DNA Sequence Variation and the Recombinational Landscape in *Drosophila pseudoobscura*: A Study of the Second Chromosome

Martha T. Hamblin and Charles F. Aquadro

Section of Genetics and Development, Cornell University, Ithaca, New York 14853

Manuscript received October 19, 1998
Accepted for publication July 1, 1999

ABSTRACT

The relationship between rates of recombination and DNA sequence polymorphism was analyzed for the second chromosome of *Drosophila pseudoobscura*. We constructed integrated genetic and physical maps of this chromosome using molecular markers at 10 loci spanning most of its physical length. The total length of the map was 128.2 cM, almost twice that of the homologous chromosome arm (3R) in *D. melanogaster*. There appears to be very little centromeric suppression of recombination, and rates of recombination are quite uniform across most of the chromosome. Levels of sequence variation (θ_W, based on the number of segregating sites) at seven loci (*tropomyosin 1*, *Rhodopsin 3*, *Rhodopsin 1*, *bicoid*, *Xanthine dehydrogenase*, *Myosin light chain 1*, and *ribosomal protein 49*) varied from 0.0036 to 0.0167. Generally consistent with earlier studies, the average estimate of θ_W at total sites is 1.5-fold higher than that in *D. melanogaster*, while average θ_W at silent sites is almost 3-fold higher. These estimates of variation were analyzed in the context of a background selection model under the same parameters of mutation rate and selection as have been proposed for *D. melanogaster*. It is likely that a significant fraction of the higher level of sequence variation in *D. pseudoobscura* can be explained by differences in regional rates of recombination rather than a larger species-level effective population size. However, the distribution of variation among synonymous, nonsynonymous, and noncoding sites appears to be quite different between the species, making direct comparisons of neutral variation, and hence inferences about effective population size, difficult. Tajima's *D* statistics for 6 out of the 7 loci surveyed are negative, suggesting that *D. pseudoobscura* may have experienced a rapid population expansion in the recent past or, alternatively, that slightly deleterious mutations constitute an important component of standing variation in this species.

BOTH *Drosophila melanogaster* and *D. pseudoobscura* are important model species for population genetic and evolutionary studies. While they are fairly closely related (their estimated divergence time is 30 million years), their evolutionary history and ecology are apparently quite different. *D. melanogaster* originated in the tropics, became commensal with humans, and has spread worldwide in the recent past (Lachaise *et al.* 1988). *D. pseudoobscura*, in contrast, originated in North America, where it lives largely apart from people in forested habitats (Dobzhansky and Epling 1944). Although *D. pseudoobscura*'s primary range extends down into Guatemala, it is a temperate, not tropical, species. Furthermore, *D. pseudoobscura*, unlike *D. melanogaster*, is not cosmopolitan. Its range appears to be limited, for reasons that are not understood, to the western half of North and Central America. Within that range, however, *D. pseudoobscura* has relatively high rates of dispersal and little population structure compared with *D. melanogaster*, which is relatively sedentary (Powell 1997). While our knowledge of the history of these species is limited, what we do know suggests that temperate populations of *D. melanogaster* may have experienced significant amounts of adaptive evolution in the recent past, while the environment of *D. pseudoobscura* may have been more stable (albeit shifting in size and location during periods of glaciation).

Levels of DNA sequence variation in these two species suggest that *D. melanogaster* has a smaller effective population size (N_e) than *D. pseudoobscura*. Schaeffer *et al.* (1987) found ~4-fold more restriction-site variation in *D. pseudoobscura* than in *D. melanogaster* in a 13-kb region around the alcohol dehydrogenase (*Adh*) locus, while sequencing of *Adh* and *Adh-dup* in a population sample of *D. pseudoobscura* (Schaeffer and Miller 1992a) revealed ~2.2-fold more variation than in a world-wide sample of *D. melanogaster* alleles (Kreitman and Hudson 1991). Restriction fragment length polymorphism (RFLP) studies of the *Xdh* locus suggested a 3-fold larger effective population size for *D. pseudoobscura* (Riley *et al.* 1989). Extensive surveys of DNA sequence polymorphism have been made only in *D. melanogaster*, however. Furthermore, the relationship between rates of recombination and polymorphism, which has been shown in *D. melanogaster* to be very important (*e.g.*, Begun and Aquadro 1992; Aquadro *et al.* 1994), has not been

Corresponding author: Martha T. Hamblin, Department of Human Genetics, University of Chicago, 924 E. 57th St., Chicago, IL 60637. E-mail: mhamblin@genetics.uchicago.edu

explored at all in *D. pseudoobscura*. Such an exploration is of great interest, as it might allow the disentanglement of some of the factors that contribute to the differences in levels of DNA sequence variation both within and between species. For example, the genetic map of *D. pseudoobscura* is known to be considerably longer than that of *D. melanogaster*, though their genomes are the same size (Powell 1997). To what extent does this larger genetic map, rather than larger species-level effective population size, account for the severalfold difference in levels of molecular variation?

Another interesting issue concerns the relative contributions of adaptive *vs.* purifying selection to the observed relationship between recombination and variation in *D. melanogaster*. That relationship is driven by the local reduction of effective population size in regions linked to targets of both positive and negative selection (Hill and Robertson 1966; Kaplan *et al.* 1989; Charlesworth *et al.* 1993; Wiehe and Stephan 1993; Hudson and Kaplan 1995; Charlesworth 1996). It has been argued that the rate of deleterious mutations has been overestimated (Keightley 1996; Fry *et al.* 1999), raising the possibility that the polymorphism in *D. melanogaster* has been shaped by significant amounts of positive selection resulting from adaptation to new environments. Given the different history of *D. pseudoobscura*, we might expect the contribution of positive selection to the relationship between recombination and variation in this species to be less important.

One of the obstacles to this type of analysis is the requirement for integrated physical and genetic maps such as are currently available for *D. melanogaster*. In the case of *D. pseudoobscura*, however, this obstacle is not huge. The presence of polytene chromosomes allows physical localization, by *in situ* hybridization, of any cloned region or PCR product. Sequence conservation of coding regions between *D. melanogaster* and *D. pseudoobscura* facilitates PCR amplification of homologous loci. Conservation of the five major linkage groups (elements A–E; Sturtevant and Novitski 1941) allows prediction of the chromosomal location of any gene that has been localized in *D. melanogaster*. We have taken advantage of these attributes to develop integrated genetic and physical maps of the second chromosome of *D. pseudoobscura*. This chromosome is homologous to chromosome 3R in *D. melanogaster*, where many genes have been surveyed, and which was used by Hudson and Kaplan (1995) to test the background selection hypothesis. We have also collected DNA sequence polymorphism data for seven loci across the chromosome. These data have been analyzed to determine the nature of the relationship between variation and recombination in *D. pseudoobscura*, and to ask whether this relationship can be explained by similar parameters (intensity and frequency of selective events, species-level effective population size) as for *D. melanogaster*.

MATERIALS AND METHODS

Fly stocks: The population sample used for this study was obtained from a collection of isofemale lines of *D. pseudoobscura* from Goldendale, Washington established by M. Noor in summer of 1996. These lines were inbred by full-sib mating to facilitate sequencing and genetic analysis. Twenty-two lines were successfully inbred for 11–15 generations. The *D. miranda* line, SP235 from Spray, Oregon, was obtained from W. Anderson.

DNA preparation: DNAs for mapping were prepared from single flies arrayed in 96-well plates (Gloor *et al.* 1993). DNAs for sequencing were prepared from groups of 10 flies from the same inbred line using the method of Ashburner (1989).

Construction of a genetic and physical map of the second chromosome: GenBank sequences of *D. pseudoobscura* genes that are homologous to genes on 3R of *D. melanogaster* were identified. These sequences were examined for the presence of microsatellites or other repeated sequences that might provide highly polymorphic, easily scored markers. Such sequences were found in or near four genes: *Glucose dehydrogenase* (*Gld*), *Rhodopsin 1* (*Rh1*), *Myosin light chain 1* (*Mlc1*), and *bicoid* (*bcd*). PCR primers were designed to amplify small products containing these repeats. In gene regions where no repeated sequences were found, a survey of sequence variation at the locus was used to identify regions that were likely to show multiple alleles by single-strand conformation polymorphism (SSCP) analysis (Orita *et al.* 1989). This approach was successful in identifying markers in four additional genes: *Rhodopsin 3* (*Rh3*), *ultrabithorax* (*Ubx*), *Xanthine dehydrogenase* (*Xdh*), and *ribosomal protein 49* (*rp49*).

In the case of *tropomyosin 1* (*trop1*), no sequence data were previously available from *D. pseudoobscura*, but the physical location was known (B. Charlesworth, personal communication). The *D. pseudoobscura* sequence of all of intron C and 713 nucleotides of intron D of the *trop1* gene was obtained by PCR using primers based on conserved exon sequence from *D. melanogaster* (1358–1379F and 2730–2710R from GenBank accession no. K03277). *D. melanogaster* contains a $(CT)_n$ microsatellite (nucleotides 2295–2324) that we found to be conserved and variable in *D. pseudoobscura*. The *D. pseudoobscura* sequence of this region has been deposited in GenBank as accession nos. AF039273 and AF039274.

These markers and their physical locations, as well as an additional microsatellite marker, *Dps2003*, that had been genetically mapped to chromosome 2 by Noor *et al.* (1999), are presented in Table 1. The marker regions were amplified from 27 partially inbred lines to identify the alleles in the population and to find suitable lines for setting up the mapping crosses. Based on the results of the marker screen, a cross was set up between lines 7 and 21, which are fixed for different alleles at 9 of the 10 markers (all but *rp49*, which only has two alleles in this sample). Examination of chromosome squashes from 7×21 F_1 larvae showed no chromosomal inversions, indicating that they share the same third chromosome inversion type. F_1 virgin females were held for 1–3 days after eclosion, then placed singly in vials of yeast-glucose food with a single F_1 male. F_1 parents were removed on the 9th day after female eclosion. All flies were maintained at 20°, on a 12-hr light/12-hr dark cycle. Although 25° is the standard temperature for mapping crosses in *D. melanogaster* (Ashburner 1989) and has been used for *D. pseudoobscura* as well (Levine and Levine 1954), ecological and behavioral studies (Dobzhansky and Eppling 1944; Taylor 1986) suggest that 20° may be a more natural temperature for this species (see discussion). We have thus chosen this temperature given our interest in the rate of recombination in natural populations.

DNA was prepared from 192 F_2 progeny for scoring each

TABLE 1

Molecular markers across chromosome 2 of *D. pseudoobscura*

Locus	GenBank accession no.	Primers	Temp.	Type of variant	Location	No. of alleles in 27 lines	Cytological location[a]
Dps2003	Not available			Length	Unknown	Not determined	43
trop1	AF039274	F:202–223 R:388–409	58°	Length: $(CT)_n$	Intron	4	43/44 border
Rh3	X65879	F:1673-1694 R:1890–1869	58°	Length	3′ UTR	6	44
Gld	M29299	F:2101-2121 R:2240-2221	65°	Length:$(CTGA)_n$	Intron	>6	45
Rh1	X65877	F:2131-2150 R:2386-2365	65°	Length:$(CAA)_n$	3′ UTR	>7	45
bcd	X55735	F:1402-1422 R:1622-1603	66°	Length:$(CAG)_n$	Coding	3	50
Ubx	X05727	F:1488-1509 R:1665-1643	59°	SSCP	Intron	4	ND
Xdh	M33977	F:1883-1902 R:2105-2086	59°	SSCP	Intron	>10	54/55 border
Mlc1	L08052	F:991-1011 R:1130-1111	58°	Length: $(CA)_n$	Intron	6	59/60 border
rp49	S59382	F:808-828 R:1163-1144	60°	SSCP	5′ UTR	2	62

ND, not determined.

[a] Location of *Dps2003* is by inference from genetic data; *trop1*, *Gld*, and *Mlc1* were determined by B. Charlesworth (personal communication); *bcd*, *Xdh*, and *rp49* are from Segarra *et al.* (1996); *Rh1* is from Carulli and Hartl (1992); *Rh3* is from this study (see text).

of the nine markers. An additional cross between lines 7 and 51 was set up in the same way to score the location of *rp49*. Inversion loops (presumably on the third chromosome, which is polymorphic for inversions in this population) were observed in F_1 larvae of this cross, raising the possibility that crossing over on the second chromosome may have been somewhat elevated in this cross, due to the interchromosomal effect (Schultz and Redfield 1951). The data were analyzed as an F_2 backcross (because there is no recombination in males) using Mapmaker (Lander *et al.* 1987) with the Kosambi mapping function.

***In situ* hybridization:** Probes for *Rh3* and *bcd* were prepared by biotinylation of the same PCR products used as sequencing templates. Hybridizations to polytene chromosome preparations were performed as described by Lim (1993), using Vectastain reagents. Maps of the polytene chromosomes were from Stocker and Kastritsis (1972).

DNA sequence variation: Approximately 1 to 1.8-kb regions were sequenced in samples of 10–12 inbred lines for seven loci whose physical and genetic locations were known. For five of these loci, one allele was also sequenced from *D. miranda*. The regions, which were chosen to include as much noncoding sequence as possible, are shown in Table 2. PCR products were sequenced directly using the Thermosequenase cycle sequencing system from Amersham (Arlington Heights, IL), after agarose gel purification using the Qiaex II system (Qiagen, Valencia, CA). Estimates of $4N_e\mu$, π, and θ_W, were calculated according to Nei (1987) and Watterson (1975), respectively. Throughout the article, θ_W refers to Watterson's estimator.

The background selection model: Physical and genetic data generated in this study (Table 3) were used in Equation 15 of the background selection model of Hudson and Kaplan (1995) to make predictions of f_0 across the second chromosome. $(1 - f_0)$ is the fractional decrease in expected variation due to the effects of background selection. Loci mapped to a polytene section were assumed to be in the middle of the section; *e.g.*, *Rh1* was assumed to be at 2.5 sections from the centromere. *Gld* and *Ubx* were not included in this analysis because we had no polymorphism data for these genes. Details of the calculations are given in Hamblin and Aquadro (1996).

The ends of the chromosomes were treated two different ways. In the first treatment (Low, Table 5), which leads to lower estimates of f_0, no additional recombination at the unmapped ends was included. *Dps2003* was assumed to be at the centromere; *rp49* was assumed to be at 19.5, with a most distal 0.5 section to the telomere having a recombination rate of zero.

In the second treatment (High, Table 5), we assumed that the unmapped ends of the chromosomes had the same rates of recombination as the adjacent mapped intervals. *Dps2003* was assumed to be at 0.5 sections from the centromere, and the most proximal 0.5 section was assumed to have the same genetic length as the interval from *Dps2003* to *trop1*: 3.8 cM. As in the first analysis, *rp49* was assumed to be at 19.5, but the most distal 0.5 section was assumed to have the same rate of recombination as the interval from *Mlc1* to *rp49*, namely 3.5 cM/0.5 section.

RESULTS

Genetic map and rates of recombination: Nine molecular markers across the second chromosome were developed based on published genomic sequence. Results of the population survey for these markers are shown in Table 1. The high level of variation at most loci made it possible to score eight of the nine markers, as well as microsatellite *Dps2003*, in F_2 progeny of a single cross.

TABLE 2

Regions sequenced and PCR and sequencing primers

Locus	Region surveyed (accession no.)	PCR primers	Annealing temp.	Additional sequencing primers
trop1	46–345 (AF039273) 1–716 (AF039274)	1358–1379F 2730–2710R (K03276)	63°	409–388R (AF039274)
Rh3	55–1000 1441–1846 (X65879)	32–51F 1890–1869R	62°	419–439F
Rh1	1015–1330 1621–2364 (X65877)	993–1014F 2386–2365R	65°	1598–1619F
bcd	473–1602 (X55735)	429–450F 1622–1603R	65°	827–844F
Xdh intron	1679–2081 (M33977)	1654–1674F 2105–2086R	58°	None
Xdh exon	2836–4298 (M33977)	2814–2835F 4321–4299R	65°	3191–3210F 3955–3936R
Mlc1	1012–1675 1991–2469 (L08052)	991–1011F 2509–2488R	60°	1272–1291F
rp49	562–1863 (S59382)	528–550F 1886–1864R	65°	808–828F 1144–1163F

The number of F_2 progeny scored was in the range of 185–192 for all markers except *Rh1*, for which 166 progeny were scored. The last marker, *rp49*, had only two alleles and was scored in a separate cross, 7 × 51. Because the physical order of the loci was already known, it was necessary to score only one other marker, *Mlc1*, in cross 7 × 51 to locate *rp49* on the genetic map. Due to technical problems, fewer progeny were scored from this cross, and the estimate of genetic distance between *Mlc1* and *rp49* is based on only 79 progeny.

The genetic map is shown in Figure 1. The order of the markers is consistent with published cytological locations, except in the case of *Rh3*. Carulli and Hartl (1992) localized all four rhodopsin genes in *D. pseudoobscura* and reported that *Rh3* is in section 53, while *Rh1* is in section 45. The genetic analysis places *Rh3* proximal to *Rh1*, closely linked to *trop1* in section 44. We performed an *in situ* hybridization using as a probe a biotinylated PCR product that was also a template for sequencing of *Rh3* and found that the probe hybridizes in section 44, which is consistent with the genetic data.

TABLE 3

Physical and genetic distances between markers

Locus	Polytene sections from centromere	Map units from *Dps2003*
Dps2003	<1	0
trop1	1.0	3.8
Rh3	1.5	4.8
Gld	2.5	10.6
Rh1	2.5	11.8
bcd	7.5	50.3
Ubx	?	56.7
xdh	12.0	83.7
Mlc1	17.0	119.5
rp49	19.5	128.2

The total length of our genetic map for this chromosome is >128 cM, as compared with the published length of 101 cM based on previously available visible and allozyme markers (Anderson 1990). The physical map of the *D. pseudoobscura* genome is divided into 100 sections, and published cytological localizations do not have the resolution of the *D. melanogaster* data. Two of our markers, *Gld* and *Rh1*, fall within section 45, and *trop1* and *Rh3* both fall within section 44. While the genetic data allow us to establish their order, we do not have precise distances between these four closely linked markers. Our most distal marker, *rp49*, is within the last section (62) of the chromosome, but is not at the very tip. Our most proximal marker, *Dps2003*, must be within the first section (43), based on its genetic location.

Physical and genetic locations are shown in Table 3 and are presented graphically in Figure 2 with similar data from chromosome arm 3R of *D. melanogaster* (Lindsley and Zimm 1992) plotted for comparison. The slope of the line in this plot is proportional to the rate of recombination. While these two chromosomal elements contain the same complement of genes and essentially the same amount of DNA (Powell 1997), the *D. pseudo-*

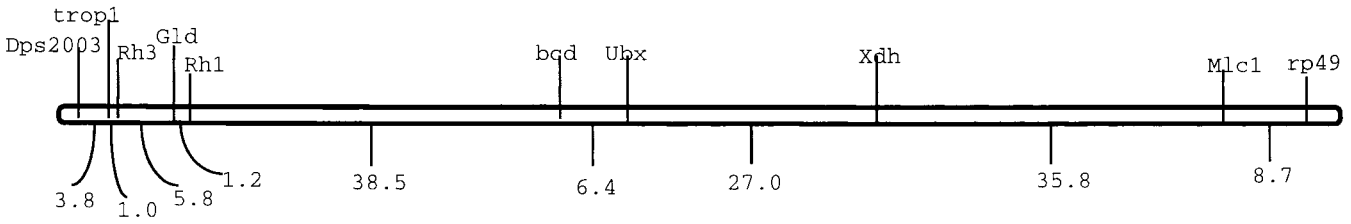

Figure 1.—Genetic map of chromosome 2. Numbers represent the distance (in centimorgans) between adjacent markers.

obscura second chromosome is genetically almost twice as large as 3R in *D. melanogaster*. Unlike *D. melanogaster*, the *D. pseudoobscura* chromosome lacks any extensive regions where recombination is drastically reduced (keeping in mind that we do not have markers at the extremes of the centromere and telomere). Rather, the rate of recombination across most of the *D. pseudoobscura* second chromosome is apparently quite uniform, which is similar to chromosome arm 3R in *D. mauritiana* (True *et al.* 1996).

Levels of DNA sequence variation in a population sample: We surveyed DNA sequence variation at seven of the loci for which we had scored genetic map position; Table 4 summarizes the data. There is a fourfold difference in θ_W at silent sites between the least variable locus, *trop1*, and the most variable, *Xdh*. Estimates of π at silent sites vary about eightfold. For *trop1*, *Rh1*, *Mlc1*, *Xdh*, and *rp49*, one allele from *D. miranda* was sequenced to obtain an estimate of divergence (Table 4). None of these five loci shows a departure from the neutral expectation when compared to each other or to *Adh* (using the Apple Hill population sample; Schaeffer and Miller 1992a) by the method of Hudson *et al.* (1987). However, because divergence to *D. miranda* is quite small (0.9–5.0%), this test has low power.

Estimates of π are lower than estimates of θ_W for all loci except *bcd*, as indicated by the negative Tajima's *D* (1989a) statistics (Table 4). Tajima's *D*'s for *Adh* (Schaeffer and Miller 1992a), *Hsp82* (Wang *et al.* 1997), and *per* (Wang and Hey 1996) from *D. pseudoobscura* are also negative. While only two of the statistics are significantly different from zero (*trop1* and *Adh* noncoding), one would expect that approximately half the statistics would be negative and half positive by chance. We say "approximately" because the distribution of Tajima's *D* is slightly skewed toward the negative. The preponderance of negative statistics ($P = 0.02$ by a sign test that does not take into account the negative skew) suggests that a population-level phenomenon may be responsible.

Fu and Li (1993) tests (D^*) are negative for five of the seven loci (all except *bcd* and *Rh1*) in our data set. Again, *trop1* is the only locus that is significantly different from zero ($D^* = -2.60$, $P < 0.02$).

Prediction of the effects of background selection: Given the recombinational map described above, we wanted to determine the expected impact of background selection on levels of neutral variation. In the absence of background selection, differences in θ among loci are due only to differences in μ, since N_e is the same across all loci. Background selection, however, causes regional differences in N_e (Charlesworth *et al.* 1993). To avoid confusion, we use the symbol $N_{e,0}$ to represent N_e in the absence of background selection (this can be thought of as species-level effective population size). Locus-specific N_e is related to $N_{e,0}$ by the parameter f_0, the fraction of variation remaining after background selection: $N_e = f_0 N_{e,0}$. At any given locus, f_0 is a function of (1) the rate of mutation to deleterious alleles (U) per genome; (2) the strength of selection against those alleles in the heterozygous state (sh); and (3) the regional rate of recombination, which determines the size of the region in which deleterious mutations will be linked to the locus.

We calculated values of f_0 using the simplified model of Hudson and Kaplan (1995; Equation 15), which assumes that rates of mutation and selective effects are uniform across the genome and that differences in the strength of background selection are solely a consequence of differences in rates of recombination.

Using $U = 1$ and $sh = 0.02$, the same values used for *D. melanogaster* (Charlesworth *et al.* 1993; Hudson

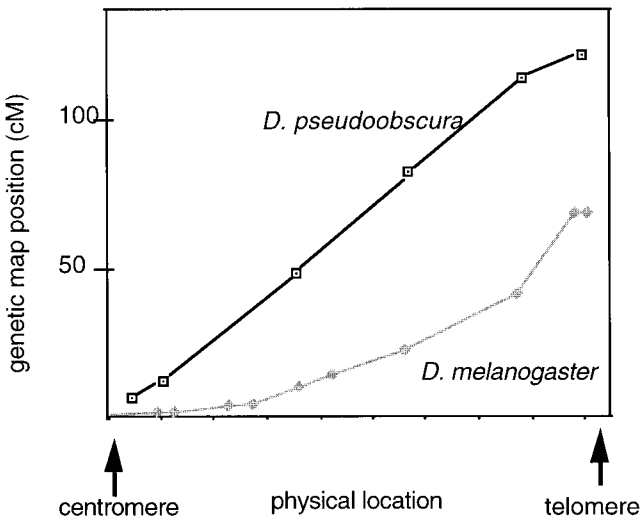

Figure 2.—Physical *vs.* genetic location on chromosome 2 of *D. pseudoobscura* and 3R of *D. melanogaster* (element E).

TABLE 4

DNA sequence variation across the second chromosome

Locus	n^a	S^b	θ^c	π^d	Tajima's *D*	Sites surveyed	Divergence (%)[e]
trop1	12	13	0.0043	0.0020	−2.104 ($P < 0.01$)	1008 noncoding	0.9
Rh3	10	22	0.0058	0.0051		1341 total	ND
		21	0.0100	0.0088	−0.550	744.3 silent	
		14	0.0089	0.0070		555 noncoding	
		7	0.0131	0.0137		189.3 synonymous	
		1	0.0006	0.0006		596.7 replacement	
Rh1	11	26	0.0069	0.0063		1191 total	
		26	0.0108	0.0102	−0.287	818.5 silent	1.6
		23	0.0110	0.0103		711 noncoding	
		3	0.0095	0.0095		107.5 synonymous	
		0	0	0		372.5 replacement	
bcd	10	19	0.0060	0.0062		1129 total	ND
		18	0.0131	0.0133	0.043	484.65 silent	
		5	0.0063	0.0063		282 noncoding	
		13	0.0227	0.0229		202.65 synonymous	
		1	0.0006	0.0008		640.35 replacement	
Xdh	11	91	0.0167	0.0148		1864 total	
		81	0.0359	0.0308	−0.669	770.55 silent	4.1
		26	0.0223	0.0174		403 noncoding	
		55	0.0483	0.0456		367.55 synonymous	
		10	0.0041	0.0030		1093.45 replacement	
Mlc1	10	38	0.0118	0.0085		1134 total	
		38	0.0126	0.0102	−0.884	1066.5 silent	1.3
		38	0.0128	0.0104		1050 intron	
		0	0	0		16.5 synonymous	
		0	0	0		67.5 replacement	
rp49	12	14	0.0036	0.0026		1302 total	
		14	0.0046	0.0037	−0.855	992.8 silent	1.4
		12	0.0044	0.0036		897 noncoding	
		2	0.0069	0.0049		95.8 synonymous	
		0	0	0		309.2 replacement	

ND, not determined due to technical problems.
[a] Number of alleles sequenced.
[b] Number of segregating sites (mutations).
[c] An estimate of $4N_e\mu$ from *S* (Watterson 1975).
[d] Nucleotide diversity (Nei 1987).
[e] Average pairwise difference between *D. pseudoobscura* alleles and one *D. miranda* allele, for all silent sites.

and Kaplan 1995), and the genetic and physical map data in Table 3, we calculated values of f_0 for the seven loci in our sequencing survey. (Because the *D. pseudoobscura* genome is divided into 20 sections, we used a value of 0.01 mutations per section, equivalent to 0.0002 mutations per polytene band in *D. melanogaster*, per generation.) Our markers cover ~95% of the physical length of the chromosome, with rates of recombination in the most proximal and most distal 0–5% being unknown. Because of this uncertainty about the ends of the chromosomes, we calculated values of f_0 under two alternative scenarios: (1) high, rates of recombination are the same as that in the adjacent segment; (2) low, there is no recombination in the unmapped segments (Table 5; see materials and methods for details). Under the first scenario, f_0 is essentially uniform across the chromosome. The second scenario predicts an ~10% difference between the most and least variable loci. This difference increases to 40% if the selection coefficient is changed to 0.005. Expected values of f_0 for these loci in *D. melanogaster* (from Figure 4 of Hudson and Kaplan 1995) are presented in Table 5 for comparison.

Expected values of f_0 are related to expected values of θ by the parameter π_0 (the level of variation in the absence of background selection, *i.e.*, $4N_{e,0}\mu$), so that $f_0 \times \pi_0 = E(\theta_W)$. Figure 3 shows the expected values of θ under the four sets of parameters, using an estimate of π_0 based on silent (noncoding and synonymous) sites that gives the best fit of the observed data to the model (see below). The high variation observed at *Xdh* is not predicted by any of the models (but note that divergence at *Xdh* is 4.1%, more than twice that at other loci; Table 4). Otherwise, the shape of the curve is best predicted by the model with $sh = 0.005$ and assuming

TABLE 5

Predictions of f_0 under the background selection model

	sh = 0.02		*sh* = 0.005		
Locus	High	Low	High	Low	mel[a]
trop1	0.748	0.716	0.647	0.590	0.13
Rh3	0.742	0.709	0.642	0.621	0.38
Rh1	0.779	0.776	0.765	0.764	0.38
bcd	0.778	0.778	0.770	0.770	0.04
Xdh	0.771	0.771	0.763	0.763	0.22
Mlc1	0.711	0.709	0.670	0.669	0.56
rp49	0.754	0.697	0.716	0.463	0.57

See materials and methods for explanation of assumptions used in calculation of high and low values.

[a] Estimate of f_0 for the same locus in *D. melanogaster*, interpolated from Figure 4 of Hudson and Kaplan (1995); *sh* = 0.005, *U* = 1.

no recombination in the unmapped segments (fourth column of Table 5).

The estimate of π_0 was found by performing regression analysis of θ_W (for all loci except *trop1* because of its significant Tajima's *D*) on the predictions of f_0. Because $E(\theta_W)/\pi_0 = f_0$, the slope of the line $\theta_W = m \times f_0$ is an estimate of π_o. (We used the "no-intercept" option of Statview, which forces the regression line to pass through the origin.) Separate regressions were performed using θ_W at total sites, silent sites, or synonymous sites only. *Xdh* is an outlier in all three data sets, so we also performed the regressions without *Xdh*, which greatly improved the fit of the data to the model. The results of the analysis, using the model with *sh* = 0.005 and assuming no recombination in the unmapped segments, are presented in Table 6. For Figure 3, we chose an estimate of π_0 based on θ_W at silent sites because it shows the highest correlation with f_0 ($r^2 = 0.91$ *vs.* $r^2 = 0.59$ and $r^2 = 0.67$ for total sites and synonymous sites, respectively).

TABLE 6

Estimates of π_0 for *sh* = 0.005 (low)

Sites	Without *Xdh*	With *Xdh*
Total	0.010	0.013
Silent[a]	0.016	0.022
Synonymous	0.020	0.031

[a] Synonymous and noncoding sites.

DISCUSSION

Genetic map of chromosome 2: Our immediate purpose in constructing a genetic map was to relate rates of recombination to levels of DNA sequence variation in natural populations (as opposed to providing a framework for identifying genetic loci). It is therefore important to consider whether our map is likely to reflect average rates of recombination in the study population from which our estimates of variation were obtained. There is genetic variation for rates of recombination in *D. pseudoobscura* (Levine and Levine 1954, 1955), and variables such as temperature, days since eclosion, and the presence of inversions are all known to affect crossover frequencies in *D. melanogaster* (Ashburner 1989); presumably these variables are important in *D. pseudoobscura* as well. The Goldendale population used in our study is polymorphic for third chromosome inversions (M. Noor, personal communication), and crossovers on the second chromosome are expected to increase when third chromosomes are heterozygous for inversion type (the Schultz-Redfield effect). Because all but one short interval (*Mlc1-rp49*) of our map was based on a cross homozygous for third-chromosome type, our inferred rates of recombination may underestimate somewhat the rates for this population in nature.

The temperature that a female *D. pseudoobscura* is likely to experience during meiosis in the wild is not known. Studies of daily activity found that flies are active

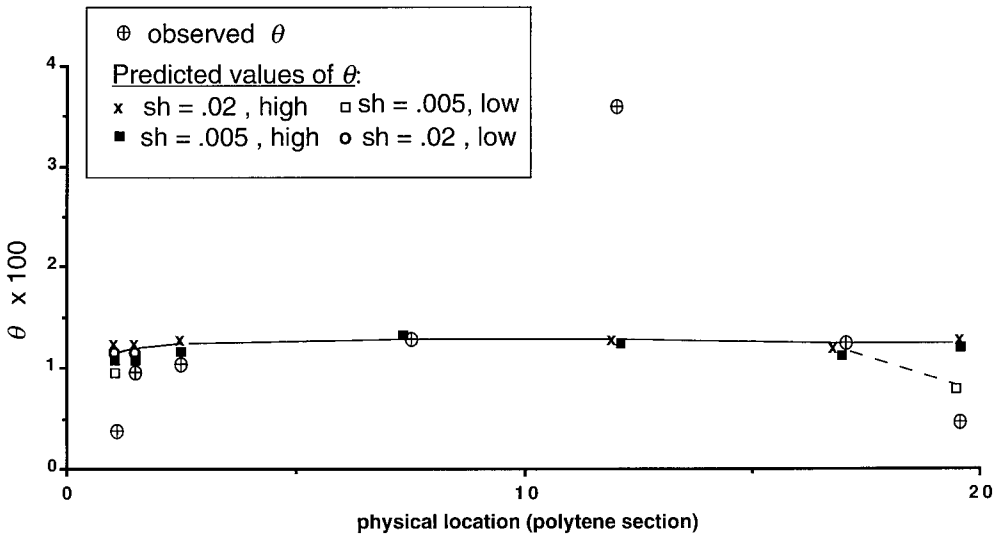

Figure 3.—Observed and predicted values of θ at silent sites (see text), under the background selection model.

at 10° to 31°, but are not usually found at baits during the hotter (>21°) parts of the day (Dobzhansky and Epling 1944). In a laboratory study of temperature choice, *D. pseudoobscura* preferred 15° over 25° (Taylor 1986). We conducted our mating experiments at 20° rather than at 25°, the standard temperature for *D. melanogaster*, because we believe this may be closer to the temperature that flies seek in nature. In any case, the effect of temperature within this range is likely to be very small, as significant temperature effects on crossover frequencies in *D. melanogaster* were observed only at temperatures >29° or <17.5° (Plough 1917).

Our map of chromosome 2, based on two genotypes chosen at random from the population, is almost 30% longer than the published map of Anderson (1990), based on visible and allozyme markers. Some of this difference may be due to the fact that we had information about cytological locations and were able to choose markers covering almost the full physical length of the chromosome. On the other hand, our map is considerably shorter (128 cM *vs.* 203.9 cM) than another independently constructed map of the second chromosome based on some of the same markers that we used (Noor *et al.* 1999). Therefore, while we have no reason to think that our map is inaccurate, it is important to realize that genetic variation, polymorphic inversions, and other variables interact to produce a distribution of crossover frequencies in natural populations, of which our map is simply one estimate.

Levels of variation: We analyzed levels of neutral variation at seven loci across the second chromosome of *D. pseudoobscura*, substantially increasing the number of estimates of sequence variation published for this species. Previous comparisons with sequence data from *D. melanogaster* have been problematic because estimates of $4N_e\mu$ from *D. melanogaster* come from many different kinds of samples (see Moriyama and Powell 1996), many of which are inappropriate for the type of analysis presented here. In addition, some of the loci surveyed were chosen with an expectation of a departure from neutrality. The most appropriate *D. melanogaster* data for comparison are those of Kindahl (1994), a collection of randomly chosen autosomal loci all surveyed in the same sample from a single North American population. Kindahl estimated total θ_W (*i.e.*, an estimate of $4N_e\mu$ at all sites, coding and noncoding) on the basis of 4-cutter variation across regions 1.9–4.6 kb in length with an average of 46% coding sequence. This is quite similar to the average of 42% coding sequence in our surveys (Table 4).

Average levels of total variation in the Goldendale population of *D. pseudoobscura* are ~1.5-fold higher than in the Maryland population of *D. melanogaster* (Table 7). Most of this difference comes from the lower end of the range: the least variable locus in *D. pseudoobscura* is 10–20-fold more variable than the least variable locus in *D. melanogaster*. The estimate of total θ_W at *Adh* in *D. pseudoobscura* was 0.015 in the most variable population sample, Gundlach-Bundshou (Schaeffer and Miller 1992b), slightly lower than our most variable locus, *Xdh*.

TABLE 7

Average levels of sequence variation compared between *D. pseudoobscura* and *D. melanogaster*

	D. melanogaster[a]	*D. pseudoobscura*[b]
Total sites		
Average θ	0.0052	0.0078
Range θ	0.0004–0.0101	0.0036–0.0167
Average π	0.0046	0.0069
Range π	0.0001–0.0084	0.0022–0.0148
Synonymous sites		
Average θ	0.0077	0.0223
Range θ	0–0.0204	0.0069–0.0483
Average π	0.0084	0.0205
Range π	0–0.0243	0.0049–0.0456

[a] Data for total sites are from Kindahl (1994).

Data for synonymous sites are from superoxide dismutase (Hudson *et al.* 1994); *transformer* (Walthour and Schaeffer 1994); *Mlc1* (Leicht *et al.* 1995); *Gld* (Hamblin and Aquadro 1997); *larval cuticle protein* Ψ (Pritchard and Schaeffer 1997); *Acp26Aa* and *Acp26Ab* (Aguade *et al.* 1992); triose phosphate isomerase (Hasson *et al.* 1998); and *rosy*, esterase 6, ornithine dehydrogenase, *larval serum protein 1*-γ (all unpublished data from V. L. Bauer and C. F. Aquadro).

[b] Data for both total and synonymous sites are from this study and Schaeffer and Miller (1992a).

For a comparison based on synonymous sites in coding sequence, we used estimates from 5 of the loci in this study (all except *trop1* and *Mlc1*, which had 0 and 16.5 synonymous sites, respectively) plus the data for *Adh* and *Adh-Dup* in the Apple Hill population (Schaeffer and Miller 1992a). For *D. melanogaster*, we used estimates from 12 autosomal loci measured in North American population samples (for details, see Table 7). Variation at synonymous sites is ~2.4–3 times higher in *D. pseudoobscura*. The greater difference in levels of synonymous variation could be due to higher variation in noncoding regions such as introns, or higher replacement polymorphism, in *D. melanogaster*. Replacement polymorphism was 12.5% of total variation in coding regions in our surveys in *D. pseudoobscura* (including *Adh* and *Adh-dup*) as compared to 26.4% reported by Moriyama and Powell (1996) for *D. melanogaster*.

Analysis in the context of regional rates of recombination: Increased overall recombination rate, a lack of substantial suppression of recombination near the centromere, and the reduced size of the linkage group (the acrocentric second chromosome of *D. pseudoobscura* contains only element E, while the metacentric third chromosome of *D. melanogaster* contains both elements D and E) all reduce the interaction of selection and linkage in *D. pseudoobscura* as compared with *D. melanogaster* (Table 5). The relative levels of silent DNA sequence variation observed for the second chromosome

of *D. pseudoobscura* (20% of the genome) can be fairly well predicted using a background selection model assuming the same average mutational and selective forces as are thought to operate in a North American population of *D. melanogaster* (Figure 3).

Note that, although we used a model that is formulated to describe background selection against deleterious mutations, any positively selected mutations that have contributed to regional reductions in effective population size will affect the fit of the model to the data. It was not our goal to discriminate between the separate effects of background selection and selective sweeps. Rather, in using the same values for U and sh as were used by Hudson and Kaplan (1995) for *D. melanogaster*, we were qualitatively testing the hypothesis that the relationship between linkage and the relative level of variation (*i.e.*, f_0) were shaped by similar *total* intensity of selection (both positive and negative) in the two species.

The relatively uniform rates of recombination across the second chromosome of *D. pseudoobscura* make most of the chromosome fairly insensitive to changes in parameters. It was therefore difficult to discriminate between the alternative models presented in Figure 3, and our qualitative assessment of fit to the models became dependent on the ends of the chromosome where our data were less reliable. It was clear, however, that a stronger, rather than a weaker, effect of selection was needed to explain the reduction in variation observed at both ends of the chromosome. Therefore, unless we assume that the genomic rate of deleterious mutation (U) is higher in *D. pseudoobscura*, our analyses provide no support for the idea that hitchhiking events have played a larger role in the recent evolutionary history of North American *D. melanogaster* than *D. pseudoobscura*.

How likely is it that U for *D. pseudoobscura* is larger than for *D. melanogaster*? It is unlikely that replication-based errors occur at a different rate between such closely related species, though densities of transposable elements (TEs), which can contribute to the background selection effect (Charlesworth 1996), can vary considerably. The distributions of TEs have not been studied extensively in *D. pseudoobscura*, but restriction enzyme surveys of three loci covering a total of 63 kb revealed no length variation of the size associated with TE insertions (Aquadro 1993). A hybridization study by Brookfield *et al.* (1984) also found few TEs in *D. pseudoobscura*. Thus there is no evidence that U is larger in *D. pseudoobscura* than in *D. melanogaster*.

It has been argued that U in *D. melanogaster* is considerably smaller, not larger, than 1 (Keightley 1996; Fry *et al.* 1999), the value that we used in all our analyses. If U were in fact much smaller than 1, the correlation between variation and recombination observed in *D. melanogaster* could not be accounted for by selection against deleterious mutations, and one would be forced to conclude that positive selection had played a major role in that relationship (Charlesworth 1996). In *D. pseudoobscura*, however, because most of the recombinational landscape of the second chromosome is quite flat, one would not need to invoke a strong role of positive selection. Rather, a lower species-level effective population size, with f_0 close to 1.0 across much of the chromosome, could explain the data. Better empirical estimates of U are needed to resolve this question.

Species-level effective population size: It has been inferred from a small number of restriction-enzyme and sequencing surveys (*e.g.*, Schaeffer *et al.* 1987; Riley *et al.* 1989; Schaeffer and Miller 1992a) that *D. pseudoobscura* has a three- to fourfold larger effective population size than *D. melanogaster*. Our larger data set of randomly chosen population samples suggests that the difference in levels of polymorphism between the two species may have been slightly overestimated. More importantly, our analysis allows us to estimate the relative contributions of differences in rates of recombination, *vs.* differences in long-term species-level effective population size, to higher variation in *D. pseudoobscura*. This can be done by comparing estimates of π_0, which directly reflects species-level effective population size, assuming similar neutral mutation rates in the two species.

For *D. melanogaster*, the estimate of $\pi_0 = 0.014$ obtained by Hudson and Kaplan (1995) is for total variation. Our estimate of π_0 at total sites in *D. pseudoobscura* is similar: 0.010–0.013 (Table 6). However, a much larger fraction of total variation in *D. melanogaster* appears to be nonsynonymous or noncoding than in *D. pseudoobscura* (Table 7). This discrepancy suggests that differences in total variation between the two species may not be a simple function of effective population size (*i.e.*, that a significant fraction of the variation may not be strictly neutral). We analyzed the relationship between recombination and variation for three classes of sites in *D. pseudoobscura* (Table 6) and found that estimates of silent variation (synonymous plus noncoding) showed the strongest relationship with the predicted effects of background selection, yielding an estimate of $\pi_0 =$ 0.016–0.022 for silent sites, which is not much higher than $\pi_0 = 0.014$ for total sites in *D. melanogaster*.

Synonymous sites are the most variable in both species and show the largest difference between the species (Table 7), so they are presumably most likely to accurately reflect differences in effective population size. Using data for seven loci on the third chromosome and the regression method described above (see results), we estimated π_0 at synonymous sites in *D. melanogaster* to be 0.026, a bit lower than the estimate of 0.03 from Hamblin and Aquadro (1997). Our estimate of $\pi_0 =$ 0.020–0.031 for synonymous sites in *D. pseudoobscura* (Table 6) completely contains the range estimated for *D. melanogaster*. While these comparisons are very crude, the result is not unreasonable and suggests little difference in species-level effective population size between *D. melanogaster* and *D. pseudoobscura*. Note that we assumed $U = 1$ in both species. If U in *D. pseudoobscura*

were actually <1 (see above), observed variation would be even closer to its maximal level, *i.e.*, species-level effective population size would be smaller.

It is quite plausible that species-level effective population sizes of these two species in North America may be more similar than had been thought. While the ecology of neither species is well understood, there is no evidence from molecular data that *D. melanogaster* has experienced a severe bottleneck in establishing its North American populations from very large ancestral African populations. In addition, *D. melanogaster*'s exploitation of abundant agricultural resources certainly provides the opportunity for high population densities.

While species-level effective population size (*i.e.*, $N_{e,0}$) may be similar in the two species, molecular evolution at any particular locus will be a function of $f_0 N_{e,0}$ at that locus, as described above. *D. pseudoobscura*'s higher rates of recombination should allow for faster, more efficient response to selection. In this light, it is interesting that *D. melanogaster*, the species with a shorter genetic map than *D. simulans* and *D. mauritiana* as well as *D. pseudoobscura* (True *et al.* 1996), is a more successful colonizer than any of them.

Excess of rare variants: No difference in the amount of selection is required to explain patterns of variation in these two species, in spite of their seemingly very different evolutionary histories. This apparent similarity may be coincidental, obscuring important differences in several underlying parameters, or it may simply reflect the limited resolution of our data. However, it may also reflect an unexpected similarity in biology suggested by the frequency distributions of variation. Our data, together with previously published results (Schaeffer and Miller 1992a; Wang and Hey 1996; Wang *et al.* 1997) show that Tajima's *D* is negative at 9 out of 10 loci in *D. pseudoobscura*. Negative Tajima's *D* statistics can be an indication of rapid population expansion (Tajima 1989a,b; Aris-Brosou and Excoffier 1996).

The possibility that *D. pseudoobscura* is not at equilibrium has been raised before: Slatkin (1994) pointed out that genetic data provide no evidence for isolation by distance in this species, yet direct estimates of dispersal would predict such an effect. This discrepancy can be explained if populations of *D. pseudoobscura* have in fact not been relatively stable but instead have recently undergone a range expansion accompanied by dramatic population growth. Such an expansion could be accompanied by adaptation to new environments, possibly comparable to the adaptive changes experienced by *D. melanogaster* in temperate regions.

A significant change in population size would violate the equilibrium assumption of the background selection model and may affect our analysis in some unknown way. Nonetheless, this reservation probably also applies to North American populations of *D. melanogaster*, which are thought to be very recently established and may be far from mutation-drift equilibrium for base-pair polymorphisms.

Alternatively, the preponderance of negative Tajima's *D*'s may be due to slightly deleterious variants being maintained at low frequencies throughout the *D. pseudoobscura* genome. At the five loci for which we have surveyed both coding and noncoding regions, there is a trend toward more negative Tajima's *D*'s in noncoding regions than at synonymous sites. If this difference were significant in a larger sample, it would support this alternative hypothesis rather than the hypothesis of population expansion.

CONCLUSIONS

Patterns of molecular variation across the second chromosome of *D. pseudoobscura* are consistent with previously published models of the effects of background selection based on data from *D. melanogaster*. Using these models, the two- to threefold higher levels of silent variation in *D. pseudoobscura* compared to *D. melanogaster* appear to be explained by the former species' twofold longer genetic map and a similar species-level effective population size. Our confidence in this conclusion will be improved by mapping and polymorphism data for more loci and evaluation of how departures from a strictly neutral, equilibrium model of background selection affect parameter estimation. In addition, better estimates of the genomic deleterious mutation rate will permit more accurate inferences about species-level effective population size and the importance of positive selection in shaping genomic patterns of variation in these species.

We thank M. Noor for providing flies, a microsatellite marker, and help with Mapmaker; W. Anderson for the *D. miranda* stock; M. Veuille, F. Depaulis, and members of the Aquadro lab for helpful discussions; and R. Hudson for comments on the manuscript. This work was supported by a grant from the National Institutes of Health to C.F.A. Some of the writing was done while M.T.H. was supported by a Chateaubriand Fellowship from the French government.

LITERATURE CITED

Aguade, M., N. Miyashita and C. H. Langley, 1992 Polymorphism and divergence in the *Mst26A* male accessory gland region in *Drosophila*. Genetics **132:** 355–362.

Anderson, W., 1990 Linkage map of the fruit fly *Drosophila pseudoobscura*, pp. 3.188–3.189 in *Genetic Maps*, edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Aquadro, C. F., 1993 Molecular population genetics of Drosophila, pp. 222–266 in *Molecular Approaches to Fundamental and Applied Entomology*, edited by J. Oakeshott and M. J. Whitten. Springer-Verlag, New York.

Aquadro, C. F., D. J. Begun and E. C. Kindahl, 1994 Selection, recombination, and DNA polymorphism in Drosophila, pp. 46–56 in *Non-neutral Evolution: Theories and Molecular Data*, edited by B. Golding. Chapman and Hall, New York.

Aris-Brosou, S., and L. Excoffier, 1996 The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol. Biol. Evol. **13:** 494–504.

Ashburner, M., 1989 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Brookfield, J. F. Y., E. Montgomery and C. H. Langley, 1984 Apparent absence of transposable elements related to the P elements of *D. melanogaster* in other species of Drosophila. Nature **310:** 330–332.

Carulli, J. P., and D. L. Hartl, 1992 Variable rates of evolution among Drosophila opsin genes. Genetics **132:** 193–204.

Charlesworth, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genet. Res. **68:** 131–149.

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

Dobzhansky, Th., and C. Epling, 1944 *Contributions to the Genetics, Taxonomy, and Ecology of Drosophila pseudoobscura and Its Relatives*. Pub. 554, Carnegie Institute, Washington, DC.

Fry, J. D., P. D. Keightley, S. L. Heinsohn and S. V. Nuzhdin, 1999 New estimates of the rates and effects of mildly deleterious mutation in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **96:** 574–579.

Fu, Y.-Y., and W.-H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

Gloor, G. B., C. R. Preston, D. M. Johnson-Schlitz, N. A. Nassif, R. W. Phillis *et al.*, 1993 Type I repressors of P element mobility. Genetics **135:** 81–95.

Hamblin, M. T., and C. F. Aquadro, 1996 High nucleotide variation in a region of low recombination in *Drosophila simulans* is consistent with the background selection model. Mol. Biol. Evol. **13:** 1133–1140.

Hamblin, M. T., and C. F. Aquadro, 1997 Contrasting patterns of nucleotide sequence variation at the glucose dehydrogenase (*Gld*) locus in different populations of *Drosophila melanogaster*. Genetics **145:** 1053–1062.

Hasson, E., I. N. Wang, L. W. Zeng, M. Kreitman and W. F. Eanes, 1998 Nucleotide variation in the triosephosphate isomerase (*Tpi*) locus of *Drosophila melanogaster* and *Drosophila simulans*. Mol. Biol. Evol. **15:** 756–769.

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

Hudson, R. R., and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics **141:** 1605–1617.

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

Hudson, R. R., K. Bailey, D. Skarecky, J. Kwiatowski and F. J. Ayala, 1994 Evidence for positive selection in the superoxide dismutase (*Sod*) region of *Drosophila melanogaster*. Genetics **136:** 1329–1340.

Kaplan, N., R. R. Hudson and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics **116:** 153–159.

Keightley, P. D., 1996 Nature of deleterious mutation load in Drosophila. Genetics **144:** 1993–1999.

Kindahl, E. C., 1994 Recombination and DNA polymorphism on the third chromosome of *Drosophila melanogaster*. Ph.D. Thesis, Cornell University, Ithaca, NY.

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.

Lachaise, D., M.-L. Cariou, J. R. David, F. Lemeunier, L. Tsacas *et al.*, 1988 Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22:** 159–255.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Leicht, B. G., S. V. Muse, M. Hanczyc and A. G. Clark, 1995 Constraints on intron evolution in the gene encoding the Myosin Alkali Light Chain in Drosophila. Genetics **139:** 299–308.

Levine, R. P., and E. E. Levine, 1954 The genotypic control of crossing over in *Drosophila pseudoobscura*. Genetics **39:** 677.

Levine, R. P., and E. E. Levine, 1955 Variable crossing over arising in different strains of *Drosophila pseudoobscura*. Genetics **40:** 399–405.

Lim, J. K., 1993 *In situ* hybridization with biotinylated DNA. Dros. Inf. Serv. **72:** 73–76.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

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

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Noor, M. A. F., M. D. Schug and C. F. Aquadro, 1999 Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genet. Res. (in press).

Orita, M., Y. Suzuki, T. Sekiya and K. Hayashi, 1989 Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics **5:** 874–879.

Plough, H. H., 1917 The effect of temperature on crossing over in Drosophila. J. Exp. Zool. **24:** 147–209.

Powell, J. R., 1997 *Progress and Prospects in Evolutionary Biology*. Oxford University Press, New York.

Pritchard, J. K., and S. W. Schaeffer, 1997 Polymorphism and divergence at a Drosophila pseudogene locus. Genetics **147:** 199–207.

Riley, M. A., M. E. Hallas and R. C. Lewontin, 1989 Distinguishing the forces controlling genetic variation at the *Xdh* locus in *Drosophila pseudoobscura*. Genetics **123:** 359–369.

Schaeffer, S. W., and E. L. Miller, 1992a Molecular population genetics of an electrophoretically monomorphic protein in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Genetics **132:** 163–178.

Schaeffer, S. W., and E. L. Miller, 1992b Estimates of gene flow in *Drosophila pseudoobscura* determined from nucleotide sequence analysis of the alcohol dehydrogenase region. Genetics **132:** 471–480.

Schaeffer, S. W., C. F. Aquadro and W. W. Anderson, 1987 Restriction-map variation in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Mol. Biol. Evol. **4:** 254–265.

Schultz, J., and H. Redfield, 1951 Interchromosomal effects on crossing over in Drosophila. CSH Symp. Quant. Biol. **16:** 175–197.

Segarra, C., G. Ribo and M. Aguade, 1996 Differentiation of Muller's chromosomal elements *D* and *E* in the Obscura group of Drosophila. Genetics **144:** 139–146.

Slatkin, M., 1994 Gene flow and population structure, pp. 3–17 in *Ecological Genetics*, edited by L. A. Real. Princeton University Press, Princeton, NJ.

Stocker, A. J., and C. D. Kastritsis, 1972 Developmental studies in Drosophila. III. The puffing patterns of the salivary gland chromosomes of *D. pseudoobscura*. Chromosoma **37:** 139–176.

Sturtevant, A. H., and E. Novitski, 1941 The homologies of the chromosome elements in the genus Drosophila. Genetics **26:** 517–541.

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

Tajima, F., 1989b The effect of change in population size on DNA polymorphism. Genetics **123:** 597–602.

Taylor, C. E., 1986 Habitat choice by *Drosophila pseudoobscura*; the roles of genotype and of experience. Behav. Genet. **16:** 271–280.

True, J. R., J. M. Mercer and C. C. Laurie, 1996 Differences in crossover frequency and distribution among three sibling species of Drosophila. Genetics **142:** 507–523.

Walthour, C. S., and S. W. Schaeffer, 1994 Molecular population genetics of sex determination genes: the *transformer* gene of *Drosophila melanogaster*. Genetics **136:** 1367–1372.

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. L., 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.

Wiehe, T. H. E. and W. Stephan, 1993 Analysis of a genetic hitchhiking model and its application to DNA polymorphism data. Mol. Biol. Evol. **10:** 842–854.

Communicating editor: R. R. Hudson