EFFECTS OF LINKAGE AND EPISTASIS ON THE EQUILIBRIUM FRE- QUENCIES OF LETHAL GENES. I. LINKAGE EQUILIBRIUM

MASATOSHI NEI

Division of Genetics, National Institute of Radiological Sciences, Chiba

Received May 27, 1964

It is well known that many, if not all, natural populations of outbreeding organisms carry numerous recessive lethal genes concealed in heterozygous condition (Dobzhansky and Wright 1941, and others). A large number of studies have been directed toward determination of the mechanism or mechanisms through which such lethal genes are retained in the population. Stern et al. (1952) and others have shown in *Drosophila melanogaster* that the viability of lethal heterozygotes is two to five percent lower than that of lethal-free homozygotes on the average, pointing to a balance between mutation and selection as an important mechanism. This balance between mutation and selection was specifically emphasized by Morton et al. (1956). On the other hand, Mukai and Burdick (1960) and others have presented data which support the so-called overdominance hypothesis or the selective advantage of heterozygotes over homozygotes.

All these results are, however, based on the study of single loci and little attention has been given to the effects of linkage and epistasis. Recently, Oshima (1963) tested the viability of heterozygous individuals for single and double lethal chromosomes in *Drosophila melanogaster* and found that some of the individuals doubly heterozygous for lethal genes are more viable than those which are singly heterozygous or heterozygous for normal genes. This suggests that there is an interaction or epistasis among lethal genes.

The purpose of this series of study is to examine the effects of linkage and epistasis on the equilibrium frequencies of lethal genes in a randomly mating population. The general mathematical solution for this problem is very complicated, so that numerical solutions have been undertaken with the aid of an electronic computer except for the case where linkage equilibrium is maintained. In this first report a general property of population dynamics of lethal genes with linkage and epistasis will be described and an algebraic solution of the equilibrium gene frequencies for the special case of linkage equilibrium will be presented together with the condition necessary for linkage equilibrium to be maintained permanently. The results of numerical solutions are to be given in the second report. No account will be taken of the effect of mutation in this investigation.

PRINCIPLE

Consider two loci each with two alleles, A_1 - A_2 and B_1 - B_2 . In a randomly mating population the genotype frequencies for these two loci at fertilization can be expressed in terms of gamete frequencies as shown in Table 1. In this Table P, Q, R, and

Table 1. Frequencies and selective values for nine genotypes

		A_1A_1	A_1A_2	A_2A_2
Selective value		a	b	0
Frequency	B_1B_1	P^2	$2PR$	R^2
Selective value		c	d	0
Frequency	B_1B_2	$2PQ$	$2(PS+QR)$	$2RS$
Selective value		0	0	0
Frequency	B_2B_2	Q^2	$2QS$	S^2

S stand for the frequencies of gametes A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 , in the parental population respectively. The gene frequencies of A_1 , A_2 , B_1 , and B_2 are, therefore, expressed as $p_1=P+Q$, $q_1=R+S$, $p_2=P+R$, and $q_2=Q+S$, respectively. Alleles A_2 and B_2 are assumed to be lethal and cause individuals homo-

zygous for either or both of them to be killed. Thus, five of the nine genotypes possible for the two loci fail to survive, the remaining four genotypes having the selective values specified in Table 1.

In this situation the changes in gamete frequencies in one generation are given by

$$\Delta P = \frac{P(aP + bR + cQ + dS - \bar{W}) - rdD}{\bar{W}}$$

$$\Delta Q = \frac{Q(cP + dR - \bar{W}) + rdD}{\bar{W}}$$

$$\Delta R = \frac{R(bP + dQ - \bar{W}) + rdD}{\bar{W}}$$

$$\Delta S = \frac{S(dP - \bar{W}) - rdD}{\bar{W}}$$

Where $\bar{W}$ denotes the mean selective value of the population, D the linkage disequilibrium or $PS - QR$, and r the recombination value between the two loci. The mean selective value $\bar{W}$ is expressed as $P(aP + 2dR + 2cQ) + 2d(PS + QR)$. The expressions given above are all special cases of the general formulas developed by Kimura (1956a) and Lewontin and Kojima (1960).

Theoretically, equilibrium values of gamete or gene frequencies can be obtained by putting $\Delta P = \Delta Q = \Delta R = \Delta S = 0$ and solving for P, Q, R, and S. In practice, however, it is not easy to obtain general solutions except for the case of $D=0$ or linkage equilibrium.

EQUILIBRIUM FREQUENCIES UNDER LINKAGE EQUILIBRIUM

Under linkage equilibrium D is equal to 0, so that the condition $\Delta P = \Delta Q = \Delta R = \Delta S = 0$ gives the following set of equations

$$\begin{aligned}
aP + bR + cQ + dS - \bar{W} &= 0 \\
cP + dR - \bar{W} &= 0 \\
bP + dQ - \bar{W} &= 0 \\
dP - \bar{W} &= 0
\end{aligned}$$

Taking note of the relations $P = p_1 p_2$, $Q = p_1 q_2$, $R = q_1 p_2$, and $S = q_1 q_2$, which always hold for linkage equilibrium, the above equations can be solved for gene frequencies as well as for gamete frequencies. The equilibrium gene frequencies thus obtained are as follows:

$$\hat{q}_1 = \frac{d-c}{2d-c} \quad \hat{q}_2 = \frac{d-b}{2d-b} \quad (2)$$

The equilibrium gamete frequencies are then obtained by the relations $P = p_1 p_2$, etc. Namely,

$$\begin{aligned}
\hat{P} &= \frac{d^2}{(2d-b)(2d-c)} & \hat{Q} &= \frac{d(d-b)}{(2d-b)(2d-c)} \\
\hat{R} &= \frac{d(d-c)}{(2d-b)(2d-c)} & \hat{S} &= \frac{(d-b)(d-c)}{(2d-b)(2d-c)}
\end{aligned} \quad (3)$$

Note that the equilibrium gene or gamete frequencies are determined independently of the value of a .

The necessary condition for the equilibrium gene frequencies to be nontrivial is obtainable directly from expressions (2). Namely, d must be larger than c and b in order that $\hat{q}_1$ and $\hat{q}_2$ be nontrivial respectively. The same condition applies to the stability of gamete frequencies. The sufficient conditions for stable equilibria for two loci can be obtained by

$$\begin{aligned}
(i) \quad & \frac{\partial^2 \bar{W}}{\partial q_1^2} < 0 \quad \frac{\partial^2 \bar{W}}{\partial q_2^2} < 0 \\
(ii) \quad & \begin{vmatrix} \frac{\partial^2 \bar{W}}{\partial q_1^2} & \frac{\partial^2 \bar{W}}{\partial q_2 \partial q_1} \\ \frac{\partial^2 \bar{W}}{\partial q_2 \partial q_1} & \frac{\partial^2 \bar{W}}{\partial q_2^2} \end{vmatrix} > 0
\end{aligned}$$

as was shown by Kimura (1956 b) and Kojima (1959). In our case the above inequalities work out to be

$$\begin{aligned}
(i) \quad & 2(bc + 2d^2) - (a + 2b + 2c)d > 0 \\
(ii) \quad & [(a + 2b + 2c)d - 2(bc + 2d^2)]^2 - 4(ad - bc)^2 > 0
\end{aligned} \quad (4)$$

As will be shown in the following, in order that the population is in linkage equilibrium for all generations, $ad - bc$ must be 0, while the term $[(a + 2b + 2c)d - 2(bc + 2d^2)]^2$ in inequality (ii) is always positive if inequality (i) is satisfied. Thus, if inequality (i) is satisfied, inequality (ii) is also satisfied automatically. Inequality (i) indicates that for the equilibrium to be stable there must be overdominance,

since it is equivalent to $\bar{W}_{A_1A_1} - 2\bar{W}_{A_1A_2} + \bar{W}_{A_2A_2} < 0$ or $\bar{W}_{B_1B_1} - 2\bar{W}_{B_1B_2} + \bar{W}_{B_2B_2} < 0$, where $\bar{W}_{A_1A_1}$, etc. are the mean selective values of genotypes A_1A_1 , etc., respectively (cf. Kimura 1956 b).

CONDITION NECESSARY FOR PERMANENT LINKAGE EQUILIBRIUM

Let us next consider the condition for linkage equilibrium to be maintained permanently or $D=0$ for all generations. Suppose that in the t th generation the frequencies of A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 gametes are P_t , Q_t , R_t , and S_t , respectively, with the linkage disequilibrium of $D_t = P_tS_t - Q_tR_t$. Then, the gamete frequencies in the next generation are given by

$$\begin{aligned} P_{t+1} &= [P_t (aP_t + cQ_t + bR_t + dS_t) - rdD_t] / \bar{W}_t \\ Q_{t+1} &= [Q_t (cP_t + dR_t) + rdD_t] / \bar{W}_t \\ R_{t+1} &= [R_t (bP_t + dQ_t) + rdD_t] / \bar{W}_t \\ S_{t+1} &= [S_t (dP_t - rdD_t) / \bar{W}_t \end{aligned} \quad (5)$$

where $\bar{W}_t$ is the mean selective value of the population in the t th generation. The linkage disequilibrium in the $(t+1)$ th generation is, therefore, obtained as follows:

$$D_{t+1} = P_{t+1}S_{t+1} - Q_{t+1}R_{t+1} = [P_t W_{P_t} S_t W_{S_t} - Q_t W_{Q_t} R_t W_{R_t} - rdD_t \bar{W}_t] / \bar{W}_t^2 \quad (6)$$

where W_{P_t} , W_{Q_t} , W_{R_t} , and W_{S_t} are the mean selective values of gametes A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 respectively in the t th generation. Thus, if $D_t = 0$, we have

$$\bar{W}_t^2 D_{t+1} = P_t S_t (W_{R_t} W_{S_t} - W_{Q_t} W_{P_t}) = P_t S_t^2 (ad - bc) \quad (7)$$

The above expression shows that the permanent linkage equilibrium is attained only when ad equals bc and the original population is in linkage equilibrium. It will be shown in the second report that if the original population is not in linkage equilibrium, the resultant equilibrium population does not necessarily have $D=0$, depending on the recombination value between the two loci, even if $ad=bc$.

COMPONENTS OF THE GENETIC VARIANCE

The genetic variance of adaptive value for two loci can be partitioned into its five components, i. e., additive (σ_A^2), dominance (σ_D^2), additive $\times$ additive (σ_{AA}^2), additive $\times$ dominance (σ_{AD}^2), and dominance $\times$ dominance (σ_{DD}^2) components (cf. Cockerham, 1954). In an equilibrium population with linkage equilibrium these components are all expressible in terms of adaptive values and they become

$$\begin{aligned} \sigma_A^2 &= 0 & \sigma_{AA}^2 &= 0 & \sigma_{AD}^2 &= 0 \\ \sigma_D^2 &= \frac{d^4 [(d-b)^2 + (d-c)^2] [(a+2b+2c)d - 2(bc+2d^2)]^2}{(2d-b)^4 (2d-c)^4} \\ \sigma_{DD}^2 &= \frac{d^4 (d-b)^2 (d-c)^2 (a-2b-2c+4d)^2}{(2d-b)^4 (2d-c)^4} \end{aligned} \quad (8)$$

Thus, the genetic variance in an equilibrium population is composed only of dominance and dominance $\times$ dominance epistatic components.

Example. If $a=1$, $b=2$, $c=2$, and $d=4$, then $ad=bc$. Therefore, the population is in linkage equilibrium permanently if the initial linkage disequilibrium is 0. The equilibrium frequencies of genes A_2 and B_2 are both $1/3$ from equations (2). Inequalities (4) indicate that these equilibrium frequencies are stable since $2(bc+2d^2) - (a+2b+2c)d = 36 > 0$. The dominance variance is computed to be 1.58, whereas the dominance $\times$ dominance epistatic variance is 0.20.

DISCUSSION

In the studies of population dynamics of lethal genes linkage equilibrium is often assumed for different loci. This investigation, however, indicates that a severe condition is necessary for permanent linkage equilibrium. It is, therefore, doubtful whether such an assumption is really satisfied for many loci in natural populations. The necessary and sufficient conditions for stable equilibria of two loci that have been developed here are valid only if a permanent linkage equilibrium obtains. The effect of linkage disequilibrium on the equilibrium gene frequencies will be examined in the second report.

Kojima (1959) has shown that for two loci to be jointly held in a stable equilibrium, overdominance is required at each locus and the geometric mean of the dominance variances from each locus must be larger than the additive $\times$ additive variance for two loci. In the case of lethal genes, however, the second of these two requirements does not seem to be necessary since the additive $\times$ additive variance is always 0 at equilibrium.

SUMMARY

As the first step in a study of the effects of linkage and epistasis on equilibrium frequencies of lethal genes, the necessary and sufficient conditions for stable equilibria of two interacting recessive lethals are worked out on the assumption of linkage equilibrium. The condition necessary for linkage equilibrium to be maintained permanently is also examined. The genetic variance in an equilibrium population is shown to be composed only of the dominance and dominance $\times$ dominance components.

LITERATURE CITED

- Cockerham, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859-882.
- Dobzhansky, Th., and S. Wright, 1941 Genetics of natural populations. V. Relations between mutation rate and accumulation of lethals in populations of *Drosophila pseudoobscura*. *Genetics* **26**: 23-51.
- Kimura, M., 1956a A model of a genetic system which leads to closer linkage by natural

- selection. *Evolution* **10**: 278-287.
- Kimura, M., 1956 b Rules for testing stability of a selective polymorphism. *Proc. Nat. Acad. Sci.* **42**: 336-340.
- Kojima, K., 1959 Role of epistasis and overdominance in stability of equilibria with selection. *Proc. Nat. Acad. Sci.* **45**: 984-989.
- Lewontin, R. C., and K. Kojima, 1960 The evolutionary dynamics of complex polymorphisms. *Evolution* **14**: 458-472.
- Mukai, T., and A. B. Burdick, 1959 Single gene heterosis associated with a second chromosome recessive lethal in *Drosophila melanogaster*. *Genetics* **44**: 211-232.
- Morton, N. E., J. F. Crow, and H. J. Muller, 1956 An estimate of the mutational damage in man from data on consanguineous marriages. *Proc. Nat. Acad. Sci.* **42**: 855-863.
- Oshima, C., 1963 The persistence of some recessive lethal genes in natural populations of *Drosophila melanogaster*. III. Difference between lethal and semi-lethal genes. *Proc. Jap. Acad.* **39**: 125-130.
- Stern, C., G. Carson, M. Kinst, E. Novitski, and D. Uphoff, 1952 The viability of heterozygotes for lethals. *Genetics* **37**: 413-449.