*Genet. Res., Camb.* (2008), **90**, *pp.* 421–431. © 2008 Cambridge University Press
doi:10.1017/S0016672308009804 Printed in the United Kingdom


# A multispecies approach for comparing sequence evolution of X-linked and autosomal sites in *Drosophila*

BEATRIZ VICOSO*, PENELOPE R. HADDRILL AND BRIAN CHARLESWORTH
*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK*

(*Received 21 May 2008 and in revised form 2 September 2008*)

**Summary**

Population genetics models show that, under certain conditions, the X chromosome is expected to be under more efficient selection than the autosomes. This could lead to 'faster-X evolution', if a large proportion of mutations are fixed by positive selection, as suggested by recent studies in *Drosophila*. We used a multispecies approach to test this: Muller's element D, an autosomal arm, is fused to the ancestral X chromosome in *Drosophila pseudoobscura* and its sister species, *Drosophila affinis*. We tested whether the same set of genes had higher rates of non-synonymous evolution when they were X-linked (in the *D. pseudoobscura*/*D. affinis* comparison) than when they were autosomal (in *Drosophila melanogaster*/*Drosophila yakuba*). Although not significant, our results suggest this may be the case, but only for genes under particularly strong positive selection/weak purifying selection. They also suggest that genes that have become X-linked have higher levels of codon bias and slower synonymous site evolution, consistent with more effective selection on codon usage at X-linked sites.

## 1. Introduction

In species with X Y or W Z sex determination, positive selection may be more effective at fixing favourable mutations that arise on the X or Z chromosome compared with the autosomes, because rare recessive or partially recessive mutations are fully expressed in the heterogametic sex, whereas in the homogametic sex they are mostly present as heterozygotes with wild-type alleles (Haldane, 1924). This can result in a higher rate of substitution of beneficial mutations at X-linked or Z-linked loci, provided that relevant conditions on dominance coefficients, dosage compensation, sex-specific mutation rates and initial frequencies of the mutations are met (Rice, 1984; Charlesworth *et al.*, 1987; Orr & Betancourt, 2001; Kirkpatrick & Hall, 2004; for a recent review, see Vicoso & Charlesworth, 2006). In contrast, under most conditions, recessive or partially recessive slightly deleterious mutations will experience a faster rate of substitution on the autosomes than on the X or Z chromosome, as a result of less effective selection against them (Charlesworth *et al.*, 1987; McVean & Charlesworth, 1999).

A major factor in determining whether X-linked loci will evolve faster or slower than autosomal loci is the fraction of mutations that are fixed by positive selection versus genetic drift (from now on, we will simply refer to X chromosomes, since the same results apply to Z chromosomes with a switch of gender). Recent studies of *Drosophila* suggest that 25–50% of divergent non-synonymous sites among related species were fixed by positive selection (Smith & Eyre-Walker, 2002; Bierne & Eyre-Walker, 2004; Welch, 2006; Andolfatto, 2007; Begun *et al.*, 2007). Accordingly, we might expect to observe faster-X evolution, provided that new beneficial mutations are, on average, at least partially recessive. There is little relevant information on the levels of dominance of beneficial mutations, although indirect evidence from the genetics of species differences in highly selfing taxa of plants (Charlesworth, 1992), and from comparisons of rates of adaptive evolution of haploid and diploid laboratory populations of yeast (Zeyl *et al.*, 2003;

* Corresponding author: Institute of Evolutionary Biology, King's Buildings, Ashworth Laboratories, West Mains Road, Edinburgh EH9 3JT, UK. Tel: +44 131 650 5476. Fax. +44 131 650 6564. e-mail: b.vicoso@ed.ac.uk
Sequence data reported in this article have been deposited in the GenBank Data Libraries under accession numbers EU931120–EU931205.

Anderson *et al*., 2004), is consistent with a predominance of at least partial recessivity of new, selectively favourable mutations.

The current availability of large DNA sequence datasets has facilitated comparative analyses of the rates of molecular evolution of the X chromosome and the autosomes. In particular, from between-species comparisons, we can estimate $K_a$ and $K_s$, the rates of non-synonymous and synonymous divergence per nucleotide site, respectively. It is commonly assumed that $K_s$ mostly reflects nearly neutral or neutral evolution, and the ratio $K_a/K_s$ is used to estimate the overall effect of selective forces. Neutral sequences evolve with $K_a/K_s \approx 1$, and negative (purifying) selection reduces this ratio, whereas recurrent positive selection increases it (Graur & Li, 2000). If positive selection is more effective on the X chromosome, we might thus expect X-linked sites to show higher $K_a/K_s$ values than autosomal sites. This test can be combined with other evidence, such as the McDonald–Kreitman test (McDonald & Kreitman, 1991), to discriminate between alternative interpretations of differences in $K_a$ and $K_s$ between X chromosomes and autosomes, such as increased positive selection, relaxation of constraint and differences in mutation rates. Currently, the evidence for higher $K_a/K_s$ on the *Drosophila* X chromosome is conflicting, with some studies finding evidence for adaptively driven faster-X effects (Thornton & Long, 2002, 2005; Clark *et al*., 2007; Singh *et al*., 2007), and others failing to detect such effects (Betancourt *et al*., 2002; Begun *et al*., 2007; Connallon, 2007).

The drawback of this approach is that the X chromosome differs considerably in its gene content from the autosomes; for example, in *Drosophila*, male-biased genes are rarely found on the X (Parisi *et al*., 2003; Sturgill *et al*., 2007), and this could lead to systematic biases in the mean sex-specificity of selection coefficients of X-linked and autosomal mutations. Since the expected values of $K_a$ and $K_s$ for the X versus the autosomes are strongly affected by such sex-specificity (Rice, 1984; Charlesworth *et al*., 1987; Vicoso & Charlesworth, 2006), differences among chromosomes in these coefficients could mask an underlying faster-X effect, and this may be of concern in some of the previous studies. If genes with similar functions have similar selection coefficients, then focusing on gene groups with similar expected sex-specific effects could reveal faster-X evolution, as in the case of mammalian sperm proteins (Torgerson & Singh, 2003) and *Drosophila melanogaster* sex-specific genes (Pröschel *et al*., 2006; Baines *et al*., 2008).

Another approach is to study the same group of genes in an autosomal and an X-linked context. The genus *Drosophila* is particularly favourable for this purpose, since its species vary both in the number and organization of their chromosomes. Muller (1940) noted that genes linked in one species also appeared to be linked in others, and proposed that all karyotypic differences among *Drosophila* species could be explained by different combinations of the six basic chromosomal arms. The chromosomal arms involved have become known as Muller's elements A–F. Comparative analyses of *Drosophila* genomes have confirmed that, despite extensive within-arm rearrangements, only small fragments of DNA have been transposed between arms (Powell, 1997; Richards *et al*., 2005; Clark *et al*., 2007).

A rearrangement that is particularly useful for investigating faster-X effects is found in species of the *Drosophila pseudoobscura* clade. Following the split from the *D. melanogaster* group and their sister *Drosophila obscura* clade, the *D. pseudoobscura* clade ancestor became fixed for a fusion between Muller's element D (the autosomal 3L arm of *D. melanogaster*) and the homologue of the *D. melanogaster* X chromosome (element A), to form, respectively, the R and L arms of the *D. pseudoobscura* X (Muller, 1940). The forces shaping the evolution of the X chromosome should, therefore, also be acting on this new R arm of the *D. pseudoobscura* X chromosome. It should be noted that the XR arm has already acquired dosage compensation, and transcription levels are therefore similar on both arms of the X chromosome (Abraham & Lucchesi, 1974; Steinemann *et al*., 1996).

This system was used to explore the question of faster-X evolution by Counterman *et al*. (2004), who compared the $K_a/K_s$ values of genes on element D for *D. pseudoobscura*/*D. melanogaster* with those for *Drosophila simulans*/*D. melanogaster*. As expected under faster-X evolution, they found a significant excess of 3L-XR genes with higher $K_a/K_s$ for the *D. pseudoobscura*/*D. melanogaster* comparison relative to the *D. simulans*/*D. melanogaster* comparison. However, Musters *et al*. (2006) found no difference between element D genes and autosomal genes with respect to $K_a$ using *D. pseudoobscura*/*D. melanogaster* genomewide data. Counterman *et al*. (2004) also examined two different pairs of species, *D. melanogaster*/*D. simulans* and *D. pseudoobscura*/*Drosophila miranda*. *D. miranda* is a close relative of *D. pseudoobscura* and shares the new XR (Muller, 1940; Steinemann & Steinemann, 1998). They found that the fraction of genes with higher $K_a/K_s$ in the *D. pseudoobscura*/*D. miranda* pair than in *D. melanogaster*/*D. simulans* was larger for 3L-XR genes, although not significantly so. Thornton *et al*. (2006) used a similar approach with a larger dataset, but found no evidence for faster-X evolution. The lack of statistically significant evidence for faster-X effects in this case could be due to the species pairs used, since they have very low levels of divergence (about 4% for *D. pseudoobscura*/*D. miranda* at synonymous sites in Bartolomé *et al*., 2005). As such they may not be ideal

Table 1. *Average rates of DNA sequence evolution for the species pairs* D. pseudoobscura/D. affinis *and* D. melanogaster/D. yakuba

| | *D. pseudoobscura*/*D. affinis* | | | *D. melanogaster*/*D. yakuba* | | |
|---|---|---|---|---|---|---|
| | $K_a$ | $K_s$ | $K_a/K_s$ | $K_a$ | $K_s$ | $K_a/K_s$ |
| 3L-XR | 0·036 (0·004)[a] | 0·253 (0·008) | 0·138 (0·015) | 0·031 (0·003) | 0·323 (0·011) | 0·096 (0·01) |
| Autosomal | 0·020 (0·003) | 0·251 (0·01) | 0·080 (0·013) | 0·018 (0·003) | 0·269 (0·015) | 0·074 (0·014) |
| X-XL | 0·037 (0·007) | 0·263 (0·017) | 0·126 (0·019) | 0·038 (0·008) | 0·298 (0·023) | 0·115 (0·021) |

The average $K_a/K_s$ is estimated from the ratio of the averages of $K_a$ and $K_s$.
[a] S.E. values are given in parentheses.

for sequence comparisons, especially as some apparent inter-species differences may reflect polymorphisms within species (Bartolomé & Charlesworth, 2006). Similarly, the evidence for faster-X evolution found in a whole genome comparison using *D. melanogaster*/*D. simulans* and *D. pseudoobscura*/*Drosophila persimilis* should be interpreted with caution, as *D. persimilis* is even closer to *D. pseudoobscura* than *D. miranda* (Singh *et al*., 2007), so that effects of polymorphism levels may be confounded with true divergence.

Accordingly, we have used sequence data from the more distant species *Drosophila affinis*, with an average $K_s$ of about 23 % from both *D. pseudoobscura* and *D. miranda* (Bartolomé *et al*., 2005; Bartolomé & Charlesworth, 2006), in order to obtain more reliable estimates of sequence divergence. We used *D. melanogaster* and *Drosophila yakuba* for the control comparison, since their mean synonymous divergence is similar to that for *D. pseudoobscura* and *D. affinis* (see Table 1 below).

## 2. Materials and Methods

### (i) *Gene selection*

Our sample consisted of *D. pseudoobscura*, *D. affinis*, *D. melanogaster* and *D. yakuba* coding sequences for 69 3L-XR genes: 39 genes that are autosomal in both groups and 27 genes that are X-linked in both groups (from the X-XL chromosomal arm).

Carolina Bartolomé provided the coding sequence for the 39 autosomal *D. affinis* genes, for seven X-XL genes and for three 3L-XR genes (Bartolomé *et al*., 2005; Bartolomé & Charlesworth, 2006). The homologous sequences from other species were retrieved directly from the published genomes of *D. melanogaster*, *D. yakuba* and *D. pseudoobscura* with the NCBI Blast algorithm (http://www.ncbi.nlm.nih.gov/BLAST/).

For the other genes, *D. melanogaster* protein coding genes were downloaded from the FlyBase website (http://www.flybase.org). To minimize possible effects of close linkage to genes under selection, they were all chosen from regions of normal recombination in *D. melanogaster* (cytological region 3C3-15F3 for the X chromosome and 62A12-71A1 for 3L, as described by Charlesworth (1996)).

For each gene, we recovered all the corresponding mRNAs in the NCBI database with the Megablast algorithm (http://www.ncbi.nlm.nih.gov/BLAST/), and verified that they had a size between 1000 and 3500 base-pairs (bp), with at least 1000 bp without introns. We identified the *D. yakuba* homologue through the UCSC BLAT server (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) and the *D. pseudoobscura* homologue through the NCBI BLAST, and kept only genes whose location was syntenic for all three species.

### (ii) *DNA extraction*

DNA was extracted from males of a *D. affinis* line originally collected from Nebraska (stock number 0141·2; Drosophila Species Resource Center) using a Qiagen DNA extraction kit (Qiagen, Crawley, West Sussex, UK).

### (iii) *DNA sequencing*

Primers were designed with the DNAstar package and the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), using the *D. pseudoobscura* sequence to amplify 1000–1300 bp of the gene from *D. affinis*. Additional internal primers were designed for sequencing. Since the *D. affinis* sequences of the 39 autosomal genes we used were provided by Carolina Bartolomé (Bartolomé *et al*., 2005; Bartolomé & Charlesworth, 2006), all the genes we sequenced were on the *D. pseudoobscura*/*D. affinis* X chromosome (66 on the 3L-XR arm and 20 on the X-XL arm). PCR products were therefore directly sequenced on both strands using the BigDye (version 3) sequencing kit and run on an ABI 3730 Genetic Analyser (Applied Biosystems, Foster City, CA, USA) by the sequencing service of the School of Biological Sciences, University of Edinburgh. The sequences generated as part of this study have been

deposited in the GenBank Data Libraries under accession numbers EU931120–EU931205.

(iv) *Estimation of $K_a$ and $K_s$*

All sequences were translated and virtual protein sequences were aligned with the European Bioinformatics Institute ClustalW interface (http://www.ebi.ac.uk/Tools/clustalw/index.html). The resulting alignment was used to align the DNA sequences with Tranalign (http://phytophthora.vbi.vt.edu/cgi-bin/emboss.pl?_action=input&_app=tranalign), which aligns coding DNA according to a protein alignment. The $K_s$ and $K_a$ were calculated using Nei & Gojobori's (1986) model of substitution (Nei & Gojobori, 1986), implemented in DnaSP version 4.50 (Rozas *et al.*, 2003; http://www.ub.es/dnasp/), with the Jukes–Cantor correction for multiple hits. Since several models of substitution can lead to artifactual biases in $K_s$ when there are differences in codon usage bias (Bierne & Eyre-Walker, 2003), we also analysed the data using the Goldman & Yang (1994) model of substitution (using the PAML software package: http://abacus.gene.ucl.ac.uk/software/paml.html). The results from tests using the Goldman & Young (1994) $d_N$ and $d_S$ measures are given in the Supplementary Material.

(v) *Codon usage*

The alignments obtained for the $K_a$ and $K_s$ analyses were used to estimate the frequency of optimal codons, Fop, using CodonW (http://codonw.sourceforge.net/). We used the *D. melanogaster* table of preferred codons (Shields *et al.*, 1988), as patterns of codon usage have been shown to be highly conserved between *Drosophila* species (Powell & Moriyama, 1997; Clark *et al.*, 2007), and we needed to have the same set of codons for the different species under comparison.

(vi) *Statistical analysis*

Descriptive statistics and statistical tests were carried out using the StatView software (version 4.5, http://www.statview.com).

## 3. Results

(i) *Within-clade comparisons*

We obtained and aligned sequences for 69 3L-XR and 66 non-3L-XR (27 X-XL and 39 autosomal) genes for the species pairs *D. pseudoobscura*/*D. affinis* and *D. melanogaster*/*D. yakuba* (see Materials and Methods section). The average $K_a$, $K_s$ and $K_a/K_s$ values are shown in Table 1 and Fig. 1 (the values for individual genes are given in the Supplementary Material). Overall, the mean values seem to support the faster-X hypothesis: both $K_a$ and $K_a/K_s$ values are higher for X-linked chromosomal arms than autosomal arms in both the *D. pseudoobscura* and *D. melanogaster* groups.

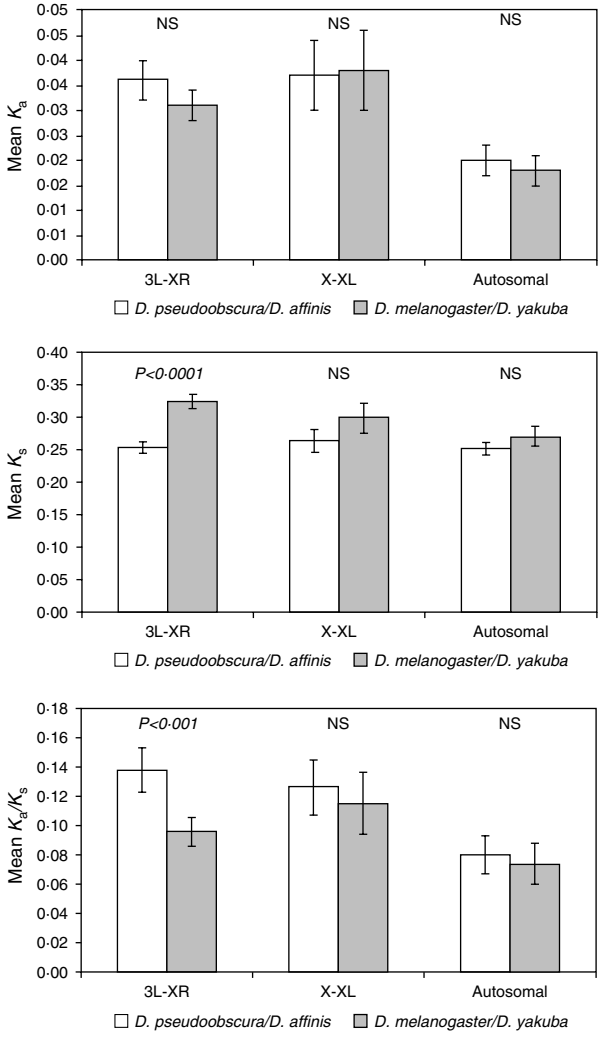


Fig. 1. Mean $K_a$, $K_s$ and $K_a/K_s$ for 3L-XR, X-XL and autosomal genes in the two clades.

It should, however, be noted that, while the higher $K_a$ and $K_a/K_s$ of the 3L-XR genes in the *D. pseudoobscura* group is in apparent agreement with the faster-X hypothesis, these genes also exhibit particularly high $K_a$ and $K_a/K_s$ values in the *D. melanogaster* group when compared with the rest of the autosomes (Table 2 and Fig. 1). This is likely to be caused by a sampling bias, as most of the autosomal genes were chosen from long, well-studied genes (Bartolomé *et al.*, 2005). In contrast, the genes that we selected for sequencing (most of the 3L-XRsample) correspond to small, unnamed (mostly unstudied) transcripts with no known function. Genes with no annotated function have been shown to be less constrained than genes with known functions (Clark *et al.*, 2007).

Table 2. *Significance values for comparisons of different chromosomal arms within species pairs*

| | Mann–Whitney test *P*-value | | | | | |
|---|---|---|---|---|---|---|
| | *D. pseudoobscura*/*D. affinis* | | | *D. melanogaster*/*D. yakuba* | | |
| | $K_a$ | $K_s$ | $K_a/K_s$ | $K_a$ | $K_s$ | $K_a/K_s$ |
| 3L-XR, XL | 0·938 | 0·594 | 0·791 | 0·968 | 0·359 | 0·651 |
| 3L-XR, auto | **0·004** | 0·941 | **0·006** | **0·005** | **0·009** | **0·036** |
| X-XL, auto | 0·099 | 0·588 | 0·065 | 0·077 | 0·306 | 0·069 |

Significant *P*-values are shown in boldface.

Table 3. *Average values of Fop for* 3L-XR, X-XL *and autosomal genes*

| | *D. affinis* | *D. pseudoobscura* | | *D. melanogaster* | *D. yakuba* |
|---|---|---|---|---|---|
| 3L-XR | **0·559 (0·01)**[a] | **0·568 (0·01)** | | 0·506 (0·009) | 0·527 (0·009) |
| | | | $P = 0·0001$ | | |
| X-XL | **0·589 (0·016)** | **0·596 (0·015)** | | **0·557 (0·019)** | **0·579 (0·015)** |
| | | | $P = 0·1301$ | | |
| Autosomes | 0·563 (0·017) | 0·562 (0·017) | | 0·540 (0·019) | 0·553 (0·019) |
| | | | $P = 0·7219$ | | |

Boldface values indicate X-linked genes. Since *D. melanogaster* has significantly reduced levels of codon usage for all the chromosomes compared with *D. yakuba* (not shown), we used *D. yakuba*/*D. pseudoobscura* to compare the Fop values in the two clades (using *D. yakuba*/*D. affinis* yields similar results). The *P*-values were obtained using Wilcoxon Signed Rank tests.
[a] S.E. values are given in parentheses.

Consistent with this, pairwise Mann–Whitney tests (Table 2) show that the autosomal sample has significantly lower $K_a$ values than the 3L-XR sample, and a lower $K_a/K_s$ in the *D. pseudoobscura*/*D. affinis* comparison. This should, however, not affect the comparison of rates of evolution on the same chromosomal arm between the two clades, since we have the same set of genes in all four species.

(ii) *Lower $K_s$ for X-linked genes*

Further examination of Fig. 1 shows that the most striking pattern is the smaller $K_s$ for the 3L-XR genes in the *D. pseudoobscura* clade compared with the *D. melanogaster* clade, whereas other chromosomal arms show no significant differences in $K_s$. A similar phenomenon was described in a human–chimpanzee comparison by Lu & Wu (2005), who found that X-linked genes had significantly lower $K_s$ values than autosomal genes. However, this is not seen for X-linked genes in comparisons among members of the *D. melanogaster* clade in our data, nor in a whole-genome analysis (Begun *et al*., 2007). Since the corresponding $K_a$ values do not differ significantly among the two clades, this comparison indicates that the faster-X effect suggested by the $K_a/K_s$ results is in fact caused by a lower $K_s$ for XR genes, not by faster non-synonymous site evolution. This could be a result of more effective selection to maintain codon usage bias in X-linked genes (McVean & Charlesworth, 1999); we examine this possibility below.

(iii) *Comparisons of codon usage*

We evaluated the frequency of optimal codons (Fop), a measure of codon usage bias (see the Materials and Methods section), for all genes in the sample (Table 3). Although X-XL genes have the highest levels of codon bias in each species, 3L-XR genes have similar levels of Fop to the autosomes in *D. pseudoobscura* and *D. affinis*. This might simply reflect sampling bias, since our *D. melanogaster* and *D. yakuba* 3L genes have lower levels of Fop than other autosomal genes, suggesting that direct comparisons between different chromosomal arms are, once again, unreliable. A more interesting result comes from comparisons between the same chromosomal arm in the two clades. *D. melanogaster* is known to have experienced a reduction in codon usage bias, thought to be due to a reduction in effective population size resulting in less efficient selection on this lineage (Akashi, 1995, 1996). We find, in agreement with previous studies, that *D. melanogaster* has significantly reduced levels of codon usage for all the

Table 4. *Proportions of genes with higher rates of evolution in the* D. pseudoobscura/D. affinis *pair compared with* D. melanogaster/D. yakuba

| | $K_a/K_s$ ($P=0{\cdot}45$) | $K_a$ ($P=0{\cdot}38$) | $K_s$ ($P=0{\cdot}002$) |
|---|---|---|---|
| 3L-XR | 70% | 54% | 19% |
| Others | 61% | 59% | 45% |

The *P*-values were obtained using Fisher's exact tests.

chromosomes compared with *D. yakuba* (not shown). We therefore used *D. pseudoobscura*/*D. yakuba* to compare the Fop values in the two clades (use of *D. affinis*/*D. yakuba* yields similar results). While Fop values are similar in the two clades for our control genes (Table 3), they are significantly higher for XR in the *D. pseudoobscura*/*D. affinis* pair than for 3L in *D. yakuba* ($P<0{\cdot}001$), consistent with the hypothesis that selection to maintain optimal codon usage is more efficient when loci are X-linked than when they are autosomal. These results will be considered further in the Discussion section.

(iv) *Pairwise comparisons*

In order to test whether the behaviour of the 3L-XR genes is different from that of the control genes, we also examined the proportion of genes that were evolving faster (genes that have higher $K_a$ or $K_a/K_s$) in the *D. pseudoobscura*/*D. affinis* pair than in *D. melanogaster*/*D. yakuba*. In the absence of a faster-X effect, this value will be similar for 3L-XR and non-3L-XR genes. If, on the other hand, there is faster-X evolution, this proportion will be higher for 3L-XR genes. The values are presented in Table 4. Unlike previous observations (Counterman *et al.*, 2004), there is no detectable faster-X effect for non-synonymous changes; despite the fact that the proportion of genes with higher $K_a/K_s$ in the *D. pseudoobscura* group is slightly higher for 3L-XR genes, this effect is not significant, and the opposite is observed for the $K_a$ values. Once again, a lower $K_s$ for 3L-XR genes in *D. pseudoobscura*/*D. affinis* is the only significant pattern.

(v) *Is there a faster-X effect in faster evolving genes?*

Various factors could have contributed to this failure to detect a faster-X effect for non-synonymous mutations (see the Introduction section). If a high $K_a/K_s$ at least partially reflects a prevalence of positive over purifying selection, then we might expect genes with high $K_a/K_s$ when autosomal to experience faster-X evolution when they become X-linked. Genes with low $K_a/K_s$, on the other hand, might experience slower-X evolution due to more efficient purifying

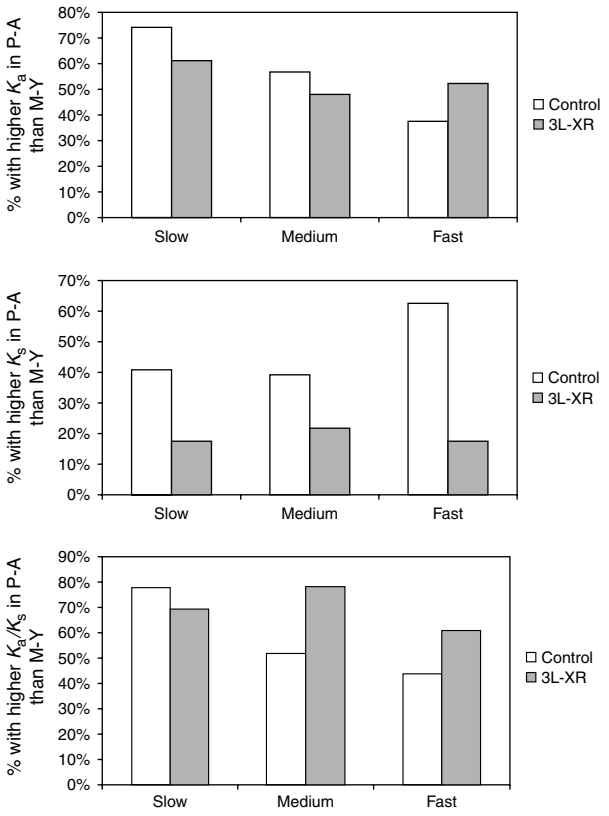


Fig. 2. 'Proportion of genes with higher $K_a$, $K_s$ and $K_a/K_s$ in the *D. pseudoobscura*/*D. affinis* pair'. P-A stands for *D. pseudoobscura-D. affinis* and M-Y for *D. melanogaster-D. yakuba*.

selection (Charlesworth *et al.*, 1987; McVean & Charlesworth, 1999). To test for this, we divided our sample into fast, medium and slow evolving genes in the following way: we ordered our 3L-XR genes according to their $K_a/K_s$ in the *D. melanogaster*/*D. yakuba* pair, and classified the first 23 as slow-evolving genes, the next 23 as medium-, and the last 23 as fast-evolving genes. We then repeated the analysis for these three classes, using as controls the non-3L-XR genes that had $K_a/K_s$ values in *D. melanogaster*/*D. yakuba* in the same range as the 3L-XR genes. The resulting sample contains 39 fast-evolving genes (23 3L-XR and 16 non-3L-XR), 46 medium (23 3L-XR and 23 non-3L-XR) and 50 slow evolving genes (23 3L-XR and 27 non-3L-XR). The proportion of genes with higher $K_a$, $K_s$ and $K_a/K_s$ in the *D. pseudoobscura*/*D. affinis* pair is shown in Fig. 2 (the numbers of genes involved is given in Table 5).

Although our sample size is now drastically reduced, the proportion of genes evolving faster in the *D. pseudoobscura*/*D. affinis* pair is behaving in the predicted direction, but the results are not statistically significant. For fast-evolving genes, 61% of the 3L-XR loci have higher $K_a/K_s$ in *D. pseudoobscura*/*D. affinis*, compared with 44% of the non-3L-XR loci,

Table 5. *Numbers of slow-, medium- and fast-evolving genes with different $K_a$, $K_s$ and $K_a/K_s$ values in the two clades*

| | *D. pseudoobscura*/ *D. affinis* | *D. melanogaster*/ *D. yakuba* | *P*-value |
|---|---|---|---|
| ***$K_a$*** | | | |
| Number of slow-evolving genes with a higher $K_a$ | | | |
| Non-3L-XR | 20 | 4 | 0·111 |
| 3L-XR | 14 | 9 | |
| Number of medium-evolving genes with a higher $K_a$ | | | |
| Non-3L-XR | 13 | 10 | 0·768 |
| 3L-XR | 11 | 12 | |
| Number of fast-evolving genes with a higher $K_a$ | | | |
| Non-3L-XR | 6 | 10 | 0·516 |
| 3L-XR | 12 | 11 | |
| ***$K_s$*** | | | |
| Number of slow-evolving genes with a higher $K_s$ | | | |
| Non-3L-XR | 11 | 16 | 0·121 |
| 3L-XR | 4 | 19 | |
| Number of medium-evolving genes with a higher $K_s$ | | | |
| Non-3L-XR | 9 | 14 | 0·337 |
| 3L-XR | 5 | 18 | |
| Number of fast-evolving genes with a higher $K_s$ | | | |
| Non-3L-XR | 10 | 6 | **0·007** |
| 3L-XR | 4 | 19 | |
| ***$K_a/K_s$*** | | | |
| Number of slow-evolving genes with a higher $K_a/K_s$ | | | |
| Non-3L-XR | 21 | 4 | 0·196 |
| 3L-XR | 16 | 9 | |
| Number of medium-evolving genes with a higher $K_a/K_s$ | | | |
| Non-3L-XR | 12 | 11 | 0·120 |
| 3L-XR | 18 | 5 | |
| Number of fast-evolving genes with a higher $K_a/K_s$ | | | |
| Non-3L-XR | 7 | 9 | 0·342 |
| 3L-XR | 14 | 9 | |

*P*-values were obtained using a Fisher's exact test.
Significant *P*-values are shown in boldface.

consistent with the faster-X hypothesis. Slow-evolving genes, on the other hand, do indeed show a moderate 'slow-X' effect: 70 % of 3L-XR genes have higher $K_a/K_s$ in the *D. pseudoobscura* group, versus 78 % for the control genes. This trend is partially caused by differences in $K_s$ (3L-XR fast-evolving genes have the lowest $K_s$ in the *D. pseudoobscura* group), but the values for $K_a$ are also consistent with the faster-X hypothesis. For fast-evolving genes, the 3L-XR arm has a higher proportion of genes with higher $K_a$ in the *D. pseudoobscura* group than the non-3L-XR sample. For slow evolving genes, we observe the opposite. None of these differences are, however, statistically significant. It should also be noted that, when we use the values of $d_N$ and $d_S$ obtained with PAML (Supplementary Material), we obtain similar patterns for $K_a$ and $K_s$, but not for $d_N/d_S$: in this case, there is a faster-X effect for all classes of gene, although it is closer to significance for fast-evolving genes.

## 4. Discussion

(i) *Is selection more efficient on the X chromosome?*

We have found that genes located on Muller's element D, which is X-linked in the *D. pseudoobscura* clade but autosomal in the *D. melanogaster* clade, have a significantly lower $K_s$ when they are X-linked than when they are autosomal (Fig. 1), whereas other genes do not differ in $K_s$ between the two clades. It is possible that this effect could be accounted for by a difference in mutation rate between X and autosomes. A higher mutation rate in males than in females leads to a lower rate of neutral evolution for the X chromosome, because the X is transmitted by females two-thirds of the time, whereas autosomes are transmitted by females only one-half of the time (Miyata *et al*., 1987; Vicoso & Charlesworth, 2006). In mammals and some lineages of *Drosophila*, there is evidence for this effect (Ebersberger *et al*., 2002; Singh

*et al*., 2007), but no statistically significant evidence supporting it has been found in X-autosome comparisons in *D. melanogaster*, *D. yakuba* or *D. pseudoobscura* (Bauer & Aquadro, 1997; Richards *et al*., 2005; Begun *et al*., 2007; Singh *et al*., 2007). While a higher rate of substitution of silent mutations has been found on the neo-*Y* chromosome of *D. miranda* compared with the neo-X, consistent with male-driven evolution (Bachtrog, 2008), this can be accounted for by the fixation of ancestral polymorphisms on the neo-*Y*, caused by its greatly reduced effective population size (Bartolomé & Charlesworth, 2006). It therefore seems unlikely that a male–female mutation rate difference could account for our observations on $K_s$.

The other possibility is that there is more effective selection to maintain codon usage on X-linked genes than on autosomal genes. Although synonymous substitutions are often treated as effectively neutral, there is ample evidence in *Drosophila* that synonymous codons are used in genes at different frequencies, because of selection for 'preferred' codons with higher efficiency or accuracy of translation (Powell & Moriyama, 1997; Clark *et al*., 2007). McVean & Charlesworth (1999) investigated the expected influence of X-linkage on codon usage bias, under the Li–Bulmer model of selection, genetic drift and reversible mutation between preferred and unpreferred codons (Li, 1987; Bulmer, 1991). If unpreferred codons are, on average, recessive or partially recessive in their effect on fitness, they will be selected out of the population more efficiently when they are on the X, leading to higher codon usage bias (McVean & Charlesworth, 1999). As we discuss in the next section, another (not mutually exclusive) possibility is that differences in effective population size between the X chromosome and the autosomes could be increasing the level of codon usage bias on the X.

Singh *et al*. (2005) estimated codon bias levels in *D. melanogaster*, *D. pseudoobscura* and *Caenorhabditis elegans* and found that these were higher on the X chromosome than on the autosomes in all three species. They excluded other factors that are correlated with codon usage bias, such as gene expression, gene length, recombination rate, gene density and protein evolution as possible causes for the X-autosome difference, suggesting that more efficient selection on the hemizygous male X is the main cause of increased codon usage bias on the X. They also compared *D. pseudoobscura* with *D. melanogaster*, and found a significant increase in codon bias for XR genes compared with their autosomal counterparts, consistent with the results in Table 3. More recently, the analysis of 12 *Drosophila* genomes has confirmed that the X chromosome has consistently higher levels of codon usage bias than the autosomes (Clark *et al*., 2007; Singh *et al*., 2007).

This evidence suggests at first sight that the reduced $K_s$ we have detected for genes on XR in the *D. pseudoobscura* clade, compared with the same genes on 3L in the *D. melanogaster* clade, reflects more effective selection to maintain codon usage bias, as a result of these genes becoming X-linked. A puzzle that arises from this observation, though, is why $K_s$ for the ancestral X is, if anything, slightly higher (but not significantly so) than for the autosomes in the *D. melanogaster*/*D. yakuba* comparison in both our data and in the genomewide comparison of the sequenced members of the *D. melanogaster* group (Begun *et al*., 2007); in particular, looking at divergence along the branch leading to *D. melanogaster* from *D. simulans*, $K_s$ for the X chromosome is approximately 1·1 times larger than for the autosomes (Begun *et al*., 2007).

One possible explanation is that codon usage in the *D. melanogaster* group is not in equilibrium, perhaps because of a historical reduction in population size, for which there is support from the genome sequence comparisons, especially for the *D. melanogaster* branch of the phylogeny (Akashi, 1995, 1996; Akashi *et al*., 2006; Begun *et al*., 2007; Nielsen *et al*., 2007). Takano-Shimizu (1999) showed that a reduction in effective population size is expected to result in a transient increase in the substitution rate above the equilibrium level, which is largest for genes with high levels of codon usage bias, since these depart the most from their final equilibrium. If the X chromosome has higher codon usage bias than the autosomes, X-linked genes would therefore show a higher transient substitution rate. The difference between XR and 3L that we have detected may, therefore, be primarily caused by an inflated rate of synonymous substitutions in the *D. melanogaster* group, with $K_s$ for XR in the *D. pseudoobscura* clade reflecting a level that is closer to the equilibrium value.

### (ii) *The effective population sizes of the X chromosome and the autosomes*

The ratio of Fop between the X and major autosomes in *Drosophila* is of the order of 1·1 (Table 3 and Singh *et al*., 2005). This is much larger than the maximum value predicted by McVean & Charlesworth (1999), who assumed that the effective population size $N_e$ for the X ($N_{eX}$) is three-quarters of that for the autosomes ($N_{eA}$). An even larger effect can, however, arise under a wide range of conditions, if the effective population size for the X chromosome is much bigger than this, as selection is less efficient on chromosomes with small effective population sizes.

Demographic effects (such as a female-biased sex ratio), increased recombination on the X chromosome, and increased variance of male reproductive success have all been shown be potential causes of

increased $N_{eX}/N_{eA}$. One way to assess if $N_e$ for the X chromosome is higher or lower than the autosomal $N_e$ is to compare neutral or nearly neutral polymorphism levels at X-linked and autosomal sites, since neutral polymorphism levels are proportional to the effective population size ($\pi$, the pairwise average diversity, is equal to $4N_e\mu$, where $\mu$ is the neutral mutation rate). Although they are not strictly neutral, synonymous sites are commonly used to approximate neutral polymorphism levels in *Drosophila*. Several studies have found that, in African populations of *D. melanogaster*, X-linked synonymous sites have higher levels of polymorphism than autosomal sites (Andolfatto, 2001; Kauer *et al.*, 2002; Mousset & Derome, 2004), suggesting that the effective population of the X is indeed higher than the expected three-fourth of the autosomal effective population size (Hutter *et al.*, 2007). This is likely to be contributing to the patterns of evolution we observe on the X chromosome, as it has a more powerful effect on the ratio of equilibrium codon usages for the X and autosomes than recessivity of deleterious mutational effects (Singh *et al.*, 2005).

### (iii) *Is there a faster-X effect for non-synonymous sites?*

Our data show a faster rate of non-synonymous site evolution for the ancestral X chromosome compared with the autosomes in both species comparisons (Tables 1 and 2), as has also been found in a genomewide comparison of the sequenced members of the *D. melanogaster* group (Begun *et al.*, 2007). For the reasons given in the first part of the Results section, however, this apparent support for faster-X evolution should be treated with some caution. The more critical test is to compare $K_a$ for 3L-XR between the two species pairs, and we found no significant effect; the significant difference in $K_a/K_s$ was entirely due to the difference in $K_s$ (Fig. 1), as discussed above. In fast-evolving genes, however, there was a suggestion of higher $K_a$ for XR-linked loci, and the opposite was observed for slow-evolving genes, but neither of these patterns was significant (Table 5 and Fig. 2).

Although they are not conclusive, these results are of interest in view of the contradictory results obtained by previous studies on faster-X evolution, as they suggest that such an effect might only be observed for genes that are under particularly strong positive selection and/or relaxed negative selection. In fact, all the studies that previously detected faster-X evolution in *Drosophila* were in some way biased towards fast-evolving genes. For example, Counterman *et al.* (2004) obtained part of their sample from a male-specific expressed sequence tag (EST) screen (Swanson *et al.*, 2001). Male-specific genes are not only expected to show an enhanced faster-X evolution, but it has also been claimed that they have a faster evolution than non-sex-biased genes in *Drosophila*, possibly as a consequence of increased positive selection (Zhang *et al.*, 2004). Consistent with this, several studies of male-biased or male reproductive genes have detected faster rates of evolution on the X (Torgerson & Singh, 2003; Wang & Zhang, 2004; Pröschel *et al.*, 2006; Baines *et al.*, 2008). Thornton *et al.* (2006), while following a similar approach to that of Counterman *et al.* (2004), chose their genes randomly, and observed no faster-X effect. This suggests that further studies on faster-X evolution should focus on fast-evolving genes, and that some of the discrepancies among different studies described in the Introduction section may arise from the use of different types of genes.

B.V. was supported by the Portuguese Foundation for Science and Technology, P.R.H. is supported by the Natural Environmental Research Council (UK) and B.C. was supported by the Royal Society (UK). We thank Dan Halligan for assistance with PAML. We also thank two anonymous referees for their comments on the manuscript.

## References

Abraham, I. & Lucchesi, J. C. (1974). Dosage compensation of genes on the left and right arms of the X chromosome of *Drosophila pseudoobscura* and *Drosophila willistoni*. *Genetics* **78**, 1119–1126.

Akashi, H. (1995). Inferring weak selection from patterns of polymorphism and divergence at 'silent' sites in *Drosophila* DNA. *Genetics* **139**, 1067–1076.

Akashi, H. (1996). Molecular evolution between *Drosophila melanogaster* and *D. simulans*: reduced codon bias, faster rates of amino acid substitution, and larger proteins in *D. melanogaster*. *Genetics* **144**, 1297–1307.

Akashi, H., Ko, W. -Y., Piao, S., John, A., Goel, P., Lin, C.-F. & Vitins, A. P. (2006). Molecular evolution in the *Drosophila melanogaster* species subgroup: frequent parameter fluctuations on the timescale of molecular divergence. *Genetics* **172**, 1711–1726.

Anderson, J. B., Sirjusingh, C. & Ricker, N. (2004). Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* **168**, 1915–1923.

Andolfatto, P. (2001). Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Molecular Biology and Evolution* **18**, 279–290.

Andolfatto, P. (2007). Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genome Research* **17**, 1755–1762.

Bachtrog, D. (2008). Evidence for male-driven evolution in *Drosophila*. *Molecular Biology and Evolution* **25**, 617–619.

Baines, J. F., Sawyer, S. A., Hartl, D. L. & Parsch, J. (2008). Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. *Molecular Biology and Evolution* **25**, 1639–1650.

Bartolomé, C. & Charlesworth, B. (2006). Evolution of amino-acid sequences and codon usage on the *Drosophila miranda* neo-sex chromosomes. *Genetics* **174**, 2033–2044.

Bartolomé, C., Maside, X., Yi, S., Grant, A. L. & Charlesworth, B. (2005). Patterns of selection on synonymous and nonsynonymous variants in *Drosophila miranda*. *Genetics* **169**, 1495–1507.

Bauer, V. L. & Aquadro, C. F. (1997). Rates of DNA sequence evolution are not sex-biased in *Drosophila melanogaster* and *D. simulans*. *Molecular Biology and Evolution* **14**, 1252–1257.

Begun, D. J., Holloway, A. K., Stevens, K., Hillier, L. W., Poh, Y. P., Hahn, M. W., Nista, P. M., Jones, C. D., Kern, A. D., Dewey, C. N., Pachter, L., Myers, E. & Langley, C. H. (2007). Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biology* **5**, e310.

Betancourt, A. J., Presgraves, D. C. & Swanson, W. J. (2002). A test for faster X evolution in *Drosophila*. *Molecular Biology and Evolution* **19**, 1816–1819.

Bierne, N. & Eyre-Walker, A. (2003). The problem of counting sites in the estimation of the synonymous and nonsynonymous substitution rates: implications for the correlation between the synonymous substitution rate and codon usage bias. *Genetics* **165**, 1587–1597.

Bierne, N. & Eyre-Walker, A. (2004). The genomic rate of adaptive amino acid substitution in *Drosophila*. *Molecular Biology and Evolution* **21**, 1350–1360.

Bulmer, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129**, 897–907.

Charlesworth, B. (1992). Evolutionary rates in partially self-fertilizing species. *American Naturalist* **140**, 126–148.

Charlesworth, B. (1996). Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genetics Research* **68**, 131–149.

Charlesworth, B., Coyne, J. A. & Barton, N. H. (1987). The relative rates of evolution of sex-chromosomes and autosomes. *American Naturalist* **130**, 113–146.

Clark, A. G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., *et al*. Drosophila 12 Genomes Consortium (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**, 203–218.

Connallon, T. (2007). Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses. *Molecular Biology and Evolution* **24**, 2566–2572.

Counterman, B. A., Ortiz-Barrientos, D. & Noor, M. A. (2004). Using comparative genomic data to test for fast-X evolution. *Evolution* **58**, 656–660.

Ebersberger, I., Metzler, D., Schwarz, C. & Paabo, S. (2002). Genomewide comparison of DNA sequences between humans and chimpanzees. *American Journal of Human Genetics* **70**, 1490–1497.

Goldman, N. & Yang, Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution* **11**, 725–736.

Graur, D. & Li, W.-H. (2000). *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates.

Haldane, J. B. S. (1924). A mathematical theory of natural and artificial selection. Part I. *Transactions of the Cambridge Philosophical Society* **23**, 19–41.

Hutter, S., Li, H., Beisswanger, S., De Lorenzo, D. & Stephan, W. (2007). Distinctly different sex ratios in African and European populations of *Drosophila melanogaster* inferred from chromosomewide single nucleotide polymorphism data. *Genetics* **177**, 469–480.

Kauer, M., Zangerl, B., Dieringer, D. & Schlotterer, C. (2002). Chromosomal patterns of microsatellite variability contrast sharply in African and non-African populations of *Drosophila melanogaster*. *Genetics* **160**, 247–256.

Kirkpatrick, M. & Hall, D. W. (2004). Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution* **58**, 437–440.

Li, W. H. (1987). Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *Journal of Molecular Evolution* **24**, 337–345.

Lu, J. & Wu, C.-I. (2005). Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. *Proceedings of the National Academy of Sciences of the USA* **102**, 4063–4067.

McDonald, J. H. & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**, 652–654.

McVean, G. T. & Charlesworth, B. (1999). A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genetics Research* **74**, 145–158.

Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K. & Yasunaga, T. (1987). Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harbor Symposia on Quantitative Biology* **52**, 863–867.

Mousset, S. & Derome, N. (2004). Molecular polymorphism in *Drosophila melanogaster* and *D. simulans*: what have we learned from recent studies? *Genetica* **120**, 79–86.

Muller, H. J. (1940). Bearing of the *Drosophila* work on systematics. In *The New Systematics* (ed. J. Huxley), pp. 185–268. Oxford, UK: Clarendon Press.

Musters, H., Huntley, M. A. & Singh, R. S. (2006). A genomic comparison of faster-sex, faster-X, and faster-male evolution between *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Journal of Molecular Evolution* **62**, 693–700.

Nei, M. & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418–426.

Nielsen, R., Bauer DuMont, V. L., Hubisz, M. J. & Aquadro, C. F. (2007). Maximum likelihood estimation of ancestral codon usage bias parameters in *Drosophila*. *Molecular Biology and Evolution* **24**, 228–235.

Orr, H. A. & Betancourt, A. J. (2001). Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157**, 875–884.

Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S. & Oliver, B. (2003). Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299**, 697–700.

Powell, J. R. (1997). *Progress and Prospects in Evolutionary Biology: The* Drosophila *Model*. New York: Oxford University Press.

Powell, J. R. & Moriyama, E. N. (1997). Evolution of codon usage bias in *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **94**, 7784–7790.

Pröschel, M., Zhang, Z. & Parsch, J. (2006). Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. *Genetics* **174**, 893–900.

Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742.

Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Meisel, R. P., Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J., van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E., Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D., Rives, C. M.,

Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D., Wheeler, D. A., Worley, K. C., Havlak, P., Durbin, K. J., Egan, A., Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y., Waldron, L., Verduzco, D., Clerc-Blankenburg, K. P., Dubchak, I., Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W., Gelbart, W., Weinstock, G. M. & Gibbs, R. A. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and *cis*-element evolution. *Genome Research* **15**, 1–18.

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.

Shields, D. C., Sharp, P. M., Higgins, D. G. & Wright, F. (1988). 'Silent' sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Molecular Biology and Evolution* **5**, 704–716.

Singh, N. D., Davis, J. C. & Petrov, D. A. (2005). X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis*. *Genetics* **171**, 145–155.

Singh, N. D., Larracuente, A. M. & Clark, A. G. (2007). Contrasting the efficacy of selection on the X and autosomes in *Drosophila*. *Molecular Biology and Evolution* **25**, 454–467.

Smith, N. G. & Eyre-Walker, A. (2002). Adaptive protein evolution in *Drosophila*. *Nature* **415**, 1022–1024.

Steinemann, M. & Steinemann, S. (1998). Enigma of Y chromosome degeneration: neo-*Y* and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **102–103**, 409–420.

Steinemann, M., Steinemann, S. & Turner, B. M. (1996). Evolution of dosage compensation. *Chromosome Research* **4**, 185–190.

Sturgill, D., Zhang, Y., Parisi, M. & Oliver, B. (2007). Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* **450**, 238–241.

Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F. & Aquadro, C. F. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **98**, 7375–7379.

Takano-Shimizu, T. (1999). Local recombination and mutation effects on molecular evolution in *Drosophila*. *Genetics* **153**, 1285–1296.

Thornton, K. & Long, M. (2002). Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. *Molecular Biology and Evolution* **19**, 918–925.

Thornton, K. & Long, M. (2005). Excess of amino acid substitutions relative to polymorphism between X-linked duplications in *Drosophila melanogaster*. *Molecular Biology and Evolution* **22**, 273–284.

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 Research* **16**, 498–504.

Torgerson, D. G. & Singh, R. S. (2003). Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Molecular Biology and Evolution* **20**, 1705–1709.

Vicoso, B. & Charlesworth, B. (2006). Evolution on the X chromosome: unusual patterns and processes. *Nature Reviews Genetics* **7**, 645–653.

Wang, X. & Zhang, J. (2004). Rapid evolution of mammalian X-linked testis-expressed homeobox genes. *Genetics* **167**, 879–888.

Welch, J. J. (2006). Estimating the genomewide rate of adaptive protein evolution in *Drosophila*. *Genetics* **173**, 821–837.

Zeyl, C., Vanderford, T. & Carter, M. (2003). An evolutionary advantage of haploidy in large yeast populations. *Science* **299**, 555–558.

Zhang, Z., Hambuch, T. M. & Parsch, J. (2004). Molecular evolution of sex-biased genes in *Drosophila*. *Molecular Biology and Evolution* **21**, 2130–2139.