Adaptive Protein Evolution of X-linked and Autosomal Genes in *Drosophila*: Implications for Faster-X Hypotheses

Tim Connallon
Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor

Patterns of sex chromosome and autosomal 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*, *Drosophila simulans*, and *Drosophila 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 and 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 and 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 and 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 Ns, where Ns is the effective population size and s is the intensity of selection; Fisher 1930; Wright 1931; Kimura 1962). Low Ns limits the rate of adaptive substitution (Betancourt and 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 and 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 autosomal 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 Ns, I analyzed coding sequence polymorphism and divergence in the *Drosophila melanogaster* lineage. Silent and amino acid replacement polymorphisms (Pn and Pn, respectively) and substitutions (Dn and Dn) were calculated for 337 protein-coding genes (80 X linked) and were used to estimate alpha, the proportion of Dn that is caused by positive selection (McDonald and Kreitman 1991; Smith and Eyre-Walker 2002). Recombination rate and gene expression data were used to mitigate biases caused by Ns reductions via low recombination (i.e., enhanced Hill–Robertson interference) and gene composition differences between chromosomes.

**Materials and Methods**

**Gene Samples**

www-biology.ucsd.edu/labs/andolfatto/), and Jody Hey’s Web page (http://lifesci.rutgers.edu/~heylab/). Orthologous sequences from *Drosophila simulans* and *Drosophila 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 5 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 genespecific 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 (True et al. 1996; Ranz et al. 2003);
- A large number of gene samples are from East African populations, which are probably closer to mutation–selection–drift equilibrium than cosmopolitan populations (David and Capy 1988; Begun and Aquadro 1993);
- Mutation rates are not sex biased in this species (Bauer and Aquadro 1997), making different X and autosome mutation rates unlikely;
- Alpha (the proportion of *D*. caused by positive selection) relies upon an assumption of neutrality for polymorphisms and silent substitutions (*P* , *P* , and *D* ; Smith and Eyre-Walker 2002). Although exclusion of low-frequency polymorphisms is usually considered suitable justification for the neutral polymorphism assumption, it has been suggested that *D*. 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 *D*. *D* ). Synonymous site selection appears to be absent in the *D. melanogaster* lineage (in contrast to *D. simulans*; e.g., McVean and Vieira 2001; Akashi et al. 2006), which suggests that the neutrality assumption is most reasonable here (this assumption is addressed in full, below).

Expression and Recombination Data

Expression data from Stolc et al. (2004) were obtained from Kevin White’s Web page (http://genome.med.yale.edu/lab/discuss.htm). For genes with data from multiple distinct probes and multiple replicates of the same probes, 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 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 gene content are reported; additional results are included in Supplementary Material online.

Five different recombination estimates per gene were obtained from Jody Hey’s Web page (http://lifesci.rutgers.edu/~heylab/; described in Hey and Kliman 2002). Most of the 337 genes analyzed here were included in the data set. When unavailable, data from the nearest gene was used (as defined by www.flybase.org). This procedure is consistent with the methodologies used to generate the recombination estimates of genes included in Hey and 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 toward 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 × 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 appear (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* . Two genes (*dpp* and *CG6495*) were polymorphic for inversions; standard alignments (i.e., noninversions) were used for all analyses. Two genes (*SR-CIII* and *SR-CIV*) included nonfunctional gene copies, which were removed from all analyses. Both exclusions do not affect the results.

*D*. and *D*. 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*. *D*. or *P*. *P*. 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 and 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 \( D_u, D_s, P_u, \) and \( P_s \) values across genes and therefore avoids a bias described in Shapiro et al. (2007).

*Drosophila melanogaster* and *D. simulans* are separated by 2 X-linked and 3 autosomal inversion differences (1 on chromosome 2R; 2 on chromosome 3R; Lemeunier and Ashburner 1976)—in which 2 X-linked and 29 autosomal genes from this data set 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. Although this X-autosome difference was not explicitly controlled, it should (if anything) bias the data set toward detecting a pattern of faster-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.

The entire list of genes with statistics is available in the Supplementary Material online.

### Results and Discussion

#### Divergence and Polymorphism

For total divergence between *D. melanogaster* and *D. simulans*, the \( D_u/D_s \) ratio is the same for the X and autosomes (table 2), consistent with previous studies (Betancourt et al. 2002; Thornton et al. 2006). However, \( D_s \) is strongly reduced, whereas \( D_u \) is the same, in the *D. simulans* relative to the *D. melanogaster* lineage (\( D_s: \chi^2 = 96.93, P < 0.00001; D_u: \chi^2 = 0.554, P = 0.457 \)), indicating relaxed selection on synonymous sites in the *D. melanogaster* lineage (see McVean and Viera (2001); Akashi et al. 2006). This pattern is stronger for X-linked compared with autosomal genes (\( \chi^2 = 14.397, P = 0.00015 \)), which consequently reduces the \( D_u/D_s \) ratio on the X relative to the autosomes in the *D. melanogaster* lineage (table 2).

X-linked genes exhibit lower \( P_u/P_s \) ratios than autosomal genes (table 2; supplementary fig. S1, Supplementary Material online; see Begun 1996; Andolfatto 2001), and this difference remains after male-biased genes, which are substantially more abundant on autosomes (table 1; Parisi et al. 2003), are removed. This pattern is consistent with 2 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 with those at high frequencies. X-autosome differences in \( P_u/P_s \) remain after singleton polymorphisms are removed (table 2), suggesting an 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 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 and 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 = 4\( N_e \mu \); where \( \mu \) is the mutation rate per nucleotide) can be

### Table 1

<table>
<thead>
<tr>
<th>X-chromosome and Autosome Gene Composition</th>
<th>Gene Expression Category</th>
<th>Total</th>
<th>Juvenile Biased</th>
<th>Adult Biased</th>
<th>Female Biased</th>
<th>Male Biased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutoff(^a)</td>
<td></td>
<td>1.5-fold</td>
<td>2-fold</td>
<td>1.5-fold</td>
<td>2-fold</td>
<td>1.5-fold</td>
</tr>
<tr>
<td>Number of genes</td>
<td>80</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Relative frequency(^b)</td>
<td>1</td>
<td>0.063</td>
<td>0.000</td>
<td>0.050</td>
<td>0.025</td>
<td>0.288</td>
</tr>
<tr>
<td>Autosomes</td>
<td>257</td>
<td>21</td>
<td>16</td>
<td>61</td>
<td>34</td>
<td>74</td>
</tr>
<tr>
<td>Number of genes</td>
<td>1</td>
<td>0.082</td>
<td>0.062</td>
<td>0.237</td>
<td>0.132</td>
<td>0.288</td>
</tr>
<tr>
<td>Relative frequency(^b)</td>
<td>1</td>
<td>0.601</td>
<td>0.027</td>
<td>0.001</td>
<td>0.0012</td>
<td>0.996</td>
</tr>
</tbody>
</table>

\(^a\) Cutoff refers to a minimum expression ratio for each category; for example, female-biased genes with a 2-fold cutoff have at least twice the expression (in terms of mRNA abundance) in females as in males.

\(^b\) Frequency of total genes for each chromosome type, that is, total X-linked or total autosomal genes.

\(^c\) Based on \( \chi^2 \) values for \( 2 \times 2 \) contingency tables.
used to study genic effective population size \( (N_e) \) as a function of recombination rate variation (e.g., Begun and Aquadro 1992; Presgraves 2005). Silent variation is positively correlated with recombination for the X and autosomes \( (r_X = 0.481; r_{aut} = 0.440; P < 0.0001 \text{ for both correlations}) \) and is particularly reduced in regions of low recombination (fig. 1; supplementary fig. S1, Supplementary Material online). Silent variability is significantly higher for X-linked genes \( (P, 0.05; 2\text{-tailed } t\text{-test}) \), but this discrepancy disappears when low-recombining genes are removed (e.g., less than \( 2/10^5 \)bp/generation; fig. 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). 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, 3 systematic biases creep into contrasts between the X and autosomes (as discussed above). Autosomes are enriched in genes with male-biased expression (table 1), which will upwardly bias autosomal estimates of alpha. Autosomes also carry a higher proportion of genes evolving in low-recombinational environments (table 3; fig. 1) and are predicted to harbor relatively higher numbers of deleterious mutations than the X chromosome (Haldane 1935); these 2 factors will tend to downwardly bias estimates of alpha (Bierne and Eyre-Walker 2004; Presgraves 2005).

To eliminate these potential sources of bias, genes inhabiting low-recombination genomic regions (i.e., recombination rate \( < 0.002 \times 10^{-5} \)bp/gen; fig. 1; supplementary fig. S1, Supplementary Material online), and genes with male-biased expression (see also supplementary fig. S2, Supplementary Material online), were removed from the analysis. Singleton polymorphisms, which are likely to include deleterious alleles (Bierne and Eyre-Walker

Table 2

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Location</th>
<th>( D_n )</th>
<th>( D_s )</th>
<th>( D_n/D_s )</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. melanogaster lineage</td>
<td>X</td>
<td>475</td>
<td>1119</td>
<td>0.424</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>Autosomes</td>
<td>1412</td>
<td>3033</td>
<td>0.466</td>
<td>0.774</td>
</tr>
<tr>
<td>D. simulans lineages</td>
<td>X</td>
<td>955</td>
<td>1882</td>
<td>0.507</td>
<td>0.774</td>
</tr>
<tr>
<td></td>
<td>Autosomes</td>
<td>2865</td>
<td>5572</td>
<td>0.514</td>
<td>0.774</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Location</th>
<th>X Chromosome</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic</td>
<td>0.0040</td>
<td>0.0056</td>
<td>0.0040</td>
<td>0.0040</td>
<td>0.0041</td>
</tr>
<tr>
<td>All gene samples</td>
<td>0.0039</td>
<td>0.0057</td>
<td>0.0041</td>
<td>0.0042</td>
<td>0.0042</td>
</tr>
<tr>
<td>African samples</td>
<td>0.0043</td>
<td>0.0063</td>
<td>0.0047</td>
<td>0.0050</td>
<td>0.0049</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Location</th>
<th>Autosomal</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0023</td>
<td>0.0023</td>
<td>0.0024</td>
</tr>
<tr>
<td>All gene samples</td>
<td>0.0027</td>
<td>0.0027</td>
<td>0.0025</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>African samples</td>
<td>0.0029</td>
<td>0.0029</td>
<td>0.0027</td>
<td>0.0025</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

To eliminate these potential sources of bias, genes inhabiting low-recombination genomic regions (i.e., recombination rate \( < 0.002 \times 10^{-5} \)bp/gen; fig. 1; supplementary fig. S1, Supplementary Material online), and genes with male-biased expression (see also supplementary fig. S2, Supplementary Material online), were removed from the analysis. Singleton polymorphisms, which are likely to include deleterious alleles (Bierne and Eyre-Walker

Table 5

<table>
<thead>
<tr>
<th>Location</th>
<th>X Chromosome</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic</td>
<td>0.6505</td>
<td>0.4596</td>
<td>0.5800</td>
<td>0.5755</td>
<td>0.5768</td>
</tr>
<tr>
<td>All gene samples</td>
<td>0.6867</td>
<td>0.4742</td>
<td>0.6075</td>
<td>0.5674</td>
<td>0.5849</td>
</tr>
<tr>
<td>African samples</td>
<td>0.6617</td>
<td>0.4560</td>
<td>0.5807</td>
<td>0.5125</td>
<td>0.5338</td>
</tr>
</tbody>
</table>

\( ^a \) The five estimates of recombination are described in Hey and Kliman (2002) and were obtained from Jody Hey’s Web site (http://lifesci.rutgers.edu/%7Eheylab/).

\( ^b \) 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.
were also removed. Estimated alpha values for the trimmed data set are quite similar between the X and autosomes (ca. 30%), with heavily overlapping confidence intervals (CIs) (fig. 2; supplementary fig. S2, Supplementary Material online). 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 Ps and decrease and/or downwardly bias Ds and increase estimates of positive selection. Current selection is expected to remove synonymous polymorphism as well as skew the distribution of variation toward 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 and Capy 1988; Begun and Aquadro 1993), D is negative for silent sites but does not differ between the X and autosomes (DX = −0.236; Daut = −0.215; P >> 0.05; 2-tailed t-test), as expected if silent sites are equally neutral on the X and autosomes. Detecting long-term selection on synonymous sites is more problematic because 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 (dSX = 0.072; dSaut = 0.061; see also Begun and Whitley 2000), as would be expected if X-linked silent sites evolve under greater constraint. Even if Ds is downwardly biased on the X, the major conclusion reported here—that X linkage does not enhance adaptive divergence—will be conservative.

Conclusion
In the D. melanogaster lineage, an equivalent proportion of adaptive nonsynonymous substitutions occur for X-linked compared with 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% of the autosomal population size (i.e., $3/4N_A = N_X$), the dominance of beneficial alleles can be estimated as $h \approx \frac{R_X}{R_X + R_A}$, where $R_X$ is the ratio of the autosomal to X-linked adaptive substitution rate (Charlesworth et al. 1987; equation [2a]). The ratio of the maximum likelihood estimates of autosomal 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 $N_e$ (fig. 1; supplementary fig. S1, Supplementary Material online; 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, $h \approx \frac{3}{3+2N_e} \approx 0.598$. CIs 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 and 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.

Supplementary Material

Supplementary table of genes, polymorphism, and divergence data and figures S1 and S2 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

Thanks to Adam Eyre-Walker for providing software, to Bryan Carstens, David Houle, Alex Kondrashov, Jeff Long, Akane Uesugi, Amanda Zellmer, and especially Lacey Knowles for valuable discussions, and to John McDonald and 3 anonymous reviewers for many good comments and suggestions on the manuscript. Funding was provided by a grant from the University of Michigan Rackham Graduate School and a training grant from the National Human Genome Research Institute (T32 HG00040).

Literature Cited


Figure S1.— 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 S2.— Alpha – the proportion of adaptive, nonsynonymous substitutions – for X-linked and autosomal genes. The X and autosomes are represented by black and white bars, respectively. Substitutions included in the analysis are those that occurred within the \textit{D. melanogaster} lineage; polymorphisms are for African and non-African gene samples. Means and 95 percent confidence intervals are shown. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; † $P < 0.0001$.