Coevolution does not slow the rate of loss of heterozygosity in a stochastic host-parasite model with constant population size

Ailene MacPherson^{1,a}, Matthew J. Keeling^b, Sarah P. Otto^a

^a*Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, Canada*

^b*Zeeman Institute of Systems Biology and Infectious Disease Research (SBIDER), University of Warwick, Coventry, United Kingdom*

Abstract

Coevolutionary negative frequency dependent selection has been hypothesized to maintain genetic variation in host and parasites. Despite the extensive literature pertaining to host-parasite coevolution, the effect of matching-alleles (MAM) coevolution on the maintenance of genetic variation has not been explicitly modelled in a finite population. The dynamics of the MAM in an infinite population, in fact, suggests that genetic variation in these coevolving populations behaves neutrally. We find that while this is largely true in finite populations two additional phenomena arise. The first of these effects is that of coevolutionary natural selection on stochastic perturbations in host and pathogen allele frequencies. While this may increase or decrease genetic variation, depending on the parameter conditions, the net effect is small relative to that of the second phenomena. Following fixation in the pathogen, the MAM becomes one of directional selection, which in turn rapidly erodes genetic variation in the host. Hence, rather than maintain it, we find that, on average, matching-alleles coevolution depletes genetic variation.

Keywords: coevolution, genetic variation, matching-alleles, negative frequency-dependent selection

1. Introduction

There is a rich history of evolutionary theory exploring the conditions under which genetic variation is maintained or depleted in finite populations. The loss of genetic variation through drift is exacerbated, for example, by fluctuations in population size (Crow, 1970, eq. 7.6.3.34), but variation is maintained through balancing selection in the form of overdominance (Crow, 1970, eq. 8.6.4) or negative frequency-dependent selection (Takahata and Nei, 1990). First suggested by Haldane (1949), one process that is often posited to maintain genetic variation is coevolution between hosts and their parasites. Coevolution, it is argued, should favour pathogens that are best at infecting the most common host genotype. This in turn should favour the spread of rare host genotypes, a form of negative frequency-dependent selection (NFDS) believed to maintain genetic variation (Clarke, 1979). Despite the long-standing interest in coevolution as a mechanism maintaining variation, few studies have explicitly analysed the dynamics of a host and parasite that are both

Email address: amacp@zoology.ubc.ca (Ailene MacPherson)

polymorphic in a fully stochastic framework.

Following Haldane's initial hypotheses, balancing selection as a result of coevolution in the form of NFDS and/or overdominance was suggested as a mechanism behind the extraordinary genetic diversity found at mammalian Major Histocompatibility Complex (MHC) loci (Bodmer, 1972). These same arguments have been used recently to explain the diversity of anti-microbial peptides in *Drosophila* (Unckless et al., 2016; Chapman et al., 2019). MHC loci, and other immune defence genes, are notable not only for their high levels of heterozygosity (> 200 alleles across three loci Klein and Figueroa, 1986; Zimmer and Emlen, 2013) but also for the long-term trans-specific persistence of these polymorphisms (Lawlor et al., 1988; Klein, 1987). Using a coalescent approach Takahata and Nei (1990) found that heterozygote advantage and NFDS are both capable of generating the observed levels of polymorphism. Importantly, however, the model they used was not explicitly coevolutionary but rather based on NFDS within a single species.

As exemplified by Takahata and Nei (1990), much of the literature on the maintenance of genetic variation alludes to yet blurs the distinction between single-species and coevolutionary NFDS (for example see Tellier et al., 2014; Otto and Michalakis, 1998; Zhao and Waxman, 2016; Llaurens et al., 2017; Ejsmond and Radwan, 2015; Rabajante et al., 2016). By the definition of NFDS in a single-species model (called direct-NFDS, Brown and Tellier, 2011), the fitness of an allele increases as its frequency declines, this in turn can favour the spread of rare alleles and the maintenance of genetic variation. In contrast two-species NFDS (called indirect-NFDS), the sort that commonly arises with host-parasite coevolution, favours alleles of the focal species that correspond to ones which are rare in the interacting species. As noted by Brown and Tellier for the gene-for-gene model (2011), we show that indirect-NFDS in matching-alleles model often has little if any impact on the maintenance of genetic variation.

Hints that indirect-NFDS does not maintain genetic variation can be found throughout the theoretical literature on matching-alleles models. Simulating coevolution in a population where genetic variation is repeatedly introduced through migration or mutation, Frank (1991, 1993) found that the dynamics were dominated by the fixation and loss of genetic variants. He attributed this effect to the repeated population bottlenecks that occur from coevolutionary driven fluctuations in population size, but it was unclear whether this same behaviour would arise in a population of constant size. Similarly, although not explicitly discussed, many individual-based models of coevolution include mutation in either one (Agrawal and Lively, 2002) or both species (Lively, 1999; Borghans et al., 2004; Ejsmond and Radwan, 2015) in order to maintain variation and hence coevolution over the long term. Modelling coevolution in an infinite population, M'Gonigle et al. (2009) found that genetic variation is not maintained at equilibrium except when mutation is very frequent. This is echoed in simulations in finite populations, where in the absence of mutation/migration, fixation in either the host or pathogen is very rapid (Gokhale et al., 2013; Schenk et al., 2018). In addition to including mutation, there are several theoretical indications that, rather than being driven by NFDS, the emergent effects of coevolution are dependent on the existence of heterozygote advantage in diploids. For example, M'Gonigle and Otto (2011) showed that the evolution of parasitism depends, not solely on NFDS, but on

whether the interaction induces heterozygote advantage, on average. Similarly, Nuismer and Otto (2004) showed that whether there is, on average, heterozygote advantage is the key determinant of how ploidy levels evolve in both hosts and parasites.

Despite this long and extensive history of verbal and theoretical models, it is unclear whether indirect-NFDS can indeed maintain genetic variation and explain the extensive heterozygosity observed at immune defence loci. Here we compare the maintenance of genetic variation in finite coevolving populations relative to that expected under neutral drift. Previous studies of finite coevolving populations have focused instead on the number of alleles maintained by mutation and coevolution (Borghans et al., 2004), the time until loss of alleles (Gokhale et al., 2013), and the impact of different model assumptions on the time until loss (Schenk et al., 2018). A direct comparison with the neutral expectation is, however, only possible analytically in simple models of host-parasite coevolution. We aim to understand the effect of host-parasite interactions on the maintenance of genetic diversity by examining one such simple single-locus model of coevolution.

2. Theoretical Background

There are two classic models of coevolution involving a single locus major-effect genes in each species, the gene-for-gene (GFG) model, which was motivated by the genetic architecture of flax-rust interactions (Flor, 1956), and the matching-alleles model (MAM), a form of host-parasite specificity that may arise from lock and key molecular interactions (Dybdahl et al., 2014). While GFG and MAM represent only two of the many possible models of host-parasite specificity, their dynamics are representative of the full range that can arise in single-locus models (Agrawal and Lively, 2002; Engelstädter, 2015). In the GFG model, hosts carry either a “susceptible” or “resistant” allele. Susceptible hosts can be infected by both “virulent” and “avirulent” parasite genotypes whereas resistant hosts can only be infected by the virulent parasite genotype. This asymmetry favours fixation of resistant hosts followed by the fixation of virulent pathogens. Deterministic models have shown that this coevolutionary arms race does not in general maintain genetic variation (Leonard, 1977, 1994). Rather stability of genetic polymorphism and the long-term maintenance of genetic variation requires the introduction of temporal delays and asynchrony between host and parasite life cycles (Brown and Tellier, 2011). As GFG is often used to model coevolution in plant-pathogen systems, this asynchrony can be introduced, for example, through seed dormancy or perenniality (Tellier and Brown, 2009).

In contrast to the coevolutionary arms-race dynamics with GFG the symmetry of MAM generates Red Queen allele frequency cycles (also known as trench-warfare dynamics), which are more amenable to the maintenance of genetic variation. MAM is often formulated, and simulated, as an interaction between a host and pathogen with discrete non-overlapping generations (Nuismer, 2017). In each generation hosts are exposed to a single parasite. If host and parasite carry the same, “matching”, allele at the coevolutionary locus then infection occurs with probability X whereas interactions between “mis-matching” host and parasite genotypes occur at a reduced rate Y . Successful infection decreases host fitness by α_H and increases

parasite fitness by α_P . The host and pathogen population sizes are often assumed to remain constant in these models such that hosts and pathogens reproduce proportionally to their fitness creating a subsequent generation of the same size.

In the deterministic (infinite population size) limit, the resulting recursion equations for the MAM are characterized by an unstable cyclic equilibrium about a host and pathogen allele frequency of $1/2$ as shown in Figure 1A (M’Gonigle et al. (2009): haploid host, haploid pathogen, Leonard (1994): diploid host, haploid pathogen). The corresponding dynamics of host heterozygosity is shown in Figure 1B. Starting from a small perturbation near the polymorphic equilibrium, heterozygosity in the host population begins near its maximum of 0.5 and declines, on average, as the allele frequencies cycle outward. As noted above, although NFDS generates Red Queen allele frequency cycles in the short term, these cycles grow in amplitude over time and are not expected to maintain genetic variation in the long term except in the presence of very high mutation rates (M’Gonigle et al., 2009). Indeed, simulating the discrete-time MAM in a finite population, we confirm that host heterozygosity in this model, defined as $H = 2p(1-p)$ where p is the host allele frequency, declines faster than expected under neutral genetic drift (see supplementary *Mathematica* notebook).

In contrast to the unstable Red Queen cycles generated in discrete time, analogous continuous-time models create neutrally stable allele frequency cycles (Woolhouse et al., 2002). Perturbations from the polymorphic equilibrium in these models lead to allele frequency cycles of constant amplitude, often referred to as a “dynamic polymorphism” (Hamilton, 1993). This dynamic polymorphism gives rise to corresponding cycles in heterozygosity also of constant amplitude (Figure 1C). This neutral stability indicates that, when averaged over the cycle, genetic variation remains constant in an infinite population, much like the behaviour of a neutral locus in a single-species model. We thus hypothesize that drift in a finite population may have similar effects in the coevolutionary model as in the neutral single-species model. Our aim here is therefore to develop an appropriate continuous-time MAM of coevolution in a finite population of constant size that, like the standard continuous-time MAM, does not lead to an inherent decline in heterozygosity in the deterministic limit. This model allows us to quantify the loss of heterozygosity with drift in the coevolutionary MAM for comparison to the standard single-species neutral model.

All the models discussed thus far make the traditional assumption that both host and parasite densities are infinite and controlled by factors independent of the host-parasite interaction. This is the case, for example, when host and parasite population sizes are fixed by a hard carrying capacity that is very large. The numbers of hosts and parasites may, however, vary dramatically in response to coevolution (Papkou et al., 2016). In addition parasites that are transmitted directly between hosts may be subject to epidemiological dynamics. While we have reason to believe that SIR dynamics can stabilize allele frequency dynamics (MacPherson et al., 2018), our goal in this work is to focus solely on the effects of NFDS that arises from coevolution. Unlike previous analyses, our goal here is to examine the stochastic nature of coevolutionary dynamics, allowing us to quantify rates of loss of genetic variation. To do so, we focus on the simplest case of strict external control of population size in both hosts and parasites without demographic or epidemiological feedback.

Figure 1: **Deterministic dynamics of matching-alleles coevolution.** Panel A: Phase-plane diagram of the unstable allele frequency cycles of the discrete-time matching-alleles model given by system (5) starting from a small initial perturbation (red point) from the polymorphic equilibrium. Parameters: $X = 0.8, Y = 0.4, \alpha = 0.4, t_{max} = 1000$. Initial conditions: Black $p_H(0) = 0.54, p_P(0) = 0.47$. Panel B: Corresponding temporal dynamics of host heterozygosity. Panel C: Phase plane of neutrally stable allele frequency cycles arising in the continuous-time MAM given by system (3) starting from three different initial conditions shown by the red points. Parameters: $X = 0.8, Y = 0.4, \kappa = 100, \delta = 1, \alpha = 0.2, \gamma = 1, t_{max} = 1000$. Initial conditions: Black/Solid $p_H(0) = 0.55, p_P(0) = 0.45$, Blue/Dashed $p_H(0) = 0.7, p_P(0) = 0.3$, Green/Dotted $p_H(0) = 0.85, p_P(0) = 0.15$. Panel D: Corresponding temporal dynamics of host heterozygosity.

3. The Model

We use a continuous-time birth-death model to describe the coevolutionary dynamics between a host and a free-living pathogen, as depicted in Figure 2. To keep the total host and pathogen population sizes constant at the same fixed value κ , we use a Moran model design with coupled birth-death events. Extension of the model to unequal host and pathogen population sizes is straightforward and the results do not differ qualitatively from those presented here (*pers. comm.*). Both host and pathogen are haploid, with coevolution depending on a single biallelic locus in each species. We represent the number of hosts and pathogens of each type by H_i and P_j where $i \in \{1, 2\}$ and $j \in \{1, 2\}$. Hosts of type i come into contact with pathogens of type j at a density-independent rate $\frac{H_i P_j}{\kappa}$. In keeping with MAM, upon contact the pathogen successfully infects the host with probability $\beta_{i,j}$. If the host and pathogen carry “matching” alleles ($i = j$) then $\beta_{i,j} = X$, whereas “mis-matching” infection occurs with a reduced probability $\beta_{i,j} = Y < X$ for $i \neq j$. If infection occurs, hosts die instantaneously with probability α . The resulting rate of fatal infection of host type i from pathogen type j is $\frac{\alpha \beta_{i,j} H_i P_j}{\kappa}$. Fatal infections result in four coupled events 1) the death of the infected host of type i , 2) birth of a random host, 3) birth of the infecting pathogen of type j representing the release of infectious particles, and 4) death of a random pathogen. Non-fatal infections do not lead to the birth or death of either the host or pathogen, effectively assuming that the pathogen mounts a very limited infection and returns to the free-living pathogen population. In addition to these four events associated with infections, natural host death-birth and free-living pathogen death-birth occur at per-capita rates δ and γ , respectively. Both events are non-selective consisting of the death and birth of random host and pathogen genotypes.

As stated above, our primary aim is to compare the maintenance of genetic variation in a finite coevolving population to that expected under neutral genetic drift alone. Focusing on genetic variation in the host, for a population of finite size κ , the expected decline in host-heterozygosity from neutral genetic drift in the haploid Moran model is given by:

$$H_{neut}(t) = H_0 e^{-\frac{2t}{\kappa}} \quad (1)$$

where H_0 is the initial heterozygosity and time, t , is measured in units of host generations, (i.e. the expected time to κ deaths, $\kappa/(\kappa d) = 1/d$ where d is the death rate). To compare the neutral single-species model to a host-parasite model, we need to define time in the host-parasite model in an equivalent way. In the coevolutionary model, the per capita host death rate, D , hence the host generation time, changes over time as the allele frequencies change.

$$D = \frac{\sum_{i,j} \beta_{i,j} \alpha H_i P_j + \delta \kappa}{\kappa} \quad (2)$$

To measure time in terms of host generations we therefore rescale absolute time τ such that $t = D\tau$ where t is again in units of host generations.

3.1. Deterministic dynamics

In the limit as the total host and pathogen population size goes to infinity ($\kappa \rightarrow \infty$) the dynamics of the host and pathogen allele frequencies $p_H = H_1/\kappa$ ($1 - p_H = H_2/\kappa$) and $p_P = P_1/\kappa$ ($1 - p_P = P_2/\kappa$) are given

Figure 2: **Schematic of the Moran MAM.** Coevolution consisting of three different types of coupled birth-death events. Infection (green/thick), natural host death (blue/thin), and free-living pathogen death (purple/standard). Solid lines denote events that occur at a rate that depends on the genotype whereas dashed arrows represent birth and death of random individuals irrespective of their genotype.

by the following system of differential equations:

$$\begin{aligned}\frac{dp_H(t)}{dt} &= \frac{(X - Y)\alpha(1 - 2p_P(t))p_H(t)(1 - p_H(t))}{X\alpha + \delta - (X - Y)\alpha(p_P(t) + p_H(t)(1 - 2p_P(t)))} \\ \frac{dp_P(t)}{dt} &= \frac{(X - Y)\alpha(1 - 2p_H(t))p_P(t)(1 - p_P(t))}{X\alpha + \delta - (X - Y)\alpha(p_P(t) + p_H(t)(1 - 2p_P(t)))}\end{aligned}\quad (3)$$

As shown in Figure 1, the behaviour of this Moran MAM in the deterministic limit is identical to that of the traditional continuous-time MAM (Woolhouse et al., 2002). Specifically, there are five equilibria of system (3). Four are unstable equilibria characterized by the fixation of one host and one pathogen genotype. The final polymorphic equilibria at $\hat{p}_H = \hat{p}_P = \frac{1}{2}$ is neutrally stable with purely imaginary leading eigenvalues, generating neutral-limit cycles in allele frequencies (see Figure 1C).

As discussed previously, the dynamic polymorphism that arises in the deterministic continuous-time MAM neither maintains nor depletes host heterozygosity (see Figure 1D). While the neutral stability of the MAM is well known, the consequences on heterozygosity are under-appreciated. Contrary to the proposed effect of coevolution and NFDS on genetic diversity (Haldane, 1949; Clarke, 1979), neutrally stable allele frequency cycles neither deplete nor restore genetic variation. Rather, in an infinite population, coevolution has no net effect on genetic diversity, averaged across a cycle. To see if this behaviour holds in finite coevolving populations, we begin by describing an analytical approach to compare the dynamics of genetic diversity in a finite population to those expected under neutral genetic drift. Because this analytical approach only applies while the cycles are small we couple this approach with an individual-based simulation describing the loss of genetic diversity more generally.

3.2. Ensemble moment dynamics

To model the dynamics of genetic diversity in a finite population we express the model depicted in Figure 2 as a continuous-time Markov chain. The complexity of the model limits the available analytical approaches. One approach we can use, however, is the ensemble moment approximation, the derivation of which is given in the Appendix. The result is an approximation for the expected host heterozygosity which we will call $\tilde{H}(t)$. As with the neutral expectation in equation (1), this is the expectation of host heterozygosity over all realizations of the stochastic process, the approximation of which is given by a Taylor series assuming the system is near the deterministic equilibrium ($H_0 \approx 1/2$) and the population size large.

$$\tilde{H}(t) \approx H_0 \left(1 - \frac{2t}{\kappa} \right) + \mathcal{O}(\epsilon^3) \quad (4)$$

where ϵ is the deviation of the stochastic process from the deterministic equilibrium and $\epsilon = 1/\kappa$. Once again t is measured in units of host generations. To this order, the ensemble moment dynamics are identical to that of the neutral expectation given by (1). Hence, as expected from the deterministic dynamics, host heterozygosity in the MAM declines as expected under neutral genetic drift. Although analytically unwieldy, numerical evaluation of $\tilde{H}$ to third order reveals slight deviations between the ensemble moment dynamics and the neutral expectation that are also observed in the individual-based simulations presented in the

following section. Specifically, $\tilde{H}$ exceeds the neutral expectation when the free-living pathogen death rate γ is less than the death rate of the host δ (negative values of $\gamma - \delta$ in Figure 3B) and falls below the neutral expectation as this rate increases.

The mechanism behind these lower-order deviations is shown schematically in Figure 3A. If we start, for example, at the equilibrium host and pathogen allele frequencies ($\hat{p}_H = 0.5$, $\hat{p}_P = 0.5$), infection, natural host death-birth, and free-living pathogen death-birth events perturb the allele frequencies. Selection then acts on those perturbations to produce small counter-clock wise allele frequency cycles. Although over the long term these allele frequency cycles have no net effect on heterozygosity, they do have an effect over the short term. For example, if only the host allele frequency is perturbed by a random host birth event then during the first quarter cycle natural selection will increase host heterozygosity, because the parasite that matches the host that, by chance, became more common, will spread in turn reducing the more common host genotype. In contrast, if only the pathogen allele frequency is perturbed, selection will decrease host heterozygosity by preferentially killing off the host that matches the parasite allele that increased by drift.

The effect of these early quarter cycle responses to selection would average out if perturbations were distributed uniformly about the equilibrium. However, natural host death affects only the host allele frequency whereas free-living pathogen death only affects the pathogen allele frequency. Hence the relative rates of these two events determine whether these early responses to selection transiently increases or decreases heterozygosity. Specifically, relatively high rates of natural host death and low rates of free-living pathogen death ($\gamma - \delta < 0$) lead to an increase in host heterozygosity relative to the neutral expectation (see Figure 3B).

3.3. Individual-based simulations

Using a Gillespie algorithm we simulated host-parasite coevolution in a finite population. For each of one-hundred randomly drawn parameter combinations, we simulated coevolution for seven different populations sizes ranging on a log scale from $\kappa = 10^2$ to $\kappa = 10^{3.5} = 3162$. The relative infection rates of matching and mis-matching genotypes were drawn such that $0 \leq Y < X \leq 1$. We restricted the probability of dying from infection, α , to lie between 0 and 1/2 as much stronger selection causes the rapid loss of alleles, limiting the amount of coevolution. Host and pathogen free-living natural death rates δ and γ were both drawn between 0.1 and 1. For each of the 700 parameter combinations, we simulated 1000 replicate populations for $t_{max} = 500$ host generations.

In tandem with each coevolutionary replicate population we simulated an analogous “neutral population” to check that it behaved correctly according to the neutral expectation given in equation (1). In these simulations, the exact same series of host and pathogen birth-death events occurred as in the coevolutionary simulations, except that the individual who was born or died due to infection was drawn at random, these simulations captured random genetic drift in both host and pathogen. As shown in Figure 4B the mean heterozygosity of these neutral simulations (shown in gray) are identical to the analytical neutral expectation given by equation (1), confirming that this is the appropriate neutral expectation for the two-species model.

Figure 3: **The effect of early perturbations.** Shown schematically in the left panel are the effects of the three processes, infection, natural host death, and natural pathogen death on host and pathogen allele frequency and the resulting effect of selection on host heterozygosity. $\tilde{H} - H_{neut}$ as a function of the relative host and pathogen death rates is shown in the right panel. Points give the deviation at time $t = 50$ host generations for 200 randomly drawn combinations of γ and δ , where $0.1 < \gamma < 1$, $0.1 < \delta < 1$, and $X = 0.9, Y = 0.7, \alpha = 0.2, \kappa = 10^5$.

Figure 4 captures the four key results of our model. First, as predicted by the second-order ensemble moment approximation given in equation (4) and shown in the inset of panel B the dynamics of heterozygosity in the coevolving populations behave neutrally, for approximately the first 10 generations, when the dynamics are very near the deterministic equilibrium. Second, as the distance from the deterministic equilibrium grows, the coevolutionary dynamics depart slightly from the neutral expectation due to the transient effects of selection on perturbations. The resulting slight increase (decrease) in heterozygosity is shown by the filled points in panel A that are either slightly greater (less) than the neutral expectation. The third key result is that heterozygosity is sometimes substantially less than expected from the neutral expectation. Following fixation, the matching-alleles genetic mechanism becomes one of directional selection, favouring the mismatching host of the remaining pathogen genotype. This directional selection (shown by red trajectories in Figure 4C) in turn rapidly erodes genetic variation within the host population. Finally, as predicted by the deterministic model, host heterozygosity in most cases is near the neutral expectation. Specifically, open points in Figure 4A) indicate parameter conditions that did not differ significantly from the neutral expectation.

4. Discussion

Contrary to theories that posit that host-parasite coevolution and the associated negative frequency-dependent selection (NFDS) should maintain genetic variation (Haldane, 1949; Clarke, 1979; Takahata and Nei, 1990), we use stochastic methods, both describing ensemble moments and conducting individual-based

Figure 4: Individual-based simulations. Panel A: Heterozygosity in the coevolving populations versus the neutral expectation at time $t = 250$ host generations. Black dashed line represents the expectation under neutrality. The red curve is the smoothed fit of the mean, light red lines are bootstrap means. The points are the mean heterozygosities for each of the 420 parameter combinations averaged across the 1000 replicate populations simulated for each. Points are filled if the mean heterozygosity for the parameter combination differs significantly from the neutral expectation (with 95% confidence with Bonferroni correction for the 100 points for each κ) versus empty when the replicate population means overlap neutrality. Panel B: Example dynamics for a single parameter condition ($X = 0.29, Y = 0.14, \alpha = 0.38, \delta = 0.99, \gamma = 0.24$, and $\kappa = 10^{2.75} = 562$). Red region represents CI of heterozygosity in the coevolving population, gray region is the CI in the analogous neutral simulation, Blue line is the neutral expectation from equation (1) and green line is heterozygosity in the second order ensemble moment approximation. Panel C: Allele frequency dynamics for example simulation. Gray trajectories represent ongoing coevolution, red are periods of directional selection, while blue are fixed host replicate populations. Parameters: $X = 0.26, Y = 0.25, \alpha = 0.35, \delta = 0.36, \gamma = 0.83, \kappa = 10^{3.5} = 1778$

simulations, to show that genetic variation is lost at nearly the same rate as in the neutral model (continuous time) or faster (discrete time). Although these results are consistent with stability analyses of the corresponding deterministic models, such stability analyses only apply near the equilibrium with equal allele frequencies. Thus, in contrast to single-species NFDS, coevolutionary NFDS does not generally maintain genetic variation. Specifically, the neutrality of the cycles ensures that for any given host allele frequency selection is as likely to be selecting for as against that allele. While the largely neutral effect of matching-alleles coevolution is recovered in our models of finite-population size, we find that stochasticity introduces two new effects. First is the effect of the initial response to selection following perturbations in allele frequency from the deterministic equilibrium, which may on average either maintain or deplete heterozygosity depending on the relative death rates of the host versus the pathogen. The second non-neutral effect is that of directional selection following pathogen fixation which rapidly erodes genetic variation.

The contrasting effects of coevolutionary indirect-NFDS and single-species direct-NFDS have been pointed out previously by Brown and Tellier (2011) in the context of the gene-for-gene coevolutionary model. Nevertheless, the distinction between these two processes and in particular their differing effects on the maintenance of genetic variation remains under-appreciated. Our finding that the MAM does not inherently maintain genetic variation contrasts with and clarifies the findings of previous computational models of the maintenance of genetic variation at MHC loci (Ejsmond and Radwan, 2015; Borghans et al., 2004). Simulating matching-alleles coevolution at multiple loci in the presence of rapid host mutation, both Ejsmond and Radwan (2015) and Borghans et al. (2004) find a weak signal of increased fitness of rare host alleles. Our results clarify, however, that this is not in fact a true signal of negative frequency-dependent selection. Rather this is a result of the fact that novel, and hence rare, host alleles may be favoured transiently if there does not yet exist genetic variation at the corresponding pathogen locus. In contrast, long term changes in host allele frequencies are not associated with corresponding changes in fitness as would be required under NFDS (see Borghans et al. (2004) Fig 4A).

Here we have focused on coevolution in a single population. Coevolution is however inherently a spatial process. Previous theoretical models suggest that spatial structure in combination with variability in the strength and nature of host-parasite coevolution across space may promote local adaptation (Gandon et al., 1996; Nuismer, 2006). Our results illustrate, however, that this is not the result of coevolution within each individual deme but rather an emergent effect of spatial structure and gene flow. Whether these spatial models maintain more genetic variation than expected under neutral drift remains a valuable topic for future work.

In the main text we have focused on host-parasite coevolution in a continuous-time model as opposed to a model of coevolution in discrete time (see equation A9). This is because, as explored in the background section, the continuous-time model is more likely to maintain genetic variation based on the behaviour of the deterministic dynamics. We confirmed that this expectation holds when extended to finite populations by simulating host-parasite coevolution in a Wright-Fisher MAM (see Appendix). As expected host-heterozygosity in such a discrete-time MAM is significantly less than the neutral expectation and less than observed in the continuous-time model (see Figure 5).

The largely neutral effects of the continuous-time MAM and variation eroding effects in discrete-time makes coevolutionary NFDS an insufficient explanation for the high levels of genetic diversity at host immunodefence loci. The models presented here make several implicit assumptions that may influence the maintenance of genetic diversity. In particular, we assume that the population size remains constant. Although this is a classic assumption in coevolutionary theory, coevolution is known to have drastic effects on population sizes of both host and pathogen in natural systems (Papkou et al., 2016). Indeed inclusion of explicit population size dynamics in the MAM results in fluctuations in host and pathogen population size (Nuismer, 2017; Frank, 1993). While these changes in population size are expected to have little to no effect on the allele frequency dynamics, they are expected to have important consequences on the maintenance of

genetic variation (Crow, 1970). We also assume that the host and parasite generations are synchronized. The extensive literature on the deterministic stability of the polymorphic equilibrium has suggested that temporal asynchrony between the host and parasite, arising for example through seed dormancy, can stabilize polymorphisms and maintain variation (Brown and Tellier, 2011; Tellier and Brown, 2009). Developing
275 an analogous MAM model, we find that here too including a temporal delay in the ability of one species to respond to the other (here by including seed dormancy) has a stabilizing effect and can help maintain variation (see Appendix).

The maintenance of genetic diversity at human MHC loci may also be the result of coevolution between hosts and pathogens that are transmitted directly from one host to another and hence subject to epidemi-
280 ological dynamics. Our previous work showed that these epidemiological dynamics stabilizes Red Queen allele frequency cycles resulting in a stable polymorphic equilibrium (MacPherson et al., 2018). In contrast to the instability of the discrete-time MAM and the neutral stability of the continuous-time MAM, the stabilizing effects of these epidemiological dynamics are expected to maintain genetic variation. While the work presented here was aimed solely at understanding the effects of coevolution on the maintenance of
285 genetic variation in finite populations, extending these results to include more realistic population size and epidemiological dynamics is necessary to fully understand the maintenance of genetic variation in natural systems.

Acknowledgements

We thank Matt Pennell, Jonathan Davies, Jeff Joy, Dan Coombs, and Aurélien Tellier for their many
290 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).

Appendix

A1. The Ensemble Moment Approximation

A1.1. General Approach

We consider the general stochastic process in Z^{+n} . Specifically, for the MAM model presented here $n = 2$, one dimension for the number of hosts of type 1 and pathogens of type 1, respectively. Let $\vec{z}(t)$ denote the state of the process at time t , for our case $\vec{z} = \{H_1, P_1\}$. We begin by transforming the variables into their deviation from the deterministic equilibrium. Specifically if $\hat{z}_i$ is the value of variable i at equilibrium, then the transformed variables are given by $\vec{x} = \vec{z} - \vec{\hat{z}}$. In addition, we only consider the case where the initial condition of the stochastic processes is the deterministic equilibrium, $\vec{x} = \vec{0}$. This transformation in variables will then allow us to easily express the moments of the stochastic process in terms of Taylor series around the deterministic equilibrium. In terms of the transformed variables, the stochastic process is described by m distinct events that occur at rates $f_e(\vec{x})$ where $e = \{1, 2, \dots, m\}$. Finally, we denote the ensemble average, the average across all sample paths, by $\langle * \rangle$.

The aim of the ensemble moment approximation is to derive a system of ODEs giving the change in the moments of the stochastic process. Beginning with the first moment (mean) we have:

$$\begin{aligned} \frac{d\langle x_i \rangle}{dt} &= \left\langle \sum_{e=1}^m (\text{event rates}) (\text{change in } x_i \text{ from event}) \right\rangle \\ &= \left\langle \sum_{e=1}^m f_e(\vec{x}) ((x_i + \Delta_{e,i}) - x_i) \right\rangle \end{aligned} \quad (\text{A1})$$

where $\Delta_{e,i}$ is the change in x_i from event e (Keeling, 2000). To take the ensemble average of the sum we express the argument as a polynomial by approximating it with a Taylor series around the deterministic equilibrium assuming $x_i \forall i$ is small. Equation (A1) is then approximated as:

$$\frac{d\langle x_i \rangle}{dt} \approx \left\langle \sum_j a_j x_j + \sum_{j,k}^{m,m} b_{j,k} x_j x_k + \sum_{j,k,l}^{m,m,m} c_{j,k,l} x_j x_k x_l + \dots \right\rangle \quad (\text{A2})$$

where a_j , $b_{j,k}$, and $c_{j,k,l}$ are constants, which are themselves functions of the deterministic equilibrium $\vec{\hat{z}}$. Rearranging we have:

$$\frac{d\langle x_i \rangle}{dt} \approx \sum_j a_j \langle x_j \rangle + \sum_{j,k}^{m,m} b_{j,k} \langle x_j x_k \rangle + \sum_{j,k,l}^{m,m,m} c_{j,k,l} \langle x_j x_k x_l \rangle + \dots \quad (\text{A3})$$

The result is an expression for the change in the first moment as a linear function of the first, second, third, and higher moments.

We use the same technique to derive the second moment ODEs:

$$\begin{aligned}\frac{d\langle x_i x_j \rangle}{dt} &= \left\langle \sum_e f_e(\vec{x}) ((x_i + \Delta_{e,i})(x_j + \Delta_{e,j}) - x_i x_j) \right\rangle \\ &= \left\langle \sum_e f_e(\vec{x}) (x_i \Delta_{e,j} + x_j \Delta_{e,i} + \Delta_{e,i} \Delta_{e,j}) \right\rangle\end{aligned}\quad (\text{A4})$$

Approximating the right-hand argument as a sum once again by a Taylor series expansion we have:

$$\frac{d\langle x_i x_j \rangle}{dt} \approx \sum_j A_j \langle x_j \rangle + \sum_{j,k}^{m,m} B_{j,k} \langle x_j x_k \rangle + \sum_{j,k,l}^{m,m,m} C_{j,k,l} \langle x_j x_k x_l \rangle + \dots \quad (\text{A5})$$

Similarly for the third moments:

$$\frac{d\langle x_i x_j x_k \rangle}{dt} \approx \sum_j \mathcal{A}_j \langle x_j \rangle + \sum_{j,k}^{m,m} \mathcal{B}_{j,k} \langle x_j x_k \rangle + \sum_{j,k,l}^{m,m,m} \mathcal{C}_{j,k,l} \langle x_j x_k x_l \rangle + \dots \quad (\text{A6})$$

where $A, B, C, \mathcal{A}, \mathcal{B}$, and $\mathcal{C}$ are all constants that depend on the deterministic equilibrium.

310 A1.2. Dynamics of host heterozygosity

We are interested in the dynamics of host heterozygosity and calculate the ensemble moment approximation for the dynamics of heterozygosity, $\tilde{H}$:

$$\begin{aligned}\frac{d\tilde{H}}{dt} &= \frac{d}{dt} (E[2p_H(1 - p_H)]) = 2 \frac{d}{dt} E[p_H - p_H^2] \\ &= 2 \left(\frac{d}{dt} \frac{\langle H_1 \rangle}{\kappa} - \frac{d}{dt} \frac{\langle H_1^2 \rangle}{\kappa^2} \right)\end{aligned}\quad (\text{A7})$$

The dynamics of heterozygosity thus depends directly on that of the first two moments. If we substitute in equation (A3) and (A5) and assume that population size is large $\kappa = 1/\epsilon$, then approximated to $\mathcal{O}(\epsilon^2)$ the dynamics of host heterozygosity simplify to equation (4) given in the main text.

A2. Maintenance of Genetic Variation in Discrete Time

315 A2.1. Model specification

We model host-parasite coevolution in discrete time with a Wright-Fisher model with a constant host and pathogen population size κ . In each generation, hosts come into contact with a single random pathogen. If host and pathogen have the same genotype i , the probability of successful infection is given by $\beta_{i,i} = X$, whereas hosts of type i are infected by pathogens of mis-matching genotype j , with probability $\beta_{i,j} = Y < X$. In the absence of infections host's and pathogen's have a fitness 1, whereas infection decreases host fitness by a factor α_H and increases pathogen fitness by α_P such that the expected fitness of host genotype i and pathogen genotype j are given by:

$$\begin{aligned}W_H(i, t) &= 1 - \alpha_H \sum_j \beta_{i,j} \frac{P_j}{\kappa} \\ W_P(j, t) &= 1 + \alpha_P \sum_i \beta_{j,i} \frac{H_i}{\kappa}\end{aligned}\quad (\text{A8})$$

A2.2. Deterministic dynamics

As we do for the continuous-time model, we begin by developing an understanding of the effect of coevolution on genetic variation in a finite population by analysing the dynamics of heterozygosity in the limit as the population size goes to infinity. In this deterministic case the coevolutionary dynamics are given by the following difference equations:

$$\begin{aligned} p_H(t+1) &= p_H + \frac{\alpha_H p_H (1 - p_H) (1 - 2p_P) (X - Y)}{\alpha_H (p_H (1 - 2p_P) + p_P) (X - Y) - \alpha_H X + 1} \\ p_P(t+1) &= p_P - \frac{\alpha_P p_P (1 - p_P) (1 - 2p_H) (X - Y)}{-\alpha_P (p_H (1 - 2p_P) + p_P) (X - Y) + \alpha_P X + 1} \end{aligned} \quad (\text{A9})$$

There are five equilibria of this model, four of these specify the fixation/loss of one host and one pathogen type, and a fifth internal equilibrium occurs at $p_H = p_P = 1/2$. Unlike the continuous time model explored in the main text this internal equilibrium is unstable and cyclic. The dynamics of the allele frequency and heterozygosity are shown in Figure 1 panels A and B, respectively. The rate at which the amplitude of the cycles grow is given by the magnitude of the leading eigenvalue, which is given by $\|\lambda\|$, a result consistent with those of (M’Gonigle et al., 2009).

$$\|\lambda\| = \sqrt{1 + \frac{\alpha_H \alpha_P (X - Y)^2}{(2 - (X + Y) \alpha_H) (2 + (X + Y) \alpha_P)}} \quad (\text{A10})$$

Due to the instability of the internal equilibrium host heterozygosity decays over time. Furthermore the larger the value of $\|\lambda\|$ the faster the heterozygosity should decay.

A2.3. Simulations of a finite population

In analogy to the continuous-time model presented in the main-text where the death of the host and birth of the pathogen from infection are coupled, for our simulations in a finite population we assume $\alpha_H = \alpha_P = \alpha$. As with the continuous-time model we simulate coevolution for 100 random parameter combinations of X, Y , and α ($0 < \alpha < 0.5$, $0 < Y < X < 1$) and 7 values of κ (ranging on a log scale between 10^2 and 10^5) for a total of 700 simulations. In contrast to the continuous-time model natural host death and free-living pathogen birth are incorporated directly in the Wright-Fisher model. For each of the simulations, sample paths were generated for 1000 independent replicate populations. All simulations are initialized at the internal polymorphic equilibrium and run for $t = 500$ generations.

In the absence of coevolution, the dynamics of host heterozygosity are given by the neutral expectation for the Wright-Fisher model:

$$H_{neut} = H_0 \left(1 - \frac{1}{\kappa}\right)^t \quad (\text{A11})$$

The difference between the simulated heterozygosity and the neutral expectation at time $t = 500$ is shown in Figure 5A. When population sizes is large ($\kappa = 10^5$ upper right hand corner) the deviation in heterozygosity between the coevolutionary simulations and neutral expectation, $\Delta H = H_{coev} - H_{neut}$, is statistically negative as expected due to the erosion of genetic variation by coevolution. The magnitude of this deviation is

however small, so while coevolution does indeed erode genetic variation the process is slow with only small effects over the parameter and time scale sampled.

As with the continuous-time model presented in the main text, directional selection (see Figure 5C) following fixation in the pathogen has a much larger effect than that of coevolutionary selection. When the population size is small (e.g. $\kappa = 10^2$), drift dominates over the effects of selection, and the heterozygosity behaves nearly neutrally. As the population size increases, the effect of this directional selection emerges. As predicted by the deterministic model variation in ΔH for a given population size is explained largely by variation in $\|\lambda\|$ (see Figure 5B). This is not only because $\|\lambda\|$ determines the effect of coevolution but because it is highly correlated with the probability of pathogen fixation and the strength of directional selection following pathogen fixation.

A3. Seed Dormancy

Here we extend the above discrete-time model to include the effects of seed dormancy, or any other process that causes a subset, s , of the host population to avoid parasite induced selection, making the host less responsive during coevolution. We analyse the effect of the resulting temporal asynchrony between host and parasite on the stability of the polymorphic equilibrium. Given a life cycle shown in Figure 6 we model evolution by following the allele frequencies among the host seedlings p_H , host juvenile seeds p_J , and within the parasite p_P . The dynamics are given by:

$$\begin{aligned} p_H(t+1) &= sp_J(t) + \frac{(1-s)p_H(t)(1-\alpha_H p_P(t)(X-Y) - \alpha_H Y)}{\alpha_H p_H(t)(1-2p_P(t))(X-Y) + \alpha_H p_P(t)(X-Y) - \alpha_H X + 1} \\ p_J(t+1) &= (1-s)p_J(t) + \frac{sp_H(t)(1-\alpha_H p_P(t)(X-Y) - \alpha_H Y)}{\alpha_H p_H(t)(1-2p_P(t))(X-Y) + \alpha_H p_P(t)(X-Y) - \alpha_H X + 1} \\ p_P(t) &= \frac{p_P(t)(\alpha_P p_H(t)(X-Y) + \alpha_P Y + 1)}{1 - \alpha_P p_H(t)(1-2p_P(t))(X-Y) - \alpha_P p_P(t)(X-Y) + \alpha_P X} \end{aligned} \quad (\text{A12})$$

There are five equilibria of system A12, four of which are trivial fixation/loss of the host and pathogen. The fifth internal equilibrium is at $\hat{p}_H, \hat{p}_J, \hat{p}_P = \frac{1}{2}$. The Jacobian matrix at this internal equilibrium is given by:

$$J = \begin{bmatrix} 1-s & s & \frac{s(X-Y)\alpha_H}{\alpha_H(X+Y)-2} \\ s & 1-s & \frac{(1-s)(X-Y)\alpha_H}{\alpha_H(X+Y)-2} \\ 0 & \frac{s(X-Y)\alpha_P}{\alpha_P(X+Y)-2} & 1 \end{bmatrix} \quad (\text{A13})$$

Focusing on the diagonal elements, we note that $\frac{dp_H(t+1)}{dp_H(t)} < 1$ and $\frac{dp_J(t+1)}{dp_J(t)} < 1$. As noted by Tellier and Brown (2009) this is indicative of direct negative frequency-dependent selection. We can evaluate the stability of system A12 by using the Routh-Hurwitz conditions on the transformed characteristic polynomial of A13. Specifically, although the Routh-Hurwitz conditions can only be applied to identify the stability criteria in continuous time, if we transform the characteristic polynomial, $p(\lambda)$, into a third order polynomial in z such that:

$$(z-1)^3 p\left(\frac{z+1}{z-1}\right) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 \quad (\text{A14})$$

then the equilibrium is stable (all eigenvalues have a magnitude less than 1) if and only if $a_i < 0 \forall i$ and $a_2 a_1 - a_3 a_0 > 0$ assuming $\alpha_H = \alpha_P = \alpha$. These criteria are satisfied in our model for a range of s , indicating the the internal equilibrium is stable under these condition:

$$\frac{\alpha^2(3X - Y)(X - 3Y) + 4 - B}{\alpha^2 XY - 1} < s < \frac{\alpha^2(3X - Y)(X - 3Y) + 4 + B}{\alpha^2 XY - 1} \quad (\text{A15})$$

$$B = \sqrt{-(\alpha(X + Y) + 2)(2 - \alpha(X + Y))(\alpha^2(9X^2 - 14XY + 9Y^2) - 4)}$$

where B must be real.

A4. Supplementary Figures

Figure 5: **Comparison between discrete- and continuous-time models.** Panel A: Mean heterozygosity in a coevolving population (at $t = 250$ generations) versus the neutral expectation for the Moran model in continuous time (red as in- see Figure 4) versus the Wright-Fisher model in discrete time (green), light green lines give smoothed fits to 1000 bootstrap samples. Points depict heterozygosity for each of the 420 parameter conditions simulated for the WF model averaged over the 1000 replicate populations. Panel B: Deviation in heterozygosity from the neutral expectation in the WF as a function of the magnitude of the leading eigenvalue. R^2 values are given for the non-linear model fit of $a + e^{-b(x-1)}$. Panel C: Example allele frequency dynamics in the WF model. Allele frequency trajectories are coloured gray when there is ongoing coevolution, red when there is directional selection, and blue when the host population is fixed. Parameters: $X = 0.9, Y = 0.26, \alpha = 0.047, \kappa = 10^{2.5} = 316$

Figure 6: **Life-cycle diagram and numerical stability analysis of MAM with seed dormancy.** Host and pathogen life-cycle (blue arrows), flux in and out of seed bank (green arrows). Shaded region of right-hand plot gives region for which the polymorphic equilibrium is stable assuming $\alpha_H = \alpha_P$ and given $X = 1, Y = 0$.

References

- Agrawal, A., Lively, C.M., 2002. Infection genetics: Gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research* 4, 79–90.
- Bodmer, W.F., 1972. Evolutionary Significance of the HL-A System. *Nature* 237, 139–146.
- Borghans, J., Beltman, J., De Boer, R., 2004. MHC polymorphism under host-pathogen coevolution. *Immunogenetics* .
- Brown, J.K.M., Tellier, A., 2011. Plant-Parasite Coevolution: Bridging the Gap between Genetics and Ecology. *Annu. Rev. of Phytopathology* 49, 345–367. doi:10.1146/annurev-phyto-072910-095301.
- Chapman, J.R., Hill, T., Unckless, R.L., 2019. Balancing Selection Drives the Maintenance of Genetic Variation in *Drosophila* Antimicrobial Peptides. *Genome Biol. Evol.* 11, 2691–2701. doi:10.1093/GBE/EVZ191.
- Clarke, B., 1979. The evolution of genetic diversity. *Proc. R. Soc. Lond. B* , 19–40.
- Crow, J.F., 1970. Genetic Loads and the Cost of Natural Selection, in: Kojima, K.i. (Ed.), *Mathematical Topics in Population Genetics*. Springer, Berlin, Heidelberg. Biomathematics, pp. 128–177.
- Dybdahl, M.F., Jenkins, C.E., Nuismer, S.L., 2014. Identifying the molecular basis of host-parasite coevolution: Merging models and mechanisms. *Am. Nat.* 184, 1–13. doi:10.1086/676591.

Ejzmond, M., Radwan, J., 2015. Red Queen Processes Drive Positive Selection on Major Histocompatibility Complex (MHC) Genes. *PLOS Comp. Biol.* .

Engelstädter, J., 2015. Host-Parasite Coevolutionary Dynamics with Generalized Success/Failure Infection Genetics. *The American Naturalist* 185, E117–E129. URL: <https://www.journals.uchicago.edu/doi/10.1086/680476>, doi:10.1086/680476.

Flor, H., 1956. The Complementary Genic Systems in Flax and Flax Rust. *Advances in Genetics* 8, 29–54. doi:10.1016/S0065-2660(08)60498-8.

Frank, S.A., 1991. Ecological and genetic models of host-pathogen coevolution. *Heredity* , 73–83.

Frank, S.A., 1993. Specificity versus detectable polymorphism in host-parasite genetics. *Proc. R. Soc. Lond. B* 254, 191–197. doi:10.1098/rspb.1993.0145.

Gandon, S., Capowiez, Y., Dubois, Y., Michalakis, Y., Olivieri, I., 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proc. R. Soc. Lond. B* 263, 1003–1009. doi:10.1098/rspb.1996.0148.

Gokhale, C.S., Papkou, A., Traulsen, A., Schulenburg, H., 2013. Lotka-Volterra dynamics kills the Red Queen: population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. *BMC Evol. Bio.* 13. doi:Artn25410.1186/1471-2148-13-254.

Haldane, J., 1949. Disease and Evolution. *Ricerca scient. Suppl.* 19, 68–76.

Hamilton, W.D., 1993. Haploid Dynamic Polymorphism in a Host with Matching Parasites: Effects of Mutation/Subdivision, Linkage, and Patterns of Selection. *Journal of Heredity* 84, 328–338. doi:10.1093/oxfordjournals.jhered.a111353.

Keeling, M.J., 2000. Metapopulation moments: coupling, stochasticity and persistence. *J. Animal Ecol.* 69, 725–736. doi:10.1046/j.1365-2656.2000.00430.x.

Klein, J., 1987. Origin of major histocompatibility complex polymorphism: The trans-species hypothesis. *Human Immunology* 19, 155–162. doi:10.1016/0198-8859(87)90066-8.

Klein, J., Figueroa, F., 1986. Evolution of the major histocompatibility complex. *Crit. Rev. Immunol.* 6, 295–386.

Lawlor, D.A., Ward, F.E., Ennis, P.D., Jackson, A.P., Parham, P., 1988. HLA-A and B polymorphisms predate the divergence of humans and chimpanzees. *Nature* 335, 268–271. doi:10.1038/335268a0.

Leonard, K.J., 1977. Selection Pressures and Plant Pathogens. *Ann. N. Y. Acad. Sci.* 287, 207–222. doi:10.1111/j.1749-6632.1977.tb34240.x.

- Leonard, K.J., 1994. Stability of equilibria in a gene-for-gene coevolution model of host-parasite interactions. volume 84. doi:10.1094/Phyto-84-70.
- Lively, C.M., 1999. Migration, virulence, and the geographic mosaic of adaptation by parasites. *Am. Nat.* 153, 34–47. doi:10.1086/303210.
- 395 Llaurens, V., Whibley, A., Joron, M., 2017. Genetic architecture and balancing selection: the life and death of differentiated variants. *Molecular Ecology* 26, 2430–2448. doi:10.1111/mec.14051.
- MacPherson, A., Otto, S.P., Nuismer, S.L., 2018. Keeping Pace with the Red Queen : Identifying the. *Genetics* 208, 779–789. doi:10.1534/genetics.117.300481/-/DC1.1.
- M’Gonigle, L.K., Otto, S.P., 2011. Ploidy and the evolution of parasitism. *Proc. R. Soc. Lond. B* 278, 2814–2822. doi:10.1098/rspb.2010.2146.
- 400 M’Gonigle, L.K., Shen, J.J., Otto, S.P., 2009. Mutating away from your enemies: The evolution of mutation rate in a host-parasite system. *Theor. Popul. Biol.* 75, 301–311. doi:10.1016/j.tpb.2009.03.003.
- Nuismer, S.L., 2006. Parasite Local Adaptation in a Geographic Mosaic. *Evolution* 60, 24. doi:10.1554/05-380.1.
- 405 Nuismer, S.L., 2017. *Introduction to Coevolutionary Theory*. First. ed., Macmillan Learning, New York.
- Nuismer, S.L., Otto, S.P., 2004. Host-parasite interactions and the evolution of ploidy. *PNAS* 101, 11036–11039. doi:10.1073/pnas.0403151101.
- Otto, S.P., Michalakis, Y., 1998. The evolution of recombination in changing environments. *Nat. Rev. Genet.* 13, 145–151.
- 410 Papkou, A., Gokhale, C.S., Traulsen, A., Schulenburg, H., 2016. Host–parasite coevolution: why changing population size matters. *Zoology* 119, 330–338. doi:10.1016/j.zool.2016.02.001.
- Rabajante, J.F., Tubay, J.M., Ito, H., Uehara, T., Kakishima, S., Morita, S., Yoshimura, J., Ebert, D., 2016. Host-parasite Red Queen dynamics with phase-locked rare genotypes. *Science Advances* 2, e1501548. URL: <https://advances.sciencemag.org/content/2/3/e1501548>, doi:10.1126/sciadv.1501548. publisher: American Association for the Advancement of Science Section: Research Article.
- 415 Schenk, H., Schulenburg, H., Traulsen, A., 2018. Suicidal Red Queen: Population dynamics and genetic drift accelerate diversity loss. *bioRxiv* , 490201doi:10.1101/490201.
- Takahata, N., Nei, M., 1990. Allelic Genealogy Under Overdominant and Frequency-Dependent Selection and Polymorphism of Major Histocompatibility Complex Loci. *Genetics* , 967–978.
- 420 Tellier, A., Brown, J.K., 2009. The influence of perenniality and seed banks on polymorphism in plant-parasite interactions. *American Naturalist* 174, 769–779. doi:10.1086/646603.

Tellier, A., Moreno-Gómez, S., Stephan, W., 2014. Speed of adaptation and genomic footprints of host-parasite coevolution under arms race and trench warfare dynamics. *Evolution* 68, 2211–2224. doi:10.1111/evo.12427.

425 Unckless, R.L., Lazzaro, B.P., Unckless, R.L., 2016. The potential for adaptive maintenance of diversity in insect antimicrobial peptides. *Phil. Trans. R. Soc. B* .

Woolhouse, M.E.J., Webster, J.P., Domingo, E., Charlesworth, B., Levin, B.R., 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics* 32, 569–577. doi:10.1038/ng1202-569.

430 Zhao, L., Waxman, D., 2016. The influence of genetic drift on the formation and stability of polymorphisms arising from negative frequency-dependent selection. *J. Theor. Biol.* 391, 51–64. doi:10.1016/j.jtbi.2015.11.011.

Zimmer, C., Emlen, D.J., 2013. *Evolution: Making sense of life*. 2nd ed., W.H. Freeman & Company.