

DYNAMICS OF CORRELATED GENETIC SYSTEMS. V. RATES OF  
DECAY OF LINKAGE DISEQUILIBRIA IN EXPERIMENTAL  
POPULATIONS OF *DROSOPHILA MELANOGASTER*

M. T. CLEGG

*Department of Botany and Department of Molecular and Population Genetics  
University of Georgia, Athens, Georgia 30602*

AND

J. F. KIDWELL AND C. R. HORCH

*Division of Biology and Medicine, Brown University,  
Providence, Rhode Island 02912*

Manuscript received December 19, 1978

Revised copy received May 21, 1979

ABSTRACT

The dynamic behavior of four-locus gametic frequency distributions was studied in five replicate cage populations of *Drosophila melanogaster* for up to 50 generations. The joint frequency distributions were resolved into gene frequencies and various disequilibrium measures. In addition,  $F$  statistics for marginal single-locus genotypic frequency distributions were followed through time. The gene frequency, disequilibrium and  $F$  statistics were obtained for four chromosome 3 enzyme marker loci [isocitrate dehydrogenase (3-27.1), esterase-6 (3-36.8), phosphoglucomutase (3-43.4) and esterase-C (3-49.0)]. The initial structure of the experimental populations featured random mating proportions, and two complementary gametic types with respect to the marker loci, thus assuring complete pairwise linkage disequilibrium among the markers.—The experimental results indicate: (1) the between-replicate variance in gene frequency varied substantially among loci, with isocitrate dehydrogenase showing the greatest between-replicate variance, and esterase-C the least. (2) The  $F$  statistics initially were strongly negative but decayed to the neighborhood of zero for all marker loci except esterase-C. The rate at which the  $F$  statistics approached zero varied among the marker loci, indicating substantial differences in the distribution of selective effects along the chromosome. The centromeric region, marked by esterase-C, shows the strongest selective effects. (3) The rate of decay of linkage disequilibrium was much faster than expected for pairs of neutral loci, averaging 1.82 times the neutral rate over all replicates and pairs of loci. This acceleration, which was observed for all six pairwise combinations of loci, was interpreted as resulting from the interaction between selection and recombination. Our experimental results are consistent with many investigations of linkage disequilibrium in natural populations of *Drosophila melanogaster* that show little or no disequilibrium among enzyme loci. (4) A fortuitous contamination of two cages revealed an apparent regulatory interaction between the migrant and nonmigrant chromosomes at the esterase-C locus. The migrant chromosomes were very rapidly absorbed into the recipient populations, despite this interaction. This result suggests that the dynamics of migration in populations may be phenomenologically richer than anticipated by simple theory.

IN recent years, there has been much interest among population geneticists in genetic structure at the multilocus level. This focus has been a natural outgrowth of a maturing of population genetic theory (*e.g.*, KARLIN 1975, for a review of two-locus selection theory). Concomitant with this expansion of multilocus theory, biochemical studies of genetic variation have revealed high levels of protein polymorphism in most populations of sexually reproducing organisms. High levels of genetic variation make a multilocus approach to the study of genetic structure both feasible and necessary.

The necessity of a multilocus perspective arises from the fact that adjacent segregating loci are, on average, very tightly linked. As a consequence, rates of recombination are low, and relatively weak epistatic selection can maintain associations among alleles at different loci. If nonrandom associations are important, then neither the statics nor the dynamics of genetic systems can be correctly described by single-locus theory. The feasibility of multilocus investigations is a direct consequence of the large number of genetic markers that can be studied in single individuals, using electrophoretic procedures.

A substantial body of literature has now accumulated describing multilocus structure in a variety of organisms. Most of these investigations involve the estimation of disequilibria from single samples drawn from natural populations. Two generalizations emerge from these studies. First, disequilibria are common in inbreeding species of plants (reviewed by BROWN 1979). Second, disequilibria between enzyme loci are quite rare in random mating species. Evidence to support this latter assertion comes largely from *Drosophila melanogaster* (reviewed by HEDRICK, JAIN and HOLDEN 1978).

Relatively few attempts have been made to study the dynamic behavior of linkage disequilibria under controlled conditions. This is not surprising because long generation times and unknown linkage relationships make experimental studies difficult, at best, in most organisms. Most of the descriptive work on linkage disequilibria comes from *Drosophila melanogaster*, an organism particularly well suited to experimental study. Published experimental studies of the behavior of linkage disequilibria between enzyme loci in *D. melanogaster* present conflicting results. RASMUSON, RASMUSON and NILSON (1967) reported the rapid decay of disequilibrium between the third chromosome loci esterase-6 and leucine aminopeptidase. These loci are more than 50 map units apart and are therefore expected to segregate independently in females. On the other hand, O'BRIEN and MACINTYRE (1971a) and BIRLEY (1974) report some degree of association among enzyme loci in experimental populations, although the association studied by O'BRIEN and MACINTYRE (1971a) was evidently unstable.

This paper reports on a set of experiments designed to measure rates of decay of linkage disequilibria in experimental populations of *Drosophila melanogaster*. The experiments, which spanned a period of nearly four years and involved sampling more than 31,700 gametes, show a marked acceleration of decay rates relative to that expected from neutral theory. This acceleration is a consequence of an interaction between selection and recombination and may aid in explaining why strong associations are uncommon in natural populations of random mating organisms.

## MATERIALS AND METHODS

The experimental populations were synthesized from four wild-type stocks. The stocks, Ames and Ottawa, were intercrossed and maintained for four months in a half-pint milk bottle population. Single pair matings were then made between this composite population and the wild-type stock, Painesville. After five days of mating activity, the parents were subjected to electrophoresis to determine their genotype at the chromosome 3 loci isocitrate dehydrogenase, *Idh* (3-27.1), esterase-6, *E<sub>6</sub>* (3-36.8), phosphoglucumutase, *Pgm* (3-43.3) and esterase-c, *E<sub>c</sub>* (3-49.0). A line homozygous for the alleles F, S, N and S at *Idh*, *E<sub>6</sub>*, *Pgm* and *E<sub>c</sub>*, respectively, was established from the progeny of one single pair mating. (F, S and N denote allozymes of fast, slow and intermediate mobility.) A second complementary stock homozygous for SFSF was established from the wild-type stock, Oregon RC, which was monomorphic for these markers.

Two sets of experimental populations, denoted I and II, were established from these lines. Two replicate populations of experiment I were begun, each with the following distribution of genotype numbers: 22 FSNS/FSNS, 44 FSNS/SFSF and 22 SFSF/SFSF, yielding an initial structure with all marker gene frequencies at 0.5 and complete pairwise linkage disequilibria among the marker genes. (Equal numbers of males and females of each genotype were used to initiate the experimental populations.) The 88 initial FSNS chromosomes represent at most four independent chromosomes, having been derived from one single-pair mating. The SFSF chromosomes were drawn from Oregon RC and represent at most 88 independent chromosomes, although the actual number of independent chromosomes is probably much less.

After 15 generations of random mating, a sample of males and virgin females was withdrawn from each population and used to initiate a large number of single-pair matings. Full-sib mating was imposed on the progeny of each mating and continued for five successive generations. After five generations of full-sib mating, two new lines, homozygous for the markers FSSF and SFNS respectively, were obtained. These new lines were each obtained from single full-sib lines. Thus, only a single independent third chromosome was probably sampled for each line. These lines were used to initiate experiment II, also with two replicate populations. The initial genotypic numbers introduced into each population were 15 FSSF/FSSF, 30 FSSF/SFNS and 15 SFNS/SFNS, again with nearly equal male and female numbers. The initial composition of the second set of experiments also featured complete initial linkage disequilibria and gene frequencies of 0.5, but with a set of gametes that were recombinants of the set introduced into experiment I.

All experimental populations were maintained in large (33 × 33 × 9 cm) plexiglass cages with 25 scintillation-vial food containers attached to the bottom of each cage. (The cages maintained approximately 1000 adult flies per generation.) The two cages of experiment I (denoted 1 and 2) were sampled on every generation through generation 24 and on alternate generations from generations 26 through 48, at which time the experiment was terminated. The cages of experiment II were sampled on every generation through generation 14 and on alternate generations through generation 50, except generations 42 and 48. On generations 26 and 4, cages 1 and 4, respectively, became contaminated. Segregation at a second chromosome marker locus known to be monomorphic in the parental lines verified the contamination event in cages 1 and 4. While the source of the contamination is unknown, it may have been from a local natural population. (The contamination event occurred at the same time, but at different generations, because experiment II was initiated more than a year after experiment I.) Because of the contamination in cage 4, cage 3 was split into two replicates on generation 14 (labeled 3A and 3B).

All populations were maintained on a 16-day discrete generation regime following a protocol like that described by CLEGG, *et al.* (1976), except that adults were cleared from the cages after five days of mating activity. Samples of 120 males were withdrawn from each cage and individually test-mated to virgin females of a homozygous tester stock. Four to eight days after test-mating, the sampled males were assayed by electrophoresis. Two weeks later, a single testcross progeny of each multiply heterozygous male was also assayed. This progeny-testing procedure permits the direct sampling of four-locus gametes unconfounded by phase relationships in multiple heterozygotes, assuming no crossing over in males; thus, the basic

data of the experiment are the gene, genotypic and gametic frequencies for five cage populations over a period spanning up to 50 generations. More than 31,700 individual gametes were sampled during the course of the experiments. Table 1 gives the minimum, maximum and average number of gametes sampled per generation for each replicate.

The electrophoretic assays were carried out on starch gels (Electrostarch) following routine methods (*e.g.*, BREWER 1970).

#### RESULTS

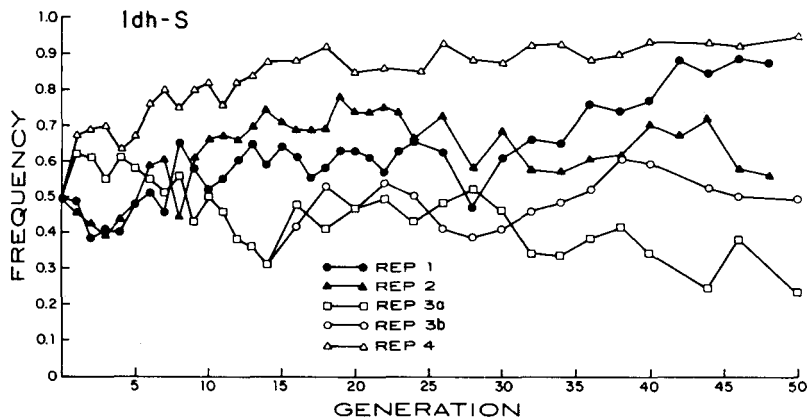
The data, too extensive to be presented in a tabular format, are displayed graphically. Figures 1a, b, c and d present gene frequencies, as a function of generation, for loci *Idh*, *E<sub>6</sub>*, *Pgm* and *E<sub>c</sub>*, respectively. Because the number of gametes sampled per replicate generation averaged about 190, the standard error for each estimated gene frequency was below 0.04 in all but five samples (generations 26, 36, 40, 44 and 46 of replicates 3A, 3B and 4). The smallest samples occurred in generation 44 of replicates 3A, 3B and 4, yielding maximum standard errors of 0.089. It is clear from Figure 1 that between-replicate variance in gene frequency is often large; this is particularly the case for *Idh* and *E<sub>6</sub>*. The gene frequency variance over replicates remains large at *Idh* even if replicate 4, which was contaminated early in the experiment, is excluded. On the other hand, exclusion of replicate 4 markedly reduces the between-replicate variance for *E<sub>6</sub>* and *Pgm*. In contrast, the *E<sub>c</sub>* locus shows a very small and relatively constant between-replicate variance throughout the experiment.

For neutral loci, the between-replicate variance in gene frequency is expected

TABLE 1

*Mean, minimum and maximum numbers of gametes sampled per generation*

	1	2	Replicate 3A	3B	4
Mean	205	197	192	191	192
Minimum	76	74	30	20	17
Maximum	286	284	240	240	238



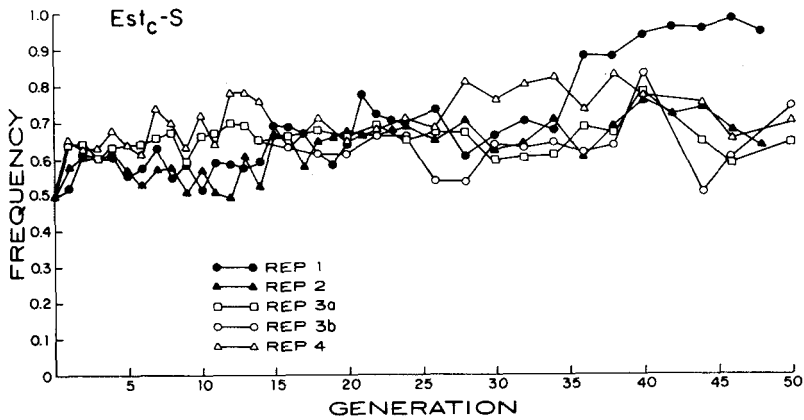
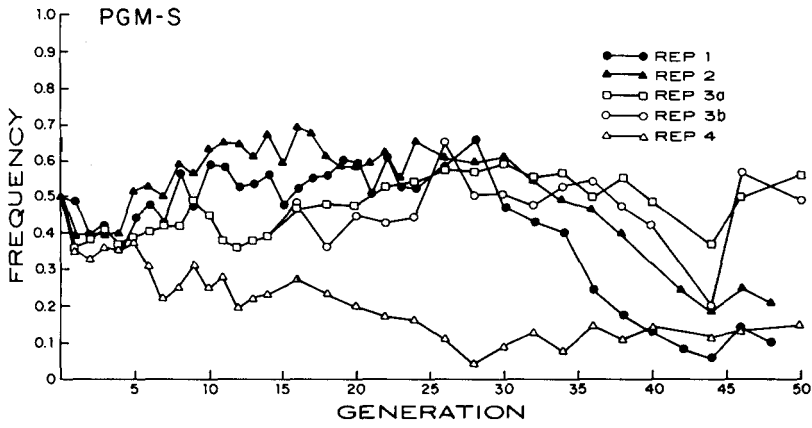
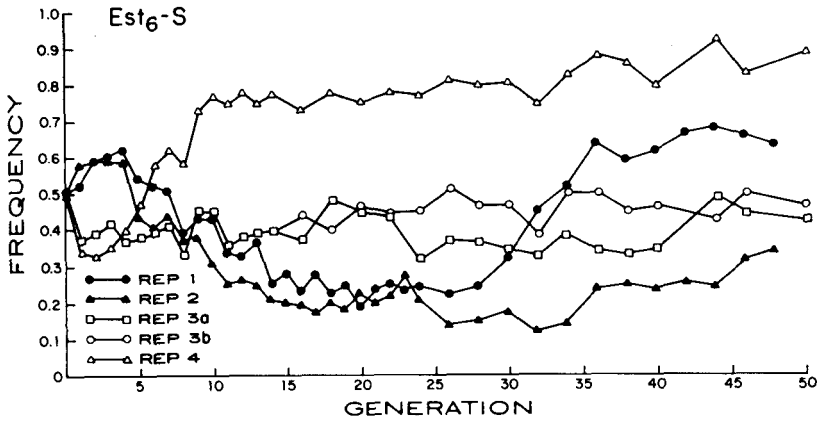


FIGURE 1.—Gene frequency as a function of generation for loci *Idh* (a), *E<sub>6</sub>* (b), *Pgm* (c), and *E<sub>c</sub>* (d).

to grow as a function of time due to the sampling of gametes in every generation. The present experiments involved moderate population sizes (approximately 1000 adult flies per generation); hence, sampling of gametes is unlikely to account for the observed between-replicate variance. Moreover, all marker loci would be expected to exhibit a similar change in variance. The more likely cause of the pattern of variation among replicates is variation in selection affecting the marker loci due to random changes in background associations as recombination proceeds.

Data from the marginal single-locus genotypic frequency distributions clearly show that each of the marker loci is strongly influenced by selection. This is best seen by considering Wright's fixation index,  $F$  [ $= 1 - h/2pq$ , where  $h$  is the relative frequency of heterozygotes and  $p = (1 - q)$  is the relative frequency of the slow allozyme]. All the genotypic frequency data derive from the population of males at the adult stage. Thus, the  $F$  statistics, shown in Figures 2a, b, c and d, measure the deviation of the genotypic frequency distribution from a random mating representation after viability selection.

The  $F$  statistics for all loci are initially negative and large in absolute value, indicating a substantial excess of heterozygotes among adults during the early generations of the experiment. As time proceeds, the  $F$  statistics decay toward zero. The time-dependent behavior of the  $F$  statistics can be readily explained by recalling that the initial chromosomal types were probably in disequilibrium for many loci besides the marker loci. Some of these loci affected fitness, and this effect was transmitted to the marker loci through the correlational structure. As recombination proceeded, the disequilibrium between the marker loci and selected loci decayed, thereby reducing the deviation of  $F$  from zero.

One additional feature of the  $F$  statistics is particularly noteworthy. The decay of  $F$  to zero is not the same for all loci. For example, *Idh* appears to have stabilized by generation 10,  $E_s$  and *Pgm* required 15 and 25 generations, respectively, to reach constant values, while  $E_c$  exhibits an erratic behavior over loci. Evidently, the distribution of selected factors is not uniform with respect to the marker loci.

The erratic behavior at  $E_c$  is a consequence of the contamination event. An interaction occurred in heterozygotes for migrant and original chromosomes resulting in the repression of the product of the  $E_c$  gene on the migrant chromosome. This appears to be an example of allelic repression (OHNO 1969; WHITE, CHILDERS and CHO 1973). The progeny-testing procedure employed in sampling gametes permitted the detection of null heterozygotes, *i.e.*, putative homozygous parents that produced testcross progeny exhibiting a different single-banded phenotype from that of the parent. Specifically, the migrant chromosome carried the S allozyme at  $E_c$ . Some phenotypically F parents crossed to a S/S tester produced phenotypically S offspring at  $E_c$ . The observed results are not consistent with an ordinary null allele because no null homozygotes were ever observed in our very extensive population samples. A lethal null can also be excluded because 589  $F_2$  progeny of the cross

$$\frac{E_s (S) \quad E_c (S^n)}{E_s (F) \quad E_c (S)} \times \frac{E_s (S) \quad E_c (S^n)}{E_s (F) \quad E_c (S)}$$

showed no difference between the proportion of  $E_s$  S/S and  $E_s$  F/F offspring, although a highly significant excess of  $E_s$  F/S heterozygotes was observed ( $S^n$  denotes the migrant gene at  $E_c$ ). If the  $S^n$  gene were lethal, the expected proportion of  $E_s$  S/S and F/F would have been  $c/(3+c)$  and  $1/(3+c)$  respectively, where  $c(=0.122)$  is the recombination fraction between  $E_s$  and  $E_c$ . (These data will be the subject of a separate report.) The net effect of the repression phenomenon was to produce spuriously high estimates of homozygous frequencies at  $E_c$  in replicates 1 and 4, and consequently large positive  $F$  values. If we consider only the uncontaminated replicates at  $E_c$ , the  $F$  statistics stabilize at markedly negative values around generation 10. The  $E_c$  marker, unlike the other markers, is influenced by very strong heterotic selection throughout the course of the experiment. This must also account for the low between-replicate variance in gene frequency observed at  $E_c$ .

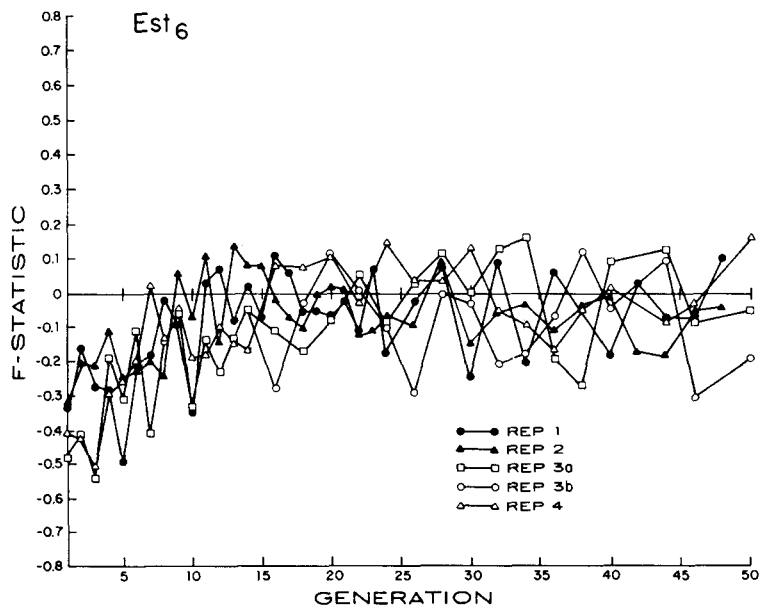
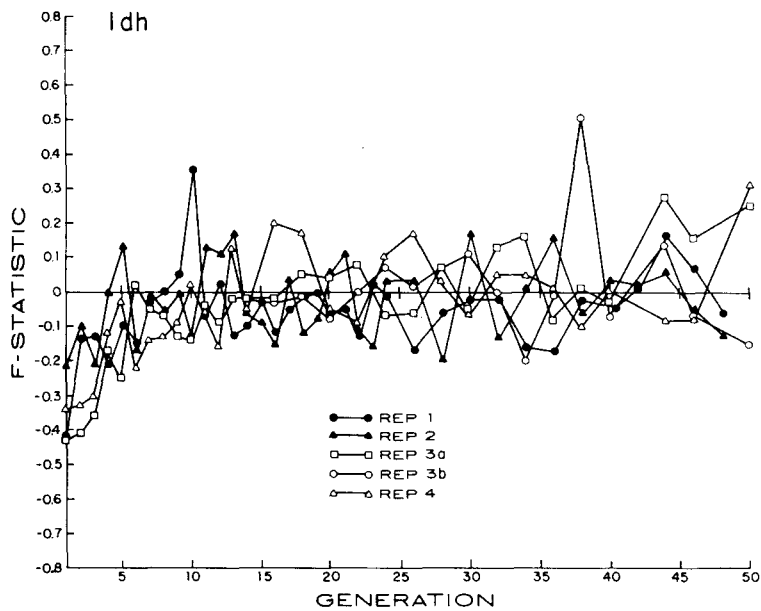
The primary objective of the experiment was to measure the rate of decay of linkage disequilibria. Figures 3a, b, c, d, e and f show the decay dynamics of the correlation in allelic state over loci ( $r$ ) for all six pairwise combinations. The statistic,  $r$ , is defined as

$$r_{AB} = \frac{\Delta_{AB}}{[p_{A_1}(1-p_{A_1})p_{B_1}(1-p_{B_1})]^{1/2}}, \text{ where}$$

$$\Delta_{AB} = g_{A_1B_1} - p_{A_1}p_{B_1}, \text{ and } g_{A_1B_1}$$

is the relative frequency of gametes carrying allele  $A_1$  at locus  $A$  and allele  $B_1$  at locus  $B$ . Marginal single-locus gene frequencies are denoted by  $p_{A_1}$ ,  $p_{B_1}$ , etc. An advantage of  $r$  in parameterizing the decay dynamics is that it is expected to decay geometrically for neutral loci in an infinite population. While the behavior of the estimates of  $r$  are dependent upon finite population size and sample size, this simple expectation provides a useful framework for analyzing the observed dynamics. Reference to Figure 3 shows that, in most instances, the observed rate of decay is markedly faster than expected for neutral loci in an infinite population. This acceleration is especially apparent for loosely linked pairs of loci (e.g., *Idh-Ec*, *Idh-Pgm*). The second point, readily apparent from Figure 3, is that most realizations of the process go to zero within the 50 generations spanned by these experiments, and this is true regardless of the orientation of the initial gametic array.

An approximate estimate of the rate of decay can be obtained from a linear regression of the log plot of the  $r$  statistic *versus* generation. The quantity  $c^* = 1 - \exp[\text{slope of linear regression}]$  can be regarded as an estimate of the effective rate of decay in the experimental populations. Table 2 reports estimates of  $c^*$ , the proportion of the variation in  $r$  statistics that can be accounted for by the linear regression ( $R^2$ ) and the generation at which a realization first crosses zero. All estimates of  $c^*$  are heterogeneous over replicates for all locus pairs (heterogeneity  $\chi^2$  tests were performed excluding replicate 3B, because estimates for replicates 3A and 3B are not independent). Nevertheless, most estimates of  $c^*$  show a markedly faster rate of recombination than would be expected based upon neutral dynamics and the known map distances between the marker loci ( $c$ ).





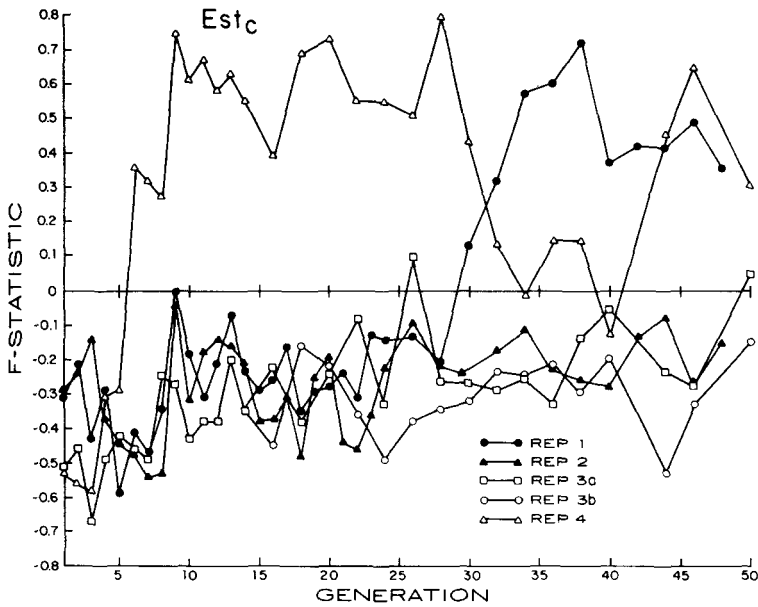
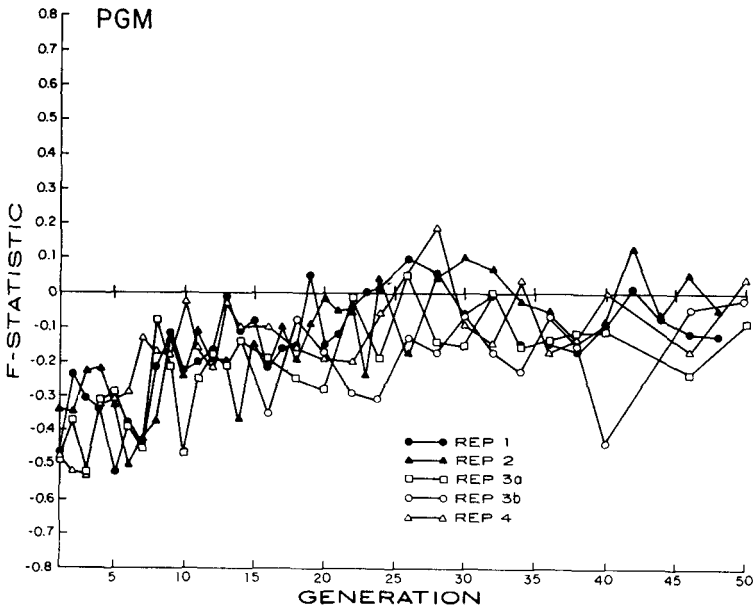
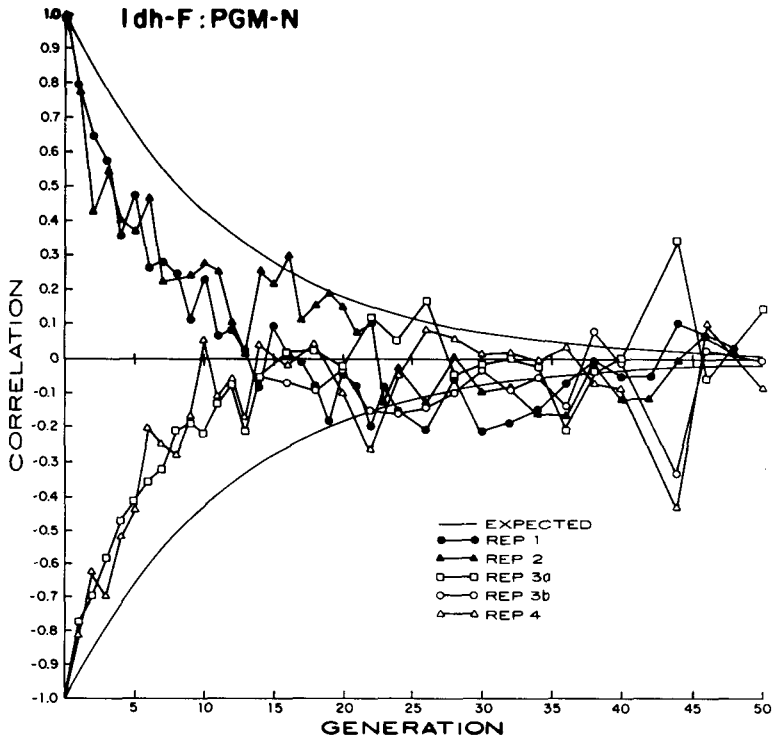
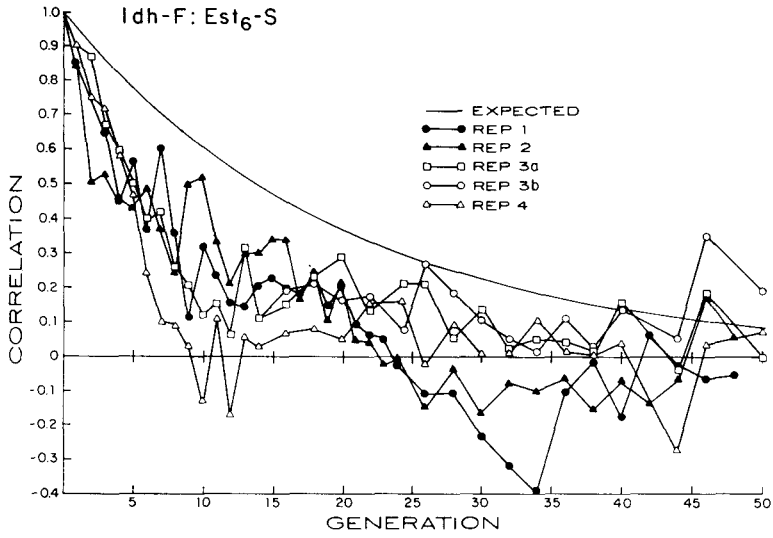
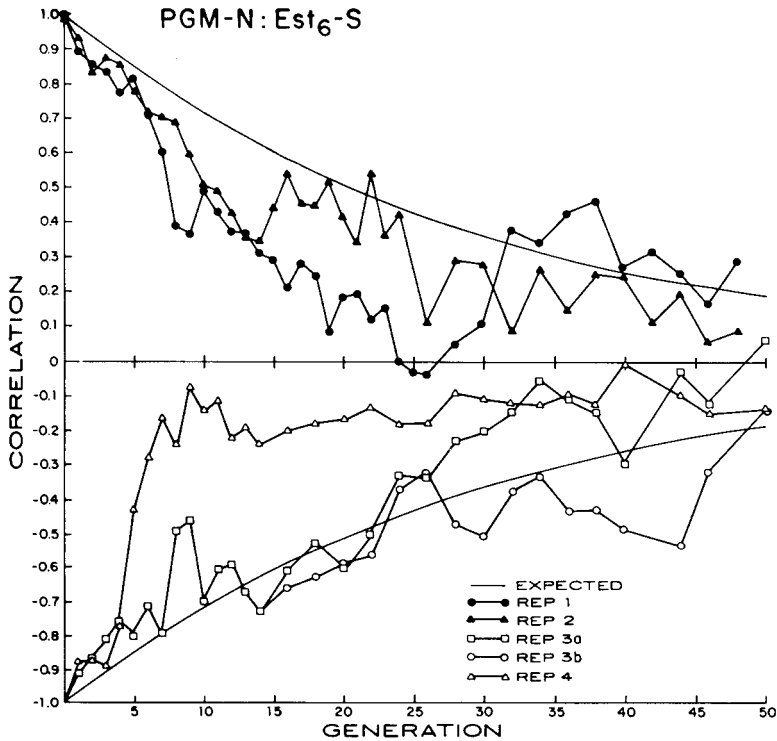
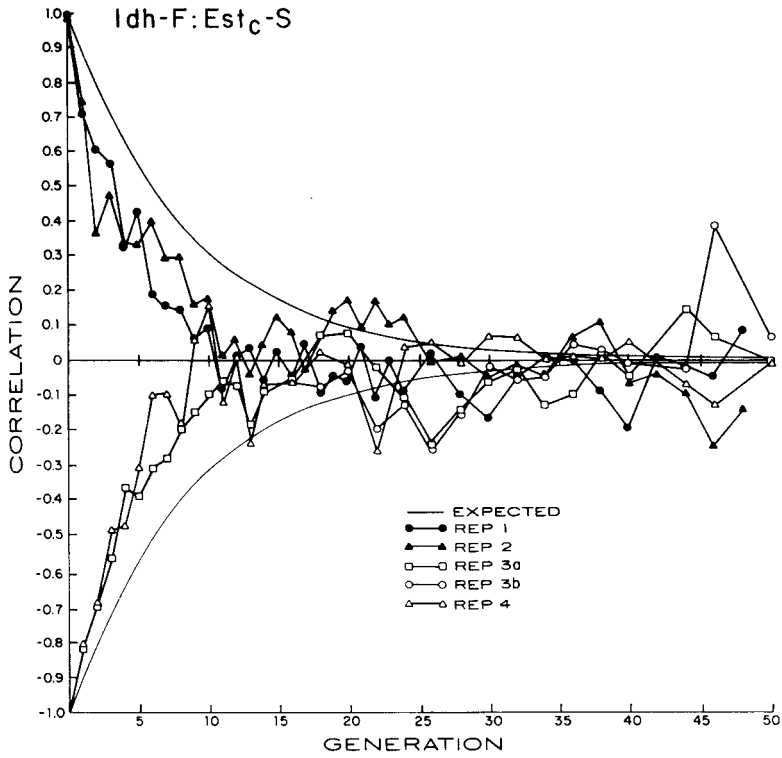


FIGURE 2.— $F$  statistics as a function of generation for loci *Idh* (a),  $E_\theta$  (b), *Pgm* (c), and  $E_c$  (d).





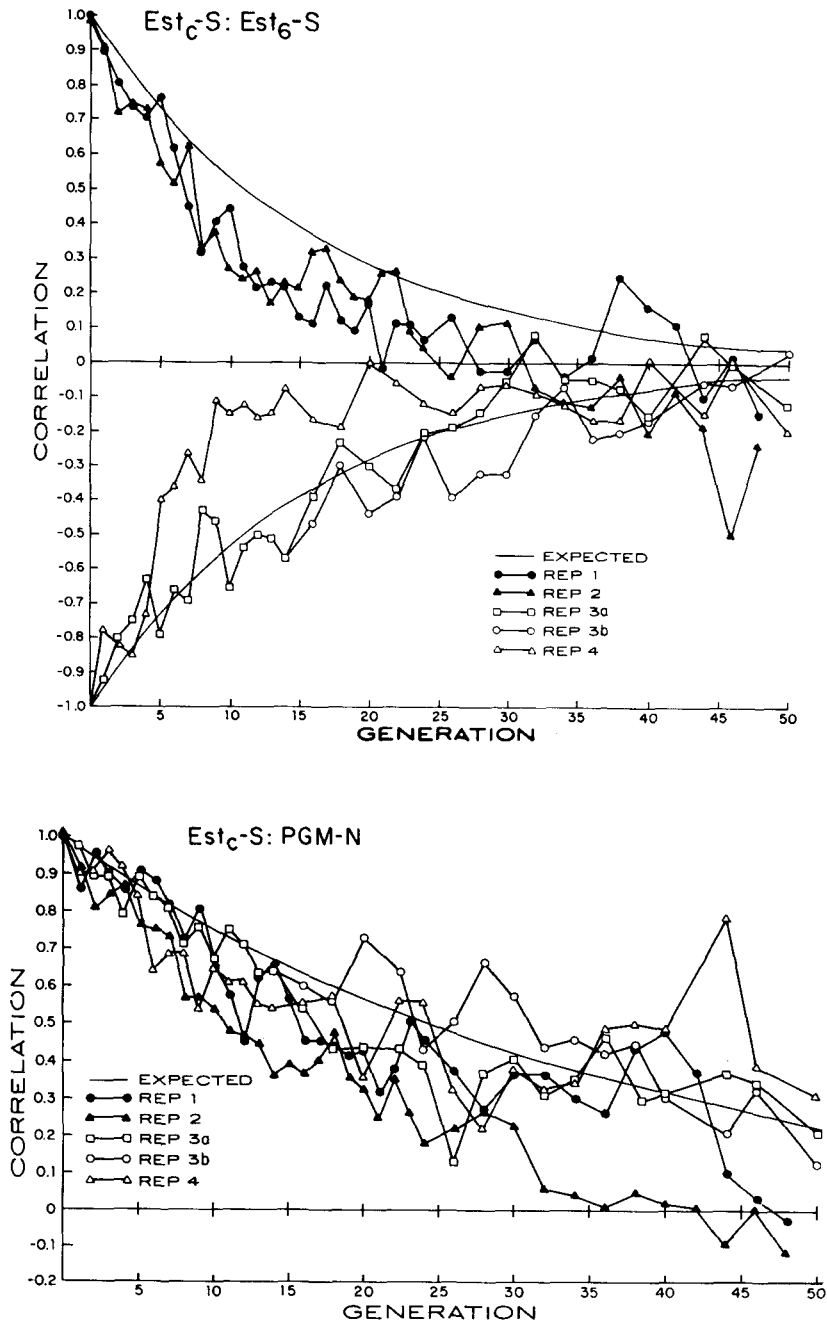


FIGURE 3.—Graphs the behavior of the correlation in allelic state among pairs of loci ( $r$ ), for all six pairwise combinations, as a function of generation. The smooth curves give the expected behavior of  $r$  for loci in an infinite population with the same linkage relationships as the markers under observation. The recombination fractions employed in computing the expected (smooth) curves are taken from O'BRIEN and MACINTYRE (1971).

TABLE 2

*Effective rate of recombination ( $c^*$ ) (with standard errors in parentheses) the proportion of the variance in the log plots accounted for by linear regression ( $R^2$ ), the number of generations until a realization crosses zero ( $n$ ), the ratio of  $c^*$  to the recombination fraction (with standard errors in parentheses), and the estimated map distances between the markers assuming no recombination in males ( $c$ )*

Locus pair	Rep	$c^*$	Statistic SE	$R^2$	$n$	$c^*/c$	SE	$c$
<i>Idh-E<sub>c</sub></i>	1	0.2376	(0.0224)	0.914	10	2.17	(0.204)	0.110
	2	0.2096	(0.0424)	0.658	12	1.91	(0.387)	
	3A	0.1692	(0.0168)	0.866	16	1.74	(0.153)	
	3B	0.0490	(0.0146)	0.348	34	0.51	(0.133)	0.097
	4	0.2695	(0.0472)	0.798	8	2.78	(0.431)	
<i>Idh-Pgm</i>	1	0.2302	(0.0286)	0.818	13	2.82	(0.351)	0.082
	2	0.0780	(0.0204)	0.401	22	0.96	(0.250)	
	3A	0.1645	(0.0171)	0.866	14	1.96	(0.210)	
	3B	0.0605	(0.0117)	0.545	20	0.72	(0.144)	0.084
	4	0.1753	(0.0240)	0.862	9	2.09	(0.294)	
<i>Idh-E<sub>6</sub></i>	1	0.0957	(0.0084)	0.853	23	1.97	(0.173)	0.049
	2	0.0885	(0.0129)	0.679	22	1.82	(0.266)	
	3A	0.0724	(0.0115)	0.611	43	1.11	(0.237)	
	3B	0.0391	(0.0096)	0.366	—	0.60	(0.198)	0.065
	4	0.3366	(0.0346)	0.898	9	5.18	(0.713)	
<i>Pgm-E<sub>c</sub></i>	1	0.0380	(0.0047)	0.652	46	1.36	(0.168)	0.028
	2	0.0800	(0.0068)	0.814	42	2.86	(0.243)	
	3A	0.0276	(0.0034)	0.695	—	2.12	(0.121)	
	3B	0.0277	(0.0025)	0.822	—	2.13	(0.089)	0.013
	4	0.0166	(0.0035)	0.440	—	1.28	(0.125)	
<i>E<sub>6</sub>-E<sub>c</sub></i>	1	0.1100	(0.0084)	0.894	20	1.80	(0.138)	0.061
	2	0.0810	(0.0099)	0.736	24	1.33	(0.163)	
	3A	0.0677	(0.0062)	0.870	30	1.83	(0.102)	
	3B	0.0496	(0.0042)	0.837	46	1.34	(0.069)	0.037
	4	0.0508	(0.0211)	0.179	18	1.37	(0.346)	
<i>Pgm-E<sub>6</sub></i>	1	0.0885	(0.0068)	0.882	23	2.68	(0.204)	0.033
	2	0.0452	(0.0040)	0.785	—	1.37	(0.121)	
	3A	0.556	(0.0057)	0.779	46	2.22	(0.173)	
	3B	0.0215	(0.0030)	0.648	—	0.86	(0.011)	0.025
	4	0.0428	(0.0090)	0.474	—	1.71	(0.273)	

There are two sources for the estimates of map distance ( $c$ ). The first source comes from the original linkage studies (summarized in O'BRIEN and MACINTYRE 1971b). The second source comes from a series of experiments conducted to search for evidence of recombination modification in replicates 3 and 4. This set of experiments involved measuring the recombination fraction among each of the marker loci, once every five generations for the first 20 generations of the experiment. The data, which show no evidence of recombination modification,

reveal some rather large differences in recombination fraction when compared to the standard map values (CLEGG, HORCH and KIDWELL, 1979). The last column of Table 2 gives the estimates of  $c$  obtained from the original linkage studies, followed by the estimates of  $c$  from replicates 3 and 4. The rate of acceleration is expressed as the ratio,  $c^*/c$  (with its standard error computed by assuming that  $c$  is a constant not subject to sampling error). In computing this ratio, the standard map distances were used for replicates 1 and 2 and the recombination fractions estimated from replicates 3 and 4 were used for estimates of  $c$  obtained from these cages. The  $c^*/c$  ratios show substantial variation among replicates within loci and over loci. Nevertheless, the general pattern is one of greatly accelerated decay. This impression is reinforced by an overall mean of 1.82 for the  $c^*/c$  values.

Higher order disequilibria statistics can also be computed from the four-locus gametic frequency distributions. One measure of three-locus association is

$$\Delta_{ABC} = g_{A_1B_1C_1} - p_{A_1}p_{B_1}p_{C_1} - p_{A_1}\Delta_{BC} - p_{B_1}\Delta_{AC} - p_{C_1}\Delta_{AB},$$

(HILL 1974; BENNETT 1954). Figure 4a shows the dynamical behavior of  $\Delta_{ABC}$  normalized to the gene frequency variance as,

$$r_{ABC} = \frac{\Delta_{ABC}}{[p_{A_1}(1-p_{A_1})p_{B_1}(1-p_{B_1})p_{C_1}(1-p_{C_1})]^{1/2}}.$$

The statistic  $r_{ABC}$  has the property that  $Nr_{ABC}^2$  is approximately distributed as  $\chi^2$  with 3 d.f. Consequently, the strength of the deviation of  $r_{ABC}$  from zero can be assessed in a probabilistic framework. The quantity  $r_{ABC}$  is zero for the initial gametic frequency distribution. Changes in the gametic frequencies early in the experiment perturb  $r_{ABC}$  away from zero, but all replicates rapidly converge back to zero and remain there. All four triplets exhibit dynamics like that displayed for  $E_c$ - $Pgm$ - $E_\delta$  in Figure 4a.

Analogous measures of four-locus association can also be computed. Figure 4b displays the behavior of  $r_{ABCD}$ , where

$$r_{ABCD} = \frac{\Delta_{ABCD}}{[p_{A_1}(1-p_{A_1})p_{B_1}(1-p_{B_1})p_{C_1}(1-p_{C_1})p_{D_1}(1-p_{D_1})]^{1/2}},$$

and where  $\Delta_{ABCD}$  is the four-locus disequilibrium function defined by BENNETT (1954) and HILL (1974). For neutral loci in an infinite random mating population,  $\Delta_{ABCD}$ , and hence  $r_{ABCD}$ , are expected to decay to zero at a rate governed by the recombination rate among all four loci (BENNETT, 1954). Figure 4b shows that the estimated values of  $r_{ABCD}$  do decay to zero for all replicates, but at a rate much faster than expected for neutral loci in an infinite population.

Another measure of association that can be applied to the four-locus gametic frequency distributions arises in multidimensional contingency analysis and is defined as (SMOUSE 1974),

$$Z_{ABCD} = \frac{g_{1111} g_{2211} g_{1112} g_{2212} g_{1121} g_{2221} g_{1122} g_{2222}}{g_{1211} g_{2111} g_{1212} g_{2112} g_{1221} g_{2121} g_{1222} g_{2122}},$$

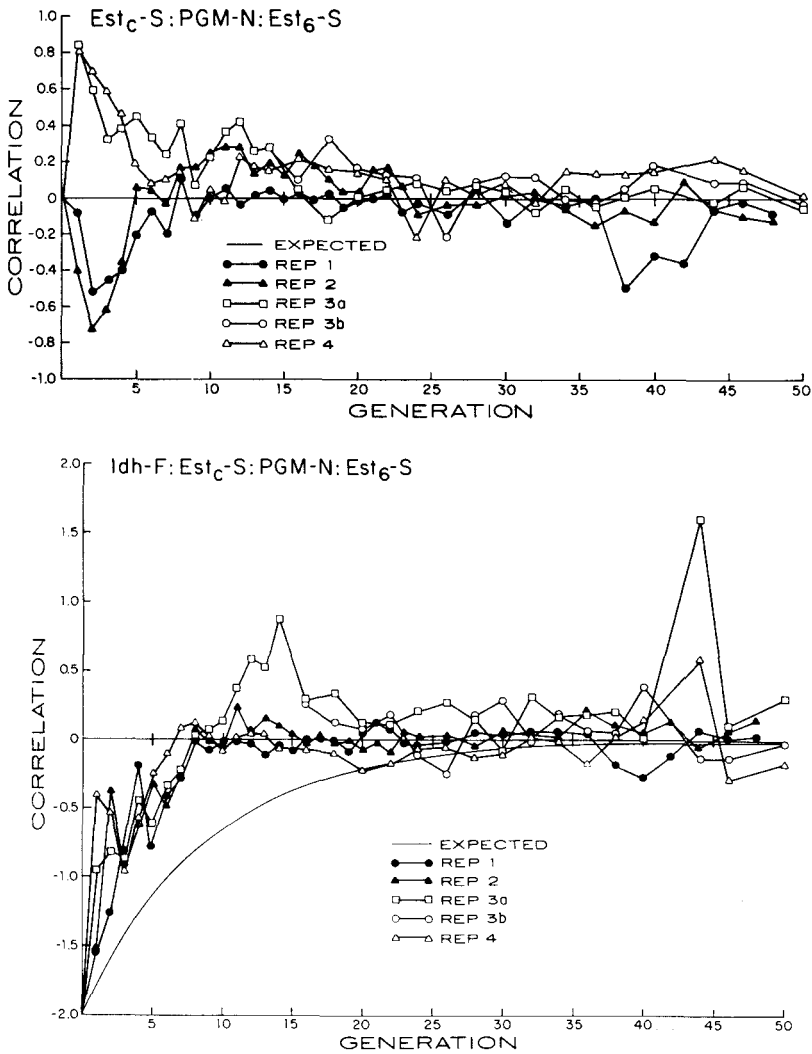


FIGURE 4.—Graphs of the dynamic behavior of three- and four-locus measures of association in 4a and 4b, respectively. The smooth curve in Figure 4b gives the expected rate at which  $r_{ABCD}$  will decay for neutral genes in an infinite population.

where  $g_{ijkl}$  is the relative frequency of a gamete carrying genes  $A_i B_j C_k D_l$  at loci  $A, B, C$  and  $D$ , respectively, ( $i, j, k, l = 1, 2$ ). The quantity  $z_{ABCD} = \log Z_{ABCD}$  can be partitioned into various measures of conditional association. For example,

$z_{AB/C_1 D_1} = \log \left[ \frac{g_{1111} g_{2211}}{g_{1211} g_{2111}} \right]$ , measures the association between  $A$  and  $B$  in gametes carrying alleles  $C_1$  and  $D_1$ . It is easy to verify that  $z_{ABCD} = z_{AB/C_1 D_1} + z_{AB/C_1 D_2} + z_{AB/C_2 D_1} + z_{AB/C_2 D_2} = z_{ABC/D_1} + z_{ABC/D_2}$ . Thus it is possible to examine the dynamic behavior of the association between a given pair or triplet in the presence of a specific background at the remaining loci. A complete analysis of all the data failed to reveal any instance where the association between a given

set of loci depended upon the state of the remaining loci. We, therefore, conclude that within the limits of the resolving power of these experiments, there is no interaction between decay rates at a specific pair or triplet of loci and genetic background.

#### DISCUSSION

The central result of the present experiment is the rapid decay of the correlation of gene frequencies to zero, or near zero, for all pairs of loci monitored. Several factors contribute to the acceleration of decay rates. One factor is gene frequency change. The expected neutral dynamics is predicated upon constant gene frequencies; however, many of the marker loci exhibit substantial changes in gene frequency during the course of the experiment. Selection is also important in accounting for the decay dynamics independent of its effect on gene frequencies. CLEGG (1978) through Monte Carlo simulations of 93-locus models, and ASMUSSEN and CLEGG (in preparation) have shown algebraically that selection and recombination interact to accelerate decay rates, provided linkage disequilibrium is high and selection is weak relative to recombination. The acceleration phenomenon arises from the effect of selection on joint genotypic frequency distributions. Selection favoring heterozygotes increases the frequency of multiple heterozygotes, thereby increasing the effective rate of recombination (CLEGG 1978). Thus, the acceleration phenomenon is not surprising, and it does have important implications in population genetics.

An especially significant implication concerns the utility of population cage experiments. A procedure sometimes employed in establishing experimental populations of *Drosophila* has been to intercross several stocks and maintain these newly synthesized populations for a year or more before initiating selection experiments. Such procedures have been criticized on the grounds that linkage associations will not have sufficient time to dissipate, especially if there is selection. The present results suggest, to the contrary, that the average segment of chromosome that remains associated with a marker locus may be much smaller in a system with selection than would obtain under neutrality. This means that marginal effects measured at the marker locus will involve a relatively small block of genes and may arise from the marker locus itself.

If there is reason to suppose that marginal effects can be measured at single loci, what are the criteria for measuring such effects? One important criterion involves the control of genetic background through experimental design. Examples of careful experiments that sought to control background effects are those of YAMAZAKI (1971) on the esterase-5 locus of *D. pseudoobscura* and of BIJLSMA and VAN DELDEN (1977) on the glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase loci of *D. melanogaster*. In the latter case, the experimental design involved synthesizing successive replicate cage experiments from an initial population whose history of recombination ranged from six months through two years. A second useful criterion for measuring marginal effects involves the manipulation of experimental environments in ways that can be directly related to the function of the gene under observation. This



method, first advanced by GIBSON (1970) and WILLS and NICHOLS (1971), has been applied to the alcohol dehydrogenase locus of *D. melanogaster* with some success (CAVENER and CLEGG 1978, and references therein). Thus, while the rapid decay of linkage disequilibria makes the measurement of selection at single loci feasible, it certainly does not obviate the need for carefully designed experiments.

A noteworthy feature of the experimental results is the markedly different behavior of the *F* statistics over loci. The intensity of selection affecting *ldh* is much less than that affecting *E<sub>c</sub>*, while *E<sub>e</sub>* and *Pgm* are intermediate in this regard. These data are consistent with previous experiments that show strong selection between *E<sub>c</sub>* and *Pgm* (CLEGG *et al.* 1976; CLEGG, KIDWELL and KIDWELL 1978). It appears that the distribution of selection intensities is not uniform along chromosome 3 and that the centromeric region marked by *E<sub>c</sub>* is under very strong selection.

While unplanned, the contamination event provided two interesting observations. First, the migrant chromosomes were very rapidly incorporated into the recipient populations. We can infer this rapid incorporation from the gene frequency behavior at all the marker loci in replicates 1 and 4, following the contamination event, even though the migrant third chromosomes were not marked by unique allozymes. The second interesting observation concerned the apparent repression of the *E<sub>c</sub>* allozyme carried on migrant chromosomes in heterozygous combinations with nonmigrant chromosomes. The mechanism of the repression phenomenon is not understood, but it may suggest regulatory differentiation between these genomes. Whatever the mode of the interaction, it did not retard the rapid increase in the migrant type.

Taken in total, these experiments support numerous studies of natural populations of *D. melanogaster* (reviewed by HEDRICK, VAIN and HOLDEN 1978) in indicating that little linkage disequilibrium is expected among enzyme loci. A rationale for this nearly universal negative finding is also provided by these results. Selection and recombination interact to rapidly dissipate nonrandom associations unless the loci are very tightly linked and/or strongly interacting. Therefore, we are led to the conclusion that most loci, even tightly linked loci, will be close to linkage equilibrium in large populations of random mating organisms. Indeed, the few cases of consistent linkage disequilibrium among loci (*e.g.*, BAKER 1975; LOUKAS and KRIMBAS 1975; ROBERTS and BAKER 1973) may constitute *prima facie* evidence of epistatic selection.

We thank R. W. ALLARD, P. HEDRICK and B. S. WEIR for commenting on the manuscript. We also wish to express our gratitude to DEBORAH CLEVINGER, NANCY DANIEL, JAMES FELLY, JUDY MOSINGER and KATHY RONES for excellent technical assistance and to MATTHEW CLEGG for computer programming. This work was supported by National Science Foundation Grants DEB76-82479 and DEB79-00822.

#### LITERATURE CITED

- BAKER, W. K., 1975 Linkage disequilibrium over space and time in natural populations of *Drosophila montana*. Proc. Natl. Acad. Sci. U.S. **72**: 4095-4099.
- BENNETT, J. H., 1954 On the theory of random mating. Ann. Eugen. **18**: 311-317.

- BIJLSMA, R. and W. VAN DELDEN, 1977 Polymorphism at the G6PD and 6PGD loci in *Drosophila melanogaster*. I. Evidence for selection in experimental populations. *Genet. Res.* **30**: 221-236.
- BIRLEY, A. J., 1974 Multi-locus polymorphism and selection in a population of *Drosophila melanogaster*. I. Linkage disequilibrium on chromosome III. *Heredity* **32**: 122-127.
- BREWER, G. J., 1970 *An Introduction to Isozyme Technique*. Academic Press, New York.
- BROWN, A. H. D., 1978 Enzyme polymorphism in plant populations. *Theoret. Pop. Biol.* **15**: 1-42.
- CAVENER, D. R. and M. T. CLEGG, 1978 Dynamics of correlated genetic systems. IV. Multilocus effects of ethanol stress environments. *Genetics* **90**: 629-644.
- CLEGG, M. T., 1978 Dynamics of correlated genetic systems. II. Simulation studies of chromosome segments under selection. *Theoret. Pop. Biol.* **13**: 1-23.
- CLEGG, M. T., C. R. HORCH and J. F. KIDWELL, 1979 Dynamics of correlated genetic systems. VI. Variation in recombination rates in experimental populations of *Drosophila melanogaster*. *J. Heredity* **70**: 297-300.
- CLEGG, M. T., J. F. KIDWELL and M. G. KIDWELL, 1978 Dynamics of correlated genetic systems. III. Behavior of chromosomal segments under lethal selection. *Genetica* **48**: 95-106.
- CLEGG, M. T., J. F. KIDWELL, M. G. KIDWELL and N. J. DANIEL, 1976 Dynamics of correlated genetic systems. I. Selection in the region of the glued locus of *Drosophila melanogaster*. *Genetics* **83**: 793-810.
- GIBSON, J., 1970 Enzyme flexibility in *Drosophila melanogaster*. *Nature* **227**: 959-960.
- HEDRICK, P., S. JAIN and L. HOLDEN, 1978 Multilocus systems in evolution. *Evol. Biol.* **11**: 101-184.
- HILL, W. G., 1974 Disequilibrium among several linked neutral genes in finite populations. I. Mean changes in disequilibrium. *Theoret. Pop. Biol.* **5**: 366-392.
- KARLIN, S., 1975 General two-locus selection models: some objectives, results and interpretations. *Theoret. Pop. Biol.* **7**: 364-398.
- LOUKAS, M. and C. B. KRIMBAS, 1975 The genetics of *Drosophila subobscura* populations. V. A study of linkage disequilibrium in natural populations between genes and inversions of the *E* chromosome. *Genetics* **80**: 331-347.
- O'BRIEN, S. J. and R. J. MACINTYRE, 1971 Transient linkage disequilibrium in *Drosophila*. *Nature* **230**: 335-336. —, 1971 A biochemical genetic map of *D. melanogaster*. *Drosophila Inform. Serv.* **46**: 89-93.
- OHNO, S., 1969 The preferential activation of maternally derived alleles in development of interspecific hybrids. pp. 137-150. In: *Heterospecific Genome Interaction*. Edited by V. DEFENDI. Wistar Inst. Press, Philadelphia.
- RASMUSON, M., B. RASMUSON and L. R. NILSON, 1967 A study of isoenzyme polymorphism in experimental populations of *Drosophila melanogaster*. *Hereditas* **57**: 263-274.
- ROBERTS, R. M. and W. K. BAKER, 1973 Frequency distribution and linkage disequilibrium of active and null esterase isozymes in natural populations of *Drosophila montana*. *Am. Naturalist* **107**: 709-726.
- SMOUSE, P., 1974 Likelihood analysis of recombinational disequilibrium in multiple-locus gametic frequencies. *Genetics* **76**: 557-565.
- WHITT, G. S., W. F. CHILDERS and P. L. CHO, 1973 Allelic expression at enzyme loci in an intertribal hybrid sunfish. *J. Heredity* **64**: 55-61.
- WILLS, C. and L. NICHOLS, 1971 Single gene heterosis in *Drosophila* revealed by inbreeding. *Nature* **233**: 123-125.
- YAMAZAKI, T., 1971 Measurement of fitness at the esterase-5 locus in *Drosophila pseudoobscura*. *Genetics* **67**: 579-603.