



The role of epistasis on the evolution of recombination in host–parasite coevolution

Roger D. Kouyos^{a,*}, Marcel Salathé^b, Sarah P. Otto^c, Sebastian Bonhoeffer^a

^a Institute of Integrative Biology, ETH Zürich, ETH-Zentrum CHN, 8092 Zürich, Switzerland

^b Department of Biology, Stanford University, Stanford, CA 94305-5020, USA

^c Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

ARTICLE INFO

Article history:

Received 5 May 2008

Available online 14 October 2008

Keywords:

Red queen

Host–parasite interactions

Coevolution

Genetic systems

Sexual reproduction

Population genetics

ABSTRACT

Antagonistic coevolution between hosts and parasites is known to affect selection on recombination in hosts. The Red Queen Hypothesis (RQH) posits that genetic shuffling is beneficial for hosts because it quickly creates resistant genotypes. Indeed, a large body of theoretical studies have shown that for many models of the genetic interaction between host and parasite, the coevolutionary dynamics of hosts and parasites generate selection for recombination or sexual reproduction. Here we investigate models in which the effect of the host on the parasite (and vice versa) depend approximately multiplicatively on the number of matched alleles. Contrary to expectation, these models generate a dynamical behavior that strongly selects against recombination/sex. We investigate this atypical behavior analytically and numerically. Specifically we show that two complementary equilibria are responsible for generating strong linkage disequilibria of opposite sign, which in turn causes strong selection against sex. The biological relevance of this finding stems from the fact that these phenomena can also be observed if hosts are attacked by two parasites that affect host fitness independently. Hence the role of the Red Queen Hypothesis in natural host parasite systems where infection by multiple parasites is the rule rather than the exception needs to be reevaluated.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

The interaction between hosts and parasites is one of the driving forces in evolution. In particular, it is well known that host–parasite interactions can affect the evolution of reproductive strategies that involve genetic shuffling. The so-called Red Queen hypothesis (RQH) claims that sex and recombination are maintained because they confer on the hosts an edge in the evolutionary arms race with various parasites (Jaenike, 1978). Although this hypothesis has been shown to work in principle (Hamilton et al., 1990; Peters and Lively, 1999; Schmid-Hempel and Jokela, 2002), the plausibility of the assumptions under which it works are hotly debated both in the theoretical (Gandon and Otto, 2007; Kouyos et al., 2007; Otto and Nuismer, 2004; Peters and Lively, 2007; Salathé et al., 2008) and the experimental literature (Fischer and Schmid-Hempel, 2005; Little, 2002; Lively, 1987; Lively et al., 1990, 1998).

It has been recognized early that the impact of Red Queen dynamics on the evolution of genetic shuffling depends crucially

on the specific type of fitness interaction between host and parasite (Bell and Smith, 1987). The subsequent theoretical discussion on this issue was dominated by the debate over the gene-for-gene and the matching-allele model (Agrawal and Lively, 2002; Frank, 1993a,b, 1994, 1996a,b; Parker, 1994, 1996): Whereas the plant-specific gene-for-gene model typically selects against genetic shuffling (Otto and Nuismer, 2004; Parker, 1994, 1996), most models in which the RQH has been shown to work belong to the matching-allele class; i.e. the fitness of host and parasite depend on the number of host loci matched by the parasite (Kouyos et al., 2007; Otto and Nuismer, 2004; Peters and Lively, 1999, 2007; Salathé et al., 2008; Schmid-Hempel and Jokela, 2002). However, even matching-allele models do not always select for higher rates of genetic shuffling; instead, the direction of evolution depends crucially on the fitness function; i.e. on how the number of matched loci determines the fitness of host and parasite (Kouyos et al., 2007; Salathé et al., 2008).

The building blocks of matching-allele models are pairs of loci (one such pair consists of one locus in the host and one corresponding locus in the parasite) that determine the outcome of infection. Usually, a host–parasite interaction is governed by several of those pairs. One important criterion to classify different matching-alleles (MA) models is the degree to which epistatic

* Corresponding author.

E-mail address: roger.kouyos@env.ethz.ch (R.D. Kouyos).

interactions between the different pairs determine the outcome of the host–parasite interaction. The most commonly used MA model assumes that the parasite can infect the host only if it matches the host at all interaction loci, thus assuming a very strong epistatic interaction between alleles at different loci. The other extreme is the so-called multiplicative matching-alleles model (MMA) where the number of matched alleles determines the fitness of host and parasite in a multiplicative way. This type of model is of particular relevance because it constitutes a “null-model” as it involves no implicit epistatic interactions between alleles at different loci. Other matching-alleles models can be obtained by adding an epistatic interaction to the MMA. Intuitively epistasis seems to play an important role for the RQH, because the hypothesis is based on the phase shift between the oscillations of epistasis and LD (Peters and Lively, 1999). Accordingly, several experimental studies have attempted to determine the amount of epistasis between the loci that govern the host parasite interaction (Kover and Caicedo, 2001; Wilfert et al., 2007a,b). Epistasis in host-parasite systems can come from two distinct sources. On one hand epistasis in the host population can be generated through statistical associations in the parasite population and vice versa. On the other hand epistasis can come from the direct genetic interactions between genes that determine the outcome of infection for an individual host and parasite combination. From a theoretical point of view, the relevance of the latter source of epistasis remains unclear and is the subject of this study.

Recently, we studied the evolution of recombination for a wide range of matching-alleles models, including both the MA and the MMA (Kouyos et al., 2007). We found that higher rates of genetic shuffling are strongly selected against in the MMA and closely related models. We related this effect to the observation that the MMA-systems seemed to converge to a state with strong positive or negative linkage disequilibrium (LD). In the example considered in this study, linkage between selected loci was strong and we showed that the sign of the LD depends on the initial conditions, but it does not change during the course of a simulation. By contrast, the only other theoretical study that considered the MMA (Otto and Nuismer, 2004), found that for this interaction model neither higher nor lower recombination rates were favored.

In multiplicative one-species models, both types of behaviors have been observed. On the one hand, Maynard-Smith (Smith, 1968) showed that in haploid systems with multiplicative selection no linkage disequilibrium is built up. Hence recombination has no effect in such systems. In diploid models with multiplicative selection, on the other hand, stable equilibria with strong LD have been described as high complementarity equilibria (HCE) (Bodmer and Felsenstein, 1967; Feldman et al., 1974; Franklin and Lewontin, 1970).

The potential relevance of HCE for the Red Queen Hypothesis is that, according to the reduction principle (Altenberg and Feldman, 1987), such strong and constant LD are expected to result in selection against genetic shuffling. Typically, host–parasite coevolution has been thought to result in oscillating LD, which can potentially favor the evolutionary maintenance of recombination. The HCE represent an alternative type of dynamic behavior, which has the opposite effect on the evolution of recombination. Assessing the impact of host–parasite interactions on the evolution of recombination therefore crucially requires knowledge of the conditions under which these opposite dynamical behaviors occur.

Although we mentioned the occurrence of such HCE in a previous publication (Kouyos et al., 2007), the conditions under which they occur has not been analyzed yet. In this article we extend (Kouyos et al., 2007) in several ways. First, we determine (i) the range of matching-alleles models where this behavior occurs and (ii) how mutation and recombination affect its occurrence and its strength. To this end we employ both an analytical approach

Table 1
Model parameters.

Symbol	Explanation
w_i^h	Fitness of host i interacting with parasite j
w_j^p	Fitness of parasite i interacting with host j
s_j	Strength of selection on the host if j loci are matched by the parasite
m_h	Mutation rate for the host
m_p	Mutation rate for the parasite
r_h	Recombination rate for the host
r_p	Recombination rate for the parasite
n_{PG}	Number of parasite generations per host generation

and numerical simulations. Second, we investigate the structural robustness of our findings, by considering both a discrete-time and a continuous-time version of our model. Third, we show that strong LD of constant sign can occur even when a host interacts with two different parasites that affect its fitness in a multiplicative way.

2. Models

Our model describes a haploid host interacting with a haploid parasite (see Table 1 for a description of the model parameters). The genomes of hosts and parasites have two biallelic loci with alleles 0 and 1. The state of the system is thus given by the host and parasite frequencies $[h; p] = [h_1, h_2, h_3, h_4; p_1, p_2, p_3, p_4]$, where the indices 1–4 correspond to the haplotypes 00, 01, 10, and 11 respectively. Statistical associations between the alleles at the two loci are measured by the linkage disequilibria (LD) $D_h = h_1h_4 - h_2h_3$ in the host and $D_p = p_1p_4 - p_2p_3$ in the parasite. Unless stated otherwise, the expression “LD” refers in the following to the LD in the host.

The processes of mutation and recombination are modeled in the following way: mutation rates (host: m_h , parasite: m_p) and recombination rates (host: r_h , parasite: r_p) are fixed. Mutation changes the allele at a given locus with a rate m (if not stated differently we assume $m = 10^{-5}$) independently of the occurrence of mutations at other loci. Forward and backward mutation rate are equal (i.e. mutation is symmetric and bidirectional). The only effect of recombination is to reduce the linkage disequilibrium ($D = h_1h_4 - h_2h_3$ for hosts and $D = p_1p_4 - p_2p_3$ for parasites) by a factor $(1 - r)$, where r is the recombination rate. The process of selection is equivalent to the one described in Kouyos et al. (2007). For the sake of completeness, a full description of selection is given below.

2.1. Discrete-time model (DM)

In the discrete-time model, one step corresponds to one parasite generation. We describe different generation times for hosts and parasites, by replacing only a fraction $\alpha = 1/n_{PG}$ of the host population after one parasite generation; n_{PG} corresponds thus to the number of parasite generations per host generation. If RMS denotes the successive action of selection (S), mutation (M), and recombination (R) on the genotype frequencies (with a subscript h for host and p for parasite), we can express the change of host and parasite genotype in a time step as:

$$\begin{aligned} h(i+1) &= (1 - \alpha)h(i) + \alpha RMS_h([h(i); p(i)]) \\ p(i+1) &= RMS_p([h(i); p(i)]) \end{aligned} \quad (DM)$$

Note that formula (DM) implies that, at each time step, the whole parasite population (but only a fraction α of the host population) experiences selection. Thus for $\alpha < 1$, parasites evolve faster than hosts. By contrast, Otto and Nuismer (2004) assume that selection acts on the entire host population.

2.2. Continuous-time model (CM)

The analogous equations in continuous time are given by

$$\begin{aligned}\frac{dh(t)}{dt} &= \alpha(RMS_h([h(t); p(t)]) - h(t)) \\ \frac{dp(t)}{dt} &= (RMS_p([h(t); p(t)]) - p(t)).\end{aligned}\quad (\text{CM})$$

These equations are obtained by taking the limit of infinitesimally small time steps dt . In each time step a fraction dt of parasites and αdt of hosts is replaced (according to the discrete time equations); i.e.

$$\begin{aligned}h(t + dt) &= (1 - \alpha dt)h(t) + \alpha dt RMS_h([h(t); p(t)]) \\ p(t + dt) &= (1 - dt)p(t) + dt RMS_p([h(t); p(t)])\end{aligned}$$

or

$$\begin{aligned}\frac{h(t + dt) - h(t)}{dt} &= \alpha(RMS_h([h(t); p(t)]) - h(t)) \\ \frac{p(t + dt) - p(t)}{dt} &= (RMS_p([h(t); p(t)]) - p(t)).\end{aligned}$$

For $dt = 1$ this corresponds to the discrete model. In the limit $dt \rightarrow 0$ (i.e. continuous time) this equation becomes the dynamical equation (CM).

2.3. Selection and interaction models

Selection is determined by the fitness matrices w_{ij}^H and w_{ij}^P . Specifically, w_{ij}^H denotes the fitness of a host-genotype i interacting with a parasite-genotype j and w_{ij}^P denotes the fitness of the parasite-genotype i interacting with a host genotype j . Since the interaction probability for host i and parasite j is proportional to their frequencies, h_i and p_j , the fitness of the host-genotype i reads

$$w_i^H = \sum_j w_{ij} p_j$$

and thus the genotype frequencies of the hosts that have undergone selection read

$$h'_i = S_h([h : p])_i = h_i \frac{w_i^H}{\sum_k w_k^H h_k}.$$

The impact of selection on the parasite frequencies is calculated analogously.

We consider host–parasite interactions of the generalized matching allele type (Kouyos et al., 2007; Salathé et al., 2008). These models assume that the outcome of the infection is determined by the number of matched alleles between host and parasite (i.e. $w_{ij}^H = w_n^H = (1 - s_n)$, where n is the number of matched loci) and that, for a given interaction, the fitness of host and parasite are inversely proportional ($w_n^P = c/w_n^H$). For this generalized matching allele model, the fitness of a host matched at no, one or two loci is given by $w_0^H = 1$, $w_1^H = 1 - s_1$ and $w_2^H = 1 - s_2$. The corresponding fitnesses of the parasite are inversely related to that of the host and are given by $w_0^P = 1 - s_2$, $w_1^P = (1 - s_2)/(1 - s_1)$, and $w_2^P = 1$. The parameters s_1 and s_2 specify a broad class of interaction types, including the matching-alleles model used in most RQH studies (Peters and Lively, 1999, 2007; Schmid-Hempel and Jokela, 2002) in which the parasite can infect the host only if it matches at all loci (i.e. $s_1 = 0$) and the multiplicative matching allele model (given by $w_2 = w_1^2$ or equivalently $s_2 = 2s_1 - s_1^2$) in which the number of matched alleles determines host and parasite fitness in a multiplicative way (see section *Models* for a more detailed description of the models).

2.4. Independent parasites

In this model, a host with two biallelic loci interacts with two (independent) parasites, A and B, each having one biallelic locus. Parasite A interacts with the first host-locus, and parasite B with the second host locus. The state of the system is given by the genotype frequencies for hosts and parasites:

$$[h; pA; pB] = [h_1, h_2, h_3, h_4; pA_1, pA_2; pB_1, pB_2],$$

where, for the hosts, the indices (1, 2, 3, 4) correspond to the genotypes (00, 10, 01, 11), and for the parasites the indices (1, 2) correspond to genotypes (0, 1). Mutation and recombination are the same as in the single species model (with the only exception that parasites do not recombine because they have only one locus). Only the selection process is different. In the one-host/one-parasite model, a host of genotype ij interacts with a parasite of genotype kl . Here, in the two-parasite model, the host ij interacts with two parasites: A of genotype k and B of genotype l . A interacts with the first host locus, B with the second; i.e. A “matches” the host if $i = k$ and B matches the host if $j = l$. The number of matched loci determines the host fitness in a similar way as for the MMA: if a host is matched by neither of the two parasites, then it has (maximal fitness) 1. If it is matched by one parasite but not by the other, it has fitness $1 - s$. If it is matched by both parasites it has fitness $(1 - s)^2$. The parasite, on the other hand, has (maximal) fitness 1 if it matches the host at the corresponding locus and fitness $1 - s$ if does not match the host. Thus, the parasite frequencies determine the host fitness in the following way

$$w_i^H = \sum_{jk} (1 - s)^{A(i,j)+B(i,k)} pA_j pB_k$$

where the function $A(i, j)$ is 1 if the parasite j matches the host i at the first locus and 0 else; and $B(i, k)$ is 1 if the parasite k matches the host i at the second locus and 0 otherwise. The host frequencies in turn determine the parasite fitness

$$w_i^{PA} = \sum_j (1 - s)^{1-A(j,i)} h_j \quad \text{and} \quad w_i^{PB} = \sum_j (1 - s)^{1-B(j,i)} h_j.$$

3. Results

In the first part of this analysis, we focus on a continuous-time model (CM, see section *Models*), because its behavior allows a simpler interpretation. In a second step, we discuss the extent to which the corresponding discrete-time model (DM) exhibits a different behavior. Unless stated otherwise, we assume that hosts have a five times longer generation time than parasites ($n_{pg} = 5$, see *Methods*). We start by studying the high complementarity (HC) behavior for the simplest case in which only selection affects the gene frequencies and then successively consider the impact of mutation and recombination.

3.1. The high-complementarity region

The HC-behavior is strongly related to the existence and stability of the equilibria of the system. Because of its perfect symmetry, the system has the obvious “central” equilibrium

$$[h_c; p_c] = \left[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}; \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right]$$

which corresponds to an LD of zero. In the absence of mutation and recombination, there are in addition two so-called “Highly-Complementary Equilibria” (HCE) given by

$$[h_{HCE+}; p_{HCE+}] = \left[\frac{1}{2}, 0, 0, \frac{1}{2}; \frac{1}{2}, 0, 0, \frac{1}{2} \right]$$

and

$$[h_{HCE-}; p_{HCE-}] = \left[0, \frac{1}{2}, \frac{1}{2}, 0; 0, \frac{1}{2}, \frac{1}{2}, 0 \right],$$

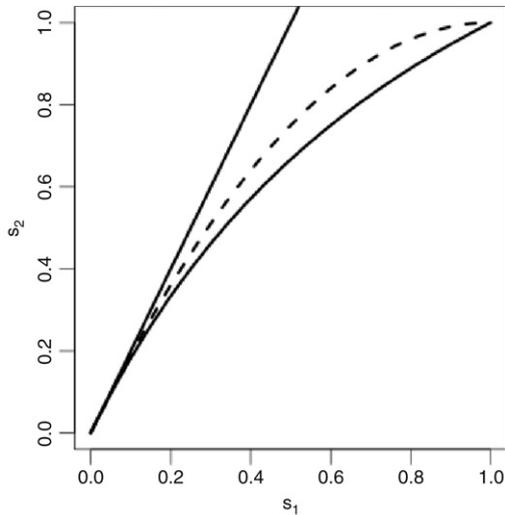


Fig. 1. Illustration of the HC region. The HC region lies between the two solid curves (upper curve: $s_2 = 2s_1$; lower curve: $s_2 = 2s_1/(1 + s_1)$). The dashed curve corresponds to the MMA, i.e. $s_2 = 2s_1 - s_1^2$ and always lies within the HC region..

which correspond to maximally positive and negative LD, respectively. There are further equilibria at the edges, e.g. $[h; p] = [1, 0, 0, 0; 0, 0, 1, 0]$ but as these are unstable and disappear as soon as mutation is introduced, we do not discuss them further.

It is instructive to consider the stability of the HCE for different interaction models (given by s_1 and s_2 ; see Models). Linearization of the continuous-time model (CM) reveals three pairs of identical eigenvalues, whose real parts read

$$\text{Re}(\lambda_{1,2}) = \frac{s_2 - 2s_1}{2 - s_2}; \quad \text{Re}(\lambda_{3,4}) = \frac{2s_1 - s_2(1 + s_1)}{(1 - s_1)(2 - s_2)};$$

$$\text{Re}(\lambda_{5,6}) = 0.$$

These expressions allow us to distinguish between two regions in the parameter-space (see Fig. 1): If $2s_1/(1 + s_1) < s_2 < 2s_1$, then the HCE are neutrally stable. By contrast, the HCE are unstable for all interaction models outside the region between the lines $s_2 = 2s_1$ and $s_2 = 2s_1/(1 + s_1)$. Thus, the HCE can persist only in the first region, which we will therefore refer to as the ‘‘HC region’’. Note that the MMA always lies in the HC region (Fig. 1). The central equilibrium, on the other hand, is neutrally stable outside the HCE region and unstable within the HCE region.

3.2. The impact of mutation

When mutation is introduced, only the central and, for some interaction models, the HC equilibria remain (see Appendix). Outside the HC region, the only equilibrium is the central equilibrium. If selection is strong in comparison to mutation (see Appendix), the system has three equilibria within the HC region: the central equilibrium and two HC equilibria, which lie in the vicinity of the HCE-points for no mutation ($m = 0$). The HC equilibria are given by

$$[h_{HCE+}; p_{HCE+}] = \left[\frac{1}{2} - \varepsilon_1, \varepsilon_1, \varepsilon_1, \frac{1}{2} - \varepsilon_1; \frac{1}{2} - \varepsilon_2, \varepsilon_2, \varepsilon_2, \frac{1}{2} - \varepsilon_2 \right]$$

and

$$[h_{HCE-}; p_{HCE-}] = \left[\varepsilon_1, \frac{1}{2} - \varepsilon_1, \frac{1}{2} - \varepsilon_1, \varepsilon_1; \varepsilon_2, \frac{1}{2} - \varepsilon_2, \frac{1}{2} - \varepsilon_2, \varepsilon_2 \right],$$

where ε_1 and ε_2 are of the order of the mutation rate m . Importantly, these equilibria are stable throughout the HC region. By contrast, the central equilibrium is stable outside but unstable

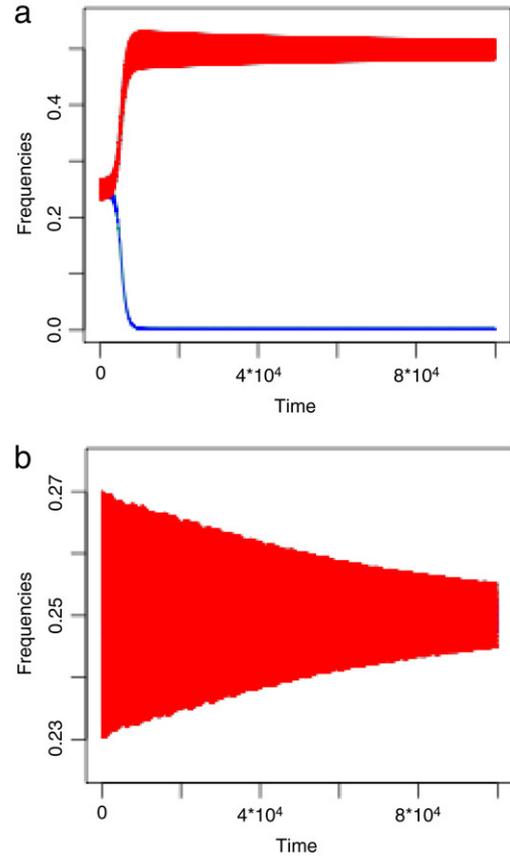


Fig. 2. HCE in the continuous-time model. Panel (a) shows the dynamics of the host-genotype frequencies (h_1 is cyan; h_2 is green; h_3 is blue; h_4 is red) for a typical realization of the MMA model (with $s = 0.1$); i.e. an interaction model within the HCE zone. Due to the overlap of the cyan (green) and the red (blue) lines only the red (blue) line is visible. For the initial conditions chosen the frequencies converge to the HCE point with positive LD. Panel (b) corresponds to a system outside the HC region ($s_1 = 0.1$; $s_2 = 0.21$). Due to the overlap of the lines, only the red line is visible. In this case the frequencies converge to the central equilibrium. Parameters: $m = 10^{-5}$; $r = 0$; $n_{PG} = 5$; initial conditions: Allele frequencies 0.51 and 0.53 at the two host loci and 0.5 at the two parasite loci. The host and the parasite population are initiated in linkage equilibrium.

within the HC region. Thus, the stability analysis suggests that the system approaches the central equilibrium outside the HC region and one of the HCE equilibria within the HC region. This is in agreement with numerical simulations (see for example Fig. 2).

The above conclusions hold only if selection is not too weak compared with mutation. If mutation is too strong, then HCE do not exist. We analyzed this phenomenon for the MMA model (which lies in the center of the HC region). There, the criterion for the existence of HCE reads approximately $s^2 > 16m$ (see Appendix); hence the squared selection coefficient must be considerably larger than the mutation rate. If selection is above the critical value then the LD at the HCE increases steeply as a function of s and converges to the maximal value $\frac{1}{4}$ (see Fig. 3). Altogether, we find that the system exhibits HCE-behavior if selection is not too weak and the interaction model lies within the HC region.

3.3. The impact of recombination

Recombination acts to reduce LD and should thus weaken the strength and range of the HC behavior, as found in one-species models (Franklin and Lewontin, 1970). For the sake of simplicity, we restrict our analysis in this part to the MMA. Numerical simulations suggest that the behavior is qualitatively

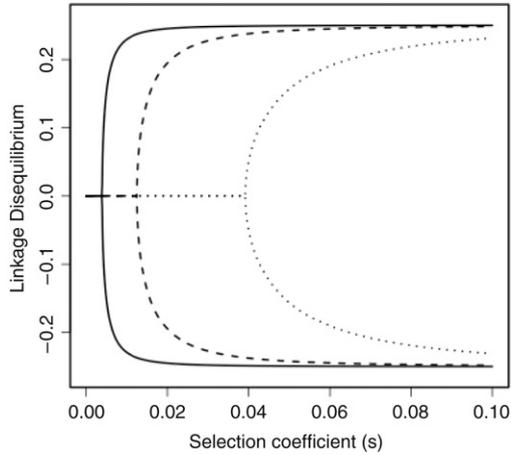


Fig. 3. Mutation destroys HC behavior for weak selection. LD at the HCE points of the MMA as a function of the strength of selection for $m = 10^{-6}$ (solid line), $m = 10^{-5}$ (dashed line), and $m = 10^{-4}$ (dotted line). If LD is 0 only the central equilibrium exists.

similar for other models within the HC-region. In the presence of recombination, the dynamical behavior becomes more complex and an equilibrium analysis does not provide a complete picture of the dynamics. Therefore, we start by describing the HC behavior numerically.

We determined the conditions that lead to HC behavior from simulation data by using the following criterion: We initiated the system at linkage equilibrium (with random allele frequencies) and iterated it for a period of 10^5 parasite generations. We then iterated it for a further 10^5 generations and counted how often the LD changes sign (Note that an initial LD develops, even for the MMA with both host and parasite population in linkage equilibrium, because the epistasis sensed by the host population is non-zero unless the allele frequencies of the parasite are exactly $\frac{1}{2}$). If the number of sign changes is zero and the magnitude of the average LD does not decrease within the next 10^5 generations, the system is said to exhibit HCE-like behavior.

Applying the numerical criterion to the MMA for the continuous-time model demonstrates the following points (Table 2). First, the sign of LD does not change during a simulation (i.e., 10^5 parasite generations) unless recombination is strong relative to selection in both hosts and parasites. Second, the average magnitude of LD decreases as a function of increasing recombination, but not as a function of time. Third, stronger selection facilitates the HC behavior, i.e. the equilibrium LD (for a given recombination rate) is stronger and the range of temporally stable (non-decreasing and constant in sign) LDs extends to higher recombination rates. Overall, the simulations exhibit two qualitatively different kinds of behavior: on one hand, there is the full HC behavior with LDs of constant sign and stable magnitude (Fig. 4a), occurring for strong selection and weak recombination. On the other hand, there is a transient HC behavior in which LD develop initially but wane over time (Fig. 4b), occurring for weak selection and strong recombination. Fig. 4 also illustrates that even for the full HC behavior, the magnitude of the LD does oscillate, although the sign and the average magnitude are constant. Thus, in a recombining population, the HC behavior consists of a limit cycle instead of an equilibrium point. In the transient HC regime, these oscillations result in sign-changes once the average LD wanes (Fig. 4b).

The impact of recombination on the HCE is qualitatively the same as on the HC oscillations, i.e. the strength of LD decreases with recombination and the HCE exist only if recombination is not too strong compared to selection (see Fig. 5). The LD at HCE plotted in Fig. 5 can be calculated analytically (see Appendix and

Table 2

HC behavior for the continuous-time model. The tables summarize the impact of recombination on host and parasite LD in the MMA. The systems were initiated in linkage equilibrium with random allele frequencies. After a burn-in period of 10^5 generations, we measured the number of sign-changes within the next 10^5 generations and the average LD over 500 generations starting at generation 2×10^5 (D_1) and starting at generation 3×10^5 (D_2). For each parameter-combination we ran 10 simulations (100 simulations for $s = 0.05$) and averaged the LDs. The first line in each cell gives the fraction of simulations for which no sign change was observed during 10^5 generations. The second line gives the average LD D_1 . The third line gives the decrease in the average LD, i.e. the average fraction of D_2/D_1 . Parameters: $m = 10^{-5}$; $n_{PC} = 5$.

s = 0.05; Host LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	100%	100%	100%	100%
	0.24	0.20	0.17	0.16
	1	1	1	1
0.001	100%	100%	100%	100%
	0.11	0.064	0.050	0.044
	1	1	0.98	0.98
0.01	100%	100%	99%	99%
	0.010	4.9×10^{-5}	3.5×10^{-5}	2.9×10^{-5}
	1	0.091	0.090	0.091
0.1	100%	99%	74%	58%
	3.5×10^{-5}	3.1×10^{-6}	2.1×10^{-6}	1.4×10^{-6}
	0.44	0.091	0.093	0.091
Parasite LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	100%	100%	100%	100%
	0.24	0.063	4.6×10^{-3}	4.4×10^{-4}
	1	1	1	0.95
0.001	100%	100%	100%	100%
	0.20	0.027	2.2×10^{-3}	2.1×10^{-4}
	1	0.99	1	0.83
0.01	100%	100%	100%	100%
	0.11	1.3×10^{-4}	9.4×10^{-6}	8.4×10^{-7}
	1	0.091	0.091	0.089
0.1	100%	100%	90%	60%
	0.039	5.2×10^{-5}	4.2×10^{-6}	3.0×10^{-7}
	0.46	0.090	0.082	0.099
s = 0.1; Host LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	100%	100%	100%	100%
	0.25 (0.248)	0.24	0.18	0.16
	1	1	1	1
0.001	100%	100%	100%	100%
	0.21	0.20	0.14	0.13
	1	1	1	1
0.01	100%	100%	100%	100%
	0.060	0.039	0.021	0.013
	1	1	1	1
0.1	100%	100%	100%	100%
	5.6×10^{-3}	3.3×10^{-5}	4.9×10^{-6}	3.4×10^{-6}
	1	0.092	0.090	0.091
Parasite LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	100%	100%	100%	100%
	0.25 (0.248)	0.17	0.024	0.0017
	1	1	1	1
0.001	100%	100%	100%	100%
	0.24	0.15	0.019	0.0015
	1	1	1	0.97
0.01	100%	100%	100%	100%
	0.18	0.073	0.0059	5.0×10^{-4}
	1	0.96	1	1
0.1	100%	100%	100%	100%

(continued on next page)

Table 2 (continued)

	0.13 1	4.7×10^{-4} 0.093	1.1×10^{-5} 0.090	7.4×10^{-7} 0.089
<i>s</i> = 0.5; Host LD				
<i>r_h</i>	<i>r_p</i>			
	0	0.001	0.01	0.1
0	100% 0.25 (0.24996) 1	100% 0.25 (0.2496) 1	100% 0.25 (0.247) 1	100% 0.22 1
0.001	100% 0.25 (0.249) 1	100% 0.25 (0.249) 1	100% 0.25 (0.245) 1	100% 0.22 1
0.01	100% 0.24 1	100% 0.24 1	100% 0.24 1	100% 0.21 1
0.1	100% 0.16 1	100% 0.16 1	100% 0.16 1	100% 0.12 1
Parasite LD				
<i>r_h</i>	<i>r_p</i>			
	0	0.001	0.01	0.1
0	100% 0.25 (0.24994) 1	100% 0.25 (0.247) 1	100% 0.22 1	100% 0.098 1
0.001	100% 0.25 (0.2496) 1	100% 0.25 (0.246) 1	100% 0.22 1	100% 0.098 1
0.01	100% 0.25 (0.246) 1	100% 0.24 1	100% 0.22 1	100% 0.094 1
0.1	100% 0.22 1	100% 0.21 1	100% 0.19 1	100% 0.064 1

supplementary *Mathematica* file); for the MMA with mutation and recombination rates equal in host and parasite, the LD is given by

$$D = \frac{\sqrt{(s^2 - (8(m - m^2)(1 - r) + 2r)(2 - 2s + s^2))}}{4s},$$

where *s* is the strength of selection, *m* the mutation rate, and *r* the recombination rate (see section Models). The HCE becomes, however, unstable when recombination differs significantly from zero and accordingly we observe stable limit cycles around the unstable HCE. Interestingly, the parameter range in which persistent HC oscillations occur is broader than the range in which

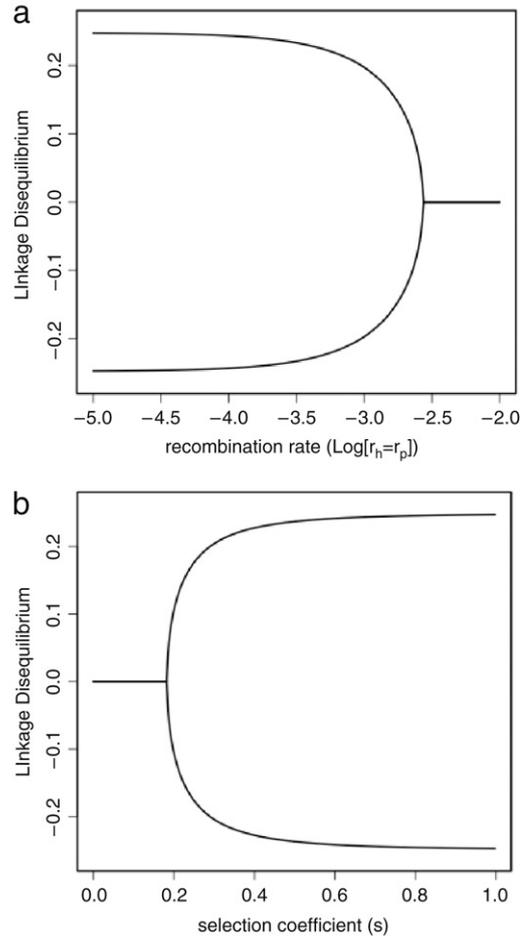


Fig. 5. LD at the HCE with recombination. The plots show the LD at the HCE points of the MMA model as a function of recombination rate for constant selection *s* = 0.1 (panel (a)) and as a function of selection for constant recombination rate *r* = 0.01 (panel (b)). If LD is 0 only the central equilibrium exists. Parameters: *m* = 10^{-5} .

HCE exist (compare Tables 2 and 3 with Fig. 5). Thus, the HC oscillations seem to be more robust against recombination than the equilibrium points.

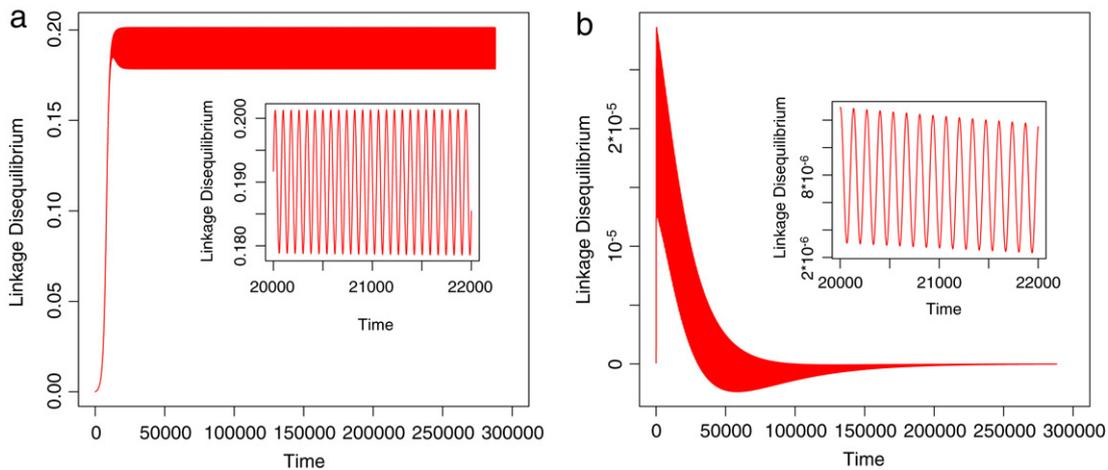


Fig. 4. HC behavior for the MMA with recombination. The figure shows the time course of the LD (generations 0– 3×10^5 ; inset: 20 000–25 000) for the MMA (*s* = 0.1) with weak recombination (*r* = 0.001) in panel (a) and strong recombination (*r* = 0.1) in panel (b). The initial conditions were chosen such that positive LD develops, but the behavior of the system is qualitatively the same with random initial conditions. Parameters: *m* = 10^{-5} ; initial conditions: Allele frequencies 0.51 and 0.53 at the two host loci and 0.5 at the two parasite loci. The host and the parasite population are initiated in linkage equilibrium.

Table 3

HC behavior for the discrete-time model. Same as Table 2, but for the discrete-time model. Only the average magnitude of the LD is plotted because we found no sign-changes and no decrease in LD in all cases. Simulations were also performed for $s = 0.01$; but no HC behavior was observed throughout the whole range of recombination rates (in accordance with the case without recombination; see dashed curve in Fig. 3). Parameters: $m = 10^{-5}$; $n_{PG} = 5$.

$s = 0.05$; Host LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.18	0.15	0.14	0.14
0.001	0.068	0.056	0.052	0.052
0.01	0.0057	0.0054	0.0053	0.0053
0.1	5.5×10^{-4}	5.5×10^{-4}	5.5×10^{-4}	5.5×10^{-4}
Parasite LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.055	0.015	0.0016	8.9×10^{-5}
0.001	0.028	0.0062	6.8×10^{-4}	4.6×10^{-5}
0.01	0.0045	8.2×10^{-4}	9.4×10^{-5}	7.0×10^{-6}
0.1	5.7×10^{-4}	9.5×10^{-5}	1.1×10^{-5}	1.0×10^{-6}
$s = 0.1$; Host LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.18	0.17	0.13	0.13
0.001	0.15	0.13	0.11	0.11
0.01	0.031	0.027	0.025	0.025
0.1	0.0026	0.0025	0.0025	0.0025
Parasite LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.061	0.044	0.0059	2.4×10^{-4}
0.001	0.055	0.037	0.0047	2.0×10^{-4}
0.01	0.019	0.0080	0.0011	6.4×10^{-5}
0.1	0.0024	9.1×10^{-4}	1.4×10^{-4}	8.0×10^{-6}
$s = 0.5$; Host LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.14	0.14	0.14	0.12
0.001	0.14	0.14	0.14	0.11
0.01	0.14	0.14	0.13	0.11
0.1	0.096	0.096	0.094	0.074
Parasite LD				
r_h	r_p			
	0	0.001	0.01	0.1
0	0.061	0.060	0.056	0.028
0.001	0.061	0.060	0.056	0.027
0.01	0.060	0.059	0.055	0.027
0.1	0.050	0.050	0.045	0.019

3.4. Differences in discrete-time

In the discrete-time model, we observe a qualitatively similar behavior: if the interaction model is close enough to the MMA (see Fig. 6), then the system converges to a state with strong LD of constant sign. There are, however, two notable differences concerning (i) the type of dynamics in an asexual population and (ii) the impact of recombination.

The main difference in asexual populations is that the HCE become unstable. Accordingly, the system does in general not approach these points, but converges to limit cycles similar to those observed in the continuous-time model for recombining populations (see Fig. 7): for $m = 0$, the system converges towards a limit cycle lying in either of the two two-dimensional subspaces

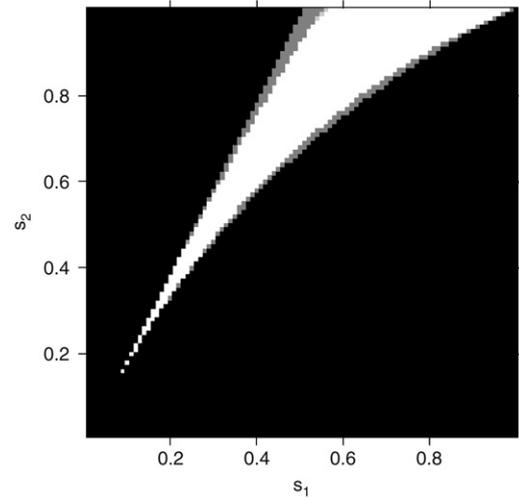


Fig. 6. Range of interaction models with HC behavior in discrete time. Every point in the figure corresponds to an interaction model (specified by s_1 and s_2). The parameters s_1 and s_2 range from 0 to 1 with 0.01 gradation; thus 10^4 (100×100) interaction models are depicted. The coloring of the point indicates whether HC behavior occurs or not and whether the point lies in the HC Region: No HC behavior occurs and point lies outside the HC Region (black); No HC behavior occurs but point lies within HC Region (gray); HC behavior occurs and point lies within HC Region (white). Every point represents an average over 10 simulations; HC behavior occurs either for all or for none of the initial values. Parameters: $m = 10^{-5}$; $r = 0$; $n_{PG} = 5$.

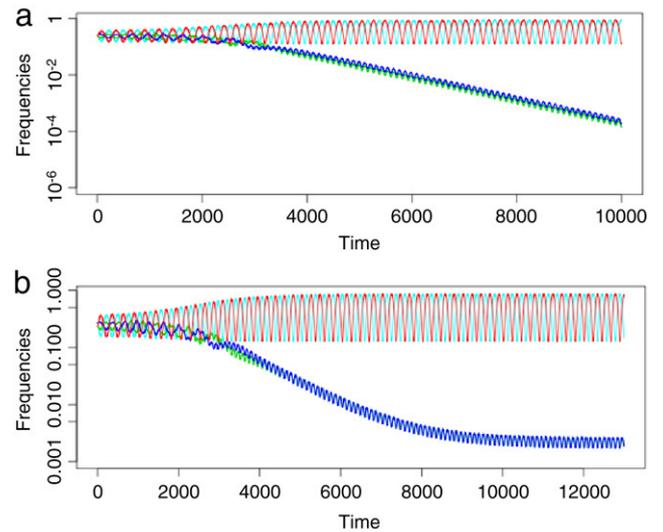


Fig. 7. HC behavior in the discrete-time model. Dynamics of the host-genotype frequencies (h_1 is cyan; h_2 is green; h_3 is blue; h_4 is red) for a typical realization of the MMA model ($s = 0.1$) without mutation (panel (a)) and with mutation rate $m = 10^{-5}$ (panel (b)). Parameters: $r = 0$; $n_{PG} = 5$; initial conditions: Allele frequencies 0.51 and 0.53 at the two host loci and 0.78 and 0.64 at the two parasite loci. The host and the parasite populations are initiated in linkage equilibrium.

S_+ or S_- , where

$$S_+ : [h_1, h_2, h_3, h_4; p_1, p_2, p_3, p_4] = (x, 0, 0, 1 - x, y, 0, 0, 1 - y)$$

and

$$S_- : [h_1, h_2, h_3, h_4; p_1, p_2, p_3, p_4] = (0, x, 1 - x, 0, 0, y, 1 - y, 0).$$

The limit cycles in S_+ and S_- correspond to strong positive and negative LD, respectively, and thus describe HC-like behavior. For $1 \gg m > 0$, the frequencies stay in the vicinity of S_+ or S_- , but the underrepresented combinations are maintained by mutation at a low steady-state level. Interestingly, the HC behavior does not occur in the discrete-time model if host and parasite have the same generation time ($n_{PG} = 1$, see Models), i.e. the system no

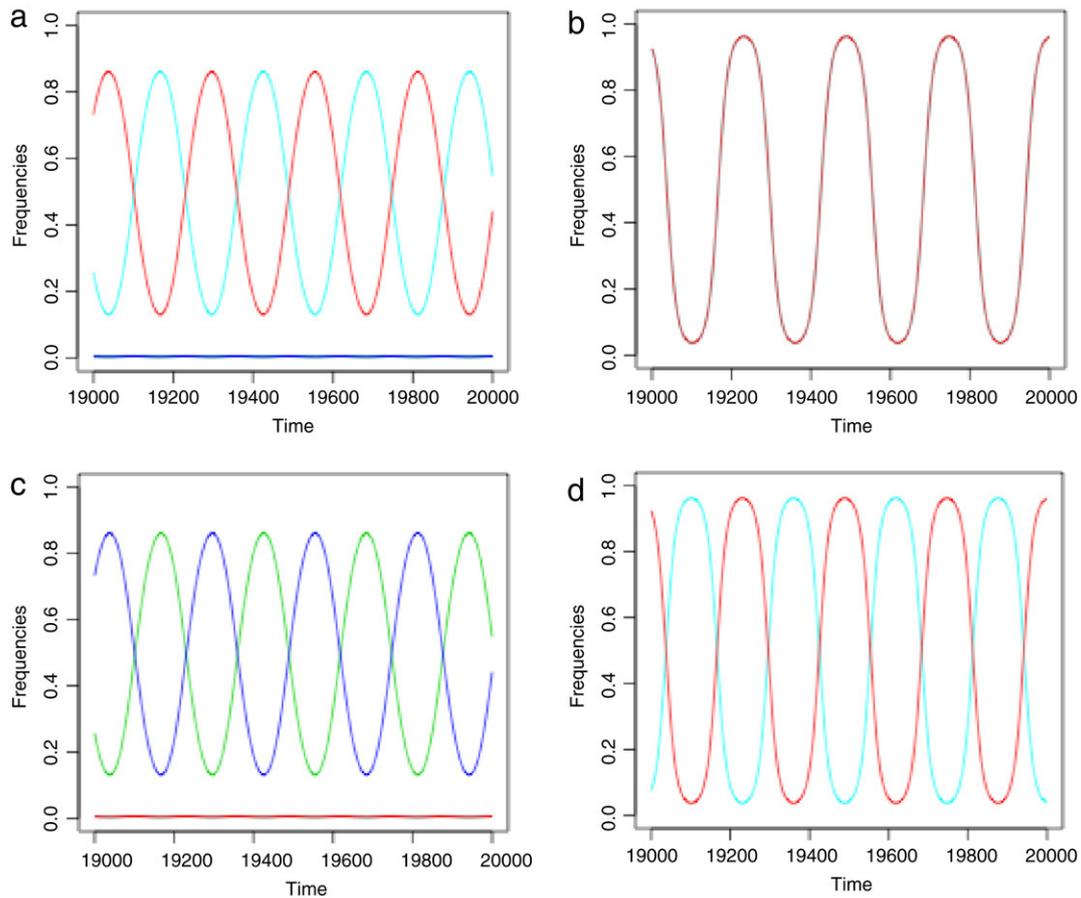


Fig. 8. HC behavior for two independent parasites in discrete time. Time-course of host (panels (a) and (c)) and parasite (panels (b) and (d)) frequencies for the model with two independent parasites with $s=0.1$ and no recombination. The color code is as follows: (h_1 is cyan; h_2 is green; h_3 is blue; h_4 is red) for panels (a) and (c); and (p_1 is cyan; p_2 is red) for panels (b) and (d). In panels (a) and (b) the initial conditions were chosen such that positive LD develops; in panels (c) and (d) such that negative LD develops. Parameters: $m = 10^{-5}$; $n_{PC} = 5$; initial conditions panel A: Allele frequencies 0.55 and 0.55 at the two host loci and 0.5 at the two parasite loci. The host and the parasite population are initiated in linkage equilibrium.; initial conditions panel B: Allele frequencies 0.55 and 0.45 at the two host loci and 0.5 at the two parasite loci. The host and the parasite population are initiated in linkage equilibrium.

longer converges towards a steady state with a non-vanishing (and strong) LD of constant sign. Smaller generation times of parasites are, however, a hallmark of host–parasite interactions and this phenomenon is thus presumably of limited biological relevance. In addition, we observed the HC behavior for $n_{PC} = 1$, when the selection on the parasite was much stronger than on the host (results not shown). Note that the way we parameterize the fitness interaction (see Models) implies that selection on hosts and parasites is of equal strength for $n_{pg} = 1$; but as n_{pg} increases the selection on the parasite relative to that of the host increases as well.

The main difference concerning the impact of recombination is that, in discrete time, persistent LD of constant sign occur throughout the whole range of recombination rates (see Table 3). The only case in which we did not find HC behavior was if selection was too weak compared with mutation (see table legend), but in that case, HC behavior does not occur even if $r = 0$. As in the continuous time model, the strength of the LD at steady state strongly depends on the recombination rate, with high recombination implying low LD. Altogether, HC behavior seems to occur in the discrete model, irrespective of the rate of recombination. We found this behavior for all parameters explored in Table 3, i.e. for selection coefficients 0.05, 0.1, and 0.5 and for recombination rates ranging from 0 to 0.1. When selection was very weak compared with mutation (see Fig. 3), the system did not exhibit HC behavior (see legend of Table 3), both with and without

recombination. The resulting (average) LD are, however, very weak if recombination is too strong compared to selection.

3.5. Two independent parasites

The above analysis of the two-species model shows that the HC behavior is very robust against recombination in the parasite, especially in the discrete-time model. Intuitively, the MMA model with high recombination in the parasite, i.e. with two almost independent parasite loci, is similar to considering two independent parasites that affect the host's fitness multiplicatively (i.e. the harm done by one parasite does not depend on the presence or absence of the other parasite). Thus, HC behavior could also occur at host loci involved in the interaction with different parasites. To test this intuition, we explicitly modeled the impact of two independent parasites; each interacting with one host locus (see Models).

Simulations of the discrete-time model with two independent parasites show that this type of system, indeed, produces HC behavior (see Fig. 8). Depending on the initial conditions, the system approaches a limit-cycle either with strong positive or with strong negative LD. Interestingly, the allele frequencies of the two parasites are in phase in the first case and out of phase (shifted by 90°) in the second case. In order to investigate the dependence of the HC phenomenon on recombination and selection, we used the numerical criterion described in the previous sections. The

Table 4

HC behavior for the discrete-time model with two independent parasites. Same as Table 3, but for the one-host vs. two-parasites model. (Here parasite-recombination is not a parameter, because parasites have only one locus. Thus, we made only one table with host-recombination (columns) and selection (rows)). Parameters: $m = 10^{-5}$; $n_{PG} = 5$.

r	s			
	0.01	0.05	0.1	0.5
0	NO HC	0.17	0.18	0.18
0.001	NO HC	0.048	0.12	0.18
0.01	NO HC	0.0052	0.024	0.16
0.1	NO HC	5.5×10^{-4}	0.0025	0.089
0.5	NO HC	1.1×10^{-4}	5.5×10^{-4}	0.023

results show that HC behavior occurs for a broad range of selection strengths and recombination rates (Table 4). The impact of these parameters is, qualitatively, the same as in the single-parasite model: (i) if selection is too weak compared with mutation, no HC behavior occurs, (ii) recombination lowers the LD in the HC state, and (iii) stronger selection makes the LD more robust against recombination.

In the continuous-time model, the HC state occurs only transiently. This means that if the system is initiated in linkage equilibrium with random allele frequencies, a strong LD of constant sign first builds up, but then decreases in strength and wanes eventually. Fig. 9 illustrates this behavior and exemplifies the impact of mutation and recombination on the strength and decay of the HC state. The example shows that although strong LD of constant sign occur only as a transient phenomenon in continuous-time, the decay rate is very small such that the LD can, nevertheless, persist for a considerable number of generations. The transient nature of the LD in the continuous-time two-parasite model, is consistent with the continuous-time one-parasite model, in which the HC behavior wanes over time, if recombination is too strong (see Table 2 and Fig. 4).

4. Discussion

The impact of host–parasite interactions on the evolution of recombination depends crucially on the type of linkage disequilibria generated by these interactions. In this article, we have shown that host–parasite interactions typically generate strong LD of constant sign, if the number of matched alleles between host and parasite determines the fitness in an approximately multiplicative way. To our knowledge, the present article is the first analysis of such a high-complementarity behavior in a coevolutionary model. Consistent with the reduction principle, such LD generate strong selection against sex and recombination (Altenberg and Feldman, 1987), as we observed previously in a model that explicitly considered modifiers of recombination (Figures 5–8 in Kouyos et al. (2007)), and as we show in the Appendix of this article. The high-complementarity behavior described in this paper is also consistent with the results of Salathé et al. (2008): There, a different type of coupling between host and parasite fitness was considered (in which the hosts fitness equals one minus the parasites fitness), and selection against recombination in the vicinity of the additive matching alleles model (AMA) was observed (Figure 3 in Salathé et al. (2008)). The AMA approaches the MMA asymptotically as selection decreases, and, indeed, we found several instances of high-complementarity behavior in the zone around the AMA (results not shown) suggesting that the observed selection against recombination in the vicinity of the AMA in Salathé et al. (2008) was due to high-complementarity phenomena.

Intuitively, one can understand how the HC behavior results in strong selection against genetic shuffling in the following way. Consider the simple case that the system is in the HCE given by host-frequencies $h = (\frac{1}{2} - \varepsilon_1, \varepsilon_1, \varepsilon_1, \frac{1}{2} - \varepsilon)$ and parasite frequencies $p = (\frac{1}{2} - \varepsilon_2, \varepsilon_2, \varepsilon_2, \frac{1}{2} - \varepsilon_2)$. Recombination in the

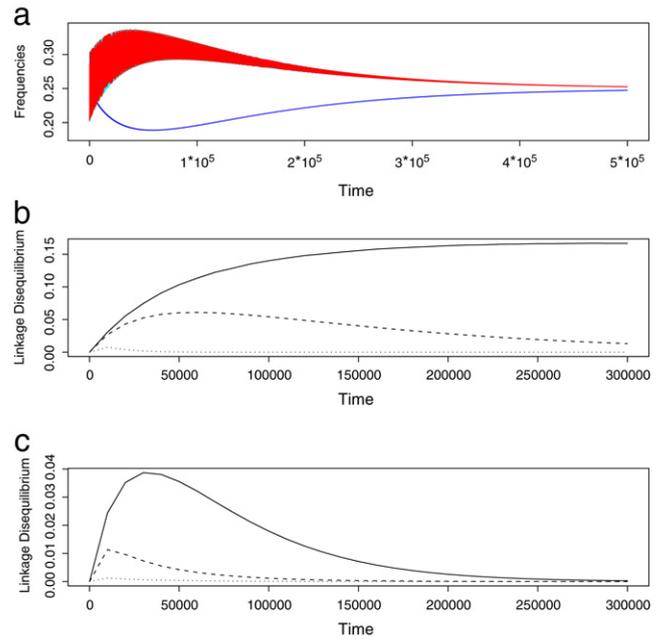


Fig. 9. HC behavior for two independent parasites in continuous time. Panel (a): Time course of the host frequencies (same color code as Fig. 8) in the continuous model with $s = 0.1$, $r = 0$, and $m = 10^{-5}$. Due to the overlap of the cyan (green) and the red (blue) lines only the red (blue) line is visible. Panel (b): Time course of the LD for mutation rates 10^{-6} (solid line), 10^{-5} (long dashed line), and 10^{-4} (dashed line); $s = 0.1$, $r = 0$. Panel (c): Time course of the LD for recombination rates 10^{-4} (solid line), 10^{-3} (long dashed line), and 10^{-2} (dashed line); $s = 0.1$, $m = 10^{-5}$. The initial conditions were chosen such that positive LD develops; but the behavior of the system is qualitatively the same with random initial conditions. Parameters: $n_{PG} = 5$; initial conditions for all panels: Allele frequencies 0.55 and 0.55 at the two host loci and 0.5 at the two parasite loci. The host and the parasite population are initiated in linkage equilibrium.

host produces the underrepresented allele combinations h_2 and h_3 at a high rate, because the LD is almost maximal. In the HC region, however, these genotypes have a lower fitness than the overrepresented h_1 and h_4 . Consider, for example, the MMA model: as ε is small, h_1 and h_4 are matched at both loci by approximately one half of the parasite population and at no locus by the other half, whereas h_2 and h_3 are matched at one locus by the whole parasite population. Thus h_1 and h_4 have fitnesses $w_1 = w_4 = \frac{1}{2} + \frac{1}{2}(1-s)^2 = 1 - s + \frac{1}{2}s^2$, and h_2 and h_3 have the lower fitness $w_2 = w_3 = 1 - s$ (see section Models). This implies that recombination always produces less fit genotypes and is thus detrimental. The effect extends beyond this particular example: the underrepresented combination of genotypes is always the one that is less fit, and since all recombination does is increase the underrepresented class at the expense of the overrepresented class, recombination is always detrimental. In principle, this short-term cost could be compensated by a long-term benefit resulting from an improved response to selection. This is, however, not the case for the HC behavior, because the underrepresented genotypes (h_2 and h_3 in the above example) always remain less fit than the overrepresented, and producing them therefore does not increase the response to selection.

We observe HC behavior as long as the interaction model is sufficiently close to the multiplicative model. Furthermore, we find that the strength of the LD in the HC state is largely independent of mutation as long as selection is above a low threshold (of the order of $\sqrt{\mu}$). Recombination, on the contrary, has a very strong impact on the existence and strength of HC states. In particular, the strength of LD in the HC state decreases with increasing recombination. Accordingly, selection against recombination should be weaker in a species that already

recombines at a high rate. Thus, while weak recombination may be invaded by asexuality, high recombination rates cannot be outcompeted by lower recombination rates.

The comparison between the continuous-time and the discrete-time model shows that most aspects of the HC behavior are structurally robust. In both cases, the system converges to strong LD of constant sign if the type of selection is approximately multiplicative. Furthermore, in both cases, recombination decreases the magnitude of the LD in the HC state, and higher selection makes the system more robust against this decreasing impact of recombination. There are two important differences between the continuous- and discrete-time model. First, in the continuous-time model, the HC behavior occurs only transiently if recombination is high, whereas, in the discrete-time model persistent HC behavior occurs independently of recombination. Second, the continuous-time model exhibits HC behavior for all values of n_{PC} , whereas the discrete-time model exhibits HC behavior only for $n_{PC} > 1$. This fact explains why Otto and Nuismer (2004) did not find this behavior for the MMA.

How plausible is a considerable effect of the HC behavior on the evolution of genetic shuffling? This depends mainly on the specific interactions found in nature. If a parasite attacks sufficiently independent targets, then the interactions may be close to the MMA and the population is predicted to exhibit HC-like behavior. This finding demonstrates that the impact of host–parasite interactions on the evolution of recombination depends crucially on the specific interaction type. In particular it underlines the necessity of epistatic interactions by showing that the “null model” with multiplicative selection does not only fail to favor the evolutionary maintenance of genetic shuffling, but in fact strongly selects against recombination. Overall, the present article highlights the need for more detailed knowledge on how the genotypes of host and parasite determine the outcome of infection. A meta-analysis of plant (Kover and Caicedo, 2001) data indeed found widespread epistasis between resistance loci, but the generality of this finding is unclear. Our analysis furthermore shows that HC behavior may also occur in the plausible case that a host is attacked by two or more independently acting parasites. This indicates that the absence of interactions between parasites could be a further obstacle for the RQH and calls for more, theoretical and experimental, investigations on the neglected issue of multi-parasite systems in the context of the RQH.

Acknowledgments

We thank Olin Silander for valuable comments. RDK, SB, and MS gratefully acknowledge support from the Swiss National Science Foundation.

Appendix A. Equilibrium points

First, notice that the discrete-time model (with any generation time) and the continuous-time model have the same equilibrium points. The calculations for obtaining the equilibrium points are performed in the supplementary *Mathematica* file “*equilibria.nb*”. Here, we merely comment on the different steps and some of the results.

A.1. Symmetry of the equilibrium

The calculations are considerably simplified by assuming that the equilibria have the following form

$$[h; p] = [h_1, h_2, h_3, h_4; p_1, p_2, p_3, p_4] = [u, x, x, u; v, y, y, v],$$

i.e. $h_1 = h_4, h_2 = h_3, p_1 = p_4, p_2 = p_3$ (this means that allele frequencies are $\frac{1}{2}$ at all involved loci). In addition, both host and

parasite frequencies must sum up to one and thus $u = \frac{1}{2} - x, v = \frac{1}{2} - y$, and

$$[h; p] = \left[\frac{1}{2} - x, x, x, \frac{1}{2} - x; \frac{1}{2} - y, y, y, \frac{1}{2} - y \right].$$

Thus instead of 6 (=8 – 2) equations with 6 variables, we only have to solve two equations with two variables. In the absence of recombination, one can show in the following way that the above symmetry must hold.

First, we show that at equilibrium $w_1^H > w_4^H$ implies $h_1 > h_4$ (and similarly for the parasite). To this end, consider the general single species two-locus/two-allele (frequencies: (f_1, f_2, f_3, f_4)) model with mutation and selection (given by fitness values (w_1, w_2, w_3, w_4)). Assume that $w_1 > w_4$ but $f_1 \leq f_4$. Let f'_i and f''_i be the frequencies after the successive action of mutation (f'_i) and selection (f''_i). We then have

$$f_1/f_4 \leq f'_1/f'_4 < (w_1/w_4)f'_1/f'_4 = f''_1/f''_4.$$

The first inequality is shown to hold in the *mathematica* file (section “auxiliary calculation”). The second inequality holds because $w_1 > w_4$ and $f'_1 > 0$ (the frequency cannot vanish after mutation). However at equilibrium we also have $f''_i = f_i$, which contradicts the above inequality. Therefore $f_4 < f_1$.

Given the above relation, the symmetry follows immediately: Assume that $h_1 > h_4$, then we obtain (for $s_2 > 0$)

$$\begin{aligned} w_1^P &= h_1 + (1 - s_2)/(1 - s_1)(h_2 + h_3) + (1 - s_2)h_4 \\ &> h_4 + (1 - s_1)(h_2 + h_3) + (1 - s_2)/(1 - s_1)h_1 = w_4^P \end{aligned}$$

thus parasite-genotype 1 is fitter than parasite genotype 4. From the paragraph above, it follows that at equilibrium $p_1 > p_4$. This implies

$$\begin{aligned} w_4^H &= p_1 + (1 - s_1)(p_2 + p_3) + (1 - s_2)p_4 \\ &> p_4 + (1 - s_1)/(1 - s_2)(p_2 + p_3) + (1 - s_2)p_1 = w_1^H, \end{aligned}$$

and thus $h_4 > h_1$, in contradiction to the assumption. Similarly, one can show that $h_1 < h_4$ is not consistent with an equilibrium and therefore $h_1 = h_4$. The other equalities follow in an analogous way.

In the presence of recombination, the above reasoning does not hold anymore because $w_1^P > w_4^P$ does not imply $p_1 > p_4$ (and similarly for the other pairs). We may still calculate the equilibria with this symmetry. We cannot, however infer that these are the only equilibria. Interestingly, numerical calculation of the equilibria (from the full set of equations) only yielded points that obeyed the symmetry.

A.2. Calculation of the equilibria

In the supplementary *Mathematica* file, the calculation is performed separately for the two cases $r = 0$ and $r > 0$ (sections “No Recombination” and “With Recombination” respectively). Here we will only comment on the case $r = 0$ (described in subsection “compute equilibria”), but the calculations for $r > 0$ are practically identical.

With the symmetry assumption discussed in the previous section, the full set of equations

$$\begin{aligned} RMS_h([h(t); p(t)]) - h(t) &= 0 \\ RMS_p([h(t); p(t)]) - p(t) &= 0, \end{aligned}$$

can be reduced to two equations with two variables (x and y). Moreover these equations can be written in the form

$$\frac{n_1(x, y)}{d_1(x, y)} = 0 \quad \text{and} \quad \frac{n_2(x, y)}{d_2(x, y)} = 0,$$

where $n_i(x, y)$ and $d_i(x, y)$ are polynomials. The denominators are equal to the mean-fitnesses, which cannot vanish. Therefore, we only have to solve the polynomial equations

$$n_1(x, y) = 0, \quad n_2(x, y) = 0.$$

Importantly, $n_1(x, y)$ is linear in y . Thus we can solve $n_1(x, y) = 0$ for y and insert the solution $y(x)$ (a rational function in x) into $n_2(x, y) = 0$, which yields an equation of the form

$$\frac{n_{red}(x)}{d_{red}(x)} = 0,$$

where $n_{red}(x)$ and $d_{red}(x)$ are polynomials. A little calculation shows that the n_{red} and d_{red} have common roots only if $s_2 = 2s_1/(1 + s_1)$. The same calculation shows that for this interaction model (i.e. $s_2 = 2s_1/(1 + s_1)$), only the central equilibrium (all genotype frequencies are $\frac{1}{4}$) solves the full equations. For all other interaction models, solving the full equations is equivalent to solving the polynomial equation $n_{red}(x) = 0$.

$n_{red}(x)$ is a fifth-order polynomial. Moreover, it can be written in the following form (see *Mathematica* file)

$$n_{red}[x] = (x - 1/4) \cdot q[-x + 2x^2],$$

where $q[.]$ is a quadratic function (we used here “[.]” to express “function of” in order to discern it from multiplication). The first term of n_{red} indicates that the central equilibrium solves (as expected) the equations. We can find all remaining equilibrium points of the full recursions by successively solving two quadratic equations. In the general case, however, the resulting solutions are very complicated. Therefore, we will interpret them by means of approximations, special cases and graphical analysis.

A.3. Approximation for small mutation rates

The polynomial $n_{red}(x)$ has five roots, that correspond to five potential solutions of the full equations. A root x_0 is a biologically relevant equilibrium point for the host parasite interaction if x_0 is real and if x_0 and $y(x_0)$ are between 0 and $\frac{1}{2}$. An obvious root is given by the central equilibrium. Here we will consider a first-order approximation in $m = m_h = m_p$ of the four remaining roots. These approximations contain the square roots $\sqrt{(2s_1 - s_2)^2}$ and $\sqrt{(2s_1/(1 + s_1) - s_2)^2}$. Thus the form of the roots depends on whether $s_2 > 2s_1$ and/or $s_2 > 2s_1/(1 + s_1)$ (notice that these inequalities correspond to the borders of the HC region). If $s_2 > 2s_1$, then the four solutions read:

$$\begin{aligned} (x_1, y_1) &= \left(\frac{2(1 - s_1)}{s_2 - 2s_1} m, \frac{1}{2} + \frac{2(1 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_2, y_2) &= \left(\frac{1}{2} - \frac{2(1 - s_2)}{s_2 - 2s_1} m, -\frac{2(1 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_3, y_3) &= \left(-\frac{2 - s_2}{s_2 - 2s_1} m, \frac{(1 - s_1)(2 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_4, y_4) &= \left(\frac{1}{2} + \frac{2 - s_2}{s_2 - 2s_1} m, \frac{1}{2} - \frac{(1 - s_1)(2 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2]. \end{aligned}$$

None of these is a biologically valid equilibrium points (either x or y lie outside the interval between 0 and $\frac{1}{2}$). For $2s_1/(1 + s_1) < s_2 < 2s_1$, the equilibria read:

$$\begin{aligned} (x_1, y_1) &= \left(\frac{2 - s_2}{2s_1 - s_2} m, \frac{(1 - s_1)(2 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_2, y_2) &= \left(\frac{1}{2} - \frac{2 - s_2}{2s_1 - s_2} m, \frac{1}{2} - \frac{(1 - s_1)(2 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_3, y_3) &= \left(-\frac{2(1 - s_2)}{2s_1 - s_2} m, \frac{1}{2} + \frac{2(1 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2] \\ (x_4, y_4) &= \left(\frac{1}{2} + \frac{2(1 - s_2)}{2s_1 - s_2} m, -\frac{2(1 - s_2)}{s_2(1 + s_1) - 2s_1} m \right) + O[m^2]. \end{aligned}$$

Here, the first and the second equilibrium point lie in the biologically relevant range; the third and the second lie outside. Finally, for $s_2 < 2s_1/(1 + s_1)$, the expressions for the equilibrium points are identical to the ones for $s_2 > 2s_1$ and thus all roots lie outside the biological range. In summary, the approximation for small m shows that for interaction models outside the HC region ($s_2 > 2s_1$ or $s_2 < 2s_1/(1 + s_1)$) only the central equilibrium exists, whereas models within the HC region ($2s_1/(1 + s_1) < s_2 < 2s_1$) have three equilibria: the central equilibrium plus two HC equilibria of the form

$$[h_{HCE+}; p_{HCE+}] = \left[\frac{1}{2} - \varepsilon_1, \varepsilon_1, \varepsilon_1, \frac{1}{2} - \varepsilon_1; \frac{1}{2} - \varepsilon_2, \varepsilon_2, \varepsilon_2, \frac{1}{2} - \varepsilon_2 \right]$$

$$\text{with } \varepsilon_1 = O[m]$$

and

$$[h_{HCE-}; p_{HCE-}] = \left[\varepsilon_1, \frac{1}{2} - \varepsilon_1, \frac{1}{2} - \varepsilon_1, \varepsilon_1; \varepsilon_2, \frac{1}{2} - \varepsilon_2, \frac{1}{2} - \varepsilon_2, \varepsilon_2 \right]$$

$$\text{with } \varepsilon_2 = O[m],$$

corresponding to the values for (x_1, y_1) and (x_2, y_2) for this range of interaction models.

A.4. Special Case: MMA

For the MMA, the expressions for the five roots of n_{red} are considerably simpler than in the general case. The roots x_3 and x_4 correspond to equilibrium points outside the biologically relevant range (in accordance with the first-order approximation in the previous section). The expression for the remaining two solutions read:

$$\begin{aligned} (x_1, y_1) &= \left(\frac{1}{4} \left(1 - \frac{SQRT}{s} \right), \frac{1}{4} \left(1 - \frac{SQRT}{s} \right) \right) \\ (x_2, y_2) &= \left(\frac{1}{4} \left(1 + \frac{SQRT}{s} \right), \frac{1}{4} \left(1 + \frac{SQRT}{s} \right) \right), \end{aligned}$$

where $SQRT = \sqrt{s^2 - 16m(1 - m)(1 - s + s^2/2)}$. If $s^2 \gg m$ then the SQRT is approximately s (it is slightly smaller) and the two roots x_1 and x_2 are close to 0 and $\frac{1}{2}$ respectively. Thus these equilibria correspond to almost maximal LD. If, however, selection is too small compared to mutation, then the second term in SQRT prevails and the roots have a non-vanishing imaginary part. In this case, the system has only the central equilibrium. The transition between the two regimes occurs at $s^2 - 16m(1 - m)(1 - s + s^2/2) = 0$ or approximately at $s^2 - 16m = 0$. Thus the system has three equilibria (the central equilibrium and two HCE) if (roughly) $s^2 > 16m$ and only the central equilibrium if $s^2 < 16m$. The interplay between mutation and selection is represented graphically in Fig. 3.

Appendix B. Selection on a modifier of recombination

We studied the selection for or against recombination in the standard way by adding a selectively neutral modifier of recombination at the edge of the genome (i.e. the gene order is MAB, where M is the modifier and A and B are the interaction loci). Here we assume that, parasites reproduce clonally while hosts reproduce by recombining their genotypes. Recombination occurs at a rate that depends on the alleles present at the modifier locus of two recombining genotypes. The recombination rate between the interaction loci is $r_{baseline}$ if both recombining genomes carry allele m . It is increased by a small amount Δr or $2\Delta r$ if one or both genomes carry allele M . The recombination rate between the modifier and the interaction loci is $r_{modifier}$ (only one recombination rate has to be defined in this case, because a recombination event between the modifier locus and the interaction loci only has an

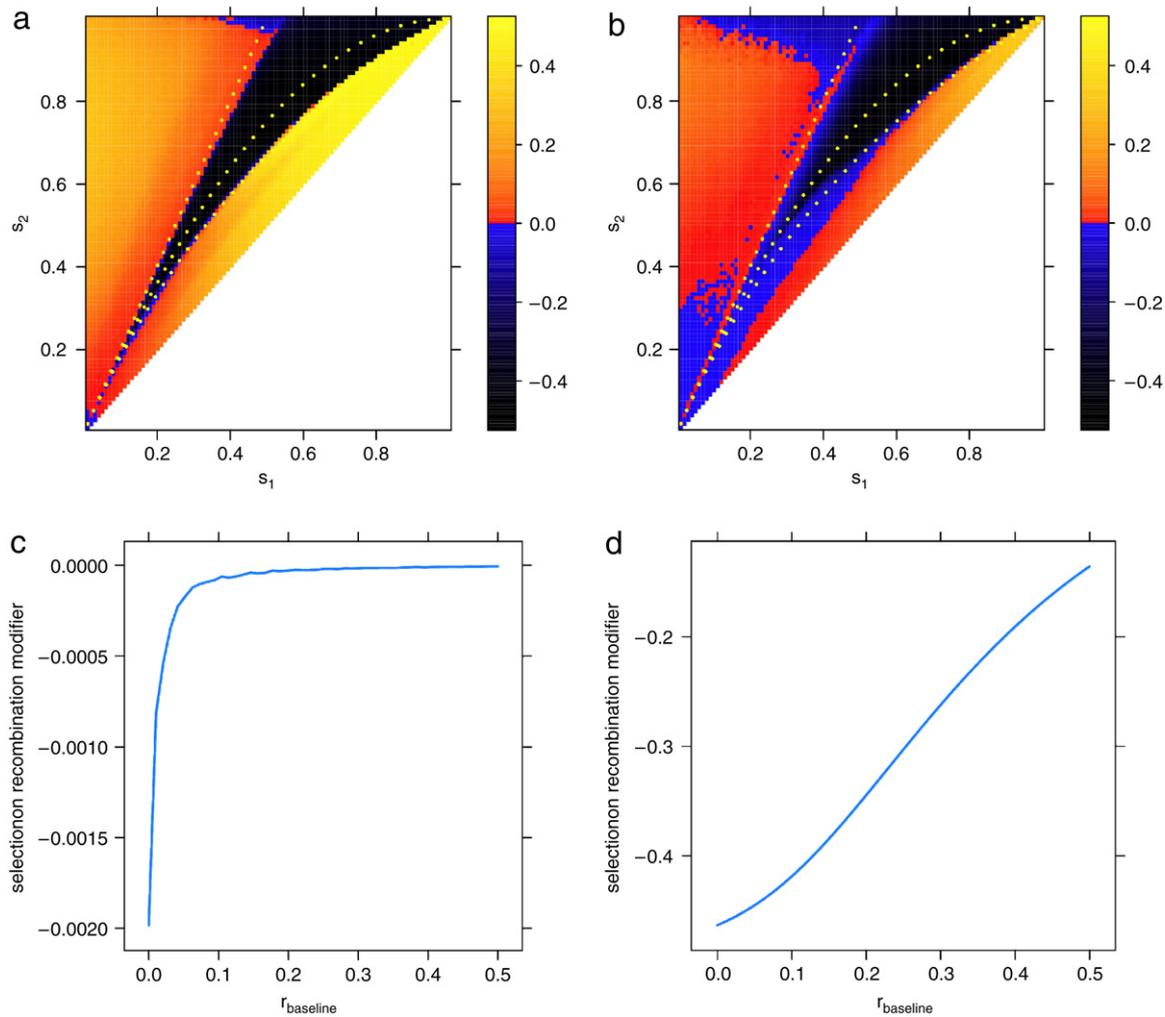


Fig. B.1. Selection on a modifier of recombination. (a and b) Selection on a modifier causing increased recombination for the generalized MA with varying coefficients s_1 and s_2 . The color of each of the points on the graph shows selection on the modifier measured as the frequency change of the modifier during a simulation run (1000 host generations) averaged over ten simulation runs with random initial allele frequencies at the interaction loci (the populations are initiated in linkage equilibrium). The recombination rate $r_{baseline}$ is 0 for a and 0.1 for b. The dotted lines indicate the HCE region and the MMA. (c and d) Selection on the recombination modifier for the MMA with $s = 0.05$ (a) and $s = 0.5$ (b) and varying recombination rates $r_{baseline}$. Parameters: $n_{PC} = 5$; $r_M = 0.05$; $\Delta r = 0.01$; $m = 10^{-5}$.

effect if the modifier alleles of the two recombining genomes are different). Since we are interested in the fate of the allele M increasing the recombination rate, $\Delta r > 0$ in all simulations. In order to investigate the fate of the modifier allele (M) increasing recombination, we introduce the M allele in 50% of the population after a burn-in phase of 1000 host generations. Fig. B.1 shows the selection on the modifier for different recombination rates and selection coefficients. In accordance with the dynamical behavior described in the main text, recombination is selected against when HC behavior occurs (compare Fig. 6 with panel a of Fig. B.1). Furthermore, the strength of the selection against recombination in the HCE region decreases with increasing recombination rate (panels b–d) and this effect of recombination becomes weaker as selection increases (panels c and d).

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tpb.2008.09.007.

References

Agrawal, A.F., Lively, C.M., 2002. Infection genetics: Gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research* 4, 79–90.

- Altenberg, L., Feldman, M.W., 1987. Selection, generalized transmission and the evolution of modifier genes. 1. The reduction principle. *Genetics* 117, 559–572.
- Bell, G., Smith, J.M., 1987. Short-term selection for recombination among mutually antagonistic species. *Nature* 328, 66–68.
- Bodmer, W.F., Felsenstein, J., 1967. Linkage and selection—Theoretical analysis of deterministic 2 locus random mating model. *Genetics* 57, 237.
- Feldman, M.W., Franklin, I., Thomson, G.J., 1974. Selection in complex genetic systems. 1. Symmetric equilibria of 3-locus symmetric viability model. *Genetics* 76, 135–162.
- Fischer, O., Schmid-Hempel, P., 2005. Selection by parasites may increase host recombination frequency. *Biology Letters* 1, 193–195.
- Frank, S.A., 1993a. Coevolutionary genetics of plants and pathogens. *Evolutionary Ecology* 7, 45–75.
- Frank, S.A., 1993b. Specificity versus detectable polymorphism in host–parasite genetics. *Proceedings of the Royal Society of London Series B-Biological Sciences* 254, 191–197.
- Frank, S.A., 1994. Recognition and polymorphism in host–parasite genetics. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 346, 283–293.
- Frank, S.A., 1996a. Problems inferring the specificity of plant–pathogen genetics—Reply. *Evolutionary Ecology* 10, 323–325.
- Frank, S.A., 1996b. Statistical properties of polymorphism in host–parasite genetics. *Evolutionary Ecology* 10, 307–317.
- Franklin, I., Lewontin, R.C., 1970. Is the gene the unit of selection. *Genetics* 65, 707.
- 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.
- Hamilton, W.D., Axelrod, R., Tanese, R., 1990. Sexual reproduction as an adaptation to resist parasites (A review). *Proceedings of the National Academy of Sciences of the United States of America* 87, 3566–3573.
- Jaenike, J., 1978. An hypothesis to account for the maintenance of sex within populations. *Evolutionary Theory* 3, 191–194.

- Kouyos, R.D., Salathé, M., Bonhoeffer, S., 2007. The red queen and the persistence of linkage-disequilibrium oscillations in finite and infinite populations. *BMC Evolutionary Biology* 7.
- Kover, P.X., Caicedo, A.L., 2001. The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites. *Molecular Ecology* 10, 1–16.
- Little, T.J., 2002. The evolutionary significance of parasitism: Do parasite-driven genetic dynamics occur *ex silico*? *Journal of Evolutionary Biology* 15, 1–9.
- Lively, C.M., 1987. Evidence from a New-Zealand snail for the maintenance of sex by parasitism. *Nature* 328, 519–521.
- Lively, C.M., Craddock, C., Vrijenhoek, R.C., 1990. Red queen hypothesis supported by parasitism in sexual and clonal fish. *Nature* 344, 864–866.
- Lively, C.M., Lyons, E.J., Peters, A.D., Jokela, J., 1998. Environmental stress and the maintenance of sex in a freshwater snail. *Evolution* 52, 1482–1486.
- Otto, S.P., Nuismer, S.L., 2004. Species interactions and the evolution of sex. *Science* 304, 1018–1020.
- Parker, M.A., 1994. Pathogens and sex in plants. *Evolutionary Ecology* 8, 560–584.
- Parker, M.A., 1996. The nature of plant-parasite specificity—Comment. *Evolutionary Ecology* 10, 319–322.
- Peters, A.D., Lively, C.M., 1999. The red queen and fluctuating epistasis: A population genetic analysis of antagonistic coevolution. *American Naturalist* 154, 393–405.
- Peters, A.D., Lively, C.M., 2007. Short- and long-term benefits and detriments to recombination under antagonistic coevolution. *Journal of Evolutionary Biology* 20, 1206–1217.
- Salathé, M., Kouyos, R.D., Regoes, R.R., Bonhoeffer, S., 2008. Rapid parasite adaptation drives selection for high recombination rates. *Evolution* 62, 295–300.
- Schmid-Hempel, P., Jokela, J., 2002. Socially structured populations and evolution of recombination under antagonistic coevolution. *American Naturalist* 160, 403–408.
- Smith, J.M., 1968. Evolution in sexual and asexual populations. *American Naturalist* 102, 469.
- Wilfert, L., Gadau, J., Baer, B., Schmid-Hempel, P., 2007a. Natural variation in the genetic architecture of a host–parasite interaction in the bumblebee *Bombus terrestris*. *Molecular Ecology* 16, 1327–1339.
- Wilfert, L., Gadau, J., Schmid-Hempel, P., 2007b. The genetic architecture of immune defense and reproduction in male *Bombus terrestris* bumblebees. *Evolution* 61, 804–815.