Contrasting Patterns of Molecular Evolution of the Genes on the New and Old Sex Chromosomes of *Drosophila miranda*

Soojin Yi and Brian Charlesworth†*

*Department of Ecology and Evolution, University of Chicago; and †Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, Scotland

In organisms with chromosomal sex determination, sex is determined by a set of dimorphic sex chromosomes that are thought to have evolved from a set of originally homologous chromosomes. The chromosome inherited only through the heterogametic sex (the Y chromosome in the case of male heterogamety) often exhibits loss of genetic activity for most of the genes carried on its homolog and is hence referred to as degenerate. The process by which the proto-Y chromosome loses its genetic activity has long been the subject of much speculation. We present a DNA sequence variation analysis of marker genes on the evolving sex chromosomes (neo-sex chromosomes) of *Drosophila miranda.* Due to its relatively recent origin, the neo-Y chromosome of this species is presumed to be still experiencing the forces responsible for the loss of its genetic activity. Indeed, several previous studies have confirmed the presence of some active loci on this chromosome. The genes on the neo-Y chromosome surveyed in the current study show generally lower levels of variation compared with their counterparts on the neo-X chromosome or an X-linked gene. This is in accord with a reduced effective population size of the neo-Y chromosome. Interestingly, the rate of replacement nucleotide substitutions for the neo-Y linked genes is significantly higher than that for the neo-X linked genes. This is not expected under a model where the faster evolution of the X chromosome is postulated to be the main force driving the degeneration of the Y chromosome.

Introduction

The presence of dimorphic sex chromosomes—a distinctive, often largely nonhomologous pair of chromosomes—is a common characteristic of chromosomal sex determination systems. Various aspects of sex chromosomes have long been subjected to numerous studies (Wilson 1925; Mittwoch 1967; Ohno 1967). Several pieces of direct and indirect evidence suggest that dimorphic sex chromosomes have evolved from an originally homologous pair of chromosomes (Mittwoch 1967; Ohno 1967; Bull 1983; Guttman and Charlesworth 1998). The chromosome that is transmitted only through the heterogametic sex (which will be referred to here as the Y chromosome) exhibits some common characteristics in different species (Bull 1983; Charlesworth 1991, 1996). It usually contains little genetic information, and often remains condensed (heterochromatic) throughout the cell cycle. A key phenomenon is that meiotic recombination between the X and Y chromosomes is usually absent or is confined to a small region of genetic homology. Transposable elements and other repetitive sequences tend to accumulate on the Y chromosome, further contributing to its differentiation from the X chromosome.

A critical step in the process of the evolution of dimorphic sex chromosomes is generally assumed to be the restriction of recombination between the X chromosome and the primitive Y chromosome. Once recombination is restricted, it facilitates the close association of primary sex-determining genes with genetic functions advantageous to males but disadvantageous to females, and thus enhances the differentiation of the formerly homologous chromosomes (Charlesworth 1991; Rice 1996). Various population genetic processes can lead to the loss of gene function from a chromosome that is prevented from crossing over (Charlesworth 1996; Rice 1996; Orr and Kim 1998). Once the Y chromosome begins to lose activity for genes that were originally homologous with those on the X chromosome, the deleterious effects associated with such functional aneuploidy favors the evolution of dosage compensation. This involves adjustment of the activity of X-linked loci in the heterogametic sex, such that the rate of transcription in males of those loci is effectively twice that in females (Bull 1983; Bashaw and Baker 1996; Cline and Meyer 1996; Jegalian and Page 1998; Lucchesi 1998). Several different mechanisms for dosage compensation have evolved independently in different lineages (Lucchesi 1994).

The independent occurrence of both the evolutionary erosion of Y chromosomes and the evolution of dosage compensation in various distantly related species suggests that simple evolutionary forces are acting to cause these phenomena. Despite ample theorizing and a long history of attention from evolutionary biologists, until recently the necessary molecular tools to study the exact mechanism of the evolution of degenerate Y chromosomes in different taxa were not available. This study is the first in a series of attempts to shed light on this problem by examining patterns of molecular evolution and sequence variation of homologous X- and Y-linked loci from the newly developing sex chromosome system of *Drosophila miranda.* As outlined by Charlesworth (1996), the different possible evolutionary mechanisms of Y chromosome degeneration make different detailed predictions about these patterns, although they all entail a considerable reduction in effective population size of the primitive Y chromosome over what is expected from its numerical abundance (see *Discussion*).

Key words: *Drosophila miranda*, DNA polymorphism, neo-sex chromosomes, dosage compensation.

Address for correspondence and reprints: Soojin Yi, Department of Ecology and Evolution, University of Chicago, 1101 East 57th Street, Chicago, Illinois 60637. E-mail: soojinyi@midway.uchicago.edu.

Mol. Biol. Evol. 17(5):703–717. 2000

The newly developing system of neo-X and neo-Y chromosomes in *D. miranda* provides a unique opportunity for this purpose. The neo-Y chromosome of *D. miranda* has resulted from the fusion of Muller's element C of the basic *Drosophila* karyotype (Powell 1997) with the ancestral Y chromosome. This fusion is absent in the two sibling species, *Drosophila pseudoobscura* and *Drosophila persimilis* (Dobzhansky 1935; MacKnight 1939), setting an upper limit to the age of this fusion as the divergence time between *D. miranda* and the other two sibling species. Several studies have shown that this system possesses the characteristics expected of developing sex chromosomes. Specifically, the neo-Y chromosome appears to be in an intermediate stage of degeneration, still possessing some active loci (MacKnight 1939; Steinemann, Steinemann, and Lottspeich 1993). On the other hand, the homologous element (X_2), which cosegregates with the true X chromosome in males, has evolved an incomplete level of dosage compensation, apparently by adopting the preexisting dosage compensation machinery (Strobel, Pelling, and Arnheim 1978; Norman and Doane 1990; Bone and Kuroda 1996; Marín et al. 1996; Steinemann, Steinemann, and Turner 1996).

This study concerns the patterns of molecular evolution and variation of the larval cuticle protein (*Lcp*) genes. These are among the first genes localized and sequenced from both of the neo-sex chromosomes in *D. miranda* (Steinemann and Steinemann 1993; Steinemann, Steinemann, and Lottspeich 1993; Steinemann, Steinemann, and Pinsker 1996). Their sequences were used to design allele-specific primers to amplify neo-X and neo-Y alleles, thus allowing us to investigate the patterns of nucleotide variation of loci on the two chromosomes. Homologous *D. pseudoobscura* alleles were also obtained for use as an outgroup. Sequence variation of the X-linked *period* (*per*) gene was also studied (see also Wang and Hey [1996]) in order to provide a reference locus with which to compare levels of variation at the *Lcp* loci.

Materials and Methods

Species and Strains

Among the 12 strains of *D. miranda* studied, 5 were acquired from the National Drosophila Species Resource Center (Bowling Green, Ohio). Mohamed Noor kindly provided an additional four lines, derived from single females captured in Mather, Calif., and Wyatt Anderson provided the last three lines (from Spray, Oreg.). All the *D. miranda* lines were subjected to *Adh* allozyme analysis to verify their species identity (Anderson, Ayala, and Michod 1977). The geographic origins of the strains are shown in table 1. A strain of the outgroup *D. pseudoobscura* was kindly provided by Jerry Coyne. Stocks were maintained in bottle culture on banana medium at 18°C.

Table 1
List of Strains of *Drosophila miranda* and *Drosophila pseudoobscura* Used in this Study, Along with Their Geographic Origins

Species	Strain ID	Geographic Origin
D. miranda	0101.3	Port Coquitlam, B.C., Canada
	0101.4	Port Coquitlam, B.C., Canada
	0101.5	Port Coquitlam, B.C., Canada
	0101.7	Port Coquitlam, B.C., Canada
	0101.9	Mather, California
	MA 28	Mather, California
	MA 32	Mather, California
	SP 138	Spray, Oregon
	SP 235	Spray, Oregon
	SP 295	Spray, Oregon
	MSH 22	Mt. St. Helena, California
	MSH 38	Mt. St. Helena, California
D. pseudoobscura . . .		Mather, California

Design of PCR Primers

Primers to amplify *D. miranda Lcp* alleles were designed from published sequences (Steinemann, Steinemann, and Pinsker 1996). They were designed so that nucleotide sites that differed between the two neo-sex chromosome alleles would be included in the primer, close to the 3′ end. Many different pairs of primers were produced initially. After preliminary PCR reactions, the primer pairs providing the most reliable PCR products were chosen. These PCR products were then sequenced to identify the allelic differences. For the *Lcp1* and *Lcp3* loci, having only the forward primer different and using the same reverse primer for neo-X and neo-Y alleles was sufficient to amplify different alleles. The actual primer sequences used are 5′-TTT TAT TGA TGC GAT TGT TTC-3′ as the forward primer for the *Lcp1* neo-X allele and 5′-GCA CTT TTA TTG ATA CGA CTG-3′ for the *Lcp1* neo-Y allele. The same reverse primer for both of the neo-sex alleles, 5′-GGT AAA GAA AGG TAT CCA GAG-3′, was used for *Lcp1* of *D. miranda*. For the *Lcp3* locus, the two primer sets were designed such that there was also a length difference between the neo-X and the neo-Y encoded alleles. The primer sequences were 5′-CAC CGC TTA TGC TAA AAA TA-3′ for the *Lcp3* neo-X forward primer and 5′-TTT GGA TAA TCT ACA ATT CG-3′ for the *Lcp3* neo-Y forward primer. Both of the *Lcp3* alleles were amplified using the same reverse primer 5′-TAT TTT CAG TGT GCT GTT CA-3′. The same logic was used to design primers for the *Lcp*2-neo-X and *Lcp*2-neo-Y alleles. Despite multiple attempts, however, we did not succeed in amplifying the *Lcp*2-neo-Y allele, so this locus was not included in the study.

Primers for the *D. pseudoobscura* genes were designed using the region that was shared by the two neo-sex-chromosome alleles; homology to *Drosophila melanogaster* (Snyder et al. 1982) was also considered where possible. The primer sequences are 5′-CGT TGT TTG GCA TTT GCT-3′ for the *Lcp1*-pseudoobscura forward primer and 5′-GAA CAG TCC ATG GTG CGA-3′ for the reverse primer. For the *Lcp3*-pseudoobscura locus, the forward primer is 5′-GCG GCT TGA TTT CGG C-3′ and the reverse primer is 5′-GCT GAT GCA ACC CGT TTT-3′.

Primers for the *per* locus were designed from the sequences reported by Wang and Hey (1996). Using the

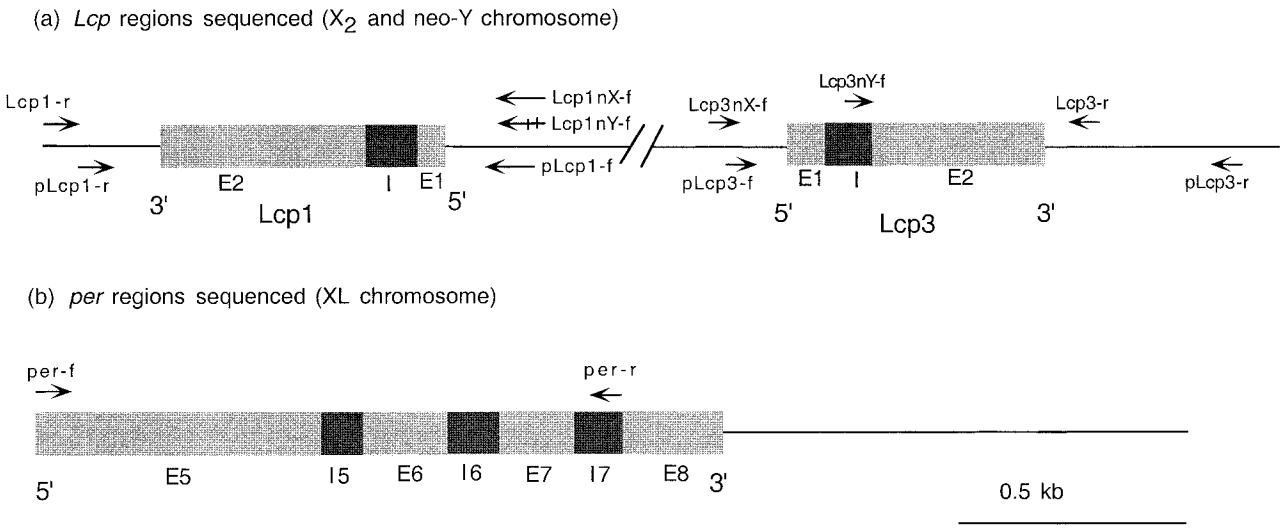

FIG. 1.—The regions surveyed in this study. Coding regions are drawn as blocks. E stands for exons (lighter blocks), and I stands for introns (darker blocks). Primers are represented as arrows. *a*, The *Lcp* region. The *Lcp1*nX-f and *Lcp1*nY-f primers cover the same region of the two alleles and differ by two nucleotide sites; the sites that differ are represented over the arrows. *b*, The *per* region.

four *D. miranda per* alleles sequenced by Wang and Hey (1996), primers providing the longest possible PCR products were designed. The resulting primer sequences were 5′-AAA ACT CTG CCT CTG GAT-3′ for the forward primer and 5′-GCA ATA GGA AAC ACA CGG-3′ for the reverse primer. Using these primers, we were able to amplify 1,480 bp of the *per* locus, extending from the latter part of exon 5 to intron 7. The structures and the primer sites of *Lcp1*, *Lcp3*, and *per* are represented in figure 1.

Sequencing

Genomic DNA from single *D. miranda* males was extracted using the Puregene DNA extraction system (Gentra). After determining the appropriate PCR conditions for each allele of each *Lcp* gene, DNA sampled from one male fly of each strain was used to amplify both of the neo-sex-chromosome alleles. One 50-μl PCR reaction usually provided enough product for subsequent direct-sequencing reactions. Both strands were sequenced using this sample. Sequencing primers were designed so that they would be spaced approximately 400 bp apart on each strand. In cases in which there was any ambiguity, the PCR was performed again using the original DNA sample. To amplify the *per* locus, basically the same method was used using single male fly genomic DNA as PCR template. Since the *per* locus is on the XL chromosome, each PCR reaction should amplify only one allele.

For the outgroup *D. pseudoobscura* allele, single-fly genomic DNA was used for the original PCR. Three separate PCR reactions using three different single-fly DNA extractions were performed, and no heterozygosity was found at the two *Lcp* loci studied. All of the sequencing was performed by the ampliTaq FS cycle sequencing method (Applied Biosystems Inc.) using an ABI 377 sequencer. Sequences obtained from this work are available in GenBank (accession numbers AF218953–AF218964 for *per*, AF219246–AF219248 and AF219253–AF219256 for *Lcp1*, and AF219248–AF219252 for *Lcp3*).

In Situ Hybridization

Preparation and in situ hybridization of salivary chromosome squashes from third-instar larvae generally followed the protocol described in Montgomery, Charlesworth, and Langley (1987), with slight modifications. PCR products from genomic DNA were extracted from 1% agarose gel using the QIAquick gel extraction kit (Qiagen) and then labeled with biotinylated dUTP (Boehringer Mannheim) by random primer extension for use as a probe. Sites of hybridization were detected by staining with diaminobenzidine and peroxidase (Vector Laboratories), with counterstaining of the polytene chromosomes in 5% Giemsa (Gurr). The banding patterns of polytene chromosomes of *D. miranda* reported by Das et al. (1982) were used for reference when determining the sites of hybridization.

Data Analysis

Sequences were first aligned using the Sequencher 3.0 program and then edited and aligned manually. Published sequences of *D. miranda* (see Steinemann, Steinemann, and Pinsker [1996] for *Lcp* sequences and Wang and Hey [1996] for *per*) allowed the sequences to be aligned unambiguously. There was no ambiguous assignment for the open reading frames for the *D. pseudoobscura* sequences. Once all of the data were aligned and trimmed to the same length, they were subject to further editing according to the purpose of the analysis. The positions and identities of nucleotide variants are given in table 2.

The longest consensus alignment was used as a reference for estimating the two commonly used measures of genetic variation from the data. The nucleotide site

Table 2
Polymorphic Nucleotide Sites of the Genes Used in this Study

	Period Locus (X chromosome)																	
Position	7 7	7 8	1 2 1	2 3 9	3 1 6	4 7 8	5 2 7	6 0 0	7 3 9	9 0 9	9 8 3	9 9 4	1 0 0 5	1 0 2 8	1 3 1 4	1 3 6 2	1 4 3 6	1 4 3 9
	R	R	S	R	S	S	R	R	S	I	R	R	R	S	R	I	I	I
0101.3.....	A	C	A	T	C	C	C	G	T	A	G	A	G	G	A	A	A	T
0101.4.....	·	·	G	·	·	·	·	·	C	T	·	C	A	A	·	·	·	·
0101.5.....	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·
0101.7.....	·	·	·	·	·	·	·	·	C	·	·	·	·	·	G	·	·	·
0101.9.....	·	·	G	·	·	·	G	C	C	T	·	C	·	A	G	T	G	A
MA28.....	G	A	·	·	·	·	G	·	C	T	·	C	·	A	G	·	·	·
MA32.....	·	·	·	·	·	T	·	C	C	·	A	·	·	A	G	·	·	·
SP138.....	·	·	·	·	·	·	·	C	C	T	·	C	A	A	G	T	G	·
SP235.....	·	·	G	·	·	·	·	C	C	T	·	C	·	A	G	T	·	·
SP295.....	·	·	G	A	·	·	G	C	C	T	·	C	A	A	G	T	·	·
MSH22.....	·	·	·	·	G	·	·	C	C	T	·	C	·	A	G	·	·	·
MSH38.....	·	·	·	A	·	·	·	C	C	T	·	C	A	A	G	T	·	·

Note.—R = replacement sites; S = synonymous sites; I = intron sites.

diversity based on the average pairwise differences among alleles present in the samples, π (Nei 1987), and the neutral mutation parameter, $\hat{\theta}$, based on the number of segregating sites in the sample (Watterson 1975), were estimated using Dna SP (Rozas and Rozas 1997). The Dna SP program was also used to calculate the chi-square value for the HKA test (Hudson, Kreitman, and Aguadé 1987) and Tajima's (1989) *D* statistic for testing whether the data showed any significant deviation from the equilibrium neutral model. Fisher's exact test was also performed using the Dna SP program to test the linkage disequilibrium status of the detected polymorphic sites. Various other analyses, including the recombination analysis, were done using the program SITES (Wakeley and Hey 1997). The permutation test of Hudson, Boos, and Kaplan (1992) was performed to detect population subdivision. The program for this was kindly provided by R. R. Hudson.

The numbers of synonymous (K_s) and nonsynonymous (K_a) substitutions per nucleotide site in the interspecies comparisons were estimated using the program K-Estimator, version 4.4 (provided by J. M. Comeron; also available from ftp.bio.indiana.edu/molbio.mswin). Confidence intervals and significance tests on K_s and K_a were computed as described elsewhere for an earlier version of the same program (Comeron and Kreitman 1998; Zeng et al. 1998).

Results

Evidence for Population Subdivision from Nucleotide Variation at the *period* Locus

While its sibling species *D. pseudoobscura* has been the subject of a number of studies of genetic variation (Powell 1997, p. 350), *D. miranda* has scarcely been investigated in this regard. To understand patterns of variation at neo-sex-chromosome loci, a measure of the ''typical'' level of nucleotide variation for this species is clearly necessary. This was the motivation for studying the level of nucleotide variation at the X-linked locus, *per*. From the 1,480 bp of this gene that were sequenced, a total of 18 polymorphic sites were found among 12 alleles. The actual base substitutions are presented in table 2. The two estimates of average nucleotide diversity, $\hat{\theta}$ and π, were around 0.5% for silent sites, comparable with the previous estimates of nucleotide diversity from this species (see below). However, one interesting feature that emerged from the polymorphic sites (table 2) was the possible existence of population subdivision. For example, none of the lines from British Columbia has the haplotype with A at site 1028 and G at site 1314, while all the other lines are fixed for this combination of variants at these two sites. Likewise, most of the lines from the locations other than British Columbia are fixed for the co-occurrence of C and C at sites 600 and 739. Some other sites show similar patterns, although there are some obvious cases of mixture.

We performed a statistical test to see whether the observed pattern of population differentiation is significant when compared with the neutral infinite-sites model with random mating, following the method of Hudson, Boos, and Kaplan (1992). Specifically, their test statistics *K* and *K** were used, since these take account of the number of nucleotide differences between different haplotypes. *K** differs from the statistic *K* in that it does not give as much weighting to large numbers of differences, due to the log transformation of the pairwise differences. These two have the most power when dealing with small sample sizes (Hudson, Boos, and Kaplan 1992). Another pair of statistics, the *H* and *H** statistics, deals with the haplotype frequencies within each locality that are present in samples of larger sizes and is considered inappropriate in a case such as this, in which none of the haplotypes are the same, to avoid any a priori lumping of haplotypes.

The results are shown in table 3. The *P* values for the tests are drawn from 10,000 permutations. Among the six pairwise comparisons, the ones between locality

Table 2
Extended

*Lcp*1-neo-X				*Lcp*3-neo-X			*Lcp*1-neo-Y	*Lcp*3-neo-Y
4	5	6	7	5	7	3	4	1
9	0	5	2	4	3	2	8	3
1	0	0	7			4	9	4
R	R	S	S	I	I	S	R	S
G	C	C	G	C	C	T	T	C
·	·	·	·	·	·	·	·	·
C	·	·	·	·	·	·	·	·
·	·	·	·	·	·	C	·	T
·	·	A	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·
·	·	A	·	·	·	C	A	T
·	·	·	·	T	·	C	·	·
·	T	·	A	·	·	·	·	·
·	·	·	·	·	A	·	·	·

1 (Port Coquitlam, B.C., Canada) and the others generally resulted in small *P* values, suggesting a certain level of population subdivision between locality 1 and the others. Other populations did not appear significantly different from each other. When the rest of the lines were pooled together and compared with the lines from locality 1, both the *K* and the *K** statistics were highly significant. F_{ST} for each comparison was also estimated using the method of Hudson, Slatkin, and Maddison (1992). When the alleles from locality 1 were compared with the rest of the samples, F_{ST} was 0.37. This value is rather high compared with other available estimates for *Drosophila* species (Charlesworth 1998). The results indicate that at the *per* locus, the British Columbia population is genetically different from other populations sampled in this study. It may be relevant to this finding that there is evidence for incipient sexual isolation between strains of *D. miranda* from the Puget Sound region, close to British Columbia, and strains from Mount Whitney, Calif. (Dobzhansky and Koller 1938).

Alternatively, the observed pattern of population differentiation could also be explained as the consequence of local selection at the *per* locus. The *per* gene in the *D. melanogaster* group encodes a fundamental property of biological clocks (Kyriacou and Hall 1994). Considering this property of the *per* gene, the observed differentiation between strains of flies from different geographic locations may be compatible with the presence of local adaptations leading to different clocks. The sequenced region does not, however, cover the Thy-Gly repeat that is known to significantly alter the properties of the PER protein and the song cycle in different species (Kyriacou and Hall 1994). In addition, an autosomal locus, *Adh*, from this species does not show a similar pattern of population subdivision (unpublished data). Obviously, further study is necessary to obtain a better understanding of any possible population structure in *D. miranda.*

Based on these observations, we pooled the samples other than those from British Columbia together and treated the samples as a mixture of two distinct populations. Specifically, we used the mean of the values for the British Columbia samples and the other populations to estimate the average nucleotide diversity for the species *D. miranda.* In the absence of any exact knowledge of the population size of each deme, this provided a rough estimate of the mutation rate scaled by the specieswide effective population size under the infinite-sites model (Charlesworth 1998; Nagylaki 1998; Stephan et al. 1998).

Table 3
Results of Population Structure Analyses of *period*

Population 1[a]	Population 2[a]	K_{ST} (*P* value)	K^*_{ST} (*P* value)	F_{ST}
1	2, 3, 4	0.186 (0.014)	0.089 (0.009)	0.37
1	2	0.073 (0.171)	0.091 (0.084)	0.18
1	3	0.344 (0.057)	0.183 (0.057)	0.52
1	4	0.340 (0.067)	0.208 (0.067)	0.42
2	3	0.063 (0.200)	0.050 (0.201)	0.04
2	4	0.0 (0.500)	−0.123 (0.900)	0.00
3	4	0.048 (0.502)	0.103 (0.209)	−0.10

[a] 1 = Port Coquitlam, B.C., Canada; 2 = Mather, Calif.; 3 = Spray, Oreg.; 4 = Mt. St. Helena, Calif.

The Level of Nucleotide Diversity at the *per* Locus

We report an estimated value of π per base of 0.0042 for all classes of sites pooled (table 4). When the arithmetic mean of the values from British Columbia and the others were used, this value was 0.0031 (table 4). This is within one standard deviation of the previous estimate from Wang and Hey (1996; based on 1.5 kb sequenced from four strains), assuming either no recombination or free recombination. This value is smaller than the estimates of variation reported from the same gene of *D. miranda*'s North American sibling species *D. pseudoobscura* and *D. persimilis,* or the cosmopolitan species *D. melanogaster* (from various sources, tabulated in Wang and Hey [1996]). The estimated π value of 0.0042 for silent sites is quite compatible with another available estimate of nucleotide variation of *D. miranda* from noncoding nucleotides located between *Est-5C* and *Est-5B* (0.0035, based on three alleles of 469 nt; Babcock and Anderson 1996). However, variation at *Hsp82* was much lower than these estimates (Wang, Wakeley, and Hey 1997). A lower level of variation at *Hsp82* was consistently observed in other species surveyed as well

Table 4
Polymorphism Summary for *period*

Sample[a]	No. of Alleles	Segregating Sites: Silent	Segregating Sites: Replacement	Total[b]: π/Site	Total[b]: $\hat{\theta}$/Site	Silent Sites[b]: π/Site	Silent Sites[b]: $\hat{\theta}$/Site
Total . . .	12	9	9	4.20	4.03	5.10	4.92
BC.	4	4	4	2.82	2.95	3.58	3.60
NBC . . .	8	8	7	3.55	3.91	5.01	5.10
Mean . . .				3.12	3.43	4.23	4.26

[a] BC = British Columbia; NBC = non-British Columbia.
[b] All values $\times 10^{-3}$.

(Wang, Wakeley, and Hey 1997), suggesting that this lower level of variation may be locus-specific. All of the available estimates of variation of *D. miranda* other than those from the present study are from just a handful of alleles.

When the polymorphic sites observed were classified into replacement sites, synonymous sites, and noncoding sites, there was greater variation at synonymous sites and noncoding sites than at nonsynonymous sites, as expected from the functional constraints on replacement sites. If we consider synonymous sites a close approximation to neutrally evolving sites, then the mean π value over the two classes of populations for an X-linked locus provides an estimate of $\theta = 3N_e\mu$, where N_e is the species effective population size for autosomal loci and μ is the neutral mutation rate, assuming that strong sexual selection is absent, such that the effective population size for males is similar to that for females (Caballero 1995; Nagylaki 1995, 1998). The $\hat{\theta}$ values can thus be used to estimate N_e if μ is known (Kreitman 1983). The neutral mutation rate μ is not an accurately known parameter, even though there are ways of estimating it from different types of data (see *Discussion*).

If we assume that the neutral mutation rate is approximately the same for all *Drosophila* species, then the level of silent site diversity can be used as an indicator of effective population size. The silent site diversity value of 0.0042 for *D. miranda* is much less than the average silent site diversity for the sibling species *D. pseudoobscura* (Moriyama and Powell 1996), consistent with the observation that *D. miranda* is much rarer in collections from the wild than is *D. pseudoobscura* (Noor 1995). When compared with the estimates from *D. melanogaster*, whose effective population size is often assumed to be in the millions (Kreitman 1983; Riley, Hallas, and Lewontin 1989), this is about half of the average silent site diversity over all genes studied, but close to the average for X-linked loci for this species (Moriyama and Powell 1996). From this result, it can be concluded that the effective population size of *D. miranda* is probably smaller than that of *D. pseudoobscura* but is comparable to that of *D. melanogaster*.

Recombination at the *per* Locus

We localized *per* by in situ hybridization to the XL chromosome arm of *D. miranda*, corresponding to the middle of band 46 in the drawn map of Das et al. (1982). This location is fairly near the base of the chromosome. We were concerned about the possibility of the existence of centromeric suppression of crossing over, as is found in *D. melanogaster* (Powell 1997, p. 362). If this were true, then *per* would reside in a low-recombination environment and hence might exhibit reduced genetic variation compared with genes in regions with high levels of recombination (Moriyama and Powell 1996). One point worth mentioning on this aspect is that the *per* locus in *D. pseudoobscura* showed less variation than did other loci surveyed in this species (Wang and Hey 1996). However, the reduction in the variability can be to a certain extent attributed to the smaller effective population size of an X-linked gene, and the authors did not find the reduction in the diversity to depart significantly from neutrality when comparisons were made with other genes using the HKA test (Wang and Hey 1996). Moreover, a recent study of chromosome 2 in *D. pseudoobscura* found little evidence of centromeric suppression of recombination (Hamblin and Aquadro 1999). Since the X chromosome in *D. melanogaster* shows less centromeric suppression of crossing over than the autosomes (Ashburner 1989, chapter 11), this suggests that there may be little or no reduction in crossing over near the centromere of the X chromosome in *D. miranda*.

We investigated the presence of recombination at the *per* locus by several different approaches. We first used Hudson and Kaplan's (1985) method to estimate the number of recombination events among the variants detected. This method is based on the presence of all four possible haplotypes in the polymorphic sites, and it detected a minimum of five different recombination events within the region investigated. We also estimated the population recombination parameter γ (Hey and Wakeley 1997), which was 0.016 per base. This value is comparable to the estimated γ value for *Adh* and *white* in *D. melanogaster* (table 2 of Hey and Wakeley 1997). In the presence of population subdivision, as seems to be the case for the *per* locus of *D. miranda*, γ tends to be underestimated (Hey and Wakeley 1997), so our estimate is at least a conservative one. When the ratio of γ to $\hat{\theta}$ is used to estimate the frequency of recombination events per mutation event per base in the population, we obtain a value close to 4. This value is again comparable to the values obtained from *Adh* and *white* in *D. melanogaster* (table 2 of Hey and Wakeley 1997).

We also examined the relationship between the linkage disequilibrium estimates for segregating sites and the distance between them. With recombination, this relationship should be negatively correlated (Schaeffer and Miller 1993). This was tested by using Spearman's rank correlation between a measure of linkage disequilibrium and physical distance between markers (Schaeffer and Miller 1993; Awadalla and Charlesworth 1999). The distances between polymorphic sites were randomized by generating a large number of data sets in which the distances between polymorphic sites were randomly assigned from the actual distances between these sites. For each randomization, the rank correlation between distance and a measure of linkage disequilibrium, r^2 (Hill and Robertson 1968), was calculated. The value from the actual sequence data was compared with the distribution of values from the randomized sets of data. We performed this test for data sets including and excluding the singleton variants. In both cases, the probability of observing the actual value of correlation was lower than 0.008. The results of the above investigation suggest that there is a considerable amount of recombination in the region surveyed, based on the polymorphism data collected.

Finally, there was no particular evidence from the variant sites detected in this study of *per* to strongly imply any recent selective sweep that might have acted on a gene in a low recombination environment (Brav-

Table 5
Polymorphism Summary for the *Lcp* Loci

Gene	No. of Alleles	Sequenced Region (bp)	Segregating Sites		Total[a]		Silent Sites[a]	
			Silent	Replacement	π/Site	$\hat{\theta}$/Site	π/Site	$\hat{\theta}$/Site
Lcp1-neo-X	12	765	2	2	1.06	1.74	1.06	1.50
Lcp1-neo-Y	12	765	0	1	0.22	0.44	0	0
Lcp3-neo-X	12	535	3	0	1.39	1.86	2.67	3.56
Lcp3-neo-Y	12	340	1	0	0.97	0.89	2.69	2.94

[a] All values $\times 10^{-3}$.

erman et al. 1995; Simonsen, Churchill, and Aquadro 1995). None of several different statistical tests of neutrality performed (Tajima's *D* statistic, the MacDonald/Kreitman test, and the Fu/Li test) detected any significant deviation from the neutral model for this locus (results not shown).

Divergence Estimates

We used all of the available *per* sequences from *D. pseudoobscura* and *D. miranda* to estimate the divergence between the two species. When all of the *per* alleles from these two species were aligned, no shared polymorphic sites were found, consistent with the idea that all of the ancient polymorphic sites have disappeared since the speciation event and that there is essentially no gene flow between the species. It is in fact known that the reproductive isolation between these two species is complete and that there are no fertile hybrids in the cross between them (Dobzhansky and Tan 1936).

Using all of the available *per* alleles of *D. pseudoobscura* reported by Wang and Hey (1996), we obtained an estimate of 0.044 per site for the net divergence (group divergence after removing the within-species variation component; Nei 1987, p. 276) between *D. pseudoobscura* and *D. miranda* for synonymous sites. Given that this value is small, such that multiple hits can be ignored, we used it as an approximation for the synonymous-site divergence. Using 1.2×10^{-8} per site per year as an estimate of the mean rate of synonymous site substitution for *Drosophila* (Wang and Hey 1996), this yields an estimated total divergence time (twice the time since the time of speciation, assuming a constant rate of evolution for both lineages since speciation) of 3.7 Myr. Wang, Wakeley, and Hey (1997) provided an estimate of divergence time of 5.25 Myr between these species based on the synonymous-site divergence of *Hsp82*. These two estimates are substantially larger than Russo, Takezaki, and Nei's (1995) estimate of 2.2 Myr, based on the third-codon-position divergence of *Adh* sequences. The disparities probably reflect the heterogeneity of the rate of evolution between different loci (Powell 1997, p. 369). There is some evidence that the rate of evolution of the *Adh* gene was slower in the *D. pseudoobscura* cluster than is typical for *Drosophila* (Russo, Takezaki, and Nei 1995). The net differences between *D. miranda* and *D. persimilis* and between *D. miranda* and *D. pseudoobscura bogotana* were also calculated from available *per* alleles by the same method. They were slightly higher than the value for the *D. miranda* and *D. pseudoobscura* comparison (0.054 and 0.067, respectively).

The Levels of Variation at the Neo-X and Neo-Y *Lcp* Loci

The organization of the *Lcp* loci on the neo-X and neo-Y chromosomes of *D. miranda* is described by Steinemann and Steinemann (1993). In a series of investigations of both functional and genomic analyses (Steinemann and Steinemann 1992, 1993; Steinemann, Steinemann, and Lottspeich 1993), the unusual structural and functional changes of the region encoding the *Lcp* cluster (including *Lcp1, Lcp2, Lcp3,* and *Lcp4*) of the neo-Y chromosome of *D. miranda* have been reported. Specifically, the neo-Y chromosome region of the *Lcp* cluster had several insertions of repetitive sequences and transposable elements that are absent in the neo-X counterpart (Steinemann and Steinemann 1993). Analyses of gene function further revealed that while the *Lcp* genes encoded by the neo-X chromosome are expressed normally, this is not the case for the neo-Y homologs. Only the *Lcp3* gene neo-Y allele was shown to display a reduced level of activity at the protein level; the other members of the *Lcp* cluster were completely inactive (Steinemann and Steinemann 1992; Steinemann, Steinemann, and Lottspeich 1993).

We surveyed a total of 1,300 bp of neo-X linked alleles and 1,105 bps of neo-Y linked alleles from *Lcp1* and *Lcp3*. We found a total of seven polymorphic sites among the neo-X alleles and only two from the neo-Y alleles (table 2). There were no shared polymorphic sites among the two neo-sex-chromosome alleles from the two loci we studied. It is, of course, expected that there would be no shared polymorphic sites that arose before the origin of the neo-sex chromosomes because of the probable unique origin of the neo-Y chromosome and its passage through males only, which lack crossing over (Powell 1997, p. 329).

The estimates of nucleotide variation for the neo-X linked and the neo-Y linked alleles at *Lcp1* and *Lcp3* are displayed separately in table 5. Both of the neo-sex chromosomes showed much lower levels of variation than the *per* locus. When estimated as the nucleotide site diversity (π), the level of reduction was anywhere between 3-fold (*Lcp3*-neo-X) and 20-fold (*Lcp1*-neo-Y). When the *Lcp1* and *Lcp3* loci were pooled, there was half the number of segregating sites for the neo-X linked alleles compared with *per*, whereas the total lengths of the genes compared are similar. The reduction of vari-

Table 6
Replacement-Site and Synonymous-Site Divergence

	Neo-X–*pseudo-obscura*	Neo-Y–*pseudo-obscura*	*P* value (see text)
Lcp1			
K_a.....	0.0036	0.0153	
K_S.....	0.0077	0.0168	
Lcp3			
K_a.....	0.0051	0.0179	
K_S.....	0.0429	0.0184	
Lcp1 + *Lcp3*			
K_a.....	0.0042	0.0173	0.012
K_S.....	0.0221	0.0164	NS

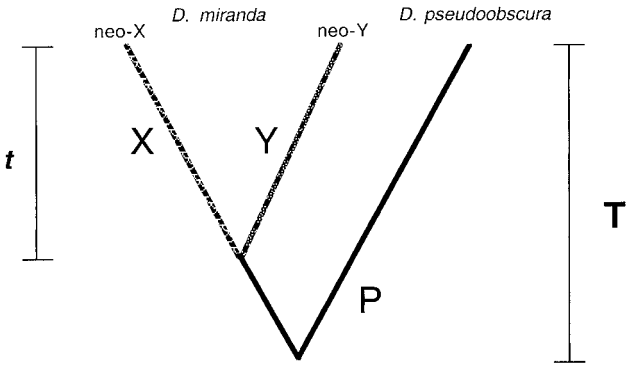

FIG. 2.—Phylogenetic relationship between the three alleles surveyed (*Drosophila miranda* neo-X and neo-Y and *Drosophila pseudoobscura*). *t* is the divergence time between the two neo-sex alleles, and *T* is the divergence time between *D. pseudoobscura* and *D. miranda.* X, Y, and P represent neo-X, neo-Y, and *D. pseudoobscura* alleles, respectively.

ation was more apparent for the neo-Y linked alleles. There was one replacement site polymorphism for the *Lcp1*-neo-Y, and one synonymous site polymorphism for the *Lcp3*-neo-Y. A reduction of the level of the variation at the neo-Y linked loci is predicted by theories of Y-chromosome degeneration (see *Discussion*). The apparent reduction of variation at the neo-X linked loci may have some implications in relation to the evolution of dosage compensation of the newly evolving X chromosomes (see *Discussion*). When the HKA test was performed using a *D. pseudoobscura* allele as the outgroup, no significant difference between *per* and the *Lcp* genes was detected, however, reflecting the low divergence between species for *Lcp1* (table 6).

Unlike the case of the *per* locus, there was little evidence for population subdivision for any of the neo-sex linked alleles. However, this may be due to the fact that there are only a small number of variant sites in the sample. Survey of a larger region may provide a more informative data set in this regard. We pooled the samples from British Columbia and the other samples for the purpose of present analyses. The diversity estimates obtained this way are likely to be upwardly biased if such a pattern of population subdivision as found is in *per* locus exists, as the presence of population subdivision increases the internal branch lengths of gene genealogy (Takahata 1991; Tajima 1993). This will make our interpretation of the low diversity at the *Lcp* loci conservative (see above).

Divergence of the Protein-Coding Sites Among the Neo-X and Neo-Y Alleles

K_a and K_s were estimated between the neo-X alleles and *D. pseudoobscura* and between the neo-Y alleles and *D. pseudoobscura* (fig. 2) using the longest consensus sequences. This comparison is informative, since the branches of the phylogenies of these alleles following the split between neo-X and neo-Y are equivalent in length. The results are presented in table 6. It is evident that there is a higher rate of nonsynonymous-site substitution between a neo-Y allele and the corresponding *D. pseudoobscura* allele than there is between a neo-X allele and *D. pseudoobscura.* For synonymous substitutions, *Lcp1* and *Lcp3* showed somewhat different patterns. For *Lcp1*, the same relationship as that for the replacement-site substitutions holds true, but K_s for *Lcp3* between neo-X and *D. pseudoobscura* was about twofold higher than that for neo-Y and *D. pseudoobscura.* Overall, the numbers were not as different as for K_a. When the two *Lcp* loci were combined, K_a was about four times as high for the neo-Y and *D. pseudoobscura* comparison as for the comparison involving neo-X, while K_s did not appear to be different.

To see whether the observed pattern of higher nonsynonymous-site substitutions for the neo-Y/*D. pseudoobscura* comparison was statistically significant, the following simulations were performed using the program K-Estimator, version 4.4 (see *Materials and Methods*). A ''hybrid'' sequence of neo-X and neo-Y alleles for both loci was made by combining the two alleles together as one sequence. Then the K_a and K_s values were estimated between the hybrid sequence and the *pseudoobscura* allele, which in this case was simply twice the length of the original one. These values were 0.0107 and 0.0192 for replacement sites and synonymous sites, respectively. We used these values as a common measure of the divergence rates for the neo-sex-chromosome alleles compared with the *pseudoobscura* alleles. From these estimated K_a and K_s values, Poisson distributions of the numbers of synonymous and replacement substitutions were generated. Using these distributions, pseudorandom coding sequences were generated with the same amino acid composition, the same GC percentage, and the same estimated transition/transversion ratios. For each iteration, the values of K_a and K_s were estimated, and a null distribution of 10,000 replicates was generated, as for the observed sequences. From this analysis, the *P* value for obtaining the observed K_a values for the neo-Y/*pseudoobscura* comparison was 0.012, supporting the conclusion that there is an accelerated rate of replacement site substitution on the branch between the neo-Y allele and the *D. pseudoobscura* allele. When only the neo-X allele was used to generate the null distribution, the probability of obtaining the observed K_a value for the neo-Y and *pseudoobscura* comparison was <0.0001.

Table 7
Numbers of Each Type of Substitution on the X, Y, and P Branches of Figure 2

	Lcp1			*Lcp3*			*Lcp1* + *Lcp3*		
Branch	*R*	*S*	*NC*	*R*	*S*	*NC*	*R*	*S*	*NC*
X	0	0	1	0	4	0	0	4	1
Y	4	1	0	3	1	1	7	2	1
P	1	1	8	1	0	0	2	1	8

Note.—See text for assignment. R = replacement sites; S = synonymous sites; NC = noncoding sites.

Evidence for Accelerated Amino Acid Substitutions at the Neo-Y Linked Loci

To further investigate whether this difference in replacement substitution rates is caused by accelerated evolution in the neo-Y lineage, the following analysis was performed. First, using the consensus sequence from each locus, alignments of the three sequences (the *D. pseudoobscura* allele, the neo-X allele, and the neo-Y allele) were made. Then, each substitution could be classified as P (the substitution occurred on the *pseudoobscura* branch of the tree connecting the three sequences), X (it occurred on the neo-X branch), or Y (it occurred on the neo-Y branch) (see fig. 2). The substitutions occurring on the P portion contained substitutions that occurred on element C of both *D. pseudoobscura* and *D. miranda* before its fusion with the Y chromosome to form the neo-Y chromosome. Considering the relatively short times separating these three sequences, it is reasonable to assume that there have been no multiple hits and/or back mutations. As a result, we are able to determine unambiguously the numbers of different types of substitutions (replacement substitutions, synonymous substitutions, and noncoding substitutions) on each branch of the lineage. The results are shown in table 7.

It is immediately apparent from table 7 that there are more replacement substitutions on the Y branch than on the X or P branches. In fact, there were no replacement-site substitutions on the X branch for both the *Lcp1* and *Lcp3* loci, while four and three replacement substitutions occurred on the Y branch for these genes, respectively. It is particularly interesting to note that eight noncoding substitutions occurred on the P branch for *Lcp1*, as opposed to none and one for the Y and X branches, respectively. The pattern of more silent and noncoding changes on the P branch was not evident for *Lcp3,* which is probably explained by the fact that the portion of the *Lcp3* gene analyzed here contains only 41 bp of noncoding DNA, compared with 343 bp for *Lcp1.*

Overall, table 7 suggests more replacement relative to synonymous and noncoding substitutions on the Y branch compared with the P and X lineages. To investigate this, we constructed a 2×2 table by pooling the synonymous and noncoding substitutions as ''silent'' changes and performed a test for equality of the ratio of silent to replacement changes on the Y and P branches for *Lcp1* and *Lcp3* together. There is a significant difference ($P < 0.05$ by Fisher's exact test). Similarly, there is a highly significant difference between X and Y ($P < 0.01$).

The information obtained in table 7 can also be used to estimate the divergence time between the two neo-sex chromosomes using the *Lcp1* noncoding sequence, assuming it is relatively free of selective pressure. If x is the fraction of the total tree length occupied by the X and Y branches together, the probability of either the observed result or no X and Y substitutions, is $x^9 + 8(1 - x)x^9$, conditioning on a total of 9 substitutions. The upper 95% confidence bound on x from this expression is 0.58. If the split between *D. miranda* and *D. pseudoobscura* occurred at time T in the past and the length of time from the common ancestor of the neo-X and neo-Y chromosomes to the present is t, then $x = (2T - t)/(2T + t)$, i.e., $t = 2T(1 - x)/(1 + x)$. Using an estimate of 2 Myr for T from the *per* locus data (see above), we can conclude that t is approximately equal to 1.06 Myr, i.e., the neo-X/Y split could not have occurred more than about one million years ago. The silent-site divergence (K_s) between the two neo-sex linked alleles is also informative in this regard. When the two *Lcp* loci were combined, K_s was 0.0299. Using a value of 1.2×10^{-8} per site per year for the rate of silent evolution (see above), the total time since divergence of the two neo-sex alleles is 2.5 Myr; in other words, $t =$ 1.25 million.

Discussion

Several models of the early steps of degeneration based on the effect of selection on properties of a nonrecombining chromosome have been proposed to explain the process of Y-chromosome degeneration (Charlesworth 1978, 1996; Rice 1987, 1996; Orr and Kim 1998; McVean and Charlesworth 2000); selection for dosage compensation and inactivation of Y-chromosome genes has also been postulated by these authors to explain the completion of the process. All of these models involve a large reduction in the effective population size of the Y chromosome relative to its numerical abundance, and hence predict a reduction in DNA sequence variation and apparent relaxation of selection at weakly selected sites located on a newly evolving Y chromosome (Charlesworth 1996; McVean and Charlesworth 2000). Studies of sequence variation and evolution on a neo-Y chromosome of the kind presented here should therefore provide a better understanding of the dynamics of the early stages of degeneration of a nonrecombining chromosome. They also have the potential to discriminate among alternative models (Charlesworth 1996).

Reduction in the Effective Population Size and Muller's Ratchet

One model that has been widely discussed is ''Muller's ratchet,'' which involves the stochastic loss of the class of chromosomes carrying the smallest number of deleterious mutations (Muller 1964; Haigh 1978). In the absence of recombination and back mutation, the lost chromosome class cannot be regenerated. The Y chromosome can gradually lose its genetic activity due

to the fixation of a deleterious mutation with each turn of the ratchet (Rice 1987; Charlesworth and Charlesworth 1997). The main problem in envisioning this process as generally responsible for the degeneration of the Y chromosomes lies in the fact that a major constraint on the speed of the ratchet is the size of the least loaded class in the distribution of the number of mutations per chromosome, at least for haploid genomes such as Y chromosomes (Stephan, Chao, and Smale 1993; Charlesworth and Charlesworth 1997). This variable is determined by the product of the population size and the frequency of the mutation-free class in an infinite population at equilibrium under mutation and selection (Haigh 1978; Stephan, Chao, and Smale 1993; Charlesworth and Charlesworth 1997).

A rough idea of the population size of *D. miranda* can be obtained as follows. We estimated the nucleotide site variation parameters $\hat{\theta}$ and π for the X-linked locus *period* in *D. miranda*. The silent-site π estimate for this gene is 0.43% (table 2), in line with previous estimates of DNA diversity in this species. This estimate is comparable with that for X-linked genes in *D. melanogaster* (Moriyama and Powell 1996), suggesting that *D. miranda* has a comparable effective population size, assuming that mutation rates are similar in the two species. Sharp and Li (1989) provide an estimate of 1.6×10^{-8} per year as an upper limit for the synonymous-substitution rate, assuming an estimated divergence time of 40 Myr between the *melanogaster* and *obscura* species groups. This value is an average for the two lineages, which probably differ in their generation times. Assuming seven generations per year as an average for the two groups, the mutation rate per generation is approximately 2×10^{-9} on the assumption of neutrality, yielding an estimate of 2 million for N_e if the expected diversity for an X-chromosomal locus is $3N_e\mu$, as expected if there is no strong sexual selection (Caballero 1995; Nagylaki 1995). Recent studies suggest the action of natural selection on synonymous sites in *Drosophila* (Akashi 1994, 1995; Akashi and Schaeffer 1997), such that the use of synonymous-site diversity as an estimate of neutral diversity could result in an underestimation of $N_e\mu$ from θ or π. On the other hand, the estimate of μ from rates of silent substitution between species is likely to be an underestimate if there is selection for preferred codon usage, resulting in a slightly larger bias in the opposite direction (McVean and Charlesworth 1999). Given the relatively weak intensity of selection on silent sites suggested by the *Drosophila* data, the net effect on the above estimate of N_e for *D. miranda* is likely to be fairly small and certainly will not affect its order of magnitude.

The effective population size for the neo-Y chromosome of *D. miranda* is thus likely to be on the order of 500,000. If standard values for the deleterious mutation rate per chromosome arm and the mean selection coefficient against a heterozygous deleterious mutation in *Drosophila* are used (Crow and Simmons 1983), the rate of movement of the ratchet in a population of this size is expected to be negligible over the span of a few million generations that has been involved in the degeneration of the *D. miranda* neo-Y (Stephan, Chao, and Smale 1993; Gordo and Charlesworth 2000). At first sight, this rules out the ratchet as a contributory factor. But Gessler (1995) and Gessler and Xu (1999) have recently argued that a wide distribution of mutational effects on fitness, such as an exponential distribution, could result in the existence of a group of mutant genotypes with a sufficiently small reduction in fitness below wild type that the size of the least loaded class will be small enough to permit the ratchet to operate, even in very large populations.

Given our current uncertainty concerning the magnitude of the deleterious mutation rate and the nature of the distribution of mutational effects on fitness (Fry et al. 1999), it is difficult to evaluate this possibility, and so the above calculation is inconclusive. It does, however, seem to rule out the possibility that the ratchet has been operating to cause the accumulation on the neo-Y of mutations with major effects on fitness when homozygous (lethals or severely detrimental mutations). The mutation rate per neo-Y chromosome for such mutations is known from data on *D. melanogaster* (Crow and Simmons 1983; Fry et al. 1999) to be too small (0.005 at most; the estimates are ≤ 0.01 for chromosome 2 of *D. melanogaster*, which is at least twice the size of the neo-Y chromosome of *D. miranda*) for the ratchet to be plausible with a population size of 500,000, following the argument of Charlesworth (1996) and assuming that such major mutations have heterozygous fitness effects on the order of 1% or more, as suggested by direct measurements (Crow and Simmons 1983).

Data on DNA sequence variation can, however, shed light on the operation of the ratchet in quite a different way. One aspect of the data which may be useful for evaluating the possible operation of the ratchet mechanism is the time back to the most recent common ancestor of the sample of neo-Y alleles, T_Y. If the ratchet is the main force causing the degeneration of the Y chromosome, then T_Y cannot exceed the time since the last turn of the ratchet, since clicks of the ratchet are accompanied by fixations of individual haplotypes in the whole population (Rice 1987; Charlesworth and Charlesworth 1997). The upper limit for the number of turns of the ratchet is thus inversely proportional to the ratio of T_Y to the time since the divergence of the neo-Y and neo-X chromosomes. Estimation of T_Y however, depends on knowledge of the shape of a neutral gene genealogy under the ratchet, which is currently poorly understood. With this limitation in mind, we provide two kinds of estimates for the number of turns of the ratchet that are likely to be upwardly biased.

First, we use the standard neutral genealogy as an approximation; this is known to be fairly accurate under the background selection model of deleterious mutations, in the absence of movement of the ratchet (Charlesworth et al. 1995). Under the neutral coalescent model, the expected time to the most recent common ancestor of a set of alleles is as follows (Hudson 1990; Griffiths and Tavaré 1996):

$$E\{T_Y\} = 4N_e\left(1 - \frac{1}{n}\right) = \frac{\theta}{\mu}\left(1 - \frac{1}{n}\right), \quad (1)$$

where n is the number of sequences in the sample, μ is the neutral mutation rate, and $\theta = 4N_e\mu$. The expected synonymous-site divergence between the two neo-sex chromosomes (k_s) is related to the time since the origin of the neo-Y chromosome by

$$T_{DIV} = \frac{E\{k_s\}}{2\mu} \quad (2)$$

The number of turns of the ratchet since the divergence time is $\leq T_{DIV}/T_Y = k_s/2\theta(1 - 1/n)$, which can be estimated using estimates of k_s and θ. In order to be conservative, we obtained a lower bound for θ that is compatible with the observed number of segregating sites by applying Kreitman and Hudson's (1991) recursion formula for the neutral case, setting the probability of obtaining two segregating sites at $\leq 5\%$. Using this value of $\hat{\theta} = 2.26 \times 10^{-4}$ and the estimated 3% divergence between the two neo-sex chromosomes, we obtain an estimate of 72 clicks of the ratchet since the origin of the neo-Y chromosome.

Another way of estimating the age of the neo-Y alleles is to assume that the fixation of a deleterious mutation caused by a turn of ratchet results in a ''star phylogeny,'' as in the case of a selective sweep (Berry, Ajioka, and Kreitman 1991). Although this assumption is not necessarily correct, it leads us to an even more conservative estimate of the depth of the gene genealogy than the above procedure. Under this assumption, T_Y can be estimated from the number of segregating sites observed in the sample of n alleles, since the expectation of the latter is now

$$S = nm\mu T_{MRCA}, \quad (3)$$

where m is the length of sequences surveyed for the neo-Y alleles (1,105 bp).

Combining this with the above formula for T_{DIV}, we thus have

$$\frac{T_{DIV}}{E\{T_Y\}} = \frac{E\{k_s\} \times m}{2S}. \quad (4)$$

We tried to provide an upwardly biased estimate of this ratio in the following way: assuming a Poisson distribution for the number of segregating sites (Berry, Ajioka, and Kreitman 1991), we obtained a 5% probability lower bound of 0.8 for the expected number of segregating sites in the sample. Using the 3% silent-site divergence between the two neo-sex alleles, this led to an estimate of at most 250 turns of the ratchet. With a divergence time between neo-X and neo-Y chromosomes of 10^6 years (see *Results*) and seven generations per year, this is approximately equivalent to one turn of ratchet per 28,000 generations, whereas the previous estimate involves one turn per 97,200 generations. The range of current estimates of mutational parameters for *Drosophila* mutations with small fitness effects is in principle compatible with both of these rates of movement of the ratchet in a haploid population of around 500,000 individuals (Gessler 1995; Gessler and Xu 1999; Gordo and Charlesworth 2000).

The implications of these possible rates of movement of the ratchet are difficult to evaluate. Given that there are probably over 2,000 genes on element C (Ashburner et al. 1999), it is at least clear that the loss of genetic activity of the majority of neo-Y genes in *D. miranda* cannot be caused directly by fixation due to the ratchet of loss-of-function mutations at each locus. Another possibility is that there has been selection to reduce the activity of genes on the neo-Y chromosome relative to that of their homologs on the neo-X chromosome in response to the decline in mean fitness of the neo-Y chromosomes by the fixation of deleterious mutations with minor effects (Charlesworth 1996; Rice 1996). The strength of selection for such a reduction is a function of the size of the fitness loss caused by such fixation. One of the conditions for a relatively rapid rate of movement of the ratchet in a very large population is a very small harmonic mean selection coefficient for the mutations that cause the ratchet to move (Gessler 1995; Gessler and Xu 1999; Gordo and Charlesworth 2000). This means that the mutations that accumulate fastest in the early stages of the ratchet are probably those with very small effects, such that their carriers would not suffer much from their presence. For example, if we assume a mean selection coefficient of 10^{-4}, as is reasonable for transposable-element insertions (Charlesworth et al. 1992), which are abundant in the neo-Y chromosome of *D. miranda* (Steinemann and Steinemann 1998), the net fitness decline from the mutations that accumulate for the first 100 turns is only 1%. It is unclear whether this could contribute significantly to the chromosomewide genetic degeneration that is seen on the neo-Y chromosome of *D. miranda*, since other factors, such as an accelerated rate of fixation of detrimental mutations due to background selection (Charlesworth 1996) or weak Hill-Robertson effects (McVean and Charlesworth 2000), may be able to cause a much more substantial fitness decline.

Reduction in the Within-Population Diversity of the Neo-Y Chromosome Predicted by Other Theories

Other theories of Y-chromosome degeneration also entail a reduction in the effective population size of the proto-Y chromosome. The hitchhiking model of Y-chromosome degeneration (Rice 1987) postulates the fixation of deleterious alleles in association with the fixation of favorable mutations on the Y chromosome. The selective sweeps induced by such fixation events on a nonrecombining chromosome are well known to reduce neutral diversity at other loci on the same chromosome (Maynard Smith and Haigh 1974; Kaplan, Hudson, and Langley 1989). The background selection model invokes an opposite form of hitchhiking: the elimination of deleterious mutations. When a nonrecombining chromosome carries one or more strongly deleterious mutations, it will be rapidly eliminated from the population; only chromosomes free of such mutations contribute to the future ancestry of the population, such that the ef-

fective population size depends largely on the size of the mutation-free class (Charlesworth, Morgan, and Charlesworth 1993). As a result, the effectiveness of selection becomes relatively weak, causing slightly deleterious mutations to experience accelerated rates of substitution due to drift, and favorable mutations to be fixed less frequently (Charlesworth 1994), contributing to Y chromosome degeneration (Charlesworth 1996; Orr and Kim 1998). This is accompanied by a reduction in the level of within-population genetic variation for nearly neutral variants. Hill-Robertson interference between weakly selected sites such that selection coefficients are on the order of $1/N_e$ (Comeron, Kreitman, and Aguadé 1999; McVean and Charlesworth 2000) can also result in a substantial reduction of within-population diversity for a nonrecombining chromosome the size of the neo-Y chromosome (McVean and Charlesworth 2000).

The observation that there is very little nucleotide variation at the neo-Y linked alleles of the *Lcp* loci is consistent with these theories. From the 1,100 bp surveyed in this study, only two sites were found to be polymorphic. The two polymorphic sites found were at very low frequencies (one and two variants, respectively, for each locus). Since there is little evidence for selective constraints on the amino acid sequences of the *Lcp* alleles on the neo-Y chromosome (see above), it is reasonable to compare the variability at both silent and replacement sites on the neo-Y chromosome with silent sites at these loci on the neo-X chromosome (a total of 441 sites for *Lcp1* and a total of 278 sites for *Lcp3*), for which five segregating sites are found. On the null hypothesis that the Y chromosome has one third the population size of the X chromosome (i.e., there is no strong effect of sexual selection or hitchhiking), we would expect the proportion of all *Lcp* variants that are found on the neo-Y chromosome to be $0.25 \times 1{,}100/(0.25 \times 1{,}100 + 0.75 \times 719) = 0.338$, compared with the observed value of 0.286. Coalescent simulations of gene trees, in which the two *Lcp* loci on the neo-X chromosome are treated as independent of each other and of the neo-Y loci (assumed to be completely linked), and where a total of nine segregating sites are assigned to the three simulated trees in proportion to their lengths, show that this difference is nonsignificant: the probability of observing a proportion of the segregating variants on the neo-Y chromosome as small as or smaller than that observed is 0.55 based on 1,000 replicates.

However, as noted above, there is a noticeably lower level of variation on the *Lcp* neo-X alleles compared with *per*, although this is not statistically significant with an HKA test. One possible explanation for the reduction of variability for the neo-X linked genes has to do with the evolution of dosage compensation. Current models of dosage compensation in *Drosophila* X chromosomes often invoke the acquirement of *cis*-acting dosage compensation determinants (Bashaw and Baker 1996). The exact nature of such determinants is not yet known. One such model suggests the presence of many ''chromatin entry sites'' on the dosage-compensated chromosome that can then attract the dosage compensation protein complex such that the adjacent area in the chromosome can also be regulated (Kelly et al. 1999). A new X chromosome such as the neo-X chromosome of *D. miranda* therefore should be in the process of acquiring such *cis*-acting determinants as an adaptive response to the degeneration of its counterpart. This scenario is in accord with the observations that the neo-X chromosome of *D. miranda* is partially associated with the same dosage compensation machinery that is used for the regulation of the real X chromosome (Bone and Kuroda 1996; Marín et al. 1996; Steinemann, Steinemann, and Turner 1996). The evolutionary acquisition of *cis*-acting determinants necessary for dosage compensation could lead to local selective sweeps which might reduce the level of variation for many genes on the neo-X chromosome.

Further data are needed to test this possibility. If this were the case, it would be more appropriate to compare *Lcp* with *per*. A comparison of variability at the *Lcp* loci with silent variation at *per* on the same lines as above shows that the expected proportion of segregating sites on the neo-Y chromosome is 0.377, compared with an observed of $2/(2 + 9) = 0.182$. Coalescent simulations of the type just described show that the probability of this is 0.08 on the null hypothesis, such that, again, the reduction in variability on the neo-Y is not significant. However, the simulations also show that the observed level of variability on the neo-Y chromosome, compared with *per*, is consistent with a substantial reduction in the effective population size of the neo-Y chromosome, which could approach a 10-fold reduction over that assumed by the null hypothesis. Again, further data on neo-Y chromosome variability are needed.

An interesting finding of this study is the observed excess of amino acid substitutions on the neo-Y–linked alleles relative to neo-X linked alleles when *D. miranda* is compared with its relatives (table 7). When the two *Lcp* genes were combined, the rate of amino acid replacement-site substitutions was more than four times as high for the comparisons involving the neo-Y genes (see above). Such an acceleration is not evident for synonymous sites. The ratio of nonsynonymous-site divergence to synonymous-site divergence was close to 1 for *Lcp1* and *Lcp3*. Our analysis showed that this effect was caused by an elevated number of replacement substitutions on the neo-Y lineage, following the split between the neo-X and neo-Y chromosomes. This is in accord with an extremely reduced effective population size of the neo-Y chromosome, allowing slightly deleterious substitutions to accumulate. Alternatively, the acceleration of amino acid substitution may have been due to a lack of functional constraints on the *Lcp*-coding region, since a previous study has shown that the neo-Y alleles are almost nonfunctional at the level of protein expression, presumably because the genes have been down-regulated (Steinemann, Steinemann, and Lottspeich 1993). On the other hand, the neo-X alleles do not show any evidence of accelerated evolution at the amino acid level, as might be expected if the degeneration of the Y chromosome is an evolutionary response to a more rapid adaptive evolution of the X chromosome (Orr and Kim 1998).

Acknowledgments

We wish to thank Dr. M. Steinemann for sharing the sequences of the neo-X and neo-Y alleles of the loci studied before publication. S.Y. wishes to thank R. R. Hudson and J. M. Comeron, S-C. Tsaur, C.-T. Ting, and P. Awadalla for their help in supplying computer codes and/or helpful discussions. This work was supported by NIH grant GM50355 to D. Charlesworth and B.C., and by NSF Doctoral Dissertation Improvement Grant DEB-9701098 to S.Y. B.C. is supported by the Royal Society.

LITERATURE CITED

AKASHI, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. Genetics **136**:927–935.

———. 1995. Inferring weak selection from patterns of polymorphism and divergence at ''silent'' sites in *Drosophila* DNA. Genetics **139**:1067–1076.

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

ANDERSON, W. W., F. J. AYALA, and R. E. MICHOD. 1977. Chromosomal and allozymic diagnosis of three species of *Drosophila. Drosophila pseudoobscura, D. persimilis,* and *D. miranda.* J. Hered. **68**:71–74.

ASHBURNER, M. 1989. *Drosophila* A laboratory handbook. Cold Spring Harbor Laboratory Press, New York.

ASHBURNER, M., S. MISRA, J. ROOTE et al. (27 co-authors). 1999. An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: the Adh region. Genetics **153**:179–219.

AWADALLA, P., and D. CHARLESWORTH. 1999. Recombination and selection at Brassica self-incompatibility loci. Genetics **152**:413–425.

BABCOCK, C. S., and W. W. ANDERSON. 1996. Molecular evolution of the sex-ratio inversion complex in *Drosophila pseudoobscura*: analysis of the esterase-5 gene region. Mol. Biol. Evol. **13**:297–308.

BASHAW, G. J., and B. S. BAKER. 1996. Dosage compensation and chromatin structure in *Drosophila.* Curr. Opin. Genet. Dev. **6**:496–501.

BERRY, A. J., J. W. AJIOKE, and M. KREITMAN. 1991. Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. Genetics **129**:1111–1117.

BONE, J. R., and M. I. KURODA. 1996. Dosage compensation regulatory proteins and the evolution of sex chromosomes in *Drosophila*. Genetics **144**:705–713.

BRAVERMAN, J. M., R. R. HUDSON, N. L. KAPLAN, C. H. LANGLEY, and W. STEPHAN. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics **140**:783–796.

BULL, J. J. 1983. Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park, Calif.

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

CHARLESWORTH, B. 1978. Model for evolution of Y chromosomes and dosage compensation. Proc. Natl. Acad. Sci. USA **75**:5618–5622.

CHARLESWORTH, B. 1991. The evolution of sex chromosomes. Science **251**:1030–1033.

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

———. 1998. Measures of divergence between populations and the effect of forces that reduce variability. Mol. Biol. Evol. **15**:538–543.

CHARLESWORTH, B., and D. CHARLESWORTH. 1997. Rapid fixation of deleterious alleles can be caused by Muller's ratchet. Genet. Res. **70**:63–73.

CHARLESWORTH, B., A. LAPID, and D. CANADA. 1992. The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. I. Element frequencies and distribution. Genet. Res. **60**: 103–114.

CHARLESWORTH, B., M. T. MORGAN, and D. CHARLESWORTH. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics **134**:1289–1303.

CHARLESWORTH, D., B. CHARLESWORTH, and M. T. MORGAN. 1995. The pattern of neutral molecular variation under the background selection model. Genetics **141**:1619–1632.

CLINE, T. W., and B. J. MEYER. 1996. Vive la difference: male vs females in flies and worms. Annu. Rev. Genet. **30**:637–702.

COMERON, J., and M. KREITMAN. 1998. The correlation between synonymous and nonsynonymous substitutions in *Drosophila*: mutation, selection or relaxed constraints? Genetics **150**:767–775.

COMERON, J., M. KREITMAN, and M. AGUADÉ. 1999. Natural selection on synonymous sites is correlated with gene length and recombination in *Drosophila*. Genetics **151**:239–249.

CROW, J. F., and M. J. SIMMONS. 1983. The mutation load in *Drosophila.* Pp. 1–35 *in* M. ASHBURNER, H. L. CARSON, and J. N. J. THOMPSON, eds. The genetics and biology of *Drosophila.* Academic Press, London.

DAS, M., D. MUTSUDDI, A. K. DUTTAGUPTA, and A. S. MUKHERJEE. 1982. Segmental heterogeneity in replication and transcription of the X_2 chromosome of *Drosophila miranda* and conservativeness in the evolution of dosage compensation. Chromosoma **87**:373–388.

DOBZHANSKY, T. 1935. *Drosophila miranda,* A new species. Genetics **20**:377–391.

DOBZHANSKY, T., and P. C. KOLLER. 1938. An experimental study of sexual isolation in *Drosophila.* Biol. Zenrtralbl. **58**: 589–607.

DOBZHANSKY, T., and C. C. TAN. 1936. Studies on hybrid sterility III. A comparison of the gene arrangement in two species, *Drosophila pseudoobscura* and *Drosophila miranda.* Z. Induk. Abstamm. Vererbungsl. **72**:88–114.

FRY, J., P. KEIGHTLEY, S. HEINNSOHN, and S. NUZHDIN. 1999. New estimates of the rates and effects of mildly deleterious mutation in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **96**:574–579.

GESSLER, D. D. 1995. The constraints of finite size in asexual populations and the rate of the ratchet. Genet. Res. **66**:241–253.

GESSLER, D., and S. XU. 1999. On the evolution of recombination and meiosis. Genet. Res. **73**:119–131.

GORDO, I., and B. CHARLESWORTH. 2000. The degeneration of haploid populations and the speed of Muller's ratchet. Genetics (in press).

GRIFFITHS, R. C., and S. TAVARÉ. 1996. Monte Carlo inference methods in population genetics. Math. Comput. Model. **23**: 141–158.

GUTTMAN, D., and D. CHARLESWORTH. 1998. An X-linked gene with a degenerate Y-linked homologue in a dioecious plant. Nature **393**:263–266.

HAIGH, J. 1978. The accumulation of deleterious genes in a population—Muller's ratchet. Theor. Popul. Biol. **14**:251–267.

HAMBLIN, M. T., and C. F. AQUADRO. 1999. DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: a study of the second chromosome. Genetics **153**:859–869.

HEY, J., and J. WAKELEY. 1997. A coalescent estimator of the population recombination rate. Genetics **145**:833–846.

HILL, W. G., and A. ROBERTSON. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. **38**:226–231.

HUDSON, R. R. 1990. Gene genealogies and the coalescent process. Pp. 1–44 *in* J. ANTONOVICS and D. FUTUYMA, eds. Oxford surveys in evolutionary biology. Oxford University Press, Oxford, England.

HUDSON, R. R., D. D. BOOS, and N. L. KAPLAN. 1992. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. **9**:138–151.

HUDSON, R. R., and N. L. KAPLAN. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111**:147–164.

HUDSON, R. R., M. KREITMAN, and M. AGUADÉ. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics **116**:153–159.

HUDSON, R. R., M. SLATKIN, and W. MADDISON. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics **132**:583–589.

JEGALIAN, K., and D. C. PAGE. 1998. A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature **394**:776–780.

KAPLAN, N. L., R. R. HUDSON, and C. H. LANGLEY. 1989. The ''hitchhiking effect'' revisited. Genetics **123**:887–899.

KELLY, R. L., V. H. MELLER, P. R. GORDADZE, G. ROMAN, R. L. DAVIS, and M. I. KURODA. 1999. Epigenetic spreading of the *Drosophila* dosage compensation complex from *roX* RNA genes into flanking chromatin. Cell **98**:513–522.

KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature **304**:412–417.

KREITMAN, M., and R. R. HUDSON. 1991. Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127**:565–582.

KYRIACOU, C. P., and J. C. HALL. 1994. Genetic and molecular analysis of *Drosophila* behavior. Adv. Genet. **31**:139–186.

LUCCHESI, J. C. 1994. The evolution of heteromorphic sex chromosomes. Bioessays **16**:81–83.

———. 1998. Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. Curr. Opin. Genet. Dev. **8**:179–184.

MACKNIGHT, R. H. 1939. The sex-determining mechanism of *Drosophila miranda.* Genetics **24**:180–201.

MCVEAN, G. A. T., and B. CHARLESWORTH. 1999. A population genetic model for the evolution of synonymous codon usage: patterns and predictions. Genet. Res. **74**:145–158.

———. 2000. The effects of Hill-Robertson interference between weakly selected sites on patterns of molecular evolution and variation. Genetics (in press).

MARÍN, I., A. FRANKE, G. I. BRADSHAW, and B. S. BAKER. 1996. The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes. Nature **383**: 160–163.

MAYNARD SMITH, J., and J. HAIGH. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. **23**:23–35.

MITTWOCH, W. 1967. Sex chromosomes. Academic Press, New York.

MONTGOMERY, E., B. CHARLESWORTH, and C. H. LANGLEY. 1987. A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. Genet. Res. **49**:31–41.

MORIYAMA, E. N., and J. R. POWELL. 1996. Intraspecific nuclear DNA variation in *Drosophila.* Mol. Biol. Evol. **13**: 261–277.

MULLER, H. 1964. The relation of recombination to mutational advance. Mutat. Res. **106**:2–9.

NAGYLAKI, T. 1995. The inbreeding effective population number in dioecious populations. Genetics **139**:473–485.

———. 1998. The expected number of heterozygous sites in a subdivided population. Genetics **149**:1599–1604.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

NOOR, M. 1995. Long-term changes in obscura group *Drosophila* species composition at Mather, California. Pan-Pac. Entomol. **71**:71–74.

NORMAN, R. A., and W. W. DOANE. 1990. Dosage compensation and dietary glucose repression of larval amylase activity in *Drosophila miranda.* Biochem. Genet. **28**:601–613.

OHNO, S. 1967. Sex chromosomes and sex-linked genes. Springer-Verlag, Berlin.

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

POWELL, J. R. 1997. Progress and prospects in evolutionary biology: the *Drosophila* model. Oxford University Press, New York.

RICE, W. R. 1987. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. Genetics **116**:161–167.

———. 1996. Evolution of the Y sex chromosome in animals. BioScience **46**:331–343.

RILEY, M., M. HALLAS, and R. LEWONTIN. 1989. Distinguishing the forces controlling genetic variation at the Xdh locus in *Drosophila pseudoobscura.* Genetics **123**:359–369.

ROZAS, J., and R. ROZAS. 1997. Dna SP version 2: a novel software package for extensive population genetic analysis. Comput. Appl. Biosci. **13**:307–311.

RUSSO, C. A. M., T. TAKEZAKI, and M. NEI. 1995. Molecular phylogeny and divergence times of drosophilid species. Mol. Biol. Evol. **12**:391–404.

SCHAEFFER, S. W., and E. L. MILLER. 1993. Estimates of linkage disequilibrium and the recombination parameter determined from segregating nucleotide sites in the alcohol dehydrogenase region of *Drosophila pseudoobscura.* Genetics **135**:541–552.

SHARP, P., and W. LI. 1989. On the rate of DNA sequence evolution in *Drosophila.* J. Mol. Evol. **28**:398–402.

SIMONSEN, K. L., G. A. CHURCHILL, and C. F. AQUADRO. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics **141**:413–429.

SNYDER, M., M. HUNKAPILLER, D. YUEN, D. SILVERT, J. FRISTROM, and N. DAVIDSON. 1982. Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. Cell **29**:1027–1040.

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. 1993. A duplication including the Y allele of Lcp2 and the TRIM retrotransposon at the Lcp locus on the degenerating neo-Y chromosome of *Drosophila miranda*: molecular structure and mechanisms by which it may have arisen. Genetics **134**:497–505.

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: X_2 and *neo-Y* of *Drosophila miranda*: I. comparison at the DNA sequence level. J. Mol. Evol. **43**: 405–412.

STEINEMANN, M., S. STEINEMANN, and B. M. TURNER. 1996. Evolution of dosage compensation. Chromosome Res. **4**:1–6.

STEPHAN, W., L. CHAO, and J. G. SMALE. 1993. The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. Genet. Res. **61**: 225–232.

STEPHAN, W., L. XING, D. A. KIRBY, and J. M. BRAVERMAN. 1998. A test of the background selection hypothesis based on nucleotide data from *Drosophila ananassae*. Proc. Natl. Acad. Sci. USA **95**:5649–5654.

STROBEL, E., C. PELLING, and N. ARNHEIM. 1978. Incomplete dosage compensation in an evolving *Drosophila* sex chromosome. Proc. Natl. Acad. Sci. USA **75**:931–935.

TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**: 585–595.

———. 1993. Measurement of DNA polymorphism. Pp. 37–59 *in* N. TAKAHATA and A. CLARK, eds. Mechanism of molecular evolution. Japan Scientific Societies Press, Sinauer, Japan.

TAKAHATA, N. 1991. Genealogy of neutral genes and spreading of selected mutations in a geographically structured population. Genetics **129**:585–595.

WAKELEY, J., and J. HEY. 1997. Estimating ancestral population parameters. Genetics **145**:847–855.

WANG, R.-L., and J. HEY. 1996. The speciation history of *Drosophila pseudoobscura* and close relatives: inferences from DNA sequence variation at the *period* locus. Genetics **144**: 1113–1126.

WANG, R., J. WAKELEY, and J. HEY. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics **147**:1091–1106.

WATTERSON, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7**:256–276.

WILSON, E. B. 1925. The cell in development and heredity. Macmillan, New York.

ZENG, L., J. COMERON, B. CHEN, and M. KREITMAN. 1998. The molecular clock revisited: the rate of synonymous vs. replacement change in *Drosophila*. Genetica **102–103**:369–382.

WOLFGANG STEPHAN, reviewing editor

Accepted January 11, 2000