Theoretical Population Biology 172 (2026) 48–55

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Sign epistasis extends the effects of balancing selection on genetic diversity

Ksenia A. Khudiakova [a,*], Nicholas H. Barton [a], Göran Arnqvist [b]

[a] Institute of Science and Technology Austria, Am Campus 1, Klosterneuburg, 3400, Austria
[b] Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, Uppsala, 752 36, Sweden

ARTICLE INFO

Dataset link: https://github.com/khudyakovaks/Sign-epistasis-extends-the-effects-of-balancing-selection-on-genetic-diversity

Keywords:
Sign epistasis
Balancing selection
Within-population genetic variation

ABSTRACT

Balancing selection, a form of selection that maintains genetic diversity, is difficult to detect, and the importance of balancing selection for the maintenance of genetic variation may be larger than often assumed. We model the possibility that the diversity-promoting effects of balancing selection extend to other loci that show sign epistasis with a locus under balancing selection. Rather than focusing on overdominance, as was done in previous efforts, we explore the effects of negative frequency dependence and show that this has important effects on the conditions under which the diversity-promoting effect of epistasis can occur in diploids. Our results show that not only recombination rate but also the dominance of sign epistasis are key parameters that determine the maintenance of polymorphism beyond the locus under direct balancing selection. We suggest that the effect we explore may play a significant role, especially when balancing selection acts on major effect loci.

1. Introduction

Efforts to intersect DNA sequencing data and population genetics theory with estimates of genetic variance suggest that the observed variability in fitness components cannot be maintained solely by mutation-selection balance (Barton and Keightley, 2002; Bonnet et al., 2022; Charlesworth, 2015). Genetic variation in life history traits is often sizable (Burch et al., 2024; Houle, 1992) and within-population polymorphism in major-effect variants such as inversions are common (Faria et al., 2019). These observations call for a better understanding of the processes that act to maintain substantial levels of non-neutral genetic diversity. One possibility is that the genome-wide effects of balancing selection are larger than typically assumed (Charlesworth, 2015; Kroymann and Mitchell-Olds, 2005). Balancing selection (BLS) is selection that acts to maintain variation, including both overdominance (i.e., heterozygote advantage) and negative frequency-dependent selection (NFDS).

Previous work has examined some of the conditions under which balancing selection may affect genetic diversity more broadly. For example, it is commonly recognized that balancing selection increases diversity across linked neutral sites (Charlesworth, 2006). Further, functional dependencies between loci under balancing selection and other sites in the genome, such that different alleles are favored on different genetic backgrounds, called sign epistasis, may play a significant role in this context. Indeed, models of haploid systems confirm that epistasis can strengthen the diversifying effects of balancing selection (Kelly, 2000; Kelly and Wade, 2000).

Sign epistasis represents inter-locus interactions between alleles such that the fitness effect of a given allele in one locus changes from beneficial to deleterious depending on the allelic state at one or more other loci (Fig. 1). Such epistatic interactions tend to generate complex fitness valleys, i.e., clusters of polygenic genotypes of low fitness, and fitness ridges or even distinct fitness peaks (Gavrilets, 2004; Saona et al., 2022; Poelwijk et al., 2011; Wolf et al., 2000). Advances in experimental techniques have highlighted a major role for this phenomenon in molecular evolution (Breen et al., 2012; Bank, 2022; Starr and Thornton, 2016). Moreover, epistasis has a central role in our understanding of incipient speciation, where evolving epistatic interactions within populations (Corbett-Detig et al., 2013; Ortiz-Barrientos et al., 2016) may result in reproductive isolation between populations through the build-up of epistatic Dobzhansky-Muller incompatibilities (Orr and Turelli, 2001). Yet, most evolutionary genetic models ignore interactions between loci (Hansen, 2013) and links between the role of epistasis in adaptation and its role in speciation, potentially connecting micro- and macroevolution, are not fully understood (Johnson, 2000; Wade, 2000). Although the contribution of physiological epistasis to additive genetic variance can be very substantial, epistasis is typically regarded as inconsequential for the short-term response to selection (Mäki-Tanila and Hill, 2014). Contrary to this view, epistasis can contribute to the long-term response to selection (Paixão and Barton, 2016; Bourg et al., 2024) and several experimental studies have unveiled an important role for interactions between loci in adaptation (Bloom et al., 2013; Good et al., 2017; Pettersson et al., 2011; Wolf et al., 2000).

* Corresponding author.
E-mail address: khudyakovaks@gmail.com (K.A. Khudiakova).

https://doi.org/10.1016/j.tpb.2026.09.001
Received 23 April 2026
Available online 10 September 2026
0040-5809/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Further, it has long been recognized that epistatic variance can be "converted" into additive genetic variance following, e.g., population bottlenecks (Barton and Turelli, 2004), and line-cross analyses have shown that the epistatic contribution to genetic variance in life history traits is often substantial (Burch et al., 2024; Roff and Emerson, 2006). However, our understanding of whether and how epistasis affects those processes that maintain allelic diversity in interacting loci within populations is relatively restricted (Hansen, 2013; Hemani et al., 2013).

Sign epistasis tends to generate co-adapted genotypes where either of the fit multilocus variants can reach highest fitness. In the absence of other forces, fit multilocus genotypes do not coexist stably but a natural extension is to ask if stable coexistence can result from the interaction between sign epistasis and balancing selection. Since recombination tends to disassociate favorable allelic combinations, the potential of epistasis to maintain genetic variation depends on the balance between the strength of selection and the rate of recombination. Hence, we expect two related general effects: (1) the efficacy of balancing and epistatic selection in maintaining variation should decrease with map distance, and (2) epistatic selection will lead to evolution of tighter linkage. Already Fisher (1930, p. 103) argued that sign epistasis can lead to decreased recombination (which, in turn, can lead to the stable coexistence of highly fit genotypes). A formal analysis was provided by Kimura (1956), who investigated the equilibria and stability of a deterministic two-locus model in which one locus exhibits heterozygote advantage and interacts with a second locus through sign epistasis. He showed that the maintenance of polymorphism at both loci is possible when linkage is sufficiently tight. Later work generalized Kimura's fitness table in a class of fitness tables called "symmetric viability models", i.e. fitness tables that are symmetric with respect to the fitness of the double heterozygote upon relabeling of loci and alleles (Bodmer and Felsenstein, 1967; Karlin and Feldman, 1970; Christiansen, 1999). Another inherent feature of the symmetric viability models is constant fitnesses, with balancing selection being modeled by heterozygote advantage. For all symmetric viability models, the conclusions remained similar to Kimura's: the diversity-promoting effects of sign epistasis are restricted to the cases when loci are tightly linked, which predicts that this effect may be of limited evolutionary significance.

However, overdominance is a specific mechanism for balancing selection and may be a rather restrictive one: first, it cannot occur in haploids and, second, the efficacy of balancing selection through overdominance is reduced by selfing (Glémin, 2021). In contrast, negative frequency dependent selection can act also in haploids, is often cited as the most widespread source of balancing selection (see recent review (Fraser and Hughes, 2026) and references therein), and its effects are much less affected by selfing (Glémin, 2021). Studying the role of negative frequency-dependent selection is also motivated by the fact that overdominance precludes other dominance relationships among loci. Since dominance of epistasis has been shown to be important in the dynamics of additive genetic variance (López-Fanjul et al., 2004; Barton and Turelli, 2004) and in the dynamics of speciation under DMIs (Sciuchetti et al., 2018), this further motivates the exploration of a model of negative frequency dependence, explicitly including the dominance of epistasis.

Here, we provide an exploration beyond symmetric viability models where we relax the assumption of constant fitnesses: instead of overdominance, we consider negative frequency dependence as the cause of balancing selection. We demonstrate that when epistatic effects show dominance, even free recombination allows for stable coexistence of polymorphic loci, challenging the long-standing view that tight linkage is required. Our results help clarify the conditions under which sign epistasis can contribute to the maintenance of within-population genetic diversity.

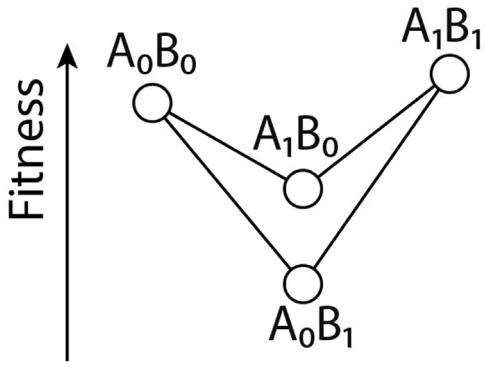


Fig. 1. Sign epistasis between two haploid biallelic loci under frequency independent selection.

Table 1
Haploid fitness table.

	B_0	B_1
A_0	$1 - tq_A$	$1 - tq_A - s$
A_1	$1 - tp_A - s$	$1 - tp_A$.

Table 2
Diploid fitness table.

	B_0B_0	B_1B_0	B_1B_1
A_0A_0	$1 - tq_A + s/2$	$1 - tq_A + \alpha$	$1 - tq_A - s/2$
A_1A_0	$1 - t/2$	$1 - t/2 + \alpha$	$1 - t/2$
A_1A_1	$1 - tp_A - s/2$	$1 - tp_A + \alpha$	$1 - tp_A + s/2$.

2. Materials and methods

We consider the deterministic dynamics of two biallelic loci under selection and recombination in an effectively infinite population. Locus A with alleles A_0 and A_1 is under symmetric negative frequency-dependent selection with selection coefficient t. Locus B with alleles B_0 and B_1 is in sign epistasis with locus A, which means that the genotypes containing a combination of conflicting alleles A_0 and B_1 or A_1 and B_0 experience an additional disadvantage proportional to the epistatic selection coefficient s. We denote the frequency of allele A_0 (resp. B_0) by q_A (q_B) and the frequency of allele A_1 (resp. B_1) by p_A (p_B).

For haploids, this results in the fitness table given by Table 1.

For diploids, two additional parameters that describe the dominance relationship within the two loci are possible. Here, we assume that allelic effects at the A locus are additive, such that the fitness of the heterozygote A_0A_1 is intermediate to the fitnesses of the two homozygotes. The dominance relationship at locus B is arbitrary, and we model it with the parameter α. When $\alpha = 0$, selection at locus B is additive. When $\alpha > 0$, this generates dominance such that the fitness of each B heterozygote is closer to the fitter two-locus homozygote (B_0 shows dominance in one background and B_1 shows dominance in the other). We term this dominance of the epistatic effect. Note that we consider $\alpha < s/2$, such that there is no direct overdominance at locus B. Nevertheless, as we will see later, $0 < \alpha < s/2$ can further stabilize double polymorphism. Thus, we consider the fitness table given by Table 2.

We assume random mating and no sex differences in fitness. Recombination rate between two loci is denoted by r, which means that proportion r of gametes in the new generation will contain combinations of alleles different from those of the parents.

We study the dynamics analytically using the multilocus algebra introduced in Barton and Turelli (1991) and Kirkpatrick et al. (2002).

3. Results

3.1. The haploid case

The haploid dynamics are described by the following set of equations:

$$\begin{aligned}
\Delta p_A &= a_A p_A q_A + a_B D + a_{AB}\left(q_A - p_A\right) D \\
\Delta p_B &= a_A D + a_B p_B q_B + a_{AB}\left(q_B - p_B\right) D \\
\Delta D &= (1-r)\Big(a_A\left(q_A - p_A\right) D + a_B\left(q_B - p_B\right) D + \\
&\quad a_{AB}\left(p_A q_A p_B q_B + \left(q_A - p_A\right)\left(q_B - p_B\right) D - D^2\right) - \Delta p_A \Delta p_B \Big) - rD,
\end{aligned} \tag{1}$$

where

$$\begin{aligned}
\bar{W} a_A &= t(q_A - p_A) - s(q_B - p_B) \\
\bar{W} a_B &= -s(q_A - p_A) \\
\bar{W} a_{AB} &= 2s \\
\bar{W} &= 1 + 2Ds - p_B s - p_A(q_B - p_B)s - t(1 - 2p_A q_A).
\end{aligned} \tag{2}$$

An example of the dynamics is shown on fig. 4a, SM. The system (1) admits a double-polymorphic equilibrium with $p_A = p_B = 1/2$ and the value of linkage disequilibrium

$$D_{eq} = \frac{1}{4}\left(\sqrt{\left(\frac{rB}{2s}\right)^2 + (1-r)} - \frac{rB}{2s}\right), \tag{3}$$

where $B = (2 - s - t)$.

This equilibrium is stable when

$$r < 2t\left(\frac{2 - s - \sqrt{s^2 + t^2}}{4 - 4s - t^2}\right). \tag{4}$$

Assuming weak selection, the stability condition reduces to $r \lesssim t$.

Numerical analyses of the selection equations confirmed these results (figs. 4b and 5, SM).

3.2. The diploid case

The diploid dynamics are described by the following set of equations.

$$\begin{aligned}
\Delta p_A &= a_A p_A q_A + a_B D + a_{AB}(q_A - p_A)D \\
\Delta p_B &= a_A D + a_B p_B q_B + a_{AB}(q_B - p_B)D,
\end{aligned} \tag{5}$$

where

$$\begin{aligned}
\bar{W} a_A &= \frac{t}{2}(q_A - p_A) - \frac{s}{2}(q_B - p_B) \\
\bar{W} a_B &= -\frac{s}{2}(q_A - p_A) + \alpha(q_B - p_B) \\
\bar{W} a_{A,B} &= \bar{W} a_{AB,\varnothing} = \frac{s}{2} \\
\bar{W} a_{BB} &= -2\alpha \\
\bar{W} &= 1 + Ds + \frac{s}{2}\left(q_A - p_A\right)(q_B - p_B) - t(1 - 2p_A q_A) + 2\alpha p_B q_B.
\end{aligned} \tag{6}$$

Note that in the selection coefficients, the terms involving s and t are half of those in the haploid case.

For linkage disequilibrium in diploids, we distinguish two types of statistical associations: those on the same chromosome will be denoted $D_{AB,\varnothing}$, and those on different chromosomes will be denoted $D_{A,B}$. Epistatic selection builds up $D_{A,B}$, while each round of free recombination brings $D_{A,B}$ towards zero. The total change in D is given by

$$\Delta D = (1-r)D^{sel}_{AB,\varnothing} + rD^{sel}_{A,B} - \Delta p_A \Delta p_B - D, \tag{7}$$

Table 3
Diploid fitness table following Bodmer and Felsenstein (1967).

	B_0B_0	B_1B_0	B_1B_1
A_0A_0	$1 - \frac{\alpha - s/2}{1 - t/2 + \alpha}$	1	$1 - \frac{\alpha + s/2}{1 - t/2 + \alpha}$
A_1A_0	$1 - \frac{\alpha}{1 - t/2 + \alpha}$	1	$1 - \frac{\alpha}{1 - t/2 + \alpha}$
A_1A_1	$1 - \frac{\alpha + s/2}{1 - t/2 + \alpha}$	1	$1 - \frac{\alpha - s/2}{1 - t/2 + \alpha}$.

where

$$\begin{aligned}
D^{sel}_{AB,\varnothing} &= D + a_A(q_A - p_A)D + a_B(q_B - p_B)D + a_{AB,\varnothing} \\
&\quad \left(p_A q_A p_B q_B + D(q_A - p_A)(q_B - p_B) - D^2\right) \\
D^{sel}_{A,B} &= a_{A,B} p_A q_A p_B q_B + a_{A,B} D^2 + a_{BB} p_B q_B D.
\end{aligned} \tag{8}$$

The system ((5), (7)) admits double-polymorphic equilibrium ($p_A = p_B = 1/2$) with

$$D_{eq} = \frac{1}{4}\left(\sqrt{\left(\frac{2rB}{s}\right)^2 + 1} - \frac{2rB}{s}\right), \tag{9}$$

where $B = (2 - t + 2\alpha)$.

With $s < t$ and $\alpha > 0$, the double-polymorphic equilibrium is stable whenever

$$r < \frac{s^2}{16\alpha t}\,\frac{\left(s\sqrt{(4t\alpha)^2 + (s(t-2\alpha))^2} - (t + 2\alpha)(s^2 - 4t\alpha)\right)}{(2 - t + 2\alpha)(s^2 - 2t\alpha)}, \tag{10}$$

and

$$r < \frac{t^2 - s^2}{4t(2-t)}, \tag{11}$$

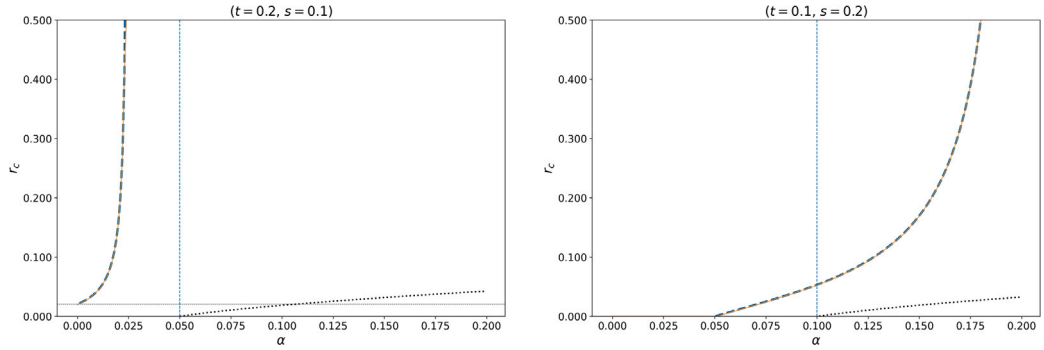
when $\alpha = 0$. The analytical results match the numerical analysis of selection equations exactly (figs. 6 and 7, SM). Thus, as in the haploid case, the condition for a double-polymorphic equilibrium in diploids under $\alpha = 0$ requires that recombination is low relative to selection. When $s > t$ such that epistatic selection is stronger than balancing selection, the double-polymorphic equilibrium is stable under condition (10) and only if α is large enough (Fig. 2b). Increasing α increases the amount of recombination that can be tolerated without destabilizing the double-polymorphic equilibrium (Fig. 2).

The increase of the parameter range with a stable double-polymorphic equilibria with an increasing α has an intuitive interpretation. When locus A is at equilibrium ($p_A = q_A = 1/2$), the genotypes $A_0A_0B_0B_0$ and $A_1A_1B_1B_1$ are the most-fit ones. However, due to their coexistence in a large population and random mating, less fit heterozygote A_1A_0 genotypes will be generated. Among those, $A_1A_0B_1B_0$ are the fittest whenever $\alpha > 0$. Therefore, the reason for the stabilizing effect of α is essentially selection in favor of double heterozygotes that manifests itself when locus A is at or close to equilibrium.

The general conclusion from previous works, which focused on the case of overdominance, is that linkage must be tight for a double polymorphic equilibrium to be stable. We can compare our findings with previous results at equilibrium for the diploid case (the focus in previous work) by setting $p_A = q_A = 1/2$. Indeed, we recover our results for D_{eq} (Eq. (9)) in Kimura (1956) if we set $\alpha = 0$, or in Karlin and Feldman (1970) and Bodmer and Felsenstein (1967) for any α.

To compare stability criteria, we will rely on the symmetric viability model studied in Bodmer and Felsenstein (1967), as it appears to be the most general treatment relevant to the case of sign epistasis. To draw a direct comparison with Table 2, we will parameterize the overdominance fitness table as detailed in Table 3.

We infer r_c for the model of overdominance by numerically checking which highest recombination rate satisfies conditions given by eqs. 27 and 28 in Bodmer and Felsenstein (1967). The results are shown in Fig. 2. As before, $\alpha > s/2$ corresponds to the overdominance at locus



(a) $s < t$ (b) $s > t$

Fig. 2. The recombination rate that can be tolerated without destabilizing the double-polymorphic equilibrium increases with the dominance of epistatic effects α. The figure shows critical r, below which the double polymorphic equilibrium is stable, against α in cases where (a) balancing selection is stronger than epistatic selection and (b) epistatic selection is stronger than balancing selection. Solid orange and dashed blue lines represent respectively numerical analyses and analytical results for the case of negative frequency-dependent selection analyzed here. Dotted line represents numerical evaluation of the conditions given by Bodmer and Felsenstein (1967). The vertical dashed line shows the value of α that marks the onset of direct overdominance at locus B. When balancing selection is stronger than epistatic selection and $\alpha = 0$, $r_c = (t^2 - s^2)/(4t(2-t))$ (black line). When epistatic selection is stronger than balancing selection, double-polymorphic equilibrium is not stable under low values of α.

B, but since this is now the only source of balancing selection, it is necessary for the stability of the double-polymorphic equilibrium. Even when the overdominance condition is satisfied, stability always requires tight linkage, contrary to the case of negative frequency-dependent selection. Therefore, the main difference with the previous modeling results appears in the stability: criteria identified in our present model of negative frequency-dependent selection are much more permissive than in previous models based on overdominance, where the stability condition for the internal equilibrium required recombination rate to be weaker than balancing selection (Kimura, 1956; Karlin and Feldman, 1970; Bodmer and Felsenstein, 1967).

4. Discussion

Whether sign epistasis, a form of interaction between alleles across loci where certain combinations confer high fitness, can promote genetic diversity within populations is a long-standing question. A necessary condition for this effect is the existence of stable polymorphism in a region under balancing selection. Because recombination tends to break up associations between favorable combinations of alleles (Hansen, 2013), recombination tends to decrease the stability of double-polymorphic equilibria and the general conclusion from the analysis of symmetric viability models is that linkage must be tight for double polymorphic equilibrium to be stable ($r \lesssim t$) (Kimura, 1956; Karlin and Feldman, 1970; Bodmer and Felsenstein, 1967).

We recovered this conclusion for the haploid case, but this differs dramatically from what we observe in diploids where the parameter α tunes between additive×additive (when $\alpha = 0$) and additive×dominant (when $\alpha > 0$) effects in the fitness table. An intuitive way to illustrate the effect of α is to note that when $\alpha > 0$, the fitness of the heterozygote for the B locus is closer to B_1B_1 in a A_1A_1 homozygote background and closer to B_0B_0 in a A_0A_0 homozygote background (Fig. 3). This illustrates why we expect $\alpha > 0$ to stabilize double polymorphism. The same stabilizing effect can also occur in general symmetric viability models (Bodmer and Felsenstein, 1967; Karlin and Feldman, 1970), where there is enough flexibility in the parameters to allow for arbitrary dominance relationships within loci. However, in symmetric viability models, the dominance of epistasis does not increase the critical recombination rate as much as in the model with frequency-dependent selection, and it seems that the critical importance of epistatic dominance was therefore overlooked in previous contributions. The increase in the range of parameter values where the internal equilibrium is stable under NFDS compared to overdominance is qualitatively similar to the effect described in Glémin (2021): frequency-dependent models of balancing selection are much less sensitive to the effects of inbreeding than are balancing selection models based on frequency-independent selection. In this sense, our study reinforces the point that NFDS is a more robust way of maintaining polymorphism than overdominance.

By analogy with a general discussion of context-specific dominance effects (Connallon and Chenoweth, 2019; Curtsinger et al., 1994), we may refer to the case $\alpha > 0$ as a beneficial reversal of dominance of epistasis because the fitness of the B heterozygote will be similar to the most fit homozygote across different genetic backgrounds. We note that this is in spirit analogous to dominance reversals across other contexts (e.g., environments, sexes, etc.), which are known to promote stable polymorphism in single locus scenarios (Connallon and Chenoweth, 2019). Drawing on parallels in single locus scenarios (Siljestam et al., 2024; Spencer and Priest, 2016), we suggest that selection should favor dominance of sign epistasis since it effectively reduces segregation load. Because the process we describe here is predicted to extend the life span of alleles across interacting loci, it may provide sufficient time to allow the evolution of elevated dominance of epistasis which, in turn, will tend to further stabilize double-polymorphism. We note that context-specific dominance effects can emerge from a variety of non-linearities at any level of the genotype-to-phenotype map (Billiard et al., 2021).

Sources of epistasis. The scenario studied here can be seen as representing a variety of mechanisms for the emergence of sign epistasis. This can appear on a genetic level in the form of co-adapted amino acids or nucleotide substitutions (Kondrashov et al., 2002) or more generally as co-adapted gene complexes (Dobzhansky and Pavlovsky, 1958). A general possibility is the emergence of sign epistasis for fitness as a result of non-linear fitness functions for polygenic traits (Fig. 3). For example, it has been suggested that major effect life history genes may frequently experience balancing selection (Arnqvist and Rowe, 2023; Christie and McNickle, 2023) and the processes we describe may act to maintain variation in loci that show epistasis with such genes. There are many potential examples of the scenario we explore (Arnqvist and Rowe, 2023; Christie and McNickle, 2023). A common case may be inversion polymorphism maintained by balancing selection

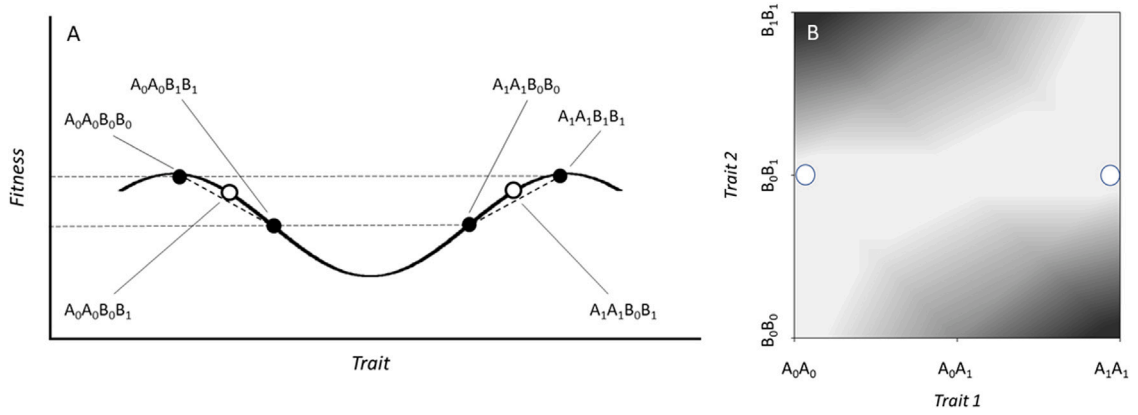


Fig. 3. For polygenic traits under non-linear selection, epistatic interactions between loci are inevitable, and curved fitness landscapes will then lead to complex allelic effects across loci. Shown are two schematic and simplified illustrations of ways in which dominance of sign epistasis could emerge. In these, genotype determines individual trait values which in turn translate into individual fitness. Under balancing selection, the relation between fitness and trait depends on allele frequencies. To simplify illustration, the frequencies at locus A are here fixed at $p_A = q_A = 1/2$. (A) When two loci have additive effects on a single phenotypic trait with two distinct fitness peaks (i.e., non-linear disruptive selection), the fitness of the heterozygotes for the B locus (open circles) may be closer to the homozygote B_0B_0 in the A_0A_0 background and closer to B_1B_1 in the A_1A_1 background, such that the dominance of the two B alleles is reversed across A genetic backgrounds (A_0A_1 heterozygotes not included here, for simplicity). (B) Alternatively, for two loci coding for distinct traits, correlational selection will also result in epistasis for fitness (light - high relative fitness, dark - low relative fitness). This may also reverse the dominance of the two B alleles across A genetic backgrounds as a result of the curvature of the fitness landscape. In this example, the fitness of the heterozygotes for the B locus (open circles) is closer to the homozygote B_0B_0 in the A_0A_0 background and closer to B_1B_1 in the A_1A_1 background.

(Berdan et al., 2022), where linkage disequilibrium between the inversion and other loci are commonly observed and actually constituted a very early demonstration of epistatic selection (Prakash and Lewontin, 1968). More specific examples include a polymorphic major effect life history locus located on the sex-chromosomes in *Drosophila mercatorum*, which interacts with several autosomal modifiers all of which are highly polymorphic in natural population (Templeton, 2000). A more familiar case is color pattern in male guppies, which shows high levels of standing genetic variation maintained by negative frequency-dependent selection. New work suggests that although a major effect locus is located on the Y-chromosome, color pattern is determined by an epistatic interaction between Y and an autosomal region in LD with the Y (Paris et al., 2022). Similarly, major effect variants that affect coloration in butterfly mimicry systems maintained by balancing selection have epistatic modifiers that apparently fine-tune coloration (Nijhout, 1991), and balancing selection on inter-locus allele combinations is likely important in maintaining polymorphism in flower morphology in plants (Kelly, 2000). A distinct but classic example may in fact be sickle-cell anemia in humans, which is controlled by the major effect HBB locus well known to be under balancing selection. It is perhaps less well known that the phenotypic manifestation of HBB genotype is determined by epistatic interactions with several other polymorphic loci (Templeton, 2000). Another illustrative example comes from supergenes (chromosomal inversions) that control social behavior in some ants and are maintained as balanced polymorphisms (Purcell and Brelsford, 2025). In the ant *Formica cinerea*, a second supergene on a different chromosome influences body size and is likewise polymorphic. Despite being unlinked, the two supergenes show strong linkage disequilibrium, which appears to be maintained by strong epistatic selection (Scarparo et al., 2026). Finally, an analogous scenario is provided by loci under sexually antagonistic selection, where different alleles are favored in males and females (Bonduriansky and Chenoweth, 2009). Such loci essentially exhibit sign epistasis for fitness with sex-determining loci under balancing selection, which can act to maintain variation in sexually antagonistic loci given permissive patterns of dominance of epistasis (Siljestam et al., 2024).

Possible scenarios for the accumulation of epistatic polymorphisms. Suppose that we start with locus A under balancing selection. Alleles at other loci favored in one genetic background but not the other may then accumulate if linkage is tight enough. This will generate indirect selection for reduced recombination rate (Otto and Feldman, 1997), a scenario analogous to the canonical model of sex chromosome evolution. Indeed, our analysis of the haploid case shows that the stability of a double polymorphism requires $r \lesssim t$.

However, we show that in diploids, the situation is significantly different because of possible dominance effects. Given balancing selection on locus A, there are at least two scenarios where these sorts of effects could be important. First, in cases where epistatic balancing selection maintains two-locus polymorphism for long enough, selection will favor the evolution of "beneficial" dominance effects (i.e., selection will favor the evolution of larger α), with effects similar to reduced recombination. This will, in turn, further stabilize the polymorphism. Second, as new mutations with epistatic effects enter into a population, those with permissive epistatic dominance effects (i.e., $\alpha > 0$) will tend to be retained by epistatic balancing selection. This way, selection may act as a sieve favoring the maintenance of, or increased lifespan of, mutations showing overall $\alpha > 0$. In both cases, we would expect the build-up of genetic polymorphism and $\alpha > 0$ to be common. Needless to say, the build-up of genetic polymorphism will be generally promoted by linkage, but as we show here, dominance effects make close linkage a less critical requirement.

How to detect the effect in real data? Uncovering the effects we discuss here using genome scans or population genomic data alone will clearly be very challenging (Arnqvist and Rowe, 2023). Co-adaptation of alleles at different loci should lead to positive linkage disequilibrium between them and the tools available for identifying regions under balancing selection are improving (Bitarello et al., 2023). Scanning for loci under balancing selection and inspecting regions or sites in positive LD with those may therefore be rewarding. For example, Stolyarova et al. (2022) found pervasive elevated short-range LD between nonsynonymous sites within divergent intra-genic haploblocks in the fungus *Schizophyllum communae*. This finding suggests that this may be

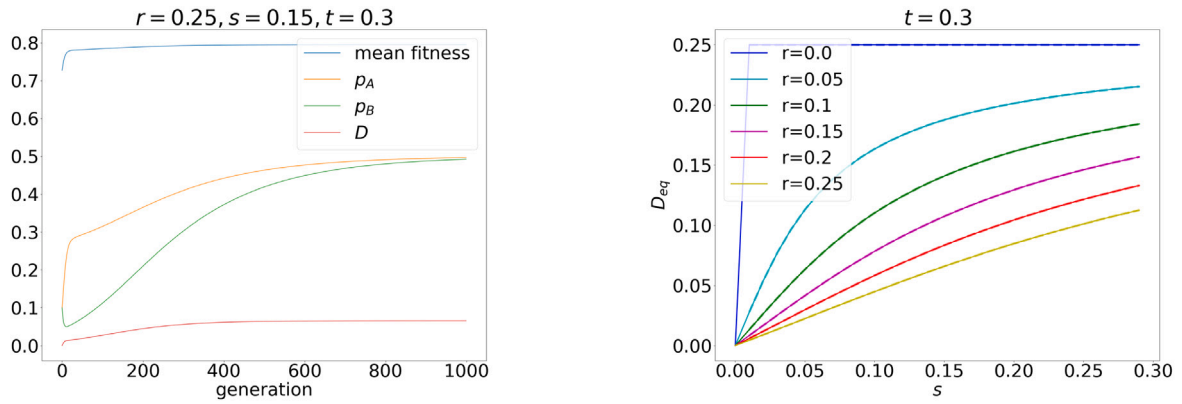


Fig. 4. (a) Change of p_A, p_B, D and mean fitness over time and (b) the relationship between the strength of epistatic selection s and linkage disequilibrium D. Solid lines represent numerical iteration of selection equations and dashed lines analytical prediction (3).

due to a combination of balancing and epistatic selection, potentially similar to the scenario we model here. Another possibility may be to empirically estimate the dominance of epistasis (i.e., α), through for example advanced line-cross analyses (Burch et al., 2024), as this is expected to be positive under stable double-polymorphism. Genomic time-series data may also offer interesting routes to assess epistatic selection, as manifested by synchronous frequency shifts in populations.

In summary, we delineate the conditions under which the diversity-promoting effect of balancing selection is expected to extend to other loci through epistasis. We find that the dominance of epistasis is central in this context, and a diversifying effect is predicted even under fairly high rates of recombination if this dominance is sizable.

CRediT authorship contribution statement

Ksenia A. Khudiakova: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Nicholas H. Barton:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Göran Arnqvist:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization.

Funding

This work was funded by grants from the Swedish Research Council (2023-03730 to G.A.) and the DOC fellowship from the Austrian Academy of Science (26293 to K.K.).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank L. Rowe for valuable discussions.

Appendix A. Haploids

See Figs. 4 and 5.

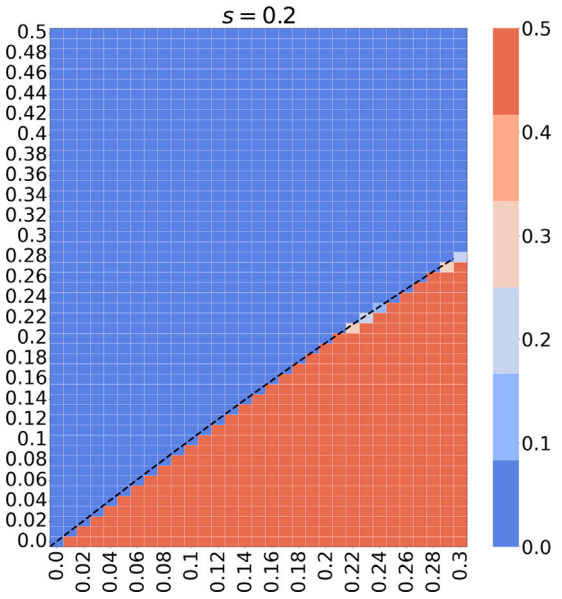


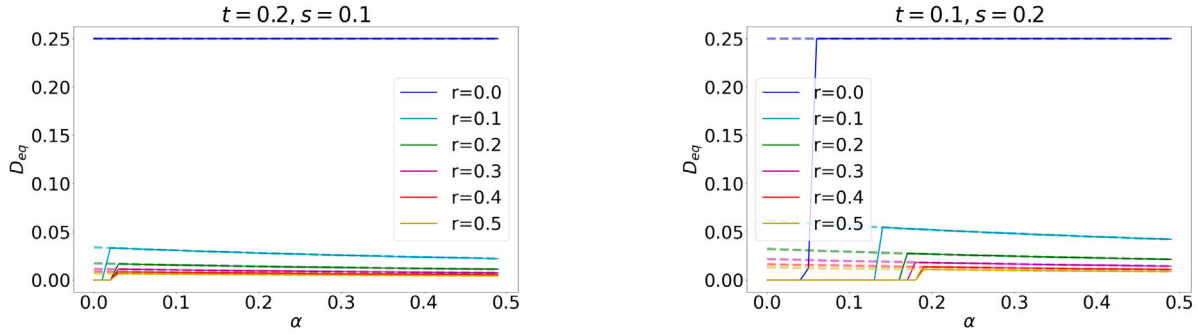
Fig. 5. Equilibrium frequency of allele B_1 in the numerical iteration of selection equations on the (r, t) space. Dashed line represents stability condition (4).

Appendix B. Diploids

See Figs. 6 and 7.

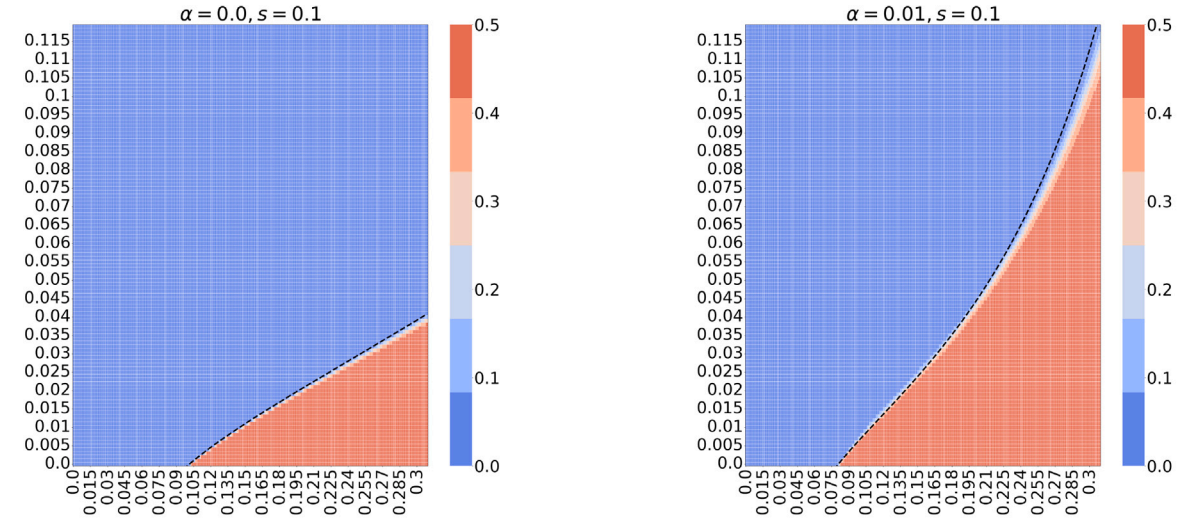
Data availability

The Python code used for the numerical analyses is available online at https://github.com/khudyakovaks/Sign-epistasis-extends-the-effects-of-balancing-selection-on-genetic-diversity.



(a) $s < t$ (b) $s > t$

Fig. 6. Linkage disequilibrium D at double-polymorphic equilibria decreases with increasing recombination rate, but is contingent upon the strength of selection and epistatic dominance α. Conditions where (a) balancing selection is stronger than epistatic selection and (b) epistatic selection is stronger than balancing selection. Solid lines represent numerical analysis and dashed lines are analytical prediction (9).



(a) $\alpha = 0$ (b) $\alpha > 0$

Fig. 7. Equilibrium p_B in the iteration of selection equations on the (r, t) space. The dashed lines are critical r below which the double polymorphism is predicted to be stable: $r_c = (t^2 - s^2)/(8t - 4t^2)$ when $\alpha = 0$ and as in (10) when $\alpha > 0$.

References

Arnqvist, G., Rowe, L., 2023. Ecology, the pace-of-life, epistatic selection and the maintenance of genetic variation in life-history genes. Mol. Ecol. 32, 4713–4724. http://dx.doi.org/10.1111/mec.17062.

Bank, C., 2022. Epistasis and adaptation on fitness landscapes. Annu. Rev. Ecol. Evol. Syst. 53, 457–479. http://dx.doi.org/10.1146/annurev-ecolsys-102320-112153.

Barton, N.H., Keightley, P.D., 2002. Understanding quantitative genetic variation. Nature Rev. Genet. 3, 11–21.

Barton, N.H., Turelli, M., 1991. Natural and sexual selection on many loci.. Genetics 127, 229–255. http://dx.doi.org/10.1093/genetics/127.1.229.

Barton, N.H., Turelli, M., 2004. Effects of genetic drift on variance components under a general model of epistasis. Evolution 58, 2111–2132. http://dx.doi.org/10.1111/j.0014-3820.2004.tb01591.x.

Berdan, E.L., Flatt, T., Kozak, G.M., Lotterhos, K.E., Wielstra, B., 2022. Genomic architecture of supergenes: connecting form and function. Philos. Trans. R. Soc. B 377, 20210192.

Billiard, S., Castric, V., Llaurens, V., 2021. The integrative biology of genetic dominance. Biol. Rev. 96, 2925–2942.

Bitarello, B.D., Brandt, D.Y.C., Meyer, D., Andrés, A.M., 2023. Inferring balancing selection from genome-scale data. Genome Biol. Evol. 15, http://dx.doi.org/10.1093/gbe/evad032.

Bloom, J.S., Ehrenreich, I.M., Loo, W.T., Lite, T.L.V., Kruglyak, L., 2013. Finding the sources of missing heritability in a yeast cross. Nature 494, 234–237.

Bodmer, W.F., Felsenstein, J., 1967. Linkage and selection: Theoretical analysis of the deterministic two locus random mating model. Genetics 57, 237–265. http://dx.doi.org/10.1093/genetics/57.2.237.

Bonduriansky, R., Chenoweth, S.F., 2009. Intralocus sexual conflict. Trends Ecol. Evolut. 24, 280–288. http://dx.doi.org/10.1016/j.tree.2008.12.005.

Bonnet, T., Morrissey, M.B., de Villemereuil, P., Alberts, S.C., Arcese, P., Bailey, L.D., Boutin, S., Brekke, P., Brent, L.J.N., Camenisch, G., Charmantier, A., Clutton-Brock, T.H., Cockburn, A., Coltman, D.W., Courtiol, A., Davidian, E., Evans, S.R., Ewen, J.G., Festa-Bianchet, M., de Franceschi, C., Gustafsson, L., Höner, O.P., Houslay, T.M., Keller, L.F., Manser, M., McAdam, A.G., McLean, E., Nietlisbach, P., Osmond, H.L., Pemberton, J.M., Postma, E., Reid, J.M., Rutschmann, A., Santure, A.W., Sheldon, B.C., Slate, J., Teplitsky, C., Visser, M.E., Wachter, B., Kruuk, L.E.B., 2022. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. Science 376, 1012–1016. http://dx.doi.org/10.1126/science.abk0853, arXiv:https://www.science.org/doi/pdf/10.1126/science.abk0853.

Bourg, S., Bolstad, G.H., Griffin, D.V., Pélabon, C., Hansen, T.F., 2024. Directional epistasis is common in morphological divergence. Evolution 78, 934–950. http://dx.doi.org/10.1093/evolut/qpae029.

Breen, M.S., Kemena, C., Vlasov, P.K., Notredame, C., Kondrashov, F.A., 2012. Epistasis as the primary factor in molecular evolution. Nature 490, 535–538. http://dx.doi.org/10.1038/nature11510.

Burch, J., Chin, M., Fontenot, B.E., Mandal, S., McKnight, T.D., Demuth, J.P., Blackmon, H., 2024. Wright was right: leveraging old data and new methods to illustrate the critical role of epistasis in genetics and evolution. Evolution 78, 624–634. http://dx.doi.org/10.1093/evolut/qpae003.

Charlesworth, D., 2006. Balancing selection and its effects on sequences in nearby genome regions. PLoS GEnetics 2, e64.

Charlesworth, B., 2015. Causes of natural variation in fitness: evidence from studies of drosophila populations. Proc. Natl. Acad. Sci. 112, 1662–1669.

Christiansen, F.B., 1999. Population Genetics of Multiple Loci. In: Wiley Series in Mathematical and Computational Biology, John Wiley & Sons, Chichester, UK; New York, NY.

Christie, M.R., McNickle, G.G., 2023. Negative frequency dependent selection unites ecology and evolution. Ecol. Evol. 13, http://dx.doi.org/10.1002/ece3.10327.

Connallon, T., Chenoweth, S.F., 2019. Dominance reversals and the maintenance of genetic variation for fitness. PLOS Biology 17, e3000118. http://dx.doi.org/10.1371/journal.pbio.3000118.

Corbett-Detig, R.B., Zhou, J., Clark, A.G., Hartl, D.L., Ayroles, J.F., 2013. Genetic incompatibilities are widespread within species. Nature 504, 135–137.

Curtsinger, J.W., Service, P.M., Prout, T., 1994. Antagonistic pleiotropy, reversal of dominance, and genetic polymorphism. Amer. Nat. 144, 210–228.

Dobzhansky, T., Pavlovsky, O., 1958. Interracial hybridization and breakdown of coadapted gene complexes in drosophila Paulistorum and drosophila willistoni. Proc. Natl. Acad. Sci. 44, 622–629. http://dx.doi.org/10.1073/pnas.44.6.622.

Faria, R., Johannesson, K., Butlin, R.K., Westram, A.M., 2019. Evolving inversions. Trends Ecol. Evolut. 34, 239–248.

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

Fraser, B.A., Hughes, K.A., 2026. Exploring negative frequency-dependent selection across levels: from genetics to ecology and back again. Philos. Trans. B 381, http://dx.doi.org/10.1098/rstb.2025.0169.

Gavrilets, S., 2004. Fitness Landscapes and the Origin of Species. Princeton University Press.

Glémin, S., 2021. Balancing selection in self-fertilizing populations. Evolution 75, 1011–1029. http://dx.doi.org/10.1111/evo.14194.

Good, B.H., McDonald, M.J., Barrick, J.E., Lenski, R.E., Desai, M.M., 2017. The dynamics of molecular evolution over 60,000 generations. Nature 551, 45–50.

Hansen, T.F., 2013. Why epistasis is important for selection and adaptation: Perspective. Evolution 67, 3501–3511. http://dx.doi.org/10.1111/evo.12214.

Hemani, G., Knott, S., Haley, C., 2013. An evolutionary perspective on epistasis and the missing heritability. PLoS Genetics 9, e1003295.

Houle, D., 1992. Comparing evolvability and variability of quantitative traits.. Genetics 130, 195–204. http://dx.doi.org/10.1093/genetics/130.1.195.

Johnson, N.A., 2000. Gene interactions and the origin of species. Epistasis and the Evolutionary Process. Oxford University Press, pp. 197–221.

Karlin, S., Feldman, M.W., 1970. Linkage and selection: Two locus symmetric viability model. Theor. Popul. Biol. 1, 39–71. http://dx.doi.org/10.1016/0040-5809(70)90041-9.

Kelly, J.K., 2000. Epistasis, linkage and balancing selection. In: Epistasis and the Evolutionary Process. Oxford University Press, pp. 146–157.

Kelly, J.K., Wade, M.J., 2000. Molecular evolution near a two-locus balanced polymorphism. J. Theoret. Biol. 204, 83–101.

Kimura, M., 1956. A model of a genetic system which leads to closer linkage by natural selection. Evolution 10, 278–287. http://dx.doi.org/10.1111/j.1558-5646.1956.tb02852.x.

Kirkpatrick, M., Johnson, T., Barton, N.H., 2002. General models of multilocus evolution. Genetics 161, 1727–1750. http://dx.doi.org/10.1093/genetics/161.4.1727.

Kondrashov, A.S., Sunyaev, S., Kondrashov, F.A., 2002. Dobzhansky–muller incompatibilities in protein evolution. Proc. Natl. Acad. Sci. 99, 14878–14883. http://dx.doi.org/10.1073/pnas.232565499.

Kroymann, J., Mitchell-Olds, T., 2005. Epistasis and balanced polymorphism influencing complex trait variation. Nature 435, 95–98. http://dx.doi.org/10.1038/nature03480.

López-Fanjul, C., Fernández, A., Toro, M.A., 2004. Epistasis and the temporal change in the additive variance-covariance matrix induced by drift. Evolution 58, 1655–1663. http://dx.doi.org/10.1111/j.0014-3820.2004.tb00451.x.

Mäki-Tanila, A., Hill, W.G., 2014. Influence of gene interaction on complex trait variation with multilocus models. Genetics 198, 355–367. http://dx.doi.org/10.1534/genetics.114.165282.

Nijhout, H.F., 1991. The Development and Evolution of Butterfly Wing Patterns. Smithsonian Institution.

Orr, H.A., Turelli, M., 2001. The evolution of postzygotic isolation: Accumulating dobzhansky-muller incompatibilities. Evolution 55, 1085–1094. http://dx.doi.org/10.1111/j.0014-3820.2001.tb00628.x.

Ortiz-Barrientos, D., Engelstädter, J., Rieseberg, L.H., 2016. Recombination rate evolution and the origin of species. Trends Ecol. Evolut. 31, 226–236. http://dx.doi.org/10.1016/j.tree.2015.12.016.

Otto, S.P., Feldman, M.W., 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. 51, 134–147. http://dx.doi.org/10.1006/tpbi.1997.1301.

Paixão, T., Barton, N.H., 2016. The effect of gene interactions on the long-term response to selection. Proc. Natl. Acad. Sci. 113, 4422–4427.

Paris, J.R., Whiting, J.R., Daniel, M.J., Ferrer Obiol, J., Parsons, P.J., van der Zee, M.J., Wheat, C.W., Hughes, K.A., Fraser, B.A., 2022. A large and diverse autosomal haplotype is associated with sex-linked colour polymorphism in the guppy. Nat. Commun. 13, 1233.

Pettersson, M., Besnier, F., Siegel, P.B., Carlborg, Ö., 2011. Replication and explorations of high-order epistasis using a large advanced intercross line pedigree. PLoS Genetics 7, e1002180.

Poelwijk, F.J., Tănase-Nicola, S., Kiviet, D.J., Tans, S.J., 2011. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. J. Theoret. Biol. 272, 141–144. http://dx.doi.org/10.1016/j.jtbi.2010.12.015.

Prakash, S., Lewontin, R., 1968. A molecular approach to the study of genic heterozygosity in natural populations. 3. Direct evidence of coadaptation in gene arrangements of drosophila.. Proc. Natl. Acad. Sci. 59, 398–405.

Purcell, J., Brelsford, A., 2025. Supergenes in organismal and social development of insects: ideas and opportunities. Curr. Opin. Insect Sci. 68, 101303.

Roff, D.A., Emerson, K., 2006. Epistasis and dominance: Evidence for differential effects in life-history versus morphological traits. Evolution 60, 1981–1990. http://dx.doi.org/10.1111/j.0014-3820.2006.tb01836.x.

Saona, R., Kondrashov, F.A., Khudiakova, K.A., 2022. Relation between the number of peaks and the number of reciprocal sign epistatic interactions. Bull. Math. Biol. 84, http://dx.doi.org/10.1007/s11538-022-01029-z.

Scarparo, G., Lopez, A., Muro, E., Brelsford, A., Purcell, J., 2026. Lethal epistasis maintains strong linkage disequilibrium between unlinked supergenes. BioRxiv.

Sciuchetti, L., Dufresnes, C., Cavoto, E., Brelsford, A., Perrin, N., 2018. Dobzhansky-muller incompatibilities, dominance drive, and sex-chromosome introgression at secondary contact zones: A simulation study: Sex-chromosome introgression in hybrid zones. Evolution 72, 1350–1361. http://dx.doi.org/10.1111/evo.13510.

Siljestam, M., Rueffler, C., Arnqvist, G., 2024. Sex-specific dominance and its effects on allelic diversity in sexually antagonistic loci. BioRxiv Prepr. http://dx.doi.org/10.1101/2024.05.06.592709v2.

Spencer, H.G., Priest, N.K., 2016. The evolution of sex-specific dominance in response to sexually antagonistic selection. Amer. Nat. 187, 658–666.

Starr, T.N., Thornton, J.W., 2016. Epistasis in protein evolution. Prot. Sci. 25, 1204–1218. http://dx.doi.org/10.1002/pro.2897.

Stolyarova, A.V., Neretina, T.V., Zvyagina, E.A., Fedotova, A.V., Kondrashov, A.S., Bazykin, G.A., 2022. Complex fitness landscape shapes variation in a hyperpolymorphic species. ELife 11, e76073. http://dx.doi.org/10.7554/eLife.76073.

Templeton, A.R., 2000. Epistasis and complex traits. In: Epistasis and the Evolutionary Process. Oxford University Press, pp. 41–57.

Wade, M.J., 2000. Epistasis as a genetic constraint within populations and an accelerant of adaptive divergence among them. In: Epistasis and the Evolutionary Process. Oxford University Press, pp. 213–231.

Wolf, J.B., Brodie, E.D., Wade, M.J., 2000. Epistasis and the Evolutionary Process. Oxford University Press.