# A MODEL OF FUNCTIONAL EPISTASIS AND LINKAGE DISEQUILIBRIUM IN POPULATIONS WITH OVERLAPPING GENERATIONS*

JAMES T. GIESEL

*Department of Zoology, University of Florida, Gainesville, Florida 32611*

Manuscript received April 15, 1976
Revised copy received March 2, 1977

ABSTRACT

A model of functional epistasis is proposed in which it is assumed that coupling and repulsion genotypes differ in metabolic efficiency and thus in development time and net fecundity. The implications of this model are investigated for iteroparous populations with fluctuating rates of increase. It is found that the fluctuations in rate of increase can lead to large fluctuations in gamete frequency and $D$, the coefficient of linkage disequilibrium, but that $D$ will almost always have a value of zero at some point during the populations' demographic cycle. Some of the model populations would be expected to be in a state of linkage disequilibrium only fleetingly: others would exhibit $D$-cycles interpretable as random fluctuation. Implications of the model for interpretations of existing data on linkage disequilibrium among enzyme loci in Drosophila are discussed.

POPULATION genetic theory predicts that epistatically interacting genes, where coupling or repulsion genotypes and gametes enjoy a fitness advantage, should reach a state of linkage disequilibrium in which the more fit gametes are present in higher frequency than their less fit counterparts. In the limit

$$D = g_{AB} \cdot g_{ab} - g_{Ab} \cdot g_{aB} \tag{1}$$

is expected to approach a maximum value near $\pm$ 0.25 where the less fit gametes are present only by virtue of recombination, *assuming that fitness differentials are directionally constant in time.*

Several studies designed to investigate states of linkage disequilibrium among gene-enzyme loci have been done. ZOUROS and KRIMBAS (1973) found linkage disequilibrium between aldehyde oxidase and xanthine dehydrogenase in two distinct populations of *Drosophila subobscura*, an observation which is strongly suggestive of epistatic selection. In addition, ROBERTS and BAKER (1973) found linkage disequilibria among a complex of esterases, a result predictable from the finding that these loci segregate for null alleles in such a way that the presence of the "nulls" should exert a strong selective organizing effect. However, the bulk of published results (MUKAI, WATANABE and YAMAGUCHI 1974; O'BRIEN

* This is the second of a series of papers on the population genetics of species with overlapping generations.

and McINTYRE 1971; O'BRIEN, McINTYRE and FINE 1968; CHARLESWORTH and CHARLESWORTH 1973; LANGLEY, TOBARI and KOJIMA 1974; PRAKASH 1974) provides no strong evidence of *selectionally derived* linkage disequilibrium or, consequently, of epistatic selection among isozyme loci. In fact, the variations in $D$ which were noted by some of these investigators have often been interpreted to lend support to "neutralist" explanations for the maintenance of genetic variation.

I contend that the lack of correspondence between the expectations of selection theory and experimental observations may commonly stem from inadequately defined and interpreted theory. It is the purpose of this communication to re-evaluate expectation by proposing a new, and perhaps more realistic, population biological model for two-locus selection, and to show that, given the model, strong epistatic interactions may be almost impossible to detect by assessment of states of linkage disequilibrium between the loci in question.

## THE MODEL SYSTEM

Assume that we are dealing with an iteroparous population with overlapping generations and that in such a population the fitness of any genotype will depend on the momentary relationship between the genotypes' fecundity distributions and the population's age distribution (GIESEL 1974). When a population reaches stable age distribution the fitnesses of its genotypes can be represented in a relative sense by:

$$W_{i,j(r)} = \sum_{x=b}^{x=d} l_{i,j(x)} m_{i,j(x)} e^{-rx} \tag{2}$$

where $e$ is the base of natural logarithms, $r$ is the population's rate of increase, $l_{i,j(x)}$ is the probability that females of the $i,j$-th genotype will be alive at reproductive age $x$, $m_{i,j(x)}$ is the number of offspring females of the $i,j$-th genotype are expected to produce at age $x$, and $W_{i,j(r)}$ is the $r$-dependent fitness of the $i,j$-th genotype (CHARLESWORTH and GIESEL 1972). Thus, if the genotypes at a locus differ in development time or age-specific distribution of fecundity, changes in the population's rate of increase will cause changes in the relative fitnesses of the genotypes. When $r$ is positive, an early reproducing genotype will be more fit than a later reproducing genotype and will increase in frequency in the population; but the same early reproducing genotype will have a relatively low fitness and a declining frequency in a declining population. If we further assume that the lifetime production of young by heterozygotes at the locus in question exceeds that of either homozygote, a system of temporally balanced polymorphism results with allele frequencies cycling between limits set by:

$$\hat{p}_r = \frac{W_{Aa(r)} - W_{aa(r)}}{2W_{Aa(r)} - W_{AA(r)} - W_{aa(r)}} \tag{3}$$

as the population's rate of increase fluctuates, and these equilibria will be stable so long as the population's proportional age distribution remains stable.

To extend this model and its conclusions to the two gene case, we need to make the following assumptions. First, assume that we have two loci and that each locus segregates for two alleles. Next, assume epistatic interaction to be of a form such that the coupling and repulsion homozygotes differ in metabolic efficiency and thus in the components of fitness. Let the coupling homozygotes have a high metabolic efficiency and consequently a short development time and great lifetime production of offspring, and the repulsion homozygotes have a lower metabolic efficiency and consequently a longer development time and lower lifetime fecundity. If we then assume, as is usual (*cf.*, LANGLEY and CROW 1974) that single and double gene heterozygotes have a higher fitness (or are slightly more efficient) than either of the sets of double gene homozygotes, we will have produced a system of temporally balanced polymorphism (GIESEL 1972, 1974). Finally, in order to have a system in which the only fitness differentials will be due to the functional epistasis, as I have defined it, assume that the metabolic efficiencies of the single gene genotypes at each locus are equal.

Thus I have defined a population which consists of three physiologically, demographically and selectionally distinct sets of genotypes: heterozygotes, coupling gamete homozygotes, and repulsion gamete homozygotes, and we can investigate the effects of population rate of increase on the equilibrium fitnesses of the genotypes and the equilibrium frequencies of the gametes.

For the purposes of the *sample* calculation presented in Table 2, net fecundity distributions of the general form depicted in Figure 1 and calculated as in Table 1 were employed. In producing the example data depicted in Table 2, an eleven-adult-age class population was imagined. The population was assumed to have an accidental death survivorship distribution with a death rate of 0.9 applying from eclosion onward. This, together with a triangular $m_x$ distribution, is shown in Table 1, where the heterozygotes' $l_x m_x$ distributions are also presented. Net reproduction ($R_o$) was adjusted to 1.33456 for the heterozygotes and to 1.00 for

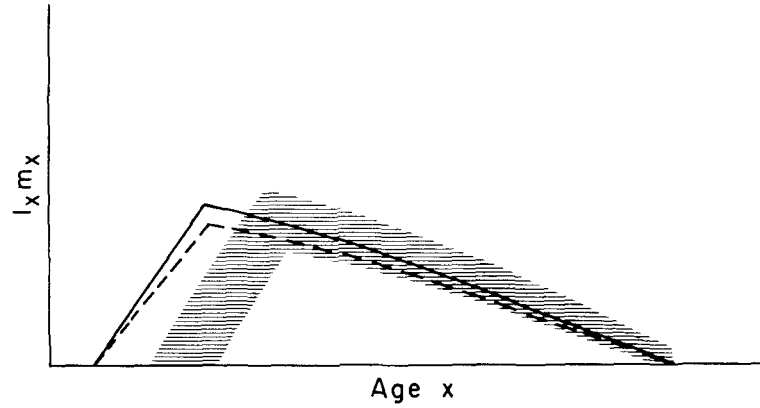


FIGURE 1.—Net fecundity distribution of: one and two gene heterozygotes (solid line), coupling gamete homozygotes (dashed line), and repulsion gamete homozygotes (various treatments are encompassed by the shaded area). In one case in each development time series net reproduction is brought to 1.0 for both CGH and RGH.

TABLE 1

*The basic life table distributions used to produce Table 2*

| $x$ | $l_x$ | 1 $m_x$ | $l_xm_x$ | 2 $m_x$ | $l_xm_x$ |
|---|---|---|---|---|---|
| 1 | 0.076 | 1 | 0.076 | | |
| 2 | 0.0684 | 3 | 0.2052 | 1 | 0.05228 |
| 3 | 0.06156 | 6 | 0.36936 | 3 | 0.14117 |
| 4 | 0.0547 | 5 | 0.2736 | 6 | 0.25097 |
| 5 | 0.04788 | 4 | 0.19152 | 5 | 0.20392 |
| 6 | 0.04104 | 3 | 0.12312 | 4 | 0.12549 |
| 7 | 0.0342 | 2 | 0.0684 | 3 | 0.07843 |
| 8 | 0.02736 | 1 | 0.02736 | 2 | 0.04182 |
| 9 | 0.02462 | 0 | 0 | 1 | 0.01569 |
| 10 | 0.02216 | | ——— | 0 | ——— |
| 11 | 0.01995 | | $\Sigma = 1.33456 = R_o$ (het) | | $\Sigma = 0.9096 = R_o$ |

An accidental probability of death of 0.9 per age class commencing with eclosion is assumed, and $m_x$ values were chosen to give triangular $m_x$ distributions. Late development was modeled by rigid translation of the $m_x$ schedule (as in column 2).

TABLE 2

*Equilibrium frequencies of repulsion gametes and* $D_r$ *values for cases in which model populations differ in development time and/ net reproduction per female*

| | | | | |
|---|---|---|---|---|
| | Development time differential is: 1 day | | | |
| Repulsion gamete net fecundity (as % of coupling gamete value) | 100 | 95 | 9096 | 8997 |
| $r_m = +0.1$ | 0.4785(0.0108) | 0.4200(0.040 ) | 0.3971(0.0515) | 0.3930(0.0535) |
| $r_m = 0.0$ | 0.5000(0.0000) | 0.4653(0.0174) | 0.4405(0.0398) | 0.4348(0.0326) |
| $r_m = -0.1$ | 0.5873(—0.044) | 0.5353(—0.018) | 0.5000(0.0000) | 0.4930(+0.004) |
| | Development time differential is: 2 days | | | |
| Repulsion gamete net fecundity (%) | 100 | | 85 | |
| $r_m = +0.1$ | 0.3981(0.0510) | | 0.3456(0.0772) | |
| $r_m = 0.0$ | 0.5000(0.0000) | | 0.4084(0.0458) | |
| $r_m = -0.1$ | 0.6828(—0.091) | | 0.5000(0.0000) | |
| | Development time differential is: 3 days | | | |
| Repulsion gamete net fecundity (%) | 100 | | 7613 | |
| $r_m = +0.1$ | 0.3247(0.0876) | | 0.2829(0.1085) | |
| $r_m = 0.0$ | 0.5000(0.0000) | | 0.3690(0.0655) | |
| $r_m = -0.1$ | 0.9405(—0.220) | | 0.5000(0.0000) | |

Equilibrium values are calculated on the assumption that the populations have reached stable age distributions appropriate to $r_m$ values of +0.10, 0.0, and —0.10 using the basic life table distribution shown in Table 1. $D_r$ values are given in brackets after the appropriate gamete frequencies.

the coupling gamete homozygotes, which began reproduction at the same age as the heterozygotes. The $l_xm_x$ distributions of the repulsion gamete homozygotes were produced by shifting their $m_x$ distribution 1, 2, or 3 age classes later than the other genotypes and their $R_o$ values were adjusted by multiplying the resultant $l_xm_x$ values by an appropriate constant (as in column 2 of the table.). Using such sets of life tables it was possible to calculate $r$-dependent equilibrium genotype fitnesses (equation 2), gamete frequencies (equation 3), and values of $\hat{D}_r$ (equation 1) for rates of population increase of +0.10, 0.0, and −0.10 for each population.

## RESULTS

The calculations presented in Table 2 demonstrate that over a wide range of development time and net fecundity differentials, $\hat{D}_r$ varies between positive and negative values—depending on the population's rate of increase and the relations of the genotypes' net fecundity distributions. When the coupling and repulsion genotypes differ only in development time, $\hat{D}_r$ varies about zero. But as the repulsion genotypes become less efficient (with longer development times and, especially, lower net fecundities) the systems lose symmetry, and $\hat{D}_r$ becomes more positive, overall. Too great efficiency differentials, expressed as losses of net fecundity, (as in example 1 − .899) can result in unconditionally positive $D$ values. Nevertheless, with the range of these examples there is usually an $r$ for which $\hat{D}_r$ equals zero, even when functional epistasis is exceedingly strong.

The calculations presented here merely exemplify expected behaviors of $D$. Obviously, actual fitnesses, gamete frequencies and values and behaviors of $\hat{D}_r$ will be dependent on actual population structures, demographics and actual relative "efficiencies" of genotypes and one can imagine an infinite number of these.

However, the ultimate meaning of such a basic pattern should be clear. Since gamete frequencies will always tend to *proceed toward* appropriate $r$-dependent equilibrium values, $D$ is expected to *cycle* between its expected equilibria. In cases where genotypes differ only in development time, the mean value of the $D$-cycle will be equal to or very near zero, and statistically significant positive or negative values of D may be found only near the peaks and nadirs of the population's numerical cycle. In fact, if population $r$ changes very rapidly, $D$ may vary only slightly from zero, even though strong epistasis exists and highly significant values of $D$ would be probable with a longer period demographic fluctuation. Then, too, if a population is demographically stable and stationary ($r = 0$), cases in which epistasis is expressed only as a development time differential, no matter how large, will be completely cryptic to elucidation by assessment of $D$. If, finally, a population has a long period, high amplitude demographic cycle, and any of a wide range of efficiency differentials, $D$ will be sometimes neagtive and sometimes positive so that investigators may be led to conclude that it is varying randomly (see MUKAI, WATANABE and YAMAGUCHI 1974). Obviously, as functional epistasis of the type discussed here increases

in strength (as in the lower RHS of Table 1 and perhaps ZOUROS and KRIMBAS 1973), its detection by analysis of linkage disequilibrium will become more probable.

## DISCUSSION

The basic assumption of this model—that naturally occurring genes and gene complexes differ in the components of net fecundity distribution and that their interactions are often, indeed commonly, of an epistatic nature—is attested to by several studies. Thus, SPASSKY, DOBZHANSKY and ANDERSON (1965) showed a positive viability synergism between second and third chromosomes in *Drosophila pseudoobscura*. TEMIN *et al.* (1969) were able to detect a slight epistatic effect on viability between second and third chromosome mild detrimentals and between mild detrimentals within a chromosome in *Drosophila melanogaster*. SPIESS (1959) presented results which indicate strong nonadditive interactions on viability between second chromosome regions of flies collected from diverse ecological regions in three species of Drosophila and noted that "epistatic effects between loci are vastly more important than additive ones." Finally, MUKAI (1969) identified synergistic interaction among newly arisen viability polygenes located in the same chromosome of *Drosophila melanogaster*. There are numerous examples of nonrandom associations of inversions which occur in natural populations and are apparently due to epistatic fitness interaction (cf. LEVITAN, 1958; LEVITAN and SALZANO 1959; STALKER 1960). Such interactions have also been shown to act with respect to the other components of fitness—another requirement of the model. Thus, the studies of TEMIN (1966), WILSON (1968), MUKAI and YAMAZAKI (1971), and BARKER and PODGER (1970) all demonstrate that there are strong correlations between the components of fitness in *Drosophila melanogaster*, and that genotypic (WILSON 1968; MUKAI and YAMAZAKI 1971) and environmental "stress" (BARKER and PODGER 1970) affect the other components of fitness proportionally more than they do viability. Data on this point are also available for at least two other species. Data collected by DOBZHANSKY, LEWONTIN and PAVLOVSKY (1964) for *Drosophila pseudoobscura* suggest that efficiency differentials are expressed primarily as differences in development time and in net fecundity. In *Drosophila persimilis* two different gene arrangements differ in development time and fecundity, and stress (in the form of crowding) affects only development time (SPIESS 1958).

Thus we find that within Drosophila there is great variation in the ways in which genotypic and environmental effects alter the net fecundity distribution, and that *in general form* the alterations and differentials are in accord with the requirements of the model I have discussed here. Nevertheless, the worth of the model is speculative in the case of electrophoresable gene-enzymes. Testing it must involve making direct comparisons of the forms of the net fecundity distributions of coupling and repulsion phase genotypes of two or more gene enzymes and then calculating, except in obvious cases, their fitnesses for a variety of rates of increase. Until such work is completed, there will be no reason to choose this model and is conclusions over those offered by the neutralist school, for any particular case.

## LITERATURE CITED

BARKER, J. S. F. and R. N. PODGER, 1970 Interspecific competition between *Drosophila melanogaster* and *Drosophila simulans*: effects of larval density on viability, developmental period, and body weight. Ecology **51**: 170–190.

CHARLESWORTH, B. and D. CHARLESWORTH, 1973 A study of linkage disequilibrium in populations of *Drosophila melanogaster*. Genetics **73**: 351–359.

CHARLESWORTH, B. and J. T. GIESEL, 1972 Selection in populations with overlapping generations. II. Relations between gene frequency and demographic variables. Am. Naturalist **106**: 388–401.

DOBZHANSKY, TH., R. C. LEWONTIN and O. PAVLOVSKY, 1964 The capacity for increase in chromosomally polymorphic and monomorphic populations of *Drosophila pseudoobscura*. Heredity **19**: 597–614.

GIESEL, J. T., 1972 The maintenance of genetic variability in natural populations—an alternative implication of the CHARLESWORTH-GIESEL hypothesis. Am. Naturalist **106**: 412–414.
——, 1974 The relative fitness of populations polymorphic for net fecundity distribution. Am. Naturalist **108**: 321–332.

LANGLEY, C. H. and J. F. CROW, 1974 The direction of linkage disequilibrium. Genetics **78**: 937–941.

LANGLEY, C. H., Y. N. TOBARI and K. KOJIMA, 1974 Linkage disequilibrium in natural populations of *Drosophila melanogaster*. Genetics **78**: 921–936.

LEVITAN, M., 1958 Non-random associations of inversions. Cold Spring Harbor Symp. Quant. Biol. **23**: 251–268.

LEVITAN, M. and F. M. SALZANO, 1959 Studies of linkage in populations. III. An association of linked inversions in Drosophila guaramunu. Heredity **13**: 243–248.

MUKAI, T., 1969 The genetic structure of natural populations and *Drosophila melanogaster*. VII. Synergistic interaction and spontaneous mutant polygenes controlling viability. Genetics **61**: 749–761.

MUKAI, T., T. K. WATANABE and O. YAMAGUCHI, 1974 The genetic structure of natural populations of *Drosophila melanogaster*. XII. Linkage disequilibrium in a large natural population. Genetics **77**: 771–793.

MUKAI, T. and T. YAMAZAKI, 1971 The genetic structure of natural populations of Drosophila melanogaster. X. Development time and viability. Genetics **69**: 385–398.

O'BRIEN, S. J. and R. J. MCINTYRE, 1971 Transient linkage disequilibrium in Drosophila. Nature **230**: 335–336.

O'BRIEN, S. J., R. J. MCINTYRE and W. FINE, 1968 A linkage disequilibrium between two gene-enzyme systems in an experimental population of *Drosophila melanogaster*. Genetics (abstr.) **60**: 208–209.

PRAKASH, S., 1974 Gene differences between the sex ratio and standard gene arrangements of the *X* chromosome and linkage disequilibrium between loci in the standard gene arrangement of the *X* chromosome in *Drosophila pseudoobscura*. Genetics **77**: 795–804.

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**: 709–727.

SPASSKY, B., T. DOBZHANSKY and W. W. ANDERSON, 1965 Genetics of natural populations. XXXVI. Epistatic interactions of the components of the genetic load in *Drosophila pseudoobscura*. Genetics **52**: 653–664.

SPIESS, E. B., 1958 Chromosomal adaptive polymorphism in *Drosophila persimilis*. II. Effects of population cage conditions on life cycle components. Evolution **12**: 234–245. ——, 1959 Effects of recombination on viability in Drosophila. Cold Spring Harbor Symp. Quant. Biol. **23**: 329–350.

STALKER, H. D., 1960 Chromosomal polymorphism in *Drosophila paramelanica* Patterson. Genetics **45**: 95–114.

TEMIN, R. G., 1966 Homozygous viability and fertility loads in *Drosophila melanogaster*. Genetics **53**: 27–46.

TEMIN, R. G., H. U. MEYER, P. S. DAWSON and J. F. CROW, 1969 The influence of epistasis on homozygous viability depression in *Drosophila melanogaster*. Genetics **61**: 497–519.

WILSON, J., 1968 Experimental determination of fitness interactions in *Drosophila melanogaster* by the method of marginal populations. Genetics **59**: 501–511.

ZOUROS, E. and C. B. KRIMBAS, 1973 Evidence for linkage disequilibrium maintained by selection in *Drosophila subobscura*. Genetics **73**: 659–674.

Corresponding editor: R. C. LEWONTIN