

**Balancing selection in species with separate sexes: Insights from Fisher's geometric model**

Tim Connallon\*, Andrew G. Clark

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

\*Corresponding author. Email: [tmc233@cornell.edu](mailto:tmc233@cornell.edu)

RUNNING TITLE: Balancing selection in species with separate sexes

KEY WORDS: intra-locus sexual conflict, sexual antagonism, heterozygote advantage, pleiotropy, maintenance of polymorphism

## ABSTRACT

How common is balancing selection, and what fraction of phenotypic variance is attributable to balanced polymorphisms? Despite decades of research, answers to these questions remain elusive. Moreover, there is no clear theoretical prediction about the frequency with which balancing selection is expected to arise within a population. Here, we use an extension of Fisher's geometric model of adaptation to predict the probability of balancing selection in a population with separate sexes, wherein polymorphism is potentially maintained by two forms of balancing selection: (1) *heterozygote advantage*, where heterozygous individuals at a locus have higher fitness than homozygous individuals; and (2) *sexually antagonistic selection* (aka intralocus sexual conflict), where the fitness of each sex is maximized by different genotypes at a locus. We show that balancing selection is common under biologically plausible conditions, and that sex differences in selection or sex-by-genotype effects of mutations can each increase opportunities for balancing selection. Although heterozygote advantage and sexual antagonism represent alternative mechanisms for maintaining polymorphism, they mutually exist along a balancing selection continuum that depends on population and sex-specific parameters of selection and mutation. Sexual antagonism is the dominant mode of balancing selection across most of this continuum.

## INTRODUCTION

One of the primary goals of population genetics is to evaluate how processes of mutation, selection, migration, recombination, and genetic drift account for the maintenance of genetic variation in natural populations. This “Great Obsession” with genetic variation (Gillespie 2004, p. 154) is reflected by a massive body of theoretical and empirical research that seeks to connect empirical patterns of population and quantitative genetic variability with the specific evolutionary processes that might account for them.

Balancing selection – selection that maintains genetic variation at a locus – could potentially account for an important fraction of observed quantitative genetic variability (Dobzhansky 1955; Lewontin 1974; Charlesworth and Hughes 1999). Although relatively few unambiguous cases of balancing selection have been documented to date (see Charlesworth 2006; Hedrick 2012; Leffler et al. 2013; Johnston et al. 2013), empirical methods for detecting their population genetic signatures are highly conservative and may fail to identify many (or most) instances of balancing selection within a genome (Charlesworth 2006; Charlesworth and Charlesworth 2010). Furthermore, balanced polymorphisms may plausibly account for several empirical observations that are not easily explained by alternative models of variation maintained by recurrent mutation. For example, the high genetic variance in life history and reproductive traits (Houle 1992; Pomiankowski and Moller 1995), and the presence of intermediate frequency alleles with large phenotypic effects (Long et al. 2000), are each potentially attributable to a subset of loci that segregate for balanced alleles (Charlesworth and Hughes 1999; Turelli and Barton 2004).

Population genetic theory provides a complementary approach for assessing the plausibility of different scenarios of balancing selection, and indeed, there are many formal models that delineate the parameter conditions required for balancing selection to operate (*e.g.*, Wright 1969;

Crow and Kimura 1970; Prout 2000). However, most theory cannot predict the *probability* of balancing selection because it does not incorporate details of the fitness effect distribution among random alleles and genotypes in a population. Such details ultimately determine the relative likelihood of different forms of selection at individual loci.

A recent study by Sellis et al. (2011) provides new insight into the probability of balancing selection by modeling the frequency of heterozygote advantage in Fisher's geometric model (FGM; Fisher 1930; Orr 2005a, b) – an influential theoretical framework for the fitness effect distribution of random mutations. Sellis et al. (2011) showed that, among loci contributing to adaptive evolutionary change, the probability of heterozygote advantage tends to increase with the “standardized size” (Orr 1998) of adaptive mutations. Under Fisher's original scaling, mutation size is given by  $x = r\sqrt{n}/(2z)$ , where  $r$  is the absolute phenotypic effect size of the mutation,  $n$  is the number of traits, and  $z$  represents the displacement of the population from a fitness optimum within phenotypic space (see Fisher 1930, pp. 38-41; Orr 1998). A mutation of size  $r$  is effectively small when the distance to the optimum is large (*e.g.* in the limit  $r/z \rightarrow 0$ ), and effectively large when the distance to the optimum is small (*e.g.*  $r \rightarrow z$ ; for a clear discussion of  $x$  and its biological meaning, see Orr 1998, pp. 937-938). Adaptation – the evolutionary movement of a population toward its optimum – causes the distance to the optimum to shrink (Orr 1998), and thereby increases the scaled sizes of random mutations. Poorly adapted populations that are far from an optimum may initially experience little opportunity for heterozygote advantage, yet adaptive evolution will ultimately overturn such initially unfavorable conditions. During the course of an adaptive walk within phenotypic space, mutations eventually become sufficiently “large” for heterozygote advantage to be likely (Sellis et al. 2011; for related results, see Manna, Martin, Lenormand 2011).

In species with separate sexes, balancing selection can potentially arise by way of heterozygote advantage, or by sexual antagonism (aka “intralocus sexual conflict”), in which the best genotype for females differs from the best genotype for males (Rice 1992; Chippindale et al. 2001). Whether balanced, sexual antagonistic alleles are common within populations has important and wide-ranging biological implications for the genetic basis of quantitative traits (Rice 1984; Turelli and Barton 2004; Bonduriansky and Chenoweth 2009), genetic loads and extinction risk (Kokko and Brooks 2003; Whitlock and Agrawal 2009; Connallon et al. 2010; Holman and Kokko 2013), mating system evolution (Seeger and Trivers 1986; Albert and Otto 2005; Blackburn et al. 2010; Roze and Otto 2012), and the evolution of genomes and genetic systems (Charlesworth and Charlesworth 1980; Day and Bonduriansky 2004; Ellegren and Parsch 2007, 2013; Mank 2009; Connallon and Clark 2010, 2011, 2013; van Doorn and Kirkpatrick 2007; 2010; Wright and Mank 2013; Kirkpatrick and Guerrero 2014; Charlesworth et al. 2014). Although prior theory clearly defines the parameter criteria for balancing selection by sexual antagonism (*e.g.*, Kidwell et al. 1977; Pamilo 1979; Patten and Haig 2009; Unckless and Herren 2009; Fry 2010; Patten et al. 2010; Arnqvist 2011; Connallon and Clark 2012; Mullon et al. 2012; Jordan and Charlesworth 2012; Patten et al. 2013), it remains unclear how often such conditions might be expected to arise in dioecious populations. A recent extension of Fisher’s geometric model provides a theoretical framework for predicting the sex-specific distribution of mutant fitness effects (Connallon and Clark 2014), yet this theory has not addressed opportunities for balancing selection.

Here, we present new theoretical predictions for the probability of balancing selection in a species with distinct sexes. Our models build upon a recent extension of Fisher’s geometric model (Connallon and Clark 2014) that incorporates two evolutionarily important features of

dioecious species: (1) sexually dimorphic fitness landscapes, which generate sex-specific patterns of selection on phenotypes expressed by each sex; and (2) sex-by-genotype interactions (a form of gene-by-environment interaction during development), in which mutations have sexually dimorphic effects on female and male phenotypes. We first extend Fisher's original scaling function by developing a mathematical expression for the standardized mutation size in a dioecious population. We show that this new scaling function can be used to calculate several interesting metrics in FGM, including: (1) the probability that random mutations experience positive, purifying, or balancing selection; and (2) the critical mutation size that maximizes the probability of balancing selection. We show that sex differences in selection on phenotypes and genotype-by-sex interactions each tend to inflate the standardized mutation size, and that these effects increase opportunities for balancing selection in the FGM framework. Finally, we demonstrate that, for most of the parameter space of sex-specific selection and gene-by-sex interactions, sexual antagonism is a pervasive feature of balancing selection.

## **MODEL**

### **Fisher's geometric model (FGM) with two sexes**

The basic model analyzed here is a diploid extension of the haploid, two-sex FGM model that was recently developed by Connallon and Clark (2014). Male and female phenotypes are each characterized by a vector of  $n$  trait values, with each vector representing a specific location in  $n$ -dimensional phenotypic space. For mathematical convenience, and following prior work in Fisher's model (Fisher 1930; Orr 1998; Sellis et al. 2011; Connallon and Clark 2014; see below), we assume that: (1) sex-specific fitness is a function of the Euclidean distance to the relevant optimum; and (2) each mutation has a random, unbiased orientation within phenotypic space.

Relaxation of these assumptions – for example, by allowing for elliptical fitness surfaces and/or mutational biases – often decreases the “effective” dimensionality in Fisher’s geometric model (see Waxman and Welch 2005; Martin and Lenormand 2006; Martin 2014), such that  $n$  may be interpreted as the effective number (rather than the total number) of independent traits.

The initial population is assumed to be fixed for a wild type genotype, for which males and females express phenotypic values of  $A_m$  and  $A_f$ , respectively, where  $A_j = \{x_{1j}, x_{2j}, \dots, x_{nj}\}$  is a vector of  $n$  trait values for the  $j$ th sex;  $j = \{f, m\}$ . Optimal trait combinations for males and females are depicted by vectors  $O_m$  and  $O_f$ , which describe the locations of sex-specific optima for the  $n$  traits. The direction of selection acting in the two sexes can be described using a pair of vectors (one for each sex) that point from  $A_j$  to  $O_j$  (Fig. 1). The lengths of the vectors are given by  $z_m$  and  $z_f$ , and represent the displacement of each sex from its optimum (*i.e.*, the Euclidean distance between  $A_j$  and  $O_j$ ). The angle between vectors is  $\theta_{sel}$ , which potentially takes any value between zero (where directions of sex-specific selection are the same in males and females) and  $\pi$  (where the sexes are selected in opposite directions). The degree of correlation between male and female orientations of directional selection is  $\rho_{sel} = \cos(\theta_{sel})$  (Rodgers and Nicewander 1988; Connallon and Clark 2014). Sex-specific fitness of females and males, respectively, follows a Gaussian function:

$$w_f = \exp(-\omega_f z_f^2) \tag{1a},$$

and

$$w_m = \exp(-\omega_m z_m^2) \tag{1b},$$

where  $\omega_f$  and  $\omega_m$  are positive constants that describe the rate of fitness decline away from each optimum.

The phenotypic effects of each mutation can be similarly summarized using sex-specific vectors (Fig. 1). Each mutation has a magnitude, with  $r_j$  representing the phenotypic (*i.e.*, Euclidean) distance between  $A_j$  and the location of individuals that are homozygous for the mutation. The magnitude in a heterozygote is  $r_j h$ , where  $h$  represents the dominance of the mutation with respect to the phenotype (*i.e.*,  $0 < h < 1$ , with  $h = 0$ ,  $h = 1/2$ , and  $h = 1$  corresponding to recessive, additive or codominant, and dominant effects, respectively). Mutational orientations are assumed to be random and unbiased in  $n$ -dimensional space. The degree of similarity between the sexes, for each mutation's orientation in phenotypic space, depends on the between-sex correlation of phenotypic effects within individual trait axes. For a mutation with magnitude  $r_m$  and  $r_f$ ,  $\rho_{mut}$  represents both the phenotypic correlation between the sexes for each trait, and the mean degree of correlation between mutant vectors ( $\rho_{mut} = E[\cos(\theta_{mut})]$ , where  $\theta_{mut}$  is the angle between the vectors of a mutation, and  $E[\ ]$  denotes the expectation; see Connallon and Clark 2014).

Taking both selection and mutation into consideration, we can calculate distances between the optimum  $O_j$ , and the phenotypic position of individuals that carry zero, one, or two copies of a given mutation. Letting  $O_j' = O_j - A_j = \{o_{1j}, o_{2j}, \dots, o_{nj}\}$ , the distance to the optimum in heterozygous males and females will be:

$$z_{m(het)} = \sqrt{z_m^2 + (hr_m)^2 - \frac{2hr_m}{\sqrt{\sum_i m_i^2}} \sum_{i=1}^n o_{im} m_i} \approx \sqrt{z_m^2 + (hr_m)^2 - \frac{2hr_m}{\sqrt{n}} \sum_{i=1}^n o_{im} m_i} \quad (2a)$$

and

$$z_{f(het)} = \sqrt{z_f^2 + (hr_f)^2 - \frac{2hr_f}{\sqrt{\sum_i f_i^2}} \sum_{i=1}^n o_{if} f_i} \approx \sqrt{z_f^2 + (hr_f)^2 - \frac{2hr_f}{\sqrt{n}} \sum_{i=1}^n o_{if} f_i} \quad (2b)$$

(Appendix A), where the values of  $m_i$  and  $f_i$  are obtained from  $n$  independent draws from a bivariate standard normal distribution with  $\text{cov}(m_i, f_i) = \rho_{mut}$  (Connallon and Clark 2014). These relationships naturally emerge from the mutation algorithm described further below (see “Simulations” section), with approximations on the right hand side of eqs. (2a) and (2b) accurate for values of  $n = 10$  or greater (Welch and Waxman 2003; Connallon and Clark 2014). Distances to the optima for homozygous carriers of the mutation,  $z_{j(hom)}$ , can be obtained by setting  $h = 1$  in eq. (2). Heterozygous and homozygous selection coefficients for mutations can be calculated directly from eqs. (1-2) (Appendix A), with homozygous selection coefficients defined as  $s_{j2} = \exp[\omega_j(z_j^2 - z_{j(hom)}^2)] - 1$ , and heterozygous selection coefficients defined as  $s_{j1} = \exp[\omega_j(z_j^2 - z_{j(het)}^2)] - 1$ .

### **Criteria for positive selection and balancing selection**

Balancing selection criteria follow the well-established theory of sex-specific selection at a diploid locus with two alleles (*e.g.*, Kidwell et al. 1977). Let  $B_1$  represent an ancestral allele, and  $B_2$  the derived allele. The relative fitnesses for the three genotypes are provided in Table 1. We assume that generations are discrete, and follow the life cycle: birth, selection prior to mating, and random mating among breeding adults. In cases where selection differs between males and females, the frequency of alleles inherited from mothers and from fathers will differ. Following convention, we define  $q_m$  and  $q_f$  as the frequency of  $B_2$  alleles among the male and female gametes that contribute to a given generation.

Following standard criteria for determining evolutionary stability (see Otto and Day 2007), a new (rare) mutation will experience selection to increase in frequency when the boundary

equilibrium  $q_m, q_f = 0$  is unstable. The criterion for selection to favor invasion of a rare mutation is:

$$0 < s_{f1} + s_{m1} \tag{3},$$

Mutations that meet this criterion can be further divided into mutually exclusive subsets: (1) positively selected mutations that are also favored to fix in the population; and (2) mutations under balancing selection, which can invade from low frequency, but are not favored to fix. For the latter case, selection favors the maintenance of ancestral (wild-type) and mutant alleles.

Balancing selection occurs when both boundary equilibria ( $q_m, q_f = 0; q_m, q_f = 1$ ) are unstable, a condition often referred to as a “protected polymorphism” (Prout 1968). The condition for balancing selection is:

$$0 < s_{f1} + s_{m1} > s_{f2}(1 + s_{m2} - s_{m1}) + s_{m2}(1 + s_{f2} - s_{f1}) \tag{4}.$$

We define positive selection when  $q_m, q_f = 0$  is unstable ( $B_2$  invades when rare), and  $q_m, q_f = 1$  is stable ( $B_1$  does not invade when rare).

## Simulations

In our subsequent analytical results, we assume that: (1) mutations have small fitness effects, as is generally expected (*e.g.*, Orr 2005a, 2005b, 2006; analytical results are appropriate for selection coefficients within the approximate range:  $s_{f1}, s_{f2}, s_{m1}, s_{m2} < 0.1$ ); and (2) dimensionality is sufficiently high so that the approximations in eq. (2) remain valid (*i.e.*, that  $n > 10$ ; Welch and Waxman 2003; Connallon and Clark 2014). Analytical results were verified using an exact simulation approach described in Connallon and Clark (2014), with adjustments for diploid inheritance. Without any loss of generality, we define the phenotypic positions of wild-type males and females, and their respective optima, so that  $A_f = A_m = \{0, 0, \dots, 0\}$ ,  $O_f = \{z_f, 0, \dots, 0\}$ , and  $O_m = \{z_m \cos(\theta_{sel}), z_m \sin(\theta_{sel}), 0, \dots, 0\}$ . Mutation magnitudes for each sex ( $r_m, r_f$ ) were

generated using a bivariate gamma distribution with equal male and female marginal distributions (values of  $r_m$  and  $r_f$  were obtained using the “mixture approach” algorithm, described in Michael and Schucany 2002). Within the  $j$ th sex, the phenotypic location of individuals homozygous for a random mutation is described by a vector,  $y_j = \{y_{1j}, y_{2j}, \dots, y_{nj}\}$ , with the elements of the vector representing a set of phenotypic changes in the  $n$  traits. For a mutation with magnitude  $r_m$  in males,  $y_{im} = r_m m_i / M$ , where  $M^2 = \sum_{i=1}^n m_i^2$ , and the  $m_i$  are independent draws from a standard normal distribution. This sampling strategy generates random, uniformly distributed points on the surface of an  $n$ -sphere with radius  $r_m$  (Muller 1959; Marsaglia 1972; Connallon and Clark 2014). For a mutation with magnitude  $r_f$  in females,  $y_{if} = r_f f_i / F$ , where  $F^2 = \sum_{i=1}^n f_i^2$ , and the  $f_i$  are independent draws from a standard normal distribution with  $\text{cov}(m_i, f_i) = \rho_{mut}$ . For a mutation with homozygous effect  $y_j = \{y_{1j}, y_{2j}, \dots, y_{nj}\}$ , the corresponding heterozygous effect is modified by dominance, so that  $y_{jh} = \{y_{1jh}, y_{2jh}, \dots, y_{njh}\}$  (following Sellis et al. 2011). Once the  $y_{im}$  and  $y_{if}$  are specified for an individual mutation, it is straightforward to calculate heterozygous and homozygous selection coefficients for each sex, by using eqs. (1-2) (see Appendix A).

## RESULTS

### Mutation size and the probability of balancing selection

The proportion of random mutations that experience positive selection, purifying selection, and balancing selection, can be directly calculated using the selection criteria outlined above (eqs. 3-4), and the distribution of fitness effects that emerges from the two-sex Fisher’s geometric model (eqs. 1-2) (see Appendix B). Following Fisher’s original concept of the standardized mutation size (Fisher 1930; Orr 1998), let  $x_A$  represent an adjusted standardized

“size” that accounts for sex differences in directional selection, and sex-by-genotype phenotypic effects of mutations. This adjusted mutation size is defined as:

$$x_A = \frac{(\omega_m r_m^2 + \omega_f r_f^2) \sqrt{n}}{2 \sqrt{(\omega_m z_m r_m)^2 + (\omega_f z_f r_f)^2 + 2 \omega_m z_m r_m \omega_f z_f r_f \rho_{mf}}} \quad (5).$$

Mutation size is now a function of sex-specific selection and mutation parameters ( $\omega_j, r_j, z_j$ ), dimensionality ( $n$ ), and the between-sex correlation of fitness effects:  $\rho_{mf} = \rho_{sel} \rho_{mut}$  (see Model section, above). Note that, by removing all possible sex differences (letting  $\rho_{mf} = 1$ ,  $\omega_m = \omega_f$ ,  $r_m = r_f$ ,  $z_m = z_f$ ), eq. (5) collapses to Fisher’s (1930) original expression:  $x = r \sqrt{(n)/(2z)}$ .

The individual probabilities of purifying, positive, and balancing selection are simple functions of the new mutation size scaling (Appendix B). The probability that selection favors invasion of a random mutation with size  $x_A$  is:

$$\Pr(s_{m1} + s_{f1} > 0) \approx \frac{1}{\sqrt{2\pi}} \int_{x_A h}^{\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi[x_A h] \quad (6),$$

where  $\Phi(hx_A)$  is the cumulative standard normal distribution, and  $h$  is the dominance of the mutation relative to the wild type allele ( $0 < h < 1$ ). In the absence of sex differences, where  $x_A = r \sqrt{(n)/(2z)}$ , eq. (6) is identical to previous, high-dimension approximations ( $n > 10$ ) for the probability of a beneficial mutation in Fisher’s geometric model (Fisher 1930; Kimura 1983; Orr 1998; Sellis et al. 2011).

Alleles selected to invade the population may go to fixation or evolve to an intermediate balanced polymorphic state. These outcomes – positive and balancing selection, respectively – are subsets of the probability space described by eq. (6), and are mutually exclusive. The probability that a random mutation evolves under balancing selection is:

$$\Pr(\text{balancing}) \approx \frac{1}{\sqrt{2\pi}} \int_{x_A h}^{x_A \frac{1-h^2}{1-h}} e^{-\frac{t^2}{2}} dt = \Phi \left[ x_A \frac{1-h^2}{1-h} \right] - \Phi [x_A h] \quad (7),$$

where  $\Phi$  again refers to the cumulative standard normal. The probability of positive selection is therefore:  $\Pr(\text{positive}) = \Pr(s_{f1} + s_{m1} > 0) - \Pr(\text{balancing}) = 1 - \Phi[x_A(1 - h^2)/(1 - h)]$ . Relative probabilities of the three modes of selection – purifying, positive, and balancing – are shown in Figure 2A. Positive selection is more common than balancing selection when scaled mutation size is small (roughly,  $x_A < 1/2$ ). For larger values of  $x_A$ , a majority of adaptive mutations evolve under balancing selection, rather than positive selection, with the probability of balancing selection ultimately converging, with increasing mutation size, to  $\Pr(\text{balancing}) = 1 - \Phi(x_A h)$ . In other words, as mutation size increases, an ever increasing share of adaptive mutations (mutations favored to invade when rare) will experience balancing selection.

As originally noted by Fisher (1930), mutations with small phenotypic effects ( $x_A \rightarrow 0$ ) are likely to improve fitness, whereas large mutations have small probabilities of doing so. A scaled mutation size near zero will therefore maximize the probability of positive selection and minimize the probability of purifying selection. In contrast, the probability of balancing selection is maximized for mutations of intermediate size. The value of  $x_A$  that maximizes the probability of balancing selection is:

$$\hat{x}_A = (1 - h) \sqrt{\frac{2}{1 - h^2(3 - 2h)} \ln \left[ \frac{(1 - h^2)}{h(1 - h)} \right]} \quad (8),$$

(Appendix C). In the absence of dominance between mutant and wild-type alleles ( $h = 1/2$ ), this reduces to  $\hat{x}_A = \sqrt{\ln(3)} \approx 1.048$ , with partial recessivity ( $h < 1/2$ ) increasing this critical mutation effect size, and partial dominance decreasing it (Fig. 2B; gray dotted line). The maximum probability of balancing selection can also be calculated by evaluating eq. (7), using the critical

mutation size of eq. (8). When phenotypic effects are codominant ( $h = 1/2$ ), a maximum of ~24% of mutations experience balancing selection; the probability of balancing selection further increases with partial recessivity, and decreases with partial dominance (Fig. 2B).

### Mean mutation size in species with separate sexes

Mutation size ( $x_A$ ) clearly influences the probability that random mutations experience purifying, positive, or balancing selection. Small-sized mutations evolve primarily under positive or purifying selection (in the limit,  $x_A \rightarrow 0$ , there is a roughly 50% probability of each type of selection). For larger  $x_A$  values, most mutations experience purifying selection, but those that do not generally evolve under balancing selection (see above; Fig. 2). How will sex-specific selection and sex-by-genotype effects influence the average magnitude of  $x_A$ ? As demonstrated below, both factors inflate the scaled mutation size, and should thereby expand opportunities for balancing selection relative to positive selection.

An intuitive way to view the scaled mutation size is by rewriting it as a function of the strength of directional selection in each sex. The strength of directional selection in the  $j$ th sex can be quantified as  $\beta_j = |\text{dln}(w_j)/\text{dz}_j| = 2\omega_j z_j$ , which is conceptually similar to the selection gradient of quantitative genetics (Lande 1980; Falconer and Mackay 1996), and is roughly proportional to the expected rate of adaptation in Fisher's geometric model (Orr 2000; Connallon and Clark 2014). Mean mutation size, as a function of the strength and orientation of sex-specific selection, is:

$$E(x_A) = E \left[ \frac{(\omega_m r_m^2 + \omega_f r_f^2) \sqrt{n}}{\sqrt{(\beta_m r_m)^2 + (\beta_f r_f)^2 + 2\beta_m \beta_f r_m r_f \rho_{mf}}} \right] \quad (9),$$

which takes into account the probability density of male and female mutational magnitudes among random mutations. To illustrate how sex differences in selection or mutational effects will affect  $E(x_A)$ , it is useful to consider two extremes of the possible values of  $\beta_m$  relative to  $\beta_f$ : (1) highly asymmetric directional selection (*i.e.*,  $\beta_m/\beta_f \rightarrow 0$ , or  $\beta_m/\beta_f \rightarrow \infty$ ), which may be plausible immediately after a change in the environment; and (2) equally strong directional selection in each sex ( $\beta_m/\beta_f \rightarrow 1$ ), which is the long-run expectation during adaptive evolution (Lande 1980; Connallon and Clark 2014).

When directional selection is highly asymmetric, the scaled mutation size reduces to  $x_A \approx (\omega_m r_m^2 + \omega_f r_f^2)(n)^{0.5}/(\beta_j r_j)$ , where  $j$  refers to the sex experiencing directional selection (there is no directional selection in the other sex if it is at its optimum, which we assume here). We arbitrarily present results for the case where males alone experience directional selection ( $\beta_m/\beta_f \rightarrow \infty$ ; similar results are obtained for the opposite case of  $\beta_m/\beta_f \rightarrow 0$ ). The average scaled size of a mutation can be approximated using a Taylor series expansion of  $x_A$ , to second order with respect to the mean mutation size of each sex. We assume that mutation size follows a bivariate gamma distribution of mutational magnitudes, with equal male and female marginal distributions [ $E(r) = E(r_m) = E(r_f)$ ;  $\text{var}(r) = \text{var}(r_m) = \text{var}(r_f)$ ], and a between sex correlation of mutational magnitudes of  $\text{corr}(r_m, r_f) = \text{cov}(r_m, r_f)/\text{var}(r)$ . The mean size of a random mutation is then:

$$\begin{aligned}
 E(x_A) &\approx \frac{E(r)(\omega_m + \omega_f)\sqrt{n}}{\beta_m} + \frac{2\omega_f\sqrt{n}}{\beta_m} \frac{\text{var}(r)[1 - \text{corr}(r_m, r_f)]}{E(r)} \\
 &= E(x_m) \left( 1 + \frac{\omega_f}{\omega_m} + \frac{\omega_f}{\omega_m} \frac{2[1 - \text{corr}(r_m, r_f)]}{k} \right)
 \end{aligned} \tag{10},$$

(Appendix D) where  $k$  is the shape parameter of the gamma distribution ( $k > 0$ ), and

$x_m = r_m \sqrt{n}/(2z_m)$  is the male-specific scaled mutation size. From eq. (10) we can visualize three factors that inflate the mean scaled mutation size.  $E(x_A)$  increases as the relative strength of

stabilizing selection in females increases ( $\omega_f/\omega_m$  increases), and as the distribution of mutation magnitudes becomes increasingly leptokurtic (where  $k$  decreases, which skews the distribution towards small values), and the correlation between sexes becomes small to negative ( $\text{corr}(r_m, r_f)$  decreases). The scaled mutation size is always elevated as a consequence of selection in females, and the magnitude of this elevation is substantial when the ratio  $\omega_f/\omega_m$  is moderate to large (e.g., for  $\omega_f/\omega_m > 1/2$  and  $\text{corr}(r_m, r_f) > 0$ ,  $E(x_A)$  is inflated by at least 50 percent relative to the case of no selection in females).

For the scenario of equally strong directional selection in each sex ( $\beta = \beta_m = \beta_f$ ), the average scaled mutation size becomes:

$$E(x_A) \approx \omega \frac{E(r)\sqrt{n}}{\beta} \left[ \frac{2}{\sqrt{2(1+\rho_{mf})}} + \frac{[1 - \text{corr}(r_m, r_f)](1 + 3\rho_{mf})}{2k(1+\rho_{mf})\sqrt{2(1+\rho_{mf})}} \right] \quad (11)$$

(Appendix D), where  $\omega = (\omega_m + \omega_f)/2$ .  $E(x_A)$  now depends on the interaction between three forms of correlation between the sexes: (1) the correlation of mutant magnitudes,  $\text{corr}(r_m, r_f)$ ; (2) the correlation of male and female selection orientations,  $\cos(\theta_{sel})$ ; and (3) the mean correlation of mutational orientations,  $E[\cos(\theta_{mut})]$  (recall that  $\rho_{mf} = \cos(\theta_{sel})E[\cos(\theta_{mut})]$ ).  $E(x_A)$  is inflated by sex differences in selection or mutational effects when the bracketed term in eq. (11) is greater than one. When mutant phenotypic effects are strongly correlated between the sexes [ $\text{corr}(r_m, r_f)$ ,  $\cos(\theta_{mut}) \rightarrow 1$ ], the bracketed term reduces to  $2[2 + 2\cos(\theta_{sel})]^{-1/2}$ , which increases as the orientations of selection become increasingly divergent between the sexes ( $\theta_{sel} > 0$ ). When phenotypic correlations between the sexes are weak [ $\text{corr}(r_m, r_f)$ ,  $\cos(\theta_{mut}) \rightarrow 0$ ],  $E(x_A)$  is elevated by a factor of  $(1 + 4k)/(2k\sqrt{2})$ , or at least  $\sim 1.4$ -fold. Figure 3 shows the magnitude of increase for  $E(x_A)$ , across a range of conditions for sex-specific selection and mutational effects.  $E(x_A)$  is inflated across the entire range of parameter space, and particularly so when the sexes are

selected in different directions ( $\cos(\theta_{sel}) \ll 1$ ) or the phenotypic effects of mutations are strongly decoupled between the sexes ( $0 < \rho_{mut} = \text{corr}(r_m, r_f) = E[\cos(\theta_{mut})] \ll 1$ ).

### **Mechanisms of balancing selection**

There are two general mechanisms that may account for individual episodes of balancing selection in the two-sex geometric model considered here. Variation may be maintained by heterozygote advantage in both sexes ( $s_{m2} < s_{m1} > 0$ , and  $s_{f2} < s_{f1} > 0$ ), or by sexually antagonistic (SA) selection, which in its broadest sense occurs when male and female fitnesses are maximized by a different genotype at a locus. Sexual antagonism models can be further divided into two subforms: (1) “directional” SA selection, where the fitness of males and females is maximized by different homozygous genotypes at the locus, leading to positive selection in one sex and purifying selection on the other ( $s_{m2} < s_{m1} < 0 < s_{f1} < s_{f2}$ , or  $s_{f2} < s_{f1} < 0 < s_{m1} < s_{m2}$ ); and (2) “mixed” SA selection, where there is positive or purifying selection in one sex, and heterozygote advantage in the other ( $s_{m2} < s_{m1} > 0$ , or  $s_{f2} < s_{f1} > 0$ ).

To estimate the relative frequency of each balancing selection mechanism, we simulated random mutations in Fisher’s geometric model, and classified the proportion of balancing selection cases that were attributable to heterozygote advantage, directional SA, or mixed SA. We first considered a scenario where females were perfectly adapted, and males were displaced from their fitness optimum ( $z_f = 0$  and  $z_m > 0$ ; equivalent results apply when  $z_f > 0$  and  $z_m = 0$ ). Under this strongly asymmetric scenario, all mutations are deleterious to females and heterozygote advantage in both sexes is not possible. Directional SA selection becomes the dominant mechanism of balancing selection when scaled mutation sizes are small or mutations have strongly correlated phenotypic effects between the sexes (*e.g.*, Fig. 4, top left panel). As

phenotypic correlations decrease and the “size” of mutations increases, the mixed SA selection scenario predominates (Figs. 4; Supplementary Material Figs. S1-S3).

We also considered scenarios where both sexes were under similarly strong directional selection ( $\beta_f = \beta_m$ ). In this case, heterozygote advantage remains rare unless the orientation of directional selection and the phenotypic effects of mutations are *both* strongly correlated between the sexes (*e.g.*,  $\rho_{sel}, \rho_{mut} \rightarrow 1$ ; Figs. 4; Supplementary Material Figs. S1-S3). Under the remaining conditions, sexual antagonism is the predominant driver of balancing selection, with the mixed SA model representing the most common mechanism across most of the parameter space. Directional SA selection dominates when males and females are selected in different directions, mutational effects are strongly correlated between the sexes, and the scaled mutation size is small (Figs. 4; Supplementary Material Figs. S1-S3).

### **The efficacy of balancing selection**

The above results (*e.g.*, eq. (7)) describe the probability of balancing selection in Fisher’s geometric model, but they do not account for evolutionary dynamics under balancing selection, in which genetic drift is expected to play an important role (Robertson 1962; Ewens and Thomson 1970; Nei and Roychoudhury 1973). To identify criteria for “effective” balancing selection – the degree to which allele frequencies in a finite population converge to the deterministic equilibrium,  $\hat{q}$  – we analyzed the stationary distribution for random mutations within the FGM framework. For a diploid locus in a Wright-Fisher population of effective size  $N_e$ , the stationary distribution of a  $B_2$  allele is proportional to:

$$f(q) \propto \frac{\exp\left[4N_e \int \frac{\Delta q}{q(1-q)} dq\right]}{q(1-q)} \quad (12)$$

(Ewens 2004; Rice 2004; Charlesworth and Charlesworth 2010), where  $q$  is the population frequency of  $B_2$ , and  $\Delta q$  is the expected frequency change of  $B_2$  over a single generation.

Assuming selection coefficients are small ( $|s_{f1}|, |s_{f2}|, |s_{m1}|, |s_{m2}| \ll 1$ ), the expected frequency change is:

$$\Delta q \approx \frac{[2(s_{m1} + s_{f1}) - (s_{f2} + s_{m2})]}{2} q(1-q)(\hat{q} - q) \quad (13),$$

where  $\hat{q} = \frac{s_{f1} + s_{m1}}{2(s_{m1} + s_{f1}) - (s_{f2} + s_{m2})}$  represents the polymorphic equilibrium under balancing selection (see Appendix E).

The relative importance of selection versus drift depends on the density of the stationary distribution near the equilibrium ( $\hat{q}$ ), relative to the density near  $q = 0$  and  $q = 1$ , with effectively strong selection leading to a high relative density near  $\hat{q}$  (Ewens 2004; pp. 26-27). The ratio of the density near the equilibrium, relative to the density near the boundaries, can be described by the ratio:

$$R = \frac{\int_{\hat{q}-\varepsilon}^{\hat{q}+\varepsilon} f(q) dq}{\int_{0.001-\varepsilon}^{0.001+\varepsilon} f(q) dq + \int_{0.999-\varepsilon}^{0.999+\varepsilon} f(q) dq} \approx \frac{1}{10^3 \hat{q}(1-\hat{q})(\exp[-N_e \alpha \hat{q}^2] + \exp[-N_e \alpha (1-\hat{q})^2])} + O(\varepsilon^3) \quad (14)$$

(Appendix F), where  $\alpha = [2(s_{m1} + s_{f1}) - (s_{f2} + s_{m2})]$  represents the net strength of selection at the locus, and  $2\varepsilon$  is an arbitrarily small allele frequency interval. Eq. (14) summarizes how the efficacy of balancing selection increases with the net strength of selection ( $\alpha$ ), as equilibrium values become increasingly intermediate ( $\hat{q} \rightarrow 1/2$ ), and with increasing  $N_e$ . While there is no specific cutoff to delineate weak from strong balancing selection, values of  $R > 10$  (assuming a sufficiently intermediate equilibrium, *i.e.*:  $\hat{q}(1-\hat{q}) > 0.001$ ) are generally sufficient for

maintaining polymorphism, and we use this cutoff as a guideline for interpreting subsequent results. (Note that this guideline shares properties with other diffusion-based metrics, such as mean heterozygosity under balancing selection, and allele fixation rates or times at polymorphic loci; *cf.* Robertson 1962; Ewens and Thomson 1970; Hedrick 1974; Nei and Roychoudhury 1973; Ewens 2004; Connallon and Clark 2012; Mullon et al. 2012).

To evaluate opportunities for effective balancing selection, and to contrast the efficacy under distinct mechanisms of sex-specific selection, we simulated random mutations and characterized the distribution of  $\hat{q}$  and  $\alpha$  for mutations meeting balancing selection criteria (eq. (4)). We again considered cases of symmetric versus asymmetric directional selection ( $\beta_f/\beta_m = 0$  and  $\beta_f/\beta_m = 1$ ) across a broad range of mutation sizes, selection and mutation correlations between the sexes ( $\rho_{sel}, \rho_{mut}$ ), and dominance ( $h$ ). Specific values of  $\hat{q}$  and  $\alpha$  vary widely among mutations subject to balancing selection, with sexually antagonistic mutations consistently more susceptible to drift than mutations under heterozygote advantage (Figs. 5, S4). Mutations subject to directional SA (positive selection in one sex, in purifying selection in the other) are associated with substantially lower values of net selection ( $\alpha$ ), on average, and are the most likely to evolve primarily by drift (Figs. 5, S4; note that the  $\alpha$  values are plotted on a logarithmic scale in Figs. 5, S4-S8).

Mutations causing heterozygote advantage in one sex (mixed SA) or both sexes (het. advantage) have similar equilibria and net selection strengths, though mixed SA mutations are somewhat more susceptible to drift. Overall, these qualitative differences among balancing selection mechanisms are robust to variability in male and female fitness landscapes ( $z_m, z_f, \rho_{sel}$ ), the strength of genetic correlations between the sexes ( $\rho_{mut}$ ), dimensionality ( $n$ ), and dominance ( $h$ ) (see Supplementary Material Figs. S4-S8).

## DISCUSSION

Our analysis yields four theoretical insights. First, there are ample opportunities for balancing selection in dioecious populations. Criteria for balancing selection at a locus are easily met for mutations of intermediate size (on the relevant scale of Fisher's geometric model:  $x_A$ , for a dioecious population), with a maximum probability typically on the order of 25 percent (conditional on dominance; Fig. 2). For all but small-effect mutations, balancing selection is more common than positive selection, and may therefore be an important feature of adaptive evolutionary change, as also recently demonstrated for the monoecious case by Sellis et al. (2011). Second, mean mutation size,  $E(x_A)$ , is elevated by divergent directional selection on male and female phenotypes, and by sexually dimorphic phenotypic effects of random mutations (sex-by-genotype effects). Consequently, the relative probability of balancing versus positive selection should increase in species with separate sexes. Third, under most of the parameter space of sex-specific selection and sex-by-genotype effects, balancing selection primarily involves sexually antagonistic alleles. Finally, in a finite population, the effective strength of sexually antagonistic balancing selection is expected to generally be weaker than that of heterozygote advantage, particularly under directional SA selection (the mixed SA model shows a smaller departure from the het. advantage model). Therefore, SA alleles under balancing selection may often exhibit population frequencies that are far from equilibrium (see Livingstone 1992; Connallon and Clark 2013), evolutionary dynamic properties that mimic those of neutral or weakly deleterious alleles (Connallon and Clark 2012), and relatively weak molecular population genetic signatures of partial selective sweeps (Connallon and Clark 2013; see Sellis et al. 2011 for an additional discussion of balancing selection signals in FGM).

## **Fitness tradeoffs and opportunities for balancing selection**

Several studies have reported sexually antagonistic genetic variation for fitness, or fitness components, in animal and plant populations (reviewed in Bonduriansky and Chenoweth 2009; van Doorn 2009; Kirkpatrick 2009; Pennell and Morrow 2013), yet the genetic basis of this quantitative genetic variation is not well known. Much of this phenotypic variation could be attributable to many loci that segregate for low-frequency sexually antagonistic alleles that are maintained at a balance between recurrent mutation, net purifying selection, and genetic drift. Alternatively, the variation could be primarily attributable to a small set of loci that segregate for intermediate frequency sexually antagonistic alleles under balancing selection.

Population genetics models are often used to evaluate the plausibility of sexually antagonism as a common mechanism for balancing selection (Kidwell et al. 1977; Pamilo 1979; Curtsinger 1980; Prout 2000; Patten and Haig 2009; Unckless and Herren 2009; Fry 2010; Patten Úbeda and Haig 2011; Arnqvist 2011; Jordan and Charlesworth 2012), and most of this theory has focused on “SA directional” selection models (but see Kidwell 1977; Gavrillets and Rice 2006; Mokkonen et al. 2011), where only a small fraction of the conceivable parameter space for sex-specific fitness and dominance will generate balancing selection at a diploid locus (Prout 2000; Patten and Haig 2009). A broader class of fitness tradeoff scenarios for balancing selection – *e.g.*, between environments that vary over time or space (Levene 1953; Haldane and Jayakar 1963), and antagonistic pleiotropy between traits (Rose 1982; Curtsinger et al. 1994) – suffer from similarly severe parameter restrictions (Prout 2000). On the other hand, fitness landscape topography biases the range of fitness effect values that beneficial alleles are likely to take (Orr 2005a, b, 2010; Sellis et al. 2011; Manna et al. 2011), with mutations that are subject to fitness tradeoffs (including sexually antagonistic alleles) preferentially occupying portions of the

conceivable parameter space that are the most amenable to balancing selection (Gillespie 1978; Fry 2010).

Our analysis of the dioecious Fisher's geometric model, coupled with the recent monoecious analysis by Sellis et al. (2011), suggests that balancing selection may arise naturally by way of the structure of multidimensional fitness landscapes. Because high-dimensional systems like FGM ensure that improvements along some trait axes are coupled with maladaptation along others (Orr 1998), adaptive mutations and alleles subject to balancing selection are nearly always subject to fitness tradeoffs. Sex differences in selection ( $\rho_{sel} < 1$ ), and sex-by-genotype effects on trait variation ( $\rho_{mut} < 1$ ), increase the effective dimensionality of Fisher's geometric model, and thereby expand opportunities for fitness tradeoffs between individuals of a population. In models that ignore the effects of sex, balancing selection by heterozygote advantage requires that the population be relatively close to a local optimum in the fitness landscape (Sellis et al. 2011). Such restrictions are alleviated in the two-sex FGM, whose net effect is to elevate the standardized mutation size, and consequently, the importance of balancing relative to positive selection during the process of adaptation.

Whereas fitness tradeoffs are expected to manifest across the full parameter space for balancing selection in Fisher's geometric model, the nature of the underlying tradeoffs varies as a function of specific properties of mutation and the phenotypic orientation of sex-specific selection. The classical heterozygote balance model (Fisher 1922) dominates when mutant phenotypic effects are strongly correlated between the sexes and directional selection is similar in strength and orientation between males and females (*i.e.*,  $\rho_{mut}, \rho_{sel} \gg 0$ ). Sexual antagonism, in the absence of heterozygote advantage in either sex (SA directional), dominates as patterns of directional selection diverge between the sexes ( $\rho_{sel}$  decreases), but mutational effects remain

tightly coupled ( $\rho_{mut} \gg 0$ ). A mixed scenario of sexual antagonism – with heterozygote advantage in one sex and positive/purifying selection in the other (SA mixed) – represents the primary mechanism for maintaining balanced polymorphisms when mutant phenotypic effects are poorly correlated between the sexes.

We can tentatively gauge opportunities for each mode of balancing selection by considering relevant empirical estimates of selection and mutational effects on phenotypes expressed by both sexes. Current estimates of sex-specific selection reveal widespread asymmetries in the strength or direction of selection on individual phenotypic traits (Cox and Calsbeek 2009; Stulp et al. 2012), and recent estimates of multivariate selection from humans and two insect species provide concrete examples of sex differential selection within multivariate trait space ( $\rho_{sel} \approx -0.61$  for three traits of the Indian meal moth: Lewis et al. 2011;  $\rho_{sel} \approx -0.73$  for seven traits in *Drosophila serrata*: Gosden et al. 2012;  $\rho_{sel} \approx 0.22$  for seven traits in a human population: Stearns et al. 2012). Although this set of studies is small, selection patterns are consistently divergent between the sexes, which implies that mutations with sexually antagonistic fitness effects should be common (Connallon and Clark 2014), and that sexual antagonism should play an important role in generating balancing selection (*i.e.*, SA directional or SA mixed mechanisms, as shown here). Estimates of quantitative genetic variability among male and female traits mostly show strong and positive phenotypic correlations between the sexes (*e.g.*, detectable, but modest gene-by-sex effects; Houle and Fierst 2012; Poissant et al. 2010; Griffin et al. 2013), which tentatively suggests an important potential role for SA directional balancing selection. There is, however, a great need for studies that measure the joint effects of random mutations on male and female traits (Houle and Fierst 2012). These should prove valuable for further evaluating predictions of

balancing selection that emerge from Fisher's geometric model, as well as the severity of genetic constraints to sex-specific adaptation and the evolution of sexual dimorphism.

## **Conclusions**

Balancing selection may be common within diploids (Sellis et al. 2011), and typically involve alleles with sexually antagonistic fitness effects. The genetic architecture of sex-specific fitness is likely to be complex, with processes of recurrent mutation, balancing selection, and genetic drift each potentially contributing to the maintenance of fitness variation (Bonduriansky and Chenoweth 2009; Connallon and Clark 2012). Our results show that the parameter criteria for sexually antagonistic balancing selection are not difficult to meet, and consequently, intermediate frequency sexually antagonistic alleles may contribute substantially to the high levels of quantitative genetic variance that are commonly observed in life-history traits and fitness components (Houle 1992; Charlesworth and Hughes 1999; Charlesworth and Charlesworth 1999). However, the common assumption that balanced polymorphisms are maintained at or near equilibrium may be unrealistic, particularly with regard to sexually antagonistic alleles, which are particularly susceptible to genetic drift (see Hesketh et al. 2013).

Finally, both theory and data suggest that the severity of sexual antagonism should generally increase during the course of adaptation, with sexually antagonistic mutations being rare in poorly adapted populations, and common in well adapted ones (*e.g.*, Long et al. 2012; Connallon and Clark 2014; Berger et al. 2014; but this is not always so: see Delcourt et al. 2009; Punzalan et al. 2014). The relative probability of balancing versus positive selection similarly increases during the course of adaptation in Fisher's geometric model (*e.g.*, because the average mutation "size" increases as the distance to the optimum shrinks; Sellis et al. 2011; see above). Taken

together, we expect that both sexual antagonism and balancing selection will be most rare following an abrupt change in the environment, in which both sexes are displaced from their fitness optima. As the population adapts, sexual antagonism and balancing selection should become increasingly important in shaping patterns of population genetic variation, leading to a reconfiguration of the sex-specific distribution of fitness effects among random mutations (Connallon and Clark 2014), and the genetic basis of phenotypic (including fitness) variance, between poorly and well-adapted populations.

## **ACKNOWLEDGEMENTS**

We thank two anonymous reviewers for valuable comments and suggestions on an earlier version of the manuscript, and members of the Clark Lab for discussion. This work was funded by NIH grant GM64590 to A. G. Clark and A. B. Carvalho.

## **LITERATURE CITED**

- Albert AYK, Otto SP. 2005. Sexual selection can resolve sex-linked sexual antagonism. *Science* 310:119-121.
- Allison AC. 1956. The sickle-cell and haemoglobin C genes in some African populations. *Annals of Human Genetics* 21:67-89.
- Arnqvist G. 2011. Assortative mating by fitness and sexually antagonistic genetic variation. *Evolution* 65:2111-2116.
- Barrett RDH, M'Gonigle LK, Otto SP. 2006. The distribution of beneficial mutant effects under strong selection. *Genetics* 174:2071-2079.
- Berger D, Grieshop K, Lind MI, Goenaga J, Maklakov AA, Arnqvist G. 2014. Intralocus sexual

conflict and environmental stress. *Evolution* doi:10.1111/evo.12439.

Blackburn GS, Albert AYK, Otto SP. 2010. The evolution of sex ratio adjustment in the presence of sexually antagonistic selection. *Am. Nat.* 176:264-275.

Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. *Trends Ecol. Evol.* 24:280-288.

Charlesworth B, Charlesworth D. 1999. The genetic basis of inbreeding depression. *Genetical Research* 74:329-340.

Charlesworth B, Charlesworth D. 2010. *Elements of Evolutionary Genetics*. Roberts & Company Publishers.

Charlesworth B, Hughes KA. 1999. The quantitative genetics of life history traits, pp. 369–392 in *Evolutionary Genetics From Molecules to Morphology*, edited by RS Singh and CB Krimbas. Cambridge University Press, Cambridge, UK.

Charlesworth B, Jordan CY, Charlesworth D. 2014. The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. *Evolution*, in press.

Charlesworth D. 2006. Balancing selection and its effects on sequences in nearby genome regions. *PloS Genetics* 2:e64.

Charlesworth D, Charlesworth B. 1980. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet Res Camb.* 35:205-214.

Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 98:1671-1675.

Connallon T, Cox RM, Calsbeek R. 2010. Fitness consequences of sex-specific selection. *Evolution* 64:1671-1682.

- Connallon T, Clark AG. 2010. Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression. *Evolution* 64:3417-3442.
- Connallon T, Clark AG. 2011. The resolution of sexual antagonism by gene duplication. *Genetics* 187:919-937.
- Connallon T, Clark AG. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. *Genetics* 190:1477-1489.
- Connallon T, Clark AG. 2013. Sex-differential selection and the evolution of X inactivation strategies. *PloS Genetics* 9:e1003440.
- Connallon T, Clark AG. 2014. The evolutionary inevitability of sexual antagonism. *Proc. Roy. Soc. B*, doi:10.1098/rspb.2013.2123.
- Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *Am. Nat.* 173:176-187.
- Crow JF, Kimura M. 1970. *An Introduction to Population Genetics Theory*. Burgess Publishing Co.
- Curtsinger JW. 1980. On the opportunity for polymorphism with sex-linkage or haplodiploidy. *Genetics* 96:995-1006.
- Curtsinger JW, Service PM, Prout T. 1994. Antagonistic pleiotropy, reversal of dominance, and genetic polymorphism. *Am Nat.* 144:210-228.
- Day T, Bonduriansky R. 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics* 167:1537-1546.
- Delcourt M, Blows MW, Rundle HD. 2009. Sexually antagonistic genetic variance for fitness in an ancestral and novel environment. *Proc R Soc B* 276:2009-2014.
- Dobzhansky, T. (1955) A review of some fundamental concepts and problems of population

- genetics. *Cold Spring Harb. Symp. Quant. Biol.* 20:1–15.
- Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.* 8:689-698.
- Ewens WJ. 2004. Mathematical population genetics. I. Theoretical introduction. 2<sup>nd</sup> Ed. Springer, New York.
- Ewens WJ, Thomson G. 1970. Heterozygote selective advantage. *Annals of Human Genetics* 33:365-376.
- Falconer DS, Mackay TFC. 1996. *Introduction to Quantitative Genetics*, 4<sup>th</sup> Ed.. Longmans Green, Harlow, Essex, UK.
- Fisher RA. 1922. On the dominance ratio. *Proc. Roy. Soc. Edinburgh* 42:321-341.
- Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Fry JD. 2010. The genomic location of sexually antagonistic genetic variation: some cautionary comments. *Evolution* 64:1510-1516.
- Gavrilets S, Rice WR. 2006. Genetic models of homosexuality: generating testable predictions. *Proc Biol Sci.* 273:3031-3038.
- Gillespie JH. 1978. A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. *Theor Pop Biol.* 14:1-45.
- Gillespie JH. 2004. *Population Genetics: a Concise Guide*. John Hopkins University Press.
- Gosden TP, Shastri KL, Innocenti P, Chenoweth SF. 2012. The B-matrix harbours significant and sex-specific constraints on the evolution of multicharacter sexual dimorphism. *Evolution* 66:2106-2116.
- Griffin RM, Dean R, Grace JL, Ryden P, Friberg U. 2013. The shared genome is a pervasive constraint on the evolution of sex-biased gene expression. *Mol. Biol. Evol.* 30:2168-2176.

- Haldane JBS, Jayakar SD. 1963. Polymorphism due to selection of varying direction. *J Genetics* 48:237-242.
- Hedrick PW. 1974. Genetic variation in a heterogeneous environment. I. Temporal heterogeneity and the absolute dominance model. *Genetics* 78:757-770.
- Hedrick PW. 1999. Antagonistic pleiotropy and genetic polymorphism: a perspective. *Heredity* 82:126-133.
- Hedrick PW. 2012. What is the evidence for heterozygote advantage selection? *Trends Ecol. Evol.* 27:698-704
- Hesketh J, Fowler K, Reuter M. 2013. Genetic drift in antagonistic genes leads to divergence in sex-specific fitness between experimental populations of *Drosophila melanogaster*. *Evolution* 67:1503-1510.
- Holman L, Kokko H. 2013. The consequences of polyandry for population viability, extinction risk and conservation. *Phil. Trans. Roy. Soc. Lond. B* 368:20120053
- Houle D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130:195-204.
- Houle D, Fierst J. 2012. Properties of spontaneous mutational variance and covariance for wing size and shape in *Drosophila melanogaster*. *Evolution* 67:1116-1130.
- Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, Slate J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* 502:93-95.
- Jordan CY, Charlesworth D. 2012. The potential for sexually antagonistic polymorphism in different genomic regions. *Evolution* 66:505-516.
- Kidwell JF, Clegg MT, Stewart FM, Prout T. 1977. Regions of stable equilibria for models of

- differential selection in the two sexes. *Genetics* 85:171-183.
- Kimura M. 1983. *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge, UK.
- Kirkpatrick M. 2009. Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136:271-284.
- Kirkpatrick M, Guerrero RF. 2014. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics*, in press.
- Kokko H, Brooks R. 2003. Sexy to die for? Sexual selection and the risk of extinction. *Annales Zoologici Fennici* 40:207-219
- Lande R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34:292-305.
- Lewontin RC. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press.
- Leffler EM, Gao Z, Pfeifer S, Segurel L, Auton A, Venn O, Bowden R, Bontrop R, Wall JD, Sella G, Donnelly P, McVean G, Przeworski M. 2013. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science* 339:1578-1582.
- Levene H. 1953. Genetic equilibrium when more than one ecological niche is available. *Am Nat.* 87:331-333.
- Lewis Z, Wedell N, Hunt J. 2011. Evidence for strong intralocus sexual conflict in the Indian meal moth, *Plodia interpunctella*. *Evolution* 65:2085-2097.
- Livingstone FB. 1992. Polymorphism and differential selection for the sexes. *Hum Biol.* 64:649-657.
- Long AD, Lyman RF, Morgan AH, Langley CH, Mackay TFC. 2000. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the

- achaete-scute complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* 154:1255-1269.
- Long TAF, Agrawal AF, Rowe L. 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Curr Biol.* 22:204-208.
- Mank JE. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *Am. Nat.* 173:141-150.
- Manna F, Martin G, Lenormand T. 2011. Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* 189:923-937.
- Marsaglia G. 1972. Choosing a point from the surface of a sphere. *Ann. Math. Statistics* 43:645-646.
- Martin G. 2014. Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. *Genetics*, in press.
- Martin G, Lenormand T. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 6:893-907.
- Michael JR, Schucany WR. 2002. The mixture approach for simulating bivariate distributions with specified correlations. *The American Statistician* 56:48-54.
- Mokkonen M, Kokko H, Koskela E, Lehtonen J, Mappes T, Martiskainen H, Mills SC. 2011. Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. *Science* 334:972-974.
- Muller ME. 1959. A note on a method for generating points uniformly on N-dimensional spheres. *Comm. Assoc. Comp. Mach.* 2:19-20
- Mullon C, Pomiankowski A, Reuter M. 2012. The effects of selection and genetic drift on the genomic distribution of sexually antagonistic alleles. *Evolution* 66:3743-3753.

- Nagylaki T. 1979. Selection in dioecious populations. *Ann. Hum. Genet.* 14:143-150.
- Nei M, Roychoudhury AK. 1973. Probability of fixation and mean fixation time of an overdominant mutation. *Genetics* 74:371-380.
- Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52:935-949.
- Orr HA. 2000. Adaptation and the cost of complexity. *Evolution* 54:13-20.
- Orr HA. 2005a. The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* 6:119-127.
- Orr HA. 2005b. Theories of adaptation: what they do and don't say. *Genetica* 123:3-13.
- Orr HA. 2006. The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. *J. Theor. Biol.* 238:279-285.
- Orr HA. 2010. The population genetics of beneficial mutations. *Phil Trans Roy Soc B.* 365:1195-1201.
- Otto SP, Day T. 2007. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution.* Princeton University Press
- Pamilo P. 1979. Genic variation at sex-linked loci: quantification of regular selection models. *Hereditas* 91:129-133
- Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* 14:83-87.
- Patten MM, Haig D. 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution* 63:2888-2895.
- Patten MM, Haig D, Ubeda F. 2010. Fitness variation due to sexual antagonism and linkage disequilibrium. *Evolution* 64:3638-3642.
- Patten MM, Ubeda F, Haig D. 2013. Sexual and parental antagonism shape genomic

architecture. *Proc. Roy. Soc. B* 280:2013117

Pennell TM, Morrow EH. 2013. Two sexes, one genome: the evolutionary dynamics of intralocus sexual conflict. *Ecology and Evolution* 3:1819-1834.

Poissant J, Wilson AJ, Coltman DW. 2010. Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. *Evolution* 64:97-107.

Pomiankowski A, Moller AP. 1995. A resolution of the lek paradox. *Proc. Roy. Soc Lond. B* 260:21-29

Prout T. 1968. Sufficient conditions for multiple niche polymorphism. *Am. Nat.* 102:493-496.

Prout T. 2000. How well does opposing selection maintain variation? Pp. 157-181 in *Evolution Genetics: from Molecules to Morphology*, edited by RS Singh and CB Krimbas. Cambridge University Press, Cambridge, UK.

Punzalan D, Delcourt M, Rundle HD. 2014. Comparing the intersex genetic correlation for fitness across novel environments in the fruit fly, *Drosophila serrata*. *Heredity* 112:143-148.

Rice SH. 2004. *Evolutionary theory: mathematical and conceptual foundations*. Sinauer Associates.

Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735-742

Rice WR. 1992. Sexually antagonistic genes: experimental evidence. *Science* 256:1436-1439.

Robertson A. 1962. Selection for heterozygotes in small populations. *Genetics* 47:1291-1300.

Rodgers JL, Nicewander WA. 1988. Thirteen ways to look at the correlation coefficient. *The American Statistician* 42:59-66.

Rose MR. 1982. Antagonistic pleiotropy, dominance, and genetic variation. *Heredity* 48:63-78.

- Roze D, Otto SP. 2012. Differential selection between the sexes and selection for sex. *Evolution* 66:558-574.
- Seeger J, Trivers R. 1986. Asymmetry in the evolution of female mating preferences. *Nature* 319:771-773.
- Sellis D, Callahan BJ, Petrov DA, Messer PW. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. *Proc. Natl. Acad. Sci. USA* 108:20666-20671.
- Stearns SC, Govindaraju DR, Ewbank D, Byars SG. 2012. Constraints on the coevolution of contemporary human males and females. *Proc. Roy. Soc. Lond. B* 279:4836-4844.
- Stulp G, Kuijper B, Buunk AP, Pollet TV, Verhulst S. 2012. Intralocus sexual conflict over human height. *Biology Letters* 23:976-978.
- Turelli M, Barton NH. 2004. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and GxE interactions. *Genetics* 166:1053-1079.
- Unckless RL, Herren JK. 2009. Population genetics of sexually antagonistic mitochondrial mutants under inbreeding. *J. Theor. Biol.* 260:132-136.
- van Doorn GS, Kirkpatrick M. 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature* 449:909-912.
- van Doorn GS, Kirkpatrick M. 2010. Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* 186:629-645
- Waxman D, Welch JJ. 2005. Fisher's microscope and Haldane's ellipse. *Am. Nat.* 166:447-457.
- Whitlock MC, Agrawal AF. 2009. Purging the genome with sexual selection: reducing mutation load through selection on males. *Evolution* 63:569-582.
- Wright AE, Mank JE. 2013. The scope and strength of sex-specific selection in genome evolution. *J. Evol. Biol.* 26:1841-1853.

Wright S. 1969. *Evolution and the Genetics of Populations*, Vol. 2. University of Chicago Press.

## Appendix A: Movements in phenotypic space and selection coefficients

Let  $A_j = \{x_{1j}, x_{2j}, \dots, x_{nj}\}$  represent the phenotypic value expressed by wild type individuals of the  $j$ th sex ( $j = \{f, m\}$ ),  $O_j = \{O_{1j}, O_{2j}, \dots, O_{nj}\}$  is the location of its optimum, and  $O'_j = O_j - A_j = \{o_{1j}, o_{2j}, \dots, o_{nj}\}$  is a rescaled optimum that becomes useful further below. Consider a mutation with a homozygous magnitude of  $r_m$  and  $r_f$ , and heterozygous magnitude of  $r_m h$  and  $r_f h$ . The phenotypic value expressed by a heterozygote will be  $A_j + h y_j$ , and a homozygote will be  $A_j + y_j$ , where  $y_j = \{y_{1j}, y_{2j}, \dots, y_{nj}\}$  is a mutation vector that describes the set of phenotypic changes in each of the  $n$  traits (this is described in the ‘‘Simulations’’ section of the main text). The Euclidean distance between a wild type individual of sex  $j$  and the respective phenotypic optimum is:

$$z_j = \sqrt{\sum_{i=1}^n (O_{ij} - x_{ij})^2}$$

The distance between a heterozygote and its optimum is:

$$\begin{aligned} z_{j(het)} &= \sqrt{\sum_{i=1}^n [O_{ij} - (x_{ij} + h y_{ij})]^2} = \sqrt{\sum_{i=1}^n (O_{ij} - x_{ij})^2 + \sum_{i=1}^n (h y_{ij})^2 - 2h \sum_{i=1}^n o_{ij} y_{ij}} \\ &= \sqrt{z_j^2 + (h r_j)^2 - 2h \sum_{i=1}^n o_{ij} y_{ij}} \end{aligned}$$

The exact result in Eq. (2) is obtained by substituting  $y_{im} = r_m m_i / M$ , where  $M^2 = \sum m_i^2$ , and  $y_{if} = r_f f_i / F$ , where  $F^2 = \sum f_i^2$  (as described in the ‘‘Simulations’’ section). With sufficiently high dimensionality ( $n > 10$  works well)  $m_i / M$  and  $f_i / F$  converge to  $m_i / \sqrt{n}$  and  $f_i / \sqrt{n}$  (Connallon and Clark 2014), which accounts for the final approximations in Eq. (2). The same approach can be taken to obtain values of  $z_{j(hom)}$ .

For a random mutation, the fitness of the three genotypes will be  $w_{j(wild\ type)} = \exp[-\omega_j z_j^2]$ ,  $w_{j(het)} = \exp[-\omega_j z_{j(het)}^2]$ , and  $w_{j(hom)} = \exp[-\omega_j z_{j(hom)}^2]$ . Selection coefficients are scaled relative to the wild type fitness. The heterozygote and homozygote selection coefficients for a random

mutation are (respectively)  $s_{j1} = w_{j(het)}/w_{j(wild\ type)} - 1 = \exp[\omega_j(z_j^2 - z_{j(het)}^2)] - 1$ , and  $s_{j2} = w_{j(hom)}/w_{j(wild\ type)} - 1 = \exp[\omega_j(z_j^2 - z_{j(hom)}^2)] - 1$ .

## Appendix B: Modes of selection and their probabilities

Define  $t = \ln(1 + s)$  as the natural logarithm of fitness for a genotype with relative fitness  $1 + s$ .

Under weak selection ( $s \ll 1$ ),  $t$  is well approximated as  $t = \ln(1 + s) \sim s$ . The criterion for balancing selection at a diploid locus can be approximated as:

$$0 < t_{mhet} + t_{fhet} > t_{mhom} + t_{fhom} \quad (A1)$$

where  $t_{mhet} = \ln(1 + s_{m1})$ ,  $t_{fhet} = \ln(1 + s_{f1})$ ,  $t_{mhom} = \ln(1 + s_{m2})$ ,  $t_{fhom} = \ln(1 + s_{f2})$ . The criterion for invasion of a new mutation is modified to:

$$0 < t_{mhet} + t_{fhet} \quad (A2).$$

Let the mutation effect size in heterozygotes be  $r_m h$  and  $r_f h$ , where  $h$  is the dominance coefficient (assumed to fall in the range  $0 < h < 1$ ), and the mutation size in homozygotes be  $r_m$  and  $r_f$ . From eq. (A1), along with eq. (2) and its surrounding text from the main text, the criteria for balancing selection can be restated as:

$$\frac{h(\omega_m r_m^2 + \omega_f r_f^2)\sqrt{n}}{2} < \omega_m r_m \sum_i o_{mi} m_i + \omega_f r_f \sum_i o_{fi} f_i < \frac{(1-h^2)(\omega_m r_m^2 + \omega_f r_f^2)\sqrt{n}}{2(1-h)} \quad (A3).$$

Terms  $m_i$  and  $f_i$  are standard normal variables with  $\text{cov}(m_i, f_i) = \rho_{mut}$ . The function in the middle is therefore normally distributed with respective mean of zero and variance of:

$$\begin{aligned}
\text{var} &= \text{var} \left( \omega_m r_m \sum_i o_{mi} m_i + \omega_f r_f \sum_i o_{fi} f_i \right) \\
&= (\omega_m r_m)^2 \left( \sum_i o_{mi}^2 \text{var}(m_i) \right) + (\omega_f r_f)^2 \left( \sum_i o_{fi}^2 \text{var}(f_i) \right) + 2\omega_m \omega_f r_m r_f \left( \sum_i o_{mi} o_{fi} \text{cov}(m_i, f_i) \right) \\
&= (\omega_m r_m)^2 \left( \sum_i o_{mi}^2 \right) + (\omega_f r_f)^2 \left( \sum_i o_{fi}^2 \right) + 2\omega_m \omega_f r_m r_f \rho_{mut} \left( \sum_i o_{mi} o_{fi} \right) \\
&= (\omega_m r_m z_m)^2 + (\omega_f r_f z_f)^2 + 2\omega_m \omega_f z_m z_f r_m r_f \rho_{mut} \cos(\theta_{sel})
\end{aligned}$$

where  $\rho_{mf} = \rho_{mut} \cos(\theta_{sel})$ . After dividing the terms of the inequality (A3) by  $\sqrt{\text{var}}$ , the function in the middle is now distributed as a standard normal  $[\mathcal{N}(0, 1)]$ , and we can now solve for the probability of balancing selection:

$$\text{Pr}(\text{bal.}) \approx \frac{1}{\sqrt{2\pi}} \int_b^a e^{-\frac{t^2}{2}} dt = \Phi(a) - \Phi(b)$$

where

$$a = \frac{(1-h^2)}{(1-h)} \frac{(\omega_m r_m^2 + \omega_f r_f^2) \sqrt{n}}{2\sqrt{(\omega_m z_m r_m)^2 + (\omega_f z_f r_f)^2 + 2\omega_m z_m r_m \omega_f z_f r_f \rho_{mf}}} = \frac{(1-h^2)}{(1-h)} x_A$$

and

$$b = h \frac{(\omega_m r_m^2 + \omega_f r_f^2) \sqrt{n}}{2\sqrt{(\omega_m z_m r_m)^2 + (\omega_f z_f r_f)^2 + 2\omega_m z_m r_m \omega_f z_f r_f \rho_{mf}}} = h x_A$$

which accounts for eq. (7) in the main text. The same basic approach can be taken to calculate the probability that selection will favor invasion of a rare mutation. In this case,  $\text{Pr}(s_{m1} + s_{f1} > 0) = 1 - \Phi(b)$ , as in eq. (6).

## Appendix C: The maximum probability of balancing selection

To find the specific value of  $x_A$  that maximizes the probability of balancing selection, we restate eq. (7) as:

$$\Phi\left(\frac{(1-h^2)}{(1-h)}x_A\right) - \Phi(hx_A) = \frac{1}{2}\left[erf\left(\frac{(1-h^2)}{(1-h)\sqrt{2}}x_A\right) - erf\left(\frac{h}{\sqrt{2}}x_A\right)\right]$$

which has a first derivative, with respect to  $x_A$ , of:

$$\frac{d}{dx_A}\left[\Phi\left(\frac{(1-h^2)}{(1-h)}x_A\right) - \Phi(hx_A)\right] = \frac{h}{\sqrt{2\pi}}\exp\left(-\frac{h^2}{2}x_A^2\right)\left[\frac{(1-h^2)}{h(1-h)}\exp\left(-x_A^2\left[\frac{1-h^2(3-2h)}{2(1-h)^2}\right]\right) - 1\right]$$

The maximum probability of balancing selection is found by setting this function to zero, and solving for  $x_A$ , which yields:

$$\hat{x}_A = (1-h)\sqrt{\frac{2}{1-h^2(3-2h)}\ln\left[\frac{(1-h^2)}{h(1-h)}\right]}$$

#### Appendix D: Mean standardized mutation size

For  $\beta_m > 0$ , and  $\beta_f = 0$  ( $z_m = 0$ ), we can approximate  $E[x_A]$  by performing a second order Taylor

Series expansion in the vicinity of  $r_m, r_f = E(r)$ , and take the expectation, which leads to:

$$\begin{aligned} E[x_A] &= \frac{E(r)\omega_m\sqrt{n}}{\beta_m} + E\left[\frac{\omega_f r_f^2\sqrt{n}}{\beta_m r_m}\right] \\ &\approx \frac{E(r)(\omega_m + \omega_f)\sqrt{n}}{\beta_m} \left[1 + \frac{\omega_f E\left\{[r_m - E(r)]^2 + [r_f - E(r)]^2 - 2[r_m - E(r)][r_f - E(r)]\right\}}{(\omega_m + \omega_f)[E(r)]^2}\right] \\ &= \frac{E(r)(\omega_m + \omega_f)\sqrt{n}}{\beta_m} \left[1 + \frac{2\omega_f[\text{var}(r) - \text{cov}(r_m, r_f)]}{(\omega_m + \omega_f)[E(r)]^2}\right] \end{aligned}$$

Eq. (10) in the text can be found by substituting  $\beta_m = 2\omega_m z_m$ ,  $E(r) = k\theta$ ,  $\text{var}(r) = k\theta^2$ , and  $\text{corr}(r_m,$

$r_f) = \text{cov}(r_m, r_f)/\text{var}(r)$ , and rearranging. Using the same approach for  $\beta = \beta_m = \beta_f$ , we get:

$$\begin{aligned}
E[x_A] &= E \left[ \frac{(\omega_m r_m^2 + \omega_f r_f^2) \sqrt{n}}{\beta \sqrt{r_m^2 + r_f^2 + 2r_m r_f \rho_{mf}}} \right] \\
&\approx \frac{E(r)(\omega_m + \omega_f) \sqrt{n}}{\beta \sqrt{2(1 + \rho_{mf})}} \left[ 1 + \frac{(1 + 3\rho_{mf}) E \{ [r_m - E(r)]^2 + [r_f - E(r)]^2 - 2[r_f - E(r)][r_m - E(r)] \}}{8(1 + \rho_{mf}) [E(r)]^2} \right] \\
&= \frac{E(r)(\omega_m + \omega_f) \sqrt{n}}{\beta \sqrt{2(1 + \rho_{mf})}} \left[ 1 + \frac{(1 + 3\rho_{mf}) [\text{var}(r) - \text{cov}(r_f, r_m)]}{4(1 + \rho_{mf}) [E(r)]^2} \right]
\end{aligned}$$

Substituting  $E(r) = k\theta$ ,  $\text{var}(r) = k\theta^2$ , and  $\text{corr}(r_m, r_f) = \text{cov}(r_m, r_f)/\text{var}(r)$ , and simplifying, yields eq. (11) in the text.

### Appendix E: Allele frequency dynamics and equilibrium of $B_2$ under balancing selection

Among zygotes, let  $q_m$  be the frequency of  $B_2$  transmitted from males, and  $q_f$  the frequency of  $B_2$  transmitted from females. The genotypic frequencies among zygotes are:  $q_m q_f = [B_2 B_2]$ ;  $q_m(1 - q_f) + q_f(1 - q_m) = [B_1 B_2]$ ;  $(1 - q_m)(1 - q_f) = [B_1 B_1]$ . Let  $q = (q_m + q_f)/2$ , and  $2\varepsilon = q_m - q_f$ . Assuming weak selection (Nagylaki 1979; Charlesworth and Charlesworth 2010), we approximate the allele frequency dynamics to first order in  $\varepsilon$ . Following selection, the expected allele frequency change in males and females will be:

$$\begin{aligned}
\Delta q_m &= \frac{(q^2 - \varepsilon^2)(1 + s_{m2}) + [q(1 - q) + \varepsilon^2](1 + s_{m1})}{1 + s_{m2}(q^2 - \varepsilon^2) + 2s_{m1}[q(1 - q) + \varepsilon^2]} - q - \varepsilon \approx \frac{q(1 - q)[s_{m1} + q(s_{m2} - 2s_{m1})]}{1 + q^2 s_{m2} + 2q(1 - q)s_{m1}} - \varepsilon \\
\Delta q_f &= \frac{(q^2 - \varepsilon^2)(1 + s_{f2}) + [q(1 - q) + \varepsilon^2](1 + s_{f1})}{1 + s_{f2}(q^2 - \varepsilon^2) + 2s_{f1}[q(1 - q) + \varepsilon^2]} - q + \varepsilon \approx \frac{q(1 - q)[s_{f1} + q(s_{f2} - 2s_{f1})]}{1 + q[s_{f2}q + 2s_{f1}(1 - q)]} + \varepsilon
\end{aligned}$$

The net change after a generation is  $(\Delta q_m + \Delta q_f)/2$ :

$$\begin{aligned}
\Delta q &= \frac{\Delta q_f + \Delta q_m}{2} \approx \frac{[2(s_{m1} + s_{f1}) - (s_{f2} + s_{m2})]q(1 - q)(\hat{q} - q)}{2} \\
\hat{q} &= \frac{s_{f1} + s_{m1}}{2(s_{m1} + s_{f1}) - (s_{f2} + s_{m2})}
\end{aligned}$$

where  $\hat{q}$  is the balanced polymorphic equilibrium when it exists.

## Appendix F: Criteria for effective balancing selection

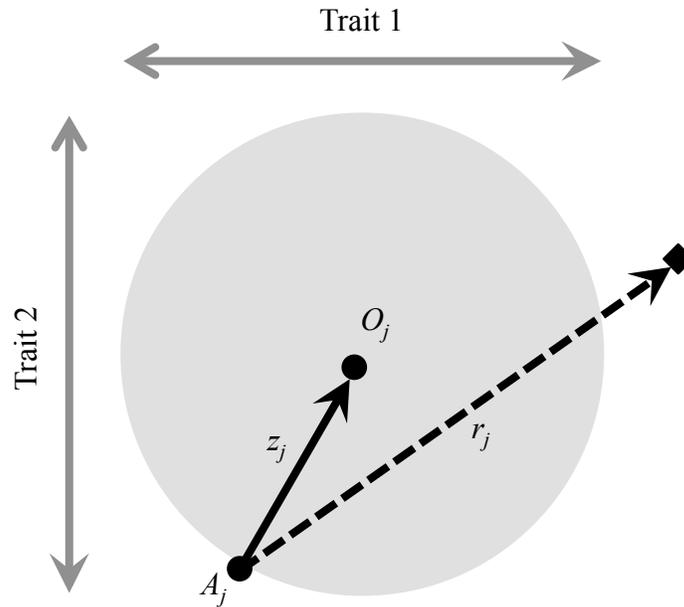
Following Ewens (2004; pp. 26-27), we wish to know the relative proportion of the density,  $f(q)$ , that is found near the deterministic equilibrium value ( $\hat{q}$ ), compared to the allele frequency boundaries:  $q = 0$  and  $q = 1$ . Using a 2<sup>nd</sup> order Taylor series expansion, we can approximate the area under the stationary distribution near arbitrary frequency point  $a$ :

$$\begin{aligned} \int_{a-\varepsilon}^{a+\varepsilon} f(q) dq &\propto \int_{a-\varepsilon}^{a+\varepsilon} \frac{\exp[-\alpha(q-\hat{q})^2]}{q(1-q)} dq \\ &\approx f(a) \int_{a-\varepsilon}^{a+\varepsilon} dq + \left( \frac{\partial f(a)}{\partial q} \right) \int_{a-\varepsilon}^{a+\varepsilon} (q-a) dq + \left( \frac{1}{2} \frac{\partial^2 f(a)}{\partial q^2} \right) \int_{a-\varepsilon}^{a+\varepsilon} (q-a)^2 dq \\ &= 2\varepsilon f(a) + \frac{2\varepsilon^3}{3} \left( \frac{1}{2} \frac{\partial^2 f(a)}{\partial q^2} \right) \end{aligned}$$

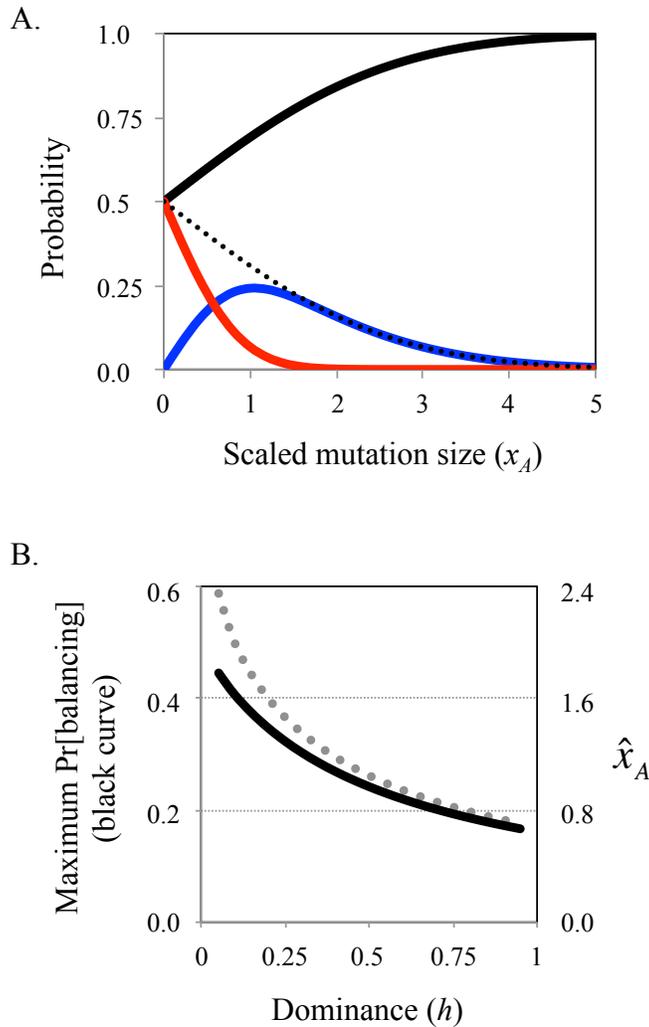
By substituting  $a = 0.001$ ,  $a = 0.999$ ,  $a = \hat{q}$  (as appropriate), we obtain eq. (14) of the main text.

**Table 1.** Sex-specific fitnesses and genotype frequencies.

	<b>Genotype</b>		
	$B_1B_1$	$B_1B_2$	$B_2B_2$
Female fitness:	1	$1 + s_{f1}$	$1 + s_{f2}$
Male fitness:	1	$1 + s_{m1}$	$1 + s_{m2}$
Zygotic frequency:	$(1 - q_m)(1 - q_f)$	$q_m(1 - q_f) + q_f(1 - q_m)$	$q_m q_f$

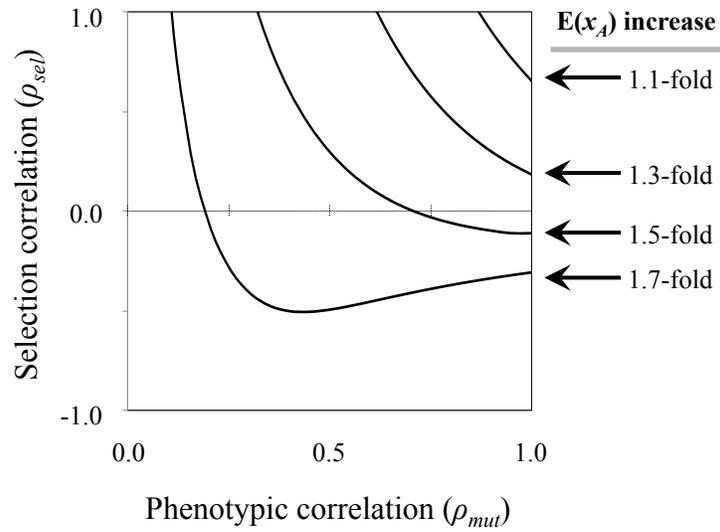


**Figure 1.** Fisher's geometric model with two sexes (an example with two traits). For the  $j$ th sex, individuals homozygous for a wild-type allele express phenotype  $A_j$ , and are separated from their optimum ( $O_j$ ) by distance  $z_j$ . Mutations alter the phenotype of their carriers, with mutant homozygotes expressing a phenotype at a distance  $r_j$  from the ancestral homozygote (an example mutant homozygote is represented by the black diamond). Heterozygotes are similarly oriented away from  $A_j$ , but have a modified distance of  $hr_j$ , where  $h$  is the dominance coefficient with respect to the mutant phenotypic effect. For the example shown, heterozygotes will express an intermediate phenotype between  $A_j$  and the black diamond, and for most values of  $h$  ( $0 < h < 1$ ), the heterozygote phenotype will be closer to the optimum than either homozygote genotype will be. In such cases, there will be heterozygote advantage for fitness in the  $j$ th sex. Females and males may be characterized by distinct fitness optima, and unique positions in phenotypic space with respect to each genotype (see Connallon and Clark 2014).

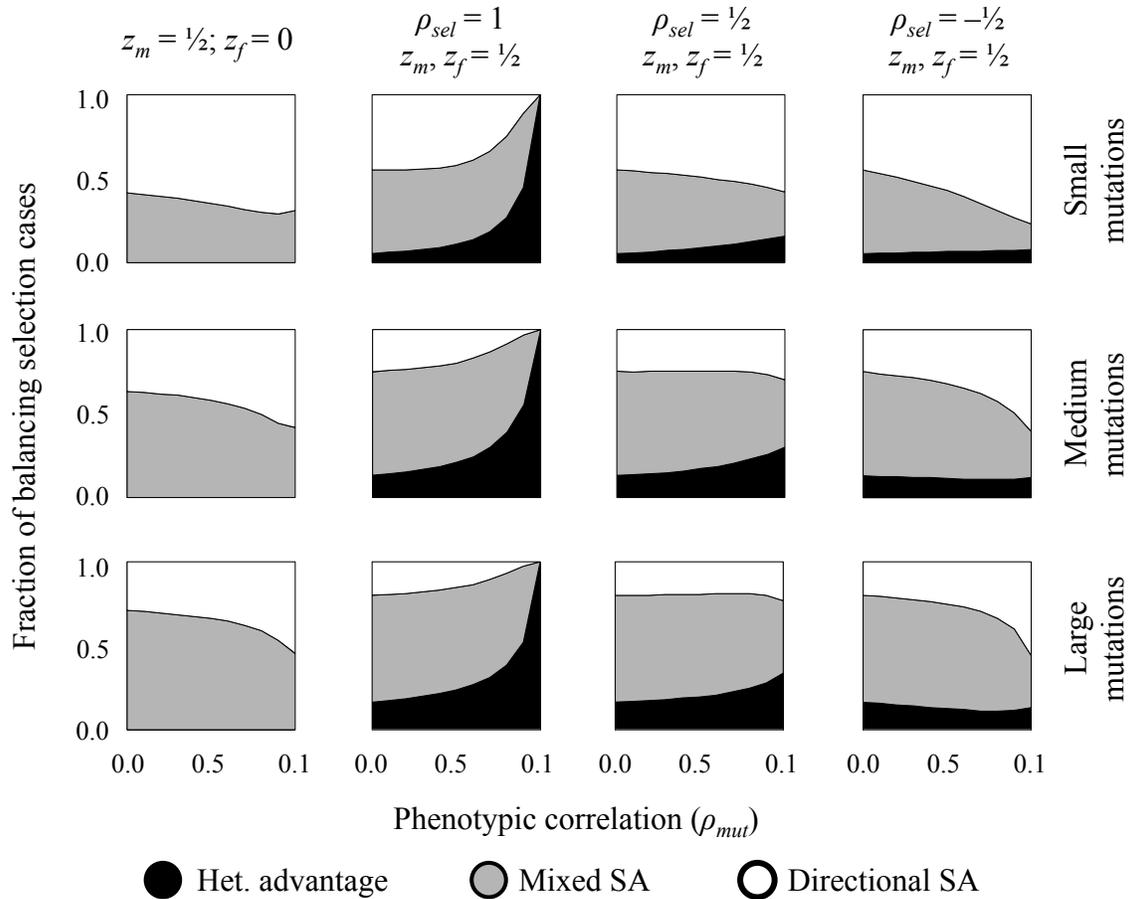


**Figure 2.** Balancing selection in Fisher’s geometric model. (A). Probabilities of purifying selection (black), positive selection (red), or balancing selection (blue), as a function of the scaled mutation of size,  $x_A$ . The dotted line represents the probability that selection will favor evolutionary invasion of a random mutation of size  $x_A$  (this is the sum of individual probabilities of purifying and balancing selection). Results are for the case of codominant expression in heterozygotes ( $h = \frac{1}{2}$ ), and curves were generated using eqs. (6-7). Simulations (not shown) confirm that these results are extremely accurate at moderate to high dimensionality (*i.e.*,  $n > 10$ ). (B). The maximum probability of balancing selection (black curve), and the critical value of

$x_A(\hat{x}_A)$  that maximizes the probability of balancing selection (dotted gray line). Results are based on eqs. (7-8).

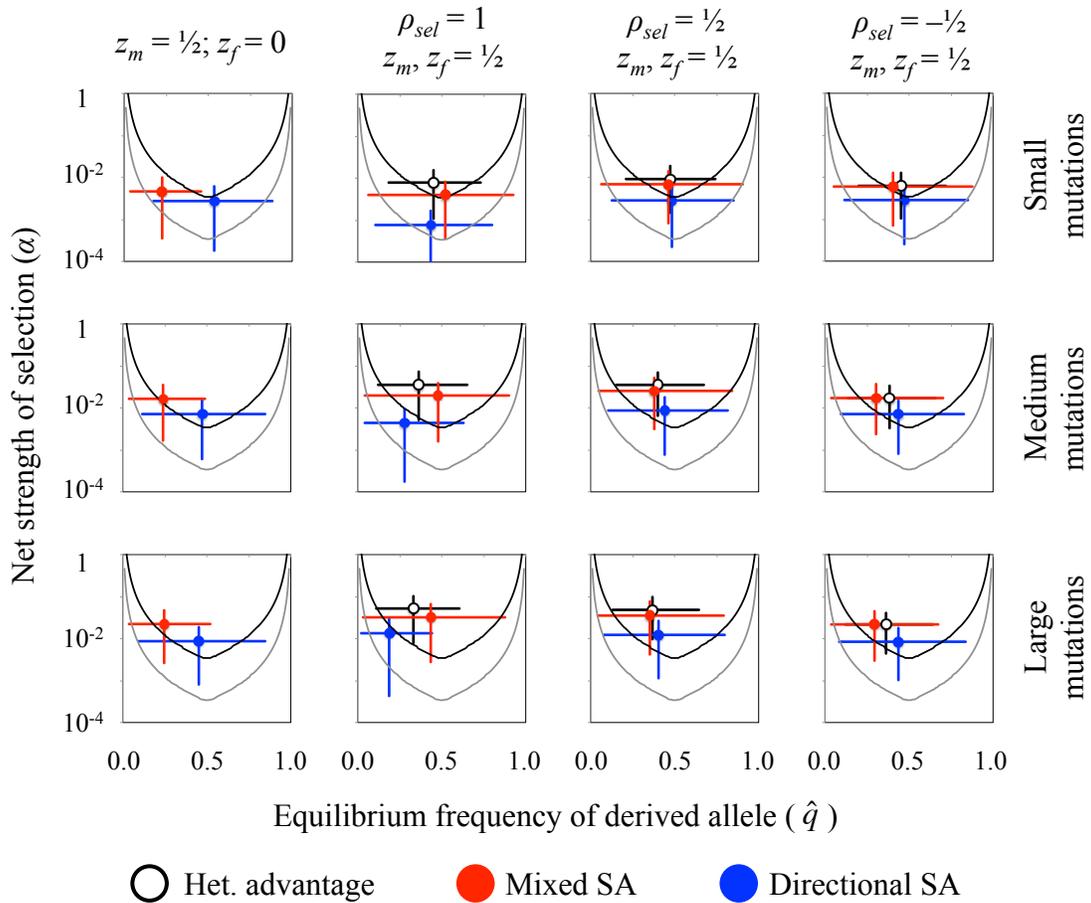


**Figure 3.** Sex differences in the direction of selection, and sex-by-genotype effects of mutations, elevate mean mutation size,  $E(x_A)$ . Solid curves show the elevation of  $E(x_A)$  relative to an idealized population in which the selection and mutational effects are identical between the sexes. The fold change to  $E(x_A)$  is based on numerical evaluation of the bracketed term in eq. (11), with  $\rho_{mut} = \text{corr}(r_m, r_f)$ .



**Figure 4.** Balancing selection mechanisms and their relative probabilities. SA selection mechanisms (white and gray shading) account for the vast majority of balancing selection cases. Heterozygote advantage becomes the dominant form of balancing selection in the restrictive case where directional selection and mutant phenotypic effects are strongly concordant between the sexes (e.g.,  $z_m \rightarrow z_f$ , and  $\rho_{sel}, \rho_{mut} \rightarrow 1$ ). For each parameter set ( $z_m, z_f, \rho_{sel}, \rho_{mut}$ ), 500,000 balanced polymorphisms were randomly generated using the simulation approach described in the text. These results were used to calculate the proportion of balancing selection arising from heterozygote advantage, mixed SA, and directional SA. For each parameter combination, mutation magnitudes were generated using a bivariate exponential distribution (gamma with shape parameter  $k = 1$ ), with equal marginal distributions, and correlation of  $\text{corr}(r_m, r_f) = \rho_{mut}$ .

Small mutations use  $E[r] = 0.05$ ; medium mutations use  $E[r] = 0.2$ ; and large mutations use  $E[r] = 0.4$  (these absolute mutation sizes correspond to male-specific scaled sizes of  $E[x_m] = E(r)(n^{0.5})/(2z_m) = 0.25$ ,  $E[x_m] = 1$ , and  $E[x_m] = 2$ , respectively). Representative results are shown for  $n = 25$  and  $\omega_m = \omega_f = 1/2$ , with additional results presented in the Supplementary Material (Figs. S1-S3).



**Figure 5.** The efficacy of selection differs between mechanisms of balancing selection. Results show the distributions (means and 75% confidence intervals) of  $\alpha$  and  $\hat{q}$  for mutations under each of the three forms of balancing selection. The areas above the continuous curves represent the parameter space for effectively strong selection, based on eq. (14) with criteria  $R > 10$  (see the text surrounding eq. (14)). Gray curves (lower) represent the parameter space for a population with effective size  $N_e = 100,000$ ; black curves (higher) are for populations with  $N_e = 10,000$ . For each balancing selection mechanism, and for each parameter set  $(z_m, z_f, \rho_{sel})$ , 10,000 balanced polymorphisms were randomly generated using the simulation approach described in the main text. These results were used to calculate means and confidence intervals. For each parameter combination, mutation magnitudes were generated using a bivariate exponential

distribution (gamma with shape parameter  $k = 1$ ), with equal marginal distributions, and correlation of  $\rho_{mut} = \text{corr}(r_m, r_f)$ . Three different mean mutation sizes were used: small mutations use  $E[r] = 0.05$ ; medium mutations use  $E[r] = 0.2$ ; and large mutations use  $E[r] = 0.4$  (these absolute mutation sizes correspond to male-specific scaled sizes of  $E[x_m] = E(r)(n^{0.5})/(2z_m) = 0.25$ ,  $E[x_m] = 1$ , and  $E[x_m] = 2$ , respectively). Representative results are shown for  $\rho_{mut} = 0.9$ ,  $n = 25$ , and  $\omega_m = \omega_f = 1/2$ , with additional results presented in the Supplementary Material (Figs. S4-S8).