THE ROYAL SOCIETY

Received 8 July 2002
Accepted 10 September 2002
Published online 13 December 2002

Modelling infection as a two-step process combining gene-for-gene and matching-allele genetics

Aneil F. Agrawal* **and Curtis M. Lively**

Department of Biology, Indiana University, Bloomington, IN 47405-3700, USA

The genetic basis of infection determines the dynamics of host–parasite coevolution and associated phenomena such as local adaptation and the evolution of sex and recombination. Here, we present parasite resistance as a two-step process in which hosts must first detect parasites and then eradicate them; failure at either step results in infection. The model incorporates 'matching-allele' (MA) genetics for detection and 'gene-for-gene' (GFG) genetics for eradication. We found that the oscillatory dynamics were similar to pure GFG genetics when the cost of 'virulence' alleles was low, but resembled pure MA genetics when the cost was high. The magnitude of the cost that switched the dynamics from GFG dominated to MA dominated depended on the genetic architecture of defence (i.e. the number of GFG and MA loci).

Keywords: host–parasite coevolution; infection; genetics

1. INTRODUCTION

Compelling evidence exists from many systems that some parasite genotypes are better at infecting particular host genotypes than are other parasite genotypes (Flor 1956; Thompson & Burdon 1992; Henter & Via 1995; Webster & Woolhouse 1998; Carius *et al.* 2001). Two major classes of models have been developed that capture this empirical observation: gene-for-gene (GFG) models and matching-alleles (MA) models. Important conceptual differences separate these two types.

Almost 50 years ago, Flor (1956) deduced the GFG model from an empirical dataset of cross-infection experiments. With respect to infection, some parasite alleles are intrinsically better than their alternatives: so-called 'virulence' alleles allow infection of a wider array of host genotypes than do their 'avirulence' counterparts. With respect to resistance, some host alleles are intrinsically better than their alternatives: 'resistance' alleles allow resistance against a wider array of parasite genotypes than do their 'susceptible' counterparts. Costs associated with virulence (resistance) alleles are required to maintain variation at parasite (host) loci (reviewed in Frank (1992)).

MA models were inspired by the notion of self–nonself recognition systems that underlie animal immune systems (Grosberg & Hart 2000). In these types of models, each parasite genotype is better than other parasite genotypes at infecting some subset of host genotypes, but is worse at infecting other host genotypes. No single parasite genotype is best at infecting all host genotypes; no single host genotype is best at resisting all parasite genotypes (Carius *et al.* 2001). Rare genotype advantage and frequency-dependent selection maintains diversity at both host and parasite loci; costs are not invoked.

Relative to GFG models, MA models typically produce allele-frequency dynamics characterized by shorter periods and higher amplitudes (Parker 1994; Agrawal & Lively 2002). This difference in allele-frequency dynamics is thought to underlie important differences in population-level phenomena produced under the different models. For example, MA models generate selection for recombination whereas GFG models do not (Parker 1994; Agrawal & Lively 2002). GFG models lead to very high values of prevalence in unstructured populations (Thrall & Burdon 2002), whereas prevalence in MA models tends to oscillate around lower values (Lively 1999). Higher levels of local adaptation are observed under MA models than under GFG models (Lively 1999).

Although much of the empirical research from plant studies has been interpreted from the GFG perspective, the validity of this approach has been debated (Frank 1996; Parker 1996). The issue of whether GFG models or MA models more accurately represent the real world remains an open question that is crucial to evaluating the role of parasites in selecting for sex and recombination.

In the present paper, we examine the possibility that both types of systems contribute to the genetic basis of infection. Previous models exploring the parameter space between pure MA models and pure GFG models have assumed that a locus may work partially like GFG and partially like MA but that all loci work the same (Parker 1994; Agrawal & Lively 2002). Here, we take the perspective that some loci work like true GFG loci whereas others work like true MA loci. Such a perspective arises naturally from viewing resistance as a two-step process whereby a host must first recognize that a parasite is present and then attempt to eradicate it. The primary goal of this model is to examine how different components of 'complex' immune interaction contribute to evolutionary dynamics.

2. THE MODEL

We consider a haploid model of host–parasite coevolution, in which we view infection as a two-step process. In the first step, hosts must be able to detect that a parasite is present. We model this step as following a MA process:

* Author for correspondence (aagrawal@bio.indiana.edu).

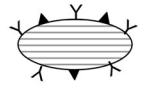
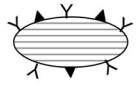
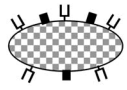
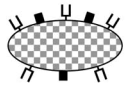
genotype	host: detector cell	host: immune-response protein	parasite: A_1V	parasite: A_1v	parasite: A_2V	parasite: A_2v
A_1R			nd ⇒ I	nd ⇒ I	d&e ⇒ R	d&ne ⇒ I
A_1r		none	nd ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne⇒ I
A_2R			d&e ⇒ R	d&ne⇒ I	nd ⇒ I	nd ⇒ I
A_2r		none	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	nd ⇒ I

Figure 1. Visualizing a two-step model of infection. Step 1: hosts have detector cells that look for foreign entities based on a self–nonself recognition system. Hosts with the A_1 genotype are depicted by striped cells with triangular projections and forked detectors. They are incapable of recognizing parasites of the A_1 genotype (striped cells with triangular projections) as foreign entities. Consequently, A_1 parasites are not detected (nd) by A_1 hosts, resulting in a successful infection (I). By contrast, A_2 parasites (chequered cells with rectangular projections) do not match A_1 hosts and are detected (d) by detector cells. Step 2: upon detection of a foreign entity by a detector cell, hosts of the R genotype release immune-response proteins. Upon binding to the forked-shaped structure on the cell surface of parasites of the V genotype, the immune-response protein eradicates (e) the parasite (R = resistance, shown in grey). In the absence of a host immune response, the forked-shaped structure enhances parasite fitness, but it is not required for parasite survival. Parasites of the V genotype do not produce this structure and thus cannot be eradicated (ne) by the immune-response protein. Note that hosts of the r genotype do not produce immune-response proteins (or at least not the protein effective against this parasite) and thus cannot eradicate (ne) this species of parasite even when it is detected. In the absence of this parasite species, there may be a cost to the ability to produce this immune-response protein if it is needlessly produced any time a detector cell finds a foreign entity. Our model is not dependent upon the hypothetical scenario described herein. It does rely on the host–parasite compatibilities shown here and in table 1. It is possible to envision other scenarios that would produce the same set of compatibilities.

if a parasite matches its host at the relevant loci, then the parasite is able to live undetected within its host; if a parasite does not match its host then the host detects that a parasite is present (figure 1). The second step occurs as the host attempts to remove an identified parasite. We model this second step as following a GFG process: hosts are able to eliminate a parasite if it has at least one resistance allele for which the parasite has an avirulence allele at the corresponding locus (figure 1). This model produces two interesting results from the perspective of parasites. When a parasite matches its host at the MA loci, the status of the GFG loci is irrelevant as the host does not recognize that the parasite is present. Alternatively, when the parasite 'wins' at the GFG loci, the status of the MA loci is irrelevant because its host is incapable of eliminating this parasite even when if it has been recognized.

We considered three versions of the model described. In the first version (model 1), there is a single locus controlling each step (i.e. one MA locus and one GFG locus). Figure 1 shows the parasite genotypes that can infect each host genotype. In the second version (model 2), there is a single locus controlling the first step and two loci determining the outcome of the second step (i.e. one MA locus and two GFG loci; see table 1*a*). In the third version (model 3), there are two loci involved in the first step and a single locus controlling the second step (i.e. two MA loci and one GFG locus; see table 1*b*).

The fitness of parasite genotype j on host genotype i is zero if the host resists infection, $W_{\mathrm{P}ji} = 0$. If parasite genotype j is able to infect host genotype i then the fitness of this parasite is $w_{\mathrm{P}ji} = (1 - k)^z$, where k is the cost of each virulence allele and z is the number of virulence alleles carried by the parasite. The fitness of host genotype i with parasite genotype j is $w_{\mathrm{H}ij} = (1 - c)^y(1 - w_{\mathrm{P}ji}s)$, where c is the cost of each resistance allele, y is the number of resistance alleles carried by the host and s is the maximum virulence of a parasite.

Assuming infinite populations of randomly mating hosts and parasites, where contact between hosts and parasites is also random, the average fitness of hosts is

$$\overline{W}_{\mathrm{H}} = \sum_{i=1}^{n_{\mathrm{H}}} h_i \sum_{j=1}^{n_{\mathrm{P}}} w_{\mathrm{H}ij} p_j,$$

where h_i is the frequency of the ith host genotype, n_{H} is the number of host genotypes, p_j is the frequency of the jth parasite genotype and n_{P} is the number of parasite genotypes. The relative fitness of the ith host genotype is

$$W_{\mathrm{H}i} = \left(\frac{1}{\overline{W}_{\mathrm{H}}}\right) \sum_{j=1}^{n_{\mathrm{P}}} w_{\mathrm{H}ij} p_j.$$

Table 1. Infection matrices for models 2 and 3.
(The reasoning underlying the entry in each cell is the same as that given in the legend for figure 1 for which infection is seen as a two-step process combining aspects of MA and GFG genetics. Abbreviations: I, infection; R, resistant (shown in bold); d, detected; nd, not detected; e, eradicated; ne, not eradicated.)

(*a*) model 2: one MA and two GFG loci

	host							
parasite	A_1rr	A_1rR	A_1Rr	A_1RR	A_2rr	A_2rR	A_2Rr	A_2RR
A_1VV	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	d&e ⇒ **R**
A_1Vv	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**
A_1vV	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	d&ne ⇒ I	d&e ⇒ **R**
A_1vv	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I
A_2VV	d&ne ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	d&e ⇒ **R**	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I
A_2Vv	d&ne ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I
A_2vV	d&ne ⇒ I	d&e ⇒ **R**	d&ne ⇒ I	d&e ⇒ **R**	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I
A_2vv	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I	nd ⇒ I

(*b*) model 3: two MA and one GFG loci

	host							
parasite	A_1B_1r	A_2B_1r	A_1B_2r	A_2B_2r	A_1B_1R	A_2B_1R	A_1B_2R	A_2B_2R
A_1B_1V	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	d&e ⇒ **R**
A_2B_1V	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	nd ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**
A_1B_2V	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	nd ⇒ I	d&e ⇒ **R**
A_2B_2V	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&e ⇒ **R**	d&e ⇒ **R**	d&e ⇒ **R**	nd ⇒ I
A_1B_1v	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&nr ⇒ I	d&nr ⇒ I
A_2B_1v	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&nr ⇒ I	d&nr ⇒ I
A_1B_2v	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&nr ⇒ I
A_2B_2v	d&ne ⇒ I	d&ne ⇒ I	d&ne ⇒ I	nd ⇒ I	d&ne ⇒ I	d&ne ⇒ I	d&nr ⇒ I	nd ⇒ I

Similarly, the mean fitness of parasites is

$$\overline{W}_{\mathrm{P}} = \sum_{j=1}^{n_{\mathrm{P}}} p_j \sum_{i=1}^{n_{\mathrm{H}}} w_{\mathrm{P}ji} h_i$$

and the relative fitness of the jth parasite genotype is

$$W_{\mathrm{P}j} = \left(\frac{1}{\overline{W}_{\mathrm{P}}}\right) \sum_{i=1}^{n_{\mathrm{H}}} w_{\mathrm{P}ji} h_i.$$

To prevent the fixation of parasite alleles under high values of virulence, we also included a mutation rate, μ, of *ca.* 3×10^{-6} to each genotype.

In our simulations, we explored 11 levels of the cost of virulence spanning the total possible range, $k = 0, 0.1, \ldots, 1$. As there is no evidence of very large costs of resistance, we restricted this parameter to a smaller range $c = 0.05$, 0.1, 0.2. The fact that such costs have been difficult to detect empirically (Frank 1992) indicates that even this confined range of c values may include costs that are larger than those that exist in reality. We also examined two levels of maximum virulence: $v = 0.5$, 0.75. Free recombination ($r = 0.5$) between loci was assumed for hosts; parasites reproduced asexually.

The simulations were initiated with both host and parasite genotypes in linkage equilibrium at random, intermediate allele frequencies for all loci. Data were collected from the simulation only after it had run for 4000 generations to allow the dynamics to stabilize. Each data point presented represents the average of three replicate simulations each initiated with a different set of randomly generated allele frequencies. There was little variation among the replicate runs for the data of interest.

Under many parameter values, allele-frequency dynamics were complex, making it difficult to characterize them with simple metrics such as period and amplitude. We therefore calculated the variance in the frequency of an allele between 4000 and 5000 generations. All else being equal, an increase in the variance in the frequency of an allele indicates an increase in the amplitude of allele-frequency fluctuations. We also measured the average absolute value of the rate of change of an allele per generation between 4000 and 5000 generations. All else being equal, an increase in the average per generation rate of allele-frequency change indicates a decrease in the period of allele-frequency fluctuations. Highly dynamic allele-frequency fluctuations were characterized by high variance and high rate. We also measured prevalence (fraction of hosts infected each generation) averaged over generations 4000–5000.

3. RESULTS

(a) *Model 1: one MA locus and one GFG locus*

The average per generation rate of host allele-frequency change over 4000–5000 generations is shown in figure 2 for both types of host loci (MA and GFG); figures 3 and

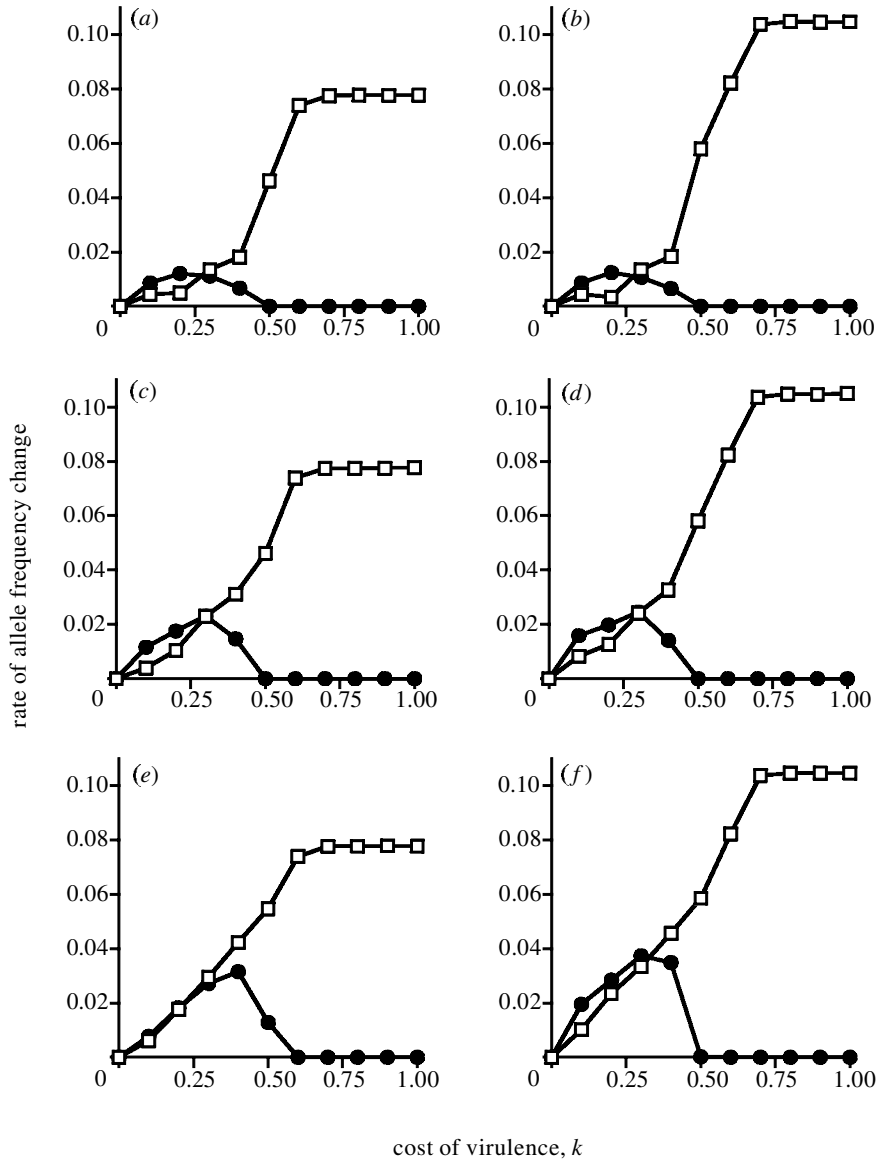


Figure 2. Results for model 1 (one MA locus and one GFG locus) showing the rate of allele-frequency change of host loci as a function of the cost of virulence, k, the maximum effect of infection (s: (*a*), (*c*) and (*e*) $s = 0.5$; (*b*), (*d*) and (*f*) $s = 0.75$) and the cost of host resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$). Open squares give the rate of change at the MA locus (detection), and closed circles give the rate of change for the GFG locus (eradication).

4 show analogous results for models 2 and 3, respectively. The variance in allele frequency displayed a similar pattern (results not shown). Taken together, the results indicate that at low values for the cost of virulence ($0 < k < 0.5$), the host GFG locus has dynamic behaviour whereas the host MA locus does not (figure 2). At higher values of k (i.e. 0.5), allele frequency at the host GFG locus is static whereas the host MA locus shows highly dynamic fluctuations. At low values of k, the virulence allele is common but typically plummets in the range $0.4 < k < 0.5$ (figure 5). Note that figure 5 shows the average frequency between generations 4000 and 5000; nonzero values always involve fluctuations around the average (as represented by figure 2). Considering the results of figures 2 and 5 together, we see that the MA locus becomes highly dynamic only when virulence alleles become rare. Increasing the level of virulence from 0.5 to 0.75 tends to increase the rate of change of dynamic loci but not the range over which the GFG locus is dynamic (figure 2).

Increasing the cost of resistance, c, increases the rate of change of the MA locus at low values of k. Previous studies on pure GFG systems have shown that increasing the cost of resistance will decrease the average frequency of virulence alleles (Frank 1992). As shown in figure 5, at low values of k, the frequency of virulence allele is indeed lower when c is higher. This reduction in the frequency of the virulence alleles is likely to be responsible for the increased rate of change of the MA locus (see § 4).

(b) *Model 2: one MA locus and two GFG loci*

The results for the rate of allele-frequency change for this model are shown in figure 3. As in the previous model, the MA locus becomes dynamic (figure 3) as the frequency of virulence alleles declines (figure 5). This tends to occur at lower values of k than in model 1. In model 2

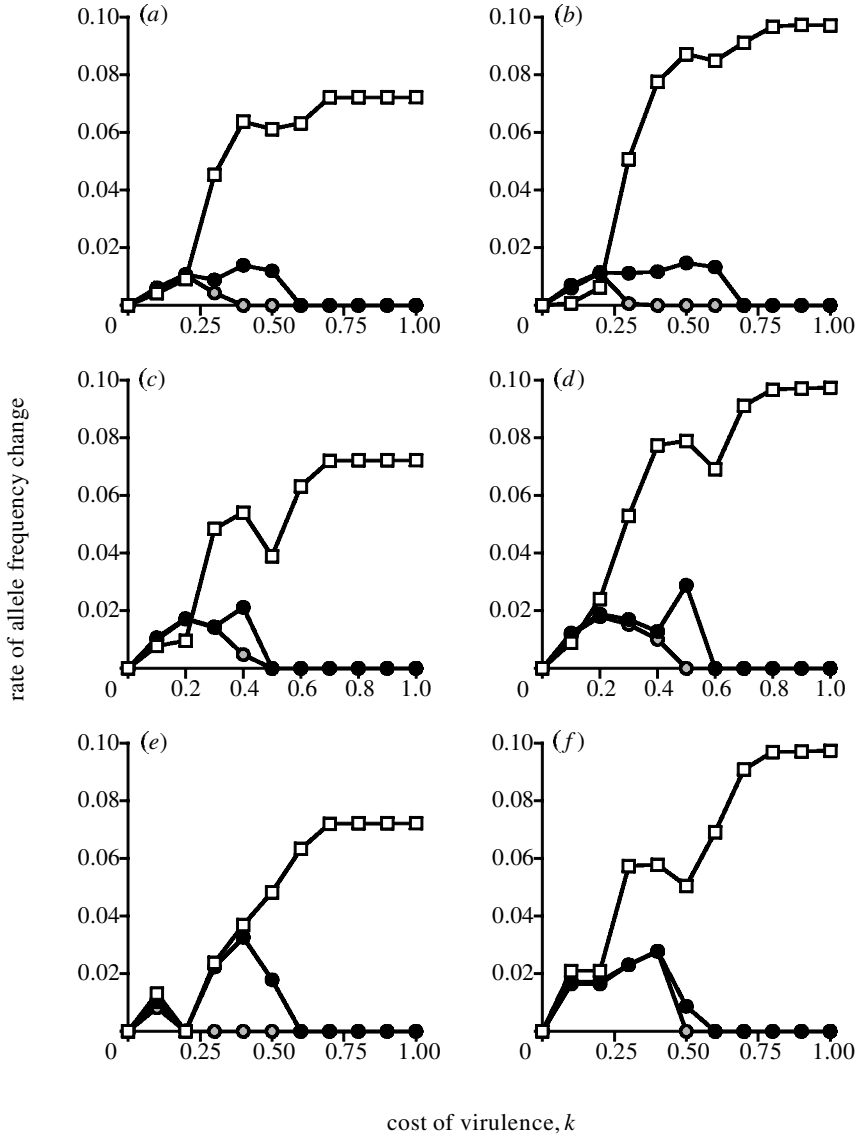


Figure 3. Results for model 2 (one MA locus and two GFG loci) showing the rate of allele-frequency change of host loci as a function of the cost of virulence, k, the maximum effect of infection (s: (*a*), (*c*) and (*e*) $s = 0.5$; (*b*), (*d*) and (*f*) $s = 0.75$) and the cost of host resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$). Open squares give the rate of change at the MA locus (detection), and black filled circles give the rate of change of the least (most) dynamic GFG locus (eradication). Data for both GFG loci are shown, but often only a single symbol or line appears because the values are equal.

the MA locus becomes highly dynamic typically by $k = 0.3$ rather than $k = 0.5$ (as seen in model 1). Dynamic behaviour from the host GFG loci disappears when $k = 0.5$. At intermediate values of k in the range *ca.* 0.4–0.6 the GFG loci often behave quite differently from each other and in ways not reflected by the corresponding loci in the other species. For example with $s = 0.5$, $k = 0.5$ and $c = 0.1$, the avirulence allele fixed at the first GFG locus while the avirulence allele at the second GFG locus oscillated rapidly between 45% and 90%. Hosts only partially reflected these frequencies: the susceptibility allele fixed at the first host GFG locus but the resistance allele fixed at the second host GFG locus. In this case, the system effectively collapsed to having one functional GFG locus. As the corresponding host GFG locus was fixed for the resistance allele, the oscillations of the avirulence allele indicate that the parasites were fluctuating between evading detection by hosts (i.e. matching them) and depending on the virulence allele to avoid eradication.

(c) *Model 3: two MA loci and one GFG locus*

The results for this model are shown in figure 4. In general, similar patterns are seen with this model as for model 1 but with one major difference. The switch from GFG-dominated dynamics to MA-dominated dynamics occurs at much higher costs of virulence ($k \sim 0.8$).

(d) *Prevalence*

The average prevalence over 4000–5000 generations for each of the three models is plotted in figure 6. Prevalence is always high at low values of k and is reduced at high values of k. Declines in prevalence correspond with the switch to MA-dominated dynamics (figures 2–4).

4. DISCUSSION

It is unclear whether natural systems tend to work like pure GFG models or pure MA models. But it is not unreasonable to propose that the interaction between

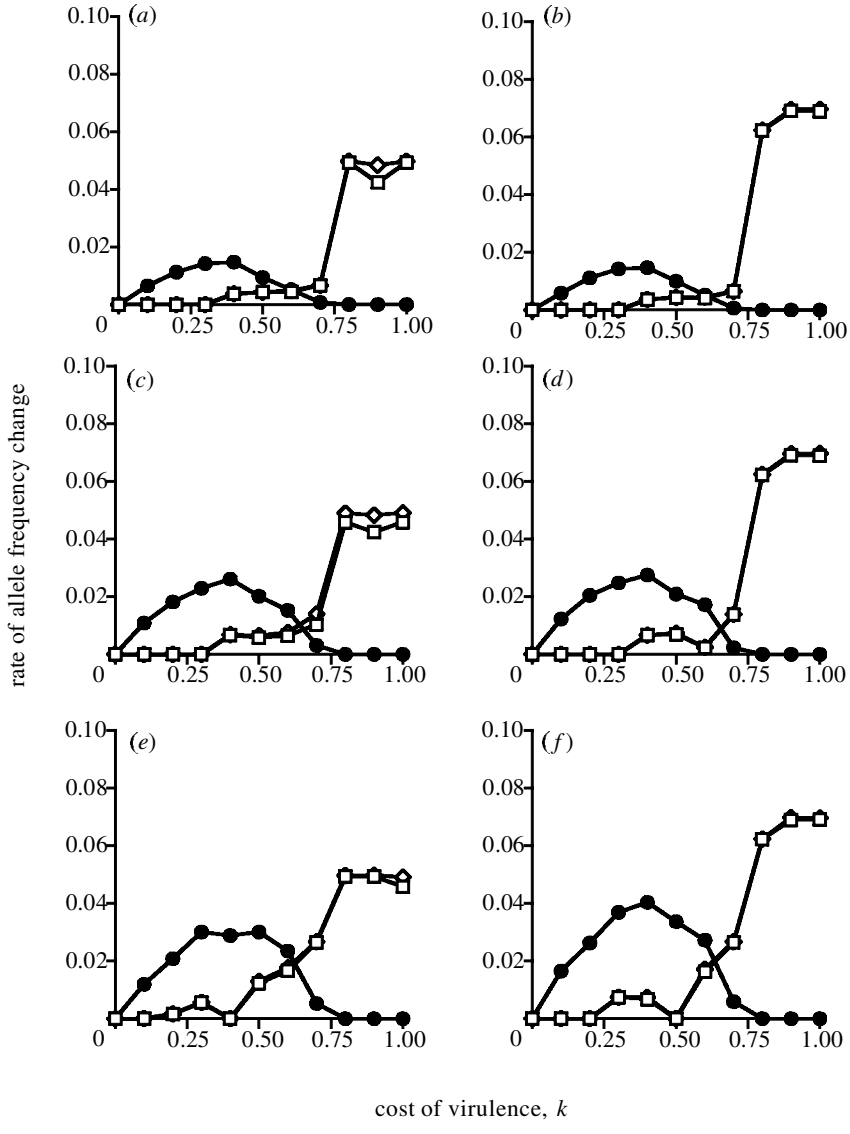


Figure 4. Results for model 3 (two MA loci and one GFG locus) showing the rate of allele-frequency change of host loci as a function of the cost of virulence, k, the maximum effect of infection (s: (*a*), (*c*) and (*e*) $s = 0.5$; (*b*), (*d*) and (*f*) $s = 0.75$) and the cost of host resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$)). Open squares (diamonds) give the rate of change at the least (most) dynamic MA locus, and closed circles give the rate of change at the GFG locus (eradication). Data for both MA loci are shown, but often only a single symbol or line appears because the values are equal.

hosts and parasites involves elements of both models. Extrapolating from the idea that hosts must first detect parasites and then attempt to eradicate them, we have built a model including both MA and GFG loci.

Surprisingly, even though our model inherently contains both MA and GFG loci, the population genetic dynamics of this host–parasite system often appear as though only one of the two types (MA or GFG) is present. When the cost of virulence is low, allele-frequency dynamics are GFG dominated, though the MA loci always remain highly polymorphic. At higher costs of virulence, the avirulence alleles fix, and the allele-frequency dynamics are MA dominated.

These results can be interpreted from the perspective of parasite virulence alleles. First, consider the typical dynamics of a pure GFG model (Sasaki 2000; Agrawal & Lively 2002). Virulence alleles are advantageous, and so become common, because they allow parasites to infect a wide array of host genotypes. If there is any cost to the host resistance alleles ($c > 0$), then resistance alleles are selected against when virulence alleles are common, because resistance alleles provide no protection against parasites carrying virulence alleles. When the frequency of resistance alleles drops to low levels, parasites no longer require costly ($k > 0$) virulence alleles to cause infection and consequently virulence alleles decline in frequency. As virulent alleles decrease in frequency, resistance alleles regain their advantage causing them to increase, which once again generates positive selection on virulence alleles.

In pure GFG models, when the cost of virulence is low, parasite allele-frequency cycles are characterized by long periods where the frequency of the virulence allele is very high punctuated by short spikes in which the virulence allele rapidly declines and then increases again. We observe this basic pattern in the mixed models presented here when the cost of virulence is low. During the long periods when virulence alleles are common, the frequency

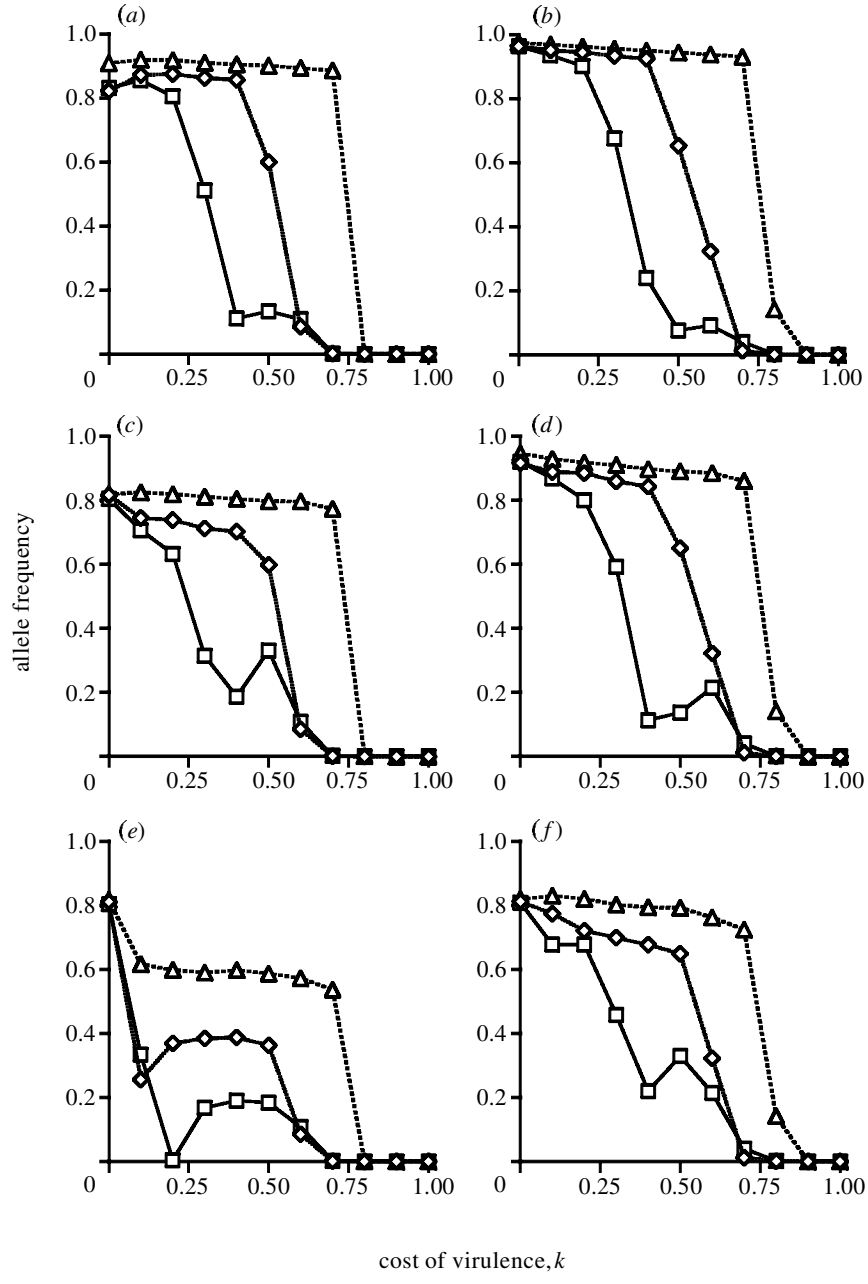


Figure 5. Frequency of the virulence allele(s) at the GFG locus (loci) as a function of function of the cost of virulence, k, the maximum effect of infection (s: (*a*), (*c*) and (*e*) $s = 0.5$; (*b*), (*d*) and (*f*) $s = 0.75$) and the cost of host resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$) for each of the three models. Diamonds show the results for model 1 (one MA and one GFG loci); squares show the results for model 2 (one MA and two GFG loci); and triangles show the results for model 3 (two MA and one GFG loci). The results from model 2 (squares) are given as the average of the virulence allele frequencies across both loci.

of the parasite genotype that has virulence alleles at all GFG loci is common. Such a parasite is capable of infecting every host genotype and this parasite genotype constitutes the so-called 'universally virulent genotype' (UVG). When the UVG becomes common, there is no selection on the MA loci as their alleles provide no benefit (or detriment) to parasites with the UVG. Consequently, the MA loci are essentially neutral much of time when the cost of virulence is low.

In pure GFG models, virulence alleles cycle even when the cost of virulence is high, provided that the cost of resistance is not too high (Sasaki 2000). Parasite populations never lose virulence alleles in pure GFG models because, when resistance alleles become common, virulence alleles become absolutely essential to parasites, regardless of their cost. Under a pure GFG model, a parasite must carry virulence alleles in order to infect a host with resistance alleles. This requirement does not hold in our mixed model. In our model, an alternative route to infection exists for parasites. By matching hosts at the MA loci, parasites can avoid detection by their hosts, making GFG loci irrelevant. However, different hosts have different MA loci making it impossible for any parasite genotype to be able to evade detection by all hosts. In this respect, matching is inferior to carrying virulence alleles (i.e. the UVG can infect all hosts whereas there is no such thing

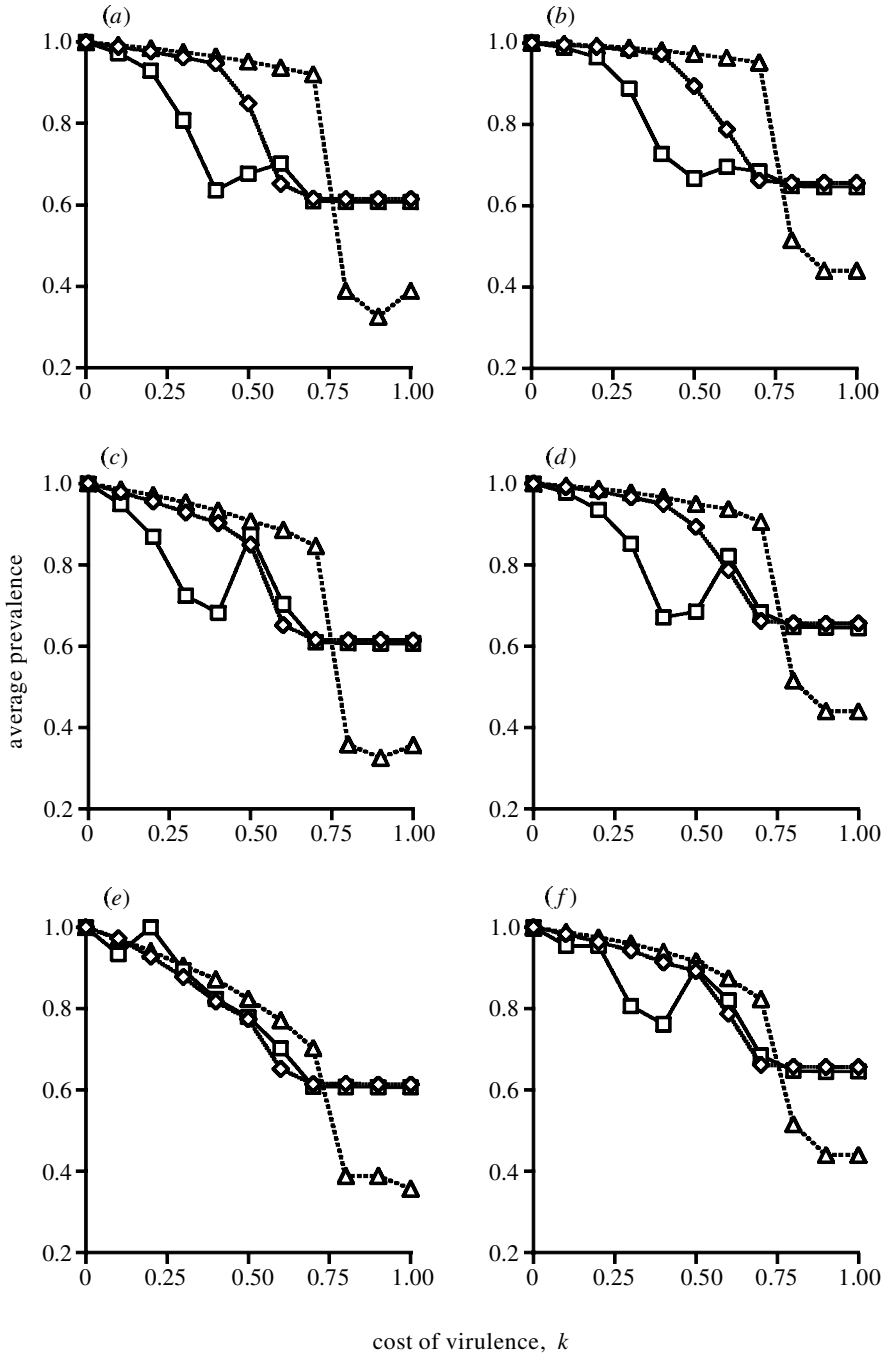


Figure 6. Prevalence of infection as a function of function of the cost of virulence, k, the maximum effect of infection (s: (*a*), (*c*) and (*e*) $s = 0.5$; (*b*), (*d*) and (*f*) $s = 0.75$) and the cost of host resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$), and the cost of host resistance (c) for each of the three models. Diamonds show the results for model 1 (one MA and one GFG loci); squares show the results for model 2 (one MA and two GFG loci); and triangles show the results for model 3 (two MA and one GFG loci).

as a universally matching parasite). However, when the cost of virulence is high, it becomes more profitable for parasites to drop their virulence alleles and attempt to match their hosts. Indeed, we find that when the cost of virulence is high, virulence alleles are driven out of the population (figure 5) and MA loci dominate the dynamics.

At low values of k, virulence alleles are profitable and tend to be common. As parasites carrying virulence alleles can resist eradication by the host immune system, avoiding detection becomes irrelevant. Selection on the MA loci is limited to brief episodes during GFG cycles where virulence becomes less common. In other words, dynamics at the MA loci can only occur when the conditions at the GFG loci permit them. However, the MA loci also exert some control over the fate of the GFG loci. As discussed in the previous paragraph, the existence of the MA loci provides parasites with an alternative route of successful infection and can drive virulence alleles out of the system if their intrinsic cost is too high.

The switch from GFG-dominated dynamics to MA-dominated dynamics can have important ecological consequences as reflected in population prevalence (figure 6). At low values of k, prevalence is high ($> 90\%$); such high prevalences are typical of pure GFG models in unstruc-

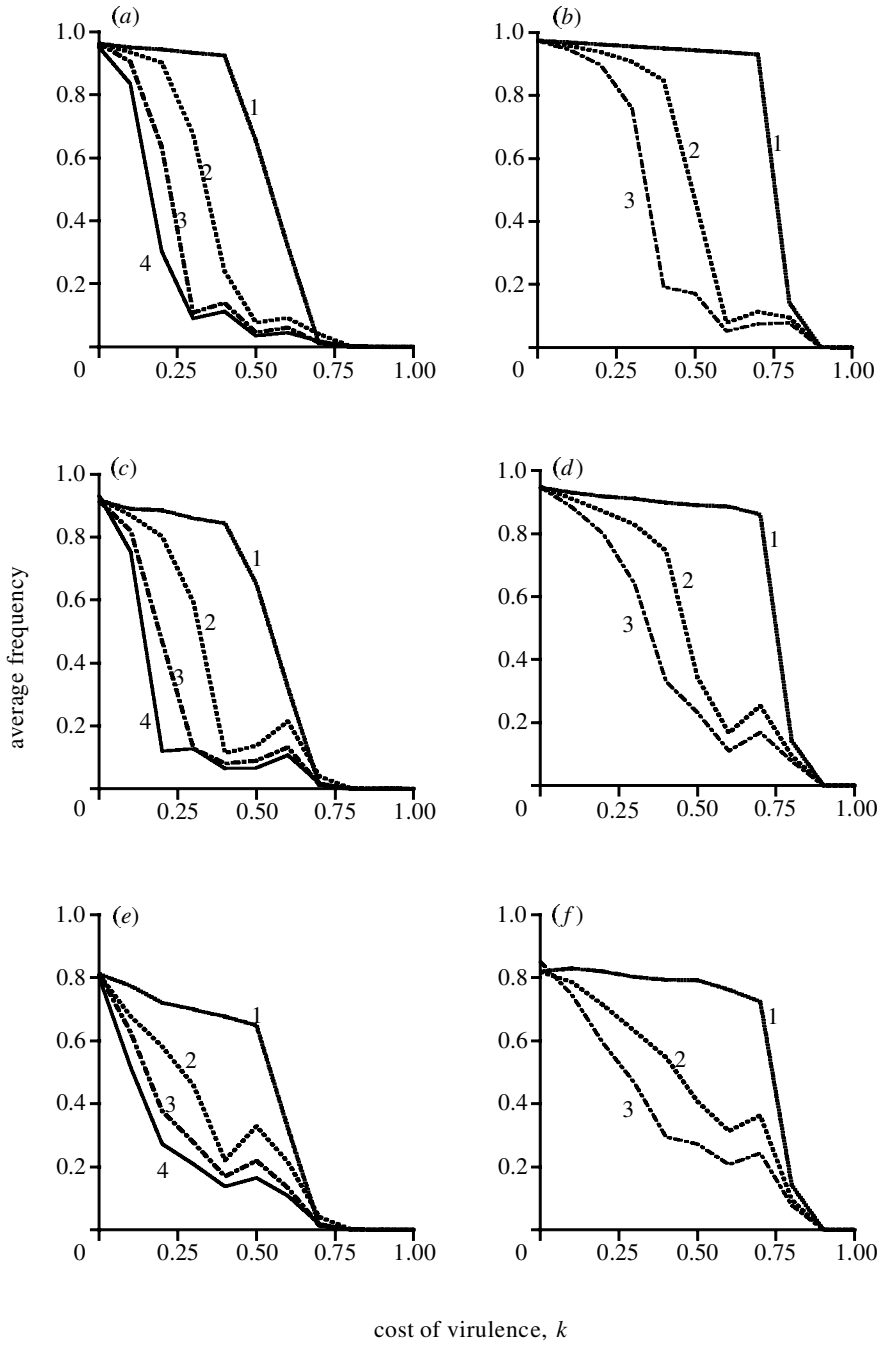


Figure 7. The effect of the number of GFG loci on the transition from GFG-dominated to MA-dominated dynamics. Following the logic underlying models 1–3 (figure 1, table 1*a*,*b*) we generated a set of models in which the number of MA loci is held constant ((*a*, *c* and *e*) one MA locus; (*b*, *d* and *f*) two MA loci) and the number of GFG loci varies. The average frequency of virulence alleles is plotted as function of the cost of virulence (*k*) for three different costs of resistance (*c*: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$) under models with different numbers of GFG loci (represented by different lines). The labels (1–4) in each panel correspond to the number of GFG loci. The average frequency of virulence alleles reflects the extent to which GFG versus MA loci dominate coevolutionary dynamics. When the average frequency of virulence alleles is high (low), GFG (MA) loci dominate coevolutionary loci. Note that the transition from GFG-dominated to MA-dominated dynamics occurs at lower values of *k* for models with more GFG loci.

tured populations (Thrall & Burdon 2002). At higher values of *k*, prevalence declines to levels typical of pure MA models with the same number of loci. Note that as the cost of resistance, *c*, increases, the transition from GFG-dominated dynamics to MA-dominated dynamics becomes more gradual and this can be observed in the decline in prevalence shown in figure 6.

In his exchange with Parker on MA versus GFG, Frank noted that the bulk of evidence supporting the GFG model came from agricultural systems (Frank 1996). He suggested that perhaps GFG appeared more prevalent in these systems because parasites could afford to carry virulence alleles with large pleiotropic costs as transmission was much easier in agricultural settings. Our results show that if infec-

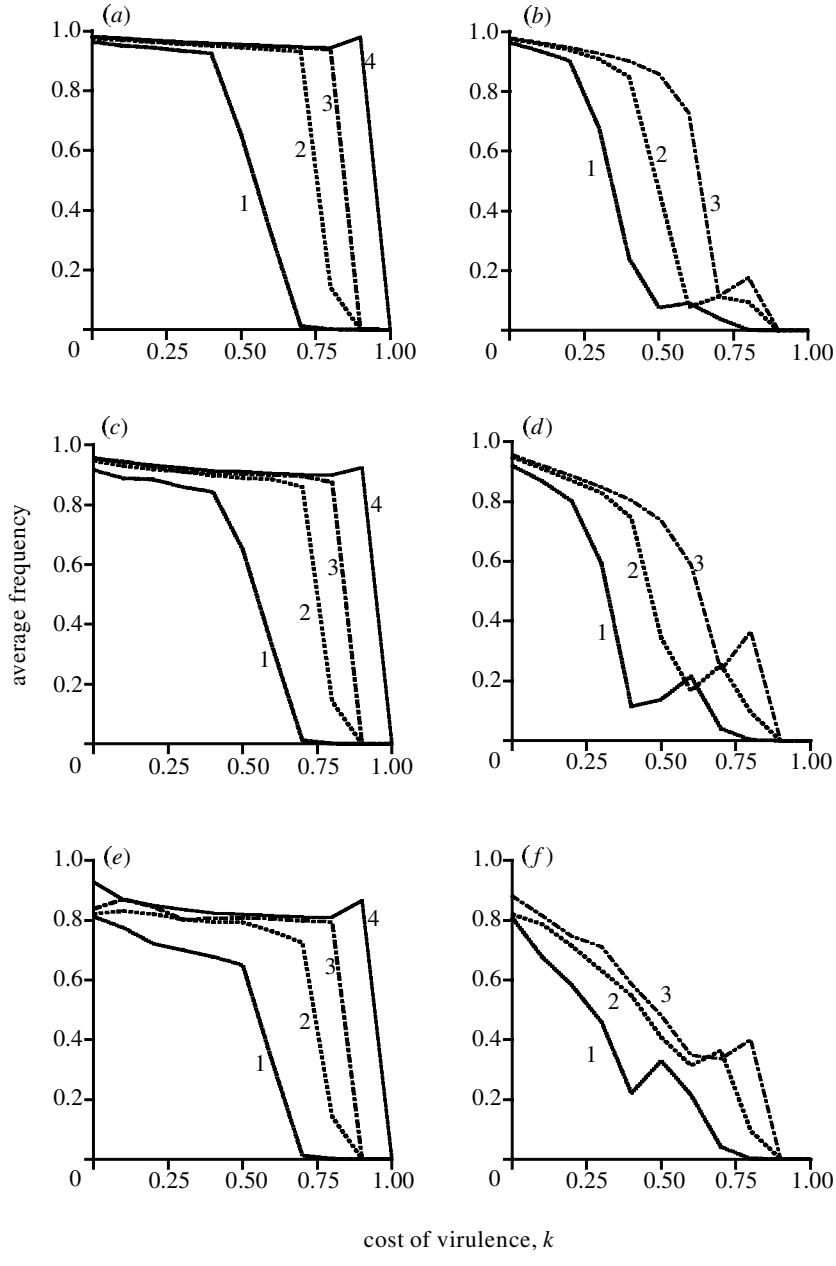


Figure 8. The effect of the number of MA loci on the transition from GFG-dominated to MA-dominated dynamics. Following the logic underlying models 1–3 (figure 1, table 1*a*,*b*) we generated a set of models in which the number of GFG loci is held constant ((*a*, *c* and *e*) one GFG locus; (*b*, *d* and *f*) two GFG loci) and the number of MA loci varies. The average frequency of virulence alleles is plotted as a function of the cost of virulence (k) for three different costs of resistance (c: (*a*,*b*) $c = 0.05$; (*c*,*d*) $c = 0.1$; (*e*,*f*) $c = 0.2$) under models with different numbers of GFG loci (represented by different lines). The labels (1–4) in each panel correspond to the number of MA loci. The transition from high frequency of virulence alleles (GFG-dominated dynamics) to low frequency of virulence alleles (MA-dominated dynamics) occurs at higher values of k for models with more MA loci.

tion is mediated by a system containing elements of both GFG and MA, then the dynamics are dominated by the GFG elements when the cost of virulence is low but are dominated by the MA elements when the cost of virulence is high. Changes in the ecology of the host–parasite system (a dramatic case being agriculture) that reduce the costs of virulence can switch the allele-frequency dynamics from being dominated by GFG to MA.

The physiology and biochemistry of host–parasite interactions are complex. These interactions are likely to involve multiple steps that work in qualitatively different ways. Although there have been many genetic analyses of host–parasite coevolution, almost all previous models depict a single-step interaction (but the type of interaction varies from model to model). A notable exception is the model by Frank (2000) that examines the evolution of a host–parasite system involving both 'specific' (qualitative) and 'nonspecific' (quantitative) forms of resistance. Frank found that some genetic and epidemiological conditions resulted in a host–parasite system that appeared to involve primarily strain-specific resistance whereas other conditions produced a system characterized by variable, but

nonspecific, resistance. The model of Frank (2000) is very different from our own with respect to the type of steps included in the interaction as well as the overall construction of the model. Nonetheless, both he and we find that complicated, multistep interactions can appear much simpler at the population level. Furthermore, this appearance can be dramatically changed with small changes in parameters that may be ecologically determined (e.g. costs of resistance or virulence).

In each of our three models the switch from GFG-dominated dynamics to MA-dominated dynamics occurs at a different level of k. Model 2 (one MA locus: two GFG loci) requires the lowest value of k, and model 3 (two MA loci: one GFG locus) requires the highest. Two forces appear to be acting to determine where this switch occurs.

First, consider the cost of having the UVG. In our model with only a single GFG locus, universal virulence is achieved through possession of a single virulence allele. In our model with two GFG loci, universal virulence requires two virulence alleles (one at each locus), which is more costly than having a single virulence allele. When there are multiple GFG loci, having a single virulence allele does not guarantee a parasite that it will be able to infect a randomly sampled host, yet the parasite still must pay a cost. Holding the cost of each virulence allele, k, constant, possession of a single virulence allele has less value relative to its cost as the number of GFG loci increases; furthermore, the total cost of universal virulence gets higher and higher. Consequently, it makes sense for parasites to exploit the alternative infection pathway of matching at lower values of k when there are more GFG loci. Indeed, we see that our model with two GFG loci (model 2) switches to MA-dominated dynamics at a lower value of k than our models with only a single GFG locus (models 1 and 3). Ironically, these results indicate that as the number of GFG loci increases, the dynamics of the system will be dominated to a greater extent by the MA loci. This conjecture is corroborated by results from a set of five-loci models (figure 7).

Second, consider how difficult it is for a parasite to match its host and evade detection. As the number of MA loci increases, it will be increasingly difficult for a parasite to correctly match its host. Consequently, the value of having virulence alleles will increase. In accordance with this logic, we see that our model with two MA loci (model 3) switches to MA-dominated dynamics at a higher k value than our models with only a single MA locus (models 1 and 2). These results indicate that as the number of MA loci increases, the GFG loci will remain important over a greater range of parameter space. An investigation of a five-locus model supports this suggestion (figure 8).

Clearly, the models presented here are simplistic and abstract representations of real host–parasite interactions. Nonetheless they may provide helpful insights into coevolutionary dynamics. Our primary goal here was to investigate how different parts of a two-step model might interact with one another to influence coevolution and related population-level phenomena such as prevalence. Contrary to our initial expectations, the dynamics of our two-step models looked very much like pure GFG systems in some regions of parameter space, but like pure MA systems in other regions. The transition from GFG-dominated dynamics to MA-dominated dynamics, however, was not always discrete. When the cost of resistance was high, $c = 0.2$, the transition tended to be more gradual (figures 5–8). This transition from GFG-dominated dynamics to MA-dominated dynamics was observed in all the models we studied and reveals that both parts of the two-step model interact to determine how the system will appear at the population level. Real systems may involve many more loci than the models presented here. The balance between which type of loci will govern allele-frequency dynamics and determine population-level phenomena probably depends on the numbers of each type of locus in a counter-intuitive manner (figures 7 and 8).

5. CONCLUSIONS

We investigate a model in which infection is a two-step process that involves detection and eradication. The detection step is governed by MA loci whereas the eradication step is governed by GFG loci. GFG loci dominate the gene-frequency dynamics when the cost of virulence is low; when the cost is high, gene-frequency dynamics are dominated by MA loci. Increasing the number of GFG loci in the system decreases the cost at which the system will be dominated by MA loci. Conversely, increasing the number of MA loci in the system increases the cost at which the system will be dominated by MA loci. Increasing the cost of resistance can make the transition from GFG-dominated dynamics to MA-dominated dynamics more gradual.

The authors thank S. Frank and M. Parker for their interesting exchange of ideas that gave the motivation for this study. Funding for this study was provided by the US National Science Foundation to C.M.L. and by fellowship from the College and Arts and Sciences of Indiana University to A.F.A.

REFERENCES

Agrawal, A. & Lively, C. M. 2002 Infection genetics: gene-for-gene versus matching-allele models, and all points in between. *Evol. Ecol. Res.* **4**, 1–12.

Carius, H.-J., Little, T. J. & Ebert, D. 2001 Genetic variation in a host–parasite association: potential for coevolution and frequency dependent selection. *Evolution* **55**, 1136–1145.

Flor, H. H. 1956 The complementary genetic systems in flax and flax rust. *Adv. Genet.* **8**, 29–54.

Frank, S. A. 1992 Models of plant-pathogen coevolution. *Trends Genet.* **8**, 213–219.

Frank, S. A. 1996 Problems inferring the specificity of plant-pathogen genetics. *Evol. Ecol.* **10**, 319–322.

Frank, S. A. 2000 Specific and non-specific defense against parasitic attack. *J. Theor. Biol.* **202**, 283–304.

Grosberg, R. K. & Hart, M. W. 2000 Mate selection and the evolution of highly polymorphic self/nonself recognition genes. *Science* **289**, 2111–2114.

Henter, H. J. & Via, S. 1995 The potential for coevolution in a host–parasitoid system. I. Genetic variation within an aphid population in susceptibility to a parasitic wasp. *Evolution* **49**, 257–265.

Lively, C. M. 1999 Migration, virulence, and the geographic mosaic of adaptation by parasites. *Am. Nat.* **153**, S34–S47.

Parker, M. A. 1994 Pathogens and sex in plants. *Evol. Ecol.* **8**, 560–584.

Parker, M. A. 1996 The nature of plant–parasite specificity. *Evol. Ecol.* **10**, 319–322.

Sasaki, A. 2000 Host–parasite coevolution in a multilocus gene-for-gene system. *Proc. R. Soc. Lond.* B **267**, 2183–2188. (DOI 10.1098/rspb.2000.1267.)

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

Thrall, P. H. & Burdon, J. J. 2002 Evolution of gene-for-gene systems in metapopulations: the effect of spatial scale of host and pathogen dispersal. *Plant Pathol.* **51**, 169–184.

Webster, J. P. & Woolhouse, M. E. J. 1998 Selection and strain specificity of compatibility between snail intermediate hosts and their parasitic schistosomes. *Evolution* **52**, 1627–1643.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.