Evolution, 48(5), 1994, pp. 1470–1477

POLYMORPHISM OF BACTERIAL RESTRICTION-MODIFICATION SYSTEMS: THE ADVANTAGE OF DIVERSITY

STEVEN A. FRANK
Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717

Abstract.—Bacterial restriction-modification systems provide defense against foreign DNA by using a self versus nonself recognition mechanism. A great diversity of recognition motifs is maintained in natural populations. Circumstantial evidence suggests that defense against bacteriophage viruses favors this diversity. (1) Bacterial restriction enzymes can destroy invading phage DNA. (2) Phage DNA can mimic the host's self-recognition mechanism. The ability of the virus to pose as a mimic favors diversification of the host's recognition motif. Other observations suggest that restriction modification (RM) does not provide any significant defensive advantages in mature communities. (1) In laboratory experiments, bacteria evolve resistance to phage by mutation and selection of the receptors to which phage adsorb. The outcome of these experiments is a community dominated by bacteria with receptor-based resistance, with a low abundance of phage and susceptible bacteria. (2) Phage are rare and receptor-based resistance is common in samples from natural communities. I present a model that shows two factors determine community composition: resources and RM diversity. Communities in resource-rich habitats are dominated by receptor-based resistance and support few phage; communities in poor habitats are dominated by restriction-modification defense and relatively abundant phage. RM diversity is itself a direct cause of community composition. As diversity increases from a low level, the abundance of phage increases and the relative abundance of receptor-based resistance declines. Further increases in diversity cause a crash in phage abundance, yielding a stable community of diverse RM types but an absence of the selective pressure—the phage—that drove the diversification. Empirical studies must sample a range of resource levels and RM diversity to analyze the forces that determine community composition.

Key words.—Bacteria, bacteriophage, host-parasite coevolution, immunity, resistance, virulence.

Received June 16, 1993. Accepted November 29, 1993.

Bacteria have a simple recognition-based immunity system that protects them from invasion by foreign DNA (Wilson and Murray 1991). There are two components to the system. Restriction enzymes cut DNA molecules that carry a particular sequence of nucleotides. Modification enzymes recognize the same nucleotide sequence but, instead of cutting the DNA, these enzymes modify the recognition site in a way that protects that molecule from restriction. A bacterial cell's own DNA is modified, otherwise the restriction enzymes would cut the DNA and kill the cell.

Restriction-modification (RM) enzymes are known for over 200 different recognition sites (Kessler and Manta 1990; Roberts 1990). Circumstantial evidence suggests that defense against bacteriophage viruses has been a powerful force promoting diversity. (1) RM can protect host cells from invading phage (Luria and Human 1952; Arber 1965). (2) Phage that develop in a bacteria with a particular RM type are modified for the associated recognition sequence. These modified phage can attack other bacteria of the same RM type but are sensitive to restriction by different RM systems. Rare RM types are favored because few phage will be modified for their recognition sequence. This frequency dependent selection promotes diversity of RM as a defense against phage (Levin 1986, 1988). (3) Phage carry a variety of antirestriction mechanisms (Kruger and Bickle 1983; Sharp 1986; Korona et al. 1993). For example, many phage lack particular RM recognition sequences. The probability of having these recognition sequences is very high if no selective pressure were acting on sequence composition.

The circumstantial evidence favors phage-mediated selection as an explanation for RM diversity. However, direct studies of interactions between phage and bacteria suggest that bacteria resist phage attack by modifying the receptor sites at which phage adsorb and enter the cell (Lenski 1984, 1988; Lenski and Levin 1985). In these studies, RM apparently has little effect on the long-term dynamics of phage and bacteria, suggesting that RM diversity may be maintained by processes other than phage-mediated selection in stable communities (Korona and Levin 1993).

Laboratory studies of phage and bacteria

maintained in chemostats provide repeatable observations about coevolution between phage and bacteria (Lenski 1988; Korona and Levin 1993). Phage and bacteria are mixed to begin the experiment. No matter what the short-term dynamics, the bacteria usually evolve a set of surface receptors that resist attack by phage. These modified receptors may reduce host growth rate because the receptors used by phage are typically the site for uptake of important nutrients.

With the appearance of receptor-based resistance, the community settles to a balance of resistant bacteria with reduced growth and sensitive bacteria with phage-induced mortality (Levin et al. 1977). In these communities, resistant bacteria typically outnumber sensitive bacteria, and there is a small phage population supported on the sensitive strain (Lenski 1988).

The outcome of evolution in laboratory communities can be summarized as follows. If receptor resistance is rare during the early phases of the experiment, RM may provide some defense against phage. As the experiment proceeds, receptor-based resistance becomes common, phage become rare, and RM loses its selective advantage.

The observations from natural populations provide conflicting evidence about the role of RM. On the one hand, phage are rare in natural isolates (Scarpino 1978), and receptor-based resistance is common (Lenski and Levin 1985). These observations support the view that phage-mediated selection is a very weak force in the maintenance of RM diversity. On the other hand, phage often carry antirestriction mechanisms, suggesting that RM is an important selective force on phage and that, in turn, phage probably influence RM diversity.

Levin (1986, 1988; Korona and Levin 1993) suggested that the conflicting evidence can be explained by a model in which RM is advantageous in colonizing new habitats, where phage are common and receptor-based resistance for the local phage has not yet evolved. As the newly established community matures, receptor-based resistance spreads and eventually dominates. Thus, RM diversity is maintained by cycles of selection that occur during colonization.

I present a model that shows phage can be a potent selective force in mature communities. The model explains the previously confusing observation that, in some mature communities, receptor-based resistance is common, phage are rare, and RM systems are diverse. The new aspect of my model is that variation in RM diversity is itself a direct cause of community structure.

The Model

I analyze a modified version of Levin's (1986) model. There are $i = 1, \ldots, N$ bacterial genotypes. Each genotype has a distinct RM type that cuts a unique DNA sequence. There are $i = 1, \ldots, N$ matching phage types, where phage type i is modified by and can therefore attack the ith RM bacterial genotype. There is an additional bacterial genotype with a modified surface receptor that prevents entry by phage. This bacterial genotype cannot be attacked by any phage.

The dynamics for the N host bacterial types, h_i, the receptor-modified host, h_r, and the N phage types, p_i, are given by

$$\frac{dh_i}{dt} = h_i\left[\gamma_s - \frac{\gamma_r h_r + \gamma_s H}{K} - \delta p_i - \delta\omega(P - p_i) - a\right], \quad (1A)$$

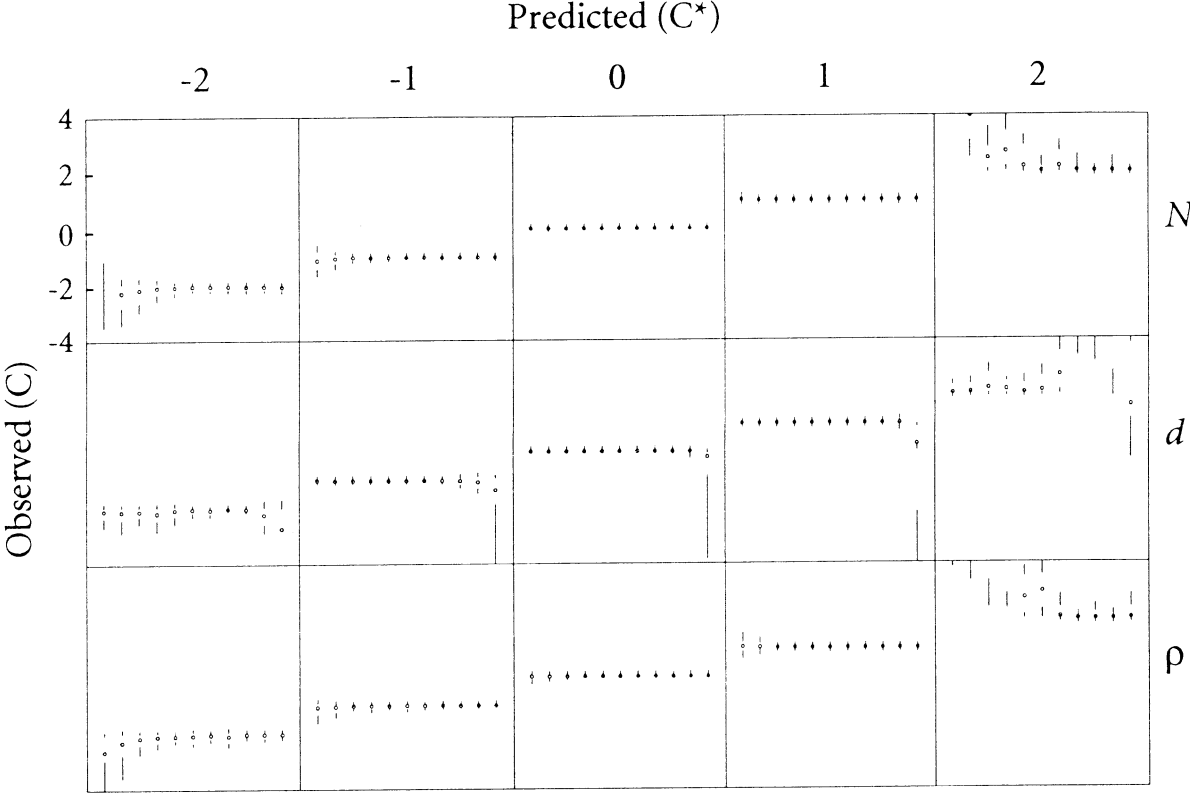
$$\frac{dh_r}{dt} = h_r\left[\gamma_r - \frac{\gamma_r h_r + \gamma_s H}{K} - a\right], \quad (1B)$$

$$\frac{dp_i}{dt} = p_i[-s + \delta\beta h_i - \delta H] + \delta\omega\beta h_i(P - p_i). \quad (1C)$$

The numbers of host cells and phage particles per unit volume are given by h_i, h_r, and p_i, with $H = \Sigma_{i=1}^{N} h_i$, and $P = \Sigma_{i=1}^{N} p_i$.

The rate of population increase for hosts with receptor-based resistance is γ_r, and for RM hosts, γ_s. Receptor-based resistance is assumed to have a cost so that $\gamma_s > \gamma_r$. The rate at which host cells are lost from the habitat by density-independent processes such as washout is a; the analogous rate for phage is s. The carrying capacity of the habitat is K, the maximum number of bacterial cells per unit volume that can be supported by the available resources. The terms with K in the denominator give the effects of density-dependent competition among the bacteria for limiting nutrients.

The rate at which phage particles adsorb to bacterial cells for a given density of cells is δ. Phage adsorb to all bacterial cells except those with receptor-based resistance. Once adsorbed, the phage enters the cell, leading to one of three outcomes. If the phage is not modified for the host's RM type, then (1) the phage DNA is cut


Predicted (C*)
-2
-1
0
1
2
Observed (C)
4
2
0
-2
-4
N
d
ρ

with probability $1 - \omega$, or (2) with probability ω, the phage is modified before it is restricted, allowing the phage to replicate, produce β progeny that are modified for the host's RM type, and kill the host. (3) If the phage is modified for the host's RM type, then β phage progeny are produced and the host is killed.

The system in equations (1) is easier to analyze when rewritten in nondimensional form (Segel 1972; Murray 1989). Nondimensional analysis focuses attention on a minimal set of parameters and highlights relative magnitudes (scaling relations) among the processes that drive the dynamics. Nondimensional analysis is accomplished without altering the dynamics or interpretation because one can translate freely between the biologically motivated formulation and the nondimensional quantities.

The system can be rewritten with the following substitutions:

$$\hat{h}_i = h_i/K, \quad \hat{h}_r = h_r/K, \quad \hat{p}_i = \delta p_i/\gamma_s,$$
$$\tau = t\gamma_s, \quad d = \delta K/\gamma_s, \quad \gamma = \gamma_r/\gamma_s,$$
$$\rho = s/\gamma_s, \quad r = a/\gamma_s. \tag{2}$$

Dropping the hats yields the nondimensional system

$$\frac{dh_i}{d\tau} = h_i[1 - \gamma h_r - H - p_i - \omega(P - p_i) - r], \tag{3A}$$

$$\frac{dh_r}{d\tau} = h_r[\gamma - \gamma h_r - H - r], \tag{3B}$$

$$\frac{dp_i}{d\tau} = p_i[-\rho + d\beta h_i - dH] + d\omega\beta h_i(P - p_i). \tag{3C}$$

The unique equilibrium of the full $2N + 1$ types in Equation (3) is obtained by setting each of the above equations to zero, yielding

$$h_i^* = \frac{\rho}{d[\beta + \omega\beta(N - 1) - N]}, \quad i = 1, \ldots, N \tag{4A}$$

$$h_r^* = 1 - Nh_i^*/\gamma - r/\gamma, \tag{4B}$$

$$p_i^* = \frac{1 - \gamma}{1 + \omega(N - 1)} \quad i = 1, \ldots, N. \tag{4C}$$

I analyze this system in two stages. First, I present the conditions under which the community tends to be composed of either purely receptor-resistant bacteria and no phage or purely RM bacteria and phage. These conditions can be established from the equilibrium properties in equations (4). Second, I show that the equilibrium equations can also predict the composition of mixed communities, where the mixtures are defined as the abundance of phage and the relative proportion of RM versus receptor-resistant bacteria.

Fixation or Loss of Receptor-Based Resistance

Receptor-based resistance is fixed as $Nh_i^* \to 0$. By contrast, receptor-based resistance is lost from the population when $h_i^* < 0$, which occurs

←

FIG. 1. Fluctuations in $C = \log_{10}(\Sigma_i h_i/h_r)$ over time. The layout of the figure is described in the text. Here I describe how the parameters were chosen. The parameter for each row was varied over 11 evenly spaced steps for the range of that parameter. The other six parameters were chosen randomly from their range, subject to the single linear constraint imposed by $C^* = \log_{10}(Nh_i^*/h_r^*)$. The ranges for the seven parameters are: $10^0 \le N \le 10^2$; $0.89 \le \gamma \le 0.99$; $10^{-6} \le \omega \le 10^{-1}$; $10^{-5} \le d \le 10^5$; $10^1 \le \beta \le 10^3$; $0.05 \le \rho \le 0.25$; $0.05 \le r \le 0.25$. For the ranges specified by powers of 10, the range was divided into uniform pieces on a logarithmic scale. All seven parameters were studied, the results for N, d, and ρ are plotted in the figure. In the six cases where β was free to vary, the other six parameters were chosen and the constraint equation was solved for β. When β was varied systematically, five parameters were chosen randomly and then d was obtained by solving the constraint equation. Each run was started by initializing each of the $2N + 1$ types with an abundance chosen by a random number from 10^{-5} to 10^{-3}, where the random distribution is uniform on a logarithmic scale. The system was further initialized by applying equations (3) for 1000 nondimensional time units, τ, where each unit is the doubling time of the RM types (see eq. 2). The ratio of the total abundance of all RM types to receptor-resistant cells was then measured after each of the following 1000 time units. The 5, 25, 50, 75, and 95 percentiles were calculated from these 1000 measurements. Each run has randomly chosen parameters and random initialization. I replicated each run 100 times for each fixed parameter value and equilibrium constraint, C^*. These replicates yielded 100 measurements for each of the five percentile categories. For example, the top of the upper vertical line is the median value over 100 replicates of the ninety-fifth percentile for the 1000 time points measured for each replicate.

when

$$N > \frac{d\beta(\gamma - r)(1 - \omega)}{(1 + \rho - \omega\beta)}. \quad (5)$$

This condition can be simplified by noting that the probability that a phage escapes restriction, ω, is likely to be small, so that $\omega\beta \ll 1$. Thus, the condition for loss of receptor-based resistance in equation (5) is approximately

$$N > d\beta(\gamma - r)/(1 + \rho). \quad (6)$$

Two predictions are worth noting. First, the relative advantage of RM defense versus receptor-based defense increases as the diversity of RM genotypes, N, increases. RM gains an advantage when diverse because each phage is specialized for only the one RM type in which it was born—a phage has only a small probability, ω, of succeeding in any of the other $N - 1$ types. As N increases, a higher proportion of phage adsorptions results in phage death rather than successful infection.

A second prediction from equation (6) is that a reduction in $d = \delta K/\gamma_s$ favors RM over receptor-based defense. A decrease in phage adsorption relative to population growth rate, δ/γ_s, favors RM by shifting the main selective pressure on the bacterial hosts away from phage resistance and toward resource competition. Likewise, a reduction in habitat quality, K, favors RM over receptor-based defense by emphasizing resource competition. In each case, the assumption that receptor-based resistance has some cost, $\gamma_r < \gamma_s$, however small, causes RM to be better suited for habitats in which resource competition is intense (Levin 1986).

The other parameters in equation (6), β, γ, r, ρ, are likely to vary over a much smaller range than N and d for a particular phage-bacteria system (see above definitions).

Mixture of Receptor-Based Resistance and RM Types

In the previous section, I used the equilibrium equations to find conditions for the fixation or loss of RM relative to receptor-based resistance. In this section, I analyze the fluctuations (stability) of the system to support those conclusions. I also extend the analysis to show that the equilibrium equations provide a good prediction for the relative abundance of RM types and receptor-based resistance in mixed communities.

The equilibrium in equations (4) can be used to predict the relative abundance of RM and receptor-resistant bacteria, $C = \log_{10}(\Sigma_i h_i/h_r)$. This prediction is useful only if fluctuations in the system are sufficiently small that the location of the equilibrium is a good guide to the community mixture.

The computer analysis presented in figure 1 shows that the equilibrium value $C^* = \log_{10}(Nh_i^*/h_r^*)$ provides an excellent prediction for community composition. The purpose of the figure is to show the match between the equilibrium prediction, C^*, listed at the top of each column of panels, and the observed distribution of values for the ratio of susceptible to resistant types, C. The observed values, measured by the scale along the y-axis of each panel, match closely the predicted values given for each column.

The remainder of this section provides details about figure 1. It is not necessary to follow these details on first reading. The only important conclusion from figure (1) is that the equilibrium equations provide a good description of community composition.

Each row of panels shows the effect of varying a particular parameter, (N, d, or ρ), with a different value of C^* in each column. Each panel shows the effect of increasing the parameter over 11 different values in a range given in the figure caption.

For example, the upper left panel of figure 1 shows the effect of varying the number of types, N, when the predicted equilibrium mixture, $C^* = -2$, is 1% RM types and 99% receptor-resistant cells. The 11 values of N along the x-axis are equally spaced values on a logarithmic scale over the range 10^0 to 10^2. Thus, the second value along the x-axis is $N = 10^{0.2} \approx 2$. Two vertical lines and a circle are shown above this value. These marks show the fluctuations in C over approximately 1000 generations. The top of the upper line is the ninety-fifth percentile of C over time, the bottom of that line is the seventy-fifth percentile, the circle is the median, and the top and bottom of the lower line are the twenty-fifth and fifth percentiles, respectively. (Further details are given in the figure caption.) Variation caused by γ, ω, β, and r is not shown, but in each case the fluctuations were smaller than in those panels shown in the figure.

The abundances of the N distinct RM systems tend to be nearly equal whenever RM types make up one-half or more of the bacterial population, $C^* > 0$. Fluctuation in the evenness of the distribution of RM types sometimes occurs when

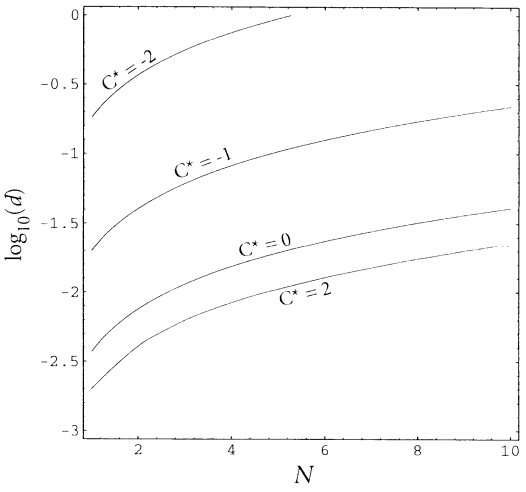


FIG. 2. Contour plot of the relative equilibrium abundance of RM and receptor-based resistance. The contours show $C^* = \log_{10}(Nh_i^*/h_r^*)$ as a function of the number of RM types, N, and the nondimensional parameter $d = \delta K/\gamma_s$. Variation in d is caused mainly by variation in resource availability, K. The other parameters were set to the center point for the plots in figure 1: $r = \rho = 0.15$; $\omega = 10^{-3.5}$; $\beta = 100$; and $\gamma = 0.94$.

RM is rare. This probably happens because, with small total abundance, shifts in absolute numbers have a greater effect on relative frequency.

RM Diversity and Community Composition

The two parameters most likely to vary widely in natural communities are N, the number of RM types, and $d = \delta K/\gamma_s$. Variation in d is caused mainly by changes in resource abundance, K. Figure 2 shows how a decline in resource quality and an increase in RM diversity increase the relative abundance of RM versus receptor-resistant bacteria.

The effects of increasing RM diversity on community composition can be divided into three stages. Each stage is labeled by a circled number in the panels of figure 3.

(1) As the number of RM types increases from $N = 1$, the abundances of phage and phage-sensitive RM types increase and the abundance of receptor-resistant bacteria decreases. A rare RM type always invades an equilibrium community with phage because none of the phage are specialized (modified) for the new type. Each new RM type increases to the point at which it maintains its own phage subpopulation that limits the further spread of that RM type. Each of the RM types is phage-limited; thus, each new RM type causes an approximately linear increase in the total abundance of phage and RM types. Because the RM types are phage limited, the resources taken by each new RM type reduce the abundance of the receptor-resistant population but do not interfere with other RM types.

(2) The receptor-resistant bacteria are eventually driven to extinction when a sufficient number of RM types have accumulated (see eq. 5). Novel RM types can continue to invade. Each new type causes a reduction in the phage population. This reduction is probably caused by the high proportion of phage deaths that result from adsorption and restriction in bacteria for which the phage DNA is unmodified, $(1 - \omega)(N - 1)/N$, and by the increase in resource competition among the RM types, which reduces the amount of bacterial productivity that can be grazed by the phage.

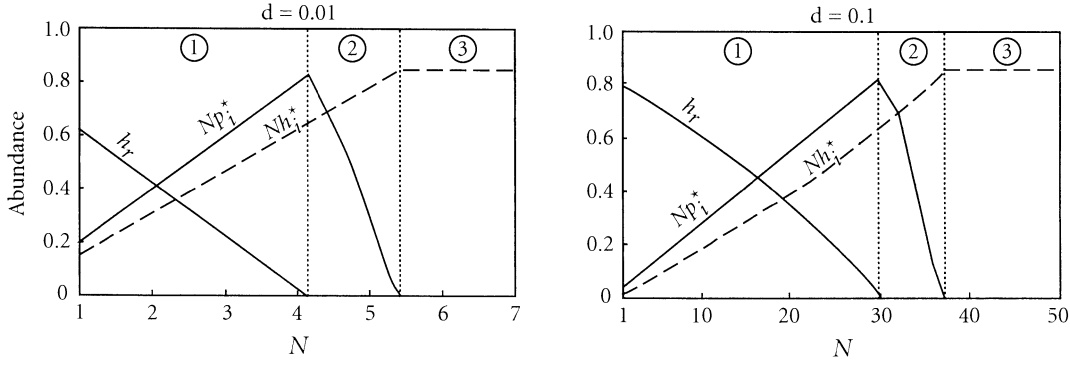


FIG. 3. Community composition as a function of the number of RM types, N. The y-axis shows abundance in nondimensional units according to equation (2), with Np_i^* scaled by 7/50 in the right panel. The three sections of each panel are described in the text. The value of γ, which controls the width of section two, is 0.8. The value of d is shown above each plot. The other parameters were set to the center point for the plots in figure 1: $r = \rho = 0.15$; $\omega = 10^{-3.5}$; and $\beta = 100$.

(3) Further increase in the number of RM types drives the phage to extinction. Extinction occurs when the condition in equation (5) is satisfied with γ set to one. The stable community at the transition between stages 2 and 3 supports a diversity of RM types but no phage.

Discussion

The model makes two predictions. First, resource-poor habitats are expected to have a higher relative abundance of RM defense versus receptor-based defense (Levin 1986, 1988). Second, RM diversity is favored in all habitats with phage because a rare RM type can always invade an equilibrium community. The interesting aspect of the model is that RM diversity can have a strong influence on community composition (figs. 2 and 3). For a fixed level of nutrients, an increase in RM diversity can cause a shift from a community with few phage and relative dominance by receptor-resistant bacteria to a community in which receptor-based resistance is rare and phage are common. A further increase in RM diversity can drive the phage to extinction. In a laboratory experiment this sequence would lead to an endpoint that, at first sight, would seem strange: a community of diverse RM types but an absence of the selective pressure—the phage.

Korona and Levin (1993) conducted a laboratory experiment to look at the other extreme of RM diversity: the case in which one RM type is present and a second type is added to the community. Receptor-based resistance was absent in the initial community. Mutations for receptor-based resistance spread in the bacterial population, leading to a community in which receptor-based resistance was common, and RM types and phage were rare or absent. This result, along with the observations from natural isolates mentioned above, led Korona and Levin to suggest that RM may not be an important defense against phage in mature communities. The model presented here shows that the role of RM diversity must be analyzed in the context of different levels of resource abundance. For example, Korona and Levin's design with lower levels of resource might allow invasion of the novel RM strain rather than dominance by resistant bacteria. In addition, the level of RM diversity is itself an important cause of community composition. For example, Korona and Levin's design with several RM types in the initial population might not be invaded by resistant bacterial strains.

Other factors probably influence RM diversity and community composition. For example, phage antirestriction mechanisms enhance host range beyond those RM types for which the virus is modified (Kruger and Bickle 1983; Sharp 1986). Such counter-measures by the virus can lead to a coevolutionary genetic system with extensive polymorphism in both the host and parasite (Chao et al. 1977; Frank 1992, 1993; Thompson and Burdon 1992). Samples of phage from wild populations show that the viruses are highly polymorphic for antirestriction mechanisms (Korona et al. 1993).

Whatever the role of viral polymorphism, it is clear that the joint effects of habitat quality and RM diversity must be analyzed to understand the maintenance of RM polymorphism.

Acknowledgments

I thank B. R. Levin for many interesting discussions, and P. Amarasekare, R. M. Bush, M. Feldgarden, and R. E. Lenski for helpful comments on the manuscript. My research is supported by National Science Foundation grant BSR-9057331 and National Institutes of Health grants GM42403 and BRSG-S07-RR07008.

Literature Cited

Arber, W. 1965. Host controlled modification of bacteriophage. Annual Review of Microbiology 19:365–368.

Chao, L., B. R. Levin, and F. M. Stewart. 1977. A complex community in a simple habitat: an experimental study with bacteria and phage. Ecology 58: 369–378.

Frank, S. A. 1992. Models of plant-pathogen coevolution. Trends in Genetics 8:213–219.

———. 1993. Coevolutionary genetics of plants and pathogens. Evolutionary Ecology 7:45–75.

Kessler, C., and V. Manta. 1990. Specificity of restriction endonucleases and DNA modification methyltransferases—a review. Gene 92:1–248.

Korona, R., and B. R. Levin. 1993. Phage-mediated selection and the evolution and maintenance of restriction–modification. Evolution 47:556–575.

Korona, R., B. Korona, and B. R. Levin. 1993. Sensitivity of naturally occurring coliphages to type I and type II restriction and modification. Journal of General Microbiology 139:1283-1290.

Kruger, D. H., and T. A. Bickle. 1983. Bacteriophage survival: multiple mechanisms for avoiding deoxribonucleic acid restriction systems of their hosts. Microbiological Review 47:345–360.

Lenski, R. E. 1984. Coevolution of bacteria and phage: are there endless cycles of bacterial defenses and phage counterdefenses? Journal of Theoretical Biology 108:319–325.

———. 1988. Dynamics of interactions between bacteria and virulent bacteriophage. Advances in Microbiology and Ecology 10:1–44.

Lenski, R. E., and B. R. Levin. 1985. Constraints on

the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. American Naturalist 125:585–602.

Levin, B. R. 1986. Restriction-modification immunity and the maintenance of genetic diversity in bacterial populations. Pp. 669–688 *in* S. Karlin and E. Nevo, eds. Evolutionary processes and evolutionary theory. Academic Press, New York.

———. 1988. Frequency-dependent selection in bacterial populations. Philosophical Transactions of the Royal Society of London, Series B 319:459–472.

Levin, B. R., F. M. Stewart, and L. Chao. 1977. Resource-limited growth, competition, and predation: a model and experimental studies with bacteria and bacteriophage. American Naturalist 111:3–24.

Luria, S. E., and M. L. Human. 1952. A non-hereditary host-induced variation in bacterial viruses. Journal of Bacteriology 64:557–559.

Murray, J. D. 1989. Mathematical biology. Springer, New York.

Roberts, R. J. 1990. Restriction enzymes and their isoschizomers. Nucleic Acids Research 18 (Suppl.): 2331–2365.

Scarpino, P. V. 1978. Bacteriophage indicators. Pp. 201–227 *in* G. Berg, ed. Indicators of viruses in water and food. Ann Arbor Science, Ann Arbor, Mich.

Segel, L. A. 1972. Simplification and scaling. SIAM Review 14:547–571.

Sharp, P. M. 1986. Molecular evolution of bacteriophages: evidence of selection against the recognition sites of host restriction enzymes. Molecular Biology and Evolution 3:75–83.

Thompson, J. N., and J. J. Burdon. 1992. Gene-for-gene coevolution between plants and parasites. Nature 360:121–125.

Wilson, G. G., and N. E. Murray. 1991. Restriction and modification systems. Annual Review of Genetics 25:585–627.

Corresponding Editor: L. Chao