LETTER

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

Scott L. Nuismer,[1]* Sarah P. Otto[2] and François Blanquart[1,3]

[1]Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA

[2]Department of Zoology, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4

[3]École 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 the parasite as well. 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

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 prefer-

entially 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, Howard & 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 segregation. 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 summarizes 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 s_H and encounters leading to resistance reduce parasite fitness by some amount s_P. With these assumptions, the fitness of a host with genotype $\mathbf{X_H}$ is:

$$W(\mathbf{X_H}) = (1 - \tau_H)^{\alpha(\mathbf{X_H})} - s_H(E_{\mathbf{X_P}}[\psi(\mathbf{X_H}, \mathbf{X_P})]), \quad (1)$$

where τ_H is the cost of resistance in the GFG model; $\alpha(\mathbf{X_H})$ is equal to 1 if genotype $\mathbf{X_H}$ 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); $E_{\mathbf{X_P}}[\cdot]$ denotes an expectation taken over the frequency

Table 1 Summary of model notation

Symbol	Meaning
$\mathbf{X}$	A genotype vector
X_j	Allele carried by an individual at position j
$f(\mathbf{X})$	The frequency of genotype vector $\mathbf{X}$
$W(\mathbf{X})$	The fitness of genotype vector $\mathbf{X}$
$\bar{W}$	Population mean fitness
$E_z[\pi]$	The expectation of π taken over the frequency distribution of z
$\mathbf{W}$	The set of positions in genotype $\mathbf{X}$ contributing to fitness W
$\mathbf{U}$	A set of positions in genotype vector $\mathbf{X}$
$p_{i,j}$	The frequency of the '1' or 'capital' allele in species i, locus j
$\zeta_{\mathbf{U}} = \prod_{j \in \mathbf{U}} (X_j - p_j)$	A measure of an individual's deviation from population expectation for set $\mathbf{U}$
$D_{i,\mathbf{U}} = E_{\mathbf{X}}[\zeta_{\mathbf{U}}]$	A measure of the statistical association among alleles in set $\mathbf{U}$ for species i
$a_{i,\mathbf{U}}$	The 'selection coefficient' acting on set $\mathbf{U}$ in species i
s_i	The fitness cost of being infected ($i = $ H) or failing to infect ($i = $ P)
τ_i	The fitness cost of carrying one or two resistance alleles ($i = $ H) or carrying two virulence alleles ($i = $ P)
$\bar{\rho}_i = p_M^2 \rho_{MM} + 2 p_M q_M \rho_{Mm} + q_M^2 \rho_{mm}$	The average level of non-random mating in species i
$\delta_{\rho,i} = p_M(\rho_{MM} - \rho_{Mm}) + q_M(\rho_{Mm} - \rho_{mm})$	The effect of the modifier allele M in species i
r_i	The recombination rate in species i
$\psi(\mathbf{X_H}, \mathbf{X_P})$	The probability that host genotype $\mathbf{X_H}$ is infected in an encounter with parasite genotype $\mathbf{X_P}$, given by Table 2
$P_i(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})$	In the plant and animal models, the probability that male genotype $\mathbf{X_{i,m}}$ is mated in an encounter with female genotype $\mathbf{X_{i,f}}$ in species i, given by Table 3
$G_{i,k}(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})$	The probability that male genotype $\mathbf{X_{i,m}}$ and female genotype $\mathbf{X_{i,f}}$ join mating group k in species i, given by Table 4
λ	The probability that a randomly selected host is infected by a randomly selected parasite
$\rho(\mathbf{X_{i,f}})$	The intensity with which a female of genotype $\mathbf{X_{i,f}}$ discriminates among male genotypes

Table 2 Genetic models of host resistance to pathogens

Pathogen genotype	Host genotype		
	BB	Bb	bb
BB	{R,I,I}	{R,I,I}	{I,I,R}
Bb	{R,R,R}	{R,R,I}	{R,I,R}
bb	{I,R,R}	{R,R,I}	{R,I,I}

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).

distribution of parasite genotypes and $\psi(\mathbf{X_H}, \mathbf{X_P})$ is the probability that parasite genotype $\mathbf{X_P}$ successfully infects host genotype $\mathbf{X_H}$ (Table 2). Similarly, the fitness of a parasite with genotype $\mathbf{X_P}$ is:

$$W(\mathbf{X_P}) = (1 - \tau_P)^{\beta(\mathbf{X_P})} - s_P(E_{\mathbf{X_H}}[1 - \psi(\mathbf{X_H}, \mathbf{X_P})]), \quad (2)$$

where τ_P is the cost of virulence in the GFG model; $\beta(\mathbf{X_P})$ is equal to 1 if genotype $\mathbf{X_P}$ 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 $E_{X_H}[\cdot]$ 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:

$$\bar{W}_H = E_{\mathbf{X_H}}[W(\mathbf{X_H})] \quad (3a)$$

$$\bar{W}_P = E_{\mathbf{X_P}}[W(\mathbf{X_P})]. \quad (3b)$$

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

Table 3 Preference matrix for 'plant' and 'animal' models

Female genotype	Male genotype		
	BB	Bb	bb
BB	$1 - (1 - \omega)\,\rho(\mathbf{X_{i,f}})$	$1 - \omega\,\rho(\mathbf{X_{i,f}})$	$1 - \omega\,\rho(\mathbf{X_{i,f}})$
Bb	$1 - \omega\,\rho(\mathbf{X_{i,f}})$	$1 - (1 - \omega)\,\rho(\mathbf{X_{i,f}})$	$1 - \omega\,\rho(\mathbf{X_{i,f}})$
Bb	$1 - \omega\,\rho(\mathbf{X_{i,f}})$	$1 - \omega\,\rho(\mathbf{X_{i,f}})$	$1 - (1 - \omega)\,\rho(\mathbf{X_{i,f}})$

Entries indicate the probability with which a female accepts a male as a mate, where the function $\rho(\mathbf{X_{i,f}})$ 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 $\omega = 1$.

dissimilar at the 'B' locus more frequently than expected by chance. Mate recognition based on 'B' locus genotypes is assumed to be perfect; incorporating errors in mate recognition would reduce the strength of selection acting on the modifier (Otto *et al.* in press). We considered three models of non-random mating, two based on random encounters and female preferences ('plant model' and 'animal model') and one based on the formation of mating groups ('grouping model').

The plant model
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:

$$W(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}) = P_i(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}), \tag{4}$$

which here equals the probability, $P_i(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})$, that a male of species i with genotype $\mathbf{X_{i,m}}$ mates with a female of species i and genotype $\mathbf{X_{i,f}}$ (Table 3). The mean fitness of mated pairs in species i is then:

$$\bar{W}_i = E_{\mathbf{X_{i,m}},\, \mathbf{X_{i,f}}}[W(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})], \tag{5}$$

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

$$f(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}) = \frac{f(\mathbf{X_{i,m}})f(\mathbf{X_{i,f}})\,W(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})}{\bar{W}_i}. \tag{6}$$

The animal model
The animal model is identical to the plant model with the exception that all female genotypes are assumed to have equal mating success (Kirkpatrick & Nuismer 2004).

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:

$$W(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}) = \frac{P_i(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})}{E_{\mathbf{X_{i,m}}}[P_i(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})]}, \tag{7}$$

where the denominator is the average mating success of a female with genotype $\mathbf{X_{i,f}}$, and all other terms are as defined for the plant model. The frequency of a mated pair consisting of male genotype $\mathbf{X_{i,m}}$ and female genotype $\mathbf{X_{i,f}}$ after non-random mating is then given by:

$$f(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}) = f(\mathbf{X_{i,m}})f(\mathbf{X_{i,f}})W(\mathbf{X_{i,m}}, \mathbf{X_{i,f}}). \tag{8}$$

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 be determined by her genotype at the modifier locus M and is denoted by $\rho_i(\mathbf{X_{i,f}})$. With these assumptions, the frequency of a mated pair consisting of male genotype $\mathbf{X_{i,m}}$ and female genotype $\mathbf{X_{i,f}}$ after non-random mating is:

$$f(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})$$
$$= f(\mathbf{X_{i,m}})f(\mathbf{X_{i,f}})\left(\rho_i(\mathbf{X_{i,f}})\sum_{k=1}^{N}\frac{G_{i,k}(\mathbf{X_{i,m}})G_{i,k}(\mathbf{X_{i,f}})}{g_{i,k}}\right.$$
$$\left. + [1 - \rho_i(\mathbf{X_{i,f}})]\right), \tag{9}$$

where $G_{i,k}(\mathbf{X_{i,m}})$ is the probability that male genotype $\mathbf{X_{i,m}}$ joins group k, which is assumed to equal $G_{i,k}(\mathbf{X_{i,f}})$ for females with the same genotype (Table 4); N is the number of groups and $g_{i,k}$ is the proportion of males of species i that join group k:

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

	Genotype		
Group	BB	Bb	bb
BB	1	0	0
Bb	0	1	0
Bb	0	0	1

$$g_{i,k} = E_{\mathbf{X_{i,m}}}[G_{i,k}(\mathbf{X_{i,m}})]. \tag{10}$$

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_i 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 ε), modifier alleles have only weak effects (order ε) 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 5 Evolution of non-random mating in host and parasite populations

Mechanism of assortment	Infection genetics					
	Outcome in hosts			Outcome in parasites		
	IMA	GFG*	MA	IMA	GFG*	MA
Plant	R	R	R	R	R	R
Animal†	D	?	A	A	?	A
Grouping‡	R	?	A	A	?	A

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.

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

*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).

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

‡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/resistance 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 that mate assortatively, because of the increased resistance of homozygous hosts, which are harder to mimic, and the increased infectivity of homozygous parasites, which are harder to recognize, relative to heterozygotes (eqns S18, 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 5×10^{-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 $\mathbf{X_H}$ after interactions with the parasite is:

$$f'(\mathbf{X_H}) = f(\mathbf{X_H}) \frac{W(\mathbf{X_H})}{\bar{W}_H} \tag{11}$$

and the frequency of parasite genotype $\mathbf{X_P}$ after interactions with the host is:

$$f'(\mathbf{X_P}) = f(\mathbf{X_P}) \frac{W(\mathbf{X_P})}{\bar{W}_P}, \tag{12}$$

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 $\alpha[\mathbf{X_H}]$ term in eqn 1 equals the number of loci heterozygous or homozygous for the resistant '1' allele and the $\beta[\mathbf{X_P}]$ term in eqn 2 equals the number of loci homozygous for the virulent '1' allele. Second, the probability that parasite genotype $\mathbf{X_P}$ infects host genotype $\mathbf{X_H}$ 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 $\mathbf{X_i}$ in the offspring generation was

$$f'''(\mathbf{X_i}) = E_{\mathbf{X_{i,m}}\mathbf{X_{i,f}}}[M(\mathbf{X_i}, \mathbf{X_{i,m}}, \mathbf{X_{i,f}})], \tag{13}$$

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

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: $\{0 \leq s_H \leq 1.0\}$, $\{0 \leq s_P \leq 1.0\}$, $\{0.05 \, s_H \leq \tau_H \leq 0.45 \, s_H\}$, $\{0.05 \, s_P \leq \tau_P \leq 0.45 \, s_P\}$, $\{0 \leq r_H \leq 0.5\}$ 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 $\{-0.10 \leq \rho_{H, MM} \leq 0.10\}$ and $\{-0.10 \leq \rho_{P, MM} \leq 0.10\}$. 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 analyt-

ical 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,

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

Mechanism of assortment	Infection genetics (%)					
	Outcome in hosts			Outcome in parasites		
	IMA	GFG	MA	IMA	GFG	MA
Plant	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0
Animal	0.0	0.0	15.1	59.0	0.0	18.6
	64.4	24.4	0.0	0.0	29.3	0.0
Grouping	0.0	41.3	80.7	96.0	46.7	82.7
	NA	NA	NA	NA	NA	NA

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

Table 7 The evolution of high levels of assortment

Mechanism of assortment	Infection genetics (%)					
	Outcome in hosts			Outcome in parasites		
	IMA	GFG	MA	IMA	GFG	MA
Animal	0.0	0.0	0.0	24.0	0.0	1.3
Grouping	0.0	77.3	92.0	98.7	60.0	85.3

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 homozygotes 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 disassortatively with respect to disease-resistance loci if such loci operate through a GFG or IMA mechanism. In a similar vein, our results suggest that diploid parasites 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

REFERENCES

Agrawal, A.F. (2006). Similarity selection and the evolution of sex: revisiting the red queen. *PLOS Biol.*, 4, 1364–1371.

Barton, N.H. & Turelli, M. (1991). Natural and sexual selection on many loci. *Genetics*, 127, 229–255.

Burdon, J.J. (1997). The evolution of gene-for-gene interactions in natural pathosystems. In: *The Gene-for-Gene Relationship in Plant-Parasite Interactions* (eds Crute, I.R., Holub, E.B. & Burdon, J.J.). CAB International, Wallingford, UK, pp. 245–262.

Bush, G.L. (1994). Sympatric speciation in animals: new wine in old bottles. *TREE*, 9, 285–288.

Craig, T.P., Itami, J.K., Abrahamson, W.G. & Horner, J.D. (1993). Behavioral evidence for host-race formation in *Eurosta solidaginis*. *Evolution*, 47, 1696–1710.

Dieckmann, U. & Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, 400, 354–357.

Dieckmann, U., Doebeli, M., Metz, J.A.J. & Tautz, D. (2004). *Adaptive Speciation*. Cambridge University Press, Cambridge, UK.

Doebeli, M. (1996). A quantitative genetic competition model for sympatric speciation. *J. Evol. Biol.*, 9, 893–909.

Doebeli, M. & Dieckmann, U. (2000). Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Am. Nat.*, 156, S77–S101.

Doebeli, M. & Dieckmann, U. (2003). Speciation along environmental gradients. *Nature*, 421, 259–264.

Dybdahl, M.F. & Lively, C.M. (1998). Host-parasite coevolution: evidence for rare advantage and time-lagged selection in a natural population. *Evolution*, 52, 1057–1066.

Ekblom, R., Saether, S.A., Grahn, M., Fiske, P., Kalas, J.A. & Hoglund, J. (2004). Major histocompatibility complex variation and mate choice in a lekking bird, the great snipe (Gallinago media). *Mol. Ecol.*, 13, 3821–3828.

Frank, S.A. (2002). *Immunology and the Evolution of Infectious Disease*. Princeton University Press, Princeton.

Froeschke, G. & Sommer, S. (2005). MHC class II DRB variability and parasite load in the striped mouse (Rhabdomys pumilio) in the southern Kalahari. *Mol. Biol. Evol.*, 22, 1254–1259.

Gandon, S. & Otto, S.P. (2007). The Evolution of Sex and Recombination in Response to Abiotic or Coevolutionary Fluctuations in Epistasis. *Genetics*, 175, 1835–1853.

Gavrilets, S. (2004). Fitness landscapes and the origin of species. Princeton University Press, Princeton.

Hamilton, W.D. (1980). Sex vs. non-sex vs. parasite. *Oikos*, 35, 282–290.

Hedrick, P.W. (2002). Pathogen resistance and genetic variation at MHC loci. *Evolution*, 56, 1902–1908.

Howard, R.S. & Lively, C.M. (1998). The maintenance of sex by parasitism and mutation accumulation under epistatic fitness functions. *Evolution*, 52, 604–610.

Howard, R.S. & Lively, C.M. (2003). Opposites attract? Mate choice for parasite evasion and the evolutionary stability of sex. *J. Evol. Biol.*, 16, 681–689.

Howard, R.S. & Lively, C.M. (2004). Good vs complementary genes for parasite resistance and the evolution of mate choice. *BMC Evol. Biol.*, 4, Article No. 48.

Kirkpatrick, M. & Nuismer, S.L. (2004). Sexual selection can constrain sympatric speciation. *Proc. R. Soc. Lond., B, Biol. Sci.*, 271, 687–693.

Kirkpatrick, M., Johnson, T. & Barton, N. (2002). General models of multilocus evolution. *Genetics*, 161, 1727–1750.

Landry, C., Garant, D., Duchesne, P. & Bernatchez, L. (2001). 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proc. R. Soc. Lond., B, Biol. Sci.*, 268, 1279–1285.

Lively, C.M., Dybdahl, M.E., Jokela, J., Osnas, E.E. & Delph, L.E. (2004). Host sex and local adaptation by parasites in a snail-trematode interaction. *Am. Nat.*, 164, S6–S18.

Milinski, M. (2006). The major histocompatibility complex, sexual selection, and mate choice. *Ann. Rev. Ecol. Evol. Syst.*, 37, 159–186.

Nagylaki, T.J. (1993). The evolution of multilocus systems under weak selection. *Genetics*, 134, 627–647.

Nuismer, S.L. (2006). Parasite local adaptation in a geographic mosaic. *Evolution*, 60, 83–88.

Nuismer, S.L. & Otto, S.P. (2004). Host-parasite interactions and the evolution of ploidy. *Proc. Natl Acad. Sci. U.S.A.*, 101, 11036–11039.

Nuismer, S.L. & Otto, S.P. (2005). Host-parasite interactions and the evolution of gene expression. *PLOS Biol.*, 3, 1283–1288.

Otto, S.P. & Nuismer, S.L. (2004). Species interactions and the evolution of sex. *Science*, 304, 1018–1020.

Otto, S.P., Servedio, M.R. & Nuismer, S.L. (2008). Frequency-dependent selection and the evolution of assortative mating. *Genetics* (in press).

Penn, D.J. & Potts, W.K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *Am. Nat.*, 153, 145–164.

Penn, D.J., Damjanovich, K. & Potts, W.K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc. Natl Acad. Sci. U.S.A.*, 99, 11260–11264.

Peters, A.D. & Lively, C.M. (1999). The red queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution. *Am. Nat.*, 154, 393–405.

Potts, W.K., Manning, C.J. & Wakeland, E.K. (1994). The role of infectious-disease, inbreeding and mating preferences in maintaining MHC genetic diversity – an experimental test. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 346, 369–378.

Sasaki, A. (2000). Host-parasite coevolution in a multilocus gene-for-gene system. *Proc. R. Soc. Lond., B, Biol. Sci.*, 267, 2183–2188.

Schwensow, N., Eberle, M. & Sommer, S. (2008). Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proc. R. Soc. Lond., B, Biol. Sci.*, 275, 555–564.

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, 53, 1446–1457.

Waxman, D. & Gavrilets, S. (2005). 20 questions on adaptive dynamics. *J. Evol. Biol.*, 18, 1139–1154.

Westerdahl, H., Waldenstrom, J., Hansson, B., Hasselquist, D., von Schantz, T. & Bensch, S. (2005). Associations between malaria and MHC genes in a migratory songbird. *Proc. R. Soc. Lond., B, Biol. Sci.*, 272, 1511–1518.

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.

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

Scott L. Nuismer[1]

Sarah P. Otto[2]

Francois Blanquart[1,3]

[1] *Department of Biological Sciences, University of Idaho, Moscow, ID 83844*

[2]*Department of Zoology, University of British Columbia, Vancouver, BC*

[3]*École 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:

$$\frac{W(\mathbf{X}_i)}{\overline{W}_i} = 1 + \sum_{\mathbf{U} \subseteq \mathbf{W}} a_{i,\mathbf{U}} (\zeta_{i,\mathbf{U}} - D_{i,\mathbf{U}}) \tag{S1}$$

where $W(\mathbf{X_i})$ is the fitness of genotype $\mathbf{X}$ in species i, $\overline{W}_i$ is the population mean fitness of species i, $\mathbf{W}$ is the set of positions that contribute to fitness, and $a_{i,\mathbf{U}}$ is the "selection coefficient" acting on the set of positions $\mathbf{U}$ 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 $\mathbf{W}$, 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 $\psi(\mathbf{X_H}, \mathbf{X_P})$ from Table 2 of the main text, to calculate selection coefficients in the host ($a_{H,\mathbf{U}}$) and parasite ($a_{P,\mathbf{U}}$) 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:

35 $\quad a_{H,B_m} = a_{H,B_f} = s_H[p_{P,B}^2 - p_{H,B}(1-2p_{P,B}q_{P,B})]+O(\varepsilon^2)$, $\qquad$ (S2a)

36 $\quad$ and the selection coefficient favoring homozygotes over heterozygotes is:

37 $\quad a_{H,B_mB_f} = -s_H(1-2p_{P,B}q_{P,B})+O(\varepsilon^2)$ $\qquad$ (S2b)

38 $\quad$ where $p_{i,B}$ is the frequency of the B allele in species i, and $q_{i,B}$ is the frequency of the b

39 $\quad$ allele in species i. Equation (S2b) shows that the sign of a_{H,B_mB_f} is always negative,

40 $\quad$ demonstrating that the IMA model favours increased heterozygosity in the host.

41 $\quad$ For the parasite, we find that the selection coefficients are:

42 $\quad a_{P,B_m} = a_{P,B_f} = s_P[p_{P,B}(1-2p_{H,B}q_{H,B}) - p_{H,B}^2]+O(\varepsilon^2)$ $\qquad$ (S3a)

43 $\quad a_{P,B_mB_f} = s_P(1-2p_{H,B}q_{H,B})+O(\varepsilon^2)$. $\qquad$ (S3b)

44 $\quad$ Equation (S3b) shows that the sign of a_{P,B_mB_f} is always positive, demonstrating that the

45 $\quad$ IMA model favours increased homozygosity in the parasite.

46 $\quad$ *Gene-for-gene*

47 $\quad$ For the gene-for-gene model, we find that the selection coefficients favoring the B allele

48 $\quad$ over the b allele in the host are:

49 $\quad a_{H,B_m} = a_{H,B_f} = q_{H,B}[s_H(1 - p_{P,B}^2) - \tau_H]+O(\varepsilon^2)$ $\qquad$ (S4a)

50 $\quad$ and the selection coefficient favoring homozygotes over heterozygotes is:

51 $\quad a_{H,B_mB_f} = -s_H(1 - p_{P,B}^2)+\tau_H +O(\varepsilon^2)$. $\qquad$ (S4b)

52 $\quad$ Equation (S4b) shows that the sign of a_{H,B_mB_f} depends on the strength of selection for

53 $\quad$ increased resistance imposed by the parasite $s_H(1- p_{P,B}^2)$ relative to the strength of

54 $\quad$ selection for reduced resistance generated by costs τ_H. Consequently, whether the

55 $\quad$ GFG model favors heterozygosity or homozygosity in the host depends on the

56 $\quad$ frequency of the parasite virulence allele.

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

58    $a_{P,B_m} = a_{P,B_f} = p_{P,B}[s_P(p_{H,B}^2 + 2p_{H,B}q_{H,B}) - \tau_P] + O(\varepsilon^2)$ (S5a)

59    $a_{P,B_m B_f} = s_P(p_{H,B}^2 + 2p_{H,B}q_{H,B}) - \tau_P + O(\varepsilon^2)$.   (S5b)

60   Equation (S5b) shows that the sign of $a_{P,B_m B_f}$ depends on the strength of selection for

61   increased virulence imposed by the host $s_P(p_{H,B}^2 + 2p_{H,B}q_{H,B})$ relative to the strength of

62   selection for reduced virulence generated by costs τ_P. Consequently, whether the GFG

63   model favors heterozygosity or homozygosity in the parasite depends on the frequency

64   of the parasite virulence allele.

65   *Matching alleles*

66   For the matching-alleles model, we find that the selection coefficients favoring the *B*

67   allele over the *b* allele in the host are:

68    $a_{H,B_m} = a_{H,B_f} = s_H[p_{H,B}(1 + 2p_{P,B}q_{P,B}) - (p_{P,B}^2 + 2p_{P,B}q_{P,B})] + O(\varepsilon^2)$ (S6a)

69   and the selection coefficient favoring homozygotes over heterozygotes is:

70    $a_{H,B_m B_f} = s_H(1 + 2p_{P,B}q_{P,B}) + O(\varepsilon^2)$.   (S6b)

71   Equation (S6b) shows that the sign of $a_{H,B_m B_f}$ is always positive, demonstrating that the

72   MA model favours increased homozygosity in the host.

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

74    $a_{P,B_m} = a_{P,B_f} = s_P[p_{P,B}(1 - 2p_{H,B}q_{H,B}) - q_{H,B}^2] + O(\varepsilon^2)$ (S7a)

75    $a_{P,B_m B_f} = s_P(1 - 2p_{H,B}q_{H,B}) + O(\varepsilon^2)$.   (S7b)

76   Equation (S7b) shows that the sign of $a_{P,B_m B_f}$ is always positive, demonstrating that the

77   MA model favours increased homozygosity in the parasite.

78

79   The selection coefficients (S2-S7) can be used to calculate the change in the

80   frequency of the modifier allele M that occurs in response to coevolutionary selection

81   (before mating has occurred). Specifically, Kirkpatrick et. al. (2002) showed that the

82   change in allele frequency at position j in species i due to selection is given by:

83    $$\Delta p_{i,j} = a_{i,j} p_{i,j} q_{i,j} + \sum_{\mathbf{U} \subseteq \mathbf{W} \neq j} a_{i,\mathbf{U}} D_{i,\mathbf{Uj}} \tag{S8}$$

84   where the notation $D_{i,\mathrm{Uj}}$ means the statistical association between positions in the set

85   that includes the set $\mathbf{U}$ and additional position j in species i. Because the modifier does

86   not directly influence the coevolutionary interactions ($a_{i,j} = 0$), equation (S8) along

87   with the selection coefficients (S2-S7) gives the change in the frequency of the modifier

88   allele M in species i:

89    $$\Delta p_{i,M} = a_{i,B_f}(D_{i,M_f B_f} + D_{i,M_m B_f}) + a_{i,B_m}(D_{i,M_m B_m} + D_{i,M_f B_m}) + a_{i,B_m B_f}(D_{i,M_m B_m B_f} + D_{i,M_f B_m B_f})$$

90   .

91   (S9)

92   Equation (S9) demonstrates that coevolutionary selection changes the frequency of the

93   modifier allele only indirectly through its statistical associations with the selected locus

94   B. There are two ways in which this can occur. First, a modifier of non-random mating

95   may change in frequency if it becomes statistically associated with the B locus allele

96   currently favoured or disfavoured by directional coevolutionary selection. This force is

97   reflected in the first two terms of (S9). Second, a modifier of non-random mating may

98   change in frequency if it becomes statistically associated with homozygosity or

99   heterozygosity at the B locus. If a particular coevolutionary model favours increased

100   heterozygosity (e.g., the host in an IMA model), then modifiers that become statistically

101   associated with heterozygous B locus genotypes will increase in frequency. This force is

102   reflected in the third term of (S9).

103   In addition to showing how the selection coefficients (e.g., S2-S7) can be used

104   to calculate changes in allele frequency, Kirkpatrick et. al. (2002) showed how they can

105   be used to calculate the change in statistical associations among a set of positions **A** in

106   species i that occur in response to selection:

107    $$\Delta D_{i,\mathbf{A}} = \sum_{\mathbf{U} \subseteq \mathbf{W}} a_{i,\mathbf{U}} (D_{i,\mathbf{AU}} - D_{i,\mathbf{A}} D_{i,\mathbf{U}})$$ (S10)

108   where the notation $D_{i,\mathbf{AU}}$ indicates the statistical association between the set of positions

109   **A** in species i after the set of positions in **U** has been removed. Combined with the

110   selection coefficients (S2-S7), equation (S10) shows that coevolutionary selection

111   generates only a single type of statistical association — deviations from Hardy-

112   Weinberg. Specifically, coevolutionary selection leads to the following change in

113   statistical associations within the B locus:

114    $$\Delta D_{i,B_m B_f} = a_{i,B_m B_f} p_{i,B} q_{i,B}$$ (S11)

115

116   **MATING**

117   Kirkpatrick et. al. (2002) showed that the changes in allele frequencies and

118   statistical associations between positions caused by non-random mating can be

119   calculated once the frequencies of mated pairs are known. As with viability selection,

120   the first step is to calculate the selection coefficients acting on positions in the genome

121   using:

122    $$\frac{f(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})}{f(\mathbf{X_{i,m}}) f(\mathbf{X_{i,f}})} = 1 + \sum_{\mathbf{U} \subseteq \mathbf{W}} a_{i,\mathbf{U}} (\zeta_{i,\mathbf{U}} - D_{i,\mathbf{U}})$$ (S12)

123   where $f(\mathbf{X_{i,m}}, \mathbf{X_{i,f}})$ is the frequency of a mated pair consisting of a male with genotype

124    $\mathbf{X_{i,m}}$ and a female with genotype $\mathbf{X_{i,f}}$ (equations 6, 8, and 9 of the main text) and $f(\mathbf{X_{i,m}})$

125   and $f(\mathbf{X_{i,f}})$ are the frequencies of male and female genotypes, respectively, prior to non-

126　random mating. Because the fitness set **W** is very large for the forms of non-random

127　mating considered here, we do not show the results for the numerous selection

128　coefficients that emerge from (S12), although a *Mathematica* package which contains

129　these formulae is available upon request. After calculating the selection coefficients

130　using (S12), the changes in allele frequencies and statistical associations between

131　positions can be calculated using equations (S8) and (S10) respectively. These results

132　are quite numerous and complex, so we do not report them here although they are

133　available upon request in the form of a *Mathematica* package.

134　　　　Non-random mating can itself cause changes in the modifier frequency for three

135　reasons. First, to the extent that there are costs of assortative mating (as in the plant

136　model, where choosier females have lower fitness), selection will act directly on the

137　modifier to reduce mating preferences. Second, non-random mating can cause

138　homozygotes and heterozygotes to differ in mating success, which can favor the

139　evolution of assortment (when homozygotes are better able to find mates) or

140　disassortment (when heterozygotes are better able to find mates). Finally, non-random

141　mating generates associations between alleles that are increasing in frequency and

142　modifiers that prefer that allele (as in models of Fisherian sexual selection); because

143　assortative mating causes a mating advantage to common alleles, this third force favors

144　the evolution of assortative mating when the allele rising in frequency at the B locus is

145　common but hinders it when the allele is rare.   More details about how these three

146　forces combine can be found in the one-species model of Otto et al. (submitted).

147

148

149

150 **SEGREGATION AND RECOMBINATION**

151   Kirkpatrick et. al. (2002) showed that the change in statistical associations

152 caused by segregation and recombination in species *i* is:

153 $$\Delta D_{i,\mathbf{A}} = \sum_{\mathbf{U}:U=A} t_{i,\mathbf{A}\leftarrow\mathbf{U}} D_{i,\mathbf{U}} , \qquad (S13)$$

154 where $t_{i,\mathbf{A}\leftarrow\mathbf{U}}$ is the probability that the positions in set **U** are transmitted into set **A** in

155 species *i*, and the notation **U**:*U=A* indicates that the sum is taken over all possible sets **U**

156 which could be transmitted into set **A** once the context information (e.g., sex of origin,

157 sex of carrier) is stripped from them. Once again, equation (S13) yields numerous

158 complex expressions and so we refrain from writing them out; they are available in a

159 *Mathematica* notebook upon request.

160

161 **QUASI-LINKAGE EQUILIBRIUM APPROXIMATION**

162 In order to derive tractable expressions for the change in modifier frequency, we made a

163 quasi-linkage equilibrium approximation. In addition to assuming that selection is

164 relatively weak (order ε), our approximation assumed: 1) non-random mating is not too

165 strong, 2) the effect of the modifier is not too great, and 3) sex and recombination are

166 relatively frequent. With these assumptions, statistical associations between positions

167 should change quickly relative to allele frequencies and approach a quasi-steady state

168 where the values of these associations are small and of order ε. Under these conditions,

169 it is possible to solve for the quasi-linkage equilibrium (QLE) values of the statistical

170 associations between positions. These QLE values can then be used in conjunction with

171 equation (S9) and the selection coefficients imposed by species interactions (S2-S7) and

172 non-random mating (available as *Mathematica* notebook) to calculate the change in the

173   frequency of the modifier allele due to species interactions and non-random mating.

174   Because we have assumed that selection is weak (order ε) and that the system is at QLE,

175   the change in modifier frequency over a single generation is simply the sum of the

176   change due to species interactions and non-random mating. In the following sections,

177   we provide results for the change in modifier frequency caused by non-random mating

178   and for the change in modifier frequency caused by species interactions for each of the

179   three models of non-random mating.

180   **QLE results for the plant model**

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

182   We find that the change in the frequency of a modifier of assortative mating in species i

183   is:

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

185   and a modifier of disassortative mating is:

186   $$\Delta p_{i,M} = -\frac{\delta_{\rho,i}}{2}[p_{i,M}q_{i,M}\chi] + O(\varepsilon^2) \qquad (S14b)$$

187   where χ is the probability of encountering a mate with a matching genotype

188   ($\chi = p_{i,B}^4 + 4p_{i,B}^2 q_{i,B}^2 + q_{i,B}^4$) and $\delta_{\rho,i}$ is the effect of the modifier in species i. Because

189   species interactions generate only indirect selection of order ε^2, but females that mate

190   non-randomly pay a direct cost in terms of reduced mating opportunities, this result

191   shows that non-random mating will never evolve in the plant model.

192   **QLE results for the animal model**

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

194   We find that the change in the frequency of a modifier due to non-random mating in

195   species i is:

196 $$\Delta p_{i,M} = \frac{1}{2}\overline{\rho}_i \delta_{\rho,i} p_{i,M} q_{i,M} p_{i,B} q_{i,B} (1-3p_{i,B}q_{i,B})[1-6p_{i,B}q_{i,B}(1-p_{i,B}q_{i,B})]+O(\varepsilon^3) \quad \text{(S15)}$$

197 where $\overline{\rho}_i$ is the average level of assortment or disassortment in species i. In contrast to

198 the plant model, this result shows that the animal model generates no direct selection on

199 the modifier locus (of order ε), instead imposing indirect sexual selection of the same

200 order (ε^2) as that caused by species interactions. This indirect selection can favor or

201 disfavor a modifier of non-random mating. Specifically, modifiers increasing the degree

202 of non-random mating (either assortment or disassortment) are favored by indirect

203 selection anytime the allele frequency at the mating trait locus lies outside of the

204 interval {0.30, 0.70} but are disfavored by indirect selection anytime the allele

205 frequency at the mating trait locus lies inside this interval. In brief, when allele

206 frequencies are intermediate, the presence of assortative mating increases the relative

207 mating success of heterozygotes compared to the average homozygote (because

208 heterozygotes are common and they prefer other heterozygotes); this heterozygous

209 mating advantage counteracts the evolution of further assortment. Similarly, the

210 presence of disassortative mating increases the relative mating success of homozygotes

211 compared to the heterozygotes (because heterozygotes are common and they prefer not

212 to mate with other heterozygotes); this homozygous mating advantage counteracts the

213 evolution of further disassortment. These arguments apply only when allele frequencies

214 are sufficiently close to 1/2. When the allele frequency at the mating trait locus is far

215 from 1/2, heterozygotes are rare and the above arguments reverse.

216

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

218 *Inverse matching alleles model* − We find that IMA interactions with the parasite lead to

219 the following change in the frequency of a modifier of non-random mating:

220 $$\Delta p_{H,M} = -\theta \delta_{\rho,H} \lambda s_H p_{H,M} q_{H,M} p_{H,B} q_{H,B} (1 - 3 p_{H,B} q_{H,B}) + O(\varepsilon^3) \qquad (S16)$$

221 where θ is an indicator variable which takes the value +1 for a modifier of assortative

222 mating and -1 for a modifier of disassortative mating, and λ is the probability that a

223 randomly selected host is infected by a randomly selected parasite. This result

224 demonstrates that interactions with the parasite favor increased disassortment.

225 Comparing result (S16) with result (S15) shows that disassortative mating is guaranteed

226 to evolve only when the average level of assortment in the population is small relative

227 to the strength of selection exerted by species interactions.

228 *Gene-for-gene model*− We find that the change in the frequency of a modifier of non-

229 random mating in the host caused by GFG interactions with the parasite is:

230 $$\Delta p_{H,M} = -\theta \delta_{\rho,H} p_{H,M} q_{H,M} p_{H,B} q_{H,B} [q_{H,B}^2 (1 - 3 p_{H,B} q_{H,B})(s_H (1 - p_{P,B}^2) - \tau_H)] + O(\varepsilon^3)$$

231 (S17)

232 This result demonstrates that interactions with the parasite favor increased disassortment

233 any time costs of resistance (τ_H) are small relative to selection for increased resistance.

234 Comparing result (S17) with result (S15) shows that disassortative mating is then

235 guaranteed to evolve only when the average level of assortment in the population is

236 small relative to the strength of selection exerted by species interactions and when costs

237 of resistance (τ_H) are small relative to selection for increased resistance.

238 *Matching alleles model* − We find that the change in the frequency of a modifier of non-

239 random mating in the host caused by MA interactions with the parasite is:

240 $$\Delta p_{H,M} = \theta \delta_{\rho,H} s_H p_{H,M} q_{H,M} p_{H,B} q_{H,B} (1 - 3 p_{H,B} q_{H,B})(1 - \lambda) + O(\varepsilon^3) \qquad (S18)$$

241   This result demonstrates that interactions with the parasite always favor increased

242   assortment in the host. Comparing result (S18) with result (S15) shows that assortative

243   mating is guaranteed to evolve only when the average level of assortment in the

244   population is small relative to the strength of selection exerted by species interactions.

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

246   *Inverse matching alleles model* − We find that the change in the frequency of a modifier

247   of non-random mating in the parasite caused by IMA interactions with the host is:

248     $$\Delta p_{P,M} = \theta \delta_{\rho,P} s_P p_{P,M} q_{P,M} p_{P,B} q_{P,B} \lambda (1 - 3 p_{P,B} q_{P,B}) + O(\varepsilon^3) \qquad \text{(S19)}$$

249   This result demonstrates that interactions with the host always favor increased

250   assortment. Comparing result (S19) with result (S15) shows that assortative mating is

251   guaranteed to evolve only when the average level of assortment in the population is

252   small relative to the strength of selection exerted by species interactions.

253   *Gene-for-gene model*− We find that the change in the frequency of a modifier of non-

254   random mating in the parasite caused by GFG interactions with the host is:

255     $$\Delta p_{P,M} = \theta \delta_{\rho,P} p_{P,M} q_{P,M} p_{P,B} q_{P,B} [p_{P,B}^2 (1 - 3 p_{P,B} q_{P,B})(s_P (1 - q_{H,B}^2) - \tau_P)] + O(\varepsilon^3) \quad \text{(S20)}$$

256   This result demonstrates that interactions with the host favor increased assortment

257   anytime costs of virulence (τ_P) are small relative to selection for increased virulence.

258   Comparing result (S20) with result (S15), shows that assortative mating is guaranteed to

259   evolve only when the average level of assortment in the population is small relative to

260   the strength of selection exerted by species interactions and when costs of virulence (τ_P)

261   are small relative to selection for increased virulence.

262

263   *Matching alleles model* − We find that the change in the frequency of a modifier of non-

264   random mating in the parasite caused by MA interactions with the host is:

265 $\quad \Delta p_{P,M} = \theta \delta_{\rho,P} s_P p_{P,M} q_{P,M} p_{P,B} q_{P,B} (1 - 3 p_{P,B} q_{P,B})[\lambda - 2 p_{H,B} q_{H,B}] + O(\varepsilon^3)$ $\qquad$ (S21)

266 It is possible to show that S21 is always positive, demonstrating that interactions with

267 the host always favor increased assortment. Comparing result (S21) with result (S15),

268 shows that assortative mating is guaranteed to evolve only when the average level of

269 assortment in the population is small relative to the strength of selection exerted by

270 species interactions.

271 **QLE results for the grouping model**

272 *Change due to assortative mating*

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

274 $\quad \Delta p_{i,M} = 0 + O(\varepsilon^3)$ $\qquad$ (S22)

275 showing that selection induced by non-random mating in the grouping model is of lower

276 order than that generated by species interactions. For this reason, the change in the

277 frequency of a modifier of assortative mating in the grouping model is, to leading order,

278 driven only by species interactions.

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

280 *Inverse matching alleles model* − We find that the change in the frequency of a modifier

281 of assortative mating in the host caused by IMA interactions with the parasite is:

282 $\quad \Delta p_{H,M} = -\dfrac{1}{4} \delta_{\rho,H} s_H p_{H,M} q_{H,M} p_{H,B} q_{H,B}[1 - 2 p_{P,B} q_{P,B}] + O(\varepsilon^3)$ $\qquad$ (S23)

283 This result shows that IMA interactions with the parasite never lead to the evolution of

284 assortative mating in the host.

285 *Gene-for-gene model* − We find that the change in the frequency of a modifier of

286 assortative mating in the host caused by GFG interactions with the parasite is:

287 $\quad \Delta p_{H,M} = -\dfrac{1}{4} \delta_{\rho,H} p_{H,M} q_{H,M} p_{H,B} q_{H,B}[s_H (1 - p_{P,B}^2) - \tau_H] + O(\varepsilon^3)$ $\qquad$ (S24)

288   This result shows that GFG interactions with the parasite lead to the evolution of

289   assortative mating in the host only if costs of resistance (τ_H) are large relative to

290   selection for increased resistance $s_H(1 - p_{P,B}^2)$.

291   *Matching alleles model* − We find that the change in the frequency of a modifier of

292   assortative mating in the host caused by MA interactions with the parasite is:

293     $$\Delta p_{H,M} = \frac{1}{4}\delta_{\rho,H}s_H p_{H,M}q_{H,M}p_{H,B}q_{H,B}[1 + 2p_{P,B}q_{P,B}] + O(\varepsilon^3) \qquad (S25)$$

294   This result shows that MA interactions with the parasite always lead to the evolution of

295   assortative mating in the host.

296   <u>*Change in parasite modifier frequency caused by interactions with host*</u>

297   *Inverse matching alleles model* − We find that the change in the frequency of a modifier

298   of assortative mating in the parasite caused by IMA interactions with the host is:

299     $$\Delta p_{P,M} = \frac{1}{4}\delta_{\rho,P}s_P p_{P,M}q_{P,M}p_{P,B}q_{P,B}[1 - 2p_{H,B}q_{H,B}] + O(\varepsilon^3) \qquad (S26)$$

300   This result shows that IMA interactions with the host always cause the evolution of

301   assortative mating in the parasite.

302   *Gene-for-gene model*− We find that the change in the frequency of a modifier of

303   assortative mating in the parasite caused by GFG interactions with the host is:

304     $$\Delta p_{P,M} = \frac{1}{4}\delta_{\rho,P}p_{P,M}q_{P,M}p_{P,B}q_{P,B}[s_P(1 - q_{H,B}^2) - \tau_P] + O(\varepsilon^3) \qquad (S27)$$

305   This result shows that GFG interactions with the host cause the evolution of assortative

306   mating in the parasite anytime the costs of virulence (τ_P) are small relative to selection

307   for increased virulence $s_P(1 - q_{H,B}^2)$.

308   *Matching alleles model* − We find that the change in the frequency of a modifier of

309   assortative mating in the parasite caused by MA interactions with the host is:

310 $$\Delta p_{P,M} = \frac{1}{4}\delta_{\rho,P}s_P p_{P,M}q_{P,M}p_{P,B}q_{P,B}[1-2p_{H,B}q_{H,B}]+O(\varepsilon^3) \qquad (S28)$$

311 This result shows that MA interactions with the host always cause the evolution of

312 assortative mating in the parasite.

313 Table S1. Percentage of three-locus simulations in which assortative mating (top entry)

314 and dissasortative mating (bottom entry) evolved in (a) the host and (b) the parasite.

Mechanism of assortment	Infection genetics			Infection genetics		
	IMA	GFG	MA	IMA	GFG	MA
Plant	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Animal	0.0%	0.0%	22.6%	37.2%	1.30%	32.4%
	39.4%	28.0%	0.0%	0.0%	25.0%	0.0%
Grouping	0.0%	35.3%	85.3%	81.3%	40.7%	81.3%
	NA	NA	NA	NA	NA	NA

315