# Protein Evolution and Codon Usage Bias on the Neo-Sex Chromosomes of *Drosophila miranda*

**Doris Bachtrog**[1]

*Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received April 21, 2003
Accepted for publication July 11, 2003

ABSTRACT

The neo-sex chromosomes of *Drosophila miranda* constitute an ideal system to study the effects of recombination on patterns of genome evolution. Due to a fusion of an autosome with the Y chromosome, one homolog is transmitted clonally. Here, I compare patterns of molecular evolution of 18 protein-coding genes located on the recombining neo-X and their homologs on the nonrecombining neo-Y chromosome. The rate of protein evolution has significantly increased on the neo-Y lineage since its formation. Amino acid substitutions are accumulating uniformly among neo-Y-linked genes, as expected if all loci on the neo-Y chromosome suffer from a reduced effectiveness of natural selection. In contrast, there is significant heterogeneity in the rate of protein evolution among neo-X-linked genes, with most loci being under strong purifying selection and two genes showing evidence for adaptive evolution. This observation agrees with theory predicting that linkage limits adaptive protein evolution. Both the neo-X and the neo-Y chromosome show an excess of unpreferred codon substitutions over preferred ones and no difference in this pattern was observed between the chromosomes. This suggests that there has been little or no selection maintaining codon bias in the *D. miranda* lineage. A change in mutational bias toward AT substitutions also contributes to the decline in codon bias. The contrast in patterns of molecular evolution between amino acid mutations and synonymous mutations on the neo-sex-linked genes can be understood in terms of chromosome-specific differences in effective population size and the distribution of selective effects of mutations.

IN natural populations, the effective population size, $N_e$, is a key variable that determines the magnitude of random sampling effects on gene frequencies (*i.e.*, genetic drift; Kimura 1983). $N_e$ varies among species and across different regions of the genome within a species. In outcrossing taxa, sex chromosomes or mitochondria, for example, have a lower $N_e$ than that of autosomes due to sex-specific inheritance and differences in ploidy levels (Caballero 1995). Moreover, the recombinational landscape of a chromosome can also have an influence on $N_e$. Under the standard neutral model of sequence evolution, the amount of neutral variability—polymorphisms with no significant fitness effects—is directly proportional to the effective population size and the mutation rate (Kimura 1983). However, a locus under directional selection can increase the magnitude of random sampling effects at closely linked loci, thereby enhancing the influence of genetic drift (Hill and Robertson 1966). This "Hill-Robertson effect" predicts lower levels of neutral variability in low-recombining regions, due to the reduction in $N_e$ imposed by linked selected sites. In accordance with this prediction, regions of reduced or no recombination in Drosophila harbor lower levels of variation at silent sites (Begun and Aquadro 1992; Zurovcova and Eanes 1999; Bachtrog and Charlesworth 2002).

Population genetics theory predicts that $N_e$ is a key determinant of the efficacy of natural selection. A mutation is effectively neutral (*i.e.*, its fate is mostly determined by random genetic drift) if the product of the effective population size $N_e$ and the selection coefficient $s$ is below unity ($|N_e s| < 1$; Kimura 1983). Thus, since selection at linked loci decreases $N_e$ (the Hill-Robertson effect), the efficacy of selection is also expected to be reduced in regions of restricted recombination. Some empirical evidence suggests that the overall level of adaptation is indeed reduced in regions of the genome where genetic recombination is reduced or absent (Weinreich and Rand 2000; Bachtrog and Charlesworth 2002; Betancourt and Presgraves 2002). Clonally transmitted genomes like organelles show reduced levels of adaptation with respect to protein sequences (Lynch and Blanchard 1998; Weinreich and Rand 2000). Codon bias—the nonrandom usage of codons encoding the same amino acid—is reduced in regions of low recombination in *Drosophila melanogaster* (Kliman and Hey 1993), consistent with Hill-Robertson effects (but see Marais and Piganeau 2002). Also, the rate of replacement substitution is reduced in low-recombining regions in the *D. melanogaster* genome, compatible with a lower rate of protein adaptation in

[1]*Present address:* Department of Molecular Biology and Genetics, 227 Biotechnology Bldg., Cornell University, Ithaca, NY 14853. E-mail: doris.bachtrog@cornell.edu

regions of restricted recombination (Betancourt and Presgraves 2002).

The Y chromosome provides an extreme example of the effects of genetic linkage on patterns of genome evolution. In species with male heterogamety, the Y chromosome lacks genetic recombination over most or all of its length. The X and Y chromosomes are thought to have descended from an ordinary pair of autosomes, with the almost complete erosion of gene function on the Y being a direct consequence of its lack of sexual recombination (Bull 1983; Charlesworth 1990). Well-studied Y chromosome systems, like those of humans or *D. melanogaster*, are ancient and have lost most of their active genes (Bachtrog and Charlesworth 2001; Carvalho *et al.* 2001; Lahn *et al.* 2001). Since the magnitude of interference between loci depends on the number of loci under selection (Li 1987; McVean and Charlesworth 2000), the opportunity for detecting the signatures of Hill-Robertson effects is considerably reduced compared to the early stages of Y chromosome evolution (Charlesworth and Charlesworth 2000). Species where a neo-Y/neo-X chromosome pair has recently been formed by a fusion between an autosome and a sex chromosome offer excellent opportunities to study the processes involved in Y chromosome degeneration (Lucchesi 1978; Charlesworth 1996; Steinemann and Steinemann 1998), since those systems still contain many functional loci under selection. As a result of such a fusion, one member of the pair of autosomes (the neo-Y) is transmitted from father to son like the true Y chromosome, while its homolog (the neo-X) cosegregates with the X (Lucchesi 1978). In Drosophila, the absence of crossing over in males ensures that such a neo-Y chromosome is genetically isolated from its partner (Gethmann 1988).

In *D. miranda*, a close relative of *D. pseudoobscura*, the Y chromosome has become fused to an autosome (element C; Muller 1940; see Figure 1). There has been partial loss of gene activity on the neo-Y, with some genes still retaining their activity (Bachtrog and Charlesworth 2002), others showing little or no expression (Steinemann *et al.* 1993), and others having been apparently completely lost (Steinemann and Steinemann 1999). This degeneration is associated with the acquisition of dosage compensation by some genes on the neo-X, by the standard Drosophila molecular mechanism (Marin *et al.* 2000). DNA sequence divergence between neo-X and neo-Y homologs indicates that they began to diverge ~1–1.5 MYA (*i.e.*, ~10 $N_e$ generations; Bachtrog and Charlesworth 2002).

Evolutionary theory predicts that deleterious mutations should accumulate on a nonrecombining chromosome (Felsenstein 1974), while beneficial mutations are less likely to be fixed (Peck 1994; Orr and Kim 1998). In a previous study based on three protein-coding genes, Bachtrog and Charlesworth (2002) demonstrated a significantly higher rate of protein evolution for one gene on the neo-X-branch of the phylogeny (*CycB*), and a significantly higher rate on the neo-Y branch for another gene (*robo*). While this pattern was suggestive, here I compare patterns of molecular evolution at 18 genes located on the neo-sex chromosomes of *D. miranda*. I investigate two aspects of molecular evolution. First, the rate of replacement substitutions is analyzed, to more generally characterize rates of protein evolution on the neo-X and neo-Y. Second, I investigate several aspects of the evolution of codon bias. We have previously shown a 30-fold reduction in $N_e$ of the neo-Y chromosome in *D. miranda* compared to that of the neo-X (Bachtrog and Charlesworth 2000; Bachtrog and Charlesworth 2002). Since the efficacy of selection is dependent on $N_e$, both lower levels of codon bias and reduced protein adaptation are expected on the neo-Y chromosome.

## MATERIALS AND METHODS

**Sequence information on the genes investigated:** Table 1 lists all the genes surveyed and the source of the sequences. The genes investigated were isolated either from a genomic library constructed from *D. miranda* (*CG11136*, *CG11159*, *CG13437*, *CG16799*, *CG30152*, *CG9025*, *Cyclin B*, *exuperantia 1*, *roundabout*, *deadpan*, *zipper*, and *maleless*) or using degenerate primers for PCR amplification (*even-skipped* and *engrailed*) or were taken from GenBank (*Lcp1*, *Lcp2*, *Lcp3*, and *Lcp4*). All genes investigated are located on the neo-sex chromosomes of *D. miranda*, as confirmed by *in situ* hybridization (primary sources; results not shown). For *CG11136*, *CG11159*, *CG13437*, *CG16799*, *CG30152*, *CG9025*, *Cyclin B*, *exuperantia 1*, *roundabout*, *deadpan*, *even-skipped*, *Lcp1*, *Lcp2*, *Lcp3*, and *Lcp4* the entire coding sequence was investigated; for the remaining genes, partial coding sequences were analyzed. All genes were also sequenced in a strain of *D. pseudoobscura*, which allows the inference of lineage-specific effects on the neo-X and the neo-Y chromosome (Figure 1). Information on library screening and PCR amplification can be found in the primary sources (Steinemann *et al.* 1996; Bachtrog and Charlesworth 2002; Bachtrog 2003) or is available from the author on request. Sequences of *D. miranda* and *D. pseudoobscura* were aligned manually. Sites with gaps were excluded from the analyses.

**Rates of protein evolution:** Several methods were used to estimate rates of protein evolution on the neo-sex chromosomes of *D. miranda*. First, rates of replacement and synonymous substitution per site, $K_a$ and $K_s$, respectively, between the neo-sex chromosomes of *D. miranda* and *D. pseudoobscura* were estimated using the method of Comeron (1995). Second, amino acid mutations and synonymous substitutions were assigned to either the neo-X or the neo-Y branch of the evolutionary tree assuming parsimony and using the *D. pseudoobscura* sequence as an outgroup (Akashi 1995). Sites with multiple substitutions were excluded from the parsimony analyses. Also, codons containing multiple substitutions were excluded from the analysis; thus, some values differ slightly from previous analyses (Bachtrog and Charlesworth 2002; Bachtrog 2003). Finally, a maximum-likelihood method was used to estimate $K_a$ and $K_s$ that accounts for unequal transition and transversion rates and unequal base and codon frequencies,

**TABLE 1**

**Genes investigated located on the neo-sex chromosome of *D. miranda***

| Locus | Full name | Frameshift/stop codons on neo-Y | Codons | Source |
|---|---|---|---|---|
| *CG11136* | *CG11136* | Functional | 804 | Bachtrog (2003) |
| *CG11159* | *CG11159* | Functional | 146 | Bachtrog (2003) |
| *CG16799* | *CG16799* | Functional | 150 | Bachtrog (2003) |
| *Cyclin B* | *Cyclin B* | Functional | 534 | Bachtrog and Charlesworth (2002) |
| *dpn* | *deadpan* | Functional | 452 | This study |
| *eve* | *even-skipped* | Functional | 360 | Bachtrog and Charlesworth (2002) |
| *exu1* | *exuperantia 1* | Functional | 481 | Bachtrog (2003) |
| *Lcp1* | *Lcp1* | Functional | 138 | Steinemann *et al.* (1996); this study |
| *Lcp3* | *Lcp3* | Functional | 112 | Steinemann *et al.* (1996); this study |
| *robo* | *roundabout* | Functional | 1398 | Bachtrog and Charlesworth (2002) |
| *zip* | *Zipper* | Functional | 372 | This study |
| *eng* | *engrailed* | Functional | 409 | This study |
| *CG13437* | *CG13437* | Frameshift mutation, large deletion | 68 | Bachtrog (2003) |
| *CG30152* | *CG30152* | Frameshift mutation | 195 | Bachtrog (2003) |
| *CG9025* | *CG9025* | Frameshift mutation, large deletion | 231 | Bachtrog (2003) |
| *mle* | *maleless lethal* | Frameshift mutation | 359 | This study |
| *Lcp2* | *Lcp2* | Start codon missing | 126 | Steinemann *et al.* (1996); this study |
| *Lcp4* | *Lcp4* | Frameshift mutation, large deletion | 112 | Steinemann *et al.* (1996); this study |
| Total | | | 6447 | |

as implemented in the *PAML* software package (http://abacus.gene.ucl.ac.uk/software/paml.html). Two approaches were used to assess differences in the rates of substitutions on the neo-X and neo-Y chromosome. First, the numbers of amino acid substitutions on the neo-sex chromosomes were compared using Tajima's relative rate test (Tajima 1993). Second, maximum-likelihood analysis of data on the neo-sex-linked genes was used to compare a model in which $K_a/K_s$ was assumed to be the same on the neo-X and neo-Y branch with a model in which it was allowed to differ, as implemented in *PAML*. Significance was assessed using the chi-square test; twice the difference in log-likelihood between models is assumed to be approximately chi-square distributed with degrees of freedom equal to the difference in the numbers of parameters.

**Codon bias analysis:** As a measure of codon bias, the effective number of codons (ENC; see Wright 1990) as well as the frequency of optimal codons (FOP; see Stenico *et al.* 1994) was used. Optimal (preferred) codons can be identified as those whose frequencies (within a synonymous family) show a positive correlation with the degree of bias at other codons in the same genes (Akashi 1995). Optimal codons have been found to be conserved between the distantly related species *D. melanogaster* and *D. pseudoobscura* (Akashi and Schaeffer 1997); thus, I used the preferred codons identified in *D. pseudoobscura* (Akashi and Schaeffer 1997) to classify synonymous mutations in *D. miranda*. A parsimony approach was used to map preferred and unpreferred synonymous substitutions to either the neo-X or the neo-Y branch of the phylogenetic tree, and the numbers were compared using Tajima's relative rate test (Tajima 1993). Sites with multiple hits were excluded from the parsimony analyses, as were codons with multiple substitutions.

## RESULTS

**Rates of protein evolution:** Table 2 and Table 3 summarize the rates of synonymous and nonsynonymous substitutions on the neo-sex chromosomes of *D. miranda*. Two genes, *exu1* and *CycB*, have an exceptionally high rate of replacement substitutions on the neo-X lineage (see Table 2 and Figure 2). For both genes, I have demonstrated that the elevated $K_a$ is due to the action of recent positive selection (Bachtrog and Charlesworth 2002; Bachtrog 2003), as indicated by significant McDonald-Kreitman (McDonald and Kreitman 1991) and Hudson-Kreitman-Aguadé tests (Hudson *et al.* 1987). Most other genes, however, show a higher rate of amino acid evolution on the branch of the phylogeny leading to the neo-Y compared to the neo-X branch (Table 2; Figure 2). Table 2 gives the pairwise estimates of synonymous and nonsynonymous divergence between the *D. miranda* neo-X and neo-Y chromosome and *D. pseudoobscura* for all loci analyzed.

**TABLE 2**

**Synonymous and nonsynonymous substitution rates between *D. pseudoobscura* and the *D. miranda* neo-X and neo-Y chromosome, using the method of Comeron (1995)**

| | Neo-X/*D. pseudoobscura* | | | Neo-Y/*D. pseudoobscura* | | |
|---|---|---|---|---|---|---|
| | $K_s$ | $K_a$ | $K_a/K_s$ | $K_s$ | $K_a$ | $K_a/K_s$ |
| *CG11136* | 5.94 | 0.15 | 0.03 | 5.96 | 1.13 | 0.19 |
| *CG11159* | 0.68 | 0.00 | 0.00 | 3.74 | 0.35 | 0.09 |
| *CG16799* | 4.44 | 1.97 | 0.44 | 4.59 | 2.76 | 0.60 |
| *CycB* | 5.80 | 2.64 | 0.46 | 4.46 | 1.19 | 0.27 |
| *dpn* | 4.43 | 0.11 | 0.03 | 4.53 | 0.66 | 0.14 |
| *eng* | 2.75 | 0.48 | 0.17 | 5.63 | 0.49 | 0.09 |
| *eve* | 3.53 | 0.00 | 0.00 | 5.80 | 0.27 | 0.05 |
| *exu1* | 6.10 | 3.59 | 0.59 | 4.58 | 2.22 | 0.48 |
| *Lcp1* | 0.77 | 0.36 | 0.47 | 1.53 | 1.68 | 1.10 |
| *Lcp3* | 3.75 | 0.82 | 0.22 | 2.49 | 1.19 | 0.48 |
| *robo* | 3.12 | 0.26 | 0.08 | 3.97 | 0.96 | 0.24 |
| *zip* | 1.63 | 0.73 | 0.45 | 1.07 | 1.22 | 1.14 |
| Functional | 3.58 (3.10)[a] | 0.93 (0.49)[a] | 0.26 (0.15)[a] | 4.03 (3.93)[a] | 1.18 (1.07)[a] | 0.29 (0.27)[a] |
| *CG13437* | 4.68 | 0.00 | 0.00 | 6.99 | 0.61 | 0.09 |
| *CG30152* | 2.37 | 0.24 | 0.10 | 3.40 | 1.68 | 0.49 |
| *CG9025* | 3.51 | 0.00 | 0.00 | 3.27 | 1.09 | 0.33 |
| *Lcp2* | 4.44 | 1.38 | 0.31 | 3.12 | 1.45 | 0.46 |
| *Lcp4* | 6.34 | 0.82 | 0.13 | 5.40 | 1.79 | 0.33 |
| *mle* | 1.94 | 0.25 | 0.13 | 3.05 | 0.35 | 0.11 |
| Nonfunctional | 3.88 | 0.45 | 0.12 | 4.20 | 1.16 | 0.28 |
| Total | 3.68 (3.39)[a] | 0.77 (0.47)[a] | 0.21 (0.14)[a] | 4.09 (4.03)[a] | 1.17 (1.10)[a] | 0.29 (0.27)[a] |

[a] Estimates of $K_a$, $K_s$, and $K_a/K_s$ excluding the *CycB* and *exu1* gene are given in parentheses.

Excluding the two genes that have undergone positive selection on the neo-X chromosome (*i.e.*, *exu1* and *CycB*), the total number of replacement substitutions is about sixfold higher on the branch of the phylogeny leading to the neo-Y compared to the neo-X (83 *vs.* 14 replacement mutations; see Table 3). This excess of replacement substitutions on the neo-Y is highly significant by Tajima's relative rate test ($P < 0.001$). Including *exu1* and *CycB*, however, still indicates a significant excess of replacement substitutions on the neo-Y chromosome (98 *vs.* 55 replacement mutations; $P = 0.015$). In contrast, the numbers of synonymous substitutions are similar on the two branches (100 *vs.* 80 mutations; $P = 0.35$; Table 3). Assuming that synonymous substitutions accumulate neutrally, this indicates that the mutation rates do not greatly differ between the neo-Y and the neo-X chromosome. The elevation in the rate of amino acid substitutions on the neo-Y is consistent with a reduction in the efficacy of selection against deleterious mutations for Y-linked genes. Six of the loci investigated appear to be nonfunctional on the neo-Y chromosome, as indicated by frameshift mutations, large deletions, or stop codons (see Table 1). Excluding these loci from the analysis, the rate of replacement accumulation is still about sixfold higher on the neo-Y than on the neo-X (64 *vs.* 11 substitutions; $P < 0.001$).

Using a likelihood-based approach reveals a similar pattern. A model with a fixed rate of $K_a/K_s$ for all branches was compared to a model with three different rates (neo-X branch, neo-Y branch, and one for the branch including *pseudoobscura* and the time in the *miranda* lineage before the split of the neo-sex chromosomes; see Figure 1). Again, most of the genes showed a higher rate of $K_a/K_s$ on the branch leading to the neo-Y chromosome (results not shown). For individual loci, however, the likelihood-ratio test is significant for only a few cases, probably due to a lack of statistical power. Maximum-likelihood analysis on the combined data set (excluding *CycB* and *exu1*) indicates that a model that allows for a different $K_a/K_s$ in each branch fits the data significantly better than a model that assumes the same rate for each branch of the phylogeny ($2\Delta L = 34.9$; $P < 0.001$). Thus, the neo-Y chromosome has, on average, a significantly higher $K_a/K_s$ ratio than the neo-X chromosome (0.26 *vs.* 0.07). Restricting the analysis to those genes that show intact open reading

**TABLE 3**

**Parsimony-based patterns of molecular evolution of genes located on the neo-sex chromosomes of *D. miranda***

| | | | Neo-X | | Neo-Y | |
|---|---|---|---|---|---|---|
| | $S^a$ | $R^a$ | $Syn^b$ | $Rep^b$ | $Syn^b$ | $Rep^b$ |
| *CG11136* | 600.3 | 1811.7 | 13 | 0 | 14 | 15 |
| *CG11159* | 97.7 | 340.3 | 0 | 0 | 2 | 1 |
| *CG16799* | 96.5 | 353.5 | 3 | 0 | 2 | 3 |
| *CycB* | 390.1 | 1211.9 | 9 | 21 | 6 | 6 |
| *dpn* | 323.8 | 1026.2 | 9 | 0 | 11 | 4 |
| *eng* | 305.6 | 903.4 | 2 | 3 | 9 | 2 |
| *eve* | 274.8 | 805.2 | 2 | 0 | 8 | 2 |
| *exu1* | 336.8 | 1106.2 | 8 | 20 | 4 | 9 |
| *Lcp1* | 97.8 | 316.2 | 0 | 0 | 1 | 4 |
| *Lcp3* | 82.4 | 253.6 | 4 | 1 | 2 | 2 |
| *robo* | 1028.5 | 3165.5 | 17 | 3 | 27 | 22 |
| *zip* | 258.9 | 857.1 | 3 | 4 | 2 | 9 |
| Functional | 3893.1 | 12150.9 | 70 (53)[c] | 52 (11)[c] | 88 (78)[c] | 79 (64)[c] |
| *CG13437* | 48.9 | 155.1 | 1 | 0 | 1 | 1 |
| *CG30152* | 135.9 | 449.1 | 1 | 0 | 3 | 5 |
| *CG9025* | 165.7 | 527.3 | 3 | 0 | 1 | 6 |
| *Lcp2* | 90.0 | 288.0 | 3 | 1 | 3 | 3 |
| *Lcp4* | 38.7 | 117.3 | 1 | 0 | 0 | 1 |
| *mle* | 272.3 | 840.7 | 1 | 2 | 4 | 3 |
| Nonfunctional | 751.4 | 2377.6 | 10 | 3 | 12 | 19 |
| Total | 4644.5 | 14528.5 | 80 (63)[c] | 55 (14)[c] | 100 (90)[c] | 98 (83)[c] |

[a] *S* and *R* refer to the number of synonymous sites and replacement sites, respectively, as computed using Nei and Gojobori (1986).

[b] Parsimony-based estimates of the number of synonymous (*Syn*) and replacements (*Rep*) changes in the neo-X and neo-Y lineage, using *D. pseudoobscura* as an outgroup sequence.

[c] Number of substitutions excluding the *CycB* and *exu1* genes is given in parentheses.

frames on the neo-Y chromosome results in the same conclusion ($K_a/K_s$ of 0.06 and 0.24 on neo-X and neo-Y, respectively; $2\Delta L = 24.4$; $P < 0.001$).

Figure 2 shows the maximum-likelihood estimates of $K_a$ and $K_s$ for the neo-X and the neo-Y branches. The variance in the number of amino acid substitutions among genes on the neo-Y is much smaller than that on the neo-X (Figure 2). Amino acid substitutions occur uniformly among neo-Y-linked genes (see Figure 2; $\chi^2 = 16.9$; $P > 0.5$), as expected if all loci on the neo-Y are suffering from a reduced effectiveness of natural selection. In addition, synonymous substitutions occur uniformly among neo-Y genes ($\chi^2 = 6.6$; $P > 0.9$). In contrast, there is significant heterogeneity in the rate of amino acid evolution among genes on the neo-X chromosome (see Figure 2; $\chi^2 = 154.5$; $P < 0.001$), yet no heterogeneity is observed for synonymous substitutions ($\chi^2 = 9.3$; $P > 0.9$). This strongly suggests that frequent positive selection is limited to only a fraction of genes on the neo-X.

**Codon bias analysis:** Reduced efficacy of natural selection for Y-linked genes predicts a reduction in level of codon bias for neo-Y- compared with neo-X-linked genes. In contrast to this expectation, the average codon bias for neo-X- and neo-Y-linked sequences is identical (mean ENC is 41 and mean FOP is 0.70 for both chromosomes; Figure 2) and only slightly higher than that in *D. pseudoobscura* (mean ENC is 41 and mean FOP is 0.71). Thus, there is no overall indication that natural selection is less effective at preventing unpreferred codons from accumulating on the neo-Y chromosome compared to the neo-X. However, for such closely related sequences as the neo-X, neo-Y, and *D. pseudoobscura*, a large number of sites are occupied by identical codons due to shared ancestry. Thus, the absolute values of codon bias will give little information about ongoing selection pressures on codon usage bias in the different lineages.

I also compared the numbers of substitutions from preferred to unpreferred synonymous codons ("unpre-

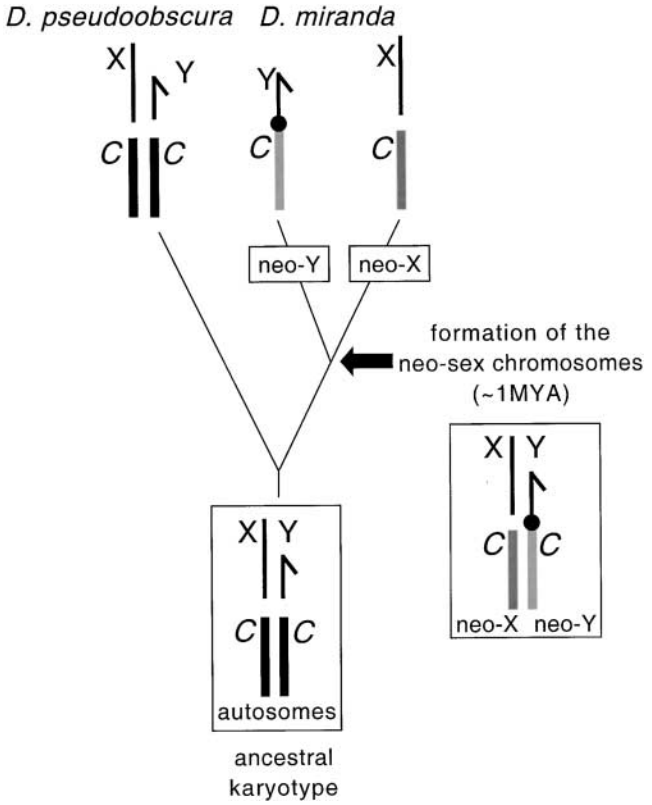


FIGURE 1.—Schematic of phylogenetic relationships of the sequences surveyed. The neo-sex chromosomes of *D. miranda* formed ~1 MYA, while *D. miranda* and *D. pseudoobscura* diverged ~2 MYA. Parsimony assumptions and outgroup sequence information from *D. pseudoobscura* were used to infer whether fixations occurred in either the neo-X or the neo-Y lineage.

ferred substitutions") to the number of unpreferred to preferred synonymous codons ("preferred substitutions"; see AKASHI 1995). If selection for codon bias has recently been reduced on the neo-Y lineage, an excess of unpreferred substitutions should be detected among neo-Y-linked genes compared to neo-X-linked genes. Surprisingly, both the neo-X and the neo-Y chromosome show a strong excess of unpreferred over preferred codon substitutions (Table 4; Figure 3). On the neo-X, 58 unpreferred and 13 preferred substitutions were detected, and on the neo-Y chromosome, 62 unpreferred and 18 preferred substitutions occurred. This excess of unpreferred over preferred substitutions on the neo-sex-linked genes is highly significant, for both chromosomes independently, as well as for the combined data set (Tajima's relative rate test, $P < 0.001$ for all cases). The same is seen if only genes that show intact open reading frames on the neo-Y are analyzed (52 *vs.* 13 unpreferred over preferred substitutions on the neo-X and 55 *vs.* 15 on the neo-Y; $P < 0.001$). The excess of unpreferred fixations on both neo-sex chromosomes suggests a genome-wide reduction in selection intensity for codon usage bias in the *D. miranda* lineage.

## DISCUSSION

As predicted by population genetics theory, the effective population size of the nonrecombining neo-Y chromosome of *D. miranda* is strongly reduced compared with its homolog, the neo-X chromosome. Both microsatellite and nucleotide polymorphism data suggest that $N_e$ of the neo-Y chromosome is ~30-fold lower than that of the neo-X (BACHTROG and CHARLESWORTH 2000, 2002). This reduction in $N_e$ is, in turn, predicted to result in a lower efficacy of both positive and negative selection on the Y chromosome. Here, I investigate two aspects of molecular evolution for 18 genes located on the neo-sex chromosomes of *D. miranda*: the rate of amino acid substitution and the level of codon bias. Apart from two loci that have undergone adaptive protein evolution on the neo-X chromosome, loci on the neo-Y chromosome generally show a higher rate of amino acid substitutions than do loci on the neo-X. Moreover, amino acid substitutions were found to be accumulating uniformly among neo-Y-linked genes. This is consistent with the observed lower effective population size of the neo-Y compared to the neo-X, reducing the efficacy of purifying selection against deleterious amino acid mutations on the neo-Y. However, neo-X and neo-Y chromosomes do not differ only in their recombination rate; other factors—such as differences in patterns of mutation or selection—could also contribute to differences in patterns of protein evolution observed between neo-X and neo-Y chromosomes.

**Reduced selective constraint on the neo-Y?** Not all of the amino acid mutations that have accumulated on the neo-Y chromosome are expected to be deleterious. There are at least two reasons why genes on the neo-Y chromosome might be under reduced selective constraint compared to their neo-X-linked homologs. First, some of the genes surveyed on the neo-Y chromosome are clearly nonfunctional, implying that amino acid mutations can accumulate in a neutral manner among those genes. Also, some genes might be dosage compensated on the neo-X, and the homologs on the neo-Y could evolve under reduced selective constraint. Measuring levels of constraint ($C$) allows us to crudely quantify the difference between the neo-X and the neo-Y in patterns of amino acid evolution. Level of constraint—following the method of EYRE-WALKER and KEIGHTLEY (1999)—was estimated to be 0.92 on the neo-X chromosome, suggesting that 92% of the amino acid mutations are deleterious enough to be eliminated by purifying selection on the neo-X. As expected, the level of constraint is much lower on the neo-Y chromosome; only 68% of all amino acid changing mutations are purged from the neo-Y. Thus, a large fraction of replacement mutations (*i.e.*, ≈24%) behave effectively neutrally on the neo-Y. However, if only genes that show intact open reading frames on the neo-Y chromosome are analyzed, the level of constraint estimated for the neo-Y-linked

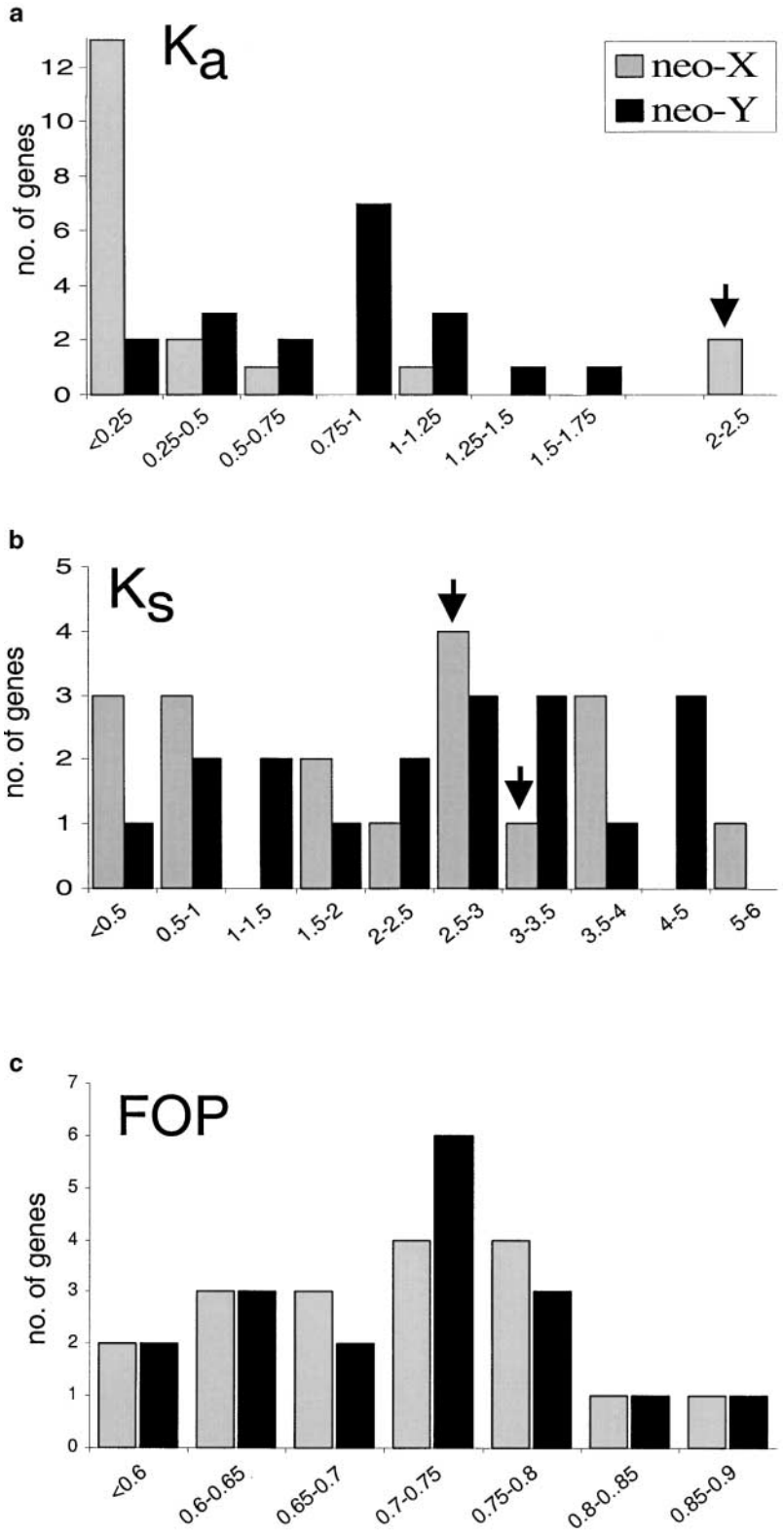


FIGURE 2.—Maximum-likelihood estimates of the rate of amino acid substitutions (a) and synonymous substitutions (b) on the neo-X and the neo-Y chromosome, as well as the frequency of optimal codons (c). The arrows indicate the positions of the *CycB* and *exu1* genes.

genes is similarly low (*i.e.*, 71%). Further, considering only those genes for which expression of the neo-Y-linked copy has been demonstrated (*i.e.*, *eve*, *robo*, and *Lcp3*), the level of constraint estimated is 76%. Thus, there is no indication that the high rate of amino acid replacements of the neo-Y-linked genes is solely the consequence of some genes containing frameshift mutations. The issue of reduced functional constraint due to dosage compensation on the neo-X is more difficult to resolve. Cytogenetic analysis has shown that parts of the neo-X chromosome in *D. miranda* may be dosage compensated (MARIN *et al.* 2000). The loci analyzed

**TABLE 4**

**Patterns of preferred (U → P) and unpreferred (P → U) substitutions on the neo-sex chromosomes of *D. miranda***

| | Neo-X[a] | | | Neo-Y[a] | | |
|---|---|---|---|---|---|---|
| | Others[b] | P → U | U → P | Others[b] | P → U | U → P |
| *CG11136* | 1 | 11 | 1 | 3 | 8 | 3 |
| *CG11159* | 0 | 0 | 0 | 1 | 1 | 0 |
| *CG16799* | 0 | 3 | 0 | 2 | 0 | 0 |
| *CycB* | 0 | 9 | 0 | 0 | 4 | 2 |
| *dpn* | 1 | 2 | 6 | 2 | 7 | 2 |
| *eng* | 0 | 1 | 1 | 3 | 6 | 0 |
| *eve* | 1 | 1 | 0 | 1 | 7 | 0 |
| *exu1* | 0 | 7 | 1 | 0 | 3 | 1 |
| *Lcp1* | 0 | 0 | 0 | 0 | 1 | 0 |
| *Lcp3* | 0 | 4 | 0 | 0 | 2 | 0 |
| *robo* | 2 | 12 | 3 | 5 | 15 | 7 |
| *zip* | 0 | 2 | 1 | 1 | 1 | 0 |
| Functional | 5 | 52 | 13 | 18 | 55 | 15 |
| *CG13437* | 1 | 0 | 0 | 1 | 0 | 0 |
| *CG30152* | 0 | 1 | 0 | 0 | 2 | 1 |
| *CG9025* | 1 | 2 | 0 | 0 | 1 | 0 |
| *Lcp2* | 1 | 2 | 0 | 0 | 1 | 2 |
| *Lcp4* | 0 | 1 | 0 | 0 | 0 | 0 |
| *mle* | 1 | 0 | 0 | 1 | 3 | 0 |
| Nonfunctional | 4 | 6 | 0 | 2 | 7 | 3 |
| Total | 9 | 58 | 13 | 20 | 62 | 18 |

[a] The *D. pseudoobscura* sequence was used to polarize changes.
[b] Others refers to changes within a synonymous class (*i.e.*, P → P or U → U).

here are spread over various locations on the neo-X chromosome; however, it is unclear which neo-Y genes might be dosage compensated on the neo-X. Thus, some fraction of the neo-Y-linked genes might evolve under reduced functional constraint due to dosage compensation on the neo-X chromosome.

**Differences in mutation bias and mutation rate?** A departure from stationarity in nucleotide composition between the neo-X and the neo-Y chromosome could also contribute to the differences observed in rates of nucleotide substitutions. To explore this possibility, intron-derived DNA was investigated. A total of ~6.4 kb of intron DNA could be aligned unambiguously between *D. pseudoobscura* and the *D. miranda* neo-X and neo-Y. AT content is identical on the neo-sex chromosomes (54.8%) and only slightly lower in *D. pseudoobscura* (54.4%). Moreover, the proportion of GC → AT *vs.* AT → GC substitutions is almost identical on the neo-X and neo-Y chromosome (2.1 *vs.* 1.9), and no significant heterogeneity in the two classes of substitutions between the neo-sex chromosomes is observed ($P$ from Fisher's exact test >0.8 for 29:14 *vs.* 49:26; Table 5). Also, synonymous substitutions show more GC → AT *vs.* AT → GC substitutions on the neo-X and neo-Y chromosome, but no significant heterogeneity between chromosomes (59:13 *vs.* 66:18; $P > 0.4$; Table 6). Thus, while there is an excess of GC → AT mutations on both neo-sex chromosomes (see below), differences in nucleotide composition between the chromosomes cannot account for the differences in rates of amino acid evolution observed. Selection on amino acid substitutions could have opposed the GC → AT mutation bias; indeed, the pattern of GC → AT *vs.* AT → GC substitutions at replacement sites differs between the neo-X and neo-Y chromosome (25:15 *vs.* 59:14; $P = 0.03$).

The number of silent mutations (*i.e.*, synonymous and noncoding) is significantly higher on the neo-Y chromosome than on the neo-X. A total of 196 silent substitutions were observed on the neo-Y, but only 149 on the neo-X (Tajima's test, $P < 0.05$). If synonymous sites and introns experience no selective constraint, this might indicate a higher mutation rate on the neo-Y chromosome than on the neo-X. Alternatively, introns and synonymous sites might be under weak selective constraints; if mutations are slightly deleterious, this also can result in a faster rate of molecular evolution on the neo-Y. If there is indeed a higher mutation rate on the neo-Y, this could partly contribute to the faster

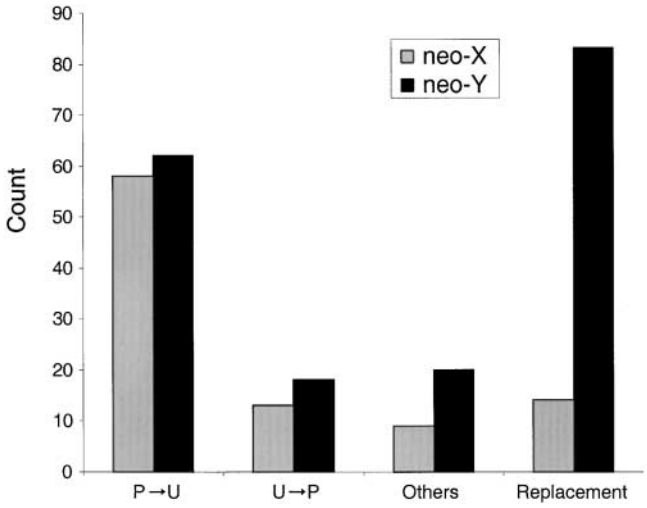


FIGURE 3.—Comparisons of patterns of codon usage bias and amino acid replacements at the neo-X and neo-Y chromosome. The number of unpreferred substitutions (P → U), the number of preferred substitutions (U → P), the number of other substitutions (*i.e.*, P → P or U → U), and the number of replacement fixations (excluding *exu1* and *CycB*) are shown.

rate of protein evolution observed, if amino acid changes are very weakly selected. However, the excess accumulation of slightly deleterious mutations on the neo-Y due to increased mutation pressure would be only 1.3-fold. This clearly cannot account for the observed 6-fold increase in amino acid substitutions on the neo-Y.

Sampling a single allele from the neo-X and the neo-Y does not allow distinguishing between fixed and segregating mutations (AKASHI 1996). The 30-fold reduction in $N_e$ of the neo-Y chromosome compared to the neo-X implies that neo-Y-derived alleles have a much shorter time back to their most recent common ancestor. While some mutations sampled from the neo-X allele might actually be segregating in the population, nearly all mutations on the neo-Y will be fixed. Thus, if slightly deleterious amino acid mutations that eventually will be eliminated by natural selection segregate on the neo-X, we would slightly underestimate the excess of protein evolution on the neo-Y chromosome.

**Reduced codon bias in the *D. miranda* lineage:** In contrast to the striking difference between the neo-X and neo-Y in patterns of protein evolution, unpreferred synonymous substitutions do not accumulate differentially on the neo-sex chromosomes. Codon usage bias is thought to be maintained by a balance of mutation, selection, and drift (SHARP and LI 1986; BULMER 1988). The strength of selection acting on synonymous codons is probably much weaker on average than that for amino acid replacements; for *D. simulans*, the selection coefficient against unpreferred codons was estimated to be on the order of $10^{-6}$ (AKASHI 1995). Assuming a statistical equilibrium (*i.e.*, if selection intensity, mutation rates, and population sizes remain constant), the number of preferred substitutions should be the same as that of unpreferred ones in a given lineage (AKASHI 1995). However, both the neo-X and the neo-Y in *D. miranda* show a significant excess of unpreferred over preferred

**TABLE 5**

**Numbers of substitutions in intron-derived DNA on the neo-X and neo-Y chromosomes**

| | | Neo-X | | | Neo-Y | | |
|---|---|---|---|---|---|---|---|
| | bp | AT → GC | GC → AT | Others[a] | AT → GC | GC → AT | Others[a] |
| *CG11136* | 261 | 1 | 3 | 1 | 3 | 1 | 2 |
| *CG11159* | 176 | 0 | 0 | 1 | 0 | 0 | 0 |
| *CG16799* | 121 | 0 | 0 | 0 | 1 | 2 | 0 |
| *CycB* | 506 | 0 | 2 | 2 | 0 | 5 | 1 |
| *dpn* | 1390 | 7 | 7 | 7 | 10 | 12 | 6 |
| *eng* | 317 | 0 | 0 | 0 | 0 | 0 | 1 |
| *eve* | 57 | 0 | 0 | 0 | 0 | 1 | 1 |
| *exu1* | 141 | 1 | 0 | 0 | 0 | 1 | 0 |
| *Lcp 1* | 69 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Lcp 3* | 61 | 0 | 0 | 0 | 0 | 0 | 0 |
| *robo* | 1064 | 0 | 4 | 5 | 4 | 14 | 2 |
| *zip* | — | — | — | — | — | — | — |
| *CG13437* | 68 | 0 | 1 | 1 | 2 | 0 | 0 |
| *CG30152* | 191 | 0 | 1 | 1 | 0 | 2 | 0 |
| *CG9025* | 1269 | 2 | 4 | 2 | 5 | 10 | 4 |
| *Lcp 2* | 72 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Lcp 4* | 63 | 0 | 0 | 0 | 0 | 0 | 0 |
| *mle* | 529 | 3 | 6 | 5 | 1 | 1 | 4 |
| Total | 6355 | 14 | 29 | 26 | 26 | 49 | 21 |

[a] Others refers to changes between G ↔ C and A ↔ T.

**TABLE 6**

**Total numbers of substitutions in all genes studied on the neo-X and neo-Y branch**

| | Neo-X | | | Neo-Y | | |
|---|---|---|---|---|---|---|
| | AT → GC | GC → AT | Others | AT → GC | GC → AT | Others |
| Noncoding | 14 | 29 | 26 | 26 | 49 | 21 |
| Synonymous | 13 | 59 | 8 | 18 | 66 | 16 |
| Replacement | 15 | 25 | 15 | 14 | 59 | 25 |

substitutions, suggesting an overall reduction in the levels of codon bias in the species. Since the intensity of selection maintaining codon bias has been estimated to be low in *D. pseudoobscura* ($N_e s \approx 1$; Akashi and Schaeffer 1997) it is likely to be even lower in *D. miranda* given the approximately sixfold lower effective population size (based on estimates of silent-site diversity; Yi *et al.* 2003). Thus, a reduction in the effective population size of *D. miranda* compared to an ancestor could have caused synonymous codons to behave as effectively neutral on both the neo-X and the neo-Y chromosome. Since the rate of substitution for neutral mutations does not depend on the population size but only on their mutation rate (Kimura 1983), unpreferred codons should accumulate similarly on both neo-sex chromosomes.

Mutation bias can also contribute to the decrease in codon bias. All preferred codons in *D. pseudoobscura* end with a G or C nucleotide (Akashi and Schaeffer 1997). A change in mutational bias toward AT mutations can be analyzed by studying substitutions in intron DNA (assuming introns are not under selective constraint). This analysis indicates that base composition of introns between *D. pseudoobscura* and *D. miranda* is similar. However, the number of GC → AT substitutions on both the neo-X and neo-Y is almost twice as large as the number of AT → GC substitutions (29 *vs.* 14 on the neo-X and 49 *vs.* 26 on the neo-Y; $P < 0.05$ for both cases; Table 5). This large excess of GC → AT substitutions indicates that nucleotide composition is not at equilibrium in *D. miranda*, but is evolving toward a higher AT content. Thus, a change in mutational bias is likely to have contributed to the decrease of codon bias in *D. miranda*.

A similar pattern for synonymous substitutions was also observed in the *D. melanogaster* and *D. simulans* lineages (Akashi 1996; McVean and Vieira 2001; Begun 2002). Both *D. simulans* and *D. melanogaster* show a genome-wide reduction in levels of codon bias compared to their ancestor (Akashi 1996; McVean and Vieira 2001). *D. melanogaster* shows no evidence of current selection on codon usage (Akashi 1996; McVean and Vieira 2001), while this relaxation in selection appears to be less extreme for *D. simulans* (McVean and Vieira 2001). This pattern is consistent with a somewhat lower $N_e$ for *D. melanogaster* than for *D. simulans*, as suggested by differences in overall levels of nucleotide variability (Andolfatto 2001). However, *D. melanogaster* also has an apparent mutation bias toward AT substitutions (Takano-Shimizu 2001), confounding the effects of selection and mutation on the evolution of codon bias.

**Reduced adaptation in regions of reduced recombination in Drosophila:** The higher rate of amino acid substitutions on the nonrecombining neo-Y chromosome and the larger variance in $K_a$ on the neo-X relate to the recent finding that the rate of amino acid substitutions and the variance in $K_a$ is elevated in regions of high recombination in *D. melanogaster* (Betancourt and Presgraves 2002). The pattern observed in *D. melanogaster* suggests that adaptive protein evolution may be constrained in low-recombining regions in the *D. melanogaster* genome, but that purifying selection against deleterious amino acid substitutions is operating genome-wide (Betancourt and Presgraves 2002). In *D. miranda*, while the average $K_a$ is higher on the neo-Y chromosome, the variance in $K_a$ among loci is much larger on the neo-X (see Figure 2). Most loci on the neo-X show very few amino acid substitutions, indicating purifying selection against deleterious amino acid changes, and two genes show rapid amino acid evolution due to positive selection (Bachtrog and Charlesworth 2002; Bachtrog 2003). The data from the neo-sex chromosomes of *D. miranda* support the hypothesis that linkage limits the rate of adaptation, with adaptive protein evolution being mainly restricted to the recombining neo-X chromosome. However, the reduced efficacy of purifying selection is more obvious for the neo-Y in *D. miranda* (*i.e.*, the general increase in $K_a$) than in regions of low recombination in *D. melanogaster* (although data for the latter are sparse; see Betancourt and Presgraves 2002). Part of this difference may be explained by the size of the regions of reduced recombination. The nonrecombining neo-Y chromosome constitutes an entire chromosome arm, with an estimate of ~2500 genes (Adams *et al.* 2000). In contrast, the various regions of reduced recombination in *D. melanogaster* each contain at most ~100 genes (Adams *et al.* 2000). Hill-Robertson effects—which are proportional to the number of linked loci under selection (Li 1987; McVean and Charlesworth 2000)—are therefore expected to be much more pronounced for the neo-Y in *D. miranda*.

I have emphasized the importance of the effective

population size $N_e$ in explaining patterns of evolution on the neo-X and neo-Y chromosome, as it is helpful to explain certain features of the data, such as a reduction in polymorphism on the neo-Y chromosome or the evolutionary dynamics of very weakly selected mutations (Charlesworth and Charlesworth 2000). However, this is clearly not the only factor involved. Noteworthy is that the effective population size inferred from nucleotide diversity on the neo-Y chromosome (*i.e.*, ~30,000) is on the same order as that estimated for humans (*i.e.*, ~10,000). Indeed, in accordance with their small population size, human populations appear to be accumulating slightly deleterious mutations (Eyre-Walker *et al.* 2002). However, a key difference between the neo-Y of *D. miranda* and human populations is the absence of sexual recombination in the former. Thus, while both might suffer from the accumulation of weakly deleterious mutations due to a small $N_e$, recombination in human populations may be sufficient to prevent the fixation of strongly deleterious mutations. In contrast, the lack of recombination on the neo-Y chromosome leaves it susceptible to processes like Muller's ratchet or hitchhiking of linked detrimental alleles with strongly advantageous mutations; both processes can fix strongly deleterious mutations (Charlesworth and Charlesworth 2000). This emphasizes the role of recombination (and not just simply differences in $N_e$) in explaining patterns of molecular evolution on the neo-X and neo-Y chromosomes.

I am grateful to P. Andolfatto, B. Charlesworth, and S. Wright for helpful discussions and comments on the manuscript. D.B. is supported by a European Molecular Biology Organization Fellowship. This research was funded by a Biotechnology and Biological Sciences Research Council grant to B. Charlesworth and by a Royal Society grant to D.B.

## LITERATURE CITED

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. Science **287:** 2185–2195.

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., and S. W. Schaeffer, 1997 Natural selection and the frequency distributions of "silent" DNA polymorphism in Drosophila. Genetics **146:** 295–307.

Andolfatto, P., 2001 Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. Mol. Biol. Evol. **18:** 279–290.

Bachtrog, D., 2003 Adaptation shapes patterns of genome evolution in sexual and asexual genomes in *Drosophila*. Nat. Genet. **34:** 215–219.

Bachtrog, D., and B. Charlesworth, 2000 Reduced levels of microsatellite variability on the neo-Y chromosome of *Drosophila miranda*. Curr. Biol. **10:** 1025–1031.

Bachtrog, D., and B. Charlesworth, 2001 Towards a complete sequence of the human Y chromosome. Genome Biol. **2:** 1016.

Bachtrog, D., and B. Charlesworth, 2002 Reduced adaptation of a non-recombining neo-Y chromosome. Nature **416:** 323–326.

Begun, D. J., 2002 Protein variation in *Drosophila simulans*, and comparison of genes from centromeric versus noncentromeric regions of chromosome 3. Mol. Biol. Evol. **19:** 201–203.

Begun, D. J., and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. Nature **356:** 519–520.

Betancourt, A. J., and D. C. Presgraves, 2002 Linkage limits the power of natural selection in *Drosophila*. Proc. Natl. Acad. Sci. USA **99:** 13616–13620.

Bull, J. J., 1983 *Evolution of Sex Determining Mechanisms*. Benjamin Cummings, Menlo Park, CA.

Bulmer, M., 1988 Codon usage and intragenic position. J. Theor. Biol. **133:** 67–71.

Caballero, A., 1995 On the effective size of populations with separate sexes, with particular reference to sex-linked genes. Genetics **139:** 1007–1011.

Carvalho, A. B., B. A. Dobo, M. D. Vibranovski and A. G. Clark, 2001 Identification of five new genes on the Y chromosome of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **98:** 13225–13230.

Charlesworth, B., 1990 The evolution of sex chromosomes. Science **251:** 1030–1033.

Charlesworth, B., 1996 The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. **6:** 149–162.

Charlesworth, B., and D. Charlesworth, 2000 The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **355:** 1563–1572.

Comeron, J. M., 1995 A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. J. Mol. Evol. **41:** 1152–1159.

Eyre-Walker, A., and P. D. Keightley, 1999 High genomic deleterious mutation rates in hominids. Nature **397:** 344–347.

Eyre-Walker, A., P. D. Keightley, N. G. Smith and D. Gaffney, 2002 Quantifying the slightly deleterious mutation model of molecular evolution. Mol. Biol. Evol. **19:** 2142–2149.

Felsenstein, J., 1974 The evolutionary advantage of recombination. Genetics **78:** 737–756.

Gethmann, R. C., 1988 Crossing over in males of higher *Diptera* (Brachycera). J. Hered. **79:** 344–350.

Hill, W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet. Res. **8:** 269–294.

Hudson, R. R., M. Kreitman and M. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kliman, R. M., and J. Hey, 1993 Reduced natural selection associated with low recombination in *Drosophila melanogaster*. Mol. Biol. Evol. **10:** 1239–1258.

Lahn, B. T., N. M. Pearson and K. Jegalian, 2001 The human Y chromosome, in the light of evolution. Nat. Rev. Genet. **2:** 207–216.

Li, W. H., 1987 Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. J. Mol. Evol. **24:** 337–345.

Lucchesi, J. C., 1978 Gene dosage compensation and the evolution of sex chromosomes. Science **202:** 711–716.

Lynch, M., and J. L. Blanchard, 1998 Deleterious mutation accumulation in organelle genomes. Genetica **102/103:** 29–39.

Marais, G., and G. Piganeau, 2002 Hill-Robertson interference is a minor determinant of variations in codon bias across *Drosophila melanogaster* and *Caenorhabditis elegans* genomes. Mol. Biol. Evol. **19:** 1399–1406.

Marin, I., M. L. Siegal and B. S. Baker, 2000 The evolution of dosage-compensation mechanisms. Bioessays **22:** 1106–1114.

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

McVean, G. A., and J. Vieira, 2001 Inferring parameters of mutation, selection and demography from patterns of synonymous site evolution in Drosophila. Genetics **157:** 245–257.

McVean, G. A. T., and B. Charlesworth, 2000 The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics **155:** 929–944.

Muller, H. J., 1940 Bearings of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by J. S. Huxley. Oxford University Press, London.

Nei, M., and T. Gojobori, 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3:** 418–426.

Orr, H. A., and Y. Kim, 1998 An adaptive hypothesis for the evolution of the Y chromosome. Genetics **150:** 1693–1698.

Peck, J. R., 1994 A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. Genetics **137:** 597–606.

Sharp, P. M., and W. H. Li, 1986 An evolutionary perspective on synonymous codon usage in unicellular organisms. J. Mol. Evol. **24:** 28–38.

Steinemann, M., and S. Steinemann, 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., S. Steinemann and F. Lottspeich, 1993 How Y chromosomes become genetically inert. Proc. Natl. Acad. Sci. USA **90:** 5737–5741.

Steinemann, M., S. Steinemann and W. Pinsker, 1996 Evolution of the larval cuticle proteins coded by the secondary sex chromosome pair: X2 and neo-Y of *Drosophila miranda*. I. Comparison at the DNA sequence level. J. Mol. Evol. **43:** 405–412.

Steinemann, S., and M. Steinemann, 1999 The *Amylase* gene cluster on the evolving sex chromosomes of *Drosophila miranda*. Genetics **151:** 151–161.

Stenico, M., A. T. Lloyd and P. M. Sharp, 1994 Codon usage in *Caenorhabditis elegans*: delineation of translational selection and mutational biases. Nucleic Acids Res. **22:** 2437–2446.

Tajima, F., 1993 Simple methods for testing the molecular evolutionary clock hypothesis. Genetics **135:** 599–607.

Takano-Shimizu, T., 2001 Local changes in GC/AT substitution biases and in crossover frequencies on *Drosophila* chromosomes. Mol. Biol. Evol. **18:** 606–619.

Weinreich, D. M., and D. M. Rand, 2000 Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. Genetics **156:** 385–399.

Wright, F., 1990 The "effective number of codons" used in a gene. Gene **87:** 23–29.

Yi, S., D. Bachtrog and B. Charlesworth, 2003 A survey of chomosomal and nucleotide sequence variation in *Drosophila miranda*. Genetics **164:** 1369–1381.

Zurovcova, M., and W. F. Eanes, 1999 Lack of nucleotide polymorphism in the Y-linked sperm flagellar dynein gene Dhc-Yh3 of *Drosophila melanogaster* and *D. simulans*. Genetics **153:** 1709–1715.

Communicating editor: D. Begun