

Evolution of Amino-Acid Sequences and Codon Usage on the *Drosophila miranda* Neo-Sex Chromosomes

Carolina Bartolomé¹ and Brian Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received July 31, 2006

Accepted for publication September 14, 2006

ABSTRACT

We have studied patterns of DNA sequence variation and evolution for 22 genes located on the neo-X and neo-Y chromosomes of *Drosophila miranda*. As found previously, nucleotide site diversity is greatly reduced on the neo-Y chromosome, with a severely distorted frequency spectrum. There is also an accelerated rate of amino-acid sequence evolution on the neo-Y chromosome. Comparisons of nonsynonymous and silent variation and divergence suggest that amino-acid sequences on the neo-X chromosome are subject to purifying selection, whereas this is much weaker on the neo-Y. The same applies to synonymous variants affecting codon usage. There is also an indication of a recent relaxation of selection on synonymous mutations for genes on other chromosomes. Genes that are weakly expressed on the neo-Y chromosome appear to have a faster rate of accumulation of both nonsynonymous and unpreferred synonymous mutations than genes with high levels of expression, although the rate of accumulation when both types of mutation are pooled is higher for the neo-Y chromosome than for the neo-X chromosome even for highly expressed genes.

A promising line of attack on the problem of understanding Y chromosome degeneration is offered by systems of so-called neo-X and neo-Y chromosomes, which have evolved in many *Drosophila* species (LUCCHESI 1978; STEINEMANN and STEINEMANN 1998; McALLISTER and CHARLESWORTH 1999; YU *et al.* 1999). These occur when an autosome becomes fused or translocated to an X or Y chromosome and spreads to fixation within the species. The classic example is *Drosophila pseudoobscura* and its relatives, which possess an X-autosome fusion that is absent from the sister group from which they diverged ~13–22 million years ago (RUSSO *et al.* 1995; TAMURA *et al.* 2004).

In *Drosophila*, the lack of crossing over in males means that a neo-Y chromosome is immediately placed in a nonrecombining genetic environment, identical to that of the true Y chromosome. Gradual degeneration can then occur, because of the cumulative effects of the reduction in the efficacy of selection in a nonrecombining region of the genome (CHARLESWORTH and CHARLESWORTH 2000), so that some neo-Y-linked genes are either reduced in activity or lost, while others are still normally active. This is exemplified by *D. miranda*, which diverged from *D. pseudoobscura* ~2 million years ago, and

has a Y-autosome fusion that has been estimated to be ~1 million years old (BACHTROG and CHARLESWORTH 2002).

Previous work on this system has shown that silent-site DNA sequence variability on the neo-Y is greatly reduced relative to the neo-X (BACHTROG and CHARLESWORTH 2002) and that the rate of accumulation of nonsynonymous substitutions is much higher on the neo-Y than the neo-X (YI and CHARLESWORTH 2000a; BACHTROG 2005). This raises the question of whether this higher rate of amino-acid sequence evolution for genes on the neo-Y, relative to their neo-X counterparts, is driven by relaxation of purifying selection or by the action of positive selection (BACHTROG 2004). Another feature of the neo-Y genes is that a substantial fraction of them have accumulated major mutational lesions, such as frameshift mutations and/or transposable element insertions (STEINEMANN and STEINEMANN 1992; BACHTROG 2003a, 2005). It is thus possible that any relaxation of selection against amino-acid mutations reflects the fact that these are neutral in genes that have been knocked out by major mutations in their coding sequences or that are poorly expressed because of regulatory mutations (STEINEMANN and STEINEMANN 1992). One goal of the work described here is to test for relaxed selection on amino-acid mutations on the neo-Y chromosome and to examine whether this is associated with reduced functionality or expression of the genes concerned.

A puzzling feature of this species is that selection seems to be acting to maintain the use of preferred codons in autosomal and X-linked genes (BARTOLOMÉ

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. DQ867535–DQ868314.

¹Corresponding author: Grupo de Medicina Xenómica (USC), Unidad de Medicina Molecular, Hospital Clínico Universitario de Santiago, Edificio Consultas (planta-2), 15706 Santiago de Compostela, Spain. E-mail: cbhusson@usc.es

et al. 2005), in contrast with the apparent accumulation of unpreferred codons detected on both neo-sex chromosomes (BACHTROG 2003b, 2005). While a relaxation of selection on codon usage bias on the neo-*Y* chromosome is in accord with theoretical predictions (CHARLESWORTH and CHARLESWORTH 2000), the reduction observed for the neo-*X* loci is unexpected and may be caused by polymorphic variants having been classed as fixed mutations (BARTOLOMÉ *et al.* 2005). Another aim of this study is to examine this question further.

MATERIALS AND METHODS

Drosophila strains: Eighteen *D. miranda* lines derived from single wild-caught females were used: 0101.3, 0101.4, 0101.5, 0101.7 (Port Coquitlam, BC, Canada), 0101.9, MA28, MA32, MA03.1, MA03.2, MA03.3, MA03.4, MA03.5, MA03.6 (Mather, CA), SP138, SP235, SP295 (Spray, OR) and MSH22, MSH38 (Mt. St. Helena, CA). Twelve of these lines were used in previous studies (Yi and CHARLESWORTH 2000b; Yi *et al.* 2003) and were obtained by S. Yi from the National Drosophila Species Resource Center (Bowling Green, OH) and from M. Noor and W. W. Anderson. The remaining six lines (MA03.1, MA03.2, MA03.3, MA03.4, MA03.5, MA03.6) were collected in July 2003 by B. Charlesworth and D. Charlesworth at Mather, California. The identity of these new strains was established by D. Bachtrog using species-specific primers (personal communication). A strain of *D. affinis* was used as an outgroup (no. 0141.2, Drosophila Species Resource Center). Stocks were maintained in bottle culture on banana medium at 18°.

Details of the genes and regions analyzed in this study are given in supplemental Table S1 at <http://www.genetics.org/supplemental/>.

Design of PCR primers and DNA extraction: Primers were designed for regions conserved between *D. pseudoobscura* and *D. melanogaster*, obtained from the publicly available genome sequences, and were used to amplify the genes from *D. miranda* and *D. affinis*. The identification of orthologous loci was performed by means of a BLAST search (<http://www.hgsc.bcm.tmc.edu/blast/blast.cgi?organism=Dpseudoobscura>). DNA was extracted separately from single male and female flies using Puregene (Gentra Systems, Minneapolis). We used standard PCR procedures (Expand High Fidelity PCR system, Roche Diagnostics, Lewes, UK) and gel purified the resulting products with Qiaquick (QIAGEN, Crawley, UK).

Cloning, sequencing, and design of allele-specific primers: Purified PCR products were cloned with TOPO-TA (Invitrogen, San Diego). The DNA sequencing of both strands was performed on an ABI3730 sequencer using Dyeamic (Amersham Biosciences UK, Little Chalfont, UK). Several plasmids for each cloning reaction were studied to exclude PCR errors and to obtain the sequences of both neo-sex copies for the initial characterization of such sequences. The products isolated from single female flies were assigned to the neo-*X* lineage, while the products isolated from males were identified as either neo-*X* or neo-*Y* sequences. The recognition of both copies enabled us to design allele-specific primers that, after optimizing the PCR conditions, were used to generate polymorphism data by direct sequencing of genomic DNA from males. Details of the primers and PCR conditions are available on request.

Sequence analyses: All readouts were first checked for accurate base calling using Sequencher (Gene Codes, Ann Arbor, MI), assembled, and then manually aligned with Se-

quence Alignment Editor (Se-AL, A. Rambaut, available at <http://evolve.zoo.ox.ac.uk/>). Exon and intron boundaries were determined from the genome of *D. melanogaster*. Introns and noncoding flanking sequences were aligned with MCALIGN (KEIGHTLEY and JOHNSON 2004). Most population genetic analyses were done with DnaSP v. 4.0 (ROZAS *et al.* 2003), except for the HKA test. This was performed with two different programs: the standard multilocus HKA test, using the program of J. Hey (<http://lifesci.rutgers.edu/hey/hey/DistributedProgramsandData.htm#HKA>), and a maximum-likelihood version of this test (WRIGHT and CHARLESWORTH 2004), available at <http://www.yorku.ca/stephenw>.

Tests of significance of Tajima's *D* for the combined set of loci on the neo-*X* chromosome were done using Hey's HKA program, assuming free recombination. For the neo-*Y* chromosome, all the loci were concatenated into one sequence, since there is no recombination. Significance for combined data was calculated with coalescent simulations (DnaSP program, ROZAS *et al.* 2003), assuming no recombination.

F_{op} , the fraction of optimal codons among all codons in a gene (IKEMURA 1981; DURET and MOUCHIROUD 1999), was obtained with a C program, kindly provided by L. Duret, using the table of optimal codons for *D. pseudoobscura* (AKASHI and SCHAEFFER 1997).

RT-PCR experiments: Total RNA from 20 *D. miranda* males was extracted with an RNeasy mini kit (QIAGEN) following the manufacturer's instructions. The product was treated with a DNase I Kit (Sigma-Aldrich, Gillingham, UK) before performing qualitative RT-PCR with the QIAGEN OneStep RT-PCR kit to avoid the amplification of any contaminant DNA. Allele-specific primers were used to compare the expression of both neo-sex copies and, when possible, the products were cloned and sequenced to check the identity of the sequences. Differences in amplification efficiency were checked by amplifying genomic DNA with the same specific primers; positive and negative (without template RNA) controls were included in each reaction. Several RNA extractions were used to ensure that the results were reproducible and independent of the original concentration of the template (see supplemental Figure S1 at <http://www.genetics.org/supplemental/>).

RESULTS

DNA sequence variation: Our tests of hypotheses concerning the evolutionary degeneration of the neo-*Y* chromosome of *D. miranda* are based on comparisons of DNA sequence diversity and divergence between the neo-*X* and neo-*Y* chromosomes. We therefore begin by summarizing the relevant results. We focus on polymorphism statistics that indicate selection against non-synonymous mutations on the neo-*X* chromosome, since the occurrence of such selection is critical for interpreting the differences between the neo-*X* and neo-*Y* chromosomes.

We sequenced >25 kb from the neo-*X* chromosome and ~23 kb from the neo-*Y*, including both coding and noncoding regions (Table 1). The unweighted mean silent nucleotide diversity estimate from pairwise differences among alleles (π_{silent} , NEI 1987), is $0.39 \pm 0.010\%$ for neo-*X* loci (average \pm SE), similar to estimates for autosomal and X-linked genes in this species (BARTOLOMÉ *et al.* 2005). These conclusions also apply to the Watterson θ -values (WATTERSON 1975) (data not shown).

TABLE 1
Nucleotide diversity on *D. miranda* neo-sex chromosomes

| Gene | No. of sites | | | | π^a | | | | θ^b | | | | |
|--------------------|--------------|--------|---------|-------|---------|-------|--------|---------|------------|-------|--------|---------|-------|
| | Syn | Nonsyn | Non-cod | Sil | Total | Syn | Nonsyn | Non-cod | Sil | Syn | Nonsyn | Non-cod | Sil |
| Polymorphism neo-X | | | | | | | | | | | | | |
| <i>Alk</i> | 164 | 553 | 374 | 538 | 1,091 | 0 | 0 | 0.09 | 0.06 | 0 | 0 | 0.16 | 0.11 |
| <i>Asx</i> | 306 | 939 | 139 | 445 | 1,384 | 0.34 | 0.03 | 0.23 | 0.31 | 0.38 | 0.06 | 0.42 | 0.39 |
| <i>CG11136</i> | 357 | 1,062 | 71 | 428 | 1,490 | 0.82 | 0 | 0.71 | 0.81 | 0.82 | 0 | 0.41 | 0.75 |
| <i>CG11159</i> | 72 | 240 | 197 | 269 | 509 | 0.16 | 0 | 0.65 | 0.52 | 0.41 | 0 | 0.74 | 0.65 |
| <i>CG13437</i> | 86 | 277 | 137 | 223 | 500 | 0.47 | 0 | 0 | 0.18 | 0.68 | 0 | 0 | 0.26 |
| <i>CG16799</i> | 103 | 354 | 117 | 220 | 573 | 1.30 | 0.13 | 1.01 | 1.15 | 1.99 | 0.08 | 1.49 | 1.72 |
| <i>CG30152</i> | 128 | 415 | 201 | 329 | 744 | 0 | 0 | 0.16 | 0.10 | 0 | 0 | 0.29 | 0.18 |
| <i>cnk</i> | 233 | 703 | 0 | 233 | 936 | 0.69 | 0.05 | — | 0.69 | 1.00 | 0.12 | — | 1.00 |
| <i>cos</i> | 220 | 674 | 126 | 346 | 1,020 | 0.17 | 0.02 | 0.09 | 0.14 | 0.13 | 0.04 | 0.23 | 0.17 |
| <i>en</i> | 261 | 771 | 0 | 261 | 1,032 | 0.09 | 0.03 | — | 0.09 | 0.23 | 0.04 | — | 0.23 |
| <i>exu1</i> | 315 | 1,029 | 78 | 393 | 1,422 | 0.05 | 0 | 0.21 | 0.09 | 0.11 | 0 | 0.43 | 0.17 |
| <i>grau</i> | 245 | 919 | 132 | 377 | 1,296 | 0.26 | 0.02 | 0.08 | 0.20 | 0.24 | 0.03 | 0.22 | 0.23 |
| <i>Lcp1</i> | 96 | 312 | 350 | 446 | 758 | 0.49 | 0.11 | 0 | 0.11 | 0.69 | 0.21 | 0 | 0.15 |
| <i>nompA</i> | 367 | 968 | 331 | 698 | 1,666 | 0.32 | 0.07 | 0.13 | 0.23 | 0.40 | 0.18 | 0.35 | 0.38 |
| <i>Pcl</i> | 233 | 838 | 73 | 306 | 1,144 | 0.14 | 0 | 0.44 | 0.21 | 0.25 | 0 | 0.80 | 0.38 |
| <i>sax</i> | 327 | 1,041 | 188 | 515 | 1,556 | 0.33 | 0 | 0 | 0.21 | 0.44 | 0 | 0 | 0.28 |
| <i>stan</i> | 418 | 1,229 | 0 | 418 | 1,647 | 1.89 | 0.05 | — | 1.89 | 1.95 | 0.14 | — | 1.95 |
| <i>Toll-7</i> | 398 | 1,207 | 0 | 398 | 1,605 | 1.20 | 0.04 | — | 1.20 | 1.02 | 0.07 | — | 1.02 |
| <i>tud</i> | 450 | 1,422 | 0 | 450 | 1,872 | 0.11 | 0.02 | — | 0.11 | 0.13 | 0.04 | — | 0.13 |
| <i>Uba1</i> | 232 | 758 | 79 | 311 | 1,069 | 0.24 | 0.03 | 0 | 0.18 | 0.50 | 0.08 | 0 | 0.37 |
| <i>Updo</i> | 189 | 573 | 54 | 243 | 816 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 300 | 937 | 133 | 433 | 1,369 | 0.17 | 0 | 0 | 0.12 | 0.10 | 0 | 0 | 0.07 |
| Total | 5,500 | 17,219 | 2,780 | 8,280 | 25,499 | 0.42 | 0.03 | 0.22 | 0.39 | 0.52 | 0.05 | 0.32 | 0.48 |
| Polymorphism neo-Y | | | | | | | | | | | | | |
| <i>Alk</i> | 165 | 553 | 317 | 482 | 1,034 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Asx</i> | 300 | 918 | 139 | 439 | 1,357 | 0 | 0.012 | 0 | 0 | 0 | 0.032 | 0 | 0 |
| <i>CG11136</i> | 357 | 1,062 | 71 | 428 | 1,490 | 0.059 | 0 | 0 | 0.049 | 0.081 | 0 | 0 | 0.068 |
| <i>CG11159</i> | 72 | 243 | 195 | 267 | 510 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG13437</i> | 40 | 119 | 68 | 108 | 227 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG30152</i> | 109 | 362 | 192 | 301 | 663 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cnk</i> | 217 | 671 | 0 | 217 | 888 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cos</i> | 219 | 672 | 126 | 345 | 1,017 | 0 | 0.017 | 0 | 0 | 0 | 0.043 | 0 | 0 |
| <i>en</i> | 332 | 979 | 0 | 332 | 1,311 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>grau</i> | 243 | 918 | 132 | 375 | 1,293 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Lcp1</i> | 95 | 313 | 350 | 445 | 758 | 0 | 0.053 | 0 | 0 | 0 | 0.106 | 0 | 0 |
| <i>nompA</i> | 381 | 997 | 331 | 712 | 1,708 | 0 | 0.011 | 0 | 0 | 0 | 0.029 | 0 | 0 |
| <i>Pcl</i> | 219 | 792 | 74 | 293 | 1,085 | 0.051 | 0 | 0 | 0.038 | 0.133 | 0 | 0 | 0.099 |
| <i>sax</i> | 293 | 916 | 67 | 360 | 1,276 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>stan</i> | 416 | 1,231 | 0 | 416 | 1,647 | 0 | 0.009 | 0 | 0 | 0 | 0.024 | 0 | 0 |
| <i>Toll-7</i> | 414 | 1,258 | 0 | 414 | 1,671 | 0 | 0.009 | 0 | 0 | 0 | 0.023 | 0 | 0 |
| <i>tud</i> | 451 | 1,421 | 0 | 451 | 1,872 | 0 | 0.015 | 0 | 0 | 0 | 0.020 | 0 | 0 |
| <i>Uba1</i> | 232 | 759 | 79 | 311 | 1,069 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Updo</i> | 190 | 572 | 54 | 244 | 816 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 299 | 940 | 133 | 432 | 1,372 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 5,044 | 15,692 | 2,328 | 7,371 | 23,064 | 0.006 | 0.006 | 0 | 0.004 | 0.011 | 0.014 | 0 | 0.008 |

Syn, synonymous; Nonsyn, nonsynonymous; Sil, silent; Non-cod, noncoding.

^aPairwise nucleotide diversity (NEI 1987).

^bNucleotide site variability is based on the number of segregating sites (WATTERSON 1975).

The neo-X chromosome exhibits high heterogeneity in levels of nucleotide variation and harbors genes that show very little diversity (*Alk*, *CG30152*, *en*, *exu1*, *Updo*) along with highly polymorphic loci, such as *CG16799*, *stan*, or *Toll-7*. We performed the HKA test (HUDSON

et al. 1987) for heterogeneity between loci by applying a combination of the standard multilocus HKA test and its maximum-likelihood version (see MATERIALS AND METHODS). We analyzed variability on the neo-X chromosome and divergence from *D. affinis* at silent sites.

After correction for multiple tests, only *stan* was found to deviate significantly from neutrality ($P = 0.001$); it is more polymorphic than expected from its divergence level. When we exclude this locus from the data set, the deviations are nonsignificant.

In contrast, all neo-*Y* loci show a large reduction in variability, compared with other *D. miranda* chromosomes, at both nonsynonymous and silent sites (mean π -values of 0.006 and 0.004%, respectively). The similarity of these values suggests an absence of strong selection on replacement mutations, in contrast to the low nonsynonymous relative to synonymous diversity on the neo-*X* chromosome [see the McDonald–Kreitman (MK) tests below].

Pairwise differences between sequences per silent site, π_{silent} , are mostly lower than the silent-site diversity estimate based on the number of segregating sites, θ_{silent} , resulting in predominantly negative Tajima's *D* statistics (TAJIMA 1989) with more rare variants than expected at neutral equilibrium (supplemental Table S2 at <http://www.genetics.org/supplemental/>). The values for the pooled neo-*X* and neo-*Y* data sets (-0.88 and -1.85 , respectively) indicate a significant excess ($P < 0.05$) of low-frequency mutations on both neo-sex chromosomes. The *exu1* gene appears to have experienced a selective sweep on the neo-*X* chromosome (BACHTROG 2003a). After removing the nine loci in region 28A–E, where this gene is located (BARTOLOMÉ and CHARLESWORTH 2006), the combined Tajima's *D* for the remaining loci is -1.25 for noncoding sites ($P < 0.0023$), but only -0.40 for synonymous sites (nonsignificant) (supplemental Table S2 at <http://www.genetics.org/supplemental/>). The values for nonsynonymous sites are significantly negative for both neo-*X* data sets (including all loci and with the nine loci removed; supplemental Table S2 at <http://www.genetics.org/supplemental/>), consistent with the action of purifying selection on replacement polymorphisms, as previously reported for autosomal and X-linked genes (BARTOLOMÉ *et al.* 2005). In addition, the comparison of the frequency distribution of synonymous and nonsynonymous mutations (SAWYER *et al.* 1987; AKASHI 1999) showed a very significant skew toward low-frequency variants for replacement changes ($P < 0.001$ on a Mann–Whitney test).

We also performed Fu and Li's *D* test (FU and LI 1993) on silent and nonsynonymous sites (supplemental Table S3 at <http://www.genetics.org/supplemental/>), again finding that these two fractions of the neo-*X* chromosome exhibit an excess of rare variants, in contrast with what was found for autosomal and X-linked loci (YI *et al.* 2003; BARTOLOMÉ *et al.* 2005). The same patterns are observed when we remove the potentially swept loci (supplemental Table S3 at <http://www.genetics.org/supplemental/>).

Silent and nonsynonymous divergence: Divergences between *D. miranda* and its relatives, using the Jukes–

Cantor correction for multiple hits (JUKES and CANTOR 1969), are shown in supplemental Table S4 at <http://www.genetics.org/supplemental/>. The mean silent divergence (K_{silent}) between the neo-*X* and neo-*Y* lineages is $3.08 \pm 0.31\%$, in line with previous estimates (YI and CHARLESWORTH 2000a; BACHTROG and CHARLESWORTH 2002). The K_{silent} length-weighted value between the neo-sex chromosomes is slightly lower (2.94%). Given the high levels of silent polymorphism shown by *D. pseudoobscura* (LOEWE *et al.* 2006; supplemental Table S1B at <http://www.genetics.org/supplemental/>), and the close relationship between this species and *D. miranda*, we applied a correction for within-species genetic variation to obtain a more accurate estimate of its divergence from *D. miranda*. This was done by subtracting the mean of the silent pairwise nucleotide diversity in both species from the average length-weighted K_{silent} between *D. miranda* and *D. pseudoobscura*.

This gave slightly lower values than the uncorrected estimates: 2.80 and 3.85% for the neo-*X* and neo-*Y* chromosomes, respectively. Applying the same correction to the silent divergence between neo-*X* and neo-*Y* chromosomes, we got a similar value (2.74%). This needs, however, to be further corrected for the contribution of ancestral polymorphisms to fixations on the neo-*Y*, which must have gone through a complete bottleneck when it was formed. We show below that these probably contributed $\sim 27\%$ of the fixations of silent mutations on the neo-*Y* since its divergence from the neo-*X* chromosome. The observed divergence between the chromosomes should therefore be reduced by a factor of $[1 - (0.5 \times 0.27)] = 0.865$, giving a final K_{silent} of 2.37%. This implies that the time since the split between the neo-sex chromosomes is $\sim 85\%$ of the time since the split of the two species. This is much greater than the estimate of $\sim 50\%$ proposed earlier (BACHTROG and CHARLESWORTH 2002).

The ratio of nonsynonymous-to-synonymous divergence (K_a/K_s) is < 1 in all cases, reflecting the existence of selective constraints on protein sequences. Comparisons between the neo-sex lineages and *D. pseudoobscura* show, on average, a lower K_a/K_s for the neo-*X* than for the neo-*Y* chromosome (0.12 ± 0.04 vs. 0.21 ± 0.04 , respectively), indicating a much faster rate of evolution of the latter, as seen previously (YI and CHARLESWORTH 2000a; BACHTROG 2005).

MK tests for the neo-*X* chromosome: We used this test to see if the patterns of polymorphism and divergence from *D. affinis* in coding regions were compatible with the equilibrium neutral model, which predicts that the ratio of replacement to silent variation should be equal to the ratio of replacement to silent divergence (MCDONALD and KREITMAN 1991). The reason for using *D. affinis* as an outgroup is the high level of ancestral polymorphism detected for *D. pseudoobscura* and *D. miranda* (CHARLESWORTH *et al.* 2005). Since polymorphism data are lacking for *D. pseudoobscura*, the use

TABLE 2

McDonald–Kreitman tests: neo-X/*D. affinis*

| Gene | Fixed | | Polymorphic | |
|----------------|--------|-------------|-------------|-------------|
| | Silent | Replacement | Silent | Replacement |
| <i>Alk</i> | 147 | 2 | 2 | 0 |
| <i>Asx</i> | 90 | 35 | 6 | 1 |
| <i>CG11136</i> | 88 | 5 | 9 | 0 |
| <i>CG13437</i> | 41 | 3 | 1 | 0 |
| <i>CG30152</i> | 59 | 4 | 2 | 0 |
| <i>cnk</i> | 47 | 5 | 8 | 3 |
| <i>cos</i> | 76 | 10 | 2 | 1 |
| <i>grau</i> | 99 | 15 | 3 | 1 |
| <i>Lcp1</i> | 50 | 1 | 2 | 2 |
| <i>nompA</i> | 125 | 11 | 9 | 6 |
| <i>Pcl</i> | 79 | 27 | 4 | 0 |
| <i>sax</i> | 151 | 4 | 5 | 0 |
| <i>stan</i> | 119 | 4 | 28 | 6 |
| <i>Toll-7</i> | 83 | 3 | 14 | 3 |
| <i>tud</i> | 133 | 94 | 2 | 2 |
| <i>Uba1</i> | 69 | 10 | 4 | 2 |
| <i>vlc</i> | 96 | 43 | 1 | 0 |
| Pooled | 1549 | 276 | 102 | 27 |

Silent and replacement are the number of silent and non-synonymous changes, respectively.

of *D. affinis*, a relatively distant species, should reduce the problem of misclassifying as fixed differences the polymorphic variants that were inherited by *D. miranda* and *D. pseudoobscura* from their common ancestor (CHARLESWORTH *et al.* 2005).

However, the use of such a distant relative increases the risk of overlooking double fixation events; underestimation of the amount of silent divergence biases our tests against detecting purifying selection. To reduce this bias, in the comparisons with *D. affinis* we multiplied the observed numbers of silent fixed differences between *D. affinis* and the neo-sex chromosomes by the ratio of Jukes–Cantor corrected to uncorrected K_{silent} estimates. Given the low level of nucleotide site diversity within *D. miranda*, such a correction is unnecessary for the polymorphism data.

Using the neo-X *vs.* *D. affinis* data, 17 genes were suitable for MK tests; *i.e.*, they had nonsynonymous and synonymous mutations in either the fixed or the polymorphic classes and at least one polymorphic site (Table 2). The highly significant excess of nonsynonymous to silent polymorphisms relative to fixations between the neo-X and *D. affinis* (mean log-odds ratio of 1.48; $z = 5.65$ on a Mantel–Haenszel test with resampling, $P < 0.001$) suggests that there has been effective purifying selection against nonsynonymous mutations on the neo-X lineage since the split with *affinis*. This is consistent with what we found previously for X-linked and autosomal genes (BARTOLOMÉ *et al.* 2005) and with the predominantly negative Tajima’s *D* values for nonsyn-

TABLE 3

McDonald–Kreitman tests: polymorphisms and fixations on the neo-X since the split with the neo-Y

| Gene | Fixed | | Polymorphic | |
|---------------|--------|-------------|-------------|-------------|
| | Silent | Replacement | Silent | Replacement |
| <i>Asx</i> | 5 | 1 | 6 | 1 |
| <i>cnk</i> | 2 | 1 | 8 | 3 |
| <i>cos</i> | 2 | 1 | 2 | 1 |
| <i>grau</i> | 3 | 5 | 3 | 1 |
| <i>Lcp1</i> | 1 | 0 | 2 | 2 |
| <i>nompA</i> | 5 | 0 | 9 | 6 |
| <i>Pcl</i> | 2 | 4 | 4 | 0 |
| <i>stan</i> | 2 | 1 | 28 | 6 |
| <i>Toll-7</i> | 2 | 0 | 14 | 3 |
| <i>tud</i> | 2 | 9 | 2 | 2 |
| <i>Uba1</i> | 6 | 0 | 4 | 2 |
| <i>vlc</i> | 4 | 6 | 1 | 0 |
| Pooled | 36 | 28 | 83 | 27 |

onymous sites (supplemental Table S2 at <http://www.genetics.org/supplemental/>).

If we use *D. affinis* to assign mutations specifically to the neo-X lineage since its split from the neo-Y (Table 3), we find a nominal but nonsignificant excess of nonsynonymous fixations on the neo-X (mean log-odds ratio of -0.43 ; not significant on a Mantel–Haenszel test). This suggests that there may have been either some positive selection on the neo-X or relaxation of purifying selection since divergence from the neo-Y.

MK tests for the neo-Y chromosome: Seven genes were suitable for comparing polymorphisms *vs.* fixations on the neo-Y chromosome since its split with the neo-X, (Table 4). There is no evidence for a significant excess of nonsynonymous fixations relative to polymorphisms, consistent with the hypothesis that the excess of nonsynonymous fixations on the neo-Y relative to the neo-X is due to relaxed purifying selection.

TABLE 4

McDonald–Kreitman tests: polymorphisms and fixations on the neo-Y since the split with the neo-X

| Gene | Fixed | | Polymorphic | |
|----------------|--------|-------------|-------------|-------------|
| | Silent | Replacement | Silent | Replacement |
| <i>Asx</i> | 4 | 5 | 0 | 1 |
| <i>CG11136</i> | 7 | 6 | 1 | 0 |
| <i>cos</i> | 3 | 6 | 0 | 1 |
| <i>Lcp1</i> | 1 | 4 | 0 | 1 |
| <i>nompA</i> | 7 | 14 | 0 | 1 |
| <i>Pcl</i> | 7 | 6 | 1 | 0 |
| <i>stan</i> | 5 | 5 | 0 | 1 |
| <i>Toll-7</i> | 9 | 10 | 0 | 1 |
| <i>tud</i> | 11 | 14 | 0 | 1 |
| Pooled | 54 | 70 | 2 | 7 |

A large excess of silent fixations was detected on the neo-*Y* branch compared with the neo-*X* (Tables 3 and 4). Partitioning mutations along the neo-*X* and neo-*Y* branches since the neo-*X*/neo-*Y* divergence, we have $K_{\text{silent}} = 0.0147$ and 0.0072 for the neo-*Y* and neo-*X* chromosomes, respectively. These values are significantly different on a Wilcoxon rank test ($z = -3.182$, $P < 0.001$). This difference is probably caused by fixations of ancestral polymorphisms on the neo-*Y* when it was originally formed. Under neutrality, the maximum expected K_{silent} along the neo-*Y* lineage due to ancestral polymorphism is the same as the ancestral π -value (CHARLESWORTH *et al.* 2005). The current *D. miranda* mean value of π for the neo-*X* is ~ 0.004 , somewhat less than the difference of ~ 0.007 between neo-*X* and neo-*Y*, and representing $\sim 27\%$ of the value of K_{silent} for the neo-*Y* lineage. The 95% confidence interval for the observed difference is 0.004–0.011 (from a paired *t*-test), which includes the predicted value. Therefore, the excess of neo-*Y* silent substitutions is probably caused by the fixation of ancestral polymorphisms.

When we compare the amount of neo-*X* *vs.* neo-*Y* polymorphisms (Tables 3 and 4), we detect a significant excess of nonsynonymous relative to silent variants on the neo-*Y* compared with the neo-*X* (mean log-odds ratio of 1.91; $z = 3.16$ on a Mantel–Haenszel test with resampling; $P < 0.001$). This provides strong evidence for a relaxation of selection against deleterious amino-acid mutations on the neo-*Y* lineage.

Gene expression on the neo-*X* and neo-*Y* chromosomes: It is desirable to test whether reduced functioning of neo-*Y* genes is associated with their higher rate of amino-acid sequence evolution (see the Introduction). One way to do this is to compare the amount of nonsynonymous divergence (K_a) between neo-*X* and neo-*Y* copies with the relative levels of gene expression on both neo-sex chromosomes. We therefore classified neo-*Y* loci into genes with high and low expression, assessed by qualitative RT–PCR, and compared their K_a values by a Mann–Whitney test. As a control, we used the K_a between *D. pseudoobscura* and *D. affinis* for the same set of genes (Table 5). This approach is different from that used by BACHTROG (2005), who determined the ratio of nonsynonymous-to-synonymous substitutions (K_a/K_s) in functional and nonfunctional genes. The advantage of our method is that K_a values are less noisy than K_a/K_s ratios (which are poorly estimated along the neo-*Y* branch).

We find a significantly higher K_a for the low-expression/disrupted genes ($P = 0.019$), but no such difference is detected between *D. pseudoobscura* and *D. affinis* ($P = 0.75$). If we remove all the genes with evidence for major mutations, such as frameshift mutations and/or deletions, the result is still significant ($P = 0.039$). This suggests that low-expression genes evolve faster for nonsynonymous fixations than high expression genes in this data set. No differences for K_s are found (data not

TABLE 5
Nonsynonymous divergence and gene expression

| | K_a (neo- <i>Y</i> /neo- <i>X</i>) | K_a (<i>pseudo/affinis</i>) |
|---------------------------------|---------------------------------------|---------------------------------|
| Low-expression genes | | |
| <i>Alk</i> | 0.005 | 0.005 |
| <i>Uba1</i> | 0.005 | 0.014 |
| <i>vlc^a</i> (1) | 0.007 | 0.044 |
| <i>Toll7^a</i> (17) | 0.008 | 0.002 |
| <i>CG30152^a</i> (12) | 0.008 | 0.013 |
| <i>cos</i> | 0.009 | 0.015 |
| <i>tud^a</i> (1) | 0.010 | 0.074 |
| <i>Lcp1</i> | 0.013 | 0.006 |
| <i>nompA</i> | 0.015 | 0.018 |
| High-expression genes | | |
| <i>grau</i> | 0.002 | 0.012 |
| <i>sax</i> | 0.003 | 0.003 |
| <i>Updo</i> | 0.003 | 0.012 |
| <i>cnk</i> | 0.003 | 0.007 |
| <i>Asx</i> | 0.005 | 0.040 |
| <i>CG13437</i> | 0.008 | 0.014 |
| <i>Pcl</i> | 0.008 | 0.033 |

Numbers of base pairs included in the deletions that caused the frameshift mutations are indicated in parentheses. Note that *grau* shows two different splice forms, with RT–PCR bands of similar intensities, on the neo-*Y* lineage (see supplemental Figure S1 at <http://www.genetics.org/supplemental/>).

^a Genes with major mutations (frameshift mutations).

shown), which implies that reduced neo-*Y* expression is associated with relaxed purifying selection. When we compare the number of neo-*X* and neo-*Y* nonsynonymous fixations among highly expressed genes, we find no significant differences, although there is a suggestion of more changes on the neo-*Y* branch (11 *vs.* 21 amino-acid substitutions, respectively).

Codon usage bias: We estimated codon usage bias from the frequency of optimal codons in a gene, F_{op} (supplemental Table S5 at <http://www.genetics.org/supplemental/>). The mean values for both chromosomes are similar to those found for autosomal and X-linked loci in the same species (BARTOLOMÉ *et al.* 2005). Codon usage bias (CUB), or selection for a preferred (P) codon, is thought to be maintained by a balance between mutation and selection against deleterious or unpreferred (U) changes (LI 1987; BULMER 1991). One way to detect selection on codon usage is to apply a modification of the McDonald–Kreitman test (MCDONALD and KREITMAN 1991), comparing the ratios of polymorphism to divergence (r_{pd}) between synonymous P and U codons (AKASHI 1995). In the absence of selection, r_{pd} for the two types of mutations (P \rightarrow U *vs.* U \rightarrow P) should be equal, unless there has been a recent change in mutational bias. In contrast, if there is selection against U codons, the ratio of polymorphism to divergence should be higher for P \rightarrow U than for U \rightarrow P changes.

As shown in Table 6, r_{pd} on the neo-*X* chromosome is much greater for P \rightarrow U changes than for U \rightarrow P

TABLE 6
Patterns of synonymous and nonsynonymous changes on the neo-sex chromosomes of
D. miranda* since the split with *D. pseudoobscura

| | Fixed | | | | | | Polymorphic | | | | | |
|--------------------------------|-------|-----|-----|-----|-------|-----|-------------|-----|-----|-----|-------|----|
| | P-U | U-P | P-P | U-U | Total | Ns | P-U | U-P | P-P | U-U | Total | Ns |
| Neo-X/ <i>D. pseudoobscura</i> | | | | | | | | | | | | |
| <i>Alk</i> | 3 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Asx</i> | 1 | 0 | 0 | 2 | 3 | 1 | 1 | 1 | 0 | 2 | 4 | 1 |
| <i>CG11136</i> | 4 | 0 | 0 | 0 | 4 | 0 | 7 | 1 | 0 | 1 | 9 | 0 |
| <i>CG11159</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG13437</i> | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG30152</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cnk</i> | 0 | 1 | 0 | 1 | 2 | 0 | 5 | 0 | 0 | 3 | 8 | 3 |
| <i>cos</i> | 1 | 0 | 0 | 3 | 4 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| <i>grau</i> | 1 | 3 | 0 | 2 | 6 | 4 | 1 | 0 | 0 | 1 | 2 | 1 |
| <i>Lcp1</i> | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 2 |
| <i>nompA</i> | 1 | 2 | 0 | 1 | 4 | 0 | 2 | 1 | 0 | 2 | 5 | 6 |
| <i>Pcl</i> | 1 | 3 | 0 | 1 | 5 | 3 | 2 | 0 | 0 | 0 | 2 | 0 |
| <i>sax</i> | 3 | 0 | 0 | 2 | 5 | 1 | 2 | 0 | 0 | 3 | 5 | 0 |
| <i>stan</i> | 1 | 3 | 0 | 1 | 5 | 2 | 15 | 4 | 0 | 9 | 28 | 6 |
| <i>Toll-7</i> | 6 | 0 | 0 | 0 | 6 | 0 | 11 | 0 | 0 | 3 | 14 | 3 |
| <i>tud</i> | 0 | 1 | 0 | 1 | 2 | 8 | 2 | 0 | 0 | 0 | 2 | 2 |
| <i>UbaI</i> | 0 | 1 | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 3 | 4 | 2 |
| <i>Updo</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 0 | 0 | 0 | 2 | 2 | 6 | 1 | 0 | 0 | 0 | 1 | 0 |
| Total | 22 | 14 | 0 | 20 | 56 | 26 | 51 | 8 | 0 | 30 | 89 | 27 |
| Neo-Y/ <i>D. pseudoobscura</i> | | | | | | | | | | | | |
| <i>Alk</i> | 2 | 0 | 0 | 2 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Asx</i> | 1 | 0 | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>CG11136</i> | 7 | 0 | 0 | 0 | 7 | 6 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG11159</i> | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG13437</i> | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG30152</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cnk</i> | 2 | 0 | 0 | 1 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cos</i> | 2 | 0 | 0 | 3 | 5 | 6 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>grau</i> | 0 | 1 | 0 | 3 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Lcp1</i> | 1 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>nompA</i> | 1 | 2 | 0 | 5 | 8 | 15 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>Pcl</i> | 4 | 3 | 0 | 3 | 10 | 6 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>sax</i> | 4 | 0 | 0 | 2 | 6 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>stan</i> | 8 | 2 | 0 | 1 | 11 | 7 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>Toll-7</i> | 11 | 0 | 0 | 3 | 14 | 12 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>tud</i> | 7 | 0 | 0 | 4 | 11 | 16 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>UbaI</i> | 4 | 2 | 0 | 1 | 7 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Updo</i> | 0 | 0 | 0 | 4 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 2 | 0 | 0 | 3 | 5 | 8 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 56 | 10 | 0 | 37 | 103 | 102 | 0 | 0 | 0 | 2 | 2 | 7 |

Ns, nonsynonymous. We used *D. pseudoobscura* for comparison and *D. affinis* to infer the ancestral state. The *D. pseudoobscura* preferences table (AKASHI and SCHAEFFER 1997) was used to establish the major codons.

mutations (2.32 vs. 0.57, $P < 0.01$, Fisher's exact test), even after removing *stan* from the analysis because of its unusually high level of polymorphism (see above) (1.71 vs. 0.36, $P < 0.05$). This is consistent with selection against deleterious synonymous mutations (P → U). Moreover, although there is a greater number of P → U than U → P fixations, the difference is not statistically significant (22 vs. 14, $P > 0.05$, χ^2 test), suggesting that

codon usage is close to equilibrium, as previously reported for autosomes and X-linked chromosomes (BARTOLOMÉ *et al.* 2005). If we pool the results from BARTOLOMÉ *et al.* (2005) with these, we get 41 (P → U):26 (U → P) fixations, which is still not significant. Using the neo-X polymorphism data to calculate $N_e s$ (BARTOLOMÉ *et al.* 2005; MASIDE *et al.* 2004), we obtain an estimate of 0.65 (two-unit support limits 0.45–0.98),

TABLE 7

Patterns of synonymous and nonsynonymous changes on the neo-sex chromosomes of *D. miranda* since their divergence

| | Fixed | | | | | Polymorphic | | | | | | |
|----------------|-------|-----|-----|-----|-------|-------------|-----|-----|-----|-----|-------|----|
| | P-U | U-P | P-P | U-U | Total | Ns | P-U | U-P | P-P | U-U | Total | Ns |
| neo-X branch | | | | | | | | | | | | |
| <i>Alk</i> | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Asx</i> | 1 | 0 | 0 | 3 | 4 | 1 | 1 | 1 | 0 | 2 | 4 | 1 |
| <i>CG11136</i> | 3 | 0 | 0 | 1 | 4 | 0 | 7 | 1 | 0 | 1 | 9 | 0 |
| <i>CG11159</i> | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG13437</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG30152</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cnk</i> | 0 | 1 | 0 | 1 | 2 | 1 | 5 | 0 | 0 | 3 | 8 | 3 |
| <i>cos</i> | 0 | 0 | 0 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| <i>grau</i> | 1 | 1 | 0 | 0 | 2 | 5 | 1 | 0 | 0 | 1 | 2 | 1 |
| <i>Lcp1</i> | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 2 |
| <i>nompA</i> | 1 | 1 | 0 | 3 | 5 | 0 | 2 | 1 | 0 | 2 | 5 | 6 |
| <i>Pcl</i> | 0 | 1 | 0 | 0 | 1 | 4 | 2 | 0 | 0 | 0 | 2 | 0 |
| <i>sax</i> | 2 | 0 | 0 | 1 | 3 | 0 | 2 | 0 | 0 | 3 | 5 | 0 |
| <i>stan</i> | 1 | 1 | 0 | 0 | 2 | 1 | 15 | 4 | 0 | 9 | 28 | 6 |
| <i>Toll-7</i> | 2 | 0 | 0 | 0 | 2 | 0 | 11 | 0 | 0 | 3 | 14 | 3 |
| <i>tud</i> | 1 | 1 | 0 | 0 | 2 | 9 | 2 | 0 | 0 | 0 | 2 | 2 |
| <i>UbaI</i> | 1 | 0 | 0 | 2 | 3 | 0 | 0 | 1 | 0 | 3 | 4 | 2 |
| <i>Updo</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 0 | 0 | 0 | 2 | 2 | 6 | 1 | 0 | 0 | 0 | 1 | 0 |
| Total | 15 | 7 | 0 | 15 | 37 | 28 | 51 | 8 | 0 | 30 | 89 | 27 |
| neo-Y branch | | | | | | | | | | | | |
| <i>Alk</i> | 1 | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Asx</i> | 2 | 0 | 0 | 0 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>CG11136</i> | 5 | 0 | 0 | 0 | 5 | 6 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>CG11159</i> | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG13437</i> | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>CG30152</i> | 1 | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cnk</i> | 2 | 1 | 0 | 1 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>cos</i> | 1 | 0 | 0 | 2 | 3 | 6 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>grau</i> | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Lcp1</i> | 1 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>nompA</i> | 1 | 1 | 0 | 4 | 6 | 14 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>Pcl</i> | 3 | 1 | 0 | 2 | 6 | 6 | 0 | 0 | 0 | 1 | 1 | 0 |
| <i>sax</i> | 6 | 0 | 0 | 1 | 7 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>stan</i> | 5 | 0 | 0 | 0 | 5 | 5 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>Toll-7</i> | 6 | 0 | 0 | 3 | 9 | 10 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>tud</i> | 7 | 0 | 0 | 4 | 11 | 14 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>UbaI</i> | 4 | 1 | 0 | 1 | 6 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>Updo</i> | 0 | 0 | 0 | 4 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| <i>vlc</i> | 2 | 0 | 0 | 3 | 5 | 7 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 47 | 4 | 0 | 30 | 81 | 97 | 0 | 0 | 0 | 2 | 2 | 7 |

Ns, nonsynonymous. We used *D. affinis* to polarize mutations. The *D. pseudoobscura* preferences table (AKASHI and SCHAEFFER 1997) was used to establish the major codons.

comparable with that reported for X-linked and autosomal genes by BARTOLOMÉ *et al.* (2005).

CUB on the neo-X and neo-Y chromosomes since their split: To test for changes in the strength of selection acting on codon usage bias after the split between the neo-sex chromosomes, we investigated each branch individually, assigning variants to each lineage by parsimony, using *D. affinis* as an outgroup (Table 7). The numbers of fixations of P → U and U → P changes on the neo-X chromosome are not significantly different

(15 *vs.* 7, $P > 0.05$), again consistent with equilibrium for codon usage. The value of r_{pd} (P → U) on the neo-X chromosome is three times higher than that for r_{pd} (U → P) (3.40 *vs.* 1.14, respectively). This difference is not statistically significant ($P = 0.104$, Fisher's exact test), although it is in the same direction as expected with selection on codon usage.

This suggests that selection may be weaker than for other chromosomes (BARTOLOMÉ *et al.* 2005). In a recent study of a large number of neo-X chromosome

genes, D. BACHTROG and P. ANDOLFATTO (personal communication) found a significantly larger number of P → U compared to U → P fixations on the neo-X chromosome, in agreement with the apparent relaxation in selection for codon bias detected here and with previous observations reported by BACHTROG (2003b, 2005). However, the difference in numbers between the two types of fixed changes of each kind may be somewhat biased by the use of *D. pseudoobscura* as an outgroup (see above).

The pattern observed for the neo-Y chromosome is quite different; when we compare P → U vs. U → P fixations between the neo-X and neo-Y lineages, we find a highly significant difference ($P = 0.01$, Fisher's exact test). This substantial accumulation of unpreferred mutations on the neo-Y chromosome is consistent with a significant reduction in selection for maintaining codon usage bias on the neo-Y lineage since its split from the neo-X, as also found by BACHTROG (2005) and D. BACHTROG and P. ANDOLFATTO (personal communication).

If genes with high expression on the neo-Y (Table 5) are compared with the neo-X, we find no significant difference in the rate of accumulation of P → U fixations between the two chromosomes, but such an effect is found for the low-expression genes. There is, however, no significant heterogeneity between the two expression classes, as might be expected from the fact that selection for maintaining codon usage is weaker for genes with low expression levels (DURET and MOUCHIROUD 1999; AKASHI 2001). When nonsynonymous fixations and P → U fixations are pooled, and compared with the neo-X and neo-Y chromosomes, we find a significant excess on the neo-Y for highly expressed genes ($P < 0.01$). It seems likely, therefore, that mildly deleterious mutations accumulate on the neo-Y even in well-expressed genes.

Evolution of base composition: The action of non-selective forces, such as biased gene conversion in favor of GC over AT (GALTIER *et al.* 2006), or changes in mutational biases may cause a departure from neutrality in the nucleotide composition of the chromosomes. These could be confounded with selection on codon usage, since most optimal codons in *D. pseudoobscura* end in G or C (AKASHI and SCHAEFFER 1997). To examine this question, we compared r_{pd} (GC → AT) to r_{pd} (AT → GC) in coding (r_c) and noncoding (r_{nc}) DNA.

If the patterns observed on the neo-sex chromosomes were caused by nonselective mechanisms, we should find no differences between these regions, but if they are due to selection on codon usage, we should detect different rates of nucleotide substitution in coding and noncoding DNA (assuming that noncoding regions are closer to neutrality). The analysis of synonymous nucleotide substitutions on the neo-X chromosome since the split with *D. pseudoobscura* (Table 8) shows that r_c is considerably greater than r_{nc} , due to a significant excess

TABLE 8

Polymorphic and fixed synonymous mutations at coding and noncoding sites on the neo-X chromosome of *D. miranda*

| | GC-AT | AT-GC | | Fisher's exact test |
|--|-------|-------|-----------------|---------------------|
| Since split with <i>D. pseudoobscura</i> | | | | |
| Coding | | | | |
| Fixed | 32 | 17 | | $P = 0.031$ |
| Polymorphic | 65 | 13 | | |
| r_{pd} | 2.03 | 0.76 | $r_c = 2.66$ | |
| Noncoding | | | | |
| Fixed | 9 | 7 | | $P = 1.000$ |
| Polymorphic | 7 | 4 | | |
| r_{pd} | 0.78 | 0.57 | $r_{nc} = 1.36$ | |
| Since split with neo-Y | | | | |
| Coding | | | | |
| Fixed ^a | 25 | 7 | | $P = 0.790$ |
| Polymorphic | 62 | 14 | | |
| r_{pd} | 2.48 | 2.00 | $r_c = 1.24$ | |
| Noncoding | | | | |
| Fixed ^a | 7 | 2 | | $P = 0.659$ |
| Polymorphic | 8 | 4 | | |
| r_{pd} | 1.14 | 2.00 | $r_{nc} = 0.57$ | |

^a“Fixed” refers to fixed differences between neo-X chromosome and *D. pseudoobscura*. *D. affinis* was used to infer the ancestral state.

^aFixed differences between neo-X and neo-Y chromosomes.

of GC → AT polymorphisms at coding sites relative to introns and noncoding flanking sequences ($P < 0.05$, Fisher's exact test), which is consistent with selection being the main force controlling codon usage bias in the ancestral state of *D. miranda* (previous to the split of both neo-sex chromosomes). In addition, when we pool the results of this study with those for autosomes and X-linked loci (BARTOLOMÉ *et al.* 2005), we find that the numbers of fixed mutations are significantly different between coding and noncoding sites (62 GC → AT and 29 AT → GC in coding regions vs. 25 GC → AT and 29 AT → GC in noncoding regions; $P < 0.01$, Fisher's exact test). Since the split between the neo-sex chromosomes, there is a much greater number of GC → AT than AT → GC fixations for exons on the neo-X chromosome ($P < 0.01$). When we compare the pattern observed on the neo-X chromosome after the split of the neo-sex chromosomes, we find that $r_c > r_{nc}$, although not significantly so, perhaps suggesting weaker selection on synonymous sites.

DISCUSSION

Variability on the neo-X chromosome: The neo-X chromosome shows a level of variability that is slightly, but not significantly, lower than that for the autosomal and X-linked loci described previously (BARTOLOMÉ *et al.* 2005), after taking the selective sweep involving

the *exu1* region into account (BACHTROG 2003a). The mean silent-site divergences from *D. pseudoobscura* and *D. affinis* for the neo-*X* (2.80 and 23.63%, respectively) are very similar to those reported by BARTOLOMÉ *et al.* (2005) for other chromosomes (3.19 and 21.76%, respectively). Overall, these comparisons suggest that the neo-*X* chromosome has similar properties to those of the other chromosomes.

The only apparent difference is the overall significantly negative Tajima's *D* and Fu and Li's *D* values for silent sites, which were not found in our previous study. However, these are associated with noncoding sites rather than with synonymous sites, and reanalysis of the data for the autosomes and the *X* chromosome show that this is true of these chromosomes as well, with combined Tajima's *D* values of -0.57 ($P < 0.05$) and -0.07 for noncoding and synonymous sites, respectively (supplemental Table S2 at <http://www.genetics.org/supplemental/>). This pattern has also been found by BACHTROG and ANDOLFATTO (2006) and is consistent with recent findings of stronger selection on noncoding sequences than on synonymous sites in coding sequences in *Drosophila* (ANDOLFATTO 2005; HADDRILL *et al.* 2005). Again, there does not seem to be anything unusual about the neo-*X* chromosome in this respect.

Variability on the neo-*Y* chromosome: As in previous studies, our data show that variability is severely reduced at both nonsynonymous and silent sites on the non-recombining neo-*Y* chromosome (BACHTROG and CHARLESWORTH 2002; BACHTROG 2004), with a neo-*X*/neo-*Y* ratio of 60 for θ at silent sites. The neo-*Y* loci surveyed here show a negative Tajima's *D* value for all sites combined (a pooled value of -1.85 , for a sample size $n = 18$, compared with a value of -1.98 found by BACHTROG 2004 for sample size $n = 12$). Pooling our data with those of BACHTROG (2004), by decreasing our data set to a sample size of 12, we get a pooled *D* value of -1.99 ($P < 0.01$). This indicates a severe distortion of the variant frequency spectrum in favor of rare variants (supplemental Figure S2 at <http://www.genetics.org/supplemental/>), suggestive of a recent selective sweep on the neo-*Y* chromosome (BACHTROG 2004).

Variability at nonsynonymous and silent sites is very similar, indicating the absence of strong selection on replacement mutations. This is consistent with the fact that most neo-*Y*-linked loci show a higher rate of protein sequence evolution than their counterparts on the neo-*X* chromosome (BACHTROG and CHARLESWORTH 2002; BACHTROG 2005). On comparing silent and nonsynonymous variants on the neo-*Y* lineage, we found no significant excess of nonsynonymous fixations relative to polymorphisms (Table 4), so there was no evidence for accelerated adaptive evolution of protein sequences on this chromosome. In addition, there is an excess of nonsynonymous *vs.* silent polymorphisms on the neo-*Y* chromosome compared to the neo-*X* (Tables 3 and 4), which strongly suggests relaxed purifying selection on

the neo-*Y* for sites whose homologs are still under selection on the neo-*X* chromosome. This is consistent with the expectation for a nonrecombining chromosome (CHARLESWORTH and CHARLESWORTH 2000).

In agreement with this result, we detected signs of degeneration in a subset of the neo-*Y* loci under analysis, as also described previously (BACHTROG 2003a, 2005). These are caused by major mutations, namely deletions and frameshift mutations (see Table 5), which result in the partial or total loss of gene function. A negative correlation between gene expression level and rate of amino-acid sequence evolution has been found in other systems (PAL *et al.* 2001; ROCHA and DANCHIN 2004; DRUMMOND *et al.* 2005), suggesting that more highly expressed genes are subject to stronger selective constraints. This might, therefore, cause neo-*Y* genes that are weakly expressed to exhibit higher rates of amino-acid sequence evolution. This raises the question of whether the higher rate of amino-acid sequence evolution of neo-*Y* genes is caused by relaxed selection on nonsynonymous mutations in defective or poorly expressed genes or by relaxed selection due to reduced effective population size (OHTA 1992; CHARLESWORTH and CHARLESWORTH 2000).

We examined this question by comparing the rate of amino-acid sequence evolution in genes with high levels of expression on the neo-*Y* chromosome (as seen by the intensity level of RT-PCR bands) and that are still free of major mutational lesions, with that for genes that are either defective or only weakly expressed on the neo-*Y* (Table 5). We found that neo-*Y* chromosome genes with low levels of expression evolve faster for nonsynonymous fixations than highly expressed loci, while no differences for K_s were found, implying that reduced neo-*Y* expression is associated with relaxed purifying selection. However, this leaves it unclear whether the accumulation of deleterious mutations on the neo-*Y* favors reduced expression on the neo-*Y* genes concerned, or whether reduced expression favors accumulation of nonsynonymous mutations (but see below).

Evolution of codon usage bias and GC content: Our analysis of codon usage bias on the neo-*X* chromosome shows that there has been selection on synonymous variants during most of its evolutionary history (the polymorphism-to-divergence ratio is significantly greater for P \rightarrow U changes than for U \rightarrow P mutations), as described previously for autosomal and *X*-linked genes (BARTOLOMÉ *et al.* 2005). Moreover, there is no departure from equality for P \rightarrow U and U \rightarrow P substitutions fixations on the neo-*X* chromosome since its split from *D. pseudoobscura*, consistent with equilibrium for codon usage (Table 6). Its estimated $N_e s$ value is similar to our previous estimate for other *D. miranda* chromosomes (BARTOLOMÉ *et al.* 2005).

However, as suggested by BACHTROG (2005), even though selection maintains some codon bias on the neo-*X* chromosome, there may have been a decline in its

strength since its split from the neo-Y; although r_{pd} ($P \rightarrow U$) is three times higher than r_{pd} ($U \rightarrow P$), this difference is not statistically significant (Table 7). A significant excess of $P \rightarrow U$ over $U \rightarrow P$ fixations on the neo-X branch was found by D. BACHTROG and P. ANDOLFATTO (personal communication), using a larger data set for the neo-X chromosome, although their use of *D. pseudoobscura* to polarize mutations may have introduced a bias toward $P \rightarrow U$ fixations, as discussed by BARTOLOMÉ *et al.* (2005).

Using the comparison of $GC \rightarrow AT$ with $AT \rightarrow GC$ mutations, we find a significant excess of synonymous $GC \rightarrow AT$ fixations in coding sequences on the neo-X chromosome since the split with *D. pseudoobscura* (Table 8), as was also found for autosomal and X-linked mutations by BARTOLOMÉ *et al.* (2005). This is not found for noncoding sequences, suggesting that there has been a relaxation of selection rather than a change in mutational bias, most probably due to a reduction in effective population size in *D. miranda* (BACHTROG 2005).

The implication is that there has been a general relaxation of selection on synonymous sites since the species split, although selection is still effective, as indicated by the significant difference in $GC \rightarrow AT$ *vs.* $AT \rightarrow GC$ polymorphisms among coding but not among noncoding sequences. The puzzle is why this is not more apparent for mutations affecting codon preferences rather than GC *vs.* AT . One possible answer is that biased gene conversion (BGC) in favor of GC *vs.* AT (MARAIS 2003) may be operating, as suggested by BARTOLOMÉ *et al.* (2005) from indirect evidence, but too weakly to produce a large excess of $GC \rightarrow AT$ *vs.* $AT \rightarrow GC$ among noncoding sequence polymorphisms. If the data for the X and autosomes are pooled with the neo-X data, there is a total of 20 $GC \rightarrow AT$ compared with 13 $AT \rightarrow GC$ polymorphisms at noncoding sites. Although nonsignificant, this is in the direction expected with BGC. Since not all $P \rightarrow U$ mutations are associated with change in base composition, there will be a weaker effect of BGC on such mutations, so there may be a weaker net effect of reduced N_e on mutations affecting codon usage than on those affecting the use of GC *vs.* AT .

We also detected a significant reduction in selection for maintaining codon usage bias on the neo-Y lineage after the split, with a large excess of $P \rightarrow U$ fixations compared with the neo-X chromosome (Table 7). Since BGC is confounded with selection at coding sites, it is conceivable that this effect could be due to the absence of BGC on the nonrecombining neo-Y chromosome. This seems very unlikely, however, given the much smaller effect of BGC than selection on codon usage on the other chromosomes.

The accumulation of $P \rightarrow U$ fixations on the neo-Y chromosome is also greater in genes with reduced expression, as would be expected given the evidence from genome analyses that lower gene expression levels are associated with lower codon usage bias (AKASHI

2001). However, when we pool nonsynonymous and $P \rightarrow U$ fixations on the two branches, we find a significant excess on the neo-Y, even in genes that are highly expressed on the neo-Y. In a study by pyrosequencing of gene expression for a much larger set of genes, BACHTROG (2006) found cases of high K_a/K_s on the neo-Y branch for genes with similar levels of expression on the two neo-sex chromosomes and detected no significant correlation between K_a/K_s and expression level.

This implies that mildly deleterious mutations are accumulating on the neo-Y chromosome, even in well-expressed genes, which is consistent with the idea that the reduced effective population size of the neo-Y chromosome is leading to mutational degeneration of codon usage and protein sequences in expressed genes (OHTA 1992; CHARLESWORTH and CHARLESWORTH 2000).

We are grateful to Xulio Maside for useful comments and discussion and to Adam Eyre-Walker and Hiroshi Akashi for helpful suggestions for improving the original manuscript. We thank Peter Andolfatto and Doris Bachtrog for showing us their unpublished results and for commenting on the manuscript and the Carnegie Institution of Washington for permission to use their field station at Mather, California, for collecting flies. This work was supported by a grant from the United Kingdom Biotechnology and Biological Sciences Research Council. B.C. is supported by the Royal Society.

LITERATURE CITED

- AKASHI, H., 1995 Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. *Genetics* **139**: 1067–1076.
- AKASHI, H., 1999 Inferring the fitness effects of DNA mutations from polymorphism and divergence data: statistical power to detect directional selection under stationarity and free recombination. *Genetics* **151**: 221–238.
- AKASHI, H., 2001 Gene expression and molecular evolution. *Curr. Opin. Genet. Dev.* **11**: 660–666.
- 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., 2005 Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**: 1149–1152.
- BACHTROG, D., 2003a Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in *Drosophila*. *Nat. Genet.* **34**: 215–219.
- BACHTROG, D., 2003b Protein evolution and codon usage bias on the neo-sex chromosomes of *Drosophila miranda*. *Genetics* **165**: 1221–1232.
- BACHTROG, D., 2004 Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. *Nat. Genet.* **36**: 518–522.
- BACHTROG, D., 2005 Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in *Drosophila*. *Genome Res.* **15**: 1393–1401.
- BACHTROG, D., 2006 Expression profile of a degenerating neo-Y chromosome in *Drosophila*. *Curr. Biol.* **16**: 1694–1699.
- BACHTROG, D., and P. ANDOLFATTO, 2006 Selection, recombination and demographic history of *Drosophila miranda*. *Genetics* **174**: 2045–2059.
- BACHTROG, D., and B. CHARLESWORTH, 2002 Reduced adaptation of a non-recombining neo-Y chromosome. *Nature* **416**: 323–326.
- BARTOLOMÉ, C., and B. CHARLESWORTH, 2006 Rates and patterns of chromosomal evolution in *Drosophila pseudoobscura* and *D. miranda*. *Genetics* **173**: 779–791.
- BARTOLOMÉ, C., X. MASIDE, S. YI, A. L. GRANT and B. CHARLESWORTH, 2005 Patterns of selection on synonymous and nonsynonymous variants in *Drosophila miranda*. *Genetics* **169**: 1495–1507.

- BULMER, M., 1991 The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129**: 897–907.
- CHARLESWORTH, B., and D. CHARLESWORTH, 2000 The degeneration of *Y* chromosomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**: 1563–1572.
- CHARLESWORTH, B., C. BARTOLOMÉ and V. NÖEL, 2005 The detection of shared and ancestral polymorphisms. *Genet. Res.* **86**: 149–157.
- DRUMMOND, D. A., J. D. BLOOM, C. ADAMI, C. O. WILKE and F. H. ARNOLD, 2005 Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. USA* **102**: 14338–14343.
- DURET, L., and D. MOUCHIROUD, 1999 Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **96**: 4482–4487.
- FU, Y. X., and W. H. LI, 1993 Statistical tests of neutrality of mutations. *Genetics* **133**: 693–709.
- GALTIER, N., E. BAZIN and N. BIERNE, 2006 GC-biased segregation of noncoding polymorphisms in *Drosophila*. *Genetics* **172**: 221–228.
- HADRILL, P. R., B. CHARLESWORTH, D. L. HALLIGAN and P. ANDOLFATTO, 2005 Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. *Genome Biol.* **6**: R67.
- HUDSON, R. R., M. KREITMAN and M. AGUADE, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.
- IKEMURA, T., 1981 Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.* **146**: 1–21.
- JUKES, T. H., and C. R. CANTOR, 1969 Evolution of protein molecules, pp. 21–132 in *Mammalian Protein Metabolism*, edited by H. N. MUNRO. Academic Press, New York.
- KEIGHTLEY, P. D., and T. JOHNSON, 2004 MCALIGN: stochastic alignment of noncoding DNA sequences based on an evolutionary model of sequence evolution. *Genome Res.* **14**: 442–450.
- LI, W. H., 1987 Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* **24**: 337–345.
- LOEWE, L., B. CHARLESWORTH, C. BARTOLOMÉ and V. NÖEL, 2006 Estimating selection on nonsynonymous mutations. *Genetics* **172**: 1079–1092.
- LUCCHESI, J. C., 1978 Gene dosage compensation and evolution of sex chromosomes. *Science* **202**: 711–716.
- MARAI, G., 2003 Biased gene conversion: implications for genome and sex evolution. *Trends Genet.* **19**: 330–338.
- MASIDE, X., A. W. LEE and B. CHARLESWORTH, 2004 Selection on codon usage in *Drosophila americana*. *Curr. Biol.* **14**: 150–154.
- MCALLISTER, B. F., and B. CHARLESWORTH, 1999 Reduced sequence variability on the neo-*Y* chromosome of *Drosophila americana americana*. *Genetics* **153**: 221–233.
- MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.
- NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- OHTA, T., 1992 The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* **23**: 263–286.
- PAL, C., B. PAPP and L. D. HURST, 2001 Highly expressed genes in yeast evolve slowly. *Genetics* **158**: 927–931.
- ROCHA, E. P., and A. DANCHIN, 2004 An analysis of determinants of amino acids substitution rates in bacterial proteins. *Mol. Biol. Evol.* **21**: 108–116.
- ROZAS, J., J. C. SANCHEZ-DEL BARRIO, X. MESSEGUER and R. ROZAS, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- RUSO, C., N. TAKEZAKI and M. NEI, 1995 Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* **12**: 391–404.
- SAWYER, S. A., D. E. DYKHUIZEN and D. L. HARTL, 1987 Confidence interval for the number of selectively neutral amino acid polymorphisms. *Proc. Natl. Acad. Sci. USA* **84**: 6225–6228.
- STEINEMANN, M., and S. STEINEMANN, 1992 Degenerating *Y* chromosome of *Drosophila miranda*: a trap for retrotransposons. *Proc. Natl. Acad. Sci. USA* **89**: 7591–7595.
- 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.
- TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- TAMURA, K., S. SUBRAMANIAN and S. KUMAR, 2004 Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol. Biol. Evol.* **21**: 36–44.
- WATTERSON, G. A., 1975 On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**: 256–276.
- WRIGHT, S. I., and B. CHARLESWORTH, 2004 The HKA test revisited: a maximum-likelihood-ratio test of the standard neutral model. *Genetics* **168**: 1071–1076.
- YI, S., and B. CHARLESWORTH, 2000a Contrasting patterns of molecular evolution of the genes on the new and old sex chromosomes of *Drosophila miranda*. *Mol. Biol. Evol.* **17**: 703–717.
- YI, S., and B. CHARLESWORTH, 2000b A selective sweep associated with a recent gene transposition in *Drosophila miranda*. *Genetics* **156**: 1753–1763.
- YI, S., D. BACHTROG and B. CHARLESWORTH, 2003 A survey of chromosomal and nucleotide sequence variation in *Drosophila miranda*. *Genetics* **164**: 1369–1381.
- YU, Y.-C., F.-J. LIN and H.-Y. CHANG, 1999 Stepwise chromosome evolution in *Drosophila albomicans*. *Heredity* **83**: 39–45.

Communicating editor: D. BEGUN