

# A century of theories of balancing selection

Filip Ruzicka<sup>\*,1,2</sup>, Martyna K. Zwoinska<sup>\*,3</sup>, Debora Goedert<sup>4</sup>, Hanna Kokko<sup>5,6</sup>, Xiang-Yi Li Richter<sup>7,8</sup>, Iain R. Moodie<sup>9</sup>, Sofie Nilén<sup>9</sup>, Colin Olito<sup>9</sup>, Erik I. Svensson<sup>9</sup>, Peter Csuppon<sup>§,10</sup>, Tim Connallon<sup>§,2</sup>

<sup>1</sup> Institute of Science and Technology Austria, Klosterneuburg, Austria.

<sup>2</sup> School of Biological Sciences, Monash University, Clayton, Australia.

<sup>3</sup> Department of Ecology & Genetics, Uppsala University, Uppsala, Sweden.

<sup>4</sup> Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway.

<sup>5</sup> Institute of Organismic and Molecular Evolution, University of Mainz, Mainz, Germany.

<sup>6</sup> Institute for Quantitative and Computational Biosciences, University of Mainz, Mainz, Germany.

<sup>7</sup> Institute of Ecology and Evolution, University of Bern, Bern, Switzerland.

<sup>8</sup> Department of Biology, University of Konstanz, Konstanz, Germany.

<sup>9</sup> Department of Biology, Lund University, Lund, Sweden.

<sup>10</sup> Institute for Evolution and Biodiversity, University of Münster, Münster, Germany.

<sup>\*,§</sup> denote shared authorship

## Abstract

Traits that affect organismal fitness are often very genetically variable. This genetic variation is vital for populations to adapt to their environments, but it is also surprising given that nature (after all) “selects” the best genotypes at the expense of those that fall short. Explaining the extensive genetic variation of fitness-related traits is thus a longstanding puzzle in evolutionary biology, with cascading implications for ecology, conservation, and human health. Balancing selection—an umbrella term for scenarios of natural selection that maintain genetic variation—is a century-old explanation to resolve this paradox that has gained recent momentum from genome-scale methods for detecting it. Yet evaluating whether balancing selection can, in fact, resolve the paradox is challenging, given the logistical constraints of distinguishing balancing selection from alternative hypotheses and the daunting collection of theoretical models that formally underpin this debate. Here, we track the development of balancing selection theory over the last century and provide an accessible review of this rich collection of models. We first outline the range of biological scenarios that can generate balancing selection. We then examine how fundamental features of genetic systems—including non-random mating between individuals, differences in ploidy, genetic drift, and different genetic architectures of traits—have been progressively incorporated into the theory. We end by linking these theoretical predictions to ongoing empirical efforts to understand the evolutionary processes that explain genetic variation.

## Introduction

Throughout nature—from birds, mammals and flies to flowering plants<sup>1–3</sup>—individuals vary genetically in their ability to survive and reproduce, supplying the raw material for evolutionary adaptation. There are two broad schools of thought to explain this abundant genetic variation for fitness<sup>2,4</sup>. One school proposes that selection continually removes genetic variation for fitness, while mutation continually replenishes it<sup>5</sup>. Under this view, the resulting equilibrium between recurrent mutation and purifying selection (“mutation-selection balance”) accounts for fitness variation. The second school proposes that genetic variation for fitness is maintained by selection—a concept known as “balancing selection”<sup>6</sup>. Balancing selection can arise from a variety of scenarios (Box 1), including selection favouring heterozygotes over homozygotes, selection for rare over common genotypes, and selective trade-offs in which genetic variants favoured in some contexts (*e.g.*, seasons, niches, sexes, life-history stages) are disfavoured in others.

Despite considerable debate during the second half of the 20<sup>th</sup> century<sup>4</sup> and recent renewed interest in balancing selection (Fig. 1A), the contribution of balancing selection to the maintenance of fitness variation remains unresolved<sup>2</sup>. Little more than a decade ago, however, the prevailing view was that balancing selection was probably of minor importance in evolution. Summarising the sentiment, Asthana *et al.*<sup>7</sup> wrote that “balancing selection [...] has not been a significant force in human evolution”, while Hedrick<sup>8</sup> similarly suggested that “a low proportion of loci in the human genome are under long-term balancing selection”.

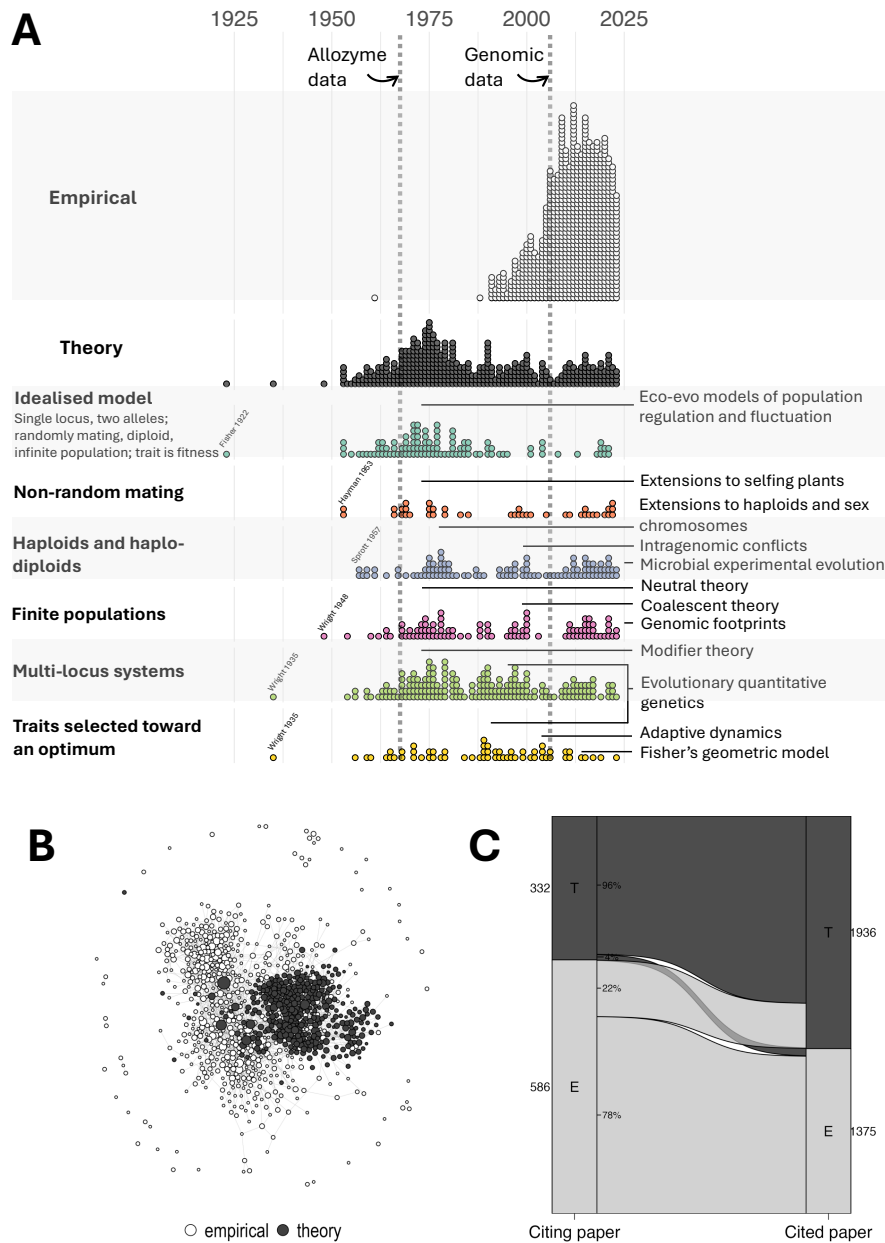
Why did this become the prevailing view? First, few empirical demonstrations of balancing selection were clearly documented at the time. The most prominent examples—including beta-globin alleles in humans<sup>9</sup>, self-incompatibility alleles in outcrossing plants<sup>10,11</sup>, chromosomal inversion polymorphisms in *Drosophila*<sup>12</sup>, and the major histocompatibility complex alleles in vertebrates<sup>13</sup>—were decades-old and viewed as outliers. When early genome scans failed to detect polymorphisms under long-term balancing selection<sup>7,14</sup>, the view that balancing selection was exceedingly rare was reinforced. Second, balancing selection hypotheses faced several theoretical objections. In particular, ubiquitous balancing selection on the abundant polymorphisms discovered in natural populations<sup>15,16</sup> seemed unlikely given the exorbitant mortality cost (“genetic load”) implied by such a hypothesis<sup>15,17</sup>, though ecological counterarguments soon tempered this criticism<sup>18–23</sup>. Meanwhile, the intuition that genetic trade-offs often generate balancing selection was undermined by theory demonstrating that trade-offs typically result in loss, rather than maintenance, of genetic polymorphism<sup>24,25</sup>. Finally, models of quantitative traits selected towards an intermediate optimum (*i.e.*, traits under “stabilizing selection”) highlighted restrictive conditions for persistent balancing selection at the genetic loci underlying the trait<sup>26–28</sup>. The notion that traits are often quantitative and selection is often stabilizing<sup>29</sup> seemed, therefore, to minimise the role of balancing selection.

Recent advances cast doubt on this consensus. New methods for detecting balancing selection in genomic sequences have now uncovered hundreds of candidate balanced polymorphisms<sup>30,31</sup>, in contrast to the small number of pre-genomic ones. Although most candidates require further validation, the current list likely represents the tip of the iceberg, given that genomic methods for inferring balancing selection are often under-powered and rely

on long-term signals<sup>31</sup>. Second, arguments about genetic loads have become less relevant to contemporary debates about balancing selection because genome-wide genetic diversity is now thought to be predominantly neutral or mildly deleterious<sup>32,33</sup>. The focus has instead shifted to the role of balancing selection in maintaining the extensive genetic variation for life-history traits and other fitness components, which is indeed too high to be explained by mutation alone<sup>2,3,34,35</sup>. Finally, new developments in theoretical modelling of adaptation give prominence to balancing selection. While earlier models emphasised the often-restrictive conditions for maintaining polymorphism indefinitely, newer models have highlighted the potential for short-lived episodes of balancing selection during the evolution of traits toward their optimum<sup>36–38</sup>. Moreover, a surge in theoretical models of fluctuating selection—inspired by recent evidence of predictable seasonal fluctuations of nucleotide polymorphism in *Drosophila*<sup>39–41</sup> and other species<sup>42</sup>—have highlighted broader conditions for balancing selection than implied by classical models<sup>42–45</sup>.

The renewed interest in balancing selection has prompted several reviews of empirical progress in detecting it<sup>13,30,31</sup>, but a review of the underlying population genetic *theories* of balancing selection is lacking. Here, we present a comprehensive and accessible overview of this theory. As is true for most models, the devil is in the details, and a deeper understanding requires some engagement with the particulars. We therefore provide a largely verbal overview in the main text, collect technical aspects in Boxes, and elaborate further in mathematical derivations presented in the Supplementary Material.

Balancing selection has a long history of study over the last century (Fig. 1A). We begin our review with the idealised conditions that characterise the very first models of balancing selection: *i.e.*, single biallelic genetic loci evolving in large, randomly-mating, diploid populations. These conditions mirror Fisher’s<sup>46</sup> classic analysis of heterozygote advantage<sup>47</sup>. As the 20<sup>th</sup> century progressed, the major technical innovations and debates that stimulated empirical evolutionary biology (*e.g.*, the advent of allozyme and genomic data; Fig. 1A) coincided with the development of theories of balancing selection. We therefore examine how various aspects of biological complexity were gradually incorporated into the theory (Fig. 1A), including non-random mating, deviations from diploidy, genetic drift, linkage and recombination between loci, and different forms of trait inheritance and phenotypic selection. We close by outlining progress in linking the theoretical predictions about balancing selection to empirical data (Fig. 1B,C).



**Figure 1. A.** A timeline of balancing selection theory and data. Each dot represents a paper on balancing selection. Papers were identified by a targeted search of the Web of Science, followed by manual curation of theory papers missed by the search ( $N = 872$  empirical papers,  $N = 402$  theory papers; see Supplementary Material for details of the collection methods). The top panel illustrates empirical research on balancing selection, which grew dramatically in the 1980s, spurred by the discovery of sequence variants (protein polymorphisms termed allozymes) and DNA variants, including molecular markers (e.g., microsatellites) and, later, genome sequences. The next panel illustrates the gradual growth of balancing selection theory, which spiked in the mid-1970s, spurred by the debate over neutral theory and a desire to account for abundant protein polymorphisms in natural populations <sup>4,15,22,48</sup>, and has gained ground in the genome era (2010s). The bottom panels zoom in on the

theoretical papers and separate them into different aspects of biological complexity (left), following the structure outlined in ‘Adding biological complexity to models of balancing selection’. Individual theory papers can span multiple aspects of complexity at once. The lines (right) highlight some particularly influential developments in the theory. **B.** Citation network of theory and empirical papers. Each node (circle) represents one article in the dataset. Node colours denote article types (empirical, theory); node size is proportional to number of times the article is cited locally within the network (*i.e.*, not the global citation count); edges show citations within the local network. Unconnected nodes do not cite articles within the network and are not cited by articles within the network. **C.** Citation patterns for each article type (E = empirical, T = theory) within the local network. Box and flow widths reflect the number of articles of a given type.

## Models of balancing selection: the basics

### The concept of balancing selection

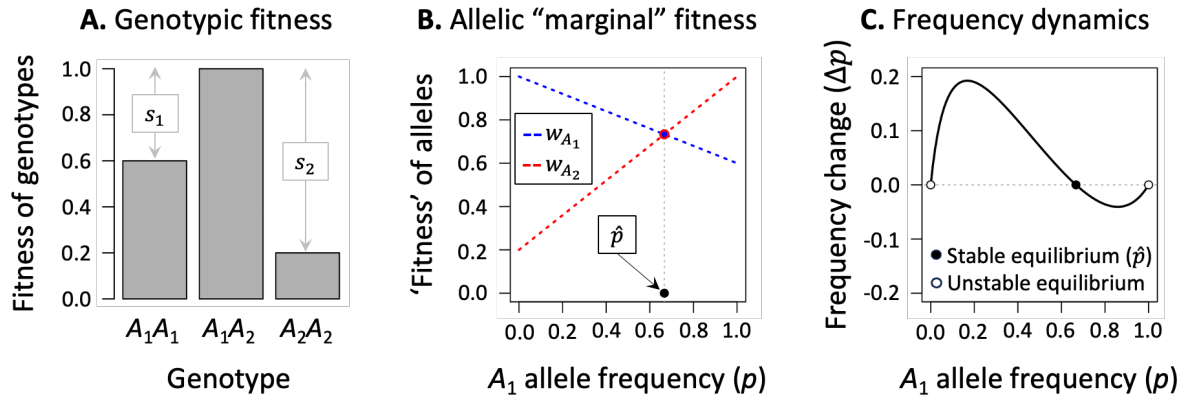
Balancing selection occurs when natural selection—without the intervention of other processes such as genetic drift, mutation, or migration—maintains genetic polymorphism. It can arise from a variety of scenarios (outlined in Box 1), but in all scenarios rare alleles have an advantage over common alleles. This aligns with the intuition that balancing selection maintains variation by opposing the loss of rare variants, as well as its operational definition of “protected polymorphism” in mathematical population genetics (<sup>49</sup>; see Box 1). Still, this definition is easy to misinterpret, and we outline some possible misconceptions in Box 2.

### Scenarios of balancing selection

Consider an idealised model in which a single genetic locus segregates for two alleles, the population is infinitely large and diploid, mating is random, and generations are discrete. One scenario that generates balancing selection under these assumptions is heterozygote advantage (*i.e.*, overdominance), in which heterozygotes have higher fitness (*i.e.*, higher survival and/or fertility) than homozygotes (Fig. 2A) <sup>46,50</sup>. Each allele, when rare, is expected to increase in frequency because rare variants are predominantly found in (fit) heterozygotes, whereas common variants are mostly found in (less fit) homozygotes. Although the fitness of each genotype is independent of its frequency, the average fitness of each *allele*—its so-called “marginal fitness”—declines with the allele’s frequency (Fig. 2B). Under heterozygote advantage, the population evolves towards a polymorphic equilibrium at which the marginal fitness of both alleles is equal (Fig. 2C). Heterozygote advantage can also maintain more than two alleles at a single locus, though conditions for maintenance of many alleles become restrictive unless there is a high degree of symmetry among homozygous and heterozygous genotypes for the set of alleles <sup>51,52</sup>.

Balancing selection is also likely, though not inevitable, when the fitness of each *genotype* declines with its frequency in the population—a scenario referred to as negative frequency-dependent selection <sup>53,54</sup>. In such cases, rare genotypes may consistently experience fitness advantages over more common ones. There is a vast ecological literature on negative

frequency-dependent selection<sup>55</sup> and myriad ways to model it (see chapter 5 of<sup>56; 57,58</sup>). In Box 1, we present a simple illustrative example that can either lead to balancing selection, or to fixation of one allele and extinction of the other (*i.e.*, to “directional selection”; see Supplementary Material).



**Figure 2.** Heterozygote advantage as an example of balancing selection. **A.** For a locus with two alleles ( $A_1$  and  $A_2$ , at frequencies  $p$  and  $q$ , respectively; Box 1), the fitness of  $A_1A_1$  homozygotes declines by  $s_1$ , and  $A_2A_2$  declines by  $s_2$ , relative to heterozygotes (fitness of the best genotype is scaled to one). **B.** Though fitness per genotype is frequency-*independent*, the average transmission rate of each allele to the next generation (its ‘marginal fitness’) is frequency-dependent. In outbred, randomly mating populations, the marginal fitness of the  $A_1$  and  $A_2$  alleles ( $w_{A_1} = 1 - ps_1$  and  $w_{A_2} = 1 - (1 - p)s_2$ , respectively) are each negative frequency-dependent (the marginal fitness of each allele declines with the frequency of that allele), becoming equal at the polymorphic equilibrium  $\hat{p} = s_2/(s_1 + s_2)$ . **C.** Whether the  $A_1$  allele increases ( $\Delta p > 0$ ) or decreases in frequency ( $\Delta p < 0$ ) depends on its current frequency relative to the equilibrium.

Balancing selection can also arise from a variety of genetic trade-offs, in which alleles that are advantageous in some contexts are harmful in others (Box 1). In the case of meiotic drive, a “driver” allele has a transmission advantage because it finds itself in more than half of the gametes produced by heterozygotes. However, driver alleles may also lower the fitness of individuals that carry them<sup>59,60</sup>. The trade-off between an allele’s transmission advantage and its fitness cost to carriers can then potentially generate balancing selection (Box 3). Trade-offs can also arise between temporally fluctuating environmental conditions<sup>42–45,61–63</sup>, different resources or “niches” used by the population (“niche antagonism”;<sup>64–67</sup>), females and males (“sexually antagonistic selection”;<sup>68–70</sup>), and life-history traits (“antagonistic pleiotropy” between *e.g.*, survival and fertility;<sup>71,72</sup>).

In each trade-off model, the balance between benefits and costs must be just right to maintain the polymorphism (Box 3), and conditions leading to directional selection are often

more permissive than those leading to balancing selection<sup>24,25</sup>. Conditions for balancing selection are particularly restrictive when selection is weak and the benefits and costs associated with each allele have similar degrees of dominance (Box 3). However, conditions for balancing selection become substantially more permissive when the cost of expressing each allele is at least partially recessive while its benefit is dominant (Box 3). Such “dominance reversals”<sup>48,52,73–76</sup>) can lead to a “net heterozygote advantage”, in which the mean fitness of heterozygotes is higher than the mean fitness of homozygotes. Net heterozygote advantage is a sufficient condition for balancing selection in randomly mating diploid populations (see Box 3 for elaboration).

### Allele frequency dynamics under balancing selection

Although the details of individual balancing selection scenarios vary, their allele frequency dynamics often simplify to a common form that highlights their conceptual unity (Box 1). For models with a single polymorphic equilibrium, allele frequency change per generation is approximately:

$$\Delta p \approx \alpha pq(\hat{p} - p) \quad (1)$$

where  $p$  and  $q = 1 - p$  refer to  $A_1$  and  $A_2$  allele frequencies,  $\alpha$  is the net strength of selection (here assumed to be weak, such that  $|\alpha| \ll 1$ ), and  $\hat{p}$  is the polymorphic equilibrium. Under balancing selection,  $\alpha > 0$  and  $0 < \hat{p} < 1$ , with values of  $\alpha$  and  $\hat{p}$  depending on the scenario of balancing selection and its underlying parameters (Box 1 and Supplementary Material).

Several insights emerge from eq. (1). First and most obviously, evolutionary change requires genetic variation at the locus ( $0 < pq < 1$ ). Second, given that  $\alpha pq$  must be positive when genetic variation is present, the term  $(\hat{p} - p)$  determines the direction of evolutionary change. The frequency of  $A_1$  increases when it is below the equilibrium, decreases when it is above the equilibrium, and remains stable at the equilibrium. Therefore, in the absence of genetic drift, balancing selection causes a genetically variable population to evolve towards the polymorphic equilibrium. This is illustrated in Fig. 2C for the case of heterozygote advantage, though similar dynamics characterise each of the scenarios outlined in Box 1.

While most models of balancing selection can be recast in the form described by eq. (1), there are notable exceptions. For example, strong negative frequency-dependent selection<sup>77</sup> and/or temporal fluctuations in the direction of selection can maintain polymorphism and exhibit predictable patterns<sup>43,63,78</sup>, but there is no polymorphic equilibrium state that the population eventually reaches. Models with multiple alleles<sup>51,79</sup>, or multiple polymorphic equilibria<sup>68,69</sup>, also cannot be described by eq. (1) (Box 1).

### Adding biological complexity to models of balancing selection

All models are simplifications, and the scenarios of balancing selection presented above obviously omit important aspects of biological complexity. In some cases, violations of their simplifying assumptions can fundamentally alter opportunities for balancing selection, or the evolutionary stability of balanced polymorphisms. For example, while many populations are indeed diploid and mate nearly randomly at most loci, others—such as haplo-diploid species, or self-fertilising plants—are not. Conditions for balancing selection that go beyond the diploid

and randomly-mating ideal were increasingly explored during the second half of the 20<sup>th</sup> century <sup>80,81</sup> (Fig. 1A). Likewise, while models that ignore genetic drift are often useful approximations of reality, drift is an inevitable feature of real populations with the potential to destabilise balanced polymorphisms. The rise of the neutral theory of molecular evolution <sup>82</sup>, coalescent theory <sup>83</sup>, along with the increasing feasibility of computer simulations, led to a burst of finite-population models of balancing selection (*e.g.*, <sup>84–86</sup>). Finally, while single-locus models may well describe traits with a simple genetic basis (*i.e.*, whose expression is dictated by few genetic variants), they apply awkwardly to continuous traits. The advent of “modifier” models, which focus on multi-locus genetic systems (*e.g.*, <sup>87,88</sup>), and various models that link genotype, phenotype and fitness (*e.g.*, <sup>27,36,89–91</sup>) extended the scope of theories of balancing selection beyond single loci (Fig. 1A). We outline the consequences of each of these aspects of biological complexity below.

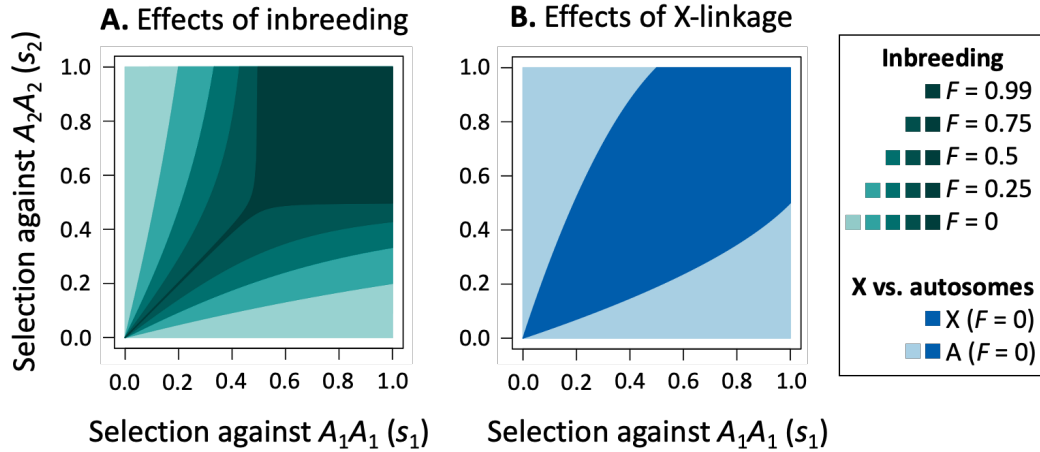
### Non-random mating

Mating patterns influence the proportion of heterozygotes, and thereby opportunities for balancing selection in scenarios where heterozygotes are favoured over homozygotes (*i.e.*, in cases involving true or net heterozygote advantage; see Boxes 1 and 3). Mating patterns that decrease the proportion of heterozygotes (*e.g.*, inbreeding or positive assortative mating by genotype) tend to decrease the range of conditions leading to balancing selection <sup>80,92</sup>, while mating patterns that increase the proportion of heterozygotes (*e.g.*, disassortative mating by genotype) tend to do the opposite <sup>93,94</sup>.

Consider the widespread case of inbreeding, which reduces heterozygote proportions relative to random mating. Models of balancing selection with inbreeding can be expressed using eq. (1), but with parameters  $\alpha$  and  $\hat{p}$  adjusted to include the population’s inbreeding coefficient ( $F$ ) where  $F = 1$  denotes complete inbreeding and  $F = 0$  denotes complete outcrossing. Under weak heterozygote advantage, the net strength of selection becomes  $\alpha = (1 - F)(s_1 + s_2)$ , the polymorphic equilibrium is  $\hat{p} = (s_2 - Fs_1)/\alpha$ , and balancing selection arises when  $Fs_1 < s_2 < s_1/F$ . Thus, inbreeding decreases the parameter range for balancing selection by a factor of  $1 - F$  when selection is weak (Fig. 3A), and somewhat less when selection is strong <sup>95</sup>. Inbreeding also depresses the scope of balancing selection resulting from trade-offs <sup>80,96–99</sup>. This can arise when there is net heterozygote advantage (*e.g.*, sexually antagonistic selection with dominance reversal <sup>100</sup>), or when self-fertilization reduces the strength of selection through the male sex function of hermaphrodites, which promotes the fixation of sexually antagonistic alleles that benefit females <sup>92,98,101</sup>. By contrast, inbreeding has little effect on balancing selection arising from negative frequency-dependent selection <sup>92,102</sup>.

Even when balancing selection is predicted to occur, inbreeding reduces the effective size of the population <sup>103</sup> and increases selective interference between genetically linked loci <sup>104</sup>. This promotes the loss of polymorphism by reducing the efficacy of balancing selection relative to genetic drift <sup>92</sup> (see section: ‘Finite populations’). On the other hand, inbreeding can sometimes facilitate balancing selection of multi-locus allele combinations <sup>101</sup> and intensify population genomic signals of long-term balancing selection <sup>105,106</sup>.





**Figure 3.** Consequences of inbreeding and X-linked (or haplo-diploid) inheritance on the maintenance of polymorphisms under heterozygote advantage. Whereas outbred and diploid populations can maintain a balanced polymorphism over the entire parameter space of selection (the full range of values for  $s_1$  and  $s_2$ ;  $0 < s_1, s_2 \leq 1$ ), inbred diploid populations (**A**) and X-linked genes or haplo-diploid populations (**B**) reduce the parameter space for balanced polymorphism (dark shaded regions). Results are based on exact criteria for balancing selection<sup>107,108</sup>.

### Haploids and haplo-diploids

Even if heterozygotes have higher fitness than homozygotes, such advantages become irrelevant in predominantly haploid populations, haploid individuals, or haploid stages of a life cycle (*i.e.*, gametes or the gametophyte stage of plants). In fully haploid populations, there is no scope for maintaining polymorphism by heterozygote advantage, meiotic drive, or antagonistic pleiotropy. Balancing selection may still arise from negative frequency-dependent selection<sup>109,110</sup> or from trade-offs between niches, seasons or sexes<sup>111–114</sup>. However, the conditions for balancing selection under such trade-offs are highly restrictive. In the case of trade-offs between sexes or niches, balancing selection conditions under haploidy are equivalent to those under diploidy with co-dominance (Box 3).

When populations experience selection in both diploid and haploid life stages, or when diploid and haploid individuals co-exist, conditions for balancing selection tend to be intermediate between purely diploid or haploid populations<sup>107,115</sup>. For example, at loci where one sex is haploid and the other is diploid (*e.g.*, haplodiploids, X-linked genes that are hemizygous in males), heterozygote advantage in the diploid context favours the maintenance of polymorphism, while the haploid context typically favours its loss. Hence, balancing selection only occurs when heterozygote advantage in the diploids outweighs directional selection in the haploids (as in the dark shaded region of Fig. 3B, where  $s_1$  and  $s_2$  are similar in magnitude, implying weak directional selection in haploids). In the rest of the parameter space, the allele that is least harmful in homozygous or haploid individuals is fixed (light shaded regions of Fig. 3B). Conditions for X-linked balancing selection are also reduced in models of female meiotic drive, antagonistic pleiotropy, and trade-offs between niches (Supplementary Material; <sup>116</sup>).

Although X-linked inheritance usually restricts conditions for balancing selection, exceptions can occur. Cases of male meiotic drive are especially complex because they are associated with skewed sex ratios, altered population size dynamics, and potentially extinction<sup>117–119</sup>. In models of sexually antagonistic selection<sup>73,120,121</sup>, the sex-specific dominance coefficients of the alleles determine whether polymorphisms are more readily maintained at autosomal (diploid) or X-linked loci. Specifically, the X is more permissive for balancing selection when male fitness costs of sexually antagonistic alleles show moderate to strong dominance ( $h_I > 1/(2 - s_I)$ , where  $h_I$  and  $s_I$  represent male dominance and selection coefficients). Autosomes are more permissive otherwise<sup>122</sup>. Comparable results arise under “ploiddally antagonistic selection”, where different alleles are favoured in the haploid and diploid stages of a life cycle<sup>115</sup>.

### Finite populations

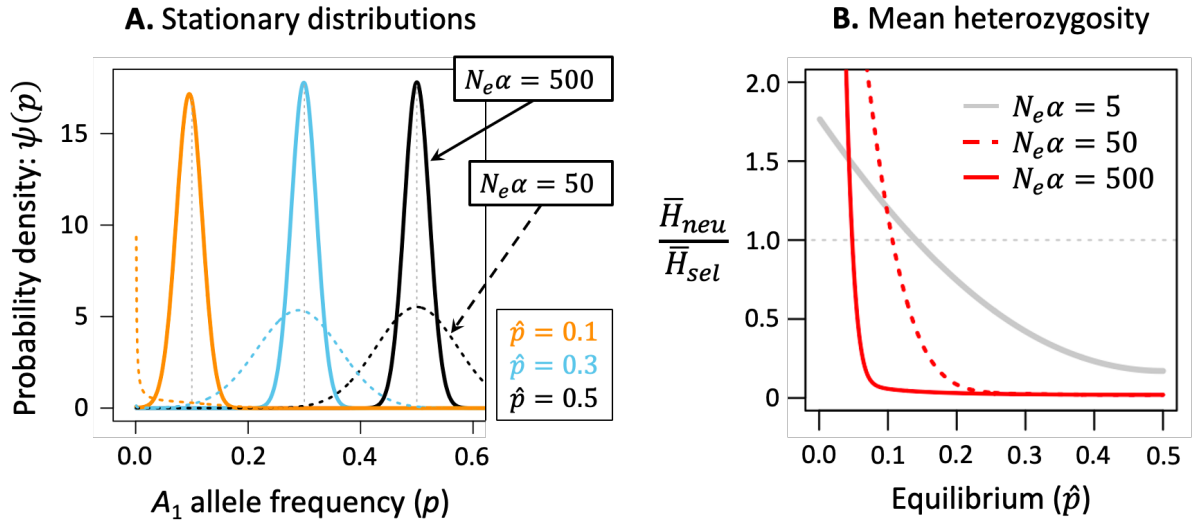
Real populations are finite and subject to genetic drift, which causes random deviations from deterministic evolutionary trajectories. Despite this randomness, population genetic models generate clear predictions for the stationary (*i.e.*, long-run) probability of observing each possible allele frequency state, as a function of the effective population size ( $N_e$ ) and the specific scenario of balancing selection (Fig. 4A). For the scenarios in Box 1, the stationary distribution for the  $A_1$  allele is:

$$\psi(p) = c[p(1 - p)]^{2kN_e u - 1} e^{-kN_e \alpha (\hat{p} - p)^2} \quad (2)$$

where  $\alpha$  is the net strength of selection (assumed to be weak),  $k$  is the number of gene copies carried by each member of the population ( $k = 2$  for diploids and  $k = 1$  for haploids),  $u$  is the mutation rate per locus (assumed to be the same for each allele), and  $c$  is a constant that ensures that the distribution integrates to one.

Analyses of the stationary distribution reveal two important consequences of drift for polymorphisms under balancing selection. First, given no additional mutations entering the population, a balanced polymorphism will eventually be lost despite selection to maintain it (see p. 165 of<sup>123</sup>). Second, and somewhat counterintuitively, drift can sometimes lead to more rapid loss of a balanced polymorphism than a neutral polymorphism<sup>84,124,125</sup>, resulting in lower genetic diversity than expected at neutrally evolving loci (Fig. 4B). This outcome is particularly likely when the population-scaled strength of selection is small (*e.g.*,  $kN_e \alpha < 10$ ) and equilibrium allele frequencies are close to zero or one<sup>85,126,127</sup>.

These counterintuitive predictions for balancing selection in finite populations can be viewed as a special case of a broader population genetic phenomenon. Intuition might lead us to predict that loci under balancing selection should exhibit the highest levels of genetic diversity, followed by neutral loci, followed by loci under positive or purifying selection. The classic theory clearly shows that balanced polymorphisms can harbour more or less diversity than neutral loci (Fig. 4B), mirroring results of recent models of genetic diversity for loci under weak positive selection<sup>128,129</sup>. Thus, there is no one-to-one mapping between balancing selection and inflation of genetic diversity relative to neutral expectations.



**Figure 4.** Interactions between balancing selection and genetic drift. **A.** Diffusion approximations are used to predict stationary (long-term) distributions (eq. (2)) for three equilibrium frequency states, and two strengths of selection. **B.** The ratio of mean genetic (*e.g.*, nucleotide) diversity for neutral loci relative to loci under balancing selection is shown. For each parameter combination, the expected heterozygosity is numerically calculated as  $\bar{H} = \int 2p(1-p)\psi(p)dp$ , where  $2p(1-p)$  is the heterozygosity and  $\psi(p)$  is the stationary distribution. When the ratio is greater than one, then expected heterozygosity is higher under neutrality than under balancing selection. Results for both panels assume  $2kN_e u = 0.01$  and  $k = 2$  (diploidy).

### Multi-locus systems with linkage

The models discussed so far have considered single loci, yet genes can mutually influence each other's potential for generating balancing selection when they are genetically linked on a chromosome. Most analyses have focused on conditions that maintain polymorphism at a pair of partially linked bi-allelic loci<sup>88</sup>. In the case of heterozygote advantage, polymorphism can be maintained at linked loci<sup>87,130</sup>. In trade-off scenarios, linkage tends to expand conditions for the maintenance of polymorphism relative to single-locus models<sup>131</sup>. For example, a haploid two-locus system can favour a balanced polymorphism in a temporally fluctuating environment<sup>132,133</sup>, but this is unlikely in otherwise similar single-locus models<sup>63,113,134</sup>. Conditions for polymorphism also expand under restricted recombination in scenarios of niche and sexual antagonism<sup>101,135–138</sup>, though the opposite is true for antagonistic pleiotropy<sup>72</sup>. Finally, loci that are not individually under balancing selection can interact epistatically to maintain a stable two-locus polymorphism, given sufficiently strong linkage between the loci<sup>139–141</sup>.

Balancing selection can also maintain allelic diversity at neutral sites that are linked to a balanced polymorphism. In cases of heterozygote advantage<sup>142</sup>, this so-called “associative overdominance” maintains polymorphism when the recombination rate between neutral and selected loci is smaller than the strength of selection<sup>143</sup>. The diversity at neutral sites increases exponentially with the number of balanced polymorphisms they are linked to<sup>141</sup>—a prediction

that applies to various scenarios, including heterozygote advantage <sup>86</sup> and temporally fluctuating selection <sup>144</sup>. Finally, just as balancing selection can sometimes *reduce* diversity at the selected site relative to neutrality (Fig. 4B), this diversity-reducing effect can extend to linked neutral loci as well <sup>144</sup>.

### **Traits selected towards an optimum**

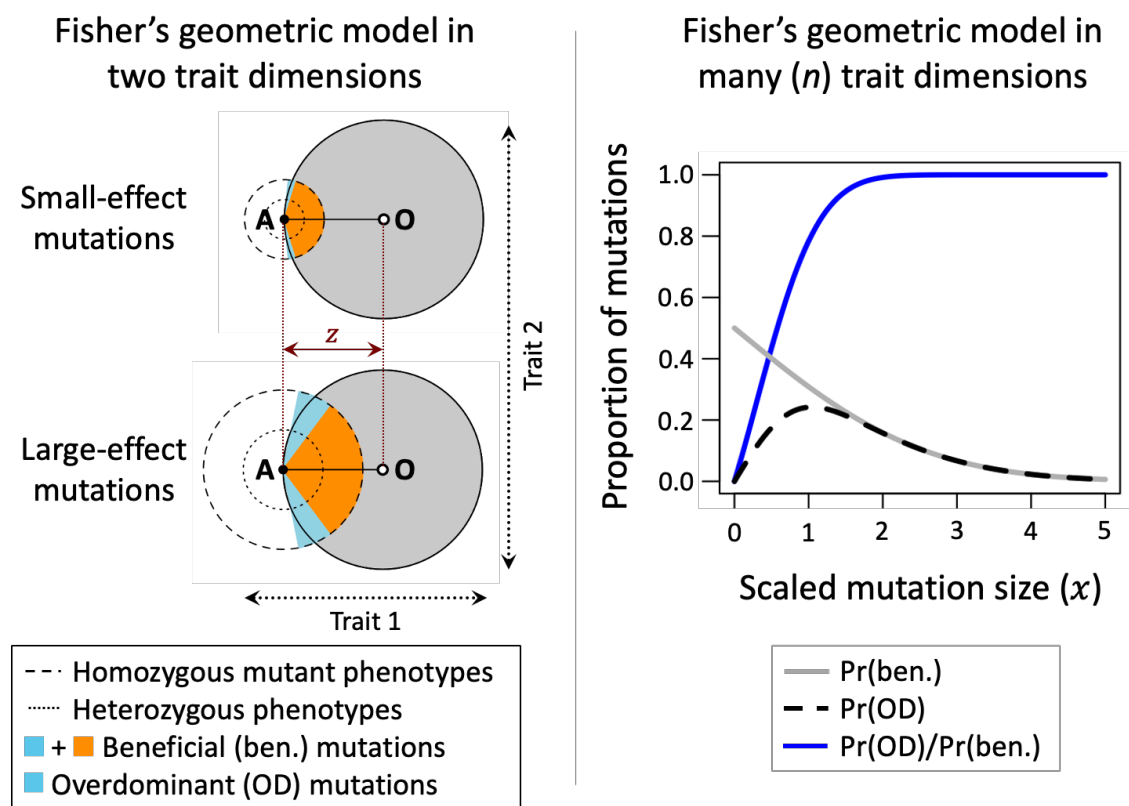
The models described so far do not explicitly include phenotypic effects; instead, genotypes are assigned fitness values that indirectly reflect selection on phenotypes. However, the lack of a genotype-phenotype map limits the range of questions we can ask about balancing selection. How often should we expect balancing selection to arise among new mutations or segregating genetic variants? Should we expect balanced polymorphisms to be transiently or persistently maintained over time? While the “parameter space” for balancing selection (Box 3) might seem, at first glance, a reasonable proxy for the likelihood of maintaining variation under a given scenario—or a framework for comparing scenarios—it does not differentiate between biologically plausible *vs.* implausible parameter values. We therefore need models that specify the mapping between genotype, phenotype and fitness.

A common way to predict the prevalence of balancing selection is to model the fitness effects of mutations affecting traits selected towards an optimum <sup>27,36,91</sup>. For example, studies using Fisher’s geometric model <sup>145</sup> usually examine trait systems in which adaptive mutations are sufficiently rare that adaptation proceeds by a temporal series of evolutionary steps (“adaptive walks”), with each step corresponding to the invasion of a new adaptive variant <sup>146</sup>. These mutation-limited dynamics of adaptation are most relevant to trait systems with small mutational targets, or where the mutations’ “scaled” phenotypic effect sizes are large (Fig. 5). Importantly, *scaled* sizes can be large even if *absolute* phenotypic effect sizes of mutations are small, provided the population is near its optimum and/or the number of pleiotropically-associated traits under selection (*i.e.*, trait dimensionality or “complexity”) is high <sup>38,145,147–149</sup>. Moreover, the very conditions that lead to mutation-limited evolution also lead to heterozygote advantage among the mutations that facilitate the adaptive walk toward the optimum <sup>36</sup>. An adaptive walk of a diploid population is therefore characterised by a series of short-lived episodes of balancing selection in which individual adaptive alleles are subject to balancing selection immediately following their spread within the population, but these balanced polymorphic states are eventually perturbed by the spread of the next adaptive allele <sup>36</sup>. These transient dynamics of balancing selection also emerge in trade-off scenarios in which mutations exhibit net heterozygote advantage (Box 3; <sup>37</sup>).

While transient episodes of balancing selection may occur during the evolutionary approach of a population to an optimum, conditions for long-term balancing selection become restrictive if there is sufficient genetic variation that allows the population to reach its optimum. In such cases, alleles whose homozygous carriers express the optimal phenotype will eventually become fixed (ch. 28 of <sup>150</sup>), and the ensuing stabilizing selection on the trait will remove rather than preserve genetic variation. If no homozygous genotype corresponds to the optimum, a single long-term balanced polymorphism can be maintained, with the remaining loci experiencing purifying selection against whichever allele is rarest <sup>26,27</sup>. Still, opportunities for long-term balanced polymorphisms affecting a polygenic trait can expand if loci contributing to the trait are tightly linked <sup>151,152</sup>, if there is pervasive dominance reversal (Box

3) or “diminishing-returns epistasis”<sup>43,153</sup>, or if there is strong “disruptive selection” favouring individuals at the trait extremes<sup>154–156</sup>.

Finally, while stabilizing selection affecting a single quantitative trait tends to remove polymorphism (with the exceptions noted above), balancing selection can still arise when variants underlying the trait have pleiotropic effects on other traits that affect fitness. In these “pleiotropic balancing selection” models<sup>27,90,157–159</sup>, selection through one fitness component promotes the maintenance of a balanced polymorphism, whereas selection through a second, pleiotropic fitness component—the quantitative trait under stabilizing selection—promotes its removal. Which component predominates depends on the relative strengths of selection through each (see Supplementary Material). Strong balancing selection generated through the first can easily offset comparatively weak stabilizing selection on the second, and vice-versa.



**Figure 5.** Balancing selection in Fisher's geometric model. There are  $n$  traits selected to a single phenotypic optimum defined by the organism's environment, with **A** denoting the ancestral phenotype (homozygotes for the ancestral allele), **O** representing the optimum phenotype and  $z$  the distance to the optimum. Fitness declines with the distance between an individual's phenotype and the optimum, and mutations have random and unbiased orientations in  $n$ -dimensional phenotypic space. Pleiotropy, which is inherent in the model, can lead to trade-offs, in which mutations causing beneficial changes in some traits simultaneously cause harmful changes in others. The left-hand panel illustrates a series of small-effect mutations (top) and large-effect mutations (bottom) for the case of  $n = 2$  traits. Beneficial mutations

(where heterozygous carriers are closer to the optimum than ancestral homozygotes) are shaded in orange and blue, with the latter representing mutations under heterozygote advantage. At high dimensions (right-hand panel), the proportion of beneficial variants that exhibit heterozygote advantage is predicted by the scaled mutation size,  $x = r\sqrt{n}/(2z)$ , which depends on the number of traits ( $n$ ), and the mutation's absolute phenotypic effect size ( $r$ ) relative to the distance to the optimum ( $z$ ). “Anisotropy” with respect to selection or mutational effects (*i.e.*, variation among traits in the strength of stabilizing selection, correlational selection, and non-random phenotypic effects of mutations) reduces the *effective* dimensionality of the system<sup>160–162</sup>.

## Linking theories of balancing selection to data

### Contributions of balancing selection to the maintenance of genetic polymorphisms

There is now ample evidence for balanced polymorphisms in natural populations<sup>13,31</sup>. For example, “top-down” approaches start with a conspicuous trait polymorphism and then investigate its genetic basis<sup>163</sup>. A classic example is sickle-cell anaemia, whose prevalence in some populations with histories of malaria exposure is explained by heterozygote advantage maintaining the beta-globin polymorphism<sup>9</sup>. Top-down approaches can identify recent and potentially ongoing balancing selection, and the ecological context of selection is sometimes known, resulting in rich portraits of the natural history of the balanced polymorphism.

“Bottom-up” approaches, by contrast, look for characteristic genomic signatures of balancing selection<sup>13,31,164</sup>. Such approaches sidestep the substantial difficulties of measuring fitness in natural populations. They are also trait-agnostic, allowing detection of balancing selection candidates affecting any trait, including those that are difficult to measure with high levels of accuracy or replication. Although bottom-up approaches are best suited for identifying signals of long-term balancing selection (*e.g.*, trans-species polymorphisms, genomic regions with elevated neutral diversity<sup>31</sup>), experimental evolution can be used to uncover ongoing (and possibly recent) selection, and even to identify the traits that drive these genetic responses. The experimental evolution approach is illustrated by classic studies of *Drosophila* inversion polymorphisms<sup>12</sup> and similar recent efforts in other species (*e.g.*,<sup>165</sup>).

Despite emerging empirical evidence for balancing selection, connections between theory and empirical research remain loose, as a citation network analysis quickly reveals (Fig. 1B,C; see also<sup>166,167</sup>). Consequently, assessing the prevalence of balancing selection in genomes remains challenging. We don't know, for instance, whether the fraction of loci influenced by balancing selection is closer to 0.1%, 1%, or 10% (though it is very unlikely to be any higher). Nonetheless, ongoing empirical research to estimate the important parameters of balancing selection models can help guide our thinking. For instance, the theory outlined in ‘Traits selected toward an optimum’ shows that transient balancing selection should be particularly common when evolution proceeds by adaptive walks involving effectively large-effect mutations (recall that mutation “size” is a function of pleiotropy and the distance of the population to the optimum; Fig. 5). Empirical work has indeed revealed many examples of

large-effect loci contributing to adaptation of traits with small mutational targets <sup>163</sup>, including balanced polymorphisms affecting pigmentation, and interactions with pathogens, in *Drosophila* <sup>168,169</sup>. Moreover, the genetic basis of polygenic traits sometimes includes a mixture of small- and large-effect loci, with the latter possibly maintained under balancing selection. This appears to be the case for cuticular hydrocarbon profiles in *Drosophila* <sup>170</sup>, age at maturity in Atlantic salmon <sup>171</sup>, and horn size in Soay sheep <sup>172</sup>. Large-effect variants at high frequencies are unlikely to be neutral or deleterious, and strongly suggest maintenance by balancing selection <sup>27</sup>, either due to direct selection on the trait of interest, or indirectly through pleiotropy and selection on another trait.

On the other hand, theory also shows that long-term stable balancing selection is unlikely when trait variation is highly polygenic, selection is stabilizing, and pleiotropy is limited (see section: ‘Traits selected toward an optimum’). In such cases, a trait’s genetic variation is predicted to be attributable to many rare alleles maintained at mutation-selection balance, with up to one locus under balancing selection <sup>26,27,154</sup>. Such predictions are consistent with the observation that evolutionary responses to artificial selection can occur over short timescales <sup>173</sup>, and with the absence of intermediate-frequency genetic variants of large-effect in many genome-wide association studies <sup>174</sup>. Nevertheless, it is important to keep in mind that the populations and traits that have been targeted for intensive study are not necessarily representative or “typical” traits, as each was chosen for personal, societal or logistical reasons. While it is certainly true that many well-studied traits are polygenic and continuously variable, others strongly deviate from the continuous polygenic ideal. In addition to visually conspicuous examples like colour patterns, traits associated with the basic molecular functions of individual genes (*e.g.*, their binding, catalytic, and other cellular functions) have small mutational targets (*e.g.*, genic mutation rates of order  $10^{-6}$ ) and are likely to display discontinuous patterns of genetic variability. Such “molecular traits” are reasonable candidates for evolution by adaptive walks <sup>175</sup>, which *are* conducive to short-term episodes of balancing selection <sup>36,37</sup>. These adaptive walk scenarios are unlikely to generate signals of long-term balancing selection (*e.g.*, inflated linked heterozygosity, gene genealogies with long internal branches, or trans-species polymorphisms), but they can generate signals of short-term balancing selection (*e.g.*, relatively low population genetic differentiation, or partial selective sweeps; <sup>13,36,164</sup>). Under this view, it is hardly surprising that our most compelling and well-understood examples of balancing selection are both short-term and mediated by selection on fundamental protein structure and function <sup>9</sup>.

Empirical research is also shedding light on other factors that affect the prevalence of balancing selection. For instance, interactions within and between loci (*e.g.*, dominance reversals; diminishing-returns epistasis) can expand conditions for balancing selection <sup>43,137,153</sup>. There is evidence for dominance reversals at major loci affecting life-history traits in Atlantic salmon, Soay sheep and *Drosophila* <sup>171,172,176</sup>, and in studies of quantitative traits such as salinity tolerance in copepods <sup>177</sup> and fitness of seed beetles <sup>178</sup>. Meanwhile, diminishing-returns epistasis for beneficial mutations appears to be common in microbial experiments <sup>179</sup>. Linkage between loci can also facilitate the establishment and maintenance of multi-locus balanced polymorphisms. Indeed, there are now several examples of large chromosome regions maintained as blocks of differentiated sequences, with evidence for long-term balancing selection <sup>180</sup>. Examples include chromosomal inversions affecting colour polymorphism in

stick insects and damselflies, and life-history traits in seaweed flies<sup>181–183</sup>. Finally, multi-locus balancing selection becomes likely when selection on traits is disruptive, rather than stabilizing<sup>154–156</sup>. Analyses of large human datasets<sup>184</sup>, as well as *Drosophila* wing morphology<sup>185,186</sup> suggest that stabilizing selection is more common than disruptive selection. However, estimates of nonlinear (*e.g.*, stabilizing and disruptive) selection are notoriously noisy<sup>29</sup>, and it remains unclear how often disruptive selection might play a role in generating balancing selection.

### **Contributions of balancing selection to genetic variation for fitness and its components**

Nucleotide diversity in the *genome* appears to be largely neutral and deleterious. Even so, balancing selection may still contribute substantially to variation in *phenotypes* and *fitness* that have been estimated in several species<sup>2,3,34,35</sup>. This is because the contribution of a single balanced polymorphism to trait and fitness variance is potentially equivalent to many hundreds of loci at mutation-selection balance<sup>2,187–189</sup>. For example, the additive genetic variance of a single locus with co-dominant, symmetric and antagonistic effects on two fitness components (*i.e.*, an antagonistic pleiotropy model with parameters  $s = s_1 = s_2$  and  $h_1 = h_2 = 0.5$ ; see Boxes 1 and 3) will be  $V_A = s^2/8$ , whereas the variance of a locus at mutation-selection balance will be  $V_A \approx 2sh\mu$  (Box 4). Given that mutation rates are typically orders of magnitude smaller than selection coefficients<sup>190,191</sup>, additive genetic variances due to a single balanced polymorphism should be orders of magnitude greater than variances of a single deleterious polymorphism. Similarly, for polygenic traits where one polymorphism is maintained by balancing selection and the rest are at mutation-selection balance (as predicted by some of the polygenic stabilizing selection models described above), the one balanced locus may account for a large fraction of the trait's genetic variance, even though most loci affecting the trait evolve under purifying selection.

A final point is that not all balancing selection scenarios maintain *additive* genetic variance for fitness at equilibrium, even though observations of high additive variation for fitness and fitness components are often used as arguments for the prevalence of balancing selection<sup>2,34,35</sup>. For example, antagonistic pleiotropy, heterozygote advantage and negative frequency-dependent selection do *not* necessarily maintain additive genetic variance for fitness, though they can contribute to the dominance variance for fitness, and the additive variance of individual fitness components (<sup>2</sup>; Box 4). Other trade-off scenarios, including antagonistic selection between niches, sexes, and temporally fluctuating environments, can maintain additive genetic variance for fitness. Thus, while balancing selection might well account for a large fraction of genetic variance for life-history traits and fitness, the different scenarios of balancing selection likely contribute to different components of the genetic variance<sup>2</sup>. There is, of course, a need for further estimates of the genetic variances of fitness traits, which are currently limited to a small number of populations. Evaluating whether such data consistently exceed what can be explained by mutation-selection balance, and identifying the specific forms of balancing selection that best explain such excesses, remains a pressing empirical challenge.



### Box 1. Balancing selection and the concept of protected polymorphism

The concept of a “protected polymorphism”<sup>49</sup> defines the parameter conditions of a model that lead to balancing selection. The idea is simple: in single locus systems with two alleles segregating in an infinitely large population, polymorphism should be maintained if selection always favours the spread of each allele when it is rare. In the models outlined in Table 1, balancing selection occurs when the  $A_1$  allele (with frequency  $p$ ) is favoured near frequencies  $p = 0$ , and the  $A_2$  allele (with frequency  $q = 1 - p$ ) is favoured near frequencies  $p = 1$ . In other words, it occurs when the boundary equilibria ( $p = 0$  and  $p = 1$ ) are “unstable”. Conditions leading to this consistent rare-allele advantage (protected polymorphism, per Prout<sup>49</sup>) are determined by stability analysis of the boundary equilibria<sup>192</sup>.

Table 1 outlines several scenarios that can lead to balancing selection (see Supplementary Material for stability analyses and conditions for balancing selection for each scenario). The parameters of these models have values between zero and one (*i.e.*,  $0 < s_1, s_2, s, h_1, h_2, h, c < 1$ ), with  $s_1, s_2$  and  $s$  representing homozygous fitness costs (*i.e.*, selection coefficients) of a particular allele and  $h_1, h_2$  and  $h$  defining dominance coefficients. Selection coefficients are positive quantities that cannot exceed one (because fitness values cannot be negative), while dominance coefficients can range from partial-to-complete recessivity ( $0 \leq h < 0.5$ ), to co-dominance (*i.e.*,  $h = h_1 = h_2 = 0.5$ ), to partial-to-complete dominance ( $0.5 < h \leq 1$ ). The strength of meiotic drive ( $c$ , where  $0 \leq c \leq 1$ ) is scaled so that  $c = 0$  corresponds to standard Mendelian segregation and  $c = 1$  to 100% transmission of the drive allele.

**Table 1. Single-locus models of balancing selection**

	$A_1A_1$	$A_1A_2$	$A_2A_2$
<b>Heterozygote advantage</b>	$1 - s_1$	1	$1 - s_2$
<b>Negative frequency-dependent selection</b>	$(1 - s_1 p)^2$	$(1 - s_1 p)(1 - s_2)$	$(1 - s_2)^2$
<b>Meiotic drive</b>			
Fitness of carrier	$1 - s$	$1 - sh$	1
$A_2$ transmission	1	$(1 + c)/2$	0
<b>Fitness trade-off</b>			
Fitness context 1	1	$1 - s_1 h_1$	$1 - s_1$
Fitness context 2	$1 - s_2$	$1 - s_2 h_2$	1

Average fitness	$1 - \frac{s_2}{2}$	$1 - \frac{s_1 h_1 + s_2 h_2}{2}$	$1 - \frac{s_1}{2}$
-----------------	---------------------	-----------------------------------	---------------------

In randomly mating diploid populations, heterozygote advantage generates balancing selection under all possible parameter conditions (*i.e.*, across the full “parameter space”;  $0 < s_1, s_2 \leq 1$ ). However, the other scenarios in Table 1 generate balancing selection across a subset of the parameter space, with the remaining space favouring removal of one allele and fixation of the other (*i.e.*, directional selection). Assuming weak selection, allele frequency change over a single generation is approximately  $\Delta p = \alpha p(1 - p)(\hat{p} - p)$  (eq. (1) in the main text), where  $\hat{p}$  denotes the polymorphic equilibrium state where it exists. Balancing selection occurs when  $\alpha \hat{p} > 0$  (the equilibrium  $p = 0$  is unstable and  $A_1$  invades when rare) and  $\alpha(1 - \hat{p}) > 0$  (the equilibrium  $p = 1$  is unstable and  $A_2$  invades when rare), in which case the population evolves to the polymorphic equilibrium. For example, with heterozygote advantage and random mating,  $\alpha = s_1 + s_2$  and  $\hat{p} = s_2/(s_1 + s_2)$ , leading to  $\alpha \hat{p} = s_2 > 0$  and  $\alpha(1 - \hat{p}) = s_1 > 0$ , both of which must be true given that  $s_1$  and  $s_2$  are positive quantities. Expressions for  $\alpha$  and  $\hat{p}$  for the other scenarios are presented in the Supplementary Material.

Various scenarios of balancing selection cannot be accommodated by eq. (1). For example, in cases where balancing selection maintains more than two alleles at a locus, the frequency dynamics of each allele usually depend on all other alleles instead of just being proportional to  $p(1 - p)$ . This is true of self-incompatibility systems, which can stably maintain many different alleles at a single locus, and also favour new alleles that may arise<sup>134,193–195</sup>. Balancing selection may also lead to perpetual change rather than convergence to a polymorphic equilibrium, including cyclic dynamics in bi-allelic systems under strong negative frequency-dependent selection<sup>77</sup>, and cycles in three-allele systems where the relative advantages of each allele relative to the others is nontransitive (*i.e.*, type A outperforms B, B outperforms C, and C outperforms A, as in rock-paper-scissors games;<sup>196</sup>). Lastly, multiple polymorphic equilibria can occur in models with multiple loci, in single-locus models with sexually antagonistic selection<sup>68,69</sup>, or with non-linear frequency-dependent selection, such as game theory models involving interactions between more than two individuals<sup>197,198</sup>.

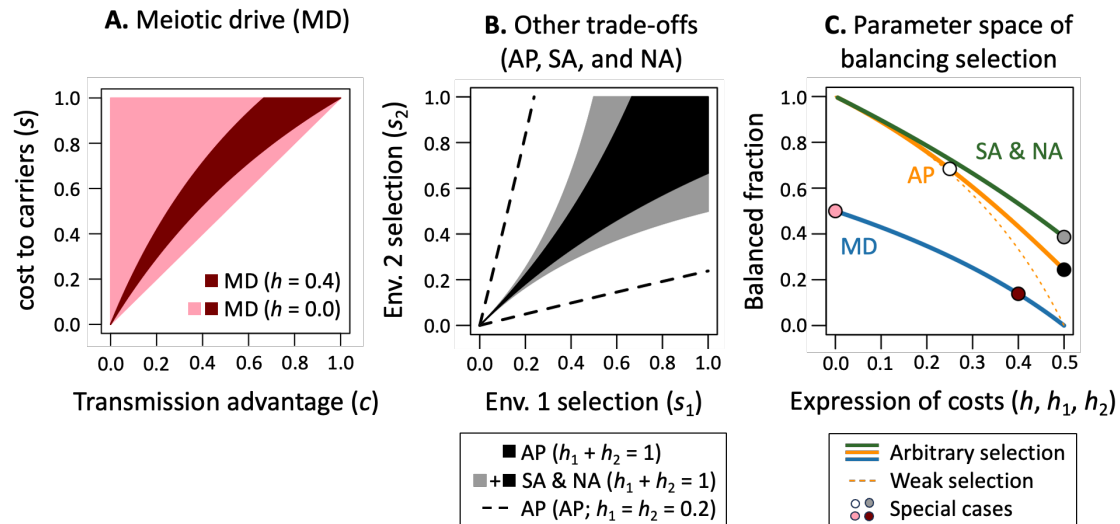
## Box 2. Clarifying misconceptions about balancing selection

The definition of balancing selection is well established in theoretical population genetics (Box 1) but is easy to misinterpret. We try to clarify some possible misconceptions below.

- ***It is not a category of trait selection.*** Natural selection refers to the differential survival or reproduction of individuals expressing different phenotypes, which can occur regardless of the genetic basis of the trait in question. In contrast, balancing selection—a concept from population genetics—is defined by the *genetic consequences* of natural selection. Specifically—in the absence of other evolutionary factors—it refers to any form of phenotypic selection that maintains polymorphism at one or more loci within the genome (Box 1). Thus, different forms of trait selection (e.g., directional, disruptive, or stabilizing; <sup>29</sup>) potentially generate balancing selection at trait-affecting loci <sup>27,154</sup>, but balancing selection does not necessarily accompany a given form of trait selection.
- ***It is not equivalent to heterozygote advantage, negative frequency-dependent selection on genotypes, or trade-offs.*** Each of these scenarios can generate balancing selection, but they need not. For example, heterozygote advantage does not always lead to balancing selection in haplodiploids and inbreeding populations <sup>92,107</sup> (Fig. 3), nor do certain models of negative frequency-dependent selection <sup>199</sup>, nor do many trade-offs (Box 3). However, balancing selection *does* imply that the “marginal fitness” of each *allele* declines with its frequency (Fig. 2B). This type of frequency-dependence should not be confused with the more common usage of the term (*i.e.*, that the fitness of each *genotype* declines with its frequency)—indeed it applies, for example, to heterozygote advantage (Fig. 2B), in which the fitness of each genotype is independent of frequency.
- ***It is not necessarily a long-term process.*** Balancing selection is often used interchangeably with long-term balancing selection, possibly because genome scans often rely on long-term rather than short-term signals (*i.e.*, long-term signals have durations greater than  $4N_e$  generations, where  $N_e$  is effective population size; <sup>31</sup>) or because the term “stability” (Box 1) is taken to imply perpetual stability. However, timescale is not part of the definition of balancing selection. It can be transient (e.g., over relatively few generations) or persistent (e.g., over many thousands of generations), with each scenario leaving a different genomic footprint.
- ***It does not always elevate polymorphism in real populations.*** When drift is strong relative to selection, balancing selection can lead to patterns of polymorphism that are indistinguishable from neutrality, both at the target of selection or at linked loci. It can even *reduce* levels of genetic diversity, relative to neutrality, when the equilibrium frequency of the selected allele is close to zero (see section: ‘Finite populations’).
- ***It is not the only evolutionary process that “maintains” genetic polymorphism.*** Balancing selection maintains polymorphism in the absence of other evolutionary factors. However, interactions between mutation, migration, genetic drift, and positive or purifying selection are also capable of maintaining polymorphism <sup>52,200,201</sup>. For example, purifying selection against harmful genetic variants is offset by mutation, which continuously introduces new harmful variants. Genetic variation can then be maintained at the equilibrium between these two opposing processes (“mutation-selection balance”).

### Box 3. Dominance and opportunities for balancing selection through trade-offs

Trade-offs can generate balancing selection, but they often will not<sup>25</sup>. The following figure outlines conditions for balancing selection under trade-off scenarios where a polymorphic equilibrium is attainable, *i.e.*: meiotic drive (MD), antagonistic pleiotropy (AP), sexually antagonistic selection (SA), and niche antagonism (NA).



Balancing selection is possible under MD (shaded regions of panel A; <sup>60</sup>) when the homozygous fitness cost (selection coefficient,  $s$ ) of a drive allele is greater than its segregation advantage ( $c$ ), and the cost is at least partially masked in heterozygous carriers ( $h < 0.5$ ; parameters as defined in Box 1). Other trade-off scenarios can generate balancing selection when fitness costs are co-dominant on average ( $h_1, h_2 = 0.5$ , shaded regions of panel B; <sup>24,60,74</sup>), though conditions are restrictive unless selection is strong. Conditions for balancing selection become more permissive when the expression of each allele is recessive in contexts where it is costly and dominant in contexts where it is beneficial <sup>48,52,73–76</sup> ( $h_1, h_2 < 0.5$ ). In the case of AP, SA and NA, this is known as a “dominance reversal” (or, more specifically, a “favourable reversal of dominance”).

Dominance reversals increase the likelihood that the average fitness of the heterozygotes is greater than the average for homozygotes (*i.e.*, that “net heterozygote advantage” arises across contexts of selection). In Table 1, within Box 1, a net heterozygote advantage occurs when  $s_1 h_1 + s_2 h_2 < \min(s_1, s_2)$ . In the extreme case of a complete dominance reversal ( $h_1 = h_2 = 0$ ), a net heterozygote advantage and balancing selection will arise across the entire parameter space for  $s_1$  and  $s_2$ . A net heterozygote advantage is a requirement for balancing selection through AP <sup>72</sup>, but not under SA or NA (*e.g.*, co-dominant costs do not lead to net heterozygote advantage but permit balancing selection; <sup>69,74</sup>).

The NA model considered above consists of two equally sized ecological niches with random dispersal and “soft selection”, which is a special case of Levene’s more general model <sup>64,65</sup>. Balancing selection through NA becomes less likely when one niche is more common than the other and selection is “hard” (<sup>62</sup>; see Supplementary Material). Conditions become more permissive when gene flow is limited between patches (*e.g.*, due to habitat selection, assortative mating by niche type, and extrinsic barriers to gene flow; <sup>65</sup>). Broader mathematical criteria for balancing selection are described in the Supplementary Material.

#### Box 4. Genetic variances due to deleterious and balanced polymorphisms

In randomly mating diploid populations, the contribution of a single biallelic locus to the additive and dominance genetic variance of a trait ( $V_A$  and  $V_D$ , respectively) is:

$$V_A = 2pq(p(T_{11} - T_{12}) + q(T_{12} - T_{22}))^2$$

$$V_D = \frac{1}{4}(2pq)^2(2T_{12} - T_{11} - T_{22})^2$$

where  $p$  and  $q$  are the population frequencies of the  $A_1$  and  $A_2$  alleles, and  $T_{11}$ ,  $T_{12}$ , and  $T_{22}$  represent the average trait (or fitness) values of individuals carrying genotypes  $A_1A_1$ ,  $A_1A_2$  and  $A_2A_2$ , respectively<sup>34,202</sup>. Thus, the locus' contribution to each variance component depends on both the proportion of the population that is heterozygous ( $2pq$ ) and the magnitude of genotypic differences in trait values.

Expressions for  $V_A$  and  $V_D$  can be evaluated under different population genetic models for the maintenance of genetic variation, allowing for comparisons of their predictions. For example, for models of mutation-selection balance, the frequency of a deleterious allele is  $\hat{q} \approx \mu/s_d$ , and its contribution to the additive genetic variance for fitness will be  $V_A \approx 2s_d\mu$ , where  $\mu$  is the mutation rate per gamete and  $s_d$  is the fitness cost to heterozygous carriers of the mutation (the approximations assume that selection is strong relative to mutation:  $s_d \gg \mu$ ). Note that the contribution of deleterious alleles to single fitness components is expected to be even lower than these approximations predict<sup>2</sup>. For balancing selection models, trade-offs between sexes, niches, or temporal environments will contribute to the additive genetic variance for overall fitness. By contrast, heterozygote advantage and negative frequency-dependent selection contribute no additive genetic variance for overall fitness at equilibrium (Table 2). The same is true for antagonistic pleiotropy (which relies on true heterozygote advantage to maintain the polymorphism; Box 3), though it contributes to additive genetic variance for individual fitness *components*<sup>2</sup>.

**Table 2.** Additive and dominance genetic variance at equilibrium under balancing selection due to heterozygote advantage, negative frequency-dependent selection (NFDS<sup>1</sup>), and a special case of antagonistic pleiotropy<sup>2</sup>.

Bal. selection scenario	$V_A$	$V_D$
Heterozygote advantage	$V_A = 0$	$V_D = (\hat{p}\hat{q})^2(s_1 + s_2)^2$
NFDS <sup>1</sup>	$V_A = 0$	$V_D = 0$
Antagonistic pleiotropy <sup>2</sup>		
Fitness component 1	$V_A = \hat{p}\hat{q}s_1^2/2$	$V_{D,1} = 0$
Fitness component 2	$V_A = \hat{p}\hat{q}s_2^2/2$	$V_{D,1} = 0$

Overall fitness

$$V_A = 0$$

$$V_D = (s_1 s_2 \hat{p} \hat{q} / 2)^2$$

---

<sup>1</sup> The model of negative frequency-dependent selection is that of Box 1.

<sup>2</sup> Pleiotropic alleles are assumed to have co-dominant effects on each fitness component ( $h_1 = h_2 = 0.5$  in Box 3) and there is a net heterozygote advantage with respect to overall fitness.

---

## Acknowledgements

We thank Brian and Deborah Charlesworth for extensive comments and suggestions. We also thank Göran Arnqvist, Jitka Polechová and Philip Hedrick for further comments. This work was supported by a H2020 Marie Skłodowska-Curie COFUND Action fellowship (#101034413, to FR), the Birgitta Sintring Foundation (#S2024-0007, to MKZ), the Research Council of Norway (302619, to DG), the Alexander von Humboldt Foundation (to HK), the Swiss National Science Foundation (#211549, to XLR), the Swedish Research Council (#2022-03603, to CO; #2020-03123, to EIS) and the European Research Council (ERC-2023-STG-#101117517, to CO). We are particularly grateful to the European Society of Evolutionary Biology for funding a Special Topics Network workshop (to TC, HK, EIS), from which this review began.

## References

1. Delph, L. F. & Kelly, J. K. On the importance of balancing selection in plants. *New Phytologist* **201**, 45–56 (2014).
2. Charlesworth, B. Causes of natural variation in fitness: evidence from studies of *Drosophila* populations. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 1662–1629 (2015).
3. Bonnet, T. *et al.* Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. *Science* **376**, 1012–1016 (2022).
4. Lewontin, R. C. *The Genetic Basis of Evolutionary Change*. *Evolution* vol. 29 (Columbia University Press, 1974).
5. Muller, H. J. Our load of mutations. *American Journal of Human Genetics* **2**, 111–176 (1950).
6. Dobzhansky, T. A review of some fundamental concepts and problems of population genetics. *Cold Spring Harbor Symposia on Quantitative Biology* **20**, 1–15 (1955).
7. Asthana, S., Schmidt, S. & Sunyaev, S. A limited role for balancing selection. *Trends Genet* **21**, 30–32 (2005).
8. Hedrick, P. W. Balancing selection. *Current Biology* **17**, R230–R231 (2007).

9. Allison, A. C. Protection afforded by sickle-cell trait against subtertian malarial infection. *British Medical Journal* **1**, 290–294 (1954).
10. Emerson, S. The genetics of self-incompatibility in *Oenothera organensis*. *Genetics* **23**, 190–202 (1938).
11. Wright, S. The distribution of self-sterility alleles in populations. *Genetics* **24**, 538–552 (1939).
12. Wright, S. & Dobzhansky, T. Genetics of natural populations. Xii. experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila Pseudoobscura*. *Genetics* **31**, 125–156 (1946).
13. Charlesworth, D. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics* **2**, e64 (2006).
14. Bubb, K., Bovee, D., Buckley, D. & Haugen, E. Scan of human genome reveals no new loci under ancient balancing selection. *Genetics* **173**, 2165–2177 (2006).
15. Lewontin, R. C. & Hubby, J. L. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* **54**, 595–609 (1966).
16. Kreitman, M. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**, 412–417 (1983).
17. Kimura, M. & Crow, J. F. The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738 (1964).
18. King, J. L. Continuously distributed factors affecting fitness. *Genetics* **55**, 483–492 (1967).
19. Milkman, R. D. Heterosis as a major cause of heterozygosity in nature. *Genetics* **55**, 493–495 (1967).



20. Sved, J. A., Reed, T. E. & Bodmer, W. F. The number of balanced polymorphisms that can be maintained in a natural population. *Genetics* **55**, 469–481 (1967).
21. Wallace, B. Hard and soft selection revisited. *Evolution* **29**, 465 (1975).
22. Clarke, B. C. The evolution of genetic diversity. *Proc. R. Soc. Lond. B.* **205**, 453–474 (1979).
23. Agrawal, A. F. & Whitlock, M. C. Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annu. Rev. Ecol. Evol. Syst.* **43**, 115–135 (2012).
24. Maynard Smith, J. & Hoekstra, R. Polymorphism in a varied environment: how robust are the models? *Genet. Res.* **35**, 45–57 (1980).
25. Prout, T. How well does opposing selection maintain variation? in *Evolutionary genetics: from molecules to morphology* 157–181 (Cambridge University Press, Cambridge, United Kingdom, 2000).
26. Wright, S. Evolution in populations in approximate equilibrium. *Journ of Genetics* **30**, 257–266 (1935).
27. Turelli, M. & Barton, N. H. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and GxE interactions. *Genetics* **166**, 1053–1079 (2004).
28. Johnson, T. & Barton, N. Theoretical models of selection and mutation on quantitative traits. *Phil. Trans. R. Soc. B* **360**, 1411–1425 (2005).
29. Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations. *The American Naturalist* **157**, 245–261 (2001).
30. Fijarczyk, A. & Babik, W. Detecting balancing selection in genomes: Limits and prospects. *Molecular Ecology* **24**, 3529–3545 (2015).

31. Bitarello, B. D., Brandt, D. Y. C., Meyer, D. & Andrés, A. M. Inferring balancing selection from genome-scale data. *Genome Biology and Evolution* **15**, evad032 (2023).
32. Leffler, E. M. *et al.* Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biology* **10**, e1001388 (2012).
33. Jensen, J. D. *et al.* The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern and Hahn 2018. *Evolution* **73**, 111–114 (2019).
34. Charlesworth, B. & Hughes, K. The maintenance of genetic variation in life history traits. in *Evolutionary Genetics: From Molecules to Morphology, Volume 1* 369–391 (Cambridge University Press, 2000).
35. Sharp, N. P. & Agrawal, A. F. An experimental test of the mutation-selection balance model for the maintenance of genetic variance in fitness components. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181864 (2018).
36. Sellis, D., Callahan, B. J., Petrov, D. A. & Messer, P. W. Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20666–20671 (2011).
37. Connallon, T. & Clark, A. G. Balancing selection in species with separate sexes: insights from Fisher’s geometric model. *Genetics* **197**, 991–1006 (2014).
38. Manna, F., Martin, G. & Lenormand, T. Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* **189**, 923–937 (2011).
39. Bergland, A. O., Behrman, E. L., O’Brien, K. R., Schmidt, P. S. & Petrov, D. A. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genetics* **10**, e1004775 (2014).
40. Machado, H. E. *et al.* Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in *Drosophila*. *eLife* **10**, e67577 (2021).

41. Rudman, S. M. *et al.* Direct observation of adaptive tracking on ecological time scales in *Drosophila*. *Science* **375**, eabj7484 (2022).
42. Johnson, O. L., Tobler, R., Schmidt, J. M. & Huber, C. D. Fluctuating selection and the determinants of genetic variation. *Trends in Genetics* **39**, 491–504 (2023).
43. Wittmann, M. J., Bergland, A. O., Feldman, M. W., Schmidt, P. S. & Petrov, D. A. Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E9932–E9941 (2017).
44. Bertram, J. & Masel, J. Different mechanisms drive the maintenance of polymorphism at loci subject to strong versus weak fluctuating selection. *Evolution* **73**, 883–896 (2019).
45. Yamamichi, M., Ellner, S. P. & Hairston, N. G. Beyond simple adaptation: Incorporating other evolutionary processes and concepts into eco-evolutionary dynamics. *Ecology Letters* **26**, (2023).
46. Fisher, R. A. On the dominance ratio. *Proc. R. Soc. Edinb.* **42**, 321–341 (1922).
47. Charlesworth, B. Fisher’s historic 1922 paper *On the dominance ratio*. *Genetics* **220**, iyac006 (2022).
48. Gillespie, J. H. & Langley, C. H. A general model to account for enzyme variation in natural populations. *Genetics* **76**, 837–848 (1974).
49. Prout, T. Sufficient conditions for multiple niche polymorphism. *The American Naturalist* **102**, 493–496 (1968).
50. Hedrick, P. W. What is the evidence for heterozygote advantage selection? *Trends Ecol Evol* **27**, 698–704 (2012).

51. Lewontin, R. C., Ginzburg, L. R. & Tuljapurkar, S. D. Heterosis as an explanation for large amounts of genic polymorphism. *Genetics* **88**, 149–169 (1978).
52. Charlesworth, B. & Charlesworth, D. *Elements of Evolutionary Genetics*. (W. H. Freeman, 2010).
53. Haldane, J. B. S. & Jayakar, S. D. Polymorphism due to selection depending on the composition of a population. *J Genet* **58**, 318–323 (1963).
54. Ayala, F. J. & Campbell, C. A. Frequency-dependent selection. *Annu. Rev. Ecol. Syst.* **5**, 115–138 (1974).
55. Gómez-Llano, M., Bassar, R. D., Svensson, E. I., Tye, S. P. & Siepielski, A. M. Meta-analytical evidence for frequency-dependent selection across the tree of life. *Ecology Letters* **27**, e14477 (2024).
56. Wright, S. *Evolution and the Genetics of Populations*. vol. 2 (The University of Chicago Press, Chicago, IL, 1969).
57. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
58. Golding, B. The prospects for polymorphisms shared between species. *Heredity* **68**, 263–276 (1992).
59. Prout, T. Some effects of variations in the segregation ratio and of selection upon the frequency of alleles. *Acta Genet, et Statist. Med* 148–151 (1953).
60. Rode, N. O., Estoup, A., Bourguet, D., Courtier-Orgogozo, V. & Débarre, F. Population management using gene drive: molecular design, models of spread dynamics and assessment of ecological risks. *Conserv Genet* **20**, 671–690 (2019).
61. Wright, S. On the roles of directed and random changes in gene frequency in the genetics of populations. *Evolution* **2**, 279 (1948).

62. Dempster, E. R. Maintenance of genetic heterogeneity. *Cold Spring Harbor Symposia on Quantitative Biology* **20**, 25–32 (1955).
63. Haldane, J. B. S. & Jayakar, S. D. Polymorphism due to selection of varying direction. *J Genet* **58**, 237–242 (1963).
64. Levene, H. Genetic equilibrium when more than one ecological niche is available. *The American Naturalist* **87**, 331–333 (1953).
65. Christiansen, F. B. Hard and soft selection in a subdivided population. *The American Naturalist* **109**, 11–16 (1975).
66. Felsenstein, J. Theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* **10**, 253–280 (1976).
67. Hedrick, P. W. Genetic polymorphism in heterogeneous environments: The age of genomics. *Annual Review of Ecology, Evolution, and Systematics* **37**, 67–93 (2006).
68. Owen, A. R. G. A genetical system admitting of two distinct stable equilibria under natural selection. *Heredity* **7**, 97–102 (1953).
69. Kidwell, J. F., Clegg, M. T., Stewart, F. M. & Prout, T. Regions of stable equilibria for models of differential selection in the two sexes under random mating. *Genetics* **85**, 171–183 (1977).
70. Bonduriansky, R. & Chenoweth, S. F. Intralocus sexual conflict. *Trends in Ecology & Evolution* **24**, 280–288 (2009).
71. Rose, M. R. Antagonistic pleiotropy, dominance, and genetic variation. *Heredity* **48**, 63–78 (1982).
72. Curtsinger, J. W., Service, P. M. & Prout, T. Antagonistic pleiotropy, reversal of dominance, and genetic polymorphism. *The American Naturalist* **144**, 210–228 (1994).

73. Fry, J. D. The genomic location of sexually antagonistic variation: some cautionary comments. *Evolution* **64**, 1510–1516 (2010).
74. Connallon, T. & Chenoweth, S. F. Dominance reversals and the maintenance of genetic variation for fitness. *PLoS Biology* **29**, e3000118 (2019).
75. Reid, J. M. Intrinsic emergence and modulation of sex-specific dominance reversals in threshold traits. *Evolution* **76**, 1924–1941 (2022).
76. Grieshop, K., Ho, E. K. H. & Kasimatis, K. R. Dominance reversals: the resolution of genetic conflict and maintenance of genetic variation. *Proc. R. Soc. B.* **291**, 20232816 (2024).
77. Chevin, L.-M., Gompert, Z. & Nosil, P. Frequency dependence and the predictability of evolution in a changing environment. *Evolution Letters* **6**, 21–33 (2022).
78. Lande, R. Expected relative fitness and the adaptive topography of fluctuating selection. *Evolution* **61**, 1835–1846 (2007).
79. Spencer, H. G. & Marks, R. W. The maintenance of single-locus polymorphism. I. Numerical studies of a viability selection model. *Genetics* **120**, 605–613 (1988).
80. Hayman, B. I. Mixed selfing and random mating when homozygotes are at a disadvantage. *Heredity* **7**, 185–192 (1953).
81. Sprott, D. A. The stability of a sex-linked allelic system. *Annals of Human Genetics* **22**, 1–6 (1957).
82. Kimura, M. *The Neutral Theory of Molecular Evolution. The neutral theory of molecular evolution* (Cambridge University Press, 1983). doi:citeulike-article-id:4441469.
83. Kingman, J. F. C. On the genealogy of large populations. *Journal of Applied Probability* **19**, 27–43 (1982).

84. Robertson, A. Selection for heterozygotes in small populations. *Genetics* **47**, 1291–1300 (1962).
85. Nei, M. & Roychoudhury, A. K. Probability of fixation and mean fixation time of an overdominant mutation. *Genetics* **74**, 371–380 (1973).
86. Kaplan, N. L., Darden, T. & Hudson, R. R. The coalescent process in models with selection. *Genetics* **120**, 819–829 (1988).
87. Bodmer, W. F. & Felsenstein, J. Linkage and selection: theoretical analysis of the deterministic two locus random mating model. *Genetics* **57**, 237–265 (1967).
88. Karlin, S. General two-locus selection models: Some objectives, results and interpretations. *Theoretical Population Biology* **7**, 364–398 (1975).
89. Bulmer, M. G. The stability of equilibria under selection. *Heredity* **27**, 157–162 (1971).
90. Gillespie, J. H. Pleiotropic overdominance and the maintenance of genetic variation in polygenic characters. *Genetics* **107**, 321–330 (1984).
91. Geritz, S. A. H. & Kisdi, E. Adaptive dynamics in diploid, sexual populations and the evolution of reproductive isolation. *Proc. R. Soc. Lond. B* **267**, 1671–1678 (2000).
92. Glémin, S. Balancing selection in self-fertilizing populations. *Evolution* **75**, 1011–1029 (2021).
93. Arnqvist, G. Assortative mating by fitness and sexually antagonistic genetic variation. *Evolution* **65**, 2111–2116 (2011).
94. Kasimatis, K. R., Ralph, P. L. & Phillips, P. C. Limits to genomic divergence under sexually antagonistic selection. *G3 (Bethesda)* **9**, 3813–3824 (2019).
95. Kimura, M. & Ohta, T. *Theoretical Aspects of Population Genetics*. (Princeton University Press, Princeton, 1971).

96. Hedrick, P. W. Maintenance of genetic polymorphism: spatial selection and self-fertilization. *The American Naturalist* **152**, 145–150 (1998).
97. Glémin, S. Surprising fitness consequences of GC-biased gene conversion: I. Mutation load and inbreeding depression. *Genetics* **185**, 939–959 (2010).
98. Jordan, C. Y. & Connallon, T. Sexually antagonistic polymorphism in simultaneous hermaphrodites. *Evolution* **68**, 3555–3569 (2014).
99. Glemin, S. Balancing selection in self-fertilizing populations. *EVOLUTION* **75**, 1011–1029 (2021).
100. Tazzyman, S. J. & Abbott, J. K. Self-fertilization and inbreeding limit the scope for sexually antagonistic polymorphism. *Journal of Evolutionary Biology* **28**, 723–729 (2015).
101. Olito, C. Consequences of genetic linkage for the maintenance of sexually antagonistic polymorphism in hermaphrodites. *Evolution* **71**, 458–464 (2017).
102. Anderson, W. W. Polymorphism resulting from the mating advantage of rare male genotypes. *Proc. Natl. Acad. Sci. U.S.A.* **64**, 190–197 (1969).
103. Charlesworth, B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* **10**, 195–205 (2009).
104. Hartfield, M., Bataillon, T. & Glémin, S. The evolutionary interplay between adaptation and self-fertilization. *Trends in Genetics* **33**, 420–431 (2017).
105. Nordborg, M., Charlesworth, D. & Charlesworth, B. Increased levels of polymorphism surrounding selectively maintained sites in highly selfing species. *Proc. R. Soc. Lond. B* **263**, 1033–1039 (1996).



106. Wiuf, C., Zhao, K., Innan, H. & Nordborg, M. The probability and chromosomal extent of trans-specific polymorphism. *Genetics* **168**, 2363–2372 (2004).
107. Pamilo, P. Genic variation at sex-linked loci: quantification of regular selection models. *Hereditas* **91**, 129–133 (1979).
108. Rocheleau, G. & Lessard, S. Stability analysis of the partial selfing selection model. *Journal of Mathematical Biology* **40**, 541–574 (2000).
109. Smouse, P. E. The implications of density-dependent population growth for frequency- and density-dependent selection. *The American Naturalist* **110**, 849–860 (1976).
110. Svensson, E. I. & Connallon, T. How frequency-dependent selection affects population fitness, maladaptation and evolutionary rescue. *Evolutionary Applications* **12**, 1243–1258 (2019).
111. Gliddon, C. & Strobeck, C. Necessary and sufficient conditions for multiple-niche polymorphism in haploids. *The American Naturalist* **109**, 233–235 (1975).
112. Czochor, R. J. & Leonard, K. J. Multiple-niche polymorphism in haploid microorganisms. *The American Naturalist* **119**, 293–296 (1982).
113. Dean, A. M. Protecting haploid polymorphisms in temporally variable environments. *Genetics* **169**, 1147–1156 (2005).
114. Débarre, F., Lenormand, T., Debarre, F. & Lenormand, T. Distance-limited dispersal promotes coexistence at habitat boundaries: reconsidering the competitive exclusion principle. *ECOLOGY LETTERS* **14**, 260–266 (2011).
115. Immler, S., Arnqvist, G. & Otto, S. P. Ploidally antagonistic selection maintains stable genetic polymorphism. *Evolution* **66**, 55–65 (2012).
116. Moody, M. Polymorphism with migration and selection. *J. Math. Biology* **8**, 73–109 (1979).

117. Edwards, A. W. F. The population genetics of “sex-ratio” in *Drosophila pseudoobscura*. *Heredity* **16**, 291–304 (1961).
118. Hamilton, W. D. Extraordinary sex ratios: A sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology. *Science* **156**, 477–488 (1967).
119. Mackintosh, C., Pomiankowski, A. & Scott, M. F. X-linked meiotic drive can boost population size and persistence. *Genetics* **217**, iyaa018 (2021).
120. Rice, W. R. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742 (1984).
121. Patten, M. M. & Haig, D. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution* **63**, 2888–2895 (2009).
122. Ruzicka, F. & Connallon, T. Is the X chromosome a hot spot for sexually antagonistic polymorphisms? Biases in current empirical tests of classical theory. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20201869 (2020).
123. Ewens, W. J. *Mathematical Population Genetics 1: Theoretical Introduction*. (Springer, 2004).
124. Connallon, T. & Clark, A. G. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. *Genetics* **190**, 1477–1489 (2012).
125. Mullon, C., Pomiankowski, A. & Reuter, M. The effects of selection and genetic drift on the genomic distribution of sexually antagonistic alleles. *Evolution* **66**, 3743–3753 (2012).

126. Carr, R. N. & Nassar, R. F. Effects of selection and drift on the dynamics of finite populations. I. ultimate probability of fixation of a favorable allele. *Biometrics* **26**, 41 (1970).
127. Ewens, W. J. & Thomson, G. Heterozygote selective advantage. *Annals of Human Genetics* **33**, 365–376 (1970).
128. Mafessoni, F. & Lachmann, M. Selective strolls: fixation and extinction in diploids are slower for weakly selected mutations than for neutral ones. *Genetics* **201**, 1581–1589 (2015).
129. Charlesworth, B. The effects of weak selection on neutral diversity at linked sites. *Genetics* **221**, iyac027 (2022).
130. Feldman, M. W., Franklin, I. & Thomson, G. J. Selection in complex genetic systems I. The symmetric equilibria of the three-locus symmetric viability model. *Genetics* **76**, 135–162 (1974).
131. Gregorius, H.-R. Establishment of allelic two-locus polymorphisms. *J. Math. Biol.* **30**, 185–197 (1991).
132. Kirzhner, V. M., Korol, A. B. & Ronin, Y. I. Cyclical environmental changes as a factor maintaining genetic polymorphism. 1. Two-locus haploid selection. *J of Evolutionary Biology* **8**, 93–120 (1995).
133. Novak, S. & Barton, N. H. When does frequency-independent selection maintain genetic variation? *Genetics* **207**, 653–668 (2017).
134. Nagylaki, T. The deterministic behavior of self-incompatibility alleles. *Genetics* **79**, 545–550 (1975).
135. Bürger, R. Polymorphism in the two-locus Levene model with nonepistatic directional selection. *Theoretical Population Biology* **76**, 214–228 (2009).

136. Patten, M. M., Haig, D. & Úbeda, F. Fitness variation due to sexual antagonism and linkage disequilibrium. *Evolution* **64**, 3638–3642 (2010).
137. Arnqvist, G., Vellnow, N. & Rowe, L. The effect of epistasis on sexually antagonistic genetic variation. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20140489 (2014).
138. Patten, M. M. Meiotic drive influences the outcome of sexually antagonistic selection at a linked locus. *J of Evolutionary Biology* **27**, 2360–2370 (2014).
139. Hastings, A. Unexpected behavior in two locus genetic systems: an analysis of marginal underdominance at a stable equilibrium. *Genetics* **102**, 129–138 (1982).
140. Kelly, J. K. & Wade, M. J. Molecular evolution near a two-locus balanced polymorphism. *Journal of Theoretical Biology* **204**, 83–101 (2000).
141. Navarro, A. & Barton, N. H. The effects of multilocus balancing selection on neutral variability. *Genetics* **161**, 849–863 (2002).
142. Ohta, T. & Kimura, M. Development of associative overdominance through linkage disequilibrium in finite populations. *Genet. Res.* **16**, 165–177 (1970).
143. Zhao, L. & Charlesworth, B. Resolving the conflict between associative overdominance and background selection. *Genetics* **203**, 1315–1334 (2016).
144. Wittmann, M. J., Mousset, S. & Hermisson, J. Modeling the genetic footprint of fluctuating balancing selection: from the local to the genomic scale. *Genetics* **223**, iyad022 (2023).
145. Fisher, R. A. *The Genetical Theory of Natural Selection*. (Clarendon Press, 1930).
146. McCandlish, D. M. & Stoltzfus, A. Modeling evolution using the probability of fixation: History and implications. *Quarterly Review of Biology* **89**, 225–252 (2014).

147. Orr, H. A. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**, 935–949 (1998).
148. McDonough, Y. & Connallon, T. Effects of population size change on the genetics of adaptation following an abrupt change in environment. *Evolution* **77**, 1852–1863 (2023).
149. McDonough, Y., Ruzicka, F. & Connallon, T. Reconciling theories of dominance with the relative rates of adaptive substitution on sex chromosomes and autosomes. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2406335121 (2024).
150. Walsh, B. & Lynch, M. *Evolution and Selection of Quantitative Traits*. (Oxford University Press, 2018).
151. Gavrilets, S. & Hastings, A. Maintenance of genetic variability under strong stabilizing selection: a two-locus model. *Genetics* **134**, 377–386 (1993).
152. Bürger, R. & Gimelfarb, A. Genetic variation maintained in multilocus models of additive quantitative traits under stabilizing selection. *Genetics* **152**, 807–820 (1999).
153. Siljestam, M., Rueffler, C. & Arnqvist, G. Sex-specific dominance and its effects on allelic diversity in sexually antagonistic loci. Preprint at <https://doi.org/10.1101/2024.05.06.592709> (2024).
154. Bürger, R. & Gimelfarb, A. The effects of intraspecific competition and stabilizing selection on a polygenic trait. *Genetics* **167**, 1425–1443 (2004).
155. Spichtig, M. & Kawecki, T. J. The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *The American Naturalist* **164**, 70–84 (2004).
156. Flinham, E., Savolainen, V., Otto, S., Reuter, M. & Mullon, C. The maintenance of genetic polymorphism underlying sexually antagonistic traits. Preprint at <https://doi.org/10.1101/2023.10.10.561678> (2023).

157. Robertson, A. The effect of selection against extreme deviants based on deviation or on homozygosis. *J Genet* **54**, 236–248 (1956).
158. Bulmer, M. G. The maintenance of the genetic variability of polygenic characters by heterozygous advantage. *Genet. Res.* **22**, 9–12 (1973).
159. Barton, N. H. Pleiotropic models of quantitative variation. *Genetics* **124**, 773–782 (1990).
160. Waxman, D. & Welch, J. J. Fisher’s microscope and Haldane’s ellipse. *The American Naturalist* **166**, 447–457 (2005).
161. Martin, G. & Lenormand, T. A general multivariate extension of Fisher’s geometrical model and the distribution of mutation fitness effects across species. *Evolution* **60**, 893–907 (2006).
162. Svensson, E. I. *et al.* Correlational selection in the age of genomics. *Nat Ecol Evol* **5**, 562–573 (2021).
163. Bomblies, K. & Peichel, C. L. Genetics of adaptation. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122152119 (2022).
164. Soni, V., Vos, M. & Eyre-Walker, A. A new test suggests hundreds of amino acid polymorphisms in humans are subject to balancing selection. *PLoS Biol* **20**, e3001645 (2022).
165. Kazancıoğlu, E. & Arnqvist, G. The maintenance of mitochondrial genetic variation by negative frequency-dependent selection. *Ecology Letters* **17**, 22–27 (2014).
166. Haller, B. C. Theoretical and empirical perspectives in ecology and evolution: A survey. *BioScience* **64**, 907–916 (2014).
167. Fitzpatrick, C. L. *et al.* Theory meets empiry: a citation network analysis. *BioScience* **68**, 805–812 (2018).

168. Unckless, R. L., Howick, V. M. & Lazzaro, B. P. Convergent balancing selection on an antimicrobial peptide in *Drosophila*. *Current Biology* (2016)  
doi:10.1016/j.cub.2015.11.063.
169. Yassin, A. *et al.* Ancient balancing selection at tan underlies female colour dimorphism in *Drosophila erecta*. *Nature Communications* **7**, 10400 (2016).
170. Rusuwa, B. B., Chung, H., Allen, S. L., Frentiu, F. D. & Chenoweth, S. F. Natural variation at a single gene generates sexual antagonism across fitness components in *Drosophila*. *Current Biology* **32**, 3161–3169.e7 (2022).
171. Barson, N. J. *et al.* Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **528**, 405–408 (2015).
172. Johnston, S. E. *et al.* Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95 (2013).
173. Sella, G. & Barton, N. H. Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annual Review of Genomics and Human Genetics* **20**, 461–493 (2019).
174. Abdellaoui, A., Yengo, L., Verweij, K. J. H. & Visscher, P. M. 15 years of GWAS discovery: realizing the promise. *The American Journal of Human Genetics* **110**, 179–194 (2023).
175. Moutinho, A. F., Eyre-Walker, A. & Dutheil, J. Y. Strong evidence for the adaptive walk model of gene evolution in *Drosophila* and *Arabidopsis*. *PLoS Biol* **20**, e3001775 (2022).
176. Karageorgi, M. *et al.* Dominance reversal protects large-effect resistance polymorphisms in temporally varying environments. Preprint at <https://doi.org/10.1101/2024.10.23.619953> (2024).

177. Posavi, M., Gelembiuk, G. W., Larget, B. & Lee, C. E. Testing for beneficial reversal of dominance during salinity shifts in the invasive copepod *Eurytemora affinis*, and implications for the maintenance of genetic variation. *Evolution* **68**, 3166–3183 (2014).
178. Grieshop, K. & Arnqvist, G. Sex-specific dominance reversal of genetic variation for fitness. *PLoS Biology* **16**, e2006810 (2018).
179. de Visser, J. A. G. M. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nat Rev Genet* **15**, 480–490 (2014).
180. Wellenreuther, M. & Bernatchez, L. Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology and Evolution* **33**, 427–440 (2018).
181. Lindtke, D. *et al.* Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Molecular Ecology* (2017) doi:10.1111/mec.14280.
182. Mérot, C., Llaurens, V., Normandeau, E., Bernatchez, L. & Wellenreuther, M. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nature Communications* **11**, 670 (2020).
183. Willink, B. *et al.* The genomics and evolution of inter-sexual mimicry and female-limited polymorphisms in damselflies. *Nat Ecol Evol* **8**, 83–97 (2023).
184. Sanjak, J. S., Sidorenko, J., Robinson, M. R., Thornton, K. R. & Visscher, P. M. Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 151–156 (2017).
185. Houle, D., Bolstad, G. H., Van Der Linde, K. & Hansen, T. F. Mutation predicts 40 million years of fly wing evolution. *Nature* **548**, 447–450 (2017).
186. Dugand, R. J., Aguirre, J. D., Hine, E., Blows, M. W. & McGuigan, K. The contribution of mutation and selection to multivariate quantitative genetic variance in an outbred



- population of *Drosophila serrata*. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026217118 (2021).
187. Crow, J. F. Dominance and overdominance. in *Heterosis. A record of researches directed towards explaining and utilizing the vigor of hybrids*. 282–297 (Iowa State University Press, Ames, 1952).
  188. Crow, J. F. Muller, Dobzhansky, and overdominance. *J Hist Biol* **20**, 351–380 (1987).
  189. Zajitschek, F. & Connallon, T. Antagonistic pleiotropy in species with separate sexes, and the maintenance of genetic variation in life-history traits and fitness. *Evolution* **72**, 1306–1316 (2018).
  190. Eyre-Walker, A. & Keightley, P. D. The distribution of fitness effects of new mutations. *Nature Reviews Genetics* **8**, 610–618 (2007).
  191. Wang, Y. & Obbard, D. J. Experimental estimates of germline mutation rate in eukaryotes: a phylogenetic meta-analysis. *Evolution Letters* **7**, 216–226 (2023).
  192. Otto, S. P. & Day, T. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. (Princeton University Press, 2007). doi:10.1515/9781400840915.
  193. Constable, G. W. A. & Kokko, H. The rate of facultative sex governs the number of expected mating types in isogamous species. *Nat Ecol Evol* **2**, 1168–1175 (2018).
  194. Czappon, P. & Constable, G. W. A. Invasion and extinction dynamics of mating types under facultative sexual reproduction. *Genetics* **213**, 567–580 (2019).
  195. Czappon, P. & Billiard, S. Revisiting the number of self-incompatibility alleles in finite populations: From old models to new results. *J of Evolutionary Biology* **35**, 1296–1308 (2022).
  196. Sinervo, B. & Lively, C. M. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature* **380**, 240–243 (1996).

197. Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics*. (Cambridge University Press, 1998). doi:10.1017/CBO9781139173179.
198. Leimar, O. & McNamara, J. M. Game theory in biology: 50 years and onwards. *Phil. Trans. R. Soc. B* **378**, 20210509 (2023).
199. Felsenstein, J. *Theoretical Evolutionary Genetics*. (1976).
200. Haldane, J. B. S. A mathematical theory of natural and artificial selection. (Part VI, Isolation.). *Math. Proc. Camb. Phil. Soc.* **26**, 220–230 (1930).
201. Haldane, J. B. S. The effect of variation on fitness. *The American naturalist* **71**, 337–349 (1937).
202. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics*. *Introduction to quantitative genetics* (Longman, 1996). doi:10.1016/j.bbabi.2008.04.029.