Contrasting Patterns of X-Linked and Autosomal Nucleotide Variation in *Drosophila melanogaster* and *Drosophila simulans*

Peter Andolfatto

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, Scotland

Surveys of molecular variation in *Drosophila melanogaster* and *Drosophila simulans* have suggested that diversity outside of Africa is a subset of that within Africa. It has been argued that reduced levels of diversity in non-African populations reflect a population bottleneck, adaptation to temperate climates, or both. Here, I summarize the available single-nucleotide polymorphism data for both species. A simple ''out of Africa'' bottleneck scenario is consistent with geographic patterns for loci on the X chromosome but not with loci on the autosomes. Interestingly, there is a trend toward lower nucleotide diversity on the X chromosome relative to autosomes in non-African populations of *D. melanogaster*, but the opposite trend is seen in African populations. In African populations, autosomal inversion polymorphisms in *D. melanogaster* may contribute to reduced autosome diversity relative to the X chromosome. To elucidate the role that selection might play in shaping patterns of variability, I present a summary of within- and between-species patterns of synonymous and replacement variation in both species. Overall, *D. melanogaster* autosomes harbor an excess of amino acid replacement polymorphisms relative to *D. simulans*. Interestingly, range expansion from Africa appears to have had little effect on synonymous-to-replacement polymorphism ratios.

Introduction

Drosophila melanogaster and *Drosophila simulans* are cosmopolitan sister species that are believed to have an African origin (Lachaise et al. 1988). The timing of their dispersal from Africa is not known with any certainty, but proposed estimates are tens of thousands of years ago for ''ancient populations'' such as those of Europe and Asia and as little as several hundred years ago for the Americas (David and Capy 1988; Lachaise et al. 1988). Several recent studies of *D. melanogaster* reveal that non-African populations harbor reduced levels of nucleotide diversity relative to African populations (Hale and Singh 1991; Begun and Aquadro 1993, 1995; Schlötterer, Vogl, and Tautz 1997; Langley et al. 2000). Similarly, a recent genome-wide study of microsatellite variation in *D. simulans* (Irvin et al. 1998) reports reduced variation in non-African populations relative to African populations. Differences between African and non-African populations have been interpreted as reflecting founder events in the history of non-African populations, directional selection for adaptation to temperate habitats, or both (David and Capy 1988; Begun and Aquadro 1993, 1995; Irvin et al. 1998; Langley et al. 2000). Hereafter, I will refer to the purely demographic hypothesis that only a limited number of African lineages gave rise to non-African populations as the ''bottleneck hypothesis.''

In *D. melanogaster*, comparisons of African and non-African levels of nucleotide diversity have focused primarily on the X chromosome (e.g., Begun and Aquadro 1993, 1995; Langley et al. 2000). Patterns of polymorphism at these loci suggest a bottleneck in non-African populations. Curiously, the handful of autosomal loci that have been studied (Clark and Wang 1997; Aguadé 1998, 1999; Tsaur, Ting, and Wu 1998; Begun et al. 1999; Andolfatto and Kreitman 2000) do not support this hypothesis. The data are more scarce for *D. simulans*, for which African and non-African single-nucleotide variation has been compared at only two-single copy nuclear loci (Begun and Aquadro 1995; Hamblin and Veuille 1999). Here, I reexamine the bottleneck hypothesis for *D. melanogaster* and *D. simulans* using single-nucleotide polymorphisms for a large number of loci scattered throughout the genome. While nucleotide variation in these two species has been summarized before (Moriyama and Powell 1996), a very different picture emerges when African and non-African populations are considered separately.

Key words: bottleneck, selection, deleterious mutations, polymorphism, inversion, synonymous, nonsynonymous, dominance.

Address for correspondence and reprints: Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom. E-mail: peter.andolfatto@ed.ac.uk.

Mol. Biol. Evol. 18(3):279–290. 2001

Materials and Methods

Test of the Bottleneck Hypothesis

I compiled the available nucleotide polymorphism data sets for which there were at least two alleles sampled from both African and non-African populations and at least one polymorphism. *In*(*2L*)*t* refers to the sequence spanning the proximal breakpoint of this *D. melanogaster* inversion (Andolfatto and Kreitman 2000). The *Adh* data set encompasses the *Adh* and *Adh-duplicate* loci. This sample was a composite of the sample of Kreitman and Hudson (1991) and unpublished data (Zimbabwe, Africa) kindly provided by S.-C. Tsaur. Inverted chromosomes were excluded for *In*(*2L*)*t* and *Adh*, since these loci are very close to the *In*(*2L*)*t* proximal breakpoint (Andolfatto and Kreitman 2000). *Adh*-Fast chromosomes were excluded, since sampling was not random with respect to this allele. *Cec-C* (Clark and Wang 1997) and *Amy-d* (Inomata et al. 1995) were chosen as representative genes from their respective clusters to avoid nonindependence issues. Data for *Pgi* in *D. melanogaster* and *D. simulans* (J. McDonald, personal communication) can be found in GenBank under accession numbers 20575–20556, 27539–27539, and 27554–27555. Unpublished data for *transformer* (R. Kulathinal,

personal communication) and *su*(*s*) and *su*(w^a) (Langley et al. 2000) were kindly provided by the authors. For *su*(*s*) and *su*(w^a), I use the European sample as the non-African sample because combined European–North American data were not available. References for other loci are as follows: *anon1A3*, *anon1E9*, and *anon1G5* (Schmid et al. 1999); *Acp26Aab* (Tsaur, Ting, and Wu 1998); *Acp29AB* (Aguadé 1999); *asense* (Hilton, Kliman, and Hey 1994); *Boss* (Ayala and Hartl 1993); *diptericin* (Clark and Wang 1997); *Dras1*, *Dras2*, and *Dras3* (Gasperini and Gibson 1999); *eve* (Ludwig and Kreitman 1995); *G6pd* (Cooke and Oakeshott 1989; Eanes et al. 1996); *period* (Kliman and Hey 1993); *Ref*(*2*)*P* (Wayne, Contamine, and Kreitman 1996); *Rh3* (Ayala, Chang, and Hartl 1993); *Tpi* (Hasson et al. 1998); *transformer* (Walthour and Schaeffer 1994); *vermilion* (Begun and Aquadro 1995); and *Yp2* and *zeste* (Hey and Kliman 1993). Included are data for 41 short expressed sequence tag (EST) markers, labeled in table 1 according to their cytological positions (8 X-linked and 33 autosomal; Teeter et al. 2000). Other loci from Teeter et al. (2000) include *achaete*, *Dms*, *Dop*, *Fog*, *frizzled*, *nina A*, *numb*, *swallow*, *sevenless*, *Tra2*, *tailless*, *thickveins*, and *Tph*. For *D. simulans*, I used the *G6pd* data set of Hamblin and Veuille (1999) and the *vermilion* data set of Begun and Aquadro (1995). I excluded alleles of unknown origin.

Individual homologous loci were compared in African and non-African samples to account for locus-specific substitution rates and effective population sizes. I considered only two-state single-base substitution variation in these analyses; nucleotide substitutions overlapping with deletions were discarded as incomplete information. I used a measure of nucleotide diversity (θ_W) based on the total number of segregating sites in a sample proposed by Watterson (1975). I included synonymous, replacement, and noncoding polymorphisms in my count of segregating sites. The summary θ_W is expected to be sensitive to changes in population size (Tajima 1989*b*). The use of another measure of nucleotide diversity (π; Tajima 1983) led to similar conclusions.

Sign tests (Sokal and Rohlf 1998, p. 444) were performed with the null hypothesis that nucleotide diversity levels are equal in non-African and African populations (i.e., that there was no bottleneck). The tests were one-tailed, since, under both the null and alternative (i.e., a bottleneck) hypotheses, we do not expect more variation outside of Africa.

The sign test assumes that loci are independent. It is unclear whether this is strictly true in the case of loci at the tip of the X chromosome in *D. melanogaster* (i.e., cytological positions 1A to 2B), which are believed to experience little crossing over (reviewed in Langley et al. 2000). This said, recombination events (cf. Hudson and Kaplan 1985) could be detected within the *su*(*s*) and *su*(w^a) loci. In addition, linkage disequilibrium has been shown to decay with distance at *su*(*s*) and *su*(w^a) on the same scale as loci in regions with higher rates of crossing over (Langley et al. 2000). Evidence for considerable recombination (perhaps gene conversion) in these data suggests that collapsing these loci into a single data point is overly conservative. Nonetheless, if the average of all six loci (weighted equally) was used as a single data point, the qualitative conclusions were unchanged.

Levels of Nucleotide Diversity by Chromosome and Geographic Locality

I compared diversities at loci in African and non-African populations of *D. melanogaster* and *D. simulans*. Average X and autosome synonymous site diversities (θ_W and π) included only synonymous sites of coding regions (i.e., all noncoding sites and loci were excluded). I report mean synonymous site diversities over all loci, weighting each of the loci equally. This weighting was not entirely appropriate, since sequenced loci varied in length and sample size, and population samples were not drawn identically. Unfortunately, this problem is inherent in analyses that combine data from many sources. I excluded the data of Teeter et al. (2000) from calculations of nucleotide diversity, since many of the loci are very short (i.e., less than several hundred base pairs) and original sequences were not available for the assignment of coding regions. DnaSPv3.0 (Rozas and Rozas 1999) was used for polymorphism analyses.

In comparisons of averages of nonhomologous loci, I wished to minimize the effects that the recombinational landscape of each chromosome may have on levels of nucleotide variation (Aquadro, Begun, and Kindahl 1994; Charlesworth 1996). To this end, I excluded *anon1E9*, *asense*, and *Dras1* from diversity calculations for *D. melanogaster*. Synonymous variation at these loci was more likely to be affected by selection at linked sites, as the crossing-over rate for these loci was estimated to be less than 5×10^{-9} per base pair per generation (cf. Comeron, Kreitman, and Aguadé [1999] and True, Mercer, and Laurie [1996] for estimated rates of crossing over).

Patterns of Synonymous and Replacement Variation

For between-species comparisons of synonymous and replacement polymorphism, I restricted the analysis to homologous loci sequenced in both species. In addition to loci referenced above, I included *Adh* and *Adh-dup* (*D. simulans*; Sumner 1991), *ci* (Berry et al. 1991), *cta* (Wayne and Kreitman 1996), *Est-6* (Cooke and Oakeshott 1989; Karotam, Delves, and Oakeshott 1993; Hasson and Eanes 1996), *Gld* (Hamblin and Aquadro 1996, 1997), *janus* (Kliman et al. 2000; F. Depaulis, personal communication), *Mlc1* (Leicht et al. 1995), *Pgd* (Begun and Aquadro 1994; Begun and Whitley 2000), *prune* (Simmons et al. 1994), *runt* (Labate, Biermann, and Eanes 1999), and *white* (Kirby and Stephan 1995; Kliman et al. 2000). For African versus non-African comparisons of synonymous and replacement polymorphism, I considered all loci with both African and non-African samples. In this part of the analysis, I included all available *Adh* and *Adh-dup* alleles for *D. melanogaster*.

Sampling Locations

This study combined data from many sources. For the majority of loci in *D. melanogaster,* African samples had a mixed sampling scheme that included one or more lines from Botswana, Kenya, Madagascar, South Africa, or Zimbabwe. Exceptions were *Boss, Ref(2)P,* and *Rh3,* for which only two West African lines were sampled. *Acp29AB, Adh, eve, Pgi,* and *Tpi* were sampled in both East and West Africa. Non-African samples were generally a mix from diverse geographic localities, with a bias toward North America. The data of Teeter et al. (2000) had a consistent sampling scheme for all loci: African populations were composed of one South African line and one (or two) Kenyan lines, and non-African lines were drawn from a worldwide sample. Generally, samples for *D. simulans* included one or more lines from East Africa. *G6pd, Boss, Pgi,* and *Rh3* were sampled in both East and West Africa; *vermilion* included only West Africa. Non-African samples were generally a mix from diverse geographic localities, with a bias toward North America. Additional details can be found in the original sources (see references above).

Results and Discussion

A Complex History for *D. melanogaster*

In table 1, I compare estimates of θ_W in African and non-African samples for 20 X-linked loci and 59 autosomal loci. I performed a sign test on the combined data to test the null hypothesis of equal levels of diversity in African and non-African populations (table 2). As expected (Begun and Aquadro 1993), X-chromosome loci had reduced levels of nucleotide diversity in non-African populations relative to African populations (17:3; $P = 0.001$). In contrast, loci on the autosomes were about as likely to be more variable outside of Africa as within Africa (28:33; table 2). Thus, as anticipated from a handful of earlier studies (Clark and Wang 1997; Tsaur, Ting, and Wu 1998; Begun et al. 1999; Andolfatto and Kreitman 2000), there existed an interesting discrepancy between the patterns observed on the X chromosome and those observed on the autosomes. Patterns for the X chromosome and the autosomes were significantly different from each other (two-tailed Fisher's exact test; $P = 0.017$). There is no detectable difference between chromosomes 2 and 3.

The data of Teeter et al. (2000) present the advantage of a consistent sampling scheme, since a similar set of lines were sampled for all markers (see *Materials and Methods*). In addition, unlike in many earlier studies, these loci were not chosen because of prior evidence for the action of natural selection. With the analysis restricted to these loci, a significant trend toward higher nucleotide diversity within Africa was apparent for the X-linked loci (10:2; $P = 0.019$), while, again, no clear trend emerged for autosomal loci (23:18; $P = 0.264$).

Demographic events such as a bottleneck are expected to affect the whole genome similarly (i.e., they are expected to result in lower levels of variation outside Africa). While genomes with a smaller effective population size will recover faster from changes in population size (e.g., Fay and Wu 1999), it is unclear how relevant this will be to the interpretation of X-autosome comparisons (see arguments below). The discrepancy between the X chromosome and the autosomes is therefore difficult to explain with a simple bottleneck in the history of non-African populations. In order to reconcile all the data with the bottleneck hypothesis, we would have to invoke post hoc differences between non-African and African populations of *D. melanogaster.*

X-Chromosome Versus Autosome Nucleotide Diversity in *D. melanogaster*

Table 3 summarizes mean synonymous site diversity levels in African and non-African samples for X-linked and autosomal loci in *D. melanogaster* and *D. simulans.* Comparisons of X and autosome levels of diversity were complicated by possible differences in their effective population sizes. Assuming equal sex ratios and no selection, a simple correction is to multiply X diversities by 4/3 (based on relative numbers of X chromosomes and autosomes). However, if sexual selection on males is prevalent in natural populations of Drosophila (cf. Andersson 1994), then the ratio of effective sizes of the X chromosome and autosomes may be closer to unity (Caballero 1995). Sex-specific life history traits (e.g., mortality rates) in the wild may further complicate the appropriate scaling of levels of variation on the X chromosome and the autosomes (B. Charlesworth, personal communication). Unfortunately, we remain virtually ignorant of the relative importance of sexual selection and sex-specific life history traits in natural populations of *D. melanogaster* and *D. simulans.* Laboratory measurements have suggested that the effective population size of females is greater than that of males (Crow and Morton 1954). Thus, the appropriate correction for X-chromosome diversity may be less than 4/3.

Whatever correction factor was used, an inescapable problem emerged when considering levels of nucleotide diversity in *D. melanogaster* (table 3). When X-chromosome variation levels were left uncorrected (as they appear in table 3), there was a trend toward lower diversity on the non-African X chromosome relative to non-African autosomes (θ_W ratio = 0.7) and somewhat elevated diversity on the African X chromosome relative to African autosomes (θ_W ratio = 1.6). If, instead, X-chromosome diversities were multiplied by 4/3, levels of variation were comparable on all chromosomes but the African X; variation on the African X chromosome was twice that of African autosomes. To establish statistical significance for these trends will require data from consistently sampled populations.

In summary, the chromosome-specific differences between African and non-African populations of *D. melanogaster* (tables 2 and 3) make a simple demographic explanation, such as a bottleneck, improbable. Additional factors must be invoked to explain the discordant geographic patterns for the X chromosome versus the autosomes. Regardless of whether or not a bottleneck occurred in the history of non-African populations, we are left to explain why X-chromosome diversity appears to

Table 1
Summary of Single-Nucleotide Polymorphism in African and Non-African Population Samples

		Non-African			African			Ratio	
	CYT	*n*	*S*	θ_W	*n*	*S*	θ_W	nAf/Af	Sign
Drosophila melanogaster									
achaete	1B	7	5	0.0045	3	4	0.0059	0.77	+
asense	1B	4	0	0.0000	2	5	0.0047	0.00	+
su(s)	1B	50	10	0.0007	50	41	0.0028	0.24	+
1C	1C	5	0	0.0000	2	2	0.0133	0.00	+
su(w)	1E	50	17	0.0019	50	49	0.0056	0.35	+
2B	2B	5	1	0.0024	2	2	0.0100	0.24	+
zeste	3A	4	1	0.0006	2	5	0.0051	0.11	+
period	3B	4	14	0.0041	2	19	0.0102	0.40	+
4C	4C	6	1	0.0026	2	1	0.0059	0.44	+
swallow	5E	8	2	0.0015	3	3	0.0040	0.39	+
8C	8C	6	2	0.0044	2	1	0.0050	0.88	+
8D	8D	6	2	0.0044	2	3	0.0150	0.29	+
Yp2	9A	4	2	0.0010	2	0	0.0000	—	−
sevenless	10A	5	3	0.0032	3	6	0.0089	0.36	+
vermilion	10A	51	43	0.0046	20	105	0.0142	0.32	+
10F	10F	6	4	0.0073	2	2	0.0083	0.88	+
13A	13A	6	2	0.0052	2	2	0.0118	0.44	+
G6pd	18E	34	31	0.0045	16	37	0.0066	0.68	+
19E	19E	6	4	0.0097	2	1	0.0056	1.75	−
Fog	20B	6	5	0.0040	3	1	0.0012	3.28	−
21D	21D	5	3	0.0096	2	3	0.0200	0.48	+
ninaA	21E	2	2	0.0040	2	6	0.0120	0.33	+
23E	23E	6	1	0.0026	2	0	0.0000	—	−
thickveins	25D	2	0	0.0000	2	5	0.0100	0.00	+
Acp26Aab	26A	39	58	0.0102	10	20	0.0052	1.94	−
27C	27C	6	4	0.0103	2	1	0.0059	1.75	−
Acp29AB	29A	12	40	0.0076	27	76	0.0113	0.67	+
30A	30A	6	1	0.0027	2	1	0.0063	0.44	+
numb	30B	2	1	0.0020	2	2	0.0040	0.50	+
ln(2L)t-St	34A	29	57	0.0152	6	31	0.0146	1.04	−
34B	34B	3	2	0.0056	2	0	0.0000	—	−
Adh-St-Slow	35B	4	25	0.0058	5	26	0.0053	1.09	−
36C	36C	6	3	0.0073	2	3	0.0167	0.44	+
Ref(2)p	37E	8	37	0.0054	2	8	0.0030	1.78	−
38B	38B	6	1	0.0022	2	1	0.0050	0.44	+
41C	41C	6	1	0.0022	2	0	0.0000	—	−
Pgi	44F	9	8	0.0012	12	11	0.0014	0.80	+
46B	46B	6	1	0.0029	2	0	0.0000	—	−
eve	46C	3	11	0.0035	2	11	0.0053	0.67	+
48C	48C	6	2	0.0052	2	1	0.0059	0.88	+
50C	50C	6	2	0.0058	2	0	0.0000	—	−
50D	50D	5	3	0.0085	2	3	0.0176	0.48	+
tra2	51B	2	1	0.0020	2	0	0.0000	—	−
51E	51E	6	1	0.0020	2	0	0.0000	—	−
Amy-d	54A	7	40	0.0110	2	8	0.0054	2.04	−
54A	54A	6	2	0.0052	2	1	0.0059	0.88	+
diptericin	56A	12	12	0.0115	3	5	0.0097	1.19	−
56B	56B	4	3	0.0117	2	2	0.0143	0.82	+
60B	60B	4	2	0.0070	2	0	0.0000	—	−
Dras3	62C	22	3	0.0015	4	3	0.0029	0.50	+
63D	63D	5	2	0.0060	2	1	0.0063	0.96	+
Dras2	64B	22	13	0.0046	4	13	0.0091	0.50	+
Tph	66A	8	4	0.0034	2	5	0.0067	0.51	+
66A	66A	5	4	0.0113	2	1	0.0059	1.92	−
67F	67F	6	4	0.0088	2	2	0.0100	0.88	+
70A	70A	5	1	0.0030	2	1	0.0063	0.48	+
frizzled	70D	7	4	0.0033	3	6	0.0080	0.41	+
anon1A3	71A	23	12	0.0033	3	6	0.0041	0.81	+
71B	71B	6	0	0.0000	2	1	0.0067	0.00	+
transformer	73A	15	4	0.0013	3	1	0.0007	1.85	−
74A	74A	5	2	0.0056	2	0	0.0000	—	−
76B	76B	6	1	0.0024	2	1	0.0056	0.44	+
77B	77B	6	1	0.0024	2	0	0.0000	—	−
80B	80B	6	0	0.0000	2	1	0.0043	0.00	+
83C	83C	6	1	0.0020	2	0	0.0000	—	−

Table 1
Continued

		Non-African			African			Ratio	
	CYT	*n*	*S*	θ_W	*n*	*S*	θ_W	nAf/Af	Sign
Gld	84D	45	9	0.0013	10	1	0.0002	5.82	−
anon1E9	85B	12	5	0.0008	3	5	0.0016	0.50	+
Dras1	85D	22	14	0.0021	4	6	0.0018	1.17	−
86C	86C	6	2	0.0044	2	0	0.0000	—	−
88B	88B	6	3	0.0055	2	0	0.0000	—	−
90A	90A	4	2	0.0064	2	0	0.0000	—	−
92B	92B	4	2	0.0084	2	2	0.0154	0.55	+
Rh3	92D	3	0	0.0000	2	2	0.0018	0.00	+
94C	94C	6	8	0.0159	2	4	0.0182	0.88	+
anon1G5	95D	14	11	0.0042	2	2	0.0024	1.73	−
Dms	96A	5	4	0.0055	3	8	0.0152	0.36	+
Boss	96F	3	6	0.0026	2	6	0.0039	0.67	+
98A	98A	6	2	0.0058	2	0	0.0000	—	−
Cec-C	99E	8	18	0.0199	5	14	0.0193	1.03	−
Tpi	99E	17	25	0.0069	8	29	0.0105	0.66	+
tailless	100B	9	6	0.0076	3	5	0.0115	0.66	+
100B	100B	4	2	0.0057	2	1	0.0053	1.09	−
Drosophila simulans									
zeste	3A	4	7	0.0052	2	13	0.0179	0.29	+
period	3B	4	28	0.0109	2	34	0.0242	0.45	+
Yp2	9A	4	3	0.0020	2	0	0.0000	—	−
vermilion	10A	24	29	0.0128	46	63	0.0236	0.54	+
G6pd	18E	21	10	0.0048	45	21	0.0083	0.58	+
Pgi	44F	11	16	0.0021	3	24	0.0063	0.33	+
eve	46C	3	10	0.0033	3	22	0.0072	0.45	+
anon1A3	71A	6	9	0.0054	6	14	0.0083	0.64	+
transformer	73A	4	17	0.0102	3	11	0.0081	1.26	−
anon1E9	85B[a]	6	62	0.0179	2	44	0.0291	0.62	+
Rh3	92D[a]	3	16	0.0094	2	16	0.0142	0.67	+
anon1G5	95D	8	28	0.0176	6	26	0.0186	0.95	+
Boss	96F	3	29	0.0126	2	24	0.0156	0.81	+

Note.—X-linked loci are indicated in bold. CYT = cytological location; Af = African; nAf = non-African; *n* = sample size; *S* = number of polymorphic sites; θ_W = Watterson's (1975) diversity estimate.

[a] Cytological position in *D. melanogaster* is given.

be elevated above that of autosomes in African populations.

A Role for Autosomal Inversions?

A possible explanation for the unexpectedly low levels of diversity for African *D. melanogaster* autosomes is the presence of common autosomal inversion polymorphisms. Based on the geographic distribution of inversions in this species (Lemeunier and Aulard 1992), it is likely that autosomal inversions are more often present in our African samples than in non-African samples. Inversions suppress crossing over within the inverted region when heterozygous. Linked neutral diversity on autosomes polymorphic for inversions may therefore be more susceptible to variation-reducing selection (and thus harbor reduced variability) than are X-linked loci (cf. Begun 1996). This explanation probably requires the persistence of inversions at appreciable frequencies. Even if this scenario explains the reduced variability on African autosomes, careful timing of inversion and bottleneck effects would be required to explain why a bottleneck pattern was not apparent in the autosomal data. Recent studies of two common inversions have suggested that they are not old and have had low historical frequencies (Wesley and Eanes 1994; Andolfatto, Wall, and Kreitman 1999). Thus, an alternative explanation is a recent change in inversion frequencies in Africa. This hypothesis predicts that nucleotide diversity of inverted chromosomes will be less than that of standard chromosomes (cf. Navarro, Barbadilla, and Ruiz 2000). However, direct comparisons of inverted and standard chromosomes (Hasson and Eanes 1996; Aguadé 1998, 1999; Andolfatto, Wall, and Kreitman 1999; Benassi et al. 1999; Depaulis, Brazier, and Veuille 1999) do not suggest markedly reduced levels of vari-

Table 2
Tests of the "No Bottleneck" Hypothesis

	Af > nAf	Af < nAf	Sign test *P*[a]
Drosophila melanogaster			
X	17	3	0.001
2	14	14	0.575
3	19	14	0.243
2 + 3	33	28	0.304
Drosophila simulans			
X	4	1	0.188
2 + 3	7	1	0.039
X + 2 + 3	11	2	0.013

Note.—Af = African, nAf = non-African.

[a] One-tailed probabilities.

Table 3
Mean Synonymous-Site Diversities for *Drosophila melanogaster* and *Drosophila simulans* by Chromosome and Geographic Locality

	Total		Non-African		African	
Locus	π	θ_W	π	θ_W	π	θ_W
D. melanogaster						
X-linked						
G6pd	0.0187	0.0242	0.0146	0.0170	0.0243	0.0260
period	0.0198	0.0206	0.0157	0.0149	0.0297	0.0297
vermilion	0.0268	0.0324	0.0207	0.0116	0.0374	0.0431
Yp2	0.0147	0.0124	0.0041	0.0044	0.0000	0.0000
zeste	0.0150	0.0182	0.0029	0.0032	0.0297	0.0297
Autosomal						
Acp26Aab	0.0171	0.0257	0.0174	0.0260	0.0107	0.0109
Acp29AB	0.0321	0.0246	0.0199	0.0230	0.0363	0.0269
Adh + Adh-dup	0.0208	0.0204	0.0160	0.0136	0.0142	0.0148
Amy-d	0.0289	0.0323	0.0334	0.0358	0.0142	0.0142
anon1A3	0.0044	0.0063	0.0037	0.0052	0.0095	0.0095
anon1G5	0.0129	0.0111	0.0113	0.0115	0.0122	0.0122
Boss	0.0170	0.0158	0.0073	0.0073	0.0165	0.0165
Cec-C	0.0304	0.0208	0.0291	0.0330	0.0248	0.0207
diptericin	0.0141	0.0286	0.0124	0.0257	0.0175	0.0176
Dras2	0.0102	0.0100	0.0050	0.0079	0.0160	0.0157
Dras3	0.0084	0.0083	0.0069	0.0065	0.0132	0.0129
eve	0.0149	0.0119	0.0165	0.0165	0.0248	0.0248
Ref(2)p	0.0039	0.0051	0.0046	0.0056	0.0024	0.0024
Rh3	0.0029	0.0034	0.0000	0.0000	0.0072	0.0072
Tpi	0.0292	0.0410	0.0235	0.0262	0.0407	0.0469
transformer	0.0051	0.0040	0.0051	0.0042	0.0046	0.0046
Means						
X-linked	0.0190	0.0216	0.0116	0.0102	0.0242	0.0257
Autosomal	0.0158	0.0168	0.0133	0.0155	0.0165	0.0161
Ratio X/autosome	1.21	1.28	0.87	0.67	1.47	1.60
D. simulans						
X-linked						
G6pd	0.0224	0.0270	0.0192	0.0281	0.0254	0.0163
Period	0.0389	0.0436	0.0282	0.0299	0.0647	0.0647
Vermilion	0.0229	0.0241	0.0226	0.0204	0.0160	0.0146
Yp2	0.0034	0.0035	0.0046	0.0043	0.0000	0.0000
Zeste	0.0295	0.0292	0.0201	0.0165	0.0532	0.0532
Autosomal						
anon1A3	0.0062	0.0079	0.0061	0.0063	0.0064	0.0084
anon1E9	0.0309	0.0317	0.0288	0.0267	0.0422	0.0422
anon1G5	0.0202	0.0323	0.0199	0.0257	0.0223	0.0264
Boss	0.0510	0.0511	0.0528	0.0527	0.0601	0.0601
Pgi	0.0082	0.0131	0.0057	0.0062	0.0173	0.0173
Rh3	0.0514	0.0569	0.0384	0.0384	0.0574	0.0574
transformer	0.0252	0.0256	0.0220	0.0190	0.0186	0.0186
Means						
X-linked	0.0234	0.0255	0.0190	0.0198	0.0319	0.0298
Autosomal	0.0276	0.0312	0.0248	0.0250	0.0320	0.0329
Ratio X/autosome	0.85	0.82	0.76	0.79	1.00	0.91

Note.—θ_W = Watterson's (1975) nucleotide diversity estimate; π = average pairwise divergence (Tajima 1983). Diversity estimates are per site.

ation in inverted chromosomes (with the exception of loci very close to inversion breakpoints).

The impact of inversions in African populations can be assessed indirectly by comparing *D. melanogaster*/*D. simulans* ratios of nucleotide diversities for the X chromosome and the autosomes. Inversions are rare in *D. simulans* (Lemeunier and Aulard 1992), so this species should not show the same pattern if inversions account for the reduced autosomal diversity in African *D. melanogaster* samples. The African *D. melanogaster*/*D. simulans* ratios of θ_W were 0.9 for the X chromosome and 0.5 for the autosomes (table 3); non-African populations had more equal ratios (0.5 and 0.6, respectively). These ratios and the estimates of mean diversities on which they were based had large standard errors. In addition, we must assume that no other factors affect X-linked and autosomal variation in *D. simulans.* This said, the pattern was consistent with the hypothesis that inversions have decreased diversity levels on the autosomes in African populations of *D. melanogaster*.

Evidence for a Bottleneck in the History of *D. simulans*

As shown in table 1, there are few data for *D. simulans* for which African/non-African comparisons of

Table 4
Summary of Synonymous and Replacement Polymorphism for Homologous Loci in *Drosophila melanogaster* and *Drosophila simulans*

	D. melanogaster					*D. simulans*				
Locus	*S*	*R*	θ_S	θ_R	θ_S/θ_R	*S*	*R*	θ_S	θ_R	θ_S/θ_R
X-linked										
asense	3	3	0.0054	0.0016		0	0	0.0000	0.0000	
G6pd	39	3	0.0242	0.0006		22	0	0.0270	0.0000	
period	19	1	0.0206	0.0003		40	6	0.0436	0.0021	
Pgd	3	1	0.0028	0.0003		15	2	0.0280	0.0012	
prune	0	1	0.0000	0.0005		8	5	0.0188	0.0038	
runt	20	3	0.0179	0.0009		8	1	0.0071	0.0003	
vermilion	42	2	0.0324	0.0005		24	4	0.0241	0.0013	
white	12	3	0.0070	0.0006		7	0	0.0090	0.0000	
Yp2	7	2	0.0124	0.0011		2	0	0.0035	0.0000	
zeste	7	0	0.0182	0.0000		11	0	0.0292	0.0000	
Autosomal										
Adh + Adh-dup	25	8	0.0173	0.0017		32	1	0.0406	0.0003	
anon1A3	5	11	0.0063	0.0040		5	11	0.0079	0.0051	
anon1E9	3	4	0.0063	0.0040		31	33	0.0317	0.0094	
anon1G5	6	4	0.0111	0.0020		17	21	0.0323	0.0111	
Boss	12	2	0.0158	0.0008		39	2	0.0511	0.0008	
ci	0	0	0.0000	0.0000		0	1	0.0000	0.0013	
cta	0	0	0.0000	0.0000		1	0	0.0017	0.0000	
eve	1	0	0.0119	0.0000		0	0	0.0000	0.0000	
Est-6	17	11	0.0166	0.0033		55	11	0.0799	0.0048	
janus	9	6	0.0141	0.0033		9	1	0.0498	0.0017	
Gld	4	5	0.0023	0.0009		26	1	0.0231	0.0003	
Mlc1	0	0	0.0000	0.0000		1	0	0.0066	0.0000	
Pgi	4	3	0.0029	0.0007		16	2	0.0131	0.0005	
relish	7	4	0.0056	0.0009		34	6	0.0255	0.0015	
Rh3	2	0	0.0034	0.0000		33	0	0.0569	0.0000	
Tpi	28	2	0.0410	0.0009		12	1	0.0257	0.0007	
transformer	2	1	0.0040	0.0006		10	3	0.0311	0.0027	
X-linked	152	19	0.0141	0.0006	22.0	137	18	0.0190	0.0009	22.1
Autosomal	125	61	0.0093	0.0014	6.8	321	94	0.0280	0.0024	11.8
Autosomal[a]	111	42	0.0096	0.0009	10.2	268	29	0.0289	0.0010	27.7

Note.—*S* = number of synonymous polymorphisms; *R* = number of replacement polymorphisms; θ_S and θ_R = Watterson's (1975) diversity estimate per site for synonymous and replacement sites, respectively. In the last three rows, *S* and *R* values are totals, and θ_S and θ_R values are averages.

[a] Excluding *anon1A3, anon1E9,* and *anon1G5* (see text).

single-nucleotide polymorphisms can be made. Combining all chromosomes, a sign test (table 2) revealed significantly lower diversity in non-African populations relative to African populations (11:2; $P = 0.013$). While this finding is consistent with the microsatellite survey of Irvin et al. (1998), a caveat in making African/non-African comparisons is the possibility of African/non-African differences in population structure. The diversity measure θ_W is sensitive to the degree of population subdivision. In particular, when populations are subdivided and both demes are sampled, θ_W will be larger than predicted for a panmictic population of the same total size (Tajima 1989*a*). Recent data from the *G6pd* locus (Hamblin and Veuille 1999) are consistent with considerable population differentiation within Africa. As an illustration, total variability at the *G6pd* locus in Africa is about twofold higher than total non-African diversity (table 1). A measure which is less sensitive to population subdivision than total diversity is the average of within-population diversities (Tajima 1989*a*). When average within-population diversities at *G6pd* were compared, African populations (θ_W within = 0.0033 per site) were more similar to non-African populations (θ_W within = 0.0036 per site). Similarly, the *vermilion* locus showed considerable between-populations differentiation within Africa (Hamblin and Veuille 1999). In summary, while the available data for *D. simulans* show a trend toward higher levels of total nucleotide diversity in Africa than outside of Africa (tables 1–3), within-population diversities are not necessarily higher in Africa. Thus, it remains unclear whether the nucleotide data support a simple bottleneck model.

Patterns of Synonymous and Replacement Polymorphism

Table 4 lists the numbers of synonymous and replacement polymorphisms and their ratios (*S*/*R*) in *D. melanogaster* and *D. simulans* for all available pairs of homologous loci. The *S*/*R* ratio for the autosomes of *D. melanogaster* was significantly lower than that for *D. simulans* autosomes (two-tailed Fisher's exact test; $P = 0.01$). In contrast, *S*/*R* ratios were surprisingly similar in the two species for the X chromosome (two-tailed Fisher's exact test; $P = 1.00$). These findings are consistent with the X-autosome contrast first noted by Begun (1996). There appears to be an excess of replacement polymorphisms and/or a deficiency of synonymous

polymorphisms on *D. melanogaster* autosomes relative to *D. simulans* autosomes. Also of interest is the pattern of within-species variability (table 4). In populations of *D. melanogaster*, the within-species *S*/*R* ratios on the X chromosome were higher than those of the autosomes. This trend was somewhat less marked in *D. simulans.* Since nonhomologous loci were compared, little can be said with certainty in X-autosome comparisons. A Fisher's exact test was performed on *S*/*R* ratios under the assumption that the mean numbers of synonymous and replacement sites do not differ between the X-linked and autosomal loci. This test revealed a significantly smaller *S*/*R* ratio on autosomes relative to the X chromosome in both species (two-tailed $P < 10^{-6}$ and $P < 0.003$ for *D. melanogaster* and *D. simulans*, respectively).

Three genes account for most of the replacement variation on *D. simulans* autosomes in table 4: *anon1A3*, *anon1E9*, and *anon1G5*. Since these genes were chosen for study specifically because they showed high rates of amino acid divergence between species (Schmid et al. 1999), they may not be representative of the ''average'' gene in the genome. Synonymous and replacement sites of the *anon* genes may experience similarly weak selection intensities. When these loci were excluded from the analyses, the conclusions changed in two ways. First, the within-species X-autosome *S*/*R* difference in *D. simulans* disappeared. Second, the between-species autosome difference in *S*/*R* ratios became much more significant (two-tailed Fisher's exact test; $P = 2 \times 10^{-6}$). While there is no a priori reason to exclude the *anon1A3*, *anon1E9*, and *anon1G5* genes from these comparisons, more confidence should be placed on the between-species X-autosome *S*/*R* difference than the within–*D. simulans* X-autosome *S*/*R* difference. The sensitivity of the results to the inclusion of the *anon* genes indicates that it may not be entirely appropriate to treat replacement and synonymous sites as discrete selected classes.

Models of Synonymous and Replacement Site Evolution

How can the above patterns of synonymous and replacement polymorphism be explained? One possibility is that the majority of amino acid replacement changes are deleterious and partially recessive (McVean and Charlesworth 1999). Allow, for the moment, the three following assumptions: (1) sites evolve independently, (2) most amino acid replacement mutations are deleterious and partially recessive, and (3) selection is stronger in *D. simulans* than in *D. melanogaster* (proportional to $N_e s$, where N_e is the effective population size and s is the mean deleterious selection coefficient). This last assumption is supported by patterns of polymorphism and divergence in these two species (Akashi 1995, 1996). Figure 1 is a schematic diagram based on figures 2 and 4 of McVean and Charlesworth (1999). The key feature is that the slope of the line relating diversity to the strength of selection is less steep on autosomes than on the X chromosome due to the recessiveness of deleterious mutations. This model neatly accounts for three

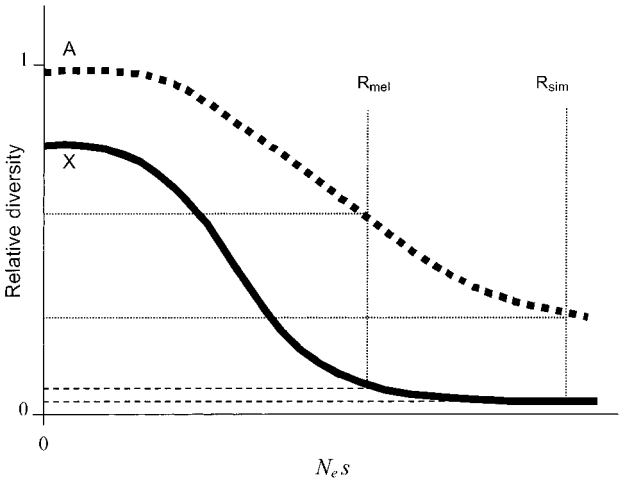

FIG. 1.—A qualitative model to account for patterns of within- and between-species patterns of synonymous and replacement polymorphism in *Drosophila melanogaster* and *Drosophila simulans* based on McVean and Charlesworth (1999). The X chromosome is represented by the heavy solid line, and autosomes are represented by the heavy dashed line. The model assumes mutation-selection-drift equilibrium, that sites evolve independently, and that deleterious mutations are partially recessive. As illustration, possible intensities of selection ($N_e s$) on amino acid replacement polymorphisms for *D. melanogaster* (R_{mel}) and *D. simulans* (R_{sim}) are indicated with vertical dotted lines. Selection intensity on synonymous polymorphisms in both species will be to the far left of the graph. The exact shape of the two lines will depend on forward and backward mutation rates and the degree to which mutations are recessive. These parameters may be different for synonymous and replacement polymorphisms.

features of the synonymous and replacement polymorphism data: (1) the within-species *S*/*R* ratio is greater on the X chromosome than on the autosomes; (2) the X-autosome *S*/*R* discrepancy is more marked in *D. melanogaster* than in *D. simulans*; and (3) as first noted by Begun (1996), the *D. melanogaster* and *D. simulans S*/*R* ratios are more discrepant on the autosomes than on the X chromosome.

A possible problem with the above model is that sites are assumed to evolve independently. An alternative explanation for the patterns of synonymous and replacement polymorphisms between species was offered by Begun (1996), who noted that transient selection on amino acid variants (cf. Gillespie 1991) may maintain amino acid variants in natural populations and cause the reduction of linked synonymous site variation. *Drosophila melanogaster* autosomes generally experience lower rates of recombination than do autosomes in *D. simulans* due to the greater extent of centromeric and telomeric suppression of crossing over (Sturtevant 1929). As a result, *D. melanogaster* autosomes should be more affected by variation-reducing selection than are those of its sister species. Transient selection on amino acid variants may also contribute to X-autosome differences within *D. melanogaster* if autosomal inversion polymorphisms reduce recombination on autosomes relative to the X chromosome. A proper evaluation of this model awaits quantification.

Both *D. melanogaster* and *D. simulans* are thought to have recently colonized Europe and the Americas.

Table 5
Summary of Synonymous and Replacement Polymorphisms by Geographic Locality

Locus	Non-African			African		
	S	*R*	*S*/*R*	*S*	*R*	*S*/*R*
Drosophila melanogaster						
X-linked						
asense	0	0		2	3	
G6pd	25	2		31	2	
period	11	0		12	1	
vermilion	14	0		41	2	
Yp2	2	0		0	0	
zeste	1	0		5	0	
Autosomal						
Acp26Aab	25	26		7	10	
Acp29AB	10	3		15	6	
Ahd + Adh-dup	14	3		19	7	
anon1A3	4	8		3	3	
anon1E9	3	4		0	0	
anon1G5	6	4		2	0	
Amy-d	31	9		5	3	
Boss	4	2		6	0	
Cec-C	4	4		2	5	
diptericin	6	5		2	3	
Dras1	1	0		0	0	
Dras2	3	0		3	0	
Dras3	3	0		3	0	
eve	1	0		1	0	
Ref(2)p	6	7		1	3	
Rh3	0	0		2	0	
Tpi	16	2		22	0	
transformer	2	1		1	0	
Totals						
X-linked	53	2	26.5	91	8	11.4
Autosomal	139	78	1.8	94	40	2.4
Drosophila simulans						
X-linked						
asense	0	0		0	0	
G6pd	10	0		21	0	
period	22	4		26	2	
vermilion	17	3		11	2	
Yp2	2	0		0	0	
zeste	5	0		9	0	
Autosomal						
anon1A3	3	6		4	8	
anon1E9	23	24		16	20	
anon1G5	11	15		10	13	
Boss	29	0		22	2	
eve	0	0		0	0	
Rh3	16	0		16	0	
transformer	5	2		4	2	
Totals						
X-linked	56	7	8.0	67	4	16.8
Autosomal	87	47	1.9	72	45	1.6

Note.—S = number of synonymous polymorphisms; R = number of replacement polymorphisms.

Thus, possible targets for recent transient or geographically localized selection are loci involved in adaptation to temperate habitats. If transient selection is shaping patterns of variability at a large number of loci, we may expect *S*/*R* ratios to be lower outside of than within Africa. As seen in table 5, *S*/*R* ratios in African and non-African populations are remarkably similar. The trend on the X toward a lower *S*/*R* ratio in non-African *D. simulans* is not significant (two-tailed Fisher's exact test; $P = 0.347$). Note that this test may lack power, since African and non-African samples share part of their genealogical histories. Nonetheless, it appears that the relatively recent range expansion from Africa, possibly accompanied by changes in effective population size, selection pressures, and changes in inversion frequencies in *D. melanogaster*, has not had a large effect on the dynamics of synonymous and replacement polymorphism in these two species.

Conclusions

Geographic patterns of nucleotide diversity in *D. melanogaster* are incompatible with a simple ''out of Africa'' bottleneck in the history of this species. In particular, the X chromosome and autosomes show distinct patterns of geographic differentiation. If a bottleneck explains the X-chromosome pattern (i.e., higher diversity in Africa than outside of Africa), a second factor is needed to explain two observations: (1) this bottleneck pattern is not observed on the autosomes and (2) diversity levels on the African autosomes are reduced relative to African X chromosomes. A possible explanation is the presence of autosomal inversions, which are generally more numerous and occur at higher frequencies in African populations. However, a quantitative assessment of the impact of inversions on autosomal diversity levels awaits more data. The pattern in *D. simulans*, although suggestive of a simple bottleneck in the history of non-African populations, could also result from ancient (i.e., African) population structure.

Drosophila melanogaster autosomes harbor an excess of amino acid replacement polymorphisms (and/or a deficiency of synonymous polymorphisms) relative to *D. simulans* autosomes, while the X chromosome shows no such pattern. Here, I have argued that several features of within- and between-species patterns of synonymous and replacement polymorphism might be explained by assuming that replacement polymorphisms are deleterious and partially recessive and that purifying selection is more efficient in *D. simulans* than in *D. melanogaster*. Generally lower recombination rates on *D. melanogaster* autosomes may also contribute to this pattern if transient selection on amino acid variants is common.

This study highlights problems encountered in comparisons of X-linked and autosomal loci. Many factors, both selective and demographic, can contribute to sex-autosome differences in levels of nucleotide diversity (Aquadro, Begun, and Kindahl 1994; Caballero 1995; Charlesworth 1996; Fay and Wu 1999). It has been proposed that X-autosome comparisons may provide a way to distinguish between background selection and positive selection in the genome (Aquadro, Begun, and Kindahl 1994). For example, if most advantageous alleles are recessive, hitchhiking models (Maynard-Smith and Haigh 1974) predict reduced diversity on the X chromosome relative to autosomes, whereas the background selection hypothesis (Charlesworth, Morgan, and Charlesworth 1993) predicts the opposite pattern (Aquadro, Begun, and Kindahl 1994). Thus, adaptation to temperate habitats may explain the larger diversity reductions on the X chromosome relative to autosomes in non-African populations of both *D. melanogaster* and

D. simulans (tables 1–3; see also Begun and Whitley 2000). However, we still have a poor understanding of how ancient population structure and recent demographic perturbations may have affected levels of nucleotide variation (in addition to the unknown impact of inversion polymorphisms in *D. melanogaster*). Comparisons of X and autosome nucleotide diversities may not be informative about the mode of selection as long as a panmictic population model remains the null hypothesis.

Acknowledgments

F. Depaulis, G. Gibson, R. Kulathinal, R. Kliman, C. Langley, K. Schmid, and S. C. Tsaur kindly provided data prior to its publication. Thanks to D. Bachtrog, B. Charlesworth, I. Gordo, G. McVean, A. Navarro, D. Rand, two anonymous reviewers, and especially M. Przeworski for helpful discussions and comments on the manuscript. P.A. was supported by an EMBO postdoctoral fellowship.

LITERATURE CITED

AGUADÉ, M. 1998. Different forces drive the evolution of the *Acp26Aa* and *Acp26Ab* accessory gland genes in the *Drosophila melanogaster* species complex. Genetics **150**:1079–1089.

———. 1999. Positive selection drives the evolution of the *Acp29AB* accessory gland protein in Drosophila. Genetics **152**:543–51.

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

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

ANDERSSON, M. 1994. Sexual selection. Princeton University Press, Princeton, N.J.

ANDOLFATTO, P., and M. KREITMAN. 2000. Molecular variation at the *In*(*2L*)*t* proximal breakpoint site in natural populations of *Drosophila melanogaster* and *D. simulans*. Genetics **154**:1681–1691.

ANDOLFATTO, P., J. D. WALL, and M. KREITMAN. 1999. Unusual haplotype structure at the proximal breakpoint of *In*(*2L*)*t* in a natural population of *Drosophila melanogaster*. Genetics **153**:1297–1311.

AQUADRO, C. F., D. J. BEGUN, and E. C. KINDAHL. 1994. Selection, recombination, and DNA polymorphism in Drosophila. Pp. 46–56 *in* B. GOLDING, ed. Non-neutral evolution: theories and molecular data. Chapman and Hall, New York.

AYALA, F. J., B. S. CHANG, and D. L. HARTL. 1993. Molecular evolution of the *Rh3* gene in Drosophila. Genetica **92**:23–32.

AYALA, F. J., and D. L. HARTL. 1993. Molecular drift of the *Bride of Sevenless* (*boss*) gene in Drosophila. Mol. Biol. Evol. **10**:1030–1040.

BEGUN, D. J. 1996. Population genetics of silent and replacement variation in *Drosophila simulans* and *D. melanogaster*: X/autosome differences? Mol. Biol. Evol. **13**:1405–1407.

BEGUN, D. J., and C. F. AQUADRO. 1993. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. Nature **365**:548–550.

———. 1994. Evolutionary inferences from DNA variation at the *6-Phosphogluconate Dehydrogenase* locus in natural populations of *Drosophila*—selection and geographic differentiation. Genetics **136**:155–171.

———. 1995. Molecular variation at the *vermilion* locus in geographically diverse populations of *Drosophila melanogaster* and *D. simulans*. Genetics **140**:1019–1032.

BEGUN, D. J., A. J. BETANCOURT, C. H. LANGLEY, and W. STEPHAN. 1999. Is the fast/slow allozyme variation at the *Adh* locus of *Drosophila melanogaster* an ancient balanced polymorphism? Mol. Biol. Evol. **16**:1816–1819.

BEGUN, D. J., and P. WHITLEY. 2000. Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. Proc. Natl. Acad. Sci. USA **97**:5960–5965.

BENASSI, V., F. DEPAULIS, G. K. MEGHLAOUI, and M. VEUILLE. 1999. Partial sweeping of variation at the *Fbp2* locus in a west African population of *Drosophila melanogaster*. Mol. Biol. Evol. **16**:347–353.

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

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. 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.

CLARK, A. G., and L. WANG. 1997. Molecular population genetics of Drosophila immune system genes. Genetics **147**:713–724.

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

COOKE, P. H., and J. G. OAKESHOTT. 1989. Amino acid polymorphisms for *Esterase 6* in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **86**:1426–1430.

CROW, J., and N. E. MORTON. 1954. Measurement of gene frequency drift in small populations. Evolution **9**:202–214.

DAVID, J. R., and P. CAPY. 1988. Genetic variation of *Drosophila melanogaster* natural populations. Trends Genet. **4**:106–111.

DEPAULIS, F., L. BRAZIER, and M. VEUILLE. 1999. Selective sweep at the *Drosophila melanogaster Suppressor of Hairless* locus and its association with the *In*(*2L*)*t* inversion polymorphism. Genetics **152**:1017–1024.

EANES, W. F., M. KIRCHNER, J. YOON, C. H. BIERMANN, I. N. WANG, M. A. MCCARTNEY, and B. C. VERRELLI. 1996. Historical selection, amino acid polymorphism and lineage-specific divergence at the *G6pd* locus in *Drosophila melanogaster* and *D. simulans*. Genetics **144**:1027–1041.

FAY, J. C., and C. I. WU. 1999. A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. Mol. Biol. Evol. **16**:1003–1005.

GASPERINI, R., and G. GIBSON. 1999. Absence of protein polymorphism in the *Ras* genes of *Drosophila melanogaster*. J. Mol. Evol. **49**:583–590.

GILLESPIE, J. H. 1991. The causes of molecular evolution. Oxford University Press, Oxford, England.

HALE, L. R., and R. S. SINGH. 1991. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. IV. Mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. Genetics **129**:103–117.

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

———. 1997. Contrasting patterns of nucleotide sequence variation at the glucose dehydrogenase (*Gld*) locus in different populations of *Drosophila melanogaster*. Genetics **145**: 1053–1062.

HAMBLIN, M. T., and M. VEUILLE. 1999. Population structure among African and derived populations of *Drosophila simulans*: evidence for ancient subdivision and recent admixture. Genetics **153**:305–317.

HASSON, E., and W. F. EANES. 1996. Contrasting histories of three gene regions associated with *In*(*3L*)*Payne* of *Drosophila melanogaster*. Genetics **144**:1565–1575.

HASSON, E., I. N. WANG, L. W. ZENG, M. KREITMAN, and W. EANES. 1998. Nucleotide variation in the Triose Phosphate Isomerase (*Tpi*) locus of *Drosophila melanogaster* and *D. simulans.* Mol. Biol. Evol. **15**:756–769.

HEY, J., and R. M. KLIMAN. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol. Biol. Evol. **10**:804–822.

HILTON, H., R. M. KLIMAN, and J. HEY. 1994. Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* complex. Evolution **48**:1900–1913.

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.

INOMATA, N., H. SHIBATA, E. OKUYAMA, and T. YAMAZAKI. 1995. Evolutionary relationships and sequence variation of *alpha-Amylase* variants encoded by duplicated genes in the *Amy* locus of *Drosophila melanogaster*. Genetics **141**:237–244.

IRVIN, S. D., K. A. WETTERSTRAND, C. M. HUTTER, and C. F. AQUADRO. 1998. Genetic variation and differentiation at microsatellite loci in *Drosophila simulans.* Evidence for founder effects in New World populations. Genetics **150**: 777–790.

KAROTAM, J., A. C. DELVES, and J. G. OAKESHOTT. 1993. Conservation and change in structural and 5′ flanking sequences of *esterase 6* in sibling Drosophila species. Genetica **88**:11–28.

KIRBY, D. A., and W. STEPHAN. 1995. Haplotype test reveals departure from neutrality in a segment of the *white* gene of *Drosophila melanogaster.* Genetics **141**:1483–1490.

KLIMAN, R. M., P. ANDOLFATTO, J. A. COYNE, F. DEPAULIS, M. KREITMAN, A. J. BERRY, J. MCCARTER, J. WAKELEY, and J. HEY. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics **156**:1913–1931.

KLIMAN, R. M., and J. HEY. 1993. DNA sequence variation at the *period* locus within and among species of the *Drosophila melanogaster* complex. Genetics **133**:375–387.

KREITMAN, M., and 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.

LABATE, J. A., C. H. BIERMANN, and W. F. EANES. 1999. Nucleotide variation at the *runt* locus in *Drosophila melanogaster* and *D. simulans.* Mol. Biol. Evol. **16**:724–731.

LACHAISE, D., L. M. CARIOU, J. R. DAVID, F. LEMEUNIER, L. TSACAS, and M. ASHBURNER. 1988. Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22**:159–225.

LANGLEY, C. H., B. P. LAZZARO, W. PHILIPS, E. HEIKINEN, and J. BRAVERMAN. 2000. Linkage disequilibria and the site frequency spectra in the *su*(*s*) and *su*(w^a) regions of the *Drosophila melanogaster* X chromosome. Genetics **156**:1837–1852.

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.

LEMEUNIER, F., and S. AULARD. 1992. Inversion polymorphism in *Drosophila melanogaster*. Pp. 339–405 *in* C. B. KRIMBAS and J. R. POWELL, eds. Drosophila inversion polymorphism. CRC Press, Boca Raton, Fla.

LUDWIG, M. Z., and M. KREITMAN. 1995. Evolutionary dynamics of the enhancer regions of *even-skipped* in Drosophila. Mol. Biol. Evol. **12**:1002–1011.

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.

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

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

NAVARRO, A., A. BARBADILLA, and A. RUIZ. 2000. Effect of inversion polymorphism on the nucleotide variability of linked chromosomal regions. Genetics **155**:685–698.

ROZAS, J., and R. ROZAS. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15**:174–175.

SCHLÖTTERER, C., C. VOGL, and D. TAUTZ. 1997. Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations. Genetics **146**:309–320.

SCHMID, K. J., L. NIGRO, C. H. AQUADRO, and D. TAUTZ. 1999. Large number of replacement polymorphisms in rapidly evolving genes of Drosophila: implications for genome-wide surveys of DNA polymorphism. Genetics **153**: 1717–1729.

SIMMONS, G. M., W. KWOK, P. MATULONIS, and T. VENKATESH. 1994. Polymorphism and divergence at the *prune* locus in *Drosophila melanogaster* and *D. simulans.* Mol. Biol. Evol. **11**:666–671.

SOKAL, R. R., and F. J. ROHLF. 1998. Biometry. W. H. Freeman, New York.

STURTEVANT, A. H. 1929. The genetics of *Drosophila simulans.* Carnegie Inst. Wash. Publ. **399**:1–62.

SUMNER, C. J. 1991. Nucleotide polymorphism in the *Alcohol Dehydrogenase* duplicate locus of *Drosophila simulans*: implications for the neutral theory. Undergraduate thesis, Princeton University, Princeton, N.J.

TAJIMA, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics **105**:437–460.

———. 1989*a*. DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-subpopulation model. Genetics **123**:229–240.

———. 1989*b*. The effect of change in population size on DNA polymorphism. Genetics **123**:597–601.

TEETER, K., M. NAEEMUDDIN, R. GASPERINI, E. ZIMMERMAN, K. P. WHITE, R. HOSKINS, and G. GIBSON. 2000. Haplotype dimorphism in a SNP collection from *Drosophila melanogaster*. J. Exp. Zool. **288**:63–75.

TSAUR, S. C., C. T. TING, and C. I. WU. 1998. Positive selection driving the evolution of a gene of male reproduction, *Acp26Aa*, of Drosophila: II. Divergence versus polymorphism. Mol. Biol. Evol. **15**:1040–1046.

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

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

WAYNE, M. L., D. CONTAMINE, and M. KREITMAN. 1996. Molecular population genetics of *Ref(2)P*, a locus which confers viral resistance in Drosophila. Mol. Biol. Evol. **13**:191–199.

WAYNE, M. L., and M. KREITMAN. 1996. Reduced variation at *concertina*, a heterochromatic locus in Drosophila. Genet. Res. **68**:101–108.

WESLEY, C. S., and W. F. EANES. 1994. Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L)Payne* in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **91**:3132–3136.

DAVID M. RAND, reviewing editor

Accepted October 25, 2000