Bistability in two-locus models with selection, mutation, and recombination

Su-Chan Park*,[a], Joachim Krug[a]

[a] *Institut für Theoretische Physik, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany*

Abstract

The evolutionary effect of recombination depends crucially on the epistatic interactions between linked loci. A paradigmatic case where recombination is known to be strongly disadvantageous is a two-locus fitness landscape displaying reciprocal sign epistasis with two fitness peaks of unequal height. Focusing on the occurrence of bistability in the equilibrium solutions, we consider here the deterministic, haploid two-locus model with reversible mutations, selection and recombination. We find analytic formulae for the critical recombination probability r_c above which two stable stationary solutions appear which are localized on each of the two fitness peaks. We also derive the stationary genotype frequencies in various parameter regimes. When the recombination rate is close to r_c and the fitness difference between the two peaks is small, we obtain a compact description in terms of a cubic polynomial which is analogous to the Landau theory of physical phase transitions.

Key words: evolution of recombination, reciprocal sign epistasis, bistability, Landau theory

1. Introduction

After more than a century of research, the evolutionary basis of sex and recombination remains enigmatic (Otto, 2009). In view of the complex evolutionary conditions faced by natural populations, the search for a single answer

*Corresponding author. Tel.: +49 221 470 3525; fax: +49 221 470 5159

Email addresses: psc@thp.uni-koeln.de (Su-Chan Park), krug@thp.uni-koeln.de (Joachim Krug)

to the question of why sexual reproduction has evolved and is maintained in the vast majority of eukaryotic species may well be futile. Nevertheless, theoretical population genetics has identified several simple, paradigmatic scenarios in which the conditions for an evolutionary advantage of sex can be identified in quantitative terms, and which are therefore open (at least in principle) to experimental verification.

One such scenario was proposed in the context of the adaptation of a population in a constant environment encoded by a fitness landscape, which assigns fitness values to all possible genotypes. In this setting, the relative advantage of sexual reproduction with respect to, say, the speed of adaptation or the mean fitness level at mutation-selection balance, depends crucially on the epistatic interactions between different loci (de Visser and Elena, 2007; Kouyos et al., 2007). In its simplest form, epistasis is associated with the curvature of the fitness surface, that is, the deviation of the fitness effect of multiple mutations, all of which are either deleterious or beneficial, compared to that predicted under the assumption of independent mutations which contribute multiplicatively or additively to fitness. It is well established that recombination speeds up the adaptation in such a fitness landscape if the epistatic curvature is negative, in the sense that the fitness of multiple mutants is lower than expected for independent mutations (Kondrashov, 1988).

Unfortunately, experimental evidence indicates that negative epistasis is not sufficiently widespread to provide a general explanation for the evolution of recombination (de Visser et al., 1997; Elena and Lenski, 1997; Bonhoeffer et al., 2004). Instead, empirical studies have highlighted the importance of a more complex form of epistasis, termed *sign epistasis*, in which not only the magnitude but also the sign of the fitness effect of a given mutation depends on the presence of mutations at other loci (Weinreich et al., 2005). A simple example of sign epistasis is shown in Fig. 1, which depicts a haploid two-locus fitness landscape. In this case the sign epistasis between the two loci is reciprocal, which leads to the appearance of two fitness peaks separated by a fitness valley (Poelwijk et al., 2007).

In the two-locus setting sign epistasis can be viewed as an extreme case of positive epistasis, and it may be expected to lead to a disadvantage of recombination. In a recent simulation study of the effect of recombination on empirically motivated five-locus fitness landscapes displaying sign epistasis (de Visser et al., 2009), we found that recombination can hamper the adaptation on such a landscape in a rather dramatic way: Instead of moving to the global fitness optimum, populations get trapped at local optima from

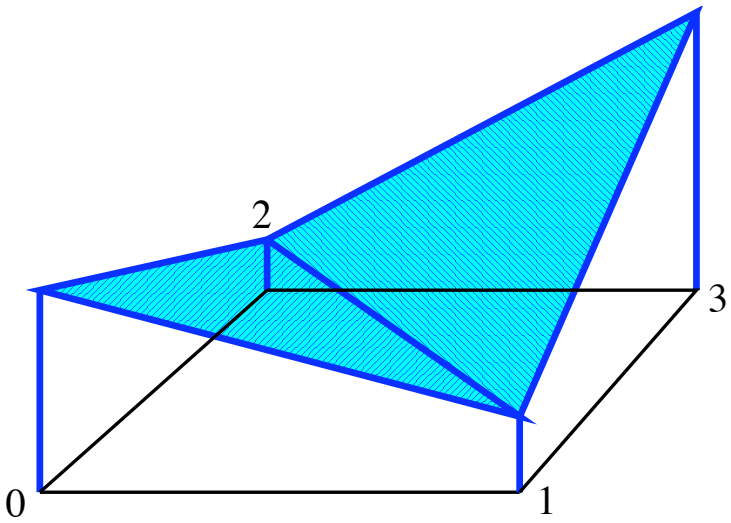

Figure 1: Schematic of a two-allele, two-locus fitness landscape with reciprocal sign epistasis and unequal peak heights. Fitness is represented by the height above the plane in which the four genotypes (labeled 0, 1, 2, 3) reside.

which (in the limit of infinite population size) they never escape. Mathematically, these numerical calculations suggest the appearance of multiple stable stationary solutions of the deterministic, infinite population dynamics above a threshold value r_c of the recombination rate.

The simplest situation where this phenomenon can occur is the two-locus landscape shown in Fig. 1, where it implies a bistability with two stationary solutions localized near each of the fitness peaks labeled by 0 and 3. Indications for the occurrence of bistability in this system can be found in several earlier studies of the two-locus problem. Crow and Kimura (1965) derived a condition for the initial increase of the high peak mutant in a population dominated by the low peak genotype. Eshel and Feldman (1970) established a sufficient condition for the low peak population to remain trapped for all times when mutations are unidirectional ($0 \to 1, 2 \to 3$), which was subsequently refined by Karlin and McGregor (1971). Feldman (1971) and Rutschman (1994) obtained conditions for the existence of multiple stationary solutions in the absence of mutations, and Bürger (1989) analyzed the problem with reversible mutations and equal peak heights. The case of symmetric fitness peaks was considered as a model for compensatory mutations in RNA secondary structure by Stephan (1996) and Higgs (1998). Finally,

bistability induced by recombination has also been observed in multilocus models in the context of quasispecies theory (Boerlijst et al., 1996; Jacobi and Nordahl, 2006) and evolutionary computation (Wright et al., 2003).

However, a comprehensive analysis of the paradigmatic two-locus system with reversible mutations, reciprocal sign epistasis and fitness peaks of unequal height does not seem to be available, and the present paper aims to fill this gap. In the next section the models used in our work are introduced in a general multilocus framework. We then specialize to the two-locus case illustrated in Fig. 1, and show that finding the stationary solutions of the model amounts to analyzing the zeros of a fourth order polynomial. The bulk of the paper is devoted to extracting useful information about the critical recombination rate and the genotype frequency distribution in various parameter regimes, and we conclude with a summary of our results and a discussion of open problems.

2. Models

2.1. Multilocus formalism

We consider a haploid population evolving on the genotype space formed by biallelic sequences with L loci. The alleles at any locus are denoted by 0 and 1, and genotype sequences will be denoted by σ, σ' etc. We assume that the population evolves under Wright-Fisher dynamics in discrete, non-overlapping generations (Fisher, 1930; Wright, 1931). Each sequence σ is endowed with a fitness w_σ, which defines the fitness landscape. The population evolves under the influence of selection, mutation, and recombination. Since we are interested in the emergence of bistability, in the sense of the existence of multiple stationary frequency distributions, the limit of infinite population size is assumed. Thus, the frequency of each genotype evolves deterministically. The frequency changes according to the following order: selection, mutation, then recombination.

Let $f_\sigma(\tau)$ denote the frequency of genotype σ at generation τ. The frequency of the genotype at generation $\tau + 1$ after selection is proportional to the product of its frequency at generation τ and its fitness w_σ. After selection, mutations can change the frequency of genotypes. We assume that the mutation probability per generation and locus, μ, depends neither on the location of the locus in the sequence nor on the alleles, and mutations are assumed to be independent of each other. Mathematically speaking, the

mutation from sequence σ' to sequence σ then occurs with probability

$$U_{\sigma,\sigma'} = (1-\mu)^{L-d(\sigma,\sigma')}\mu^{d(\sigma,\sigma')}, \tag{1}$$

where $d(\sigma,\sigma')$ is the Hamming distance between two sequences σ and σ', i.e. the number of loci at which the two sequences differ. After selection and mutation, the frequency distribution will be

$$\tilde{p}_\sigma = \sum_{\sigma'} U_{\sigma\sigma'} \frac{w_{\sigma'}}{\bar{w}\,(\tau)} f_{\sigma'}\,(\tau)\,, \tag{2}$$

where $\bar{w}(\tau) \equiv \sum_\sigma w_\sigma f_\sigma(\tau)$ is the mean fitness at generation t.

After selection and mutation, recombination follows. At first, two sequences, say σ' and σ'', are selected with probability $\tilde{p}_{\sigma'}\tilde{p}_{\sigma''}$. With probability $1-r$, no recombination happens. With probability r, however, the two sequences recombine in a suitable way and generate two recombinants, which may be identical to or different from σ' and σ'' depending on the recombination scheme and the properties of the two initial sequences. One may interpret the case $r<1$ as a model for organisms which can reproduce sexually as well as asexually.

Within our scheme, the recombination probability $R_{\sigma|\sigma'\sigma''}$ that one of the recombinants which is chosen randomly is the sequence σ takes the form

$$R_{\sigma|\sigma'\sigma''} = (1-r)\frac{\delta_{\sigma\sigma'}+\delta_{\sigma\sigma''}}{2} + rW_{\sigma|\sigma'\sigma''}, \tag{3}$$

where δ_{xy} is the Kronecker delta symbol which is 1 (0) if $x=y$ ($x\neq y$). The first term stands for the result of an asexual reproduction event. The second term which describes the generation of a possible new sequence takes a model-dependent form. For example, for the uniform crossover scheme (Sywerda, 1989), i.e., the random shuffling of alleles at each locus from two parental sequences, one has

$$W_{\sigma|\sigma'\sigma''} = F_{\sigma|\sigma'\sigma''}2^{-d(\sigma',\sigma'')}, \tag{4}$$

where $F_{\sigma|\sigma'\sigma''}$ is 1 if σ can be a recombinant of σ' and σ'' and 0 otherwise. If crossover occurs by breakage of a randomly selected juncture among the $L-1$ links followed by rejoining, the probability becomes

$$W_{\sigma|\sigma'\sigma''} =? \frac{\delta_{\sigma\sigma'}+\delta_{\sigma\sigma''}}{2}\left(1-(1-\delta_{\sigma'\sigma''})\,\frac{\ell_d-\ell_1}{L-1}\right) + \frac{C_{\sigma|\sigma'\sigma''}}{2(L-1)}, \tag{5}$$

where d (although we omit the arguments) is the Hamming distance between σ' and σ'', ℓ_i $(i = 1, \ldots, d)$ stands for the location of the i^{th} different locus between σ' and σ'' counted from the left ($\ell_i < \ell_j$ if $i < j$), and $C_{\sigma|\sigma'\sigma''} = \ell_{j+1} - \ell_j$ if the crossover occurring between loci ℓ_{j+1} and ℓ_j can result in a genotype σ; if $d < 2$, C is set to be zero, see below. The factor 2 occurring in the denominators reflects that there are two possible recombinants that can arise from a single crossover. The first term shows that if the breakage occurs outside the region between the first and last locus at which the two sequences differ, no new sequence appears the recombination. Note that when $d(\sigma', \sigma'') \leq 1$, we get

$$W_{\sigma|\sigma'\sigma''} = \frac{\delta_{\sigma\sigma'} + \delta_{\sigma\sigma''}}{2} \quad \text{if} \quad d(\sigma', \sigma'') \leq 1 \tag{6}$$

for both schemes. This is not a coincidence; it just shows that new sequences arise by recombination only if $d(\sigma', \sigma'') \geq 2$.

There are two sum rules for any recombination scheme. One is

$$\sum_{\sigma} R_{\sigma|\sigma'\sigma''} = 1, \tag{7}$$

which simply means that some sequence has to be produced by the recombination event. The other is

$$\sum_{\sigma', \sigma''} R_{\sigma|\sigma'\sigma''} = 2^L \tag{8}$$

for any σ. To prove (8), we introduce the transformation T_i which changes the allele at locus i (for convenience, locus 1 is the leftmost letter in the sequence and i increases as we move from left to right, but the conclusion is independent of this convention). For example, $T_1(0, 0, 0) = (1, 0, 0)$ and $T_2(1, 1, 1) = (1, 0, 1)$ when $L = 3$. Clearly $T_i T_j = T_j T_i$ and $T_i^2 = I$, where I is the identity transformation. Let T be any composition of T_i's. Then

$$R_{T\sigma|T\sigma'T\sigma''} = R_{\sigma|\sigma'\sigma''} \tag{9}$$

is a trivial identity because the recombination probability only depends on whether the alleles at a given locus are the same in the two recombining sequences or not, and T does not change this equality relation. Since $\sigma' \neq \sigma''$ implies $T\sigma' \neq T\sigma''$ for any T, we get

$$G_\sigma \equiv \sum_{\sigma', \sigma''} R_{\sigma|\sigma'\sigma''} = \sum_{T\sigma', T\sigma''} R_{T\sigma|T\sigma'T\sigma''} = \sum_{\sigma', \sigma''} R_{T\sigma|\sigma'\sigma''} = G_{T\sigma}, \tag{10}$$

where the first equality is from Eq. (9), the second equality is a change of summation variables, and third is the definition of G. Since T is arbitrary, we proved that G_σ is same for all sequences. Using Eq. (7), we can prove Eq. (8) by observing that

$$2^L G_0 = \sum_\sigma G_\sigma = \sum_{\sigma'\sigma''} \sum_\sigma R_{\sigma|\sigma'\sigma''} = 2^{2L}, \tag{11}$$

where $\sigma = 0$ is the sequence with 0's at all loci and Eq. (7) has been used. One easily checks that the sum rule is valid for $r = 0$. If the number of alleles at the i^{th} locus is n_i rather than 2, 2^L in (8) is replaced by $\prod_i n_i$. The proof can be easily generalized to this case.

After selection, mutation, and recombination, the frequency of a sequence σ in the next generation becomes

$$\begin{aligned} p_\sigma &= \sum_{\sigma',\sigma''} R_{\sigma|\sigma'\sigma''} \tilde{p}_{\sigma'} \tilde{p}_{\sigma''} = \tilde{p}_\sigma + r \left(\sum_{\sigma' \neq \sigma''} W_{\sigma|\sigma'\sigma''} \tilde{p}_{\sigma'} \tilde{p}_{\sigma''} - \tilde{p}_\sigma (1 - \tilde{p}_\sigma) \right) \\ &= \tilde{p}_\sigma + r \left(\sum_{d(\sigma',\sigma'')\geq 2} W_{\sigma|\sigma'\sigma''} \tilde{p}_{\sigma'} \tilde{p}_{\sigma''} - \tilde{p}_\sigma \sum_{d(\sigma',\sigma)\geq 2} \tilde{p}_{\sigma'} \right), \end{aligned} \tag{12}$$

where we have used Eq. (6). In Eq. (12), $d(\sigma', \sigma'') \geq 2$ signifies the summation over all ordered pairs of sequences with Hamming distance larger than 1. Likewise $d(\sigma', \sigma) \geq 2$ means the summation over all σ' whose Hamming distance from σ is larger than 1.

The evolution Eq. (12) becomes trivial if $\mu = \frac{1}{2}$. In this case, $U_{\sigma,\sigma'}$ becomes constant with value 2^{-L}, and Eqs. (2), (8), and (12) yield $p_\sigma = \tilde{p}_\sigma = 2^{-L}$ regardless of $f_\sigma(\tau)$, the fitness landscape, and any other parameters. From this trivial evolution one can learn that the equilibrium distribution depends on the order of the three processes selection, mutation, and recombination. For example, if mutation precedes selection, the frequency distribution at the next generation is not constant, though stationarity is still attained in one generation. We therefore believe that the qualitative features of the model are insensitive to the order of three evolutionary forces. In the following we always exclude the trivial case $\mu = 1/2$, and restrict the analysis to $\mu < 1/2$.

2.2. Two loci

It has already been shown by Bürger (1989) and Higgs (1998) that the haploid, two-locus setting is sufficient to obtain multiple equilibrium solu-

tions. Hence, in what follows, we will restrict ourselves to the two-locus problem ($L = 2$). We translate σ into a binary number, numbering the four genotypes according to $00 \to 0,\ 01 \to 1,\ 10 \to 2,\ 11 \to 3$. Equation (12) with Eq. (5) (single crossover scheme) then becomes

$$\begin{aligned} p_0 &= \frac{1}{\bar{w}(\tau)}\left[f_0 w_0(1-\mu)^2 + (f_1 w_1 + f_2 w_2)\mu(1-\mu) + f_3 w_3 \mu^2\right] \\ &\quad -r(1-2\mu)^2 \frac{\tilde{D}}{\bar{w}(\tau)^2}, \\ p_1 &= \frac{1}{\bar{w}(\tau)}\left[f_1 w_1(1-\mu)^2 + (f_0 w_0 + f_3 w_3)\mu(1-\mu) + f_2 w_2 \mu^2\right] \\ &\quad +r(1-2\mu)^2 \frac{\tilde{D}}{\bar{w}(\tau)^2}, \end{aligned} \tag{13}$$

where $\tilde{D}(\tau) = f_0(\tau) f_3(\tau) w_0 w_3 - f_1(\tau) f_2(\tau) w_1 w_2$ and the equation for f_3 (f_2) can be obtained by exchanging $0 \leftrightarrow 3$ ($1 \leftrightarrow 2$) in the first (second) equation above. If instead we use the uniform crossover scheme Eq. (4), r in the above equations is replaced by $r/2$. Note that the case of free recombination frequently considered in the literature corresponds to $r = \frac{1}{2}$ in Eq. (13), which can also be interpreted as uniform crossover with probability 1.

By rearranging (13), we get

$$\frac{\bar{w}(\tau)}{1-2\mu}(p_0 - p_3) = f_0 w_0 - f_3 w_3, \tag{14}$$

$$\frac{\bar{w}(\tau)}{1-2\mu}(p_1 - p_2) = f_1 w_1 - f_2 w_2, \tag{15}$$

$$p_0 p_3 - p_1 p_2 = \frac{(1-r)(1-2\mu)^2}{\bar{w}(\tau)^2}\tilde{D}. \tag{16}$$

A nontrivial conclusion one can draw from Eq. (16) is that linkage equilibrium, $p_0 p_3 = p_1 p_2$, is attained in one generation if $r = 1$, which corresponds to the single crossover scheme with recombination probability 1.

The stationary distribution is calculated by setting $p_\sigma = f_\sigma$ in above three equations, which gives

$$\frac{f_0}{f_3} = \frac{\bar{w} - (1-2\mu)w_3}{\bar{w} - (1-2\mu)w_0} \equiv A, \tag{17}$$

$$\frac{f_1}{f_2} = \frac{\bar{w} - (1-2\mu)w_2}{\bar{w} - (1-2\mu)w_1}, \tag{18}$$

$$\frac{f_0 f_3}{f_1 f_2} = \frac{\bar{w}^2 - (1-r)(1-2\mu)^2 w_1 w_2}{\bar{w}^2 - (1-r)(1-2\mu)^2 w_0 w_3} \equiv B, \tag{19}$$

where $\bar{w}$ is the mean fitness at stationarity. With the two additional conditions

$$f_0 + f_1 + f_2 + f_3 = 1 \tag{20}$$

and

$$\bar{w} = w_0 f_0 + w_1 f_1 + w_2 f_2 + w_3 f_3 \tag{21}$$

the steady state solution is fully determined. Without loss of generality, w_3 is set to be largest (with possible degeneracy). For simplicity, we set $w_1 = w_2$, which by (18) implies that $f_1 = f_2 \equiv f$. Since the dynamics is invariant under multiplication of all fitnesses by the same factor, we can choose

$$w_3 = 1, w_0 = 1 - t, w_1 = w_2 = 1 - t - s \tag{22}$$

with $0 < t < 1$ and $-t < s < 1 - t$. For the sake of brevity, however, we will sometimes use w_0 and w_1 rather than s and t in what follows.

Note that the problem studied by Bürger (1989) and Higgs (1998) corresponds to the case with $t = 0$ and $0 < s$. In this paper, we will assume that $t > 0$ and $s > 0$, that is, the fitness landscape has a unique global optimum and reciprocal sign epistasis (Fig. 1). Unlike the fitness landscape with symmetric peaks, it is difficult to find the solution exactly (see also Appendix B). We will investigate the approximate solutions by assuming that some of parameters are very small compared to others.

3. Bistability

3.1. General behavior of solutions

Since the frequency of each genotype is strictly positive in the steady state, Eq. (17) excludes the mean fitness $\bar{w}$ from being in the range between $(1-2\mu)w_0$ and $(1-2\mu)w_3$. For obvious reasons, we will refer to a solution with $\bar{w} > (1-2\mu)w_3$ as high-fitness solution (HFS) and a solution with $\bar{w} < (1-2\mu)w_0$ as low-fitness solution (LFS). As an immediate consequence of Eq. (17), we see that a HFS (LFS) implies $f_3 > f_0$ ($f_3 < f_0$). The largest fitness being w_3, the existence of a HFS is expected regardless of the strength of the recombination probability (we will see later that this is indeed the case). Hence we are mainly interested in the conditions which allow for a LFS. Later, we will see that if they exist, there are actually two LFS's, only

one of which is locally stable. Since we are interested in stable solutions, in what follows LFS refers exclusively to the locally stable solution. Intuitively, the HFS should be locally stable, so the emergence of a LFS naturally implies bistability.

Without much effort, one can easily find necessary conditions for the bistability. First note that $\bar{w}$ is larger than w_1 by definition. Therefore, a LFS is possible only if $w_1 < (1-2\mu)w_0$, or

$$\mu < \frac{s}{2(1-t)}. \tag{23}$$

Since $\mu = \frac{1}{2}$ gives a unique steady state ($f_i = \frac{1}{4}$ for all i) regardless of r, it could have been anticipated that bistability requires a restriction on μ. A necessary condition on r can be obtained from Eq. (19). While the numerator of the expression defining B is always positive, the denominator would be negative for the LFS if $w_0 - (1-r)w_3 < 0$. Hence a necessary condition for bistability is

$$r > 1 - \frac{w_0}{w_3} = t, \tag{24}$$

which appears also in earlier work (Crow and Kimura, 1965; Eshel and Feldman, 1970; Karlin and McGregor, 1971; Feldman, 1971; Rutschman, 1994). Most of the calculations in this paper are devoted to refining the conditions for bistability. To this end, we will reduce the five equations (17,18,19,20,21) to a single equation for $\bar{w}$.

Equations (17) and (19) along with the normalization (20) yield the relations

$$f = \frac{\sqrt{A}}{2\sqrt{A} + \sqrt{B}(1+A)}, \quad f_3 = \frac{\sqrt{B}}{2\sqrt{A} + \sqrt{B}(1+A)}, \quad f_0 = Af_3. \tag{25}$$

and the definition (21) of the mean fitness $\bar{w}$ implies that

$$2\left(\bar{w} - w_1\right) = \sqrt{\frac{B}{A}}\left(w_3 + w_0 A - (1+A)\bar{w}\right). \tag{26}$$

Since our main focus is on the LFS, it is convenient to define the auxiliary variable x through

$$\bar{w} = (1-2\mu)(w_0 - x), \tag{27}$$

which implies that $x < -t$ and $x > 0$ for the HFS and the LFS, respectively. We note for future reference that with this reparametrization, Eq.(17) takes the simple form

$$A = 1 + \frac{t}{x}. \tag{28}$$

Taking the square on both sides of Eq. (26) results in a quartic equation

$$h(x) \equiv h_0(x) + r h_1(x) = 0, \tag{29}$$

where the polynomials $h_0(x)$ and $h_1(x)$ are *independent of* r (see Appendix A and the Mathematica file in the online supplement for explicit expressions).

We can draw some general conclusions concerning solutions of the quartic equation by evaluating $h(x)$ at selected points. Since $h(x)$ is negative at $x = 0$, $x = -t$, and at the point $x = x_1$ defined by $(1 - 2\mu)(w_0 - x_1) = w_1$, and the coefficient of fourth order term is positive (see Appendix A), there always exist solutions in $x > x_1$ and $x < -t$ which correspond to $\bar{w} < w_1$ and $\bar{w} > (1 - 2\mu)w_3$, respectively. The meaningless solution, $\bar{w} < w_1$, has appeared because we took squares on both sides of Eq. (26). In the online supplement, we show that $h(x)$ is positive when $x = 1 - t - 1/(1 - 2\mu)$, or $\bar{w} = w_3 = 1$. This proves, as anticipated, that the HFS is present for all values of r.

Hence the condition for bistability is recast as the condition for $h(x)$ to have a positive solution which is smaller than x_1. Because $h(0)$ and $h(x_1)$ are negative, the existence of a solution in the range $0 < x < x_1$ always entails two solutions in the same region if we count the number of degenerate solutions as 2. Let us assume that $h(x) = 0$ has two degenerate solutions at $x = x_c$ when $r = r_c$. This means that x_c is the simultaneous solution of two equations $h(x_c) = h'(x_c) = 0$, where the prime denotes the derivative with respect to the argument. Later, this simple relation will play a crucial role in finding r_c as well as x_c. Now let us change r infinitesimally from r_c to $r_c + \varepsilon_r$, and let the solutions of $h(x) = 0$ for $r = r_c + \varepsilon_r$ take the form $x_c + \varepsilon_x$. Note that we are only interested in solutions with $\varepsilon_x \to 0$ as $\varepsilon_r \to 0$ in the complex x plane. Since two other solutions exist outside of the range $0 < x < x_1$ and both $h(0)$ and $h(x_1)$ are negative, $h(x)$ with $r = r_c$ has local maximum at $x = x_c$, that is, $h''(x_c; r_c) < 0$. Figure 2 illustrates this situation.

Using Eq. (29) we get

$$h(x_c + \varepsilon_x; r_c + \varepsilon_r) \approx -\frac{1}{2}|h''(x_c; r_c)|\varepsilon_x^2 + \varepsilon_r h_1(x_c) = 0. \tag{30}$$

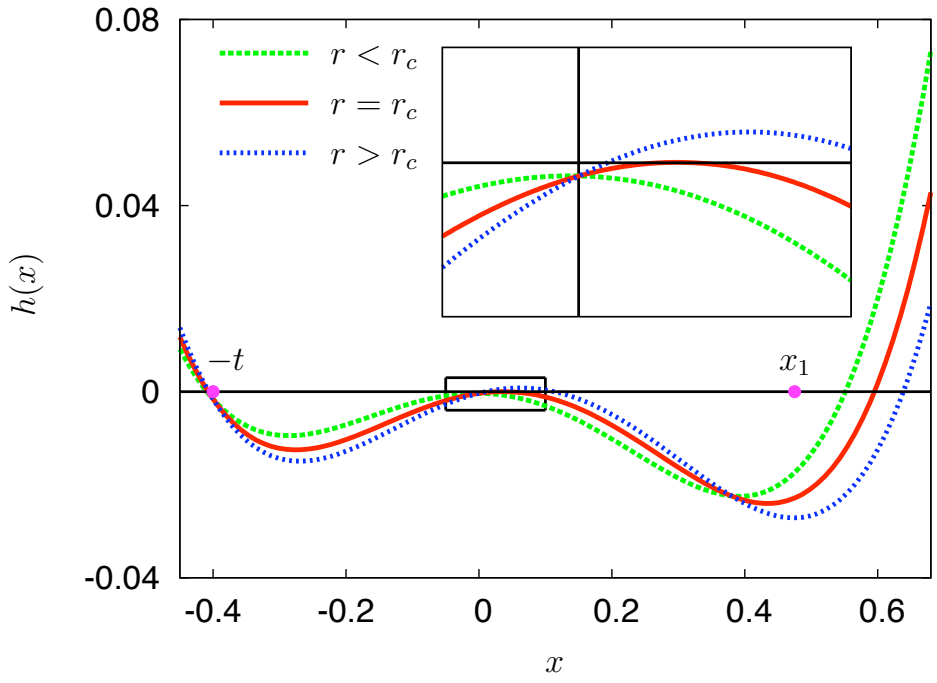

Figure 2: Behavior of the function $h(x)$ around the critical recombination probability. In this example, $s = 0.5$, $t = 0.4$, and $\mu = 0.01$ are used, which yields $r_c \approx 0.965$. The curves meet at (three) points where $h_1(x)$ vanishes. The locations of the points $x = -t$ and $x = x_1$ are indicated by filled circles. Low fitness solutions correspond to zero crossings in $0 < x < x_1$, and high fitness solutions to those with $x < -t$. The zero crossings with $x > x_1$ are spurious. Inset: Close-up view of the boxed area. For clarity, the vertical line at $x = 0$ is also drawn. One of the solutions of $h_1(x) = 0$ happens to be close to $x = 0$.

Hence real solutions are possible if the second term is positive. By definition, $r_c h_1(x_c) = -h_0(x_c)$. For $r = 0$, Eq. (29) reduces to $h_0(x) = 0$, and we may conclude from the condition Eq. (24), which is violated for $r = 0$, that this equation does not have a solution in the region $0 < x < x_1$. Hence $h_0(x_c) < 0$ because $h_0(0) < 0$ and $h_0(x_1) < 0$ (see Appendix A), which, in turn, implies that $h_1(x_c) > 0$. To summarize, we have proved that if x_c is the degenerate solution of $h(x) = 0$ when $r = r_c$, there are two solutions in the region $0 < x < x_1$ when $r > r_c$. One should note that r_c, if it exists, is unique, as otherwise a contradiction to Eq. (30) would arise.

To establish the existence of bistability for $r > r_c$, it remains to find r_c. Even though general solutions of quartic equations are available, it is difficult to extract useful information from them. Rather, we will look for approximate solutions by assuming that one of the three parameters μ, s, and t is much smaller than the other two. In fact, it follows from the condition (23) that there cannot be any bistability when $s \ll \mu$ (unless $t \to 1$). Hence, in this paper, we will not pursue the case that s is the smallest parameter.

Before turning to the derivation of the approximate solutions, we further exploit the linear r-dependence of $h(x)$ in Eq.(29). It implies that the two equations $h(x_c; r_c) = h'(x_c; r_c) = 0$ with two unknowns can be reduced to a single equation for x_c which does not involve r_c. To be specific, from $h_0(x_c) + r_c h_1(x_c) = h_0'(x_c) + r_c h_1'(x_c) = 0$ we obtain

$$r_c = -\frac{h_0(x_c)}{h_1(x_c)} = -\frac{h_0'(x_c)}{h_1'(x_c)}, \tag{31}$$

or

$$H(x_c) \equiv h_0(x_c)h_1'(x_c) - h_1(x_c)h_0'(x_c) = 0. \tag{32}$$

Thus, rather than finding r_c and x_c simultaneously, we will first find x_c by solving Eq. (32), which in turn, gives r_c from Eq. (31). Equation (32) will be analyzed below, after we have introduced one more useful concept.

3.2. Critical mutation probability

As evidenced by Eq. (23), a finite critical recombination probability r_c can exist only if μ is sufficiently small. Mathematically, this implies that r_c diverges as μ approaches a certain critical mutation probability μ_c, such that bistability is not possible for $\mu > \mu_c$. Although r cannot, strictly speaking, exceed unity, we will see later that this definition of μ_c will be of great use to find an accurate expression for r_c. Setting $r_c = \infty$ in Eq. (31), we see that μ_c

is the solution of the equations $h_1(x_c) = h_1'(x_c) = 0$. Since $h_1(x)$ is a cubic function taking the form $h_1(x) = -C_3x^3 - C_2x^2 + C_1x - C_0$, x_c is also the solution of the equation

$$G(x) = xh_1'(x) - 3h_1(x) = C_2x^2 - 2C_1x + 3C_0 = 0. \tag{33}$$

¿From $G(x)$ and $h_1'(x)$, we can construct two linear equations for x such that

$$G_1(x) = C_2h_1'(x) + 3C_3G(x) = -2(3C_1C_3 + C_2^2)x + C_1C_2 + 9C_0C_3 = 0, \tag{34}$$

$$G_2(x) = \frac{1}{x}(C_1G(x) - 3C_0h_1'(x)) = (C_1C_2 + 9C_0C_3)x - 2(C_1^2 - 3C_0C_2) = 0, \tag{35}$$

where we have used the fact that $x_c \neq 0$. Hence, the value of x_c for $r_c = \infty$ is given by

$$x_c^\infty = \frac{C_1C_2 + 9C_0C_3}{2(C_2^2 + 3C_1C_3)} = \frac{2(C_1^2 - 3C_0C_2)}{C_1C_2 + 9C_0C_3}. \tag{36}$$

Note that C_i's depend on μ, s, and t, which means we have an equation for μ_c such that

$$(C_1C_2 + 9C_0C_3)^2 - 4(C_1^2 - 3C_0C_2)(C_2^2 + 3C_1C_3) = 0, \tag{37}$$

which, in turn, provides x_c^∞ by inserting μ_c in Eq. (36). In fact, Eq. (37) is equivalent to the discriminant of cubic polynomials (see online supplement).

We now assume that $s, t \ll 1$, which also implies $\mu_c \ll 1$ by Eq. (23). Then to leading order the C_i's become

$$C_3 = 2s + t,\ C_2 = (t + 2\mu)(2s + t) - s^2,\ C_1 = t(s^2 - 2\mu(2s + t)),\ C_0 = \mu^2t^2, \tag{38}$$

and Eq. (37) becomes (see the Mathematica file in online supplement)

$$3s^{10}\nu^2(1 - t)\left(1 + \nu - z(2 + \nu)\right)^2 M(z) = 0, \tag{39}$$

where $z = \mu/s$, $\nu = t/s$, and $M(z)$ is a cubic polynomial with

$$\begin{aligned} M(z) &= 32(\nu + 2)z^3 - (13\nu^2 + 48\nu + 48)z^2 \\ &\quad +2(2\nu^3 + 7\nu^2 + 9\nu + 6)z - (1 + \nu)^2. \end{aligned} \tag{40}$$

Note that $z = (1+\nu)/(2+\nu)$ cannot be a solution because of Eq. (23). Hence the critical mutation probability is the solution of the equation $M(z) = 0$. As shown in the online supplement, $M(z)$ is an increasing function of z, which,

along with $M(0) < 0$, allows only one positive real solution of $M(z) = 0$ unless $\nu = 0$. One can find the exact solution in the Mathematica file, which is not suitable to present here. We will just present the asymptotic behavior of the solution for later purposes (see the Mathematica file). When $t \ll s$ ($\nu \ll 1$),

$$\mu_c = \frac{s}{4}\left[1 - 3\left(\frac{t}{4s}\right)^{2/3}\right] \text{ and } x_c^\infty = \left(\frac{st^2}{16}\right)^{1/3}, \tag{41}$$

and when $s \ll t$ ($\nu \gg 1$)

$$\mu_c \approx \frac{s^2}{4t} \text{ and } x_c^\infty = \frac{s^2}{4t}. \tag{42}$$

Although we derived the above two expressions from the exact solution (see Mathematica file), it is not difficult to find the asymptotic behavior without resorting to the exact solution. When ν is finite, it would be more useful to have a numerical value. In the case $\nu = 1$ ($t = s$) we get

$$\mu_c \approx 0.107s \text{ and } x_c^\infty \approx 0.0616s. \tag{43}$$

Below we will see how μ_c can be used to derive improved approximations for r_c.

3.3. Approximation for small mutation probability

Now we will move on to finding the critical recombination probability. We begin with the investigation of the approximate solutions for small mutation probability ($\mu \ll s, t$). Let us assume that $x_c = x_0 + a_\mu \mu + O(\mu^2)$, which should be justified self-consistently. For later purposes, we introduce the parameters

$$\alpha = (1-t)(s+t)^2 - s^2, \quad \beta = (1-t)(s+t)^2 + s^2. \tag{44}$$

Note that α is positive if $s < \sqrt{1-t} + 1 - t$, which is automatically satisfied because s is smaller than $1 - t$ by definition.

The leading behavior of $H(x_c)$ becomes

$$H(x_c) = x_0^2(t + x_0)^2(a_2 x_0^2 + 2a_1 x_0 + a_0) + O(\mu), \tag{45}$$

where the a_i's are parameters satisfying $a_1^2 - a_0 a_2 < 0$ (see the Mathematica file in online supplement). Since x_c must be positive, the only possible

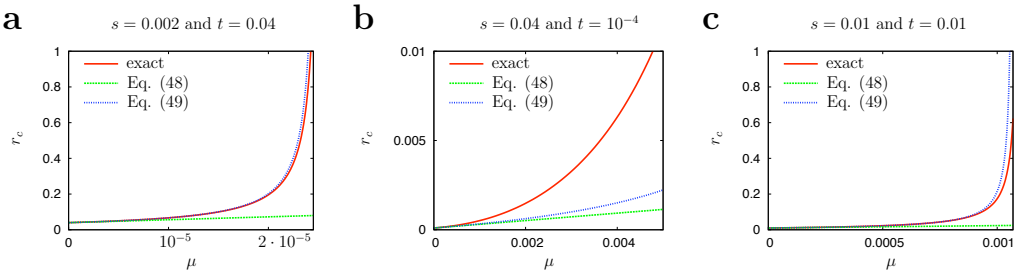

Figure 3: Comparison of the exact r_c with the approximate solutions Eqs. (48) and (49) for three cases: (a) $s \ll t$, (b) $t \ll s$, and (c) $s = t$. Eq. (49) shows an excellent agreement with the exact r_c when t is not too small compared to s. Eq. (48) is generally poor in predicting r_c.

solution for x_0 is $x_0 = 0$. Accordingly, the actual leading behavior of $H(x_c)$ becomes

$$H(x_c) = -\mu^2 s^2 t^2 \left((1-t)^2\alpha - a_\mu^2\beta\right) + O(\mu^3) = 0, \tag{46}$$

which gives

$$a_\mu = (1-t)\left(\frac{\alpha}{\beta}\right)^{1/2}. \tag{47}$$

By putting $x_c = a_\mu \mu$ into Eq. (31) and keeping terms up to $O(\mu)$, we find

$$r_c = t + 2\frac{1-t}{s^2}\left(\alpha + \sqrt{\alpha\beta}\right)\mu \equiv t + c_\mu \mu, \tag{48}$$

which clearly satisfies the bound Eq. (24), and shows that this bound is saturated when $\mu \to 0$.

The approximation for r_c can be significantly improved by matching the approximation for small μ with the behavior of r_c when μ is close to μ_c. Since r_c becomes infinite at $\mu = \mu_c$, we write

$$r_c = t\frac{1+\rho\mu}{1-\mu/\mu_c}, \qquad \text{with} \qquad \rho = 2\frac{1-t}{s^2 t}\left(\alpha + \sqrt{\alpha\beta}\right) - \frac{1}{\mu_c}. \tag{49}$$

where ρ is determined such that the correct leading behavior is reproduced when μ is small.

In Fig. 3, the exact values of r_c obtained from numerical calculations are compared with the two approximations Eqs. (48) and (49). One can find more such graphs in the Mathematica file. As expected, both expressions give a reliable prediction when μ is sufficiently small. However, as the exact

value of r_c becomes larger, Eq. (48) does not give an accurate estimation, which is not surprising. The more surprising observation is that Eq. (49) gives an excellent prediction for r_c in almost all ranges. However, Eq. (49) becomes a poor guidance for r_c as t gets smaller, which suggests that the case with small t should be treated separately. In the next subsection, we will study the two-locus model for small t.

3.4. Approximation for small t

Now let us move on to the case with small t ($t \ll \mu, s$) which connects the present study to the problem with symmetric fitness peaks considered by Bürger (1989) and Higgs (1998). As before, we will find x_c from Eq. (32).

As shown by Bürger (1989) and Higgs (1998) (see also Appendix B), x_c should approach zero as $t \to 0$. This can be rigorously shown from Eq. (32). If we assume that x_c is finite as $t \to 0$, the leading order of Eq. (32) becomes

$$2s^2(2-s)(1-2\mu)\left[\{(1-2\mu)x - 2(\mu_{c0}-\mu)\}^2 + 4(1-s)\mu_{c0}^2\right]x^4 = 0, \quad (50)$$

where

$$\mu_{c0} = \frac{s}{2(2-s)} \quad (51)$$

is the critical mutation probability for the symmetric problem derived in Appendix B. Obviously, the only real solution is $x = 0$.

It might seem natural to assume that $x_c = c_1 t + O(t^2)$ as in Sec. 3.3. However, this turns out to be wrong. Not to be misled by an incorrect intuition, let us expand $H(x_c)$ only with the assumption that x_c is small, i.e., we do not specify how small x_c is compared to t. Then, to leading order, the equation $H(x_c) = 0$ becomes

$$2x_c^4 + 4x_c^3 t - 2a_t x_c t^2 - a_t t^3 = 0, \quad (52)$$

where

$$a_t = \frac{(s-2\mu)^2\mu^2}{2s(1-2\mu)(2\mu^2 + \mu_{c0}(s-4\mu))}. \quad (53)$$

Note that terms of order x_c^5 and x_c^6 are neglected compared to x_c^4. Actually, there is a term of order x_c^2 in $H(x_c)$, but its coefficient is $O(t^2)$, so it is neglected compared to $x_c t^2$. Let us assume that $x_c \sim t^\delta$. If $\delta \geq 1$, the solution of Eq. (52) is $x_c = -t/2$ which does not fall into the range $0 < x_c < x_1$. If $\delta < 1$, the leading behavior of Eq. (52) should be $x_c^4 - a_t t^2 x_c = 0$ which gives

$x_c \approx (a_t t^2)^{1/3}$. A more accurate estimate of x_c is derived in the Mathematica file, which reads

$$x_c \approx (a_t t^2)^{1/3} - \frac{t}{2}. \tag{54}$$

Note that the power $\frac{2}{3}$ was already observed when we calculated μ_c for $t \ll s$ in Eq. (41). ¿From Eq. (31) along with Eq. (54) we get

$$r_c \approx r_{c0}\left(1 + \frac{3\mu_{c0}(2\mu^2 + \mu_{c0}(s-4\mu))}{2s\mu^2(\mu_{c0}-\mu)}\left(a_t t^2\right)^{1/3} - \frac{2\mu_{c0}^2(1+s)}{s^2(\mu_{c0}-\mu)}t\right), \tag{55}$$

where r_{c0} is the critical recombination for $t = 0$ given by (see Appendix B for derivation)

$$r_{c0} = \frac{2\mu^2}{(1-2\mu)(\mu_{c0}-\mu)}. \tag{56}$$

We will use the same technique as in Sec. 3.3 to improve the quality of the approximation for r_c. Since r_c diverges at $\mu = \mu_c$ rather than at μ_{c0} (note that $\mu_{c0} > \mu_c$), we can associate the behavior for small r_c with that for larger r_c in such a way that

$$r_c = \frac{2\mu^2}{(1-2\mu)(\mu_c-\mu)}(1 + \tilde{\rho}_0 t^{2/3} + \tilde{\rho}_1 t), \tag{57}$$

where the coefficients $\tilde{\rho}_0$ and $\tilde{\rho}_1$ are determined by requiring that the leading behavior of r_c in Eq. (57) is same as that in Eq. (55). This yields

$$\tilde{\rho}_0 = \frac{3\mu_{c0}}{2s(\mu_{c0}-\mu)}\left\{\frac{2\mu^2 + \mu_{c0}(s-4\mu)}{\mu^2}a_t^{1/3} - (1-s)(2\mu_{c0})^{1/3}\right\}, \quad \tilde{\rho}_1 = 0, \tag{58}$$

where we have used a more accurate expression for μ_c than Eq. (41), derived in the Mathematica file, which reads

$$\mu_c = \mu_{c0} - \frac{3(1-s)}{4(2-s)}(2\mu_{c0}t^2)^{1/3} + \frac{2\mu_{c0}^2(1+s)}{s^2}t. \tag{59}$$

For completeness, we present the corresponding x_c^∞ which reads

$$x_c^\infty = \left(\frac{\mu_{c0}t^2}{4}\right)^{1/3} - \frac{t}{2}. \tag{60}$$

Figure 4 compares the exact values with the approximations Eqs. (55) and (57). s and t in Fig. 4a are same as those in Fig. 3b. For sufficiently small

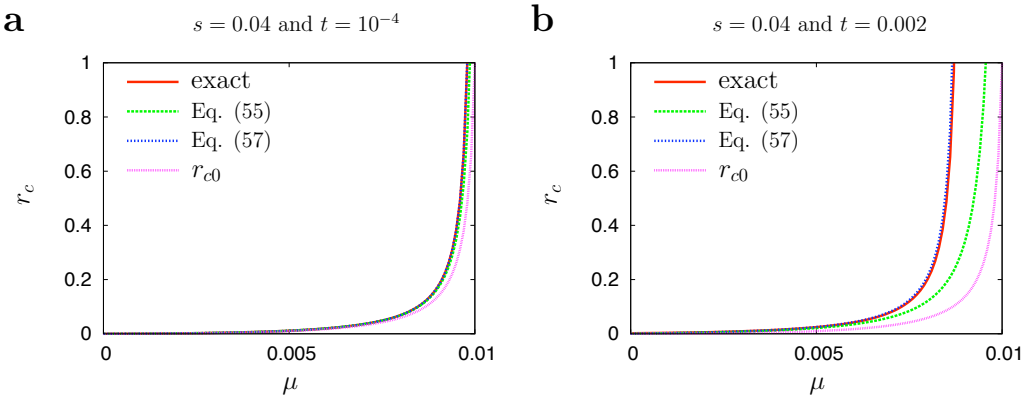

Figure 4: Comparison of the exact r_c with the approximate solutions Eqs. (55) and (57) for $s = 0.04$ and (a) $t = 10^{-4}$, (b) $t = 0.002$. For comparison, we also draw the critical recombination probability when $t = 0$ (r_{c0}). For $t = 10^{-4}$, both approximations are in good agreement with the exact solutions. As t increases, Eq. (55) starts to deviate from the exact solution, but the improved approximation still has predictive power.

t, both approximations show a good agreement. As anticipated, Eq. (55) becomes worse as t increases, even though Eq. (57) is still in good agreement with the exact solution. Needless to say, Eq. (57) fails when t is not much smaller than s. Although Fig. 4 seems to suggest that the agreement is good even for very small μ, this is an artifact because r_c itself is too small. In this regime, one should use the approximation developed in Sec. 3.3.

To summarize our findings up to now, we have provided approximate expressions for r_c which are valid in the specified regimes. Taken together, these expressions cover essentially the full range of biologically relevant parameters.

3.5. Frequency distributions

So far we have investigated the critical recombination and mutation probabilities. To complete the analysis, we need to determine the frequency distribution at stationarity. For the LFS, this can be readily done using Eq. (30). ¿From Eq.(28) we see that the solution with the smaller $x > 0$ will confer the larger value of f_0. Below we will argue that the solution with the larger f_0 is stable. Hence we limit ourselves to the study of the solution with smaller x. We also present approximate expressions for the HFS.

As before, we have to conduct separate analyses depending on which parameter is the smallest. For these analyses we will use the expansions

Eqs. (48) and (55) for r_c rather than the improved approximations Eqs. (49) and (57), which are not suited for a systematic perturbative solution.

We begin with the case when $\mu \ll s, t$. The functions $h''(x_c, r_c)$ and $h_1(x_c)$ in (30) then become

$$-h''(x_c, r_c) \approx 2t\beta,\ h_1(x_c) \approx a_\mu s^2 t\mu, \tag{61}$$

which gives

$$\varepsilon_x \approx -x_c \left(\frac{s^2 \varepsilon_r}{(1-t)\sqrt{\alpha\beta}\mu} \right)^{1/2} \equiv -x_c \epsilon, \tag{62}$$

where $x_c = a_\mu \mu$ and the definition of ϵ is clear. If we set $x = x_c + \varepsilon_x$ and $r = r_c + \varepsilon_r = t + c_\mu \mu + \varepsilon_r$ from (48), we get

$$A = 1 + \frac{t}{x} \approx \frac{t}{x_c(1-\epsilon)}, \quad B = \frac{\alpha}{t\left(\varepsilon_r - 2\varepsilon_x + (c_\mu - 2a_\mu)\mu\right)}, \tag{63}$$

which are large. Note that A becomes negative when $\epsilon > 1$, which means that the regime of validity of the above approximation is quite narrow. To leading order, the population frequencies are then obtained from (25) as

$$\begin{aligned} f &\approx \frac{1}{\sqrt{AB}} \approx \sqrt{\alpha}\frac{1-t}{s}\mu \left(1 + \sqrt{\frac{\alpha}{\beta}}(1-\epsilon)\right), \\ f_3 &\approx \frac{1}{A} \approx \frac{1-t}{t}\sqrt{\frac{\alpha}{\beta}}(1-\epsilon)\mu, \end{aligned} \tag{64}$$

and, by normalization, $f_0 = 1 - f_3 - 2f$. The second (unstable) solution is obtained by setting $\epsilon \mapsto -\epsilon$.

To find the HFS, we shift the variable x to $y = -(x+t)$ and look for a solution of $g(y) \equiv h(-t-y) = 0$. As shown in the Mathematica file, the HFS is located at $y = 0$ ($x = -t$) for $\mu \to 0$. This is a consequence of the fact that when $\mu = 0$, the stationary fitness of the HFS is $\bar{w} = w_3$. If we now set $y = \sum_{n\geq 1} y_n \mu^n$ and expand $g(y)$ as a power series in μ, the equation $g(y) = 0$ gives (see Mathematica file)

$$\begin{aligned} y_1 &= 0,\ y_2 = \frac{r(1-s-t)^2 + (s+t)(2-s-t)}{(s+t)^2(r(1-t)+t)}t, \\ y_3 &= 2y_2 \frac{t(s+t)^2(1-r) + r(2s+t)}{(r(1-t)+t)(s+t)^2} \end{aligned} \tag{65}$$

Up to $O(\mu^2)$, the genotype frequencies for the HFS become

$$f \approx \frac{\mu}{s+t}(1-y_2\mu), \quad f_0 \approx A \approx \frac{y_2}{t}\mu^2, \tag{66}$$

and $f_3 = 1-2f-f_0$.

One can see qualitative differences between Eqs. (66) and (64). First, the frequency of the less populated fitness peak genotype is different in the two cases. In Eq. (64), $f_3 = O(\mu)$, but in Eq. (66), $f_0 = O(\mu^2)$. However, this tendency cannot persist when r is large. For example, if $r = 1$, Eq. (19) suggests that f_3 should be $O(\mu^2)$ provided f is still $O(\mu)$. Hence, this qualitative difference only occurs when r is close to r_c. Second, the leading behavior of the frequency f of valley genotypes does not depend on r in Eq. (66), which is not true in Eq. (64) because of the dependence on ϵ.

To analyze the stability of the solutions, we linearize Eq. (13) at the steady state frequency. For the stability analysis, we assume that $f_1(\tau) = f_2(\tau)$ for all τ, which is true if they are equal initially. The linearization then yields a square matrix with rank 2, whose largest eigenvalue (in absolute value) determines the stability. For the HFS, the eigenvalues up to $O(\mu^0)$ are $1-s-t$ and $(1-t)(1-r)$ which are smaller than 1. Hence, the HFS is always stable. At $r = r_c$, the largest eigenvalue for the LFS is expected to be 1. Since we are restricted to an approximation up to $O(\mu)$, all we can show is that the largest eigenvalue of the LFS is $1+O(\mu^2)$ at $r = r_c$. In the Online Supplement, we show that the largest eigenvalue for $s = t$ becomes $1+O(\mu^2)$ and the smaller one is $(1-s-t)/(1-t)$. When treated numerically, it is easy to see that the stable solution indeed corresponds to the smaller x (details not shown).

The next step we will take is to find the frequency distribution of the LFS in the case that t is much smaller than s and μ. ¿From Eqs. (54) and (55), we get (up to leading order)

$$\begin{aligned} -h''(x_c, r_c) &= \frac{6s(2\mu^2 + \mu_{c0}(s-4\mu))}{\mu_{c0}-\mu}(a_t t^2)^{1/3}, \\ h_1(x_c) &= 2s(2-s)(1-2\mu)(\mu_{c0}-\mu)(a_t t^2)^{2/3}, \end{aligned} \tag{67}$$

thus from Eq.(30)

$$\varepsilon_x = -(\mu_{c0}-\mu)(a_t t^2)^{1/3}\left(\frac{2(2-s)(1-2\mu)}{3(2\mu^2+\mu_{c0}(s-4\mu))}\frac{\varepsilon_r}{(a_t t^2)^{1/3}}\right)^{1/2} \equiv -x_c\eta, \tag{68}$$

where η has an obvious meaning. Accordingly, A and B becomes

$$A \approx 1 + A_t, \tag{69}$$

$$\sqrt{B} \approx \frac{s-2\mu}{2\mu}\left(1 + B_t x_c + B_r \varepsilon_r + B_x \varepsilon_x\right), \tag{70}$$

where

$$A_t = \frac{t}{x_c}(1-\eta)^{-1}, \quad B_t = \frac{8\mu^2 - 4s\mu(1+\mu) - s^2(1-4\mu)}{8a_t(s^2+8\mu^2-4s\mu(1+\mu))},$$
$$B_r = -\frac{2s^2\mu^2}{\mu_{c0}(s-2\mu)^2 r_{c0}^2}, \quad B_x = \frac{s(1-2\mu)(s-4\mu)(\mu_{c0}-\mu)}{2(s-2\mu)^2\mu^2}. \tag{71}$$

The above approximation is valid only when $\eta \ll 1$ ($\varepsilon_r \ll t^{2/3}$). Note that unlike the previous case, A is close to 1 ($A_t \sim t^{1/3}$). Hence the frequency distribution for the LFS becomes

$$f \approx \frac{\mu}{s}\left(1 - 2\left(\frac{1}{2} - \frac{\mu}{s}\right)\left(B_t x_c + B_r \varepsilon_r + B_x \varepsilon_x + \frac{A_t^2}{8}\right)\right), \tag{72}$$

$$f_3 \approx \left(\frac{1}{2} - \frac{\mu}{s}\right)\left(1 - \frac{A_t}{2}\right), \tag{73}$$

$$f_0 \approx \left(\frac{1}{2} - \frac{\mu}{s}\right)\left(1 + \frac{A_t}{2}\right), \tag{74}$$

where we have kept the leading order of each term. Since the above approximation requires that $\varepsilon_r = r - r_c \ll t^{2/3}$, it cannot reproduce the symmetric solution in Appendix B, which applies when $t \to 0$ at fixed r.

3.6. Landau theory

In this final subsection, we develop an approximation that is valid when r is close to r_c and the asymmetry of the fitness landscape is small, in the sense that t is smaller than all other parameters. This approximation is inspired by the Landau theory from the physics of phase transitions, and it will allow us to represent both the LFS and the HFS in a simple, compact form.

We start from the observation that, according to Eqs. (72) ,(73), and (74), the valley genotype frequency $f \approx \mu/s$ in the regime of interest, with the peak frequencies f_3 and f_0 symmetrically placed around $1/2 - \mu/s \approx 1/2 - f$.

Moreover, the difference $f_0 - f_3 \approx A_t \sim t^{1/3}$ becomes small for $t \to 0$. This motivates the parametrization

$$f_0 = \left(\frac{1}{2} - f\right)(1-u)\,, f_3 = \left(\frac{1}{2} - f\right)(1+u) \tag{75}$$

which defines the new variable u. Inserting this into Eq. (14) with $p_\sigma = f_\sigma$ we obtain

$$\bar{w} = (1-2\mu)\left(1 + (1-u)\,\frac{t}{2u}\right). \tag{76}$$

On the other hand, from the definition (21) of $\bar{w}$, we find a relation between f and u such that

$$f = -\frac{t}{2u}\frac{(1-u)(1+u-2\mu)}{2s + t(1+u)} + \frac{2\mu}{2s + t(1+u)}. \tag{77}$$

Up to now, everything is exact. Note that when $u \ll 1$ and $t \ll u$, the leading behavior of Eq. (77) is μ/s which is consistent with the LFS frequency distribution in Eq. (72). Moreover, as the mean fitness for the HFS in the case of small t is not expected to deviate much from $1 - 2\mu$, the HFS also requires that $t \ll u$. So for all solutions, the leading behavior of f is μ/s. This is rather different from the case when μ is the smallest parameter.

By keeping the leading terms under the assumption that $t \ll \mu \ll s \ll 1$ and $t \ll u \ll 1$, from Eqs. (16) and (77) we obtain the equation

$$t - (r_0 - r)u - ru^3 = 0 \tag{78}$$

for u, where $r_0 = 8\mu^2/s$. If we interpret r as the (inverse) temperature, t as the external magnetic field, and u as the magnetization, this has precisely the form of the Landau equation for the para- to ferromagnetic phase transition (Plischke and Bergersen, 2006).

The general solution of Eq. (78) can be written in a compact form. Let

$$\Delta = \left(\frac{t}{2r}\right)^2 - \left(\frac{r - r_0}{3r}\right)^3 \tag{79}$$

denote the discriminant of Eq. (78). When $\Delta > 0$, there is only one real solution which reads

$$u_{\mathrm{HFS}} = \left(\frac{t}{2r} + \sqrt{\Delta}\right)^{1/3} + \left(\frac{t}{2r} - \sqrt{\Delta}\right)^{1/3}. \tag{80}$$

For r sufficiently far below r_0, in the sense that $r_0 - r \gg (t^2 r)^{2/3}$, this reduces to

$$u_{\rm HFS} \approx \frac{t}{r_0 - r}, \tag{81}$$

which is the solution of Eq. (78) with the cubic term omitted. When $\Delta < 0$, there are three real solutions

$$u_{\rm HFS} = 2\left(\frac{r - r_0}{3r}\right)^{1/2} \cos\frac{\theta}{3}, \quad u = -2\left(\frac{r - r_0}{3r}\right)^{1/2} \sin\left(\frac{\pi}{6} \mp \frac{\theta}{3}\right), \tag{82}$$

where $\tan\theta = 2r\sqrt{|\Delta|}/t$ with $0 \le \theta \le \pi/2$. The stable LFS corresponds to the smallest value of u,

$$u_{\rm LFS} = -2\left(\frac{r - r_0}{3r}\right)^{1/2} \sin\left(\frac{\pi}{6} + \frac{\theta}{3}\right). \tag{83}$$

One can easily see that for $t \to 0$ ($\theta \to \pi/2$) the solutions (82) and (83) approach the symmetric peak solutions

$$u_{\rm HFS} = \sqrt{1 - r_0/r}, \quad u_{\rm LFS} = -\sqrt{1 - r_0/r}. \tag{84}$$

The critical recombination probability can be found from $\Delta = 0$ which gives

$$r_c = 8\frac{\mu^2}{s}\left(1 + \frac{3}{4}\left(\frac{st}{2\mu^2}\right)^{2/3} + O(t^{4/3})\right). \tag{85}$$

Note that this agrees with Eq. (55) only up to order $t^{2/3}$.

Although the LFS in Eqs. (72), (73), and (74) is valid only when $\epsilon_r \ll t^{2/3}$, the approximate solutions Eqs. (80) and (82) turn out to be in good agreement with the exact solutions, provided r_0 is replaced by the exact critical recombination rate r_{c0} for the symmetric peak problem [see Eq.(56)]. In Fig. 5, we compare the exact values of u with the approximate solutions for $t = 10^{-6}$, $s = 10^{-2}$, and $\mu = 10^{-3}$. For these parameters, η becomes larger than 1 when $\varepsilon_r \approx 3 \times 10^{-5}$.

4. Discussion

In this work we have presented a detailed analysis of a deterministic, haploid two-locus model with a fitness landscape displaying reciprocal sign

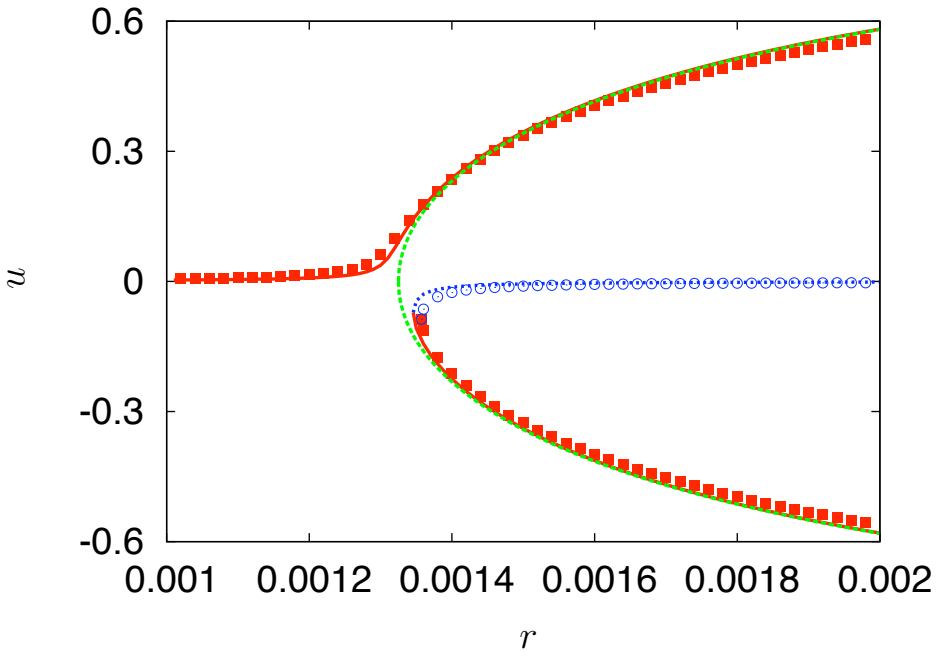

Figure 5: Plots of $u = (f_3 - f_0)/(1 - 2f)$ obtained from the exact numerical solutions (symbols) and from Eq. (78) (full curves) as a function of r for $t = 10^{-6}$, $s = 10^{-2}$, and $\mu = 10^{-3}$. For these parameters $r_0 = 8 \times 10^{-4}$ and $r_{c0} \approx 1.325 \times 10^{-3}$. The filled squares are stable solutions and open circles are unstable solutions. The dashed line shows the symmetric peak solutions (84) for $r > r_{c0}$. The approximate solution is seen to be valid well beyond the regime where $\eta < 1$.

epistasis. We have established the conditions for the occurrence of bistability, and derived accurate approximations for the critical recombination rate r_c at which bistability sets in. Specifically, Eqs. (48) and (55) are based on a systematic expansion of r_c in terms of the mutation probability μ and the peak asymmetry t, respectively, while the interpolation formulae Eq. (49) and Eq. (57) provide numerically accurate values of r_c over a wide range of parameters. In particular, our results show that the lower bound (24) on r_c becomes an equality for $\mu \to 0$, which is consistent with earlier results obtained either directly at $\mu = 0$ (Feldman, 1971; Rutschman, 1994) or using a unidirectional mutation scheme (Eshel and Feldman, 1970; Karlin and McGregor, 1971). Clearly the limiting behavior for $\mu \to 0$ should not depend on the mutation scheme employed. Approximate results for the stationary frequency distributions are found in Eqs. (64), (66) and Eqs. (72), (73), (74). Of particular interest are the simple formulae derived from the cubic equation (78), which are remarkably accurate close to the bistability threshold and for small t.

It would be of considerable interest to extend the present study to finite populations. In the presence of genetic drift true bistability cannot be maintained, and the focus is instead on the time scale on which a transition from the LFS to the HFS (or vice versa) is induced by stochastic fluctuations. For the case of symmetric peaks ($t = 0$) this question has been addressed by Higgs (1998) in the framework of a diffusion approximation. A key step in his analysis was the reduction to a one-dimensional problem by fixing the frequency of the valley genotypes at its stationary value $f = \mu/s$. However, we have seen above that in the case when s and t are large in comparison to μ, f cannot be treated as a constant and therefore the reduction to a one-dimensional diffusion equation is generally not possible. Some progress could be made in the regime where the (effectively one-dimensional) Landau equation (78) applies, and we intend to pursue this approach in the future.

Acknowledgments

We wish to thank to Alexander Altland, Reinhard Bürger, Arjan de Visser, Paul Higgs and Kavita Jain for useful discussions. Support by Deutsche Forschungsgemeinschaft within SFB 680 *Molecular Basis of Evolutionary Innovations* is gratefully acknowledged. In addition, J.K. acknowledges support by NSF under grant PHY05-51164 during a visit at KITP, Santa Barbara, where this work was begun.

A. The function $h(x)$ and its values at selected points

All information in this appendix can be found in the Mathematica file in the online supplementary material, and is provided here only for completeness.

By squaring both sides of Eq. (26), the steady state mean fitness becomes the solution of the equation $h(x) = 0$ where

$$\begin{aligned} &4(\bar{w} - w_1)^2 - \frac{B}{A}(w_3 + w_0 A - (1+A)\bar{w})^2 \\ &= \frac{4(1-t-x)h(x)}{x(t+x)((1-t-x)^2 - (1-r)(1-t))}. \end{aligned} \tag{86}$$

As shown in the Mathematica file, $h(x)$ takes the form $h_0(x) + r h_1(x)$ with

$$h_0(x) = b_4 x^4 + b_3 x^3 + b_2 x^2 + b_1 x - b_0, \tag{87}$$

$$h_1(x) = -(1-2\mu)^2 c_3 x^3 - (1-2\mu) c_2 x^2 + c_1 x - c_0, \tag{88}$$

where

$$\begin{aligned} b_4 &= (1-2\mu)(2s+t), \quad c_3 = 2s + t - (s+t)^2, \\ b_3 &= (t^2 + 2st - 2s^2)(1-2\mu) + \mu^2(4c_3 + t^2), \\ c_2 &= (t + 2\mu - 4t\mu)c_3 - s^2, \\ b_2 &= -3s^2 t(1-2\mu) - \mu^2\left[4(1-2t)c_3 + 3t^2(1-t)\right], \\ c_1 &= t(1-2\mu)(s^2 - 2\mu(1-t)c_3) + \mu^2 t^2 (1-s-t)^2, \\ b_1 &= -(1-2\mu)s^2 t^2 - \mu^2 t\left[(4-5t)c_3 - t(1-t)(2-3t)\right], \\ c_0 &= (1-t)(1-s-t)^2 t^2 \mu^2, \quad b_0 = (1-t)(2-s-2t)st^2\mu^2. \end{aligned} \tag{89}$$

Note that $b_4 > 0$ if $\mu < \frac{1}{2}$. The values of $h(x)$ at $x = 0$, $x = -t$, and $x = x_1 \equiv w_0 - w_1/(1-2\mu)$ are

$$h(0) = -b_0 - r c_0, \tag{90}$$

$$h(-t) = -t^2\mu^2\left[1 - (1-r)(1-s-t)^2\right], \tag{91}$$

$$\begin{aligned} h(x_1) &= -\frac{w_1}{(1-2\mu)^3}\left[1 - (1-r)(1-2\mu)^2\right] \times \\ &\quad \times (s(s+t)(1-\mu) - c_3\mu)^2, \end{aligned} \tag{92}$$

which are all negative if $0 < \mu < \frac{1}{2}$.

B. Solution for symmetric fitness peaks ($t = 0$)

When $t = 0$, our problem is reduced to that (approximately) solved by Higgs (1998). For the paper to be self-contained, we solve the case with $t = 0$ exactly in this appendix.

When $t = 0$, Eq. (17) suggests either $f_3 = f_0$ or $\bar{w} = 1 - 2\mu$. Let us first consider $\bar{w} = 1 - 2\mu$. Needless to say, this solution is impossible if $\mu \geq \frac{1}{2}$. Since $\bar{w} = (f_0 + f_3) + 2f(1 - s) = 1 - 2fs$, we get

$$f = \frac{\mu}{s}, \ f_0 + f_3 = 1 - 2\frac{\mu}{s}. \tag{93}$$

¿From Eq. (19) with $\bar{w} = 1 - 2\mu$, we get

$$(f_0 - f_3)^2 \ = \ (f_0 + f_3)^2 - 4f_0 f_3 = \frac{2}{rs}(1 - 2\mu)\xi, \tag{94}$$

where $\xi = (2 - s)(\mu_{c0} - \mu)(r - r_{c0})$ with r_{c0} and μ_{c0} given in Eqs.(51) and (56).

To have an asymmetric solution ($f_0 \neq f_3$), ξ should be nonnegative. Because $\mu > \mu_{c0}$ implies $r_{c0} < 0$, ξ is nonnegative only if $r \geq r_{c0}$ and $\mu < \mu_{c0}$ (note that when $\mu = \mu_{c0}$, ξ is nonzero and negative, because Eq. (56) diverges as $1/(\mu_{c0} - \mu)$ for $\mu \to \mu_{c0}$). Since r_{c0} cannot be larger than $\frac{1}{2}$ for the uniform crossover and 1 for the single crossover, the more restrictive condition on μ becomes $\mu < \mu_c$ where μ_c is the solution of the equation $r_{c0} = \frac{1}{2}$ or 1 given by

$$\mu_c = \begin{cases} \dfrac{s}{2(1 + \sqrt{1 + 2s - s^2})} & \text{for uniform crossover,} \\ \dfrac{s}{4} & \text{for single crossover.} \end{cases} \tag{95}$$

Note that μ_c is the same as μ_{c0} up to leading order of s.

Now let us find the solution with $f_3 = f_0$. ¿From $1 - 2fs = \bar{w}$ and Eq. (19) along with the substitution $\bar{w} = (1 - 2\mu)(1 + y)$, we obtain

$$g_s(y) \equiv y^2 + (r(1 - s) + s + \xi)\, y + \xi = 0. \tag{96}$$

Since

$$g_s(-s) = -(1 - s)(2 - s)\left[\frac{2\mu^2}{1 - 2\mu} + r(\mu + \mu_{c0})\right]$$

is negative, there is only one solution with $y > -s$ which is

$$y = \frac{-2\xi}{r(1-s)+s+\xi+((r(1-s)+s+\xi)^2-4\xi)^{1/2}}. \tag{97}$$

When $\xi < 0$ (either $\mu \geq \mu_{c0}$ or $\mu < \mu_{c0}$ together with $r < r_{c0}$), the larger solution is nonnegative. On the other hand, if $\xi > 0$ ($\mu < \mu_{c0}$ and $r > r_{c0}$), the larger solution which is still larger than $-s$ is negative.

References

Boerlijst, M. C., Bonhoeffer, S., Nowak, M. A., 1996. Viral quasi-species and recombination. Proc. R. Soc. Lond. Ser. B 263, 1577–1584.

Bonhoeffer, S., Chappey, C., Parkin, N. T., Whitcomb, J. M., Petropoulos, C. J., 2004. Evidence for positive epistasis in hiv-1. Science 306, 1547–1550.

Bürger, R., 1989. Linkage and the maintenance of heritable variation by mutation-selection balance. Genetics 121, 175–184.

Crow, J. F., Kimura, M., 1965. Evolution in sexual and asexual populations. Am. Nat. 99, 439–450.

de Visser, J. A. G. M., Elena, S. F., 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. Nature Reviews Genetics 8, 139–149.

de Visser, J. A. G. M., Hoekstra, R. H., van den Ende, H., 1997. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. Evolution 51, 1499–1505.

de Visser, J. A. G. M., Park, S.-C., Krug, J., 2009. Exploring the effect of sex on empirical fitness landscapes. Am. Nat. 174, S15–S30.

Elena, S. F., Lenski, R. E., 1997. Test of synergistic interactions among deleterious mutations in bacteria. Nature 390, 395–398.

Eshel, I., Feldman, M. W., 1970. On the evolutionary effect of recombination. Theor. Popul. Biol. 1, 88–100.

Feldman, M. W., 1971. Equilibrium studies of two locus haploid populations with recombination. Theor. Popul. Biol. 2, 299–318.

Fisher, R. A., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Higgs, P. G., 1998. Compensatory neutral mutations and the evolution of rna. Genetica 102/103, 91–101.

Jacobi, M. N., Nordahl, M., 2006. Quasispecies and recombination. Theor. Popul. Biol. 70, 479–485.

Karlin, S., McGregor, J., 1971. On mutation selection balance for two-locus haploid and diploid populations. Theor. Popul. Biol. 2, 60–70.

Kondrashov, A. S., 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336, 435–440.

Kouyos, R. D., Silander, O. K., Bonhoeffer, S., 2007. Epistasis between deleterious mutations and the evolution of recombination. Trends in Ecology and Evolution 22, 308–315.

Otto, S. P., 2009. The evolutionary enigma of sex. Am. Nat. 174, S1–S14.

Plischke, M., Bergersen, B., 2006. Equilibrium Statistical Physics. World Scientific, Singapore.

Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., Tans, S. J., 2007. Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445, 383–386.

Rutschman, D. A., 1994. Dynamics of the two-locus haploid model. Theor. Popul. Biol. 45, 167–176.

Stephan, W., 1996. The rate of compensatory mutation. Genetics 144, 419–426.

Sywerda, G., 1989. Uniform crossover in genetic algorithm. In: Schaffer, J. D. (Ed.), Proceedings of the third international conference on genetic algorithms. Morgan Kaufmann, San Francisco, pp. 2–9.

Weinreich, D. M., Watson, R. A., Chao, L., 2005. Perspective: Sign epistasis and genetic constraints on evolutionary trajectories. Evolution 59, 1165–1174.

Wright, A. H., Rowe, J. E., Stephens, C. R., Poli, R., 2003. Bistability in a gene pool GA with mutation. In: DeJong, K., Poli, R., Rowe, J. (Eds.), Foundations of Genetic Algorithms (FOGA-7). Morgan Kaufmann, San Francisco, pp. 63–80.

Wright, S., 1931. Evolution in mendelian populations. Genetics 16, 97–159.