

## **Chapter 1**

### **Introduction**

Most sexually reproducing species consist of phenotypically distinct males and females, which each exhibit extensive sexual differentiation for morphological, behavioral, physiological, and life history traits (Darwin 1859, 1871; Andersson 1994). Less obvious, but equally important, is the fact that males and females each make equal genetic contributions to the next generation, yet often exhibit different reproductive variance: male fitness variance is generally greater than female variance (Bateman 1948; Trivers 1972; Andersson 1994; but see Snyder & Gowaty 2007). These observations are trademarks of sexual selection and sexual reproduction, and suggest that patterns of selection on individual mutations might also commonly differ between the sexes.

Given the potential ubiquity of sex-specific selection, it is surprising that the subject has traditionally received little attention from theoretical or experimental perspectives. This has abruptly changed during the last several years, perhaps due to two recent but already well-established findings. The first, from the classical and quantitative genetics literature, is that opposing selection pressures between males and females (or “sexual antagonism” *sensu* Rice 1984) are commonly observed within animal populations, and appear to be important features of genome wide genetic variation for fitness (Chippindale et al. 2001; Gibson et al. 2002; Foerster et al. 2007; Cox & Calsbeek 2009). The second finding, from the field of molecular evolution, is that sexual dimorphism (long thought to be a signature of adaptive sexual differentiation with respect to morphological traits) is a common characteristic of many individual genes. That is, sexual dimorphism for gene expression characterizes up to 50 percent of genes within insect and nematode genomes, as well as a major proportion of tissue-specific expression profiles from birds and mammals (Ellegren & Parsch 2007). Furthermore, the genomic distribution of “sex-

biased genes” follows several interesting, nonrandom patterns of linkage with respect to sex chromosomes and autosomes, which some researchers have argued is consistent with theories of sex-specific selection (Parisi et al. 2003; Connallon & Knowles 2005; Mank & Ellegren 2009). The implication of these discoveries is that differential selection between the sexes: (A) is more common than previously assumed, (B) produces genome-wide signatures of genetic and gene expression variability, and (C) has different evolutionary consequences on sex chromosomes and autosomes.

This dissertation was undertaken with two related goals in mind. The first goal is to describe processes of differential selection between the sexes, and to infer proximate population genetic consequences of sex-specific selection, including those related to the maintenance of genetic variation for fitness, and the heritability of fitness between parents and offspring. The second goal is to better understand the long-term adaptive consequences of sexual reproduction by considering how both sex-specific selection and recombination might influence processes of population adaptation.

Sex-specific selection occurs when the strength and/or the direction of selection acting on DNA substitutions differs between males and females. To illustrate, consider a single locus ( $A$ ), two-allele model of genetic variation ( $A_1, A_2$ , with frequencies  $p$  and  $q$ , respectively):

	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness:	1	$1 + h_f s_f$	$1 + s_f$
Male fitness:	1	$1 + h_m s_m$	$1 + s_m$

where the rate of allele frequency change per generation is approximated by:

$$\frac{dq}{dt} = \frac{1}{2} s_f q(1-q)[q + h_f(1-2q)] + \frac{1}{2} s_m q(1-q)[q + h_m(1-2q)] \quad (1)$$

(see Ewens 2004). For an autosomal gene, which spends half of its evolutionary history within male and half within female bodies, evolutionary dynamics depend on the net affect of selection averaged across the sexes. For selection coefficients that are both positive ( $s_m, s_f > 0$ ), or both negative ( $s_m, s_f < 0$ ), the rate of response to selection depends upon the relative contributions of males and females to population adaptation. Conditions

where selection is stronger in males than females (e.g.,  $|s_m| > |s_f|$ ) cause sexually reproducing species to adapt faster than asexuals, can generate long-term adaptive benefits of sexual reproduction, and will cause a relative reduction in standing genetic variation in sexual versus asexual populations (Kodric-Brown & Brown; Agrawal 2001; Siller 2001). For selection in opposing directions (“sexual antagonism”;  $s_m \propto -s_f$ ), adaptation is constrained by a population genetic “tug-of-war” (Lande 1980; Rice & Chippindale 2001), generates a cost associated with sexual reproduction, and will increase levels of standing fitness variation maintained via balancing selection (Owen 1952, 1953; Haldane 1962; Kidwell et al. 1977) or increased mutation-selection equilibria for alleles with a net negative fitness effect when averaged across the sexes (see Chapter 6).

A peculiar, and empirically useful case of sex-specific selection also arises under conditions of haplodiploid inheritance, which include instances of sex linkage in which the sex-limited chromosome (the W or the Y chromosome) is degenerate. Dominance is eliminated in the haploid or “hemizygous” sex and genotypic fitness is described as:

	$A_1$	$A_2$
Haploid fitness:	1	$1 + t$

which leads to an evolutionary response to selection at a rate:

$$\frac{dq}{dt} = tq(1 - q) \tag{2}$$

(Charlesworth et al. 1987). With haplodiploidy, the response to selection will generally differ between the sexes except under an extremely restrictive condition, i.e.

$t/s = h + q(1 - 2h)$ . Sex-linked genes are also asymmetrically inherited, and spend two-thirds of their evolutionary history within bodies of the “diploid sex” and the remaining one-third within the “haploid sex”. This inheritance asymmetry, coupled with the expectation that selection intensity will often differ between the sexes, suggests that sex-linked inheritance should represent an interesting context to study sex-specific selection.

The dissertation is subdivided into three subsections. Within the first section – **Molecular Signatures of Adaptation on X chromosomes and Autosomes** – population genetic and genomic data from related *Drosophila* species is used to infer patterns of dominance for beneficial mutations. As such, the research contrasts patterns of molecular genetic variation and interspecific divergence between chromosomes and considers these patterns within a theoretical framework of sex-specific directional selection that arises from haplodiploid inheritance. The results are interpreted within a context of adaptive constraint caused by nonadditive allelic effects, and are germane to evolutionary phenomena such as “the faster-X” (Charlesworth et al. 1987) and “Haldane’s sieve” (Haldane 1927; Orr & Betancourt 2001).

The second subsection – **Sex-Linkage, Sex-Specific Selection, and the Heritability of Fitness** – integrates population genetic and quantitative genetics theory to consider how genome architecture and sex-specific selection might influence the heritability of fitness between parents and offspring. The new theory has implications for the field of sexual selection, where a great deal of effort has been directed towards identifying indirect genetic effects of nonrandom mating (Andersson 1994; Møller & Alatalo 1999). As discussed in Chapters 3-5, the theory, a meta-analysis of published research, and new data from an experimental population of *Drosophila melanogaster*, support the idea that sex chromosomes alter fitness relationships between males and females, and fathers and offspring. Estimates of indirect genetic effects in species with different sex chromosome systems are therefore likely to yield different results even if the underlying genetic basis of fitness variation is both substantial and equally abundant within different populations.

The final subsection – **Population Benefits and Costs of Sexual Reproduction** – addresses the adaptive significance of two consequences of sexual reproduction. The first consequence, sex-specific selection, can hypothetically improve or constrain population productivity (a measure of “adaptation”). I consider the joint action of concordant directional selection ( $s_m/s_f > 0$ ) and opposing selection between the sexes ( $s_m/s_f < 0$ ), and analyze how the relative proportion of each type of sex-specific selection influences the net benefits and costs of sexual reproduction. An extensive review of male and female selection estimates from wild and laboratory populations are used to estimate likely parameters from the model. The analysis suggests that differential selection upon the

sexes generates a net adaptive constraint, and therefore adds to other costs of sexual reproduction (e.g., the twofold cost of sex; Maynard Smith 1978).

Recombination represents a second consequence of sexual reproduction and is thought to improve population adaptation by eliminating or reducing genetic associations caused by physical linkage (Fisher 1930; Muller 1932; Crow & Kimura 1965). In particular, tight linkage can inflate stochastic reproductive success variance associated with finite population size, and thereby influence the relative importance of genetic drift versus natural selection (this effect of linkage is also known as “Hill-Robertson” interference; Hill & Robertson 1966; Felsenstein 1974). The adaptive consequences of linkage are therefore predicted to be most profound for mutations of small effect, and less important for mutations of large effect (where the intensity of selection is much greater than the inverse of gene-specific effective population size; Ohta 1973). Using yeast genomic resources, I analyzed correlations between recombination rate and divergence within genes subject to relatively weak versus relatively strong selection and interpret these with respect to theoretical benefits of recombination.

## REFERENCES

- Agrawal AF. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature* 411:692-695.
- Andersson M. 1994. *Sexual selection*. Princeton University Press, Princeton.
- Bateman AJ. 1948. Intra-sexual selection in *Drosophila*. *Heredity* 2:349-368.
- Connallon T, Knowles LL. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends in Genetics* 21:495-499.
- Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* 173:176-187.
- Charlesworth B, Coyne JA, Barton N. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat.* 130:113-146.
- Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proceedings of the National Academy of Sciences USA* 98:1671-1675.
- Crow JF, Kimura M. 1965. Evolution in sexual and asexual populations. *American Naturalist* 99:439-450.
- Darwin C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, London.
- Darwin C. 1871. *The descent of man, and selection in relation to sex*. John Murray, London.
- Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics* 8:689-698.
- Ewens WJ. 2004. *Mathematical population genetics. I. Theoretical introduction*. Springer Science & Business Media, Inc., New York.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.
- Fisher RA. 1930. The distribution of gene ratios for rare mutations. *Proceedings of the Royal Society of Edinburgh* 50:204-219.
- Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH, Kruuk LEB. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447:1107-1110.

- Gibson JR, Chippindale AK, Rice WR. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London B* 269:499-505.
- Haldane JBS. 1927. A mathematical theory of natural and artificial selection, part V: selection and mutation. *Proceedings of the Cambridge Philosophical Society* 28:838-844.
- Haldane JBS. 1962. Conditions for stable polymorphism at an autosomal locus. *Nature* 193:1108.
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genetical Research* 8:269-294.
- Kidwell JF, Clegg MT, Stewart FM, Prout T. 1977. Regions of stable equilibria for models of differential selection in two sexes under random mating. *Genetics* 85:171-183.
- Kodric-Brown A, Brown JH. 1987. Anisogamy, sexual selection, and the origin and maintenance of sex. *Evolutionary Ecology* 1:95-105.
- Lande R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34:292-305.
- Mank JE, Ellegren H. 2009. Sex linkage of sexually antagonistic genes is predicted by female, but not male effects in birds. *Evolution*, in press.
- Maynard Smith J. 1978. *The evolution of sex*. Cambridge University Press, Cambridge, UK.
- Møller AP, Alatalo RV. 1999. Good genes effects in sexual selection. *Proceedings of the Royal Society of London B* 266:85-91.
- Muller HJ. 1932. Some genetic aspects of sex. *American naturalist* 66:118-138.
- Ohta T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature* 246:96-98.
- Orr HA, Betancourt AJ. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics*. 157:875-884.
- Owen ARG. 1952. A genetical system admitting of two stable equilibria. *Nature* 170:1127-1128.
- Owen ARG. 1953. A genetical system admitting of two distinct stable equilibria under natural selection. *Heredity* 7:97-102.

- Parisi M, et al. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science*. 299:697-700.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735-742.
- Rice WR, Chippindale AK. 2001. Intersexual ontogenetic conflict. *Journal of Evolutionary Biology* 14:685-693.
- Siller S. 2001. Sexual selection and the maintenance of sex. *Nature* 411:689-692.
- Snyder BF, Gowaty PA. 2007. A reappraisal of Bateman's classic study of intrasexual selection. *Evolution* 61:2457-2468.
- Trivers RL. 1972. Parental investment and sexual selection. Pages 136-179 in B. Campbell, ed. *Sexual selection and the descent of man, 1871-1971*. Aldine, Chicago.

## **Part I: Molecular Signatures of Adaptation on X chromosomes and Autosomes**

### **Chapter 2**

#### **Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses**

##### **ABSTRACT**

Patterns of sex chromosome and autosome evolution can be used to elucidate the underlying genetic basis of adaptive change. Evolutionary theory predicts that X-linked genes will adapt more rapidly than autosomes if adaptation is limited by the availability of beneficial mutations, and if such mutations are recessive. In *Drosophila*, rates of molecular divergence between species appear to be equivalent between autosomes and the X chromosome. However, molecular divergence contrasts are difficult to interpret because they reflect a composite of adaptive and nonadaptive substitutions between species. Predictions based on faster-X theory also assume that selection is equally effective on the X and autosomes; this might not be true because the effective population sizes of X-linked and autosomal genes systematically differ. Here, population genetic and divergence data from *Drosophila melanogaster*, *D. simulans* and *D. yakuba* are used to estimate the proportion of adaptive amino acid substitutions occurring in the *D. melanogaster* lineage. After gene composition and effective population size differences between chromosomes are controlled, X-linked and autosomal genes are shown to have equivalent rates of adaptive divergence, with approximately 30 % of amino acid substitutions driven by positive selection. The results suggest that adaptation is either unconstrained by a lack of beneficial genetic variation, or that beneficial mutations are

not recessive and are thus highly visible to natural selection whether on sex chromosomes or on autosomes.

## INTRODUCTION

The rate of adaptive evolution is expected to differ for X-linked and autosomal genes under several scenarios. If adaptation relies upon new mutations, X-linked genes will adapt faster than autosomes when beneficial mutations are (on average) recessive, and adapt slower than autosomes when beneficials are dominant (Charlesworth et al. 1987). In species with male-biased mutation rates, X-linkage decreases the beneficial mutation rate, thereby decreasing the opportunity for X-linked adaptive divergence (Kirkpatrick & Hall 2004). If adaptation proceeds by the fixation of standing genetic variation (i.e., formerly deleterious alleles), autosomal genes will adapt faster than X-linked genes because autosomes harbor a much larger pool of potential beneficial alleles (Orr & Betancourt 2001).

A growing number of studies have tested these scenarios by comparing the evolutionary rates of X-linked and autosomal genes (reviewed in Vicoso & Charlesworth 2006; Mank et al. 2007). To the extent that genetic divergence between species is driven by positive selection, ‘faster X’ evolution might indicate that adaptation proceeds by fixation of new, recessive ( $h < 0.5$ ) mutations, and consequently, that evolutionary divergence throughout most of the genome (consisting mostly of autosomes) is constrained by reduced expression of beneficial alleles. However, evolutionary divergence between species is caused by a combination of genetic drift and natural selection (Kimura 1968; Ohta 1973; Gillespie 1991), and the relative contribution of each process can potentially differ between the X and autosomes, thereby rendering divergence rate contrasts uninformative.

The strength of natural selection relative to genetic drift increases as the effective population size increases (selection is a function of  $N_e s$ , where  $N_e$  is the effective population size and  $s$  is the intensity of selection; Fisher 1930; Wright 1931; Kimura 1962). Low  $N_e$  limits the rate of adaptive substitution (Betancourt & Presgraves 2002; Presgraves 2005) and enhances the fixation rate of slightly deleterious mutations due to genetic drift (Haddrill et al. 2007). Genomic regions with different effective population

sizes might exhibit different rates of adaptive substitution, but these will often be obscured by correlated nearly-neutral substitution differences. This is likely to affect divergence rate contrasts between X-linked and autosomal genes, which experience different effective population sizes as a result of differing census size (i.e.,  $4/3 N$  autosome copies per  $N$  X chromosome copies), sexual selection (Charlesworth 2001), or interactions between linkage and directional selection (i.e., ‘Hill-Robertson interference’; Hill & Robertson 1966; Felsenstein 1974; Charlesworth 1994; Barton 1995; Betancourt et al. 2004).

By combining population genetic and divergence data, the individual contributions of genetic drift and positive selection to X chromosome and autosome evolution can be teased apart. To address whether the rate of adaptation is higher on the X chromosome, and to control for potential X-autosome differences in  $N_e$ , I analyzed coding sequence polymorphism and divergence in the *Drosophila melanogaster* lineage. Silent and amino acid replacement polymorphisms ( $P_s$  and  $P_n$ , respectively) and substitutions ( $D_s$  and  $D_n$ ) were calculated for 337 protein coding genes (80 X-linked), and were used to estimate alpha, the proportion of  $D_n$  that is caused by positive selection (McDonald & Kreitman 1991; Smith & Eyre-Walker 2002). Recombination rate and gene expression data were used to mitigate biases caused by  $N_e$  reductions via low recombination (i.e., enhanced Hill-Robertson interference), and gene composition differences between chromosomes.

## **METHODS**

### **Gene Samples**

Population genetic data for *Drosophila melanogaster* was obtained from Genbank (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=nucleotide>), the *Drosophila* Polymorphism Database (DPDB; <http://dpdb.uab.es/dpdb/dpdb.asp>), Peter Andolfatto’s webpage (<http://www-biology.ucsd.edu/labs/andolfatto/>), and Jody Hey’s webpage (<http://lifesci.rutgers.edu/~hey/lab/>). Orthologous sequences from *D. simulans* and *D. yakuba* were obtained via Genbank, when available. When multiple copies of a gene were available, one was chosen arbitrarily. For the remaining genes, *D. melanogaster* sequences were used to BLAST against the *D. simulans* and *D. yakuba* genome assembly (<http://flybase.org/blast/>). Of the 350 genes that had five or more *D.*

*melanogaster* sequences, both *D. simulans* and *D. yakuba* outgroup sequences were available for 337; these genes were used in all analyses.

The *D. melanogaster* lineage served as the focus of this study for several reasons:

- Extensive genomic data is available to control for gene-specific variation in gene expression and local recombination rates, both of which influence adaptive protein evolution in *Drosophila* (Presgraves 2005; Pröschel et al. 2006; Shapiro et al. 2007), and vary between species (Ranz et al. 2003; True et al. 1996);
- A large number of gene samples are from East African populations, which are probably closer to mutation-selection-drift equilibrium than cosmopolitan populations (David & Capy 1988; Begun & Aquadro 1993);
- Mutation rates are not sex-biased in this species (Bauer & Aquadro 1997), making different X and autosome mutation rates unlikely;
- Alpha (the proportion of  $Dn$  caused by positive selection) relies upon an assumption of neutrality for polymorphisms and silent substitutions ( $Pn$ ,  $Ps$ ,  $Ds$ ; Smith & Eyre-Walker 2002). While exclusion of low-frequency polymorphisms is usually considered suitable justification for the neutral polymorphism assumption, it has been suggested that  $Ds$  is subject to stronger constraint on the X relative to the autosomes (Comeron et al. 1999; Singh et al. 2005), which will artificially enhance the signature of X-linked positive selection (by elevating  $Dn/Ds$ ). Synonymous site selection appears to be absent in the *D. melanogaster* lineage (in contrast to *D. simulans*; e.g., Mcvean & Viera 1999; Akashi et al. 2006), which suggests that the neutrality assumption is most reasonable here (this assumption is addressed in full, below).

### **Expression & Recombination Data**

Expression data from Stolc et al. (2004) were obtained from Kevin White's webpage (<http://genome.med.yale.edu/lab/discuss.htm>). For genes with data from multiple distinct probes and multiple replicates of the same probe(s), the average of each probe was calculated before averaging across probes. Three expression variables were obtained from the data: (1) the average expression across all life stages and between the sexes; (2) the expression ratio between juvenile and adult life stages, and (3) the expression ratio between adult males and females. Average expression levels do not differ between X-

linked and autosomal genes analyzed in this study. However, the frequency of male-biased and adult-biased genes is significantly lower on the X chromosome (Table 2.1). A previous study indicates that male-biased expression substantially elevates the signature of positive selection in *Drosophila* (Pröschel et al. 2006). Results controlling for X and autosome male-biased and adult-biased gene content are reported.

Five different recombination estimates per gene were obtained from Jody Hey's webpage (<http://lifesci.rutgers.edu/~hey/lab/>; described in Hey & Kliman 2002). Most of the 337 genes analyzed here were included in the dataset. When unavailable, data from the nearest gene was used (as defined by [www.flybase.org](http://www.flybase.org)). This procedure is consistent with the methodologies used to generate the recombination estimates of genes included in Hey & Kliman (2002). X-linked estimates were multiplied by 4/3 to control for the lack of recombination in male *Drosophila* and the biased transmission of X-linked genes towards females; this 4/3 correction assumes an equal sex ratio among parents, which is necessarily true. Results reported here use the KH93 estimator, though results with different estimators are equivalent. "Low recombination" regions were classified as having a crossover frequency less than  $0.002 \times 10^{-5}$  per base pair, per generation. This cutoff was used because genes in higher regions of recombination showed similar patterns of silent nucleotide variability between the X and autosomes, and because previous studies indicate that this is approximately the region in which signatures of positive selection disappear (Presgraves 2005; Shapiro et al. 2007). Use of a higher cutoff produces the same general result.

### **Statistical Analyses**

Genes were aligned with ClustalX online (<http://www.ch.embnet.org/software/ClustalW-XXL.html>) and adjusted by hand. McDonald-Kreitman and Tajima's D statistics were calculated with DnaSP, Version 4.10 (Rozas et al. 2003). Watterson's estimate of silent nucleotide diversity ( $\theta$ ) was calculated by hand (as described in Hein et al. 2005), and used to estimate genic effective population size, or  $N_e$ . Two genes (*dpp* & *CG6495*) were polymorphic for inversions; standard alignments (i.e., noninversions) were used for all analyses. Two genes (*SR-CIII*

& *SR-CIV*) included nonfunctional gene copies, which were removed from all analyses. Both exclusions do not affect the results.

$D_n$  and  $D_s$  values (nonsynonymous and synonymous substitutions between species) were calculated using an arbitrarily selected sequence from *D. melanogaster*, and orthologs from *D. simulans* and *D. yakuba*. This procedure permits divergence estimates to be independent of polymorphism (a concern if X and autosome polymorphism systematically differs). The number of substitutions in the *D. melanogaster* lineage for each gene was estimated with the equation,  $A - (A + B - C) / 2$ , where  $A$  is the number of substitutions separating *D. melanogaster* and *D. yakuba*,  $B$  is the number between *D. simulans* and *D. yakuba*, and  $C$  is the number between *D. simulans* and *D. melanogaster*.

Differences between the X and autosomes for  $D_n/D_s$  or  $P_n/P_s$  ratios were determined by G-tests, calculated with DnaSP. The proportion of substitutions between species that were fixed via positive selection (or alpha, using the method of Bierne & Eyre-Walker 2004) was estimated with the software package DoFe (using default settings), kindly provided by Adam Eyre-Walker (described in Eyre-Walker 2006). DoFe uses a maximum likelihood approach that maximizes the number of genes that can be analyzed (i.e., even those with very little polymorphism), and does not sum  $D_n$ ,  $D_s$ ,  $P_n$ , and  $P_s$  values across genes and therefore avoids a bias described in Shapiro et al. (2007)

*D. melanogaster* and *D. simulans* are separated by two X-linked and three autosomal inversion differences (one on chromosome 2R; two on chromosome 3R; Lemeunier & Ashburner 1976) – in which 2 X-linked and 29 autosomal genes from this dataset reside. All analyses involving interspecific divergence data were performed with and without these genes. The reported results include genes within the inversions, but the patterns and conclusions are the same whether these are included or excluded. It should also be noted that *D. melanogaster* autosomes harbor a greater number of inversion polymorphisms than X chromosomes do (Aulard et al. 2002), which could differentially depress  $N_e$  on autosomes. While this X-autosome difference was not explicitly controlled, it should (if anything) bias the dataset towards detecting a pattern of fast-X adaptive evolution and make the results reported below conservative. Furthermore, patterns of silent nucleotide variability (reported below and reflecting  $N_e$ ) indicate that this does not cause a

systematic bias between the chromosomes, after recombinational differences are controlled.

## RESULTS AND DISCUSSION

### Divergence and Polymorphism

For total divergence between *D. melanogaster* and *D. simulans*, the  $D_n/D_s$  ratio is the same for the X and autosomes (Table 2.2), consistent with previous studies (Betancourt et al. 2002; Thornton et al. 2006). However,  $D_s$  is strongly reduced, whereas  $D_n$  is the same, in the *D. simulans* relative to the *D. melanogaster* lineage ( $D_s \chi^2 = 96.93$ ;  $P < 0.00001$ ;  $D_n \chi^2 = 0.554$ ;  $P = 0.457$ ), indicating relaxed selection on synonymous sites in the *D. melanogaster* lineage (see Mcvean & Viera 1999; Akashi et al. 2006). This pattern is stronger for X-linked compared to autosomal genes ( $\chi^2 = 14.397$ ;  $P = 0.00015$ ), which consequently reduces the  $D_n/D_s$  ratio on the X relative to the autosomes in the *D. melanogaster* lineage (Table 2.2).

X-linked genes exhibit lower  $P_n/P_s$  ratios than autosomal genes (Table 2.2; see Begun 1996; Andolfatto 2001), and this difference remains after male-biased genes, which are substantially more abundant on autosomes (Table 2.1; Parisi et al. 2003), are removed. This pattern is consistent with two hypotheses: that selection is more efficient at removing partially recessive deleterious mutations on the X relative to the autosomes (e.g., Haldane 1935), and that a larger fraction of nucleotides are evolving neutrally on the autosomes relative to the X. These hypotheses can be distinguished by removing low frequency (e.g., singleton) polymorphisms, which are much more likely to be strongly deleterious compared to those at high frequencies. X-autosome differences in  $P_n/P_s$  remain after singleton polymorphisms are removed (Table 2.2), suggesting enhanced purifying selection against weakly deleterious mutations on the X chromosome (i.e., less neutrality on the X).

### Silent Nucleotide Variation, Recombination, and $N_e$

Because X-linked loci, on average, spend two-thirds of their evolutionary history in female genomes, and *Drosophila* males do not recombine (Morgan 1912), recombinational discrepancies are predicted to arise between the X and autosomes (Table

2.3; note that the relative rates of recombination for X and autosomal genes analyzed here do not differ from the genome-wide pattern). Hill-Robertson interference is enhanced and effective population size ( $N_e$ ) decreases in regions of reduced recombination (Hill & Robertson 1966; Felsenstein 1974). This limits the effectiveness of natural selection on weakly selected genetic variation; recombinational differences between the X and autosomes must be controlled to detect differences between X and autosome adaptive divergence that are attributable to hemizygous expression.

Patterns of synonymous nucleotide variability ( $\theta = 4N_e\mu$ , where  $\mu$  is the mutation rate per nucleotide) can be used to study genic effective population size ( $N_e$ ) as a function of recombination rate variation (e.g., Begun & Aquadro 1992; Presgraves 2005). Silent variation is positively correlated with recombination for the X and autosomes ( $r_X = 0.481$ ;  $r_{aut} = 0.440$ ;  $P < 0.0001$  for both correlations), and is particularly reduced in regions of low recombination (Figs. 2.1-2.2). Silent variability is significantly higher for X-linked genes ( $P < 0.05$ ; two-tailed t test), but this discrepancy disappears when low-recombining genes are removed (e.g., less than  $2 \times 10^{-8}$  /bp /generation; Fig. 2.1), indicating that effective population size does not differ between the X and autosomes for genes in moderate to high regions of recombination.

### **The Proportion of Positively Selected Substitutions**

Polymorphism and divergence data were used to estimate the proportion of substitutions driven by positive selection (i.e., alpha; Fig. 2.3). When all genes and polymorphisms are included, alpha is estimated at 0.18 for X-linked and 0.00 for autosomal genes (X-linked alpha is not significantly greater than zero at  $P < 0.05$ ). However, three systematic biases creep into contrasts between the X and autosomes (as discussed above). Autosomes are enriched in genes with male-biased expression (Table 2.1), which will upwardly bias autosomal estimates of alpha. Autosomes also carry a higher proportion of genes evolving in low recombinational environments (Table 2.3; Fig. 2.1), and are predicted to harbor relatively higher numbers of deleterious mutations than the X chromosome (Haldane 1935); these two factors will tend to downwardly bias estimates of alpha (Bierne & Eyre-Walker 2004; Presgraves 2005).

To eliminate these potential sources of bias, genes inhabiting low-recombination genomic regions (i.e., rec. rate  $< 0.002 \times 10^{-5}$  /bp/gen; Figs. 2.1-2.2), and genes with male-biased expression, were removed from the analysis. Singleton polymorphisms, which are likely to include deleterious alleles (Bierne & Eyre-Walker 2004), were also removed. Estimated alpha values for the trimmed dataset are quite similar between the X and autosomes (approximately 30 percent), with heavily overlapping confidence intervals (Figs. 2.3-2.4). There is no discernible “faster-X effect” for adaptive substitutions.

Previous studies report higher codon usage bias on the X relative to the autosomes (Comeron et al. 1999; Singh et al. 2005). If X-linked selection on synonymous sites is relatively strong, it may downwardly bias  $P_s$  and decrease, and/or downwardly bias  $D_s$  and increase estimates of positive selection. Current selection is expected to remove synonymous polymorphism as well as skew the distribution of variation towards rare alleles (e.g., singletons). The Tajima’s  $D$  statistic reflects the polymorphic frequency distribution for a gene and decreases as the proportion of rare alleles increases (i.e., becomes negative as rare alleles make up a larger fraction of variation; Hein et al. 2005). In genes sampled from African populations (which are more likely to reflect selection rather than demographic factors; David & Cappy 1988; Begun & Aquadro 1993),  $D$  is negative for silent sites, but does not differ between the X and autosomes ( $D_X = -0.236$ ;  $D_{aut} = -0.215$ ;  $P \gg 0.05$ ; two-tailed ttest), as expected if silent sites are equally neutral on the X and autosomes. Detecting long-term selection on synonymous sites is more problematic since the exact mutation rates of the X and autosomes are not known. However, the rate of silent site divergence in the *D. melanogaster* lineage is not lower on the X ( $dS_X = 0.072$ ;  $dS_{aut} = 0.061$ ; see also Begun & Whitley 2000), as would be expected if X-linked silent sites evolve under greater constraint. Even if  $D_s$  is downwardly biased on the X, the major conclusion reported here – that X-linkage does not enhance adaptative divergence – will be conservative.

## CONCLUSION

In the *Drosophila melanogaster* lineage, an equivalent proportion of adaptive nonsynonymous substitutions occur for X-linked compared to autosomal genes. It is possible that a pattern of faster-X evolution might be obscured by unidentified random or

systematic differences between X-linked and autosomal genes, though this is not particularly likely, as several major factors known to influence protein adaptation were controlled (see above). Given that the rate of total (i.e., adaptive plus neutral substitutions) divergence is the same on the X and autosomes (Betancourt et al. 2002; Thornton et al. 2006; but see Counterman et al. 2004), X-linkage does not appear to elevate the rate of protein adaptation in *Drosophila*.

Two models of adaptation are consistent with the patterns reported here. If adaptive divergence proceeds by the fixation of new beneficial mutations, and X-linked population size is 75 percent of the autosome population size (i.e.,  $3/4 N_A = N_X$ ), the dominance of beneficial alleles can be estimated as  $h \approx \frac{R_X}{2(2 - R_X)}$ , where  $R_X$  is the ratio of the autosome to X-linked adaptive substitution rate (Charlesworth et al. 1987; equation (2a)). The ratio of the maximum likelihood estimates of autosome and X-linked alpha, which can be used to approximate  $R_X$ , is 0.817 (excluding singleton polymorphism, low recombination and 1.5-fold and above male-biased genes), and suggests that  $h \approx 0.345$ . However, in *D. melanogaster*, the X-linked effective population size is similar to autosomal  $N_e$  (Figs. 2.1-2.2; Andolfatto 2001). The number of new beneficial mutations arising each generation will therefore be proportionally similar between the X and autosomes. Assuming equal X and autosome population size,  $h \approx \frac{R_X}{3 - 2R_X} \approx 0.598$ . Confidence intervals surrounding estimates of alpha are fairly broad, which reduces the precision of the dominance coefficient estimate. Nevertheless, it is clear, assuming a model of adaptation using new mutations, that beneficial mutations would be roughly additive in expression.

A second possibility is that multiple sources of variation – new mutations and standing genetic variation – jointly contribute to adaptive substitutions between species. Adaptation using previously deleterious genetic variation is expected to cause a pattern of “faster-autosome” evolution, independent of the dominance of beneficial alleles (Orr & Betancourt 2001). If both new mutations and previously deleterious genetic variation contribute to adaptive divergence, and the fitness effects of beneficial alleles are recessive, then X-linked and autosomal genes should evolve at similar rates. Teasing

apart these alternative possibilities will be an interesting challenge for future research on the genetic basis of adaptation.

Table 2.1. X chromosome and autosome gene composition

		Gene Expression Category									
		Total	Juvenile-biased		Adult-biased		Female-biased		Male-biased		
	cutoff <sup>(1)</sup>	--	1.5- fold	2-fold	1.5- fold	2-fold	1.5- fold	2-fold	1.5- fold	2-fold	
X	# genes	80	5	0	4	2	23	11	4	2	
	relative freq. <sup>(2)</sup>	1	0.063	0.000	0.050	0.025	0.288	0.138	0.050	0.025	
Autosomes	# genes	257	21	16	61	34	74	43	58	45	
	relative freq. <sup>(2)</sup>	1	0.082	0.062	0.237	0.132	0.288	0.167	0.226	0.175	
	<i>P</i> <sup>(3)</sup>	--	0.601	0.027	0.001	0.012	0.996	0.587	0.002	0.002	

1. Cutoff refers to a minimum expression ratio for each category; e.g. female-biased genes with a 2-fold cutoff have at least twice the expression (in terms of mRNA abundance) in females as in males.
2. Frequency of total genes for each chromosome type, i.e. total X-linked or total autosomal genes.
3. Based on chi-square values for 2 x 2 contingency tables.

Table 2.2. Divergence and polymorphism data for X-linked and autosomal genes.<sup>(1)</sup>

Comparison		Location	$D_n$	$D_s$	$D_n/D_s$	$P^{(2)}$
<i>D. melanogaster</i> lineage		X	475	1119	0.424	0.145
		autosomes	1412	3033	0.466	
<i>D. mel.</i> and <i>D. sim.</i> Lineages		X	955	1882	0.507	0.774
		autosomes	2865	5572	0.514	
Populations		Location	$P_n$	$P_s$	$P_n/P_s$	$P$
Worldwide	with singletons	X	268	997	0.269	< 0.0001
		autosomes	1277	2838	0.450	
	no singletons	X	123	615	0.200	< 0.0001
		autosomes	631	1758	0.359	
African	with singletons	X	191	797	0.240	< 0.0001
		autosomes	702	1800	0.390	
	no singletons	X	98	396	0.247	0.08

autosomes	322	1040	0.310
-----------	-----	------	-------

---

<sup>1</sup> Results from the entire set of genes are shown: worldwide autosomes,  $n = 257$ ; African autosomes,  $n = 148$ ; worldwide X-linked,  $n = 80$ ; African X-linked,  $n = 57$ . Equivalent results were obtained when genes were partitioned into gene expression categories.

<sup>2</sup> G test (d.f. = 1).

Table 2.3. Average recombination rates for X-linked and autosomal genes<sup>(1,2)</sup>.

		Estimators of Recombination				
		KH93	ACE	R_TE	R	R_P
Autosomes	Genomic	0.0026	0.0026	0.0023	0.0023	0.0024
	All gene samples	0.0027	0.0027	0.0025	0.0024	0.0025
	African samples	0.0029	0.0029	0.0027	0.0025	0.0026
X Chromosome	Genomic	0.0040	0.0056	0.0040	0.0040	0.0041
	All gene samples	0.0039	0.0057	0.0041	0.0042	0.0042
	African samples	0.0043	0.0063	0.0047	0.0050	0.0049
A:X ratio	Genomic	0.6505	0.4596	0.5800	0.5755	0.5768
	All gene samples	0.6867	0.4742	0.6075	0.5674	0.5849
	African samples	0.6617	0.4560	0.5807	0.5125	0.5338

<sup>1</sup> The five estimates of recombination are described in Hey & Kliman (2002) and were obtained from Jody Hey's website (<http://lifesci.rutgers.edu/%7Eheylab/>). X chromosome rates were adjusted by multiplying by 4/3 (see methods). Results reported in units x 10<sup>-5</sup>/bp /gen.

---

<sup>2</sup> Genomic sample size: X-linked genes = 2242, autosomes = 10757; all genes sampled in this study: X-linked = 80, autosomes = 257; African derived samples: X-linked = 57, autosomes = 148.

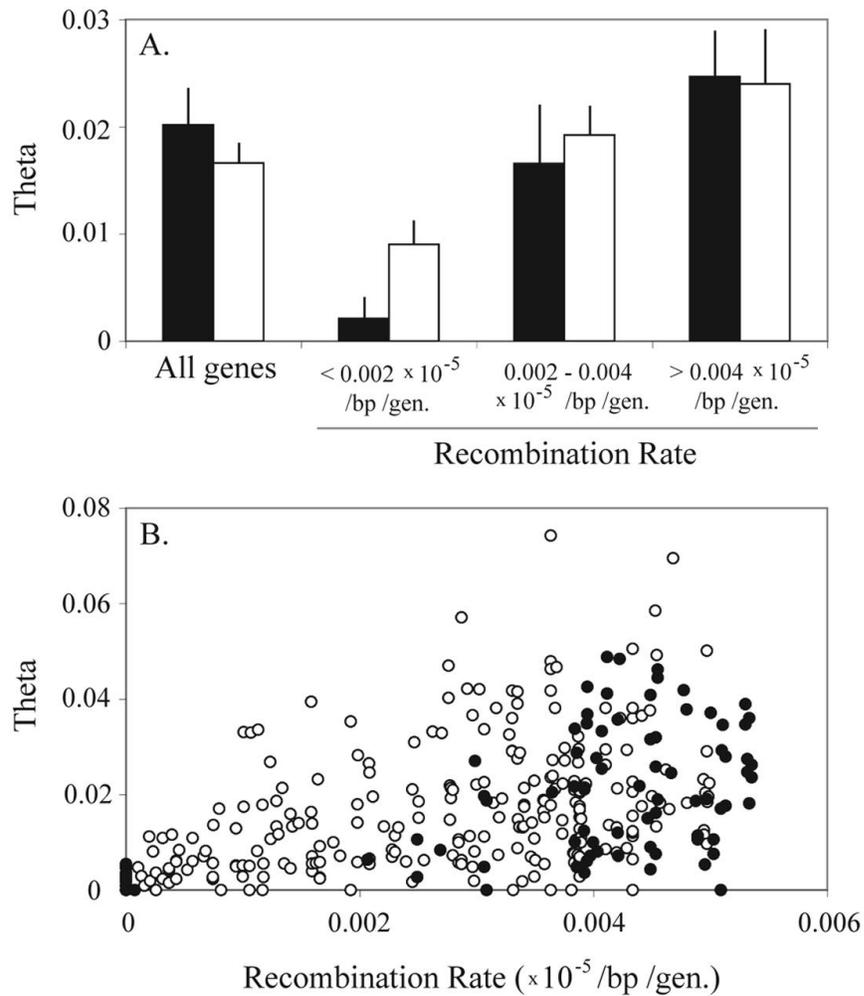


Figure 2.1. Patterns of silent nucleotide diversity (theta per nucleotide) for X-linked and autosomal genes as a function of recombination rate. The X and autosomes are represented by black and white bars, respectively. (A) Means and 95 percent confidence intervals are shown for theta across three intervals of recombination. (B) Theta across a continuous gradient of recombination. X-linked and autosomal data are represented by black and white bars (or dots), respectively. Theta values include singleton polymorphism. All X-autosome contrasts do not differ statistically, with the exception of theta using all genes, which is higher on the X, and theta at low recombination, which is larger on autosomes ( $P < 0.05$ ; two-tailed t-test). Data are from both African and cosmopolitan *D. melanogaster* population samples.

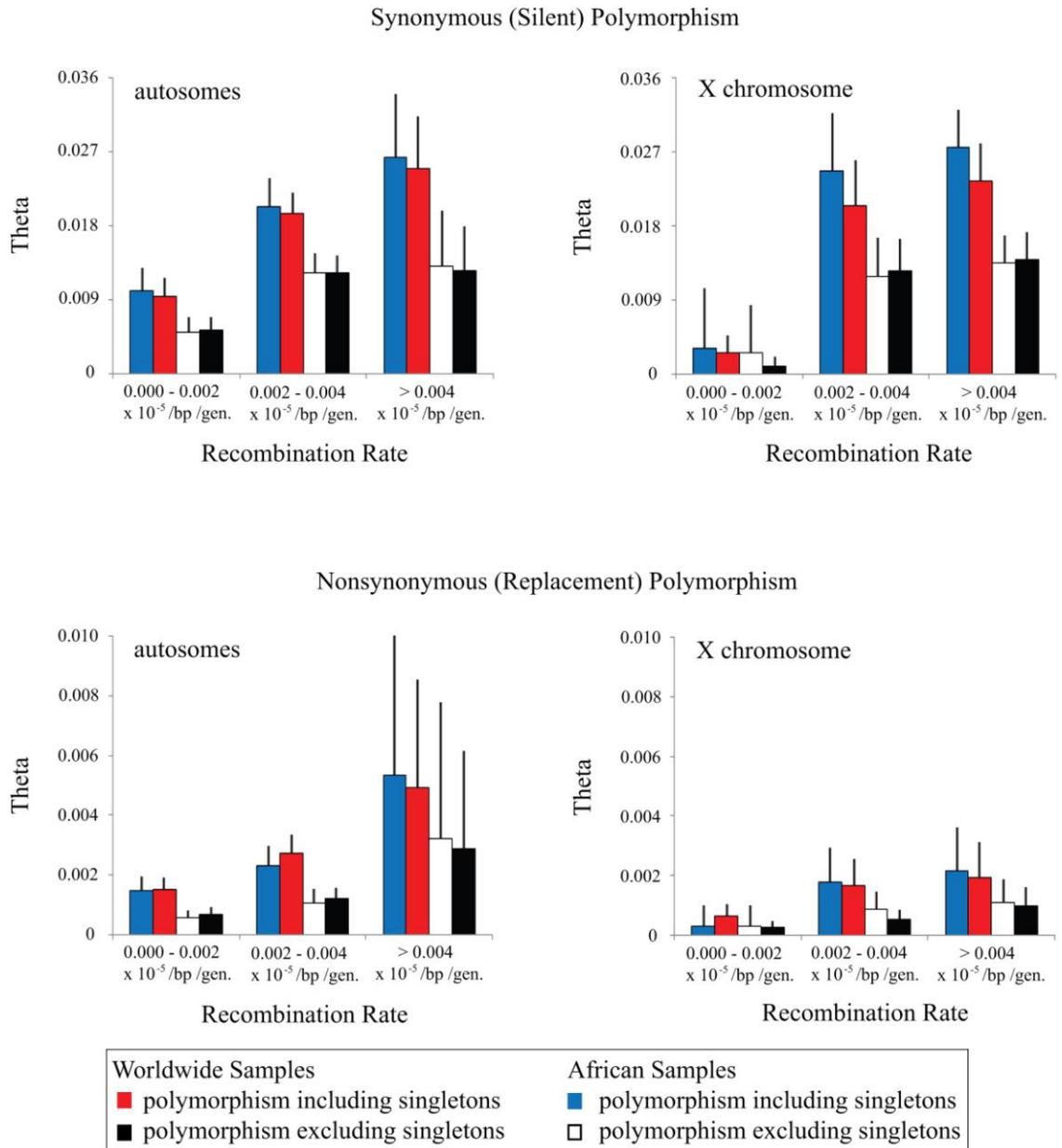


Figure 2.2. A comparison of silent and replacement polymorphism (theta per nucleotide) from the entire dataset (Worldwide) and African-derived samples. Means and 95 percent confidence intervals are shown for theta across three intervals of recombination. All contrasts between populations do not differ statistically ( $P > 0.05$ ; two-tailed t-test, no correction for multiple comparisons).

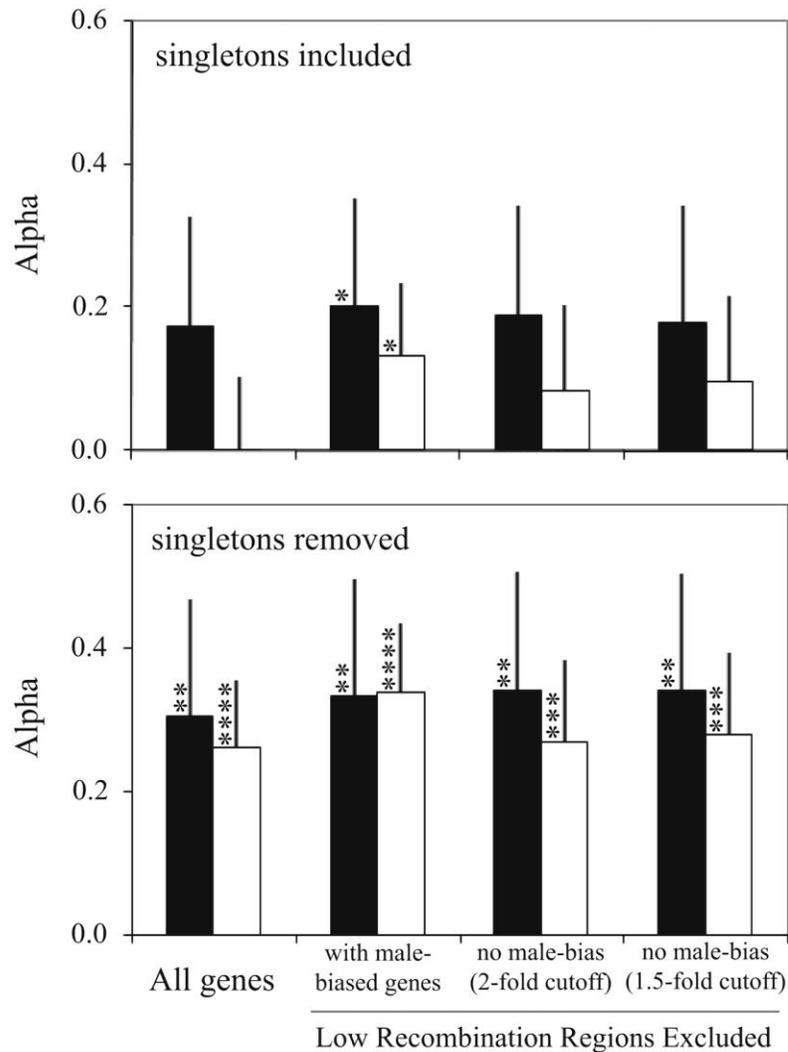


Figure 2.3. Alpha – the proportion of adaptive, nonsynonymous substitutions – for X-linked and autosomal genes. Genes are partitioned based on sex-specific gene expression profiles. The X and autosomes are represented by black and white bars, respectively. Substitutions included in the analysis are those that occurred within the *D. melanogaster* lineage; all polymorphism data is included (worldwide samples). Means and 95 percent confidence intervals are shown. \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$ ; significance based on a likelihood ratio test of the null hypotheses,  $\alpha = 0$  (implemented with the DoFe software).

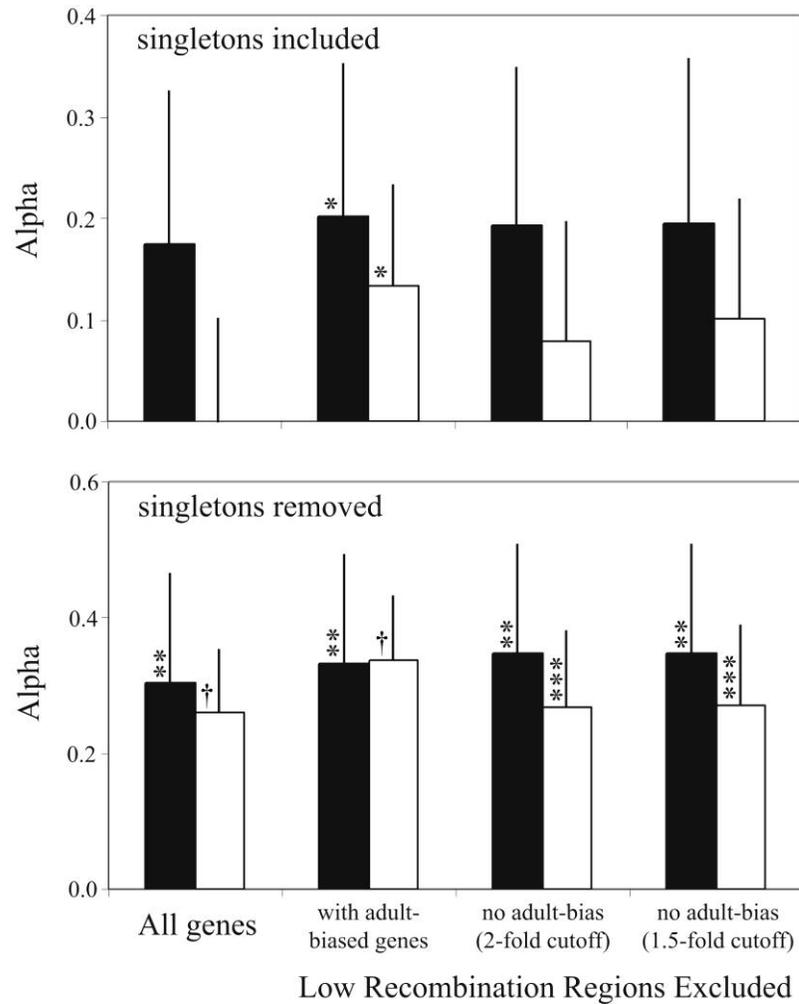


Figure 2.4. Alpha – the proportion of adaptive, nonsynonymous substitutions – for X-linked and autosomal genes. Genes are partitioned based on age-specific gene expression profiles. The X and autosomes are represented by black and white bars, respectively. Substitutions included in the analysis are those that occurred within the *D. melanogaster* lineage; polymorphisms are for African and non-African gene samples. Means and 95 percent confidence intervals are shown. \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; †  $P < 0.0001$ ; significance based on a likelihood ratio test of the null hypotheses,  $\alpha = 0$  (implemented with the DoFe software).

## REFERENCES

- Akashi H, Ko WY, Piao S, John A, Goel P, Lin CF, Vitins AP. 2006. Molecular evolution in the *Drosophila melanogaster* species subgroup: frequent parameter fluctuations on the timescale of molecular divergence. *Genetics* 172:1711-1726.
- Andolfatto P. 2001. Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Mol Biol Evol.* 18:279-290.
- Aulard S, David JR, Lemeunier F. 2002. Chromosomal inversion polymorphism in Afrotropical populations of *Drosophila melanogaster*. *Genet Res Camb.* 79:49-63.
- Barton NH. 1995. Linkage and the limits to natural selection. *Genetics.* 140:821-841.
- Bauer VL, Aquadro CF. 1997. Rates of DNA sequence evolution are not sex-biased in *Drosophila melanogaster* and *D. simulans*. *Mol Biol. Evol.* 14:1252-1257.
- Begun DJ. 1996. Population genetics of silent and replacement variation in *Drosophila simulans* and *D. melanogaster*: X/autosome differences? *Mol Biol Evol.* 13:1405-1407.
- Begun DJ, Aquadro, CF. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356:519-520.
- Begun DJ, Aquadro, CF. 1993. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature.* 365:548-550.
- Begun DJ, Whitley P. 2000. Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. *Proc Natl Acad Sci USA* 97:5960-5965.
- Betancourt AJ, Kim Y, Orr HA. 2004. A pseudohitchhiking model of X vs. autosomal diversity. *Genetics.* 168:2261-2269.
- Betancourt AJ, Presgraves, DC. 2002. Linkage limits the power of natural selection in *Drosophila*. *Proc Natl Acad Sci USA* 99:13616-13620.
- Betancourt AJ, Presgraves DC, Swanson WJ. 2002. A test for faster X evolution in *Drosophila*. *Mol Biol Evol.* 19:1816-1819.
- Bierne & Eyre-Walker A. 2004. The genomic rate of adaptive amino acid substitution in *Drosophila*. *Mol Biol Evol.* 21:1350-1360.
- Charlesworth B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet Res.* 63:213–228.

- Charlesworth B, Coyne JA, Barton N. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat.* 130:113-146.
- Charlesworth B. 2001. The effect of life-history and mode of inheritance on neutral genetic variability. *Genet Res Camb.* 77:153-166.
- Comeron JM, Kreitman M, Aquade M. 1999. Natural selection on synonymous sites is correlated with gene length and recombination in *Drosophila*. *Genetics* 151:239-249.
- Counterman BA, Ortiz-Barrientos D, Noor MA. 2004. Using comparative genomic data to test for faster-X evolution. *Evolution.* 58:656–660.
- David JR, Capy P. 1988. Genetic variation of *Drosophila melanogaster* natural populations. *Trends Genet.* 4:106-111.
- Eyre-Walker A. 2006. The genomic rate of adaptive evolution. *Trends Ecol Evol.* 21:569-575.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics.* 78:737-756.
- Fisher RA. 1930. The distribution of gene ratios for rare mutations. *Proc Roy Soc Edinburgh.* 50:204-219.
- Gillespie JH. 1991. *The Causes of Molecular Evolution.* Oxford University Press, Oxford.
- Haddrill PR, Halligan DL, Tomaras D, Charlesworth B. 2007. Reduced efficacy of selection in regions of the *Drosophila* genome that lack crossing over. *Genome Biology* 8:R18.
- Hein J, Schierup MH, Wiuf C. 2005. *Gene Genealogies, Variation and Evolution: A Primer in Coalescent Theory.* Oxford University Press, New York.
- Hey J, Kliman RM. 2002. Interactions between natural selection, recombination and gene density in the genes of *Drosophila*. *Genetics.* 160:595-608.
- Haldane JBS. 1935. The rate of spontaneous mutation of a human gene. *J Genet.* 31:317-326.
- Hill WG, Robertson A. 1966. The effect of linkage on the limits of artificial selection. *Genet Res.* 8:269-294.
- Kimura M. 1962. On the probability of fixation of mutant genes in a population. *Genetics.* 47:713-719.
- Kimura M. 1968. Evolutionary rate at the molecular level. *Nature.* 217:624-626.

- Kirkpatrick M, Hall DW. 2004. Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution*. 58:437-440.
- Lemeunier F, Ashburner M. 1976. Relationships within the *Melanogaster* species subgroup of the genus *Drosophila* (Sophophora). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. *Proc Roy Soc Lond B* 193:275-294.
- Mank JE, Axelsson E, Ellegren H. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Res*. 17:618-624.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*. 351:652-654.
- McVean GAT, Vieira J. 2001. Inferring parameters of mutation, selection and demography from patterns of synonymous site evolution in *Drosophila*. *Genetics* 157:245-257.
- Morgan T. 1912. Complete linkage in the second chromosome of the male of *Drosophila*. *Science*. 36:719-720.
- Ohta T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature*. 246:96-98.
- Orr HA, Betancourt AJ. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics*. 157:875-884.
- Parisi M, et al. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science*. 299:697-700.
- Presgraves DC. 2005. Recombination enhances protein adaptation in *Drosophila melanogaster*. *Curr Biol*. 15:1651-1656.
- Pröschel M, Zhang Z, Parsch J. 2006. Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. *Genetics*. 174:893-900.
- Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* 300:1742-1745.
- Rozas J, et al. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. 19:2496-2497.
- Shapiro JA, et al. 2007. Adaptive genic evolution in the *Drosophila* genomes. *Proc Nat Acad Sci USA*. 104:2271-2276.

- Singh ND, Davis JC, Petrov DA. 2005. X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis*. *Genetics*. 171:145-155.
- Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in *Drosophila*. *Nature* 415:1022-1024.
- Stolc V, et al. 2004. A gene expression map for the euchromatic genome of *Drosophila melanogaster*. *Science*. 306:655-660.
- Thornton K, Bachtrog D, Andolfatto P. 2006. X chromosomes and autosomes evolve at similar rates in *Drosophila*: no evidence for faster-X protein evolution. *Genome Res.* 16:498-504.
- True JR, Mercer JM, Laurie CC. 1996. Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* 142:507-523.
- Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. *Nat Rev Genet.* 7:645-653.
- Wright S. 1931. Evolution in Mendelian populations. *Genetics*. 16:97-159.

## **Part II: Genome Architecture and the Heritability of Fitness**

### **Chapter 3**

#### **Sex Linkage, Genic Capture, and the Heritability of Fitness**

##### **ABSTRACT**

Several sexual selection models predict that females will obtain indirect genetic benefits by preferentially mating with males that improve the genetic quality and fitness of their offspring. However, despite widespread observations of additive population genetic variation for fitness as well as for male sexually selected traits, estimated fitness associations between fathers and offspring are often weak. Perhaps more puzzling, the strength of these associations differs drastically between species, leading many researchers to question the relevance of genetic benefits for processes of sexual selection. Here, I show that the sex chromosome system of a species can strongly influence the genetic architecture of male and female fitness variation, and consequently, the heritability of components of fitness between fathers and their offspring. Using models of sex-specific genetic variation, I show that indirect genetic benefits are much less likely to be detected in species with X chromosomes, relative to species with undifferentiated sex chromosomes, environmental sex determination, or Z chromosomes. Empirical observations from the sexual selection literature are consistent with predictions of the models, though additional studies will be required to circumvent phylogenetic non-independence between sex determination systems. This study strongly suggests that inferences about genetic benefits of female choice must be considered within a species-specific genomic context, and has several implications for the evolution of female preferences as well as the genomic consequences of sex and sexual selection.

## INTRODUCTION

The question of whether female mating preferences lead to indirect genetic benefits for offspring has major population genetic implications. By preferentially mating with a nonrandom subset of the male population, females can potentially reinforce natural selection by favoring males carrying genetic variation that enhances non-reproductive components of fitness. Such choice for “good genes” can improve the fitness of her offspring (Fisher 1915; Williams 1966; Trivers 1972; Zahavi 1975) and enhance adaptation in sexually reproducing populations (Kodric-Brown & Brown 1987; Whitlock 2000; Agrawal 2001; Siller 2001; Candolin & Heuschele 2008; Whitlock & Agrawal 2009). Empirical observations of indirect genetic benefits might also help to resolve the “lek paradox”, which predicts that strong female choice will erode population genetic variation for fitness, and eliminate genetic benefits as a potential explanation for the evolutionary maintenance of female preference (Borgia 1979; Andersson 1994; Rowe & Houle 1996; Kotiaho et al. 2008). Finally, indirect benefits are a reflection of the evolutionary forces maintaining population genetic variation for fitness. Therefore, the presence of indirect genetic benefits or costs may indicate the specific evolutionary forces that maintain fitness variation within natural populations. This is perhaps the major unresolved question in evolutionary biology (Lewontin 1974; Lewontin et al. 1981; Gillespie 2004).

When preferred male traits convey information about genetic variation for fitness (Zahavi 1975, 1977; Berglund et al. 1996), offspring of these males can potentially inherit genes that increase survival, fecundity, or enhance male mating success. The Genic Capture Model (Rowe & Houle 1996) offers an elegant rationale for why sexually selected traits should convey information about genetic quality. If sexually selected traits are condition-dependent and individual condition is a property of genome-wide fitness, then these traits will indirectly “capture” fitness variability, and male trait expression will be positively correlated with male genome quality. Females favoring males with well-developed ornaments or weapons can potentially improve offspring fitness by obtaining high-quality genes from their mates.

The basic requirements of the Genic Capture Model – condition dependence of sexually selected traits, and high additive genetic variation for condition and for sexually

selected traits – have strong empirical support (Houle 1992; Pomiankowski & Møller 1995; Jennions et al. 2001), yet there is surprisingly little consensus as to whether preferred males generally improve offspring fitness (Kokko et al. 2003; Kotiaho & Puurtinen 2007). Current evidence is inconsistent across study systems, with some species yielding strong evidence for genetic benefits for offspring (e.g., Petrie 1994; Forsman & Hagman 2006), and others showing sex-specific fitness benefits (e.g., Jones et al. 1998; Wedell & Tregenza 1999; Fedorka & Mousseau 2004), sex-specific costs (Pischedda & Chippindale 2006; Oneal et al. 2007; Delcourt et al. 2009; Connallon & Jakubowski 2009), or no fitness effects at all (e.g., Boake 1985). The fact that fitness heritability between fathers and offspring differs between study systems is puzzling, and suggests that the nature of fitness variation might fundamentally differ between populations or species. While this is an intriguing possibility, few studies have explicitly analyzed how species-specific attributes might account for empirical patterns of indirect fitness effects between species.

Sex chromosome systems represent a species-specific genomic feature that can potentially influence fitness heritability between fathers and offspring. First, sex-linked genetic variation differentially contributes to male and female fitness variation (James 1973; Cowley & Atchley 1988; Long & Rice 2007; Wayne et al. 2007). In species with heteromorphic sex chromosomes (where the Y or W chromosome is degenerate), X- and Z-linked variation is completely expressed in the heterogametic (haploid or hemizygous) sex, and can strongly contribute to fitness variance among individuals. Sex-linked variation is expected to have a lower impact on fitness variance within the homogametic (diploid) sex because sex-linked alleles are only partially expressed due to heterozygosity. Sex-linked genes also follow unique patterns of maternal and paternal inheritance between species (White 1973; Bull 1983). In species with X chromosomes, females inherit X-linked variation from both parents, while male X chromosomes are entirely maternally derived. This situation is completely reversed in species with Z chromosomes, where male Z chromosomes are biparentally inherited and the female Z is paternally inherited. Finally, sex chromosomes and autosomes can favor the accumulation of different types of fitness variation. “Good genes” variation (e.g., variation maintained at mutation-selection balance; Haldane 1937) is expected to be

relatively abundant on autosomes, whereas “sexually antagonistic” variation – where alleles favored by one sex are deleterious when expressed in the other – is expected to have a higher relative abundance on sex chromosomes (Rice 1984; Patten & Haig 2009).

Sex determination systems, and the proportion of the genome that is sex-linked, are highly variable between taxonomic groups (reviewed in White 1973; Bull 1983; Solari 1994; Devlin & Nagahama 2002). This variability might consequently generate species-specific genetic architecture for fitness and species-specific genetic benefits and costs to offspring. Though these ideas are implicit in previous theoretical and empirical studies (e.g., Cowley & Atchley 1988; Long & Rice 2007; Wayne et al. 2007), no prior models consider total, genome-wide fitness heritability as a function of the sex chromosome system. It is therefore unclear how strongly sex chromosomes might influence indirect fitness consequences of nonrandom mating.

Here, I analyze models of fitness heritability between fathers and offspring of each sex, in species with different sex chromosome systems (i.e., X-linkage, Z-linkage, undifferentiated or no sex chromosomes), and with arbitrary genomic proportions of sex-linked inheritance (i.e., zero to 100 percent). I specifically consider indirect fitness benefits and costs under models of “sexually concordant” genetic variation, where alleles good for males are also good for females, and “sexually antagonistic” variation, where alleles good for males are deleterious when present in female genomes. Published empirical estimates of genetic benefits and costs are reconsidered within this new theoretical framework. The results have several implications for female choice as well as the population genetic consequences of sexual selection.

## **MODEL**

### **Models of Fitness Variation**

Theoretical fitness associations between fathers and offspring are considered under two empirically well-supported models of genetic variation for fitness. Under a model of sexually concordant selection, genetic variation is maintained at a balance between deleterious mutations and purifying selection (Haldane 1937), as is likely to characterize selection at many (if not most) loci within animal genomes (e.g., Houle et al. 1996; Haag-Liautard et al. 2007; Mitchell-Olds et al. 2007; Morrow et al. 2008). Following recent

studies that demonstrate the presence of sexually antagonistic variation and unresolved sexually antagonistic selection (Chippindale et al. 2001; Foerster et al. 2007; Brommer et al. 2007; Cox & Calsbeek 2009), I also consider the impact of sex chromosomes on fitness heritability, when sexual antagonism maintains population genetic variation for fitness.

It is assumed throughout that gene content on sex chromosomes and autosomes is the same with respect to mutant fitness effects. The biological reality is that this assumption might often be violated, as a variety of studies indicate that sexually selected and sexually dimorphic traits sometimes show nonrandom linkage patterns between sex chromosomes and autosomes (e.g., Reinhold 1998; Ellegren & Parsch 2007; Mank 2009a; but see Fitzpatrick 2004; Fairbairn & Roff 2006). However, nullification of the theoretical predictions developed below would require that genes with male functions (that influence male fitness) are completely absent on sex chromosomes. This situation appears highly unlikely, and there is considerable evidence to suggest that this is not the case in well-studied species such as *Drosophila*, where X-linked mutation accumulation studies report relatively high male fertility and viability declines (Lindsley & Lifschytz 1972; Gong et al. 2005). Thus, the general assumption of equivalent gene composition between sex chromosomes and autosomes should generally be reasonable, yet is subject to modification if contradictory data emerges.

### **Sexually Concordant Genetic Variation (“Good Genes”)**

Under a model of sexually concordant selection, each locus has two possible alleles: a beneficial allele,  $A_2$ , and a deleterious allele,  $A_1$ . The fitness scheme, per diploid locus, is as follows:

Genotype:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness:	$1 - s_1$	$1 - s_1 h$	1
Male fitness:	$1 - s_2$	$1 - s_2 h$	1

where  $s_1$  and  $s_2$  are the female and male selection coefficients, and  $h$  is the dominance coefficient ( $0 < h, s_1, s_2 \leq 1$ ). For an X- or Z-linked locus, the scheme is slightly modified in the heterogametic sex, where heterozygosity is impossible:

Genotype:	$A_1$	$A_2$
XY male fitness:	$1 - s_3$	1
ZW female fitness:	$1 - s_4$	1

Under X-linked sex determination, and assuming multiplicative epistasis between loci, fitness of males and females, respectively, is:

$$w_m = \exp(-j_A s_2 - k_A s_2 h - k_{mX} s_3 - \varepsilon_m) \quad (1a)$$

$$w_f = \exp(-j_A s_1 - k_A s_1 h - j_X s_1 - k_X s_1 h - \varepsilon_f) \quad (1b)$$

where the variables  $j_A$  and  $k_A$  refer to the number of autosomal loci that are homozygous and heterozygous for a deleterious mutation, respectively;  $j_X$  and  $k_X$  (in females) refer to the number of X-linked loci that are homozygous and heterozygous for a deleterious mutation, respectively;  $k_{mX}$  refers to the number of X-linked mutations per male; and  $\varepsilon_m$  and  $\varepsilon_f$  refer to male and female environmental (non-genetic) components of fitness.

Because individual deleterious mutations are rare, each of these five variables is approximately independent with variances:  $Var(j_A) \approx L_A \hat{q}_A^2$ ,  $Var(k_A) \approx 2L_A \hat{q}_A$ ,  $Var(j_X) \approx L_X \hat{q}_X^2$ ,  $Var(k_X) \approx 2L_X \hat{q}_X$ , and  $Var(k_{mX}) \approx L_X \hat{q}_X$  (see Appendix 3.1 for variance derivations), where  $L_A$  and  $L_X$  refer to the numbers of autosomal and X-linked loci, and  $\hat{q}_A$  and  $\hat{q}_X$  are the equilibrium deleterious mutation frequencies per autosomal and X-linked locus, respectively.

Under Z-linked sex determination, sex-specific fitness functions are:

$$w_m = \exp(-j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h - \varepsilon_m) \quad (2a)$$

$$w_f = \exp(-j_A s_1 - k_A s_1 h - k_{fZ} s_4 - \varepsilon_f) \quad (2b)$$

where  $j_Z$  and  $k_Z$  (in males) refer to the number of Z-linked loci that are homozygous and heterozygous for a deleterious mutation, and  $k_{fZ}$  refers to the number of Z-linked mutations per female. As before, all variables are approximately independent with variances:  $Var(j_Z) \approx L_Z \hat{q}_Z^2$ ,  $Var(k_Z) \approx 2L_Z \hat{q}_Z$ , and  $Var(k_{fZ}) \approx L_Z \hat{q}_Z$  (see Appendix 3.1),

where  $L_Z$  refers to the number of Z-linked loci, and  $\hat{q}_Z$  is the equilibrium deleterious mutation frequency per Z-linked locus.

The approach to modeling frequencies of deleterious mutations follows previous theory by Haldane (1937). Ignoring cases of complete recessivity (i.e.,  $h = 0$ ) and strong selection ( $s > 0.05$ ), approximate equilibrium frequencies for autosomal, X-linked, and Z-linked deleterious alleles are, respectively:

$$\hat{q}_A \approx \frac{\mu_m + \mu_f}{h(s_1 + s_2)} \quad (3a)$$

$$\hat{q}_X \approx \frac{2\mu_f + \mu_m}{2s_1h + s_3} \quad (3b)$$

and

$$\hat{q}_Z \approx \frac{2\mu_m + \mu_f}{2s_2h + s_4} \quad (3c)$$

where  $\mu_m$  and  $\mu_f$  are male- and female-specific mutation rates per locus (because the  $A_1$  allele is rare, approximately all mutations are from  $A_2$  to  $A_1$ ; back-mutation is assumed to be negligible).

The diploid genomic mutation rate is  $U = 2L\mu = L(\mu_m + \mu_f)$ , where  $L$  is the total number of loci, and  $\mu$  is the mutation rate, averaged across the sexes. This can be reframed to obtain expressions for the number of autosomal and sex-linked loci –  $L_A = U(1 - P)/(\mu_m + \mu_f)$  and  $L_X = L_Z = LP/(\mu_m + \mu_f)$ , respectively – where  $P$  is the proportion of loci that are sex-linked (i.e., X- or Z-linked).

The fitness effects of autosomal relative to sex-linked mutations can be described under conditions with and without dosage compensation, with dosage influencing the relative size of sex-linked relative to autosomal selection coefficients ( $s_2$  vs.  $s_3$ , and  $s_1$  vs.  $s_4$ ). Following the approach of Charlesworth et al. (1987),  $s_2 = s_3$  and  $s_1 = s_4$  under conditions of dosage compensation. When there is no dosage compensation,  $s_2 = 2s_3$  and  $s_1 = 2s_4$ .

For simplicity, the environmental variance components ( $\varepsilon_m$  and  $\varepsilon_f$ ) are defined without gene by environment interactions, such that environmental variance does not systematically vary between species with different sex chromosome systems. In other words,  $E[\varepsilon_m]$ ,  $E[\varepsilon_f]$ ,  $Var(\varepsilon_m)$  and  $Var(\varepsilon_f)$  are constant across species. Nevertheless, the

introduction of gene-environment interactions into the model tends to reinforce the patterns described below, and do not affect the conclusions presented (see Appendix 3.3; discussion below).

### **Offspring Fitness as a Function of Paternal Fitness**

Parent-offspring regression coefficient ( $\beta$ ) equations were developed to describe the linear relationship between male fitness ( $w_m$ ) and expected fitness of their offspring,  $E[w_o|w_m]$ . Following Lynch & Walsh (1998; eq. 3.14b) regression coefficients are defined as:

$$\beta = \frac{\text{Cov}(w_m, E[w_o|w_m])}{\text{Var}(w_m)} \quad (4)$$

This approach was used for three reasons. First, regression coefficients, which describe the slope of the relationship between offspring fitness and the fitness of fathers, approximate fitness benefit or fitness cost effect sizes. Second, these coefficients are analytically tractable because they are simple functions of the fitness variance of fathers and the fitness covariance between fathers and offspring, and can easily be partitioned into sex-specific fitness heritabilities to sons and to daughters. Finally, estimates of the regression coefficients, variances, and/or covariances between relatives, can potentially be obtained for quantitative (e.g., Long & Rice 2007) and life-history traits (e.g., Møller & Alatalo 1999) that are closely associated with fitness.

For mathematical convenience, and because the relationship between the natural logarithm of fitness and number of deleterious or beneficial mutations is linear under conditions of multiplicative epistasis (Shnol & Kondrashov 1993; Gillespie 2004), analytical solutions are presented for the regression of  $\ln(E[w_o|w_m])$  on  $\ln(w_m)$ . The theoretical results are presented below as the regression coefficient for a species with sex chromosomes, standardized against the regression coefficient for a species without sex chromosomes (i.e.,  $\beta_X/\beta_{none}$  and  $\beta_Z/\beta_{none}$ ), which permits an interpretation of the relative increase or decrease of father-to-offspring heritability caused by the presence of sex chromosomes.

### Fitness Variance and Covariance Under Sexually Concordant Selection

In species with X-linked inheritance, covariance between fathers and their sons and daughters (respectively) follow the functions:

$$\text{Cov}[\ln(w_m), \ln(w_{\text{sons}})] = \frac{s_2^2}{2} [h + \hat{q}_A(1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] \quad (5a)$$

and

$$\begin{aligned} \text{Cov}[\ln(w_m), \ln(w_{\text{daughters}})] &= \frac{s_1 s_2}{2} [h + \hat{q}_A(1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] \\ &\quad + s_1 s_3 [h + \hat{q}_X(1-2h)] \text{Var}(k_{mX}) \end{aligned} \quad (5b)$$

and fitness variance among fathers is:

$$\text{Var}[\ln(w_m)] = s_2^2 [\text{Var}(j_A) + h^2 \text{Var}(k_A)] + s_3^2 \text{Var}(k_{mX}) + \text{Var}(\varepsilon_m) \quad (6)$$

(see Appendix 3.2 for details).

In species with Z-linked inheritance, covariance between fathers and their sons and daughters (respectively) follow the functions:

$$\begin{aligned} \text{Cov}[\ln(w_m), \ln(\bar{w}_{\text{sons}})] &= \frac{s_2^2}{2} [h + \hat{q}_A(1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] \\ &\quad + \frac{s_2^2}{2} [h + \hat{q}_Z(1-2h)] [2\text{Var}(j_Z) + h\text{Var}(k_Z)] \end{aligned} \quad (7a)$$

and

$$\begin{aligned} \text{Cov}[\ln(w_m), \ln(\bar{w}_{\text{daughters}})] &= \frac{s_1 s_2}{2} [h + \hat{q}_A(1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] \\ &\quad + \frac{s_2 s_4}{2} [2\text{Var}(j_Z) + h\text{Var}(k_Z)] \end{aligned} \quad (7b)$$

and fitness variance among fathers is:

$$\text{Var}[\ln(w_m)] = s_2^2 [\text{Var}(j_A) + h^2 \text{Var}(k_A) + \text{Var}(j_Z) + h^2 \text{Var}(k_Z)] + \text{Var}(\varepsilon_m) \quad (8)$$

(see Appendix 3.2 for details).

### Regression Coefficients Under Sexually Antagonistic Genetic Variation

Previous theory shows that sexual antagonistic variation can have multiple possible equilibria per locus (Kidwell et al. 1977) and complex distributions on sex chromosomes relative to autosomes (Pamilo 1979; Rice 1984; Hedrick & Parker 1997; Patten & Haig 2009). Since the key parameters affecting sexually antagonistic allele frequencies and genomic distributions are currently unknown, the specific case where sexually

antagonistic alleles have additive fitness allelic effects is considered, i.e., where for diploid loci (autosomal or sex-linked in the homogametic sex):

Genotype:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness:	1	$1 - t_1/2$	$1 - t_1$
Male fitness:	$1 - t_2$	$1 - t_2/2$	1

For haploid loci (sex-linked in the heterogametic sex):

Genotype:	$A_1$	$A_2$
X-linked male fitness:	$1 - t_3$	1
Z-linked female fitness:	1	$1 - t_4$

Under the assumption that selection coefficients are relatively small ( $t < 0.1$ ), stable polymorphic equilibria occur when  $t_1 \approx t_2$  for autosomal loci, and  $t_1 \approx t_3$  and  $t_2 \approx t_4$  for sex-linked loci. These conditions lead to equilibrium sexually antagonistic allele frequencies of  $\hat{q}_{SA} \approx 0.5$ , and predict similar levels of sexually antagonistic polymorphism at sex-linked and autosomal loci (Patten & Haig 2009).

Under these starting conditions, the sexual antagonism model yields several simple analytical results (see Appendix 3.4). For species with X-linked sex-determination, regression coefficients for sons and daughters, respectively, are:

$$\beta_{sons} = \frac{(1-P)}{2(1+P)} \quad (9a)$$

and

$$\beta_{daughters} = -\frac{(1+2P)}{2(1-t)(1+P)} \approx -\frac{(1+2P)}{2(1+P)} \quad (9b)$$

The sex-specific regression coefficients under Z-linkage are:

$$\beta_{sons} = \frac{1}{2} \quad (10a)$$

and

$$\beta_{daughters} = -\frac{(1+P)}{2(1-t)} \approx -\frac{(1+P)}{2} \quad (10b)$$

Dominance effects for sexually antagonistic alleles are expected to greatly increase the amount of sexually antagonistic variation on sex chromosomes relative to autosomes (Rice 1984; Patten & Haig 2009), and consequently, the overall contribution of sex-linked variation to total fitness variance among individuals. Therefore, by assuming an additive sexually antagonistic model, the minimum effect of the sex chromosome system on fitness heritability between fathers and offspring is quantified. Incorporating dominance into the models merely enhances heritability differences between species with different sex chromosome systems.

## RESULTS

### Genetic Benefits Under Sexually Concordant Selection

Sex chromosomes decouple the genetic basis of male and female fitness variance, despite the same alleles being favored by selection in both sexes. For the homogametic sex (females in species with X-linked sex determination and males in species with Z-linked sex determination) the entire genome is diploid, and fitness is determined by the number of mutations in the genome, independent of linkage. Since most mutational variation is located on autosomes (Haldane 1937; Hedrick & Parker 1997), autosome quality will be the primary genetic determinant of total fitness. In contrast, the heterogametic sex is haploid for a portion of the genome, which causes sex chromosome genetic quality to have a disproportionately high influence on fitness. The magnitude of this effect depends upon the relative dominance of mutations as well as the proportion of genomic diversity on the sex chromosome.

This decoupling of fitness between the sexes can generate substantial heritability differences between species with different sex chromosome systems (Fig. 3.1). The presence of X chromosomes weakens the genetic benefits of having a high-fitness father (represented here as a reduction in the regression coefficient of offspring fitness on paternal fitness; see Fig. 3.1). This reduction in fitness heritability applies to both sons and daughters, and can be quite strong. For example, under conditions of equally strong selection in males and females ( $s_1 = s_2 = s_3$ ), and partially recessive deleterious mutations

( $h < 0.5$ ), X-linkage at five percent of the genome (as is typical for mammal species; Solari 1994) can reduce offspring fitness benefits by over 40 percent (top panel of Fig. 3.1). Higher genomic proportions of X-linkage further decrease the magnitude of offspring fitness benefits.

The presence of Z chromosomes marginally influences genetic benefits to sons, and enhances fitness benefits for daughters (Fig. 3.1). This effect arises because the fitness heritability between fathers and daughters is higher for a Z-linked compared to an autosomal locus (each female Z chromosome is paternally inherited), and the fitness effect of Z-linked mutations is particularly strong in hemizygous females. The benefit of inheriting a relatively mutation-free Z will be strong for daughters if mutations generally have recessive fitness effects.

The impact of the sex chromosome system on fitness heritability remains qualitatively the same whether there is dosage compensation or not. In either case, X chromosomes reduce the paternal-offspring fitness relationship and Z chromosomes enhance the relationship between fathers and daughters. However, a lack of dosage compensation dampens the magnitude of these X and Z chromosome effects (Fig. 3.2; top left panel). This dampening effect is relatively weak in species with Z chromosomes, but is strong in species with X-linkage. Incidentally, dosage compensation appears to be ubiquitous in species with X chromosomes, but not in species with Z chromosomes (Mank 2009b). This biological pattern is expected to maximize fitness heritability differences between species with X-linkage compared to those with Z-linkage.

Many animal species exhibit stronger selection and higher mutation rates in males relative to females (Ellegren 2007; Whitlock & Agrawal 2009; Hollis et al. 2009). Under these conditions, the proportion of fitness variation that is sex-linked is expected to increase in species with Z chromosomes, and decrease in species with X chromosomes (see Hedrick 2007; Appendix 3.5). This should, in principle, enhance fitness heritability between fathers and offspring in species with X chromosomes, and increase the fitness association between fathers and daughters in species with Z chromosomes. However, sex-biased mutation and selection do not strongly influence the impact of sex chromosomes on fitness heritability (Fig. 3.2; top right and bottom left panels). A single exception arises for the case Z-linkage, where stronger selection in males can greatly

improve fitness heritability between fathers and daughters. For example, when selection is three times stronger in males than females ( $s_2 = 3s_1 = 3s_d$ ), Z-linkage across ten percent of the genome (as is typical for birds; Solari 1994) will enhance fitness benefits to daughters by up to 50 percent. The magnitude of this effect increases with the strength of selection in males relative to females, and the proportion of the genome that is Z-linked.

Environmental contributions to paternal fitness variance will serve to reduce the association between male fitness of mean fitness of their offspring. The question is whether environmental variance differentially influences fitness heritability between different species. When environmental variance is independent of genotype, the fitness variance due to environment will be constant across individuals and between species, and tends to decrease the effects of the sex chromosome system on fitness heritability (Fig 3.2; bottom right panel). However, this result only characterizes “fixed” environmental variance that is independent of genotype. Fitness variance due to gene-by-environment (GxE) interactions tends to reinforce and strengthen fitness heritability differences between species with different sex chromosome systems (see Appendix 3.3 for a full analysis). To the extent that GxE interactions are common (as appears to be the case for a wide variety of quantitative traits; Mackay 2001; Mackay et al. 2009), fixed environmental variance is unlikely to eliminate fitness heritability differences between sex chromosome systems.

### **Sexually Antagonistic Fitness Variation**

Though sexually concordant directional selection (“good genes”) probably characterize evolution at a large majority of loci within animal genomes (Morrow et al. 2008), sexually antagonistic variation is also likely to contribute strongly to overall fitness variation (Chippindale et al. 2001; Prasad et al. 2007; Foerster et al. 2007; Cox & Calsbeek 2009). Under all forms of sex determination, sexual antagonism is expected to generate a positive fitness association between fathers and sons and a negative association between fathers and daughters. Nevertheless, the magnitude of these effects can potentially be influenced by sex linkage, and either enhance or diminish overall patterns of fitness heritability between fathers and offspring.

Fitness heritability due to sexually antagonistic variation is affected by each species' sex chromosome system (Fig 3.3; top panel). There is a negative fitness relationship between fathers and daughters in all species, with both X and Z chromosomes exacerbating the negative effect to daughters. Since “good genes” fitness benefits to daughters are relatively weak in species with X chromosomes, the addition of sexually antagonistic variation is more likely to overwhelm positive genetic associations between fathers and daughters caused by non-sexually antagonistic fitness variation. In other words, the likelihood of empirically observing a net negative fitness association between fathers and daughters should be greatest for species with X chromosomes, particularly those with a large X.

Fitness of sons is positively associated with paternal fitness, yet this association is considerably weaker in species with X chromosomes compared to those with other sex determination systems. Sexually antagonistic fitness benefits to sons will be exactly the same in species with Z chromosomes as those without sex chromosomes (Fig 3.3; bottom panel). X-linkage rapidly decreases the association between fathers and sons, with moderately sized X chromosomes (e.g.,  $P = 0.1$ ) reducing the regression coefficient for fitness by about 15 percent. Large X chromosomes ( $P > 0.2$ ), which are often observed in insect species, can reduce the regression coefficient between fathers and sons by over 25 percent.

It is worth reemphasizing that the sexual antagonism model analyzed here assumes that alleles have additive fitness effects; consequently, sex-linked and autosomal loci are expected to harbor similar levels of sexually antagonistic genetic variation (Pamilo 1979; Hedrick & Parker 1997). A violation of this assumption is expected to redistribute most sexually antagonistic variation to the sex chromosomes (Rice 1984; Patten & Haig 2009), which should further inflate the contribution of the X to total male fitness variance, and will consequently reduce the empirical detectability of genetic benefits to sons, and enhance detectability of sexually antagonistic fitness costs to daughters.

## **DISCUSSION**

Mutation accumulation studies indicate that deleterious mutations carry an average heterozygous fitness cost of about one percent (Shabalina et al. 1997), and that the

deleterious mutation rate is typically on the order of one, per genome, per generation (Haag-Liautard et al. 2007; Baer et al. 2007). Rice (1988) argued that such parameters should generate relatively high fitness variance at mutation-selection equilibrium, resulting in substantial genetic benefits to offspring of nonrandomly mating females. The results presented here suggest that that genetic benefits, estimated by the fitness correlation between fathers and offspring, is highly sensitive to the genomic architecture of fitness, which in turn, is strongly influenced by each species' sex chromosome system. When one considers fitness heritability to both sons and to daughters, species with X chromosomes are expected to show reduced fitness heritability between fathers and their offspring, particularly if deleterious mutations are partially recessive and sexually antagonistic variation contributes to fitness variance among individuals. Both of these conditions have strong empirically support (Charlesworth & Charlesworth 1999; Bonduriansky & Chenoweth 2009).

The results of these genetic models raise four questions. Do sex chromosome systems, in practice, differentially affect patterns of fitness heritability between fathers and offspring in different species? Do most studies of indirect fitness benefits focus on species with X chromosomes, and if so, is this likely to explain the currently modest empirical support for genetic benefits? What do theoretical patterns of fitness heritability, as a function of sex chromosomes, imply about the evolution of female mating biases? Finally, what do empirical patterns of fitness heritability tell us about the maintenance of genetic variation for fitness? Each of these questions is addressed below.

### **Sex Chromosomes and Viability Benefits to Offspring**

Genetic benefits of mate choice require that the genetic association between fathers and offspring for total fitness be positive. A test of whether such associations differ between species with different sex chromosome systems would require estimates of total fitness heritability for a large number of species, and across a range of sex chromosome systems. Estimation of total fitness is currently not feasible for most species, yet several studies have estimated associations between male mating success and offspring preadult viability, traits that are likely to be strongly linked to overall fitness.

To assess whether genetic benefit effect sizes might be influenced by sex chromosome systems, I conducted a search of the sexual selection literature (relevant data were previously compiled by Møller and Alatalo 1999; I extended their dataset by searching the ISI Web of Knowledge using a variety of keywords associated with sexual selection and genetic benefits) and identified a total of 28 species where the target of female choice was known (preferred male traits were identified), and the association between paternal phenotype and offspring viability was measured (see Appendix 3.6). Effect sizes (presented as Pearson correlation coefficients) were obtained directly from the literature, or calculated using transformations described in Rosenthal (1994). In a few cases, species data was available from more than one independent study (i.e., *Acrocephalus arundinaceus*), more than one male trait influencing female choice (i.e., *Dendrobates leucomelas*, *Epipedobates tricolor*), or multiple conditions in which offspring viability was tested (i.e., *Tribolium castaneum*). Effect sizes were averaged in such cases to obtain species-specific coefficients. Sex chromosome data was obtained from White (1973), Bull (1983), Solari (1994), Devlin & Nagahama (2002), and Mank et al. (2006).

Estimates of genetic benefits vary substantially between studies, yet partitioning species according to their sex chromosome system reveals a striking pattern: offspring viability benefits are systematically lower in species with X chromosomes (Fig 3.4), consistent with the theoretical predictions given above. Most of the evidence against good genes benefits comes from species with X chromosomes – i.e., those species that are predicted by theory to show the weakest fitness correlations between fathers and offspring. When one considers data from species without X chromosomes, the evidence for good genes effects is quite strong. Considering the 14 species (from Appendix 3.6) with undifferentiated sex chromosomes or with Z-linked sex determination, every heritability estimate is positive ( $\chi^2 = 14$ ;  $p = 0.0001$ ), and the average effect size is relatively large ( $\bar{r} = 0.403$ ; lower 95% CI: 0.237; upper 95% CI: 0.569).

Though sex chromosome variability appears to partially account for species-specific genetic benefits, this interpretation comes with a major caveat. When the data are considered within a phylogenetic framework, it becomes clear that there is little phylogenetic independence between species with similar sex chromosome systems.

Though individual branches of the phylogeny remain consistent with model predictions, the sample size of five independent lineages precludes a statistical test of significance. This current limitation of the data can be alleviated by focusing in the future on related species with different sex chromosome systems. For example, fish and reptile species exhibit a diversity of environmental, X- and Z-linked chromosomal sex determination (Devlin & Nagahama 2002; Organ & Janes 2008). Mosquitoes also represent a potentially interesting group, as some species have completely undifferentiated sex chromosomes, while others have relatively large, heteromorphic sex chromosomes (Presgraves and Orr 1998).

Additional experimental support for model predictions comes from *Drosophila*, which has a relatively large X chromosome (approximately 20 percent of the genome). In most studies, paternal mating success is a relatively poor predictor of offspring viability (e.g., Partridge 1980; Schaeffer et al. 1984; Pitnick 1991). In contrast, Taylor et al. (1987) studied genetic variation on autosomes, and discovered an extremely strong positive genetic correlation between male mating success and larval viability under competition conditions ( $r^2 \approx 0.85$ ; see their Fig. 2). In their study, all flies shared the same X chromosome, and genetic variance among males was entirely based on autosomal differences. When considered together, these studies suggest that the *Drosophila* X normally obscures autosomal genetic quality. Thus, genetic benefits are high in studies with a controlled X chromosome (i.e., Taylor et al. 1987) but weak in studies considering variation across the entire genome (e.g., Partridge 1980; Schaeffer et al. 1984).

### **Heritability of Adult Fitness Components**

Data on the inheritance of adult fitness components is relatively sparse and difficult to interpret. In idealized experimental systems (e.g., pedigree studies using well-characterized natural populations or direct estimation using lab-adapted populations of *Drosophila*) adult fitness can be measured (Chippindale et al. 2001; Qvarnström et al. 2006; Brommer et al. 2007; Foerster et al. 2007; Kruuk et al. 2008). Such studies have only been conducted using a small number of populations and yield mixed results (Qvarnström et al. 2006; Brommer et al. 2007; Kirkpatrick 2009).

Most studies estimate the heritability of adult phenotypes that are correlated with fitness. This latter approach is more taxonomically inclusive and can potentially yield information about whether or not there is a net benefit or net cost to offspring fitness. Several of these studies have yielded evidence that males with high mating success sire reproductively successful sons (Reynolds & Gross 1992; Jones et al. 1998; Wedell & Tregenza 1999; Fedorka & Mousseau 2004; Head et al. 2005; Taylor et al. 2007; but see Pischedda & Chippindale 2006; Oneal et al. 2007; Connallon & Jakubowski 2009). However, since the strength of the correlation between the measured trait and “fitness” is unknown, these studies cannot be used to contrast fitness effect sizes between species.

The current data cannot address whether and how strongly sex chromosomes influence the heritability of adult fitness components. However, it is worth pointing out that most published studies have utilized insect species with X chromosomes. The focus on species with X-linkage can potentially undermine attempts to discover evidence for genetic benefits, since theory predicts that X-linkage will reduce fitness benefits between fathers and offspring and enhance sexually antagonistic fitness costs to daughters.

### **Indirect Benefits and the Evolution of Mating Biases**

Previous theory shows that sex determination systems can differentially influence the evolution of female choice for indirect benefits. These studies examine the conditions under which female preference alleles invade the population as a consequence of nonrandom associations (linkage disequilibrium) with alleles improving viability (Hastings 1994; Kirkpatrick & Hall 2004), male attractiveness (Reeve & Pfennig 2003; Kirkpatrick & Hall 2004), or alleles with sexually antagonistic fitness effects (Albert & Otto 2005).

Kirkpatrick & Hall (2004) showed that the evolution of female choice is facilitated when female preference alleles are X-linked, a finding that appears to be at odds with the results presented above, that X-linkage reduces the magnitude of indirect fitness benefits between fathers and their offspring. This apparent difference is important and sheds light on a common mismatch between theory and experiments in sexual selection. The strength of selection on mate choice depends upon the genetic correlation between alleles influencing the female preference and the genetic variation for fitness. Such correlations

accumulate over many generations, and might be more easily detected with pedigrees (e.g., by calculating sex specific breeding values; Kokko et al. 2003; Foerster et al. 2007), or by artificial selection experiments (e.g., Wilkinson & Reillo 1994). In contrast, most sexual selection studies test for fitness associations between fathers and offspring across two generations. While such data can be extremely informative about the genetic architecture of fitness or sexually selected traits, it can provide misleading evidence regarding the mechanisms driving the evolution of female choice, and may be more misleading in some species relative to others.

Discussion here of genetic benefits and costs has focused on models where male quality follows an absolute ranking scheme, such that, if genetic quality of male A is greater than male B, and male B is greater than male C, then A must also be greater than C. This general assumption is appropriate regarding genetic variation under directional selection in males, as expected for deleterious and sexually antagonistic mutations. On the other hand, “genetic quality” can be context-dependent when fitness is influenced by genetic interactions (e.g., genetic compatibility; Zeh & Zeh 1996, 1997; Neff & Pitcher 2005), or frequency-dependent selection (e.g., Clark et al. 1999). Both types of fitness variation (absolute and context-dependent) can potentially influence female mating decisions. However, context-dependent genetic benefits are not necessarily expected to differentially influence species with different sex determination mechanisms (though overdominant fitness interactions between alleles is expected to be rare for sex-linked compared to autosomal loci; Pamilo 1979; Hedrick & Parker 1997).

### **Heritability versus Population Genetic Variation for Fitness**

The opportunity for indirect genetic benefits has major population genetic implications for sexually reproducing species. As a consequence of sexual selection, fitness variance among males is typically greater than among females, which suggests that genetic variation underlying male traits is subject to strong selection (e.g., Bateman 1948; Andersson 1994). However, the relationship between genetic variation for male fitness and genetic variation for female fitness is not well known (Whitlock & Agrawal 2009; Bonduriansky & Chenoweth 2009; Kirkpatrick 2009). It is unclear whether genetic

variation typically has sex-limited, sexually concordant, or sexually antagonistic fitness consequences.

The good genes hypothesis predicts that sexual selection often follows the same evolutionary trajectory as natural selection. Such reinforcement of natural selection by sexual selection has two general implications. First, it implies that adaptation within sexually reproducing populations will be enhanced by sexual selection (Candolin & Heuschele 2008; Whitlock & Agrawal 2009). It also implies that conflicting selection pressures between life-history stages or between the sexes (i.e., sexual antagonism or antagonistic pleiotropy) – processes that potentially maintain fitness variation (Prout 1999) – are unlikely to characterize much population genetic variation for fitness. By default, other hypotheses, including variation maintained at mutation-selection balance, become more plausible.

This general prediction is complicated by sex chromosomes, which can alter the indirect fitness consequences of mating with high-fitness males. Sex chromosome differences may permit some species to show strong evidence for indirect benefits, while others will display patterns of weak benefits, or perhaps net costs to daughters caused by sexually antagonistic variation. Such heritability differences between species may be deceptive. Instead of arising because species harbor different amounts or types of genetic variation, they may simply arise from species differences in the genomic architecture of fitness, as caused by sex chromosome differences. This explanation might account for apparently conflicting observations in *Drosophila* species, which have particularly large X chromosomes. Despite most mutations appearing to have negative effects on both sexes (Morrow et al. 2008), mate choice tends to generate a net fitness cost to daughters consistent with a model of sexual antagonism (Pischedda & Chippindale 2006; Oneal et al. 2007; Connallon & Jakubowski 2009). That X-linkage exacerbates sexually antagonistic costs and minimizes genetic benefits might explain this discrepancy.

## **Conclusion**

Sex-specific genetic architecture for fitness can play a major role in determining species-specific patterns of fitness heritability. This sex-specificity is strongly influenced by the sex chromosome system – i.e., whether males or females are heterogametic,

whether sex chromosomes are homomorphic or heteromorphic, and whether X or Z chromosomes are small or large. Consequently, fitness heritability differences between animal species can arise via differences in genome structure.

The theory presented here suggests that evidence for positive fitness associations between fathers and offspring is least likely to come from species with X chromosomes, yet such species are more likely to reveal negative fitness associations between fathers and daughters – a consequence of sexually antagonistic genetic variation. Current empirical data are consistent with predictions of the theory. This data does not prove that sex chromosomes are responsible for heritability differences between species, but it emphasizes the necessity of considering how genomic attributes of species might influence our inferences about genetic variation and adaptation.

Figure 3.1. Sex chromosome size and “good genes” benefits to offspring.  $\beta_D$  is the regression coefficient of daughter fitness as a function of paternal fitness;  $\beta_S$  is the regression coefficient for the fitness of sons;  $P$  represents the proportion of the diploid genome that is sex-linked. Regression coefficients for species with sex chromosomes are standardized by the corresponding coefficients for species without sex chromosomes (scaled genetic benefits =  $\beta_X/\beta_{\text{none}}$  or  $\beta_Z/\beta_{\text{none}}$ ). Results shown above are under conditions of dosage compensation and equal selection and mutation between the sexes ( $s_1 = s_2 = s_3 = s_4 = 0.01$ ;  $\mu_f = \mu_m = 10^{-6}$ ). To emphasize genetic effects, environmental variance is set to zero.

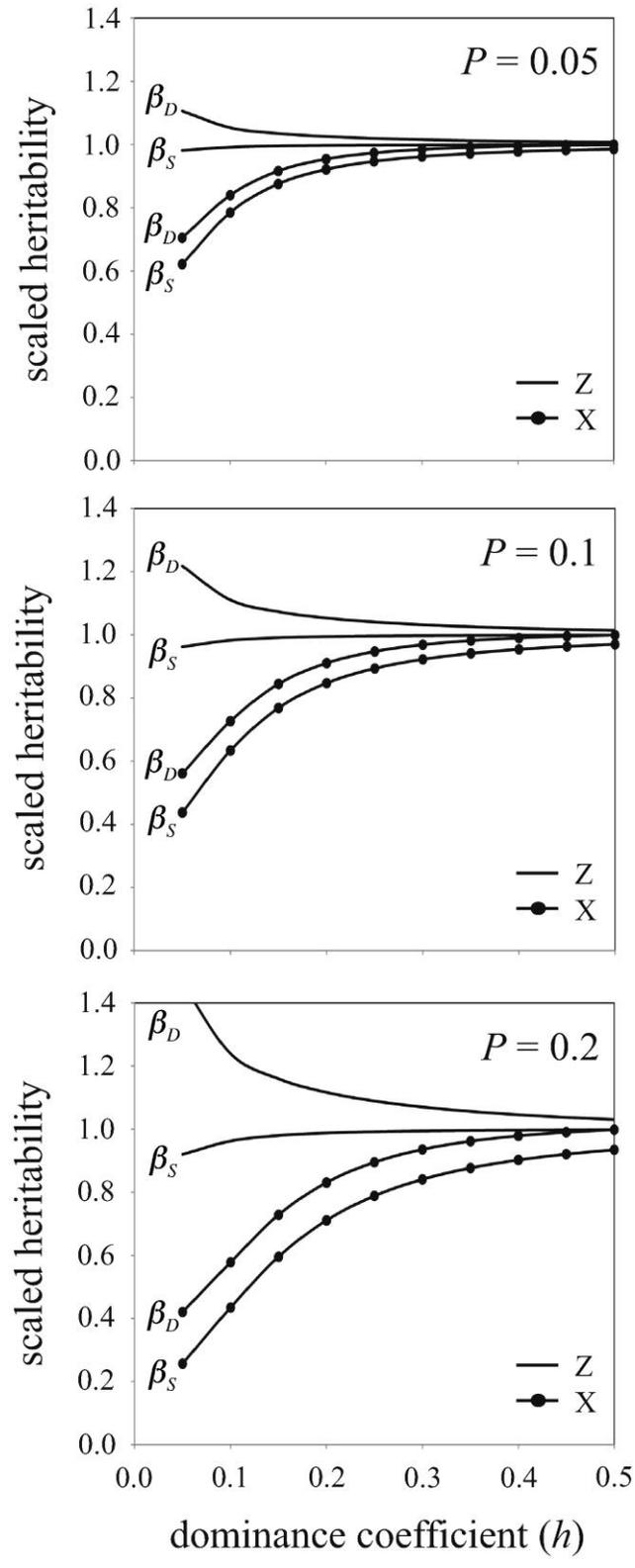
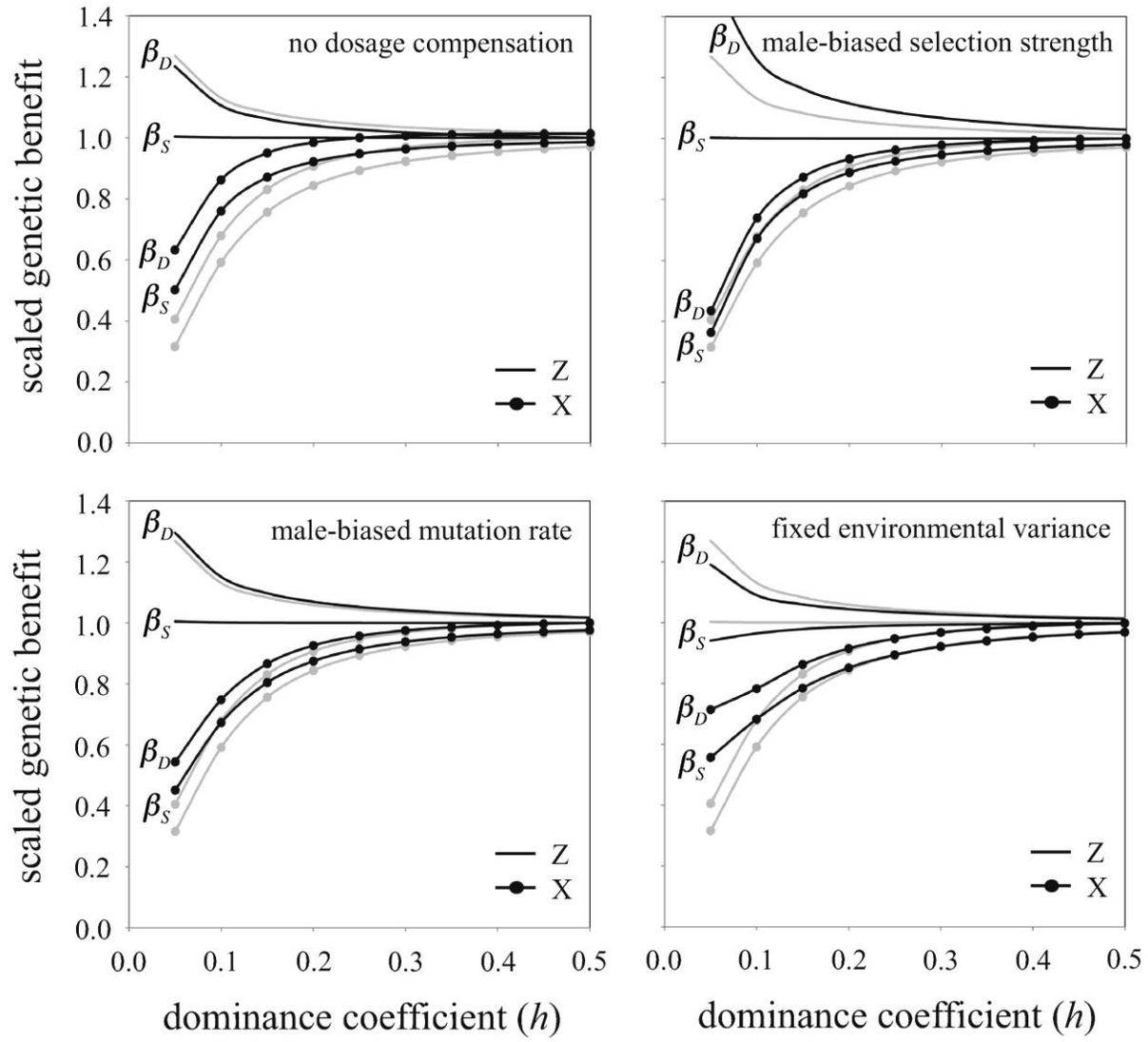


Figure 3.2. Representative effects of dosage compensation, environmental variance, and unequal mutation and selection between males and females. Presentation of results follows that of Figure 1, with sex chromosomes representing ten percent of the genome ( $P = 0.1$ ). Results in gray correspond to theoretical results from Figure 1, second panel ( $P = 0.1$ ;  $s_1 = s_2 = s_3 = s_4 = 0.01$ ;  $\mu_f = \mu_m = 10^{-6}$ ;  $\text{Var}(\varepsilon) = 0$ ). Results in black correspond to four conditions: (A) dosage compensation is absent; (B) selection coefficients are four times larger in males than females; (C) mutation rates per locus are four times larger in males than females; (D) environmental variance is a constant across species (no gene-by-environment interactions), arbitrarily defined here as  $\text{Var}(\varepsilon_m) = 1.4 \times 10^{-3}$ .



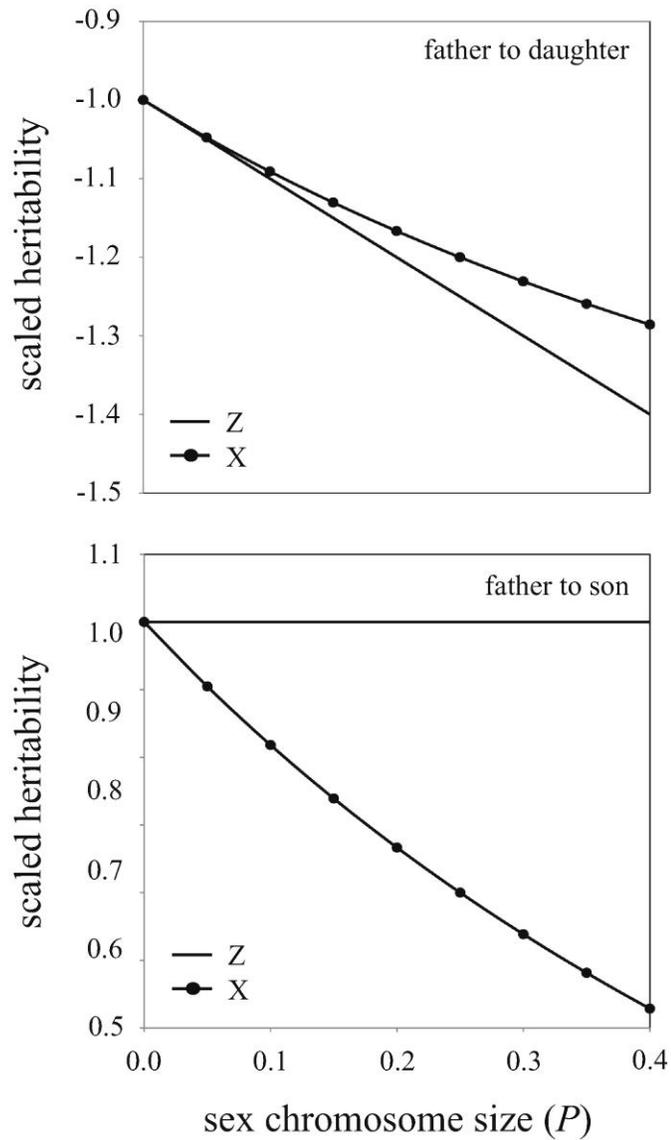
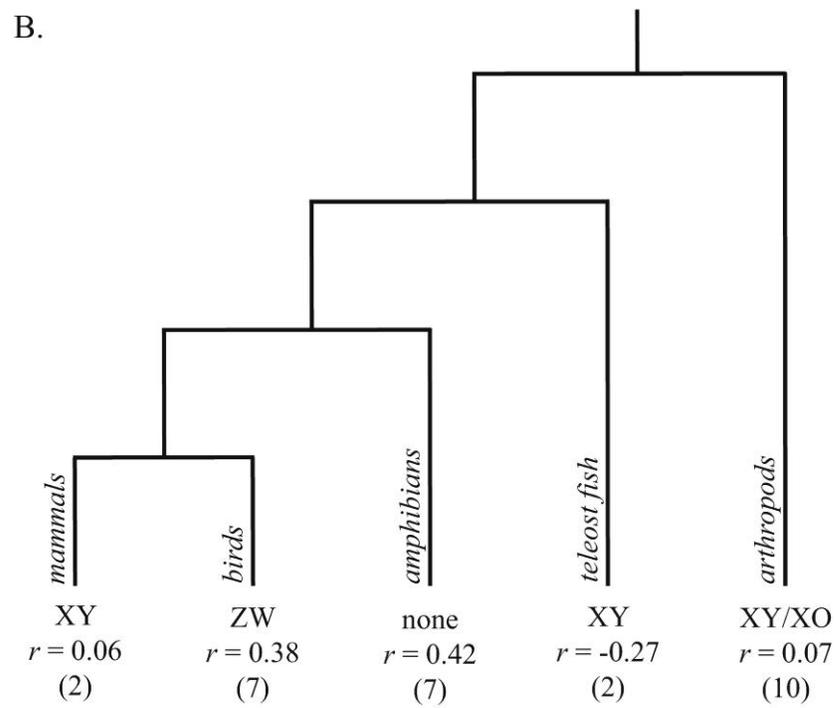
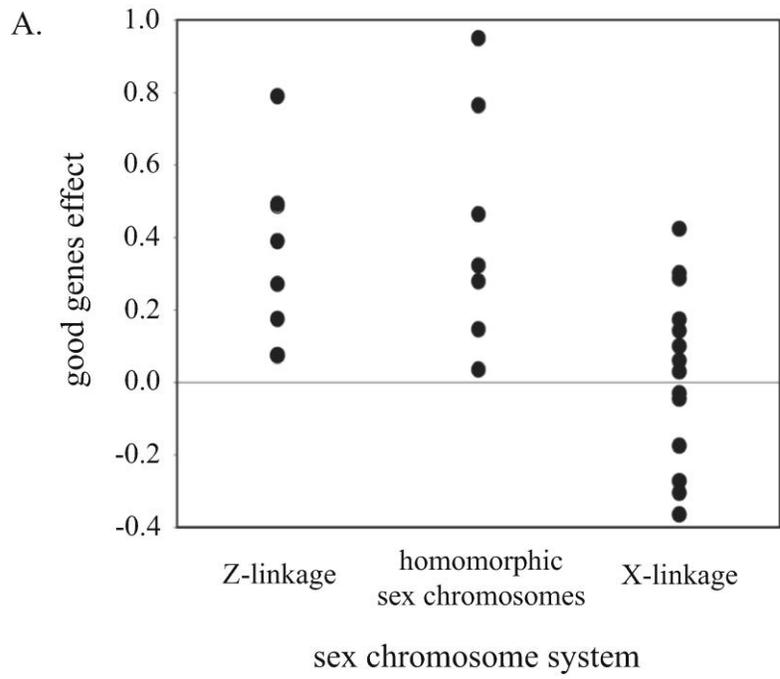


Figure 3.3. Father-to-offspring fitness heritability under X- and Z-linked sex determination and sexually antagonistic variation. Heritability is standardized by fitness heritability in a species without sex chromosomes. Results are based on equations (9) and (10), assuming small selection coefficients ( $t < 0.05$ ).

Figure 3.4. X-linkage is associated with reduced genetic benefits in nature. Good genes viability benefits were collected for 28 species with diverse sex chromosome systems (see Appendix 3.6). For each species, effect sizes are presented as Pearson correlation coefficients ( $r$ ) between male traits conferring increased mating success and offspring survival. **A**). Z-linked and X-linked taxa are those with heteromorphic sex chromosomes (W or Y chromosomes are at least partially degenerate or are absent). Species with genetic sex determination but with homomorphic sex chromosomes are classified as 'none'. Sample sizes are X:  $n = 14$ , Z:  $n = 7$ , and 'none':  $n = 7$ . Groups were compared with two-tailed Mann-Whitney U-tests (calculated online at <http://faculty.vassar.edu/lowry/utest.html>), with X-linked species having significantly reduced good genes effects relative to the other species (X vs. Z:  $p = 0.008$ ; X vs. none:  $p = 0.0124$ ). **B**). the mean effect size (sample size is in parentheses) in species partitioned into independent evolutionary lineages.



### Appendix 3.1. Genomic variance and covariance functions for genotypes at sex-linked and autosomal loci

Variables  $j_A$  and  $k_A$  are drawn from a multinomial distribution with parameters  $n = L_A$ ,  $p_j = \hat{q}_A^2$ , and  $p_k = 2\hat{q}_A(1-\hat{q}_A)$ . Because  $\hat{q}_A$  is expected to be small ( $\hat{q}_A \ll 0.01$ ),  $j_A$  and  $k_A$  have the following properties of variance and covariance:

$$\text{Var}(k_A) = L_A 2\hat{q}_A(1-\hat{q}_A)[1-2\hat{q}_A(1-\hat{q}_A)] \approx 2L_A\hat{q}_A$$

$$\text{Var}(j_A) = L_A\hat{q}_A^2(1-\hat{q}_A^2) \approx L_A\hat{q}_A^2$$

and

$$\text{Cov}(j_A, k_A) = -L_A\hat{q}_A^2 2\hat{q}_A(1-\hat{q}_A) \approx 0$$

Variables  $j_X$  and  $k_X$  are drawn from a multinomial distribution with parameters  $n = L_X$ ,  $p_j = \hat{q}_X^2$ , and  $p_k = 2\hat{q}_X(1-\hat{q}_X)$ , with covariance and variances:

$$\text{Var}(k_X) = L_X 2\hat{q}_X(1-\hat{q}_X)[1-2\hat{q}_X(1-\hat{q}_X)] \approx 2L_X\hat{q}_X$$

$$\text{Var}(j_X) = L_X\hat{q}_X^2(1-\hat{q}_X^2) \approx L_X\hat{q}_X^2$$

and

$$\text{Cov}(j_X, k_X) = -L_X\hat{q}_X^2 2\hat{q}_X(1-\hat{q}_X) \approx 0$$

The number of X-linked deleterious mutations,  $k_{mX}$ , is binomially distributed with variance  $\text{Var}(k_{mX}) = L_X\hat{q}_X(1-\hat{q}_X) \approx L_X\hat{q}_X$ .

## Appendix 3.2. Variance and covariance functions for fathers and expected fitness of their offspring

### I. X-linked sex determination

Using the linearity property for independent variables, the expected fitness of offspring, given the fitness of their fathers is:

$$E[\ln(w_{\text{sons}})|j_A, k_A, k_{mX}, \varepsilon_m] = E[-j_A^s s_2 | j_A, k_A, k_{mX}, \varepsilon_m] + E[-k_A^s s_2 h | j_A, k_A, k_{mX}, \varepsilon_m] + E[-k_{mX}^s s_3 | j_A, k_A, k_{mX}, \varepsilon_m] + E[-\varepsilon_m^s | j_A, k_A, k_{mX}, \varepsilon_m]$$

$$E[\ln(w_{\text{daughters}})|j_A, k_A, k_{mX}, \varepsilon_m] = E[-j_A^d s_1 | j_A, k_A, k_{mX}, \varepsilon_m] + E[-k_A^d s_1 h | j_A, k_A, k_{mX}, \varepsilon_m] + E[-j_X^d s_1 | j_A, k_A, k_{mX}, \varepsilon_m] + E[-k_X^d s_1 h | j_A, k_A, k_{mX}, \varepsilon_m] + E[-\varepsilon_f^d | j_A, k_A, k_{mX}, \varepsilon_m]$$

where superscripts refer to sons (“s”) and daughters (“d”).

Assuming independence between paternal and maternal genotype, and between paternal and offspring environment, the individual terms can be rewritten as:

$$E[-j_A^s s_2 | j_A, k_A, k_{mX}, \varepsilon_m] = -\left(j_A + \frac{k_A}{2}\right) \hat{q}_A s_2$$

$$E[-k_A^s s_2 h | j_A, k_A, k_{mX}, \varepsilon_m] = -\left(j_A + \frac{k_A}{2}\right) (1 - 2\hat{q}_A) s_2 h - L_A \hat{q}_A s_2 h$$

$$E[-j_A^d s_1 | j_A, k_A, k_{mX}, \varepsilon_m] = -\left(j_A + \frac{k_A}{2}\right) \hat{q}_A s_1$$

$$E[-k_A^d s_1 h | j_A, k_A, k_{mX}, \varepsilon_m] = -\left(j_A + \frac{k_A}{2}\right) (1 - 2\hat{q}_A) s_1 h - L_A \hat{q}_A s_1 h$$

$$E[-j_X^d s_1 | j_A, k_A, k_{mX}, \varepsilon_m] = -k_{mX} \hat{q}_X s_1$$

$$E[-k_X^d s_1 h | j_A, k_A, k_{mX}, \varepsilon_m] = -k_{mX} (1 - 2\hat{q}_X) s_1 h - L_X \hat{q}_X s_1 h$$

$$E[-k_{mX}^s s_3 | j_A, k_A, k_{mX}, \varepsilon_m] = -L_X \hat{q}_X s_3$$

$$E[-\varepsilon_m^s | j_A, k_A, k_{mX}, \varepsilon_m] = -E[\varepsilon_m]$$

$$E[-\varepsilon_f^d | j_A, k_A, k_{mX}, \varepsilon_m] = -E[\varepsilon_f]$$

Substitution and simplification yields:

$$E[\ln(w_{sons})|j_A, k_A, k_{mX}, \varepsilon_m] = \ln(\bar{w}_{sons}) = -\frac{s_2}{2}[h + \hat{q}_A(1-2h)](2j_A + k_A) - L_A \hat{q}_A s_2 h - L_X \hat{q}_X s_3 - E[\varepsilon_m]$$

$$E[\ln(w_{daughters})|j_A, k_A, k_{mX}, \varepsilon_m] = \ln(\bar{w}_{daughters}) = -\frac{s_1}{2}[h + \hat{q}_A(1-2h)](2j_A + k_A) - s_1[h + \hat{q}_X(1-2h)]k_{mX} - L_A \hat{q}_A s_2 h - L_X \hat{q}_X s_1 h - E[\varepsilon_f]$$

**Covariance between fathers and sons:**

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{sons})] &= \text{Cov}\left\{-j_A s_2 - k_A s_2 h - k_{mX} s_3 - \varepsilon_m, -\frac{s_2}{2}[h + \hat{q}_A(1-2h)](2j_A + k_A) - L_A \hat{q}_A s_2 h - L_X \hat{q}_X s_3 - E[\varepsilon_m]\right\} \\ &= \frac{s_2^2}{2}[h + \hat{q}_A(1-2h)][2\text{Var}(j_A) + h\text{Var}(k_A)] \end{aligned}$$

**Covariance between fathers and daughters:**

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{daughters})] &= \text{Cov}\left\{\begin{array}{l} -j_A s_2 - k_A s_2 h - k_{mX} s_3 - \varepsilon_m, \\ -s_1[h + \hat{q}_A(1-2h)]j_A - \frac{s_1}{2}[h + \hat{q}_A(1-2h)]k_A - s_1[h + \hat{q}_X(1-2h)]k_{mX} - L_A \hat{q}_A s_1 h - L_X \hat{q}_X s_1 h - E[\varepsilon_f] \end{array}\right\} \\ &= \frac{s_1 s_2}{2}[h + \hat{q}_A(1-2h)][2\text{Var}(j_A) + h\text{Var}(k_A)] + s_1 s_3[h + \hat{q}_X(1-2h)]\text{Var}(k_{mX}) \end{aligned}$$

**Fitness variance among fathers:**

$$\begin{aligned} \text{Var}[\ln(w_{father})] &= \text{Var}(-j_A s_2 - k_A s_2 h - k_{mX} s_3 - \varepsilon_m) \\ &= s_2^2[\text{Var}(j_A) + h^2\text{Var}(k_A)] + s_3^2\text{Var}(k_{mX}) + \text{Var}(\varepsilon_m) \end{aligned}$$

## II. Z-linked sex determination

Using the same approach as above, we have:

$$E[\ln(w_{sons})|j_A, k_A, j_Z, k_Z, \varepsilon_m] = E[-j_A^s s_2 | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-k_A^s s_2 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-j_Z^s s_2 | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-k_Z^s s_2 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] \\ + E[-\varepsilon_m^s | j_A, k_A, j_Z, k_Z, \varepsilon_m]$$

$$E[\ln(w_{daughters})|j_A, k_A, j_Z, k_Z, \varepsilon_m] = E[-j_A^d s_1 | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-k_A^d s_1 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-k_{fZ}^d s_4 | j_A, k_A, j_Z, k_Z, \varepsilon_m] + E[-\varepsilon_f^d | j_A, k_A, j_Z, k_Z, \varepsilon_m]$$

with individual terms:

$$E[-j_A^s s_2 | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_2 \left( j_A + \frac{k_A}{2} \right) \hat{q}_A$$

$$E[-j_Z^s s_2 | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_2 \left( j_Z + \frac{k_Z}{2} \right) \hat{q}_Z$$

$$E[-k_A^s s_2 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_2 h \left( j_A + \frac{k_A}{2} \right) (1 - 2\hat{q}_A) - s_2 h L_A \hat{q}_A$$

$$E[-k_Z^s s_2 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_2 h \left( j_Z + \frac{k_Z}{2} \right) (1 - 2\hat{q}_Z) - s_2 h L_Z \hat{q}_Z$$

$$E[-j_A^d s_1 | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_1 \left( j_A + \frac{k_A}{2} \right) \hat{q}_A$$

$$E[-k_{fZ}^d s_4 | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_4 \left( j_Z + \frac{k_Z}{2} \right)$$

$$E[-k_A^d s_1 h | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -s_1 h \left( j_A + \frac{k_A}{2} \right) (1 - 2\hat{q}_A) - s_1 h L_A \hat{q}_A$$

$$E[-\varepsilon_m^s | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -E[\varepsilon_m]$$

$$E[-\varepsilon_f^d | j_A, k_A, j_Z, k_Z, \varepsilon_m] = -E[\varepsilon_f]$$

Substitution and simplification yields:

$$E[\ln(w_{sons})|j_A, k_A, j_Z, k_Z, \varepsilon_m] = -\frac{s_2}{2} [h + \hat{q}_A (1 - 2h)] [2j_A + k_A] - \frac{s_2}{2} [h + \hat{q}_Z (1 - 2h)] [2j_Z + k_Z] - s_2 h [L_A \hat{q}_A + L_Z \hat{q}_Z] - E[\varepsilon_m]$$

$$E[\ln(w_{daughters})|j_A, k_A, j_Z, k_Z, \varepsilon_m] = -\frac{s_1}{2} [h + \hat{q}_A (1 - 2h)] [2j_A + k_A] - \frac{s_4}{2} [2j_Z + k_Z] - s_1 h L_A \hat{q}_A - E[\varepsilon_f]$$

**Covariance between fathers and sons:**

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{sons})] &= \text{Cov} \left\{ \begin{array}{l} -j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h - \varepsilon_m, \\ -s_2 [h + \hat{q}_A (1-2h)] j_A - \frac{s_2}{2} [h + \hat{q}_A (1-2h)] k_A - s_2 [h + \hat{q}_Z (1-2h)] j_Z - \frac{s_2}{2} [h + \hat{q}_Z (1-2h)] k_Z - s_2 h L_A \hat{q}_A - s_2 h L_Z \hat{q}_Z - E[\varepsilon_m] \end{array} \right\} \\ &= \frac{s_2^2}{2} [h + \hat{q}_A (1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] + \frac{s_2^2}{2} [h + \hat{q}_Z (1-2h)] [2\text{Var}(j_Z) + h\text{Var}(k_Z)] \end{aligned}$$

**Covariance between fathers and daughters:**

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{daughters})] &= \text{Cov} \left\{ \begin{array}{l} -j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h - \varepsilon_m, \\ -s_1 [h + \hat{q}_A (1-2h)] j_A - \frac{s_1}{2} [h + \hat{q}_A (1-2h)] k_A - s_4 j_Z - \frac{s_4}{2} k_Z - s_1 h L_A \hat{q}_A - E[\varepsilon_f] \end{array} \right\} \\ &= \frac{s_1 s_2}{2} [h + \hat{q}_A (1-2h)] [2\text{Var}(j_A) + h\text{Var}(k_A)] + \frac{s_2 s_4}{2} [2\text{Var}(j_Z) + h\text{Var}(k_Z)] \end{aligned}$$

**Fitness variance among fathers:**

$$\begin{aligned} \text{Var}[\ln(w_m)] &= \text{Var}(-j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h - \varepsilon_m) \\ &= s_2^2 [\text{Var}(j_A) + h^2 \text{Var}(k_A)] + s_2^2 [\text{Var}(j_Z) + h^2 \text{Var}(k_Z)] + \text{Var}(\varepsilon_m) \end{aligned}$$

### Appendix 3.3. Incorporating gene by environment interactions

With gene by environment (GxE) interactions, the environmental variance term might not be uniform across different genotypes. To account for such interactions, environmental variance is reconsidered under a distribution with constant means,  $E[\varepsilon_m] = m$  and  $E[\varepsilon_f] = f$ , and environmental variance as a function of the genotype:

$$\text{With X-linkage: } \text{Var}(\varepsilon_m) = C_m \text{Var}(-j_A s_2 - k_A s_2 h - k_{mX} s_3)$$

$$\text{With Z-linkage: } \text{Var}(\varepsilon_m) = C_m \text{Var}(-j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h)$$

$$\text{No sex chromosomes: } \text{Var}(\varepsilon_m) = C_m \text{Var}(-j_A s_2 - k_A s_2 h)$$

where  $C_m$  represents a positive constant that relates the environmental variance to the genotypic variance. Since the means are constant across genotypes, the fitness covariance between environment and genotype within individuals and between fathers and offspring will be zero. The fitness variance among fathers is therefore modified as follows...

#### *Variance among fathers (with X-linkage):*

$$\begin{aligned} \text{Var}[\ln(w_m)] &= \text{Var}(-j_A s_2 - k_A s_2 h - k_{mX} s_3 - \varepsilon_m) \\ &= \text{Var}(-j_A s_2 - k_A s_2 h - k_{mX} s_3) + \text{Var}(-\varepsilon_m) \\ &= (1 + C_m) \text{Var}(-j_A s_2 - k_A s_2 h - k_{mX} s_3) \\ &= (1 + C_m) [s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_3^2 \text{Var}(k_{mX})] \end{aligned} \quad (\text{S1})$$

#### *Variance among fathers (with Z-linkage):*

$$\begin{aligned} \text{Var}[\ln(w_m)] &= \text{Var}(-j_A s_2 - k_A s_2 h - j_Z s_2 - k_Z s_2 h - \varepsilon_m) \\ &= (1 + C_m) [s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_2^2 \text{Var}(j_Z) + s_2^2 h^2 \text{Var}(k_Z)] \end{aligned} \quad (\text{S2})$$

#### *Variance among fathers (no heteromorphic sex chromosomes):*

$$\begin{aligned} \text{Var}[\ln(w_m)] &= \text{Var}(-j_A s_2 - k_A s_2 h - \varepsilon_m) \\ &= (1 + C_m) [s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A)] \end{aligned} \quad (\text{S3})$$

What effect will GxE interactions have on the regression coefficients between fathers and offspring? Consider the case where there is no fixed or GxE environmental variance.

Using equations (6) and (8) from the text, the regression coefficients for species with an X, with a Z, and without heteromorphic sex chromosomes (respectively) are:

$$\beta_X = \frac{\text{Cov}_X(w_m, E[w_O|w_m])}{\text{Var}_X(w_m)} = \frac{\text{Cov}_X(w_m, E[w_O|w_m])}{s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_3^2 \text{Var}(k_{mX})}$$

$$\beta_Z = \frac{\text{Cov}_Z(w_m, E[w_O|w_m])}{\text{Var}_Z(w_m)} = \frac{\text{Cov}_Z(w_m, E[w_O|w_m])}{s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_2^2 \text{Var}(j_Z) + s_2^2 h^2 \text{Var}(k_Z)}$$

$$\beta_{\text{none}} = \frac{\text{Cov}_{\text{none}}(w_m, E[w_O|w_m])}{\text{Var}_{\text{none}}(w_m)} = \frac{\text{Cov}_{\text{none}}(w_m, E[w_O|w_m])}{s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A)}$$

By incorporating GxE interactions into the paternal variance – equations (S1), (S2), and (S3), respectively – the denominator is modified by a factor of  $(1 + C_m)$ . This reduces each species' regression coefficient, but the reduction is the same across species. Thus, the relative magnitude of regression coefficients remains unchanged between species.

Suppose there are both GxE and fixed environmental effects. Since the GxE fitness variance is a function of the genotype and is independent of fixed environmental effects, the regression coefficients can be modified to:

$$\beta_X = \frac{\text{Cov}_X(w_m, E[w_O|w_m])}{(1 + C_m)[s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_3^2 \text{Var}(k_{mX})] + \text{Var}(\varepsilon_{\text{fixed}})}$$

$$\beta_Z = \frac{\text{Cov}_Z(w_m, E[w_O|w_m])}{(1 + C_m)[s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A) + s_2^2 \text{Var}(j_Z) + s_2^2 h^2 \text{Var}(k_Z)] + \text{Var}(\varepsilon_{\text{fixed}})}$$

$$\beta_{\text{none}} = \frac{\text{Cov}_{\text{none}}(w_m, E[w_O|w_m])}{(1 + C_m)[s_2^2 \text{Var}(j_A) + s_2^2 h^2 \text{Var}(k_A)] + \text{Var}(\varepsilon_{\text{fixed}})}$$

Again, environmental variability (GxE as a function of  $C_m$ , and fixed effects as a function of  $Var(\epsilon_{fixed})$ ) reduces the regression coefficients. To the extent that GxE variance is large relative to the fixed environmental variance, the relative magnitude of the regression coefficients remains constant between species with different sex chromosome systems.

### Appendix 3.4. Variation maintained by sexually antagonistic selection

#### I. Additive Allele Model

The conditions that lead to a stable, sexually antagonistic polymorphism depend on the relative magnitude of female relative to male selection coefficients, and are therefore unaffected by dosage compensation. Under moderate to weak selection ( $s_i < 0.1$ ), polymorphic equilibria are stable when  $t_1 \approx t_2 \approx t_3 \approx t_4$ , and converge to frequency  $\hat{q}_{SA} \approx 0.5$ . We can therefore alter the sexually antagonistic genotypic fitness model as follows:

Genotype:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness:	$1 + s_f$	$1 + s_f/2$	1
Male fitness:	$1 - s_m$	$1 - s_m/2$	1

for diploid loci (autosomal or sex-linked in the homogametic sex), and

Genotype:	$A_1$	$A_2$
X-linked male fitness:	$1 - s_m$	1
Z-linked female fitness:	$1 + s_f$	1

where  $s_m = t_1 = t_2 = t_3 = t_4$  and  $s_f = \frac{t_1}{1-t_1} = \frac{t_2}{1-t_2} = \frac{t_3}{1-t_3} = \frac{t_4}{1-t_4}$

Extending to a multilocus model with multiplicative epistasis between loci yields the fitness expressions for sexually antagonistic loci:

$$w_m = \exp(-j_{*A}s_m - k_{*A}s_m/2 - k_{*mX}s_m)$$

$$w_f = \exp(j_{*A}s_f + k_{*A}s_f/2 + j_{*X}s_f + k_{*X}s_f/2)$$

under X-linked sex determination, and:

$$w_m = \exp(-j_{*A}s_m - k_{*A}s_m/2 - j_{*Z}s_m - k_{*Z}s_m/2)$$

$$w_f = \exp(j_{*A} s_f + k_{*A} s_f / 2 + k_{*Z} s_f)$$

under Z-linked sex determination. The sexual antagonism variables are defined as follows:

- $j_{*A}$  and  $k_{*A}$  refer to the number of autosomal sexually antagonistic loci with genotype  $A_1A_1$  and  $A_2A_2$  (respectively) and follow a multinomial distribution with  $Var(j_{*A}) = L_A \hat{q}_{SA}^2 (1 - \hat{q}_{SA}^2) = \frac{3L_A}{16}$ ,  $Var(k_{*A}) = L_A 2\hat{q}_{SA}(1 - \hat{q}_{SA})[1 - 2\hat{q}_{SA}(1 - \hat{q}_{SA})] = \frac{L_A}{4}$ , and  $Cov(j_{*A}, k_{*A}) = -L_A 2\hat{q}_{SA}(1 - \hat{q}_{SA})\hat{q}_{SA}^2 = -\frac{L_A}{8}$ , where  $L_A$  is the number of autosomal sexually antagonistic loci.
- $j_{*X}$  and  $k_{*X}$  refer to the number of sex-linked sexually antagonistic loci with genotype  $A_1A_1$  and  $A_2A_2$  (respectively) and follow a multinomial distribution with  $Var(j_{*X}) = L_X \hat{q}_{SA}^2 (1 - \hat{q}_{SA}^2) = \frac{3L_X}{16}$ ,  $Var(k_{*X}) = L_X 2\hat{q}_{SA}(1 - \hat{q}_{SA})[1 - 2\hat{q}_{SA}(1 - \hat{q}_{SA})] = \frac{L_X}{4}$ , and  $Cov(j_{*X}, k_{*X}) = -L_X 2\hat{q}_{SA}(1 - \hat{q}_{SA})\hat{q}_{SA}^2 = -\frac{L_X}{8}$ , where  $L_X$  is the number of X-linked sexually antagonistic loci. (The same results apply for homozygous and heterozygous loci on the Z chromosome:  $j_{*Z}$  and  $k_{*Z}$ ).
- $k_{*mX}$  and  $k_{*fZ}$  refer to the number of sex-linked sexually antagonistic loci with genotype  $A_1$  (in heterogametic males and females, respectively) and follow binomial distributions with  $Var(k_{*mX}) = L_X \hat{q}_{SA}(1 - \hat{q}_{SA}) = \frac{L_X}{4}$ , and  $Var(k_{*fZ}) = L_Z \hat{q}_{SA}(1 - \hat{q}_{SA}) = \frac{L_Z}{4}$ .
- Assuming that  $N$  sexually antagonistic loci are randomly distributed throughout the genome,  $L_A = N(1 - P)$  and  $L_X = L_Z = NP$  (as before,  $P$  is the proportion of the genome that is sex-linked)

## II. X-linked sex determination.

### *Sexually antagonistic fitness scheme:*

$$w_{father} = \exp(-s_m j_{*A} - s_m k_{*A} / 2 - s_m k_{*mX})$$

$$E[w_{sons} | j_{*A}, k_{*A}, k_{*mX}] = \bar{w}_{sons} = \exp[-s_m j_{*A}^s - s_m k_{*A}^s / 2 - s_m k_{*mX}^s]$$

$$E[w_{daughters} | j_{*A}, k_{*A}, k_{*mX}] = \bar{w}_{daughters} = \exp[s_f j_{*A}^d + s_f k_{*A}^d / 2 + s_f j_{*X}^d + s_f k_{*X}^d / 2]$$

The offspring terms are defined as:

$$j_{*A}^s = j_{*A}^d = (j_A + k_A / 2) \hat{q}_A = (j_A + k_A / 2) / 2$$

$$k_{*A}^s = k_{*A}^d = j_A + k_A / 2 + L_A \hat{q}_A - 2(j_A + k_A / 2) \hat{q}_A = L_A / 2$$

$$k_{*mX}^s = L_X \hat{q}_X = L_X / 2$$

$$j_{*X}^d = k_{mX} \hat{q}_X = k_{mX} / 2$$

$$k_{*X}^d = k_{mX} (1 - \hat{q}_X) = k_{mX} / 2$$

Substitution and rearrangement yields:

$$\bar{w}_{sons} = \exp\left[-\frac{s_m}{2} (j_A + k_A / 2) - s_m L_A / 4 - s_m L_X / 2\right]$$

$$\bar{w}_{daughters} = \exp\left[\frac{s_f}{2} (j_A + k_A / 2) + s_f L_A / 4 + 3s_f k_{mX} / 4\right]$$

**Covariance:**

The covariance between fathers and sons is:

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{sons})] &= \text{Cov}\left[-s_m j_A - s_m k_A / 2, -\frac{s_m}{2} j_A - \frac{s_m}{4} k_A\right] \\ &= \text{Cov}(s_m j_A, \frac{s_m}{2} j_A) + \text{Cov}(s_m j_A, \frac{s_m}{4} k_A) + \text{Cov}(\frac{s_m}{2} k_A, \frac{s_m}{2} j_A) + \text{Cov}(\frac{s_m}{2} k_A, \frac{s_m}{4} k_A) \\ &= \frac{s_m^2}{2} \text{Var}(j_A) + \frac{s_m^2}{8} \text{Var}(k_A) - \frac{s_m^2 L_A}{16} \\ &= \frac{s_m^2 N(1-P)}{16} \end{aligned}$$

The covariance between fathers and daughters is:

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{daughters})] &= \text{Cov}\left[-s_m j_A - \frac{s_m}{2} k_A - s_m k_{mX}, \frac{s_f}{2} j_A + \frac{s_f}{4} k_A + \frac{3s_f}{4} k_{mX}\right] \\ &= \text{Cov}(-s_m j_A, \frac{s_f}{2} j_A) + \text{Cov}(-s_m j_A, \frac{s_f}{4} k_A) + \text{Cov}(-\frac{s_m}{2} k_A, \frac{s_f}{2} j_A) + \text{Cov}(-\frac{s_m}{2} k_A, \frac{s_f}{4} k_A) + \text{Cov}(-s_m k_{mX}, \frac{3s_f}{4} k_{mX}) \\ &= -\frac{s_m s_f}{2} \text{Var}(j_A) - \frac{s_m s_f}{8} \text{Var}(k_A) - \frac{3s_m s_f}{4} \text{Var}(k_{mX}) + \frac{s_m s_f L_A}{16} \\ &= -\frac{s_m s_f L_A}{16} - \frac{3s_m s_f L_X}{16} = -\frac{s_m s_f N(1+2P)}{16} \end{aligned}$$

**Variance among fathers:**

$$\begin{aligned}
 \text{Var}[\ln(w_{\text{father}})] &= \text{Var}(-s_m j_A - s_m k_A / 2) + \text{Var}(-s_m k_{mX}) \\
 &= s_m^2 \text{Var}(j_A) + \frac{s_m^2}{4} \text{Var}(k_A) + s_m^2 \text{Var}(k_{mX}) + s_m^2 \text{Cov}(j_A, k_A) \\
 &= s_m^2 \text{Var}(j_A) + \frac{s_m^2}{4} \text{Var}(k_A) + s_m^2 \text{Var}(k_{mX}) - \frac{s_m^2 L_A}{8} \\
 &= \frac{2s_m^2 L_A}{16} + \frac{s_m^2 L_X}{4} = \frac{s_m^2 N(1+P)}{8}
 \end{aligned}$$

**III. Z-linked sex determination.**

**Sexually antagonistic fitness scheme:**

$$\begin{aligned}
 w_{\text{father}} &= \exp(-s_m j_A - s_m k_A / 2 - s_m j_Z - s_m k_Z / 2) \\
 E[w_{\text{sons}} | j_A, k_A, j_Z, k_Z] &= \bar{w}_{\text{sons}} = \exp[-s_m i_{*A}^s - s_m k_{*A}^s / 2 - s_m i_{*Z}^s - s_m k_{*Z}^s / 2] \\
 E[w_{\text{daughters}} | j_A, k_A, j_Z, k_Z] &= \bar{w}_{\text{daughters}} = \exp[s_f j_{*A}^d + s_f k_{*A}^d / 2 + s_f k_{*Z}^d]
 \end{aligned}$$

The offspring terms are defined as:

$$\begin{aligned}
 i_{*A}^s &= j_{*A}^d = (j_A + k_A / 2) \hat{q}_A = (j_A + k_A / 2) / 2 \\
 k_{*A}^s &= k_{*A}^d = j_A + k_A / 2 + L_A \hat{q}_A - 2(j_A + k_A / 2) \hat{q}_A = L_A / 2 \\
 k_{*Z}^d &= j_Z + k_Z / 2 \\
 i_{*Z}^s &= (j_Z + k_Z / 2) \hat{q}_Z = (j_Z + k_Z / 2) / 2 \\
 k_{*Z}^s &= L_Z / 2
 \end{aligned}$$

Substitution and rearrangement yields:

$$\bar{w}_{sons} = \exp\left[-\frac{s_m}{2} j_A - \frac{s_m}{4} k_A - \frac{s_m}{4} L_A - \frac{s_m}{2} j_Z - \frac{s_m}{4} k_Z - \frac{s_m}{4} L_Z\right]$$

$$\bar{w}_{daughters} = \exp\left[\frac{s_f}{2} j_A + \frac{s_f}{4} k_A + \frac{s_f}{4} L_A + s_f j_Z + \frac{s_f}{2} k_Z\right]$$

**Covariance:**

The covariance between fathers and sons is:

$$\begin{aligned} \text{Cov}[\ln(w_{father}), \ln(\bar{w}_{sons})] &= \text{Cov}\left[-s_m j_A - \frac{s_m}{2} k_A - s_m j_Z - \frac{s_m}{2} k_Z, -\frac{s_m}{2} j_A - \frac{s_m}{4} k_A - \frac{s_m}{2} j_Z - \frac{s_m}{4} k_Z\right] \\ &= \text{Cov}\left(-s_m j_A, -\frac{s_m}{2} j_A\right) + \text{Cov}\left(-s_m j_A, -\frac{s_m}{4} k_A\right) + \text{Cov}\left(-\frac{s_m}{2} k_A, -\frac{s_m}{2} j_A\right) + \text{Cov}\left(-\frac{s_m}{2} k_A, -\frac{s_m}{4} k_A\right) + \\ &\quad \text{Cov}\left(-s_m j_Z, -\frac{s_m}{2} j_Z\right) + \text{Cov}\left(-s_m j_Z, -\frac{s_m}{4} k_Z\right) + \text{Cov}\left(-\frac{s_m}{2} k_Z, -\frac{s_m}{2} j_Z\right) + \text{Cov}\left(-\frac{s_m}{2} k_Z, -\frac{s_m}{4} k_Z\right) \\ &= \frac{s_m^2}{2} \text{Var}(j_A) + \frac{s_m^2}{8} \text{Var}(k_A) + \frac{s_m^2}{2} \text{Var}(j_Z) + \frac{s_m^2}{8} \text{Var}(k_Z) - \frac{s_m^2 L_Z}{16} - \frac{s_m^2 L_A}{16} \\ &= \frac{s_m^2 L_A}{16} + \frac{s_m^2 L_Z}{16} = \frac{s_m^2 N}{16} \end{aligned}$$

The covariance between fathers and daughters is:

$$\begin{aligned}
\text{Cov}[\ln(w_{father}), \ln(\bar{w}_{daughters})] &= \text{Cov}\left[-s_m j_A - \frac{s_m}{2} k_A - s_m j_Z - \frac{s_m}{2} k_Z, \frac{s_f}{2} j_A + \frac{s_f}{4} k_A + s_f j_Z + \frac{s_f}{2} k_Z\right] \\
&= \text{Cov}\left(-s_m j_A, \frac{s_f}{2} j_A\right) + \text{Cov}\left(-s_m j_A, \frac{s_f}{4} k_A\right) + \text{Cov}\left(-\frac{s_m}{2} k_A, \frac{s_f}{2} j_A\right) + \text{Cov}\left(-\frac{s_m}{2} k_A, \frac{s_f}{4} k_A\right) + \\
&\quad \text{Cov}\left(-s_m j_Z, s_f j_Z\right) + \text{Cov}\left(-s_m j_Z, \frac{s_f}{2} k_Z\right) + \text{Cov}\left(-\frac{s_m}{2} k_Z, s_f j_Z\right) + \text{Cov}\left(-\frac{s_m}{2} k_Z, \frac{s_f}{2} k_Z\right) \\
&= -\frac{s_m s_f}{2} \text{Var}(j_A) - \frac{s_m s_f}{8} \text{Var}(k_A) - s_m s_f \text{Var}(j_Z) - \frac{s_m s_f}{4} \text{Var}(k_Z) + \frac{s_m s_f L_Z}{8} + \frac{s_m s_f L_A}{16} \\
&= -\frac{s_m s_f L_A}{16} - \frac{s_m s_f L_Z}{8} = -\frac{s_m s_f N(1+P)}{16}
\end{aligned}$$

**Variance among fathers:**

$$\begin{aligned}
\text{Var}[\ln(w_{father})] &= \text{Var}\left(-s_m j_A - s_m k_A / 2\right) + \text{Var}\left(-s_m j_Z - s_m k_Z / 2\right) \\
&= s_m^2 \text{Var}(j_A) + \frac{s_m^2}{4} \text{Var}(k_A) + s_m^2 \text{Var}(j_Z) + \frac{s_m^2}{4} \text{Var}(k_Z) - \frac{s_m^2 L_Z}{8} - \frac{s_m^2 L_A}{8} \\
&= \frac{2s_m^2 L_A}{16} + \frac{2s_m^2 L_Z}{16} = \frac{s_m^2 N}{8}
\end{aligned}$$

**Appendix 3.5. Consequences of sex-specific mutation and selection on the relative frequency of deleterious mutations on sex chromosomes and autosomes.**

Consider the case where there is dosage compensation (similar results can be obtained without dosage compensation, but this assumption removes two parameters, since  $s_2 = s_3$  and  $s_1 = s_4$ ). The frequencies of autosomal, X-linked, and Z-linked mutations are:

$$\hat{q}_A \approx \frac{\mu_m + \mu_f}{h(s_1 + s_2)} = \frac{\mu_f(1 + \alpha_u)}{hs_1(1 + \alpha_s)}$$

$$\hat{q}_X \approx \frac{2\mu_f + \mu_m}{2s_1h + s_3} = \frac{\mu_f(2 + \alpha_u)}{s_1(2h + \alpha_s)}$$

and

$$\hat{q}_Z \approx \frac{2\mu_m + \mu_f}{2s_2h + s_4} = \frac{\mu_f(2\alpha_u + 1)}{s_1(2h\alpha_s + 1)}$$

where  $\alpha_s = s_2/s_1$  and  $\alpha_u = \mu_m/\mu_f$ .

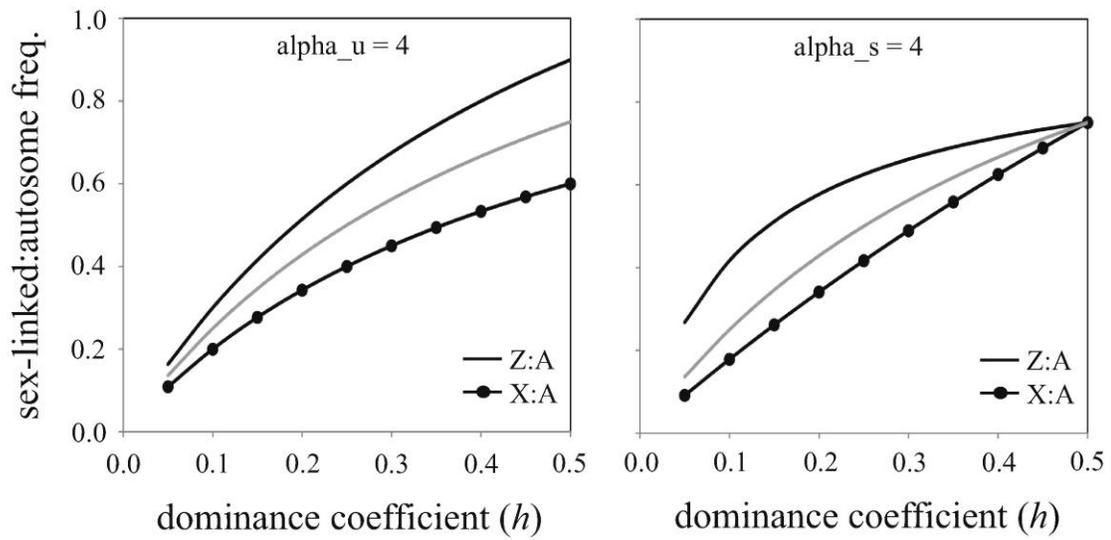
The relative frequency of sex-linked compared to autosomal mutations is:

$$\frac{\hat{q}_X}{\hat{q}_A} = \frac{h(2 + \alpha_u)(1 + \alpha_s)}{(1 + \alpha_u)(2h + \alpha_s)}$$

and

$$\frac{\hat{q}_Z}{\hat{q}_A} = \frac{h(2\alpha_u + 1)(1 + \alpha_s)}{(1 + \alpha_u)(2h\alpha_s + 1)}$$

Under a range of dominance, sex-biased mutation ( $\alpha_u > 1$ ) and sex-biased selection ( $\alpha_s > 1$ ), increase the relative frequencies of Z-linked mutations, and decrease the frequencies of X-linked mutations:



**Figure A3.5.1.** Frequencies of X-linked, Z-linked and autosomal deleterious mutations are influenced by male-biased mutation rate and selection strength. The gray lines show the condition where mutation rates and selection coefficients are exactly equal between males and females (i.e.  $\alpha_u = \alpha_s = 1$ ).

**Appendix 3.6. Data from the meta-analysis**

Family	Genus	Species	Common name	Sex chromosome		Ref.
				system	$r$	
Bufonidae	<i>Bufo</i>	<i>americanus</i>	American toad	undifferentiated	0.147	[1,2]
			southern			
Pelobatidae	<i>Scaphiopus</i>	<i>multiplacatus</i>	spadefoot toad	undifferentiated	0.036	[2,6]
Hylidae	<i>Hyla</i>	<i>versicolor</i>	gray treefrog	undifferentiated	0.323	[2,5]
			yellow-banded			
Dendrobatidae	<i>Dendrobates</i>	<i>leucomelas</i>	poison frog	undifferentiated	0.765	[4]
			phantasmal			
Dendrobatidae	<i>Epipedobates</i>	<i>tricolor</i>	poison frog	undifferentiated	0.95	[4]
			southwestern			
Bufonidae	<i>Bufo</i>	<i>woodhousei</i>	Woodhouse's toad	undifferentiated	0.279	[2,3]
			Couch's spadefoot			
Pelobatidae	<i>Scaphiopus</i>	<i>couchi</i>	toad	undifferentiated	0.464	[2,7]
Poeciliidae	<i>Poecilia</i>	<i>reticulata</i>	guppy	X-linked	-0.364	[18]

Cricetidae	<i>Clethrionomys glareolus</i>	bank vole	X-linked	-0.304	[2,10]
Salmonidae	<i>Salmo trutta</i>	brown trout	X-linked	-0.175	[19]
Pseudostigmatidae	<i>Megaloprepus coerulatus</i>	giant damselfly	X-linked	-0.03	[17]
	<i>undecimpunctata</i>	spotted cucumber			
Chrysomelidae	<i>Diabrotica howardi</i>	beetle	X-linked	0.287	[12]
		southern ground			
Gryllidae	<i>Allonemobius socius</i>	cricket	X-linked	0.03	[8]
Lycosidae	<i>Hygrolycosa rubrofasciata</i>	wolf spider	X-linked	0.143	[2,16]
Coelopidae	<i>Coelopa frigida</i>	seaweed fly	X-linked	0.301	[2,11]
Antilocapridae	<i>Antilocapra americana</i>	pronghorn	X-linked	0.424	[9]
Tenebrionidae	<i>Tribolium castaneum</i>	red flour beetle	X-linked	-0.045	[21]
Sepsidae	<i>Sepsis cynipsea</i>	dung fly	X-linked	0.1	[2,20]
Drosophilidae	<i>Drosophila montana</i>	fruit fly	X-linked	0.173	[2,13]
Chrysomelidae	<i>Galerucella nymphaeae</i>	water-lily beetle	X-linked	-0.272	[2,14]
		two-spotted			
Gryllidae	<i>Gryllus bimaculatus</i>	cricket	X-linked	0.061	[2,15]

Sylviidae	<i>Acrocephalus arundinaceus</i>	great reed warbler	Z-linked	0.175	[2,22,23]
Phasianidae	<i>Pavo cristatus</i>	Indian peafowl	Z-linked	0.79	[2,28]
Muscicapidae	<i>Ficedula albicollis</i>	collared flycatcher	Z-linked	0.075	[2,24]
Paridae	<i>Parus caeruleus</i>	blue tit	Z-linked	0.272	[2,26]
Hirundinidae	<i>Hirundo rustica</i>	barn swallow	Z-linked	0.487	[2,25]
Phasianidae	<i>Phasianus colchicus</i>	common pheasant	Z-linked	0.39	[2,29]
Paridae	<i>Parus major</i>	great tit	Z-linked	0.493	[2,27]

---

[1] Howard et al (1994) *Evolution* 48:1286-1300;

[2] Møller & Alatalo (1999) *Proc Roy Soc Lond B* 266:85-91;

[3] Mitchell (1990) *Evolution* 44:502-519;

[4] Forsman & Hagman (2006) *Evolution* 2148-2157;

[5] Welch et al (1998) *Science* 280:1928-1930;

[6] Woodward (1986) *Am Nat* 128:58-65;

[7] Woodward (1987) *Oecologia* 73:626-629;

[8] Fedorka & Mousseau (2004) *Nature* 429:65-67;

[9] Byers & Waits (2006) *Proc Natl Acad Sci USA* 103:16343-16345;

[10] Horne (1998) *Evolution of female choice in the bank vole*. PhD thesis, University of Jyväskylä, Finland;

[11] Crocker & Day (1987) *Behav Ecol Sociobiol.* 20:295-301;

- 
- [12] Tallamy et al (2003) *Proc Roy Soc Lond B* 270:77-82;
- [13] Hoikkala et al (1998) *Proc Roy Soc Lond B* 265:503-508;
- [14] Parri et al (1998) *Oecologia* 114:127-132;
- [15] Simmons (1987) *Behav Ecol Sociobiol.* 21:313-322;
- [16] Alatalo et al (1998) *Proc Roy Soc Lond B* 265:2297-2301;
- [17] Fincke and Hadrys (2001) *Evolution* 55:762-772;
- [18] Brooks (2000) *Nature* 406:67-70;
- [19] Jacob et al (2007) *BMC Evol Biol* 7:207;
- [20] Blanckenhorn et al (1998) *J Evol Biol* 11:735-753;
- [21] Edvardsson & Arnqvist (2006) *Behavior Genet* 36:775-782;
- [22] Hasselquist et al (1996) *Nature* 381:229-232;
- [23] Forstmeier & Leisler (2004) *Behav Ecol* 15:555-563;
- [24] Gustafsson et al (1995) *Nature* 375:311-313;
- [25] Møller (1994) *Proc Natl Acad Sci USA* 91:6929-6932;
- [26] Kempnaers et al (1992) *Nature* 357:494-496;
- [27] Norris (1993) *Nature* 362:537-539;
- [28] Petrie (1994) *Nature* 371:598-599;
- [29] von Schantz et al (1998) *Nature* 337:166-169.

## REFERENCES

- Agrawal AF. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature* 411:692-695.
- Albert AYK, Otto SP. 2005. Sexual selection can resolve sex-linked sexual antagonism. *Science* 310:119-121.
- Andersson, M. 1994. *Sexual selection*. Princeton University Press, Princeton.
- Baer CF, Miyamoto MM, Denver D. 2007. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Reviews Genetics* 8:619-631.
- Bateman AJ. 1948. Intra-sexual selection in *Drosophila*. *Heredity* 2:349-368.
- Berglund A, Bisazza A, Pilastro A. 1996. Armaments and ornaments: an evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society* 58:385-399.
- Boake CRB. 1985. Genetic consequences of mate choice: a quantitative genetic method for testing sexual selection theory. *Science* 227:1061-1063.
- Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. *Trends in Ecology & Evolution* 24:280-288.
- Borgia G. 1979. Sexual selection and the evolution of mating systems. Pages 19-80 in M. Blum and A. Blum, eds. *Sexual Selection and Reproductive Competition*. Academic Press, New York.
- Brommer JE, Kirkpatrick M, Qvarnström A, Gustafsson L. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. *PLoS ONE* 2:e744.
- Bull JJ. 1983. *Evolution of sex determining mechanisms*. Benjamin Cummings, Menlo Park.
- Candolin U, Heuschele J. 2008. Is sexual selection beneficial during adaptation to environmental change? *Trends in Ecology & Evolution* 23:446-452.
- Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist* 130:113-146.
- Charlesworth B, Charlesworth D. 1999. The genetic basis of inbreeding depression. *Genetical Research* 74:329-340.

- Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proceedings of the National Academy of Sciences USA 98:1671-1675.
- Clark AG, Begun DJ, Prout T. 1999. Female x male interactions in *Drosophila* sperm competition. Science 283:217-220.
- Connallon T, Jakubowski E. 2009. Association between sex ratio distortion and sexually antagonistic fitness consequences of female choice. Evolution, in press.
- Cowley DE, Atchley WR. 1988. Quantitative genetics of *Drosophila melanogaster*. II. Heritabilities and genetic correlations between sexes for head and thorax traits. Genetics 119:421-433.
- Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. American Naturalist 173:176-187.
- Delcourt M, Blows MW, Rundle HD. 2009. Sexually antagonistic genetic variance for fitness in an ancestral and a novel environment. Proceedings of the Royal Society, B 276:2009-2014.
- Devlin RH, Nagahama Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208:191-364.
- Ellegren H. 2007. Characteristics, causes and evolutionary consequences of male-biased mutation. Proceedings of the Royal Society, B 274:1-10.
- Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nature Reviews Genetics 8:689-698.
- Fairbairn DJ, Roff DA. 2006. The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity 97:319-328.
- Fedorka KM, Mousseau TA. 2004. Female mating bias results in conflicting sex-specific offspring fitness. Nature 429:65-67.
- Fisher RA. 1915. The evolution of sexual preference. Eugenics Review 7:184-192.
- Fitzpatrick MJ. 2004. Pleiotropy and the genomic location of sexual selected genes. American Naturalist 163:800-808.

- Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH, Kruuk LEB. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447:1107-1110.
- Forsman A, Hagman M. 2006. Calling is an honest indicator of paternal genetic quality in poison frogs. *Evolution* 60:2148-2157.
- Gillespie J. 2004. Population genetics: A concise guide. 2<sup>nd</sup> Ed. Johns Hopkins University Press, Baltimore.
- Gong Y, Woodruff RC, Thompson JR, Jr. 2005. Deleterious genomic mutation rate for viability in *Drosophila melanogaster* using concomitant sibling controls. *Biology Letters* 1:492-495.
- Haag-Liautard C, Dorris M, Maside X, Macaskill S, Halligan DL, Charlesworth B, Keightley PD. 2007. Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* 445:82-85.
- Haldane JBS. 1937. The effect of variation on fitness. *American Naturalist* 71:337-349.
- Hastings IM. 1994. Manifestations of sexual selection may depend on the genetic basis of sex determination. *Proceedings of the Royal Society of London B* 258:83-87.
- Head ML, Hunt J, Jennions MD, Brooks R. 2005. The indirect benefits of mating with attractive males outweigh the direct costs. *PLoS Biology* 3:e33.
- Hedrick PW. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution* 61:2750-2771.
- Hedrick PW, Parker JD. 1997. Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annual Review of Ecology and Systematics* 28:55-83.
- Hollis B, Fierst JL, Houle D. 2009. Sexual selection accelerate the elimination of a deleterious mutant in *Drosophila melanogaster*. *Evolution* 63:324-333.
- Houle D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130:195-204.
- Houle D, Morikawa B, Lynch M. 1996. Comparing mutational variabilities. *Genetics* 143:1467-1483.
- James TW. 1973. Covariances between relatives due to sex-linked genes. *Biometrics* 29:584-588.

- Jennions MD, A. P. Møller AP, Petrie M. 2001. Sexually selected traits and adult survival: a meta-analysis. *Quarterly Review of Biology* 76:3-36.
- Jones TM, Quinnell RJ, Balmford A. 1998. Fisherian flies: benefits of female choice in a lekking sandfly. *Proceedings of the Royal Society of London, B* 265:1651-1657.
- Kidwell JF, Clegg MT, Stewart FM, Prout T. 1977. Regions of stable equilibria for models of differential selection in two sexes under random mating. *Genetics* 85:171-183.
- Kirkpatrick M. 2009. Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136:271-284.
- Kirkpatrick M, Hall DH. 2004. Sexual selection and sex linkage. *Evolution* 58:683-691.
- Kodric-Brown A, Brown JH. 1987. Anisogamy, sexual selection, and the origin and maintenance of sex. *Evolutionary Ecology* 1:95-105.
- Kokko H, Brooks R, Jennions M, Morley J. 2003. The evolution of mate choice and mating biases. *Proceedings of the Royal Society of London B* 270:653-664.
- Kotiaho JS, Lebas NR, Puurtinen M, Tomkins JL. 2008. On the resolution of the lek paradox. *Trends in Ecology & Evolution* 23:1-3.
- Kotiaho JS, Puurtinen M. 2007. Mate choice for indirect genetic benefits: scrutiny of the current paradigm. *Functional Ecology* 21:638-644.
- Kruuk LEB, Slate J, Wilson AJ. 2008. New answers to old questions: the evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology, Evolution, and Systematics* 39:525-548.
- Lewontin RC. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Lewontin RC, Moore JA, Provine WB, Wallace B. 1981. *Dobzhansky's Genetics of Natural Populations I-XLIII*. Columbia University Press, New York.
- Lindsley DL, Lifschytz E. 1972. The genetic control of spermatogenesis in *Drosophila*. Pages 203-222 *in* R. A. Beatty and S. Gluecksohn-Waelsh, eds. *Proceedings of the International Symposium on the Genetics of the Spermatozoan*. Bogtrykkeriet Forum, Copenhagen.

- Long TAF, Rice WR. 2007. Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of *Drosophila melanogaster*. *Proceedings of the Royal Society, B* 274:3105-3112.
- Lynch M, Walsh JB. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc., Sunderland, MA.
- Mackay TFC. 2001. The genetic architecture of quantitative traits. *Annual Review of Genetics* 35:303-339.
- Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, in press.
- Mank JE. 2009a. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *American Naturalist* 173:141-150.
- Mank JE. 2009b. The W, X, Y and Z of sex-chromosome dosage compensation. *Trends in Genetics* 25:226-233.
- Mank JE, Hall DW, Kirkpatrick M, Avise JC. 2006. Sex chromosomes and male ornaments: a comparative evaluation in ray-finned fishes. *Proceedings of the Royal Society of London B* 273:233-236.
- Mitchell-Olds T, Willis JH, Goldstein DB. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* 8:845-856.
- Møller AP, Alatalo RV. 1999. Good genes effects in sexual selection. *Proceedings of the Royal Society of London B* 266:85-91.
- Morrow EH, Stewart AD, Rice WR. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *Journal of Evolutionary Biology* 21:1046-1054.
- Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology* 14:19-38.
- Oneal E, Connallon T, Knowles LL. 2007. Conflict between direct and indirect benefits of female choice in desert *Drosophila*. *Biology Letters* 3:29-32.
- Organ L, Janes DE. 2008. Evolution of sex chromosomes in Sauropsida. *Integrative and Comparative Biology* 48:512-519.
- Pamilo P. 1979. Genetic variation at sex-linked loci: quantification of regular selection models. *Hereditas* 91:129-133.

- Partridge L. 1980. Mate choice increases a component of offspring fitness in fruit flies. *Nature* 283:290-291.
- Patten MM, Haig D. 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution*, in press.
- Petrie M. 1994. Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* 371:598-599.
- Pischedda A, Chippindale AK. 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. *PLoS Biology* 4:2099-2103.
- Pitnick S. 1991. Male size influences mate fecundity and remating interval in *Drosophila melanogaster*. *Animal Behaviour* 41:735-745.
- Pomiankowski A, Møller AP. 1995. A resolution of the lek paradox. *Proceedings of the Royal Society of London B* 26:21-29.
- Prasad NG, Bedhomme S, Day T, Chippindale AK. 2007. An evolutionary cost of separate genders revealed by male-limited evolution. *American Naturalist* 169:29-37.
- Presgraves DC, Orr HA. 1998. Haldane's rule in taxa lacking a hemizygous X. *Science* 282:952-954.
- Prout T. 1999. How well does opposing selection maintain variation? Pages 157-181 in R. S. Singh and C. B. Krimbas, eds. *Evolutionary Genetics: from Molecules to Morphology*. Cambridge University Press, UK.
- Qvarnström A, Brommer JE, Gustafsson L. 2006. Testing the genetics underlying the coevolution of mate choice and ornament in the wild. *Nature* 441:84-86.
- Reeve HK, Pfennig DW. 2003. genetic biases for showy males: are some genetic systems especially conducive to sexual selection? *Proceedings of the National Academy of Sciences USA* 100:1089-1094.
- Reinhold K. 1998. Sex linkage among genes controlling sexually selected traits. *Behavioral Ecology and Sociobiology* 44:1-7.
- Reynolds JD, Gross MR. 1992. Female mate preference enhances offspring growth and reproduction in a fish, *Poecilia reticulata*. *Proceedings of the Royal Society of London, B* 250:57-62.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735-742.

- Rice WR. 1988. Heritable variation in fitness as a prerequisite for adaptive female choice: the effect of mutation-selection balance. *Evolution* 42:817-820.
- Rowe L, Houle D. 1996. The lek paradox and the capture of genetic variance by condition-dependent traits. *Proceedings of the Royal Society of London B* 263:1416-1421.
- Rosenthal R. 1994. Parametric measures of effect size. Pages 231-244 in H. Cooper, and L. V. Hedges, eds. *The handbook of research synthesis*. Russel Sage Foundation, New York.
- Schaeffer SW, Brown CJ, Andersson WW. 1984. Does mate choice affect fitness? *Genetics* 107:s94.
- Shabalina SA, Yu L, Kondrashov AS. 1997. Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed selection. *Proceedings of the National Academy of Sciences USA* 94:13034-13039.
- Shnol EE, Kondrashov AS. 1993. The effect of selection on the phenotypic variance. *Genetics* 134:995-996.
- Siller S. 2001. Sexual selection and the maintenance of sex. *Nature* 411:689-692.
- Solari AJ. 1994. *Sex chromosomes and sex determination in vertebrates*. CRC Press, Boca Raton.
- Taylor CE, Pereda AD, Ferrari JA. 1987. On the correlation between mating success and offspring quality in *Drosophila melanogaster*. *American Naturalist* 129:721-729.
- Taylor ML, Wedell N, Hosken DJ. 2007. The heritability of attractiveness. *Current Biology* 17:R959-960.
- Trivers RL. 1972. Parental investment and sexual selection. Pages 136-179 in B. Campbell, ed. *Sexual selection and the descent of man, 1871-1971*. Aldine, Chicago.
- Wayne ML, Telonis-Scott M, Bono LM, Harshman L, Kopp A, Nuzhdin SV, McIntyre LM. 2007. Simpler mode of inheritance of transcriptional variation in male *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA*. 104:18577-18582.
- Wedell N, Tregenza T. 1999. Successful fathers sire successful sons. *Evolution* 53:620-625.

- White MJD. 1973. *Animal cytology and evolution*, 3<sup>rd</sup> Edition. Cambridge University Press, Cambridge.
- Whitlock MC. 2000. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution* 54:1855-1861.
- Wilkinson GS, Reillo PR. 1994. Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly. *Proceedings of the Royal Society of London B* 255:1-6.
- Williams GC. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press, Princeton.
- Whitlock MC, Agrawal AF. 2009. Purging the genome with sexual selection: reducing mutation load through selection in males. *Evolution* 63:569-582.
- Zahavi A. 1975. Mate selection – selection for a handicap. *Journal of Theoretical Biology* 53:205-214.
- Zahavi A. 1977. Cost of honesty: Further remarks on the handicap principle. *Journal of Theoretical Biology* 67:603-605.
- Zeh JA, Zeh DW. 1996. The evolution of polyandry I: Intragenomic conflict and genetic incompatibility. *Proceedings of the Royal Society of London B* 263:1711-1717.
- Zeh JA, Zeh DW. 1997. The evolution of polyandry II: Post-copulatory defences against genetic incompatibility. *Proceedings of the Royal Society of London B* 264:69-75.

## Chapter 4

### Association Between Sex Ratio Distortion and Sexually Antagonistic Fitness Consequences of Female Choice

#### ABSTRACT

Genetic variation can be beneficial to one sex yet harmful when expressed in the other – a condition referred to as sexual antagonism. Because X chromosomes are transmitted from fathers to daughters, and sexually antagonistic fitness variation is predicted to often be X-linked, mates of relatively low-fitness males might produce high-fitness daughters, while mates of high-fitness males produce low-fitness daughters. Such fitness consequences have been predicted to influence the evolution of female mating biases and the offspring sex ratio. Females might evolve to prefer mates that provide good genes for daughters or might adjust offspring sex ratios in favor of the sex with the highest relative fitness. We test these possibilities in a lab-adapted population of *Drosophila melanogaster*, and find that females preferentially mate with males carrying genes that are deleterious for daughters. Preferred males produce equal numbers of sons and daughters, whereas unpreferred males produce female-biased sex ratios. As a consequence, mean offspring fitness of unpreferred males is higher than offspring fitness of preferred males. This observation has several interesting implications for sexual selection and the maintenance of population genetic variation for fitness.

#### INTRODUCTION

The presence of sexually antagonistic variation – where alleles increasing male fitness are deleterious when expressed in female genomes – poses a dilemma for the evolution of

female choice (Chippindale et al. 2001). Females face a tradeoff of producing relatively fit sons and unfit daughters, or relatively fit daughters and unfit sons. If sexually antagonistic variation is sufficiently abundant and X-linked, as suggested by theory and data (Rice 1984; Gibson et al. 2002; Pischedda & Chippindale 2006), there will be no benefit of mating with relatively fit males, but there can be a cost associated with the production of low-fitness daughters.

Two possible evolutionary responses are predicted when sexually antagonistic fitness variation predominates, and has a strong X-linked component. Albert & Otto (2005) argued that female mating biases will evolve to favor males providing good genes to daughters. Female choice might therefore reverse the direction of selection acting on males and resolve the sexual antagonism. Calsbeek & Sinervo (2004) proposed an alternative response, that females should modify the sex ratio of their offspring in order to minimize indirect costs of mating with relatively high-fitness males.

The fruit fly *Drosophila melanogaster* is a suitable species for testing these hypotheses. Lab-adapted populations of flies, where the environmental context of adaptation is known, are amenable to the measurement of traits closely associated with fitness. Previous studies show that sexually antagonistic genetic variation influences adult fitness variation (Chippindale et al. 2001; Long & Rice 2007), and much of this variation appears to be X-linked (Gibson et al. 2002; Pischedda & Chippindale 2006). Furthermore, females can adjust the sex ratio of their offspring (Mange 1970; Long & Pischedda 2005; Fuller & Mousseau 2007), suggesting that adaptive sex ratio adjustment with respect to sexually antagonistic variation is at least possible.

I conducted experiments using a lab-adapted population of fruit flies to address two questions:

- (1) Do female mating biases lead to sexually antagonistic fitness consequences for offspring?
- (2) Is the offspring sex ratio skewed in favor of the sex with the highest relative fitness?

## **METHODS**

### ***Drosophila* stocks**

Female choice and adult fitness components were estimated using the IV population, a lab-adapted population of *Drosophila melanogaster* that is described in Houle & Rowe (2003). The IV population and the competitor population, IVe – a lab-adapted population that is homozygous for the *ebony* mutation – were kindly provided by David Houle.

### **Mate choice trials**

Female mating biases were ascertained by two approaches. A series of tournament-style mating trials were used to identify males differing in mating success (Fig. 4.1). Trials used two- and three-day-old virgin males and females, sampled from the IV population. Each trial was conducted with a pair of males and a single virgin female. The first male to successfully mate was designated as the winner. Winners or losers from each round of trials were arbitrarily paired with each other and retested. Males that lost or won three successive trials were used for sex ratio and offspring fitness assays.

Male-male competition, or intra-sexual selection, might influence the outcome of the tournament-style assays, and could potentially override female mating biases. To test whether females preferentially mate with males that perform well in the tournament setting, I conducted a series of female mating latency experiments. A single male and virgin female were placed in a vial and observed until copulation occurred. The relative success of each male was estimated by the time required for him to mate. Successful males from the tournament trials were able to achieve copulation more quickly than unsuccessful males (Fig. 4.2), indicating that the tournament trials capture information about female mating biases and are not driven by male-male competition. Females bias matings towards relatively successful (3 wins, 0 losses) and away from relatively unsuccessful (3 losses, 0 wins) males. Such males are hereafter referred to as “preferred” and “unpreferred”, respectively.

### **Offspring Sex Ratio & Adult Fitness Components**

During the day after mating success trials were concluded, preferred and unpreferred males were assigned at random to virgin females from the IV population. Previous

research using *D. melanogaster* indicates that male fertility quickly recovers during a lag period of this duration (Markow et al. 1978), and thus, seminal fluid limitation should not adversely affect the mates of preferred males. Mates of experimental males were then permitted to lay eggs in vials for 12 hours. From these eggs, sets of 40 to 50 eggs were transferred to 8 ounce bottles, each containing standard cornmeal medium, 20 adult *ebony* females, and 20 *ebony* males. This is a typical adult density for IV and IVe flies, creating a typical larval environment for this lab-adapted population. The relatively low density of introduced eggs per bottle also minimizes interactions between the experimental offspring.

Offspring sex ratio was examined in three independent experimental trials. The first two trials each followed offspring of 40 preferred and 40 unpreferred males. The third trial followed offspring of 70 preferred and 70 unpreferred males. Offspring from the third trial were used for fitness assays. Adult offspring were collected on the 14<sup>th</sup> day, consistent with the evolutionary history of the population, which has been continuously reared on a 14-day generation cycle since 1975, representing over 800 generations of adaptation to the lab environment (Houle & Rowe 2003).

Sex-specific selection in *Drosophila melanogaster* may influence the evolution of juvenile growth traits (Prasad et al. 2007). Juvenile growth differences can therefore potentially underlie sexually antagonistic fitness effects that manifest at the adult stage. However, juvenile sex-specific selection does not appear to give rise to sexually antagonistic viability selection. Indeed, there is a strong, positive intersexual genetic correlation for *Drosophila* juvenile viability (e.g., Chippindale et al. 2001). Since our major concern here is with sexual antagonism, the results focus on adult fitness-related traits – female fecundity and male mating success – that are potentially influenced by sexually antagonistic variation. However, the overall conclusions do not rely upon an emphasis on adult-stage fitness. Estimates of egg-to-adult viability for preferred and unpreferred males reveal no mortality differences between treatments (unpreferred offspring:  $n = 1248$  eggs, 70.8 % survival; preferred offspring:  $n = 1082$  eggs, 69.8 % survival; two-tailed Fisher exact test  $p = 0.82$ ).

Adult-stage female fitness was estimated by the number of eggs produced on the 14<sup>th</sup> day of the life cycle. Female offspring were placed in pairs, along with two randomly

assigned *ebony* males, into vials containing standard cornmeal medium and were permitted to lay eggs for 24 hours. Houle & Rowe (2003) previously showed that this is the critical time period during which egg laying contributes to adult female fitness.

Male fitness was estimated by mating success experiments, conducted during the 14<sup>th</sup> day of the life cycle. Male offspring were individually transferred to mate competition vials, each containing a randomly selected (and unrelated) male and female. These females and competitor males were each heterozygous for an *ebony* allele, and expressed the wild-type pigmentation pattern. Each competition vial was observed until the female mated with one of the males. The female was then isolated and permitted to lay eggs in a fresh vial. Paternity was assigned 14 days later by the presence or absence of *ebony*-expressing offspring. This measure of male fitness eliminates postcopulatory sexual selection, but is not expected to bias the results because precopulatory mating success is positively correlated with sperm competition success (Bangham et al. 2002; Hosken et al. 2008; though the two traits have different genetic bases: Zhang et al. 2008). The measure also assumes that mating success on day 14 is correlated with overall male success, which will be a function of success on or before day 14. This assumption could potentially be violated if there is an extreme reversal in relative male mating success during the span of a couple of days, but there is no a priori reason to expect such an extreme reversal, nor is there any such precedent in *Drosophila melanogaster*.

Chi-square tests were used to detect sex ratio deviations from unity (1:1). Fisher exact tests were used to examine whether preferred and unpreferred fathers had sons with different degrees of mating success. Two-tailed t-tests were used to test whether egg production rates differed between daughters of preferred and unpreferred fathers.

## **RESULTS AND DISCUSSION**

### **Offspring Fitness Estimates**

Preferred fathers had sons with slightly higher mating success, though the difference was small and not statistically significant (percent paternity unpreferred = 0.627,  $n = 193$ ; percent paternity preferred = 0.647,  $n = 207$ ; Fisher exact test  $p = 0.679$ ). Daughters of preferred males had decreased fecundity compared to the daughters of unpreferred males (Fig. 4.1; unpreferred daughters produced a mean of 70.20 eggs per vial,  $n = 203$  vials;

preferred daughters produced 56.85 eggs per vial,  $n = 164$  vials; two-tailed ttest  $p = 0.0000004$ ). The estimated fitness gain of 3 percent to sons is substantially lower than the estimated fitness drop of 19 percent to daughters of preferred males.

A strong, negative fitness correlation between fathers and daughters, coupled with marginal father-son fitness heritability cannot be explained by autosome-linked sexually antagonistic variation, which predicts that costs and benefits will be symmetrical between sons and daughters (Kidwell et al. 1977). It is also possible that females differentially provision eggs fertilized by unpreferred males. However, a “paternal effect” such as this should reduce the fitness of both sons and daughters – this prediction is difficult to reconcile with the data. The offspring fitness pattern is consistent with prior theory and data suggesting that adult fitness traits are strongly influenced by X-linked sexually antagonistic variation (Rice 1984; Gibson et al. 2002; Pischedda & Chippindale 2006; Oneal et al. 2007).

### **Offspring Sex Ratio**

Offspring sex ratios differed between experimental treatments, with preferred males producing sons and daughters at equal ratios, and unpreferred males producing daughter-biased sex ratios (Table 4.1). The single exception in trial 2 (equal sex ratio for preferred and unpreferred fathers) can be attributable to its markedly reduced sample size relative to sampling in trails 1 and 3. Indeed, the overall sex ratio reduction in unpreferred male broods is relatively strong and highly significant (male : female ratio = 0.874;  $p = 0.0056$ ).

There are three possible mechanistic explanations for the sex ratio biases observed here. The sex ratio might be equal among fertilized eggs, but viability selection might differentially remove males and females from the adult population. This mechanism would require that son and daughter viability be decoupled for unpreferred but not for preferred males. This scenario appears unlikely in *D. melanogaster*, where juvenile viability is strongly positively correlated between the sexes (Chippindale et al. 2001).

Females might adjust progeny sex ratios in response to their mate. This is a possibility in *D. melanogaster*, where females have been shown to adjust offspring sex ratio in a mating context-dependent manner (Mange 1970; Long & Pischedda 2005;

Fuller & Mousseau 2007). When variation is sexually antagonistic, this hypothesis predicts that females mated to preferred males will produce male-biased offspring sex ratios, while unpreferred males will sire female-biased broods. Only part of this pattern is supported, as preferred male broods have a sex ratio near unity. Nevertheless, the direction of skew in offspring of unpreferred males is adaptive since it is biased towards the sex with highest relative fitness (daughters).

Males with low mating success might have sex ratio distorting X chromosomes, which are common in *Drosophila* populations (Jaenike 2001). Associations between mating success and male meiotic drive have been reported in studies using mice and stalk-eyed flies (Lenington 1991; Wilkinson et al. 1998), though it is not known whether “driving” X chromosomes are associated with sexually antagonistic variation in these species. Such linkage disequilibrium might be expected. Males carrying a driving X with female-beneficial alleles will have higher quality offspring than males carrying driving, female-detrimental X chromosomes. The effect could promote the development of linkage disequilibrium between sex ratio and sexually antagonistic alleles. This is an area of population genetics theory that has yet to be formally explored (see Burt & Trivers 2006 for verbal discussion of meiotic drive and sexual antagonism).

## CONCLUSION

The results indicate that female mating biases in *Drosophila* might cause a net decrease to offspring fitness. This is not simply an artifact of the tournament mate-trial setting because mating latency trials independently confirm that females mate more readily with males designated as “preferred” (Fig. 4.2). Mating biases that reduce offspring fitness seem paradoxical, yet three factors could potentially explain the persistence of such a pattern of female choice. First, males might provide a direct benefit to their mates, which could potentially counteract any indirect costs (e.g., Oneal et al. 2007). Although current evidence from *Drosophila melanogaster* suggests that direct effects of female preferences are not beneficial and instead could be costly (Friberg & Arnqvist 2003), this scenario would be an interesting topic for future research. A second possibility is that *Drosophila* female mating biases are passive rather than active. For example, male mating success is at least partially a function of locomotor activity, with

high-activity males encountering and consequently mating with more females than low-activity males. A passive female mating bias of this variety is not directly costly to females because there is no energetic cost of searching for a mate (Kokko et al. 2006; Kotiaho & Puurtinen 2007), and is likely to have sexually antagonistic fitness consequences for offspring (Long & Rice 2007). Furthermore, females are only expected to evolve to resist males carrying female-detrimental genes if the cost of active mate choice is less than the indirect costs of having less fit offspring. Lastly, multiple mating might mitigate the indirect fitness costs that are observed here, in single mating contexts. Fitness estimates are for offspring from females that were singly-mated to a preferred or unpreferred male. In natural contexts, *Drosophila* females mate multiply, which could help to eliminate sexually antagonistic consequences for offspring if females preferentially use X-bearing sperm from “unpreferred” males and Y-bearing sperm from “preferred” males.

The sex ratio bias observed for unpreferred fathers has adaptive consequences for both parents. Unpreferred males and their mates have high-fitness daughters and benefit by producing daughters in excess of sons. By producing offspring with a higher mean fitness, unpreferred fathers might make a greater long-term genetic contribution to the population than might be expected based on their relative mating success. In other words, male mating success variance might be considerably higher than the actual fitness variance among males.

Sexually antagonistic selection was largely ignored experimentally until the last several years, but steadily mounting evidence now indicates that it is an important mechanism maintaining population genetic variation for fitness (e.g., Chippindale et al. 2001; Foerster et al 2007; Cox & Calsbeek 2008). The potential ubiquity of sexually antagonistic variation, coupled with a variety of sex ratio distortion mechanisms (e.g., Clutton-Brock & Iason 1986; Jaenike 2001), suggests that the results reported here might commonly occur in nature. To our knowledge, only one other such report, in an *Anolis* lizard species (Calsbeek & Bonneaud 2008), has been published. Future work in additional animal species might benefit by integrating female choice assays with analyses of sex ratio modification and sex-specific offspring fitness.

Table 4.1. Male mating success and offspring sex ratios

	Males	Females	Sex Ratio	$P^1$
Trial 1				
unpreferred	296	366	0.809	0.0065
preferred	255	260	0.981	0.83
Trial 2				
unpreferred	78	72	1.08	0.62
preferred	87	86	1.01	0.94
Trial 3				
unpreferred	417	467	0.893	0.093
preferred	377	378	0.997	0.97
All Trials				
unpreferred	791	905	0.874	0.0056
preferred	719	724	0.993	0.90

<sup>1</sup> Significance based on chi-squared tests with 1 d.f.

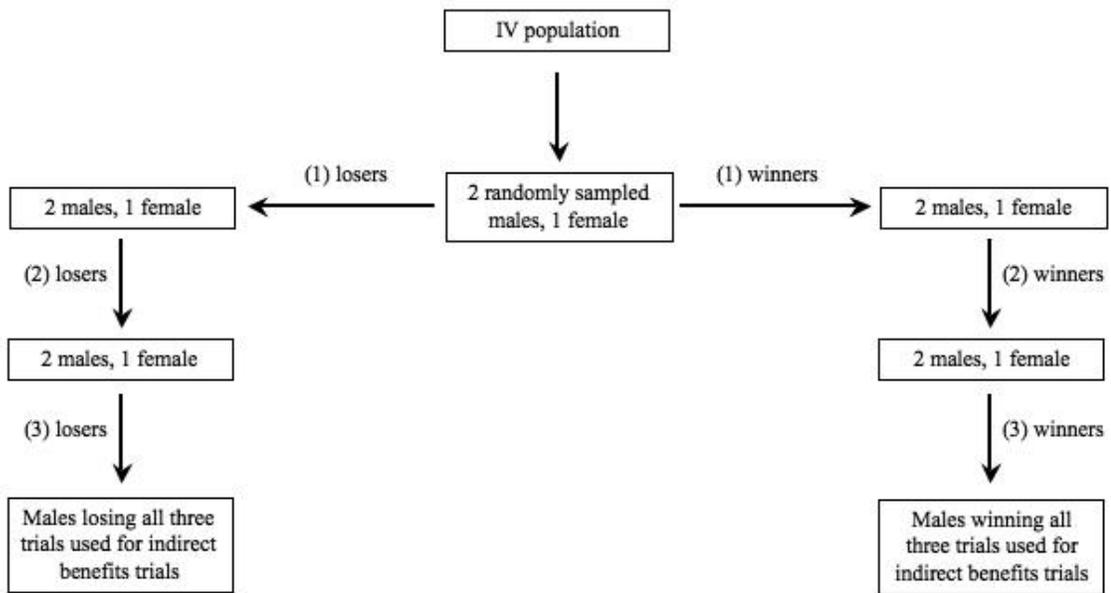


Figure 4.1. The male mating success trial design includes three rounds of competition during which successful males (right) or unsuccessful males (left) are retained for breeding experiments.

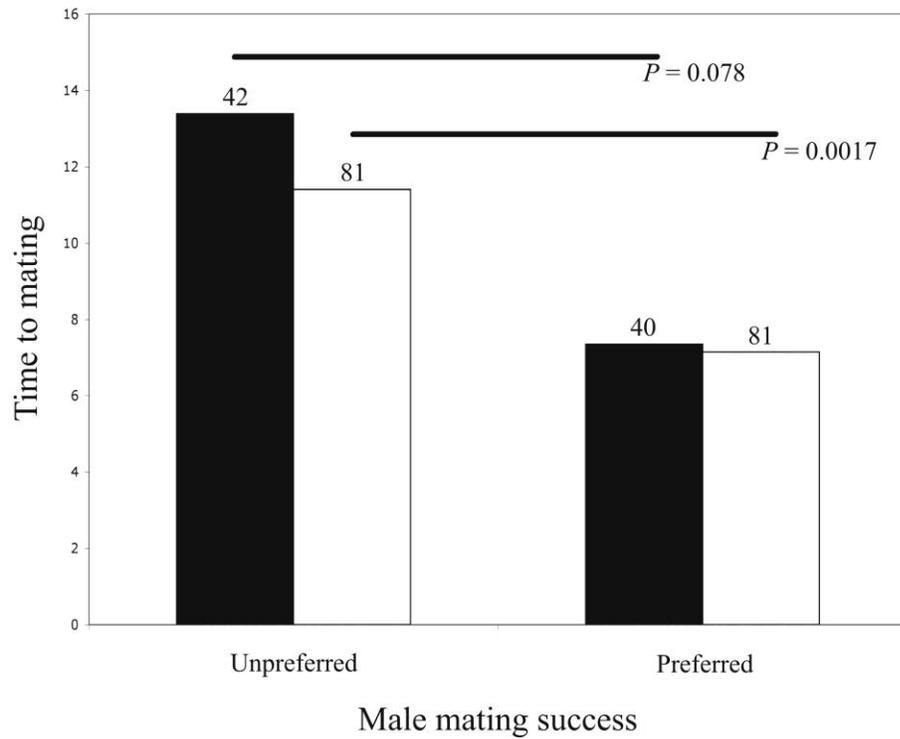


Figure 4.2. Relationship between different female preference assays. Black bars represent males that won or lost all three tournament trials. Open bars represent males that won or lost the first two trials. Means (time to mating in minutes) and sample sizes are shown. *P* values were calculated from two-tailed Mann-Whitney U tests.

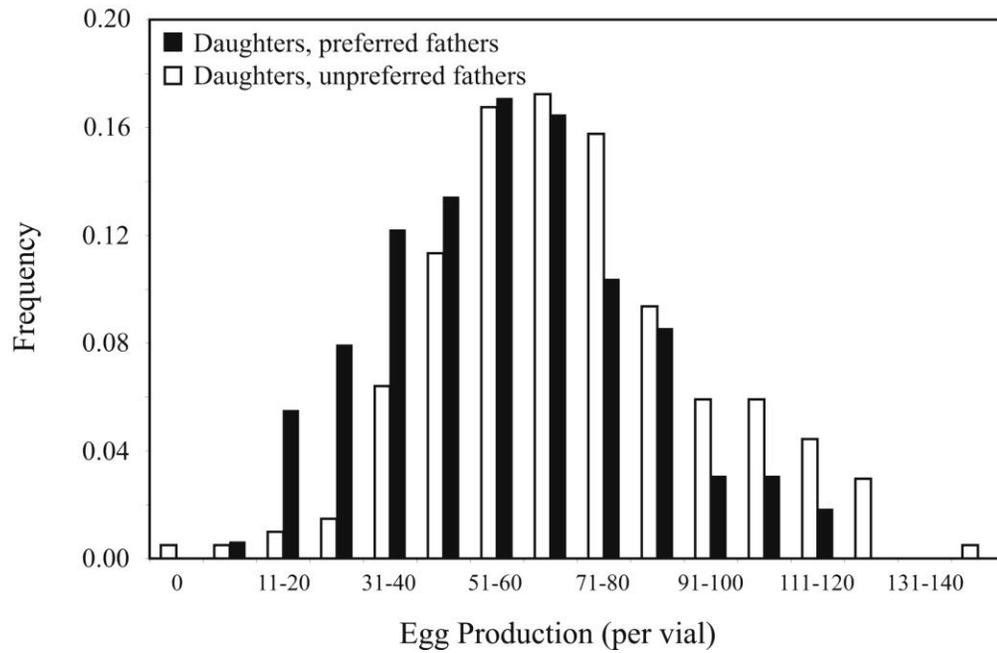


Figure 4.3. Paternal mating success influences daughter fecundity. The egg production distribution is based on egg counts per vial, each containing two experimental females and two randomly selected males (see methods for details).

## REFERENCES

- Albert AYK, Otto SP. 2005. Sexual selection can resolve sex-linked sexual antagonism. *Science* 310:119-121.
- Bangham J, Chapman T, Partridge L. 2002. Effects of body size, accessory gland and testis size and pre- and postcopulatory success in *Drosophila melanogaster*. *Animal Behavior* 64:915-921.
- Burt A, Trivers R. 2006. *Genes in Conflict: The Biology of Selfish Genetic Elements*. Harvard University Press, Cambridge.
- Calsbeek R, Bonneaud C. 2008. Postcopulatory fertilization bias as a form of cryptic sexual selection. *Evolution* 62:1137-1148.
- Calsbeek R, Sinervo B. 2004. Within-clutch variation in offspring sex determined by differences in sire body size: cryptic mate choice in the wild. *Journal of Evolutionary Biology* 17:464-470.
- Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proceedings of the National Academy of Sciences USA* 98:1671-1675.
- Clutton-Brock TH, Iason GR. 1986. Sex ratio variation in mammals. *Quarterly Review of Biology* 61:339-374.
- Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* 173:176-187.
- Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH, Kruuk LEB. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447:1107-1110.
- Fuller BA, Mousseau TA. 2007. Precision in sex allocation is influenced by mate choice in *Drosophila melanogaster*. *Journal of Evolutionary Biology* 20:1700-1704.
- Friberg U, Arnqvist G. 2003. Fitness effects of female mate choice: preferred males are detrimental for *Drosophila melanogaster* females. *Journal of Evolutionary Biology* 16:797-811.
- Gibson JR, Chippindale AK, Rice WR. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London B* 269:499-505.

- Hosken DJ, Taylor ML, Hoyle K, Higgins S, Wedell N. 2008. Attractive males have greater success in sperm competition. *Current Biology* 18:R553-R554.
- Houle D, Rowe L. 2003. Natural selection in a bottle. *American Naturalist* 161:50-67.
- Jaenike J. 2001. Sex chromosome meiotic drive. *Annual Review of Ecology, Evolution and Systematics* 32:25-49.
- Kokko H, Jennions MD, Brooks R. 2006. Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution and Systematics* 37:43-66.
- Kotiaho JS, Puurtinen M. 2007. Mate choice for indirect genetic benefits: scrutiny of the current paradigm. *Functional Ecology* 21:638-644.
- Lenington S. 1991. The t-complex: a story of genes, behavior, and populations. *Advances in the Study of Behavior* 20:51-86.
- Long TAF, Pischedda A. 2005. Do female *Drosophila melanogaster* adaptively bias offspring sex ratios in relation to the age of their mate? *Proceedings of the Royal Society of London B* 272:1781-1787.
- Long TAF, Rice WR. 2007. Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of *Drosophila melanogaster*. *Proceedings of the Royal Society of London B* 274:3105-3112.
- Mange AP. 1970. Possible nonrandom utilization of X- and Y-bearing sperm in *Drosophila melanogaster*. *Genetics* 65:95-106.
- Markow TA, Quaid M, Kerr S. 1978. Male mating experience and competitive courtship in *Drosophila melanogaster*. *Nature* 276:821-822.
- Oneal E, Connallon T, Knowles LL. 2007. Conflict between direct and indirect benefits of female choice in desert *Drosophila*. *Biology Letters* 3:29-32.
- Pischedda A, Chippindale AK. 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. *PLoS Biology* 4:2099-1203.
- Prasad NG, Bedhomme S, Day T, Chippindale AK. 2007. An evolutionary cost of separate genders revealed by male-limited evolution. *American Naturalist* 169:29-37.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735-742.
- Wilkinson GS, Presgraves DC, Crymes L. 1998. Male eye span in stalk-eyed flies indicates genetic quality by meiotic drive suppression. *Nature* 391:276-279.

Zhang R, Amah L, Fiumera AC. 2008. Autosomal variation for male body size and sperm competition phenotypes is uncorrelated in *Drosophila melanogaster*. *Biology Letters* 4:500-503.

## Chapter 5

### Evidence for cryptic genetic benefits of female choice in *Drosophila melanogaster*

#### INTRODUCTION

A major goal in evolutionary biology is to understand why females exhibit strong mating preferences when their mates do not provide direct benefits, such as resources or parental care (Trivers 1972; Borgia 1979; Andersson 1994). Females might receive an indirect benefit by choosing mates with ‘good genes’ that enhance offspring survival, although the importance of the good genes mechanism of female choice critically depends upon the availability of genetic variation for fitness, the ability of females to discriminate among males of variable genetic quality, and the transmission of heritable fitness variation between males and their offspring (Rowe & Houle 1996).

Theory shows that genome architecture can strongly influence the magnitude and empirical detectability of genetic benefits (see Chapter 3). Evidence for genetic benefits is expected to be particularly weak in species with male heterogametic sex determination, where males are hemizygous for X-linked genes. This situation arises because X chromosomes generate several asymmetries between fitness variation among fathers and the expected fitness of their offspring. When recessive deleterious mutations contribute to fitness variation, male fitness is disproportionately influenced by X chromosome genetic quality, whereas female fitness is mainly influenced by autosomal quality. Consequently, fitness heritability between fathers and daughters is low because the fitness of each is primarily influenced by genetic variation at different genomic locations. Fitness heritability to sons will necessarily be low because the father-son heritability of X-linked genetic variation is zero. Sexually antagonistic variation, which appears to be

preferentially X-linked (Rice 1984; Gibson et al. 2002), compounds the issue because alleles associated with paternal mating success are not passed to sons, and reduce offspring fitness when transmitted to daughters. Sexual antagonism is therefore expected to dampen and reduce an already weak fitness correlation between fathers and offspring.

The obscuring effect of the X chromosome is particularly acute in the *Drosophila* model system, which is amenable to the measurement of fitness (via lab-adapted populations; Houle & Rowe 2003; Rice et al. 2005), but also has a large X chromosome, typically ranging between approximately 20 percent (as in *D. melanogaster*) to 40 percent of the genome (as in *D. pseudoobscura*). Several previous studies report weak or no viability benefits of nonrandom mating in *Drosophila* (e.g., Partridge 1980; Schaeffer et al. 1984; Pitnick 1991). However, a single study using experimental males with standardized, homogeneous X chromosomes reports a very strong relationship between paternal mating success and offspring viability (Taylor et al. 1987). The comparative results of these studies support the idea that X-linked variation obscures underlying heritable fitness variation on autosomes, yet these experiments were conducted using different experimental protocols, with different populations, and within different laboratories. Thus, the results warrant further investigation.

Good genes might more easily be detected experimentally via manipulation of X chromosome transmission. If male fitness is strongly influenced by the X chromosome, good genes benefits will become magnified when sons rather than daughters inherit the paternal X. Paternal inheritance of high-quality X chromosomes is more beneficial to sons than to daughters because X-linked mutations are completely expressed in sons, but only partially expressed in daughters. If sexually antagonistic variation is present, daughters also benefit by not receiving a paternal X chromosome; by removing X-linked sexually antagonistic variation, otherwise obscured genetic benefits might be revealed.

To test this idea, I conducted a series of mate choice trials and measured viability of offspring of “preferred” and “unpreferred” males in a lab-adapted population of *Drosophila melanogaster*. Each male was crossed to a wild-type female to produce offspring inheriting his autosomes and daughters inheriting his X chromosome. These males were also crossed to mutant females (XXY, compound-X genotype), which permit normal autosome inheritance, with the transmission of the paternal X to sons rather than

daughters.

## **METHODS**

### ***Drosophila* stocks**

Female choice and indirect benefits were estimated in the *IV* population, a lab-adapted population of *Drosophila melanogaster* that is described in Houle & Rowe (2003). The *IV* population and the competitor population, *IVe* - a lab-adapted population that is homozygous for the *ebony* mutation – were kindly provided by David Houle. A compound-X stock, *C(1)DX*, was obtained from the Bloomington Stock Center. The genetic background of the *IV* population was introgressed into the *C(1)DX* stock for over 20 generations.

### **Female choice trials**

Following the same method as fully described in Chapter 5, series of tournament-style male mating success trials were conducted to distinguish relatively “unpreferred” and “preferred” males.

### **Offspring survival**

Individual preferred and unpreferred males were mated to both *IV* and *C(1)DX* females. In both crosses, offspring inherit paternal autosomes. Paternal X chromosomes pass to daughters in crosses to *IV* females. The X is passed from father to son in the *C(1)DX* cross. Females were permitted to lay eggs in vials for 12 hours. Sets of approximately 50 eggs were transferred to 8 ounce bottles with standard cornmeal medium, and 20 ebony females and 20 ebony males (from the *IVe* population). This is a typical adult density for *IV* flies, and should create a typical larval competition environment for this lab-adapted population. Because the population has been evolving under a 14-day generation cycle for over 800 generations, egg to adult viability can be estimated as the proportion of individuals emerging as an adult by the end of the 14-day cycle. Three replicates of the experiment were carried out during December 2007, January 2008, and May 2008. For the first two experimental trials, offspring survival was estimated for 40 preferred and unpreferred males. 70 preferred and unpreferred males

were used for the third trial.

### Statistical analysis

The probability of survival ( $p$ ) for each male type and each cross was estimated by maximum likelihood. Following Rice (2006), the likelihood function is:

$$L(p) \propto p^x (1-p)^{n-x}$$

where  $n$  is the number of eggs produced by the mates of preferred or unpreferred males, and  $x$  is the number of adults emerging by day 14. The maximum likelihood estimator for the proportion of surviving offspring is:

$$\hat{p} = \frac{x}{n}$$

The MLE estimator is approximately normal with mean  $E[\hat{p}] = p$  and variance:

$$Var(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n}$$

Estimates and standard errors of  $p$  were adjusted so that the male category with the highest offspring survival was scaled to one. In other words:

$$\hat{p}_{adjusted} = \frac{\hat{p}_i}{\hat{p}_{max}}$$

where  $\hat{p}_i$  the maximum likelihood estimator for male category  $i$  ( $i =$  unpreferred or preferred).

The adjusted standard error is:

$$SE(\hat{p}_{adjusted}) = \frac{1}{\hat{p}_{max}} \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}$$

Results for the three experimental trials are reported in Table 1. Standardized pooled data are shown in Fig. 1.

## RESULTS AND CONCLUSION

Results of the experiment confirm theoretical predictions that X-linked genetic variation strongly influences male fitness variation, and consequently, species with X chromosomes should exhibit a low fitness heritability between fathers and offspring, despite high levels of genetic variation affecting fitness (see Chapter 4). In the standard cross, offspring viability was nearly identical between preferred and unpreferred males

(pref. = 73.7 % survival; unpref. = 73.4 % survival; two-tailed Fisher exact test  $p = 0.944$ ). The manipulated cross revealed an otherwise masked benefit of mate choice, with the offspring of preferred fathers having improved survival relative to those of unpreferred fathers (pref. = 22.8 % survival; unpref. = 26.0 % survival; two-tailed Fisher exact test  $p = 0.0197$ ). The viability difference within the manipulated treatment was approximately 12 percent (Fig. 1).

These patterns highlight the fact that each species' genome architecture can potentially have a profound affect on fitness heritability between relatives. Heritability between fathers and offspring has received strong emphasis within the sexual selection literature, because it has strong implications for the evolution and adaptive significance of nonrandom mating (Møller & Alatalo 1999). Negative evidence for genetic benefits might seem, on the surface, to imply that male mating success variance is almost entirely based on environmental variance, or gene-by-environment interactions, which both predict trivial fitness heritability to offspring (e.g., Greenfield & Rodriguez 2004). On the contrary, the presence of sex chromosomes can profoundly alter patterns of heritability between species, independent of the relative abundance of additive genetic variation for fitness. Consequently researchers must incorporate information about an animal species' genome in order to properly interpret the relationship between heritability, the nature of genetic variation affecting fitness, and its consequences for mate choice evolution.

Table 5.1. Female choice and offspring egg-to-adult viability data for the three experimental trials<sup>1</sup>

Trial	Manipulated cross (X to sons)		Normal cross (X to daughters)	
	Unpreferred fathers	Preferred fathers	Unpreferred fathers	Preferred fathers
1	0.2268 (0.0190)	0.2787 (0.0213)	0.8054 (0.0138)	0.8085 (0.0156)
2	0.2102 (0.0154)	0.2597 (0.0136)	0.6224 (0.0312)	0.7238 (0.0289)
3	0.2350 (0.0094)	0.2561 (0.0095)	0.7083 (0.0129)	0.6978 (0.0140)

<sup>1</sup> Proportion of offspring emerging as adults (standard errors are in parentheses).

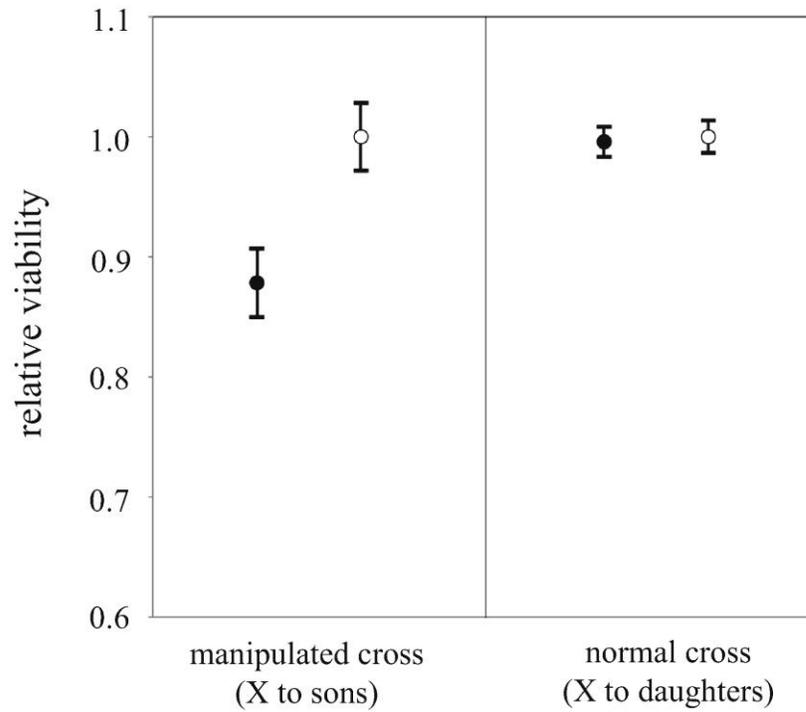


Figure 5.1. Genetic benefits of female choice under two crossing schemes. Open circles refer to offspring of “preferred” fathers. Closed circles are offspring of “unpreferred” fathers. Data are standardized against the survival of offspring of preferred fathers (see Methods section for details).

## REFERENCES

- Andersson M. 1994. Sexual selection. Princeton University Press, Princeton.
- Borgia G. 1979. In Sexual selection and reproductive competition in the insects. Blum, M. S. & Blum, N. A. (eds.). Academic Press, London, pp. 19-80.
- Gibson JR, Chippindale AK, Rice WR. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. Proceedings of the Royal Society of London B 269:499-505.
- Rice JA. 2006. Mathematical Statistics and Data Analysis, 3 Ed. Thompson. Duxbury Press.
- Greenfield MD, Rodriguez RL. 2004. Genotype-environment interaction and the reliability of mating signals. Animal Behaviour 68:1461-1468.
- Houle D, Rowe L. 2003. Natural selection in a bottle. American Naturalist 161:50-67.
- Møller AP, Alatalo RV. 1999. Good genes effects in sexual selection. Proceedings of the Royal Society of London B 266:85-91.
- Partridge L. 1980. Mate choice increases a component of offspring fitness in fruit flies. Nature 283:290-291.
- Pitnick S. 1991. Male size influences mate fecundity and remating interval in *Drosophila melanogaster*. Animal Behaviour 41:735-745.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735-742.
- Rice WR, Linder JE, Friberg U, Lew TA, Morrow EH, Stewart A. 2005. Interlocus antagonistic coevolution as an engine of speciation: assessment with hemiclinal analysis. Proceedings of the National Academy of Science USA 102:6527-6534.
- Rowe L, Houle D. 1996. The lek paradox and the capture of genetic variance by condition-dependent traits. Proceedings of the Royal Society of London B 263:1416-1421.
- Schaeffer SW, Brown CJ, Andersson WW. 1984. Does mate choice affect fitness? Genetics 107:s94.
- Taylor CE, Pereda AD, Ferrari JA. 1987. On the correlation between mating success and offspring quality in *Drosophila melanogaster*. American Naturalist 129:721-729.

Trivers RL. 1972. Parental investment and sexual selection. Pages 136-179 in B. Campbell, ed. *Sexual selection and the descent of man, 1871-1971*. Aldine, Chicago.

## **Part III: Population Benefits and Costs of Sexual Reproduction**

### **Chapter 6**

#### **Sexually Antagonistic Selection and the Twofold Cost of Sex**

##### **ABSTRACT**

Sexual reproduction is extremely widespread, yet it is expected to carry with it a twofold cost relative to asexuality. Theory suggests that sex-specific selection may enhance purifying selection against deleterious mutations and help to eliminate the twofold cost of sex. However, recent experiments indicate that sex-specific selection can also be costly owing to sexual antagonism: alleles harmful to one sex can accumulate within a population because they are favored in the other sex. Whether sex-specific selection provides a net fitness benefit or cost depends, in part, on the relative frequency and strength of sexually concordant versus sexually antagonistic selection throughout a species' genome. Here, we model the net fitness consequences of sex-specific selection to describe its costs and benefits while explicitly considering both sexually concordant and sexually antagonistic selection. The model shows that, even when sexual antagonism is rare, the fitness costs that it imposes will generally overwhelm the fitness benefits of sexually concordant selection. Furthermore, the cost of sexual antagonism is, at best, only partially resolved by the evolution of sex-limited gene expression. To evaluate the key parameters of the model, we analyze an extensive dataset of sex-specific selection gradients from wild populations, along with data from the literature on experimental evolution. These empirical data suggest that sex-specific selection likely imposes a net cost, though additional research will be required to confirm this conclusion.

## INTRODUCTION

The ubiquity of sexual reproduction has been described as “the outstanding puzzle in evolutionary biology” (G. C. Williams, as cited by Ridley 1996). All else being equal, an asexual population should grow at twice the rate of a sexual population because the male half of a sexual population does not contribute to its fecundity. In other words, sex bears an inherent “twofold cost” relative to asexuality (Maynard Smith 1978). Nevertheless, sexual reproduction is widespread, and must therefore confer benefits that outweigh this twofold cost (Bell 1982; Rice 2002).

One feature that distinguishes sexual populations from their asexual counterparts is the potential for natural and sexual selection to differ in strength and/or direction between the sexes. Such sex-specific selection has recently been proposed as a potential resolution to the paradox of sex (Sharp & Agrawal 2008; Whitlock & Agrawal 2009). Mathematical theory shows that selection can compensate for the twofold cost of sex if it acts in the same direction on males and females, but is relatively stronger in males (Manning 1984; Koeslag & Koeslag 1994; Whitlock 2000; Agrawal 2001; Siller 2001; Lorch et al. 2003; Hadany & Beker 2007). This occurs because selection on males does not affect the reproductive rate of a population, which depends only on the survival and fecundity of females. However, strong selection on males can provide a benefit by purging deleterious mutations from the population. Thus, a sexual population can experience strong purifying selection without suffering extreme reductions in its reproductive rate via increased mortality or variable fecundity of females. The benefits of this mechanism can theoretically arise from sex differences in selection arising from differential survival, fecundity, and/or mating success.

For sex-specific selection to provide a benefit, selection must favor the same alleles in each sex and differ only in its relative strength. Whether selection has this sexually concordant effect is difficult to assess at the genotypic level, but it is increasingly recognized that many phenotypic traits are subject to opposing selection pressures in each sex (Rice & Chippindale 2001; Cox & Calsbeek 2009). This sexual antagonism can lead to the accumulation of mutations that are beneficial to males, but detrimental to females, thereby reducing the fitness of sexual populations (Prasad et al. 2007; Bonduriansky &

Chenoweth 2009). Under such a scenario, sex-specific selection can generate an additional cost of sexual reproduction.

Given these two highly divergent outcomes, it remains unclear whether sex-specific selection typically reduces or adds to the twofold cost of sex. This uncertainty stems from two major limitations of existing theory and data. First, theoretical models have yet to consider the effects of sexual antagonism when assessing the benefits of sexual selection. Second, the paucity of sex-specific selection estimates has precluded rigorous empirical tests of these models (for discussion, see Whitlock & Agrawal 2009; Hollis et al., 2009). Here, we extend the theoretical and empirical scope of previous studies by incorporating both deleterious mutations and sexually antagonistic selection pressures into a new model that contrasts fitness between sexual and asexual populations. We then discuss empirical data from experimental evolution and field-based studies of sex-specific selection with respect to the key parameters of this model. By combining a new mathematical model with several independent lines of empirical data, we show that sex-specific selection is likely to impose additional costs on sexual species, contrary to the conclusions of previous theoretical studies.

## **MODEL**

Sex-specific selection introduces two consequences for adaptation throughout the genome. A proportion of mutations (hereafter  $\zeta$ ), formerly deleterious to all individuals, becomes beneficial to one sex and remains deleterious to the other. Such mutations are subject to sexually antagonistic selection. All other mutations ( $1 - \zeta$ ) remain deleterious and are subject to sexually concordant selection, although the strength of purifying selection may differ between the sexes. Thus, we use the term ‘sexually antagonistic’ to refer to sex-specific selection in which the direction of selection differs between the sexes, and ‘sexually concordant’ to refer to sex-specific selection in which the direction of selection is the same in both sexes, despite differences in its relative strength. Below, we separately address the population genetic consequences of sexually concordant and sexually antagonistic selection, and then consider how both processes jointly contribute to genome-wide fitness.

### Sexually Concordant Selection.

For the case of sexually concordant selection at a locus with two alleles,  $A_1$  and  $A_2$ , sex-specific fitness follows the scheme:

Genotype:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness ( $w_f$ ):	$1 - s_f$	$1 - s_f h$	1
Male fitness ( $w_m$ ):	$1 - s_m$	$1 - s_m h$	1

where the selection coefficients  $s_m$  and  $s_f$ , and the dominance coefficient  $h$ , are all positive. Here, the  $A_1$  allele is disfavored in both sexes, but will persist at low frequency due to mutations from  $A_2$  to  $A_1$ . Under sexually concordant selection, the deleterious allele  $A_1$  evolves to a single equilibrium, at mutation-selection balance:

$$\hat{p} \approx \frac{2\mu}{s_f h (1 + \alpha_C)} \quad (1)$$

(Whitlock & Agrawal 2009), where  $\alpha_C = s_m/s_f$  and  $\mu$  refers to the mutation rate from  $A_2$  to  $A_1$  (backmutation is assumed to be negligible because  $A_1$  is rare). Assuming multiplicative epistasis between mutations, and following the approach of Kimura & Maruyama (1966) to extend across all sexually concordant loci in the genome (see also Maynard Smith 1978; Agrawal 2001), the number of deleterious mutations per individual is expected to follow a Poisson distribution with parameter:

$$\lambda_C \approx \frac{2U(1 - \zeta)}{s_f h (1 + \alpha_C)} \quad (2)$$

where  $U$  represents the genomic mutation rate per zygote, per generation.

### Sexually Antagonistic Selection

For the case of sexual antagonism at a locus with two alleles, sex-specific fitness follows the scheme:

Genotype:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Female fitness ( $w_f$ ):	$1 - s_f$	$1 - h_f s_f$	1
Male fitness ( $w_m$ ):	1	$1 - h_m t_m$	$1 - t_m$

where sexually antagonistic selection coefficients  $s_f$  and  $t_m$ , and dominance coefficients  $h_f$  and  $h_m$ , are all positive. Here, the  $A_I$  allele increases male fitness, thereby generating conflicting selection pressures between the sexes.

Sexual antagonism generates three possible evolutionary equilibria, which depend upon the strength of selection in males relative to females (Kidwell et al. 1977; Appendix 6.1). When selection is stronger in females than in males, the male-beneficial allele,  $A_I$ , evolves to a low-frequency equilibrium at mutation-selection balance. This condition specifically occurs when:

$$s_f > \frac{t_m(1-h_m)}{h_f(1-t_m)} \quad (3)$$

When selection is stronger in males than in females, the female beneficial allele,  $A_2$ , occurs at low frequency at mutation-selection balance. This condition occurs when:

$$s_f < \frac{t_m h_m}{1-h_f+t_m h_m} \quad (4)$$

When selection in males and females is similar in strength, sexual antagonism can maintain a stable polymorphism, under the condition:

$$\frac{t_m h_m}{1-h_f+t_m h_m} < s_f < \frac{t_m(1-h_m)}{h_f(1-t_m)} \quad (5)$$

When  $A_I$  is equally dominant in males and females (i.e.,  $h_f = 1 - h_m$ ), and selection coefficients are relatively small ( $t_m, s_f < 0.1$ ), the conditions favoring a balanced polymorphism are quite restrictive (Kidwell et al. 1977; Prout 1999; Appendix 3.1), and there are essentially two equilibria: one in which selection on females dominates and  $A_I$  remains at low frequency (condition 3; hereafter “weak” sexual antagonism), and one where selection on males dominates and  $A_I$  approaches fixation (condition 4; hereafter “strong” sexual antagonism).

### **Net Fitness of Females Under Weak Sexual Antagonism**

When sexual antagonism is weak, selection in females for modifiers that mitigate the harmful effects of  $A_I$  will also be weak (i.e., on the order of the mutation rate; Wright

1929), and sexually antagonistic fitness effects should persist over evolutionary time. The equilibrium frequency of such female-deleterious mutations is approximated by:

$$\hat{p} \approx \frac{2\mu}{s_f h_f (1 - \alpha_A)} \quad (6)$$

where  $\alpha_A$  is the relative strength of sexual antagonism, defined as  $\alpha_A = t_m/s_f$  (see Appendix 6.2). Extending across all sexually antagonistic loci (a fraction  $\zeta$  of total loci), the number of male-beneficial/female-detrimental mutations per female follows a Poisson distribution with parameter:

$$\lambda_A \approx \frac{2U\zeta}{s_f h_f (1 - \alpha_A)} \quad (7)$$

which represents the average number of male-beneficial (female-detrimental) mutations per individual.

By combining sexually antagonistic and sexually concordant loci, mean female fitness can be calculated as a function of the number of harmful mutations per female (i.e., female-detrimental plus universally deleterious alleles):

$$\bar{W}_f = \left\{ \sum_{k=0}^{\infty} (1 - s_f h_f)^k \exp(-\lambda_A) \frac{(\lambda_A)^k}{k!} \right\} \left\{ \sum_{j=0}^{\infty} (1 - s_f h)^j \exp(-\lambda_C) \frac{(\lambda_C)^j}{j!} \right\} \quad (8)$$

When combined with equations (2) and (7), this simplifies to:

$$\bar{W}_f = \exp \left\{ -2U \left[ \frac{(1-\zeta)}{1+\alpha_C} + \frac{\zeta}{1-\alpha_A} \right] \right\} \quad (9)$$

### Net Female Fitness Under Strong Sexual Antagonism

Under strong sexual antagonism ( $\alpha_A > 1$ ), deterministic invasion of a male-beneficial allele can lead to two potential outcomes. First, selection for modifiers can lead to the evolution of sex-limited expression (Lande 1980; Rice 1984; Otto & Bourguet 1999). This means of resolving sexual antagonism generates loci with sex-specific effects on fitness (Bonduriansky & Chenoweth 2009), and which carry deleterious mutations at frequency:

$$\hat{p} \approx \frac{2\mu}{s_f h_f} \quad (10)$$

For such loci, mutations may occur during every generation, while selection is limited to 50 percent of generations (i.e., each mutation has a 0.5 probability of occurring in a male or a female genome during each generation). These loci will therefore harbor at least twice as much deleterious variation as loci that are under sexually concordant selection (where  $\alpha_C \geq 1$ ). Under a scenario where sexual antagonism is completely resolved by sex-limited expression, mean fitness of females at equilibrium will be:

$$\bar{W}_f = \exp\left\{-2U\left(\frac{1+\psi\alpha_C}{1+\alpha_C}\right)\right\} \quad (11)$$

where  $\psi$  represents the genome-wide proportion of female-expressed loci with sex-limited fitness effects.

If the cost of silencing expression of a sexually antagonistic gene is greater than the cost of expressing the male-beneficial variant, the evolution of sex-limited expression will be constrained. In such cases, sex-limitation might be disadvantageous and the sexual antagonism will not be resolved. To examine the consequences of strong and unresolved sexually antagonistic selection with respect to the fitness of females, we must define a new term that refers to the number of loci that are under sexually antagonistic selection (hereafter  $L_A$ ). Since the genomic mutation rate is  $U = 2L\mu$ , where  $L$  is the number of loci in the genome (2 reflects the fact that a mutation can be paternally or maternally inherited for a diploid locus), the number of sexually antagonistic loci will be:

$$L_A = \zeta L = \frac{\zeta U}{2\mu}.$$

For strong sexual antagonism ( $\alpha_A > 1$ ) and relatively small selection coefficients (i.e.,  $s_f, t_m < 0.1$ , as previously assumed), female-detrimental alleles are nearly fixed within the population (Kidwell et al. 1977). At mutation-selection equilibrium, the frequency of a female-beneficial allele ( $A_2$ ) is:

$$\hat{p} \approx \frac{2\mu}{s_f(\alpha_A - 1)(1 - h_f)} \quad (12)$$

(see Appendix 6.2). Since  $A_2$  is rare, the number of loci that are homozygous for the female-beneficial allele will be negligible. The number of heterozygous, sexually antagonistic loci is Poisson-distributed with parameter:

$$\lambda_{A_1A_2} = 2\hat{p}(1-\hat{p})L_A \approx \frac{2U\zeta}{s_f(\alpha_A - 1)(1-h_f)} \quad (13)$$

Mean fitness of females with respect to the sexually antagonistic loci will be:

$$\begin{aligned} \bar{W}_{fA} &= \sum_{k=0}^{\infty} (1-s_f)^{L_A-k} (1-s_f h_f)^k \exp(-\lambda_{A_1A_2}) \frac{(\lambda_{A_1A_2})^k}{k!} \\ &= (1-s_f)^{\left(\frac{\zeta U}{2\mu}\right)} \exp\left(\frac{2U\zeta}{(1-s_f)(\alpha_A - 1)}\right) \end{aligned} \quad (14)$$

Mean fitness of females, considering both sexually antagonistic and sexually concordant loci, is:

$$\begin{aligned} \bar{W}_{sex} &= (1-s_f)^{\left(\frac{\zeta U}{2\mu}\right)} \exp\left(\frac{2U\zeta}{(1-s_f)(\alpha_A - 1)} - \frac{2U(1-\zeta)}{1+a_C}\right) \\ &\approx \exp\left\{-U\left[\frac{2(1-\zeta)}{1+a_C} + \zeta\left(\frac{s_f}{2\mu} - \frac{2}{\alpha_A - 1}\right)\right]\right\} \end{aligned} \quad (15)$$

## Simulations

The analytical solutions presented here rely on two important assumptions about sexually antagonistic variation. Selection coefficients are assumed to be relatively small and dominance is assumed to be identical between the sexes (i.e.,  $A_I$  is equally dominant in males and females, which requires that  $h_f = 1 - h_m$ ). Under these assumptions, conditions favoring a stable polymorphism are so restrictive that they can largely be ignored. However, as selection increases in strength, or dominance becomes sex-specific, sexual antagonism is expected to increasingly favor the maintenance of genetic variation and conclusions based on analytical models become less reliable. We therefore performed a series of simulations to account for strong genic-selection and sex-specific dominance, and show that equations (9) and (15) represent lower and upper limits for the costs of sexual antagonism, respectively (see Appendix 6.3 for details). As the strength of selection increases ( $t_m, s_f \rightarrow 1$ ) or under dominance reversal conditions (e.g.,  $h_f = h_m$ ), the parameter space where sexual selection provides a net benefit, which we present in Fig. 1 (based on equation (9)), becomes reduced relative to the analytical approximation. Our main conclusion below, that even a small proportion of sexually antagonistic selection

generates a cost to females that overwhelms benefits of sexually concordant selection, is therefore conservative.

## RESULTS

When sex-specific selection is absent, the average fitness of an asexual individual is:

$$\bar{W}_{asex} = e^{-U} \quad (16)$$

(e.g., Kimura & Maruyama 1966; Kondrashov & Crow 1988). Sex-specific selection can substantially alter mean fitness of females in a sexually reproducing population.

Consistent with previous work (Agrawal 2001; Siller 2001), we find that sex-specific selection is beneficial, and will mitigate the twofold cost of sex, in the absence of sexual antagonism ( $\zeta = 0$ ) and when deleterious mutations are more harmful to males than to females ( $\alpha_C > 1$ ; Fig. 6.1).

When there is sexual antagonism ( $\zeta > 0$ ), sex-specific selection can also generate a net fitness benefit for females, but only when  $\alpha_C$  increases much more rapidly than  $\alpha_A$ . Under weak sexual antagonism ( $\alpha_A < 1$ ), sex-specific selection improves the fitness of females under the condition:

$$1 > \frac{2(1-\zeta)(1-\alpha_A) + 2\zeta(1+\alpha_C)}{(1+\alpha_C)(1-\alpha_A)} \quad (17)$$

Otherwise, sex-specific selection adds to the cost of sex. An analysis of condition (17) shows that sex-specific selection generates a net fitness cost for females when even a small fraction of the genome is exposed to sexually antagonistic selection (Fig. 6.1).

When antagonism is strong (i.e.,  $\alpha_A > 1$ ), the costs of sexual antagonism generally overwhelm any benefits generated by sexually concordant selection. Sex-specific selection produces a net benefit under the condition:

$$1 > \frac{2(1-\zeta)}{1+\alpha_C} - \frac{\zeta 2}{\alpha_A - 1} + \frac{\zeta s_f}{2\mu} \quad (18)$$

where the mutation rate per locus ( $\mu$ ) and female selection coefficient ( $s_f$ ) no longer drop out of the analysis (as they do under weak sexual antagonism). Estimates of  $\mu$  from the literature indicate that it is on the order of  $10^{-8}$  (e.g., Haag-Liautard et al. 2007). While estimates of  $s$  are less clear, it is likely within a range between  $s = 10^{-5}$  (based on analysis of segregating polymorphism; Loewe et al. 2006, Andolfatto 2007) and  $s = 10^{-2}$  (based on

mutation-accumulation; Shabalina et al. 1997). Given a rough estimate of  $s_f/\mu$  between 1000 and 1,000,000, the term  $\zeta s_f/2\mu$  will often severely inflate the right side of condition (18), even when sexual antagonism is rare ( $\zeta \ll 1$ ) and sexually concordant selection is strong ( $\alpha_C$  goes to infinity). In other words, strong sexually antagonistic selection produces a cost that is so severe that it is unlikely to be offset by benefits of enhanced purifying selection in males.

Under both weak and strong sexual antagonism, the transition points that define the boundary where sex-specific selection is beneficial versus costly are independent of the genomic mutation rate,  $U$ . However, the magnitude of the net costs or benefits of sex-specific selection is closely tied to  $U$ . Within the parameter space producing a net cost of sex-specific selection, high values of  $U$ , which previous models suggest should increase the benefits of sex (Kondrashov 1982; Charlesworth 1990; Agrawal 2001; Siller 2001; Keightley & Otto 2006), can have the opposite effect and actually increase the net cost of sex.

Lastly, we can ask how the evolution of sex-limited expression, as a long-term consequence of sexual antagonism, will influence the adaptation of females. Under the assumption that sexually antagonistic selection is completely resolved by the evolution of sex-limitation, sex-specific selection produces a net benefit when:

$$\alpha_C > \frac{1}{1-2\psi} \quad (19)$$

(recall that  $\psi$  refers to the genome-wide proportion of female-expressed loci that are sex-limited). This result suggests that fitness costs of sexual antagonism cannot be completely resolved by the evolution of sex-limitation. Rather, sex-limitation can generate a net cost of sex-specific selection, despite strong sexually concordant selection at loci expressed in both sexes (Fig. 6.2).

## DISCUSSION

Our model derives the costs and benefits of sex-specific selection with respect to three key parameters: the relative strength of sexually concordant selection in males and females ( $\alpha_C = s_m/s_f$ ), the relative strength of sexually antagonistic selection in males and females ( $\alpha_A = t_m/s_f$ ), and the proportion of the genome exposed to sexual antagonism ( $\zeta$ ).

To empirically test whether sexual selection provides a long-term benefit or an additional cost of sex, we would ideally analyze sex-specific fitness estimates for a large number of mutations throughout the genome. At present, such data do not exist (Whitlock & Agrawal 2009). However, several other lines of existing evidence have strong implications for this issue. Below, we analyze three independent empirical approaches, discuss their current implications and limitations with respect to quantifying net benefits and costs of sex-specific selection, and propose additional experiments that might shed further light on the subject.

### **Sex-Specific Selection Coefficients from Visible Mutations in *Drosophila***

Experiments using the fruit fly *Drosophila melanogaster* indicate that visible mutations typically produce different fitness consequences in males relative to females (Whitlock & Bourguet 2000; Pischedda & Chippindale 2005; Sharp & Agrawal 2008). These mutations have effects ranging from sexual concordance to sexual antagonism (Table 6.1). When selection is sexually concordant, it tends to be stronger in males relative to females (with mean  $\alpha_C$  between 2.3 and 2.4, approximately). However, sexually antagonistic selection is also relatively common ( $\zeta$  is approximately 0.3). This combination of sex-specific fitness effects is expected to generate a net fitness cost of sex-specific selection, even when sexual antagonism is weak (i.e.,  $\alpha_A \ll 1$ ).

However, it is unlikely that these data can be interpreted in such a straightforward manner. Not only is this sample size of 14 mutations small and limited to a single species, but these mutations are also potentially not representative of most spontaneous mutations within *Drosophila*. Each mutation produces a marked phenotypic and fitness effect, with selection coefficient estimates that are considerably larger than those estimated from mutation-accumulation experiments (e.g., Shabalina et al. 1997). Since most of the mutations affect eyes, wings, or body pigmentation, they are also likely to directly influence mating interactions between the sexes. This should exacerbate the deleterious fitness effects of these mutations in males, thus upwardly biasing estimates of the intensity of concordant selection in males relative to females (Hollis et al. 2009). Furthermore, mutations with large fitness effects are expected to be deleterious to both sexes (i.e., as an extension of Fisher's Geometric Model; Fisher 1930). Thus, visible

mutations might have a higher incidence of sexually concordant selection, relative to a random collection of mutations with smaller phenotypic and fitness effects.

Though visible-mutation data imply that sex-specific selection might generate a net fitness cost for females, the caveats listed above currently preclude any definitive conclusions. Future studies in *Drosophila* could greatly improve upon the strength of our inferences by selecting larger samples of random mutations for analysis. This could be accomplished via gene knockout-deficiency genotypes (readily available through *Drosophila* stock centers; Presgraves 2003), or by using RNA interference libraries to silence random sets of genes (Dietzl et al. 2007). Detecting sex-specific fitness effects of individual mutations might also be approached indirectly via experimental evolution within population cages, as shown in two recent studies (Stewart et al. 2005; Hollis et al. 2009).

### **Mutation Accumulation and Sex-Specific Fitness**

A recent study using *Drosophila melanogaster* took a novel and promising approach towards quantifying the sex-specific effects of individual mutations (Morrow et al. 2008). Adopting a sex-limited “Middle Class Neighborhood” design, the authors permitted spontaneous mutations to accumulate in both sexes, but limited selection to either males or females. In each experimental lineage, they observed a net fitness decline in both sexes, with a faster rate of decline in the sex not subject to selection (hereafter referred to as the “unselected sex”). This pattern was symmetric: the rate of fitness decline in females was twice as high when they were the unselected sex relative to when they were the selected sex. Likewise, fitness of males declined at a twofold higher rate when they were the unselected sex relative to when they were the selected sex.

These results have three major implications for inferring sex-specific fitness effects. First, they suggest that a majority of mutations are deleterious to both sexes (i.e., they are sexually concordant). Second, a faster rate of fitness decline for the unselected sex indicates that a nontrivial proportion of the genome has sex-limited and/or sexually antagonistic effects on fitness. Third, the pattern of symmetry (i.e., that fitness declines about twice as fast for the unselected sex, whether male or female) suggests that strength of purifying selection does not generally differ between the sexes (i.e.,  $\alpha_C \approx 1$ ).

This final point can be illustrated with a simple model. Consider a scenario of sexually concordant selection with sex-specific selection coefficients:  $s_f \alpha_C = s_m$ . When selection is experimentally limited to males, the rate of fitness decline in sex  $i$  (where  $i = f$  in the case of females;  $i = m$  in males) is expected to be:

$$\Delta w_i(M) = \frac{2s_i h \mu \left[ 1 - \frac{\alpha_C}{1 + \alpha_C} \right]}{1 - 2\hat{p}s_i h} \quad (20)$$

(see Appendix 6.4 for details). Similarly, when selection is limited to females, the fitness decline in each sex occurs at the rate:

$$\Delta w_i(F) = \frac{2s_i h \mu \left[ 1 - \frac{1}{1 + \alpha_C} \right]}{1 - 2\hat{p}s_i h} \quad (21)$$

The rate of fitness decline when selection is limited to males relative to the case where selection is limited to females is therefore equal to:

$$\frac{\Delta w_i(M)}{\Delta w_i(F)} = \frac{1}{\alpha_C} \quad (22)$$

If  $\alpha_C > 1$ , females will benefit by being the unselected sex, whereas males will suffer a cost of being the unselected sex. In females, this should at least partially offset fitness costs arising from the accumulation of female-deleterious variation at sexually antagonistic and sex-limited loci. In males, the cost of being the unselected sex will be enhanced by sexually antagonistic and sex-limited loci. The results of Morrow et al. (2008) clearly show that both males and females suffer approximately identical, twofold costs of being the unselected sex. This pattern suggests that the fitness consequences of sexually concordant mutations are approximately equal between males and females. Fitness benefits to females of strong purifying selection in males do not appear to offset costs of sex-specific selection.

The mutation accumulation approach provides a nice complement to studies of individual mutations of large effect in *Drosophila*. An added benefit of the approach is that it can be applied to non-model species. To date, only two additional studies have examined whether sex-specific selection reduces the fitness costs of mutation accumulation. These studies (both using the bulb mite, *Rhizoglyphus robini*) yield conflicting evidence: one reports a benefit to embryo viability (Radwan 2004), whereas

the other reports no effect on female fecundity (Radwan et al. 2004). Future studies, using spontaneous and induced mutation accumulation in a wider variety of animal populations, could further illuminate patterns of sex-specific selection across the genome.

### **Sex-Specific Selection Gradients Estimated Within Natural Populations**

The *Drosophila* data discussed above indicate that benefits of sex-specific selection are unlikely to offset its costs. However, this inference is drawn solely from laboratory-adapted populations of a single insect species. Though the genetic resources available for *Drosophila* are generally lacking in non-model organisms, studies of selection in the field have produced a large dataset of sex-specific selection estimates from a wide variety of natural animal populations. Below, we use this extensive sex-specific selection gradient dataset and present an analysis of sex-specific selection in the wild. The resulting analyses are tentatively used as a proxy for inferring patterns of sex-specific selection throughout non-model animal genomes.

We used a large dataset of 423 sex-specific measures of selection acting on 90 traits from 34 animal species (the full dataset is presented by Cox & Calsbeek 2009) to estimate model parameters ( $\alpha_C$ ,  $\alpha_A$ ,  $\zeta$ ) from several different subsets of these data. First, we treated each reported selection gradient or differential as an independent observation (Appendix A of Cox & Calsbeek 2009). This approach maximized the inclusion of available data (423 estimates), but many of these estimates comprise spatial or temporal replicates of the same traits measured in the same species, and therefore cannot be considered independent observations. Thus, we repeated our estimates of model parameters using a smaller dataset in which a single mean value of selection was derived for any replicated measures (Appendix B of Cox & Calsbeek 2009). This yielded a smaller dataset (203 estimates), but one free from multiple counting due to replicated measures. This dataset also includes estimates of net selection obtained by treating gradients and differentials from individual fitness components (i.e., viability, fecundity, mating success) as additive (see Cox & Calsbeek 2009 for details).

Our selection gradient analysis yields two interesting results. First, the strength of selection in males is approximately equal to the strength of selection in females for both sexually concordant and antagonistically selected traits (Table 6.2). This pattern is

consistent across different classes of traits, and under different types of selection (e.g., viability, fecundity, mating success, and total “net” fitness). Furthermore, the relative strength of concordant and antagonistic selection remains approximately the same for traits exposed to statistically significant directional selection in both sexes (i.e., those with selection gradient estimates that are significantly different than zero; Table 6.2).

Secondly, sexually antagonistic selection is common relative to sexually concordant selection, with the proportion of traits subject to sexual antagonism ranging between 25 and 55 percent (Table 6.2). Because interest in sexually antagonistic processes has expanded rapidly during the past decade (Rice & Chippindale 2001; Bonduriansky & Chenoweth 2009), it is possible that recent publications are biased towards cases of sexual antagonism. However, when the analysis is restricted to studies published before the year 2000, we found that reports of sexual antagonism were slightly more frequent ( $\zeta = 0.47$  for pre-2000 traits;  $\zeta = 0.53$  for pre-2000 traits under statistically significant directional selection)

Patterns of sex-specific selection in the wild are generally consistent with the idea that benefits of strong sexually concordant selection are relatively weak. The prevalence of unresolved sexual antagonism is therefore expected to generate a net cost of sex-specific selection. This conclusion comes with the caveat that selection gradients are based on quantitative traits, which have complicated polygenic and environmental developmental bases, and therefore limit our ability to make direct connections between genotype and phenotype. Although selection gradient data should be considered as complimentary and not necessarily equivalent to genotype-based fitness estimates, they appear to exhibit similar patterns of sex-specific selection as genetic data from *Drosophila* (see above). Both lines of evidence suggest that there might be a net cost of sex-specific selection.

## **Conclusion**

Our mathematical results show that a relatively small proportion of unresolved sexual antagonism will often overwhelm any fitness benefits arising from strong sexually concordant selection. Furthermore, the resolution of sexual antagonism via sex-limited gene expression is also likely to generate long-term fitness costs for sexually reproducing

populations. This suggests that sex-specific selection should generally reduce the fitness of females and add to the inherent twofold cost of sex, but whether this is actually true of most populations is currently unclear. However, a growing body of research suggests that sexually antagonistic selection is ongoing and detectable within animal genomes (see above; Chippindale et al. 2001; Rice & Chippindale 2001; Prasad et al. 2007; Foerster et al. 2007; Brommer et al. 2007; Cox & Calsbeek 2009; Bonduriansky & Chenoweth 2009). Unresolved sexual antagonism is clearly costly and should represent an adaptive constraint for sexually reproducing species. Evidence demonstrating benefits of stronger sexually concordant selection in males than females is substantially weaker. Such benefits have been found in some cases (e.g., Promislow 1998; Dolgin et al. 2006; Whitlock & Agrawal 2009; Hollis et al. 2009), but most evidence suggests that such benefits are relatively weak (see above; Holland 2002; Rundle et al. 2006; Fricke & Arnqvist 2007; Candolin & Heuschele 2008; Maklakov et al. 2009). In line with predictions from our model and the current lack of support for strong adaptive benefits of sexually concordant selection, we conclude that sex-specific selection is unlikely to yield a net benefit to most sexual species. Instead, the fitness costs of sexual antagonism may often increase the inherent twofold cost of sex.

This does not imply that sexual reproduction is incapable of providing benefits that balance multiple severe costs. Indeed, several possible evolutionary mechanisms, including the Red Queen (Hamilton 1980; Agrawal 2006) or interactions between recombination and purifying selection (Kimura & Maruyama 1966; Felsenstein 1974; Kondrashov 1982; Keightley & Otto 2006), can provide long-term benefits to sexual populations. Nevertheless, benefits derived from these possible mechanisms must be substantial to outweigh high costs associated with sexual reproduction.

Finally, the balance of benefits and costs of sex-specific selection can potentially differ among sexually reproducing species. Opportunities for sexually antagonistic selection might vary between species where the sexes have similar strategies for maximizing fitness, relative to species with fitness landscapes that are highly discordant between males and females. Most studies emphasize sex-specific selection related to mating success, with males selected to maximize their number of mates, and females selected to increase mate quality (Trivers 1972), or decrease mating frequency (Holland

& Rice 1998). Sex differences in selection can also arise from ecological differences in species where males and females systematically inhabit different environments (e.g., Trivers & Willard 1973; Bull & Charnov 1977), encounter different sources of mortality (Magnhagen 1991), or exploit different foraging strategies (Shine 1989). While previous authors have speculated that such ecological differentiation might promote population growth by reducing intra-specific competition (e.g., Selander 1966), ecological differences between males and females might instead exacerbate sexually antagonistic selection or promote genomic expansion of sex-limited loci. From the perspective of population genetics, these consequences of sex-specific selection can constrain adaptation, and potentially reduce population productivity.

Table 6.1. Point estimates of sex-specific selection on visible mutations in *Drosophila*. Single estimates of selection are shown when fitness is measured in a single environment. Ranges are reported when selection was estimated in multiple environments.

---

Concordantly Selected mutations

Mutation	$s_f$	$s_m$	$a_c$	Source
<i>b</i>	0.08	0.34	4.43	[1]
<i>Dr</i>	0.13	0.75 - 0.80	5.79 - 6.15	[3]
<i>e/sr</i>	0.46	0.76	1.65	[1]
<i>Frd</i>	0.22	0.20 - 0.22	0.92 - 1.00	[3]
<i>Gla</i>	0.53	0.78 - 0.84	1.46 - 1.58	[3]
<i>L</i>	0.26	0.62 - 0.68	2.39 - 2.63	[3]
<i>Ly</i>	0.53	0.80 - 0.88	1.51 - 1.67	[3]
<i>nub<sup>1</sup></i>	0.38	0.74	1.95	[2]
<i>Pin</i>	0.22	0.04 - 0.06	0.18 - 0.26	[3]
<i>Sb</i>	0.19	0.48 - 0.52	2.50 - 2.73	[3]

---

---

Antagonistically Selected mutations

Mutation	$s_f$	$t_m$	$a_A$	Source
<i>ca</i>	0.018	0.38	21.37	[1]
<i>h</i>	0.52	0.12	0.23	[1]
<i>px/sp</i>	0.46	0.61	1.33	[1]
<i>U</i>	0.20	0.12 - 0.19	0.59 - 0.97	[3]

---

[1] Selection coefficients based on data from Whitlock & Bourguet (2000; their Table 1); productivity (fecundity & offspring survival) is the measure of female fitness; competitive mating success is the measure of male fitness.

[2] Selection coefficients based on total sex-specific fitness estimates reported by Pischedda & Chippindale (2005).

[3] Selection coefficients based on data from Sharp & Agrawal (2008; their Table 3); total fitness assumes juvenile viability and adult reproductive success interact multiplicatively.

Table 6.2. Parameter estimates (means and 95% confidence intervals) based on sex-specific selection gradient data from natural animal populations.

	$n$	$a_C^{(1)}$	$a_A^{(1)}$	$z^{(2)}$
Traits Analyzed <sup>(3)</sup>				
All	423	1.20 (1.05, 1.37)	0.98 (0.83, 1.15)	0.41 (0.36, 0.46)
Statistically Sig.	33	1.28 (1.04, 1.58)	0.87 (0.66, 1.14)	0.39 (0.25, 0.56)
Fitness Component <sup>(4)</sup>				
Net Selection	47	1.03 (0.69, 1.55)	1.02 (0.62, 1.69)	0.47 (0.33, 0.61)
Fecundity	39	0.97 (0.66, 1.42)	0.86 (0.45, 1.57)	0.26 (0.15, 0.41)
Mating Success	28	1.34 (0.86, 2.13)	1.06 (0.57, 1.99)	0.54 (0.36, 0.71)

Viability	71	0.81 (0.57, 1.14)	1.21 (0.72, 2.07)	0.34 (0.24, 0.45)
-----------	----	-------------------	-------------------	-------------------

---

<sup>(1)</sup> Averages and 95% confidence intervals were calculated for the variable  $x = S_m/(S_m + S_f)$ , where  $S_m$  and  $S_f$  are the male and female selection gradients, respectively. Means and 95% confidence intervals for  $x$  were used to calculate equivalent statistics for  $S_m/S_f$  by substituting  $S_f\alpha_C$  for  $S_m$ , which leads to the relationship:  $\alpha_C = S_m/S_f = x/(1 - x)$ . The same approach was applied to cases of sexually antagonistic selection, with absolute values of selection gradients.

<sup>(2)</sup>  $\zeta$  is estimated as the proportion of traits where male and female selection gradients had opposite signs (i.e.,  $S_m/S_f < 0$ ). Confidence intervals for  $\zeta$  were obtained by hand, using the chi-square distribution.

<sup>(3)</sup> Data are from “dataset A”, as described in Cox & Calsbeek (2009); the “significant” category refers to traits with statistically significant selection gradients in both sexes.

<sup>(4)</sup> Data are from “dataset B”, as described in Cox & Calsbeek (2009).

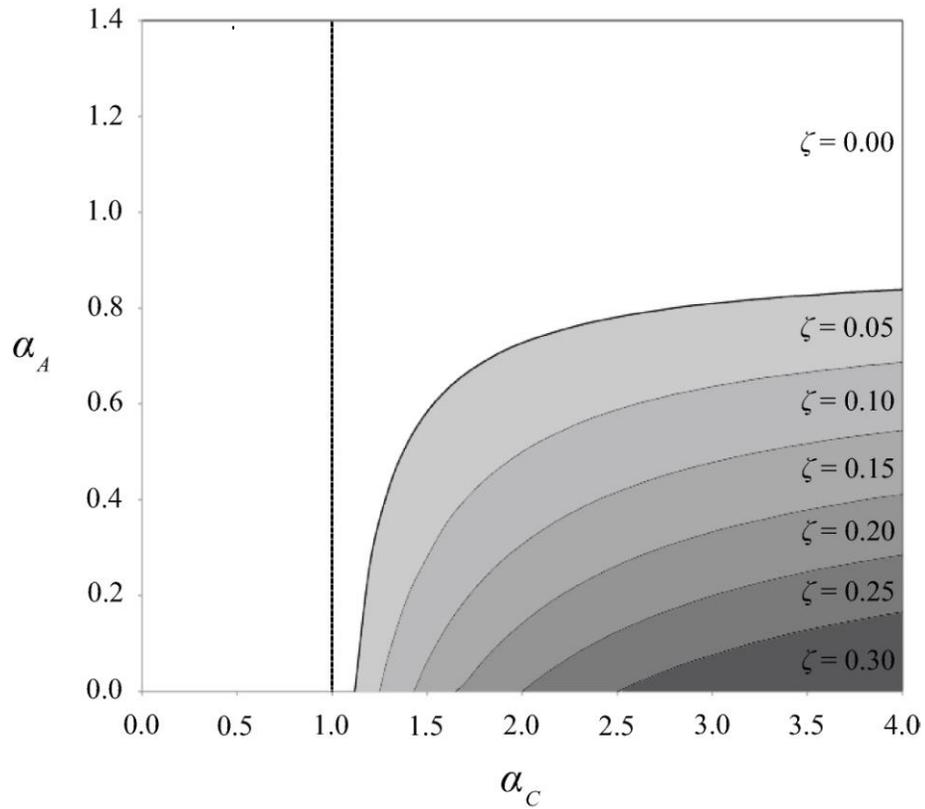


Figure 6.1. Benefits of sex-specific selection are severely constrained by sexual antagonism. The parameters  $\alpha_A$  and  $\alpha_C$  refer (respectively) to the strength of sexually antagonistic and sexually concordant selection in males relative to females (details provided within the text). Space below each curve corresponds to conditions under which sex-specific selection provides a net benefit of sex. When all genes evolve under sexually concordant selection ( $\zeta = 0$ ), mean female fitness increases when selection is stronger in males, as indicated by the dotted line at  $\alpha_C = 1$ . Shading indicates the proportion of the genome that experiences sexual antagonism ( $\zeta$ ). Results were obtained with equation (17) from the text.

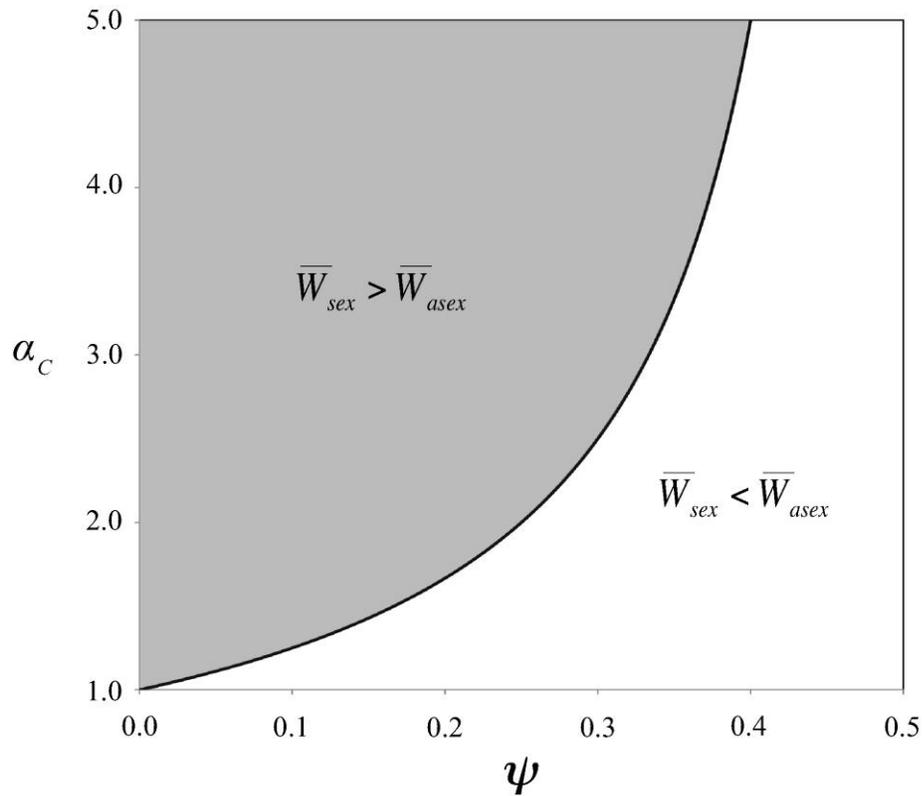


Figure 6.2. Sex-limitation reduces opportunities for benefits of sex-specific selection. Conditions where sex-specific selection provides a net fitness benefit to females will become reduced as the sex-limited proportion of the genome ( $\psi$ ) increases. Results were obtained with equation (19), under the assumption that sexual antagonism is completely resolved by sex-limitation. The addition of unresolved sexual antagonism further reduces the parameter space where sex-specific selection is beneficial.

## APPENDIX 6.1. Equilibrium conditions for sexually antagonistic alleles

Under sexually antagonistic selection, genetic variation is maintained at three possible equilibria: (A) when selection in females is stronger than selection in males, the male-beneficial allele,  $A_1$ , occurs at low frequency at mutation-selection balance; (B) when selection in males is stronger than selection in females, the female beneficial allele,  $A_2$ , occurs at low frequency at mutation-selection balance; (C) when selection is similarly strong in males and females, male- and female-beneficial alleles are maintained at intermediate frequencies by balancing selection.

Following Kidwell et al. (1977), the zygotic frequencies of the three genotypes are:  $[A_1A_1] = p_m p_f$ ,  $[A_1A_2] = p_m(1 - p_f) + p_f(1 - p_m)$ , and  $[A_2A_2] = (1 - p_m)(1 - p_f)$ , respectively. Note that  $p_m$  and  $p_f$  refer to the frequency of the  $A_1$  allele among breeding males and females. After selection, the frequency of the  $A_1$  allele in males is:

$$p_m' = \frac{2p_m p_f + [p_m(1 - p_f) + p_f(1 - p_m)](1 - h_m t_m)}{2\bar{w}_m} \quad (1a)$$

where

$$\bar{w}_m = 2p_m p_f + [p_m(1 - p_f) + p_f(1 - p_m)](1 - h_m t_m) + (1 - p_m)(1 - p_f)(1 - t_m) \quad (1b)$$

The frequency of the  $A_1$  allele in females, after selection, is:

$$p_f' = \frac{2p_m p_f(1 - s_f) + [p_m(1 - p_f) + p_f(1 - p_m)](1 - h_f s_f)}{2\bar{w}_f} \quad (2a)$$

where

$$\bar{w}_f = 2p_m p_f(1 - s_f) + [p_m(1 - p_f) + p_f(1 - p_m)](1 - h_f s_f) + (1 - p_m)(1 - p_f) \quad (2b)$$

By ignoring mutation from  $A_1 \rightarrow A_2$ , we can examine stability of the equilibrium  $\hat{p}_m = \hat{p}_f = 0$  (i.e., the condition where the male beneficial allele is absent from the population). The Jacobian matrix is:

$$\begin{pmatrix} \frac{(1 - t_m h_m)}{2(1 - t_m)} & \frac{(1 - t_m h_m)}{2(1 - t_m)} \\ \frac{1 - s_f h_f}{2} & \frac{1 - s_f h_f}{2} \end{pmatrix}$$

which has a dominant eigenvalue:

$$\lambda = \frac{(1 - t_m h_m)}{2(1 - t_m)} + \frac{1 - s_f h_f}{2} \quad (3)$$

Female selection dominates and the male beneficial allele will not invade the population when:

$$s_f > \frac{t_m(1-h_m)}{h_f(1-t_m)} \quad (4a)$$

Following the same approach for the equilibrium  $\hat{p}_m = \hat{p}_f = 1$ , male selection dominates and the male-beneficial approaches fixation when:

$$s_f < \frac{t_m h_m}{1-h_f+t_m h_m} \quad (4b)$$

A balanced polymorphism occurs (and  $A_I$  remains at intermediate frequency) when:

$$\frac{t_m h_m}{1-h_f+t_m h_m} < s_f < \frac{t_m(1-h_m)}{h_f(1-t_m)} \quad (4c)$$

Under relatively weak selection, and when the sexually antagonistic alleles show the same dominance in males and females (under our parameterization, when  $h_f = 1 - h_m$ ), there is virtually no opportunity for a balanced polymorphism. As described in the main document, the male-beneficial allele remains at low frequency, at mutation-selection balance, when  $\alpha_A < 1$ . The male-beneficial allele is nearly fixed (the female-beneficial is maintained at mutation-selection balance) when  $\alpha_A > 1$ .

**APPENDIX 6.2. Sexually antagonistic mutation-selection equilibria under weak selection ( $s_f, t_m < 0.1$ ) and equal dominance for  $A_I$  between the sexes ( $h_f = 1 - h_m$ )**

Under weak sexual antagonism ( $\alpha_A < 1$ ), the evolution of a rare male-beneficial/female-detrimental allele ( $A_I$ ) is described by the recursion:

$$p_{x+1} = \frac{p_x(1 - s_f h_f) + p_x(1 - t_m h_m)/(1 - t_m)}{2} + \mu \quad (1)$$

where  $x$  refers to the generation and  $\mu$  refers to the mutation rate from  $A_2$  to  $A_I$  (backmutation is negligible because  $A_I$  is rare). The frequency, at equilibrium, is:

$$\hat{p} = \frac{2\mu(1 - t_m)}{s_f h_f - t_m s_f h_f - t_m h_f} \approx \frac{2\mu}{s_f h_f(1 - \alpha_A)} \quad (2)$$

Under strong sexual antagonism ( $\alpha_A > 1$ ), the evolution of a rare female-beneficial/male-detrimental allele ( $A_2$ ) is described by the recursion:

$$p_{x+1} = \frac{p_x(1 - t_m h_m) + p_x(1 - s_f h_f)/(1 - s_f)}{2} + \mu \quad (3)$$

where  $\mu$  refers to the mutation rate from  $A_I$  to  $A_2$ . The equilibrium frequency is:

$$\hat{p} = \frac{2\mu(1 - s_f)}{[s_f(\alpha_A - 1) - \alpha_A s_f^2](1 - h_f)} \approx \frac{2\mu}{s_f(\alpha_A - 1)(1 - h_f)} \quad (4)$$

### APPENDIX 6.3. Simulation results.

To examine the impact of strong selection and sex-specific dominance, we calculated exact equilibrium allele frequencies under sexually antagonistic selection, symmetrical mutation rates (i.e., mutation rate from  $A_1 \rightarrow A_2 = A_2 \rightarrow A_1$ ), and a wide range of male and female selection and dominance coefficients.

We simulated evolution of a female-deleterious allele that evolves according to the recursion:

$$p_{x+1} = (p_f + p_m)(1 - \mu)/2 + [1 - (p_f + p_m)/2]\mu \quad (10)$$

where  $x$  refers to the generation,

$$p_f = \frac{p_x^2(1 - s_f) + p_x(1 - p_x)(1 - s_f h_f)}{p_x^2(1 - s_f) + 2p_x(1 - p_x)(1 - s_f h_f) + (1 - p_x)^2} \quad (11a)$$

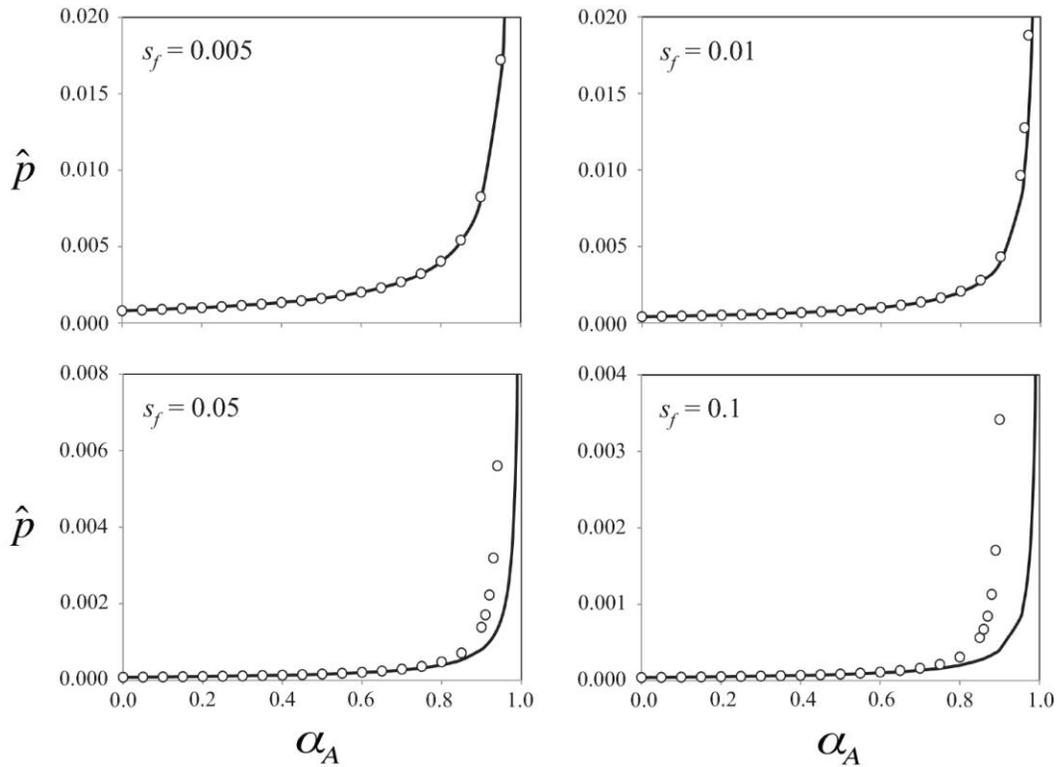
and

$$p_m = \frac{p_x^2 + p_x(1 - p_x)(1 - t_m h_m)}{p_x^2 + 2p_x(1 - p_x)(1 - t_m h_m) + (1 - p_x)^2(1 - t_m)} \quad (11b)$$

Equilibrium conditions were calculated by carrying out recursions until  $p_x = p_{x+1}$ .

**Strong Selection.** We first examine equilibrium conditions under additive expression ( $h_m = h_f = 0.5$ ) and compare these results to our weak selection approximation (Eq. 6 of the main manuscript). Representative results are presented in Figure A6.3.1.

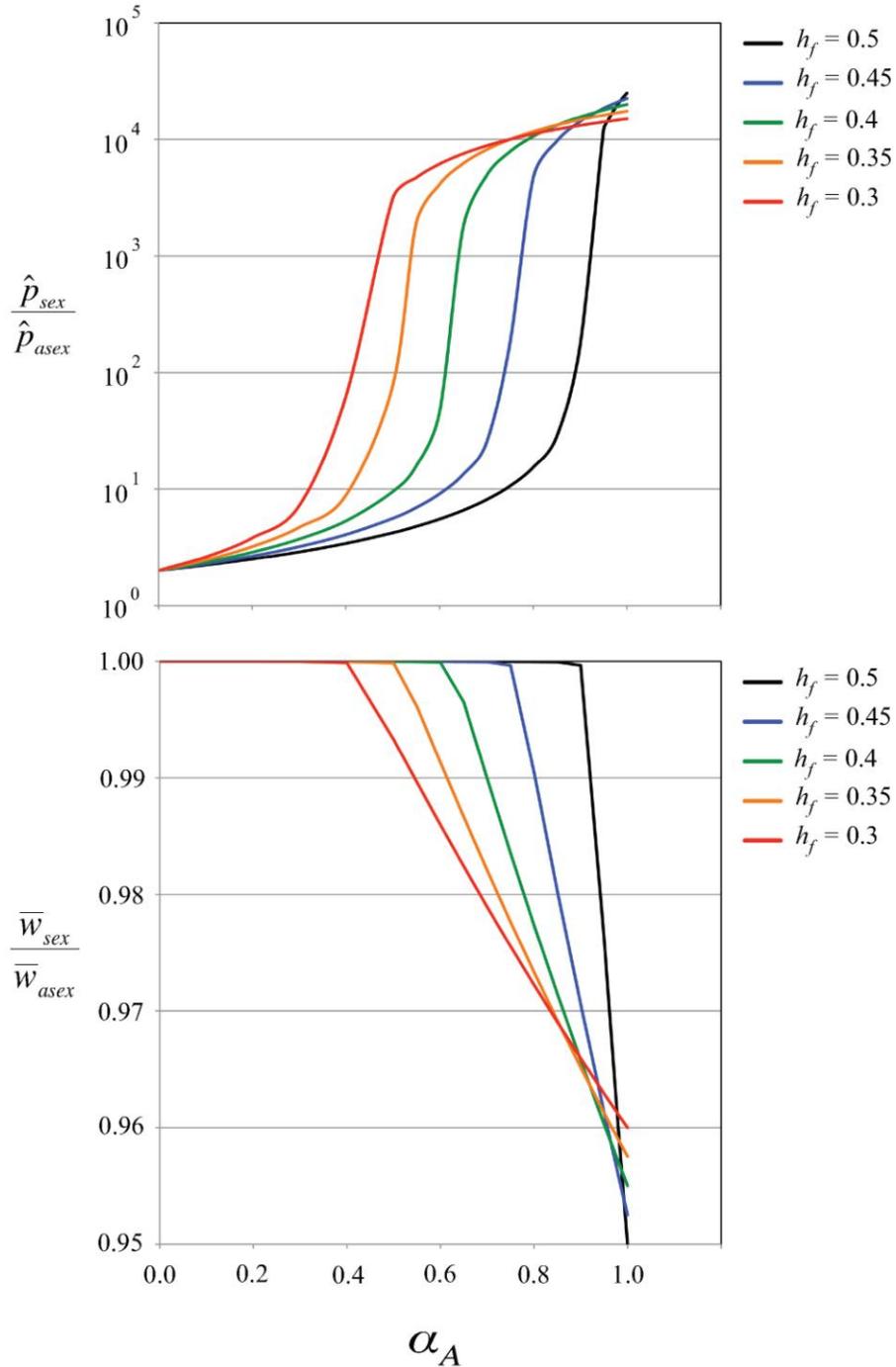
Approximations of equilibrium conditions are extremely good when selection coefficients are small (e.g.,  $s_f < 0.05$ ), or when  $\alpha_A \leq 0.8$ . When these conditions are violated, male-beneficial alleles increase within the population to higher frequencies than predicted by the mutation-selection approximation. Consequently, the parameter conditions that provide a net benefit of sexual selection (those represented in Fig. 1 from the main manuscript) will be constricted as the strength of selection increases. Thus, the conclusions presented in the manuscript are conservative.



**Figure A6.3.1.** Comparison of exact simulated equilibrium conditions with analytical approximations of the frequency ( $\hat{p}$ ) of a male-beneficial/female-deleterious allele under additive sexually antagonistic selection. Approximations are represented as solid lines and are calculated with Eq. (6) from the main manuscript. Simulated results are represented as circles. Results shown use a mutation rate of  $\mu = 10^{-6}$ , but do not depend upon the specific mutation rate.

**Sex-Specific Dominance.** The outcome of sexually antagonistic selection can be complex when sexually antagonistic alleles have nonadditive fitness effects. To examine the effect of sex-specific dominance we considered a dominance reversal scenario, where male-beneficial alleles are dominant in males and recessive in females. Dominance reversals occur in our model when  $h_m = h_f$  (i.e., the  $A_1$  allele is dominant in males, and  $A_2$  is dominant in females). Such conditions favor invasion of the male-beneficial allele (Kidwell et al. 1977). Our simulations show that dominance reversals exacerbate the cost of sexual antagonism by facilitating invasion of the male-beneficial alleles (Figure A6.3.2), even when sexual antagonism is weak:  $\alpha_A < 1$ . Invasion of male-

beneficial/female-detrimental alleles exacerbates sexually antagonistic costs to female fitness, making our main conclusions conservative.



**Figure A6.3.2.** Representative equilibrium conditions for sexually antagonistic selection with dominance reversals between the sexes. The upper panel shows the relative increase in frequency arising from positive selection in males of a female-deleterious allele. The

asexual frequency is maintained by mutation-selection balance. The lower panel shows mean female fitness in a sexual population subject to sexual antagonism, relative to mean fitness in a population of asexuals. Dominance reversals, where  $A_I$  is dominant in males but recessive in females, occur (and are symmetrical) when  $h_m = h_f$ . The displayed simulations were carried out with a homozygous female selection coefficient of  $s_f = 0.1$  and a mutation rate of  $u = 10^{-6}$ . Mean fitness is for a single locus.

**Appendix 6.4. Sex-specific fitness consequences of sex-limited mutation accumulation.**

When purifying selection is permitted to occur in males but not females, the rate of increase of a sexually concordant deleterious allele is:

$$\Delta p(M) = \mu - \frac{s_f h \alpha_C}{2} \hat{p}(1 - \hat{p}) \quad (1a)$$

Substituting  $\hat{p} \approx \frac{2\mu}{s_f h(1 + \alpha_C)}$ , this reduces to:

$$\Delta p(M) = \mu - \frac{\alpha_C \mu}{(1 + \alpha_C)} \left( 1 - \frac{2\mu}{s_f h(1 + \alpha_C)} \right) \approx \mu \left( 1 - \frac{\alpha_C}{(1 + \alpha_C)} \right) \quad (1b)$$

When purifying selection is permitted to occur in females but not males, the rate of increase of a sexually concordant deleterious allele is:

$$\Delta p(F) = \mu - \frac{s_f h}{2} \hat{p}(1 - \hat{p}) \quad (2a)$$

Substituting  $\hat{p} \approx \frac{2\mu}{s_f h(1 + \alpha_C)}$ , this reduces to:

$$\Delta p(F) = \mu - \frac{\mu}{(1 + \alpha_C)} \left( 1 - \frac{2\mu}{s_f h(1 + \alpha_C)} \right) \approx \mu \left( 1 - \frac{1}{(1 + \alpha_C)} \right) \quad (2b)$$

Fitness declines at a sex-specific rate of:

$$\Delta W_i = \frac{\bar{W}_{i(t)} - \bar{W}_{i(t+1)}}{\bar{W}_{i(t)}} \quad (3)$$

where  $i = f$  in the case of females,  $i = m$  in the case of males,  $t$  refers to the generation, and:

$$\bar{w}_{i(t)} = p_i^2(1 - s_i) + 2p_i(1 - p_i)(1 - s_i h) + (1 - p_i)^2 \approx 1 - 2p_i s_i h \quad (4)$$

The approximation (right side of eq. 4) assumes that deleterious mutations occur at low frequency within the population, as seems reasonable.

By letting  $p_t = \hat{p}$  and  $p_{t+1} = \hat{p} + \Delta p$ , and substituting eq. (4) into eq. (3), the sex-specific rate of fitness decline reduces to:

$$\Delta w_i = \frac{2\Delta p s_i h}{1 - 2\hat{p} s_i h} \quad (5)$$

The rate of female fitness decline when selection is limited to males, compared to when selection is limited to females, is:

$$\frac{\Delta w_f(M)}{\Delta w_f(F)} = \frac{1}{\alpha_C} \quad (6)$$

The rate of male fitness decline when selection is limited to females, compared to when selection is limited to males, is:

$$\frac{\Delta w_m(F)}{\Delta w_m(M)} = \alpha_C \quad (7)$$

## REFERENCES

- Agrawal AF. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature* 411:692-695.
- Agrawal AF. 2006. Similarity selection and the evolution of sex: revisiting the Red Queen. *PLoS Biology* 4:e265.
- Andolfatto P. 2007. Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genome Research* 17:1755-1762.
- Bell G. 1982. *The masterpiece of nature*. University of California Press, Berkeley.
- Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. *Trends in Ecology and Evolution* 24:280-288.
- Brommer JE, Kirkpatrick M, Qvarnström A, Gustafsson L. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. *PLoS One* 2:e744.
- Bull JJ, Charnov EL. 1977. When is sex environmentally determined? *Nature* 266:828-830.
- Candolin U, Heuschele J. 2008. Is sexual selection beneficial during adaptation to environmental change? *Trends in Ecology & Evolution* 23:446-452.
- Charlesworth B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genetical Research* 55:199-221.
- Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proceedings of the National Academy of Science USA* 98:1671-1675.
- Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* 173:176-187.
- Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oettel S, Scheiblauer S, Couto A, Marra V, Keleman K, Dickson BJ. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448:151-156.
- Dolgin ES, Whitlock MC, Agrawal AF. 2006. Male *Drosophila melanogaster* have higher mating success when adapted to their thermal environment. *Journal of Evolutionary Biology* 19:1894-1900.

- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.
- Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press, London.
- Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton Brock TH, Kruuk LEB. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447:1107-1119.
- Fricke C, Arnqvist G. 2007. Rapid adaptation to a novel host in a seed beetle (*Callosobruchus maculatus*): the role of sexual selection. *Evolution* 62:440-454.
- Haag-Liautard C, Dorris M, Maside X, Macaskill S, Halligan DL, Charlesworth B, Keightley PD. 2007. Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* 445:82-85.
- Hadany L, Beker T. 2007. Sexual selection and the evolution of obligatory sex. *BMC Evolutionary Biology* 7:e245.
- Hamilton WD. 1980. Sex versus non-sex versus parasite. *Oikos* 35:282-290.
- Holland B. 2002. Sexual selection fails to promote adaptation to a new environment. *Evolution* 56:721-730.
- Holland B, Rice WR. 1998. Perspective: Chase-away sexual selection: antagonistic seduction versus resistance. *Evolution* 52:1-7.
- Hollis B, Fierst JL, Houle D. 2009. Sexual selection accelerates the elimination of a deleterious mutant in *Drosophila melanogaster*. *Evolution* 63:324-333.
- Keightley PD, Otto SP. 2006. Interference among deleterious mutations favors sex and recombination in finite populations. *Nature* 443:89-92.
- Kidwell JF, Clegg MT, Stewart FM, Prout T. 1977. Regions of stable equilibria for models of differential selection in two sexes under random mating. *Genetics* 85:171-183.
- Kimura M, Maruyama T. 1966. The mutational load with epistatic gene interactions in fitness. *Genetics* 54:1337-1351.
- Koeslag PD, Koeslag JH. 1994. Koinophilia stabilizes bi-gender sexual reproduction against asex in an unchanging environment. *Journal of Theoretical Biology* 166:251-260.

- Kondrashov AS, Crow JF. 1988. King's formula for the mutation load with epistasis. *Genetics* 120:853-856.
- Kondrashov AS. 1982. Selection against harmful mutations in large sexual and asexual populations. *Genetical Research* 40:325-332.
- Lande R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34:292-305.
- Loewe L, Charlesworth B, Bartolome C, Noel V. 2006. Estimating selection on nonsynonymous mutations. *Genetics* 172:1079-1092.
- Lorch PD, Proulx S, Rowe L, Day T. 2003. Condition-dependent sexual selection can accelerate adaptation. *Evolutionary Ecology Research* 5:867-881.
- Magnhagen C. 1991. Predation risk as a cost of reproduction. *Trends in Ecology & Evolution* 6:183-185.
- Maklakov AA, Bonduriansky R, Brooks RC. 2009. Sex differences, sexual selection and ageing: an experimental evolution approach. *Evolution*, in press.
- Manning JT. 1984. Males and the advantage of sex. *Journal of Theoretical Biology* 108:215-220.
- Maynard Smith J. 1978. *The evolution of sex*. Cambridge University Press, Cambridge, UK.
- Morrow EH, Stewart AD, Rice WR. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *Journal of Evolutionary Biology* 21:1046-1054.
- Otto SP, Bourguet D. 1999. Balanced polymorphisms and the evolution of dominance. *American Naturalist* 153:561-574.
- Pischedda A, Chippindale AK. 2005. Sex, mutation and fitness: asymmetric costs and routes to recovery through compensatory evolution. *Journal of Evolutionary Biology* 18:1115-1122.
- Prasad NG, Bedhomme S, Day T, Chippindale AK. 2007. An evolutionary cost of separate genders revealed by male-limited evolution. *American Naturalist* 169:29-37.
- Presgraves DC. 2003. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics* 163:955-972.
- Promislow DEL, Smith EA, Pearse L. 1998. Adult fitness consequences of sexual

- selection in *Drosophila melanogaster*. Proceedings of the National Academy of Science USA 95:10687-10692.
- Prout T. 1999. How well does opposing selection maintain variation? Pages 157-182 in Lewontin RC, Singh RS, Krimbas CB, Krimpas KV, eds. Evolutionary genetics: From molecules to morphology. Cambridge University Press, Boston.
- Radwan J. 2004. Effectiveness of sexual selection in removing mutations induced with ionizing radiation. Ecology Letters 7:1149-1154.
- Radwan J, Unrug J, Snigorska K, Gawronska K. 2004. Effectiveness of sexual selection in preventing fitness deterioration in bulb mite populations under relaxed natural selection. Journal of Evolutionary Biology 17:94-99.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735-742.
- Rice WR. 2002. Experimental tests of the adaptive significance of sexual recombination. Nature Reviews Genetics 3:241-251.
- Rice WR, Chippindale AK. 2001. Intersexual ontogenetic conflict. Journal of Evolutionary Biology 14:685-693.
- Ridley M. 1996. Evolution, 2<sup>nd</sup> Ed. Blackwell Sciences, Inc., Cambridge, MA.
- Rundle HD, Chenoweth SF, Blows MW. 2006. The roles of natural and sexual selection during adaptation to a novel environment. Evolution 60:2218-2225.
- Selander RK. 1966. Sexual dimorphism and differential niche utilization in birds. Condor 68:113-151
- Shabalina SA, Yampolsky LY, Kondrashov AS. 1997. Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed natural selection. Proceedings of the National Academy of Science USA 94:13034-13039.
- Sharp NP, Agrawal AF. 2008. Mating density and the strength of sexual selection against deleterious alleles in *Drosophila melanogaster*. Evolution 62:857-867.
- Shine R. 1989. Ecological causes for the evolution of sexual dimorphism: a review of the evidence. Quarterly Review of Biology 64:419-461.
- Stewart AD, Morrow EH, Rice WR. 2005. Assessing putative interlocus sexual conflict in *Drosophila melanogaster* using experimental evolution. Proceedings of the National Academy of Sciences USA 102:2029-2035.

- Siller S. 2001. Sexual selection and the maintenance of sex. *Nature* 411:689-692.
- Trivers RL. 1972. Parental investment and sexual selection. Pages 136-179 *in* B. Campbell, ed. *Sexual selection and the descent of man, 1871-1971*. Aldine, Chicago.
- Trivers RL, Willard DE. 1973. Natural selection of parental ability to vary the sex ratio of offspring. *Science* 179:90-92.
- Whitlock MC. 2000. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution* 54:6:1855-1861.
- Whitlock MC, Agrawal AF. 2009. Purging the genome with sexual selection: reducing mutation load through selection in males. *Evolution* 63:569-582.
- Whitlock MC, Bourguet D. 2000. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution* 54:1654-1660.
- Wright S. 1929. Fisher's theory of dominance. *American Naturalist* 63:274-279.

## Chapter 7

### Recombination Rate and Protein Evolution in Yeast

#### ABSTRACT

Theory and artificial selection experiments show that recombination can promote adaptation by enhancing the efficacy of natural selection, but the extent to which recombination affects levels of adaptation across the genome is still an open question. Because patterns of molecular evolution reflect long-term processes of mutation and selection in nature, interactions between recombination rate and genetic differentiation between species can be used to test the benefits of recombination. However, this approach faces a major difficulty: different evolutionary processes (i.e., negative versus positive selection) produce opposing relationships between recombination rate and genetic divergence, and obscure patterns predicted by individual benefits of recombination. We use a combination of polymorphism and genomic data from the yeast *Saccharomyces cerevisiae* to infer the relative importance of nearly-neutral (i.e., slightly deleterious) evolution in different gene categories. For genes with high opportunities for slightly deleterious substitution, recombination substantially reduces the rate of molecular evolution, whereas divergence in genes with little opportunity for slightly deleterious substitution is not strongly affected by recombination. These patterns indicate that adaptation throughout the genome can be strongly influenced by each gene's recombinational environment, and suggest substantial long-term fitness benefits of enhanced purifying selection associated with sexual recombination.

## INTRODUCTION

Genetic drift is expected to overpower natural selection when selection is weak and effective population size ( $N_e$ ) is small (Fisher 1930; Wright 1931; Kimura 1962). Recombination increases the effective population size in which genes evolve by reducing interference between linked loci under selection (Hill & Robertson 1966; Felsenstein 1974). As a result, recombination is expected to facilitate the spread of beneficial mutations and the elimination of deleterious mutations (Rice 2002; Otto & Lenormand 2002). Because recombination rates vary between different regions of a genome (e.g., yeast: Gerton et al. 2000; *Drosophila*: Hey & Kliman 2002; Mammals: Jensen-Seaman et al. 2004; plants: Mezard 2006), adaptation at the molecular level might be strongly affected by each gene's recombinational environment – genes evolving in low recombination regions are expected to be poorly adapted relative to those in high recombination regions (Barton 1995; Presgraves 2005).

Comparative genome analyses are potentially useful for assessing whether recombination promotes adaptation because long-term evolutionary processes are reflected in patterns of genetic divergence between species (Pal et al. 2001; Presgraves 2005; Betancourt & Presgraves 2002; Bachtrog 2004; Marais et al. 2004; Zhang & Parsch 2005; Paland & Lynch 2006). However, genomic approaches face a major challenge – multiple processes can contribute to evolutionary divergence between species and each predicts a different relationship between protein evolution and recombination rate (Table 7.1). The rate of neutral substitution is unaffected by recombination and will tend to reduce correlations between recombination rate and total nucleotide divergence between species (Birky & Walsh 1988), mildly deleterious (i.e., nearly neutral) substitutions will generate a negative correlation between recombination and divergence (Charlesworth 1994), and adaptive substitutions generate a positive correlation between recombination and divergence (Peck 1994).

The relationship between recombination and divergence will be shaped by the predominant process of molecular evolution (i.e., neutral; slightly deleterious; adaptive). To test whether recombination facilitates adaptation throughout the genome (by enhancing purifying and positive selection), genes evolving under purifying selection and those evolving via positive selection should be analyzed separately, as each predicts a

different relationship between divergence and recombination rate. Unfortunately, inferring the processes causing molecular divergence has traditionally been problematic without detailed within- and between-species genetic data (Eyre-Walker 2006), which limits the extent of the genome that can be analyzed.

Here I take an alternative approach. By capitalizing on an extensive volume of yeast (*Saccharomyces* spp.) genomic and polymorphism data, individual genes can be partitioned by their ‘opportunity’ for slightly deleterious evolution, and benefits of recombination can be tested. Furthermore, direct estimates of local recombination rates are available for most genes in the *S. cerevisiae* genome, whole-genome sequencing projects provide data for estimating protein evolutionary rates, and previous studies have revealed that the average strength of selection predictably varies between genes with different functional attributes (reviewed in Pal et al. 2006; see below). Evolutionary theory predicts that mildly deleterious substitutions will accumulate readily in genes subject to weak selection, but not in those subject to strong selection (see Fig. 7.1). To the extent that substitutions differentiating species are often deleterious (e.g., genes with weakly-selected mutations; Fig. 7.1 – small  $|N_e s_d|$ ), increased recombination is expected to decrease the rate of protein evolution by inhibiting the spread of deleterious mutations. However, in genes with little opportunity for slightly deleterious divergence (e.g., genes subject to strong purifying selection; Fig. 7.1 – large  $|N_e s_d|$ ), recombination will increase the rate of divergence by enhancing the spread of beneficial mutations (to the extent that such mutations frequently arise).

## **METHODS**

### **Data**

Publically available polymorphism data was obtained via Genbank (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=Popset>) and the Polymorphix Database (<http://pbil.univ-lyon1.fr/polymorphix/>). Genes with at least four samples from *S. cerevisiae* and at least one polymorphic site were included in the analysis, resulting in a dataset of 35 genes (lower 25% expression,  $n = 11$ ; lower 50% expression,  $n = 12$ ; upper 50% expression,  $n = 23$ ; upper 25% expression,  $n = 17$ ; essential genes,  $n = 7$ ; nonessential genes,  $n = 28$ ), comprising 34443 nonsynonymous, and 9975 synonymous

nucleotide sites. The mean number of samples per gene was  $\bar{n} = 22$ . Orthologous sequences from *S. paradoxus* were obtained via BLAST search at The *Saccharomyces* Genome Database (<http://www.yeastgenome.org/>).

Per gene recombination rates for *S. cerevisiae* are those reported by Gerton et al. (2000; data available at <http://derisilab14.ucsf.edu/hotspots/>). These estimates refer to recombination rates per sexual generation. The total, per generation recombination rate during the evolutionary history of each gene is the product of the rate under sexual reproduction ( $R_{sex}$ ) and the frequency of outcrossing ( $O_c$ );  $R_{TOT} = R_{sex} O_c$  (modified from [43]). Because the exact value of  $O_c$  will be the same for all genes in the genome,  $R_{sex}$  per gene  $i$ , relative to  $R_{sex}$  for other genes, will be the same as  $R_{TOT}$  per gene  $i$ , relative to  $R_{TOT}$  for other genes  $\left( i.e. \frac{R_{TOT}(i)}{R_{TOT}(\max)} = \frac{R_{sex}(i)O_c}{R_{sex}(\max)O_c} = \frac{R_{sex}(i)}{R_{sex}(\max)} \right)$ , and  $R_{sex}$  should accurately capture relative rates of total recombination for each gene. Furthermore, because  $0 < O_c < 1$  (Ruderfer et al. 2006), the critical transition between  $R_{TOT} \approx 0$  to  $R_{TOT} > 0$ , predicted to most strongly impact the efficacy of selection (McVean & Charlesworth 2000), will be represented with yeast genes.

Protein divergence data for *S. cerevisiae* and *S. paradoxus* orthologous genes were kindly provided by D. Allan Drummond (see Drummond et al. 2005 for details). Gene expression values for *S. cerevisiae* were calculated from seven time period estimates during the diauxic shift by DeRisi et al. (1997; data available at <http://cmgm.stanford.edu/pbrown/explore/array.txt>). Average and maximum expression levels for the seven time periods were calculated using the method of Kliman et al. (2003). Results for average gene expression ( $E_{avg}$ ) across the time periods are presented here; results do not differ when maximum expression values ( $E_{max}$ ) are used. Essential genes were identified through the GeneMerge database (Castillo-Davis & Hartl 2003; <http://www.oeb.harvard.edu/hartl/lab/publications/GeneMerge/GeneMerge.html>). Gene length, dispensability, protein-protein interactions and genome map positions were obtained from the *Saccharomyces* Genome Database. Genes with no known interaction partner were excluded from analyses involving PPI as a variable. Space between genes (SBG) was calculated by the method of Hey & Kliman (2002). Recombination, expression, length, SBG and divergence estimates were available for 4786 genes in total,

including essential genes with two or more PPI ( $n = 329$ ), essential genes with 1 PPI ( $n = 195$ ), nonessential genes with two or more PPI ( $n = 762$ ), and nonessential genes with 1 PPI ( $n = 835$ ).

## Analysis

Population samples for each gene were aligned with ClustalW online (<http://www.ch.embnet.org/software/ClustalW-XXL.html>), available online, and manually adjusted.  $P_n$ ,  $P_s$ ,  $D_n$ , and  $D_s$  values were calculated with DnaSP, Version 4.10 (Rozas et al. 2003). Watterson's estimate of silent nucleotide diversity ( $\theta$ ) was calculated by hand (as described in Hein et al. 2005). The complete polymorphism dataset is provided in Appendix 7.1.

Genes were classified into low and high expression level categories, based on quantile partitions. These correspond with ranges of lower 25%: 2.25 to 3.15 ( $n = 1197$ ); lower 50%: 2.25 to 3.31 ( $n = 2393$ ); upper 50%: 3.31 to 3.53 ( $n = 2393$ ); and upper 25%: 3.53 to 4.54 ( $n = 1197$ ) log mRNA abundance.

All divergence estimates were  $\log_{10}$  transformed to facilitate linear comparisons, which are presented (values of  $dN = 0$  were converted to  $dN = 0.0001$  prior to log transformation); the results are robust and also obtained with nonparametric comparisons (TC & LLK unpub.). Partial correlation analysis was used to compare recombination rate with  $dN$  (the rate of nonsynonymous substitutions); the same results were obtained for the comparison between recombination rate and  $dN/dS$ . The partial  $r$  statistic reported here reflects the association between recombination rate and protein divergence after associations between gene expression, gene length, and SBG were removed. These factors are known to influence patterns of protein evolution (Pal et al. 2006; TC unpublished data), are all correlated with one another (i.e., recombination is positively correlated with expression and gene density, but negatively correlated with length), and can therefore give rise to spurious correlations between the variables of interest. All statistical analyses were carried out with JMP (SAS Institute). Statistical comparisons between  $r$  for different gene categories were carried out with software available online (<http://department.obg.cuhk.edu.hk/researchsupport/Correlation.asp>). Bonferroni

corrections for multiple comparisons ( $\alpha/5$ ; because of the 5 categories explored in Fig. 8.4) were used to adjust  $P$  values of statistical significance.

## RESULTS AND DISCUSSION

### Inferring the fitness effects of deleterious mutations

Highly expressed genes appear to evolve under stronger purifying selection than low-expressed genes (although the mechanistic basis of this pattern is still debated; Drummond et al. 2005; Pal et al. 2006; McInerney 2006). Consequently, gene expression level is a good predictor of the average fitness effect of deleterious mutations. Experimental gene knockouts have also identified suites of genes that are essential for survival, while many others are nonessential. To the extent that whole gene knockout phenotypes reflect the fitness effects of individual mutations, mutations in essential genes are predicted to have larger fitness effects than mutations in nonessential genes (Pal et al. 2006; Hirsch et al. 2001; Wall et al. 2005; Zhang & He 2005). Lastly, proteins have variable numbers of interaction partners (protein-protein interactions per gene – PPI – range up to nearly 300 PPI in yeast; TC unpub.), which indicate the level of constraint due to pleiotropy (Pal et al. 2006). Because individual mutations are likely to disrupt more cellular processes in genes with many PPI compared to those with few PPI, purifying selection is expected to be stronger in genes with many PPI (He & Zhang 2006).

Previous inferences of the strength of purifying selection acting on different gene categories are based on an observed elevated rate of nonsynonymous substitution in low-expressed nonessential genes with few PPI (Pal et al. 2006), which assumes that elevated rates of substitution are caused by genetic drift. An alternative possibility is that rapidly evolving genes undergo frequent bouts of positive selection. To test this assumption, I analyzed available polymorphism data from *S. cerevisiae* genes (see Methods; Appendix 7.1). Under a neutral/nearly-neutral model, patterns of within species polymorphism are expected to mirror patterns of interspecific substitution (i.e., genes with high substitutions rates also exhibit high levels of polymorphism; Kimura 1983). Positive selection decouples patterns of polymorphism and divergence and is expected to increase the number of substitutions relative to polymorphisms (McDonald & Kreitman 1991).

Low-expressed genes harbor more nonsynonymous polymorphisms (represented by the  $P_n/P_s$  ratio;  $P_n$  and  $P_s$  refer to nonsynonymous to synonymous polymorphisms, respectively) and substitutions ( $D_n/D_s$ ), than highly expressed genes, consistent with the neutral/nearly-neutral model (Fig. 7.2). Low-expressed genes also harbor higher levels of moderate- to high-frequency polymorphism (i.e., “non-singleton” polymorphism). High frequency polymorphisms are not expected to be strongly deleterious (see Bierne & Eyre-Walker 2004), but rather, will consist of neutral and slightly deleterious mutations – mutations that can potentially become fixed via genetic drift.  $D_n/D_s$  ratios are significantly lower than  $P_n/P_s$  ratios for all gene expression categories (G test; singletons included:  $P < 0.0001$ ; singletons excluded:  $P < 0.005$ , except for upper 25 % expressed genes:  $P > 0.1$ ), indicating that the predominant pattern of selection on yeast genes is purifying. Similar results are reached by partitioning the data into gene essentiality categories (Fig. 7.3; a meaningful statistical analysis based on PPI was not possible due to small sample size). The results strongly support previous inferences of selection intensity based on divergence data, and suggest that nonessential genes with low expression are relatively likely to evolve under a nearly-neutral process.

### **Recombination and protein divergence**

Analysis of 4786 genes in the yeast species *Saccharomyces cerevisiae* and *S. paradoxus* shows that nonsynonymous divergence is weakly negatively correlated with recombination rate (partial  $r = -0.052$ ,  $P < 0.01$  after Bonferroni correction; as previously reported by Pal et al. 2001). This negative relationship is markedly stronger for low-expressed genes, particularly nonessential genes with few PPI (Fig. 7.4). In contrast, the divergence of highly expressed genes tends to correlate positively with recombination rate, though all such associations are weak. For both classes of nonessential genes as well as for the entire dataset, the relationship between recombination and divergence in low-expressed genes is more negative than that of highly expressed genes (Upper vs. lower 25% expression quartiles: all genes,  $n_{low} = 1197$ ,  $n_{high} = 1197$ ,  $P = 0.041$ ; nonessential PPI = 1,  $n_{low} = 246$ ,  $n_{high} = 218$ ,  $P < 0.0001$ ; nonessential PPI > 1,  $n_{low} = 222$ ,  $n_{high} = 142$ ,  $P = 0.041$ . Upper vs. lower 50% expression quartiles: all genes,  $n_{low} = 2393$ ,  $n_{high} = 2393$ ,  $P =$

0.148; nonessential PPI = 1,  $n_{low} = 424$ ,  $n_{high} = 411$ ,  $P = 0.046$ ; nonessential PPI > 1,  $n_{low} = 435$ ,  $n_{high} = 327$ ,  $P = 0.005$ ).

These results clearly show that recombination can influence the rate of protein evolution at a genome wide scale and that the impact of recombination rate variation is strongest for low-expressed, nonessential genes with few PPI. Associations between recombination and divergence rate cannot be explained by covariation between recombination rate and several variables that independently affect protein evolution (the effects were controlled for; see Methods). Estimates of the relative recombination rate between genes are coarse and limited by the quality of the *S. cerevisiae* recombination map, and there are potential evolutionary changes in recombination between *S. cerevisiae* and *S. paradoxus*. However, both of these factors will decrease the strength of associations between divergence and recombination, and will cause our test to be conservative.

Mutation bias is also unlikely to account for the effect of recombination on protein evolution. I present associations between recombination and divergence at nonsynonymous sites ( $dN$ ) rather than between recombination and  $dN/dS$  ratios because synonymous sites are under selection in yeast (e.g., Kliman et al. 2003). Indeed, codon usage bias ( $F_{op}$ ) is positively correlated with recombination (as previously reported by Kliman et al. 2003), most strongly for highly expressed genes, which presumably have stronger selection for optimal codons (highest 50% expression partial  $r_{rec-Fop} = 0.214$  vs. lowest 50%  $r_{rec-Fop} = 0.100$ ; highest 25% expression partial  $r_{rec-Fop} = 0.249$  vs. lowest 25%  $r_{rec-Fop} = 0.089$ ;  $P < 0.0001$  for both comparisons). As a consequence,  $dS$  is negatively correlated with recombination (highest 50% expression partial  $r_{rec-Fop} = -0.131$  vs. lowest 50%  $r_{rec-Fop} = -0.077$ ,  $P = 0.031$ ; highest 25% expression partial  $r_{rec-Fop} = -0.122$  vs. lowest 25%  $r_{rec-Fop} = -0.075$ ,  $P = 0.123$ ). Furthermore, direct estimates indicate higher mutation rates in regions of high recombination (Datta & Jinksrobertson 1995; Holbeck & Strathern 1997), which should make our tests conservative. Despite these caveats with respect to using  $dS$  to estimate underlying mutational dynamics across the genome,  $dN/dS$  produces nearly-identical patterns of covariation with recombination (unpublished results).

The results are consistent with evolutionary theory suggesting that recombination enhances the efficacy of selection (e.g., Rice 2002; Barton 1995). Mutations with weak fitness effects respond to selection when the effective population size ( $N_e$ ) is large, but evolve via genetic drift when  $N_e$  is small. By increasing  $N_e$ , recombination enhances the power of selection and minimizes genetic drift. Furthermore, the adaptive consequences of recombination may be extreme in yeast since most genes in the yeast genome can be defined by weak purifying selection (~75% of genes are nonessential; ~50% have one PPI; TC unpublished). Such genes also tend to reside in genomic regions with relatively low recombination frequencies (Fig. 7.5). The correlations revealed by the data are particularly striking when one considers the method by which the genes are partitioned. Functional genomic data permits classification of genes according to their relative opportunities for slightly deleterious evolution. However, multiple types of substitutions (i.e., slightly deleterious, neutral, and beneficial) are likely to contribute to each gene's total genetic divergence between species. This plurality should dampen patterns predicted by any single processes, and will cause the conclusions presented here to be conservative.

Why does highly-expressed gene divergence show no correlation with recombination rate? There are two major possibilities. If genes under stronger selection tend to experience fewer beneficial mutations, their overall divergence rate might be relatively unaffected by local recombination. This might occur because strongly-selected genes are closer to perfection than weakly-selected genes and therefore have less opportunity for improvement, or because tradeoffs via pleiotropy (as indicated by high PPI; He & Zhang 2006) limit the opportunity for beneficial mutations (Fisher 1930; Orr 1998; Otto 2004). Secondly, selection for beneficial mutations might be very strong. The adaptive impact of varying recombination rate (and thus  $N_e$ ) is expected to decrease with the strength of selection (i.e.,  $s$ ; Kimura 1962; Hedrick 1983). If beneficial mutations tend to be strongly advantageous, they will tend to become fixed in low or in high recombinational environments.

## CONCLUSION

This study shows that recombination reduces evolutionary divergence in genes under relatively weak purifying selection (e.g., low-expressed, nonessential, few PPI), and at

best, marginally increases divergence in genes under strong purifying selection (e.g., highly expressed, essential, many PPI). This pattern suggests that enhanced purifying selection is a primary long-term benefit of recombination in nature. The efficient removal of deleterious mutations might increase the competitive ability of sexual species and contribute to the observed ubiquity of sexual reproduction in eukaryotes (Rice 2002; Bell 1982).

While this interpretation is appealing and supported by both theory and data from other taxa (e.g., Barton 1995; Presgraves 2005), it should be noted that inferences about the processes driving nucleotide divergence between species are tentative and reflect a major limitation of molecular divergence data. Future studies using entire-genome polymorphism and divergence data can add resolution by estimating the proportion of adaptive substitutions per gene (see Eyre-Walker 2006). Such estimates, combined with inferences about the slightly deleterious substitution rate (derived from expression, essentiality and PPI data), will permit a much improved analysis of the benefits of recombination.

Table 7.1. Different processes of molecular evolution produce different correlations between recombination rate and nucleotide divergence between species<sup>1</sup>.

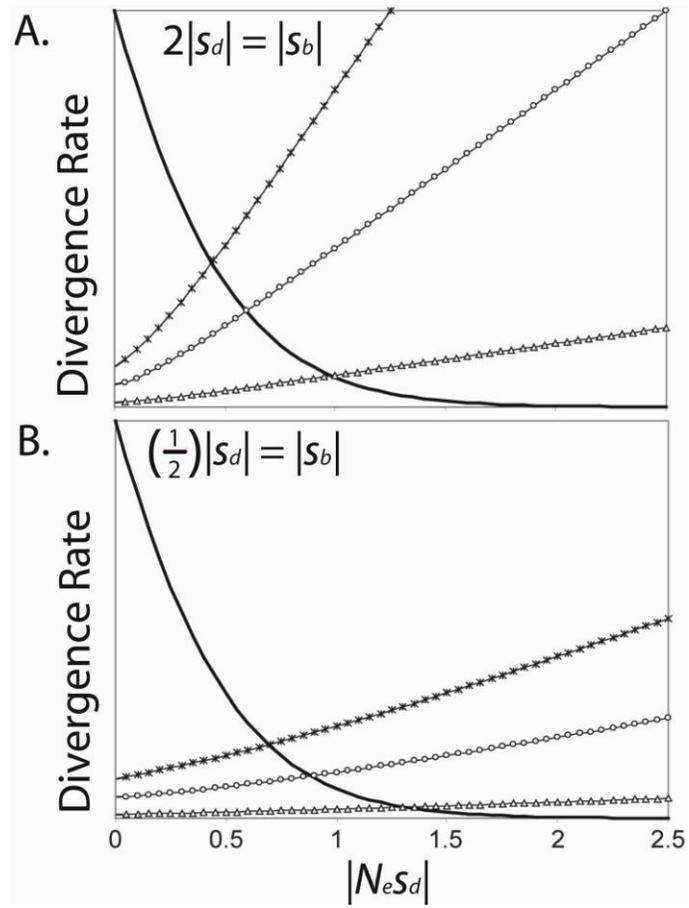
Model of evolution	Fitness effect of substitutions <sup>2</sup>	Correlation between recombination rate and divergence
Neutral	$s = 0$	None
Slightly deleterious	$-1/(2N_e) < s < 0$	Negative
Adaptive	$s > 0$	Positive

<sup>1</sup> See text for details

<sup>2</sup> The selection coefficient,  $s$  ( $-1 \leq s < \infty$ ), reflects the strength and direction of selection acting upon new mutations during the process of fixation

Figure 7.1. Substitution rate as a function of the effective strength of selection. The bold solid curve represents the slightly deleterious divergence rate between species.

Remaining curves represent adaptive divergence under three scenarios: open triangles when the beneficial mutation rate ( $u_b$ ) is 1% of the deleterious mutation rate ( $u_d$ ); open circles when  $u_b$  is 5% of  $u_d$ ; 'X-marks' when  $u_b$  is 10% of  $u_d$ . The strength of selection against deleterious mutations ( $/N_e s_d$ ) is shown on the x-axis, where  $s_d$  is the average strength of purifying selection. A. The average strength of positive selection ( $s_b$ ) is equal to the average strength of purifying selection ( $s_d$ ); B.  $s_b$  is twice as strong as  $s_d$ ; C.  $s_d$  is twice as strong as  $s_b$ . When selection is weak, mildly-deleterious substitutions outnumber adaptive substitutions; divergence in weakly-selected genes is therefore predicted to be negatively correlated with local recombination rate. As the strength of selection increases, adaptive substitutions predominate; divergence in strongly-selected genes is therefore predicted to be positively correlated with recombination rate. The adaptive divergence rate is  $u_b(1 - e^{-4Nsp})/(1 - e^{-4Ns})$ , where  $s$  is the average benefit conferred by each mutation, and  $p$  is the initial frequency of each mutation (results for  $p = 0.0001$  are shown). The slightly deleterious rate is  $u_d(1 - e^{-4Nsp})/(1 - e^{-4Ns})$ , where  $u_d$  is the deleterious mutation rate, and  $s$  is the average cost of each mutation (equations modified from Kimura 1962;  $N_e = N$ ). The assumption that the beneficial mutation rate is much smaller than the deleterious mutation rate is supported by theory and mutation accumulation experiments (Fisher 1930; Wloch et al. 2001; Keightley & Lynch 2003; Joseph & Hall 2004; but see Shaw et al. 2002, 2003).



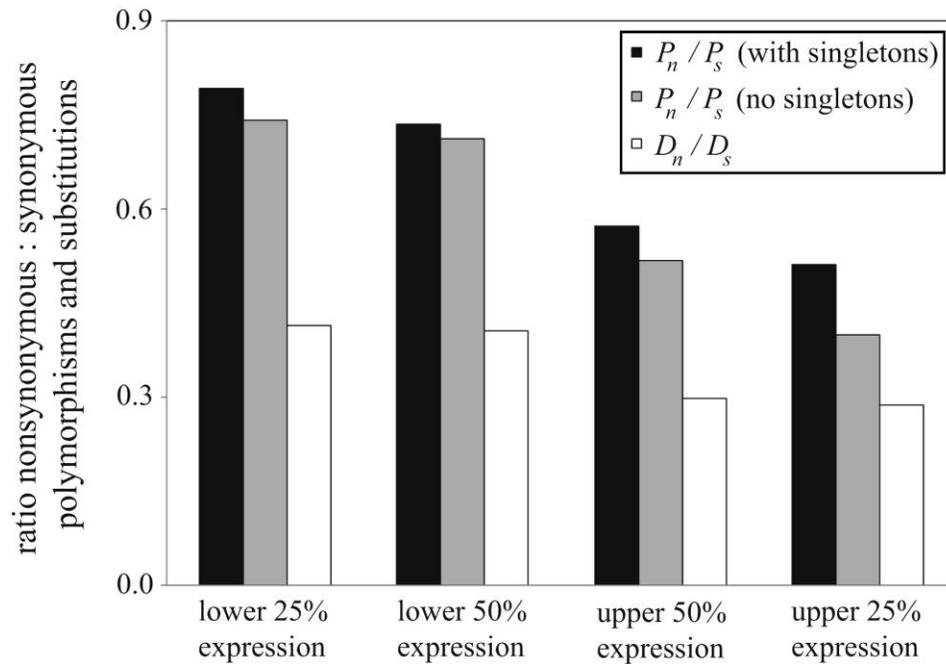


Figure 7.2. Ratios of replacement to silent polymorphism ( $P_n/P_s$ ) in *S. cerevisiae*, and substitutions ( $D_n/D_s$ ) between *S. cerevisiae* and *S. paradoxus*. Results were obtained by pooling polymorphism and divergence data for multiple genes within each expression category (see Fig. S1 for similar results using a different approach).  $P_n/P_s$  ratios are lower in highly expressed genes: upper vs. lower 50% with singletons,  $P = 0.172$ ; upper vs. lower 25% with singletons,  $P = 0.026$ ; upper vs. lower 50% without singletons,  $P = 0.203$ ; upper vs. lower 25% without singletons,  $P = 0.023$ .  $D_n/D_s$  ratios are lower for highly expressed genes for all comparisons ( $P < 0.0001$ ).

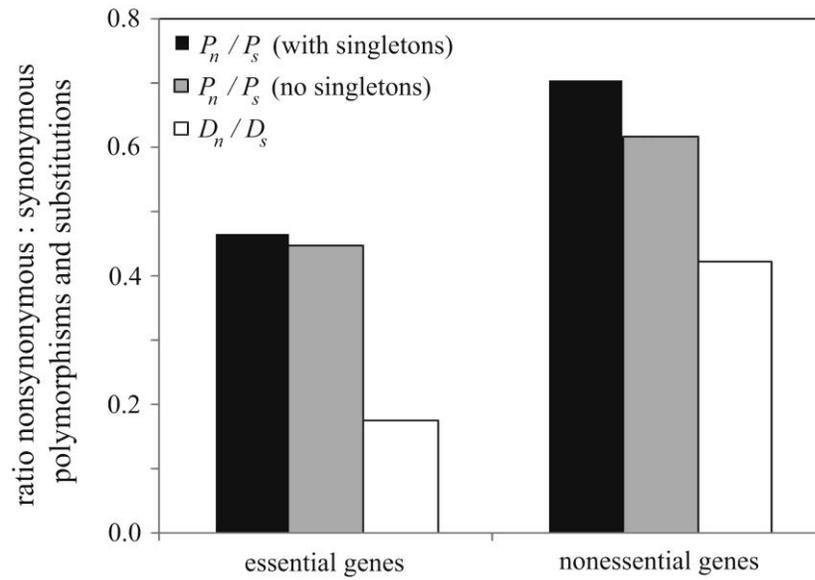


Fig. 7.3. Ratios of replacement to silent polymorphism ( $P_n/P_s$ ) in *S. cerevisiae*, and substitutions ( $D_n/D_s$ ) between *S. cerevisiae* and *S. paradoxus*; genes are partitioned according to knockout viability phenotype. Results were obtained by pooling polymorphism and divergence data for multiple genes within each of the two categories.  $P_n/P_s$  ratios are lower in essential genes: with singletons,  $P = 0.067$ ; without singletons,  $P = 0.806$ .  $D_n/D_s$  ratios are lower for essential genes ( $P < 0.00001$ ).

Figure 3

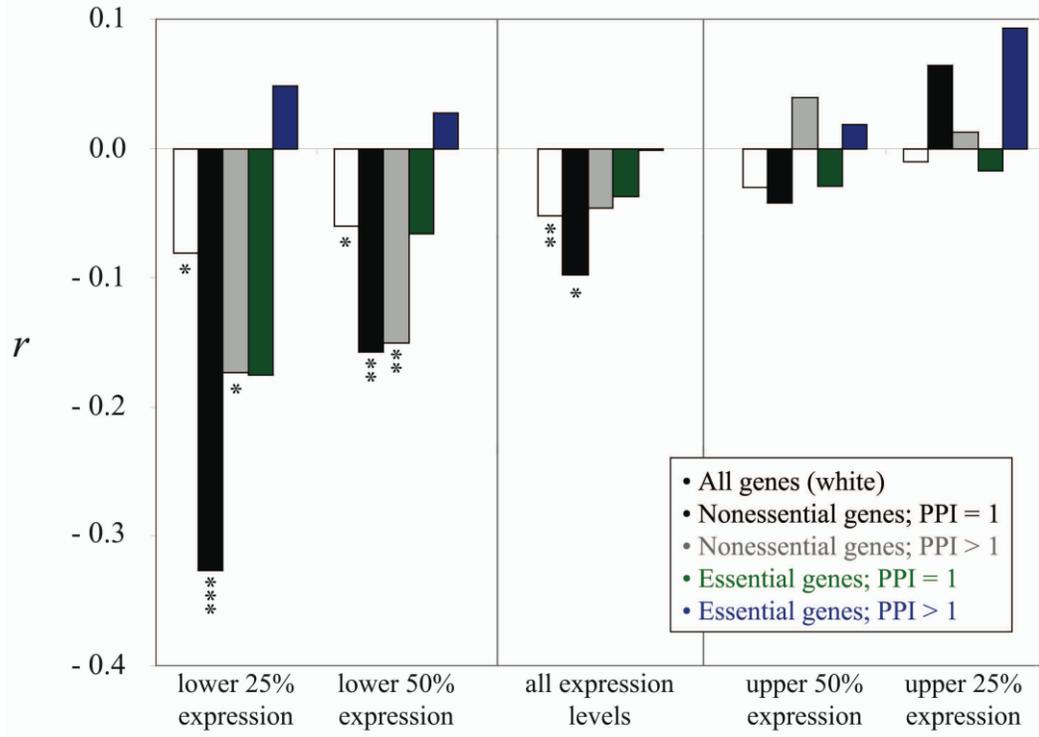


Figure 7.4. Recombination and protein evolutionary rate. The relationship ( $r$  = partial correlation coefficient; see materials and methods) between recombination rate and  $dN$  for five gene expression intervals: the lower gene expression quartile (2.25 to 3.15 log mRNA abundance), the lower 50% expressed genes (2.25 to 3.31 log mRNA abundance), the entire range of expression, the upper 50% expression (3.31 to 4.54 log mRNA abundance), and the upper gene expression quartile (3.53 to 4.54 log mRNA abundance). \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$  (after Bonferroni correction for five comparisons).

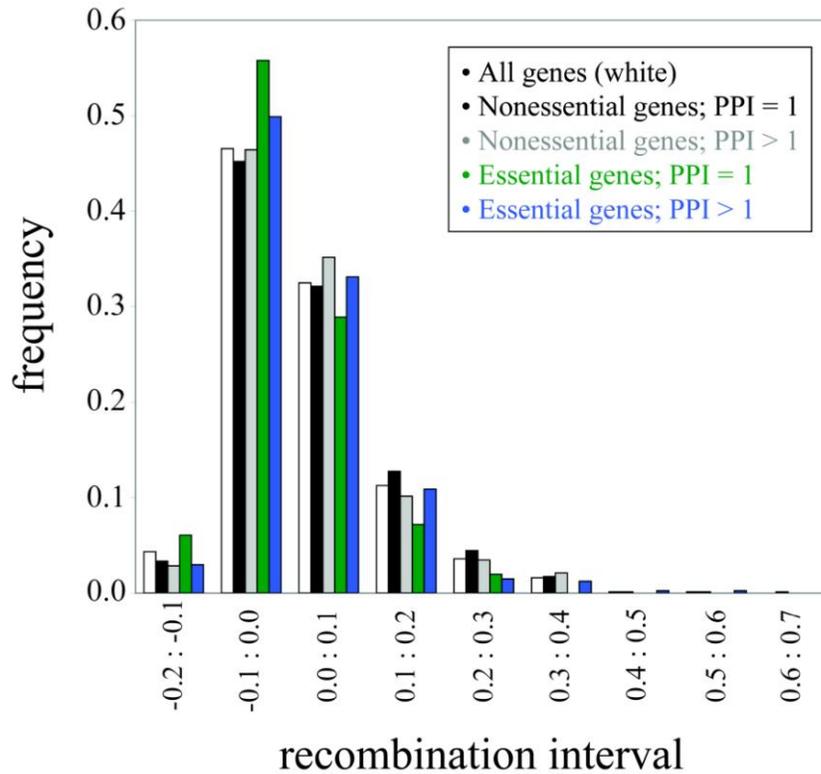


Fig. 7.5. The distribution of recombination rates for *S. cerevisiae* genes. Recombination rate per gene was estimated with a microarray experiment as described in Gerton et al. [8]. Recombination rates are plotted as the logarithm of the mean hybridization ratio of recombinant probes ( $P2$ ) relative to a total genomic probe ( $P1$ ); higher values of  $\log_{10}(P2/P1)$  reflect higher frequencies of recombination.

## APPENDIX 7.1. Yeast polymorphism and Divergence Data

Gene	$N$	With singletons		No singletons		$D_s$	$D_n$	Expression Quantile <sup>(1)</sup>	Essential?
		$P_s$	$P_n$	$P_s$	$P_n$				
ACT1	10	1	0	1	0	14	0	4	yes
CCA1	73	11	1	4	1	53	10	2	yes
CDC19	29	3	0	3	0	22	6	4	yes
CWP1	15	7	5	6	1	57	29	4	no
FIG1	20	8	8	6	2	86	11	1	no
FZF1	29	6	4	6	4	87	85	1	no
GCN4	19	5	6	2	1	55	17	4	no
HHT2	9	4	0	4	0	15	0	4	no
HIS3	15	7	4	2	2	no BLAST hit		3	no
MBP1	9	7	4	1	2	69	12	4	yes
MKT1	4	1	2	1	2	185	52	4	no

MLH1	14	6	14	5	8	192	74	1	no
MLS1	77	12	4	6	2	65	7	4	no
NEW1	6	3	1	3	0	23	8	4	no
PDC1	6	0	4	0	1	34	11	4	no
PDR10	75	6	7	4	5	66	28	4	no
PEA2	17	18	14	7	5	111	35	1	no
PHD1	29	6	7	5	6	71	35	3	no
PMS1	5	3	1	0	0	214	117	1	no
RME1	13	3	4	0	2	65	68	4	no
RNQ1	12	2	1	1	1	40	24	3	no
SEC53	16	2	2	2	0	52	3	4	yes
SPT3	17	13	5	5	3	80	3	1	no
SSU1	29	5	9	5	7	96	25	3	no
STE2	19	18	7	8	5	100	18	4	no
SUP35	9	7	11	2	4	71	23	4	yes

TAO3	13	46	18	16	6	499	83	1	yes
TRP1	19	3	3	2	2	53	19	3	no
URA1	15	15	1	8	1	57	7	4	no
URE2	26	7	2	2	2	62	4	3	no
VMA1	9	47	14	44	12	65	3	4	no
YAR023C	17	5	11	4	8	25	36	1	no
YCR007C	16	5	3	2	1	49	48	1	no
YHL044W	19	7	16	4	5	52	75	1	no
ZDS2	59	8	5	7	4	44	28	1	no

---

1. Quantile 1: lowest 25% expression; Quantile 2: 25-50%; Quantile 3: 50-75%; Quantile 4: highest 25% expression.

## REFERENCES

- Bachtrog D. 2004. Protein evolution and codon usage bias on the neo-sex chromosomes of *Drosophila miranda*. *Genetics* 165:1221-1232.
- Barton NH 1995. Linkage and the limits to natural selection. *Genetics* 140:821-841.
- Bell G. 1982. The masterpiece of nature. University of California Press, Berkeley.
- Betancourt AJ, Presgraves DC. 2002. Linkage limits the power of natural selection in *Drosophila*. *Proceedings of the National Academy of Science USA* 99:13616-13620.
- Bierne N, Eyre-Walker A. 2004. The genomic rate of adaptive amino acid substitution in *Drosophila*. *Molecular Biology and Evolution* 21:1350-1360.
- Birky CW, Walsh JB. 1988. Effects of linkage on rates of molecular evolution. *Proceedings of the National Academy of Science USA* 85:6414–6418.
- Castillo-Davis CI, Hartl DL. 2003. GeneMerge – post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics*. 19:891-892.
- Charlesworth B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetical Research* 63:213-227.
- Datta A, Jinksrobertson S. 1994. Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science* 268:1616-1619.
- DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
- Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH. 2005. Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Science USA* 102:14338–14343.
- Eyre-Walker A. 2006. The genomic rate of adaptive evolution. *Trends in Ecology and Evolution* 21:569-575.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.
- Fisher RA. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.
- Fisher RA. 1930. The distribution of gene ratios for rare mutations. *Proceedings of the Royal Society of Edinburgh* 50:204-219.
- Gerton JL, DeRisi J, Schroff R, Lichten M, Brown PO, Petes TD. 2000. Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Science USA* 97:11383-11390.

- He X, Zhang J. 2006. Toward a molecular understanding of pleiotropy. *Genetics* 173:1885-1891.
- Hedrick PW. 1983. *Genetics of populations*. Science Books International, Boston.
- Hein J, Schierup MH, Wiuf C. 2005. *Gene Genealogies, Variation and Evolution: A Primer in Coalescent Theory*. Oxford University Press, New York.
- Hey J, Kliman RM. 2002. Interactions between natural selection, recombination and gene density in the genes of *Drosophila*. *Genetics* 160:595-608.
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genetical Research* 8:269-294.
- Hirsh AE, Fraser HB. 2001. Protein dispensability and rate of evolution. *Nature* 411:1046-1049.
- Holbeck SL, Strathern JN. 1997. A role for REV3 in mutagenesis during double stranded break repair in *Saccharomyces cerevisiae*. *Genetics* 147:1017-1024.
- Jensen-Seaman MI, Furey TS, Payseur BA, Lu YT, Roskin KM, Chen CF, Thomas MA, Haussler D, Jacob HJ. 2004. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research* 14:528-538.
- Joseph SB, Hall DW. 2004. Spontaneous mutations in diploid *Saccharomyces cerevisiae*: More beneficial than expected. *Genetics* 168:1817-1825.
- Keightley PD, Lynch M. 2003. Toward a realistic model of mutations affecting fitness. *Evolution* 57:683-685.
- Kimura M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47:713-719.
- Kimura M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- Kliman RM, Irving N, Santiago M. 2003. Selection conflicts, gene expression, and codon usage trends in yeast. *Journal of Molecular Evolution* 57:98-109.
- Marais G, Domazet-Lošo T, Tautz D, Charlesworth B. 2004. Correlated evolution of synonymous and nonsynonymous sites in *Drosophila*. *Journal of Molecular Evolution* 59:771-779.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351:652-654.

- McInerney JO. 2006. The causes of protein evolutionary rate variation. *Trends in Ecology and Evolution* 21:230-232.
- McVean GAT, Charlesworth B. 2000. The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* 155:929-944.
- Mezard C. 2006. Meiotic recombination hotspots in plants. *Biochem Soc Trans* 34:531-534.
- Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52:935-949.
- Otto SP. 2004. Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proceedings of the Royal Society of London B* 271:705-714.
- Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination. *Nature Reviews Genetics* 3:252-261.
- Pal C, Papp B, Hurst LD. 2001. Does the recombination rate affect the efficiency of purifying selection? The yeast genome provides a partial answer. *Molecular Biology and Evolution* 18:2323-2326.
- Pal C, Papp B, Lercher MJ. 2006. An integrated view of protein evolution. *Nature Reviews Genetics* 7:337-348.
- Paland S, Lynch M. 2006. Transitions to asexuality result in excess amino acid substitutions. *Science* 311:990-992.
- Peck JR. 1994. A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. *Genetics* 137:597-606.
- Presgraves DC. 2005. Recombination enhances protein adaptation in *Drosophila melanogaster*. *Current Biology* 15:1651-1656.
- Rice WR. 2002. Experimental tests of the adaptive significance of sexual recombination. *Nature Reviews Genetics* 3:241-251.
- Rocha EPC. 2006. The quest for the universals of protein evolution. *Trends in Genetics* 22:412-416.
- Rozas J, Sánchez-DelBarrio JC, Messegyer X, Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496-2497.

- Ruderfer DM, Pratt SC, Seidel HS, Kruglyak L. 2006. Population genomic analysis of outcrossing and recombination in yeast. *Nature Genetics* 38:1077-1081.
- Shaw FH, Geyer CJ, Shaw RG. 2002. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* 56:453-463.
- Shaw RG, Shaw FH, Geyer, C. 2003. What fraction of mutations reduces fitness? A reply to Keightley and Lynch. *Evolution* 57:686-689.
- Wall DP, Hirsh AE, Fraser HB, Kumm J, Giaever G, Eisen MB, Feldman MW. 2005. Functional genomic analysis of the rates of protein evolution. *Proceedings of the National Academy of Science USA* 102:5483-5488.
- Wloch DM, Szafranec K, Borts RH, Korona R. 2001. Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics* 159:441-452.
- Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97-159.
- Zhang J, He X. 2005. Significant impact of protein dispensability on the instantaneous rate of protein evolution. *Molecular Biology and Evolution* 22:1147-1155.
- Zhang Z, Parsch J. 2005. Positive correlation between evolutionary rate and recombination rate in *Drosophila* genes with male-biased expression. *Molecular Biology and Evolution*. 22:1945-1947.