When do host–parasite interactions drive the evolution of non-random mating?

Scott L. Nuismer,1* Sarah P. Otto2 and François Blanquart1,3
1Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA
2Department of Zoology, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4
3École Normale Supérieure, 45 rue d’Ulm, Paris, 75005, France
*Correspondence: E-mail: snuismer@uidaho.edu

Abstract

Interactions with parasites may promote the evolution of disassortative mating in host populations as a mechanism through which genetically diverse offspring can be produced. This possibility has been confirmed through simulation studies and suggested for some empirical systems in which disassortative mating by disease resistance genotype has been documented. The generality of this phenomenon is unclear, however, because existing theory has considered only a subset of possible genetic and mating scenarios. Here we present results from analytical models that consider a broader range of genetic and mating scenarios and allow the evolution of non-random mating in both the parasite and host. Our results confirm results of previous simulation studies, demonstrating that coevolutionary interactions with parasites can indeed lead to the evolution of host disassortative mating. However, our results also show that the conditions under which this occurs are significantly more fickle than previously thought, requiring specific forms of infection genetics and modes of non-random mating that do not generate substantial sexual selection. In cases where such conditions are not met, hosts may evolve random or assortative mating. Our analyses also reveal that coevolutionary interactions with hosts cause the evolution of non-random mating in parasites as well. In some cases, particularly those where mating occurs within groups, we find that assortative mating evolves sufficiently to catalyze sympatric speciation in the interacting species.

Keywords

Assortative mating, coevolution, disassortative mating, major histocompatibility complex, sympatric speciation.

INTRODUCTION

Parasites have frequently been invoked as a source of selection favouring host individuals that generate genetically diverse offspring. For instance, the Red Queen hypothesis posits that interactions with parasites favour host individuals that reproduce sexually. This hypothesis has been extensively explored both theoretically (e.g., Hamilton 1980; Howard & Lively 1998; Peters & Lively 1999; Agrawal 2006; Gandon & Otto 2007) and empirically (e.g., Dybdahl & Lively 1998; Lively et al. 2004) and shown to work within a narrow range of conditions. Specifically, for interactions with parasites to favour sexual reproduction in hosts, selection must be very strong, and resistance to infection must be mediated by only a few genetic loci (Otto & Nuismer 2004).

Host individuals could also produce genetically diverse and potentially more resistant offspring by mating preferentially with genetically dissimilar individuals (e.g., Potts et al. 1994; Penn & Potts 1999; Hedrick 2002; Howard & Lively 2004; Milinski 2006). Much of the empirical evidence for this possibility comes from studies of mate preference in vertebrates with respect to major histocompatibility complex (MHC) genotypes. Specifically, studies of mate preference in multiple vertebrate species have demonstrated that females prefer to mate with males whose MHC genotype is different from their own (Penn & Potts 1999; Landry et al. 2001; Schwensow et al. 2008). Because heterozygosity at MHC loci has been shown to be associated with resistance to parasites and pathogens (e.g., Penn et al. 2002; Froeschke & Sommer 2005; Westerdahl et al. 2005), these studies suggest a potential role for parasites in observed patterns of disassortative mating in host populations.

Unlike the Red Queen hypothesis, the conditions under which coevolution with parasites drives the evolution of
disassortative mating in host populations have been subject
to only limited theoretical investigation. Specifically, How-
ard & Lively (2003) used simulation models to study the
evolution of disassortative mating in coevolving populations
of hosts and parasites. Their results showed that alleles
increasing female preference for genetically dissimilar males
spread when rare in host populations. Because their
simulations assumed that coevolution and mate choice are
mediated by haploid loci and that all females are guaranteed
to mate, the generality of their result remains unclear.

Virtually no effort has been spent evaluating whether
coevolution with hosts drives the evolution of assortative
mating in parasites. Previous studies have demonstrated
host–parasite interactions favour parasites that reduce the
number of allelic copies they express, either by reducing
their ploidy (Nuismer & Otto 2004) or by expressing only a
single gene copy (Nuismer & Otto 2005), which suggests
that interactions with hosts may favour parasites that mate
assortatively. Specifically, assortative mating is an additional
mechanism through which parasites could conceivably
decrease the number of allelic copies expressed, albeit
indirectly through the formation of homozygous offspring.

Our goal here is to develop and analyse a set of
mathematical models to study the evolution of non-random
mating in coevolving populations of hosts and parasites. By
considering multiple genetic systems of pathogen resistance,
diverse mating ecologies and multiple genetic mechanisms
of mate recognition, we hope to generalize previous
theoretical studies suggesting that parasites can drive the
evolution of disassortative mating in host populations.
Because our models allow non-random mating to evolve in
the parasite as well as in the host, we expect our models to
yield novel predictions for the conditions under which
coevolution with a host population promotes assortative
mating and potentially sympatric speciation in parasite
populations.

MODEL DESCRIPTION
We tracked evolutionary change in host and parasite
genotype frequencies over a life cycle where species
interactions were followed by mating and formation of
the next generation through recombination and segrega-
tion. Both host and parasite were assumed to be
hermaphroditic organisms characterized by two diallelic
diploid loci and population sizes sufficiently large for the
effects of genetic drift to be ignored. The first locus ‘M’
was assumed to modify the intensity of non-random
mating and to have two alleles M and m. The second locus
‘B’ was assumed to mediate species interactions and mating
preferences and to have two alleles B and b. Because the
loci involved in mate choice and host–parasite interactions
are one and the same (so that associations between these
two processes cannot be broken apart by recombination),
our modelling framework focuses on a scenario that is
particularly conducive to the evolution of non-random
mating (Gavrilets 2004). There is evidence, however, that
this may be the case for some loci involved in pathogen
recognition and mate choice, such as vertebrate MHC loci
e.g. Penn & Potts 1999). Our analyses rely heavily on the
work of Kirkpatrick et al. (2002), and we have attempted to
maintain consistency with their notation. Table 1 summari-
zes the key parameters and variables and their biological
interpretation; a companion Mathematica notebook is
available upon request.

Species interactions
Host and parasite were assumed to encounter one another
at random, with the outcome of encounters (either infection
or resistance) depending on the diploid genotypes of host
and parasite at the ‘B’ locus. We assume that B locus
genotypes of host and parasite interact following one of the
three commonly used models of host resistance to parasites
(Table 2). The inverse matching-alleles (IMA) model is
predicated on hosts having a suite of recognition molecules
capable of binding to a particular suite of pathogen antigens,
in a manner similar to the vertebrate MHC system (Frank
2002). The gene-for-gene (GFG) model has been shown to
be common in interactions between plants and pathogens,
and is based on a system where avirulent parasites produce
an elicitor that can be recognized by resistant hosts but not
by susceptible hosts and where virulent parasites do not
produce the elicitor (Burdon 1997). The matching-alleles
(MA) model is based on self/non-self recognition, where
parasites characterized by surface proteins different from
those of the host are recognized as non-self and an immune
response is mounted. The MA model has been suggested to
play a role in invertebrate resistance to pathogens and in the
maturation of the adaptive immune response of vertebrates
(Frank 2002).

We assume that each host encounters at most a single
parasite per generation. Encounters leading to infection
reduce host fitness by some amount $q_h$ and encounters
leading to resistance reduce parasite fitness by some amount
$s_p$. With these assumptions, the fitness of a host with
genotype $X_{Hh}$ is:

$$W(X_{Hh}) = (1 - \tau_{hh})^{2g(X_{Hh})} - \eta_1(\mathcal{E}_{X_{H}}[\psi(X_{Hh}, X_{P})]) + \tau_{hh}^{2g(X_{Hh})} - \eta_1(\mathcal{E}_{X_{H}}[\psi(X_{Hh}, X_{P})])$$

where $\tau_{hh}$ is the cost of resistance in the GFG model; $g(X_{Hh})$
is equal to 1 if genotype $X_{Hh}$ is heterozygous or homozygous
for the resistant B allele and 0 otherwise, making the
empirically motivated assumption that the B resistance allele
is dominant for both resistance and costs (Burdon 1997);
$\mathcal{E}_{X_{H}}[\cdot]$ denotes an expectation taken over the frequency
Table 1 Summary of model notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>A genotype vector</td>
</tr>
<tr>
<td>$X_j$</td>
<td>Allele carried by an individual at position $j$</td>
</tr>
<tr>
<td>$f(X)$</td>
<td>The frequency of genotype vector $X$</td>
</tr>
<tr>
<td>$W(X)$</td>
<td>The fitness of genotype vector $X$</td>
</tr>
<tr>
<td>$\bar{w}$</td>
<td>Population mean fitness</td>
</tr>
<tr>
<td>$E_2[\pi]$</td>
<td>The expectation of $\pi$ taken over the frequency distribution of $\zeta$</td>
</tr>
<tr>
<td>$W$</td>
<td>The set of positions in genotype vector $X$ contributing to fitness $\bar{w}$</td>
</tr>
<tr>
<td>$U$</td>
<td>A set of positions in genotype vector $X$</td>
</tr>
<tr>
<td>$p_{i,j}$</td>
<td>The frequency of the ‘I’ or ‘capital’ allele in species $i$, locus $j$</td>
</tr>
<tr>
<td>$\zeta_{U}$</td>
<td>A measure of an individual’s deviation from population expectation for set $U$</td>
</tr>
<tr>
<td>$D_{i,U} = E_X[s_{i,U}]$</td>
<td>The ‘selection coefficient’ acting on set $U$ in species $i$</td>
</tr>
<tr>
<td>$s_i$</td>
<td>The fitness cost of being infected ($i = H$) or failing to infect ($i = P$)</td>
</tr>
<tr>
<td>$\tau_i$</td>
<td>The fitness cost of carrying one or two resistance alleles ($i = H$) or carrying two virulence alleles ($i = P$)</td>
</tr>
<tr>
<td>$\delta_{i,j} = \beta_m(p_{MM} - p_{MM}) + \gamma_{MM}(p_{MM} - p_{MM})$</td>
<td>The recombination rate in species $i$</td>
</tr>
<tr>
<td>$\psi(X_{H}, X_{P})$</td>
<td>The probability that host genotype $X_{H}$ is infected in an encounter with parasite genotype $X_{P}$, given by Table 2</td>
</tr>
<tr>
<td>$P_i(X_{im}, X_{id})$</td>
<td>In the plant and animal models, the probability that male genotype $X_{im}$ is mated in an encounter with female genotype $X_{id}$ in species $i$, given by Table 3</td>
</tr>
<tr>
<td>$G_{id}(X_{im}, X_{id})$</td>
<td>The probability that male genotype $X_{im}$ and female genotype $X_{id}$ join mating group $k$ in species $i$, given by Table 4</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>The probability that a randomly selected host is infected by a randomly selected parasite</td>
</tr>
<tr>
<td>$\rho(X_{id})$</td>
<td>The intensity with which a female of genotype $X_{id}$ discriminates among male genotypes</td>
</tr>
</tbody>
</table>

Table 2 Genetic models of host resistance to pathogens

<table>
<thead>
<tr>
<th>Pathogen genotype</th>
<th>Host genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BB$</td>
<td>${R, I, I}$</td>
</tr>
<tr>
<td>$Bb$</td>
<td>${R, R, R}$</td>
</tr>
<tr>
<td>$bb$</td>
<td>${I, R, R}$</td>
</tr>
</tbody>
</table>

Entries indicate whether a particular host genotype is infected by (I) or resistant to (R) particular pathogen genotypes. The first entry in each vector is for the inverse matching-alleles model, the second for the gene-for-gene model, and the third for the matching-alleles model. The pattern of dominance in the gene-for-gene model is motivated by empirical data (Burdon 1997).

The fitness of a parasite with genotype $X_{P}$ is:

$$W(X_{P}) = (1 - \tau_{P})^{b(X_{P})} - \psi(E_{Xi} - \psi(X_{H}, X_{P})). \quad (2)$$

where $\tau_{P}$ is the cost of virulence in the GFG model; $b(X_{P})$ is equal to 1 if genotype $X_{P}$ is homozygous for the virulent B allele and 0 otherwise, making the empirically motivated assumption that the B virulence allele is recessive for both virulence and costs (Burdon 1997) and $E_{Xi}$ [ ] denotes an expectation taken over the frequency distribution of host genotypes.

Equations 1 and 2 can be used to calculate the mean fitness of host and parasite populations:

$$W_{H} = E_{X_{H}}[W(X_{H})] \quad (3a)$$

$$W_{P} = E_{X_{P}}[W(X_{P})]. \quad (3b)$$

The first section of the Supporting Information (eqns S1–S11) shows how eqns 1–3 can be used to calculate the changes in allele frequencies and statistical associations within and between loci that result from species interactions.

Mating

Following interactions between species, mating occurs. We assume that the phenotypic effects of the modifier locus, M, are restricted to individuals acting as females. Modifiers of assortative mating cause females to mate with males genetically similar at the ‘B’ locus more frequently than expected by chance, whereas modifiers of disassortative mating cause females to mate with males genetically similar at the ‘H’ locus.

© 2008 Blackwell Publishing Ltd/CNRS
and the male

The plant model assumes that individuals encounter one another at random and mate with a probability determined by the female’s M locus genotype, the female’s B locus genotype and the male’s B locus genotype (Table 3). Because the plant model assumes that females who decline to mate with a non-preferred male do not have another chance to mate (as might be the case in a pollen-limited plant), the plant model generates strong direct selection against choosy females. Following Kirkpatrick & Nuismer (2004), we first determine the fitness of a mated pair in species $i$:

$$W_i = E_{X_{im}, X_{if}}[W_i(X_{im}, X_{if})],$$

where the notation $E_{X_{im}, X_{if}}[\cdot]$ indicates an expectation taken over the frequency distribution of male genotypes, $f(X_{im})$, and female genotypes, $X_{if}$, in species $i$. The frequency of a mated pair consisting of male genotype $X_{im}$ and female genotype $X_{if}$ after non-random mating is then given by:

$$f(X_{im}, X_{if}) = \frac{f(X_{im})f(X_{if})W_i(X_{im}, X_{if})}{W_i}.$$

The grouping model

Instead of being based on female preferences for males encountered at random, as are the plant and animal models, the grouping model is based on a female’s decision to mate either: (i) within a group or (ii) at random among all groups. The grouping model does not impose strong direct costs on choosy females nor does it generate strong sexual selection on B locus genotypes (Otto et al. in press). The probability that a female decides to mate within a group is assumed to equal $\frac{1}{q}$ and all other terms are as defined for the plant model. The frequency of a mated pair consisting of male genotype $X_{im}$ and female genotype $X_{if}$ after non-random mating is then given by:

$$f(X_{im}, X_{if}) = \frac{f(X_{im})f(X_{if})W_i(X_{im}, X_{if})}{W_i}.$$

Consequently, it does not impose direct costs on choosy females as does the plant model, although it does generate strong sexual selection on B locus genotypes. Following Kirkpatrick & Nuismer (2004), we define the fitness of a mated pair in species $i$ for the animal model as:

$$W_i(X_{im}, X_{if}) = \frac{P_i(X_{im}, X_{if})}{E_{X_{im}}[P_i(X_{im}, X_{if})]},$$

Table 3 Preference matrix for ‘plant’ and ‘animal’ models

<table>
<thead>
<tr>
<th>Female genotype</th>
<th>Male genotype</th>
<th>$BB$</th>
<th>$Bb$</th>
<th>$bb$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BB$</td>
<td>$1 - (1 - \omega) \rho(X_{im})$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - (1 - \omega) \rho(X_{im})$</td>
</tr>
<tr>
<td>$Bb$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - (1 - \omega) \rho(X_{im})$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - (1 - \omega) \rho(X_{im})$</td>
</tr>
<tr>
<td>$bb$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - \omega \rho(X_{im})$</td>
<td>$1 - (1 - \omega) \rho(X_{im})$</td>
</tr>
</tbody>
</table>

Entries indicate the probability with which a female accepts a male as a mate, where the function $\rho(X_{im})$ measures the intensity with which a female discriminates among male genotypes as determined by her genotype at the modifier locus. For disassortative mating, $\omega = 0$; for assortative mating $\omega = 1$.

© 2008 Blackwell Publishing Ltd/CNRS
The second section of the Supporting Information (eqn S12) shows how eqns 6, 8 and 9 can be used to calculate the changes in allele frequencies and statistical associations within and between loci that result from non-random mating.

Recombination and segregation

Following the formation of mated pairs, haploid gametes are formed via segregation and recombination. Recombination between the M and B loci is assumed to occur at rate \( r \) in species \( i \). Gametes then unite at random within each mated pair to produce diploid offspring. Under these conditions, segregation and recombination do not change allele frequencies, although they do change the statistical associations within and between loci. The third section of the Supporting Information (eqn S13) shows how these statistical associations change as a consequence of recombination and segregation.

MODEL ANALYSIS

Quasi-linkage equilibrium approximation

In order to derive analytical solutions for the conditions leading to the evolution of non-random mating in coevolving host–parasite interactions, we assumed that coevolutionary selection is relatively weak (\( s \) of order \( \varepsilon \)), modifier alleles have only weak effects (order \( \varepsilon \)) and mating is initially approximately random. We were then able to use quasi-linkage equilibrium approximations (Barton & Turelli 1991; Nagylaki 1993; Kirkpatrick et al. 2002) to derive relatively simple expressions for the change in the frequency of modifier alleles in host and parasite (eqns S14–S28). These results demonstrate that species interactions can, at least under some conditions, drive the evolution of non-random mating in both host and parasite. However, our results also show that the type of mating that evolves (i.e. random, assortative, disassortative) depends on the mating behaviour of the species and the genetic mechanism of infection/resistance mediating interspecific interactions (Table 5).

Table 4 Probability of joining mating group \( j \) for the grouping model

<table>
<thead>
<tr>
<th>Group</th>
<th>Genotype</th>
<th>( BB )</th>
<th>( Bb )</th>
<th>( bb )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( BB )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>( Bb )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>( Bb )</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

\[ g_{i,k} = E \left[ G_{i,k}(X_{i,m}) \right]. \quad (10) \]

The first form of mating that evolves in the host is random (R) mating evolves in the host population and parasite population, based on the analytical results described in the supporting material. IMA, inverse matching alleles; GFG, gene-for-gene, MA, matching alleles.

Entries indicate whether assortative (A), disassortative (D) or random (R) mating evolves in the host population and parasite population, based on the analytical results described in the Supporting material.

**Table 5** Evolution of non-random mating in host and parasite populations

<table>
<thead>
<tr>
<th>Outcome in hosts</th>
<th>Outcome in parasites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mechanism of assortment</strong></td>
<td><strong>Mechanism of assortment</strong></td>
</tr>
<tr>
<td>IMA</td>
<td>GFG*</td>
</tr>
<tr>
<td>Plant</td>
<td>R</td>
</tr>
<tr>
<td>Animal( \dagger )</td>
<td>D</td>
</tr>
<tr>
<td>Grouping( \ddagger )</td>
<td>R</td>
</tr>
</tbody>
</table>

\*For the animal and grouping model, the outcome of the GFG model depends on the costs of resistance and virulence relative to selection imposed by species interactions (Supporting material).

\( \dagger \)Results for the animal model assume that assortative mating is weak relative to selection imposed by species interactions (Supporting material).

\( \ddagger \)Because mating within groups only generates assortative mating, disassortment was not considered.

?Indicates that analytical predictions are equivocal.

For both host and parasite, the evolution of non-random mating is more likely with some mating behaviours than others. Specifically, our results show that non-random mating will never evolve in the plant model. The reason being that the plant model generates strong direct costs of non-random mating that can never be overcome by any indirect benefits that accrue through species interactions (eqn S14). In contrast, the animal model does not generate direct costs of non-random mating and thus is significantly more conducive to the evolution of non-random mating. However, the animal model does induce sexual selection, which can generate indirect selection on non-random mating capable of overwhelming indirect selection generated by species interactions (eqn S15). The last form of mating behaviour considered, the grouping model, is the most conducive to the evolution of non-random mating because it generates no direct costs and only very weak indirect effects (eqn S22). However, because the grouping model assumes mating occurs within groups, it can only be used to gain insight into the evolution of assortative mating, not disassortative mating.

Although the potential for non-random mating to evolve is largely determined by mating behaviour, our results suggest that whether assortative or disassortative mating evolves in the host and parasite is primarily determined by the genetic basis of infection/resistance. Specifically, we find that the IMA model favours the evolution of disassortative or random mating in the host, but assortative mating in the...
parasite (eqns S16, S19, S23 and S26). This result arises because heterozygous hosts tend to be more resistant than the average homozygous host but homozygous parasites tend to be more infectious, on average, than heterozygous parasites. Although this argument holds for the GFG model as well, the costs of carrying resistance and virulence alleles can, in some cases, cause homozygous hosts to have, on average, greater fitness than heterozygous hosts and heterozygous parasites to have greater fitness than homozygous parasites, on average. Consequently, whether the GFG model favours assortative or disassortative mating in host and parasite depends on the relative magnitudes of fitness gains accruing through infection/virulence and fitness losses accruing through costs of resistance and virulence, which in turn depend on the allele frequencies in hosts and parasites (eqns S17, S20, S24 and S27). Finally, our results show that the MA model favours hosts and parasites (eqns S17, S21, S25 and S28).

**Numerical simulation and extension to multiple loci**

We developed deterministic multi-locus simulations to evaluate whether our analytical results were robust to strong selection, modifiers of large effect, and cases where mate choice and species interactions were mediated by multiple loci. An additional goal of these simulations was to clarify the pattern of non-random mating expected to evolve in the GFG model, for which our analytical results were inconclusive. Simulations assumed that both host and parasite had genomes consisting of a single modifier locus with alleles $M$ and $m$, located at the terminus of the chromosome, along with an additional set of $n$ loci with alleles 0’ or ‘1’, which mediated interactions with the other species as well as mate choice. Recombination between adjacent loci was assumed to occur at a rate of $r_i$ in species $i$, and loci mediating mate choice and species interactions were assumed to mutate with probability $5 \times 10^{-6}$ (mutation at the modifier locus was ignored). As with our analytical model, the life cycle was: (i) species interactions, (ii) mating and (iii) offspring production. Because simulations were developed using many of the same assumptions and equations as the analytical model, we describe in detail only those aspects of the simulations that differed significantly from our analytical model. Simulation source code (C++) is available upon request.

The frequency of host genotype $X_{HH}$ after interactions with the parasite is:

$$f'(X_{HH}) = f(X_{HH}) \frac{W'(X_{HH})}{W_H},$$

(11)

and the frequency of parasite genotype $X_P$ after interactions with the host is:

$$f'(X_P) = f(X_P) \frac{W'(X_P)}{W_P},$$

(12)

where all terms are as defined for the analytical model but extended to multiple loci with the following assumptions. First, costs of resistance and virulence in the GFG model were assumed to be multiplicative such that the $\alpha[X_H]$ term in eqn 1 equals the number of loci heterozygous or homozygous for the resistant ‘1’ allele and the $\beta[X_P]$ term in eqn 2 equals the number of loci homozygous for the virulent ‘1’ allele. Second, the probability that parasite genotype $X_P$ infects host genotype $X_H$ was determined in the following way. For each locus other than the modifier, the outcome of the interaction (infection or resistance) was determined using Table 2. Only if the outcome of the interaction was infected for all loci did the parasite succeed in infecting the host. Consequently, the fraction of host genotypes resistant to any particular parasite genotype generally increased with increasing numbers of loci. We did not consider other biologically plausible scenarios such as the case where each locus contributes only partial resistance (Sasaki 2000).

Among individuals that have survived host–parasite selection (eqns 11 and 12), the frequency of mated pairs was calculated using eqns 6, 8 and 9 by making the following assumptions regarding multi-locus interactions. First, we assumed that non-random mating depended on the exact diploid genotype of potential mates, where individuals discriminated between their own genotype and all other genotypes equally, with no distinction made based on quantitative similarity. That said, we assumed that individuals did not discriminate among differences based solely on the parent of origin. For example, an individual with paternal haplotype $\{0,0\}$ and maternal haplotype $\{0,1\}$ would recognize as different a mate consisting of paternal and maternal haplotypes $\{1,0\}$ and $\{0,0\}$ but not a mate carrying haplotypes $\{0,1\}$ and $\{0,0\}$.

Offspring were produced from mated pairs such that the frequency of genotype $X_i$ in the offspring generation was

$$f''(X_i) = E_{X_{im}X_{if}}[M(X_i, X_{im}, X_{if})],$$

(13)

where $E_{X_{im}X_{if}}[\cdot]$ indicates an expectation taken over the frequency distribution of mated pairs, and $M(X_i, X_{im}, X_{if})$ is the probability that a mating between a male of genotype $X_{im}$ and a female of genotype $X_{if}$ produces an offspring of genotype $X_i$.

We used these simulations to explore the evolution of modifiers of non-random mating across a broad range of parameter conditions, restricting our attention to cases where non-random mating evolved in only one species at a time. Specifically, we ran simulations for genomes consisting
of one or two loci mediating species interactions (plus a modifier of mating), where parameters were chosen at random from uniform distributions on the following intervals: \( \{0 \leq q_{H} \leq 1.0\} \), \( \{0 \leq q_{P} \leq 1.0\} \), \( \{0.05 \leq q_{H} \leq 0.45\} \), \( \{0.05 \leq q_{P} \leq 0.45\} \), \( \{0 \leq r_{H} \leq 0.5\} \) and \( \{0 \leq r_{P} \leq 0.5\} \). In all cases, we assumed modifiers of non-random mating had additive effects, such that genotype \( mm \) mated at random, genotype \( MM \) mated non-randomly and genotype \( Mm \) exhibited exactly intermediate levels of non-random mating. For the plant and animal models, we considered modifiers with effects drawn from uniform distributions on \( \{0.0 \leq \rho_{H,MM} \leq 0.10\} \) and \( \{0.0 \leq \rho_{P,MM} \leq 0.10\} \). For the grouping model, we considered only modifiers of assortative mating, and thus the effects of the modifier were drawn from uniform distributions on: \( \{0.0 \leq \rho_{H,MM} \leq 0.10\} \) and \( \{0.0 \leq \rho_{P,MM} \leq 0.10\} \).

At the beginning of each simulation, the frequency of the modifier allele was set to 0.01, and allele frequencies at loci involved in species interactions and mate recognition were chosen at random from the interval \( \{0,1\} \). The frequency of the modifier allele was then tracked over 10 000 generations. If the average modifier frequency over the final 1500 generations was > 0.011, the modifier was considered to have increased in frequency or spread. For each of the nine models (three models of infection genetics and three models of non-random mating), we ran 300 simulations. Thus the results reported in Table 6 for the two locus case are based on a total of 2700 simulation runs. Because the results for the case of three loci are quite similar, we report the results from these additional simulations in the Supporting Information (Table S1).

Results of simulations are in broad qualitative agreement with our analytical predictions for the IMA and MA models (Table 6), and generate predictions for the GFG model (Table 6). However, simulations also reveal that our analytical predictions are not perfect, with non-random mating failing to evolve in some simulation runs (compare Table 5 with Table 6). This discrepancy is most notable for the MA and IMA models of resistance when non-random mating is mediated by the animal model. There are at least two explanations for this discrepancy. First, the MA model generates underdominant selection in both species and does not efficiently maintain genetic polymorphism at the mating/interaction loci in diploid models (eqns S6b and S7b and Nuismer 2006). Consequently, in many cases, polymorphism may be eroded before any significant modifier evolution occurs. Second, the summary provided in Table 5 assumes that mating is nearly random (i.e. sexual selection is weak, so that eqn S15 is small relative to eqns S16–S21). In the animal model, modifier alleles causing substantial levels of non-random mating can induce sufficient sexual selection to violate this assumption, in which case the more detailed results in the Supporting Information should be used.

We next used simulations to evaluate the potential for very strong assortative mating, and thus incipient sympatric speciation, to evolve in the host and parasite. We restricted these simulations to the animal and grouping models as non-random mating never evolves in the plant model. For each combination of infection genetics and mating behaviour, we ran 75 simulations with modifiers inducing very strong host assortment \( \{0.96 \leq \rho_{H,MM} \leq 1.00\} \) and an additional 75 simulations with modifiers inducing very strong parasite assortment \( \{0.96 \leq \rho_{P,MM} \leq 1.00\} \). In each simulation, we evaluated whether the modifier approached fixation (average modifier frequency >0.99 over the final 500 generations) by generation 10 000. In all other respects, these simulations were identical to those previously described.

The results of our speciation simulations are reported in Table 7, and demonstrate that interactions between hosts and parasites can promote sympatric speciation. This is particularly likely when assortative mating occurs through a mechanism of group formation. In this case, parasite speciation occurs for all models of infection/resistance, and in the animal model it occurs for all models of infection/resistance.

### Table 6 Percentage of two-locus simulations in which assortative mating (top entry) and disassortative mating (bottom entry) evolved in the host and parasite

<table>
<thead>
<tr>
<th>Mechanism of assortment</th>
<th>Infection genetics (%)</th>
<th>Outcome in hosts</th>
<th>Outcome in parasites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IMA</td>
<td>GFG</td>
<td>MA</td>
</tr>
<tr>
<td>Plant</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Animal</td>
<td>0.0</td>
<td>0.0</td>
<td>15.1</td>
</tr>
<tr>
<td></td>
<td>64.4</td>
<td>24.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Grouping</td>
<td>0.0</td>
<td>41.3</td>
<td>80.7</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

IMA, inverse matching alleles; GFG, gene-for-gene; MA, matching alleles.

### Table 7 The evolution of high levels of assortment

<table>
<thead>
<tr>
<th>Mechanism of assortment</th>
<th>Infection genetics (%)</th>
<th>Outcome in hosts</th>
<th>Outcome in parasites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IMA</td>
<td>GFG</td>
<td>MA</td>
</tr>
<tr>
<td>Animal</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Grouping</td>
<td>0.0</td>
<td>77.3</td>
<td>92.0</td>
</tr>
</tbody>
</table>

Entries give the percentage of two-locus simulations that led to incipient speciation for the host and the parasite.

IMA, inverse matching alleles; GFG, gene-for-gene; MA, matching alleles.
and host speciation occurs readily for the GFG and MA models. If, in contrast, assortative mating is mediated by the animal model, speciation occurs with significant frequency only in the parasite when infection/resistance follows an IMA model. The reason the grouping model is so permissive of sympatric speciation is that it does not impose direct costs of assortative mating and leads to only very weak sexual selection (S22). Consequently, any genetic mechanism of infection/resistance that causes heterozygotes to be consistently less fit than the average homozygote, and maintains genetic polymorphism, will ultimately cause speciation.

**DISCUSSION**

Overall, the results of our models show that coevolution between hosts and parasites can promote the evolution of non-random mating. However, our results also reveal that significant levels of non-random mating are likely to evolve only in those species where mating does not impose strong direct costs or frequency-dependent effects. In addition, our results demonstrate that whether assortative or disassortative mating evolves depend primarily on the genetic mechanism of resistance. As a gross generalization, disassortative mating is more likely to evolve in hosts than in parasites, and assortative mating is more likely to evolve in parasites than in hosts. Where mating imposes no direct costs and only very weak frequency dependence, as when non-random mating arises from mating within groups of similar genotypes, we find that assortative mating can evolve to high enough levels to cause sympatric speciation.

Whether a particular type of infection genetics favours the evolution of assortative or disassortative mating depends on its consequences for the expected fitness of heterozygotes relative to homozygotes. For the IMA model, heterozygous hosts tend to be more resistant than homozygous hosts, but homozygous parasites tend to be more infectious than heterozygous parasites. Consequently, the IMA model favours disassortative mating in the host population but assortative mating in the parasite population. Although this argument also holds for the GFG model when constitutive costs of resistance/virulence are absent, incorporating such costs causes increased resistance/infectivity to become negatively associated with fitness under some conditions. When this occurs, the fitness of homozygous hosts may exceed that of heterozygous hosts, and the fitness of heterozygous parasites may exceed that of homozygous parasites. Consequently, whether the GFG model favours assortative or disassortative mating depends on the magnitude of fitness costs of resistance/infectivity relative to the magnitude of fitness gains accrued through increased resistance/infectivity. In contrast to the IMA and GFG models, the MA model causes homozygous hosts to be more resistant than heterozygous hosts (because heterozygotes carry fewer alleles to mimic), and thus favours the evolution of assortative mating in both host and parasite.

Our results for the MA model conflict with those of Howard & Lively (2003, 2004) who found that coevolution in an MA model commonly favoured the evolution of disassortative mating in host populations. This discrepancy likely arises because the model of Howard & Lively (2003, 2004) assumes selection acts on haploids whereas our models assume selection acts on diploids. Consequently, non-random mating evolves in response to epistatic interactions in the model of Howard & Lively (2003, 2004), but in response to dominance interactions in our two-locus analytical model and two-locus simulations. Because our three-locus simulations yield qualitatively similar results to our two-locus simulations, however, it seems that – at least for modest numbers of loci – it is dominance interactions that determine the type of non-random mating that evolves.

In addition to demonstrating the importance of the genetic basis of resistance and infection, our results show that the mating behaviour of the interacting species determines if, and to what extent, non-random mating evolves. Specifically, our results show that non-random mating never evolves in the plant model, evolves with modest frequency in the animal model and readily evolves in the grouping model (Table 6). The primary difference between these models is the extent to which they generate costs and impose sexual selection, with the plant model generating the strongest costs and the grouping model the weakest. Sexual selection can inhibit the evolution of non-random mating by reversing the fitness difference between heterozygotes and homozygotes or by eliminating polymorphism at mating trait loci (as in the animal model) (Otto et al. in press). That mating behaviour plays an important role in determining the fate of modifiers of disassortative mating may explain why some studies of mate preference based on disease-resistance genotype have found evidence for disassortative mating (Penn & Potts 1999; Milinski 2006) whereas others have not (Ekblom et al. 2004; Milinski 2006).

Our results also suggest that host–parasite interactions promote sympatric speciation in some cases. Specifically, we find that incipient sympatric speciation evolves readily in both host and parasite when assortative mating is mediated by the grouping model. These results provide further evidence that ecological interactions between individuals and species may be important catalysts for sympatric speciation (e.g. Doebeli 1996; Dieckmann & Doebeli 1999; Doebeli & Dieckmann 2000, 2003; Dieckmann et al. 2004) but also that direct fitness costs and frequency-dependent selection caused by sexual selection can inhibit sympatric speciation (e.g. Otto et al. in press; Gavrilets 2004;
Kirkpatrick & Nuismer 2004; Waxman & Gavrilets 2005). It may thus be little surprise that some of the best studied cases of putative sympatric speciation in parasitic taxa occur in insect parasites of plants that follow a grouping model where mating occurs on the host plant, minimizing the importance of sexual selection (Craig et al. 1993; Bush 1994; Via 1999).

Together, the results of our models demonstrate that coevolution between hosts and parasites can be an important force driving the evolution of non-random mating. In addition to this very broad prediction, our models generate several more specific and empirically testable predictions. One of the most interesting is that plant species which are not pollen limited should frequently mate assortatively at those loci involved in immune recognition, at least when such immunity is based on an MA or IMA mechanism and costs of assortative mating are not strong. Testing these and other predictions of our models should help to clarify the role coevolution plays in shaping patterns of non-random mating in hosts and parasites and may also provide valuable insights into the genetic basis of infection and resistance.

ACKNOWLEDGEMENTS

We thank Ben Ridenhour, Peter Thrall and three anonymous referees for their helpful comments. Funding was provided by NSF grants DEB 0343023 and DMS 0540392 to SLN, by the Natural Sciences and Engineering Research Council of Canada (SPO) and by the National Evolutionary Synthesis Center (NSF #EF-0423641; SPO).

REFERENCES


**Supporting Information**

Additional Supporting Information may be found in the online version of this article.

**Appendix S1** Species interactions, mating, segregation and recombination, quasi-linkage equilibrium approximation.

**Table S1** Percentage of three-locus simulations in which assortative mating (top entry) and disassortative mating (bottom entry) evolved in (a) the host and (b) the parasite.

Please note: Blackwell publishing is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Editor, Peter Thrall
Manuscript received 14 March 2008
First decision made 9 April 2008
Manuscript accepted 23 April 2008

© 2008 Blackwell Publishing Ltd/CNRS
When do host-parasite interactions drive the evolution of non-random mating?

Scott L. Nuismer¹
Sarah P. Otto²
Francois Blanquart¹,³

¹Department of Biological Sciences, University of Idaho, Moscow, ID 83844
²Department of Zoology, University of British Columbia, Vancouver, BC
³École Normale Supérieure, 45 rue d’Ulm, Paris, France
Supplementary Material

Appendix S1

SPECIES INTERACTIONS

Kirkpatrick et al. (2002) demonstrated that the strength of selection acting on a set of positions U, can be calculated using:

\[ \frac{W(X_i)}{\bar{W}_i} = 1 + \sum_{u \in W} a_{i,u}(\xi_{i,u} - D_{i,u}) \]  

(S1)

where \( W(X_i) \) is the fitness of genotype \( X \) in species \( i \), \( \bar{W}_i \) is the population mean fitness of species \( i \), \( W \) is the set of positions that contribute to fitness, and \( a_{i,U} \) is the "selection coefficient" acting on the set of positions \( U \) in species \( i \). Here, we define positions as do Kirkpatrick et al. as a particular locus in a particular context (e.g., the M locus in a female inherited from a male). Because coevolution is mediated only by host and parasite "B" locus genotypes, and because males and females are assumed to be equivalent at this stage in the life cycles, there are only two positions within the fitness set \( W \), which we will denote as \( B_m \) and \( B_f \) where the subscript indicates whether a "B" locus allele in a particular individual was inherited paternally (\( m \) subscript) or maternally (\( f \) subscript). We used equations (1-3), in conjunction with the particular values for \( \psi(X_H, X_P) \) from Table 2 of the main text, to calculate selection coefficients in the host (\( a_{H,U} \)) and parasite (\( a_{P,U} \)) to leading order for each of the three models of coevolution.

Inverse matching alleles

For the inverse matching-alleles model, we find that the selection coefficients favoring the \( B \) allele over the \( b \) allele in male (\( a_{H,B_m} \)) and female (\( a_{H,B_f} \)) hosts are:
and the selection coefficient favoring homozygotes over heterozygotes is:

\[ a_{H,B_m} = a_{H,B_f} = s_H[p_{P,B}^2 - p_{H,B}(1 - 2p_{P,B}q_{P,B})] + O(\epsilon^2), \]  

(S2a)

where \( p_{i,B} \) is the frequency of the \( B \) allele in species \( i \), and \( q_{i,B} \) is the frequency of the \( b \) allele in species \( i \). Equation (S2b) shows that the sign of \( a_{H,B,B_f} \) is always negative, demonstrating that the IMA model favours increased heterozygosity in the host.

For the parasite, we find that the selection coefficients are:

\[ a_{P,B_m} = a_{P,B_f} = s_P[p_{P,B}(1 - 2p_{H,B}q_{H,B}) - p_{H,B}^2] + O(\epsilon^2) \]  

(S3a)

\[ a_{P,B,B_f} = s_P(1 - 2p_{H,B}q_{H,B}) + O(\epsilon^2). \]  

(S3b)

Equation (S3b) shows that the sign of \( a_{P,B,B_f} \) is always positive, demonstrating that the IMA model favours increased homozygosity in the parasite.

**Gene-for-gene**

For the gene-for-gene model, we find that the selection coefficients favoring the \( B \) allele over the \( b \) allele in the host are:

\[ a_{H,B_m} = a_{H,B_f} = q_{H,B}[s_H(1 - p_{P,B}^2) - \tau_H] + O(\epsilon^2) \]  

(S4a)

and the selection coefficient favoring homozygotes over heterozygotes is:

\[ a_{H,B,B_f} = -s_H(1 - p_{P,B}^2) + \tau_H + O(\epsilon^2). \]  

(S4b)

Equation (S4b) shows that the sign of \( a_{H,B,B_f} \) depends on the strength of selection for increased resistance imposed by the parasite \( s_H(1 - p_{P,B}^2) \) relative to the strength of selection for reduced resistance generated by costs \( \tau_H \). Consequently, whether the GFG model favors heterozygosity or homozygosity in the host depends on the frequency of the parasite virulence allele.
For the parasite, we find that the selection coefficients are:

\[ a_{P,B_p} = a_{P,B_f} = p_{P,B_f} \left[ s_P \left( p_{H,B}^2 + 2p_{H,B}q_{H,B} \right) - \tau_P \right] + O(e^2) \]  
\[ (S5a) \]

\[ a_{P,B_B} = s_P \left( p_{H,B}^2 + 2p_{H,B}q_{H,B} \right) - \tau_P + O(e^2) \].  
\[ (S5b) \]

Equation (S5b) shows that the sign of \( a_{P,B_B} \) depends on the strength of selection for increased virulence imposed by the host \( s_P \left( p_{H,B}^2 + 2p_{H,B}q_{H,B} \right) \) relative to the strength of selection for reduced virulence generated by costs \( \tau_P \). Consequently, whether the GFG model favors heterozygosity or homozygosity in the parasite depends on the frequency of the parasite virulence allele.

Matching alleles

For the matching-alleles model, we find that the selection coefficients favoring the \( B \) allele over the \( b \) allele in the host are:

\[ a_{H,B_p} = a_{H,B_f} = s_H \left\{ p_{H,B} \left( 1 + 2p_{P,B}q_{P,B} \right) - \left( p_{P,B}^2 + 2p_{P,B}q_{P,B} \right) \right\} + O(e^2) \]  
\[ (S6a) \]

and the selection coefficient favoring homozygotes over heterozygotes is:

\[ a_{H,B_B} = s_H \left( 1 + 2p_{P,B}q_{P,B} \right) + O(e^2) \].  
\[ (S6b) \]

Equation (S6b) shows that the sign of \( a_{H,B_B} \) is always positive, demonstrating that the MA model favors increased homozygosity in the host.

For the parasite, we find that the selection coefficients are:

\[ a_{P,B_p} = a_{P,B_f} = s_P \left[ p_{P,B} \left( 1 - 2p_{H,B}q_{H,B} \right) - q_{H,B}^2 \right] + O(e^2) \]  
\[ (S7a) \]

\[ a_{P,B_B} = s_P \left( 1 - 2p_{H,B}q_{H,B} \right) + O(e^2) \].  
\[ (S7b) \]

Equation (S7b) shows that the sign of \( a_{P,B_B} \) is always positive, demonstrating that the MA model favors increased homozygosity in the parasite.
The selection coefficients (S2-S7) can be used to calculate the change in the frequency of the modifier allele $M$ that occurs in response to coevolutionary selection (before mating has occurred). Specifically, Kirkpatrick et. al. (2002) showed that the change in allele frequency at position $j$ in species $i$ due to selection is given by:

$$\Delta p_{i,j} = a_{i,j}p_{i,j}q_{i,j} + \sum_{U \subseteq W \neq j} a_{i,U}D_{i,U} \quad (S8)$$

where the notation $D_{i,U}$ means the statistical association between positions in the set that includes the set $U$ and additional position $j$ in species $i$. Because the modifier does not directly influence the coevolutionary interactions ($a_{i,j} = 0$), equation (S8) along with the selection coefficients (S2-S7) gives the change in the frequency of the modifier allele $M$ in species $i$:

$$\Delta p_{i,M} = a_{i,B_f}(D_{i,M_fB_f} + D_{i,M_mB_f}) + a_{i,B_m}(D_{i,M_mB_m} + D_{i,M_fB_m}) + a_{i,B_mB_f}(D_{i,M_mB_mB_f} + D_{i,M_fB_mB_f}) \quad (S9)$$

Equation (S9) demonstrates that coevolutionary selection changes the frequency of the modifier allele only indirectly through its statistical associations with the selected locus $B$. There are two ways in which this can occur. First, a modifier of non-random mating may change in frequency if it becomes statistically associated with the $B$ locus allele currently favoured or disfavoured by directional coevolutionary selection. This force is reflected in the first two terms of (S9). Second, a modifier of non-random mating may change in frequency if it becomes statistically associated with homozygosity or heterozygosity at the $B$ locus. If a particular coevolutionary model favours increased heterozygosity (e.g., the host in an IMA model), then modifiers that become statistically associated with heterozygous $B$ locus genotypes will increase in frequency. This force is reflected in the third term of (S9).
In addition to showing how the selection coefficients (e.g., S2-S7) can be used to calculate changes in allele frequency, Kirkpatrick et. al. (2002) showed how they can be used to calculate the change in statistical associations among a set of positions $A$ in species $i$ that occur in response to selection:

\[ \Delta D_{i,A} = \sum_{u \in W} a_{i,u} (D_{i,Au} - D_{i,A} D_{i,u}) \]  

(S10)

where the notation $D_{i,Au}$ indicates the statistical association between the set of positions $A$ in species $i$ after the set of positions in $U$ has been removed. Combined with the selection coefficients (S2-S7), equation (S10) shows that coevolutionary selection generates only a single type of statistical association — deviations from Hardy-Weinberg. Specifically, coevolutionary selection leads to the following change in statistical associations within the $B$ locus:

\[ \Delta D_{i,B,B'} = a_{i,B,B'} P_{i,B} q_{i,B} \]  

(S11)

**MATING**

Kirkpatrick et. al. (2002) showed that the changes in allele frequencies and statistical associations between positions caused by non-random mating can be calculated once the frequencies of mated pairs are known. As with viability selection, the first step is to calculate the selection coefficients acting on positions in the genome using:

\[ \frac{f(X_{im}, X_{if})}{f(X_{im}) f(X_{if})} = 1 + \sum_{u \in W} a_{i,u} (\xi_{i,u} - D_{i,u}) \]  

(S12)

where $f(X_{im}, X_{if})$ is the frequency of a mated pair consisting of a male with genotype $X_{im}$ and a female with genotype $X_{if}$ (equations 6, 8, and 9 of the main text) and $f(X_{im})$ and $f(X_{if})$ are the frequencies of male and female genotypes, respectively, prior to non-
random mating. Because the fitness set $W$ is very large for the forms of non-random mating considered here, we do not show the results for the numerous selection coefficients that emerge from (S12), although a Mathematica package which contains these formulae is available upon request. After calculating the selection coefficients using (S12), the changes in allele frequencies and statistical associations between positions can be calculated using equations (S8) and (S10) respectively. These results are quite numerous and complex, so we do not report them here although they are available upon request in the form of a Mathematica package.

Non-random mating can itself cause changes in the modifier frequency for three reasons. First, to the extent that there are costs of assortative mating (as in the plant model, where choosier females have lower fitness), selection will act directly on the modifier to reduce mating preferences. Second, non-random mating can cause homozygotes and heterozygotes to differ in mating success, which can favor the evolution of assortment (when homozygotes are better able to find mates) or disassortment (when heterozygotes are better able to find mates). Finally, non-random mating generates associations between alleles that are increasing in frequency and modifiers that prefer that allele (as in models of Fisherian sexual selection); because assortative mating causes a mating advantage to common alleles, this third force favors the evolution of assortative mating when the allele rising in frequency at the B locus is common but hinders it when the allele is rare. More details about how these three forces combine can be found in the one-species model of Otto et al. (submitted).
Kirkpatrick et. al. (2002) showed that the change in statistical associations caused by segregation and recombination in species \( i \) is:

\[
\Delta D_{i,\mathcal{A}} = \sum_{\mathcal{U}:\mathcal{U}=\mathcal{A}} t_{i,\mathcal{A} \leftarrow \mathcal{U}} D_{i,\mathcal{U}},
\]  

(S13)

where \( t_{i,\mathcal{A} \leftarrow \mathcal{U}} \) is the probability that the positions in set \( \mathcal{U} \) are transmitted into set \( \mathcal{A} \) in species \( i \), and the notation \( \mathcal{U}:\mathcal{U}=\mathcal{A} \) indicates that the sum is taken over all possible sets \( \mathcal{U} \) which could be transmitted into set \( \mathcal{A} \) once the context information (e.g., sex of origin, sex of carrier) is stripped from them. Once again, equation (S13) yields numerous complex expressions and so we refrain from writing them out; they are available in a Mathematica notebook upon request.

QUASI-LINKAGE EQUILIBRIUM APPROXIMATION

In order to derive tractable expressions for the change in modifier frequency, we made a quasi-linkage equilibrium approximation. In addition to assuming that selection is relatively weak (order \( \varepsilon \)), our approximation assumed: 1) non-random mating is not too strong, 2) the effect of the modifier is not too great, and 3) sex and recombination are relatively frequent. With these assumptions, statistical associations between positions should change quickly relative to allele frequencies and approach a quasi-steady state where the values of these associations are small and of order \( \varepsilon \). Under these conditions, it is possible to solve for the quasi-linkage equilibrium (QLE) values of the statistical associations between positions. These QLE values can then be used in conjunction with equation (S9) and the selection coefficients imposed by species interactions (S2-S7) and non-random mating (available as Mathematica notebook) to calculate the change in the
frequency of the modifier allele due to species interactions and non-random mating.

Because we have assumed that selection is weak (order $\epsilon$) and that the system is at QLE, the change in modifier frequency over a single generation is simply the sum of the change due to species interactions and non-random mating. In the following sections, we provide results for the change in modifier frequency caused by non-random mating and for the change in modifier frequency caused by species interactions for each of the three models of non-random mating.

**QLE results for the plant model**

*Change in modifier frequency due to non-random mating*

We find that the change in the frequency of a modifier of assortative mating in species $i$ is:

$$\Delta p_{i,M} = -\frac{\delta_{\rho,i}}{2} [p_{i,M} q_{i,M} (1 - \chi)] + O(\epsilon^2)$$  \hspace{1cm} (S14a)

and a modifier of disassortative mating is:

$$\Delta p_{i,M} = -\frac{\delta_{\rho,i}}{2} [p_{i,M} q_{i,M} \chi] + O(\epsilon^2)$$  \hspace{1cm} (S14b)

where $\chi$ is the probability of encountering a mate with a matching genotype ($\chi = p_{1,b}^4 + 4 p_{1,b}^2 q_{1,b}^2 + q_{1,b}^4$) and $\delta_{\rho,i}$ is the effect of the modifier in species $i$. Because species interactions generate only indirect selection of order $\epsilon^2$, but females that mate non-randomly pay a direct cost in terms of reduced mating opportunities, this result shows that non-random mating will never evolve in the plant model.

**QLE results for the animal model**

*Change in modifier frequency due to non-random mating*

We find that the change in the frequency of a modifier due to non-random mating in species $i$ is:
where $\bar{\rho}_i$ is the average level of assortment or disassortment in species $i$. In contrast to the plant model, this result shows that the animal model generates no direct selection on the modifier locus (of order $\varepsilon$), instead imposing indirect sexual selection of the same order ($\varepsilon^2$) as that caused by species interactions. This indirect selection can favor or disfavor a modifier of non-random mating. Specifically, modifiers increasing the degree of non-random mating (either assortment or disassortment) are favored by indirect selection anytime the allele frequency at the mating trait locus lies outside of the interval $\{0.30, 0.70\}$ but are disfavored by indirect selection anytime the allele frequency at the mating trait locus lies inside this interval. In brief, when allele frequencies are intermediate, the presence of assortative mating increases the relative mating success of heterozygotes compared to the average homozygote (because heterozygotes are common and they prefer other heterozygotes); this heterozygous mating advantage counteracts the evolution of further assortment. Similarly, the presence of disassortative mating increases the relative mating success of homozygotes compared to the heterozygotes (because heterozygotes are common and they prefer not to mate with other heterozygotes); this homozygous mating advantage counteracts the evolution of further disassortment. These arguments apply only when allele frequencies are sufficiently close to 1/2. When the allele frequency at the mating trait locus is far from 1/2, heterozygotes are rare and the above arguments reverse.

\begin{equation}
\Delta p_{i,M} = \frac{1}{2} \bar{\rho}_i \delta \rho_{i,M} p_i q_i (1 - 3 p_{i,a} q_{i,a}) [1 - 6 p_{i,a} q_{i,a} (1 - p_{i,a} q_{i,a})] + O(\varepsilon^3)
\end{equation}

\textit{Change in host modifier frequency caused by interactions with parasite}
Inverse matching alleles model – We find that IMA interactions with the parasite lead to
the following change in the frequency of a modifier of non-random mating:

\[ \Delta p_{H,M} = -\theta \delta_{p,H} \lambda s_{H,H,M} p_{H,M} p_{H,b} q_{H,b} (1 - 3p_{H,b}q_{H,b}) + O(\epsilon^3) \] (S16)

where \( \theta \) is an indicator variable which takes the value +1 for a modifier of assortative
mating and -1 for a modifier of disassortative mating, and \( \lambda \) is the probability that a
randomly selected host is infected by a randomly selected parasite. This result
demonstrates that interactions with the parasite favor increased disassortment.

Comparing result (S16) with result (S15) shows that disassortative mating is guaranteed
to evolve only when the average level of assortment in the population is small relative
to the strength of selection exerted by species interactions.

Gene-for-gene model – We find that the change in the frequency of a modifier of non-
random mating in the host caused by GFG interactions with the parasite is:

\[ \Delta p_{H,M} = -\theta \delta_{p,H} \lambda s_{H,H,M} p_{H,M} p_{H,b} q_{H,b} [q_{H,b}^2 (1 - 3p_{H,b}q_{H,b}) (s_{H} (1 - p_{H,b}^2) - \tau_{H})] + O(\epsilon^3) \] (S17)

This result demonstrates that interactions with the parasite favor increased disassortment
any time costs of resistance (\( \tau_{H} \)) are small relative to selection for increased resistance.

Comparing result (S17) with result (S15) shows that disassortative mating is then
guaranteed to evolve only when the average level of assortment in the population is
small relative to the strength of selection exerted by species interactions and when costs
of resistance (\( \tau_{H} \)) are small relative to selection for increased resistance.

Matching alleles model – We find that the change in the frequency of a modifier of non-
random mating in the host caused by MA interactions with the parasite is:

\[ \Delta p_{H,M} = \theta \delta_{p,H} s_{H} p_{H,M} q_{H,M} p_{H,b} q_{H,b} (1 - 3p_{H,b}q_{H,b})(1 - \lambda) + O(\epsilon^3) \] (S18)
This result demonstrates that interactions with the parasite always favor increased
assortment in the host. Comparing result (S18) with result (S15) shows that assortative
mating is guaranteed to evolve only when the average level of assortment in the
population is small relative to the strength of selection exerted by species interactions.

Change in parasite modifier frequency caused by interactions with host

Inverse matching alleles model – We find that the change in the frequency of a modifier
of non-random mating in the parasite caused by IMA interactions with the host is:

\[ \Delta p_{p,m} = \theta \delta_{\rho,p} \rho^s p_{p,m} q_{p,b} (1 - 3 p_{p,b} q_{p,b}) + O(\epsilon^3) \]  

(S19)

This result demonstrates that interactions with the host always favor increased
assortment. Comparing result (S19) with result (S15) shows that assortative mating is
guaranteed to evolve only when the average level of assortment in the population is
small relative to the strength of selection exerted by species interactions.

Gene-for-gene model – We find that the change in the frequency of a modifier of non-
random mating in the parasite caused by GFG interactions with the host is:

\[ \Delta p_{p,m} = \theta \delta_{\rho,p} \rho^s p_{p,m} q_{p,b} (1 - 3 p_{p,b} q_{p,b}) (1 - q_{p,b}^2 - \tau_p) + O(\epsilon^3) \]  

(S20)

This result demonstrates that interactions with the host favor increased assortment
anytime costs of virulence (\( \tau_p \)) are small relative to selection for increased virulence.

Comparing result (S20) with result (S15), shows that assortative mating is guaranteed to
evolve only when the average level of assortment in the population is small relative to
the strength of selection exerted by species interactions and when costs of virulence (\( \tau_p \))
are small relative to selection for increased virulence.

Matching alleles model – We find that the change in the frequency of a modifier of non-
random mating in the parasite caused by MA interactions with the host is:
It is possible to show that $S_{21}$ is always positive, demonstrating that interactions with the host always favor increased assortment. Comparing result (S21) with result (S15), shows that assortative mating is guaranteed to evolve only when the average level of assortment in the population is small relative to the strength of selection exerted by species interactions.

**QLE results for the grouping model**

**Change due to assortative mating**

We find that the change in the frequency of a modifier of assortative mating is

$$\Delta p_{i,M} = O(\varepsilon^3)$$  \hspace{1cm} (S22)

showing that selection induced by non-random mating in the grouping model is of lower order than that generated by species interactions. For this reason, the change in the frequency of a modifier of assortative mating in the grouping model is, to leading order, driven only by species interactions.

**Change in host modifier frequency caused by interactions with parasite**

**Inverse matching alleles model** – We find that the change in the frequency of a modifier of assortative mating in the host caused by IMA interactions with the parasite is:

$$\Delta p_{H,M} = -\frac{1}{4} \delta_{p,H} s_H p_{H,M} q_{H,M} p_{H,B} q_{H,B} [1 - 2 p_{P,B} q_{P,B}] + O(\varepsilon^3)$$  \hspace{1cm} (S23)

This result shows that IMA interactions with the parasite never lead to the evolution of assortative mating in the host.

**Gene-for-gene model** – We find that the change in the frequency of a modifier of assortative mating in the host caused by GFG interactions with the parasite is:

$$\Delta p_{H,M} = -\frac{1}{4} \delta_{p,H} p_{H,M} q_{H,M} p_{H,B} q_{H,B} [s_H (1 - p_{P,B}) - \tau_H] + O(\varepsilon^3)$$  \hspace{1cm} (S24)
This result shows that GFG interactions with the parasite lead to the evolution of assortative mating in the host only if costs of resistance ($\tau_h$) are large relative to selection for increased resistance $s_h(1-p_{p,b}^2)$.

Matching alleles model – We find that the change in the frequency of a modifier of assortative mating in the host caused by MA interactions with the parasite is:

$$\Delta p_{H,M} = \frac{1}{4} \delta_{p,H} s_H p_{H,M} q_{H,M} p_{H,B} q_{H,B} [1 + 2p_{p,b} q_{p,b}] + O(\epsilon^3)$$  \hspace{1cm} (S25)

This result shows that MA interactions with the parasite always lead to the evolution of assortative mating in the host.

Change in parasite modifier frequency caused by interactions with host

Inverse matching alleles model – We find that the change in the frequency of a modifier of assortative mating in the parasite caused by IMA interactions with the host is:

$$\Delta p_{P,M} = \frac{1}{4} \delta_{p,P} s_P p_{P,M} q_{P,M} p_{P,B} q_{P,B} [1 - 2p_{p,b} q_{p,b}] + O(\epsilon^3)$$  \hspace{1cm} (S26)

This result shows that IMA interactions with the host always cause the evolution of assortative mating in the parasite.

Gene-for-gene model – We find that the change in the frequency of a modifier of assortative mating in the parasite caused by GFG interactions with the host is:

$$\Delta p_{P,M} = \frac{1}{4} \delta_{p,P} s_P p_{P,M} q_{P,M} p_{P,B} q_{P,B} [s_P(1-q_{H,B}^2) - \tau_P] + O(\epsilon^3)$$  \hspace{1cm} (S27)

This result shows that GFG interactions with the host cause the evolution of assortative mating in the parasite anytime the costs of virulence ($\tau_P$) are small relative to selection for increased virulence $s_P(1-q_{H,B}^2)$.

Matching alleles model – We find that the change in the frequency of a modifier of assortative mating in the parasite caused by MA interactions with the host is:
This result shows that MA interactions with the host always cause the evolution of assortative mating in the parasite.

Table S1. Percentage of three-locus simulations in which assortative mating (top entry) and dissasortative mating (bottom entry) evolved in (a) the host and (b) the parasite.

<table>
<thead>
<tr>
<th>Mechanism of assortment</th>
<th>Infection genetics</th>
<th>Infection genetics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IMA</td>
<td>GFG</td>
</tr>
<tr>
<td>Plant</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Animal</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>39.4%</td>
<td>28.0%</td>
</tr>
<tr>
<td>Grouping</td>
<td>0.0%</td>
<td>35.3%</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>