

# Polymorphism in Multilocus Host–Parasite Coevolutionary Interactions

Aurélien Tellier<sup>1</sup> and James K. M. Brown

*Department of Disease and Stress Biology, John Innes Centre, Colney, Norwich, NR4 7UH, United Kingdom*

Manuscript received April 11, 2007

Accepted for publication September 14, 2007

## ABSTRACT

Numerous loci in host organisms are involved in parasite recognition, such as major histocompatibility complex (MHC) genes in vertebrates or genes involved in gene-for-gene (GFG) relationships in plants. Diversity is commonly observed at such loci and at corresponding loci encoding antigenic molecules in parasites. Multilocus theoretical models of host–parasite coevolution predict that polymorphism is more likely than in single-locus interactions because recurrent coevolutionary cycles are sustained by indirect frequency-dependent selection as rare genotypes have a selective advantage. These cycles are stabilized by direct frequency-dependent selection, resulting from repeated reinfection of the same host by a parasite, a feature of most diseases. Here, it is shown that for realistically small costs of resistance and virulence, polycyclic disease and high autoinfection rates, stable polymorphism of all possible genotypes is obtained in parasite populations. Two types of epistatic interactions between loci tend to increase the parameter space in which stable polymorphism can occur with all possible host and parasite genotypes. In the parasite, the marginal cost of each additional virulence allele should increase, while in the host, the marginal cost of each additional resistance allele should decrease. It is therefore predicted that GFG polymorphism will be stable (and hence detectable) when there is partial complementation of avirulence genes in the parasite and of resistance genes in the host.

HOST–PARASITE interactions are recognized as a major evolutionary force producing biological diversity. Genetic variation for resistance reduces the probability that an individual parasite can infect an individual host (MAY and ANDERSON 1990) and conversely, genetic diversity at parasite recognition loci increases the range of potentially susceptible hosts. Spatial and temporal genetic polymorphism is commonly found in nature at loci involved in host–parasite recognition such as the major histocompatibility complex (MHC) in vertebrates (APANUS *et al.* 1997; HILL 2001) or genes involved in gene-for-gene (GFG) relationships, a common feature of plant–parasite interactions (THRALL *et al.* 2001; LAINE 2004). In both the MHC and the GFG systems, hosts and parasites may have multiple interacting loci (APANUS *et al.* 1997; HILL 2001; PALOMINO *et al.* 2002). Interactions among several plant resistance (*RES*) genes and parasite avirulence (*AVR*) genes have been documented for numerous diseases, of which the best studied include barley powdery mildew (JORGENSEN 1994), flax rust (THRALL *et al.* 2001), and rice blast (DEWIT 1992), as well as several diseases of the model plant *Arabidopsis thaliana* (HOLUB 2001).

In multilocus systems of host–parasite interactions, negative indirect frequency-dependent selection (FDS)

is thought to account for the great polymorphism found in MHC genes (APANUS *et al.* 1997; HILL 2001; BORGHANS *et al.* 2004) and GFG genes (FRANK 1993a; SASAKI 2000; SALATHE *et al.* 2005; SEGARRA 2005). In this hypothesis, host and parasite genotypes have a selective advantage when they are rare in coevolving populations. This leads to sustained coevolutionary cycles because when a parasite rare allele is selected, its frequency increases, selecting in turn for the corresponding resistant host genotype. It is hypothesized that these regular cycles of genotype frequencies prevent invasion by a single genotype, especially when mutations introduce new alleles in populations (SASAKI 2000; BORGHANS *et al.* 2004). An important question about coevolution is therefore whether or not multilocus interactions are sufficient to maintain polymorphism by themselves or if other ecological and biological factors are required.

In GFG relationships in plants, resistance is induced if the plant has a resistance (*RES*) gene enabling recognition of a specific parasite avirulence (*AVR*) protein (DANGL and JONES 2001). The parasite is not detected by the host and resistance is not induced if the host has a susceptibility allele (*res*) or the parasite has a virulence allele (*avr*). The asymmetry of the GFG interaction implies that in the absence of other factors, there will be an “arms race,” as successive pairs of *RES* and *AVR* alleles are driven to fixation in host and parasite populations, respectively (BERGELSON *et al.* 2001; HOLUB 2001). Accounting for the diversity observed at host and parasite

<sup>1</sup>Corresponding author: Section of Evolutionary Biology, Biocenter, University of Munich, 82152 Planegg-Martinsried, Germany.  
E-mail: tellier@zi.biologie.uni-muenchen.de

GFG loci (THRALL *et al.* 2001; LAINE 2004) is a significant challenge because a universally virulent pathogen genotype with an *avr* allele at each locus might be expected to become fixed as it can infect all plant genotypes (FRANK 1993a; SEGARRA 2005).

Conditions for the maintenance of polymorphism in GFG interactions have been studied in a single-locus system, with a single matching pair of a host *RES* gene and a parasite *AVR* gene (TELLIER and BROWN 2007). Coevolution implies the existence of indirect FDS, because the rate of natural selection on *RES* depends on the frequency of *avr* and vice versa. Polymorphism can be maintained only if there is also negative, direct FDS, such that the strength of natural selection for the host resistance allele or the parasite virulence allele or both declines with increasing frequency of that allele itself (TELLIER and BROWN 2007). Thus, while costs of *RES* and *avr* are necessary to maintain polymorphism, they are not sufficient to do so in a single-locus system (TELLIER and BROWN 2007) or in multilocus GFG interactions (SASAKI 2000; SEGARRA 2005). In a single-locus GFG interaction, direct FDS is generated if the parasite passes through more than one generation in the same host individual, a feature that is common to most plant diseases. Such polycyclic diseases are characterized by an autoinfection rate, the percentage of parasite spores reinfesting the same host from one parasite generation to the next (BARRETT 1980). In single-locus GFG interactions, stable long-term polymorphism can be most readily maintained in host and parasite populations at high autoinfection rates (TELLIER and BROWN 2007).

With respect to the stability of polymorphism, multilocus systems have behavior similar to that of single-locus GFG systems. For monocyclic diseases on annual plants (one parasite generation per host generation), long-term stable polymorphism with all possible host and parasite genotypes cannot be obtained (SASAKI 2000; SEGARRA 2005). However, multilocus interactions and high mutation rates increase the variance of the lifetime of a mutation (SASAKI 2000; SEGARRA 2005). Note that the dynamics of genotype frequencies in multilocus models are highly affected by stochastic processes (mutation and drift) when many loci are considered (FRANK 1993a; SASAKI 2000; SALATHE *et al.* 2005). Here, we investigate whether the existence of multiple GFG loci further stabilizes a system in which epidemiological and ecological factors generate direct FDS (TELLIER and BROWN 2007) or further increases the variance of the lifetime of transiently polymorphic alleles (HOLUB 2001; SALATHE *et al.* 2005; SEGARRA 2005).

A key issue is to discover the epidemiological and genetic factors that cause polymorphism at host and parasite multiple loci to be transient or stable (arms race or trench warfare models) (STAHL *et al.* 1999; HOLUB 2001). We extend the results of TELLIER and BROWN (2007) to a multilocus GFG system with polycyclic

disease, assuming realistically small costs of *RES* and *avr*. Mutation is included in this model because stochastic processes (mutation and drift) have been shown to play an important role in multilocus GFG coevolution (FRANK 1993a; SASAKI 2000; SALATHE *et al.* 2005). We show that polymorphism can be maintained at several host and parasite loci when the autoinfection rate is high. Moreover, compared to a single-locus GFG relationship, multilocus interactions diminish the minimum constitutive cost of each *RES* and *avr* allele necessary for polymorphism to be maintained.

A complication when moving from a single-locus to a multilocus model is the possible existence of interactions between loci. Functional studies of *avr* or *RES* alleles show increasing experimental evidence of epistatic effects between loci in parasites (BAI *et al.* 2000; WICHMANN and BERGELSON 2004; KAY *et al.* 2005; MUDGETT 2005). In microbial parasites of plants, *AVR* proteins have a dual role: as well as being triggers for induction of host defenses upon recognition by *RES* proteins, some of them at least are pathogenicity effectors (DANGL and JONES 2001; ALFANO and COLLMER 2004; SKAMNIOTI and RIDOUT 2005; RIDOUT *et al.* 2006). There may be partial complementation between *AVR* genes (BAI *et al.* 2000; WICHMANN and BERGELSON 2004; KAY *et al.* 2005; MUDGETT 2005; SKAMNIOTI and RIDOUT 2005), so increasing the number of mutations of *AVR* genes to *avr* alleles that have lost pathogenicity effector activity may have a synergistically negative effect on parasite fitness. In plants, on the other hand, *RES* genes induce the expression of similar defense processes (BROWN 2003). Significant costs might therefore arise when one *RES* gene is expressed (TIAN *et al.* 2003), but expression of numerous genes would not necessarily increase cost of defense very much (BERGELSON and PURRINGTON 1996). The models analyzed here incorporate functions to describe epistasis between the costs of multiple *RES* and *avr* alleles, depending on their number. For example, the marginal cost of adding a single new *RES* allele may decrease as the number of existing *RES* alleles increases, while the marginal cost of adding an *avr* allele may increase with the number of existing *avr* alleles. These epistatic cost functions are shown here to increase the parameter space in which stable polymorphism can be maintained in both host and parasite. This supports the hypothesis that multilocus GFG systems favor the maintenance of polymorphism at individual loci with the assumption of realistically small costs of *RES* and *avr* alleles.

#### TWO-LOCUS GFG MODEL WITH MONOCYCLIC DISEASE (MODEL A)

**The model:** Model A describes a GFG system for two interacting loci in the host and the parasite. Both organisms reproduce clonally. Two alleles, *RES* and *res* in the host and *AVR* and *avr* in the parasite, are present

**TABLE 1**  
**Host and parasite fitnesses for monocyclic disease and two loci in each species interacting by gene-for-gene relationships**

	Parasite fitness				Host fitness			
	<i>AA</i>	<i>Aa</i>	<i>aA</i>	<i>aa</i>	<i>AA</i>	<i>Aa</i>	<i>aA</i>	<i>aa</i>
<i>RR</i>	0	0	0	1 - <i>b</i> <sub>2</sub>	1 - <i>u</i> <sub>2</sub>	1 - <i>u</i> <sub>2</sub>	1 - <i>u</i> <sub>2</sub>	(1 - <i>s</i> )(1 - <i>u</i> <sub>2</sub> )
<i>Rr</i>	0	0	1 - <i>b</i> <sub>1</sub>	1 - <i>b</i> <sub>2</sub>	1 - <i>u</i> <sub>1</sub>	1 - <i>u</i> <sub>1</sub>	(1 - <i>s</i> )(1 - <i>u</i> <sub>1</sub> )	(1 - <i>s</i> )(1 - <i>u</i> <sub>1</sub> )
<i>rR</i>	0	1 - <i>b</i> <sub>1</sub>	0	1 - <i>b</i> <sub>2</sub>	1 - <i>u</i> <sub>1</sub>	(1 - <i>s</i> )(1 - <i>u</i> <sub>1</sub> )	1 - <i>u</i> <sub>1</sub>	(1 - <i>s</i> )(1 - <i>u</i> <sub>1</sub> )
<i>rr</i>	1	1 - <i>b</i> <sub>1</sub>	1 - <i>b</i> <sub>1</sub>	1 - <i>b</i> <sub>2</sub>	1 - <i>s</i>	1 - <i>s</i>	1 - <i>s</i>	1 - <i>s</i>

at each locus and are coded 1 and 0, respectively (FRANK 1993a; SASAKI 2000; THRALL and BURDON 2002). For example, a plant genotype with a *RES* allele at the first locus and a *res* allele at the second locus is described as *10*, as is the parasite genotype with *AVR* at the first locus and *avr* at the second locus. An incompatible interaction occurs when the host *RES* allele matches the parasite *AVR* allele at least at one of the two interacting loci (1 matching at one or more loci; FRANK 1993a). Following a common assumption of GFG relationships, an incompatible interaction results in the parasite being unable to infect the host successfully (Table 1; DANGL and JONES 2001). This model is based on TELLIER and BROWN (2007) and is a GFG system for monocyclic disease slightly simplified from SEGARRA (2005). Fitnesses of host and parasite genotypes are given in Table 1, with the following parameters: *s* is the cost to a plant of being diseased, *u*<sub>1</sub> (*u*<sub>2</sub>) is the cost of one (or two) *RES* alleles, and *b*<sub>1</sub> (*b*<sub>2</sub>) is the cost of one (or two) *avr* alleles. For instance, in Table 1, the fitness of a *10* parasite is 0 on plant genotypes *11* and *10* but (1 - *b*<sub>1</sub>) on plant genotypes *00* and *01*. Recurrence equations for the frequencies of the genotypes are given in the APPENDIX.

**Existence of equilibrium points:** Each genotype has an equilibrium frequency; for example, that of host genotype *11* is defined as  $\widehat{RR} = RR_g = RR_{g+1}$ . There are trivial equilibria defined by fixation of one or two host or parasite genotypes. Host equilibria are thus fixation of double-*RES* plants ( $\widehat{RR} = 1$ ), fixation of double-susceptibility ( $\widehat{rr} = 1$ ), and fixation of both single-resistant genotypes ( $\widehat{Rr} + \widehat{rR} = 1$ ). Similar conditions are found for parasite genotypes:  $\widehat{AA} = 1$ ,  $\widehat{aa} = 1$ , and  $\widehat{Aa} + \widehat{aA} = 1$ . The conditions for stability of these trivial equilibria are given in SASAKI (2000) and SEGARRA (2005).

The main point of interest here is the existence of multilocus polymorphism, defined as an equilibrium state with three or four host and parasite genotypes, as can be found in natural populations (THRALL *et al.* 2001; LAINE 2004). This occurs at the nontrivial equilibrium, where host and parasite genotype frequencies are as follows:

$$\widehat{RR} = \frac{b_2 + b_1 b_2 - 2b_1}{1 - b_1}, \quad \widehat{rR} = \widehat{Rr} = \frac{b_1(1 - b_2)}{1 - b_1}, \quad \widehat{rr} = 1 - b_2 \tag{1}$$

$$\widehat{AA} = \frac{(1 - s)(2u_1 - u_1 u_2 - u_2)}{s(1 - u_1)(1 - u_2)}, \quad \widehat{Aa} = \widehat{aA} = \frac{(1 - s)(u_2 - u_1)}{s(1 - u_1)(1 - u_2)}, \tag{2}$$

$$\widehat{aa} = \frac{s - u_2}{s(1 - u_2)}$$

(analysis of recurrence equations is in the APPENDIX with Mathematica 5.0; WOLFRAM RESEARCH 2003). Note that the equilibrium frequencies of host genotypes depend on costs of *avr* alleles (*b*<sub>2</sub> and *b*<sub>1</sub>) while parasite equilibrium frequencies are functions of the costs of *RES* alleles (*u*<sub>2</sub> and *u*<sub>1</sub>) and the cost of disease (*s*) (FRANK 1992).

In Equation 1 the conditions for all host genotypes to exist simultaneously are: 0 <  $\widehat{rr}$  < 1 (this is always true because 0 < *b*<sub>2</sub> < 1); 0 <  $\widehat{rR}$ ,  $\widehat{Rr}$  < 1  $\Rightarrow$  *b*<sub>2</sub> < 1 and *b*<sub>1</sub> < 1/(2 - *b*<sub>2</sub>), which is reasonable, as costs of virulence tend to be small (BERGELSON and PURRINGTON 1996; BROWN 2003); and

$$0 < \widehat{RR} < 1; \widehat{RR} > 0 \Leftrightarrow b_2 > \frac{2b_1}{(b_1 + 1)}. \tag{3}$$

Previous GFG models have assumed multiplicative costs of two single *avr* alleles where *b*<sub>2</sub> = 1 - (1 - *b*<sub>1</sub>)<sup>2</sup> (SEGARRA 2005). Condition (3) is fulfilled if the cost of having two *avr* alleles (*b*<sub>2</sub>) is greater than or equal to the multiplicative cost of two single *avr* alleles because 1 - (1 - *b*<sub>1</sub>)<sup>2</sup> > 2*b*<sub>1</sub>/(*b*<sub>1</sub> + 1).

In Equation 2 the conditions for all parasite genotypes to exist simultaneously are: 0 <  $\widehat{aa}$  < 1  $\Rightarrow$  *u*<sub>2</sub> < *s* < 1 (otherwise virulent parasites are eliminated from the population); 0 <  $\widehat{Aa}$ ,  $\widehat{aA}$  < 1  $\Rightarrow$  *u*<sub>2</sub> > *u*<sub>1</sub> [*i.e.*, the cost of having two *RES* alleles must be larger than that of one *RES* allele (similarly  $\widehat{aA} > 0 \Rightarrow u_2 > u_1$ )]; and

$$0 < \widehat{AA} < 1; \widehat{AA} > 0 \Leftrightarrow u_2 < \frac{2u_1}{(u_1 + 1)}. \tag{4}$$

In previous GFG models, the costs of having two *RES* alleles have been multiplicative: *u*<sub>2</sub> = 1 - (1 - *u*<sub>1</sub>)<sup>2</sup>. Condition (4) is not satisfied if the cost of having two *RES* alleles (*u*<sub>2</sub>) is equal to the multiplicative cost of two single *RES* alleles because 1 - (1 - *u*<sub>1</sub>)<sup>2</sup> > 2*u*<sub>1</sub>/(*u*<sub>1</sub> + 1).

Equation 4 shows that when multiplicative costs are assumed (as in SASAKI 2000; SALATHE *et al.* 2005; SEGARRA 2005), double-*AVR* parasites cannot be main-

tained in populations. Epistasis of fitness costs of *RES* and *avr* alleles is thus essential for the existence of an interior equilibrium point with all four host and all four parasite genotypes.

**Stability of the equilibrium point:** Following analysis in TELLIER and BROWN (2007), we use a logit transformation of genotype frequencies in model A (APPENDIX and supplemental Section 1 at <http://www.genetics.org/supplemental/>), which simplifies considerably the analysis of the genotype dynamics. At host generation  $g$ :

$$f_{Aa} = \log \frac{Aa_g}{aA_g}, \quad f_{aa} = \log \frac{aa_g}{AA_g}, \quad f_{rR} = \log \frac{rR_g}{Rr_g},$$

$$\text{and } f_{rr} = \log \frac{rr_g}{RR_g}. \quad (5)$$

The change ( $\Delta$ ) in the ratio of parasite genotype *10* and *01* frequencies between generation  $g$  and  $g + 1$  is then

$$\Delta f_{Aa} = \log \frac{Aa_{g+1}}{aA_{g+1}} - \log \frac{Aa_g}{aA_g}. \quad (6)$$

Thus the system of equations of model A (see APPENDIX and supplemental Section 1) can be rewritten as

$$\begin{pmatrix} \Delta f_{Aa} \\ \Delta f_{aa} \\ \Delta f_{rR} \\ \Delta f_{rr} \end{pmatrix} = J_A \begin{pmatrix} f_{Aa} \\ f_{aa} \\ f_{rR} \\ f_{rr} \end{pmatrix}, \quad (7)$$

where  $J_A$  is the Jacobian matrix of the system. The dynamics of the system are determined by analysis of the eigenvalues of  $J_A$  (see APPENDIX). For a model with four variables, two pairs of eigenvalues ( $\lambda_{1,2}$  and  $\lambda_{3,4}$ ) are solutions of the characteristic polynomial equation of  $J_A$ . The pairs of eigenvalues can be real  $\lambda_{1,2} = \alpha_1 \pm \sqrt{\beta_1}$  (and  $\lambda_{3,4} = \alpha_2 \pm \sqrt{\beta_2}$ ) or complex  $\lambda_{1,2} = \alpha_1 \pm i\sqrt{\beta_1}$  (and  $\lambda_{3,4} = \alpha_2 \pm i\sqrt{\beta_2}$ ), with

$$\alpha_1 = \frac{\partial \Delta f_{Aa}}{\partial f_{Aa}} + \frac{\partial \Delta f_{rR}}{\partial f_{rR}} \quad \text{and} \quad \alpha_2 = \frac{\partial \Delta f_{aa}}{\partial f_{aa}} + \frac{\partial \Delta f_{rr}}{\partial f_{rr}} \quad (8)$$

(see APPENDIX). An exact condition for stability of an interior equilibrium of this dynamical system with four variables (Equation 7) is that the four eigenvalues of  $J_A$  must lie within a unit circle centered on  $(-1, 0)$  in the complex plane (ROUGHGARDEN 1996; KOT 2001). The following condition is derived from the Routh–Hurwitz criterion for stability of a dynamical system:

$$-1 < \alpha_1 < 0 \quad \text{and} \quad -1 < \beta_1 < 1 \quad \text{and} \quad -1 < \alpha_2 < 0$$

$$\text{and} \quad -1 < \beta_2 < 1 \quad (9)$$

(ROUGHGARDEN 1996; KOT 2001). For model A, the eigenvalues are

$$\begin{cases} \lambda_{1,2} = \pm i\sqrt{-\beta_1} \\ \lambda_{3,4} = \pm \sqrt{\beta_2} \end{cases} \text{ if } \beta_1 < 0 \quad \text{or} \quad \begin{cases} \lambda_{1,2} = \pm \sqrt{\beta_1} \\ \lambda_{3,4} = \pm \sqrt{\beta_2} \end{cases} \text{ if } \beta_1 > 0 \quad (10)$$

with  $\beta_1 = (b_2 - 2b_1(1 - b_2))(u_2 - u_1(2 - u_2))(s - u_2)/s(1 - b_1)(1 - u_1)(1 - u_2)$  and  $\beta_2 = (1 - s)(1 - b_2)b_1^2(u_1 - u_2)^2/s(1 - b_1)(1 - u_2)$ . Consequently, there is always at least one eigenvalue that does not verify condition (9), and the interior, nontrivial equilibrium (Equations 1 and 2) is always unstable. The mathematical reason for this is that all diagonal elements of  $J_A$  are zero, and therefore  $\alpha_1 = \alpha_2 = 0$  (Equation 8, APPENDIX, and supplemental Section 1 at <http://www.genetics.org/supplemental/>). The elements of  $J_A$  are the rates of natural selection on the ratio of genotype frequencies. For example,  $\partial \Delta f_{aa} / \partial f_{aa}$  is the rate of selection on the double-*avr* parasite genotype (*00*) as a function of its own frequency, *i.e.*, the rate of direct FDS on *00* parasites (TELLIER and BROWN 2007). Therefore,  $\alpha_1$  and  $\alpha_2$  are the sums of the direct FDS coefficients for the four ratios of genotype frequency (Equation 8). Equation 10 demonstrates that for monocyclic diseases, there is no direct negative FDS for host or parasite genotypes, and a polymorphic state with three or four host and parasite genotypes is always unstable.

#### TWO-LOCUS GFG MODEL WITH POLYCYCLIC DISEASE

**Model description:** Model B is a multilocus GFG system with polycyclic disease, where polycyclic pathogens undergo several ( $G$ ) multiplicative generations during one host generation. Here, the simplest case of  $G = 2$  parasite generations per host generation is considered. The autoinfection rate ( $\psi$ ) is the percentage of infectious spores that reinfect the same host plant in the second parasite generation (BARRETT 1980; TELLIER and BROWN 2007). The cost to a plant of being diseased increases with the number of successive parasite infections, with a maximum fitness loss of  $\phi$  after  $G$  parasite generations (CAMPBELL and MADDEN 1990; TELLIER and BROWN 2007). The loss of plant reproductive output caused by disease increases disproportionately with  $\pi$ , the number of successful parasite generations on a host plant ( $\pi \leq G$ ) because, as the parasite grows multiplicatively, corresponding damage is done to the host (CAMPBELL and MADDEN 1990). The plant fitness ( $F$ ) is a decreasing function of  $\pi$  where  $z$  is a parameter defining the shape of the disease curve ( $z > 1$ ):

$$F = 1 - \phi \left( \frac{\pi}{G} \right)^z \quad (11)$$

(TELLIER and BROWN 2007).  $\varepsilon$  is the decrease of plant fitness after  $\pi = 1$  infection ( $\varepsilon = \phi(1/2)^z$ ; Equation 11,  $G = 2$ ,  $\pi = 1$ ). For simplicity, the parasite reproductive fitness does not depend on  $\pi$ . Deterministic equations for evolution of genotype frequencies in time are given below and can be obtained from fitnesses given in supplemental Tables S1 and S2 (supplemental Section 2 at <http://www.genetics.org/supplemental/>). Table 2 is a summary of equations in model B with only autoinfection ( $\psi = 1$ ).

TABLE 2

Fitness of hosts and parasites in model B for interactions with two parasite generations per host generation ( $G = 2$ ) and only autoinfection between parasite generations ( $\psi = 1$ )

Parasite genotypes (frequencies) within host generation $g$		Fitness at beginning of host generation $g + 1$	
First generation	Second generation	Fitness of second parasite infection	Host fitness
Host genotype $11 (RR_g)$			
$11 (AA_g)$	$11 (AA_1)$	0	$1 - u_2$
$10 (Aa_g)$	$10 (Aa_1)$	0	$1 - u_2$
$01 (aA_g)$	$01 (aA_1)$	0	$1 - u_2$
	$00 (aa_1)$	$1 - b_2$	$(1 - u_2)(1 - \epsilon)$
$00 (aa_g)$	$00 (aa_g)$	$1 - b_2$	$(1 - u_2)(1 - \phi)$
Host genotype $00 (rr_g)$			
$00 (aa_g)$	$00 (aa_g)$	$1 - b_2$	$(1 - \phi)$
$01 (aA_g)$	$10 (Aa_g)$	$1 - b_1$	$(1 - \phi)$
$10 (Aa_g)$	$10 (Aa_g)$	$1 - b_1$	$(1 - \phi)$
$11 (AA_g)$	$11 (AA_g)$	1	$(1 - \phi)$
Host genotype $01 (rR_g)$			
$11 (AA_g)$	$11 (AA_1)$	0	$1 - u_1$
$01 (aA_g)$	$01 (aA_1)$	0	$1 - u_1$
	$00 (aa_1)$	$1 - b_2$	$(1 - u_1)(1 - \epsilon)$
	$10 (Aa_1)$	$1 - b_1$	$(1 - u_1)(1 - \epsilon)$
$00 (aa_g)$	$00 (aa_g)$	$1 - b_2$	$(1 - u_1)(1 - \phi)$
$10 (Aa_g)$	$10 (Aa_g)$	$1 - b_1$	$(1 - u_1)(1 - \phi)$
Host genotype $10 (Rr_g)$			
$11 (AA_g)$	$11 (AA_1)$	0	$1 - u_1$
$10 (Aa_g)$	$10 (Aa_1)$	0	$1 - u_1$
	$00 (aa_1)$	$1 - b_2$	$(1 - u_1)(1 - \epsilon)$
	$01 (aA_1)$	$1 - b_1$	$(1 - u_1)(1 - \epsilon)$
$00 (aa_g)$	$00 (aa_g)$	$1 - b_2$	$(1 - u_1)(1 - \phi)$
$01 (aA_g)$	$10 (Aa_g)$	$1 - b_1$	$(1 - u_1)(1 - \phi)$

As an example, the outcome of infection on plant genotype  $11$  is described below (supplemental Table S1 in supplemental Section 2 at <http://www.genetics.org/supplemental/>). In the first parasite generation ( $\pi = 1$ ), a  $11$  plant can encounter parasite genotypes  $11$ ,  $01$ , or  $10$  that cannot infect it successfully. In the second parasite generation ( $\pi = 2$ ), that plant can then either (i) encounter spores from the same genotypes (frequencies  $AA_1$ ,  $Aa_1$ ,  $aA_1$ ) that cannot infect it (the plant's fitness is then  $1 - u_2$ ) or (ii) be infected by the supervirulent genotype  $00$  (frequency  $aa_1$ ), so its fitness is  $(1 - u_2)(1 - \epsilon)$ .

On the other hand, when  $11$  plants are infected by a supervirulent ( $00$ ) parasite at  $\pi = 1$  (frequency  $aa_g$  at the start of generation  $g$ ) the following occurs at  $\pi = 2$ :

- i. A proportion  $\psi$  of these plants remain infected by the same parasite genotype (autoinfection). A proportion  $\psi aa_g RR_g$  of all the plants in the population has fitness  $(1 - u_2)(1 - \phi)$  after two consecutive successful infections.

- ii. A proportion  $1 - \psi$  are allo-infected by virulent parasites (here, only those with the supervirulent genotype  $00$ ) produced in the first parasite generation with frequency  $aa_1$  [proportion  $(1 - \psi)aa_gaa_1RR_g$ ]. These plants are also infected twice and have fitness  $(1 - u_2)(1 - \phi)$ .
- iii. A proportion  $1 - \psi$  may encounter spores from the first parasite generation of the genotypes  $01$ ,  $10$ , or  $11$  (frequencies  $AA_1$ ,  $Aa_1$ , and  $aA_1$ ) that cannot infect. Their fitness is  $(1 - u_2)(1 - \epsilon)$ .

**Formulas:** The following are deterministic equations for a two-locus GFG system with two parasite generations per host generation with independent fitness costs of host resistance or parasite virulence alleles at different loci and no mutation. Frequencies of parasite genotypes after the first parasite generation are identical to those in model A (see Table 2 and supplemental Section 1 at <http://www.genetics.org/supplemental/>):

$$\frac{aa_1}{AA_1} = \frac{aa_g(1 - b_2)}{AA_g rr_g}$$

$$\frac{Aa_1}{aA_1} = \frac{Aa_g(1 - b_1)(rr_g + rR_g)}{aA_g(1 - b_1)(rr_g + Rr_g)}$$

After the second parasite generation (*i.e.*, at the start of the next host generation,  $g + 1$ ), parasite genotype frequencies are as follows:

Ratio of parasite genotype frequencies  $00$  to  $11$ :

$$\frac{aa_{g+1}}{AA_{g+1}} = \frac{(1 - b_2) \left( \begin{aligned} &RR_g[aa_1(1 - aa_g) + aa_g\psi(1 - b_2) + aa_g(1 - \psi)aa_1] \\ &+ rr_g[aa_g\psi(1 - b_2) + aa_1(1 - \psi)] \\ &+ Rr_g[aa_1(AA_g + Aa_g) + aA_g(1 - \psi)aa_1 \\ &+ aa_g\psi(1 - b_2) + aa_g(1 - \psi)aa_1] \\ &+ rR_g[aa_1(AA_g + aA_g) + Aa_g(1 - \psi)aa_1 \\ &+ aa_g\psi(1 - b_2) + aa_g(1 - \psi)aa_1] \end{aligned} \right)}{rr_g(\psi AA_g + (1 - \psi)AA_1)}$$

Ratio of parasite genotype frequencies  $10$  to  $01$ :

$$\frac{Aa_{g+1}}{aA_{g+1}} = \frac{(1 - b_1) \left( \begin{aligned} &rR_g[Aa_1(AA_g + aA_g) + Aa_g(\psi(1 - b_1) \\ &+ Aa_1(1 - \psi)) + aa_g(Aa_1(1 - \psi))] \end{aligned} \right)}{(1 - b_1) \left( \begin{aligned} &Rr_g[aA_1(AA_g + Aa_g) + aA_g(\psi(1 - b_1) + aA_1(1 - \psi))] \\ &+ aa_g(aA_1(1 - \psi)) \\ &+ rr_g[\psi aA_g(1 - b_1) + aA_1(1 - \psi)] \end{aligned} \right)}$$

The ratios of host genotype frequencies at the end of host generation  $g$  and the start of generation  $g + 1$  are as follows:

Ratio of host genotype frequencies  $01$  to  $10$ :

$$\frac{rR_{g+1}}{rR_{g+1}} = \frac{rR_g(1 - u_1) \left( \begin{aligned} &(AA_1 + aA_1)(AA_g + aA_g) + (1 - \epsilon) \\ &\times [(AA_g + aA_g)(Aa_1 + aa_1) \\ &+ (1 - \psi)(Aa_g + aa_g)(AA_1 + aA_1)] \\ &+ (1 - \phi)[Aa_g\psi + Aa_g(1 - \psi)(Aa_1 + aa_1) \\ &+ aa_g\psi + aa_g(1 - \psi)(aa_1 + Aa_1)] \end{aligned} \right)}{Rr_g(1 - u_1) \left( \begin{aligned} &(AA_1 + Aa_1)(AA_g + Aa_g) + (1 - \epsilon) \\ &\times [(AA_g + Aa_g)(aA_1 + aa_1) \\ &+ (1 - \psi)(aA_g + aa_g)(AA_1 + Aa_1)] \\ &+ (1 - \phi)[aA_g\psi + aA_g(1 - \psi)(aA_1 + aa_1) \\ &+ aa_g\psi + aa_g(1 - \psi)(aa_1 + aA_1)] \end{aligned} \right)}$$

Ratio of host genotype frequencies  $00$  to  $11$ :

$$\frac{r_{g+1}}{RR_{g+1}} = \frac{r_g(1-\phi)}{RR_g(1-u_2) \left( \frac{(1-\varepsilon)[aa_1(1-aa_g) + aa_g(1-\psi)(1-aa_1)]}{(1-aa_g)(1-aa_1) + (1-\phi)[aa_g(\psi + aa_1(1-\psi))]} \right)}$$

**Existence of equilibrium point ( $\psi = 1$ ):** The same trivial equilibrium points exist as in model A. The complexity of the equations in model B constrains analysis of the nontrivial equilibrium point to the case when there is only autoinfection ( $\psi = 1$ ). Owing to quadratic terms and the nonlinear behavior of model B, the following assumptions were made: (i)  $\widehat{aa}_1 = \widehat{aa}(1-b_2)/(1-b_2\widehat{aa})$  and (ii)  $b_1$  is small so that  $(1-b_1) \approx 1$  and  $(1-b_1)(1-b_2) \approx (1-b_2)$ . The accuracy of these approximations and of the following equilibrium frequencies was tested numerically across a wide range of parameter values. Theoretical values obtained with Mathematica 5.0 (WOLFRAM RESEARCH 2003) were compared to numerically simulated values, calculated as the mean of genotype frequencies over the last 100 generations of 5000 simulated host generations. These approximations (Equations 12 and 14) are accurate for moderate to high values of  $\phi$ , but less accurate for values of  $\widehat{r}$  when  $\phi < u_1$  or values of  $\widehat{AA}$  when  $u_1 < \phi < u_2$  (supplemental Section 3 at <http://www.genetics.org/supplemental/>), as

$$\widehat{aa} = \frac{\omega - \sqrt{\omega^2 - 4\varepsilon(\phi - u_2)}}{2(1-u_2)(\varepsilon + b_2(\phi - \varepsilon))} \tag{12}$$

$$\widehat{aA} = \widehat{Aa} = \frac{\omega - \sqrt{\omega^2 - 4(u_2 - u_1)(\phi - u_2)(1 + \phi)}}{2(1-u_1)(1-u_2)(\varepsilon + b_2(\phi - \varepsilon))}$$

$$\widehat{AA} = \frac{\omega - \sqrt{\omega^2 - 4\varepsilon(1 + \phi)(2u_1 + u_1u_2 - u_2)}}{2(1-u_1)(1-u_2)(\varepsilon + b_2(\phi - \varepsilon))},$$

where  $\omega = (1-b_2)(\phi + \varepsilon) + b_2(\phi - \varepsilon - u_2(1-\varepsilon))$ . In Equation 12 the conditions for all parasite genotypes to exist simultaneously are  $\widehat{aa} > 0 \Leftrightarrow \phi > u_2$  and  $\widehat{Aa} > 0 \Leftrightarrow u_2 > u_1$  (similarly for  $\widehat{aA} > 0$ ):

$$\widehat{AA} > 0 \Leftrightarrow u_2 < \frac{2u_1}{(1-u_1)}. \tag{13}$$

In contrast to model A (Equation 4), condition (13) is fulfilled if the cost of having two *RES* alleles ( $u_2$ ) is lower than or equal to the multiplicative cost of two single *RES* alleles. The equilibrium frequency of double-*AVR* parasites increases when the difference between  $u_2$  and  $u_1$  diminishes:

$$\widehat{r} = \frac{(3 - \widehat{aa} - 2b_2\widehat{aa})(1-b_2)^2}{3 - \widehat{aa} - 2b_2} \tag{14}$$

$$\widehat{rR} = \widehat{Rr} = \frac{b_1(3 - \widehat{aa})(1-b_2)^2}{(1-b_1)(3 - \widehat{aa} - 2b_2)}$$

$$\widehat{RR} = \frac{(b_2 - 2b_1 + b_1b_2)(1-b_2)^2}{3 - \widehat{aa} - 2b_2}.$$

In Equation 14 the conditions for all host genotypes to exist simultaneously are  $\widehat{r} > 0$ ,  $\widehat{rR} > 0$ , and  $\widehat{RR} > 0$  because  $3 > \widehat{aa} + 2b_2\widehat{aa}$  and  $3 > \widehat{aa} + 2b_2$  by definition:

$$\widehat{RR} > 0 \Leftrightarrow b_2 > \frac{2b_1}{(1+b_1)}. \tag{15}$$

Condition (15) is identical to Equation 3 and is fulfilled if the cost of having two *avr* alleles ( $b_2$ ) is greater than or equal to the multiplicative cost of two single *avr* alleles. The equilibrium frequency of double-*RES* plants becomes higher with a greater difference between  $b_2$  and  $b_1$ .

**Stability of the equilibrium point:** The local stability of the nontrivial equilibrium is analyzed when there is only auto-infection ( $\psi = 1$ ) because in this situation, stable polymorphism occurs over a wider parameter space (TELLIER and BROWN 2007). The main differences from model A are that the following coefficients are not zero (supplemental Section 4 at <http://www.genetics.org/supplemental/>):

$$\frac{\partial \Delta f_{Aa}}{\partial f_{Aa}} = \frac{-Rr_g Aa_g aA_g}{1 - b_2aa_g + Rr_g(AA_g + Aa_g)} = x_5. \tag{16}$$

This coefficient is always negative as  $1 + Rr_g(AA_g + Aa_g) > -b_2aa_g$ :

$$\frac{\partial \Delta f_{aa}}{\partial f_{aa}} = \frac{-(1-r_g)(1-b_2)aa_gAA_g}{(1-b_2aa_g)(2-2b_2aa_g + (1-r_g)(1-aa_g))} = x_6. \tag{17}$$

This coefficient is also negative as  $2 + (1-r_g)(1-aa_g) > 2b_2aa_g$ .

The Jacobian matrix for model B,  $J_B$ , can thus be rewritten:

$$J_B = \begin{pmatrix} x_5 & 0 & x_3 & 0 \\ 0 & x_6 & 0 & x_4 \\ x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \end{pmatrix}. \tag{18}$$

Approximations for the elements  $x_1, x_2, x_3, x_4$  are derived in supplemental Section 4 at <http://www.genetics.org/supplemental/>.  $J_B$  is diagonalizable and has two pairs of eigenvalues ( $\lambda_{1,2}$  and  $\lambda_{3,4}$ ) that can be real,

$$\lambda_{1,2} = \frac{1}{2} \left( x_5 \pm \sqrt{x_5^2 + 4x_1x_3} \right) \text{ if } x_5^2 + 4x_1x_3 > 0 \tag{19}$$

and

$$\lambda_{3,4} = \frac{1}{2} \left( x_6 \pm \sqrt{x_6^2 + 4x_2x_4} \right) \quad \text{if } x_6^2 + 4x_2x_4 > 0,$$

or complex,

$$\lambda_{1,2} = \frac{1}{2} \left( x_5 \pm i\sqrt{-x_5^2 - 4x_1x_3} \right) \quad \text{if } x_5^2 + 4x_1x_3 < 0 \quad (20)$$

and

$$\lambda_{3,4} = \frac{1}{2} \left( x_6 \pm i\sqrt{-x_6^2 - 4x_2x_4} \right) \quad \text{if } x_6^2 + 4x_2x_4 < 0.$$

A necessary condition for stability (Equation 9) is verified because  $x_5 = \partial\Delta f_{Aa}/\partial f_{Aa}$  and  $x_6 = \partial\Delta f_{aa}/\partial f_{aa}$  are both negative. However, a second condition for stability is that both absolute values of the discriminants of the characteristic polynomial,  $|x_5^2 + 4x_1x_3|$  and  $|x_6^2 + 4x_2x_4|$ , must be  $<1$  (Equation 9). Analytical derivation of this second condition is not possible because of the nonlinearity of equations in model B and because only approximations of the equilibrium genotype frequencies can be obtained.

**Simulation methods:** Host and parasite genotype frequencies were therefore simulated numerically for different values of  $\psi$  and  $\phi$ . Simulations were run in Matlab version 7.0 (Release 14) for 15,000 host generations, by which time stable behavior (or genotype fixation) was achieved, with different sets of initial host and genotype frequencies (all host and parasite genotypes were present at the beginning of each simulation). The system was considered to be stable when the amplitude of the fluctuations of each genotype frequency decreased in time and converged toward an equilibrium value for any of the initial allele frequencies tested.

Mutations, especially with high mutation rates, regularly introduce new rare genotypes into host and parasite populations (KIRBY and BURDON 1997; SASAKI 2000; SALATHE *et al.* 2005). We therefore compared results of simulations with and without mutation. One set of simulations was done with a mutation rate of  $10^{-5}$ . A second set of simulations assumed two different mutation rates:  $10^{-5}$  if a mutation results in a loss of function (from *RES* to *res* and from *AVR* to *avr*) and  $10^{-8}$  for a gain-of-function mutation in the reverse direction (from *res* to *RES* or from *avr* to *AVR*) (KIRBY and BURDON 1997). A host or parasite genotype was considered lost from a population when its frequency was  $<10^{-6}$ , but could be subsequently reintroduced by mutation (if any). On the other hand, a genotype was fixed in a population when its frequency was  $>1 - 10^{-6}$ . For a given set of parameter values, the results of the different types of simulations (with and without mutations or with different initial genotype frequencies) were compared. The description of results follows with models B1 and B2.

#### MULTIPLICATIVE CONSTITUTIVE COSTS (MODEL B1)

Model B1 is a two-locus GFG system with multiplicative costs of *RES* and *avr* alleles, *i.e.*, no epistatic interactions among loci for fitness values (FRANK 1993b; SASAKI 2000; SALATHE *et al.* 2005; SEGARRA 2005). Recurrence equations for genotype frequencies are those of model B, where  $b_2$  and  $u_2$  are the costs of having two *RES* or *avr* alleles:

$$b_2 = 1 - (1 - b_1)^2 \quad \text{and} \quad u_2 = 1 - (1 - u_1)^2. \quad (21)$$

Simulations were run with  $b_1 = u_1 = 5\%$  and  $b_2 = u_2 = 9.75\%$ , these values being chosen to allow comparison with single-locus results (TELLIER and BROWN 2007). When there is only auto-infection ( $\psi = 1$ ), the double-*RES* genotype has a very low equilibrium frequency ( $<10^{-4}$ , Equation 14). Model B1 is tested numerically to determine if the equilibrium point with all host and parasite genotypes exists for different values of  $\psi$  and to discover the range of parameter values of  $\psi$  and  $\phi$  for which the equilibrium point is stable.

Results of simulations are summarized in Figure 1 by the state of the system (stable or unstable) and the genotypes maintained in the host (Figure 1a) and parasite (Figure 1b) populations. Figure 2a shows the dynamics of host genotype frequencies in one simulation typical of area B of Figure 1a with stable polymorphism of three host genotypes (*00*, *10*, *01*). Similarly, Figure 2b shows the dynamics of parasite gene frequencies in area I of Figure 1b where there is stable polymorphism of all four parasite genotypes, with genotype *11* at a very low frequency. Typical simulation results for each area of Figure 1 are provided in supplemental Section 4 at <http://www.genetics.org/supplemental/>.

If disease severity is smaller than the cost of one *RES* allele, there is fixation of host *res* alleles (*00*) because there is no net advantage to resistance (area A in Figure 1a). As a result, *AVR* alleles (*11*) are fixed in the parasite population because virulence is costly (area G in Figure 1b). In mathematical terms, areas A and G correspond to the situation with zero (or negative) equilibrium frequencies of parasite genotypes *01*, *10*, and *11* because  $\phi < u_1 < u_2$  (Equations 12). The limit of areas A and G is thus the cost of one *RES* allele ( $\phi = u_1$ ).

At medium to high autoinfection ( $\psi$ ) and low to medium disease severity ( $\phi$ ), there is stable polymorphism of three host genotypes (*00*, *10*, *01*) (Figure 1a, area B, and Figure 2a). Double-resistant plants (*11*) are eliminated from the population because the benefit of being superresistant (not being infected) is not large enough to overcome the cost of having two *RES* alleles. As a consequence, when  $u_1 < \phi < u_2$  and autoinfection rates are intermediate to high, because *avr* alleles are costly, double-*avr* parasites are eliminated from the parasite population (area H in Figure 1b). Stable polymorphism

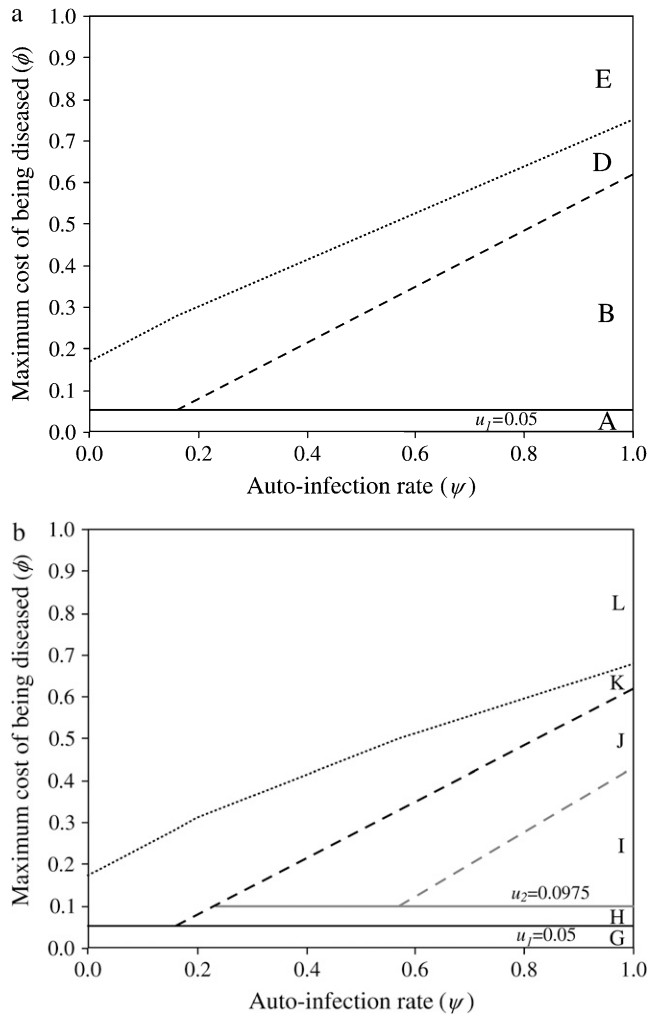


FIGURE 1.—Stability area plots for a two-locus model with independent costs of alleles at different loci (model B1), in relation to values of the autoinfection rate ( $\psi$ ) and the cost to a plant of being diseased by two parasite generations ( $\phi$ ).  $u_1$ ,  $u_2$ , costs of host resistance;  $b_1$ ,  $b_2$ , costs of parasite virulence;  $u_1 = b_1 = 0.05$ ,  $u_2 = b_2 = 0.0975$  (subscript 1, cost of one allele; subscript 2, cost of two alleles). (a) Stability of polymorphism in the host population. Area A, fixation of double-susceptible genotype  $00$ ; area B, stable polymorphism with  $00$  and the single-resistant genotypes  $10$  and  $01$ ; area D, unstable polymorphism with  $00$ ,  $10$ ,  $01$ ; area E, fixation of  $00$ . (b) Stability of polymorphism in the parasite population. Area G, fixation of the double-avirulent genotype  $11$ ; area H, stable polymorphism with  $11$  and the single-avirulent genotypes  $01$  and  $10$ ; area I, stable polymorphism with all four genotypes; area J, stable polymorphism with  $10$ ,  $01$ , and the double-virulent genotype  $00$ ; area K, unstable polymorphism with  $10$ ,  $01$ ,  $00$ ; area L, fixation of  $00$ .

with parasite genotypes  $11$ ,  $10$ , and  $01$  occurs. Mathematically, area H corresponds to a situation where the double-*avr* equilibrium frequency is zero (or negative) as  $\phi < u_2$  (Equation 12).

With increasing  $\phi$  and high autoinfection rates, all four parasite genotypes ( $00$ ,  $01$ ,  $10$ ,  $11$ ) coexist in stable polymorphism (area I, Figure 1b) because there is direct FDS acting on parasite genes (Equations 16 and 17),

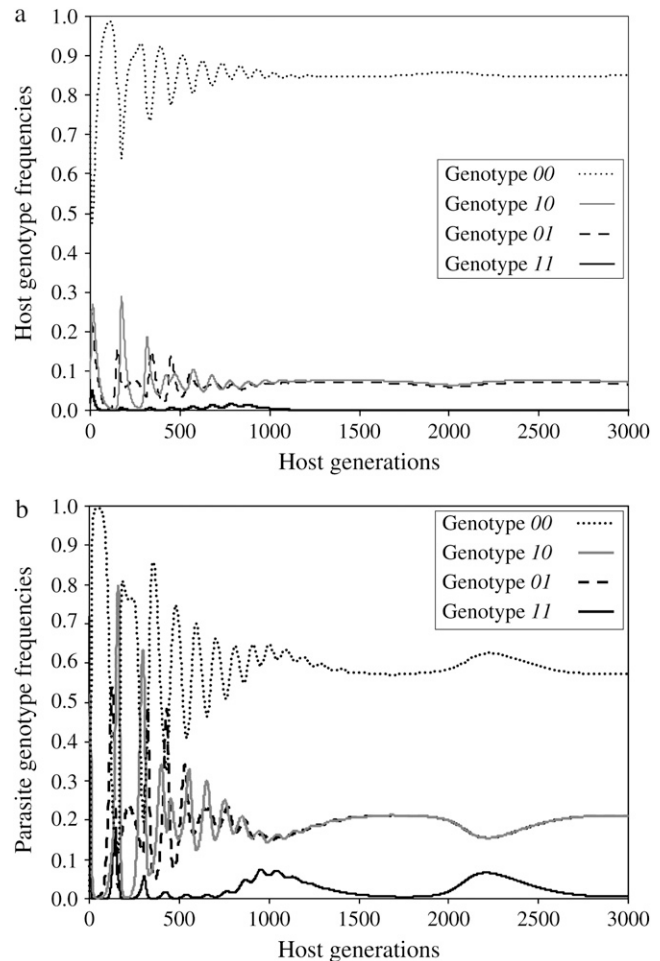


FIGURE 2.—Dynamics of host and parasite genotype frequencies in a two-locus model with independent costs of alleles at different loci (model B1) defined by two loci as a function of the number of host generations. There is no mutation. Autoinfection rate  $\psi = 0.9$  and maximum cost of disease  $\phi = 0.2$ .  $u_1$ ,  $u_2$ , costs of host resistance;  $b_1$ ,  $b_2$ , costs of parasite virulence;  $u_1 = b_1 = 0.05$ ,  $u_2 = b_2 = 0.0975$ . (a) Maintenance of host genotypes with one or both susceptibility alleles but the double-resistant genotype  $11$  is eliminated (area B in Figure 1a). (b) Maintenance of all four parasite genotypes (area I in Figure 1b).

and the interior equilibrium for all possible parasite genotypes exists (Equation 12; Figure 2b). The parameter space in which polymorphism is stable diminishes with increasing  $\phi$ , because resistant genotypes ( $10$  and  $01$ ) are selected more strongly, in turn selecting for double-*avr* parasites ( $00$ ). Therefore, at intermediate to high  $\psi$ , increasing  $\phi$  favors double-*avr* parasites and counterselects double-*AVR* genotypes ( $11$ ) (area J). In area J, values of  $\psi$  and  $\phi$  do not allow the existence of an interior equilibrium frequency for the double-*AVR* parasite (Equation 12).

A key result is that the size of areas H, I, and J in Figure 1b matches that of area B in Figure 1a. This is because stability of polymorphic state depends on the strength of direct FDS against strength of indirect FDS, both of



which are determined by  $\psi$  and  $\phi$ . Therefore, conditions for stability are identical for host and parasite populations, as shown for single-locus interactions (TELLIER and BROWN 2007), and only the existence of equilibrium frequencies of the various genotypes discriminates between the different dynamics in host and parasite populations. Moreover, the equilibrium frequency of the double-*avr* genotype (*00*) increases with  $\phi$  (Equation 12) and always has the highest frequency in the parasite population. This is in agreement with observations from surveys in natural populations (DINOOR and ESHED 1987; BEVAN *et al.* 1993; THRALL *et al.* 2001) (Figure 3b).

When there is a high cost of disease (high  $\phi$ ) (Figure 1, a and b), there is first strong selection for resistant host genotypes *01*, *10*, *11*. They select strongly for virulent parasite genotypes and especially for the double-*avr* genotype (*00*). Very high frequencies of double-*avr* parasite (area K, Figure 1b) then lead to a long-term increase of the double-susceptible genotype frequency (*00*) because *RES* alleles are costly. At very high  $\phi$ , this results in the fixation of the double-*avr* parasite genotype and the double-susceptible host genotype (areas E and L in Figure 1). The dynamical system is unstable when  $\phi$  increases, because the indirect FDS overrides the direct frequency-dependent stabilizing effect (areas D and K).

Realistic mutation rates do not affect the behavior of the model (stability or instability) or frequencies at equilibrium in stable areas (B in Figure 1a and H–J in Figure 1b). Without mutation, when the system is unstable (areas D and E in Figure 1a and K and L in Figure 1b) there is fixation of the double-*avr* parasite genotype. Mutations can sustain stochastic coevolutionary cycles by recurrent introduction of new rare genotypes in areas D and K, following an arms race model. However, as the system has unstable behavior (areas D and E in Figure 1a and K and L in Figure 1b) each coevolutionary cycle results in the fixation of the double-*avr* parasite until a new mutation arises.

EPISTATIC INTERACTIONS AMONG LOCI (MODEL B2)

In model B2 epistatic interactions are assumed between loci both in host and in parasite. General expressions are shown here for the costs of multiple *avr* (or *RES*) alleles in a multilocus GFG system with  $n$  loci. The maximum cost of having  $n$  *avr* (or *RES*) alleles is  $b_{\max}(u_{\max})$ .

The cost  $b_k$  of having  $k$  *avr* alleles is thus

$$b_k = b_{\max} \left( \frac{k}{n} \right)^\theta \quad \text{with } 0 \leq k \leq n. \quad (22)$$

The marginal cost of each additional mutation from *AVR* to *avr* increases exponentially with the number of existing *avr* alleles, such that the loss of two *AVR* func-

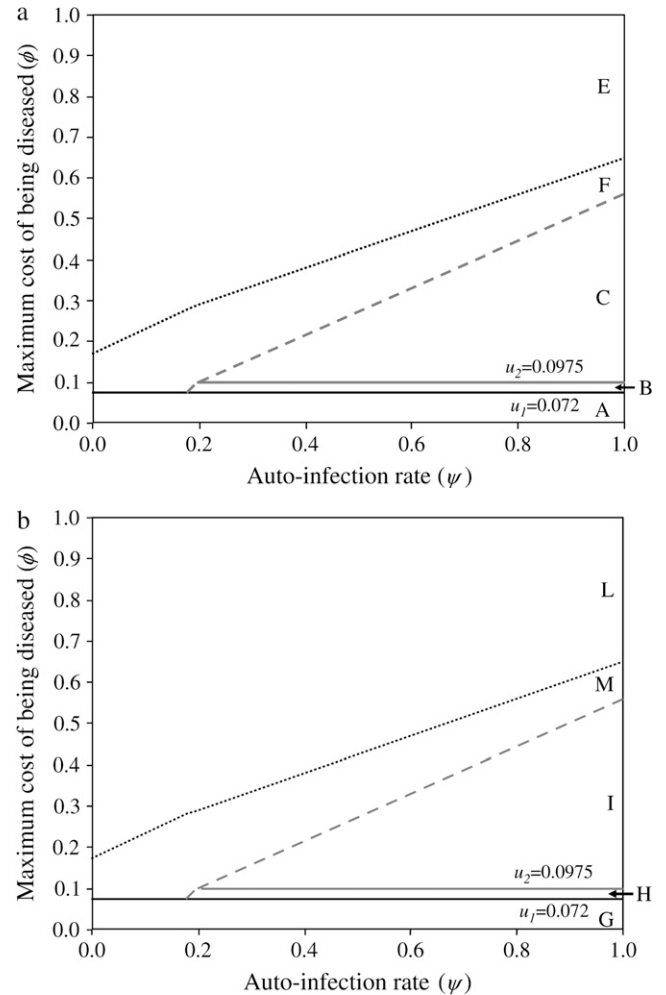


FIGURE 3.—Stability area plots for a two-locus model with epistasis in costs of alleles at different loci (model B2), in relation to values of the autoinfection rate ( $\psi$ ) and the cost to a plant of being diseased by two parasite generations ( $\phi$ ).  $u_1$ ,  $u_2$ , costs of host resistance;  $b_1$ ,  $b_2$ , costs of parasite virulence;  $u_1 = 0.072$  and  $b_1 = 0.024$ ;  $u_2 = b_2 = 0.0975$ . (a) Stability of polymorphism in the host population. Area A, fixation of double-susceptible genotype *00*; area B, stable polymorphism with *00* and the single-resistant genotypes *10* and *01*; area C, stable polymorphism with all four genotypes (*10*, *01*, *00*, *11*); area F, unstable polymorphism with all four genotypes; area E, fixation of *00*. (b) Stability of polymorphism in the parasite population. Area G, fixation of the double-avirulent genotype *11*; area H, stable polymorphism with *11* and the single-avirulent genotypes *01* and *10*; area I, stable polymorphism with all four genotypes (*10*, *01*, *00*, *11*); area M, unstable polymorphism with all four genotypes; area L, fixation of the double-virulent genotype *00*.

tions is more costly to the pathogen than expected if the costs were independent; therefore, we choose  $\theta > 1$ , and the cost curve has a convex shape. In model B2,  $n = 2$  so  $b_{\max} = b_2$ .

On the other hand, the marginal cost of each additional *RES* allele diminishes with increasing number of existing *RES* alleles, so the cost of two *RES* alleles is lower than expected if the costs at different loci were

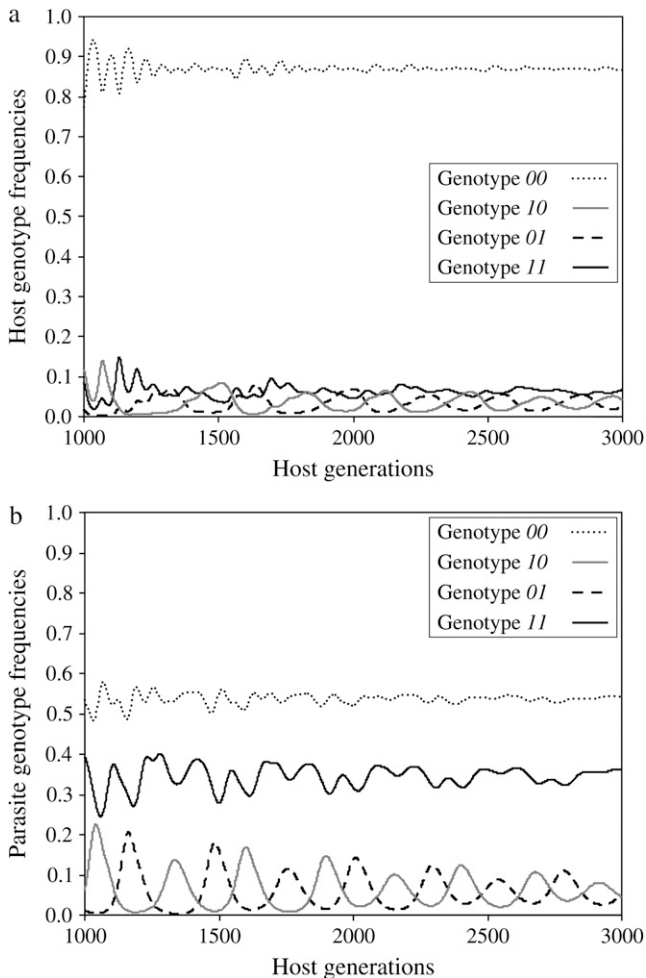


FIGURE 4.—Dynamics of host and parasite genotype frequencies in a two-locus model with epistasis in costs of alleles at different loci (model B2) as a function of the number of host generations. There is no mutation. Autoinfection rate  $\psi = 0.9$  and maximum cost of disease  $\phi = 0.2$ .  $u_1, u_2$ , costs of host resistance;  $b_1, b_2$ , costs of parasite virulence;  $u_1 = 0.072$  and  $b_1 = 0.024$ ;  $u_2 = b_2 = 0.0975$ . (a) Maintenance of all four host genotypes (area C in Figure 3a). (b) Maintenance of all four parasite genotypes (area I in Figure 3b).

independent. The cost curve has a concave shape when  $\xi < 1$ . The cost  $u_k$  of having  $k$  *RES* alleles is thus

$$u_k = u_{\max} \left( \frac{k}{n} \right)^\xi \quad \text{with } 0 \leq k \leq n. \quad (23)$$

In model B2,  $n = 2$  so  $u_{\max} = u_2$ . To compare results from models B1 and B2, the cost of having two alleles is fixed to 0.0975 ( $u_2 = b_2 = 0.0975$ , Equation 6). The only difference between models B1 and B2 is then the cost of having one allele with  $u_1 = 0.072$  ( $\xi = 0.4$ ) and  $b_1 = 0.024$  ( $\theta = 2$ ). Results for each area of Figure 3 can be seen in supplemental Section 6 at <http://www.genetics.org/supplemental/>.

In model B2, stable polymorphism with all four genotypes is maintained in host (area C in Figures 3a

and 4a) and parasite populations (area I in Figures 3b and 4b) with medium to high  $\psi$ . This occurs because the equilibrium frequency of double-*RES* hosts increases in proportion to the difference between  $b_1$  and  $b_2$  (Equation 14). In other words, increasing the cost of having two *avr* alleles compared to the cost of one *avr* allele favors parasite genotypes 01 and 10, thus enhancing selection for double-*RES* host genotypes and increasing the value of  $\widehat{RR}$  (Figure 4a). In simulations conducted with epistasis between parasite loci and multiplicative costs of *RES* alleles, there is also stable polymorphism with all four host and parasite genotypes (data not shown). Moreover, epistasis among host loci that decreases the difference between  $u_1$  and  $u_2$  has the effect of diminishing the equilibrium frequency of double-*avr* parasites and increasing that of double-*AVR* parasites. The diminution of the cost of having two *RES* alleles compared to that of one *RES* allele decreases selection for host genotypes 01 and 10, thus enhancing selection for double-*AVR* genotypes and increasing the value of  $\widehat{AA}$  (Figure 4b).

The total stability area for the host in model B2 (C and B in Figure 3a) has a comparable size but is not identical to area B in Figure 1a (model B1). For the parasite, stability areas H and I together in Figure 3b have a comparable size but are not identical to areas H, I, and J together in Figure 1b (model B1). The areas for stable polymorphism are comparable because conditions for stability depend on the strength of direct FDS, which mainly depends on  $\psi$  and  $\phi$  (see Equations 19 and 20). However, stability conditions also depend on costs of *RES* and *avr* alleles (Equations 16 and 17), which differ between models B1 and B2, which is why stability areas do not overlap exactly between Figures 1 and 2. Although all four host and parasite genotypes can be maintained, double-*res* plants and double-*avr* parasites have higher equilibrium frequencies than the other genotypes, in agreement with results from natural populations (THRALL *et al.* 2001) (Figure 4, a and b).

In model B2, the equilibrium point with the four host and parasite genotypes exists. However, when  $\phi$  increases, because indirect FDS overrides the stabilizing effect of direct FDS, this equilibrium state becomes unstable in areas F (Figure 3a) and M (Figure 3b).

Other results from model B2 are similar to those from model B1. If  $\phi < u_1$ , host genotype 00 (area A in Figure 3a) and parasite genotype 11 (area G in Figure 3b) become fixed. When  $u_1 < \phi < u_2$ , and  $\psi$  is intermediate to high, double-*RES* genotypes and double-*avr* parasites are eliminated, respectively, from the host (area B in Figure 3a) and parasite (area H in Figure 3b) populations. In areas B and H, there is stable polymorphism with three host and three parasite genotypes. Finally, at very high  $\phi$ , there is fixation of the 00 host and parasite genotypes (areas E and L in Figure 3). Simulations for more than two loci generalize our conclusions from models B, showing the generality of the approach ( $n = 3$

in supplemental Section 7 at <http://www.genetics.org/supplemental/>).

## DISCUSSION

As in the single-locus model of TELLIER and BROWN (2007), polycyclic disease gives rise to conditions that stabilize polymorphism at equilibrium (model B, cf. model A). Stable polymorphism with three or four host genotypes and four parasite genotypes occurs at intermediate to high rates of autoinfection (model B). For monocyclic disease (model A), there is no direct FDS in the host ( $\partial\Delta f_{rR}/\partial f_{rR} = \partial\Delta f_{rr}/\partial f_{rr} = 0$ ) or parasite ( $\partial\Delta f_{Aa}/\partial f_{Aa} = \partial\Delta f_{aa}/\partial f_{aa} = 0$ ) populations (Equation 10). Polymorphism is stabilized in polycyclic diseases, because the stability of GFG systems depends on the outcome of infection in the first parasite generation ( $g, 1$ ) influencing the second parasite generation ( $g, 2$ ). Model B extends the principle of single-locus GFG coevolution (TELLIER and BROWN 2007) to multiple loci. Polycyclic disease generates direct FDS for parasite virulence ( $\partial\Delta f_{Aa}/\partial f_{Aa} < 0$  and  $\partial\Delta f_{aa}/\partial f_{aa} < 0$ ) but not for host resistance ( $\partial\Delta f_{rR}/\partial f_{rR} = \partial\Delta f_{rr}/\partial f_{rr} = 0$ ). Increasing alloinfection (decreasing  $\psi$ ) tends to make successive parasite generations on the same plant independent of one another, causing selection against parasite genotypes to tend to become independent of their own frequency (TELLIER and BROWN 2007) and decreasing the parameter space in which polymorphism is stable. This can be explained as follows.

The coefficient  $\partial\Delta f_{aa}/\partial f_{aa}$  tends to zero when  $\psi$  is small. When the frequency ( $AA$ ) of the double-*AVR* parasite (genotype  $II$ ) is high and  $\psi$  is low, most double-*RES* plants infected by double-*avr* parasites in ( $g, 1$ ) then encounter a double-*AVR* parasite in ( $g, 2$ ). Increasing  $\psi$ , however, increases the probability of these double-*RES* plants remaining infected with a double-*avr* parasite in ( $g, 2$ ). Hence at higher frequencies of double-*AVR* parasites and increasing autoinfection, the strength of natural selection for the double-*avr* genotype and against double-*AVR* becomes greater ( $\partial\Delta f_{aa}/\partial f_{aa}$  is more negative). Similarly, the coefficient  $\partial\Delta f_{Aa}/\partial f_{Aa}$  tends to zero when  $\psi$  is low. When the frequency ( $Aa$ ) of parasite genotype  $IO$  and  $\psi$  are both low, most  $OI$  plants infected by  $IO$  parasites in ( $g, 1$ ) then encounter a  $OI$  parasite in ( $g, 2$ ). Increasing  $\psi$ , however, increases the probability of these  $OI$  plants remaining infected with a  $IO$  parasite in ( $g, 2$ ). Hence, natural selection for parasite genotype  $IO$  and against  $OI$  is stronger ( $\partial\Delta f_{Aa}/\partial f_{Aa}$  is more negative) when the frequency of  $IO$  parasites is lower, and this effect is stronger as  $\psi$  increases.

The absence (model B1, Figures 1 and 2) or presence (model B2, Figures 3 and 4) of epistatic interactions between fitness costs at different GFG loci has a considerable influence on the maintenance of multiple genotypes in host and parasite populations. Epistatic

interactions between virulent loci allow the existence of stable equilibrium frequencies of plant genotypes with multiple *RES* alleles. In two-locus systems, double-resistant plants ( $II$ ) are maintained if *avr* alleles have a negative synergistic effect on parasite fitness (model B2), and these results extend to three-locus interactions (model C in supplemental Section 7 at <http://www.genetics.org/supplemental/>). In biological terms, increasing exponentially the cost of each *avr* allele as a function of the total number of *avr* alleles counterselects genotypes with multiple *avr* alleles. The resulting higher parasite genotypic diversity then favors multiple-*RES* plant genotypes. Moreover, when the cost of each *RES* allele diminishes as a function of the total number of *RES* alleles, this counterselects host genotypes with intermediate numbers of *RES* alleles and favors double-resistant plants. Parasites with few *AVR* alleles are then favored against super-*avr* parasites that have high costs of *avr*. As a consequence, the equilibrium point with all four parasite genotypes is more likely to exist.

Epistatic interactions between *RES* and *avr* loci are interesting in relation to current advances in research on the function of *RES* and *avr* genes. To date, the great majority of experiments on plant and parasite fitness have been conducted on single genes (*RES* or *avr*). Some (LEONARD 1969; VERA CRUZ *et al.* 2000; THRALL and BURDON 2003; TIAN *et al.* 2003) but not all (BERGELSON and PURRINGTON 1996; VERA CRUZ *et al.* 2000; BROWN 2003) experiments have detected such costs. The structure of fitness costs emerging from current research in molecular biology supports the hypothesis that epistasis is of the type that leads to GFG polymorphism being stable (and hence detectable).

Avirulence genes in plant parasites have a dual role. The proteins they encode are recognized by the host plant's defense machinery, and hence their avirulence function, similar to the antigenicity of parasites of vertebrates. However, many *AVR* proteins also have effector activity, promoting infection, colonization, or pathogenicity (SKAMNIOTI and RIDOUT 2005; JONES and DANGL 2006; RIDOUT *et al.* 2006). Parasites generally have numerous avirulence/effector genes (KAY *et al.* 2005) and there is evidence for redundancy between *AVR* proteins (BAI *et al.* 2000; WICHMANN and BERGELSON 2004; KAY *et al.* 2005; MUDGETT 2005; SKAMNIOTI and RIDOUT 2005). Increasing the number of mutations of *AVR* genes to *avr* alleles may therefore have a synergistic negative effect on parasite fitness (infectivity, growth, reproduction, etc.). In experiments on multiple knock-outs of avirulence/effector genes in *Xanthomonas axonopodis*, the loss of function of one or two of four avirulence genes did not affect significantly bacterial growth, but a significant effect was observed when three or four genes were knocked out (WICHMANN and BERGELSON 2004). Several parasite *AVR* genes exist as gene families and are thus predicted to complement each other's effector function to some extent (*e.g.*, *Blumeria graminis*; SKAMNIOTI

and RIDOUT 2005; RIDOUT *et al.* 2006). Whether such a situation is general is not yet known, but if the cost of having one or very few *avr* alleles is low (2% in model B2), it might explain the lack of experimental evidence for a high cost of a single virulence allele (VERA CRUZ *et al.* 2000; THRALL and BURDON 2003). It is also predicted that polymorphism would be commonly observed in multi-gene families where there is synergistic epistasis of the costs of different *avr* alleles, as in models B2 and C. An experiment to test this prediction would estimate the fitness costs of combinations of various numbers of *AVR* genes in a common background (WICHMANN and BERGELSON 2004).

Antagonistic interactions between *RES* alleles in their effect on host fitness are not essential to maintain polymorphism in parasite genotypes, but favor maintenance of genotypes with a low to intermediate number of *avr* alleles. We assume here that the cost of having one *RES* allele is high (7% in our simulations compared to 9% found by TIAN *et al.* 2003), but that the marginal cost of each new *RES* allele added at other loci diminishes (here only 2% more cost for the second allele). High costs of *RES* have been found experimentally (TIAN *et al.* 2003), but if different *RES* genes each had such a cost, the fitness of a plant with several *RES* genes would be severely depressed (BROWN 2003). However, if the marginal cost of adding a new *RES* gene is small, the fitness load of many *RES* genes may not be much greater than that of one (BERGELSON and PURRINGTON 1996; PALOMINO *et al.* 2002). Functional data on resistance reactions show that many *RES* genes activate similar defense proteins (JONES and DANGL 2006). The cost of expressing host defenses may be similar whether they are triggered by a single *RES*-*AVR* interaction or by several pairs of *RES* and *AVR* genes.

Other theoretical models have suggested that multi-locus GFG systems for monocyclic disease enhance polymorphism maintenance (FRANK 1993a, 1997; SASAKI 2000; THRALL and BURDON 2002; SALATHE *et al.* 2005; SEGARRA 2005). In a multilocus GFG coevolutionary system with  $n$  interacting loci, there are  $2^n$  genotypes in host and parasite. An increase in  $n$  also diminishes the expected frequency of each host and parasite genotype to a mean equilibrium frequency of  $1/2^n$  (FRANK 1993a, 1997). In finite and spatially structured populations, allele frequencies in a high-dimension system (high  $n$ ) are thus more sensitive to random processes (mutations, genetic drift, and migrations) counteracting the frequency-dependent selection process (FRANK 1993a, 1997; THRALL and BURDON 2002). High mutation rates (SASAKI 2000; SALATHE *et al.* 2005; SEGARRA 2005) introduce new genotypes at high frequencies and sustain successive stochastic frequency-dependent selection cycles. As there is no direct frequency-dependent selection in simple models of monocyclic disease, mutation lead to arms race coevolutionary dynamics, with recurrent fixation of alleles, rather than trench warfare

dynamics, with stable polymorphism. Interestingly, in our models, realistic rates of mutation do not affect the outcome of coevolution in terms of the stability of polymorphism or the number of genotypes maintained but merely increase the time to genotype fixation and the lifetime of a mutation (this was also shown by SEGARRA 2005).

Polycyclic disease and autoinfection are important features of many diseases of plants and animals and have been shown to favor stable long-term maintenance of polymorphism at host and parasite loci (TELLIER and BROWN 2007). Here, a similar outcome is observed in multilocus GFG interactions with realistic costs of *RES* and *avr* alleles (model B1), in contrast to monocyclic disease (model A). Models B2 and C indicate the importance of epistatic interactions between host and parasite loci for costs of multiple *RES* and *avr* alleles and predict that GFG polymorphism will be stable (and hence detectable) when there is precisely the structure of costs that seems to be emerging from current discoveries in molecular biology.

#### LITERATURE CITED

- ALFANO, J. R., and A. COLLMER, 2004 Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu. Rev. Phytopathol.* **42**: 385–414.
- APANUS, V., D. PENN, P. R. SLEV, L. R. RUFF and W. K. POTTS, 1997 The nature of selection on the major histocompatibility complex. *Crit. Rev. Immunol.* **17**: 179–224.
- BAL, J. F., S. H. CHOI, G. PONCIANO, H. LEUNG and J. E. LEACH, 2000 *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. *Mol. Plant-Microbe Interact.* **13**: 1322–1329.
- BARRETT, J. A., 1980 Pathogen evolution in multilines and variety mixtures. *Z. Pflanzenk. Pflanzens.- J. Plant Dis. Prot.* **87**: 383–396.
- BERGELSON, J., and C. B. PURRINGTON, 1996 Surveying patterns in the cost of resistance in plants. *Am. Nat.* **148**: 536–558.
- BERGELSON, J., G. DWYER and J. J. EMERSON, 2001 Models and data on plant-enemy co-evolution. *Annu. Rev. Genet.* **35**: 469–499.
- BEVAN, J. R., I. R. CRUTE and D. D. CLARKE, 1993 Variation for virulence in *Erysiphe Fischeri* from *Senecio vulgaris*. *Plant Pathol.* **42**: 622–635.
- BORGHANS, J. A. M., J. B. BELTMAN and R. J. DE BOER, 2004 MHC polymorphism under host-pathogen co-evolution. *Immunogenetics* **55**: 732–739.
- BROWN, J. K. M., 2003 A cost of disease resistance: Paradigm or peculiarity? *Trends Genet.* **19**: 667–671.
- CAMPBELL, C. L., and L. V. MADDEN, 1990 *Introduction to Plant Disease Epidemiology*. Wiley Interscience, New York.
- DANGL, J. L., and J. D. G. JONES, 2001 Plant pathogens and integrated defence responses to infection. *Nature* **411**: 826–833.
- DEWIT, P., 1992 Molecular characterization of gene-for-gene systems in plant-fungus interactions and the application of avirulence genes in control of plant-pathogens. *Annu. Rev. Phytopathol.* **30**: 391–418.
- DINOOR, A., and N. ESHED, 1987 The analysis of host and pathogen populations in natural ecosystems, pp. 75–88 in *Populations of Plant Pathogens*, edited by M. S. WOLFE and C. E. CATEN. Blackwell Scientific Publications, Oxford.
- FRANK, S. A., 1992 Models of plant pathogen co-evolution. *Trends Genet.* **8**: 213–219.
- FRANK, S. A., 1993a Co-evolutionary genetics of plants and pathogens. *Evol. Ecol.* **7**: 45–75.
- FRANK, S. A., 1993b Specificity versus detectable polymorphism in host-parasite genetics. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **254**: 191–197.

- FRANK, S. A., 1997 Spatial processes in host-parasite genetics, pp. 325–352 in *Metapopulation Biology: Ecology, Genetics and Evolution*, edited by I. HANSKI and M. GILPIN. Academic Press, New York.
- HILL, A. V. S., 2001 The genomics and genetics of human infectious disease susceptibility. *Annu. Rev. Genomics Hum. Genet.* **2**: 373–400.
- HOLUB, E. B., 2001 The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* **2**: 516–527.
- JONES, J. D. G., and J. L. DANGL, 2006 The plant immune system. *Nature* **444**: 323–329.
- JØRGENSEN, J. H., 1994 Genetics of powdery mildew resistance in barley. *Crit. Rev. Plant Sci.* **13**: 97–119.
- KAY, S., J. BOCH and U. BONAS, 2005 Characterization of AvrBs3-like effectors from a *Brassicaceae* pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Mol. Plant-Microbe Interact.* **18**: 838–848.
- KIRBY, G. C., and J. J. BURDON, 1997 Effects of mutation and random drift on Leonard's gene-for-gene co-evolution model. *Phytopathology* **87**: 488–493.
- KOT, M., 2001 *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge, UK.
- LAINE, A. L., 2004 Resistance variation within and among host populations in a plant-pathogen metapopulation: implications for regional pathogen dynamics. *J. Ecol.* **92**: 990–1000.
- LEONARD, K. J., 1969 Selection in heterogeneous populations of *Puccinia graminis* f. sp. *avenae*. *Phytopathology* **59**: 1845–1850.
- MAY, R. M., and R. M. ANDERSON, 1990 Parasite host co-evolution. *Parasitology* **100**: S89–S101.
- MUDGETT, M. B., 2005 New insights to the function of phytopathogenic bacterial type III effectors in plants. *Annu. Rev. Plant Biol.* **56**: 509–531.
- PALOMINO, M. M., B. C. MEYERS, R. W. MICHELMORE and B. S. GAUT, 2002 Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res.* **12**: 1305–1315.
- RIDOUT, C. J., P. SKAMNIOTI, O. PORRITT, S. SACRISTAN, J. D. G. JONES *et al.*, 2006 Multiple avirulence paralogs in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *Plant Cell* **18**: 2402–2414.
- ROUGHGARDEN, J., 1996 *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Prentice Hall, New York.
- SALATHE, M., A. SCHERER and S. BONHOEFFER, 2005 Neutral drift and polymorphism in gene-for-gene systems. *Ecol. Lett.* **8**: 925–932.
- SASAKI, A., 2000 Host-parasite co-evolution in a multilocus gene-for-gene system. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **267**: 2183–2188.
- SEGARRA, J., 2005 Stable polymorphisms in a two-locus gene-for-gene system. *Phytopathology* **95**: 728–736.
- SKAMNIOTI, P., and C. J. RIDOUT, 2005 Microbial avirulence determinants: guided missiles or antigenic flak? *Mol. Plant Pathol.* **6**: 551–559.
- STAHL, E. A., G. DWYER, R. MAURICIO, M. KREITMAN and J. BERGELSON, 1999 Dynamics of disease resistance polymorphism at the Rpm1 locus of *Arabidopsis*. *Nature* **400**: 667–671.
- TELLIER, A., and J. K. M. BROWN, 2007 Stability of genetic polymorphism in host-parasite interactions. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **274**: 809–817.
- THRALL, P. H., and J. J. BURDON, 2002 Evolution of gene-for-gene systems in metapopulations: the effect of spatial scale of host and pathogen dispersal. *Plant Pathol.* **51**: 169–184.
- THRALL, P. H., and J. J. BURDON, 2003 Evolution of virulence in a plant host-pathogen metapopulation. *Science* **299**: 1735–1737.
- THRALL, P. H., J. J. BURDON and A. YOUNG, 2001 Variation in resistance and virulence among demes of a plant host-pathogen metapopulation. *J. Ecol.* **89**: 736–748.
- TIAN, D., M. B. TRAW, J. Q. CHEN, M. KREITMAN and J. BERGELSON, 2003 Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* **423**: 74–77.
- VERA CRUZ, C. M., J. F. BAL, I. ONA, H. LEUNG, R. J. NELSON *et al.*, 2000 Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc. Natl. Acad. Sci. USA* **97**: 13500–13505.
- WICHMANN, G., and J. BERGELSON, 2004 Effector genes of *Xanthomonas axonopodis* pv. *vesicatoria* promote transmission and enhance other fitness traits in the field. *Genetics* **166**: 693–706.
- WOLFRAM RESEARCH, 2003 *Mathematica*. Wolfram Research, Champaign, IL.

Communicating editor: M. VEUILLE

#### APPENDIX: TWO-LOCUS GFG MODEL WITH MONOCYCLIC DISEASE

**The model:** The parasite genotype frequencies at host generation  $g$  are  $AA_g$ ,  $Aa_g$ ,  $aA_g$ , and  $aa_g$  for the genotypes  $11$ ,  $10$ ,  $01$ , and  $00$ , respectively. Similarly,  $r_r_g$ ,  $rR_g$ ,  $Rr_g$ , and  $RR_g$  stand for the frequencies in generation  $g$  of the respective host genotypes  $11$ ,  $10$ ,  $01$ , and  $00$ . The frequency of parasite genotype  $11$  in host generation  $g + 1$  is thus  $AA_{g+1} = AA_g r_r_g / \bar{w}_A$ . Similarly,

$$aa_{g+1} = \frac{aa_g(1 - b_2)}{\bar{w}_A},$$

$$aA_{g+1} = \frac{aA_g(1 - b_1)(r_r_g + Rr_g)}{\bar{w}_A},$$

and

$$Aa_{g+1} = \frac{Aa_g(1 - b_1)(r_r_g + rR_g)}{\bar{w}_A}.$$

Where  $\bar{w}_A$  is the overall parasite population fitness,

$$\bar{w}_A = AA_g r_r_g + aa_g(1 - b_2) + aA_g(1 - b_1)(r_r_g + Rr_g) + Aa_g(1 - b_1)(r_r_g + rR_g).$$

Host  $11$  genotype frequency at host generation  $g + 1$  is then

$$RR_{g+1} = \frac{RR_g(1 - u_2)(aa_g(1 - s) + aA_g + Aa_g + AA_g)}{\bar{w}_R}$$

and

$$Rr_{g+1} = \frac{Rr_g(1 - u_1)((1 - s)(aa_g + aA_g) + Aa_g + AA_g)}{\bar{w}_R}$$

$$rr_{g+1} = \frac{rr_g(1 - u_1)((1 - s)(aa_g + Aa_g) + aA_g + AA_g)}{\bar{w}_R}$$

$$r_{g+1} = \frac{r_g(1 - s)}{\bar{w}_R}.$$

Where  $\bar{w}_R$  is the overall host population fitness,

$$\bar{w}_R = \left[ \frac{RR_g(1 - u_2)(aa_g(1 - s) + aA_g + Aa_g + AA_g) + Rr_g(1 - u_1)((1 - s)(aa_g + aA_g) + Aa_g + AA_g)}{+ rR_g(1 - u_1)((1 - s)(aa_g + Aa_g) + aA_g + AA_g) + r_g(1 - s)} \right].$$

**Stability of the equilibrium state:** Using logit transformations of the above equations (Equations 5 and 6 in the text), the Jacobian matrix  $J_A$  can be rewritten:

$$J_A = \begin{pmatrix} \partial\Delta f_{Aa}/\partial f_{Aa} & \partial\Delta f_{Aa}/\partial f_{aa} & \partial\Delta f_{Aa}/\partial f_{rR} & \partial\Delta f_{Aa}/\partial f_{rr} \\ \partial\Delta f_{aa}/\partial f_{Aa} & \partial\Delta f_{aa}/\partial f_{aa} & \partial\Delta f_{aa}/\partial f_{rR} & \partial\Delta f_{aa}/\partial f_{rr} \\ \partial\Delta f_{rR}/\partial f_{Aa} & \partial\Delta f_{rR}/\partial f_{aa} & \partial\Delta f_{rR}/\partial f_{rR} & \partial\Delta f_{rR}/\partial f_{rr} \\ \partial\Delta f_{rr}/\partial f_{Aa} & \partial\Delta f_{rr}/\partial f_{aa} & \partial\Delta f_{rr}/\partial f_{rR} & \partial\Delta f_{rr}/\partial f_{rr} \end{pmatrix}.$$

The coefficients of  $J_A$  are the rates of natural selection of the ratio of genotype frequencies. For example,  $\partial\Delta f_{aa}/\partial f_{rr}$  is the rate of selection on the double-*avr* parasite genotype ( $00$ ) as a function of the frequency of the double-susceptible host genotype ( $00$ ). Close to the equilibrium point, the Jacobian matrix coefficients are approximately

$$J_A = \begin{pmatrix} 0 & 0 & \frac{b_1^2(1-b_2)^2}{1-b_1} & 0 \\ 0 & 0 & 0 & \frac{2b_1-b_2-2b_1b_2}{1-b_1} \\ \frac{-(1-s)(u_2-u_1)^2}{s(1-u_2)} & 0 & 0 & 0 \\ 0 & \frac{(s-u_2)(2u_1-u_2-u_1u_2)}{s(1-u_2)(1-u_1)} & 0 & 0 \end{pmatrix}$$

(supplemental Section 1 at <http://www.genetics.org/supplemental/>). The Jacobian matrix  $J_A$  is diagonalizable and thus has four eigenvalues ( $\lambda_{1-4}$ ):

$$\begin{cases} \lambda_{1,2} = \pm i\sqrt{-\beta_1} \\ \lambda_{3,4} = \pm \sqrt{\beta_2} \end{cases} \text{ if } \beta_1 < 0 \quad \text{or} \quad \begin{cases} \lambda_{1,2} = \pm \sqrt{\beta_1} \\ \lambda_{3,4} = \pm \sqrt{\beta_2} \end{cases} \text{ if } \beta_1 > 0$$

with

$$\beta_1 = \frac{(b_2 - 2b_1(1 - b_2))(u_2 - u_1(2 - u_2))(s - u_2)}{s(1 - b_1)(1 - u_1)(1 - u_2)}$$

and

$$\beta_2 = \frac{(1 - s)(1 - b_2)b_1^2(u_1 - u_2)^2}{s(1 - b_1)(1 - u_2)}.$$

The sign of  $\beta_1$  depends on values of costs ( $u_1$ ,  $u_2$ ,  $b_1$ ,  $b_2$ ,  $s$ ), and  $\beta_2$  is always positive.