Restriction modification systems as engines of diversity

Kim Sneppen[1], Szabolcs Semsey[1], Aswin S. N. Seshasayee[2] and Sandeep Krishna[1,2,3]*

[1] Center for Models of Life, Niels Bohr Institute, Copenhagen, Denmark, [2] National Centre for Biological Sciences, Bangalore, India, [3] Simons Centre for the Study of Living Machines, National Centre for Biological Sciences, Bangalore, India

Restriction modification (RM) systems provide protection against a broad spectrum of phages. However, the likelihood of a phage permanently bypassing this can be as high as 0.1 per infection (Korona et al., 1993) which makes for a relatively weak defense. Here we argue that, apart from providing such transient defenses, RM systems can facilitate long-term coexistence of many bacterial strains. We show that this diversity can be as large as the burst size of the phage but no larger—a curious correspondence between a number at the level of species and another number at the level of individuals. Such a highly diverse and stably coexisting ecosystem is robust to substantial variation in both bacterial growth rates and strength of their RM systems, which might be one reason why quite weak RM systems exist in the wild.

Keywords: bacteriophage defense, epigenetic labeling, phage-bacteria interaction, ecosystem diversity, mathematical modeling, population dynamics

1. Introduction

Restriction-modification (RM) systems (Arber, 1965; Rambach and Tiollais, 1974; Kruger and Bickle, 1983; Kessler and Manta, 1990) are often discussed in connection with bacterial defense against phage predation both in bacteria (Bickle and Kruger, 1993) and archaea (Vasu and Nagaraja, 2013; Makarova et al., 2014), and are the most widespread phage defense system in bacterial genomes (Makarova et al., 2014). In addition, RM systems have evolved "moonlighting roles" in recombination, metabolism, and gene regulation (Vasu and Nagaraja, 2013). An RM system consists of two enzymatic functions, that both recognize and act on a certain DNA sequence. One of the enzymes cleaves the sequence if it is not methylated, whereas the other enzyme adds methyl groups to the sequence to protect it from cleavage. Thereby, foreign non-methylated DNA entering the cell, including DNA of invading phages, will be preferentially cleaved. This provides immunity to infection, to an extent that depends on the phage as well as on the particular RM system (Korona and Levin, 1993; Korona et al., 1993).

The immunity imposed by RM systems is not perfect. Typically, a given RM-system only recognizes a few DNA segments in a given phage, and for each of these there is a chance that it gets methylated by the bacterial modification enzyme before it is cleaved. Thereby, the phage sometimes can complete a successful infection and produce a burst of progeny phage. Furthermore, when a phage survives an infection cycle, the DNA of its progeny will be methylated and, upon infection of the same bacterial strain, will be able to bypass the RM system (Arber, 1965; Korona and Levin, 1993; Korona et al., 1993). Korona et al. (1993) suggest that this transient advantage of RM systems is useful when the host first invades an environment containing new phage, because it allows time to reach a larger population size, and thus more chance for the emergence of immune mutants and/or for the formation of a biofilm or other density-dependent mechanisms that enhance phage-bacteria

coexistence (Heilmann et al., 2012). However, the fact remains that the bacteria effectively "teach" the successful phage how to bypass the defense on all the siblings of the bacteria, and it is therefore a challenge to understand the long-term advantage of this widespread class of defense systems.

Here, we explore idealized conditions where RM systems could have a long-term advantage as quantified in terms of bacterial populations under steady state growth and dilution. In agreement with more complex models (Frank, 1994; Pagie and Hogeweg, 2000) we find that even a weak RM system provides a long-term advantage to a bacterial strain when there is a phage that infects both this and other strains of bacteria. Further, we will see that the presence of a RM system becomes necessary when competing bacteria also have an RM system. Further, we argue that a phage with burst size β can facilitate stable coexistence of up to β different bacterial strains, provided that they have unique RM systems—each RM system provides an epigenetic variant of the free phage, and this overcomes the typically strong limits on diversity in mutually supporting phage-bacteria ecosystems (Jover et al., 2013; Haerter et al., 2014).

2. Materials and Methods

We consider a system of many bacterial strains (b_i, $i = 1, \ldots, D$) with different RM systems, all preyed on by a single species of phage (Levin et al., 1988; Frank, 1994). The DNA of each individual phage particle is assumed to carry the methylation obtained in the host from which it emerged, and thus there are also D different phage types but they are all simply differently "labeled" versions of the same phage species. We therefore assume that all these different types infect a bacterial strain b_i with the same rate constant η_i and produce the same lytic burst β_i, both dependent on the bacterial strain only. The imperfection of b_i's RM system is parametrized by ω_i, the probability that an attacking phage p_j, with a non-protective methylation, completes a lytic cycle and thereby adds β_i phages to the pool of phage p_i. Each bacterial strain also has a specific growth rate γ_i and a common dilution/death rate α, which together set the maximal total bacterial density, or carrying capacity. This implicitly assumes that all bacteria compete for the same limiting resource. Phage also decay spontaneously at a rate δ_i. The resultant equations are:

$$\frac{db_i}{dt} = \gamma_i b_i \cdot (1 - B) - \alpha b_i - \eta_i b_i \cdot p_i - \eta_i \omega_i \cdot (P - p_i) \cdot b_i, \tag{1}$$

$$\frac{dp_i}{dt} = \beta_i \eta_i b_i \cdot p_i + \beta_i \eta_i \omega_i \cdot (P - p_i) \cdot b_i - \eta_i B \cdot p_i - \delta_i p_i, \tag{2}$$

where $B = \sum b_i$ and $P = \sum p_i$ are the total bacterial and phage populations, both measured (as are each b_i and p_i) in units of the bacterial carrying capacity at the available food supply rate. Notice that phage species that are completely unmodified could simply be modeled as phage produced by lysis of a bacteria having an RM system with $\omega_i = 1$.

The figures show dynamics obtained by numerically solving these equations using simple Euler integration implemented in Fortran, but the results have been cross-checked with more accurate integration methods.

3. Results

3.1. Bacteria Having RM Systems Can Invade and Out-Compete Bacteria Without RM Systems

We first investigate a special case, consisting of two bacterial strains: a "red" strain having an RM system (b_1) and a "blue" strain without (b_0), i.e., having $\omega_0 = 1$ (see **Figure 1A**). **Figure 1B** shows what happens when the initial steady state, consisting only of the "blue" bacterial strain and its phage, is invaded by the "red" strain of bacteria, containing a weak RM system. Although the invading strain is introduced at a density 10^{-9} and only has $\omega_1 = 0.1$, it takes over 99% of the biomass relatively fast. Whenever $\gamma_1 < \gamma_0$ (the presence of an RM system should result in slower replication), the non-RM strain is not eliminated completely. However, even if the invading bacteria with an RM system has a substantially lower growth rate, it typically does much better, see **Figure 1C**. Also note that even a weak RM system has a huge impact on the steady-state biomass of its host—the gain is only proportional to $1 - \omega_1$, therefore, the steady state difference between, for instance, $\omega_1 = 0.1$ and $\omega_1 = 10^{-8}$ is only 10%.

It is important to notice that these simple considerations also apply to the case where one bacterial strain has a subset of the RM systems present in the other strain. Thus, our simple considerations imply that even a weak new RM system will allow for invasion of microbial ecosystems.
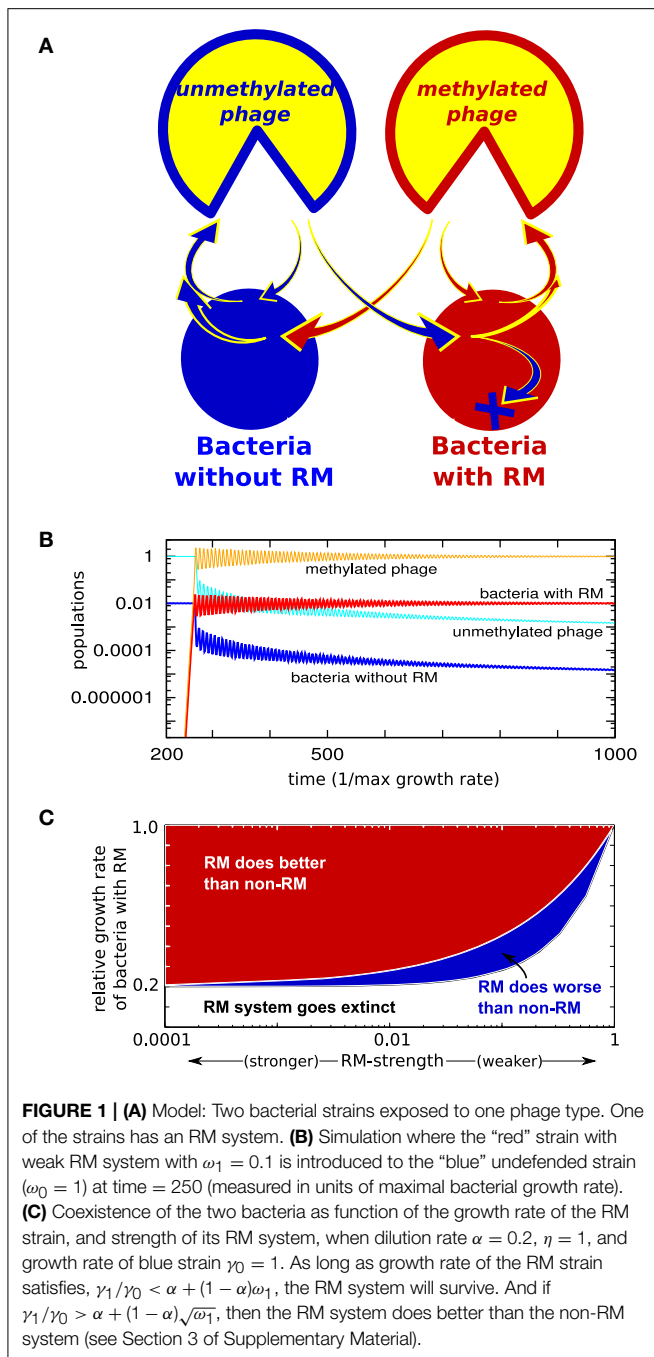
Our main predictions are that: (i) RM systems provide not only a transient advantage against phage, but also a long-term benefit when competing with bacteria lacking RM systems; (ii) Even weak RM systems are enough to provide this long-term benefit.

These predictions are in agreement with the observation that the majority of bacterial genomes carry at least one RM system (Murray, 2002). Bacteria that do not have any RM systems are typically intracellular pathogens (e.g., Chlamydia, Rickettsia), which are not exposed to phages.

3.2. The Phage Burst Size Limits the Diversity of RM Systems

The selective pressure on RM systems increases when there are multiple competing bacterial strains. **Figure 2** examines a system based on Equations (1) and (2) above, where the parameters β, η (and of course α) do not vary from bacteria to bacteria, but the RM strength ω_i and the bacterial growth rate γ_i are different for different bacterial strains. Note that our equations would not change if a bacterial strain has several RM systems provided they are distinct from the RM systems in other strains. Multiple RM systems in one strain are equivalent, in this case, to a single RM system whose effective ω is the product of the individual ωs.

The main predictions from **Figure 2** are: (i) The number of surviving bacterial strains, D, is about equal to the burst size β of the phage—an intriguing connection between two

FIGURE 1 | (A) Model: Two bacterial strains exposed to one phage type. One of the strains has an RM system. **(B)** Simulation where the "red" strain with weak RM system with $\omega_1 = 0.1$ is introduced to the "blue" undefended strain ($\omega_0 = 1$) at time = 250 (measured in units of maximal bacterial growth rate). **(C)** Coexistence of the two bacteria as function of the growth rate of the RM strain, and strength of its RM system, when dilution rate $\alpha = 0.2$, $\eta = 1$, and growth rate of blue strain $\gamma_0 = 1$. As long as growth rate of the RM strain satisfies, $\gamma_1/\gamma_0 < \alpha + (1 - \alpha)\omega_1$, the RM system will survive. And if $\gamma_1/\gamma_0 > \alpha + (1 - \alpha)\sqrt{\omega_1}$, then the RM system does better than the non-RM system (see Section 3 of Supplementary Material).

FIGURE 2 | Final distribution of surviving bacteria when starting with 500 bacterial strains, each having an RM system with strength $\omega_i \in [0.0001 : 1]$**, growth rate** $\gamma_i \in [0.2 : 1]$**.** Phage adsorption and decay rates are $\eta_i = 1$ and $\delta_i = 0.2$. The burst sizes vary as indicated. **(A)** Shows bacterial populations vs. RM strengths. **(B)** Shows bacterial populations vs. RM strengths and bacterial growth rates.

apparently unrelated numbers; (ii) All surviving strains have $\omega \stackrel{<}{\sim} 1/\beta$, implying that weak RM systems are under increasing evolutionary pressure in ecosystems with large diversity; (iii) The population size of each bacterial strain is nearly independent of the strength of its RM system, as long as this is stronger that the critical threshold of $\approx 1/\beta$. This "equipartition" of bacterial biomass is also maintained between bacteria with different growth rates, see **Figure 2B**, reflecting the stabilizing effect of preferential killing of the leading bacterial strain by phages (Thingstad and Lignell, 1997).
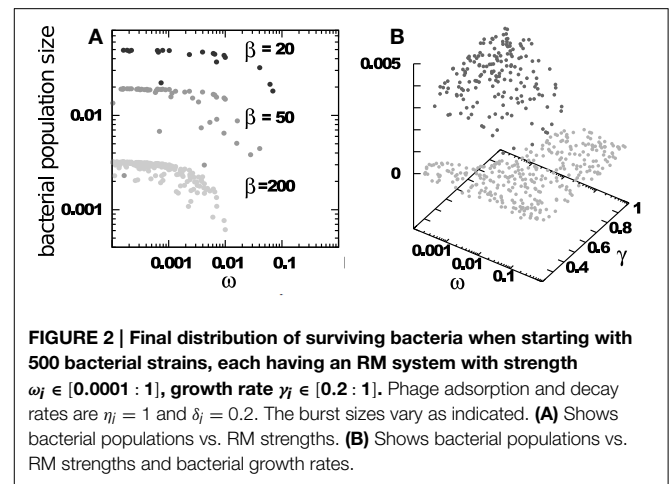
Intuitively, these observations are connected: A sustainable coexistence is only possible if, on average, one phage from a burst of β finds a non-immune host cell. Because bacterial biomass is nearly equipartitioned amongst the D strains, only a fraction $1/D$ of phages can reach a non-immune host. Therefore, coexistence requires that $\beta/D > 1$, i.e., $D < \beta$.

The $D \approx \beta$ "rule" can be understood in more detail by considering the simpler case where γ_i are the same for all bacteria, and only ω_i varies. Then the total bacterial density of D coexisting bacteria is (see Section 2 of Supplementary Material):

$$B = \frac{(\delta/\eta) \sum_i \frac{1}{1-\omega_i}}{\beta - \sum_i \frac{1-\beta\omega_i}{(1-\omega_i)}} \approx \frac{\delta}{\beta\eta} \cdot \frac{D}{1 - D/\beta}, \quad (3)$$

with the last approximation becoming exact when all the RM systems are sufficiently strong, i.e., $\omega_i \ll 1/\beta$. Here $\delta/(\beta\eta)$ is the population of one isolated bacterial strain due to predation from the phage (Campbell, 1960), measured in units of the carrying capacity. In order to be consistent, we must have $0 \leq B \leq \gamma^*$ (where $\gamma^* = (\gamma - \alpha)/\gamma$). Accordingly, one can only have coexistence of

$$D < \beta \cdot \frac{\gamma^*}{\gamma^* + \delta/\eta} \quad (4)$$

distinct bacterial RM systems.

Measured values for phages which infect *E. coli* in good growth conditions give $\delta/\eta \ll 1$ (De Paepe and Taddei, 2006), and in this situation we would predict $D \approx \beta$. In oceans with estimated $\delta = 0.04 - 0.11\,\text{h}^{-1}$ (Noble and Fuhrman, 1997), more phages may die/degrade before reaching a bacteria, and δ/η may become substantial. Then we would predict that less than β strains would be able to coexist. In both cases, β is an upper bound on the number of strains.

3.3. Different Regimes of Selection in an Open System
Above we examined a case where we start with a certain number of RM systems and see how many survive. Such a system would represent a case where multiple strains attempt to invade a

new niche, and the predictions from this can be tested in the laboratory using a closed experimental system. Below, we extend this analysis to a situation where new RM systems regularly invade the system. This kind of open system would represent realistic bacterial communities, for example in oceans or soil. **Figure 3** shows a simulation starting with $D = 5$ bacterial strains and iterated for 500 time units before introducing a new strain with new ω_i and γ_i values. Notice that some of the initial bacteria may be eliminated during these 500 time-units. The simulation subsequently proceeds by adding a new bacterial strain every 500 time-units.

Two "selection regimes" can be identified. First, when D is much less than β, the bacterial diversity increases (blue curve in **Figure 3A**), while at the same time the total bacterial biomass B also increases (gray curve in **Figure 3A**). Bacteria with RM systems gain, both individually and in total, by the presence of more bacteria with different RM systems.

In this regime, the RM systems are under increasing selective pressure, with their average strength increasing (their average ω decreases, as shown by the blue curve in **Figure 3B**), however, the growth rates of bacteria are *not* under any selection pressure (yellow curve in **Figure 3B**). Finally, the overall phage abundance

FIGURE 3 | Development of a system where bacterial strains with new RM systems are added sequentially (500 time steps). For the new strain, $\beta_i = 50$, $\eta_i = 1$, $\delta_i = 0.2$, ω_i is chosen from $[0.0001 : 1]$, and γ_i from $[0.2 : 1]$. The bacterial dilution rate is fixed at $\alpha = 0.2$. The simulation starts with few strains, and at the end we let the simulation run for 100000 time units without adding new bacterial strains. **(A)** Shows diversity/burst-size (blue), total bacterial population (gray) and total phage population (red) vs. time. **(B)** Shows RM strengths (blue) and bacterial growth rates (yellow) vs. time.

also increases with the number of RM systems (red curve in **Figure 3A**).

The second selection regime appears when the total bacterial density approaches carrying capacity, the diversity D becomes close to β, and the weakest of the RM systems has strength $< 1/\beta$. In this resource-limited regime, the phage population starts declining and bacteria are now selected primarily for high growth rates.

The behavior of the total phage population can again be understood by examining the simple case where only RM strengths vary from bacteria to bacteria. The steady state value of the total phage population then turns out to be related to diversity and bacterial population as follows (see Section 2 of Supplementary Material):

$$P \approx D \cdot \frac{\gamma(1 - B) - \alpha}{\eta} \tag{5}$$

This equation teaches us that P first increases with D, and then subsequently decreases as B approaches the carrying capacity of the available food sources.

This open system with increasing diversity augments the previous main predictions: (i) RM systems provide not just a transient defense against phage, but a long-term benefit both at the level of individual strains and total bacterial biomass; (ii) When there are many bacterial strains with different RM systems, selection favors those which carry RM systems with strength $< 1/\beta$ (or multiple weaker RM systems with a combined efficiency of $< 1/\beta$); (iii) The diversity of coexisting RM systems is limited by the phage burst size; (iv) As diversity approaches its upper limit, the phages die out and the biomass of the bacteria becomes limited by resource, not by phage.

4. Discussion

RM systems are intriguing because of their abundance and because they provide easily propagated epigenetic variants of virulent phages. In the absence of such variation, there are strong constraints on the number of phage and bacteria strains that can coexist. For instance, at most two bacterial strains can coexist with a single phage species (Haerter et al., 2014) and even a single bacterial strain might require density-dependent protective mechanisms to enable robust coexistence (Heilmann et al., 2012). As we have shown above, in the presence of RM systems many more bacterial strains ($\sim$ β) can coexist with a single phage species, even when these strains have quite different growth rates. The RM-derived epigenetic phage variants overcome such constraints because they limit the growth of their specific host. This levels the playing field for the bacterial strains, resulting in an equipartition of the bacterial biomass even amongst strains with very different growth rates. In contrast, other phage defense systems fail to produce as strong a cooperation between bacterial strains, although they may provide stronger individual defenses against phage. For instance, receptor mutants that prevent phage recognition do not act as a sink for phages and therefore do not help other strains. The CRISPR system, which can be thought of as providing "adaptive
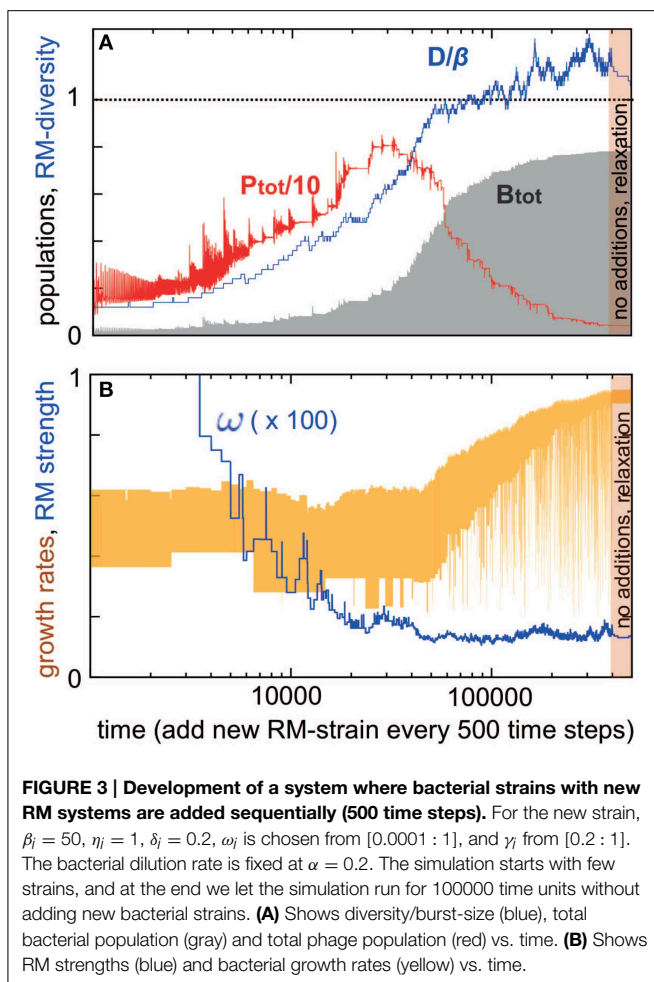
immunity" as opposed to the "innate immunity" of RM systems (Makarova et al., 2014), also provides a strong defense. Bacteria having it act as a sink for phages, but do not create phage variants at all (Haerter and Sneppen, 2012). Thus, a community of bacterial strains with only CRISPR defenses and one phage predator will not be able to sustain more than two coexisting strains.

The central prediction from our models is that the sustainable diversity of co-existing RM systems is strongly coupled to the phage burst size, as formulated in Equation (4). We find this intriguing because diversity is a population-level parameter, while the burst size is a biophysical parameter that is set by things like the composition and energy content of a single cell. A priori we would not have expected the diversity to be limited so directly by such a biophysical parameter, but our model reveals why RM systems lead to this prediction. Interestingly, the number of reported RM specificities (~300) (Roberts et al., 2010; Seshasayee et al., 2012) is similar to the median burst size of various known phage that infect *E. coli* (De Paepe and Taddei, 2006). Of course, the burst size of a given phage species may vary from host to host and with growth conditions (Weitz and Dushoff, 2008) and is exposed to variation due to occasional mutations. A smaller burst requires a shorter lysis time (Wang, 2006) and should give faster exponential growth at high host density (Abedon et al., 2001). In contrast, abundant diversity of RM systems favors large burst sizes, thus providing an opposing selection pressure. In other words, if a mutant with burst size $\beta_{small} < D$ arises in an ecosystem with D different RM systems, on average only $\beta_{small}/D < 1$ phage from each burst will find a permissive host, so its population will steadily decline until it goes extinct.

Overall, this paper explored the costs and benefits of bacterial RM-systems within the simple framework of mass-action kinetics of a well mixed system. Our investigation was motivated by the expectation, at first glance, that RM systems cannot provide a bacterial strain with a long-term population gain when it is the only host of the phage. Instead, our models show that presence of an RM system allows a bacterial population to use a phage to invade the ecological niche of another bacterial strain that does not have appropriate restriction mechanisms. Furthermore, we suggest RM-systems act as an engine of diversity, helping multiple bacterial strains to co-exist by generating epigenetic variations of the same virulent phage species. Thereby, much of the microbial diversity observed in nature may be supported by this relatively simple DNA recognition and elimination toolbox.

Author Contributions

All authors contributed to conception and design of the study, simulations and analysis, and writing of the manuscript.

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00528/abstract

References

Abedon, S. T., Herschler, T. D., and Stopar, D. (2001). Bacteriophage latent-period evolution as a response to resource availability. *Appl. Environ. Microbiol.* 67, 4233–4241. doi: 10.1128/AEM.67.9.4233-4241.2001

Arber, W. (1965). Host specificity of DNA produced by *escherichia coli*, v. the role of methionine in the production of host specificity. *J. Mol. Biol.* 11, 247–256. doi: 10.1016/S0022-2836(65)80055-9

Bickle, T. A., and Kruger, D. H. (1993). Biology of DNA restriction. *Microbiol. Rev.* 57, 434–450.

Campbell, A. M. (1960). Conditions for the existence of bacteriophage. *Evolution* 15, 153–165. doi: 10.2307/2406076

De Paepe, M., and Taddei, F. (2006). Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol.* 4:e193. doi: 10.1371/journal.pbio.0040193

Frank, S. A. (1994). Restriction-modification systems: the advantage of diversity. *Evolution* 48, 1470–1477. doi: 10.2307/2410241

Haerter, J. O., Mitarai, N., and Sneppen, K. (2014). Phage and bacteria support mutual diversity in a narrowing staircase of coexistence. *ISME J.* 8, 2317–2326. doi: 10.1038/ismej.2014.80

Haerter, J. O., and Sneppen, K. (2012). Spatial structure and lamarckian adaptation explain genetic diversity at crispr locus. *mBio* 3:e00126–12. doi: 10.1128/mBio.00126-12

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

Jover, L. F., Cortez, M. H., and Weitz, J. S. (2013). Mechanisms of multi-strain coexistence in host-phage systems with nested infection networks. *J. Theor. Biol.* 332, 65–77. doi: 10.1016/j.jtbi.2013.04.011

Kessler, C., and Manta, V. (1990). Specificity of restriction endonucleases and DNA modification methyltransferases - a review. *Gene* 92, 1–248. doi: 10.1016/0378-1119(90)90486-B

Korona, R., Korona, B., and Levin, B. R. (1993). Sensitivity of naturally occuring coliphages to type i and type ii restriction and modification. *J. Gen. Microbiol.* 139, 1283–1290. doi: 10.1099/00221287-139-6-1283

Korona, R., and Levin, B. R. (1993). Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution* 47, 565–575. doi: 10.2307/2410071

Kruger, D. H., and Bickle, T. A. (1983). Bacteriophage survival: multiple mechanisms for avoiding deoxyribonucleic acid restriction systems of their host. *Microbiol. Rev.* 47, 345–360.

Levin, B. R., Antonovics, J., and Sharma, H. (1988). Frequency-dependent selection in bacterial populations. *Proc. R. Soc. Lond. B* 319, 459–472. doi: 10.1098/rstb.1988.0059

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2014). Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* 41, 4360–4377. doi: 10.1093/nar/gkt157

Murray, E. N. (2002). Immigration control of DNA in bacteria: self versus non-self. *Microbiology* 148, 3–20. Available online at: http://mic.sgmjournals.org/content/148/1/3.full

Noble, R. T., and Fuhrman, J. A. (1997). Virus decay and its causes in coastal waters. *Appl. Environ. Microbiol.* 63, 77–83.

Pagie, L., and Hogeweg, P. (2000). Individual- and population-based diversity in restriction-modification systems. *Bull. Math. Biol.* 62, 759–774. doi: 10.1006/bulm.2000.0177

Rambach, A., and Tiollais, P. (1974). Bacteriophage lambda having ecori endonuclease sites only in the nonessential region of the genome. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3927–3930. doi: 10.1073/pnas.71.10.3927

Roberts, R. J., Vince, T., Posfai, J., and Macelis, D. (2010). Rebasea database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 38, D234–D236. doi: 10.1093/nar/gkp874

Seshasayee, A. S. N., Singh, P., and Krishna, S. (2012). Context-dependent conservation of DNA methyltransferases in bacteria. *Nucleic Acids Res.* 40, 7066–7073. doi: 10.1093/nar/gks390

Thingstad, T. F., and Lignell, R. (1997). Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.* 13, 19–27. doi: 10.3354/ame013019

Vasu, K., and Nagaraja, V. (2013). Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol. Mol. Biol. Rev.* 77, 53–72. doi: 10.1128/MMBR.00044-12

Wang, N. (2006). Lysis timing and bacteriophage fitness. *Genetics* 172, 17–26. doi: 10.1534/genetics.105.045922

Weitz, J. S., and Dushoff, J. (2008). Alternative stable states in hostphage dynamics. *Theor. Ecol.* 1, 13–19. doi: 10.1007/s12080-007-0001-1