}$
(e) Cyclic dominance, with an alternation of targeting and non-targeting subpopulations, if $P_A^* < p_{B|A} < P_A^{**}$ and $P_B^{**} < p_{A|B} < P_B^*$; or else, if $P_A^{**} < p_{B|A} < P_A^*$ and $P_B^* < p_{A|B} < P_B^{**}$

Two important observations are pertinent. First, with the exception of the last regime, the population always reaches an equilibrium state in which all viruses of a given type (A or B) either possess or lack the mini-CRISPR array. Second, for some parameter combinations, the equilibrium is not unique, that is, the system displays bistability. Specifically, if $P_A^* < p_{B|A} < P_A^{**}$ and $P_B^* < p_{A|B} < P_B^{**}$, unidirectional targeting, either from A to B or from B to A (but not mutual targeting), will evolve. Which of the two possible equilibria is reached in that bistable regime, depends on which virus acquired the spacer first. The existence of this regime implies that instances of unidirectional targeting can be found even if viruses A and B are identical in all their parameters. Alternatively, if $P_A^{**} < p_{B|A} < P_A^*$ and $P_B^{**} < p_{A|B} < P_B^*$, the population can evolve toward the regimes of mutual targeting or no targeting, and both outcomes are stable.

It follows from Eq. 9 that, in order to maintain a mini-CRISPR array, the probability of coinfection must be greater than the cost of the array. Moreover, possession of mini-CRISPR arrays is more advantageous for strains with low yields in mixed coinfections (low w_A) because those would benefit the most from curtailing other, more efficient coinfecting strains.

Competition Between Two Lytic Viruses

When considering 2 lytic viruses, the model can be simplified by noting that any of them can be the first to enter the host cell with the same probability, that is, $q = 1/2$. Moreover, empirical data indicate that viruses that target each other are typically closely related (Faure et al., 2019; Medvedeva et al., 2019). Therefore, it seems reasonable to assume equal fitness costs ($c_A = c_B = c$). Similarly, unless one of the viruses has specialized to exploit the other (which is unlikely given the close relatedness), similar coinfection yields ($w_A = w_B = w$) can be expected. With these simplifications, the threshold coinfection probabilities required for the maintenance of mini-CRISPR arrays (Eqs 9–10) become:

$$P^* = \frac{c}{1 - w} \tag{11}$$

$$P^{**} = \frac{2c}{1 + c} \tag{12}$$

The dependencies of these thresholds on the fitness cost and the productivity in mixed coinfections are plotted in **Figure 2**. Clearly, in most realistic scenarios, the critical coinfection probabilities have the same order of magnitude as the fitness cost (the exception would be for viruses whose productivity remains largely unaffected when they coinfect with a second virus, that is, if $w \approx 1$). Furthermore, the greater the degree of interference (that is, the smaller the value of w), the lower the critical coinfection threshold.

If both viruses can use the molecular machinery synthesized by the other virus, and in the absence of other sources of
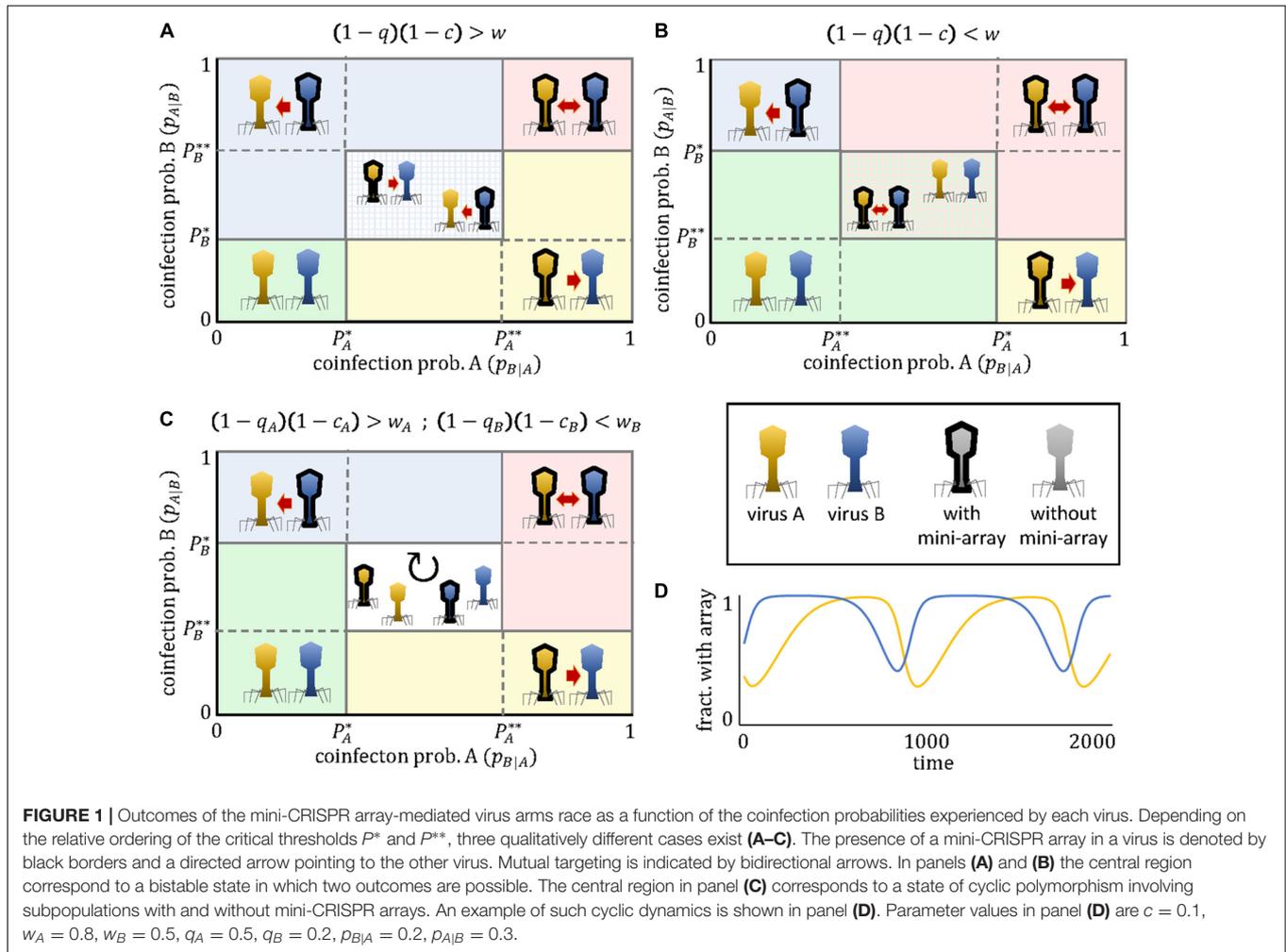
FIGURE 1 | Outcomes of the mini-CRISPR array-mediated virus arms race as a function of the coinfection probabilities experienced by each virus. Depending on the relative ordering of the critical thresholds P^* and P^{**}, three qualitatively different cases exist **(A–C)**. The presence of a mini-CRISPR array in a virus is denoted by black borders and a directed arrow pointing to the other virus. Mutual targeting is indicated by bidirectional arrows. In panels **(A)** and **(B)** the central region correspond to a bistable state in which two outcomes are possible. The central region in panel **(C)** corresponds to a state of cyclic polymorphism involving subpopulations with and without mini-CRISPR arrays. An example of such cyclic dynamics is shown in panel **(D)**. Parameter values in panel **(D)** are $c = 0.1$, $w_A = 0.8$, $w_B = 0.5$, $q_A = 0.5$, $q_B = 0.2$, $p_{B|A} = 0.2$, $p_{A|B} = 0.3$.
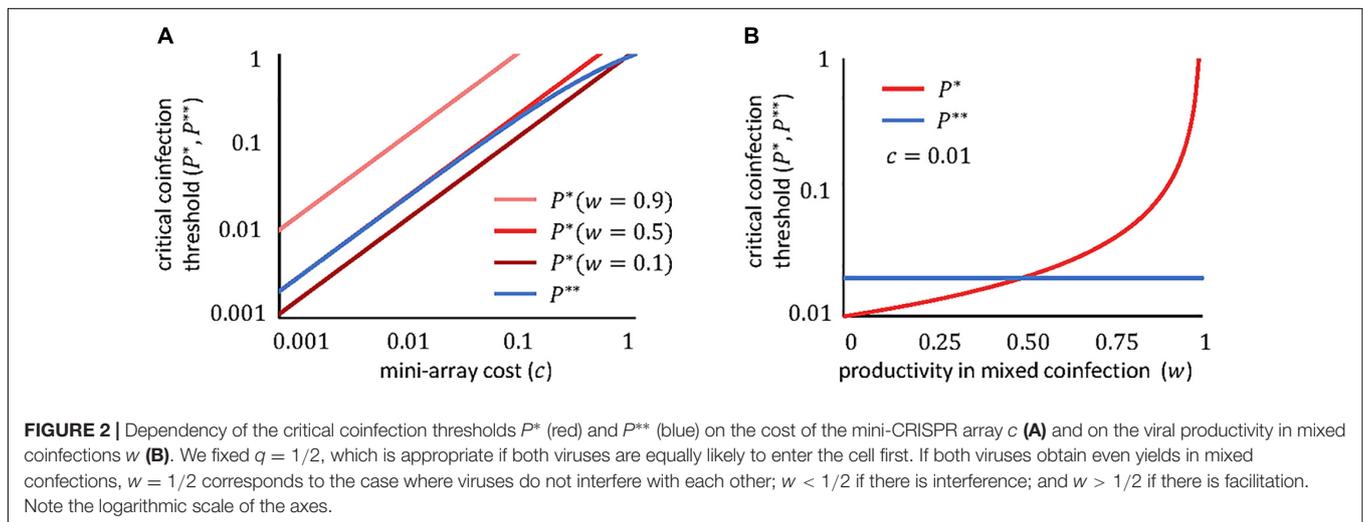

FIGURE 2 | Dependency of the critical coinfection thresholds P^* (red) and P^{**} (blue) on the cost of the mini-CRISPR array c **(A)** and on the viral productivity in mixed coinfections w **(B)**. We fixed $q = 1/2$, which is appropriate if both viruses are equally likely to enter the cell first. If both viruses obtain even yields in mixed confections, $w = 1/2$ corresponds to the case where viruses do not interfere with each other; $w < 1/2$ if there is interference; and $w > 1/2$ if there is facilitation. Note the logarithmic scale of the axes.

interference, the productivity of each virus in a mixed coinfection will be half of the productivity in a pure infection (that is, $w \approx 1/2$). If we also assume that the cost of the array is small ($c \ll 1$), the critical coinfection probabilities in Eqs 11–12 tend

to the same value, $P^{**} \approx P^* = 2c$. The possible evolutionary outcomes under this scenario become much simpler, with each virus maintaining a mini-CRISPR array against the other virus only if its coinfection probability is greater than twice the cost

of the array (**Figure 3**). In contrast, strong interference between coinfecting viruses ($w < 1/2$) introduces additional outcomes and broadens the range of conditions in which unidirectional targeting can evolve (**Figure 1A**).

Competition Between a Lytic and a Temperate Viruses

Let us now consider a scenario with one lytic (L) and one temperate (T) virus. By temperate virus, we refer both to proviruses integrated in the host genome and to non-lytic viruses that produce non-lethal chronic infections. Although the latter are particularly common in archaea (Prangishvili et al., 2017), actively replicating temperate phages also constitute some of the most abundant viruses in the human gut virome (Shkoporov et al., 2018). Based on the different lifestyles of lytic and temperate viruses, the most likely scenario for a mixed coinfection corresponds to the superinfection of a cell that already hosts a temperate virus by an incoming lytic virus. In this scenario, the mini-CRISPR arrays carried by the temperate virus, combined with the *cas* genes from the host, act as a heterotypic superinfection exclusion mechanism that protects the host from the lytic virus. Importantly, such a mechanism remains effective even if the lytic virus harbors a mini-CRISPR array against the "resident" temperate virus. In our model, this can be captured by setting $q_L = 0$ and $q_T = 1$. Note, however, that unidirectional targeting of the temperate virus by the lytic virus leads to the elimination of the temperate virus and production of pure offspring of the lytic virus. If the temperate virus is present as a provirus, its destruction will also lead to the degradation of the host genome; in that case, we assume that replication of the lytic virus remains unaffected by host genome degradation. Because the cost of carrying a spacer can differ among viruses with distinct lifestyles, we consider virus-specific cost parameters, c_L and c_T. Provirus induction is implicitly modeled by the parameter w_T, which jointly represents the probability of induction and the fraction of viral particles that successfully insert as a provirus in a new host.

As in the previous cases, the outcome of the evolutionary process depends on the relative values of the coinfection probabilities $p_{L|T}$ and $p_{T|L}$ with respect to three critical thresholds:
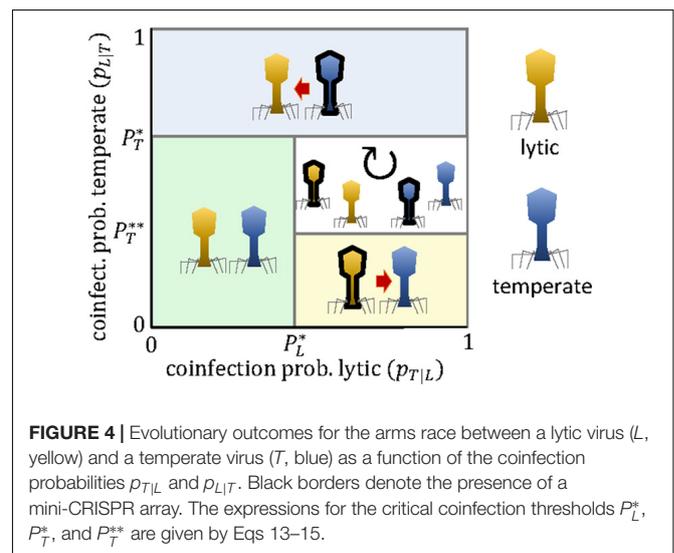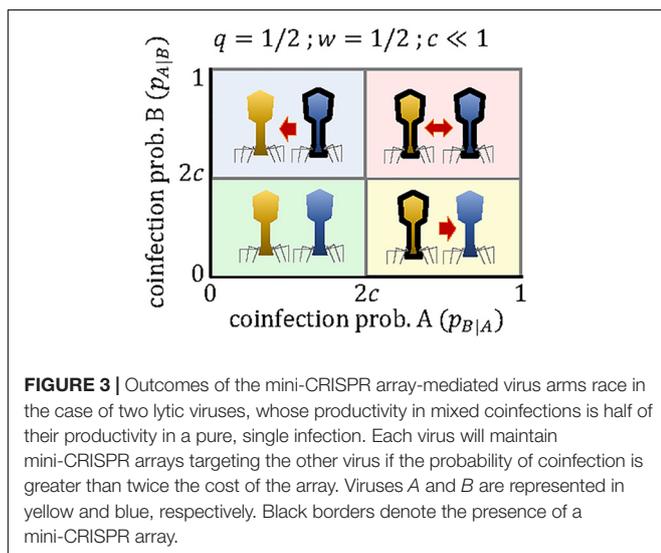
$$P_T^* = \frac{c_T}{1 - w_T} \tag{13}$$

$$P_T^{**} = c_T \tag{14}$$

$$P_L^* = \frac{c_L}{1 - w_L} \tag{15}$$

(there is an additional threshold for the lytic virus at $P_L^{**} = 1$, but it is irrelevant because, by definition, $p_{T|L} \leq 1$). Once again, P^* and P^{**} represent the critical coinfection probabilities beyond which it becomes profitable to target the other virus that either does (P^{**}) or does not engage (P^*) in the arms race via its mini-CRISPR array.

Figure 4 shows that competition among lytic and temperate viruses can lead to unidirectional targeting or cyclic polymorphism, but not to stable, reciprocal targeting. The reason is that, once the temperate virus engages in the arms race, the lytic virus does not benefit anymore from carrying the mini-CRISPR array.

The outcome of the competition between a lytic and a temperate virus depends on (i) whether the latter is an actively replicating non-lytic virus or a provirus. For a provirus, we must distinguish (ii) whether or not it provides the host with CRISPR-independent immunity against subsequent infection by the lytic virus, and (iii) whether such immunity can be overcome by a lytic virus that carries a mini-array against the temperate virus. Let us start with the case of an actively replicating non-lytic virus. Because the long-term productivity of chronic viruses is tightly linked to the survival of the host, superinfection by lytic viruses severely reduces their fitness ($w_T \ll 1$). In terms of the model, this implies that temperate viruses can maintain mini-arrays against lytic viruses as long as the coinfection probability

FIGURE 3 | Outcomes of the mini-CRISPR array-mediated virus arms race in the case of two lytic viruses, whose productivity in mixed coinfections is half of their productivity in a pure, single infection. Each virus will maintain mini-CRISPR arrays targeting the other virus if the probability of coinfection is greater than twice the cost of the array. Viruses A and B are represented in yellow and blue, respectively. Black borders denote the presence of a mini-CRISPR array.

FIGURE 4 | Evolutionary outcomes for the arms race between a lytic virus (L, yellow) and a temperate virus (T, blue) as a function of the coinfection probabilities $p_{T|L}$ and $p_{L|T}$. Black borders denote the presence of a mini-CRISPR array. The expressions for the critical coinfection thresholds P_L^*, P_T^*, and P_T^{**} are given by Eqs 13–15.

is greater than the cost of the mini-array ($P_T^* \approx P_T^{**} = c_T$). Thus, frequent encounters between the temperate and the lytic virus will result in unidirectional targeting from the former toward the latter. The same reasoning applies to those proviruses that do not provide immunity against a given lytic virus. However, many proviruses encode superinfection exclusion mechanisms that prevent infection by closely related viruses and, sometimes, even by distantly related ones (Dedrick et al., 2017; Gentile et al., 2019; Mavrich and Hatfull, 2019). Some of these mechanisms act at a very early stage of infection (e.g., inhibiting virus attachment or blocking the release of viral DNA into the cell), so that mini-CRISPR arrays would not help the incoming virus to superinfect the host. In those cases, CRISPR-mediated targeting becomes useless both for the temperate virus (which does not need it) and for the lytic virus (which does not benefit from it). This situation is not accounted for by the game matrix of the model (**Table 2**) but it is easy to see that it leads to no targeting among the competing viruses. Some superinfection exclusion mechanisms act at later stages of the infection process, which could allow the lytic virus to use mini-CRISPR arrays to target and destroy key proviral genes and escape immunity. This scenario equals to setting null productivity for the lytic virus upon coinfection in the absence of CRISPR-mediated targeting ($w_L = 0$, $w_T \approx 1$), and leads to a cyclic targeting dynamics (no targeting, lytic to provirus, mutual targeting, provirus to lytic, and no targeting again) if the coinfection probabilities for both viruses are greater than the cost of the mini-array ($P_T^* = \infty$, $P_T^{**} = c_T$, $P_L^* = c_L$).

Competition Between Two Temperate Viruses

Let us consider a scenario in which a temperate virus (A) infects a cell that already hosts another temperate virus or provirus (B). Unlike in the previous section, which dealt with viruses that qualitatively differ in their lifestyles, here, the reverse scenario in which B infects a cell that already hosts A is also possible and, once the population reaches equilibrium, equally likely. Accordingly, to apply the general model to 2 temperate viruses, we set the parameter $q = 1/2$. Additionally, to study competition among temperate viruses, it is necessary to distinguish between the cases of actively replicating non-lytic viruses and proviruses.

The model for the competition between two actively replicating temperate viruses is formally the same as for two lytic viruses, with the critical coinfection thresholds given by Eqs 11–12 (**Figure 2**). By contrast, when considering temperate viruses that become proviruses upon infection, unidirectional targeting of a resident provirus by an incoming virus leads to degradation of the host genome which, obviously, prevents the incoming virus from inserting into the host genome. This peculiarity of the competition between proviruses requires modifying the payoff matrix in **Table 2**, such that the fitness of virus A when it targets virus B becomes $(1 - c_A)(1 - p_{B|A} + q_A\, p_{B|A})$, regardless of whether virus B targets A or not. With this modification, after substituting $q = 1/2$, the critical threshold P^* for the competition of 2 proviruses becomes

$$P^* = \frac{2c}{1 + c - 2w} \qquad (16)$$

whereas P^{**} remains the same as in Eq. 12.

For both proviruses and actively replicating non-lytic viruses, the possible evolutionary outcomes will depend on the effect of coinfection for the long-term productivity of the temperate virus. In the case of actively replicating non-lytic viruses, it seems reasonable to assume that coinfection substantially decreases viral productivity. To get an approximate idea of how that affects the evolutionary outcome, if the productivity of the resident virus decreases to half of its original value, the evolutionary regimes will coincide with those shown for lytic viruses in **Figure 3**. Specifically, each virus will target the other if the coinfection probability is greater than twice the cost of the mini-array. In the case of proviruses, the quantitative effect of coinfection on long-term survival is more difficult to assess. In the absence of superinfection exclusion, as long as the arrival of a second provirus does not severely affect survival of the resident provirus ($w \geq 1/2$), the possible evolutionary outcomes would be those shown in **Figure 1B**, with $P^{**} \approx 2c$ and $P^* \geq 1$. Under this assumption, we would expect either no targeting or mutual targeting between proviruses (but not unidirectional targeting), with mutual targeting evolving only if the coinfection probability of both viruses is greater than twice the cost of the mini-array. The same outcomes are predicted if one or both proviruses possess superinfection exclusion mechanisms that can be overcome through CRISPR-mediated targeting (this scenario corresponds to setting $w = 1/2$, combining the facts that only the resident provirus can replicate and that the probability of arriving first is $1/2$). In contrast, if superinfection exclusion occurs and cannot be counteracted by CRISPR-mediated targeting, the benefit of harboring the mini-array vanishes for every provirus that encodes superinfection exclusion mechanisms (if only one does, the other can still benefit from the mini-array and evolve unidirectional targeting if the coinfection probability is greater than twice the cost of the mini-array).

Loss or Decay of CRISPR Spacers in Mini-Arrays

In the model discussed so far, the only factor that opposes the engagement of viruses into CRISPR-mediated arms races is the fitness cost of mini-arrays. Motivated by the fact that prokaryotic genome evolution displays a consistent deletion bias (Mira et al., 2001; Kuo and Ochman, 2009) and by specific observation on the loss of CRISPR spacers (Gudbergsdottir et al., 2011; Erdmann et al., 2014; Liu et al., 2016), here we extend the model to explore the contribution of non-adaptive loss or decay of mini-array spacers to the outcome of CRISPR-mediated interviral conflicts. Although the general solution becomes more complicated, the main finding is that, for a realistic range of parameter values, fitness cost and mini-array loss rate are exchangeable. Thus, all statement made above in terms of the mini-array cost, can be reinterpreted in terms of the spontaneous loss rate of mini-arrays.

Mini-array loss can be easily incorporated into the present framework as follows. If spacers are lost at a fixed rate d, Eq. 2 becomes a replicator-mutator equation:

$$\begin{cases} \dot{x}_A = x_A\left((1 - d)\,f_{A,1}(x_B) - \phi_A(x_A, x_B)\right) \\ \dot{x}_B = x_B\left((1 - d)\,f_{B,1}(x_A) - \phi_B(x_A, x_B)\right) \end{cases} \qquad (17)$$

with the same fitness parameters $f_{A,1}$, $f_{A,2}$, $f_{B,1}$, $f_{B,2}$, ϕ_A, ϕ_B as before. For simplicity, we present the equations for $w = 1/2$ and

omit subindices unless necessary to prevent ambiguity. The first critical coinfection threshold becomes

$$P_A^* = P_B^* = 2\left(c + d - cd\right) \tag{18}$$

The general expression for the second critical threshold becomes more complicated because it indirectly depends, for a given virus, on the coinfection probability for the other virus. Specifically,

$$P_A^{**} = \frac{2\left(c + d - cd\right)}{1 + \left(1 - d/s_B\right)\left(1 - 2\left(1 - c\right)\left(1 - q\right)\left(1 - d\right)\right)} \tag{19}$$

where

$$s_B = \frac{p_{A|B}/2 - c}{1 - c} \tag{20}$$

In these expressions, s_B is the relative reduction in the fitness of virus B associated with the loss of the CRISPR mini-array (i.e., the selection coefficient against CRISPR loss), provided that virus A does not engage in the arms race. Similar expressions are obtained for P_B^{**} and s_A by replacing subindices A and B. Due to their dependency on $p_{A|B}$ and $p_{B|A}$, the critical coinfection thresholds P_A^{**} and P_B^{**} are no longer straight lines (as represented in **Figures 1**, **3**, **4**) but curves in the $p_{A|B}$ vs. $p_{B|A}$ parameter space. Nevertheless, the parameter combinations that lead to no targeting, unidirectional targeting, and mutual targeting can still be obtained by applying the rules from **Figure 1**. Moreover, if the loss rate is small compared to the strength of selection ($d \ll s_A$, $d \ll s_B$), **Figure 1** also provides an approximate description of the bistable regimes (in that case, **Figure 1A** applies if $\left(1 - q\right)\left(1 - c\right)\left(1 - d\right) > 1/2$ and **Figure 1B** if $\left(1 - q\right)\left(1 - c\right)\left(1 - d\right) < 1/2$. For larger values of the loss rate, the bistable regimes are determined by a more convoluted set of equations (not shown here), and the range of parameters that allow for bistability shrinks as the loss rate grows.

The expressions from Eqs 18–20 fully apply to a loss-driven scenario where the mini-CRISPR array entails no cost ($c = 0$), and spacer loss is the only force counteracting selection for the mini-CRISPR array. To further simplify the final expressions, we assume that the loss rate is small compared to the strength of selection (note that, if $c = 0$, such condition becomes $2d \ll p_{A|B}$, $2d \ll p_{B|A}$ and it is expected to hold if coinfections are frequent enough to promote the maintenance of the mini-CRISPR array). These considerations lead to the critical thresholds $P^* = 2d$ and $P^{**} = d/(q - qd + d)$, which coincide with the expressions in the original model (Eqs 9–10) but with the cost of the array replaced by the loss rate.

To summarize, the qualitative results of the simple model are robust to the introduction of modest rates of spacer loss. An obvious quantitative difference is that, because loss of spacers continuously produces spacer-free viruses, the populations that engage in the arms race will be polymorphic. For example, in a unidirectional targeting scenario where strain A carries the anti-B spacer, only a fraction $x_A = 1 - d/s_A$ of virus A will have the spacer.

DISCUSSION

Competition among bacterial and archaeal viruses can lead to a mini-CRISPR array-mediated interviral arms race if coinfections are frequent enough. The model described here predicts how frequent coinfections must be so that it becomes advantageous for a virus to carry a mini-CRISPR array targeting a competitor. Precise quantification of such threshold requires measuring the values of two key parameters: (i) the cost of maintaining a mini-CRISPR array with the virus-targeting spacer, and (ii) the productivity of a virus lacking a mini-CRISPR array in a mixed coinfection relative to the productivity of that same virus in isolation. For two lytic viruses, a necessary (and often sufficient) condition for the emergence of mutual targeting is that the coinfection probability is greater than twice the cost of having the array. The same condition holds for the emergence of mutual targeting between two temperate viruses (either proviruses or actively replicating non-lytic viruses). For the competition between lytic and temperate viruses, the model predicts evolution of unidirectional targeting, from the temperate virus to the lytic virus, or cyclic targeting dynamics if the coinfection probability exceeds the cost of the mini-array. Mutual targeting between lytic and temperate viruses is unstable because lytic viruses only benefit from targeting a temperate virus if the latter does not already target the lytic virus. These results emphasize the interplay between superinfection exclusion and CRISPR-mediated targeting in interviral conflicts that involve prophages. Indeed, the outcome of the mini-CRISPR array-mediated arms race critically depends on whether the competing viruses encode mechanisms to prevent superinfection and whether such mechanisms can be vanquished by viruses that carry mini-CRISPR arrays.

Our analysis sheds light on the effect of viral interference and facilitation on interviral conflicts, and predicts that mini-CRISPR arrays are more likely to evolve in pairs of closely related viruses that strongly interfere with each other when coinfecting the same host cell. Another prediction that arises from the model is that mini-CRISPR arrays should be more frequent in "specialist" viruses (viruses with narrow host and geographical ranges) than in "generalist" ones, especially if the virus-host interaction network is nested as suggested by empirical data (Flores et al., 2013). The reason is that specialist viruses in a nested virus-host interaction network are subject to a higher frequency of coinfection and have more predictable coinfecting partners than generalist viruses. This prediction of the model will be put to test as data on virus host ranges accumulate.

Empirically observed instances of virus pairs with mutually targeting CRISPR-arrays often involve closely related viruses (based on the similarity of the terminase large subunit) with high sequence identity in their protospacer regions (Faure et al., 2019). As a result, a major contribution to the cost of the array could result from primed adaptation (Datsenko et al., 2012) induced by the spacer against its own carrier. The effects of primed adaptation on the propagation of phages and plasmids strongly depend on the model system, the spacer sequence, and the genetic background of the host (Datsenko et al., 2012; Fineran et al., 2014; Xue et al., 2015; Strotskaya et al., 2017). Overall,

the available experimental evidence suggests that mobile genetic elements could incur a substantial fitness cost (on the order of 0.1 or greater) for carrying an array with imperfect self-targeting spacers (Fineran et al., 2014; Xue et al., 2015). If these observations can be extrapolated to natural conditions in hosts other than *E. coli*, they would imply that the CRISPR-mediated arms race among closely related viruses can only be supported if coinfections with a competing virus represent a substantial fraction (about 10% or greater) of the infections by a given virus, or if spacers are divergent enough from the homologous sequence in the virus' own genome to avoid primed adaptation. An alternative possibility is that viruses that employ CRISPR arrays to target competitors possess mechanisms to inhibit primed adaptation, as has been suggested for bacteriophage T7 (Strotskaya et al., 2017). In the absence of primed adaptation, the cost of mini-CRISPR arrays is probably small. In that case, the spread of mini-arrays can still be limited by spacer loss due to recombination between the repeats. When modified to include spacer loss, the conclusions of our study remain qualitatively unchanged, in the sense that, for practical purposes, the loss rate and the cost of the mini-array are exchangeable parameters.

The model described here is simple, yet capable to produce a broad variety of evolutionary outcomes. Specifically, equilibrium states with mutual targeting, unidirectional targeting, no targeting, cyclic polymorphism, and bistability are recovered. The asymmetry that leads to unidirectional targeting most often results from different coinfection probabilities for each virus. Thus, if virus *A* encounters virus *B* more frequently than *B* encounters *A* (e.g., because virus *B* has a broader range or greater abundance), it is more likely that only *A* would carry a spacer against *B*. Other causes that could lead to asymmetric outcomes would be unequal fitness costs and different yields in coinfection (Refardt, 2011), with lower yields promoting the maintenance of CRISPR mini-arrays. Furthermore, unidirectional targeting could evolve among viruses that are functionally and ecologically equivalent (i.e., equal in all their parameters) if the cost of the CRISPR array is outweighed by the benefit of targeting an array-free, but not an array-carrying, virus.

The model can be readily generalized to include three or more competing viruses with different coinfection probabilities. In that case, the outcome of the competition depends on how the fitness cost scales with the number of targeted viruses. In particular, if the cost increases linearly, the problem can be reduced to a combination of pairwise interactions, each of which can be separately analyzed with the 2-virus model. More specifically, the equations of the 2-virus model can be simply applied to each pair of viruses to determine if they carry spacers against the other. A notable aspect of the intervirus arms race becomes manifest in the regime of frequent coinfections with mutual targeting. In this regime, viruses engage in an analog of the Prisoner's Dilemma (Axelrod and Hamilton, 1981; Maynard Smith, 1982), where having and not having a mini-CRISPR array targeting the competing virus becomes analogous to defecting and "cooperating" (not defecting), respectively. When coinfections are frequent, the best strategy for each virus is to target the competitor (defect), which leads to the evolution of reciprocal targeting. Paradoxically, the overall fitness that both

viruses obtain when they defect and target each other is lower than the fitness that they would obtain if they both pursued their mutual benefit and did not engage in an arms race. Qualitatively similar outcomes (although based on different mechanisms) have been previously observed in the experimental evolution of phages at high multiplicity of infection (Turner and Chao, 1999) or in metapopulations with unrestricted migration (Kerr et al., 2006; Eshelman et al., 2010), suggesting that evolutionary dilemmas leading to suboptimal overall fitness could be a characteristic (although not universal, see Turner and Chao, 2003) feature of interviral conflicts.

Notwithstanding the possible sampling bias, most empirical instances of viral mini-CRISPR arrays targeting other viruses involve prophages or non-lytic archaeal viruses. Because the formal conditions required for the evolution of mini-CRISPR arrays in lytic and temperate viruses are not dramatically different, we propose that the higher prevalence of mini-CRISPR arrays in temperate viruses results from a greater frequency of coinfection in these viruses. Indeed, single-cell sequencing of a hyperthermophilic archaeal community has recently shown that most archaeal cells harbor one or more viruses (Munson-McGee et al., 2018) which are, presumably, temperate. We expect that future studies quantifying the frequency of coinfection in environmental samples will contribute to clarify whether lytic and temperate viruses substantially differ in their heterotypic coinfection rates. A second factor that could help explaining the higher prevalence of mini-CRISPR arrays in temperate viruses is their closer and temporally extended association with the host, which would give temperate viruses more opportunities to integrate spacers through recombination with the host CRISPR array and/or via CRISPR adaptation. One or the other of these spacer acquisition mechanisms are likely to operate in different virus-host systems. In mini-array carrying phages, the leader sequences in front of the mini-arrays differ from those in the hosts, presumably precluding adaptation and limiting spacer acquisition to the recombinational route (Faure et al., 2019). By contrast, in archaeal viruses, the leaders are identical to those of the host suggesting that adaptation could occur directly in the virus mini-array (Medvedeva et al., 2019).

From an evolutionary perspective, our model addresses the role of selection and neutral loss in the fate of mini-CRISPR arrays. However, it does not include other sources of genotypic variation that fuel the evolutionary process, such as mini-array acquisition and mutation of target sequences. The acquisition of spacers was not included in our model because of the uncertainty about the mechanism and also because it is unclear whether spacer capture is a limiting factor for the observed distribution of mini-CRISPR arrays in viral taxa. This aspect will need to be revisited as more empirical evidence becomes available. Another relevant process that is not covered in this work is the evolutionary escape from CRISPR targeting through mutations in viral proto-spacers. Because the long-term efficacy of proto-spacer mutation as an escape strategy depends on the rate at which new spacers are acquired, the design of more realistic coevolutionary models will also require a better understanding of the dynamics of spacer and mini-array acquisition in viral populations. Further refinements to the

model should also include an explicit consideration of the host and virus population dynamics and how changes in abundance reflect on coinfection probabilities. Such an explicit modeling of time-dependent coinfection dynamics, which is not part of the present model, can shed light on whether the larger fluctuations of host populations associated with the proliferation of lytic viruses lead to fundamental differences in the fate and efficiency of mini-CRISPR arrays, as predicted for "regular," cell-borne arrays (Iranzo et al., 2013).

There could be additional potential benefits that mini-CRISPR arrays might provide to the carrier virus and that are independent of interviral conflict but rather have to do with overcoming the host defenses. For example, high levels of transcription of the mini-CRISPR array could potentially lead to competitive exclusion of the host spacers from the CRISPR-Cas effector modules, protecting the virus from targeting. Additionally, repeats in the mini-CRISPR array could promote the insertion of the virus within the host CRISPR array, thus impairing the host immune response; indeed, insertion of proviruses into CRISPR arrays has been observed in several bacterial genomes (Faure et al., 2019; Varble et al., 2019). These scenarios could be immediately incorporated into the present framework as a negative-signed contribution to the fitness cost of the mini-array (parameter c in the model). Nevertheless, the fact that the spacers in CRISPR mini-arrays typically target closely related viruses and the observed instances of mutual targeting strongly suggest that the main role of mini-CRISPR arrays is related to interviral conflict.

Finally, it is worth noting that the modeling framework developed here can be generalized beyond viruses, to plasmids that might carry mini-arrays (Lillestol et al., 2006) and beyond the CRISPR mini-arrays, to any means of intervirus competition.

Several recent studies suggest that viruses indeed employ multiple, diverse mechanisms in the arms races with their competitors (Berngruber et al., 2010; Gentile et al., 2019; Mavrich and Hatfull, 2019; Montgomery et al., 2019), and investigation of these mechanisms using the game theoretical framework can enhance our understanding of the evolution of inter-virus conflicts.

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

JI and GF conceived the study. JI constructed the model. JI, GF, YW, and EK analyzed the results. JI and EK wrote the manuscript that was edited and approved by all authors.

FUNDING

REFERENCES

Axelrod, R., and Hamilton, W. D. (1981). The evolution of cooperation. *Science* 211, 1390–1396. doi: 10.1126/science.7466396

Berngruber, T. W., Weissing, F. J., and Gandon, S. (2010). Inhibition of superinfection and the evolution of viral latency. *J. Virol.* 84, 10200–10208. doi: 10.1128/JVI.00865-10

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3:945. doi: 10.1038/ncomms1937

Dedrick, R. M., Jacobs-Sera, D., Bustamante, C. A., Garlena, R. A., Mavrich, T. N., Pope, W. H., et al. (2017). Prophage-mediated defence against viral attack and viral counter-defence. *Nat. Microbiol.* 2:16251. doi: 10.1038/nmicrobiol. 2016.251

Delbruck, M. (1945). Interference between bacterial viruses: III. The mutual exclusion effect and the depressor effect. *J. Bacteriol.* 50, 151–170. doi: 10.1128/jb.50.2.151-170.1945

Duerkop, B. A. (2018). Bacteriophages shift the focus of the mammalian microbiota. *PLoS Pathog.* 14:e1007310. doi: 10.1371/journal.ppat.1007310

Erdmann, S., Le Moine Bauer, S., and Garrett, R. A. (2014). Inter-viral conflicts that exploit host CRISPR immune systems of *Sulfolobus. Mol. Microbiol.* 91, 900–917. doi: 10.1111/mmi.12503

Eshelman, C. M., Vouk, R., Stewart, J. L., Halsne, E., Lindsey, H. A., Schneider, S., et al. (2010). Unrestricted migration favours virulent pathogens in experimental metapopulations: evolutionary genetics of a rapacious life history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 2503–2513. doi: 10.1098/rstb.2010.0066

Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019). CRISPR-Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* 17, 513–525. doi: 10.1038/s41579-019-0204-7

Fineran, P. C., Gerritzen, M. J., Suarez-Diez, M., Kunne, T., Boekhorst, J., van, S. A., et al. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E1629–E1638. doi: 10.1073/pnas.140007 1111

Flores, C. O., Valverde, S., and Weitz, J. S. (2013). Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *ISME J.* 7, 520–532. doi: 10.1038/ismej.2012.135

Gentile, G. M., Wetzel, K. S., Dedrick, R. M., Montgomery, M. T., Garlena, R. A., Jacobs-Sera, D., et al. (2019). More evidence of collusion: a new prophage-mediated viral defense system encoded by mycobacteriophage sbash. *mBio* 10:e00196-19. doi: 10.1128/mBio.00196-19

Gudbergsdottir, S., Deng, L., Chen, Z., Jensen, J. V., Jensen, L. R., She, Q., et al. (2011). Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* 79, 35–49. doi: 10.1111/j.1365-2958.2010.07 452.x

Heilmann, S., Sneppen, K., and Krishna, S. (2012). Coexistence of phage and bacteria on the boundary of self-organized refuges. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12828–12833. doi: 10.1073/pnas.1200771109

Iranzo, J., Lobkovsky, A. E., Wolf, Y. I., and Koonin, E. V. (2013). Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *J. Bacteriol.* 195, 3834–3844. doi: 10.1128/JB.00 412-13

Iranzo, J., Lobkovsky, A. E., Wolf, Y. I., and Koonin, E. V. (2015). Immunity, suicide or both? Ecological determinants for the combined evolution of anti-pathogen defense systems. *BMC Evol. Biol.* 15:43. doi: 10.1186/s12862-015-0324-2

Kerr, B., Neuhauser, C., Bohannan, B. J., and Dean, A. M. (2006). Local migration promotes competitive restraint in a host-pathogen 'tragedy of the commons'. *Nature* 442, 75–78. doi: 10.1038/nature04864

Kuo, C. H., and Ochman, H. (2009). Deletional bias across the three domains of life. *Genome Biol. Evol.* 1, 145–152. doi: 10.1093/gbe/evp016

Lillestol, R. K., Redder, P., Garrett, R. A., and Brugger, K. (2006). A putative viral defence mechanism in archaeal cells. *Archaea* 2, 59–72. doi: 10.1155/2006/542818

Liu, G., She, Q., and Garrett, R. A. (2016). Diverse CRISPR-Cas responses and dramatic cellular DNA changes and cell death in pKEF9-conjugated *Sulfolobus* species. *Nucleic Acids Res.* 44, 4233–4242. doi: 10.1093/nar/gkw286

Lu, M. J., and Henning, U. (1994). Superinfection exclusion by T-even-type coliphages. *Trends Microbiol.* 2, 137–139. doi: 10.1016/0966-842x(94)90601-7

Mavrich, T. N., and Hatfull, G. F. (2019). Evolution of superinfection immunity in cluster a mycobacteriophages. *mBio* 10:e00971-19. doi: 10.1128/mBio.00971-19

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge, MA: Cambridge University Press.

Medvedeva, S., Liu, Y., Koonin, E. V., Severinov, K., and Krupovic, M. (2019). Virus-borne mini-CRISPR arrays are involved in intervial conflicts. *Nat. Commun.* 10:5204. doi: 10.1038/s41467-019-13205-2

Mira, A., Ochman, H., and Moran, N. A. (2001). Deletional bias and the evolution of bacterial genomes. *Trends Genet.* 17, 589–596. doi: 10.1016/s0168-9525(01)02447-7

Montgomery, M. T., Guerrero Bustamante, C. A., Dedrick, R. M., Jacobs-Sera, D., and Hatfull, G. F. (2019). Yet more evidence of collusion: a new viral defense system encoded by gordonia phage CarolAnn. *mBio* 10, e2417–e2418. doi: 10.1128/mBio.02417-18

Munson-McGee, J. H., Peng, S., Dewerff, S., Stepanauskas, R., Whitaker, R. J., Weitz, J. S., et al. (2018). A virus or more in (nearly) every cell: ubiquitous networks of virus-host interactions in extreme environments. *ISME J.* 12, 1706–1714. doi: 10.1038/s41396-018-0071-2

Nowak, M. A., and May, R. M. (1994). Superinfection and the evolution of parasite virulence. *Proc. Biol. Sci.* 255, 81–89. doi: 10.1098/rspb.1994.0012

Perez Sepulveda, B., Redgwell, T., Rihtman, B., Pitt, F., Scanlan, D. J., and Millard, A. (2016). Marine phage genomics: the tip of the iceberg. *FEMS Microbiol. Lett.* 363:fnw158. doi: 10.1093/femsle/fnw158

Prangishvili, D., Bamford, D. H., Forterre, P., Iranzo, J., Koonin, E. V., and Krupovic, M. (2017). The enigmatic archaeal virosphere. *Nat. Rev. Microbiol.* 15, 724–739. doi: 10.1038/nrmicro.2017.125

Pratama, A. A., and van Elsas, J. D. (2018). The 'Neglected' soil virome - potential role and impact. *Trends Microbiol.* 26, 649–662. doi: 10.1016/j.tim.2017.12.004

Raper, A. T., Stephenson, A. A., and Suo, Z. (2018). Functional insights revealed by the kinetic mechanism of CRISPR/Cas9. *J. Am. Chem. Soc.* 140, 2971–2984. doi: 10.1021/jacs.7b13047

Refardt, D. (2011). Within-host competition determines reproductive success of temperate bacteriophages. *ISME J.* 5, 1451–1460. doi: 10.1038/ismej.2011.30

Rohwer, F., Prangishvili, D., and Lindell, D. (2009). Roles of viruses in the environment. *Environ. Microbiol.* 11, 2771–2774. doi: 10.1111/j.1462-2920.2009.02101.x

Shibata, M., Nishimasu, H., Kodera, N., Hirano, S., Ando, T., Uchihashi, T., et al. (2017). Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nat. Commun.* 8:1430. doi: 10.1038/s41467-017-01466-8

Shkoporov, A. N., and Hill, C. (2019). Bacteriophages of the human gut: the "Known Unknown" of the microbiome. *Cell Host Microbe* 25, 195–209. doi: 10.1016/j.chom.2019.01.017

Shkoporov, A. N., Khokhlova, E. V., Fitzgerald, C. B., Stockdale, S. R., Draper, L. A., Ross, R. P., et al. (2018). PhiCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides* intestinalis. *Nat. Commun.* 9:4781. doi: 10.1038/s41467-018-07225-7

Singh, D., Sternberg, S. H., Fei, J., Doudna, J. A., and Ha, T. (2016). Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* 7:12778. doi: 10.1038/ncomms12778

Strotskaya, A., Savitskaya, E., Metlitskaya, A., Morozova, N., Datsenko, K. A., Semenova, E., et al. (2017). The action of *Escherichia coli* CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Res.* 45, 1946–1957. doi: 10.1093/nar/gkx042

Sun, X., Gohler, A., Heller, K. J., and Neve, H. (2006). The ltp gene of temperate Streptococcus thermophilus phage TP-J34 confers superinfection exclusion to Streptococcus thermophilus and Lactococcus lactis. *Virology* 350, 146–157. doi: 10.1016/j.virol.2006.03.001

Suttle, C. A. (2007). Marine viruses–major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Turner, P. E., and Chao, L. (1998). Sex and the evolution of intrahost competition in RNA virus phi6. *Genetics* 150, 523–532.

Turner, P. E., and Chao, L. (1999). Prisoner's dilemma in an RNA virus. *Nature* 398, 441–443. doi: 10.1038/18913

Turner, P. E., and Chao, L. (2003). Escape from Prisoner's Dilemma in RNA phage phi6. *Am. Nat.* 161, 497–505. doi: 10.1086/367880

Varble, A., Meaden, S., Barrangou, R., Westra, E. R., and Marraffini, L. A. (2019). Recombination between phages and CRISPR-cas loci facilitates horizontal gene transfer in staphylococci. *Nat. Microbiol.* 4, 956–963. doi: 10.1038/s41564-019-0400-2

Vink, J. N. A., Martens, K. J. A., Vlot, M., McKenzie, R. E., Almendros, C., Estrada Bonilla, B., et al. (2019). Direct visualization of native CRISPR target search in live bacteria reveals cascade DNA surveillance mechanism. *Mol. Cell.* 77, 39–50. doi: 10.1016/j.molcel.2019.10.021

Wagner, R. R. (1960). Viral interference. Some considerations of basic mechanisms and their potential relationship to host resistance. *Bacteriol. Rev.* 24, 151–166. doi: 10.1128/mmbr.24.1.151-166.1960

Wigington, C. H., Sonderegger, D., Brussaard, C. P., Buchan, A., Finke, J. F., Fuhrman, J. A., et al. (2016). Re-examination of the relationship between marine virus and microbial cell abundances. *Nat. Microbiol.* 1:15024. doi: 10.1038/nmicrobiol.2015.24

Xue, C., Seetharam, A. S., Musharova, O., Severinov, K., Brouns, S. J., Severin, A. J., et al. (2015). CRISPR interference and priming varies with individual spacer sequences. *Nucleic Acids Res.* 43, 10831–10847. doi: 10.1093/nar/gkv1259