The influence of disease prevalence on polymorphism maintenance in host-parasite interactions

Stefany Moreno

Supervisors: Dr. Aurélien Tellier and Dr. Wolfgang Stephan

September 1, 2012

1 Introduction

Polymorphism in loci associated to host resistance and parasite virulence is commonly found in natural populations [1, 2, 3, 4, 5, 6]. Indeed, host-parasite interactions have been early on identified as important selective forces promoting the maintenance of genetic diversity in biological systems [2, 3, 4]. Two alternative scenarios have been proposed to explain the observed amount of polymorphism at loci associated to these antagonistic interactions. In the "arms race" model, there is recurrent fixation of alleles in both the host and the parasite, so polymorphism in virulence and resistance exists but is transient in time [2, 7]. This model is mainly driven by mutation which introduces novel alleles in the population which then become fixed by selective sweeps (positive selection). On the other hand, in the "trench warfare" model, the frequencies of resistance and virulence alleles are maintained at intermediate values by balancing selection. In this case polymorphism is stable in time and is predicted to be ancient in natural populations [7]. One of the main aims of current research in host-parasite interactions is to determine the ecological, genetical and epidemiological factors that can favor each of these two models of coevolution.

The previous scenarios have been studied using discrete time models based on gene-for-gene (GFG) interactions between hosts and parasites or Matching-Allele models (MAM) [3]. GFG models are a common mechanism of infection in plant and invertebrate diseases [8, 9]. In the classic GFG model proposed by Leonard [10], it is assumed that both the host and the parasite have one locus. The host has two alleles, for resistance and susceptibility and the parasite has two alleles for virulence and avirulence. Infection occurs if the host is susceptible or if the parasite is virulent. Otherwise, if the host is resistant and the parasite avirulent, the host produces a resistance reaction avoiding infection. Since disease is costly, hosts are selected to be resistant which in turn imposes a pressure

for parasites to become virulent and overcome host defenses. This leads to coevolutionary cycles where the allelic frequencies of one species determine the fitness of the alleles in the other species (virulence is favored when resistance is common and resistance is favored when virulence is rare), which is known as indirect frequency dependence selection (iFDS) [11]. In models where only iFDS occurs, cycles are unstable, driving alleles to fixation recurrently so the system behaves like an arms race scenario. Polymorphism can become stable in a GFG model, however, if there is negative direct frequency dependence selection (ndFDS) [11]. This mechanism imposes a fitness cost in alleles that are in high frequency in the population preventing them to be fixed. Ecological and epidemiological factors like polycyclic disease or spatial heterogeneity in the costs of virulence and resistance have the potential to introduce ndFDS in the system producing a trench warfare scenario for variation of the allelic frequencies over time [11, 12, 13].

One of the main assumptions of the previous models is that the host encounters the parasite every generation. However, it is often the case in natural populations that potential hosts are never exposed to the parasite during their lifetime [14, 15, 16]. Both epidemiological and environmental conditions can affect the prevalence of the parasite in a population over space and time. For instance, microclimatic conditions, dormancy periods of the parasite and mechanisms of parasite dispersion have been reported as factors that can modify the frequency of parasite encounters in plant and invertebrate populations [17, 18, 15].

Spatiotemporal variation in the encounter rates with the parasite can change the intensity of selection for resistance traits in the host [14, 20]. When the host does not encounter the parasite, the effect of iFDS disappears since the fitness of the host is independent of the frequency of virulence in the parasite population. The importance of variation in the parasite prevalence on the dynamics of host-parasite interactions has been shown experimentally in studies on plant and bacterial populations[14, 21]. These studies have focused mainly in heterogeneity at a spatial scale, a very common feature of natural systems where species occur as separate demes connected through migration. In these scenarios, high variation in the parasite encounter rates have been observed even for neighboring populations [14, 22]. These findings have been integrated into the theory of the 'geographic mosaic of coevolution' which suggests that variation in ecological conditions across spatially structured populations can lead to differences in selection pressures between demes [23]. Demes where there is strong reciprocal selection are usually referred as "hotspots" of coevolution while demes with weak reciprocal selection are referred as "coldspots".

The simple GFG model has been already studied in spatially structured populations where it was shown that differences in costs of resistance and virulence between demes can promote stability

of genetic polymorphism by introducing ndFDS through migration [13]. However, the effect of differences in parasite encounter rates between populations has never been explored in the framework of GFG interactions in hosts metapopulations.

This project has the following objectives. First, to relax the assumption of permanent exposure to the parasite in the simple GFG model to determine the effect of parasite encounter rates on the dynamics of the system. We want to analyze the effect of parasite encounter rates on the frequency of virulence and resistance alleles and its potential to introduce stability to the polymorphism. Second, we will introduce variation in parasite encounter rates in spatially structured populations. This would allow us to evaluate if differences only in the prevalence of the parasite between demes can introduce ndFDS through migration and stabilize the genetic polymorphism. Third, we aim to determine what is the effect of fitness differences between resistant and susceptible hosts when they do not encounter the parasite, on the evolution of virulence and resistance. This is based on the observation that in natural populations adaptations to parasite infection can be expressed constantly in host populations even if the parasite is not present [24, 25, 26, 27]. This could disfavor resistant hosts if the prevalence of the parasite is low.

2 Model description

2.1 One deme scenario

Both the host and the parasite are haploid with discrete generations on time. A GFG model is assumed so that the host has two alleles, resistance (RES) and susceptibility (res) and the parasite has two alleles, virulence (avr) and avirulence (AVR). Infection occurs if the parasite is virulent or if the host is susceptible. Each generation the hosts encounters the parasite with probability ρ. If the parasite is virulent it can infect both types of hosts but it faces a cost of virulence b. Avirulent parasites do not bear this cost but if they encounter a resistant host they have a cost c of not being able to infect it. Following Tellier and Brown [11], and their biological supported assumption that parasites not able to infect the host have a very large fitness cost, we assume that $c \approx 1$.

Being diseased has a cost s for the host. Based on recent molecular findings on gene expression and disease resistance network activation, we assume that resistant hosts encountering parasites trigger a resistance reaction which incurs a cost of fitness [28, 29]. Such resistance reaction can be overcomed by virulent parasites who switch off defenses via effectors, where avirulent parasites cannot. The cost of being resistant for the host when exposed to a parasite is u. If the host does not

encounter the parasite, which happens with probability $1-\rho$, susceptible hosts have no fitness cost because they are not infected. Resistant hosts have a cost u^*, which corresponds to the basal cost of having the allele for resistance even when the parasite is not present. It is assumed that $u = u^* + \epsilon$, where $\epsilon \geq 0$ can be interpreted as the cost of activating and expressing the disease resistance pathway mentioned before. When the parasite does not encounter the host, it does not survive, regardless of its genotype. The fitness costs for the host and parasite populations are summarized in Table 1. As a simplifying assumption, all the costs and the encounter rate are assumed to be constant through time.

In generation g, the frequency of *avr* in the parasite population is a_g and the frequency of *RES* in the host population is R_g. A_g and r_g correspond to the frequencies of *AVR* and *res* respectively. The recurrence equations are:

Host

$$\begin{aligned} \frac{R_{g+1}}{r_{g+1}} &= \frac{R_g}{r_g}\left(\frac{\rho(1-u)(1-s+sA_g)+(1-\rho)(1-u^*)}{1-\rho s}\right) \\ &= \frac{R_g}{r_g}\left(\frac{(1-u)(1-\rho s+\rho sA_g)+(1-\rho)\epsilon}{1-\rho s}\right) \end{aligned} \tag{1}$$

Parasite

$$\frac{A_{g+1}}{a_{g+1}} = \frac{A_g}{a_g}\left(\frac{1-cR_g}{1-b}\right) \tag{2}$$

Note that the fitness of the virulence allele in the parasite population is independent of ρ. This means that the probability of encountering a host is assumed to be equal for virulent and avirulent parasites, so ρ does not introduce fitness differences between the two type of parasites. This system has four trivial equilibriums $(\hat{R}, \hat{A})$ in (0,0), (0,1), (1,0) and (1,1). The non-trivial equilibrium for this system is:

$$\hat{R} = \frac{b}{c} \tag{3}$$

$$\begin{aligned} \hat{A} &= \frac{u^* - s\rho u^* - s\rho\epsilon + \rho\epsilon}{\rho s(1-u^*-\epsilon)} \\ &= \frac{\epsilon(1-s)}{s(1-u)} + \frac{u^*(1-\rho s)}{\rho s(1-u)} \end{aligned} \tag{4}$$

2.1.1 Analysis of the equilibrium

Note that the frequency of host alleles depends on the parasite parameters and vice versa (see [11, 30]). From (3) this means that resistance increases with the cost of virulence which is expected since resistant plants are favored when virulence is not common. An internal equilibrium point exists i.e. $0 < \hat{R} < 1$ since all the costs are assumed to be less than 1 and $c > b$ which is necessary for virulence to evolve in the first place. The equation for $\hat{A}$ can be written in terms of u as,

$$\hat{A} = \frac{u(1-\rho s) - \epsilon(1-\rho)}{\rho s(1-u)}$$

$\hat{A} > 0$ when $u^*(1-\rho s) + \epsilon\rho(1-s) > 0$, this is when ρ, u, s and ϵ are between 0 and 1. Note that $\hat{A} < 1$ if $u^* + \rho\epsilon < \rho s$. This means that the expected cost of resistance over plants that encounter and not the parasite, must be lower than the expected cost of being diseased. In the simple GFG model, this corresponds to the condition $s > u$ which is necessary for resistance to evolve in the first place.

If we fix the value of u, at the equilibrium the frequency of AVR decreases when ϵ increases. That is, the frequency of virulent parasites increases when the basal cost of resistance decreases. This happens because lower cost of basal resistance selects for higher frequency of resistant plants, which in turn favors virulent parasites. Note that the effect of ϵ is proportional to $1-\rho$, which is the rate at which hosts do not encounter parasites.

The equation for $\hat{A}$ can also be written as,

$$\hat{A} = \frac{\epsilon - su}{s(1-u)} + \frac{u^*}{\rho s(1-u)}$$

This means that ρ has a negative effect on the frequency of avirulent parasites. In other words, when the encounter rate with the parasite increases, virulent parasites are favored. This result can be explained in terms of the intensity of selection for resistance in the host population with respect to ρ. When ρ is low, susceptible hosts are favored because they have no fitness cost in contrast to resistant hosts that have a cost u^*. Then, avirulent parasites are favored. When ρ is high, resistance can provide advantages since disease can be avoided so the frequency of *res* increases, which favors virulent parasites.

Note that the effect of ρ in the equilibrium value depends on the existence of a cost u^* for resistant plants that do not encounter the parasite. Otherwise, if $u^* = 0$, the equilibrium of the system does not depend on ρ because when the parasite is not present selection is neutral in the host since

both *RES* and *res* have the same fitness. This can have the potential to maintain polymorphism in the host, because if the prevalence of the parasite is low and selection for resistance is mainly neutral when the parasite is not present, most of the change in the allelic frequencies for resistance will be driven by stochastic effects like drift and mutation. We will explore this, by accounting for stochasticity using simulations.

2.1.2 Stability analysis

The dynamics of the system was analyzed using its Jacobian matrix. The logit transformations $f_R = \log \frac{R_g}{r_g}$ and $f_A = \log \frac{A_g}{a_g}$ were made to simplify the analysis and the discriminant Q_A of the characteristic equation of the Jacobian matrix J_A was calculated. Under the reasonable assumptions that costs are between 0 and 1 and that $u^* + \rho\epsilon < \rho s$ and $b < c$, which are necessary for resistance and virulence to evolve in the first place (see above), the sign of Q_A is negative, which means that the frequencies rotate anti clockwise in a plot of (R,a). See Figure 1.

The model does not introduce negative frequency dependence selection (ndFDS) because the diagonal elements of J_A, $d\Delta f_R/df_R$ and $d\Delta f_A/df_A$ equal zero. This means that the fitness of an allele is independent of its own frequency. Since the diagonal elements of J_A are zero, the eigenvalues have no real parts. Then, the complex parts of the eigenvalues should be zero for the equilibrium to be stable which is not the case for this system. Then, the equilibrium is unstable. See Appendix I for the calculations.

2.2 Two demes scenario

As it was discussed before, heterogeneity in spatial structured populations has the potential to introduce stability to the system promoting the long-term maintenance of genetic polymorphism. In this model the influence of the parasite prevalence is studied in the case of two demes that are linked by migration of both the host and the parasite. Selection happens inside each deme as in the one deme model and migration takes place after selection. Each deme has its own costs and parasite encounter rates which are noted by u_i, b_i, s_i, c_i, ρ_i, u_i^* and $\epsilon_i (= u_i - u_i^*)$ for deme i. As in the one deme scenario, these correspond to cost of resistance, cost of virulence, cost of being diseased for the host, cost for the parasite of not being able to infect the host, encounter rate with the parasite, basal cost of resistance and cost of expression of a resistant allele respectively. For simplicity it is assumed that migration for both the host and the parasite is symmetrical between the two demes so there is the same number of immigrants and emigrants per generation in each deme. The migration

rates for the host and the parasite are noted as m_H and m_P respectively. To illustrate the system of equations for each deme, the equations for deme 1 will be shown. Before migration, when just selection has acted, the allele frequencies are as in the one deme model (See (1) and (2)):

$$\tilde{R}_1 = R_1((1-u_1)(1-\rho_1 s_1 + \rho_1 s_1 A_1) + (1-\rho_1)\epsilon_1)$$

$$\tilde{r}_1 = r_1(1-\rho_1 s_1)$$

$$\tilde{A}_1 = A_1(1-c_1 R_1)$$

$$\tilde{a}_1 = a_1(1-b_1)$$

where R_1, r_1, A_1 and a_1 correspond to the allele frequencies of *RES*, *res*, *AVR* and *avr* in deme 1 before selection and $\tilde{R_1}$, $\tilde{r_1}$, $\tilde{A_1}$ and $\tilde{a_1}$ indicates frequencies after selection. The frequencies in deme 2 after selection are noted by $\tilde{R}_2$, $\tilde{r}_2$, $\tilde{A}_2$ and $\tilde{a}_2$ and are calculated in the same way as for deme 1 but with the parameters u_2, b_2, s_2, c_2, ρ_2, u_2^* and ϵ_2 and the starting frequencies R_2 and A_2. Once migration occurs the allelic frequencies in the next generation for deme 1 are given by:
Host:

$$\frac{R_1'}{r_1'} = \frac{(1-m_H)\tilde{R}_1 + m_H\tilde{R}_2}{(1-m_H)\tilde{r_1} + m_H\tilde{r_2}} \tag{5}$$

Parasite:

$$\frac{A_1'}{a_1'} = \frac{(1-m_P)\tilde{A}_1 + m_P\tilde{A}_2}{(1-m_P)\tilde{a_1} + m_P\tilde{a_2}} \tag{6}$$

If there is no migration between the two demes ($m_H = m_P = 0$) the dynamics in each of them collapses to the one deme model that was analyzed in the previous section. As it was shown analytically for the one deme model, the internal equilibrium of the system is unstable.

2.2.1 Calculation and analysis of the equilibrium

When there is migration of both the host and the parasite, the internal equilibrium ($\hat{R}_i$, $\hat{A}_i$) in deme i depends on the equilibrium value in the other deme. This makes the analytical calculation of the equilibrium in each deme very difficult because the frequencies of both $\hat{R}_i$ and $\hat{A}_i$ depend on the selection coefficients of the two demes and on the migration rates. Some special cases can be, however, computed analytically when assuming that either $m_P = 0$ or $m_H = 0$.

a) m_P=0

If the parasite does not migrate (m_P=0) the frequency of virulence in each population behaves as in the simple one deme model and $A_1' = \tilde{A}_1$ and $A_2' = \tilde{A}_2$ (to see it replace m_P=0 in (6)). Then, the frequencies of the virulent allele in the next generation are the same as the frequencies after selection because migration does not occur. Since the allelic frequencies in one species determine the equilibrium in the other species, this means that $\hat{R}_i$ for both demes is independent of the migration rate and corresponds to the value in the simple one deme model. Then, for $m_P = 0$,

$$\hat{R}_1 = \frac{b_1}{c_1} \text{ and } \hat{R}_2 = \frac{b_2}{c_2}$$

The frequencies of $\hat{A}_1$ and $\hat{A}_2$ are then calculated based on $\hat{R}_1$ and $\hat{R}_2$. For $\hat{A}_1$ this is,

$$\hat{A}_1 = \hat{A_1^{od}} + \frac{m_H}{(1-m_H)} \frac{1}{(1-u_1)\rho_1 s_1} \left(\frac{\hat{r}_2}{\hat{r}_1}(1-\rho_2 s_2) - \frac{\hat{R}_2}{\hat{R}_1}((1-u_2)(1-\rho_2 s_2 + \rho_2 s_2 \hat{A}_2) + (1-\rho_2)\epsilon_2) \right) \tag{7}$$

where $\hat{A_1^{od}}$ corresponds to the equilibrium for A_1 in the simple one deme model (See previous section, equation (4)). First, note that the effect of migration is to deviate the equilibrium from the value expected in the case of a single population, $\hat{A_1^{od}}$. As expected, higher migration rates have a larger effect on the equilibrium frequency since $\frac{m_H}{1-m_H}$ increases. There are two possible scenarios: If the frequency of resistant hosts is higher in deme 1 than in deme 2, $\hat{R}_1 > \hat{R}_2$, the term inside the parenthesis in (7) is positive and migration will favor avirulent parasites in deme 1. This happens because if there are more resistant hosts in deme 1 than in deme 2, migration will decrease the frequency of resistance in deme 1 which favors in return avirulent parasites. On the other hand, if the number of susceptible hosts is higher in deme 1, $\hat{R}_1 < \hat{R}_2$, the term inside the parenthesis is negative, so migration will increase the frequency of virulent parasites in deme 1 since resistant hosts are entering the population of deme 1.

The encounter rate with the parasite modifies the influence of migration on the equilibrium. More specifically, when ρ_1 increases, the effect of migration on the frequency of virulence in deme 1 decreases (See (7)). In other words, a deme with a high prevalence of the parasite is less affected by migration than a deme with low prevalence of the parasite. This asymmetry in the effects of "hot spots" (high parasite prevalence) and "cold spots" (low parasite prevalence) has been observed

experimentally and has been attributed to the fact that hot spots act as net sources of migrants and cold spots as net recipients [21]. Our result suggests that even if the contribution to the migrant pool is identical for both type of demes, cold spots are more likely to be affected by migration than hot spots (when $m_P = 0$). Finally, note that the effect of the encounter rate on migration is independent of u^*. Even if the basal cost of resistance is zero, the encounter rate can still modify the equilibrium point, which is not the case for the one deme model.

b) $m_H = 0$

If the host does not migrate (m_H=0), the frequency of resistance in each deme behaves as in the one deme model. Then, $R_1' = \tilde{R_1}$ and $R_2' = \tilde{R_2}$ (to see it replace m_H=0 in (5)). The equilibrium values for $\hat{A_1}$ and $\hat{A_2}$ are given as in the one deme model by,

$$\hat{A_1} = \frac{u_1(1-\rho_1 s_1) - \epsilon_1(1-\rho_1)}{\rho_1 s_1(1-u_1)} \text{ and } \hat{A_2} = \frac{u_2(1-\rho_2 s_2) - \epsilon_2(1-\rho_2)}{\rho_2 s_2(1-u_2)}$$

The frequencies of $\hat{R_1}$ and $\hat{R_2}$ are then calculated based on $\hat{A_1}$ and $\hat{A_2}$. For $\hat{R_1}$ this is,

$$\hat{R_1} = \hat{R_1^{od}} - \frac{m_P}{(1-m_P)}\frac{1}{c_1}\left(\frac{\hat{a}_2}{\hat{a}_1}(1-b_2) - \frac{\hat{A_2}}{\hat{A_1}}(1-c_2\hat{R}_2)\right) \tag{8}$$

As before, $\hat{R_1^{od}}$ corresponds to the equilibrium for R_1 in the simple one deme model (See previous section, equation (3)) and the effect of migration is to move the equilibrium point away from this value. There are two possible scenarios: If the frequency of virulence is higher in deme 1 than in deme 2, $\hat{a}_1 > \hat{a}_2$, the term in the parenthesis in (8) is negative so migration favors resistance in deme 1. This happens because migration increases the frequency of avirulent parasites which in turn favor resistant hosts. If the frequency of virulence is lower in deme 1 than in deme 2, $\hat{a}_1 < \hat{a}_2$, the term in the parenthesis is positive, so migration decreases the frequency of resistant hosts since virulent parasites are entering the population and then susceptible hosts become favored. Note that in this scenario, as opposed to the case when m_P=0, the encounter rate of the parasite does not have a direct effect on the influence of migration in the allele frequencies (See (8)). However, since $\hat{A_1}$ and $\hat{A_2}$ are defined as in the one deme model because the host does not migrate, they depend on ρ. As it was shown higher encounter rates favor virulent parasites, which means that differences between $\hat{A_1}$ and $\hat{A_2}$ could be potentially caused by differences in the encounter rates between the demes. Differences between $\hat{A_1}$ and $\hat{A_2}$ can in turn modify the frequency of resistance in each deme through migration of the parasite according to (8).

Finally, note that if $u_1^* = u_2^* = 0$, the encounter rate has no effect on the frequencies of $\hat{A}_1$ and $\hat{A}_2$. This means that if there is no cost of resistance for the host when the parasite is not present and the host does not migrate, the encounter rates have no effect on the equilibrium of both the resistance and virulence alleles in each deme. This result is explained because selection for virulence is equivalent in a deme with high prevalence than in one with low prevalence *as long as* the prevalence does not affect the allelic frequencies in the host. This is the case if $u_1^* = u_2^* = 0$ because selection in the host is neutral when the parasite is not present, so the encounter rate has no effect on the frequencies of resistance. Then, if only the parasite migrates, differences in parasite prevalence in each deme will not affect the equilibrium of the system. Further discussion will be made about this point, however note that this does not mean that the encounter rate does not have an effect in the trajectories of the allele frequencies in general but only on the equilibrium point.

See the Appendix for all the calculations.

2.2.2 Stability analysis

The system was analyzed using the Jacobian matrix of each deme. The logit transformations $f_{Ri} = \log \frac{R_i}{r_i}$ and $f_{Ai} = \log \frac{A_i}{a_i}$ were made to simplify the analysis and the discriminant of the characteristic equation of the Jacobian matrix J_i for each deme was calculated.

The diagonal elements of the Jacobian matrix for deme i are given by (see the appendix for calculations),

$$\frac{d\Delta f_{Ri}}{df_{Ri}} = R_i r_i \left(\frac{((1-u_i)(1-\rho_i s_i a_i) + (1-\rho_i)\epsilon_i)(1-m_H)}{(1-m_H)\tilde{R}_i + m_H \tilde{R}_j} + \frac{(1-\rho_i s_i)(1-m_H)}{(1-m_H)\tilde{r_i} + m_H \tilde{r_j}} \right) - 1 \tag{9}$$

$$\frac{d\Delta f_{Ai}}{df_{Ai}} = A_i a_i \left(\frac{(1-c_i R_i)(1-m_P)}{(1-m_P)\tilde{A}_i + m_P \tilde{A}_j} + \frac{(1-b_i)(1-m_P)}{(1-m_P)\tilde{a_i} + m_P \tilde{a_j}} \right) - 1 \tag{10}$$

After simplifying these expressions it is clear that both differentials are negative as long as there is migration between both demes and at least some of the costs differ between the demes. These results were already described in [13], where the authors showed that differences in u, b, s or c have the potential to generate ndFDS through migration. What our results suggest is that even if all those costs are equal between both demes, differences only in the parasite prevalence can introduce ndFDS to the system. This is because even if all the other costs are constant, ρ can introduce differences in selection for resistance affecting both $\tilde{R}_i$ and $\tilde{R}_j$ and in turn, A_i and A_j. Since u^* can also introduce fitness differences between resistant and susceptible hosts it can also introduce differences between $\tilde{R}_i$ and $\tilde{R}_j$. The effect of u^* on stability is lower, however, since its importance

into the model depends on ρ because this cost only appears if the parasite is not present.

Note that in the case that $m_P = 0$ there is no ndFDS in the parasite but only in the host. The opposite applies when $m_H = 0$. The equilibrium can be stable, however, since the condition $\frac{d\Delta f_{Ri}}{df_{Ri}} + \frac{d\Delta f_{Ai}}{df_{Ai}} < 0$ holds even if ndFDS is introduced in only one species.

We showed that differences in the prevalence of the parasite can introduce ndFDS through migration. This means that they have the potential to stabilize genetic polymorphism in the long term. Analytically we could not derive conditions for the stability of the equilibrium (see appendix) so we will explore the system with simulations.

3 Simulation studies

3.1 Simulation description

So far this model has been explored deterministically in both the one and two deme versions. To determine the effect of genetic drift and mutation in the context of this model, stochastic simulations were implemented for both scenarios. The additional parameters in these simulations are the population sizes, N_H and N_P of both the host and the parasite, and the mutation rates between the alleles. For all the results that will be shown it is assumed that the size of the parasite population equals the size of the host population. In the two deme model it is further assumed that the two demes have the same size.

At the start of each generation the allele frequencies for resistance and virulence, R and a, are obtained using the recursion equations that were explained before for each of the models (See (1), (2), (5) and (6)). Then, the number of resistant individuals is sampled from a binomial distribution with parameters $N = N_H$ and $p = R$ and the number of virulent individuals is sampled from a binomial distribution with parameters $N = N_P$ and $p = a$. In this way genetic drift is introduced to the model.

Once drift has occurred there is mutation in both the host and the parasite populations. Mutation occurs in both directions, this is, resistant individuals can become susceptible and *vice versa* and virulent individuals can become avirulent and *vice versa*. All the mutation rates are assumed to be equal. For the results that will be shown they were set to 1×10^{-5}. This corresponds to the probability per individual per generation of mutating.

After mutation the frequency of resistant and virulence individuals is calculated and a new generation starts.

3.1.1 Statistics calculated from the simulations

The statistics presented were calculated in simulations for 10000 generations. Over this period of time the following statistics were measured to assess the stability properties of the system and the speed of evolution. Percentage of time in fixation for R, r, A and a, which is the number of generations when the allele was fixed in the population over the total number of generations. Based on this measure it is possible to determine parameter regions for which there is stability because when the equilibrium is stable none of the alleles is fixed. For large population sizes mutation can avoid complete fixation of one allele even if the equilibrium is unstable, for this reason the percentage of time when the frequencies were near the boundaries (> 0.95) was also measured for all the alleles. The speed of evolution was measured by counting the total number of cycles for both the resistance and virulence allele over time. This was done by fitting a smooth spline curve to the trajectory of the allele frequencies over time. The function used to fit the spline curve is *smooth.spline* from the R package 'stats' and the smoothing parameter was 0.15.

Since the starting allele frequencies R_0 and A_0 can affect the behavior of the system, all the statistics are averaged over 100 runs where R_0 and A_0 are sampled from a uniform distribution over the interval $[0.1, 0.5]$.

3.2 Results

3.2.1 One deme scenario

As it was shown analytically, simulations confirmed that the equilibrium is unstable, so recurrent fixation of alleles was observed as in the arms race scenario (Figure 1).

High encounter rates favor virulent parasites because they select for resistant hosts. This was observed in the simulations where the time near fixation (> 0.95) of the virulence allele increased with the encounter rate (Figure 2). Since resistance hardly gets fixed even if the host encounters the parasite every generation (unpublished results), the effect of the encounter rate can be seen in the time near fixation (> 0.95) of the allele for susceptibility. This time decreases with higher encounter rates which then implies that resistance is favored (Figure 2).

When the prevalence of the parasite is low, the basal cost of resistance u^* plays a more important role in the dynamics of the system. Indeed, one of the questions we raised before was if polymorphism in resistance could be maintained when $u^* = 0$ because selection in the host is neutral when the parasite is not present. If the strength of positive selection acting on the host is weak i.e. when the parasite has low prevalence, polymorphism in host populations could be maintained by the effect of

drift and mutation. This prediction was confirmed in the simulations where stable polymorphism was observed when $u^* = 0$ (Figure 3).

Finally, it has been suggested that in GFG models resistance costs must exist, otherwise all the individuals would evolve to be resistant and polymorphism in natural populations could not be explained. We explored the case when the costs are zero, $u^* = u = 0$ and we found that for low parasite prevalence, polymorphism for the host can be stable in time even if there are no costs of resistance (Figure 4a). This result points again to the balance between neutral and positive selection in the host which depends on the encounter rate with the parasite. If the host population is frequently exposed to the parasite, there is strong selection for resistance which eventually gets fixed in the population (Figure 4b). On the other hand, if the host encounters the parasite sporadically, selection on resistant hosts will be mainly neutral which then can maintain polymorphism in resistance over time (Figure 4a). Note that stable polymorphism is only observed in the host population since the parasite is always virulent.

Speed of evolution: High parasite prevalence speeds up coevolutionary cycles over time (Figure 5). This was expected since iFDS, which is the main driver of change in allelic frequencies in this model, occurs only when the parasite encounters the host.

3.2.2 Two deme scenario

As it was already shown analytically, ndFDS can be introduced through migration between demes with different prevalence of the parasite even when all the other costs are equal. To determine under which conditions this can stabilize genetic polymorphism we explored the effect of modifying ρ and u^* in deme 1, on the stability properties of the system. In deme 2, these parameters were fixed to $\rho = 0.5$ and $u^* = 0.05$. All the other costs were equal in both demes to $u = b = 0.1, s = 0.4$ and $c = 0.95$. This was done for three different migration rates (0.01, 0.05, 0.1) and two different population sizes ($N = 1000, 5000$). In most of of the simulations fixation of avirulence and resistance was not observed (even with low population size), thereafter, we investigate stability properties of the system based on the alleles for virulence and susceptibility. All the graphs shown are for statistics measured in deme 1. Since the demes are coupled by migration it is enough to evaluate if there is stability in only one of them, to asses stability in both.

Stability increases with population size since the times near fixation decrease considerably for $N = 5000$ in comparison to $N = 1000$ (Figure 7). This is consistent with the observation that genetic drift can pulled frequencies away from the equilibrium even if this is stable. Besides, stability decreases with higher migration rates, which is expected since the oscillations for the two demes are

synchronized with high migration. Migration rates under 0.01 were not explored but its effect is expected to be the opposite, namely, that populations become independent. In both cases the system behaves as in the one deme scenario which is unstable.

Differences in the prevalence of the parasite can stabilize polymorphism even when u^* is equal between the two demes (see Figure 6 and Figure 7, red dotted line). This is an important observation because even if u^* is unknown in natural populations a large difference in the parasite prevalence between demes connected by migration can be enough to stabilize polymorphism. The main reason why differences in prevalence stabilize polymorphism is because the strength of selection for both virulence and resistance changes with the parasite encounter rate (Figure 2). Then, migration between demes with different prevalence can bring the frequencies to intermediate values introducing ndFDS in the system. The deme with lower parasite prevalence would act as coldspot - where evolution happens at a lower speed and the effect of iFDS is weak -, while the deme with higher parasite prevalence would be a hotspot. On the other hand, when ρ is constant differences only in u^* can also stabilize polymorphism but to a much lower extent. When ρ is constant, fixation times can decrease for some values of $u_1^* - u_2^*$ but they are never below 20% (Figure 7, black dotted line).

When both ρ and u^* differ between demes, stability was mainly found in three regions of the parameter range (Figure 8): a) $\rho_1 < \rho_2, u_1^* > u_2^*$ which corresponds to the lower right corner of the parameter space; b) $\rho_1 < \rho_2, u_1^* < u_2^*$, which is the lower left corner; c) $\rho_1 \approx 1, \rho_1 > \rho_2$, which is the upper part. In these three regions, the alleles for virulence and susceptibility are not fixed which means that there is long term polymorphism of the allele frequencies. The characteristics of the three regions are discussed next:

a) $\rho_1 < \rho_2, u_1^* > u_2^*$

In this case, there is a low prevalence of the parasite in deme 1 and carrying the allele for resistance is costly for the host. In deme 2, the conditions are the opposite because there is a higher prevalence of the parasite and carrying the allele for resistance is not very costly. This generates two opposite selection pressures in each of the demes: There is strong selection for hosts in deme 1 to become susceptible and for hosts in deme 2 to become resistant. Migration then maintains polymorphism because it introduces ndFDS in each deme avoiding fixation of susceptible hosts in deme 1 and resistant hosts in deme 2. This is the region of the parameter space where there is always stability, even for low population size and high migration rates (Figure 7).

b) $\rho_1 < \rho_2, u_1^* < u_2^*$

In this case, there is a low prevalence of the parasite in deme 1 but carrying the allele for resistance

is cheap for the host while in deme 2, the parasite has a high prevalence and carrying resistance is costly. These conditions generate fitness differences between the host and the parasite populations. In deme 1, the host population has a higher mean fitness than the parasite population, contrary to deme 2 where the parasite does best.

c) $\rho_1 \approx 1, \rho_1 > \rho_2$

In this case deme 1 behaves as the simple GFG model (parasite encounter rate is assumed to be one) which selects for virulent parasites and susceptible hosts. Migration from a deme with lower prevalence of the parasite (where the frequency of avirulence is higher) prevents virulence to be fixed and favors resistant hosts in deme 1 stabilizing polymorphism.

Determining the biological systems in which each of the scenarios mentioned can occur requires a better understanding of the genetic system of host populations. Then, it would be possible to determine whether there is large variation in the basal cost of resistance between populations and which type of factors (e.g. genetical, ecological) are responsible for this variation.

Speed of evolution: It was found that migration from demes with high prevalence of the parasite can increase the speed of evolution in demes with low prevalence of the parasite (Figure 9). This is consistent with the analytical results for the equilibrium frequency that predicted a larger effect of migration for demes with lower ρ.

4 Discussion

Most of the models [11, 13, 31, 32] in host-parasite coevolution assume that the host is constantly exposed to the parasite, although there is much empirical evidence showing that this is not the case in natural populations [17, 18, 19, 22]. Parasite survival, dispersal and transmission are highly dependent on environmental conditions [14, 15], so infection rates in the host population can be very small. Indeed, it is not rare that a potential host never encounters a parasite during its lifetime [16]. When the host is not exposed to the parasite there is no indirect FDS because the fitness of the host does not depend on the frequency of virulence in the parasite population. For this reason, the parasite prevalence can have a large effect on the coevolutionary dynamics of host-parasite interactions.

By accounting for parasite prevalence in the simple GFG model we showed that the parasite encounter rate affects the strength of selection for resistance and virulence alleles in the host and parasite populations, respectively. High prevalence of the parasite selects for virulent parasites by

favoring resistance, while low prevalence of the parasite selects for avirulent parasites by favoring susceptibility (Figure 2). This is consistent with results of laboratory experiments in plant populations where resistance was found to be positively correlated with the rate of parasite encounter [14].

If the rate of parasite encounter is constant in time (as we assume), negative direct FDS is not introduced in the simple GFG model. This means that if the prevalence of the parasite does not vary over time, the model predicts an arms race scenario where there is recurrent fixation of alleles in both the parasite and the host (Figure 1). Theoretical models have shown that temporal variation of selection coefficients due to changes in the environment can increase stability of genetic polymorphism in a single population [33]. Since temporal variability in the parasite encounter rate is a common feature of natural populations [15], we suggest that exploring the stability properties of a single population model with heterogeneity in parasite prevalence over time is an important topic for further research.

4.1 Statistical polymorphism in resistance in a single population

Genetic polymorphism can be maintained in the absence of ndFDS, however, if there is weak selection for resistance in the host and genetic drift has a large effect [3]. This is known as "statistical polymorphism" and can emerge in multilocus GFG systems where there are no costs of resistance or virulence [34], or where there are epistatic interactions between loci associated with the costs [38]. Ecological factors such as spatial heterogeneity and the existence of pacemaker demes can also account for the existence of this type of polymorphism [13, 32, 37].

Here, we proposed a simpler scenario where statistical polymorphism can occur in a single population with one locus. This corresponds to a population where the parasite is in low prevalence and the cost of basal resistance is zero. In this scenario, selection does not occur in the host population when the parasite is absent because carrying the resistance allele has no cost, so susceptible and resistant hosts have the same fitness. Since the parasite is in low prevalence, the effect of neutral selection in the host is higher than the effect of positive selection for resistance and genetic polymorphism can be maintained over time by drift and mutation (Figure 3).

Studies in plant-pathogen systems have found that there are high levels of resistance in populations that have not been exposed to the parasite for many generations [22]. These results are consistent with the scenario proposed before because if carrying resistance genes would be costly, resistant individuals would be eliminated by natural selection in the absence of the parasite. Finally,

note that this does not imply that the cost of expression is also zero. For this reason we suggest that more empirical studies should be done to measure separately the cost of carrying and expressing resistance genes.

4.2 Spatial heterogeneity in parasite prevalence

Spatial heterogeneity has been recognized as an important mechanism for stabilizing genetic polymorphism in host-parasite interactions [3, 13, 32]. Selection pressures acting on host and parasite populations can change between populations with different environments which has the potential of introducing ndFDS through migration. Heterogeneity has been mainly studied in the context of variation in costs, with models accounting for differences in the cost of being diseased [32], and in the costs of resistance and virulence between demes [13].

Our results suggest an alternative scenario that does not require differences in the costs between populations, which are in general hard to prove in nature [35]. We showed that even if all the costs are equal between demes, heterogeneity in parasite prevalence can introduce ndFDS through migration. This has the potential of generating long-term balanced genetic polymorphism if the differences in prevalence are large enough between the populations (Figure 6 and see Figure 7, red dotted line). If polymorphism is maintained, variation of allele frequencies over time would be consistent with the trench warfare model of evolution. This result implies that even if the costs of resistance, virulence or disease are unknown in a natural system, only large heterogeneity in the prevalence of the parasite would be enough to stabilize genetic polymorphism.

Large differences in parasite prevalence are plausible in natural populations connected by migration. In fact, spatial variation in parasite prevalence has been observed at a regional scale [15, 19, 36] and even at smaller scales for neighboring plant populations [14]. This has been explained because the prevalence of the parasite, at least in plant-pathogen systems, can be very sensitive to microclimatic conditions [14]. In the case of species that perform long migrations, variation in the parasite encounter rate between sites can also be common. For instance, in birds, it has been found that diversity and abundance of blood parasites is higher in migrant than non-migrant individuals [5].

If spatial heterogeneity in the costs could be demonstrated in natural populations, considering the prevalence of the parasite would be still important. As it was shown for the cost of basal resistance, if both u^* and the encounter rate differ between populations, stability of genetic polymorphism can be achieved in a larger area of the parameter space compared to the case when only one of the two changes (Figure 7 and Figure 8). Some preliminary simulations showed that this is also the case for

coupled variation of the encounter rate and the costs u, s, b and c.

4.3 Speed of evolution

The speed of evolution in host-parasite systems has raised some attention because it could be useful to distinguish between contrasting evolutionary scenarios in natural populations. Indeed, it has been proposed that coevolution in the trench warfare scenario is in general faster than in the arms race model [39]. What our results suggest is that one additional factor that should be taken in consideration when studying the speed of evolution in natural populations is the prevalence of the parasite in the population. Here, we showed that parasite encounter rate has a large effect on the speed of evolution in a GFG model. Higher prevalence of the parasite speeds up coevolutionary cycles even when all the other parameters are constant (Figure 5), due to the effect of iFDS which is not present when the host does not encounter the parasite. Since variation in parasite prevalence is common in the wild [14, 15, 19, 22, 36], parasite encounter rates are likely to have an important effect on the speed of evolution observed in natural populations.

In terms of the speed of evolution in spatial heterogenous landscapes, our results showed that migration from demes with high prevalence of the parasite, increased the speed of evolution in demes with low prevalence (Figure 9). This is consistent with experimental studies of the effect of migration between demes with different strength of selection in bacteria populations [21, 20]. There, it was shown that the speed of evolution in coldspots (demes where reciprocal selection is weak) can increase through migration from hotspots (demes where reciprocal selection is strong). This pattern has also been predicted by theoretical models on coevolution in heterogeneous landscapes [23, 40]. In our model hotspots would correspond to demes with high parasite prevalence, and coldspots to demes with low prevalence.

In conclusion, accounting for parasite prevalence in the simple GFG model can stabilize genetic polymorphism. We showed that "statistical polymorphism" is observed when the cost of resistance in the absence of the parasite is zero in a single population. In the case of spatially structured populations, differences in parasite prevalence between demes can stabilize polymorphism by introducing ndFDS through migration. Further empirical studies should be done to determine how selection in the host population operates when the parasite is not present and which conditions can account for spatiotemporal variation of the parasite encounter rate. Parasite prevalence is a factor that can influence the stability and the general dynamics of host-parasite interactions as it was shown in

this work. For this reason, it should not be ignored when studying host-parasite coevolution both theoretically and in natural populations.

References

[1] Stahl E., Dwyer G., Mauricio R., Kreitman M., Bergelson J. 1999. *Dynamics of disease resistance polymorphism at theRpm1 locus of Arabidopsis* Nature, Vol. 400.

[2] Holub E. 2001. *The arms race is ancient history in Arabidopsis, the wildflower*. Nat Rev Genet., 2:516 - 527.

[3] Brown J., Tellier A. 2011. Plant-Parasite Coevolution: Bridging the Gap between Genetics and Ecology. Annu. Rev. Phytopathol. 49:34567

[4] Bockhurst M. Rainey P., Buckling A. 2004. *The effect of spatial heterogeneity and parasites on the evolution of host diversity* Proc. R. Soc. Lond. B. 271, 107 - 111

[5] Jenkins T., Thomas G., Hellgren O., Owens I. 2011. *Migratory behavior of birds affects their coevolutionary relationship with blood parasites*. Evolution $66 - 3 : 740 - 751$.

[6] Tschirren B., Andersson M., Scherman K., Westerdahl H.,Raberg L. 2011. *Contrasting patterns of diversity and population differentiation at the innate immunity gene toll-like receptor 2 (TLR2) in two sympatric rodent species*. Evolution $66 - 3 : 720 - 731$

[7] Bergelson J, Dwyer G, Emerson JJ. 2001. *Models and data on plant-enemy coevolution*. Annu. Rev. Genet. 35:469 - 99

[8] Thompson JN, Burdon JJ. 1992. Gene-for-gene coevolution between plants and parasites. Nature 360:121 - 25

[9] Wilfert L, Jiggins FM. 2010. *Host-parasite coevolution: genetic variation in a virus population and the interaction with a host gene*. J. Evol. Biol. 23:144755

[10] Leonard, K. J. 1977 Selection pressures and plant pathogens. Ann. NYAcad. Sci. USA 287, 207 - 222.

[11] Tellier A., Brown J. 2007. *Stability of genetic polymorphism in host - parasite interactions*. Proc. R. Soc. B. 274, $809 - 817$.

[12] Tellier A., Brown J. 2009. *The Influence of Perenniality and Seed Banks on Polymorphism in Plant-Parasite Interactions*. The American Naturalist. Vol. 174, no. 6.

[13] Tellier A., Brown J. 2011. *Spatial heterogeneity, frequency-dependent selection and polymorphism in host-parasite interactions* BMC Evolutionary Biology 11:319

[14] Laine A. 2006. *Evolution of host resistance: looking for coevolutionary hotspots at small spatial scales.* Proc. R. Soc. B. 273, 267 – 273

[15] Soubeyrand S., Laine A., Hanski I., Penttinen A. 2009. *Spatiotemporal Structure of Host-Pathogen Interactions in a Metapopulation.* The American Naturalist. Vol 174, No 3.

[16] Laine A, Tellier A. 2008. *Heterogeneous selection promotes maintenance of polymorphism in host-parasite interactions* Oikos, 117:1281 - 1288.

[17] Arneberg P., Skorping A., Grenfell B., Read A. 1998. *Host densities as determinants of abundance in parasite communities* Proc. R. Soc. Lond. B. 265, 1283 – 1289

[18] Little T., Ebert D. 2000. *The cause of parasitic infection in natural populations of Daphnia (Crustacea: Cladocera): the role of host genetics* Proc. R. Soc. Lond. B. 267,2037 – 2042

[19] Thrall P., Laine A., Ravensdale M., Nemri A., Dodds P., Barrett L., Burdon J. 2012. *Rapid genetic change underpins antagonistic coevolution in a natural host - pathogen metapopulation* Ecology Letters, 15 : 425 – 435

[20] Brockhurst M., Buckling A., Poullain V., Hochberg M. 2006. *The impact of migration from parasite - free patches on antagonistic host - parasite coevolution* Evolution 61- 5: 1238 - 1243

[21] Vogwill T., Fenton A., Buckling A., Hochberg M.., Brockhurst M. 2009. *Source Populations Act as Coevolutionary Pacemakers in Experimental Selection Mosaics Containing Hotspots and Coldspots.* The American Naturalist. Vol. 173, no. 5.

[22] Laine A. 2004. *Resistance variation within and among host populations in a plantpathogen metapopulation: implications for regional pathogen dynamics* Journal of Ecology 92, 990 - 1000

[23] Thompson JN. 2005. The Geographic Mosaic of Coevolution. Chicago, IL. Univ. Chicago Press

[24] Olson M., Wapstra E., Madsen T., Ujvari B., Rugfelt C. 2005. *Costly parasite resistance: a genotype-dependent handicap in sand lizards?* Biol. Lett. 1,375 – 377

[25] Langand J., Jourdane J., Coustau C., Delay B., Morand S. 1998. *Cost of resistance, expressed as a delayed maturity, detected in the host - parasite system Biomphalaria glabrata/ Echinostoma.* Heredity 80, 320 - 325

[26] Ebert D. 2005. *Host Adaptations against the Costs of Parasitism.* Chapter 6. In: Ebert D. 2005. Ecology, Epidemiology, and Evolution of Parasitism in Daphnia. Bethesda (MD) National Center for Biotechnology information. ISBN - 10: 1 - 932811 - 06 - 0

[27] Jokela J, Lively C. 1995. *Parasites, sex, and early reproduction in a mixed population od freshwater snails.* Evolution Int J Org Evolution. 49:1268 - 1271.

[28] Tsuda K, Sato M, Stoddard T, Glazebrook J, Katagiri F. 2009. *Network properties of robust immunity in plants.* PLoS Genet. 5:e1000772

[29] Katagiri F, Tsuda K. 2010. *Understanding the plant immune system.* Mol. Plant-Microbe Interact. 23:1531 - 36

[30] Frank S. 1993. *Coevolutionary genetics of plants and pathogens.* Evol Ecol, 7: 45 - 75.

[31] Nuismer S. 2006. *Parasite local adaptation in a geographic mosaic* Evolution 60, 24 - 30.

[32] Gavrilets S., Michalakis Y. 2008. *Effects of environmental heterogeneity on victim-exploiter coevolution* Evolution 62 - 12: 3100 - 3116

[33] Mostowy R., Engelstadter J. 2011. *The impact of environmental change on hostparasite coevolutionary dynamics* Proc. R. Soc. B. 278,2283 − 2292

[34] Salathe M., Scherer A., Bonhoeffer S. 2005. *Neutral drift and polymorphism in gene-for-gene systems* Ecology Letters, 8 : 925 − 932

[35] Bergelson J, Purrington C. 1996. *Surveying patterns in the cost of resistance in plants.* Am Nat, 148:536 - 558.

[36] Laine A., Hanski I. 2006. *Large-scale spatial dynamics of a specialist plant pathogen in a fragmented landscape* J. Ecol. 94: 217 - 226

[37] Damgaard C. 1999. *Coevolution of a Plant Host - Pathogen Gene-for-gene System in a Metapopulation Model without Cost of Resistance or Cost of Virulence* J. Theor. Biol. 201,1 − 12.

[38] Tellier A, Brown J. 2007. *Polymorphism in multilocus host-parasite coevolutionary interactions.* Genetics 177:1777 - 90

[39] Ebert D. 2008. *Host-parasite coevolution: insights from the Daphnia-parasite model system.* Curr. Opin. Microbiol. 11: 290 - 301.

[40] Gomulkiewicz, R., Thompson J., Holt R., Nuismer S, Hochberg M. 2000. *Hot spots, cold spots, and the geographic mosaic theory of coevolution.* American Naturalist 156 : 156 − 174.

5 Figures and tables

Table 1: Fitness of hosts and parasites for this model. When the host encounters the parasite, which happens with probability ρ, the model corresponds to the simple GFG model. When the host does not encounter the parasite, the fitness of both species is independent of the frequencies in the other species. It is assumed that if the parasite does not encounter the host, it has fitness zero. *na* stands for not applicable (because there is no encounter of both species).

	Genotypes (frequency)		Fitness	
Parasite encounter	Host	Parasite	Host	Parasite
ρ	*RES* (R_g)	*AVR* (A_g)	$1-u$	$1-c$
		avr (a_g)	$(1-u)(1-s)$	$1-b$
	res (r_g)	*AVR* (A_g)	$1-s$	1
		avr (a_g)	$1-s$	$1-b$
$1-\rho$	*RES* (R_g)		$1-u^*$	*na*
	res (r_g)		1	*na*
		AVR (A_g)/*avr* (a_g)	*na*	0

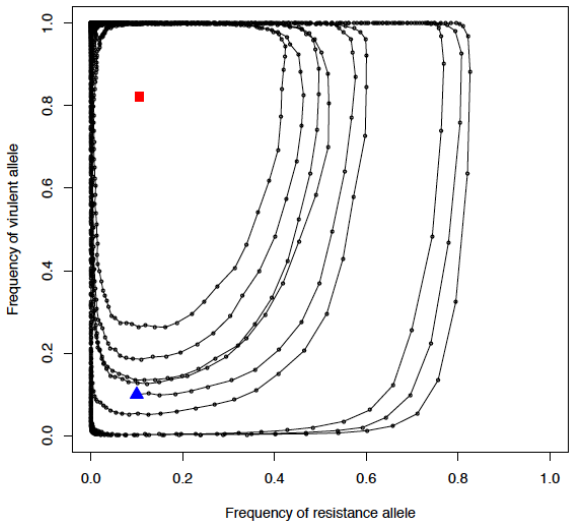


Figure 1: Phase plot of the frequencies of virulence and resistance. The blue triangle indicates the starting allele frequencies (R_0,a_0) and the red square corresponds to the equilibrium frequencies ($\hat{R}$,$\hat{a}$). The equilibrium is unstable so recurrent fixation of alleles is observed. The parameters are N=10000, $\rho = 0.7$, $u^* = 0.01$ $u = b = 0.1, s = 0.4, c = 0.95$.

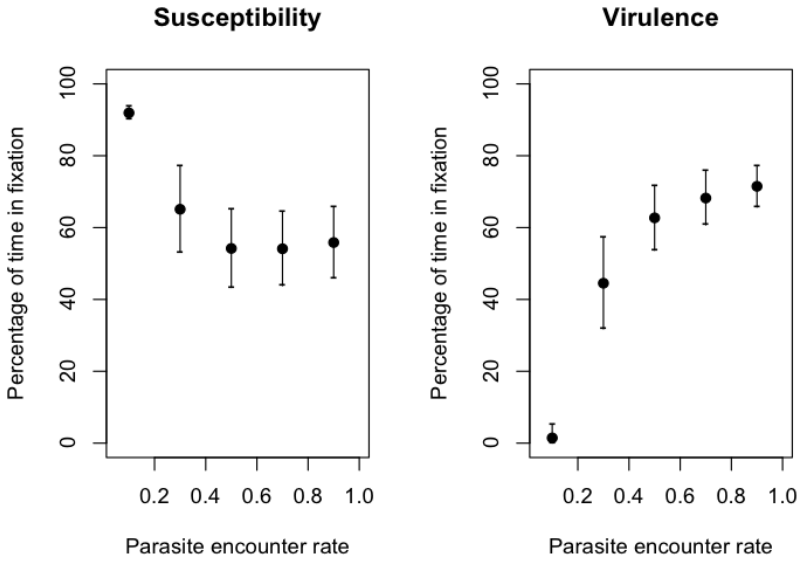


Figure 2: Effect of the parasite encounter rate in the fixation time of the alleles for susceptibility and virulence. The mean and standard deviation were calculated over 100 runs with the same parameters but different initial allele frequencies. The other parameters were set to $N = 1000, u^* = 0.05, u = b = 0.1, s = 0.4, c = 0.95$.

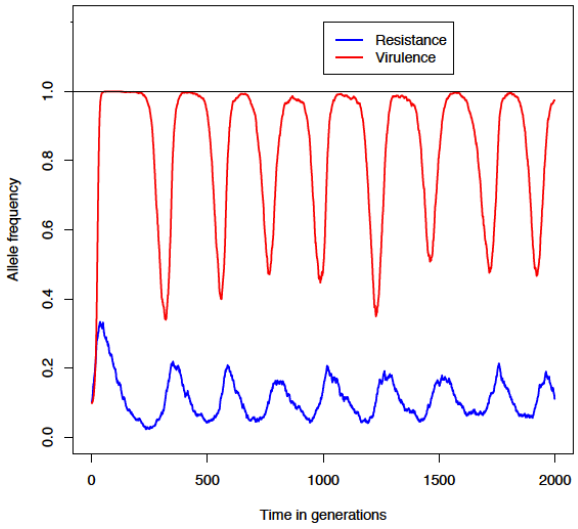
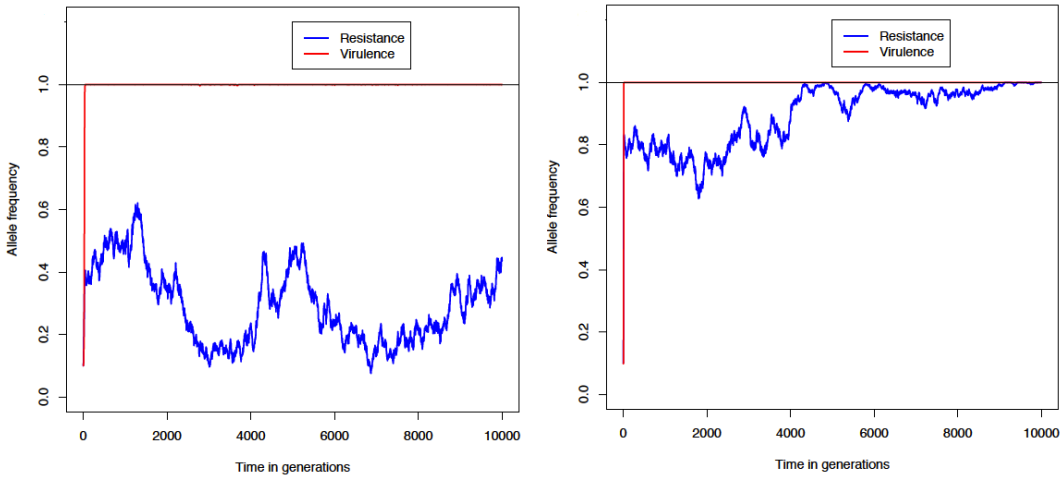


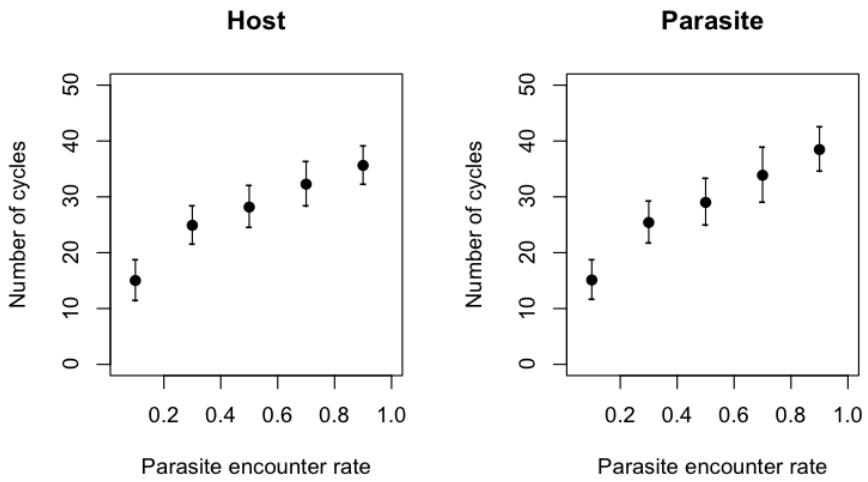
Figure 3: Polymorphism of the allele for resistance in the host population when there is low prevalence of the parasite and the basal cost of resistance is zero ($\rho = 0.2$, $u^* = 0$). Since selection in the host does not occur when the parasite is absent, most of the change in the frequencies of resistance is driven by drift and mutation. The other parameters were fixed to N=10000, $u = b = 0.1, s = 0.4, c = 0.95$.



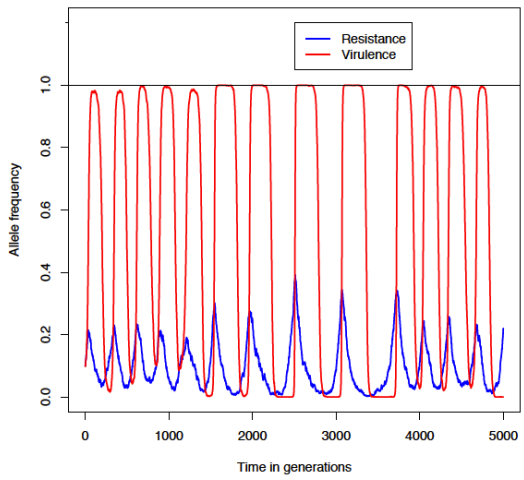
(a) Low parasite prevalence, $\rho = 0.2$

(b) High parasite prevalence, $\rho = 1$

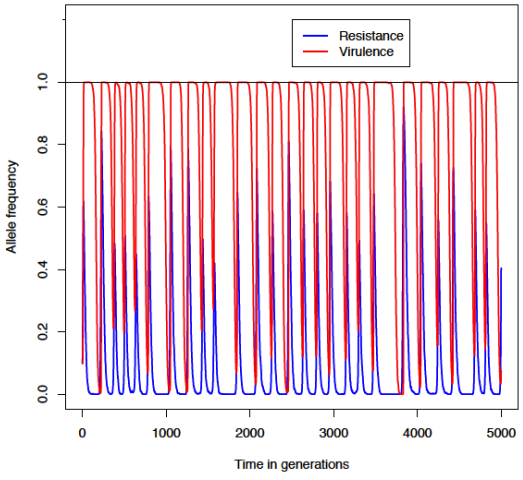
Figure 4: Low prevalence of the parasite can maintain the polymorphism in the host even when there is no cost associated to resistance $u^* = u = 0$. Otherwise, if the prevalence is high, resistance becomes fixed in the host population. Note that virulence is always fixed in the parasite population. The parameters used were $N = 10000$, $b = 0.1, s = 0.4, c = 0.95$.



(a)

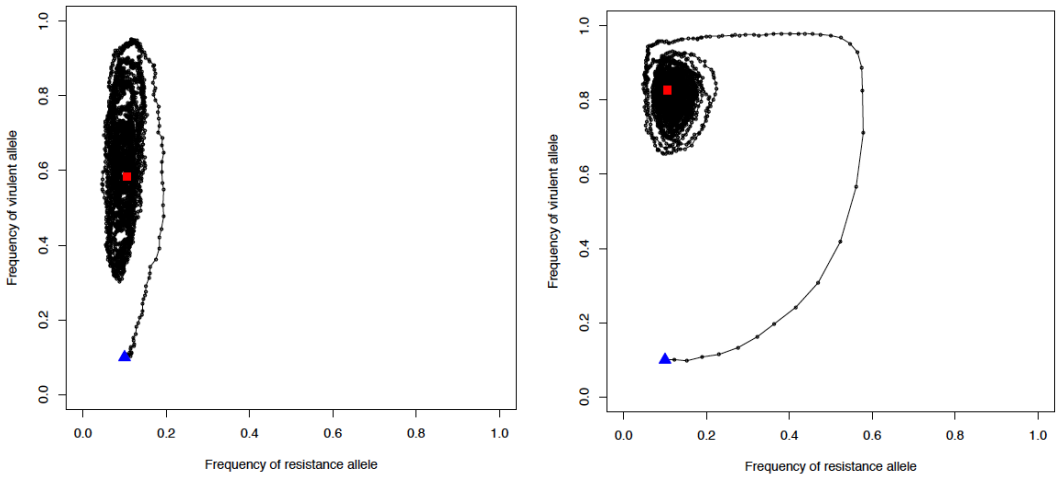


(b) Low parasite prevalence, $\rho = 0.2$



(c) High parasite prevalence, $\rho = 0.8$

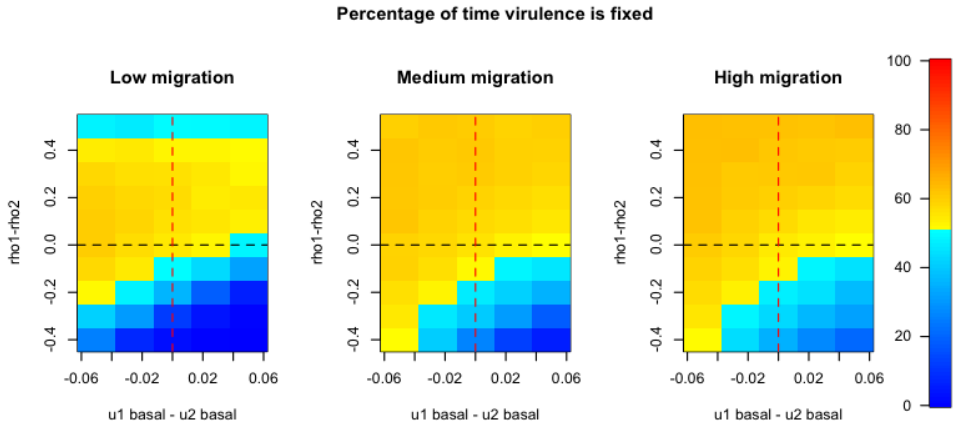
Figure 5: The number of cycles of the allele frequencies increases when the prevalence of the parasite is higher (a). Then, the parasite prevalence speeds up co-evolution. For instance, in b) there is low prevalence of the parasite, while in c) the prevalence is high. In figure a), the mean and standard deviation were calculated over 100 runs with the same parameters but different initial allele frequencies. All the other parameters were set to N=10000, $u^* = 0.01$ $u = b = 0.1, s = 0.4, c = 0.95$ in the three figures.



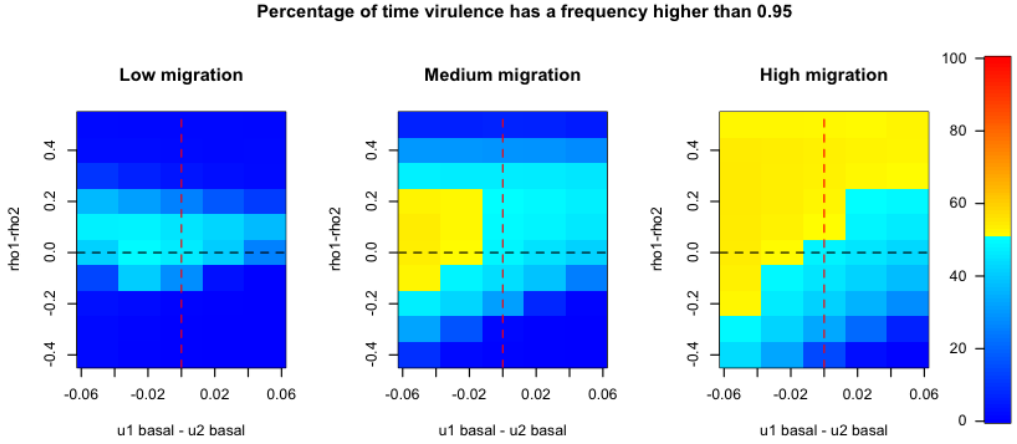
(a) Parasite prevalence, $\rho_1 = 0.4$

(b) Parasite prevalence, $\rho_2 = 0.8$

Figure 6: Example of a scenario where only differences in parasite prevalence stabilize genetic polymorphism. In deme 1 prevalence is $\rho_1 = 0.4$ and in deme 2 prevalence is $\rho_2 = 0.8$. Migration rates are $m_H = m_P = 0.01$ All the other parameters are equal to N=10000, $u^* = 0.01$ $u = b = 0.1, s = 0.4, c = 0.95$ in both demes.



(a) N=1000



(b) N=5000

Figure 7: Effect of the magnitude of differences in parasite prevalence (ρ_1- ρ_2) and cost of basal resistance ($u_1^* - u_2^*$) on the stability of genetic polymorphism. $N = 1000$ in the first figure and $N = 5000$ in the second figure. The percentage of time above 0.95 is plotted when $N = 5000$ because due to mutation ($N\mu$) alleles are hardly fixed even if the equilibrium is unstable (as opposed to when N=1000 where the mutational input is lower and complete fixation is observed). Low, medium and high migration rates correspond to 0.01, 0.05 and 0.1 respectively ($m_H = m_P$). The red dotted line indicates the case where only the parasite encounter rate differs between demes. The black dotted line indicates the case where only the basal cost of resistance, u^* differs between demes. The parameters in deme 2 are fixed to $\rho_2 = 0.5$, $u_2^* = 0.05$ so only ρ_1 and u_1^* change. All the values correspond to the mean over 100 runs with the same parameter combination but different initial allele frequencies. The other parameters are equal between demes and fixed to $u_1 = u_2 = b_1 = b_2 = 0.1$ $s_1 = s_2 = 0.4, c_1 = c_2 = 0.95$.

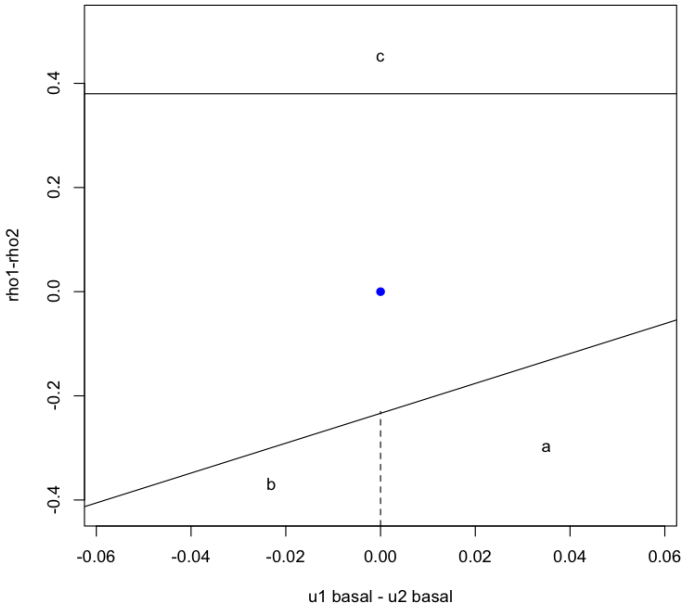
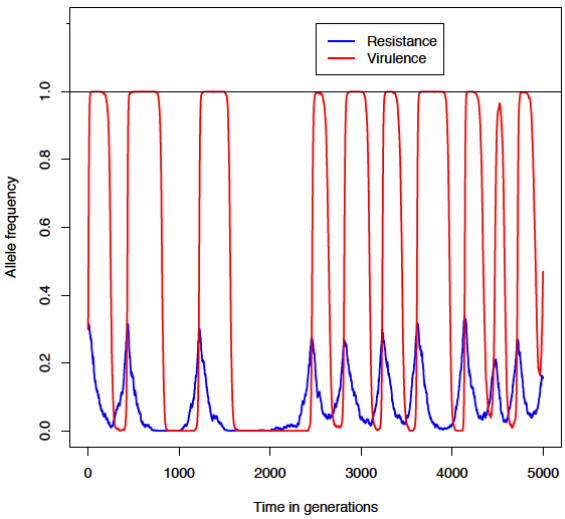
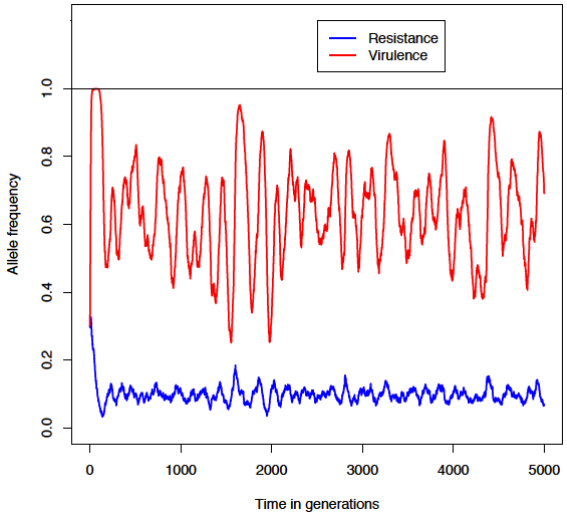


Figure 8: Regions where stability increases when both ρ and u^* differ between the populations. a) $\rho_1 < \rho_2, u_1^* > u_2^*$, b) $\rho_1 < \rho_2, u_1^* < u_2^*$ and c) $\rho_1 \approx 1, \rho_1 > \rho_2$.



(a) Single population, $\rho = 0.2$



(b) Same population with migration from a hotspot where $\rho = 0.8$

Figure 9: In the single population (graph a) the prevalence of the parasite is low ($\rho = 0.2$) so reciprocal selection is weak and the deme behaves as a coldspot. When this deme receives migration from a hotspot ($\rho = 0.8$) the speed of evolution increases with respect to the case without migration (graph a). In both the hotspot and coldspot, $N = 10000$, $u^* = 0.01$, $u = b = 0.1, s = 0.4, c = 0.95$. Migration rates are $m_H = m_P = 0.01$.

6 Appendix: Analytical calculations

6.1 Stability analysis for the one deme model

The recurrence equations are,

Host

$$\frac{R_{g+1}}{r_{g+1}} = \frac{R_g}{r_g}\left(\frac{\rho(1-u)(1-s+sA_g)+(1-\rho)(1-u^*)}{1-\rho s}\right) \tag{11}$$

Parasite

$$\frac{A_{g+1}}{a_{g+1}} = \frac{A_g}{a_g}\left(\frac{1-cR_g}{1-b}\right) \tag{12}$$

The dynamics of the system was analyzed using its Jacobian matrix. The following logit transformations were made:

$$f_R = \log\frac{R_g}{r_g} \text{ and } f_A = \log\frac{A_g}{a_g}$$

Then, the system can be written as,

$$\begin{aligned}
\Delta f_R &= f_{R+1} - f_R \\
&= \log\frac{R_{g+1}}{1-R_{g+1}} - \log\frac{R_g}{1-R_g} \\
&= \log\frac{R_{g+1}r_g}{R_g r_{g+1}} \\
&= \log(\rho(1-u)(1-s+sA_g)+(1-\rho)(1-u^*)) - \log(1-\rho s)
\end{aligned}$$

$$\begin{aligned}
\Delta f_A &= f_{A+1} - f_A \\
&= \log\frac{A_{g+1}}{1-A_{g+1}} - \log\frac{A_g}{1-A_g} \\
&= \log(1-cR_g) - \log(1-b)
\end{aligned}$$

From the definition of f_A and f_R it follows that,

$$\frac{dR}{df_R} = R_g r_g$$
$$\frac{dA}{df_A} = A_g a_g$$

Using the chain rule, the coefficients of the Jacobian matrix are calculated as follows:

$$\frac{d\Delta f_R}{df_R} = \frac{d\Delta f_R}{dR}\frac{dR}{df_R} = 0$$
$$\frac{d\Delta f_R}{df_A} = \frac{d\Delta f_R}{dA}\frac{dA}{df_A} = \frac{A_g a_g (1-u)\rho s}{\rho(1-u)(1-s+sA_g)+(1-\rho)(1-u^*)}$$
$$\frac{d\Delta f_A}{df_R} = \frac{d\Delta f_A}{dR}\frac{dR}{df_R} = -\frac{cR_g(1-R_g)}{1-cR_g}$$
$$\frac{d\Delta f_A}{df_A} = \frac{d\Delta f_A}{dA}\frac{dA}{df_A} = 0$$

First, note that this model does not introduce direct-frequency dependence since $d\Delta f_R/df_R = d\Delta f_A/df_A = 0$. The discriminant of the characteristic equation of the matrix near the equilibrium is:

$$\begin{aligned} Q_A &= 4\left(\frac{\hat{A}_g(1-\hat{A}_g)(1-u)\rho s}{(1-u)(1-\rho s+\rho s\hat{A}_g)+(1-\rho)\epsilon}\right)\left(\frac{-c\hat{R}_g(1-\hat{R}_g)}{1-c\hat{R}_g}\right) \\ &= -4\left(\frac{(\rho s-u+\epsilon(1-\rho))(u(1-\rho s)-\epsilon(1-\rho))}{\rho s(1-u)(1-\rho s)}\right)\left(\frac{b(c-b)}{c(1-b)}\right) \end{aligned} \tag{13}$$

If $u^*+\rho\epsilon < \rho s$, which is equivalent to the condition $s < u$ in the simple GFG model, the expression $(\rho s - u + \epsilon(1-\rho))$, in the first term of (13) is positive and $Q_A < 0$. Note that $(u(1-\rho s)-\epsilon(1-\rho))$ is positive if all the costs are assumed to be less than 1 since $u > \epsilon$ and $1-\rho s > 1-\rho$. The second term in (13) is also positive because the costs are less than 1 and also because $c > b$ which is necessary for virulence to evolve in the first place. Since the diagonal elements of J_A are zero, and both eigenvalues are complex with no real parts the internal equilibrium is unstable.

6.2 Calculation of the equilibrium for the two deme model

a) $m_P = 0$

Replacing m_P=0 in (6), $A_1' = \tilde{A}_1$ and $A_2' = \tilde{A}_2$. This is, the frequencies in the next generation for the parasite population are equal to the frequencies after selection because the parasite does not migrate. Since the allelic frequencies in one species determine the equilibrium in the other species, $\hat{R}_i$ for both demes is independent of the migration rate and corresponds to the value in the simple one deme model. Then,

$$\hat{R}_1 = \frac{b_1}{c_1} \text{ and } \hat{R}_2 = \frac{b_2}{c_2}$$

At the equilibrium for R_i (See (5)),

$$\frac{\hat{R}_i}{1-\hat{R}_i} = \frac{(1-m_H)\tilde{R}_i + m_H\tilde{R}_j}{(1-m_H))(1-\tilde{R}_j) + m_H(1-\tilde{R}_j)}$$

where,

$$\tilde{R}_i = \hat{R}_i((1-u_i)(1-\rho_i s_i + \rho_1 s_i \hat{A}_i) + (1-\rho_i)\epsilon_i)$$

$$\tilde{r}_i = \hat{r_i}(1-\rho_i s_i)$$

$\hat{A}_i$ can be obtained from the previous equation, since it appears in the term for $\tilde{R}_i$. After simplifying and reorganizing the equation, we obtained,

$$\hat{A}_i = \frac{u_i(1-\rho_i s_i) - \epsilon_i(1-\rho_i)}{(1-u_i)\rho_i s_i} + \frac{m_H}{(1-m_H)}\frac{1}{(1-u_i)\rho_i s_i}\left(\frac{\hat{r}_j}{\hat{r}_i}(1-\rho_j s_j) - \frac{\hat{R}_j}{\hat{R}_i}((1-u_j)(1-\rho_j s_j + \rho_j s_j \hat{A}_j) + (1-\rho_j)\epsilon_j)\right)$$

Note that the first term corresponds to the equilibrium calculated for the one deme model (See the expression for the equilibrium in section 2.1.1). We will call this term $\hat{A_i^{od}}$.

a) $m_H = 0$

Replacing m_H=0 in (5), $R_1' = \tilde{R}_1$ and $R_2' = \tilde{R}_2$. This is, the frequencies in the next generation for the host population are equal to the frequencies after selection because the host does not migrate. Since the allelic frequencies in one species determine the equilibrium in the other species, $\hat{A}_i$ for both demes is independent of the migration rate and corresponds to the value in the simple one deme model. Then,

$$\hat{A}_i = \frac{u_i(1-\rho_i s_i) - \epsilon_i(1-\rho_i)}{\rho_i s_i(1-u_i)} \text{ and } \hat{A}_j = \frac{u_j(1-\rho_j s_j) - \epsilon_j(1-\rho_j)}{\rho_j s_j(1-u_j)}$$

At the equilibrium for A_i (See (6)),

$$\frac{\hat{A}_i}{1-\hat{A}_i} = \frac{(1-m_P)\tilde{A}_i + m_P\tilde{A}_j}{(1-m_P)(1-\tilde{A}_j) + m_P(1-\tilde{A}_j)}$$

where,

$\tilde{A}_i = \hat{A}_i(1 - c_i\hat{R}_i)$

$\tilde{a_i} = \hat{a}_i(1 - b_i)$

$\hat{R}_i$ can be obtained from the previous equation, since it appears in the term for $\tilde{A}_i$. After simplifying and reorganizing the equation, we obtained,

$$\hat{R}_i = \frac{b_i}{c_i} - \frac{m_P}{(1-m_P)}\frac{1}{c_i}\left(\frac{\hat{a}_j}{\hat{a}_i}(1-b_j) - \frac{\hat{A}_j}{\hat{A}_i}(1-c_j\hat{R}_j)\right)$$

The first term corresponds to the equilibrium calculated for the one deme model (See (3)). We will call this term $\hat{R_i^{od}}$.

6.3 Stability analysis for the two deme model

The recurrence equations are,

Host

$$\frac{R_i'}{r_i'} = \frac{(1-m_H)\tilde{R}_i + m_H\tilde{R}_j}{(1-m_H)\tilde{r_i} + m_H\tilde{r_j}} \tag{14}$$

Parasite

$$\frac{A_i'}{a_i'} = \frac{(1-m_P)\tilde{A}_i + m_P\tilde{A}_j}{(1-m_P)\tilde{a_i} + m_P\tilde{a_j}} \tag{15}$$

where,

$\tilde{R}_i = R_i((1-u_i)(1-\rho_i s_i + \rho_i s_i A_i) + (1-\rho_i)\epsilon_i)$

$\tilde{r}_i = r_i(1-\rho_i s_i)$

$\tilde{A}_i = A_i(1-c_i R_i)$

$\tilde{a}_i = a_i(1-b_i)$

The following transformations were made to simplify the calculations,

$$f_{Ri} = \log\frac{R_i}{r_i} \text{ and } f_{Ai} = \log\frac{A_i}{a_i}$$

Then the changes of f_{Ri} and f_{Ai} are,

$$\begin{aligned}\Delta f_{Ri} &= f_{R'i} - f_{Ri} \\ &= \log((1-m_H)\tilde{R}_i + m_H\tilde{R}_j) - \log((1-m_H)\tilde{r_i} + m_H\tilde{r_j}) - f_{Ri} \\ \Delta f_{Ai} &= f_{A'i} - f_{Ai} \\ &= \log((1-m_P)\tilde{A}_i + m_P\tilde{A}_j) - \log((1-m_P)\tilde{a_i} + m_P\tilde{a_j}) - f_{Ai}\end{aligned}$$

From the definition of f_{Ai} and f_{Ri} it follows that,

$$\begin{aligned}\frac{dR_i}{df_{Ri}} &= R_i r_i \\ \frac{dA_i}{df_{Ai}} &= A_i a_i\end{aligned}$$

The diagonal elements of the Jacobian matrix for deme i are:

$$\begin{aligned}\frac{d\Delta f_{Ri}}{df_{Ri}} &= \frac{d\Delta f_{Ri}}{dR_i}\frac{dR_i}{df_{Ri}} \\ &= R_i r_i \left(\frac{((1-u_i)(1-\rho_i s_i a_i) + (1-\rho_i)\epsilon_i)(1-m_H)}{(1-m_H)\tilde{R}_i + m_H\tilde{R}_j} + \frac{(1-\rho_i s_i)(1-m_H)}{(1-m_H)\tilde{r_i} + m_H\tilde{r_j}} \right) - 1\end{aligned}$$

This can be simplified as,

$$\frac{d\Delta f_{Ri}}{df_{Ri}} = -\frac{\tilde{R}_i R_i (1-m_H)\tilde{r_j} m_H + \tilde{r_i} r_i (1-m_H)\tilde{R}_j m_H + \tilde{R}_j \tilde{r_j} m_H^2}{((1-m_H)\tilde{R}_i + m_H\tilde{R}_j)((1-m_H)\tilde{r_i} + m_H\tilde{r_j})}$$

If $m_H \neq 0$, $\frac{d\Delta f_{Ri}}{df_{Ri}} \neq 0$ because $\tilde{R}_i, \tilde{r_i}, \tilde{R}_j, \tilde{r_i} \geq 0$ and $\tilde{R}_i + \tilde{r_i} = 1$, $\tilde{R}_j + \tilde{r_j} = 1$. Indeed, the sum of the terms in the numerator is positive, which means that the expression is negative (because of the minus sign) and there is ndFDS.

$$\begin{aligned}\frac{d\Delta f_{Ai}}{df_{Ai}} &= \frac{d\Delta f_{Ai}}{dA_i}\frac{dA_i}{df_{Ai}} \\ &= A_i a_i \left(\frac{(1-c_i R_i)(1-m_P)}{(1-m_P)\tilde{A}_i + m_P\tilde{A}_j} + \frac{(1-b_i)(1-m_P)}{(1-m_P)\tilde{a_i} + m_P\tilde{a_j}} \right) - 1\end{aligned}$$

This can be simplified as,

$$\frac{d\Delta f_{Ai}}{df_{Ai}} = -\frac{\tilde{A}_i A_i (1-m_P)\tilde{a_j} m_P + \tilde{a_i} a_i (1-m_P)\tilde{A}_j m_P + \tilde{A}_j \tilde{a_j} m_P^2}{((1-m_P)\tilde{A}_i + m_P\tilde{A}_j)((1-m_P)\tilde{a_i} + m_P\tilde{a_j})}$$

If $m_P \neq 0$, $\frac{d\Delta f_{Ai}}{df_{Ai}} \neq 0$ because $\tilde{A}_i, \tilde{a}_i, \tilde{A}_j, \tilde{a}_i \geq 0$ and $\tilde{A}_i + \tilde{a}_i = 1$, $\tilde{A}_j + \tilde{a}_j = 1$. Indeed, the expression is negative so there is ndFDS.

The previous analysis shows that is possible to introduce ndFDS through migration between demes that have different coefficients of selection. To determine if the equilibrium is stable it is necessary to determine the terms outside the diagonal of the Jacobian matrix. These are given by,

$$\frac{d\Delta f_{Ri}}{df_{Ai}} = \frac{d\Delta f_{Ri}}{dA_i}\frac{dA_i}{df_{Ai}} = A_i a_i \left(\frac{R_i \rho_i s_i (1-u_i)(1-m_H)}{(1-m_H)\tilde{R}_i + m_H \tilde{R}_j} \right)$$

$$\frac{d\Delta f_{Ai}}{df_{Ri}} = \frac{d\Delta f_{Ai}}{dR}\frac{dR_i}{df_{Ri}} = R_i r_i \left(\frac{-c_i A_i (1-m_P)}{(1-m_P)\tilde{A}_i + m_P \tilde{A}_j} \right)$$

The discriminant of the Jacobian matrix can be approximated near the equilibrium point by,

$$Q_i = \left(\frac{d\Delta f_{Ri}}{df_{Ri}} - \frac{d\Delta f_{Ai}}{df_{Ai}} \right)^2 + 4 \left(\frac{d\Delta f_{Ri}}{df_{Ai}} \right) \left(\frac{d\Delta f_{Ai}}{df_{Ri}} \right)$$

However, we could not derive analytically conditions for the stability of the equilibrium point so we explored the system by simulations.