doi:10.1111/evo.14165

Feedback between coevolution and epidemiology can help or hinder the maintenance of genetic variation in host-parasite models

Ailene MacPherson,[1,2] Matthew J. Keeling,[3] and Sarah P. Otto[1]

[1] *Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*

[2] *E-mail: amacp@zoology.ubc.ca*

[3] *Zeeman Institute of Systems Biology and Infectious Disease Research (SBIDER), University of Warwick, Coventry CV4 7AL, United Kingdom*

Received April 10, 2020
Accepted December 7, 2020

Antagonistic coevolution has long been suggested to help maintain host genetic variation. Although ecological and epidemiological feedbacks are known to have important consequences on coevolutionary allele-frequency dynamics, their effects on the maintenance of genetic variation remains poorly understood. Here, we extend previous work on the maintenance of genetic variation in a classic matching alleles coevolutionary model by exploring the effects of ecological and epidemiological feedbacks, where both allele frequencies and population sizes are allowed to vary over time. We find that coevolution rarely maintains more host genetic variation than expected under neutral genetic drift alone. When and if coevolution maintains or depletes genetic variation relative to neutral drift is determined, predominantly, by two factors: the deterministic stability of the Red Queen allele-frequency cycles and the chance of allele fixation in the pathogen, as this results in directional selection and depletion of genetic variation in the host. Compared to purely coevolutionary models with constant host and pathogen population sizes, ecological and epidemiological feedbacks stabilize Red Queen cycles deterministically, but population fluctuations in the pathogen increase the rate of allele fixation in the pathogen, especially in epidemiological models. Our results illustrate the importance of considering the ecological and epidemiological context in which coevolution occurs when examining the impact of Red Queen cycles on genetic variation.

KEY WORDS: **Coevolution, eco-evolutionary feedbacks, epidemiology, genetic variation, negative frequency dependent selection.**

First suggested by Haldane (1949), antagonistic coevolution between hosts and their parasites is thought to maintain genetic variation through negative frequency dependent selection (NFDS). Arising from the interaction between hosts and parasites, such reciprocal natural selection must occur within an ecological, or in the case of an infectious pathogen, an epidemiological context. Feedbacks, either ecological (Song et al. 2015; Ashby et al. 2019) or epidemiological (Penman et al. 2013; Ashby and Gupta 2014; MacPherson et al. 2018), can have important impacts on the resulting long-term coevolutionary dynamics. Coevolutionary can, in turn, influence epidemiological dynamics (MacPherson et al. 2018) and traits, for example, susceptibility, tolerance, virulence (Best et al. 2008, 2010, 2017; Ashby and Boots 2015). The importance of ecological and epidemiological feedbacks, however, on the maintenance of host genetic variation is not yet well understood.

Coevolution was originally suggested to maintain genetic variation, for example, in the matching-alleles model (MAM), due to selection for rare genotypes (Haldane 1949; Clarke 1979). Specifically, because parasites are better at infecting hosts that carry a "matching" genotype, natural selection will thus favor host (parasite) genotypes that are rare (common) in the

interacting species. Despite resembling NFDS within a species, which does help to maintain genetic variation (Takahata and Nei, 1990), MacPherson et al. (2020) showed that coevolution between a host and free-living pathogen does not maintain genetic variation relative to an equivalent model of neutral genetic drift when the host and pathogen population sizes are forced to remain constant. By assuming that host and parasite population sizes remain constant, this previous work ignored the effects of ecological and epidemiological feedbacks.

Exemplified by biological pest control (Myers and Bazely 2003), parasites can be an important form of population regulation in their hosts. Depending on the genetic basis of the interaction and the form of population regulation (interspecific regulation only: Song et al. 2015; inter- and intraspecific regulation: Ashby et al. 2019; Nuismer 2017), ecological feedbacks can change the stability of coevolutionary equilibria. For example, in the case of the MAM, the combination of inter- and intraspecific interactions can lead to a stable polymorphism at a coevolving locus (Ashby et al. 2019). Similarly, epidemiological feedbacks, for example, those arising in a susceptible infected recovered compartmental model, have been shown to disrupt classic Red Queen allele-frequency cycles (Ashby and Gupta 2014; MacPherson et al. 2018), resulting instead in a stable polymorphism.

These previous models consider host and parasite population sizes deterministically, effectively examining their effects on host and parasite population densities. To fully understand the effect of ecological and epidemiological feedbacks on the maintenance of genetic variation, however, we must model eco- and epi-coevolutionary feedbacks on host and parasite populations in a fully stochastic context. As in MacPherson et al. (2020), we quantify the effect of matching-alleles coevolution on the maintenance of genetic variation by comparing host heterozygosity in a host-parasite coevolutionary model to the analogous single-species neutral expectation. To fully understand both the effect of ecological feedbacks in free-living parasites and epidemiological feedbacks in directly transmitted pathogens, we present and contrast three coevolutionary models. First, we review the maintenance of genetic variation in the absence of feedbacks (MacPherson et al. 2020). We then consider the maintenance of genetic variation in a model of a host and free-living pathogen with ecological feedbacks that result from both inter- and intraspecific population regulation (analogous to the model considered by Ashby et al. 2019; Frank 1993; Gandon et al. 1996; Nuismer 2017). Finally, we examine epidemiological feedbacks in a susceptible infected susceptible (SIS) compartmental model of a directly transmitted pathogen (as in Ashby and Gupta 2014; Ashby and Boots 2015).

Theoretical Background

There exists an extensive theoretical literature exploring when and how evolution either maintains or depletes genetic variation. In infinitely large populations, genetic variation is eroded by purifying and directional selection. On the other hand, some forms of natural selection can help maintain genetic variation, for example heterozygote advantage (Crow 1970, Section 9.7) or NFDS (Takahata and Nei 1990). In addition to these deterministic processes, genetic variation is lost continuously through genetic drift in finite populations. The rate at which drift occurs is proportional to the effective population size, which in turn depends on the census population size, sex ratio (Crow 1970), population subdivision (Whitlock and Barton 1997), or age structure (Hill 1972). In addition to these effects on the maintenance of genetic variation within a single species, biotic interactions in the form of coevolution may influence the maintenance of genetic variation (Haldane 1949; Clarke 1979). As coevolution favors genotypes that are rare in the interacting species, these hypotheses are rooted in the view that NFDS helps maintain variation.

There exists an important distinction, however, between traditional direct frequency dependent selection, which favors genotypes when their own frequency is rare, and indirect frequency dependent selection that arises via coevolution, as reviewed for models of gene-for-gene (GFG) coevolution between plants and their pathogens (Brown and Tellier 2011). As in the GFG model, indirect frequency dependent selection in the MAM does not maintain genetic variation relative to neutral genetic drift (MacPherson et al. 2020). This previous work focused exclusively on the effects of coevolution (changes in allele frequency) on the maintenance of genetic variation by assuming that the host and parasite population sizes remained constant over time, a common assumption of many MAM (Nuismer 2017). Here and in this work, we examine the impact of drift on genetic variation in the host, by tracking the probability that two alleles chosen at random differ, that is, $2p_H(1-p_H)$ for a haploid host with allele-frequency p_H (a measure known as "effective heterozygosity," which we refer to simply as "heterozygosity"). By comparing host heterozygosity in a coevolutionary model to the neutral expectation under genetic drift alone, we found that genetic variation with coevolution is almost always less than or equal to the neutral expectation. Although the deterministic stability of the coevolutionary model would suggest that heterozygosity should evolve neutrally, decreases in heterozygosity relative to the neutral expectation occurs, in part, because chance fixation in the pathogen in finite populations turns symmetric coevolutionary selection into directional selection, which rapidly erodes genetic variation in the host.

Although assuming that the host and parasite population size remain constant allowed us to focus on the effects of coevolution, this assumption is violated in many systems (Papkou et al. 2016). Specifically, host population size fluctuations resulting from coevolution are expected to reduce the effective host population size, thereby increasing the rate of genetic drift. Gokhale et al. (2013), for example, attempted to examine the effect of such population size fluctuations on genetic variation by comparing fixation times in a coevolutionary model to those in a neutral model. However, their model comparison focused on holding the arithmetic mean population size constant across models rather than the harmonic mean population size which is the appropriate effective population size comparison for genetic drift (Crow 1970). Given the dynamics of effective population size alone, coevolution and the associated population dynamics are thus expected to deplete genetic variation. However, as shown by Ashby et al. (2019), the effective population size is not the only thing that changes with the introduction of ecological feedbacks. In particular, modeling host and parasite population dynamics with density-dependent population growth allows the growth rate of the host to increase whenever parasitism results in more deaths. This introduces an additional form of feedback between the host and parasite, which slightly stabilizes the allele-frequency dynamics. What then, is the net effect of ecological feedbacks on the maintenance of genetic variation?

To disentangle this, we explore two models that include density-dependent growth in the host. In the first, which we label the "ecological" MAM, the parasite is again free living, which a population size that varies depending on the infection dynamics. Although this assumption accurately captures the life history of many host-parasite systems, for example, the infection of *Daphnia magna* by one of its many water-borne parasites (Ebert 2005), many diseases are directly transmitted and cannot live independent of their hosts. We thus also develop a model of a directly transmitted pathogen using compartmental models from epidemiology, where hosts are categorized by their infection status, for example as "susceptible" (S) or "infected" (I), which we refer to as the "epidemiological" MAM. As a result the evolution of directly transmitted pathogens is subject to yet another feedback mechanism between host and pathogen resulting from the depletion of susceptible hosts from infection.

Epidemiological dynamics may also be expected to affect the maintenance of genetic variation. Specifically, epidemiological dynamics disrupt the neutral stability of Red Queen allele-frequency cycles observed in the constant-size model, resulting instead in a stable polymorphic equilibrium (MacPherson et al. 2018). One consequence of the strong stabilizing effects of epidemiology is that, in comparison to free-living parasites where the polymorphic equilibrium is only weakly stable (Ashby et al. 2019), directly transmitted infections should be better able to maintain genetic variation in their hosts. However, in the SIS model considered here the effective population size of a directly transmitted pathogen is restricted to being less than or equal to that of its host, because the number of infected hosts is less than or equal to the total host population size and because we assume strong bottlenecks during transmission (i.e., new infection begin from a single parasite genotype). As a result of having a smaller effective population size, alleles in the parasite are more likely to be lost or fixed by chance, increasing the likelihood of directional selection on the host and hence reducing the amount of host genetic variation.

Here we explore the maintenance of genetic variation in three models: a constant-size model (as in MacPherson et al. 2020), an ecological model, and a third SIS epidemiological model, the latter two incorporating density dependence and allowing population size fluctuations in both the host and parasite. All the models explored in this article consider coevolution between a haploid host and haploid parasite with a single-locus biallelic matching-alleles genetic basis. For each model we begin by exploring the deterministic coevolutionary dynamics, using the stability of the polymorphic equilibrium to predict ability of each model to maintain genetic variation in a finite population. We then test these predictions and compare the models using individual-based simulations of coevolution in finite populations. To evaluate the ability of different models to maintain variation we compare host heterozygosity in each of the three coevolutionary models H_{coev} relative to the heterozygosity expected under neutral drift H_{neut}, first by fixing the equilibrium host population size to a value K ("Stochastic Dynamics" section), then in "Model Comparison" section by fixing both host and pathogen equilibrium population sizes across models.

Deterministic Dynamics

CONSTANT-SIZE MODEL: REVIEW

MacPherson et al. (2020) considered a coevolutionary model of a host and free-living pathogen, assuming that the population sizes of both species remain constant over time and equal to one another. We begin by generalizing the model allowing the total host population size to be $f_H\kappa$ and pathogen population size $f_P\kappa$, both of which are constant and proportional to κ, which sets the system size. We can thus assess the effect of host population size on the maintenance of variation by varying $f_H\kappa$ (which we refer to as K, the host population size), while varying the relative population sized by changing f_H and f_P. The dynamics of this model are exclusively evolutionarycompared to ecological, characterized solely by the dynamics of host and pathogen allele frequency. Hosts of type $i = \{1, 2\}$ are infected by pathogens of type $j = \{1, 2\}$ at a rate of $\beta_{i,j} = X/\kappa$ if host and pathogen are "matching" ($i = j$) and at a reduced rate Y/κ if "mis-matching" ($i \neq j$).

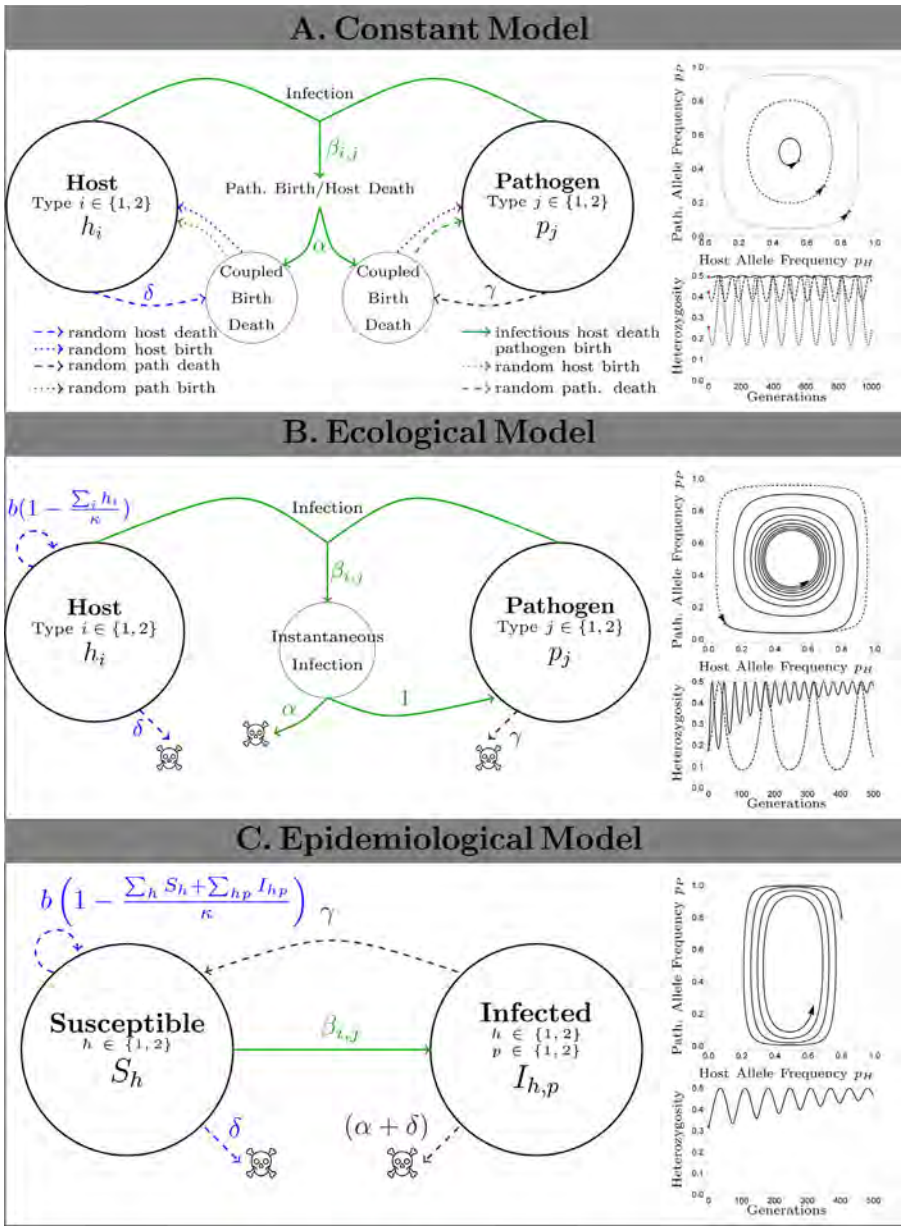


Figure 1. **Model design and deterministic dynamics of the three models explored in this article. Flow diagrams illustrate the design of the (A) constant-size, (B) ecological, and (C) epidemiological models. Green lines indicate infection, blue lines host demographic dynamics, and purple pathogen demographic dynamics. Solid lines indicate processes that are genotype specific, whereas dashed lines indicate events that occur at random with respect to genotype. Top right hand panels show phase-plane dynamics of (A) neutrally stable cycles for three different initial conditions (shown as red points), (B) second-order stable cycles for two-parameter conditions with the same initial condition, and (C) first-order stable cycles. Lower right-hand panels depict the corresponding deterministic dynamics of heterozygosity.**

Shown graphically in Figure 1A, with probability α infection in the constant-size model leads to host death and pathogen birth. To keep host and pathogen population sizes constant, we use a Moran model structure that couples host death with the birth of a host drawn at random with respect to the current allele frequency; similarly, we couple pathogen birth resulting from infection with the death of a randomly selected pathogen.

In addition to infection, host and pathogen population turnover is determined by the death (and coupled random birth) of hosts, which occurs at rate δ, and free-living pathogen death

(and coupled birth), which occurs at rate γ. The resulting dynamics of host and pathogen allele frequencies are given by:

$$\begin{aligned}\frac{dp_H}{dt} &= f_P\alpha(X-Y)(1-2p_P)p_H(1-p_H)\\ \frac{dp_P}{dt} &= -f_H\alpha(X-Y)(1-2p_H)p_P(1-p_P),\end{aligned} \tag{1}$$

where $p_H = \frac{h_1}{h_1+h_2}$ and $p_P = \frac{p_1}{p_1+p_2}$ are the frequencies of type 1 hosts and pathogens, respectively. System (1) is equivalent to the constant-size model presented previously (MacPherson et al. 2020), except that time is no longer normalized in terms of host generations (see supplementary *Mathematica* notebook available on Dryad at https://doi.org/10.5061/dryad.cz8w9gj2c). System (1) has two types of equilibria: first, fixation of one host and one pathogen genotype and second, an internal polymorphic equilibria at $p_H = p_P = 1/2$ (see supplementary *Mathematica* notebook). The fixation equilibria are unstable and the polymorphic equilibria characterized by neutral stability. With a purely imaginary leading eigenvalue, this model is characterized by neutral Red Queen cycles in host and pathogen allele frequencies. Shown in Figure 1A, these Red Queen allele-frequency cycles correspond to cycles in host heterozygosity, the measure of genetic variation that we will use throughout. Although host heterozygosity fluctuates in this coevolutionary model, it neither increases nor decreases when averaged over time. Hence, the neutral stability of the polymorphic equilibrium in this deterministic MAM suggests that in a finite population heterozygosity might decline, on average, due to random genetic drift in a similar manner seen in a classic one-species neutral model.

ECOLOGICAL MODEL

We explicitly incorporate eco-evolutionary feedbacks using a model of host-parasite coevolution first presented by Frank (1993). Shown graphically in Figure 1B, hosts are born at a per capita density-dependent rate $b(1-\frac{\sum_i h_i}{\kappa})$, where b is the intrinsic growth rate of the pathogen, h_i is the number of hosts of type i and κ is the "growth-limit" of the host, a quantity that is proportional to the carrying capacity. Hosts die from natural causes at rate δ. In the absence of parasitism the host reaches an equilibrium population size (carrying capacity) of $\sum_i h_i = \frac{\kappa(b-\delta)}{b}$. As in the constant-size model, infection of hosts of type i by pathogens of type j occurs at rate $\beta_{i,j}$ as defined above. Infections occur instantaneously in this model such that infected hosts die instantly with probability α (and survive with probability $1-\alpha$). All infections lead to the instantaneous birth of a pathogen, while free-living pathogens die at rate γ. The resulting dynamics are given by:

$$\begin{aligned}\frac{dh_i}{dt} &= bh_i\left(1-\frac{\sum_k h_k}{\kappa}\right) - \alpha h_i\sum_j \beta_{i,j}p_j - \delta h_i\\ \frac{dp_j}{dt} &= p_j\sum_j \beta_{i,j}h_i - \gamma p_j.\end{aligned} \tag{2}$$

There are four types of equilibria of system (2): host and parasite extinction, disease-free hosts, disease persistence with the fixation of one host and one pathogen type, and finally disease persistence with a polymorphic host and polymorphic pathogen with allele frequencies at 1/2 (see supplementary *Mathematica* notebook). Host extinction occurs whenever $b < \delta$. The system reaches the disease-free equilibrium whenever $b > \delta$ and $\gamma > \frac{(X-Y)(b-\delta)}{2b} \equiv \gamma^*$, defining a critical pathogen death rate, γ^*, above which the pathogen will not persist. As in the constant model, the equilibria with pathogen persistence and fixation of one host and one pathogen type are never stable, therefore when $b > \delta$ and $\gamma < \gamma^*$ the dynamics of system (2) is described by the stability of the following internal polymorphic equilibrium:

$$\begin{aligned}\hat{h}_i &= \frac{\gamma\kappa}{X+Y} = K\\ \hat{p}_j &= \frac{\kappa(b(-2\gamma+X+Y)-\delta(X+Y))}{\alpha(X+Y)^2},\end{aligned} \tag{3}$$

where K here is defined as the total equilibrium host population size, for comparison with the constant population size model. The linear stability analysis gives two complex conjugate pairs of eigenvalues. The first, leading, pair of eigenvalues is purely imaginary. Corresponding to eigenvectors in the direction of the host and pathogen allele frequencies (leaving the total population sizes constant). These leading eigenvalues predict the existence of neutrally stable allele-frequency cycles, although as we will describe next this neutral stability is not robust. The second complex conjugate pair may be real or complex, but in either case, the real part is always negative when $\gamma < \gamma^*$. Corresponding to eigenvectors altering the total host and pathogen population sizes (leaving allele frequencies constant), these eigenvalues describe the, often rapid, stabilization of host and pathogen population size.

Despite the often drastic transient effects of host-pathogen interactions on population dynamics, the neutral stability of both this ecological model and the constant population size model has been used to argue that population dynamics have little effect on coevolutionary dynamics (Nuismer 2017). Numerical exploration of the model reveals, however, that despite the neutral linear stability of system (2), the dynamics cycle inwards toward the polymorphic equilibrium (Frank 1993; Ashby et al. 2019). As illustrated in Figure 1B, allele frequencies spiral inward toward the

equilibrium revealing second-order stability not captured by the first-order (linear) stability analysis. As a result mean host heterozygosity increases, on average, over time in this deterministic model.

In contrast to the constant-size model, the second-order stability arises from the additional feedback between host and parasite dynamics that arises as a result of density-dependent population growth. This is illustrated by the contrast between the dynamics observed here relative to those of Song et al. (2015), whose model does not include explicit intraspecific density dependence and does not exhibit second-order stability (neutrally stable cycles are found around the equilibrium). The second-order stability of system (2) has remained largely uncharacterized in previous explorations of similar models, as the stabilization is inherently slow and difficult to observe when dynamics are only examined numerically over short time scales (Nuismer 2017). In another similar ecological model, Rabajante et al. (2016) explore the effect of different types of ecological responses (e.g., form of the functional response) on the persistence of cycles involving multiple alleles. However, they do not explicitly quantify the maintenance of genetic variation.

Unlike a first-order stable equilibrium, it is difficult to obtain an analytical expression for the magnitude of this stabilizing effect. Instead, we compute a numerical one by fitting the host allele-frequency dynamics with the sinusoidal function using the NonlinearModelFit function in *Mathematica* (see supplementary files):

$$p_H(t) \approx ae^{rt}\sin\left(\sqrt{\frac{\gamma(b(-2\gamma+X+Y)-\delta(X+Y))}{X+Y}}t+t_0\right), \quad (4)$$

where a, r, and t_0 are fitted parameters, the quantity r approximating the rate of stabilization. The coefficient of t in this expression is the period of the allele-frequency cycles as predicted from the linear stability analysis.

EPIDEMIOLOGICAL MODEL

To model coevolution between a host and a directly transmitted pathogen we use a SIS compartmental model. In this model, hosts of type $h \in \{1, 2\}$ are characterized by their infection status as either susceptible, denoted by S_h or infected, $I_{h,p}$, where $p \in \{1, 2\}$ is the genotype of the infecting pathogen. Shown schematically in Figure 1C, all hosts give birth regardless of infection status at a per capita density-dependent rate $b(1-\frac{\sum_k S_k+\sum_{k,l} I_{k,l}}{\kappa})$ with all individuals susceptible at birth. Similarly, hosts die of natural causes at a constant per capita rate δ regardless of their infection status. κ is once again proportional to the host carrying capacity in the absence of infection. Hosts of type h are infected with pathogens of type p at rate $\beta_{h,p}$ defined as above. Once infected hosts die from virulent causes at a rate α. Finally, infected hosts recover at a rate γ. As with the free-living death rate, γ in the epidemiological model captures pathogen death (turnover) that is independent of death (turnover) of the host. The resulting SIS model is given by the following system of differential equations:

$$\begin{aligned}\frac{dS_h}{dt} &= b\left(S_h+\sum_l I_{h,l}\right)\left(1-\frac{\sum_k S_k+\sum_{k,l} I_{k,l}}{\kappa}\right)-\delta S_h \\ &\quad -\sum_{k,l}\beta_{h,l}S_h I_{k,l}+\gamma\sum_l I_{h,l} \\ \frac{dI_{h,p}}{dt} &= \beta_{h,p}S_h\sum_k I_{k,p}-(\delta+\alpha)I_{h,p}-\gamma I_{h,p}. \end{aligned} \quad (5)$$

The total host population size in this system is given by $\sum_h S_h + \sum_{h,p} I_{h,p}$. In the absence of coinfection and mutation, as assumed here, there is no within-host evolution and hence from an evolutionary perspective the pathogen population size is equal to that of the infected hosts, $\sum_{h,p} I_{h,p}$. Hence, the pathogen population size in this epidemiological model is always less than that of the host.

System (5) has the same four types of equilibria as the ecological model; host extinction, disease-free hosts, endemic equilibria with fixation in both the host and pathogen, and an endemic-polymorphic equilibrium with a host and pathogen allele frequency of 1/2 (see supplementary *Mathematica* notebook and Appendix A1 for the complete equilibrium and stability analysis). The stability of each equilibrium, as well as the transient dynamics of system (5), is determined by the basic reproductive number, $R_0(\phi)$, which is the number of secondary infections per infected host which in a coevolutionary model depends on the fraction of matching hosts ϕ (see Appendix), where $R_0(\phi)$ is:

$$R_0(\phi) = \frac{(X\phi+Y(1-\phi))(b-\delta)}{b(\alpha+\gamma+\delta)}. \quad (6)$$

In particular, the fourth polymorphic equilibrium is determined by setting the frequency of matching hosts to 1/2 and is stable whenever $R_0(1/2) > 1$. The stability of the endemic-polymorphic equilibrium of this SIS model reiterates the results of MacPherson et al. (2018), which found that epidemiological dynamics stabilize Red Queen cycles.

For our purposes of exploring of the maintenance of genetic variation, we will focus exclusively on parameter conditions for which the endemic-polymorphic is stable, $R_0(1/2) > 1$. The first-order stability of this equilibrium establishes an interesting contrast between the epidemiological model, the second-order stability of the ecological model, and finally the neutral stability of the constant-size model. Developing predictions for the maintenance of genetic variation in finite populations from the stability of the deterministic models alone we should expect that each additional form of feedback between host and pathogen (i.e., coevolutionary feedback, density-dependence ecological feedback, and epidemiological feedback) would introduce an additional degree of

stability to the polymorphic equilibrium, which in turn would help maintain genetic variation. Before we can contrast the maintenance of genetic variation across models, however, we must understand how the parameters of each individual model affect the dynamics of heterozygosity in finite populations.

Stochastic Dynamics

In this section, we simulate drift in finite population within the three models considered above to determine the impact of coevolutionary feedback on genetic variation. Run in C++, the individual-based simulations for each model were based on a Gillespie algorithm (Gillespie 1977) with rates given by the flow diagrams shown in Figure 1. The effect of coevolution on the maintenance of genetic variation is examined by comparing the expected heterozygosity in a model of host-parasite coevolution, H_{coev}, measured as the mean heterozygosity averaged across 1000 replicate sample paths, to the expected heterozygosity simulated under neutral genetic drift H_{neut} (see Appendix A2). In contrast to the constant-size model for which there is an expression for neutral drift that allows the analytical exploration of the dynamics of heterozygosity (see MacPherson et al. 2020), there is no such expression for the ecological and epidemiological birth-death processes. Given the results of the constant-size model, however, we expect the tendency to maintain genetic variation, measured as $\Delta H = H_{coev} - H_{neut}$, to be a function of four quantities: the total system size (K), the stability of the polymorphic equilibrium, the relative turnover rates in the host versus the pathogen, and the probability of fixation in the pathogen. Summarized by the top row of Table 1, for each model we then sample parameters pseudo-randomly across parameter space in a manner designed to most effectively test the effect of these four factors. All simulations were run for 250 host generations with the results presented in the main text focused on explaining variation in ΔH at the last host generation.

When population size varies over time, the effective population size, N_e, is given by the harmonic mean of the total population size (Crow 1970). For the ecological and epidemiological model, the effective population size is therefore determined by two factors, the equilibrium population size and variation about this equilibrium due to stochasticity, density dependence, and Red Queen cycles. To separate these two effects, we fix the expected equilibrium host population sizes to a constant value, K, with the remaining variation in host effective population size due solely to variation about this equilibrium size. Although not possible in the epidemiological model, in the constant-size and ecological models we also fix pathogen population size to K to limit any effect of variation in parasite effective population size. Sampling 100 random parameter combinations for a given value of K we expect a positive correlation between ΔH and N_e, with the latter is estimated from the dynamics of the simulation. We then explore the effect of equilibrium population size by repeating all 100 simulations for seven values of K ranging on a log scale from 10^2 to $10^{3.5}$. To facilitate comparisons across parameter space the individual-based simulations are analyzed in terms of host generations where one host generations is defined by the death of K individuals in the simulation. As emphasized in our exploration of the deterministic models, the stability of the polymorphic equilibrium differs across models but also varies across parameter space. In particular, the constant-size model is neutrally stable and hence exhibits no variability in stability across parameter space. The ecological model exhibits second-order stability with variation in the degree of stability across parameter space approximated by the variation in the value of r (see equation 4). Finally, the epidemiological model is first-order stable with variation in stability given by the variation in the leading eigenvalue λ, which can be computed numerically.

In addition to total host population size and stability, the analysis of the maintenance of genetic variation in the constant-size model identified two additional factors influencing the dynamics of heterozygosity in finite coevolving populations (MacPherson et al. 2020). The first is the effect of natural selection on the stochastic perturbations in host and pathogen allele frequency. Discussed in detail in MacPherson et al. (2020), reciprocal natural selection between the two species increases (decreases) genetic variation when turnover in the host population is faster (slower) than turnover in the pathogen population. The second effect of finite population size is the rapid erosion of genetic variation in the host following stochastic fixation in the pathogen. Specifically, following fixation in the pathogen, the MAM shifts from a model of symmetric reciprocal natural selection to one of directional selection favoring the "mis-matching" host of the remaining pathogen genotype. The importance of directional selection is determined primarily by the probability of allele fixation in the pathogen.

CONSTANT MODEL

As in the previous paper, we begin by considering the case where the host and pathogen population sizes are equal, $f_H = f_P = 1$ and $\kappa = K$. The relative turnover in the host and parasite population is determined by $\delta - \gamma$. By drawing both δ and γ randomly between 0.1 and 1, we ensure that the turnover rate is not exceedingly small and that the distribution of relative turnover rates in the host versus the pathogen is symmetric. Throughout we consider the strength of selection to be determined by the extent of matching $0 < Y < X < 1$ and pathogen virulence α, which in this case is a probability that we let vary between 0 and 1. At generation 250, we use a loess smoothed mean fit and an accompanying 100 bootstrap fits (resampling with replacement from the 1000 sample paths per parameter set), to discern the average

Table 1. Individual-based simulation parameter selection.

	Constant-size	Ecological	Epidemiological
Random Sampling	Time scale and Population Size $f_H = f_P = 1$ equal host and path. $\kappa = K$ population size Turnover $\delta \sim U[0.1, 1]$ Generation time not infinitesimally small $\gamma = [0.1, 1]$ symmetric perturbations Selection $0 < Y < X \leq 1$ $\alpha = [0, 1]$	Time scale and Population Size $b = \frac{(X+Y)(\alpha Y+\delta)}{X+Y-2\gamma}$ $\sum_i \hat{h}_i = \sum_j \hat{p}_j$ $\kappa = K\frac{X+Y}{2\gamma}$ Turnover $\delta \sim U[0.1, 1]$ comparable to constant-size $\gamma = [0.1, \gamma^*]$ stable equilibrium note: perturbations are symmetric by inspection Selection $0 < Y < X \leq 1$ $\alpha = [0, 1]$	Time Scale and Population Size $b = 1$ sets time scale κ such that $\sum_h \hat{S}_k + \sum_{k,l} \hat{I}_{k,l} = K$ Turnover $\delta \sim U[0.1, 1]$ non-zero turnover $\gamma = \left[0, \frac{X+Y-2(X+Y)\delta}{2}\right]$ stable equilibrium: $R_o(1/2) > 1$ note: perturbations are symmetric when $\gamma = 0$ Selection $0 < Y < X \leq 1$ $\alpha = \left[0, \frac{X+Y-2(X+Y)\delta-\gamma}{2}\right]$ stable equilibrium: $R_o(1/2) > 1$
Model Comparison	Time scale and Population Size $f_H = \frac{\hat{h}^*}{\kappa}$ $f_P = \frac{\hat{p}^*}{\kappa}$ $\kappa = K$ Turnover $\delta = \delta^*$ $\gamma = \delta^* \frac{\hat{h}^*}{\hat{p}^*}$ equal perturbations Selection $X = X^*$ $Y = Y^*$ $\alpha = \frac{\alpha^*}{\alpha^*+\delta^*}$ equal probability of death from infection	Time scale and Population Size $\gamma = \frac{\hat{h}^*(X^*-Y^*)}{2\kappa}$ $\kappa = K\frac{X^*+Y^*}{2\gamma}$ Turnover $\delta = \frac{\hat{p}^*(X^*-Y^*)}{2\kappa}$ equal perturbations Selection $X = X^*$ $Y = Y^*$ $\alpha = \frac{\alpha^*}{\alpha^*+\delta^*}$ equal probability of death from infection	Time scale and Population Size $b = 1$ sets time scale κ such that $\sum_k \hat{S}_k + \sum_{k,l} \hat{I}_{k,l} = K$ $\sum_k \hat{S}_k + \sum_{k,l} \hat{I}_{k,l} \equiv \hat{h}^* \quad \sum_{k,l} \hat{I}_{k,l} \equiv \hat{p}^*$ Turnover $\delta \sim U[0.1, 1] \equiv \delta^*$ non-zero turnover $\gamma = 0$ equal perturbations Selection $0 < Y < X \leq 1 \quad X = X^* \, Y = Y*$ $\alpha = \left[0, \frac{X+Y-2(X+Y)\delta}{2}\right] \equiv \alpha^*$ stable equilibrium: $R_o(1/2) > 1$

Row 1: Parameter conditions and constraints used to randomly sample across parameter space. *U*[*a*, *b*] denote uniformly distributed random variables. For the epidemiological model, κ is chosen such that the equilibrium host population size equals *K*. Results of corresponding individual-based simulations are shown in Figures 2 and 3. Row 2: Parameter conditions and constraints used to sample parameter combinations for model comparison. Results of corresponding simulations are shown in Figure 4. Note, all simulations were run for 250 generations, with one generation defined as *K* host death events.

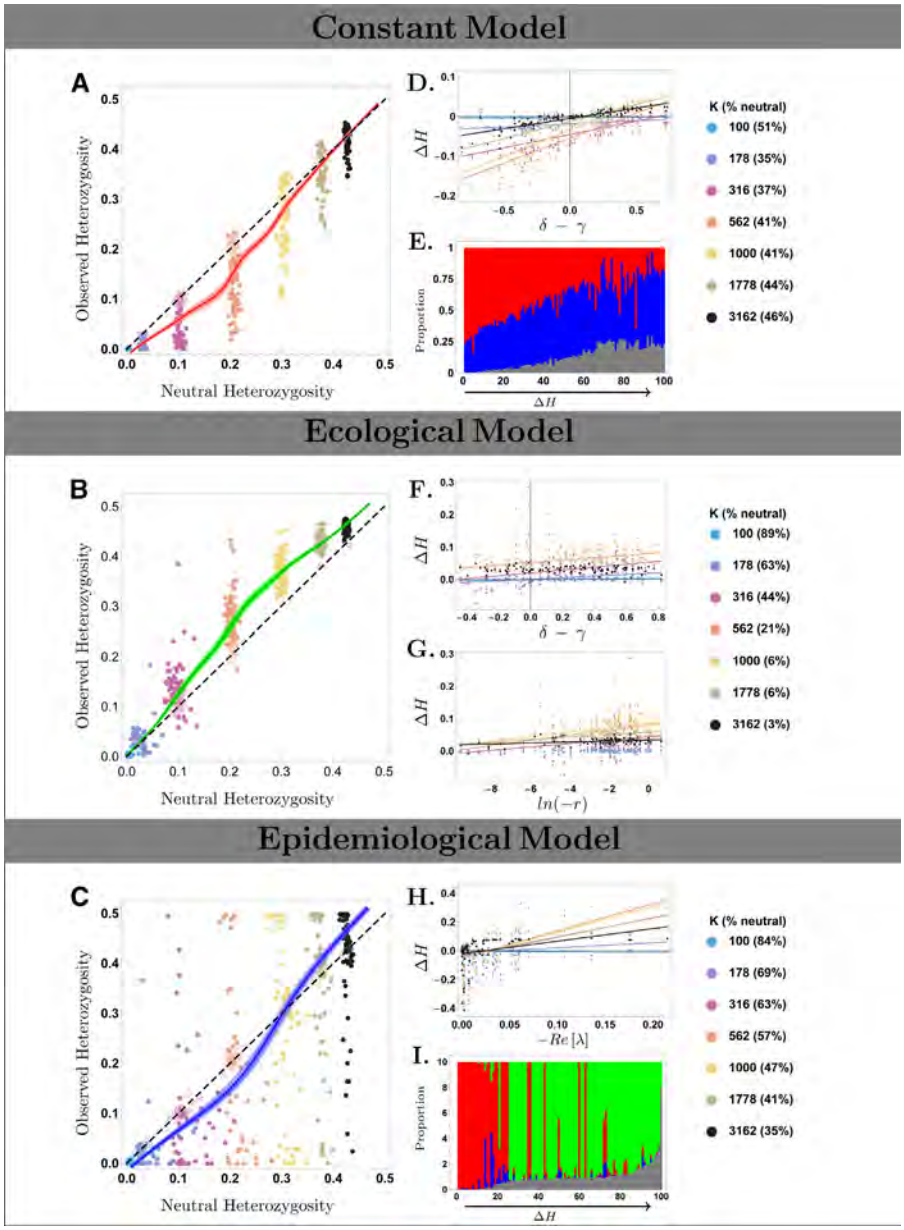


Figure 2. **Observed heterozygosity within a coevolutionary model versus the neutral expectation. Left-hand panels: observed heterozygosity versus (simulated) neutral expectation for (A) constant-size, (B) ecological, and (C) epidemiological models. Each point represents the mean of 1000 replication simulations for a given parameter set (solid if different from the neutral expectation, hallow if not). Black dashed line gives the neutral expectation, dark colored lines gives loess smoothed fits to the observed data whereas light colored lines are smoothed fit for 100 bootstrap samples. Panels (D, F): Effect of relative host versus pathogen turnover on ΔH (shown as lines when slopes differ significantly from 0). Panels (G, H): Effect of the deterministic stabilizing force (measured by r, equation 4) in (F) and by the real part of the leading eigenvalue $-Re[\lambda]$ in G. Panels (E, I): Bar charts showing the proportion of simulation replicates resulting in directional selection following fixation in the pathogen (red), pathogen extinction (green), fixation in the host (blue), or ongoing coevolution (gray) for the 100 parameter conditions for $K = 10^{2.5}$ arranged in order of increasing ΔH. All plots are shown at generation 250. Legends give value of the equilibrium host population size K and the percentage of parameter sets for which the mean of the 1000 replicate simulations did not differ significantly from the neutral expectation.**

difference in heterozygosity between the coevolutionary model and the neutral expectation (see supplementary *Mathematica* file). As predicted by the neutral stability of the deterministic dynamics, 42% of points in the constant-size model (Fig. 2A) do not differ significantly from the neutral expectation (MacPherson et al. 2020).

Given that host and parasite population sizes remain constant, there exists no variation in effective population size for a

given value of K. Similarly, as the constant-size model exhibits true neutral stability, there exists no variation in stability across parameter space, leaving only two factors to explain the variation in ΔH, the relative turnover rates and the probability of allele fixation in the pathogen. Explored in greater analytical detail in MacPherson et al. (2020), genetic variation increases as a function of $\delta - \gamma$, with higher host turnover ($\delta > \gamma$) helping maintain genetic variation (Fig. 2D). In addition to the effect of turnover, simulations that maintain more variation (higher values of ΔH) tend to be those where sample paths tended not to fix for one parasite type but remained polymorphic in both the host and the parasite (we refer to such cases as exhibiting "ongoing coevolution"); see Figure 2E.

ECOLOGICAL MODEL

The pseudo-random sampling of parameters in the ecological model closely resembles that of the constant-size model. In particular, we focus on the case when $\sum_i \hat{h}_i = \sum_j \hat{p}_j = K$. As with the constant-size model we draw γ and δ from the same distribution so that relative turnover rates in the host and pathogen are, on average, symmetric and not exceedingly small.

We find that variation is maintained more often in the ecological model, relative to the constant-size model (Fig. 2B), although the total amount of variation maintained is only slightly greater on average than expected in a neutral model with an equal host population size. The greater ability of the ecological model to maintain variation is consistent with the behavior of the deterministic models where, in contrast to the neutral stability of the constant-size model, the ecological model exhibits second-order stability.

Relative to the constant-size model, there exists less variability in the amount of genetic variation maintained (ΔH) across parameter sets, especially when compared to the variability in genetic variation observed among the 1000 stochastic sample paths simulated with the same parameters (see Fig. A1). This significantly reduces the power to identify correlations between ΔH and factors such as the relative turnover rates (Fig. 2F) or the deterministic stability r (Fig. 2G). Nevertheless the inferred stability of the internal equilibrium in the ecological model, r, is a significant predictor of how much variation is maintained. In addition to the effects of stability, relative turnover rates, as measured by $\delta - \gamma$, affect the extent to which genetic variation is maintained, as in the constant-size model.

Despite the a priori expectation that fluctuations in population size, and a corresponding decrease in effective population size, would be a strong determinant of the maintenance of genetic variation, we in fact observe little variation in the effective population size due to the stabilizing effect of density dependence on the host and consequently on the parasite population sizes (as captured by the strongly negative nonleading eigenvalues that involve the population sizes, see Appendix A3). Finally, like the constant-size model, genetic fixation in the pathogen followed by directional selection can occur in the ecological model, although the likelihood of this occurring is reduced due to the second-order stability. When an allele fixes in the pathogen, this was often rapidly followed by pathogen extinction, which was not possible in the constant-size model and which resulted in neutral drift in the host rather than in directional selection. As a result, we observe little effect of directional selection on ΔH in this model.

EPIDEMIOLOGY MODEL

One key distinction between the ecological and epidemiological model is the free-living versus directly transmitted nature of the pathogen. One consequence of direct transmission and the absence of coinfection is that the effective parasite population size, $\sum_{h,p} I_{h,p}$, is restricted to being a subset of the host population, $\sum_h S_h + \sum_{h,p} I_{h,p}$. Therefore, unlike the ecological and constant population size model above we are unable to consider the case of equal host and pathogen population sizes. Instead we focus on the case where the equilibrium host population size is fixed at K, letting the equilibrium pathogen population size vary. A second consequence of direct transmission is that pathogen turnover is no longer independent of host turnover. Specifically, the per capita rate of pathogen turnover, $(\gamma + \delta)$, can never be less than that of the host, δ, and will exceed that of the host if $\gamma > 0$, with the special case of $\gamma = 0$ and hence equal turnover rates explored in more detail in the next section.

Relative to the previous models, ΔH in the epidemiological model is highly variable for a given value of K (Fig. 2C). On average however, heterozygosity is reduced relative to the neutral expectation when K is small and above the neutral expectation when K is large. This is a result of the balance between equilibrium stability and directional selection following fixation in the pathogen. In the absence of pathogen fixation, the deterministic model predicts that heterozygosity should exceed the neutral expectation, which is borne out when K is large and when fixation in the pathogen is rare. As the host, and more importantly the pathogen, population size declines as K decreases, the probability of pathogen fixation followed by the loss of host genetic variation increases.

These same processes explain variation across parameter space for a given value of K. Increasing the stability of the internal equilibrium increases genetic variation (Fig. 2H). Similarly, cases for which genetic variation far exceeds the neutral expectation (large positive ΔH) are characterized primarily by pathogen extinction, and hence largely neutral host evolution, whereas cases where genetic variation falls below the neutral expectation (negative ΔH) are dominated by allele fixation in the pathogen and sustained directional selection in the host (Fig. 2I). As with the ecological model, the effective population size is a

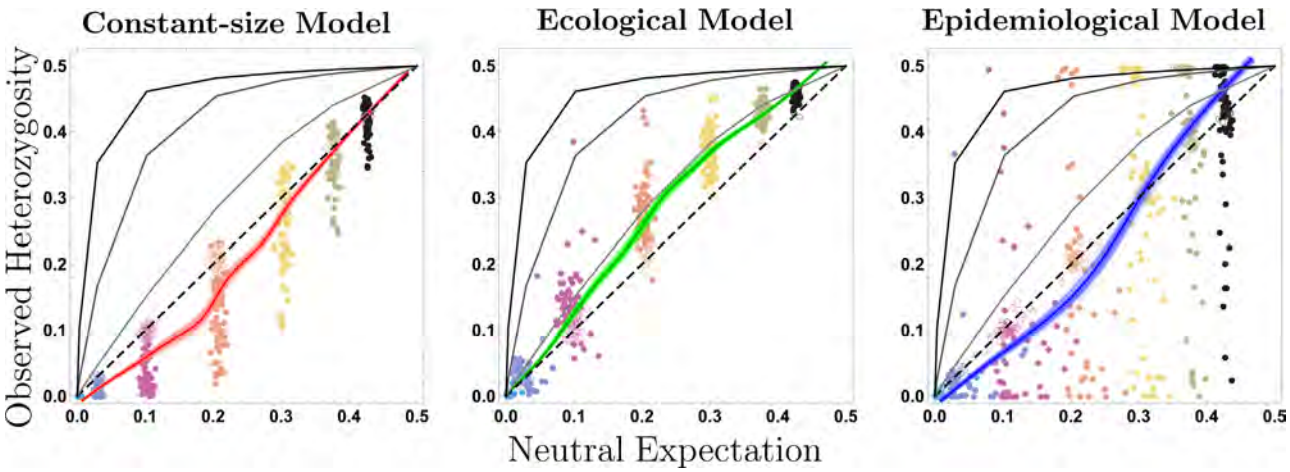


Figure 3. **Maintenance of genetic variation relative to overdominance. The data in Figure 2 are compared to the expected amount of variation at generation 250 in a host-population subject to overdominance in the absence of parasites (calculated using the Moran model with exact matrix iteration, see supplementary *Mathematica* file). Overdominance model: $W_{AA} = 1 - s, W_{Aa} = 1, W_{aa} = 1 - s$, where $s = 0.01$ light gray, $s = 0.05$ gray, and $s = 0.1$ black.**

weak predictor of ΔH heterozygosity in both neutral and coevolutionary models are impacted by the effective population size. The final potential determinant of ΔH we considered is the relative turnover rates in the host versus pathogen. Unlike the constant-size and ecological models we find no significant effect of relative turnover rates, here measured by $-\gamma$ (i.e., the host turnover rate, δ, minus that of the parasite, $\delta - \gamma$). Even if we similarly constrain the constant-size and ecological models such that host turnover rats are greater than parasite turnover rats, the effect of relative turnover rates remains significant for those models but not for the epidemiological model (see Fig. A3). This is perhaps not surprising given that the epidemiological model is more complex, with hosts structured into susceptible and infectious classes, leading to first-order stability of the endemic equilibrium.

In summary, host heterozygosity in these coevolutionary models evolves nearly neutrally, particularly in contrast to the maintenance of genetic variation due to other process, for example, overdominance (see Fig. 3). Nevertheless we find that genetic variation in coevolving populations can be maintained slightly more often than expected in a neutral model when host populations are large, as a result of the stabilizing processes such as density-dependent growth or epidemiological feedback. By contrast, genetic variation is often rapidly depleted when host populations are small, as a result of directional selection following fixation in the pathogen.

Model Comparison

In this section we explore in more detail the effect of stability and directional selection by comparing the maintenance of genetic variation across models. One important determinant of the maintenance of genetic variation is host and pathogen population sizes. Direct comparison between the simulation results from the models above (Fig. 2) is challenging because the models place different constraints on the population size. With a directly transmitted parasite (as in the epidemiological model) the parasite population size is restricted to being less than that of the host population size (ignoring coinfection and within-host dynamics). In contrast, for the constant-size and ecological models, the host and population sizes were equal, as defined by K. In this section our aim is to sample parameters in such a manner as to optimize the comparison across models (see Table 1). Hence in contrast to the simulations above we allow the value of K to differ between hosts and parasites. We then equate host and parasite population sizes across models. As a result, the model comparison below focuses solely on cases where the host population size is greater than that of the pathogen.

Another factor affecting the maintenance of genetic variation is the relative turnover rates in the host and pathogen. Hence we eliminate this effect by setting the host and pathogen turnover rates equal to one another. Finally, whereas α in the epidemiological model is the rate of host death from infection, α in the ecological and constant-size models is the probability of death from infection. Hence, we equate the value of this parameter across models accordingly such that the probability of death from infection remains constant. As we did in the previous section we then sampled randomly across this constrained parameter space running simulations for 100 random parameter combinations each replicated for seven values of K. Once again each of these 700 simulations consists of 1000 replicate stochastic sample paths run over 250 host generations.

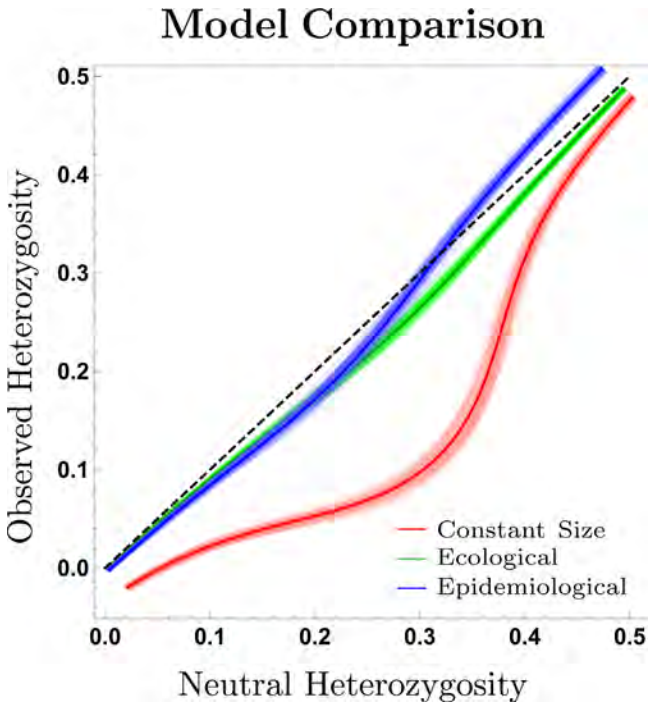


Figure 4. Comparison of observed heterozygosity across models. The observed heterozygosity relative to the expectation under neutral genetic drift for the constant population size (red), ecological (green), and epidemiological (blue) models. Black dashed line shows neutral expectation, dark colored lines give the loess smoothed fit to the simulated data across parameters, whereas light colored lines give smoothed fit to 100 bootstrap samples. Parameter conditions for simulations were drawn to optimize model comparison as given in the second row of Table 1.

Figure 4 gives an across model comparison of the smoothed mean fit (and bootstrap confidence intervals) for the observed heterozygosity versus the expectation under simulated neutral drift. The foremost result is that deterministic stability is an overall good predictor of the relative maintenance of genetic variation, with the constant-size model less able to maintain variation than the ecological model, which in turn is less able to maintain variation than the epidemiological model. In addition, as we concluded previously, coevolutionary systems rarely exhibit more genetic variation than would be expected under neutrality, and when they do this increase is small. Equated in this way, in fact, we only observe an increase in genetic variation in the highly stable epidemiological model at large population sizes.

The overall decrease in genetic variation across models relative to the neutral expectation is due to directional selection following allele fixation in the pathogen (see Fig. A4). By constraining pathogen population size to be small relative to that of the host (as required by the epidemiological model), the probability of fixation in the pathogen, and resulting directional selection, is increased. This is exemplified by the constant population size model for which the vast majority of simulations experience directional selection. This comparison across models demonstrates therefore the overall importance of directional selection and deterministic stability on the internal equilibrium in determining the maintenance of genetic variation in systems exhibiting host-parasite coevolution.

Discussion

Beginning with Haldane (1949), reciprocal natural selection between hosts and pathogens has been suggested as a mechanism to maintain genetic variation. Antagonistic selection, it was argued, should favor genotypes that are rare in the interacting species resulting in a form of indirect NFDS. MacPherson et al. (2020) found that this is not the case, at least in a model where the host and pathogen population sizes are forced to remain constant. This result reinforces the previous work in other coevolutionary systems (Brown and Tellier 2011) that found that unlike NFDS within a single species (Takahata and Nei 1990), indirect NFDS does not maintain genetic variation. Instead genetic variation declines at a rate similar to that seen under neutral genetic drift alone. This finding is consistent with the neutral stability of the allele-frequency cycles in the infinite population size limit. Population sizes of hosts and parasites are not, however, expected to remain constant. Coevolution in natural systems must be modeled in a broader ecological and epidemiological context, both of which have been previously shown to stabilize Red Queen allele-frequency cycles (Ashby et al. 2019; MacPherson et al. 2018). By examining and comparing the maintenance of host genetic variation across a range of ecological and epidemiological contexts we examine if and when we are likely to observe the maintenance of polymorphism in both hosts and parasites needed to ensure ongoing coevolution in nature.

One key finding of our models is that, regardless of ecological and epidemiological context, coevolutionary systems, on average, behave nearly neutrally. More specifically, summarized by the rough comparison to the maintenance of genetic variation in a model of overdominance (Fig. 3), the maintenance of genetic variation is, at best, equivalent to very weak overdominance across the three coevolutionary models explored here. Specifically, when variation was maintained at a higher rate than neutral, the effect was equivalent to selective advantage of heterozygotes of $<1\%$ across parameter sets, on average, despite the much stronger coefficients used in the coevolutionary models. Hence, like neutral drift, genetic variation will only persist in populations on the order of N generations, making it unlikely to observe ongoing coevolution in small isolated populations.

Despite this average neutral behavior, there exists substantial variation in the maintenance of genetic variation across models and across parameter space within each model, allowing us

to predict the conditions under which host and parasite polymorphism is most likely to persist. We are most likely to observe reciprocal natural selection maintained in systems with large parasite effective population sizes, where the parasite is unlikely to fix and generate directional selection in the host. We thus expect, in general, more variation to be maintained in both host and parasite populations where there are large free-living pathogen populations than with directly transmitted diseases whose population sizes are limited by the host, all else being equal. For directly transmitted infections, pathogen population size increases with transmission rate and decreases with virulence. Note however, that reducing virulence reduces the strength of reciprocal natural selection and subsequently the frequency of Red Queen allele-frequency cycles. As a result, even in populations where coevolutionary dynamics are likely to persist they may be difficult to observe. To observe rapid coevolutionary changes in these systems there must be strong genetic specificity, $X \gg Y$. For a given host and parasite effective population size, we are most likely to find ongoing coevolution between hosts and their directly transmitted infections. In contrast to the free-living or indirectly transmitted parasites of most model systems of coevolution, this result provides an argument for why directly-transmitted infections may be especially informative examples of coevolution.

Another key result of these models is that the outcome of host-parasite interactions depends qualitatively on the ecological and epidemiological context in which that interaction occurs. Two coevolving systems with the same equilibrium population sizes and allele frequencies can have fundamentally different effects on the long-term maintenance of genetic variation and the persistence of coevolutionary dynamics, depending on the eco-evolutionary processes that stabilize the small perturbations from that equilibrium.

Although the models presented here capture different forms of host-parasite feedback, they represent only a small set of all possible ecological and epidemiological contexts in which host-parasite interactions can occur. For example, many pathogens, such as malaria, are vector-transmitted. Vector transmission introduces temporal delays into the epidemiological system that can lead to the allele frequencies in the host and pathogen to changing asynchronously. Such asynchrony has been shown to help maintain genetic variation in other contexts (Brown and Tellier 2011). Similarly, we have assumed that only a single host species can be infected, but many pathogens are capable of infecting multiple hosts (e.g., this is true for 60% of human pathogens (Taylor et al. 2001)). In particular, the classic natural system for studying coevolution between the trematode parasite *Microphallus* sp. and the New Zealand mud snail *Potamopyrgus antipodarum* involves a multihost life cycle requiring the infection of both snails and waterfall (King et al. 2011). Sharing a common pathogen can result in a form of apparent competition among hosts and is known to influence the evolution of epidemiological traits such as susceptibility, virulence, and resistance (Woolhouse et al. 2001). Finally, we have ignored coinfection and competition within a host. Allowing coinfection would allow the parasite population size to be greater than that of the host in the epidemiological model and could alter parasite allele-frequency dynamics via within-host selection. Understanding how our results extend to this broader range of life cycles would be valuable for predicting the overall contribution of antagonistic coevolution to the maintenance of genetic variation in nature.

As host and pathogen polymorphism rarely persists over long periods of time within isolated populations, processes such as mutation and migration likely play a predominant role in shaping patterns of host-parasite evolution and genetic variation in nature. Consistent with the geographic mosaic theory of coevolution (Thompson 1994, 2005), loss of variation due to genetic drift coupled with its reintroduction through migration will generate spatial variability in the presence/absence of ongoing coevolution. Spatial patterns of host and parasite local adaptation will be determined by the balance between coevolutionary selection, genetic drift and migration all of which themselves will depend on the ecological and epidemiological dynamics within individual demes. Taken together our results reinforce that indirect NFDS in the MAM is generally not a strong force for maintaining variation in and of itself. Nevertheless, the contrast between the constant-size and ecological and epidemiological models exemplifies the importance of considering the full ecological and epidemiological context in which interaction occurs.

AUTHOR CONTRIBUTIONS

AM conceived of the project, built and analyzed the models, and wrote the manuscript. MJK contributed to model development and analysis. SPO assisted with model development, analysis, and with manuscript edits.

ACKNOWLEDGMENTS

We thank Matt Pennell, Jonathan Davies, Jeff Joy, Dan Coombs, Aurélien Tellier, and two anonymous reviewers for their many helpful suggestions that improved this manuscript. This project was supported by a fellowship from the Godfrey-Hewitt Mobility Award and the AAUW dissertation fellowship and a fellowship from the University of British Columbia to A.M. as well as a Natural Sciences and Engineering Research Council of Canada grant to S.P.O. (RGPIN-2016-03711).

DATA ARCHIVING

Supplementary materials are available at https://doi.org/10.5061/dryad.cz8w9gj2c.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

LITERATURE CITED

Ashby, B., and M. Boots. 2015. Coevolution of parasite virulence and host mating strategies. Proc. Natl. Acad. Sci. U.S.A. 112:13290–13295.

Ashby, B., and S. Gupta. 2014. Parasitic castration promotes coevolutionary cycling but also imposes a cost on sex. Evolution 68:2234–2244.

Ashby, B., R. Iritani, A. Best, A. White, and M. Boots. 2019. Understanding the role of eco-evolutionary feedbacks in host-parasite coevolution. J. Theor. Biol. 464:115–125.

Best, A., A. White, and M. Boots. 2008. Maintenance of host variation in tolerance to pathogens and parasites. Proc. Natl. Acad. Sci. U.S.A. 105:20786–20791.

Best, A., A. White, and M. Boots. 2010. Resistance is futile but tolerance can explain why parasites do not always castrate their hosts. Evolution 64:348–357.

Best, A., B. Ashby, A. White, R. Bowers, A. Buckling, B. Koskella, and M. Boots. 2017. Host– parasite fluctuating selection in the absence of specificity. Proc. Royal Soc. B 284:20171615.

Brown, J. K. M., and A. Tellier. 2011. Plant-parasite coevolution: bridging the gap between genetics and ecology. Annu. Rev. Phytopathol. 49:345–367.

Clarke, B. 1979. The evolution of genetic diversity. Proc. R. Soc. Lond. B 205:19–40.

Crow, J. F. 1970. Genetic loads and the cost of natural selection. Pp. 128–177 *in* K.-I. Kojima, ed. Mathematical topics in population genetics, biomathematics. Springer, Berlin.

Ebert, D. 2005. Ecology, epidemiology, and evolution of parasitism in *Daphnia*. National Library of Medicine (US), National Center for Biotechnology Information, Bethesda, MD.

A Frank, S. 1993. Specificity versus detectable polymorphism in host-parasite genetics. Proc. R. Soc. Lond. B 254:191–197.

Gandon, S., Y. Capowiez, Y. Dubois, Y. Michalakis, and I. Olivieri. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. Proc. R. Soc. Lond. B 263:1003–1009.

T Gillespie, D. 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81:2340–2361.

Gokhale, C. S., A. Papkou, A. Traulsen, and H. Schulenburg. 2013. Lotka-Volterra dynamics kills the Red Queen: population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. BMC Evol. Biol. 13:254.

Haldane, J. 1949. Disease and evolution. Sci. Ricerca 19:68–76.

Hill, W. 1972. Effective size of populations with overlapping. Theor. Popul. Biol. 289:278–289.

King, K. C., J. Jokela, and C. M. Lively. 2011. Trematode parasites infect or die in snail hosts. Biol. Lett. 7:265–268.

MacPherson, A., S. P. Otto, and S. L. Nuismer. 2018. Keeping pace with the Red Queen: identifying the. Genetics 208:779–789.

MacPherson, A., M. J. Keeling, and S. P. Otto. 2020. Coevolution does not slow the rate of loss of heterozygosity in a stochastic host-parasite model with constant population size. https://doi.org/10.1101/2020.04.07.024661.

P Moran, P. A. 1958. Random processes in genetics. Math. Proc. Camb. Philos. Soc. 54:60–71.

Myers, J. H., and D. R. Bazely. 2003. Ecology and control of introduced plants. Cambridge Univ. Press, Cambridge.

Nuismer, S. L. 2017. Introduction to coevolutionary theory. 1st ed. Macmillan Learning, New York.

Papkou, A., C. S. Gokhale, A. Traulsen, and H. Schulenburg. 2016. Host–parasite coevolution: why changing population size matters. Zoology 119:330–338.

Penman, B. S., B. Ashby, C. O. Buckee, and S. Gupta. 2013. Pathogen selection drives nonoverlapping associations between HLA loci. Proc. Natl. Acad. Sci. U.S.A. 110:19645–19650.

Rabajante, J. F., J. M. Tubay, H. Ito, T. Uehara, S. Kakishima, S. Morita, J. Yoshimura, and D. Ebert. 2016. Host-parasite Red Queen dynamics with phase-locked rare genotypes. Sci. Adv. 2:e1501548.

Song, Y. X., C. S. Gokhale, A. Papkou, H. Schulenburg, and A. Traulsen. 2015. Host-parasite coevolution in populations of constant and variable size. BMC Evol. Biol. 15. https://doi.org/10.1186/s12862-015-0462-6.

Takahata, N., and M. Nei. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124:967–978.

Taylor, L. H., S. M. Latham, and M. E. Woolhouse. 2001. Risk factors for human disease emergence. Philos. Trans. R. Soc. B 356:983–989.

Thompson, J. 1994. The coevolutionary process. Univ. of Chicago Press, Chicago, IL.

Thompson, J. 2005. Geographic mosaic of coevolution. Univ. of Chicago Press, Chicago, IL.

Wakeley, J. 2009. Coalescent theory: an introduction. Roberts and Company Publishers, Greenwood Village, CO.

Whitlock, M. C., and N. H. Barton. 1997. The effective size of subdivided population. Genetics 146:427–441.

Woolhouse, M. E. J., L. H. Taylor, and D. T. Haydon. 2001. Population biology of multihost pathogens. Science 292:1109–1112.

Associate Editor: J. W. Busch
Handling Editor: A. G. McAdam

Appendix A

Stability and Transient Dynamics of the Epidemiological Model

Shown schematically in Figure A1, and derived in the supplementary *Mathematica* notebook, $R_0(\phi)$, as given by equation (6), determines the transient dynamics of system (5), given the frequency of matching-hosts, ϕ. As noted in the text, $R_0(1/2)$ determines the stability of the endemic equilibrium (stable when $R_0(1/2) > 1$). In addition $R_0(0)$ and $R_0(1)$ determine the ability of the parasite to spread where no hosts match or all hosts match, respectively, as determined by the local stability analysis (see supplementary *Mathematica* file).

When $R_0(1) < 1$ (dark pink area), the pathogen is never able to invade even in a population consisting completely of matching susceptible hosts. When $R_0(1) > 1$ but $R_0(1/2) < 1$, at least one of the two pathogen types will be able to spread initially. This spread will, however, result in host evolution, leading to an increase in the number of resistant hosts and eventual extinction of the disease. This is in contrast to when $R_0(1)$ and $R_0(1/2)$ are both greater than 1 but $R_0(0) < 1$ (light green area) where, although only the matching pathogen type will be able to spread initially, the resulting change in host allele frequency eventually

Table A1. Variation in ΔH across parameter space versus among replicates.

Log(K)	Constant size			Ecological			Epidemiologica		
	W	B	Ratio	W	B	Ratio	W	B	Ratio
2	0.0003	0.001	4.14	0.0011	0.003	2.27	0.0013	0.007	5.74
2.25	0.0013	0.011	8.84	0.0034	0.019	5.62	0.003	0.04	13.21
2.5	0.0029	0.04	13.57	0.0057	0.043	7.5	0.0047	0.08	17.05
2.75	0.0046	0.079	16.97	0.0059	0.048	8.11	0.0053	0.111	21.02
3	0.0054	0.103	19.06	0.0043	0.032	7.41	0.0047	0.123	26.43
3.25	0.0049	0.098	19.85	0.0027	0.017	6.47	0.0035	0.116	33.22
3.5	0.0038	0.073	19.52	0.0017	0.009	5.6	0.0024	0.091	37.29

Column 1: The average variation in the observed heterozygosity H_{coev} among the 1000 replicates within (W) each of the 100 simulations. Column 2: The variation between (B) the average observed heterozygosity for the 100 simulations across parameter sets. Column 3: The ratio of the variation between versus within (B/W) simulations.

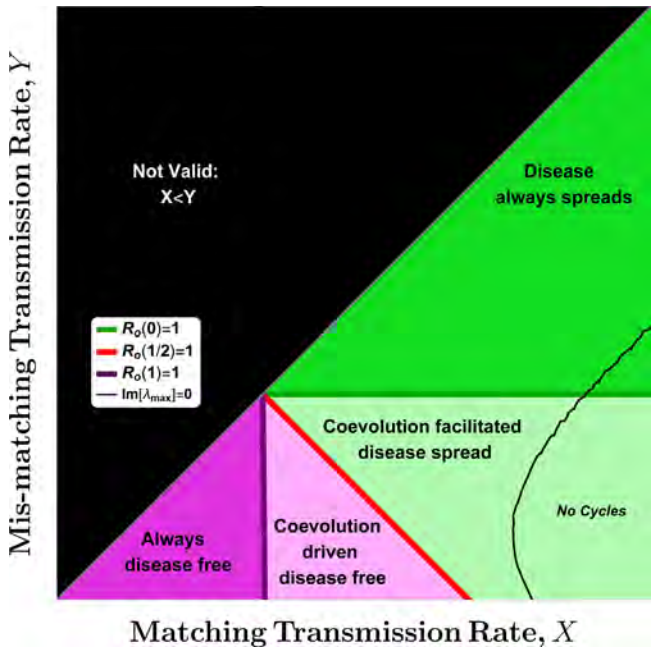


Figure A1. Epidemiological model: deterministic model stability and the role of $R_0(\phi)$. Transient and equilibrium dynamics of the deterministic epidemiological model (system (5)) as a function of $R_0(\phi)$ and the relative matching (X) and mis-matching (Y) transmission rates. Considering only areas where $X > Y$ (i.e., excluding black area), the pathogen is either lost (pink areas) or maintained (green areas) at equilibrium. Dark pink area: $R_0(\phi) < 1$ for all ϕ, disease never spreads. Light pink area: $R_0(\phi) < 1$ except when ϕ is near 1, disease never maintained. Light green area: $R_0(\phi) > 1$ for all ϕ greater than 1/2, polymorphic endemic sustained by coevolution. Dark green area: $R_0(\phi) > 1$ for all ϕ, both pathogen types spread when rare.

allows the invasion of the second pathogen type resulting in a stable endemic equilibrium. Finally, when $R_0(\phi) > 1$ for all ϕ both pathogen types are able to invade initially, and the endemic equilibrium is stable regardless of the coevolutionary dynamics. Although for the majority of parameter space the dynamics of system (5) are characterized by cyclic dynamics, when transmission rates are high and virulence low, the cyclic dynamics dissipate resulting in a rapid and smooth approach to the endemic equilibrium. Given this behavior is rare with parameters we consider only parameters with cyclic dynamics.

Neutral Genetic Drift in the Individual-Based Simulations

Although the single-species Moran model is a good approximation to neutral genetic drift in the ecological and epidemiological model, it is not in fact the exact true neutral expectation for the birth-death processes considered here because of the changes in host population size. For the ecological and epidemiological models considered the Moran model provides a good approximation as the dynamics of host population size were near constant (only slightly changing the geometric mean size, see Appendix A3). Nevertheless to eliminate any biases, rather than using a comparison between the observed heterozygosity and the expectation from the neutral Moran model, unless otherwise noted, we compare the heterozygosity in coevolutionary simulations to the expected heterozygosity from an equivalent neutral simulation. Specifically, for every coevolutionary simulation we ran a complementary neutral simulation with the same sequence of birth and death events except where the genotype of the individual who is born or died was drawn at random, allowing the dynamics of the population to change accordingly. The one exception to this is the few rare cases where hosts went extinct within a replicate, making it impossible to measure host heterozygosity. To eliminate possible bias from excluding these replicates, in the few cases where host extinction occurs prior to generation 250 we set the coevolutionary heterozygosity to 0 and the neutral expectation to that expected under the Moran model.

In contrast to the model presented in MacPherson et al. (2020), where time was normalized with respect to host

Table A2. Linear model fits shown in Figure 2.

K	Constant size			Ecological						Epidemiological		
	$\delta - \gamma$			$\delta - \gamma$			$\ln(-r)$			$-Re[\lambda]$		
	R^2	Slope	CI	R^2	Slope	CI	R^2	Slope	CI	R^2	Slope	CI
100	**0.12**	**0.002**	**0.001**	**0.25**	**0.005**	**0.002**	0	0	0	**0.06**	**−0.048**	**0.039**
178	**0.24**	**0.016**	**0.006**	**0.14**	**0.025**	**0.012**	**0.03**	**0.002**	**0.002**	**0.1**	**0.346**	**0.214**
316	**0.45**	**0.068**	**0.015**	**0.1**	**0.044**	**0.027**	**0.06**	**0.005**	**0.004**	**0.35**	**1.357**	**0.37**
562	**0.58**	**0.127**	**0.022**	**0.06**	**0.038**	**0.03**	**0.12**	**0.008**	**0.004**	**0.37**	**1.893**	**0.499**
1000	**0.69**	**0.131**	**0.018**	0.04	0.019	0.02	**0.21**	**0.007**	**0.003**	**0.28**	**1.795**	**0.58**
1778	**0.7**	**0.086**	**0.011**	0.0	−0.003	0.011	**0.23**	**0.004**	**0.001**	**0.17**	**1.304**	**0.569**
3162	0.72	0.052	0.006	0.01	−0.003	0.006	**0.12**	**0.002**	**0.001**	**0.13**	**0.879**	**0.453**

Note: Fitted slope, associated confidence intervals, and R2 values for linear model fits shown in Figure 2. Slopes that differed significantly from 0 (95% confidence) are given in bold.

generations, the deterministic models above are specified in terms of absolute time. However to facilitate comparisons across parameter space the individual-based simulations are analyzed in terms of host generations where one host generations is defined by the death of K individuals in the simulation. Analytical normalization of the constant-size model in MacPherson et al. (2020) was required for us to compare the ensemble moment dynamics analytically to the expectation under the one-species neutral Moran model. Although analogous ensemble moment dynamics can be obtained for the ecological and epidemiological model, there exists no concise analytical neutral expectation with which to compare them.

Effect of N_e in the Ecological Model

Fluctuations and stochastic variation in population size in the ecological model is expected to decrease effective population size N_e, which, as mentioned above, is the harmonic mean population size over time (Crow 1970). These decreases in effective population size should increase the rate of loss of genetic variation via genetic drift and result in a reduction in the maintenance of genetic variation relative to the expectation in the neutral Moran model. Due to the absence of an exact analytical neutral expectation for the ecological and epidemiological models, in the main text we focused on the comparison between the observed heterozygosity under coevolution and the simulated neutral expectation, $\Delta H = H_{coev} - H_{neut}$. As the simulated neutral expectation is also subject to population fluctuations and stochasticity, we need to instead compare the observed heterozygosity under coevolution to the expectation under the constant-size Moran model $\Delta H_{\text{Moran}} = H_{coev} - H_{\text{Moran}}$ (Moran 1958; Wakeley 2009):

$$H_{\text{Moran}} = \frac{1}{2}\left(1 - \frac{2}{\sum_i \hat{H}_i}\right)^g, \tag{A1}$$

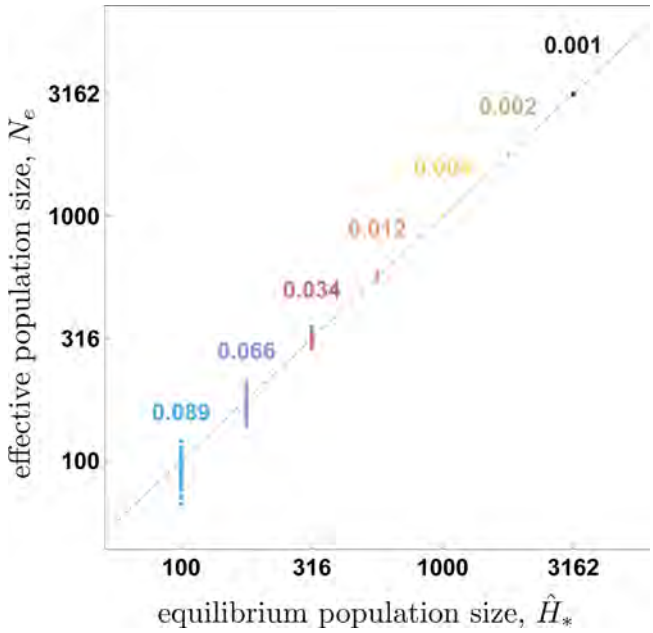
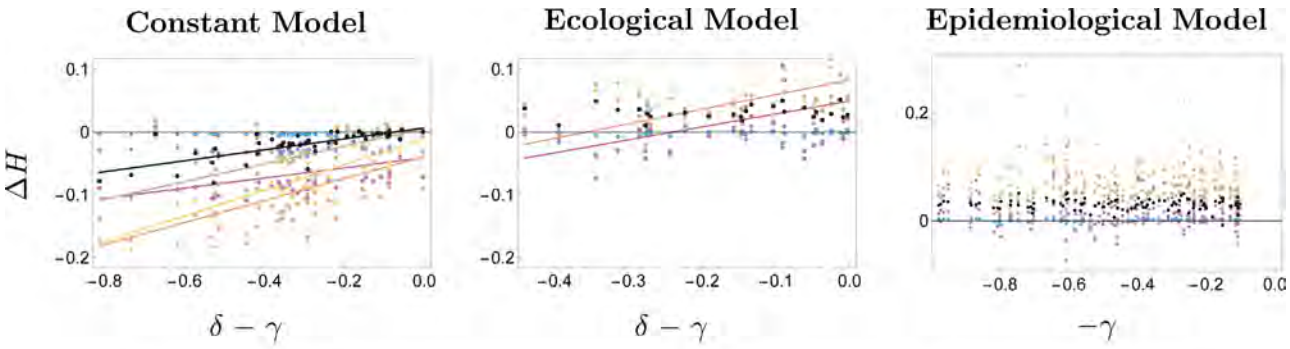


Figure A2. Variation in effective population size in the ecological model. Effective population size, N_e averaged across the 1000 replicate demes for each of the 100 randomly drawn parameter combinations (Row 1, column 2 of Table 1), as a function of equilibrium population size $\sum_i \hat{h}_i$. Numbers above each value of $\sum_i \hat{h}_i$ give the coefficient of variation.

where g is the host generation. We find, however, only a weak relationship between the relative effective population size $\tilde{N}_e = \frac{N_e - \sum_i \hat{H}^i}{\sum_i \hat{H}^i}$ and ΔH_{Moran} as shown in the supplementary *Mathematica* notebook. This is likely the result of low variation in N_e (shown in Fig. A2). As addressed in detail in the section on the deterministic ecological model the dynamics of population size are determined by a pair of nonleading eigenvalues (see supplementary*Mathematica* notebook), which typically had substantially negative real parts, leading to rapid stabilization of the total population size.



K	Constant-size $\delta - \gamma$			Ecological $\delta - \gamma$			Epidemiological $-\gamma$		
	R^2	Slope	CI	R^2	Slope	CI	R^2	Slope	CI
100	0.03	0.001	0.002	0	0	0.013	0	-0.001	0.003
178	0.05	0.015	0.018	0.09	0.0783	0.1	0	-0.003	0.16
316	**0.16**	**0.081**	**0.054**	**0.16**	**0.207**	**0.199**	0.01	-0.016	0.036
562	**0.29**	**0.176**	**0.081**	**0.2**	**0.232**	**0.192**	0	-0.001	0.003
1000	**0.46**	**0.204**	**0.064**	0.12	0.103	0.115	0.01	-0.015	0.026
1778	**0.51**	**0.141**	**0.04**	0.04	0.031	0.063	0.01	-0.006	0.014
3162	**0.57**	**0.087**	**0.022**	0.01	0.009	0.035	0.01	-0.003	0.008

Figure A3. Effect of turnover when constrained such that pathogen turnover greater than that of the host. Plots show ΔH as a function of the relative turnover rates in the host versus the pathogen. Relative turnover is measured as the rate of host minus pathogen turnover which equals $\delta - \gamma$ in the constant-size and ecological models and $-\gamma$ in the epidemiological model. Note that virulent death α affects hosts and pathogens equally and hence is not included. Given that the relative turnover rate in the epidemiological model is always negative ($-\gamma < 0$) we similarly constrain the constant-size and ecological models by considering only the subset of parameter sets for which $\gamma > \delta$. The resulting linear relationships with significantly nonzero slopes are shown. Fitted slope, associated confidence intervals, and R^2 values for linear model fits are given in the table.

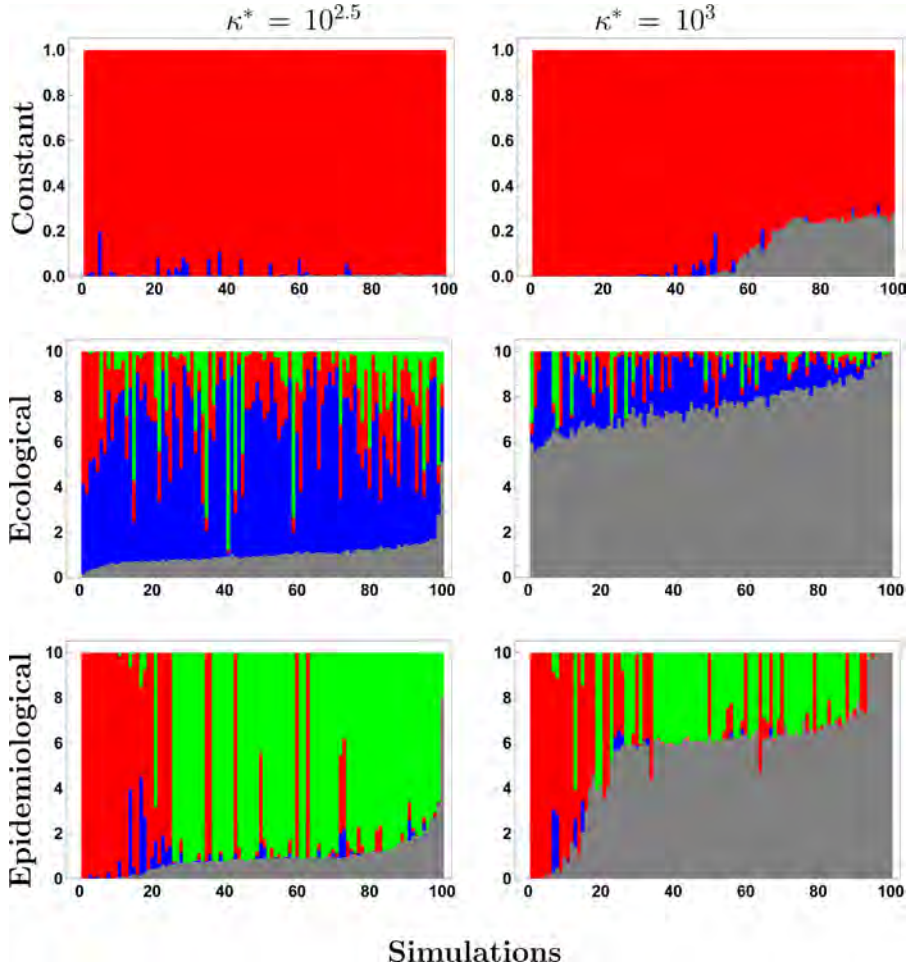


Figure A4. Effect of directional selection across models. Bar charts showing the frequency of directional selection following fixation in the pathogen (red), pathogen extinction (green), fixation in the host (blue), and ongoing coevolution (gray) for each model where parameters are drawn to optimize model comparison (see Table 1). In all panels parameter sets are arranged in order of increasing average ΔH over the 1000 replicates, with each vertical bar giving the proportion of replicates exhibiting each type of behavior at generation 250.