SEX LINKAGE, SEX-SPECIFIC SELECTION, AND THE ROLE OF RECOMBINATION IN THE EVOLUTION OF SEXUALLY DIMORPHIC GENE EXPRESSION

Tim Connallon¹ ² and Andrew G. Clark¹

¹Department of Molecular Biology & Genetics, Cornell University, Ithaca, New York 14853-2703 ²E-mail: tmc233@cornell.edu

Sex-biased genes—genes that are differentially expressed within males and females—are nonrandomly distributed across animal genomes, with sex chromosomes and autosomes often carrying markedly different concentrations of male- and female-biased genes. These linkage patterns are often gene- and lineage-dependent, differing between functional genetic categories and between species. Although sex-specific selection is often hypothesized to shape the evolution of sex-linked and autosomal gene content, population genetics theory has yet to account for many of the gene- and lineage-specific idiosyncrasies emerging from the empirical literature. With the goal of improving the connection between evolutionary theory and a rapidly growing body of genome-wide empirical studies, we extend previous population genetics theory of sex-specific selection by developing and analyzing a biologically informed model that incorporates sex linkage, pleiotropy, recombination, and epistasis, factors that are likely to vary between genes and between species. Our results demonstrate that sex-specific selection and sex-specific recombination rates can generate, and are compatible with, the gene- and species-specific linkage patterns reported in the genomics literature. The theory suggests that sexual selection may strongly influence the architectures of animal genomes, as well as the chromosomal distribution of fixed substitutions underlying sexually dimorphic traits.

KEY WORDS: Antagonistic pleiotropy, epistasis, sex-biased genes, sex chromosomes, sexual antagonism.

Sexual dimorphism is common among animal species, and is thought to reflect sexually discordant natural or sexual selection (Darwin 1871; Andersson 1994). The idea that sexual dimorphism reflects differential adaptation—that sex-specific selection drives evolutionary divergence between the sexes—is clearly articulated by Trivers (1972), who states:

One can, in effect, treat the sexes as if they were different species, . . . female “species” usually differ from male species in that females compete among themselves for such resources as food but not for members of the opposite sex, whereas males ultimately compete only for members of the opposite sex, all other forms of competition being important only insofar as they affect this ultimate competition. (Trivers 1972, p. 153)

From a population genetics perspective, the “sexes as species” analogy is less apt. Excluding the few genes that are limited to a single sex (e.g., Y-linked genes in males; W-linked genes in females), males and females carry the same set of genes and make equal reproductive contributions to each generation. Males and females may be exposed to different patterns of natural and sexual selection, yet “gene flow” between the sexes is unrestricted and should constrain intersexual divergence. Although it is clear that sexual dimorphism does evolve, the genetic basis and sequence of evolutionary events that permit males and females to diverge is not well known.

The evolution of sexual dimorphism has been conceptualized with two models (Darwin 1871; Fisher 1958; Rhen 2000; Coyne
Under a “sex-limited” model, sexual dimorphism evolves by selection of mutations with sex-limited phenotypic effects. If sex-limited mutations are available, the evolution of sexual dimorphism is not problematic. For most mutations, however, at least some degree of expression is expected in both sexes (see Lande 1980; Rice 1984; Rhen 2000; Morrow et al. 2008; Poissant et al. 2010; though the magnitude may be sex-specific: e.g., Mackay 2001). Consequently, the evolution of sexual dimorphism might require genetic substitutions at multiple loci, with genetic interactions underlying intersexual divergence. Such a multistep process could potentially involve sexually antagonistic divergence—correlated evolution between the sexes that increases adaptation of one sex and reduces it for the other—followed by the evolution of sex-limited expression of sexually antagonistic traits (Lande 1980; Rice 1984; Bonduriansky and Chenoweth 2009; van Doorn 2009; Stewart et al. 2010).

Large proportions of animal genomes are differentially expressed between the sexes (Parisi et al. 2003; Ellegren and Parsch 2007), with “sex-biased” gene expression potentially decoupling male and female development and permitting adaptive sexual differentiation (Williams and Carroll 2009). The genomic distribution (i.e., chromosomal linkage patterns) of sex-biased genes may also provide clues about the evolutionary processes and genetic basis underlying sex-specific divergence (Mank 2009). Despite substantial variability between species, there is an emerging consensus that sex-biased genes are nonrandomly distributed between sex chromosomes and autosomes (e.g., Reinke et al. 2000; Wang et al. 2001; Parisi et al. 2003; Ranz et al. 2003; Khil et al. 2004; Kaiser and Ellegren 2006; Storchova and Divina 2006; Sturgill et al. 2007; Mueller et al. 2008; Mank and Ellegren 2009; Meisel et al. 2009; Moškovský et al. 2010). Such patterns may suggest an important role for sexual antagonism during the evolution of sexual dimorphism (e.g., Parisi et al. 2003; Rogers et al. 2003; Oliver and Parisi 2004; Connallon and Knowles 2005; Ellegren and Parsch 2007; Gurbich and Bachtrog 2008; Mank 2009; Innocenti and Morrow 2010)—an interpretation inspired by population genetics theory that predicts unique evolutionary fates for sex-linked versus autosomal sexually antagonistic mutations (mutations beneficial to one sex and deleterious to the other; Pamilo 1979; Rice 1984; Patten and Haig 2009; Fry 2010).

Nevertheless, the connection between sexually dimorphic gene expression patterns and population genetics theory is incomplete for several reasons. First, although sexual antagonism is often invoked, it does not by itself generate sexual dimorphism, and previous theory has mostly neglected epistasis between sexually antagonistic alleles and mutations that modify intersexual genetic correlations (but see Rice 1984). The role of epistasis is likely to be profound: recent research reveals that interactions between different cis-regulatory motifs often underlie sexually dimorphic expression (Williams and Carroll 2009). Second, genetic correlations between different traits will also constrain adaptive evolution, yet it is unclear how strongly pleiotropy might hinder the evolution of sexual dimorphism or whether such constraints might differentially impact sex chromosomes and autosomes.

Third, the current theory is almost entirely deterministic and focuses on conditions in which selection favors the invasion of a rare allele, rather than the probability and rate of sex-specific divergence, which is governed by selection, mutation, and genetic drift (Crow and Kimura 1970; Ohta 1992). Contrasts between chromosomes should be based on the relative invasion probabilities of individual mutations, and waiting times for the evolution of sexual dimorphism. Finally, the relationship between phenotype and fitness has not been consistently articulated by theory, in some cases leading to opposing predictions about the relationship between sex linkage and sexual antagonism (e.g., Rice 1984; Patten and Haig 2009; Fry 2010). Fortunately, there is a considerable body of research linking gene expression variation to fitness variation (see below), which can be used to refine sex-specific selection models and incorporate biologically plausible parameterization.

With the goal of integrating empirical patterns of sex-biased gene expression and population genetics theory, we developed and analyzed a series of two-locus models for the evolution of sex-biased expression in response to selection for gene expression divergence. We separately consider three general evolutionary routes toward sexually dimorphic gene expression: (1) the fixation of sex-limited and tissue-specific mutations that are unconstrained by pleiotropy or sexual antagonism and directly generate sexual dimorphism; (2) the sequential fixation of sexually antagonistic or pleiotropically constrained alleles and alleles at modifier loci, which interact epistatically to generate sexual dimorphism or tissue-specific divergence; and (3) the co-invasion and simultaneous fixation of linked alleles that are not individually favored by selection, but are beneficial in combination. For each scenario, we analytically characterize the necessary conditions and timescale of sex-biased gene evolution under X-, Z- and autosomal linkage. The theory provides a general framework for the evolution of sex-biased gene expression and generates a set of specific and testable predictions.

**Model**

To accommodate the potential impact of both linkage and epistasis, we first developed two-locus, biallelic population genetic models with arbitrary fitness assigned to each genotype and sex (Table 1). Because we are interested in contrasting patterns of evolution on sex chromosomes and autosomes, the loci are assumed to be both X-linked, both Z-linked, or both autosomal. We discuss the potential role of interchromosomal interactions within the discussion. The distance between interacting loci is presented as a function of the recombination rate between them. Because
recombination is potentially sexually dimorphic, sex-specific recombination rates between the loci are given by parameters $r_f$ and $r_m$ for females and males, respectively. Recursions for the four possible haplotypes ($A_1B_1, A_2B_1, A_1B_2, A_2B_2$) are developed in Appendix 1, and follow the sequence of (a) birth, (b) selection, (c) recombination, and (d) syngamy.

Because the population mutation rate in animals—for example, $4N_e u$ per autosomal locus, per generation—is expected to be less than one, we assume that each locus will initially be fixed or nearly fixed for a common allele (Crow and Kimura 1970). Consequently, evolutionary outcomes of selection at single loci or pairs of loci will be governed by patterns of selection acting on rare genetic variants. Under this assumption, the strength of selection, and probability of invasion for rare alleles or haplotypes can be addressed analytically with local linearized stability criteria.

A common genotype is susceptible to the evolutionary invasion of a rare allele or haplotype when the leading eigenvalue for the system, that is the largest eigenvalue of the Jacobian matrix for the set of recursions, is greater than one. Furthermore, the leading eigenvalue provides information about the probability of invasion versus loss of the rare allele, as described further below. For each system of inheritance—autosomal, X- and Z-linkage—there are three potential leading eigenvalues, which are each presented in Table 2. Details of the recursions, as well as the stability analysis, can be found in Appendices 1 and 2.

**Fitness as a Function of Gene Expression**

Throughout the analysis, we focus on evolutionary divergence in gene expression due to *cis*-regulatory substitutions. This particular focus is empirically justified by the disproportionate contribution of *cis*-relative to *trans*-regulatory changes during species divergence for gene expression and morphology (e.g., Wittkopf et al. 2004; Wray 2007; Carroll 2008; Wittkopf et al. 2008; Graze et al. 2009), as well as the important role of *cis*-regulatory interactions during sex-biased gene regulation and morphological divergence between the sexes (Williams and Carroll 2009; also see Loehlin et al. 2010a,b). An additional justification is concep-
presented here permits asymmetries within or between male and female fitness landscapes. Landscapes (i.e., deviations above and below expression optima are equally costly), this need not necessarily be the case. The theory of the sexes within some contexts (context 1), and are divergent in others (context 2). Note that although we display symmetrical fitness gene is expressed in multiple tissues or during multiple time periods during development. Expression-fitness functions overlap between
two-locus models of sex-specific selection, with alleles $A_1$ and $B_1$ fixed.

<table>
<thead>
<tr>
<th>Locus A</th>
<th>Locus B</th>
<th>Both loci (A and B interacting)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autosome Linkage: $\lambda_A = \frac{f_1}{2f_1} + \frac{m_{1f}}{2m_{11}}$</td>
<td>$\lambda_B = \frac{f_1}{2f_1} + \frac{m_{1f}}{2m_{11}}$</td>
<td>$\lambda_{AB} = \frac{f_{2c}(1-r_f)}{2f_1} + \frac{m_{2c}(1-r_f)}{2m_{11}}$</td>
</tr>
<tr>
<td>X-Linkage: $\lambda_A = \frac{f_1}{4f_1} (1 + \frac{1 + 8 f_{2c}m_{11}}{f_{2c}m_{11}})$</td>
<td>$\lambda_B = \frac{f_1}{4f_1} (1 + \frac{1 + 8 f_{2c}m_{11}}{f_{2c}m_{11}})$</td>
<td>$\lambda_{AB} = \frac{f_{2c}(1-r_f)}{4f_1} (1 + \frac{1 + 8 f_{2c}m_{11}}{f_{2c}m_{11}})$</td>
</tr>
<tr>
<td>Z-Linkage: $\lambda_A = \frac{m_{1f}}{m_{11}} (1 + \frac{1 + 8 m_{1f} f_1}{m_{1f} f_1})$</td>
<td>$\lambda_B = \frac{m_{1f}}{m_{11}} (1 + \frac{1 + 8 m_{1f} f_1}{m_{1f} f_1})$</td>
<td>$\lambda_{AB} = \frac{m_{2c}(1-r_f)}{4m_{11}} (1 + \frac{1 + 8 m_{1f} f_1}{m_{1f} f_1})$</td>
</tr>
</tbody>
</table>

**Figure 1.** Divergent expression-fitness landscapes generate conflicting selection over gene expression. The vertical gray lines represent the wild-type gene expression at the time point when the fitness landscapes diverge between the sexes. Arrows indicate the net direction of selection: stabilizing selection in the top panels, and selection for increased expression in the bottom panels. (A) Sex-specific selection without pleiotropy: selection favors a different level of gene expression in males and females. (B) Sex-specific selection with pleiotropy. A gene is expressed in multiple tissues or during multiple time periods during development. Expression-fitness functions overlap between the sexes within some contexts (context 1), and are divergent in others (context 2). Note that although we display symmetrical fitness landscapes (i.e., deviations above and below expression optima are equally costly), this need not necessarily be the case. The theory presented here permits asymmetries within or between male and female fitness landscapes.

on the premise that fitness benefits of gene expression follow a saturating function. Because expression of a gene is expected to also carry one or more costs—that is, energetic and metabolic trade-offs (Kacser and Beeby 1984; Hurst and Randerson 2000; Wagner 2005, 2007), competition for translation (Gout et al. 2010), interference between molecular pathways (Lion et al. 2004), or toxicity (Clark 1991)—it is a mathematical necessity that any composite fitness function that incorporates both benefits and costs of gene expression will be concave within the vicinity of the gene expression optimum (see Supporting information). A balance between benefits and costs is particularly relevant for the expression and evolution of sex-biased genes, which appear to be condition-dependent, implying a significant cost of gene expression (Wyman et al. 2010).

(2) Direct experimental measurement of benefits, costs, and total fitness as a function of the gene expression phenotype, supports a concave fitness surface (Dean et al. 1986; Dykhuizen et al. 1987; Dean 1989; Papp et al. 2003; Dekel and Alon 2005). Although the concavity assumption is ultimately an issue that can only be resolved by additional experimental work, the current evidence strongly suggests that fitness is a concave function of a gene’s expression level. We therefore adopt this concave relationship throughout our analysis, with the caveat that our predictions are subject to revision following future empirical findings.

If we assume that the fitness topologies of males and females diverge by small increments relative to time, and that the majority of mutations have small effects on gene expression relative to the landscape, then we can model the fitness effects of mutations along a concave fitness surface. Figure 2 presents a conceptual relationship between concave fitness surfaces and their relationship to selection and dominance parameters. Fitness ($w$) as a function of gene expression ($x$) can be formally developed with the relationship:

$$w(x) = \exp \left\{-c |\tilde{x} - x|^k \right\},$$
Beneficial mutations

Consider a wild-type allele $A_1$, which is fixed within a population and which causes gene expression at a level $x_1$. Following a shift in the environment, $x_1$ no longer resides at the fitness optimum such that $x_1 \neq \hat{x}$. $A_2$ is a rare mutation that changes gene expression from $x_1$ to $x_1 + \Delta x$ when heterozygous, and to $x_1 + 2\Delta x$ when homozygous. For simplicity, assume that wild-type expression is below the optimum, causing selection to favor alleles that increase expression (because the model is symmetrical, the opposite case yields the same results). Assuming that the homozygous genotype does not overshoot the expression optimum, such that $\hat{x} \geq x_1 + 2\Delta x$ and $\Delta x > 0$, the fitness of the three genotypes will be: $w(A_1A_1) = \exp[-c(\hat{x} - x_1)^2]$, $w(A_1A_2) = \exp[-c(\hat{x} - x_1 - \Delta x)^2]$, and $w(A_2A_2) = \exp[-c(\hat{x} - x_1 - 2\Delta x)^2]$. The dominance coefficient of $A_2$ over $A_1$ is therefore:

$$h_b = \frac{1 - \exp [c\Delta x (2\hat{x} - 2x_1 - \Delta x)]}{1 - \exp [4c\Delta x (\hat{x} - x_1 - \Delta x)]} \approx \frac{2(\hat{x} - x_1 - \Delta x) + \Delta x}{4(\hat{x} - x_1 - \Delta x)} = \frac{1}{2} + \frac{\Delta x}{4(\hat{x} - x_1 - \Delta x)},$$

which can hypothetically range between 0.5 < $h_b$ < 0.75 (in this parameterization, the additive case has $h_b = \frac{1}{2}$). Extending the theory for an arbitrary $k$-th order fitness function, dominance of beneficial mutations will range between 0.5 < $h_b$ < $(2^k - 1)/2^k$.

Deleterious mutations

When the population resides at the fitness peak (i.e., the wild type genotype produces optimal gene expression), all mutations are deleterious because they cause gene expression to deviate from the optimum. Fitness of a deleterious mutation in heterozygous and homozygous state (respectively) follows the functions: $w(A_1A_2) = \exp \{ -c(\Delta x)^2 \}$, and $w(A_2A_2) = \exp \{ -c(2\Delta x)^2 \}$. For this second-order function, dominance of each deleterious mutation is $h_d \approx \frac{1}{2}$. For a generalized $k$-th order function, dominance of deleterious mutations is $h_d \approx 1/2^k$.

Pleiotropy

Previous models generally assume that sexually antagonistic selection arises from opposing directional selection between the sexes (e.g., Fig. 1A). An alternative possibility arises for genes expressed in multiple contexts or tissues (e.g., Chintapalli et al. 2007), yet the population genetic consequences of pleiotropy have yet to be explicitly analyzed within the theoretical context of sex-specific selection (Fitzpatrick 2004; Mank 2009). To incorporate pleiotropy into the theory, we follow Curtsinger et al. (1994) and model total fitness per genotype as a multiplicative function of fitness within each context of expression (e.g., tissue or developmental stage). Thus, for a genotype “g” expressed in $n$ contexts or tissues, total fitness is $w_g = \Pi_{i=1}^{n} w_i$. The model is roughly equivalent to a linear fitness model in which the fitness benefits and costs of a mutation are small for each context (1 - $w_i < 1$). For example, a diallelic locus with one allele ($A_1$) favored in $j$ contexts and the other favored in one context ($A_2$), will follow the fitness scheme: $w(A_1A_1) = 1 - x;$
w(A_1A_2) = (1 - sh) \prod_{i=1}^{L} (1 - t_i h_i) \approx 1 - sh - \sum_{i=1}^{L} t_i h_i; and 
w(A_2A_1) = \prod_{i=1}^{L} (1 - t_i) \approx 1 - \sum_{i=1}^{L} t_i, where s and h represent 
selection and dominance coefficients with respect to the tissue that 
is under directional selection, and t_i and h_i are the selection 
and dominance coefficients with respect to the j tissues that are 
under stabilizing selection. Note that this is probably the simplest 
model for pleiotropy, and is limited to the case in which the 
genotype is the unit of selection.

PROBABILITY OF INVASION
For a population with common alleles A_1 and B_1 fixed or nearly 
fixed (thus, stability is assessed at the equilibrium \( [A_1] = [B_1] = 1 \)), the leading 
eigenvalue for the system of recursions provides information about the strength and direction of selection acting on 
rare alleles or haplotypes. The leading eigenvalue minus one 
(\( \lambda_L - 1 \)) provides the equivalent of a selection coefficient at the 
particular point at which the system is analyzed (e.g., for a rare 
mutation; Otto and Bourguet 1999; Otto and Yong 2002). For 
our purposes, leading eigenvalues can be used to approximate the 
probability of invasion for new mutations. Such probabilities can 
potentially refer to individual alleles (A_2 or B_2), or of haplotypes 
(A_2B_2). For \( 1 \gg \lambda_L - 1 \gg 1/N_e \), where \( N_e \) is the effective 
population size, the probability of establishment versus loss of a new 
mutation is:

\[
P = \frac{2(\lambda_L - 1)}{\lambda_L^2} \approx 2(\lambda_L - 1)
\]

(i.e., by branching process; Haldane 1927; Otto and Day 2007). 
The approximation is used in subsequent results, and its accuracy 
is verified by simulation (see Figs. S1 and S2).

TEMPORAL DYNAMICS
To describe the relative rate at which sexual dimorphism evolves 
under different modes of inheritance, we develop expressions to 
characterize the mean waiting time until sex-specific expression 
divergence. Under the assumption of strong selection and weak 
mutation (SSWM; Gillespie 1984, 1991), where the strength of 
selection acting on a rare allele or haplotype is much greater than 
the reciprocal of the population size (\( 1 \gg \lambda_L - 1 \gg 1/N_e \), 
as described above), and the rate of mutation is much smaller 
than \( 1/N_e \) (i.e., \( N_e u \ll 1 \)), the temporal dynamics of individual 
substitutions are dominated by the waiting time until invasion of 
an allele or haplotype (the time required for an allele to arise 
within the population and then escape stochastic loss by genetic 
drift), and the transit time for each selective sweep is negligible.

For models involving single-step evolutionary transitions 
(e.g., the invasion of alleles or haplotypes with sex-limited and 
tissue-specific effects), we calculate the mean waiting time until 
the invasion of a derived allele or haplotype. This basic approach 
can also be applied to sequential invasion models with epistasis 
(Weinreich and Chao 2005)—in our case, where initial invasion of 
a sexually antagonistic or antagonistic pleiotropic allele generates 
selection for a modifier allele causing sex- or tissue-specific 
divergence. As with the single-step scenario, SSWM assumptions 
permit us to approximate the two-step fixation process by analyzing 
the waiting time for each invasion event. The expected waiting 
time until both substitutions is equal to the sum of individual waiting 
times until invasion. These approaches are used widely in the 
thoretical literature, and are accurate under SSWM conditions 
(e.g., Stephan 1996; Weinreich and Chao 2005; Kim 2007).

Results

EVOLUTIONARY ROUTES TOWARD SEXUALLY 
DIMORPHIC GENE EXPRESSION
Sexually dimorphic gene expression can evolve by three basic 
evolutionary pathways. For mutations with sex-limited and tissue-
specific effects on gene expression, sexual dimorphism can evolve 
by single, unconstrained genetic substitutions, provided that the 
mutation alters expression within the specific context where di-
vergence is favored. For mutations that are constrained by sexual 
antagonism or pleiotropy, sex-specific adaptation may arise as 
a consequence of multiple genetic substitutions involving muta-
tions whose epistatic effects generate a sex- and/or tissue-specific 
phenotypic response.

The fixation of combinations of mutations, although more 
restrictive than the fixation of sex-limited alleles, can occur either 
sequentially or simultaneously. If one of the mutations at the two 
loci is individually favored in the wild-type genetic background, 
it can increase from low initial frequency. Formally, this occurs 
when \( \lambda_A > 1 \) or \( \lambda_B > 1 \) (see Table 2). As one allele increases in 
frequency, selection will favor alleles at the other locus that inter-
act positively with the invading allele such that the combination 
of alleles improves net fitness across the sexes. Such two-step pro-
cess leads to the eventual fixation of positively interacting pairs 
of alleles that optimize expression divergence between the sexes.

If mutations at the two loci are individually disfavored by 
selection (i.e., \( \lambda_A < 1 \) and \( \lambda_B < 1 \)), but are favored in com-
bination, sexual dimorphism can potentially evolve if the positive 
epistatic interactions between mutations are strong relative to the 
recombinational distance between the loci (Crow and Kimura 
1965; Weinreich and Chao 2005). Formally, this will occur when 
\( \lambda_{AB} > 1 \) (Table 2). In terms of gene expression evolution, an allele 
combination might be favored because it decouples gene expres-
sion variation between the sexes or between different tissues, 
and thereby reduces evolutionary constraints imposed by sexu-
ally antagonistic selection or antagonistic pleiotropy. Both types 
of epistasis could arise from fitness interactions between muta-
tions at sex- or tissue-specific and nonsex-specific cis-regulatory 
elements (see Williams and Carroll 2009).
The opportunity for each of these evolutionary transitions can systematically differ between sex chromosomes and autosomes. Differential accumulation of sex-linked and autosomal substitutions is a function of several factors: (1) whether males or females are selected to diverge, (2) whether alleles are sex-limited in expression or expressed in both sexes, (3) whether mutations have pleiotropic effects or instead alter expression within single tissues, and (4) whether the relative rate of recombination differs between sex chromosomes and autosomes. Below, we present these results by contrasting X-linked and autosomal inheritance. These results also apply to Z versus autosome contrasts, which represent a mirror image, with results reversed across sexes.

THE FIXATION OF SEX-LIMITED, TISSUE-SPECIFIC MUTATIONS

If a mutation only alters gene expression within the tissue and sex where divergence is beneficial, there will be no constraint to its invasion and eventual fixation. With recurrent mutation to the favored allele, the substitution will eventually become fixed, with the waiting time determined by the mutation rate and strength of selection in favor of the mutation. For an autosomal allele, the mean waiting time until fixation will be:

$$T_{A_{fix}} = \frac{1}{N_A u_A \Pr(fix) + \tau}, \quad (2)$$

where $N_A$ is the autosomal effective size ($N_A \approx 2N_e$), $u_A$ is the autosomal rate of mutation to the sex-limited allele, $\Pr(fix)$ is the probability that the allele eventually becomes fixed, and $\tau$ represents the transit time of the allele—that is, the time it takes for the allele to change from frequency $1/N_A$ to frequency one. Given the simplifying assumption of strong selection/weak mutation (SSWM: $1 \gg h_b \gg 1/N_e \gg u$; Gillespie 1984, 1991), the second term has little impact and can be ignored (e.g., Stephan 1996; Gillespie 2000; Weinreich and Chao 2005; Kim 2007). Furthermore, the probability of fixation for a sex-limited beneficial mutation is approximately equal to the heterozygous selection coefficient in favor of that mutation ($sh_b$, the result for sex-limited selection on the allele). Equation (2) can therefore be modified to:

$$T_{A_{fix}} = \frac{1}{N_A u_A s h_b}, \quad (3a)$$

Under the same conditions, except with X-linked inheritance, the mean waiting times for male-limited and for female-limited beneficial substitutions (respectively) are

$$T_{X_{fix}(f)} = \frac{3}{4N_X u_X s h_b}, \quad (3c)$$

where $N_X$ represents the X-linked effective size, and $u_X$ the X-linked mutation rate per generation with respect to the allele. These results represent reciprocals of the instantaneous rate of adaptive substitution, which are often used in theories of molecular evolution, including contrasts between sex chromosomes and autosomes (Charlesworth et al. 1987; Kirkpatrick and Hall 2004; Vicoso and Charlesworth 2009a).

The specific waiting times until substitution on the X and autosomes is sensitive to the effective mutation rate on each chromosome (the ratio $N_X u_X/N_A u_A$ is expected to be 3/4, but may vary as a result of demography, life-history, or breeding system; Charlesworth 2001, 2009; Ellegren 2007, 2009; Hedrick 2007). Nevertheless, the relative rate of fixation for male-beneficial mutations will always be more sensitive to X-linkage and dominance. Under the null expectation, $4N_X u_X = 3N_A u_A$, the relative waiting time until female-beneficial substitution is the same on the X and autosomes: $T_{X_{fix}(f)}/T_{A_{fix}} = 1$. The relative time until male-beneficial substitution is $T_{X_{fix}(m)}/T_{A_{fix}} = 2h_b$, where the substitution rate is always more rapid on the autosomes under the concave fitness landscape model (i.e., $h_b > 1/2$ when $k > 1$). X-linkage facilitates female-beneficial substitution as long as $4N_X u_X > 3N_A u_A$, and male-beneficial substitution when $2N_X u_X > 3N_A u_A h_b$, with the latter requirement always more restrictive for $h_b > 1/2$. These results are in agreement with several previous studies concerned with “faster-X” molecular evolution (i.e., Charlesworth et al. 1987; Kirkpatrick and Hall 2004; Vicoso and Charlesworth 2009a).

INVASION OF PLEIOTROPIC MUTATIONS, FOLLOWED BY MODIFIERS OF PLEIOTROPY

Adaptive divergence may be constrained if expression-altering mutations also affect tissues evolving under stabilizing selection. Consider the evolution of a pleiotropic locus $A$, at which a derived mutation ($A_2$: $A_1$ represents the wild type) is beneficial within one tissue and deleterious within another. For simplicity, we assume that mutations have sex-limited effects on gene expression; the evolutionary genomic consequences of sexual antagonism are considered separately. Table 3 presents the fitness parameterization under a model of antagonistic pleiotropy, where coefficients $s$ and $t$ are positive and small ($0 < s, t \ll 1$), and the dominance coefficient, $h_d = 1/2^k$ ($0 < h_d < 0.5$ for $k > 0$).

Under autosomal linkage and sex-specific selection in either sex, a derived allele $A_2$ will invade when:

$$s > \frac{th_d}{1 - h_d}, \quad (4a)$$
Table 3. Fitness parameterization under the antagonistic pleiotropy model.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>$A_1, A_1A_1$</th>
<th>$A_1A_2$</th>
<th>$A_2, A_2A_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female-specific fitness effects</td>
<td>$f_{11} = 1 - s$</td>
<td>$f_{21} = (1 - h_d)(1 - h_d)$</td>
<td>$f_{31} = 1 - t$</td>
</tr>
<tr>
<td>Male fitness (A)</td>
<td>$m_{11} = 1$</td>
<td>$m_{21} = 1$</td>
<td>$m_{31} = 1$</td>
</tr>
<tr>
<td>Male fitness (X)</td>
<td>$m_1 = 1$</td>
<td>$-$</td>
<td>$m_2 = 1$</td>
</tr>
<tr>
<td>Male-specific fitness effects</td>
<td>$f_{11} = 1$</td>
<td>$f_{21} = 1$</td>
<td>$f_{31} = 1$</td>
</tr>
<tr>
<td>Male fitness (A)</td>
<td>$m_{11} = 1 - s$</td>
<td>$m_{21} = (1 - h_d)(1 - h_d)$</td>
<td>$m_{31} = 1 - t$</td>
</tr>
<tr>
<td>Male fitness (X)</td>
<td>$m_1 = 1 - s$</td>
<td>$-$</td>
<td>$m_2 = 1 - t$</td>
</tr>
</tbody>
</table>

will become fixed when:

$$ s > \frac{t(1 - h_d)}{h_d}, $$

and will remain polymorphic when:

$$ \frac{t h_d}{1 - h_d} < s < \frac{t(1 - h_d)}{h_d}. $$

(4b)

(4c)

(these results are in agreement with those of Curtsinger et al. 1994). If $A_2$ remains polymorphic, it will converge to the equilibrium frequency:

$$ \hat{p} = \frac{s(1 - h_d) - t h_d}{(s + t)(1 - 2 h_d)}. $$

(4d)

These conditions of invasion, fixation, and polymorphism are the same under female-specific selection at an X-linked locus. Under male-limited selection and X-linkage, the conditions for invasion and eventual fixation are the same ($s > t$), and there is no parameter combination permitting a stable polymorphism.

When $A_2$ is favored, a single copy will invade the population with probability $\sim 2(\lambda_A - 1)$, where $\lambda_A$ is presented in the first row of Table 2. Under SSWM conditions, the mean waiting time until $A_2$ invades on an autosome will be:

$$ T_{wa}(A_2) = \frac{1}{2 N_A \mu_A (\lambda_A - 1)}. $$

(5a)

The expected waiting time under X-linkage will be:

$$ T_X(A_2) = \frac{1}{2 N_{AXX}(\lambda_A - 1)}. $$

(5b)

The invasion of mutations with deleterious pleiotropic effects can subsequently select for a modifier allele ($B_2$) that limits expression divergence to the appropriate tissue, thereby maximizing fitness with respect to gene expression across different tissues. Assuming allelic interactions in cis-, an $A_2B_2$ (coupling) haplotype will cause tissue-specific expression divergence, and $A_2B_1$ haplotypes will be associated with both beneficial and costly expression divergence (i.e., $A_2B_1$ haplotypes eliminate the pleiotropic cost, whereas $A_2B_2$ haplotypes do not; for a complete parameterization of the two-locus system, see Table S1). As before, invasion criteria for the $B_2$ allele can be addressed with local linearized stability criteria. Autosomal and X-linked Jacobian matrices were recalculated for an arbitrary equilibrium frequency of $A_2$, and stability was determined with the modifier allele absent ($B_2 = 0$; Appendix 3). Under autosomal inheritance and tight linkage between $A$ and $B$ ($r_{m_2} > r_{f} < 1/2$, as expected for two cis-regulatory loci for the same gene), the leading eigenvalue with respect to the $B$ locus is:

$$ \lambda_B \approx \frac{(1 - \hat{p}) f_{21}c(1 - r_f) + \hat{p} f_{31}}{2 \bar{w}_f} + \frac{(1 - \hat{p}) m_{21}c(1 - r_m) + \hat{p} m_{32}}{2 \bar{w}_m}, $$

(6a)

where $\bar{w}_f$ and $\bar{w}_m$ represent mean female and male fitness as a function of the frequency of $A_2$, with $B_1$ fixed. For X-linked loci, the leading eigenvalue is:

$$ \lambda_B \approx \frac{(1 - \hat{p}) f_{22}C(1 - r_f) + \hat{p} f_{32}}{4 \bar{w}_f} \times \left(1 + \sqrt{1 + \frac{8 \bar{w}_m}{w_m(1 - \hat{p}) f_{22}C(1 - r_f) + \hat{p} f_{32}}}ight). $$

(6b)

An analysis of these eigenvalues indicates that, given invasion of $A_2$ prior to the introduction of $B_2$, and if an $A_1B_2$ haplotype is not strongly deleterious, selection will generally favor the invasion of the modifier allele. As the net rate of recombination between the loci increases relative to the strength of selection, invasion of a modifier mutation is not guaranteed unless it is neutral in an $A_1$ genetic background, or $A_2$ is common within the population (see Appendix 4 for details). Given our focus on epistatic interactions in cis-, we focus our analysis of the sequential invasion model under a scenario of relatively tight linkage. We revisit this issue within the discussion.

The waiting time until $B_2$ invades (following invasion and equilibrium of the $A_2$ allele) is sensitive to the equilibrium frequency of $A_2$, the rate at which $A_2B_2$ haplotypes are formed by...
mutation and recombination, and the strength of selection in favor of the \(A_1B_2\) haplotype, once created. Accounting for each of these factors (see Appendix 5), the mean waiting time until \(B_2\) invades will fall within the boundary:

\[
\frac{N_A \hat{p}(1 - \hat{p})(r_m + r_f) + 2}{2N_A \hat{p} \nu_A(\lambda_B - 1)[N_A(1 - \hat{p})(r_m + r_f) + 2]}
< T_{\text{auto}}(B_2) < \frac{1}{2N_A \hat{p} \nu_A(\lambda_B - 1)} \tag{7a}
\]

under autosomal inheritance, and:

\[
\frac{2N_X \hat{p}(1 - \hat{p})r_f + 3}{2N_X \hat{p} \nu_X(\lambda_B - 1)[2N_X(1 - \hat{p})r_f + 3]}
< T_X(B_2)
< \frac{1}{2N_X \hat{p} \nu_X(\lambda_B - 1)} \tag{7b}
\]

under X-linked inheritance (where \(N_A\) and \(N_X\) refer to the autosomal and X-linked effective size, \(\nu_A\) and \(\nu_X\) represent chromosome-specific mutation rates at the \(B\) locus, and \(\lambda_B\) is characterized by eqs. 6a or 6b). When \(A_2\) is rare and \(A_1B_2\) is neutral, the waiting time approaches the term on the left. As the equilibrium frequency of \(A_2\) and the deleterious effect of \(A_1B_2\) increase, the waiting time approaches the term on the right (Appendix 5).

Permissive conditions for gene expression divergence, along with the expected waiting times until \(A_2\) and \(B_2\) invasion, are presented in Figure 3. These representative results show that X-linkage will constrain male-specific evolutionary divergence, but will not constrain female-specific divergence. As with the sex- and tissue-limited model presented previously, X-linkage will actually enhance the rate of female-specific adaptation as long as \(4N_Xu_X > 3N_Au_A\), and will be equal when \(4N_Xu_X = 3N_Au_A\).

These patterns are a natural consequence of evolution along a concave fitness surface. Concave fitness landscapes generate dominance reversals: deleterious fitness effects act recessively, whereas beneficial effects are dominant (see above; Fig. 2). In females, which are diploid throughout the genome, the expression of beneficial and deleterious fitness effects of individual mutations does not differ between the X and autosomes. However, female-limited selection can be more effective on the X because each X-linked locus evolves within a female genome with two-thirds probability, in contrast to one-half probability for autosomes. In males, benefits of expression divergence are marginally enhanced by X-linked inheritance because they are dominant, whereas costs of expression divergence are strongly enhanced by X-linkage because they act recessively. For male-specific adaptive divergence, pleiotropy should generate a strong bias off of the X because pleiotropic constraints in males are exacerbated by hemizygous expression. The severity of the bias will increase with the curvature of the fitness landscape (i.e., as \(k\) increases).

**Figure 3.** Sequential invasion conditions and waiting times until tissue-specific gene expression divergence under antagonistic pleiotropy. Pleiotropically expressed alleles (\(A_1\) and \(A_2\)) follow the fitness parameterization from Table 3. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until derived alleles invade at the pleiotropic and modifier loci, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage. Results were obtained using the SSWM approximations presented in the text. Results are shown for \(k = 2\), \(t = 0.01\), and use the permissive left hand term of equations (7a and b); mutation rates per locus and per sex are \(u = \frac{r}{2}\) and \(v = 10^{-6}\); X and autosomal effective sizes are \(N_X = 10^6 = 4N_A/3\). Results use dimorphic recombination parameters \(r_f = 0.001\) and \(r_m = 0\), yet sexually monomorphic recombination yields a nearly identical pattern.

**INVASION OF SEXUALLY ANTAGONISTIC MUTATIONS, FOLLOWED BY MODIFIERS OF SEX-LIMITATION**

Mutations that are expressed in both sexes can give rise to sexual antagonism if they are favored in one sex but disfavored in the
When females are selected to diverge, the same results apply, with
this feature generating selection for alleles at secondary loci that limit expression divergence to the sex exposed to directional selection. Because invasion at these secondary loci is limited by prior invasion at the sexually antagonistic locus, the opportunity for sequential invasion depends on whether each sexually antagonistic allele can invade and trigger the two-step process of divergence.

The specific conditions for invasion have been described in several previous studies (e.g., Kidwell et al. 1977; Pamilo 1979; Rice 1984; Albert and Otto 2005; Patten and Haig 2009; Fry 2010). We summarize these results below, with relevant eigenvalues (Table 2) evaluated using the fitness landscape model described above and in Table 4.

When divergence is favored in males, an autosomal male-beneficial allele can invade when

\[ s_m > \frac{s_f h_d}{1 - h_d(1 - s_f)} . \]  

(8a)

Such an allele will go to fixation when

\[ s_m > \frac{s_f(1 - h_d)}{h_d(1 - s_f)} . \]  

(8b)

The system will remain polymorphic when

\[ \frac{s_f h_d}{1 - h_d(1 - s_f)} < s_m < \frac{s_f(1 - h_d)}{h_d(1 - s_f)} . \]  

(8c)

Given our constraint that \( h_d < 0.5 \), the male-beneficial polymorphism will converge to the equilibrium frequency

\[ \hat{p}_m \approx \frac{s_m(1 - h_d) - s_f h_d}{s_m + s_f(1 - 2h_d)} . \]  

(8d)

When females are selected to diverge, the same results apply, with \( s_m \) and \( s_f \) substituted.

The invasion of an X-linked, male-beneficial mutation can occur when

\[ s_m > \frac{2s_f h_d}{1 + s_f h_d} . \]  

(9a)

These alleles will go to fixation when

\[ s_m > \frac{2s_f(1 - h_d)}{1 - s_f h_d} . \]  

(9b)

An X-linked, female-beneficial allele can invade when

\[ s_f > \frac{s_m}{2 - h_d(2 - s_m)} . \]  

(9c)

and will become fixed when:

\[ s_f > \frac{s_m}{h_d(2 - s_m)} . \]  

(9d)

X-linked polymorphism will be maintained under the condition

\[ \frac{2s_f h_d}{1 + s_f h_d} < s_m < \frac{2s_f(1 - h_d)}{1 - s_f h_d} . \]  

(9e)

with the male- and female-beneficial alleles converging to the respective equilibrium frequencies

\[ \hat{p}_m \approx \frac{s_m - 2s_f h_d}{2s_f(1 - 2h_d)} . \]  

(9f)

and

\[ \hat{p}_f \approx \frac{2s_f(1 - h_d) - s_m}{2s_f(1 - 2h_d)} . \]  

(9g)

The equilibria each require that \( h_d < 0.5 \), as expected for the concave fitness landscape model.

The expected time until invasion of a derived, sexually antagonistic allele (\( A_2 \)) is given by equations (5a) and (5b), with the relevant eigenvalues \( (\lambda_A) \) parameterized using fitness values from Table 4. Coupling haplotypes with derived sexually antagonistic and modifier alleles (\( A_2B_2 \)) will cause sex-specific expression divergence, whereas repulsion haplotypes (\( A_2B_1 \)) generate expression divergence in both sexes (for a complete parameterization of the two-locus system, see Table S2). As with antagonistic pleiotropy (again, assuming that \( A_1B_2 \) haplotypes are not strongly deleterious), most conditions favoring invasion of a sexually antagonistic allele will also subsequently favor invasion of a linked, \( \phi \)-regulatory modifier of sex-limitation (see Appendix 4).
The general stability conditions and waiting times until the $B_2$ allele invades (conditional on $A_1$ reaching an equilibrium frequency greater than zero) are given by equations (6a–7b); see Appendix 3 and 4 for the additional case of relatively loose linkage.

Results for the sexually antagonistic modifier model show that gene expression divergence, involving the sequential invasion of sexually antagonistic and modifier alleles, will generally be constrained by X-linked inheritance (Figure 4). The conditions permitting such a process are broader on the autosomes. Furthermore, the mean waiting time until the resolution of sexual antagonism (i.e., the invasion and fixation of $A_1$ and $B_2$ alleles) will be abbreviated on the autosomes relative to the X. These results apply to both male- and female-specific expression divergence. The underlying mechanism for the bias off the X is similar to the case of male-specific divergence under antagonistic pleiotropy: sexually antagonistic fitness costs (to the sex exposed to stabilizing selection) are exacerbated by X-linkage and minimized by autosomal linkage, which limits opportunities for the invasion of sexually antagonistic alleles on the X. These results are in agreement with a recent analysis by Fry (2010), which focused on opportunities for stable polymorphism on the X and autosomes, and runs counter to the intuition that sexual antagonism will generally lead to an enrichment of sex-linked, sex-biased genes.

**SIMULTANEOUS INVASION OF EPISTATICALLY BENEFICIAL MUTATIONS**

Strong selective constraints can limit opportunities for gene expression divergence. If individual mutations evolve under a net purifying selection (averaged across tissues or across the sexes; formally, when $\lambda_A < 1$ and $\lambda_B < 1$; Table 2), and purifying selection is strong enough to prevent the fixation of individual mutations by genetic drift ($1 - \lambda_A$, $1 - \lambda_B \gg 1/N$), then sex-specific evolutionary divergence requires the simultaneous substitution of positively interacting pairs of mutations. These pairs of mutations can simultaneously invade when selection in favor of a double-mutant haplotype is strong relative to the probability of recombination between the mutations (Crow and Kimura 1965; Weinreich and Chao 2005; Kim 2007).

The specific parameter conditions that are conducive to simultaneous invasion can systematically differ between sex chromosomes and autosomes, with invasion conditions for the derived haplotype depending on: (1) whether the haplotype is favored in males or in females; (2) whether it is X-linked or autosomal; and (3) the degree of linkage ($1 - r_f$, $1 - r_m$) between the two loci. The eigenvalue $\lambda_{AB}$ (Table 2) characterizes whether an $A_2B_2$ haplotype can invade a population fixed for $A_1B_1$. Assuming that each $A_2B_2$ haplotype causes expression divergence within the appropriate sex and tissue (i.e., sexually antagonistic or pleiotropic fitness costs are absent in the beneficial haplotype $A_1B_2$; under male-specific selection, fitness is defined as $m_{11} = m_1 = 1 - s_m$,

$$m_{22} = 1 - h_d s_m , \text{ and } m_2 = 1$$

under female-specific selection, fitness is defined as $f_{11} = 1 - s_f$, and $f_{22} = 1 - h_d s_f$; these are used to evaluate $\lambda_{AB}$, simultaneous invasion of a male-beneficial haplotype will be favored on the autosomes when

$$s_m > \frac{r_m + r_f}{1 + r_f - h_d(1 - r_m)}.$$  \hspace{1cm} (10a)

Simultaneous invasion of an autosomal, female-beneficial haplotype can occur when

$$s_f > \frac{r_m + r_f}{1 + r_m - h_d(1 - r_f)}.$$  \hspace{1cm} (10b)

Under X-linkage, simultaneous invasion of a male-beneficial haplotype can occur when

$$s_m > \frac{2r_f}{1 + r_f}.$$  \hspace{1cm} (10c)

Invasion of a female-beneficial haplotype is favored when

$$s_f > \frac{r_f}{1 - h_d(1 - r_f)}.$$  \hspace{1cm} (10d)

The waiting time until simultaneous invasion of a pair of linked alleles depends on both the probability of invasion of an $A_2B_2$ haplotype, as well as the rate at which $A_2B_2$ haplotypes are created by mutation or recombination. Assuming $1 \gg \lambda_{AB} - 1 \gg 1/N$, the probability of invasion and eventual fixation of a male-beneficial haplotype will be

$$\Pr(A_2B_2|m, \text{ aut.}) \approx 2(\lambda_{AB} - 1) = \frac{(1 - h_d s_m)(1 - r_m) - (1 - s_m)(1 + r_f)}{(1 - s_m)}.$$  \hspace{1cm} (11a)

on the autosomes, and:

$$\Pr(A_2B_2|m, X) \approx \frac{1 - r_f}{2} \left( 1 + \sqrt{1 + \frac{1}{(1 - s_m)(1 - r_f)}} \right) - 2$$  \hspace{1cm} (11b)

on the X. The invasion probability of a female-beneficial haplotype will be

$$\Pr(A_2B_2|f, \text{ aut.}) \approx \frac{(1 - h_d s_f)(1 - r_f) - (1 - s_f)(1 + r_m)}{(1 - s_f)}.$$  \hspace{1cm} (11c)

on the autosomes, and

$$\Pr(A_2B_2|f, X) \approx \frac{(1 - h_d s_f)(1 - r_f)}{2(1 - s_f)} \times \left( 1 + \sqrt{1 + \frac{1 - s_f}{(1 - h_d s_f)(1 - r_f)}} \right) - 2$$ \hspace{1cm} (11d)

on the X.

To calculate the rate at which $A_2B_2$ haplotypes are generated, we assume (following previous theory; see Weinreich and Chao
2005; Kim 2007) that the rate of recombination between repulsion haplotypes, $A_2B_1$ and $A_1B_2$, is negligible (as expected when $r_m$, $r_f \ll 1$, a requirement for invasion, and when $A_2B_1$ and $A_1B_2$ are rare, as expected under purifying selection), and that parameters of mutation and purifying selection are the same at locus $A$ and $B$. Given these simplifications, the rate at which $A_2B_2$ haplotypes are created is $2u^2/\omega$, where $u$ is the mutation rate at each locus, per generation, and $\omega$ is the net strength of purifying selection acting on each deleterious haplotype. Incorporating the rate of creation and probability of fixation for $A_2B_2$, the mean waiting time until invasion will be

$$T(A_2B_2|i, j) = \frac{\omega_j}{2N_f u^2 \Pr(A_2B_2|i, j)}.$$

(12)

where $j$ can refer to $X$ or autosomal linkage, and $i$ refers to male or female selection (subscripts $m$ and $f$). Characterizing the net strength of purifying selection on deleterious haplotypes requires a detailed specification of the nature of sex-specific selection acting on $A_2B_1$ and $A_1B_2$ haplotypes. Although space limitation precludes a detailed analysis of all possibilities for the $X$ and autosomes, we present waiting time results under different ratios: $R_m = \omega_X/\omega_A$. This is the most salient factor affecting the relative rate of $A_2B_2$ creation on each chromosome.

Invasion conditions and waiting times, under the simultaneous substitution model, are presented in Figures 5 and 6. In species where recombination only occurs in females (e.g., *Drosophila*), the invasion conditions of female-beneficial haplotypes are roughly the same between the $X$ and autosomes, whereas male-beneficial haplotypes are constrained by $X$-linkage (Fig. 5). In species with recombination in both sexes (e.g., mammals), $X$-linkage expands the invasion conditions of both female- and male-beneficial haplotypes (Fig. 6). These species-specific results reflect an underlying constraint that will often prevent pairs of mutations from invading simultaneously: rare, beneficial allelic combinations can spread if they tend to be co-inherited. Recombination between loci breaks apart beneficial genetic combinations, thereby preventing the invasion of adaptive genetic complexes. A closer look at conditions (10a–10d) provides an explanation for the effect of recombination on the relative haplotype invasion conditions on the $X$ and autosomes. When there is no recombination in males ($r_m = 0$), the minimum selection coefficient required for a female-beneficial haplotype to invade is identical between the $X$ and autosomes; with male-specific selection, autosomal linkage is more conducive to invasion when $h_f < 0.5$, approximately (assuming that $r_f \ll 1$). As $r_m$ increases, recombination becomes effectively higher on the autosomes, the minimum conditions necessary for invasion will increase for autosomal haplotypes, and $X$-linkage will facilitate the invasion of both male- and female-beneficial allelic combinations.

The waiting time until beneficial haplotypes become fixed is sensitive to the strength of positive selection, the effective rate of recombination, and the pattern of net purifying selection against individual derived alleles (Figs. 5 and 6). When the strength of purifying selection against disfavored alleles is greater on the $X$...
Sex linkage and sexual dimorphism

Figure 5. Simultaneous invasion conditions and waiting times for epistatically beneficial haplotypes when males do not recombine. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until double-mutant haplotypes invade, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage. The net autosomal purifying selection against $A_2B_2$ or $A_1B_1$ haplotypes is $\omega_A = 0.001$, with results for two $\omega_X/\omega_A$ ratios shown. Results are shown for $k = 2$ ($h_d = \frac{1}{4}$), $s_f = 0.01$ for the case of directional selection in males, and $s_m = 0.01$ for directional selection in females; mutation rates per locus and per sex are $u = v = 10^{-7}$; X and autosome effective sizes are $N_A = 10^6 = 4N_X/3$. Results were obtained using the SSMW approximations presented in the text.

(i.e., $\omega_X/\omega_A > 1$), the underlying conditions permitting invasion will be unaffected. However, the waiting time until invasion will increase on the X, due to a decrease in the rate at which $A_2B_2$ haplotypes arise within the population.

Because our model invokes an interaction between two pre-specified loci, and the ancestral population ($A_1B_1$ fixed) is initially two mutation steps away from a beneficial haplotype, the waiting time for simultaneous invasion can be relatively long. Nevertheless, the actual rate of epistatic coevolution might still be substantial. Simultaneous invasion is limited by the rate at which beneficial haplotypes arise within a population. This, in turn, depends on two factors: (1) the number of genes, at any given time, that are exposed to sex-specific selection for expression divergence; and (2) the number of epistatically interacting mutational combinations that influence sex-specific expression variation at each gene. If there are a large number of genes under sex-specific disruptive selection, or many epistatically beneficial allele combinations for each gene, then epistatic transitions toward sexually dimorphic expression may be common.

Discussion

The ubiquity of sexual reproduction (Bell 1982; Rice 2002), along with widespread observations of selection differences between males and females (Andersson 1994; Arqvist and Rowe 2005), has inspired a large and growing body of evolutionary theory. Sex-specific selection has implications for several important topics in evolutionary biology, including the maintenance of genetic variation for fitness, the genomic architecture of species differences and sex-specific traits, the opportunity for and rate of adaptation in sexually reproducing populations, and the evolution of female mating biases (e.g., Manning 1984; Kondrashov 1988; Koeslag and Koeslag 1994; Whitlock 2000; Agrawal 2001; Rice and Chippindale 2001; Siller 2001; Lorch et al. 2003; Fedorka...
and Mousseau 2004; Albert and Otto 2005; Fischetta and Chippindale 2006; Scotti and Delph 2006; Hadany and Beker 2007; Calsbeek and Bonneau 2008; Candolin and Heuschele 2008; Bonduriansky and Chenoweth 2009; Cox and Calsbeek 2009; Van Doorn 2009; Whitlock and Agrawal 2009; Blackburn et al. 2010; Connallon 2010; Connallon et al. 2010; Cox and Calsbeek 2010).

The theory presented here builds upon several independent contributions, including the population genetics of sexual antagonism (e.g., Owen 1953; Haldane 1962; Kidwell et al. 1977; Pamilo 1979; Patten and Haig 2009; Fry 2010; and especially Rice 1984) and antagonistic pleiotropy (e.g., Curtsinger et al. 1994; Prout 1999), X versus autosome molecular evolution (Charlesworth et al. 1987; Kirkpatrick and Hall 2004; Vicoso and Charlesworth 2009a), the evolution of peak shifts (e.g., Crow and Kimura 1965; Weinreich and Chao 2005; Kim 2007), and the evolutionary and physiological basis of allelic dominance (e.g., Wright 1934; Kacser and Burns 1981; Dekel and Alon 2005; Gout et al. 2010).

Below, we discuss the potential role of sex-specific selection during the evolution of sex-biased gene expression, by: (1) outlining the predictions of our models within the context of previous theory; and (2) contrasting observed genomic patterns of sex-biased gene expression with theoretical predictions.

**SEX-SPECIFIC SELECTION AND THE EVOLUTION OF SEX-BIASED GENE EXPRESSION**

Sex-biased gene expression might evolve through adaptive shifts by males or by females. For example, a male-biased gene might evolve in response to stabilizing selection in males and selection for decreased expression in females, or through stabilizing selection in females and selection for increased expression in males. In other words, the observation of sex-biased expression is consistent with selection for either male or female gene expression divergence, and additional information is required to equate sex-biased expression with male- or female-specific adaptive divergence (Connallon and Knowles 2005). To interpret genomic patterns of sex-biased expression within the theoretical framework developed here, the frequency of each evolutionary route...
toward male- and female-biased gene expression must be known. Within *Drosophila*, the answer appears to be clear: the evolution of male-biased gene expression typically involves expression increases in males (male divergence); female-biased gene expression involves expression increases in females (female divergence; Vicoso and Charlesworth 2009b). Although we acknowledge that additional research will be required to assess the generality of this pattern, the following discussion is based on this clear pattern from *Drosophila*. That is, our interpretation of the data assumes that male-biased genes evolve by directional selection for male expression divergence, and female-biased genes evolve by directional selection for female expression divergence. Our theoretical results are valid whether this assumption holds true, yet interpretation of the empirical data, in light of the theory, is subject to revision pending future research.

Given the empirical relationship between sex-biased gene expression and sex-specific divergence described above, sex-specific selection theory makes five predictions about the genomic distribution of male- and female-biased genes.

1. When mutations underlying sex-specific expression divergence have sex-limited and tissue-specific effects, genes that are highly expressed in the heterogametic sex (males in species with an X, females with a Z) will more readily evolve on autosomes. Genes that are highly expressed in the homogametic sex (XX females or ZZ males) may become preferentially sex linked when the effective mutation rate is relatively large per sex-linked locus ($4N_uA > 3N_Au_A$ or $4N_Zu_Z > 3N_Au_A$)—a condition that is likely to vary between different animal lineages (for recent reviews of the theory and data, see Charlesworth 2009; Ellegren 2009). This prediction has obvious parallels with those of sex-linked versus autosomal molecular evolution theory (see Kirkpatrick and Hall 2004; Vicoso and Charlesworth 2009a).

2. Pleiotropy greatly accentuates the patterns predicted under a model of sex- and tissue-limited mutation. Genes that are highly expressed within the homogametic sex are expected to accumulate at similar rates on the sex chromosomes and autosomes, and may disproportionately accumulate on the X or Z when $4N_uA > 3N_Au_A$ or $4N_Zu_Z > 3N_Au_A$. Sex linkage severely constrains adaptive gene expression divergence in the heterogametic sex, by reducing the range of parameters that are conducive to evolutionary divergence and by extending the mean waiting time until favored alleles invade in the population. To our knowledge, this result has not previously been reported in the population genetics literature.

3. Sexual antagonism over a gene’s expression level can represent a greater evolutionary constraint for sex-linked loci. Under a model of sequential coevolution between sexually antagonistic alleles and modifiers of sex-limited expression, sex chromosomes are less hospitable to both male- and female-biased genes. This result contradicts the widespread intuition that sex linkage should generally promote the evolution of sex-biased expression from an initially sexually antagonistic state—an expectation based on the assumption of constant allelic dominance for beneficial and deleterious mutations (see Rice 1984; Patten and Haig 2009; Fry 2010). Concave fitness landscapes represent an interesting and possibly widespread example of context-dependent dominance, where deleterious alleles will be partially recessive and beneficial alleles will be partially dominant. Fry (2010) recently demonstrated that such dominance reversals will accommodate an excess of balanced sexually antagonistic polymorphism on the autosomes relative to the X. Our study extends this result to show that, if sex-biased expression divergence involves the sequential invasion of sexually antagonistic alleles, followed by modifiers of sex-limited expression divergence, sex-biased genes are likely to accumulate more readily on autosomes (i.e., the invasion conditions are more permissive, and the waiting time until resolution is shorter).

4. The relative rate of recombination in males versus females can have a major influence on the evolution and genomic distribution of sex-biased genes. In species where recombination occurs in one sex only (i.e., the homogametic sex; Haldane 1922; Huxley 1928; Lenormand and Dutheil 2005), the probability of invasion for sex-linked relative to autosomal genetic combinations that cause sex-biased expression will decrease when they benefit the nonrecombining sex, and increase when they benefit the recombining sex. When both sexes recombine, sex linkage will facilitate the evolution of both male- and female-biased gene expression. This finding also provides a novel prediction for evolutionary theories of adaptive peak shifts, which had not previously considered sex-linked inheritance (e.g., Crow and Kimura 1965; Weinreich and Chao 2005; Kim 2007).

5. The relative contribution of linkage and epistasis to genomic patterns of sexually dimorphic expression will ultimately depend on the strength of evolutionary constraints influencing the evolution of interacting loci. As such, our results can be viewed along a continuum of evolutionary constraint. Under weak constraint, individual derived alleles can reach a high population frequency or become fixed in the population. This initial invasion event can also generate selection for epistatic modifiers, and physical linkage between epistatically interacting loci becomes less relevant to the evolutionary dynamics. Under strong constraints, antagonistically selected alleles persist as rare balanced A
polymorphisms or as ephemeral deleterious mutations. Coadaptation between these alleles and expression modifiers is facilitated by, and in some cases may require, tight physical linkage between the interacting loci. This suggests that regional genomic patterns of linkage and recombination will most strongly influence the evolution of sex-biased genes from strongly constrained precursors. Weakly constrained genes may show less sensitivity to physical linkage.

**SEX-SPECIFIC SELECTION AND THE OBSERVED CHROMOSOMAL DISTRIBUTION OF SEX-BIASED GENES**

Observed chromosomal distributions of sex-biased genes are compatible with the theory outlined here. One of the most striking empirical patterns is the positive association between sex linkage and preferential expression within the homogametic sex. Such linkage patterns have been reported in species of *Caenorhabditis elegans*, *Drosophila*, mouse, silkworm, and chicken (e.g., Ranz et al. 2003; Khil et al. 2004; Reinke et al. 2004; Kaiser and Ellegren 2006; Arunkumar et al. 2009). This particular pattern is exclusively predicted by models of sex-specific selection, and can potentially arise via multiple evolutionary pathways (e.g., through invasion of sex-limited mutations and coevolution of epistatically interacting, linked mutations).

Genes that are preferentially expressed within the heterogametic sex exhibit inconsistent chromosomal distributions between species, yet these lineage-specific patterns are both informative and compatible with models of sex-specific selection. In *Drosophila*, where males do not recombine, there is a deficit of X-linkage for genes with male-biased expression and male-specific function (e.g., accessory gland proteins; Swanson et al. 2001; Parisi et al. 2003; Mueller et al. 2005; Sturgill et al. 2007). In birds and mammals, where recombination rates are relatively similar between the sexes (Groenen et al. 2000; Stauss et al. 2003; Hansson et al. 2005; Lenormand and Dutheil 2005; Akesson et al. 2007; Backström et al. 2008; Hale et al. 2008; Stapley et al. 2008; Jaari et al. 2009), genes that are preferentially expressed in the heterogametic sex are enriched on the X or Z (Wang et al. 2001; Lercher et al. 2003; Khil et al. 2004; Mořkovský et al. 2010). These lineage-specific patterns suggest an important role for epistatic co-evolution in the evolution of sex-biased expression, and also agree nicely with widespread observations of cis-interactions mediating sex-specific regulation during development (Williams and Carroll 2009).

Whether interacting mutations are fixed sequentially or simultaneously, a negative relationship is expected between sex linkage and the relative pleiotropy of male-biased genes. To date, two studies have been published that are relevant to this prediction. Fitzpatrick (2004) analyzed the chromosomal distribution of 63 *Drosophila* “sexually selected” genes associated with courtship behavior, mating receptivity and seminal fluid, and found that the number of X-linked genes was proportional to the size of the X (as a function of euchromatin), suggesting a random chromosomal distribution of sexually selected genes. However, many genes in the study were detected through their mutant phenotypes, and X-linked mutations are often easier to detect (e.g., Haldane 1935). Using a random set of visible mutations to establish a baseline expectation of X-linkage (Table A1 of Fitzpatrick 2004), the proportion of sex-linked, sexually selected genes appears to be lower than expected by chance (Table A1 set: 39% X-linked; sexually selected set: 27% X-linked). Because mutations at these sexually selected genes are generally pleiotropic, the pattern agrees with the theoretical prediction regarding X-linkage and pleiotropy. In a human study using gene expression data from 14 different tissues, Lercher et al. (2003) found that expression breadth (as a metric of pleiotropy) was significantly lower for X-linked relative to autosomal genes. An analysis of male-specific tissues (genes exclusively expressed in the prostate) enhances this bias toward low pleiotropy on the X (again, as predicted by the theory).

Two additional properties of X and Z chromosomes are likely to influence linkage patterns of sex-biased genes. Sex chromosomes are inactivated during male meiosis in mammals and *Drosophila* (Lifschytz and Lindsay 1972; Solari 1974; Hense et al. 2007; Turner 2007), and during female meiosis in the chicken (Schoenmakers et al. 2009). There is now considerable evidence from *Drosophila* that this process of meiotic sex chromosome inactivation (MSCI) constrains the evolution and retention of sex-linked male-biased genes, particularly those expressed in testis (Arbeitman et al. 2002; Bétran et al. 2002; Parisi et al. 2003; Wu and Xu 2003; Meisel et al. 2009; Vibranovski et al. 2009a,b). Constraints against sex-biased gene expression may also arise because the heterogametic sex carries a single copy of each X- or Z-linked gene. To the extent that gene expression levels are constrained by dosage compensation, highly expressed genes within the heterogametic sex are expected to be disproportionately autosomal (Rogers et al. 2003; Vicoso and Charlesworth 2006), a prediction with some support from *Drosophila* (Vicoso and Charlesworth 2009b; Bachtrog et al. 2010). Because these processes tend to skew male-biased genes toward autosomes, evolutionary constraint imposed by MSCI and dosage compensation will tend to reinforce linkage patterns driven by sex-specific selection.

A distinct lack of nonrandomness between X and autosome gene content has been reported for the mosquito *Anopheles gambiense* (Hahn and Lanzaro 2005). This result may be shaped by a suite of lineage-specific factors, including MSCI (reported to occur in Anopheles; McKee and Handel 1993), dosage compensation (unconfirmed), as well as linkage and recombination. Recombination occurs in male and female Anopheles—possibly at similar rates (Benedict et al. 2003)—which could effectively depress the recombination rate per generation on the X and
generate a more favorable environment for male-biased genes compared to Drosophila. On the other hand, sample size may affect these conclusions, as relatively few genes included in the analysis are X-linked. Considering the entire dataset and using a twofold gene expression cut-off to define male- and female-biased expression (M/F > 2 and M/F < 0.5, respectively; data obtained from the SEBIDA database: Gnad and Parch 2006), A. gambiae shows a slight enrichment in female-biased genes (X = 8.3%; n = 782 female-biased genes) and a deficit of male-biased genes (X = 5.0%; n = 481 male-biased genes) relative to the genome-wide expectation (X = 7.1% for the entire dataset; n = 4281 total genes). However, given the low proportion of X-linkage, neither difference is statistically significant.

THE SHAPE OF THE EXPRESSION-FITNESS LANDSCAPE

Predictions of the theory depend on the shape of the relationship between gene expression and fitness, which we have assumed follows a concave function, at least within the region of trait space where the evolution of sexually dimorphic expression occurs. If most gene expression, evolution occurs within the vicinity of a local fitness optimum, then the assumption of fitness concavity does not appear to be controversial, and indeed, several lines of empirical and theoretical evidence support it (e.g., Wright 1934; Charlesworth 1979; Kacser and Burns 1981; Dekel and Alon 2005; Lunzer et al. 2005; Phadnis and Fry 2005; Kalisky et al. 2007; Bedford and Hartl 2009; Zhang et al. 2009; see above). This assumption is further justified by analogy to mathematical theories of DNA adaptation, which suggest that beneficial substitutions will follow a diminishing returns function (Gillespie 1984; see Orr 2005 for a clear discussion of this body of work). The specific degree of fitness concavity, which we incorporate into our model with parameter k (where the fitness function is concave whenever k > 1, and the degree of concavity increases with k), cannot currently be inferred empirically. We simply note that the actual value of k has little impact on qualitative predictions of the models as long as k > 1 (see Figures S3–S6 for analyses with different values of k). Thus, to the extent that selection for sex-specific divergence occurs on a concave expression-fitness function, our qualitative predictions will be robust to the actual degree of concavity.

The concavity assumption may be violated for genes that, prior to adaptation, are severely distant from a local fitness optimum. Although concavity near the optimum is expected, linear (k = 1) or convex (0 < k < 1), fitness landscapes could potentially arise within other regions of trait space. Given our analysis above, convex fitness surfaces will cause a shift toward recessive fitness effects of beneficial mutations (in terms of our model, h0 < 0.5 when 0 < k < 1). Under this condition, the probability of invasion for a male-beneficial mutation will be enhanced by X-linkage, a scenario that is more in line with Rice’s (1984) classic analysis of sexually antagonistic alleles.

Although we focus on cis-regulatory substitutions, our results can easily be generalized to incorporate a possible role of trans-regulatory mutations during gene expression evolution. As with cis-regulatory modifiers of sex- or tissue-limitation, the invasion of pleiotropic or sexually antagonistic alleles can potentially generate selection in favor of trans-acting modifier alleles (this is similar to the scenario studied by Rice 1984). To the extent that trans-acting modifiers are unconstrained by pleiotropy, sequential coevolution between cis- and trans-alleles may be likely. This also opens up the possibility of epistatic coevolution between loci on different chromosomes (including coevolution between the X, Y, and autosomes), which may be important during the evolution of reproductive isolation between species (see Ortiz-Barrientos et al. 2007). However, simultaneous invasion of positively interacting cis- and trans-mutations will generally be impossible if such mutations are not closely linked. Thus, cis-cis-epistasis should permit a broader range of evolutionary transitions relative to cis-trans- or trans-trans-interactions. Because trans-regulatory mutations are also expected to be more pleiotropic, they should provide little opportunity for adaptation compared to cis-regulatory mutations (Stern 2000; Prud’homme et al. 2007; Wray 2007). Nevertheless, current genomic analyses of sex-biased gene expression cannot rule out a role of trans-substitutions during the evolutionary origin of sex-biased expression, and future empirical research will be required to further address this issue.

Conclusion

Sex-specific selection, particularly selection for gene expression divergence between males and females, can play a prominent role in shaping the genomic distributions of sex-biased genes and genes underlying sexual dimorphism. We show that, although the underlying patterns of selection can be complex, the genomic predictions of sex-specific selection hypotheses are generally consistent with empirical distributions of sex-biased genes, although the current data apply to a relatively small (but growing) set of well-characterized species. Future genome-wide analyses of sex-specific genetic architecture, including the distributions of genes with sex-specific expression patterns in nonmodel organisms, will permit a better evaluation of these theoretical predictions.

One notable pattern emerging from this theory is that chromosomal distributions of sex-biased genes are likely to have both gene- and species-specific attributes. These predictions may help to inform future genomic analyses, and may point toward the specific evolutionary processes underlying a diversity of linkage patterns between gene categories and species.

Finally, we focus on selection acting on gene expression variation in an attempt to explain the evolution of a gene expression
phenotype (i.e., sex-biased gene expression) that is ubiquitous at the molecular genomic scale. Intersexual divergence in gene expression may play a general role during the evolution of sexually dimorphic phenotypes, including sex-specific morphology, behavior, and life history. On the other hand, the genetic basis of sexual dimorphism may critically depend on the types of genetic substitutions that are differentially favored between males and females. In principle, disruptive selection might favor protein sequence divergence between the sexes, yet it is difficult to see how conflicts over coding sequence would necessarily become resolved by the evolution of dimorphic gene expression alone. Additional theory, particularly theory involving the evolution of alternative splicing and gene duplication, will be required to address how sexual dimorphism might evolve from an initial condition of coding sequence conflict between the sexes. Experimental research will also be required to better understand the degree to which coding and noncoding mutations contribute to sex-specific fitness variation (for a promising step toward this difficult goal, see Innocenti and Morrow 2010).

ACKNOWLEDGMENTS
This work benefited greatly from discussions with members of the Clark Lab, especially R. Meisel, and comments from three anonymous reviewers. Funding was provided by an NIH grant (GM64590) to AGC and A. Bernardo Carvalho.

LITERATURE CITED


Appendix 1: Development of The Recursions

Haplotypes and sex-specific frequencies

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency in Sperm</th>
<th>Frequency in Eggs</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1B_1</td>
<td>y_1</td>
<td>x_1</td>
</tr>
<tr>
<td>A_1B_2</td>
<td>y_2</td>
<td>x_2</td>
</tr>
<tr>
<td>A_2B_1</td>
<td>y_3</td>
<td>x_3</td>
</tr>
<tr>
<td>A_2B_2</td>
<td>y_4</td>
<td>x_4</td>
</tr>
</tbody>
</table>

Following the convention of much of the sex-specific selection literature of tallying allele (or in our case haplotype) frequencies within the eggs and sperm (e.g., Owen 1953; Kidwell et al. 1977; Patton and Haig 2009), we developed the following recursion equations for the autosomes. Expected frequencies of each haplotype within male and female gametes in the next generation are
Within a given generation, mean fitness in males and females, respectively, at the autosomal loci will be

\[
\tilde{w}_m = x_1 y_1 m_{11} + (x_1 y_2 + x_2 y_1) m_{12} + x_2 y_2 m_{13} + (x_1 y_3 + x_3 y_1) m_{21} + (x_2 y_3 + x_3 y_2) m_{22} + (x_1 y_4 + x_4 y_1) m_{31} + (x_2 y_4 + x_4 y_2) m_{32} + (x_1 y_5 + x_3 y_3) m_{33} + (x_2 y_5 + x_4 y_3) m_{43} + (x_5 y_1 + x_1 y_5) m_{51} + (x_5 y_2 + x_2 y_5) m_{52} + (x_5 y_3 + x_3 y_5) m_{53} + (x_5 y_4 + x_4 y_5) m_{54}.
\]

\[
\tilde{w}_f = x_1 y_1 f_{11} + (x_1 y_2 + x_2 y_1) f_{12} + x_2 y_2 f_{13} + (x_1 y_3 + x_3 y_1) f_{21} + (x_2 y_3 + x_3 y_2) f_{22} + (x_1 y_4 + x_4 y_1) f_{31} + (x_2 y_4 + x_4 y_2) f_{32} + (x_1 y_5 + x_3 y_3) f_{33} + (x_2 y_5 + x_4 y_3) f_{34} + (x_5 y_1 + x_1 y_5) f_{43} + (x_5 y_2 + x_2 y_5) f_{44} + (x_5 y_3 + x_3 y_5) f_{53} + (x_5 y_4 + x_4 y_5) f_{54}.
\]

Appendix 2: Stability with \( A_1 \) and \( B_1 \) Fixed

To evaluate stability of the equilibrium \( x_1 = y_1 = 1 \), we calculated the following Jacobian matrices for the system of X-linked and autosomal recursions. These matrices, evaluated at the aforementioned state \( (x_1 = y_1 = 1) \), and rearranged to block-triangular form, are

\[
J_A|_{x_1=y_1=1} = \begin{pmatrix}
0 & 0 & \frac{2x_1 y_1 f_{11}}{2\tilde{w}_f} & \frac{2x_1 y_2 f_{12}}{2\tilde{w}_f} & \frac{2x_2 y_2 f_{13}}{2\tilde{w}_f} & \frac{2x_1 y_3 f_{21}}{2\tilde{w}_f} & \frac{2x_1 y_4 f_{22}}{2\tilde{w}_f} & \frac{2x_1 y_5 f_{31}}{2\tilde{w}_f} & \frac{2x_1 y_6 f_{32}}{2\tilde{w}_f} & \frac{2x_1 y_7 f_{33}}{2\tilde{w}_f} & \frac{2x_2 y_5 f_{34}}{2\tilde{w}_f} & \frac{2x_5 y_1 f_{51}}{2\tilde{w}_f} & \frac{2x_5 y_2 f_{52}}{2\tilde{w}_f} & \frac{2x_5 y_3 f_{53}}{2\tilde{w}_f} & \frac{2x_5 y_4 f_{54}}{2\tilde{w}_f}
\end{pmatrix}
\]

Mean fitness in males and females, respectively, for X-linked loci will be

\[
\tilde{w}_m = x_1 m_1 + x_2 m_2 + x_3 m_3 + x_4 m_4 + x_5 m_5.
\]
Because both matrices follow the block triangular form, their eigenvalues can be determined from each submatrix along the diagonal (Otto and Day 2007). Excluding the first $2 \times 2$ matrix of zeros, the leading eigenvalues of the remaining three submatrices are each candidates for the leading eigenvalue of the overall matrix. These three candidate eigenvalues are reported in Table 2, for each mode of inheritance.

**Appendix 3: Stability with $A_2$ Present and $B_2$ Absent**

**(1) Tight Linkage Between Locus A and B**

To evaluate stability of the modifier locus under the sequential invasion model, we recalculated the Jacobian matrix with $B_1$ fixed and $A_2$ at arbitrary frequency $x_3$, which represents the equilibrium for $A_2$ when $B_1$ is fixed. The analysis yields an $8 \times 8$ block triangular matrix, with one $4 \times 4$ matrix representing the resident allele $B_1$ and another $4 \times 4$ representing the nonresistant allele $B_2$. By definition, all eigenvalues of the resident submatrix be less than one (the system is stable with $B_1$ fixed). Thus, we focus on the second submatrix.

For autosomal inheritance, the relevant matrix is:

\[
J_A|_{B_1=1} = \begin{pmatrix}
0 & \lambda_x \phi_j & 0 & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & \lambda_x \phi_j & 0 & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 \\
0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j & \lambda_y \phi_j & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j & \lambda_y \phi_j \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j
\end{pmatrix}
\]

When $A_2$ is fixed ($x_3 = 1$), the leading eigenvalue (associated with $A_2B_2$ haplotype) is

\[
\lambda_B = \frac{m_{32}}{2w_m} + \frac{f_{32}}{2w_f}.
\]

When $A_2$ is polymorphic (equilibrium frequency $\hat{p}$), exact eigenvalues are more complex. However, because both loci ($A$ and $B$) are cis-acting, we can approximate them by assuming tight linkage and ignoring squared recombinational terms ($r_m^2$, $r_f^2$, $r_mr_f \approx 0$). This yields a leading eigenvalue:

\[
\lambda_B \approx (1 - \hat{p})f_{32c}(1 - r_f) + \hat{p}f_{32} + \frac{(1 - \hat{p})m_{32c}(1 - r_m) + \hat{p}m_{32}}{2w_f}.
\]

For X-linked inheritance, the relevant matrix is:

\[
J_X|_{B_1=1} = \begin{pmatrix}
0 & \lambda_x \phi_j & 0 & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & \lambda_x \phi_j & 0 & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda_x \phi_j & 0 & \lambda_y \phi_j & 0 \\
0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j & \lambda_y \phi_j & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j & \lambda_y \phi_j \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_x \phi_j
\end{pmatrix}
\]

with characteristic polynomial:

\[
\text{det}(J_X - \lambda I) = \left[\lambda + \frac{m_5}{w_m}\right][1 - \hat{p}]f_{32c}(1 - r_f) + \frac{(1 - \hat{p})m_{32c}(1 - r_m) + \hat{p}m_{32}}{2w_f} - \lambda^2
\]

When $A_2$ is fixed, the leading eigenvalue is

\[
\lambda_B = f_{32c} \frac{\left(\frac{1}{4w_f} + \sqrt{1 + \frac{4w_fm_4}{w_m f_{32}}}ight)}{4w_f}.
\]

When $A_2$ is polymorphic (again, assuming $r_f^2 \approx 0$)

\[
\lambda_B \approx (1 - \hat{p})f_{32c}(1 - r_f) + \hat{p}f_{32} + \frac{(1 - \hat{p})m_{32c}(1 - r_m) + \hat{p}m_{32}}{2w_f}.
\]
(2) LOOSE LINKAGE BETWEEN LOCUS A AND B

For strong recombination relative to selection \((r_{xy} \gg \text{selection})\), and small selection coefficients, we can approximate the eigenvalue associated with a \(B_2\) mutation, assuming linkage equilibrium between \(A\) and \(B\) (for similar analyses, see Otto and Bourguet 1999; Takahasi and Tajima 2005). When both loci are autosomal:

\[
\lambda_B = \frac{(1 - \hat{p})^2 m_{12} + \hat{p}(1 - \hat{p})(m_{22C} + m_{22R}) + \hat{p}^2 m_{32}}{2\hat{w}_m} \\
+ \frac{(1 - \hat{p})^2 f_{12} + \hat{p}(1 - \hat{p})(f_{22C} + f_{22R}) + \hat{p}^2 f_{32}}{2\hat{w}_f}.
\]

For two loosely linked X-linked loci

\[
\lambda_B = \frac{(1 - \hat{p}) m_3 + \hat{p} m_4}{3\hat{w}_m} \\
+ \frac{(1 - \hat{p})^2 f_{12} + \hat{p}(1 - \hat{p})(f_{22C} + f_{22R}) + \hat{p}^2 f_{32}}{3\hat{w}_f}.
\]

Appendix 4: Conditions Favoring Invasion of \(B_2\), Following Invasion of \(A_2\)

(1) ANTAGONISTIC PLEIOTROPY; \(A_2\) FIXED

For an autosomal locus and \(A_2\) fixed, the leading eigenvalue (which is associated with the \(A_2B_2\) haplotype) is

\[
\lambda_B = \frac{1}{2} + \frac{1 - h_d t}{2(1 - r)}
\]

which is greater than one as long as \(h_d < 1\). Since this is always true, \(B_2\) is always favored. For \(A_2\) fixed at an X-linked locus, invasion is favored when:

\[
\frac{2\hat{w}_f - f_{32}}{f_{32}} < \frac{m_4}{\hat{w}_m}
\]

When selection occurs in males, this reduces to \(r > 0\), which is always true. When selection occurs in females, it reduces to \(h_d < 1\), which is always true.

(2) ANTAGONISTIC PLEIOTROPY; \(A_2\) POLYMORPHIC; TIGHT LINKAGE BETWEEN \(A\) AND \(B\)

For autosomal linkage, \(B_2\) is favored under male- and female-specific selection, respectively, when:

\[
1 < \frac{1 - r_f (1 - \hat{p})}{2} + \frac{(1 - \hat{p})(1 - h_d t)(1 - r_m) + \hat{p}(1 - h_d t)}{2\hat{w}_m}
\]

As \(A_2\) goes to fixation \((p \rightarrow 1)\), the previous results apply. Alternatively, as \(p \rightarrow 0\), the minimal requirement for invasion is \(r_m + r_f < s(1 - h_d)\), which is fairly permissive for tight linkage, particularly given our precondition that \(s(1 - h_d) > h_d\)—the minimum requirement for \(A_2\) to invade (eq. 4a). For X-linked inheritance and male-specific selection, \(B_2\) is always favored when \(A_2\) invades because \(A_3\) always becomes fixed. For female-specific selection, invasion of \(B_2\) will occur as long as \(s(1 - h_d) > r_f\), which is likely to occur given our precondition for \(A_2\) invasion: \(s(1 - h_d) > h_d\) (eq. 4a).

(3). ANTAGONISTIC PLEIOTROPY; \(A_2\) POLYMORPHIC; LOOSE LINKAGE BETWEEN \(A\) AND \(B\)

In contrast to the tight linkage scenario, where invasion is primarily determined by the \(A_2B_2\) interaction (relative fitness \(f_{22c}\) and \(m_{22R}\)), a rare \(B_2\) allele will experience selection in a variety of genetic backgrounds. If \(A_2\) is rare, and \(B_2\) is sufficiently deleterious in an \(A_1\) genetic background (e.g., \(f_{12} < f_{11}\) and/or \(f_{22R} < f_{21}\)), invasion may be impossible.

Under the most permissive scenario, where \(B_2\) is neutral on an \(A_1\) background and \(A_2B_2\) is beneficial (i.e., the parameterization of Table S1), the invasion of \(B_2\) will always be favored as long as \(A_2\) is present within the population. For the purpose of illustrating this result, we use the parameterization from Table S1, assuming male-specific selection and autosomal linkage for both loci. Invasion of \(B_2\) is favored when:

\[
0 < \frac{\hat{p}(1 - \hat{p}) h_d t + \hat{p}^2 (1 - h_d) t}{\hat{w}_m}
\]

which simply requires that \(A_2\) be present within the population.

(4). SEXUAL ANTAGONISM; \(A_2\) FIXED

Following fixation of a male-beneficial allele, a modifier of sex-limitation will invade when:

\[
1 < \frac{1 - h_d s_f}{1 - s_f}
\]

which simply requires that \(h_d < 1\), which is always true. The same results apply to modifiers following the invasion of a female-beneficial, autosomal mutation.

For X-linked inheritance, invasion of a modifier is favored when:

\[
\frac{2\hat{w}_f - f_{32}}{f_{32}} < \frac{m_4}{\hat{w}_m}
\]

which requires that \(h_d < 1\) following fixation of a male-beneficial allele and \(s_m > 0\) following fixation of a female-beneficial allele. Both will always be true.

(5). SEXUAL ANTAGONISM; \(A_2\) POLYMORPHIC; TIGHT LINKAGE BETWEEN \(A\) AND \(B\)

Under autosomal inheritance, selection always favors the modifier when \(s_m(1 - h_d) > r_m + r_f\) (following invasion of a male-beneficial allele) and \(s_f(1 - h_d) > r_m + r_f\) (following
invasion of a female-beneficial allele). These minimum criteria are removed as the \( A_2 \) equilibrium increases in frequency.

Under X-linked inheritance, and following invasion of a male beneficial allele, the modifier always invades when \( s_m > 2\gamma \), with this condition relaxed as the equilibrium frequency of \( A_2 \) increases. Following female-beneficial invasion, a modifier will invade as long as \( s_f(1 - h_2) > \gamma \), with this restriction removed for higher frequencies of \( A_2 \).

**Appendix 5: Waiting Times for Modifier Invasion, Following Invasion of \( A_2 \) and Convergence to Equilibrium**

When \( A_1 \) and \( A_2 \) are polymorphic, there are two invasion routes for \( B_2 \):

1. If a \( B_2 \) mutation arises on an \( A_2 \) chromosome (with probability \( \hat{\rho} \)), the \( A_2B_2 \) haplotype will invade with probability \( 2(\lambda_B - 1) \). The mean waiting time for such an event on the autosomes and X, respectively, is

\[
T_{A_2}(\text{inv.} | A_2 B_2) = \frac{1}{2N_A \hat{\rho} v_A (\lambda_B - 1)}
\]

and

\[
T_X(\text{inv.} | A_2 B_2) = \frac{1}{2N_X \hat{\rho} v_X (\lambda_B - 1)}.
\]

where \( v_A \) and \( v_X \) refer to the chromosome-specific rate of mutation from \( B_1 \) to \( B_2 \), and \( \lambda_B \) is the leading eigenvalue for the B locus, assuming \( A_2 \) at frequency \( \hat{\rho} \).

2. If \( B_2 \) initially arises with \( A_1 \) (with probability \( 1 - \hat{\rho} \)), recombination with an \( A_2B_1 \) haplotype can produce \( A_2B_2 \), which once again invades with probability \( 2(\lambda_B - 1) \). The probability of such an event strongly depends on the marginal fitness of an \( A_2B_2 \) haplotype within a population otherwise fixed for \( B_1 \). In the most permissive case, \( A_1B_2 \) is effectively neutral (i.e., \( A_1B_1 \) and \( A_1B_2 \) have the same marginal fitness, as parameterized in Tables S1 and S2). Without recombination, the number of individuals inheriting a neutral \( A_1B_2 \) haplotype destined to be lost from the population will be geometric with mean \( N_A(1 - \hat{\rho}) \) (García-Dorado et al. 2003).

Thus, the probability that a \( B_2 \) allele on an \( A_1 \)-bearing chromosome will escape to an \( A_2 \)-bearing chromosome prior to loss from the population is:

\[
Pr(\text{escape}) = 1 - \sum_{t=1}^{\infty} \frac{[1 - \hat{\rho}(r_m + r_f)/2]^{t-1}}{N_A(1 - \hat{\rho})} \times \frac{N_A \hat{\rho}[1 - \hat{\rho}(r_m + r_f)]}{N_A \hat{\rho}[1 - \hat{\rho}(r_m + r_f) + 2]},
\]

where the approximation assumes that \( \frac{\hat{\rho}(r_m + r_f)}{N_A(1 - \hat{\rho})} \approx 0 \). For the X-linked scenario, the probability of escape is:

\[
Pr(\text{escape}) = 1 - \sum_{t=1}^{\infty} \frac{[1 - 2\hat{\rho}r_f/3]^{t-1}}{N_X(1 - \hat{\rho})} \times \frac{2N_X \hat{\rho}[1 - \hat{\rho}r_f]}{2N_X \hat{\rho}[1 - \hat{\rho}r_f + 3]}.
\]

The waiting time until \( A_2B_2 \) invades, when \( B_2 \) is initially linked to an \( A_1 \) allele, is

\[
T_{A_2}(\text{inv.} | A_1 B_2) = \frac{1}{2N_A \hat{\rho} v_A (\lambda_B - 1)}
\]

and

\[
T_X(\text{inv.} | A_1 B_2) = \frac{1}{2N_X \hat{\rho} v_X (\lambda_B - 1)}.
\]

As selection against \( A_1B_2 \) haplotypes increases (i.e., for \( A_1B_2 \) deleterious), or the \( A_2 \) equilibrium approaches a relatively high frequency within the population (\( \hat{\rho} \rightarrow 1 \), virtually all invasion events proceed from an initial \( A_1B_2 \) genetic background. This suggests that the mean waiting time until \( B_2 \) invasion will fall within the range is

\[\frac{1}{1/T(\text{inv.} | A_1 B_2) + 1/T(\text{inv.} | A_2 B_2)} < T(B_2) < T(\text{inv.} | A_2 B_2).\]

Substituting the autosomal and X-linked terms, the range for each chromosome will be

\[
\frac{N_A \hat{\rho}(r_m + r_f) + 2}{2N_A \hat{\rho} v_A (\lambda_B - 1)[N_A(1 - \hat{\rho})(r_m + r_f) + 2]} < T_{A_2}(B_2) < \frac{1}{2N_A \hat{\rho} v_A (\lambda_B - 1)}
\]
and

\[
\frac{2N_X \hat{p}(1 - \hat{p}) r_f + 3}{2N_X \hat{p}v_X(\lambda_B - 1)(2N_X(1 - \hat{p}) r_f + 3)} < T_X(B_2) < \frac{1}{2N_X \hat{p}v_X(\lambda_B - 1)}.
\]

When \( A_2 \) is rare and \( A_1B_2 \) is neutral, the waiting time approaches the term on the left. As the equilibrium frequency of \( A_2 \) and the deleterious effect of \( A_1B_2 \) increase, the waiting time approaches the term on the right.

**Supporting Information**

The following supporting information is available for this article:

**Figure S1.** Invasion probabilities for sexually antagonistic alleles.
**Figure S2.** Fixation probabilities for haplotypes with sex-specific benefits.
**Figure S3.** Antagonistic pleiotropy, modifier model.
**Figure S4.** Sexual antagonism, modifier model.
**Figure S5.** Double-mutant haplotype invasion model with no recombination in males.
**Figure S6.** Double-mutant haplotype invasion model with equal recombination in males and females.
**Table S1.** Full fitness parameterization for the two-locus pleiotropy modifier model, with \( B_2 \) neutral on an \( A_1 \) genetic background.
**Table S2.** Full fitness parameterization for the two-locus sexual antagonism modifier model, with \( B_2 \) neutral on an \( A_1 \) genetic background.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
I. Concavity of the Fitness Function when Benefits of Gene Expression Saturate

Consider a model of fitness variation based on the balance of benefits and costs of gene expression. Benefits and costs are designated by functions $B(x)$ and $C(x)$, respectively, where the variable $x$ is the gene’s expression level. Fitness, as a function of expression level, is proportional to the difference between benefits and costs: $w(x) \propto B(x) - C(x)$.

Suppose that benefits of gene expression follow a saturating function of gene expression level, as suggested by both theory and data. One possible saturating function is:

$$B(x) = \frac{bx}{1 + ax}$$

where $a$ and $b$ are positive constants. Within the region of $ax \ll 1$, the benefit function will be approximately linear, but will saturate as $ax$ goes to infinity. Fitness will be proportional to:

$$w(x) \propto \frac{bx}{1 + ax} - C(x)$$

(A). Linearly increasing costs. The cost of gene expression is expected to, at minimum, increase linearly as a result of metabolic or energetic allocation tradeoffs. Thus, for a linear cost model fitness is proportional to:

$$w(x) \propto \frac{bx}{1 + ax} - cx$$

Gene expression is optimized at:

$$\hat{x} = \frac{\sqrt{b/c} - 1}{a}$$

which is greater than zero when $b > c$. At the optimum:

$$\frac{d^2w(x)}{dx^2} \bigg|_{x=\hat{x}} = \frac{-2ac}{\sqrt{b/c}}$$

which is negative, and implies a concave fitness function.
(B). **Faster-than-linear cost function.** If costs accumulate faster than linearly, fitness can be redefined as:

\[ w(x) \propto \frac{bx}{1 + ax} - \frac{cx}{1 - x/d} \]

(e.g., Dekel and Alon 2005; Gout et al. 2010) where expression is constrained to fall below some limit: \( 0 < x < d \). The optimal expression level (\( \hat{x} \)) is found at the maximum of the fitness function:

\[ \frac{dw(x)}{dx} = 0 = \frac{b}{(1 + a\hat{x})^3} - \frac{c}{(1 - \hat{x}/d)^2} \]

\[ \hat{x} = \frac{d(\sqrt{b} - \sqrt{c})}{\sqrt{b} + ad\sqrt{c}} \]

which has a nonzero fitness optimum when \( b > c \). The second derivative at the optimum is:

\[ \left. \frac{d^2w(x)}{dx^2} \right|_{x=\hat{x}} = -\frac{2(\sqrt{b} + ad\sqrt{c})^4}{d^4bc(ad + 1)^3} \]

which is always negative, implying concavity.
II. Supplementary Tables

Table S1. Full fitness parameterization for the two-locus pleiotropy modifier model, with $B_2$ neutral on an $A_1$ genetic background.

I. Selection for male expression divergence

Maternal Haplotype:

<table>
<thead>
<tr>
<th>Paternal Haplotype:</th>
<th>$A_1B_1$</th>
<th>$A_1B_2$</th>
<th>$A_2B_1$</th>
<th>$A_2B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1B_1$</td>
<td>$f_{11} = 1$</td>
<td>$f_{12} = 1$</td>
<td>$f_{21} = 1$</td>
<td>$f_{22C} = 1$</td>
</tr>
<tr>
<td>$A_1B_2$</td>
<td>$f_{12} = 1$</td>
<td>$f_{13} = 1$</td>
<td>$f_{22R} = 1$</td>
<td>$f_{23} = 1$</td>
</tr>
<tr>
<td>$A_2B_1$</td>
<td>$f_{21} = 1$</td>
<td>$f_{22R} = 1$</td>
<td>$f_{31} = 1$</td>
<td>$f_{32} = 1$</td>
</tr>
<tr>
<td>$A_2B_2$</td>
<td>$f_{22C} = 1$</td>
<td>$f_{23} = 1$</td>
<td>$f_{32} = 1$</td>
<td>$f_{33} = 1$</td>
</tr>
<tr>
<td>$A_1B_1$</td>
<td>$m_{11} = 1 - s$</td>
<td>$m_{12} = 1 - s$</td>
<td>$m_{21} = 1 - h_s(s + t)$</td>
<td>$m_{22C} = 1 - h_s$s</td>
</tr>
</tbody>
</table>
\[
\begin{align*}
A_1B_2 & \quad m_{12} = 1 \cdot s \quad m_{13} = 1 \cdot s \quad m_{22R} = 1 \cdot h_d(s + t) \quad m_{23} = 1 \cdot h_d s \\
A_2B_1 & \quad m_{21} = 1 \cdot h_d(s + t) \quad m_{22R} = 1 \cdot h_d(s + t) \quad m_{31} = 1 \cdot t \quad m_{32} = 1 \cdot h_d t \\
A_2B_2 & \quad m_{22C} = 1 \cdot h_d s \quad m_{23} = 1 \cdot h_d s \quad m_{32} = 1 \cdot h_d t \quad m_{33} = 1 \\
-- & \quad m_{1} = 1 \cdot s \quad m_{2} = 1 \cdot s \quad m_{3} = 1 \cdot t \quad m_{4} = 1
\end{align*}
\]

II. Selection for female expression divergence

**Maternal Haplotype:**

<table>
<thead>
<tr>
<th>Paternal Haplotype:</th>
<th>(A_1B_1)</th>
<th>(A_1B_2)</th>
<th>(A_2B_1)</th>
<th>(A_2B_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1B_1)</td>
<td>(f_{11} = 1 \cdot s)</td>
<td>(f_{12} = 1 \cdot s)</td>
<td>(f_{21} = 1 \cdot h_d(s + t))</td>
<td>(f_{22C} = 1 \cdot h_d s)</td>
</tr>
<tr>
<td>(A_1B_2)</td>
<td>(f_{12} = 1 \cdot s)</td>
<td>(f_{13} = 1 \cdot s)</td>
<td>(f_{22R} = 1 \cdot h_d(s + t))</td>
<td>(f_{23} = 1 \cdot h_d s)</td>
</tr>
<tr>
<td>(A_2B_1)</td>
<td>(f_{21} = 1 \cdot h_d(s + t))</td>
<td>(f_{22R} = 1 \cdot h_d(s + t))</td>
<td>(f_{31} = 1 \cdot t)</td>
<td>(f_{32} = 1 \cdot h_d t)</td>
</tr>
<tr>
<td></td>
<td>( f_{22C} = 1 - h ds )</td>
<td>( f_{23} = 1 - h ds )</td>
<td>( f_{32} = 1 - h dt )</td>
<td>( f_{33} = 1 )</td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>( A_2B_2 )</td>
<td>( m_{11} = 1 )</td>
<td>( m_{12} = 1 )</td>
<td>( m_{21} = 1 )</td>
<td>( m_{22C} = 1 )</td>
</tr>
<tr>
<td>( A_1B_1 )</td>
<td>( m_{12} = 1 )</td>
<td>( m_{13} = 1 )</td>
<td>( m_{22R} = 1 )</td>
<td>( m_{23} = 1 )</td>
</tr>
<tr>
<td>( A_1B_2 )</td>
<td>( m_{21} = 1 )</td>
<td>( m_{22R} = 1 )</td>
<td>( m_{31} = 1 )</td>
<td>( m_{32} = 1 )</td>
</tr>
<tr>
<td>( A_2B_1 )</td>
<td>( m_{22C} = 1 )</td>
<td>( m_{23} = 1 )</td>
<td>( m_{32} = 1 )</td>
<td>( m_{33} = 1 )</td>
</tr>
<tr>
<td>( A_2B_2 )</td>
<td>( -- )</td>
<td>( m_1 = 1 )</td>
<td>( m_2 = 1 )</td>
<td>( m_3 = 1 )</td>
</tr>
</tbody>
</table>
Table S2. Full fitness parameterization for the two-locus sexual antagonism modifier model, with $B_2$ neutral on an $A_1$ genetic background.

I. Selection for male-specific expression divergence

<table>
<thead>
<tr>
<th>Paternal Haplotype:</th>
<th>$A_1B_1$</th>
<th>$A_1B_2$</th>
<th>$A_2B_1$</th>
<th>$A_2B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1B_1$</td>
<td>$f_{11} = 1$</td>
<td>$f_{12} = 1$</td>
<td>$f_{21} = 1 - h_d s_f$</td>
<td>$f_{22C} = 1$</td>
</tr>
<tr>
<td>$A_1B_2$</td>
<td>$f_{12} = 1$</td>
<td>$f_{13} = 1$</td>
<td>$f_{22R} = 1 - h_d s_f$</td>
<td>$f_{23} = 1$</td>
</tr>
<tr>
<td>$A_2B_1$</td>
<td>$f_{21} = 1 - h_d s_f$</td>
<td>$f_{22R} = 1 - h_d s_f$</td>
<td>$f_{31} = 1 - s_f$</td>
<td>$f_{32} = 1 - h_d s_f$</td>
</tr>
<tr>
<td>$A_2B_2$</td>
<td>$f_{22C} = 1$</td>
<td>$f_{23} = 1$</td>
<td>$f_{32} = 1 - h_d s_f$</td>
<td>$f_{33} = 1$</td>
</tr>
<tr>
<td>$A_1B_1$</td>
<td>$m_{11} = 1 - s_m$</td>
<td>$m_{12} = 1 - s_m$</td>
<td>$m_{21} = 1 - h_d s_m$</td>
<td>$m_{22C} = 1 - h_d s_m$</td>
</tr>
</tbody>
</table>
\[ A_1B_2 \quad m_{12} = 1 - s_m \quad m_{13} = 1 - s_m \quad m_{22R} = 1 - h_d s_m \quad m_{23} = 1 - h_d s_m \]

\[ A_2B_1 \quad m_{21} = 1 - h_d s_m \quad m_{22R} = 1 - h_d s_m \quad m_{31} = 1 \quad m_{32} = 1 \]

\[ A_2B_2 \quad m_{22C} = 1 - h_d s_m \quad m_{23} = 1 - h_d s_m \quad m_{32} = 1 \quad m_{33} = 1 \]

\[ -- \quad m_1 = 1 - s_m \quad m_2 = 1 - s_m \quad m_3 = 1 \quad m_4 = 1 \]

II. Selection for female-specific expression divergence

Maternal Haplotype:

<table>
<thead>
<tr>
<th>Paternal Haplotype:</th>
<th>(A_1B_1)</th>
<th>(A_1B_2)</th>
<th>(A_2B_1)</th>
<th>(A_2B_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1B_1)</td>
<td>(f_{11} = 1 - s_f)</td>
<td>(f_{12} = 1 - s_f)</td>
<td>(f_{21} = 1 - h_d s_f)</td>
<td>(f_{22C} = 1 - h_d s_f)</td>
</tr>
<tr>
<td>(A_1B_2)</td>
<td>(f_{12} = 1 - s_f)</td>
<td>(f_{13} = 1 - s_f)</td>
<td>(f_{22R} = 1 - h_d s_f)</td>
<td>(f_{23} = 1 - h_d s_f)</td>
</tr>
<tr>
<td>(A_2B_1)</td>
<td>(f_{21} = 1 - h_d s_f)</td>
<td>(f_{22R} = 1 - h_d s_f)</td>
<td>(f_{31} = 1)</td>
<td>(f_{32} = 1)</td>
</tr>
</tbody>
</table>
\[ A_2B_2 \quad f_{22C} = 1 - h_\delta s_f \quad f_{23} = 1 - h_\delta s_f \quad f_{32} = 1 \quad f_{33} = 1 \]

\[ A_1B_1 \quad m_{11} = 1 \quad m_{12} = 1 \quad m_{21} = 1 - h_\delta s_m \quad m_{22C} = 1 \]

\[ A_1B_2 \quad m_{12} = 1 \quad m_{13} = 1 \quad m_{22R} = 1 - h_\delta s_m \quad m_{23} = 1 \]

\[ A_2B_1 \quad m_{21} = 1 - h_\delta s_m \quad m_{22R} = 1 - h_\delta s_m \quad m_{31} = 1 - s_m \quad m_{32} = 1 - h_\delta s_m \]

\[ A_2B_2 \quad m_{22C} = 1 \quad m_{23} = 1 \quad m_{32} = 1 - h_\delta s_m \quad m_{33} = 1 \]

\[ -- \quad m_1 = 1 \quad m_2 = 1 \quad m_3 = 1 - s_m \quad m_4 = 1 \]
III. Supplementary Figures

**Figure S1. Invasion probabilities for sexually antagonistic alleles.** Solid lines are based on the eigenvalue-based invasion probability (*i.e.*, the branching process approximation described in the main text) under a model of sexually antagonistic selection. Each filled circle is the result of 500,000 replicate simulations, with the probability of invasion calculated as the proportion of mutations that remain within the population after 10,000 generations of selection and sampling with replacement. Results shown for $N_e = 200,000$ with sex ratio of one (#males = #females = 100,000). In each simulation run, a single copy of a derived allele $A_2$ is introduced into a population fixed for the ancestral allele $A_1$. Fitness parameterization follows Table 4.
Figure S2. Fixation probabilities for haplotypes with sex-specific benefits. Solid lines are based on the eigenvalue-based invasion probability (i.e., the branching process approximation described in the main text). Circles represent the fraction of 100,000 simulation runs where an $A_2B_2$ haplotype (initially one copy within the population) goes to fixation. Results shown for $N_e = 200,000$ with sex ratio of one (#males = #females = 100,000). $A_2B_2$ haplotypes improve fitness by fraction $s_{b}h_{b} = 0.0075$ in heterozygotes ($A_2B_2/A_1B_1$).
and by $s_b = 0.01$ in hemizygotes. Deleterious haplotypes ($A_2B_1; A_1B_2$) reduce fitness by $s_d h_d = 0.001$ in heterozygotes and by $s_d = 0.002$ in hemizygotes.
Figure S3. Antagonistic pleiotropy, modifier model. The results use the same parameterization as Figure 3, with decreased concavity of the fitness function. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until derived alleles invade at the pleiotropic and modifier loci, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage.
Figure S4. Sexual antagonism, modifier model. The results use the same parameterization as Figure 4, with decreased concavity of the fitness function. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until derived alleles invade at the pleiotropic and modifier loci, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage.
Figure S5. Double-mutant haplotype invasion model with no recombination in males. The results use the same parameterization as Figure 5, with decreased concavity of the fitness function, and $\omega_X/\omega_A = 1$. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until derived alleles invade at the pleiotropic and modifier loci, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage.
Figure S6. Double-mutant haplotype invasion model with equal recombination in males and females. The results use the same parameterization as Figure 6, with decreased concavity of the fitness function, and $\omega_X/\omega_A = 1$. Opportunities for invasion are described by the shaded bar at the base of each panel. The curves represent the mean time until derived alleles invade at the pleiotropic and modifier loci, with the uniform curve representing the X-linked scenario and the circle-bearing curve representing autosomal linkage.