Local Recombination and Mutation Effects on Molecular Evolution in Drosophila

Toshiyuki Takano-Shimizu

Department of Population Genetics, National Institute of Genetics, Mishima, Shizuoka-ken 411-8540, Japan

Manuscript received January 12, 1999
Accepted for publication July 6, 1999

ABSTRACT

I studied the cause of the significant difference in the synonymous-substitution pattern found in the *achaete-scute complex* genes in two Drosophila lineages, higher codon bias in *Drosophila yakuba*, and lower bias in *D. melanogaster*. Besides these genes, the functionally unrelated *yellow* gene showed the same substitution pattern, suggesting a region-dependent phenomenon in the *X*-chromosome telomere. Because the numbers of A/T → G/C substitutions were not significantly different from those of G/C → A/T in the *yellow* noncoding regions of these species, a AT/GC mutational bias could not completely account for the synonymous-substitution biases. In contrast, we did find an ~14-fold difference in recombination rates in the *X*-chromosome telomere regions between the two species, suggesting that the reduction of recombination rates in this region resulted in the reduction of the efficacy of selection in *D. melanogaster*. In addition, the *D. orena yellow* showed a 5% increase in the G + C content at silent sites in the coding and noncoding regions since the divergence from *D. erecta*. This pattern was significantly different from those at the *orena Adh* and *Amy* loci. These results suggest that local changes in recombination rates and mutational pressures are contributing to the irregular synonymous-substitution patterns in Drosophila.

DELETERIOUS mutations arise at much higher rates than advantageous and compensatory mutations (Chao 1990; Duarte *et al.* 1992; Rice 1994; Andersson and Hughes 1996), indicating the importance of the effective elimination of deleterious mutations from a population. Linkage associations between selective loci reduce the efficacy of natural selection (the so-called Hill-Robertson effect; Hill and Robertson 1966; Felsenstein 1974), and thus recombination rates affect accumulation rates of deleterious and advantageous mutations (Birky and Walsh 1988; Charlesworth 1994; Peck 1994; Barton 1995). Indeed, a positive relationship exists between within-species DNA variation and recombination rates (Aguadé *et al.* 1989; Stephan and Langley 1989, 1998; Begun and Aquadro 1992; Nachman 1997; Dvořák *et al.* 1998). However, most of the studies suggesting the significant role of recombination in between-species DNA divergence are based on comparisons among different genes or among very divergent organisms or genomes (Kliman and Hey 1993; Lynch 1997; Munté *et al.* 1997), which does not necessarily guarantee a constant selective pressure. It is highly desirable to examine the role of recombination in molecular evolution more directly.

Three findings support the hypothesis that most synonymous changes in unicellular organisms are not strictly neutral but under weak selection. First, the usage of synonymous codons is clearly nonrandom and species specific (Grantham *et al.* 1980). Second, the most frequently used codons in *Escherichia coli* and yeast correspond to the most abundant isoaccepting tRNA species (Ikemura 1981a, 1982), and the degree of codon bias of a gene is positively correlated with the cellular amount of its mRNA or protein (Ikemura 1981b; Bennetzen and Hall 1982). Third, the synonymous-substitution rate in Enterobacteriaceae is negatively correlated with the degree of codon bias (Sharp and Li 1986). Although mutational bias affects base composition to some extent (Carulli *et al.* 1993; Kliman and Hey 1994; Kliman and Eyre-Walker 1998), the data in Drosophila also provide evidence for selective constraints acting on codon usage (Shields *et al.* 1988; Sharp and Li 1989; Akashi 1994, 1995; Moriyama and Powell 1997). In particular, Akashi (1995) estimated the product of the degree of selection intensity and the effective population size as about 2 in *Drosophila simulans*.

Two recent articles reported that significant locus-lineage interaction exists in the synonymous-substitution rates among Drosophila lineages (Takano 1998; Zeng *et al.* 1998). What is more, Takano's findings suggest a higher codon bias in the *yakuba achaete-scute complex* (*AS-C*) genes than in their *melanogaster–simulans* homologues. Following Akashi's method (Akashi 1994), Takano (1998) calculated the numbers of homologous codons at which *D. yakuba* encodes a major codon and the common ancestor of *D. melanogaster* and *D. simulans* encodes a nonmajor codon (y_{maj}–ms_{non}), and similarly for the opposite configuration (y_{non}–ms_{maj}). Major co-

Author e-mail: totakano@lab.nig.ac.jp

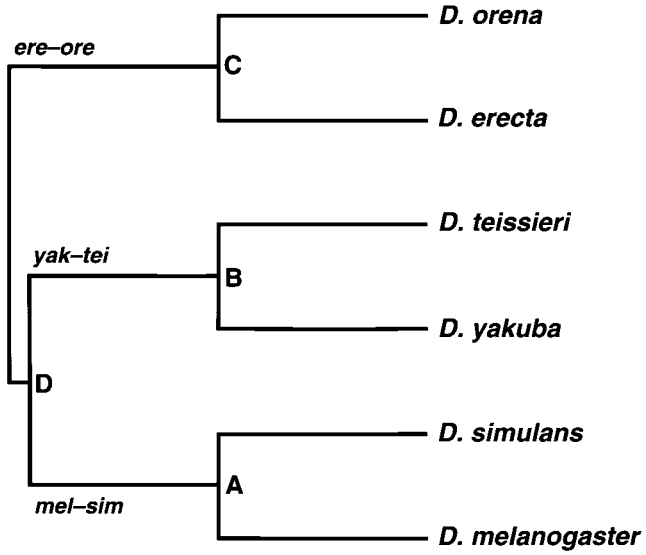

Figure 1.—Assumed phylogenetic relationships (cladogram) among the six Drosophila species. There are nine branches: three internal branches, the *mel–sim* (node D to A), *yak–tei* (node D to B), and *ere–ore* (node D to C) and six terminal branches, the *melanogaster*, *simulans*, *yakuba*, *teissieri*, *erecta*, and *orena*.

dons are the ones that appear in greater frequency in highly biased genes (Akashi 1995). These numbers were then used to obtain a ratio that would indicate codon-bias difference between the species. The ratio (y_{maj}–ms_{non} / y_{non}–ms_{maj}) was significantly higher in the four *AS-C* genes in the telomeric region of the *X* chromosome than in the other nine genes studied (Takano 1998).

A lack of outgroup species sequences, however, prevented the inference of direction of the substitutions; it thus could not be determined if there was a relaxation of codon bias in the *melanogaster* lineage or an increase in bias in the *yakuba* lineage. Another question is whether the change in selection intensity is specific to these functionally related *AS-C* genes or specific to this region. To answer these questions, I determined the DNA sequences of the two *AS-C* genes of *D. erecta*, and the *yellow* (*y*) sequences of *D. simulans*, *D. yakuba*, *D. teissieri*, *D. erecta*, and *D. orena*. Most of the previous studies suggest that *D. erecta* and *D. orena* are more distantly related to *D. melanogaster* than *D. yakuba*, as illustrated in Figure 1 (Solignac *et al.* 1986; Cariou 1987; Jeffs *et al.* 1994; Shibata and Yamazaki 1995). The *y* gene is located distally to the *AS-C* genes and is functionally unrelated to the *AS-C* genes. The results suggest a region-dependent substitution pattern. I also present evidence that *D. yakuba* has ~14 times higher recombination rates in the telomeric region of the *X* chromosome than *D. melanogaster*. It is posited that this difference results in very biased substitutions in the two lineages—a relaxation of purifying selection in the *melanogaster* lineage and a manifestation of positive selection in the *yakuba* lineage. Furthermore, this study reveals extremely GC-biased substitutions at the *orena y* locus. The *yakuba* lineage showed a significant difference in the AT/GC bias between the *y* coding and noncoding silent substitutions; the *orena* branch showed no significant difference in the substitution pattern between the coding and noncoding regions, but there were significant differences between *y* and two autosomal genes (*Adh* and *Amy*), suggesting a local mutation effect in the *orena* branch.

MATERIALS AND METHODS

DNA sequence: Directly from the products of the polymerase chain reaction (PCR), I determined the following sequences: the *yellow* (*y*) and *extramacrochaetae* (*emc*) sequences of a single male fly of *D. yakuba* (stock no. 14021-0261.0 from the National Drosophila Species Resource Center at Bowling Green, Ohio); the *y, lethal of scute* (*l'sc*), and *achaete* (*ac*) of a single male of *D. erecta* (a stock was provided by N. Inomata); the *y* sequences of a highly inbred line of *D. simulans* [Sim-5 (G20)], a single male fly of *D. teissieri* (a stock was provided by N. Inomata), and a single male of *D. orena* (a stock was provided by C. C. Laurie). The primer sequences are available upon request. The sequence data appear in the DNA Data Bank of Japan, European Molecular Biology Laboratory, and GenBank sequence databases with the following accession numbers: from AB017569 through AB017577 and from AB026336 through AB026344. The other sequences used were derived from previous research (Geyer *et al.* 1986; Geyer and Corces 1987; Villares and Cabrera 1987; Alonso and Cabrera 1988; Garrell and Modolell 1990; Martín-Campos *et al.* 1992; Takano 1998). Four *ac* and four *l'sc* sequences from *D. melanogaster*, *D. simulans*, *D. yakuba*, and *D. erecta* and six *y* sequences including *D. teissieri* and *D. orena* were aligned using CLUSTALW (Thompson *et al.* 1994), and the alignments were modified by eye in some cases. Gaps and regions with alignment difficulty were excluded from the analysis. Lengths of alignments are 1290 bp for the *y* second exon sequences, 525 bp for the 5′ flanking sequences of the *y*, 205 bp for the *y* intron sequences, 771 bp for the entire *l'sc* coding sequences, 606 bp for the entire *ac* coding sequences, and 535 bp for the partial *emc* coding sequences. The *Adh* and *Amy* proximal sequences of the same six species were analyzed with the published alignments except for two cases of the 3′ regions of the *Adh* and *Amy* (Jeffs *et al.* 1994; Shibata and Yamazaki 1995).

The numbers of substitutions on each branch were counted on the basis of the following parsimonious assumptions, without correction. The phylogenetic relationships among the six species are depicted in Figure 1, which is supported by most previous studies (*e.g.*, Jeffs *et al.* 1994). The nucleotides at the nodes were estimated as those that require the smallest number of total substitutions, and sites at which any node nucleotide could not be determined uniquely were excluded. The numbers of silent sites excluded were 3 at *ac*, 0 at *l'sc*, 14 at *y*, 21 at *Adh*, and 10 at *Amy*. In addition, two substitutions in the *mel–sim* branch occurred in one codon of the *y* gene. Because two different pathways had roughly equal weights calculated according to Miyata and Yasunaga (1980; 0.54 for the pathway to require one synonymous and one replacement substitution *vs.* 0.46 for the other to require two replacement substitutions), these substitutions were also excluded from the analysis.

I used Akashi's classification of synonymous substitutions: preferred substitution (which means substitutions from a non-

major codon to a major codon), unpreferred substitution from a major codon to a nonmajor codon, and the others (Akashi 1995).

***In situ* hybridization:** *In situ* hybridization with digoxigenin-labeled probes was performed on the salivary gland chromosomes. The genomic DNA of the three species was labeled via PCR using the following primers: 5′-TTTACTTGCGGCGA TGGTCA-3′ and 5′-AGGTGGCTTATGCTGTTCCC-3′ for the *y* locus, 5′-CGACTGCGACAAGGTGAGCAC-3′ and 5′-GGA AGTGGAGGCTGCCGACTG-3′ for *su(w^a)*, and 5′-AGGAGG ACGACATCACCAAGA-3′ and 5′-CAACGGTCTCGGCGGC AACCT-3′ for *sta*. These primers were designed from the published sequences, GenBank accession nos. X06481 (Geyer and Corces 1987), X06589 (Chou *et al.* 1987), and M90422 (Melnick *et al.* 1993).

Measuring recombination frequency: I used three lines for each species: *y white* (*w*) stock and two isofemale lines for *D. melanogaster* (one collected in Australia and the other collected in West Africa); *y w* stock and two inbred lines for *D. simulans* (one collected in Congo and the other collected in Madagascar); and three isofemale lines for *D. yakuba*, one from the National Drosophila Species Resource Center at Bowling Green, Ohio (stock no. 14021-0261.0, yak-1 in this article) and two from Umeå Drosophila Stock Center, Umeå, Sweden (stock nos. S180 and S181). Using these lines, two different G0 crosses with reciprocal crosses were done to make F_1 females for each of the three species. For *D. melanogaster* and *D. simulans*, these F_1 females were backcrossed to the parental *y w* stocks with two replications; for *D. yakuba*, they were crossed separately to two out of the three stocks used as parental stocks. The same G1 crosses were repeated twice on different days. Thus, I made 48 crosses: 2 different G0 crosses × 2 reciprocal crosses × 2 replications × 2 different days × 3 species, within 5 days. Five-day-old virgin F_1 females were crossed to 2- to 6-day-old males in a vial. The next day, all the parental flies were transferred to a new bottle and discarded 5 days later. All the emerging flies from the bottles were counted every day after the 11th or 12th day from transfer, except for 2 days. Because of fecundity differences among the three species, each cross had a different number of flies: 5 pairs of flies for *D. melanogaster*, 10 pairs for *D. simulans*, and 8 pairs for *D. yakuba*. The mean numbers of F_2 flies per bottle were 881 for *D. melanogaster*, 670 for *D. simulans*, and 795 for *D. yakuba*. All the crosses were done at 22°.

Single-strand conformation polymorphism (SSCP; Orita *et al.* 1989) and length-polymorphism markers were developed for the *y*, *su(w^a)*, and *sta* loci, but visible mutations were used for the *y* and *w* loci of *D. melanogaster* and *D. simulans*. An SSCP marker was also made for the *w* locus of *D. melanogaster* to confirm genotypes of some males. Table 1 summarizes the primer sequences and the genotype-determining methods. Because of failure of amplification of the *yakuba sta*-1 fragments for some template DNAs, the *sta*-2 primers were designed and used for amplification of these DNAs and the genotype of all recombinants was determined using the *sta*-2 primers as well as the *sta*-1. Each of the nine parental fly stocks was originated from single-pair mating and homozygous for the markers used. The *su(w^a)* primers for *D. melanogaster* were designed from the published sequence, GenBank accession number X06589 (Chou *et al.* 1987), the *sta* primers except for the *yakuba sta*-2 from M90422 (Melnick *et al.* 1993), and the others from our sequence data (data not shown). The polymerase chain reaction was done in a total volume of 10 μl including 1 unit of AmpliTaq Gold (Roche Molecular Systems,

TABLE 1

Summary of the markers

Gene	Primers	Method
D. melanogaster		
y		Visible marker
su(w^a)-1	CGACTGCGACAAGGTGAGCAC GGAAGTGGAGGCTGCCGACTG	Length variation
su(w^a)-2	GACGATGCTTCACTCAAACTA CGGATGTGCCTGACCAGACAG	SSCP of *Hin*dIII digests
sta	GGAGGACTGGAACGAGGACAC GCATAATAAGGTGGAAGCACG	SSCP
w		Visible marker
w	CGACAGGCGAGTGACAATAAA GACAACAAAGAAACGGCAATG	SSCP
D. simulans		
y		Visible marker
su(w^a)	CGACTGCGACAAGGTGAGCAC TAAACCGACATTTTCCCTGTG	SSCP
sta	GGCAAGACCTGGGAGAAGC GCGACGAGATGACGAAGAT	SSCP
w		Visible marker
D. yakuba		
y	GCAACTGCGTCTGACATCATT GCTTGTGGTTTTCTGGTGGTG	Length variation
su(w^a)	AGTAACAAACCCAAGTAGCAA CGAATCAGAGCGAGAGGAACG	Length variation
sta-1	the same primers used for *D. melanogaster*	SSCP
sta-2	ATGGCACTACATCAGCACAGC GCAGAAAATGAAAACAATAGG	SSCP

Branchburg, NJ). SSCP of amplification products was analyzed on 5–20% gradient polyacrylamide gels run at ~5° in a buffer circulation system (AE-6370; Atto, Tokyo) as in Hongyo *et al.* (1993); the length polymorphism was analyzed on 3% Agarose-21 gels (Nippon Gene, Toyama, Japan).

I determined only F_2 male genotypes from the molecular markers because typing of *X*-chromosome markers was easier for males than females. This probably did not affect the results seriously because no significant F_2 sex difference was found in the recombination fraction for the *y* and *w* loci for *D. melanogaster* (83/6854 in males and 105/7245 in females) and *D. simulans* (97/5159 in males and 104/5566 in females) and because there was no significant difference between the numbers of F_2 males and F_2 females for *D. yakuba* (6302 males and 6422 females). Genotypes were determined for all *melanogaster* and *simulans* recombinants between the *y* and *w* loci, except for four accidental losses of *simulans* recombinants and neglected double crossovers between the *y* and *w*; genotypes were determined for randomly chosen *yakuba* F_2 males, 40 to 47 males from each bottle.

RESULTS

Synonymous-substitution bias in the *melanogaster* and *yakuba* lineages: Using the *D. erecta* sequences as an outgroup, I estimated the number of synonymous substitutions along the *melanogaster* and *yakuba* lineages at the *y*, *ac*, and *l'sc* loci. The results indicated an accumulation both of unpreferred synonymous substitutions in the *mel–sim* and *melanogaster* branches and of preferred ones in the *yak–tei* and *yakuba* branches (Table 2). Although there was no evidence for a relaxation of codon bias in the *melanogaster* lineage at *y*, the ratio of the number of preferred substitutions to that of unpreferred substitutions was significantly higher in the *yak–tei* + *yakuba* branches than the *mel–sim* + *melanogaster* branches (P from Fisher's exact test < 0.01, for 18 and 1 *vs.* 11 and 8). The same result was obtained for the pooled data of the *ac* and *l'sc* genes ($P < 0.001$, G' with Williams' correction = 17.6, for 18 and 5 *vs.* 7 and 25). Together with a lack of substitution bias in the nine genes from other genomic regions in the (*yakuba* + *yak–tei* + *mel–sim*; Takano 1998), the findings suggest region-specific changes in substitution pattern.

The substitutions at the *emc* gene on the third chromosome did not support a function-dependent-change hypothesis either. *emc* functions as an antagonist to the *AS-C* genes and there are dosage-sensitive interactions between *emc* and the proneural genes such as *ac* and *sc* (Moscoso del Prado and Garcia-Bellido 1984; Ellis *et al.* 1990; Garrell and Modolell 1990). Selection intensity acting on the *AS-C* genes might have changed between the two species, still keeping a balance between the expression levels of the *AS-C* and *emc* genes. Under this hypothesis, we expect the same substitution pattern at the *emc* gene as the *AS-C* genes; however, the number of homologous codons at which *D. melanogaster* encodes a major codon and *D. yakuba* encodes a nonmajor codon was almost the same as the number of codons in the

TABLE 2

Biased synonymous-substitution pattern in *D. melanogaster*, *D. yakuba*, and *D. erecta*

Gene	*mel–sim* + *melanogaster*			*simulans*			*yak–tei* + *yakuba*			*ere–ore* + *erecta*		
	Preferred	Unpreferred	Others	Preferred	Unpreferred	Others	Preferred	Unpreferred	Others	Preferred	Unpreferred	Others
yellow	11	8	8				18	1	8	13	9	6
	(3)[a]	(0)[a]	(3)[a]	2	4	6						
achaete	3	9	4				5	2	1	7	1	3
	(1)[a]	(4)[a]	(0)[a]	5	1	1						
lethal of scute	4	16	4				13	3	4	4	1	1
	(1)[a]	(2)[a]	(1)[a]	1	1	0						
Sum	18	33	16				36	6	13	24	11	10
	(5)[a]	(6)[a]	(4)[a]	8	6	7						

A highly significant difference exists in the numbers of preferred and unpreferred synonymous substitutions between (*mel–sim* + *melanogaster*) and (*yak–tei* + *yakuba*) ($P < 0.001$, G' with Williams' correction = 25.4, for 18 and 33 in *mel–sim* + *melanogaster vs.* 36 and 6 in *yak–tei* + *yakuba*). A significant difference also exists in the same numbers between (*mel–sim* + *melanogaster*) and (*ere–ore* + *erecta*) ($P < 0.005$, G' with Williams' correction = 9.2).

[a] The number of substitutions in the *melanogaster* branch.

TABLE 3

Lack of mutational bias in the noncoding regions at the *yellow* locus

Region	*mel–sim + melanogaster + simulans*			*yak–tei + yakuba + teissieri*		
	A/T → G/C	G/C → A/T	Others	A/T → G/C	G/C → A/T	Others
5′ flanking and intron	21 (13 + 3 + 5)[a]	27 (13 + 8 + 6)[a]	14 (7 + 5 + 2)[a]	13 (5 + 4 + 4)[b]	15 (4 + 7 + 4)[b]	8 (2 + 5 + 1)[b]
Synonymous substitutions in coding	14 (8 + 3 + 3)[a]	16 (10 + 0 + 6)[a]	9 (3 + 3 + 3)[a]	32 (17 + 3 + 12)[b]	7 (3 + 0 + 4)[b]	6 (3 + 1 + 2)[b]
	$G' = 0.1$ ($P > 0.50$)[c]			$G' = 9.2$ ($P < 0.005$)[c]		

[a] The numbers of substitutions in the *mel–sim*, *melanogaster*, and *simulans* branches are given individually.
[b] The numbers of substitutions in *yak–tei*, *yakuba*, and *teissieri*.
[c] The results of a *G*-test with Williams' correction for the numbers of two classes of substitutions, A/T → G/C and the reverse, between the coding synonymous substitutions and the noncoding silent ones are given. Because none of 12 tests of independence [3 combinations of branches (*e.g.*, *mel–sim vs. mel*, *mel–sim vs. sim*, and *mel vs. sim*) × 2 regions (noncoding and coding) × 2 (*melanogaster–simulans* and *yakuba–teissieri*)] was significant, the number of the substitutions in the three branches (*mel–sim*, *melanogaster*, and *simulans*; *yak–tei*, *yakuba*, and *teissieri*) were pooled.

opposite configuration (nine and eight, respectively), implying a lack of substitution bias at this locus.

Taken together, the present results suggest that the biased synonymous-substitution pattern was a region-dependent phenomenon in the *X*-chromosome telomere and not specific to functionally related genes.

AT/GC mutational bias at the *y* locus: Because all the major codons are C- or G-ending codons (Akashi 1995), a hypothesis to explain the above results is that there is A/T → G/C (A or T to G or C) mutational bias in the *yakuba* lineage and G/C → A/T mutational bias in the *melanogaster* lineage, specifically in the telomeric region of the *X* chromosome. Comparing AT/GC biases between the coding and noncoding regions of the *y* gene, I examined mutational biases and their relationship with codon-usage biases. Table 3 reveals an absence of AT/GC-biased substitutions in the 5′ flanking region and the intron both in the *mel–sim*, *melanogaster*, and *simulans* branches and in *yak–tei*, *yakuba*, and *teissieri*. This table also reveals a significant heterogeneity in the numbers of A/T → G/C silent substitutions and the reverse between the coding and the noncoding regions in the *yakuba* and *teissieri* lineages. The data thus contradicted the AT/GC-biased mutation hypothesis.

Unexpectedly, the *orena y* locus showed extremely GC-biased substitutions, which has increased the G + C content at silent sites by ~5% since the divergence from *D. erecta* (Table 4). The ratio of the number of A/T → G/C substitutions to the number of the reverse in the *orena* branch was significantly different from those at *Adh* (P from Fisher's exact test <0.01, for 54 and 3 *vs.* 13 and 6) and *Amy* loci (P from Fisher's exact test $<10^{-6}$, for 54 and 3 *vs.* 3 and 9). This ratio was also significantly different from those at *y* in the *ere–ore* branch (P from Fisher's exact test <0.002, for 54 and 3 *vs.* 15 and 8) and in *erecta* (P from Fisher's exact test $<10^{-4}$, for 54 and 3 *vs.* 5 and 8). This suggests that the local mutation pressure changed in the *orena* branch markedly at the *y* locus. However, because the substitutions in the *Adh* noncoding regions also showed a significant A/T to G/C bias ($P < 0.005$, G' with Williams' correction = 9.4, for 11 *vs.* 1), the change is not specific to the *y* locus, but the degrees of bias may depend on regions.

This change, in turn, probably led to the higher substitution rate in the *orena* branch compared with the *erecta*. Indeed, the Tajima's 1*D* test (Tajima 1993) showed a significant departure from the molecular clock hypothesis ($P < 0.001$, $\chi^2 = 20.0$, for 60 in *orena vs.* 20 in *erecta*; $P < 0.001$, $\chi^2 = 14.5$, for 41 *vs.* 13 synonymous substitutions in the coding region). Consistently with a lack of strong mutational bias, the same tests for the *Adh* and *Amy* data showed no significant deviation ($P > 0.1$, $\chi^2 = 2.7$, for 28 in *orena vs.* 17 in *erecta* at *Adh*; $P > 0.5$, $\chi^2 = 0.4$, for 14 in *orena vs.* 11 in *erecta* at *Amy*).

Difference in recombination rates in the telomeric region of the *X* chromosome between *D. melanogaster* and *D. yakuba*: A difference in recombination rates is another possible cause of the interspecific difference in codon biases. Because of linkage to deleterious mutations and advantageous mutations that occur in the vicinity, a reduction of recombination rate leads to a reduction of effective population size in the region involved (Hill and Robertson 1966; Maynard Smith and Haigh 1974; Kaplan *et al.* 1989; Begun and Aquadro 1992; Charlesworth *et al.* 1993; Charlesworth 1994; Peck 1994; Barton 1995; Santiago and Caballero 1998). The lower the recombination rate, the smaller the effective size. The telomere of the *X* chromosome of *D. melanogaster* (the *y-AS-C* region) is a region of severely reduced recombination (Lindsley and Sandler 1977).

The interspecific difference in the degree of codon bias suggested that natural selection at the *AS-C* and *y* genes acts more effectively in *D. yakuba* than in *D. melanogaster*. To explain the observed region-specific substitution patterns in the *melanogaster* and *yakuba* lin-

TABLE 4

Mutational bias at the *yellow* locus on the *D. orena* branch

Gene	Region	*ere–ore*			*erecta*			*orena*		
		A/T → G/C	G/C → A/T	Others	A/T → G/C	G/C → A/T	Others	A/T → G/C	G/C → A/T	Others
yellow	Noncoding	4	5	2	2	1	4	16	2	1
	Synonymous substitutions									
	in coding	11	3	1	3	7	3	38	1	2
	Sum	15	8	3	5	8	7	54	3	3
Adh	Noncoding	6	11	8	5	0	5	11	1	8
	Synonymous substitutions									
	in coding	1	5	2	4	2	1	2	5	1
	Sum	7	16	10	9	2	6	13	6	9
Amy-p	Noncoding	3	2	0	2	1	2	2	5	2
	Synonymous substitutions									
	in coding	0	0	1	3	1	2	1	4	0
	Sum	3	2	1	5	2	4	3	9	2

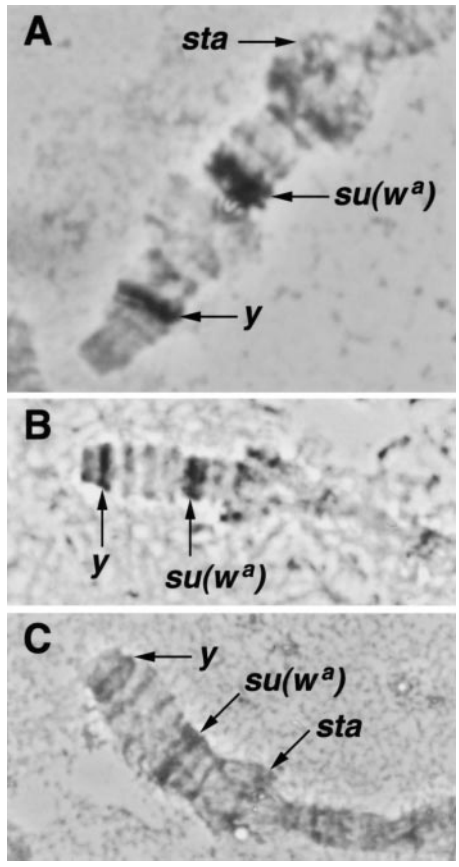

Figure 2.—Localization of the *y*, *sta*, and *su(w^a)* genes on the polytene chromosomes of *D. yakuba* (A), *D. erecta* (B), and *D. melanogaster* (C).

eages, we expected a higher recombination rate of this region in *D. yakuba* than in *D. melanogaster*. To test this, I studied recombination frequency in the regions of *y*, *suppressor of apricot* [*su(w^a)*], and *stubarista* (*sta*) genes, where the three genes of *D. yakuba* are all in the tip of the *X* chromosome and in the same order as in *D. melanogaster*, as shown in Figure 2. The standard genetic map positions in *D. melanogaster* are 0.0 for *y*, 0.1 for *su(w^a)*, 0.4 for *sta*, and 1.5 for *w* (Lindsley and Zimm 1992). The fulfillment of the prediction is shown in Table 5. The recombinant fraction in *D. melanogaster* was consistent with the published map positions, but the magnitude of recombination frequency in *D. yakuba* was >10 times higher than that in *D. melanogaster* both in the *y-su(w^a)* and in the *su(w^a)-sta* region. The recombination frequency in the *y-sta* region of *D. yakuba* did not differ between the two F_1 genotypes (12/337 for yak-1/S181 and 13/347 for S180/S181). The estimates for *D. simulans* were even lower than those of *D. melanogaster*.

Some autosomal multiple inversions in the heterozygous condition are known to increase crossing-over frequencies at the tip of the *X* chromosome (interchromosomal effect). For instance, the simultaneous presence of the heterozygous autosomal inversions, *In(2L)Cy* + *In(2R)Cy* and *In(3L)Payne* + *In(3R)Payne*, increases the

TABLE 5

Recombination frequencies and their 95% confidence limits in the telomeric region of the *X* chromosome

Species	*y-w*	*y-su(w^a)*	*su(w^a)-sta*	Entire *y-sta* region
D. melanogaster	0.0121 (83/6854) [0.0097–0.0150]	0.0015 (10/6854) [0.0007–0.0027]	0.0012 (8/6854) [0.0005–0.0023]	0.0026 [0.0016–0.0041]
D. simulans	0.0188 (97/5159) [0.0153–0.0229]	0.0008 (4/4946^a) [0.0002–0.0021]	0.0002 (1/4946^a) [0.0 – 0.0011]	0.0010 [0.0003–0.0023]
D. yakuba		0.0219 (15/684) [0.0123–0.0359]	0.0146 (10/684) [0.0070–0.0267]	0.0365 [0.0238–0.0535]

[a] Because of accidental losses of four recombinants between the *y* and *w* loci, the effective number was obtained by (93/97) × the total number of emerging flies (5159). The confidence limits were calculated as in Stevens (1942) and are shown in brackets.

recombination frequency at the tip of the *X* chromosome by about three times (Schultz and Redfield 1951). Cytogenetic study revealed an inversion on the second chromosome [presumably *In(2L)m* in Lemeunier and Ashburner 1976] in yak-1/S181 heterozygotes, but none on the second and third chromosomes in S180/S181 heterozygotes. A lack of difference in recombination frequencies between the yak-1/S181 and S180/S181 heterozygotes is consistent with the previous finding that the heterozygous, naturally occurring inversions *In(2L)t* and *In(3R)P* do not increase recombination frequencies at the tip of the *X* chromosome (Sniegowski *et al.* 1994). Consequently, the interchromosomal effect of heterozygous autosomal inversions cannot explain ~14 times higher recombination frequencies in *D. yakuba*.

In sum, these results are best explained by the hypothesis that the degree of recombination rate affects the effective population sizes of local chromosomal regions and consequently the efficacy of natural selection.

Effect of changes in recombination rates on efficacy of selection and substitution rates: Using a stabilizing-selection model, Kimura (1981) first studied codon-usage bias and synonymous-substitution rates within the framework of the neutral theory. Li (1987) and Bulmer (1991) subsequently studied the same problem by using selection-mutation-drift models. Both studies concluded that a very weak selective difference between synonymous codons produces a strong bias. These models predict higher evolutionary rates of genes with weaker selection intensity (Gillespie 1994; Eyre-Walker and Bulmer 1995). On the other hand, despite the great difference in the patterns of synonymous substitution between the *melanogaster* and *yakuba* lineages, the total number of synonymous substitutions observed was not significantly different between the two lineages (67 in the *mel–sim* and *melanogaster* branches and 55 in the *yak–tei* and *yakuba* branches from Table 2). It may be that not enough time has elapsed to be able to detect effects of the changes in recombination rates of the two lineages on the synonymous-substitution rates. To examine this possibility, I studied the dynamics of codon bias in a population that was not in equilibrium.

The above expectation of higher evolutionary rates of weakly biased genes is not always the case when a population is not in statistical equilibrium. To illustrate this point, consider a population of N diploid individuals at mutation-selection-drift equilibrium and assume two states, major and nonmajor, at each codon. For simplicity, assume that the actual population size (N) is equal to the effective size (N_e). Assume genic selection, and let s be the selective advantage of major codons over nonmajor codons in the homozygous condition. Write u for the mutation rate from a major codon to a nonmajor codon and v for the reverse mutation rate, and further assume $N_e(u + v) \ll 1$. The substitution rate per codon divided by the mutation rate ($u + v$), k, is given by Eyre-Walker and Bulmer (1995),

$$k = \{2Npu\phi(-S) + 2N(1 - p)v\phi(S)\}/(u + v), \quad (1)$$

where p is the frequency of major codons, $S = 2N_e s$, and $\phi(-S)$ ($\phi(S)$)is the ultimate fixation probability of nonmajor (major) codons whose initial frequency is $1/2N$. Note that the first term is the number of substitutions from major codons to nonmajor codons and the second term is the number of the reverse substitutions. $\phi(S)$ is given by $\phi(S) = S/(2N(1 - e^{-S}))$ (Fisher 1930; Wright 1931) and the equilibrium frequency of major codons ($\hat{p}$) by $\hat{p} = e^S/(e^S + u/v)$ (Li 1987; Bulmer 1991). Figure 3A shows the substitution rate at equilibrium as a function of S when $u/v = 1.5$. Greater selection intensity (S) generally results in lower substitution rates as shown in Gillespie (1994) and Eyre-Walker and Bulmer (1995). The equilibrium frequency is a monotonically increasing function of S; however, a nonzero S gives the maximum substitution rate when $u/v > 1$. This finding may yield a lower limit for the magnitude of S. Akashi (1996) studied the base composition at twofold degenerate sites for weakly biased genes and estimated the ratio of G/C → A/T mutation rate to the reverse mutation rate as 1.5. As mentioned above, all the major codons of *D. melanogaster* are G- or C-ending

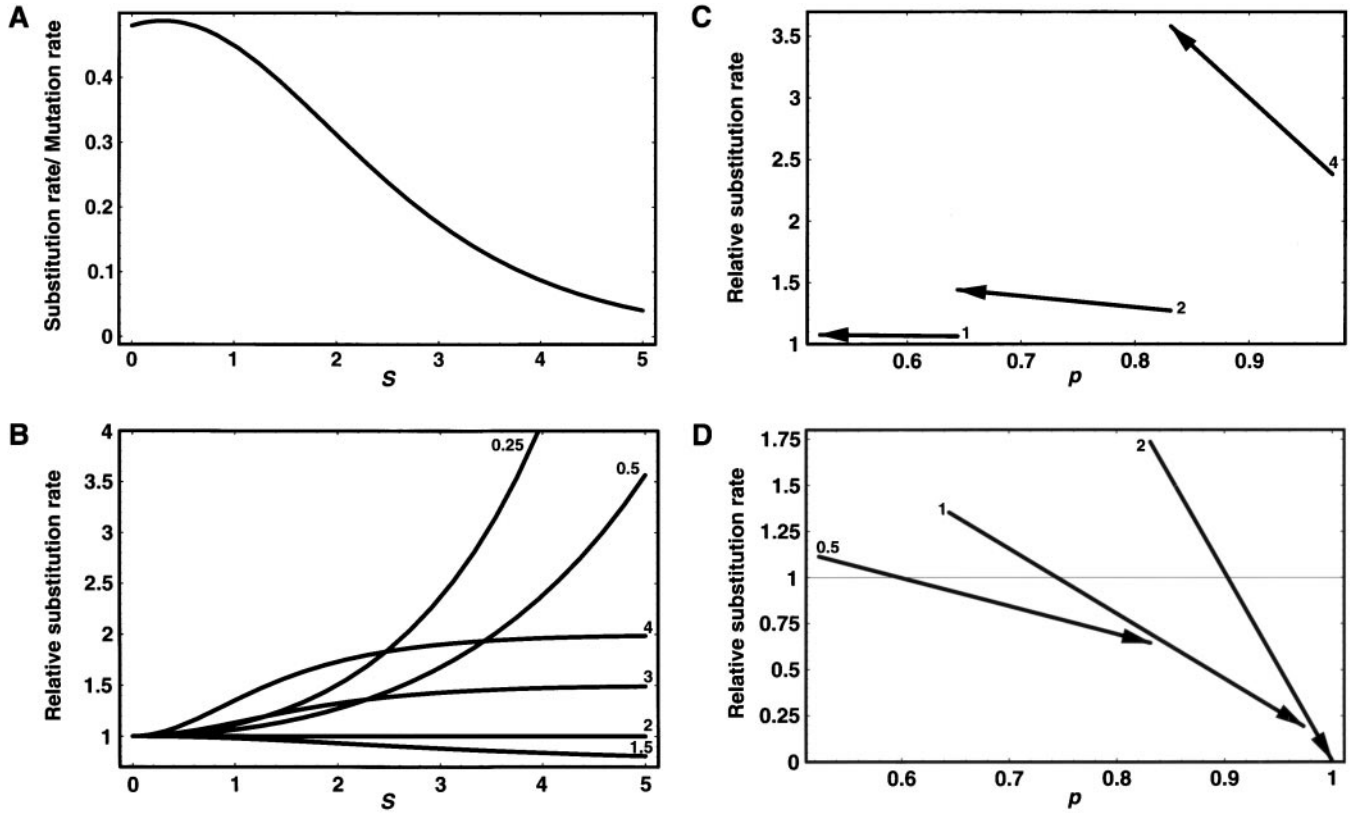

Figure 3.—Substitution rates in equilibrium and nonequilibrium conditions. The ratio of mutation rate from a major to a nonmajor codon (u) to the reverse mutation rate (v) is 1.5 in all figures. (A) Substitution rate in equilibrium state as a function of twice the product of the effective population size and selection coefficient ($S = 2N_e s$). The substitution rate is divided by the mutation rate ($u + v$). (B) Expected number of ultimate fixations of mutations that occurred in the generation when the strength of selection changed from S to aS divided by the equilibrium substitution rate before the change when $a = 0.25$, 0.5, 1.5, 2, 3, and 4. (C and D) Substitution rates as functions of the major-codon frequency (p) when $a = 0.5$; $S = 1$, 2, and 4 (C), and when $a = 4$; $S = 0.5$, 1, and 2 (D). The substitution rate is divided by the equilibrium substitution rate before the change. The frequency p decreases when $a = 0.5$ and increases when $a = 4$. The directions of the changes are shown by arrowheads.

codons, and thus natural selection acts against the mutational pressure. The substitution rate at equilibrium for $u/v = 1.5$ is maximized at about $S = 0.3$. The significant negative relationship between the substitution rate and codon-usage bias in the Drosophila genes (Shields *et al.* 1988; Sharp and Li 1989; Moriyama and Gojobori 1992) may indicate that the selection intensity (S) with respect to major and nonmajor codons is >0.3 for most genes.

Next, consider a case in which the strength of selection suddenly changes from S to aS (a is a constant). By substituting aS for S in Equation 1, we have

$$k = \frac{aS}{1 + u/v} \times \left[\left(\frac{e^{aS} - u/v}{1 - e^{aS}}\right)p + \frac{e^{aS}}{e^{aS} - 1}\right], \qquad (2)$$

where k is a function of p. Because we assume that the population is generally monomorphic [$N_e(u + v) \ll 1$], this provides an approximate substitution rate when the frequency of major codons is p. If this is not the case, a more rigorous formulation is required to compare with the actual data because the number of substitutions is a composite of mutations that have arisen at a variety of times.

Putting $\hat{p}$ before the change $[e^S/(e^S + u/v)]$ for p, we obtain the expected number of ultimate fixations when the strength of natural selection changed. The ratios of this quantity to the original substitution rate in equilibrium for $a = 0.25$, 0.5, 1.5, 2, 3, and 4 are graphically presented in Figure 3B, which shows an increase in the number of substitutions after the changes in the magnitude of S for $a < 1$ and $a > 2$. Interestingly, when $a = 2$, the quantity given by Equation 2 is exactly the same as the equilibrium substitution rate before the change. The increase when $a < 1$ is attributed to a faster accumulation of disadvantageous mutations (nonmajor codons); the increase when $a > 2$ is attributed to an excess of advantageous substitutions that leads to a new higher equilibrium frequency of major codons. A transient increase of substitution rate is also found with changes in mutation pressure (Sueoka 1993). When $u/v = 1.5$ and $S > 0.3$, the rate for $a > 1$ decreases as the frequency of major codons approaches a new equilibrium frequency, and the new equilibrium substi-

tution rate for $a > 1$ is lower than the original one as expected (Figure 3D). The substitution rates divided by the equilibrium substitution rates before the changes for $a = 0.5$ are illustrated as functions of p in Figure 3C, and those for $a = 4$ in Figure 3D. The relative substitution rate for $a = 0.5$ is always >1 (Figure 3C); that for $a = 4$ is >1 for a limited period (Figure 3D). This does not necessarily mean that a high substitution rate for $a > 2$ lasts for only a very short time. As an example, assume $S = 1$, $a = 4$, and $u/v = 1.5$. The equilibrium frequency of major codons before the change is 0.64. The relative substitution rate is >1 until p reaches 0.74, meaning a 10% increase of p. In addition, when $S = 1$ and $a = 0.5$, the relative substitution rate immediately after the change is 1.06 and that at the new equilibrium state ($aS = 0.5$) is 1.08; the relative substitution rate immediately after the change for $S = 1$ and $a = 4$ is 1.35. These transient-rate effects may account for a lack of significant difference in the number of substitutions between the *melanogaster* and *yakuba* lineages, despite a greater selection intensity in the *yakuba* lineage compared with the *melanogaster*.

DISCUSSION

The significant role of recombination in molecular evolution is suggested by three lines of evidence: (i) a significantly reduced codon bias in the regions of very reduced recombination rates compared with the other regions in Drosophila (Kliman and Hey 1993; Comeron *et al.* 1999); (ii) higher evolutionary rates, probably due to higher accumulation rates of mildly deleterious mutations in tRNA genes in organelles with no effective recombination compared with nuclear tRNA genes (Lynch 1997), although the recombination effects could not be uncoupled from the population-size effects because the organelles probably have lower N_e than the nuclear genomes; and (iii) nonrecombining genomes that experience multiple rounds of bottlenecks exhibit lower fitness than recombining genomes under the influence of the same level of random genetic drift (Rice 1994). In addition, a recent study reveals that *y* and *scute* (one member of the *AS-C*) have the highest synonymous substitution rates between *D. melanogaster* and *D. subobscura* among 18 genes compared, and that the codon bias for these genes is much higher in *D. subobscura* than in *D. melanogaster* (Munté *et al.* 1997). Because of chromosome rearrangements, the *y* and *AS-C* genes of *D. subobscura* seem to be in regions of normal recombination rates.

The above data are all consistent with the hypothesis that reduction of recombination rates and population sizes leads to relaxation of selective pressure. Most of these studies, however, are based on comparisons among different genes or among very divergent organisms or genomes, which do not necessarily guarantee a constant selective pressure. The higher substitution rates in nonrecombining genomes, for instance, may be due to a change in the magnitude of selection coefficients. In contrast, we have compared the codon biases of the same genes between closely related species, in which much less environmental and background effects are expected. More important, these findings suggest, simultaneously, the action of positive selection for synonymous substitutions in the *yakuba* lineage and relaxation of purifying selection against these in the *melanogaster* lineage. One caveat is that we do not know the recombination rate in the ancestral population, which should presumably lie somewhere between those of the two species.

Major and nonmajor codons are defined for the *D. melanogaster* genes but not for the *D. yakuba* genes; thus, some between-species differences in major codons may exist. This potential difference, if any, cannot explain the region-dependent difference in codon-usage biases between the two species. Indeed, the number of (y_{maj}–ms_{non}) codons for the genes located in other genomic regions is about the same in the opposite configuration (y_{non}–ms_{maj}); a significant heterogeneity in these numbers exists between the *AS-C* genes and the others studied (Takano 1998). A previous study also shows an absence of differences in codon usage between *D. melanogaster* and *D. pseudoobscura*, where the latter species is more distantly related to the former species than *D. yakuba* (Akashi and Schaeffer 1997). Therefore, the absence of a detailed knowledge of codon usage in *D. yakuba* does not seriously affect the conclusion.

As shown in Figure 3, the synonymous substitution rate depends on the efficacy of natural selection ($S = 2N_e s$) which, in turn, depends on the recombination rate because of linkage to deleterious and advantageous mutations. Charlesworth (1994) has pointed out that changes in the recombinational environment of a locus due to chromosome rearrangements may lead to an evolutionary rate variation among different lineages. In addition, the recombination rate can change drastically without obvious structural changes (True *et al.* 1996; this study). This hypothesis differs from a model of deleterious mutations and changing population size (*e.g.*, Takahata 1987) in two ways. First, the recombination-dependent substitution-rate hypothesis does not assume deleterious effects of all mutations but allows for an excess of advantageous-mutation substitutions relative to deleterious-mutation substitutions after changes in selection intensity. Second, and more important, changes in the population size affect all genes in the genome in the same direction, but the recombination rate may change depending on the region. Actually, the pattern predicted by the recombination-dependent substitution-rate hypothesis is similar to that of the house-of-cards models studied by Ohta (1992), Iwasa (1993), and Gillespie (1994). Gillespie (1994) rejected his house-of-cards model as a mechanism for

mammalian molecular evolution because the model allows only a narrow range of selection intensity to affect the substitution rate. Akashi's estimate of 2.2 for $N_e s$ of *D. simulans* (Akashi 1995) lies in this range, and thus the model may be applied to synonymous substitutions in Drosophila. However, the recombination effects cannot explain the locus-lineage interaction effects observed in the synonymous substitutions fully. A significant interaction effect also exists in the four *AS-C* genes (Takano 1998), but there seems to be a lack of gene-dependent difference in the recombination rates among these genes.

I previously studied the substitution-rate variation among the Drosophila lineages with special attention to ancestral polymorphism (Takano 1998). In that article, I assumed a very low recombination rate in the telomeric region of the *yakuba X* chromosome, as in the *melanogaster X*, but this assumption has turned out to be clearly wrong. Thus, the quantitative evaluation of effects of ancestral polymorphism on rate variation in Drosophila also remains to be accomplished.

Surprisingly, this study reveals strongly GC-biased silent substitutions at the *orena y* locus, which has increased the G + C content at silent sites by 5% since the divergence from *D. erecta*. Although the study of the band profiles of Drosophila DNA in buoyant density gradient has shown much lower heterogeneity than mammalian DNAs (Thiery *et al.* 1976), recent studies at fine scales have revealed significant heterogeneities in G + C content among 140- to 340-kb fragments (Carulli *et al.* 1993) and in intron G + C content among genes (Kliman and Hey 1993, 1994). The present finding further suggests that local mutation pressures can change among closely related species. This local mutation effect is very likely to be responsible for the higher substitution rate in the *orena* branch compared with the *erecta*. Sueoka (1993) has shown that changes in mutational bias are followed by transient increases in substitution rates. The ratio of the total number of silent substitutions assigned for the *orena* branch to that for the *erecta* was 3.0. As for changes in selection intensity, the maximum substitution rate is obtained immediately after changes in mutation pressure. Under the assumption of $u + v =$ constant, the ratio of this maximum substitution rate to the original equilibrium rate is only 1.25 even after a change from $u/v = 1.5$ to $u/v = 10^{-6}$, suggesting an increase of v itself on the *orena* branch.

The findings on local changes in recombination rates and mutation pressures do not imply that the changes occurred only in the telomeric region of the *X* chromosome. Significant locus-lineage interaction exists in the synonymous-substitution rates among genes located in other genomic regions (Takano 1998). We badly need more data on interspecific variation in recombination rates and mutation pressures, but it is conceivable that these factors fluctuate over time in other regions as well. Indeed, we did detect a significant A/T → G/C mutational bias in the *orena Adh* noncoding regions, where, although the difference was not statistically significant, the number of substitutions in the *orena* branch was larger than that in *erecta* (28 *vs.* 17).

In sum, this study provides evidence for both local recombination and mutation effects as causes of the irregular molecular evolution at synonymous sites in Drosophila.

I thank T. Ohta and H. Tachida for suggestions, Y. Ishii for technical assistance, and L. Gilner for improving the manuscript. I also thank two anonymous reviewers and J. Hey for many helpful comments, especially one of the reviewers for pointing out an error in the original version. I thank the National Drosophila Species Resource Center, Umeå Drosophila Stock Center, N. Inomata, and C. C. Laurie for fly stocks. This work was supported by the Ministry of Education, Science, Sports and Culture of Japan.

LITERATURE CITED

Aguadé, M., N. Miyashita and C. H. Langley, 1989 Reduced variation in the *yellow-achaete-scute* region in natural populations of *Drosophila melanogaster*. Genetics **122:** 607–615.

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

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

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

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

Alonso, M. C., and C. V. Cabrera, 1988 The *achaete-scute* gene complex of *Drosophila melanogaster* comprises four homologous genes. EMBO J. **7:** 2585–2591.

Andersson, D. I., and D. Hughes, 1996 Muller's ratchet decreases fitness of a DNA-based microbe. Proc. Natl. Acad. Sci. USA **93:** 906–907.

Barton, N. H., 1995 Linkage and the limits to natural selection. Genetics **140:** 821–841.

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.

Bennetzen, J. L., and B. D. Hall, 1982 Codon selection in yeast. J. Biol. Chem. **257:** 3026–3031.

Birky, C. W., Jr., and J. B. Walsh, 1988 Effects of linkage on rates of molecular evolution. Proc. Natl. Acad. Sci. USA **85:** 6414–6418.

Bulmer, M., 1991 The selection-mutation-drift theory of synonymous codon usage. Genetics **129:** 897–907.

Cariou, M. L., 1987 Biochemical phylogeny of the eight species in the *Drosophila melanogaster* subgroup, including *D. sechellia* and *D. orena*. Genet. Res. **50:** 181–185.

Carulli, J. P., D. E. Krane, D. L. Hartl and H. Ochman, 1993 Compositional heterogeneity and patterns of molecular evolution in the Drosophila genome. Genetics **134:** 837–845.

Chao, L., 1990 Fitness of RNA virus decreased by Muller's ratchet. Nature **348:** 454–455.

Charlesworth, B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. **63:** 213–227.

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

Chou, T.-B., Z. Zachar and P. M. Bingham, 1987 Developmental expression of a regulatory gene is programmed at the level of splicing. EMBO J. **6:** 4095–4104.

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.

Duarte, E., D. Clarke, A. Moya, E. Domingo and J. Holland, 1992 Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc. Natl. Acad. Sci. USA **89:** 6015–6019.

Dvořák, J., M.-C. Luo and Z.-L. Yang, 1998 Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing Aegilops species. Genetics **148:** 423–434.

Ellis, H. M., D. R. Spann and J. W. Posakony, 1990 *extramacrochaetae*, a negative regulator of sensory organ development in Drosophila, defines a new class of helix-loop-helix proteins. Cell **61:** 27–38.

Eyre-Walker, A., and M. Bulmer, 1995 Synonymous substitution rates in Enterobacteria. Genetics **140:** 1407–1412.

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

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Garrell, J., and J. Modolell, 1990 The Drosophila *extramacrochaetae* locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. Cell **61:** 39–48.

Geyer, P. K., and V. G. Corces, 1987 Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the *yellow* locus in *Drosophila melanogaster*. Genes Dev. **1:** 996–1004.

Geyer, P. K., C. Spana and V. G. Corces, 1986 On the molecular mechanism of gypsy-induced mutations at the *yellow* locus of *Drosophila melanogaster*. EMBO J. **5:** 2657–2662.

Gillespie, J. H., 1994 Substitution processes in molecular evolution. III. Deleterious alleles. Genetics **138:** 943–952.

Grantham, R., C. Gautier, M. Gouy, R. Mercier and A. Pavé, 1980 Codon catalog usage and the genome hypothesis. Nucleic Acids Res. **8:** r49–r62.

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

Hongyo, T., G. S. Buzard, R. J. Calvert and C. M. Weghorst, 1993 'Cold SSCP': a simple, rapid and non-radioactive method for optimized single-strand conformation polymorphism analyses. Nucleic Acids Res. **21:** 3637–3642.

Ikemura, T., 1981a Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. **146:** 1–21.

Ikemura, T., 1981b Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J. Mol. Biol. **151:** 389–409.

Ikemura, T., 1982 Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes: differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. J. Mol. Biol. **158:** 573–597.

Iwasa, Y., 1993 Overdispersed molecular evolution in constant environments. J. Theor. Biol. **164:** 373–393.

Jeffs, P. S., E. C. Holmes and M. Ashburner, 1994 The molecular evolution of the alcohol dehydrogenase and alcohol dehydrogenase-related genes in the *Drosophila melanogaster* species subgroup. Mol. Biol. Evol. **11:** 287–304.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics **123:** 887–899.

Kimura, M., 1981 Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc. Natl. Acad. Sci. USA **78:** 5773–5777.

Kliman, R. M., and A. Eyre-Walker, 1998 Patterns of base composition within the genes of *Drosophila melanogaster*. J. Mol. Evol. **46:** 534–541.

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

Kliman, R. M., and J. Hey, 1994 The effects of mutation and natural selection on codon bias in the genes of Drosophila. Genetics **137:** 1049–1056.

Lemeunier, F., and M. Ashburner, 1976 Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*Sophophora*). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc. R. Soc. Lond. Ser. B **193:** 275–294.

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

Lindsley, D. L., and L. Sandler, 1977 The genetic analysis of meiosis in female *Drosophila melanogaster*. Philos. Trans. R. Soc. Lond. Biol. Sci. **277:** 295–312.

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

Lynch, M., 1997 Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes. Mol. Biol. Evol. **14:** 914–925.

Martín-Campos, J. M., J. M. Comerón, N. Miyashita and M. Aguadé, 1992 Intraspecific and interspecific variation at the *y-ac-sc* region of *Drosophila simulans* and *Drosophila melanogaster*. Genetics **130:** 805–816.

Maynard Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. **23:** 23–35.

Melnick, M. B., E. Noll and N. Perrimon, 1993 The Drosophila *stubarista* phenotype is associated with a dosage effect of the putative ribosome-associated protein D-p40 on spineless. Genetics **135:** 553–564.

Miyata, T., and T. Yasunaga, 1980 Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. J. Mol. Evol. **16:** 23–36.

Moriyama, E. N., and T. Gojobori, 1992 Rates of synonymous substitution and base composition of nuclear genes in Drosophila. Genetics **130:** 855–864.

Moriyama, E. N., and J. R. Powell, 1997 Codon usage bias and tRNA abundance in *Drosophila*. J. Mol. Evol. **45:** 514–523.

Moscoso del Prado, J., and A. Garcia-Bellido, 1984 Genetic regulation of the *achaete-scute* complex of *Drosophila melanogaster*. Roux's Arch. Dev. Biol. **193:** 242–245.

Munté, A., M. Aguadé and C. Segarra, 1997 Divergence of the *yellow* gene between *Drosophila melanogaster* and *D. subobscura*: recombination rate, codon bias and synonymous substitutions. Genetics **147:** 165–175.

Nachman, M. W., 1997 Patterns of DNA variability at *X*-linked loci in *Mus domesticus*. Genetics **147:** 1303–1316.

Ohta, T., 1992 The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. **23:** 263–286.

Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi and T. Sekiya, 1989 Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. USA **86:** 2766–2770.

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

Rice, W. R., 1994 Degeneration of a nonrecombining chromosome. Science **263:** 230–232.

Santiago, E., and A. Caballero, 1998 Effective size and polymorphism of linked neutral loci in populations under directional selection. Genetics **149:** 2105–2117.

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

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

Sharp, P. M., and W.-H. Li, 1989 On the rate of DNA sequence evolution in *Drosophila*. J. Mol. Evol. **28:** 398–402.

Shibata, H., and T. Yamazaki, 1995 Molecular evolution of the duplicated *Amy* locus in the *Drosophila melanogaster* species subgroup: concerted evolution only in the coding region and an excess of nonsynonymous substitutions in speciation. Genetics **141:** 223–236.

Shields, D. C., P. M. Sharp, D. G. Higgins and F. Wright, 1988 "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. Mol. Biol. Evol. **5:** 704–716.

Sniegowski, P. D., A. Pringle and K. A. Hughes, 1994 Effects of autosomal inversions on meiotic exchange in distal and proximal regions of the X chromosome in a natural population of *Drosophila melanogaster*. Genet. Res. **63:** 57–62.

Solignac, M., M. Monnerot and J.-C. Mounolou, 1986 Mitochon-

drial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. J. Mol. Evol. **23:** 31–40.

Stephan, W., and C. H. Langley, 1989 Molecular genetic variation in the centromeric region of the *X* chromosome in three *Drosophila ananassae* populations. I. Contrasts between the *vermilion* and *forked* loci. Genetics **121:** 89–99.

Stephan, W., and C. H. Langley, 1998 DNA polymorphism in Lycopersicon and crossing-over per physical length. Genetics **150:** 1585–1593.

Stevens, W. L., 1942 Accuracy of mutation rates. J. Genet. **43:** 301–307.

Sueoka, N., 1993 Directional mutation pressure, mutator mutations, and dynamics of molecular evolution. J. Mol. Evol. **37:** 137–153.

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

Takahata, N., 1987 On the overdispersed molecular clock. Genetics **116:** 169–179.

Takano, T. S., 1998 Rate variation of DNA sequence evolution in the Drosophila lineages. Genetics **149:** 959–970.

Thiery, J.-P., G. Macaya and G. Bernardi, 1976 An analysis of eukaryotic genomes by density gradient centrifugation. J. Mol. Biol. **108:** 219–235.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

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.

Villares, R., and C. V. Cabrera, 1987 The *achaete-scute* gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. Cell **50:** 415–424.

Wright, S., 1931 Evolution in Mendelian populations. Genetics **16:** 97–159.

Zeng, L.-W., J. M. Comeron, B. Chen and M. Kreitman, 1998 The molecular clock revisited: the rate of synonymous vs. replacement change in *Drosophila*. Genetica **102/103:** 369–382.

Communicating editor: J. Hey