Molecular and Phenotypic Variation of the *white* Locus Region in *Drosophila melanogaster*

Naohiko Miyashita and Charles H. Langley

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Manuscript received December 1, 1987
Revised copy accepted May 19, 1988

ABSTRACT

Restriction site and insertion/deletion polymorphism in a 45-kb region of the *white* locus on the *X* chromosome in *Drosophila melanogaster* was investigated for 64 *X* chromosome lines with six 6-cutter and ten 4-cutter restriction enzymes. A total of 109 polymorphisms were detected (54 restriction sites and 55 insertions/deletions). Estimated heterozygosity per nucleotide for this region (0.004–0.008) was similar to those of the *Adh* and 87A heat-shock locus regions located on the autosomes in *D. melanogaster*. This is contrary to a simple prediction based on the theory of mutation selection-balance of partially recessive deleterious mutants which predicts less variation on *X* chromosomes. Large linkage disequilibria between pairs of polymorphisms (including insertions and deletions) within the transcriptional unit (especially the 3′ end of the 1st intron) were observed. As expected from population genetics theory, linkage disequilibria between these polymorphisms were greater for those pairs that are physically closer on the restriction map. Linkage equilibrium was typically observed when the pairs of sites were separated by 2 kb or more. Although significant between-line variation in eye pigment was observed ($P < 0.05$), there is little evidence for strong associations between this phenotype and the polymorphisms at the DNA level.

RECENT reports show that population genetic variation at the DNA level in *Drosophila melanogaster* is quite high. On the average 1 out of every 200 nucleotides is heterozygous for alternative nucleotides. An obvious question concerns the significance of such DNA sequence polymorphisms to phenotypic variation, natural selection and long-term evolution. Since a significant portion of the DNA of metazoans may be involved in the regulation of gene expression, a segment of the genome with known regulatory roles is worthy of study. The *white* locus provides an opportunity to examine DNA sequence polymorphism in coding, noncoding and flanking regulatory regions and to relate these to phenotypic variation.

The *white* locus is the first gene for which a mutation was reported in *Drosophila melanogaster* (MORGAN 1910; BRIDGES and MORGAN 1923). Completely amorphic *white* mutants lack pigment in the compound eye, ocelli, testes sheath and Malphigian tubules, while retaining good fertility and viability in the laboratory. However, population cage experiments have suggested that the *white* locus can be a subject of natural selection [see JONES and PROBERT (1980) and references therein]. The locus is found on the *X* chromosome (1.5 and 3C2) (BRIDGES 1938; LINDSLEY and GRELL 1968). *white* has played an important role in the advance of general and molecular genetics [see reviews by JUDD (1987) and HAZELRIGG (1987)]. The *white* locus region has been cloned (BINGHAM, LEVIS and RUBIN 1981), and a 14-kb region which includes the structural gene has been sequenced (O'HARE *et al.* 1984). Genetic experiments have identified sites within the complex *white* locus (centromere proximal and 5′ of the transcriptional unit) that are involved in dosage compensation, interaction with the *trans*-acting regulator encoded by the *zeste* locus, pairing-dependent regulation and tissue-specific expression [see JUDD (1976, 1987) for references]. Recent transformation experiments have located some of these functions in the 5′-flanking DNA (LEVIS, HAZELRIGG and RUBIN 1985; BINGHAM and ZACHAR 1985; PIRROTTA, STELLER and BOZZETTI 1985). Despite the abundance of information on this locus, the functional nature of the *white* gene product, which apparently consists of 705 amino acids, has not been established, although involvement in *trans*-membrane pigment transport has been suggested (SULLIVAN and SULLIVAN 1975; MOUNT 1987).

GREEN (1959a) mapped a significant difference in eye pigment between two strains of wild-type flies. This suggested the existence of variation at the *white* locus affecting adult eye pigment deposition in natural populations of *D. melanogaster*. Recent studies of natural populations of *D. melanogaster* have shown an enormous amount of DNA polymorphism in and around several loci. Some of this molecular variation seems to be associated with variation in gene expression (AQUADRO *et al.* 1986; LAURIE-AHLBERG and

STAM 1987; LANGLEY *et al.* 1988). It has been found that the *white* locus region is also highly polymorphic (LANGLEY and AQUADRO 1987).

A primary purpose of the present study was to find the relationship between the variation of eye pigment level (a phenotype of the *white* locus) and the naturally occurring variation at the DNA level of the *white* locus region. This type of information is critical to understanding of the maintenance and significance of genetic variation in natural populations. From the viewpoint of population genetics, genes on the *X* chromosome in *D. melanogaster* are expected to show less variation than autosome genes for two reasons. The force of random genetic drift in reducing standing variation is expected to be somewhat stronger, since the effective population size for genes on the *X* chromosome is ¾ that of genes on the autosomes. A potentially greater reduction in the relative amounts of standing variation can be expected from natural selection against mildly deleterious mutations. The hemizygosity of *X* chromosomal genes in males every third generation will lead to a more rapid elimination and lower equilibrium frequency of partially recessive deleterious mutants (HALDANE 1927). To the extent that molecular variation is indeed mildly deleterious (OHTA 1972, 1987; KIMURA 1983), genes on the *X* chromosome, such as *white*, should be less variable. Random genetic drift may also cause linkage disequilibrium (HILL and ROBERTSON 1968; OHTA and KIMURA 1969). If linkage disequilibrium among molecular polymorphisms is due to random genetic drift, then the level of linkage disequilibrium among *X* chromosome polymorphisms should be comparable to that observed among polymorphic sites in and around autosomal loci. It has already been reported that the *white* locus region contains amount of variation at the DNA level comparable to that of the *Adh* region located on the second chromosome (LANGLEY, MONTGOMERY and QUATTLEBAUM 1982; AQUADRO *et al.* 1986) and a heat-shock locus on the third chromosome (LEIGH BROWN 1983), and that some of the polymorphic variations in the *white* locus region are in linkage disequilibrium (LANGLEY and AQUADRO 1987). However, the size and structure of the sample of LANGLEY and AQUADRO were inappropriate to draw any general conclusions other than that there are significant amounts of restriction site and insertional variation in the *white* locus region. Here these problems were investigated in more detail.

MATERIALS AND METHODS

Lines: Sixty-four *X* chromosome isogenic lines, that originated from three different collections (North Carolina, Texas and Fukuoka, Japan), were used in this experiment. The construction of these lines was described in MIYASHITA *et al.* (1986). These *X* chromosome lines have a co-isogenic background for the second and third chromosomes derived from a single highly inbred line (Ho-R). Therefore, any variation in eye pigment can be attributed to the differences among the *X* chromosomes. The effect of the fourth chromosome was assumed to be negligible.

Experimental design and eye pigment measurement: For each line, 25 pairs of flies were placed in a half-pint milk bottle, and allowed to lay eggs for 4 days. Two bottles were used for each line. Emerging flies were aged for 6 days in separate vials for each sex. After aging, flies were frozen in liquid nitrogen. Two samples of 10 heads from each sex were collected by vortexing frozen flies, and kept at −70° until the eye pigment assay. The setup of the bottles and collection of head samples were repeated twice (2 blocks). Samples from each block were assayed in one day. A total of 507 samples were used, because 5 samples were lost during the experiment.

Eye pigment was measured by a slight modification of the method of EPHRUSSI and HERALD (1944) and HAZELRIGG, LEVIS and RUBIN (1984). Only the brown eye pigment was measured, since it has been shown that the amounts of the two eye pigments (red and brown) are correlated (GREEN, 1959a). A sample of 10 fly heads from each line in a 1.5-ml Eppendorf tube was homogenized in 20 μl of TE (10 mM Tris-HCl, pH 8, and 1 mM EDTA), from which 5 μl were taken for protein assay. To the remaining homogenate, 100 μl of acidified methanol (0.1% HCl in absolute methanol) was added and homogenized again. After adjusting the volume to 1 ml with the acidified methanol, the sample was vortexed for 30 min, spun for 5 min in an Eppendorf centrifuge, and then the supernatant was transferred to a new tube. To the sample, 40 μl of 0.5% hydrogen peroxide was added, then the sample was spun for 10 min. The supernatant was used for OD measurement at 480 nm. The amount of total protein in 10 heads was determined by the method of LOWRY *et al.* (1951).

Analysis of variance: The model of analysis of variance (ANOVA) for the *X* chromosome effect on eye pigment variation followed (model 1):

$$Y_{ijkl} = u + B_i + S_j + (B*S)_{ij} + X_k + (B*X)_{ik} + (S*X)_{jk} + (B*S*X)_{ijk} + e_{ijkl},$$

where Y was the OD measurement at 460 nm, u was the overall mean, B_i was the ith block effect ($i = 1, 2$), S_j was the effect of the jth sex (j = male or female), X_k was the kth X chromosome effect ($k = 1$–64), and e_{ijkl} was the error term. The terms in parentheses represent the interaction effects. The effect of sex was assumed to be fixed, and the other effects were assumed to be random.

In order to test the effect of polymorphic variation at the DNA level on eye pigment variation, the following model was used for each polymorphism detected (model 2):

$$Y_{ijkl(k)m} = u + B_i + S_j + (B*S)_{ij} + V_k + (B*V)_{ik} + (S*V)_{jk} + (B*S*V)_{ijk} + X_{l(k)} + (B*X)_{il(k)} + (S*X)_{jl(k)} + (B*S*X)_{ijl(k)} + e_{ijkl(k)m},$$

where V_k was the effect of the kth allele (k = 0:absence or 1:presence) and $X_{l(k)}$ was the effect of the lth ccromosome within the kth allele. The other terms were the same as in model 1. Further, two-locus interaction effects were tested

with a modified model 2, where another polymorphism effect ($V'_{k'}$) was added:

$$Y_{ijkk'l(kk')m} = u + B_i + S_j + (B*S)_{ij} + V_k + (B*V)_{ik} + (S*V)_{jk} + (B*S*V)_{ijk} + V'_{k'} + (B*V')_{ik'} + (S*V')_{jk'} + (B*S*V')_{ijk'} + (V*V')_{kk'} + X_{l(kk')} + (B*X)_{il(kk')} + (S*X)_{jl(kk')} + (B*S*X)_{ijl(kk')} + e_{ijkk'l(kk')m},$$

where $V'_{k'}$ was the effect of k'th allele of the second polymorphisms ($k' = 0$:presence, or 1:absence).

For the calculation of the sums of squares for both models 1 and 2, a "GLM" procedure of the "SAS" statistical system was used (SPEED, HOCKING and HACKNEY 1978; HELWIG and COUNCIL 1979). The *F*-tests were performed according to NETER and WASSERMAN (1974).

Restriction map analysis: Two different levels of analyses were performed: one with six 6-cutter restriction enzymes (*Eco*RI, *Hin*dIII, *Bam*HI, *Pst*I, *Sac*I and *Sal*I) of standard SOUTHERN (1975) blotting technique and the other with ten 4-cutter restriction enzymes (*Alu*I, *Dde*I, *Hae*III, *Hha*I, *Hin*fI, *Msp*I, *Rsa*I, *Sau*3A, *Sau*96I and *Taq*I) as described by KREITMAN and AGUADÉ (1986a). Genomic DNA from each of 64 *X* chromosome lines was purified in CsCl gradient (BINGHAM, LEVIS and RUBIN 1981).

The purpose of the analysis with 6-cutter restriction enzymes was to get a rough picture of restriction map variation by studying a large region, in this case a 45-kb region, and to confirm a previous study on the same region of the *white* locus (LANGLEY and AQUADRO 1987). The procedures for digestion, electrophoresis, blotting and nick-translation of probes were essentially the same as described before (LANGLEY, MONTGOMERY and QUATTLEBAUM 1982; AQUADRO *et al.* 1986; LANGLEY and AQUADRO 1987), except that ZetaBind (Cuno) or GeneScreen (DuPont) was used as the blotting membrane. Six single digestions and two double digestions (*Eco*RI and *Bam*HI, *Hin*dIII and *Sac*I) were performed. The probes used in the present study were λ phage clones: λm11B.1, λm1.2 and λm11A.1 described by LEVIS, BINGHAM and RUBIN (1982).

Since the sequence of a 14-kb *white* locus region encompassing the transcriptional unit was available (O'HARE *et al.* 1984), it was possible to apply the technique of KREITMAN and AGUADÉ (1986a). By this technique, detailed variation in the restriction map within and around the *white* locus transcriptional unit was determined. The procedure of digestion, blotting, probe preparation, hybridization and nick-translation of probes was described in KREITMAN and AGUADÉ (1986a). Probes were prepared by purifying fragments from 4 plasmid subclones kindly provided by ROBERT LEVIS: pm11.5 (*Eco*RI site +6611–*Hin*dIII site +3171), pm12.5 (*Hin*dIII site +3172–*Sal*I site −670), pm12.3 (*Sal*I site −671–*Sal*I site −1532) and pm12.8 (*Sal*I site −1533–*Sal*I site −3050). As a result, 9661 bp of the *white* locus region, including the entire transcriptional unit, were studied.

RESULTS

Restriction map variation: Figure 1 summarizes the restriction map variation for the 64 *X* chromosome lines of the *white* locus region in *D. melanogaster*. The locations of polymorphisms were numbered according to the coordinate system of LEVIS, BINGHAM and RUBIN (1982), in which the insertion point of the *copia* element in w^a was assigned to be zero. A total of 327 restriction sites were scored (56 sites in the 6-cutter restriction enzyme study and 271 sites in the 4-cutter restriction enzyme study). Nineteen 6-cutter restriction sites and 42 4-cutter restriction sites were polymorphic. Estimates of variation at the DNA level from these data accommodated the fact that three of the 4-cutter restriction enzymes (*Alu*I, *Sau*3A and *Taq*I) share recognition sequences with four of the 6-cutter restriction enzymes (*Hin*dIII, *Bam*HI, *Sac*I and *Sal*I) and some insertion/deletion polymorphisms were associated with restriction site polymorphisms. This latter problem was resolved by comparing the restriction patterns of these enzymes with those predicted by the published sequence of the *white* locus region, and assuming that restriction site polymorphisms that were consistently associated with insertion/deletion polymorphisms in the same region of the DNA were caused by the deletion of part or all of the recognition sequence. A total of 109 polymorphisms (54 restriction sites and 55 insertions/deletions) were detected (Table 1). An interesting observation was that one polymorphic *Bam*HI site (location +4435) could be attributed to two differences: (1) the gain of a *Taq*I site #58, deletion of a G at +4436 or 4437, or alternatively insertion of an A between 4436 and 4437) and (2) the loss of a *Sau*3A site and the gain of a *Hin*fI site #59, both +4434, substitution of C for T). From the 54 restriction site polymorphisms, three measures of nucleotide variation were calculated. The individual population estimates were very similar to those obtained from pooling the data. The heterozygosity per nucleotide (which is approximately equal to $\theta = 4N_e\mu$, under selective neutrality, where N_e is the effective population size and μ is the mutation rate to selectively neutral alleles) was estimated to be 0.008 ± 0.005 for the region studied with 6-cutter restriction enzymes and 0.004 ± 0.002 for the region studied with 4-cutter restriction enzymes (EWENS, SPIELMAN and HARRIS 1981; HUDSON 1982). These estimated standard deviations were based on the model of no recombination (HUDSON, 1982). The indices of nucleotide diversity [π of NEI and LI (1979) and NEI and TAJIMA (1983)] were 0.009 ± 0.005 and 0.004 ± 0.002 for the 6-cutter and 4-cutter regions, respectively. The heterozygosity per nucleotide calculated according to ENGELS (1981) gave similar estimates. The estimates from the 6-cutter restriction enzyme study were in agreement with those ($\theta = 0.013 \pm 0.007$ and $\pi = 0.011 \pm 0.006$) of the previous investigation of the *white* locus region (LANGLEY and AQUADRO 1987). As found in a previous report (LANGLEY and AQUADRO 1987), these estimates of variation in the *white* locus region were similar to those reported for the *Adh* locus region located on the second chromosome (LANGLEY, MONTGOMERY

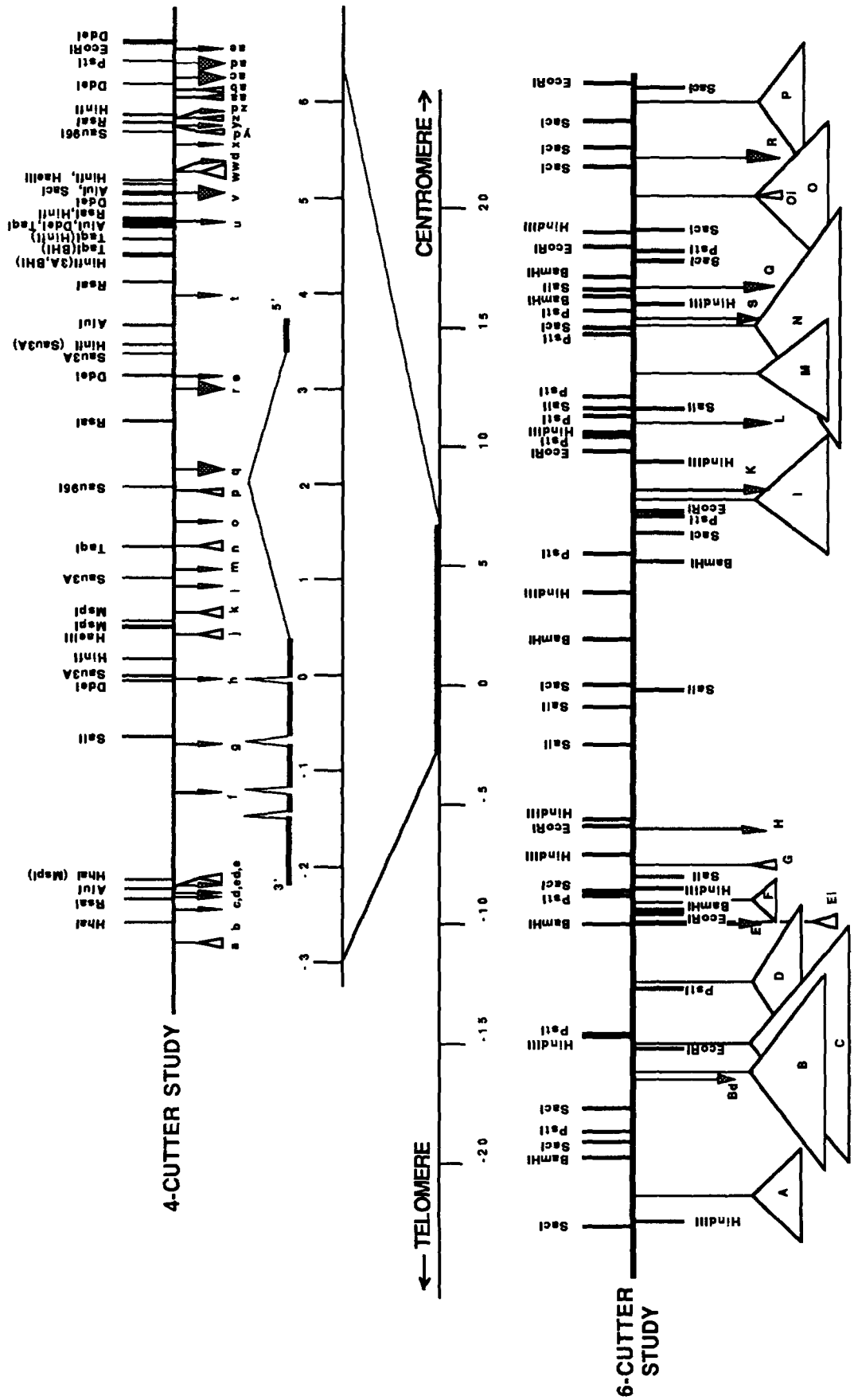


FIGURE 1.—Summary of restriction map variation in the *white* locus region of *D. melanogaster* for 64 X chromosomes. The upper map is the summary of 4-cutter restriction enzyme study on a 9.7-kb region. Only polymorphic variations are shown in the upper map (236 monomorphic restriction sites not shown). The summary of the 6-cutter study on a 45-kb region is shown in the lower map. Monomorphic sites are above the line; polymorphic variations below. Approximate sizes of insertion/deletion polymorphisms are as follows: 6-cutter study, A = 4–5 kb; B = 5–8 kb; Bd = 0.1 kb; C = 10 kb; D = 2–7 kb; E = 0.05 kb; Ei = 0.1 kb; F = 2 kb; G = 2 kb; H = 0.05 kb; I = 5.5 kb; K = 0.1 kb; L = 0.05 kb; M = 3–7 kb; N = 10 kb; O = 6 kb; Oi = 0.2 kb; P = 5 kb; Q = 0.05 kb; R = 0.05 kb; S = 0.02 kb; 4-cutter study, a = 30 bp; b = 1–2 bp; c = 1–2 bp; d = 1–2 bp; e = 1–2 bp; ed = 1–2 bp; f = 10 bp; g = 1–2 bp; h = 1–2 bp; j = 10 bp; k = 3–5 bp; l = 10 bp; m = 10 bp; n = 5 bp; o = 30 bp; p = 1–2 bp; q = 30 bp; r = 30 bp; s = 10 bp; t = 10 bp; u = 15 bp; v = 30 bp; w = 30 bp; wd = 10 bp; x = 5–20 bp; y = 1–10 bp; yd = 1–10 bp; z = 5 bp; zd = 5 bp; aa = 1–5 bp; ab = 1–2 bp; ac = 200 bp; ad = 300 bp; ae = 10 bp.

and QUATTLEBAUM 1982; KREITMAN 1983; AGUADRO *et al.* 1986), and for the 87A heat-shock locus located on the third chromosome (LEIGH BROWN 1983). Given the large standard errors there was little power to discern differences among these estimates.

Table 1 lists the haplotypes of the 64 *X* chromosome lines used in this study. As expected from the large number of polymorphisms scored (109), only one pair of lines shared the same haplotype (*i.e.*, only two chromosomes, RL23AA and RL26A, out of the 64 sampled were identical). There were still 59 haplotypes when only the 54 restriction site polymorphisms were considered; five restriction site haplotypes occurred twice in the sample. In no case were members of these pairs of haplotypes found in different geographic subsamples.

Table 2 shows the summary of χ^2 tests for the heterogeneity in the number of polymorphisms in three different regions studied with both 4- and 6-cutter restriction enzymes. The expected numbers in the comparison labeled "Restriction Sites Only" were based on the number of possible single nucleotide changes that would have been detectable in the study, *i.e.*, "site equivalents" of KREITMAN and AGUADÉ (1986a). The expected values for "Insertions/Deletions" and the total "Restriction Sites and Insertions/Deletions" were calculated by simply assuming that the number of polymorphisms were distributed proportionally to the size of the region. The test results indicated that the transcriptional unit was less variable than the 5′-flanking region. The number of polymorphisms in the 3′-flanking region was small and appeared to be intermediate. This overall pattern held for restriction sites considered separately. The number of insertion/deletion polymorphisms showed a similar tendency, although it was not statistically significant. The smaller number of observed polymorphic restriction sites in the transcriptional unit could be attributed to the selective constraints on amino acid replacement substitutions (KIMURA 1983). To test for heterogeneity in the number of silent changes, the transcriptional unit was divided into the exons and introns. In the exons, there were six restriction site polymorphisms [*Alu*I (#24, −2184), *Hha*I & *Msp*I (#25, −2089), *Sal*I (#30, −670), *Hin*fI (#34, 96), *Hin*fI & *Sau*3A (#54, 3499) and *Alu*I (#55, 3599)], of which *Hin*fI & *Sau*3A (3499) was definitely a replacement change and *Alu*I (−2184) and *Alu*I (3599) were in nontranslated regions. The expected values in the 3′-flanking, exons, introns and 5′-flanking region were calculated based on "silent site equivalents." The heterogeneity test was not statistically significant ($\chi^2 = 4.08$ with d.f. = 2, $P > 0.1$). A similar test in which noncoding exon sequences were pooled with the intron sequences was also not statistically significant ($\chi^2 = 3.83$ with d.f. = 2, $P > 0.1$). Thus, these results indicated that silent polymorphisms occurred in proportion to the size of each region, and that polymorphism of the amino acid sequence of the *white* locus was significantly lower. In the larger region surveyed with the 6-cutter restriction enzymes, no heterogeneity in the distribution of polymorphic restriction sites was detected. Nor was the distribution of detectable small insertion/deletion polymorphisms heterogeneous. But as Figure 1 clearly shows, the large insertions (>100 bp) were found exclusively in the flanking regions.

Figure 2 is the frequency spectrum of 109 polymorphic variations detected in this study. Restriction site and small insertion/deletion polymorphisms were distributed continuously from high to low frequency, while large insertion polymorphisms were (with one exception, insertion *D*) limited to very low frequencies. This observation was confirmed by significant heterogeneity of type of variation and frequency class ($\chi^2 = 16.4$ with d.f. = 6, $P < 0.05$). This test was conducted after classifying the three types of polymorphisms into four frequency classes.

Linkage disequilibrium between polymorphic variations: Figure 3 shows the significance levels of χ^2 tests for linkage disequilibrium between all the pairs of polymorphisms (including insertion/deletion variations) in which the rare alleles occurred more than once in the sample. There were 85 such polymorphisms. Since heterogeneities of both allele frequencies and two-locus haplotype frequencies were detected among the three populations, the χ^2 values calculated from each population were pooled, and this pooled χ^2 was tested with the corresponding degrees of freedom. Although significant linkage disequilibria were scattered, strong linkage disequilibria were clustered in the transcriptional unit and especially in the 1st intron. Since the power of the statistical test of linkage disequilibrium (χ^2) was affected by the marginal allele frequencies, it is important to note that the frequency distribution of polymorphisms in the transcriptional unit was not significantly different from that of 5′-flanking region, where linkage disequilibria did not seem to be clustered. Although the statistical power of tests were decreased when each of three populations was analyzed separately (because of smaller samples size in the tests), the clustering of significant linkage disequilibrium in the transcriptional unit was still observed in each population (analysis not shown).

All the pair-wise linkage disequilibria within each population in the 10-kb region studied with both 4-cutter and 6-cutter restriction enzymes were expressed as R^2 (square of correlation coefficient of variation frequencies), and ordered into 50 groups based on the physical distance between the sites. The average R^2 values and average distances were calcu-

TABLE 1

Genetic variation in the *white* locus region of *Drosophila melanogaster*

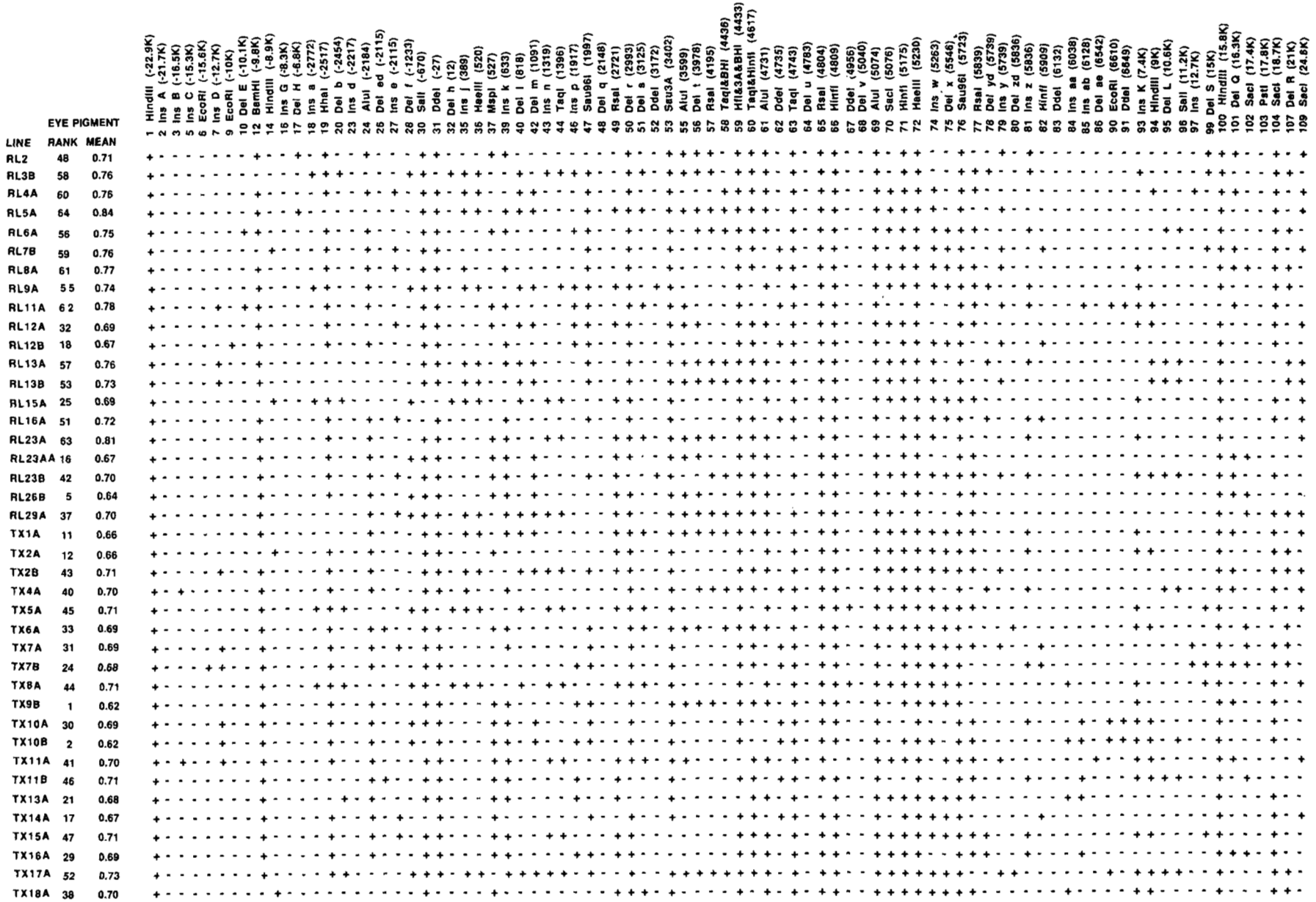
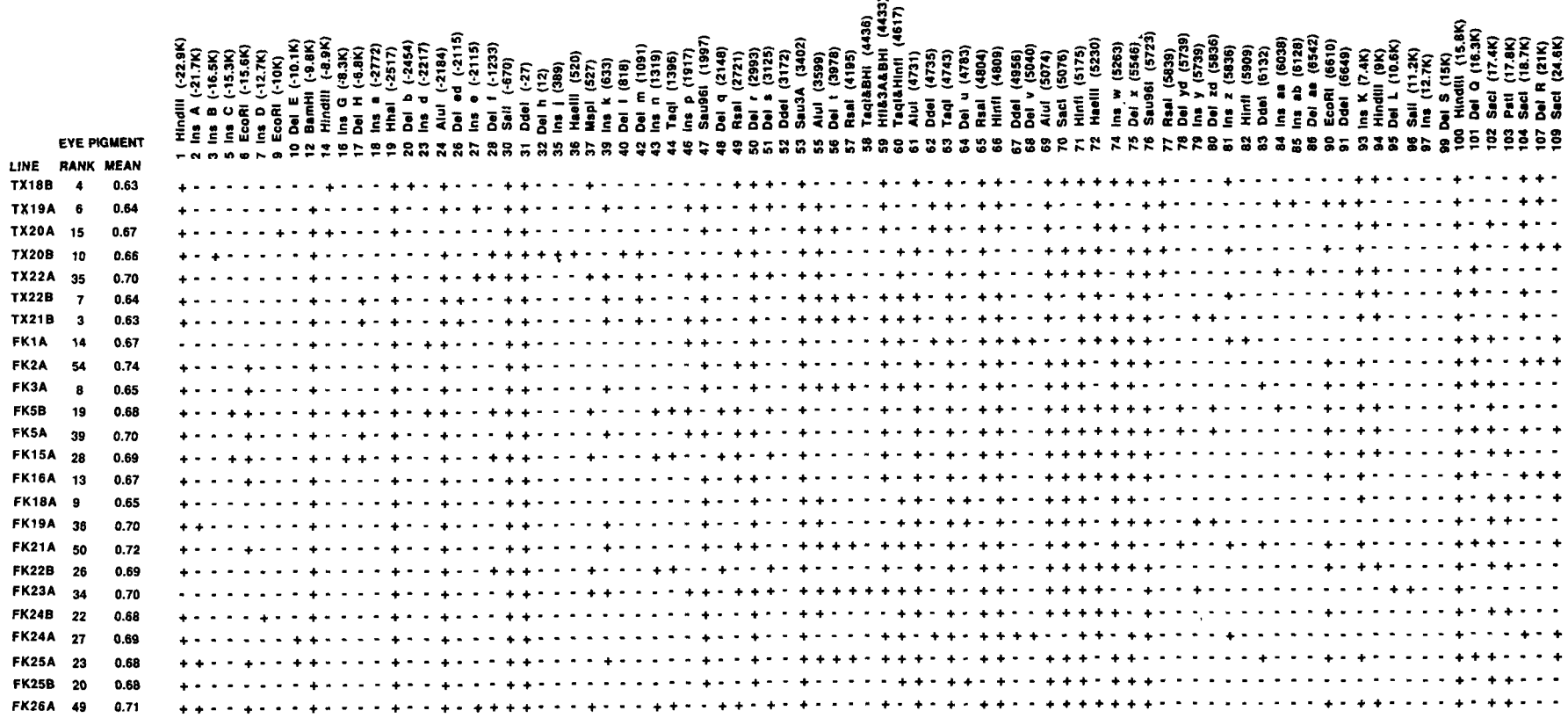


TABLE 1—*Continued*



Polymorphic restriction sites, insertions and deletions are numbered in the 3′ to 5′ order along the *white* region restriction map. The coordinate position of each polymorphism is indicated (K = kilobases). The presence is indicated by a "+" and absence by a "−." A deletion is indicated by "Del" and an insertion by "Ins." Three populations are represented: RL, North Carolina; TX, Texas; FK, Fukuoka, Japan. The following polymorphisms were unique to the indicated lines: RL3B—*Pst*I (+, #8, −13.1K), Del ad (+, #87, 6241); RL4A—Del Bd (+, #4, −16.5K); RL5A—Del ad (+, #88, 6400), *Pst*I (−, #89, 6491); RL9A—Del o (+, #45, 1591), Ins Oi (+, #106, 20.1K); RL12A—*Hha*I & *Msp*I (+, #25, −2089); RL12B—*Sal*I (+, #15, −8.4K), *Hin*fI & *Sau*3A (−, #54, 3499), Ins O (+, #105, 20.1K); RL23A—*Sau*3A (−, #41, 1021); Tx2B—Ins I (+, #92, 7.2K), Ins N (+, #98, 14.7K); TX13A—Del c (+, #22, −2301); TX19A, *Rsa*I (+, #21, −2311), Ins P (+, #108, 24K); TX20A—Ins Ei (+, #11, −10.1K), *Msp*I (−, #38, 549); Del wd (+, #73, 5265); TX20B—*Sau*3A (+, #33, 19), *Hin*fI (+, #34, 96); FK22B—Ins F (+, #13, −9.4K); FK26A—Del g (+, #29, −736).

TABLE 2

Number of polymorphisms at the *white* locus region studied with both 4- and 6-cutter restriction enzymes

	3′-Flanking region (817 bp)		Transcriptional unit (5971 bp)		5′-Flanking region (2873 bp)			
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Total	χ^2 (d.f. = 1)
Restriction site and insertion/deletion	5	6.3	33	45.7	36	22.0	74	12.7 ($P < 0.005$)
Restriction site only	2	2.4	17	25.8	21	11.6	40	10.5 ($P < 0.005$)
Insertion/deletion only	3	2.9	16	21.0	15	10.1	34	3.6 (NS)

Obs, observed; Exp, expected; NS, not significant.

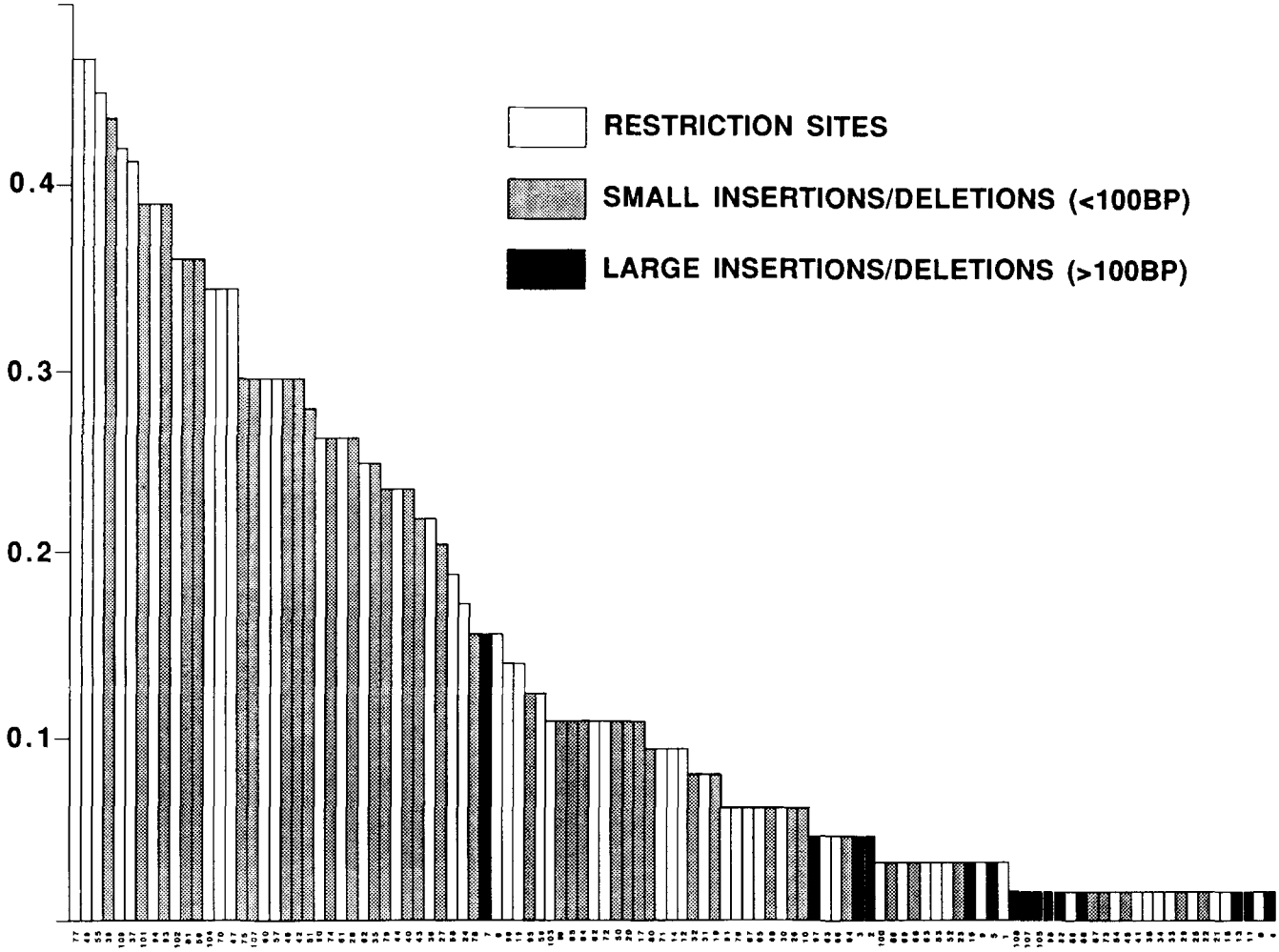


FIGURE 2.—Frequency spectrum of polymorphic variations in the *white* locus region of *D. melanogaster*. Polymorphic variations are labeled by the numbers in Table 1.

lated for each group. Figure 4 shows the relationship between the average R^2 and the average distance. As expected from population genetics theory, linkage disequilibrium declines as distance increases. Linkage equilibrium was reached at about 2 kb. Expected values in the plot were obtained by fitting the observed values (average R^2 values) to the expectation of R^2 with a correction for sample size (WEIR and HILL 1986; R. R. HUDSON, personal communication). The relatively poor fit of the expected curve to the observed points was probably due to the marginal constraints of the expectation of R^2 and/or heterogeneity among regions of the *white* locus in the overall amount of linkage disequilibrium (this nonlinear regression was done with a "NLIN" procedure of "SAS"). The same pattern of decrease in linkage disequilibrium was observed for the entire 45-kb region.

Eye pigment variation: Since the eye pigment was not significantly correlated with the amount of total protein in the head, only the raw values of eye pigment measurement were used for all the analyses. It was shown by the ANOVA model 1 that there was a

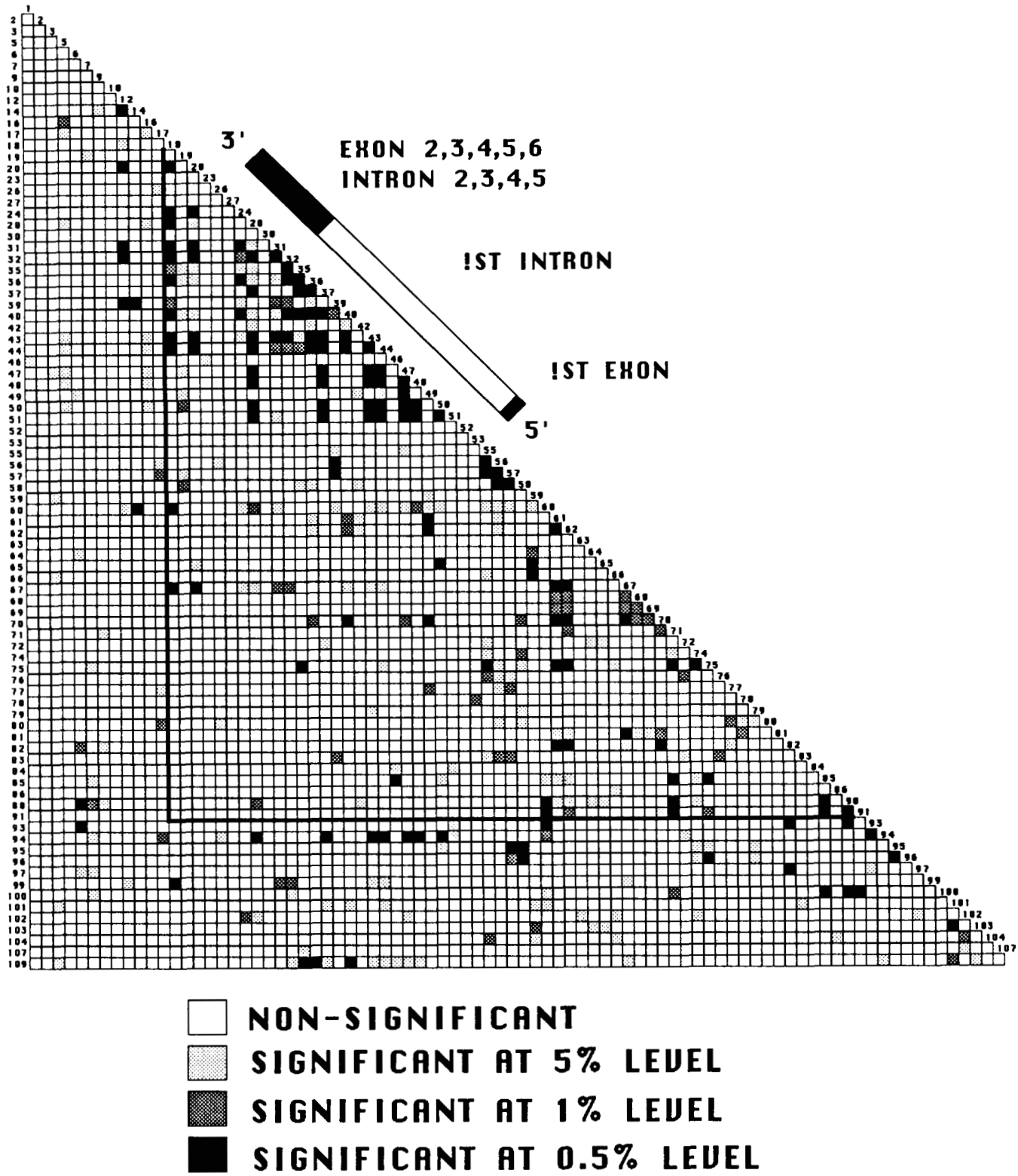


FIGURE 3.—Pair-wise linkage disequilibria between polymorphic variations in the *white* locus region of *D. melanogaster*. Significance of χ^2 test is shown for each pair. The box outside indicates the transcriptional unit of the *white* locus. Polymorphic variations are labeled by the numbers in Table 1. The solid line delineates the region studied with both 4- and 6-cutter restriction enzymes.

significant, but not very strong, *X* chromosome effect on the variation of the eye pigment ($P < 0.05$). This result suggests that there exists segregating genetic factor(s) on the *X* chromosome that influence the amount of eye pigment. The amount of eye pigment was not different between the sexes. Neither was there

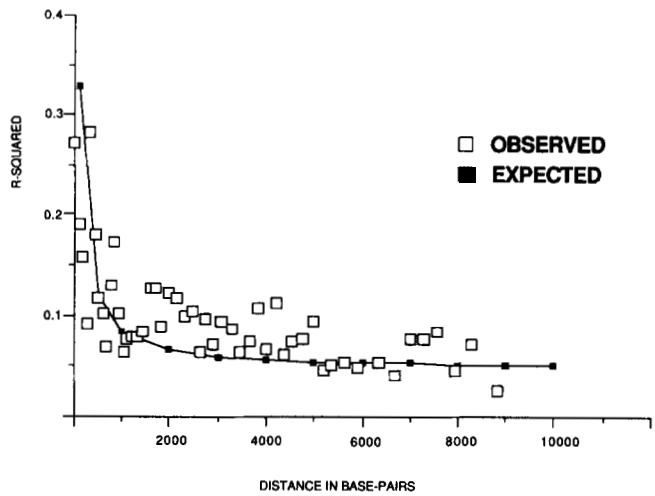


FIGURE 4.—Relationship between map distance and linkage disequilibrium in the 10-kb region studied with both 4-cutter and 6-cutter restriction enzymes. Degree of linkage is expressed in terms of square of correlation in frequency (R^2). Expected values in the plot were obtained by fitting the observed values (averages) to the expectation of R^2 with a correction for sample size (WEIR and HILL 1986). $E(R^2) = 1/(1 + 4C*d) + 1/n_a$, where d = distance in base pairs and n_a = average of sample size. 'C', the product of effective size of population and recombination rate per base pair, was obtained by fitting. Fitting was also done with a more complex model based on the results of HUDSON (1985) and R. R. HUDSON (personal communication). However, the result from fitting this second model was very similar to that shown.

any interaction effect between sex and *X* chromosome line. Thus genetic variation in dosage compensation was not found.

From the analyses with the ANOVA model 2, it was found that only 6 out of 109 polymorphisms showed a significant association with eye pigment variation ($P < 0.05$). Five of them were in intermediate frequencies, *Taq*I (#44, 1395), *Rsa*I (#57, 4195), *Taq*I & *Bam*HI (#58, 4436), x (#75, 5546) and yd (#78, 5739), and the other, *Sau*3A (#41, 1021) was unique. It should be noted that none of these polymorphisms were located within an exon. Polymorphisms detected within exons, even the replacement substitution (*Hin*fI & *Sau*3A #54, 3499), did not seem to be associated with the amount of eye pigment. But three of the five polymorphisms that appear to be associated with variation in eye pigment were within clusters of sites in strong linkage disequilibrium in the large intron or immediately 5′ of the transcriptional unit. None of the large insertion/deletion polymorphisms were associated with pigment variation. However, the frequencies of the large insertions were too low for the tests to be very meaningful. From this analysis, it was not possible to determine if any of the detected variation in the *white* locus region causes phenotypic variation in eye pigment, since 6 significant test results from 109 random tests might simply have occurred by chance. Most of the polymorphic variations seen at the DNA level in this study could not be related to the variation of eye pigment.

It has been shown by *P* element transformation experiments that normal *white* locus expression is controlled by several 5′ *cis*-acting sequences. One region which is thought to be important for eye pigmentation is the 5′ region from 654 to 990 bp upstream from the first exon. This is the region with which the *su*(w^{sp}) product interacts in the suppression of w^{sp} (DAVISON *et al.* 1985; CHAPMAN and BINGHAM 1985). In this region, three polymorphic restriction sites were found in this sample. One of the three sites was the *Taq*I & *Bam*HI (#58, 4436) site that showed a significant ($P < 0.01$) association with eye pigment. The means of eye pigment (OD measurement) were 0.69 and 0.73 for lines in which the site was absent or present, respectively. Another region from 1.1 to 1.9 kb upstream from the 5′ end of the *white* transcript has an important role on the eye pigmentation and interaction of *white* with *zeste*[1] (GEHRING *et al.* 1984; LEVIS, HAZELRIGG and RUBIN 1985; PIRROTTA, STELLER and BOZZETTI 1985). This region contains *cis*-regulatory sequences essential for *white* locus expression in the testes of adult males, and can influence the level of eye pigment. From Figure 1, it can be seen that the region was very polymorphic (11 polymorphisms in this 0.8-kb region). However, the analyses by the ANOVA model 2 indicated that none of these polymorphisms in this 0.8-kb region showed a significant association with eye pigment variation. The region of 216 bp upstream from the first exon is thought to be responsible for dosage compensation of the *white* locus. No polymorphism was detected in this small region, which was consistent with the observation that no effect of sex was detected by the ANOVA 1.

Epistasis between sites can result in linkage disequilibrium (KIMURA 1956; LEWONTIN and KOJIMA 1960). Therefore, it is of interest to compare the occurrence of significant interaction effects on a phenotype (eye pigment variation) with the linkage disequilibria clustered in the transcriptional unit. Despite the meager observed association between individual polymorphisms and eye pigment variation, there was a possibility that significant interactions between polymorphisms could be present. In order to infer epistasis from the ANOVA, it was assumed that there were no dominance effects. Twenty-four polymorphisms that were found in the transcriptional unit and near flanking regions (#18, #19, #20 and #24–#58 in Figure 3) were subjected to a modification of ANOVA model 2 in order to detect two-locus interaction effects on eye pigment variation. A total of 221 combinations out of 276 possible pairs were tested, since the required four possible two-locus haplotypes did not exist in the sample for 55 of the combinations. Only 9 combinations were significant for a two-locus interaction effect

on eye pigment variation. Since so few were statistically significant (9 out of 221) and eight with $P < 0.05$, there was little evidence for extensive epistasis. Four of the 9 significant combinations corresponded to pairs of sites that showed significant linkage disequilibria in Figure 3. Three were in the first intron, between #35 and #43, #35 and #44, and #36 and #43, and the other was between #18 and #24, which were located at the 3′ end of the transcriptional unit and flanking region, respectively. None of these were located side by side. Three of the six polymorphic variations that were found to have significant effects were included in this analysis, *Taq*I (#44, 1396), *Rsa*I (#57, 4195) and *Taq*I & *Bam*HI (#58, 4436). None of these pairs of polymorphisms had significant interaction effects. For the 55 combinations which were eliminated from the above analysis, haplotype effect on eye pigment was tested by a modified ANOVA 1, in which X_k was substituted with H_k (the kth haplotype effect, $k = 1$–3 for all the 55 combinations). In only 3 out of 55 tests was a significant haplotype effect detected. The three significant combinations did not correspond to pairs of sites that showed significant linkage disequilibria. Thus there was little correspondence between the clustering of linkage disequilibria and the distribution of detected two site interactions or haplotype effects on eye pigment variation.

DISCUSSION

The amount of detectable genetic variation in eye pigment among the 64 *X* chromosome lines in this study was not large and was only significant at the 5% level. This contrasts with the enormous amount of molecular variation detected among the 64 lines examined. There were 109 polymorphisms in the region. Many of these were located in regions known to be involved with the regulation of *white* locus expression. Since the contribution of *X* chromosomes to eye pigment variation was only marginally detectable, it was not particularly surprising that few of the molecular polymorphisms (individually or as pairs) could be associated with variation in pigment. This result contrasts with that reported for *Adh* (AQUADRO *et al.* 1986; LAURIE-AHLBERG and STAM 1987) and *Amy* (LANGLEY *et al.* 1988) where many molecular polymorphisms in the vicinity of the loci were correlated with one another and with substantial variation in enzyme activity. Since eye pigment deposition is the terminal state of a pathway, of which the *white* locus is only one step, it may be reasonable to expect that little variation in eye pigment variation would be attributable to the *white* locus. There is little evidence in this study for any functional significance of the molecular variations detected.

The level of nucleotide variation found in the *white* locus region on the *X* chromosome was similar to that in the *Adh* locus region on the second chromosome or in the 87A heat-shock gene region on the third chromosome. This result is contrary to the simple prediction based on stronger selection pressure against partially recessive deleterious variation on the *X* chromosome compared to that on the autosomes. It also implies that selection pressure, if any, at the DNA level is not any stronger on the *X* chromosome. Of course, this conclusion may be too simple and premature. Other loci on the *X* chromosome may give different results. For example, the level of variation in the *pgd* region and the *Zw* region, both on the *X* chromosome, are lower than those of the *white*, *Adh* and heat-shock locus regions (N. MIYASHITA and C. H. LANGLEY, unpublished data). SCHAEFFER, AQUADRO and LANGLEY (1988) found that the level of molecular genetic variation in the *X*-linked *Notch* locus region is similar to that of *Adh* and *white*. Therefore, it seems that inter-locus comparison of variability will require a much larger sample of loci and careful consideration of effects such as levels of recombination and density of transcriptional units.

The results of this study are consistent with purifying (negative) selection acting on several types of polymorphism in the *white* locus region. The entire transcriptional unit, exons and introns, was less variable than the 5′-flanking region with respect to the number of polymorphisms (restriction site and insertion/deletion together). This was also true when only the number of restriction site polymorphisms was considered. However, there was no evidence for different levels of silent polymorphisms in the various regions of the *white* locus. Small insertion/deletion polymorphisms were equally common within and outside the *white* transcriptional unit. But none were found in coding sequence. Large insertions were only found in the more distant flanking regions at very low frequencies. These observations are all consistent with natural selection eliminating variation at the *white* locus. The pattern of silent site polymorphism detected in this study was different from that of *Adh* region (KREITMAN and AGUADÉ 1986b). In their study, an excess number of silent polymorphic restriction sites in the coding sequence was detected, compared to the flanking sequence. It is interesting to compare the patterns of silent site polymorphism between the two regions studied with the same technique. A comparison of nucleotide variation between the *white* and *Adh* regions is summarized in Table 3. Although a proper statistical test to compare variation at the DNA level is not established yet, it is clear that the *Adh* coding region was exceptionally variable, as indicated by the large value of θ for silent site equivalents. Also, it should be noted that the other regions were relatively homogeneous with respect to the level of variation. This result again indicates the peculiarly

TABLE 3

Summary comparison of levels of silent site variation at the DNA level in the *white* and *Adh* locus regions

	Adh regions				*white* regions			
	3′-Flanking	Translated	Nontranslated	5′-Flanking	3′-Flanking	Translated	Nontranslated	5′-Flanking
Size	880	756	1093	4276	817	2118	3853	2873
Site equivalents	142	258 (63)	187	455	134	691 (167)	696	628
Proportion of polymorphic sites	2/18	8/47	6/26	10/62	2/11	4/84	12/77	18/83
θ per silent site equivalents ($\times 10^{-3}$)	2.80	25.21	6.37	4.42	3.33	5.39	3.74	6.34

The number of polymorphic sites and site equivalents in the *Adh* locus region were obtained from the consensus sequences in KREITMAN (1983) and KREITMAN and AGUADÉ (1986b). Included in this calculation are 63 bp of 5′ flanking region in KREITMAN (1983), which were not considered in KREITMAN and AGUADÉ (1986b). Site equivalents in the *white* locus region were calculated for both 4- and 6-cutter restriction enzymes. The estimates of θ were calculated utilizing the method of WATTERSON (1975). Silent site equivalents for the translated regions are shown in the parentheses. In the other regions silent site equivalents are the same as site equivalents.

high level of variation in the *Adh* coding region, especially the fourth exon, as noted previously (KREITMAN 1983; HUDSON, KREITMAN and AGUADÉ 1987).

The analysis of linkage disequilibrium clearly showed that nonrandom associations quickly disappeared as the distances between polymorphisms increased, as expected (HILL and ROBERTSON 1968). It is surprising that linkage disequilibrium in the *white* locus region was evident only within 2 kb. A reasonable estimate of recombination rate at the DNA level in this region is about 0.002% per kb (JUDD 1987). These results suggest that a recombination rate of 0.0027% is enough to randomize DNA sequences in natural populations of *D. melanogaster*. From the fitted curve of average R^2 to distance it was possible to obtain an estimate of effective population size, N_e, on the order of 10^5. However, because the "observed values" were not independently sampled, no confidence interval for this estimate of N_e is available.

The observed clustering of linkage disequilibria, in the transcriptional unit, was not an obvious prediction of the neutral theory. A "cold spot" of crossing over in the transcriptional region could be put forth to explain the observed distribution. Genetic studies on recombination among several *white* mutations have given little evidence for nonuniformity in the map of the *white* locus (GREEN 1959b; JUDD 1964). The recombination rates in the larger chromosomal region around *white* (from *zeste* to *Notch*) do not seem to be higher per kilobase than that estimated from within the *white* locus (RUDKIN 1965; B. H. JUDD, personal communication). It should be mentioned that actual recombination in nature could differ from these estimated values, since some of these estimates came from the studies between mutants caused by transposon insertions and/or were carried out in a somewhat contrived genetic background. Under the assumption of uniform recombination in the studied region the clustering of linkage disequilibria in the transcriptional unit might be explained by epistatic natural selection among polymorphisms and concomitant hitchhiking of surrounding regions. Obviously recent transient (*i.e.*, nonequilibrium) changes in frequency of one or a few sites under natural selection could lead to clustering of linkage disequilibria. Most of the polymorphisms detected in this region, even those found in the transcriptional unit, did not seem to influence the deposition of eye pigment. In addition, an analysis of interaction effects on eye pigment variation indicates that only a few combinations of polymorphic sites in the transcriptional unit showed significant interaction effects. These occurrences of significant interaction effects did not correspond well to the distribution of the linkage disequilibria. At the present, it is not clear how this clustering of linkage disequilibria in and around the *white* locus transcriptional unit came about. To answer this question, further investigation at the molecular level in this region and other regions of the genome seems to be necessary. At the very least, it is necessary to determine if the observed clustering is only unique to the *white* region and/or to these population samples.

The authors express their thanks to M. AGUADÉ, T. EMIGH, W. HILL, R. HUDSON, M. KREITMAN, B. LANGE, G. SIMMONS, M. WATADA, and B. WEIR for their comments and help, and to C. C. LAURIE for allowing us to use her laboratory equipment at North Carolina State University for eye pigment measurement.

LITERATURE CITED

AQUADRO, C. F., S. F. DEESE, M. M. BLAND, C. H. LANGLEY and C. C. LAURIE-AHLBERG, 1986 Molecular population genetics of alcohol dehydrogenase gene region of *Drosophila melanogaster*. Genetics **114:** 1165–1190.

BINGHAM, P. M., and Z. ZACHAR, 1985 Evidence that two mutations, w^{DZL} and w^{1}, affecting synapsis-dependent genetic behavior of *white* are transcriptional regulatory mutations. Cell **40:** 819–825.

BINGHAM, P. M., R. LEVIS and G. M. RUBIN, 1981 Cloning of DNA sequences from the *white* locus of *Drosophila melanogaster* by a novel and general method. Cell **25:** 693–704.

BRIDGES, C. B., 1938 A revised map of the salivary gland *X* chromosome of *Drosophila melanogaster*. J. Hered. **29:** 11–13.

BRIDGES, C. B., and T. H. MORGAN, 1923 *The Third-Chromosome Group of Mutant Characters of Drosophila melanogaster.* Carnegie Inst. Wash. Publ. 327.

CHAPMAN, C. H., and P. M. BINGHAM, 1985 Evidence that the locus of a novel type of suppressor mutation regulates transcription of the *white* locus. Drosophila Inform. Serv. **61:** 48–49.

DAVISON, D., C. H. CHAPMAN, C. WEDEEN and P. M. BINGHAM, 1985 Genetic and physical studies of a portion of the *white* locus participating in transcriptional regulation and in synapsis-dependent inter-actions in *Drosophila* adult tissues. Genetics **110:** 479–494.

ENGELS, W. R., 1981 Estimating genetic divergence and genetic variability with restriction endonucleases. Proc. Natl. Acad. Sci. USA. **78:** 6329–6333.

EPHRUSSI, B., and J. L. HERALD, 1944 Studies of eye pigments of *Drosophila.* I. Methods of extraction and quantitative estimation of the pigment components. Genetics **29:** 148–175.

EWENS, W. J., R. S. SPIELMAN and H. HARRIS, 1981 Estimation of genetic variation at the DNA level from restriction endonuclease data. Proc. Natl. Acad. Sci. USA **78:** 3748–3750.

GEHRING, W. J., R. KLEMENZ, U. WEBER and U. KLOTTER, 1984 Functional analysis of the *white*$^+$ gene of *Drosophila* by P-factor-mediated transformation. EMBO J. **3:** 2077–2085.

GREEN, M. M., 1959a The discrimination of wild-type isoalleles at the *white* locus of *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **45:** 549–553.

GREEN, M. M., 1959b Spatial and functional properties of pseudoalleles at the white locus in *Drosophila melanogaster.* Heredity **13:** 302–315.

HALDANE, J. B. S., 1927 A mathematical theory of natural and artificial selection. Part V. Selection and Mutation. Proc. Camb. Philos. Soc. **23:** 838–844.

HAZELRIGG, T., 1987 The *Drosophila white* gene: a molecular update. Trends Genet. **3:** 43–47.

HAZELRIGG, T., R. LEVIS and G. M. RUBIN, 1984 Transformation of *white* DNA in *Drosophila*: dosage compensation, *zeste* interaction, and position effects. Cell **36:** 469–481.

HELWIG, J. T., and K. A. COUNCIL (Editors), 1979 *SAS User's Guide.* SAS Institute, Inc., Raleigh, N.C.

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

HUDSON, R. R., 1982 Estimating genetic variability with restriction endonucleases. Genetics **100:** 711–719.

HUDSON, R. R., 1985 The sampling distribution of linkage disequilibrium under an infinite allele model without selection. Genetics **109:** 611–631.

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

JONES, J. S., and R. F. PROBERT, 1980 Habitat selection maintains a deleterious allele in a heterogeneous environment. Nature **287:** 632–633.

JUDD, B. H., 1964 The structure of intralocus duplication and deficiency chromosomes produced by recombination in *Drosophila melanogaster*, with evidence for polarized pairing. Genetics **49:** 253–265.

JUDD, B. H., 1976 Complex loci. pp. 767–799 In: *The Genetics and Biology of Drosophila, 1b*, Edited by M. ASHBURNER and E. NOVITSKI. Academic Press, London.

JUDD, B. H., 1987 The *white* locus of *Drosophila melanogaster.* pp. 81–94. In: *Results and Problems in Cell Differentiation 14. Structure and Function of Eukaryotic Chromosomes*, Edited by W. HENNING. Springer-Verlag, Berlin.

KIMURA, M., 1956 A model of a genetic system which leads to closer linkage by natural selection. Evolution **10:** 278–287.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, London.

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

KREITMAN, M., and M. AGUADÉ, 1986a Genetic uniformity in two populations of *Drosophila melanogaster* revealed by filter hybridization of four-nucleotide-recognizing restriction enzyme digests. Proc. Natl. Acad. Sci. USA **83:** 3562–3566.

KREITMAN, M., and M. AGUADÉ, 1986b Excess polymorphism at the *Adh* locus in *Drosophila melanogaster.* Genetics **114:** 93–110.

LANGLEY, C. H., and C. F. AQUADRO, 1987 Restriction map variation in natural populations of *Drosophila melanogaster: white* locus region. Mol. Biol. Evol. **4:** 651–663.

LANGLEY, C. H., E. A. MONTGOMERY and W. F. QUATTLEBAUM, 1982 Restriction map variation in the Adh region of *Drosophila.* Proc. Natl. Acad. Sci. USA **79:** 5631–5635.

LANGLEY, C. H., A. E. SHRIMPTON, T. YAMAZAKI, N. MIYASHITA, Y. MATSUO and C. F. AQUADRO, 1988 Naturally occurring variation in the restriction map of the *Amy* region of *Drosophila melanogaster.* Genetics **119:** 619–629.

LAURIE-AHLBERG, C. C., and L. F. STAM, 1987 Use of *P*-element-mediated transformation to identify the molecular basis of naturally occurring variants affecting *Adh* expression in *Drosophila melanogaster.* Genetics **115:** 129–140.

LEIGH BROWN, A. J., 1983 Variation at the 87A heat-shock locus in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **80:** 5350–5354.

LEVIS, R., P. M. BINGHAM and G. M. RUBIN, 1982 Physical map of the *white* locus of *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **79:** 564–568.

LEVIS, R., T. HAZELRIGG and G. M. RUBIN, 1985 Separable *cis*-acting control elements for expression of the *white* gene of *Drosophila.* EMBO J. **4:** 3488–3499.

LEWONTIN, R. C., and K. KOJIMA, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14:** 458–472.

LINDSEY, D. L., and E. H. GRELL, 1968 *Genetic Variations of Drosophila melanogaster.* Carnegie Inst. Wash. Publ. 627.

LOWRY, O. H., N. J. ROSEBROUGH, A. S. FARR and R. J. RANDALL, 1951 Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193:** 265–275.

MIYASHITA, N., C. C. LAURIE-AHLBERG, A. N. WILTON and T. H. EMIGH, 1986 Quantitative analysis of *X* chromosome effects on the activities of the glucose 6-phosphate and 6-phosphogluconate dehydrogenases of *Drosophila melanogaster.* Genetics **113:** 321–335.

MORGAN, T. H., 1910 Sex limited inheritance in *Drosophila.* Science **32:** 120–122.

MOUNT, S. M., 1987 Sequence similarity. Nature **325:** 487.

NEI, M., and W-H. LI, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:** 5269–5273.

NEI, M., and F. TAJIMA, 1981 DNA polymorphism detectable by restriction endonucleases. Genetics **97:** 145–163.

NETER, J., and W. WASSERMAN, 1974 *Applied Linear Statistical Methods: Regression, Analysis of Variance and Experimental Design.* R. D. Irwin, Inc., Homewood, Ill.

O'HARE, K., C. MURPHY, R. LEVIS and G. M. RUBIN, 1984 DNA sequence of the *white* locus of *Drosophila melanogaster.* J. Mol. Biol. **180:** 437–455.

OHTA, T., 1972 Evolutionary rate of cistrons and DNA divergence. J. Mol. Evol. **1:** 150–157.

OHTA, T., 1987 Very slightly deleterious mutations and the molecular clock. J. Mol. Evol. **26:** 1–6.

OHTA, T., and M. KIMURA, 1969 Linkage disequilibrium due to random genetic drift. Genet. Res. **13:** 47–55.

PIRROTTA, V., H. STELLER and M. P. BOZZETTI, 1985 Multiple upstream regulatory elements control the expression of the *Drosophila white* gene. EMBO J. **4:** 3501–3508.

RUDKIN, G. T., 1965 The relative mutability of DNA in regions

of the *X* chromosome of *Drosophila melanogaster*. Genetics **52:** 665–681.

SCHAEFFER, S. W., C. F. AQUADRO and C. H. LANGLEY, 1988 Restriction-map variation in the *Notch* region of *Drosophila melanogaster*. Mol. Biol. Evol. **5:** 30–40.

SOUTHERN, E., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 537–544.

SPEED, F. M., R. R. HOCKING and O. P. HACKNEY, 1978 Methods of analysis of linear models with unbalanced data. J. Am. Stat. Assoc. **73:** 105–112.

SULLIVAN, D. T., and M. C. SULLIVAN, 1975 Transport defects as the physiological basis for eye color mutants of *Drosophila melanogaster*. Biochem. Genet. **13:** 603–613.

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

WEIR, B. S., and W. G. HILL, 1986 Nonuniform recombination within the human β-globin gene cluster (Letter to the Editor). Am. J. Hum. Genet. **38:** 776–778.

Communicating editor: M. TURELLI