THE EFFECT OF EPISTASIS AND GAMETIC UNBALANCE ON GENETIC LOAD¹

P. E. HANSCHKE, S. K. JAIN AND R. W. ALLARD

Departments of Pomology, Agronomy and Genetics, University of California, Davis

Received June 16, 1966

GENETIC load has been defined as the extent to which an equilibrium population is impaired because not all individuals are of optimum genotype. The concept, first discussed by HALDANE (1937), is the basis of theory developed by CROW (1958) and others for assessing the relative impact of various facets of the genetic system (especially mutation, segregation and inbreeding) on the average fitness of populations in genetic equilibrium. LEVENE (1963) recently discussed various theoretical and experimental difficulties in utilizing the theory of genetic loads in the analysis of the genetic structure of populations. In its simplest form, this theory involves several assumptions among which those of no epistasis on the fitness scale and linkage equilibrium are particularly restrictive. The purpose of this study is to investigate the effects of epistasis and gametic unbalance (linkage disequilibrium) on the manifestation of genetic loads when inbreeding is imposed on random-mating populations. Since the measures of joint inbreeding at two or more loci are influenced by gametic unbalance, the expressions developed by HALDANE (1949) for joint inbreeding are first generalized to take into account gametic and zygotic associations. The loads maintained at interacting loci are then evaluated in terms of genetic models involving several kinds of epistasis.

INBREEDING LOAD IN POPULATIONS WITHOUT EPISTASIS

Genetic load has been explicitly defined, for equilibrium populations, as the fitness of the optimal genotype (by convention taken to be unity) minus the average population fitness (CROW 1958). Thus, if genotypic frequencies are denoted by $p_1, p_2, p_3, \dots, p_n$ ($\sum_{i=1}^n p_i = 1$) and the associated fitness values, relative to the optimal value of 1, are denoted by $w_1, w_2, w_3, \dots, w_n$, genetic load is given by (MORTON, CROW, and MULLER 1956)

$$L = 1 - \sum_{i=1}^n p_i w_i = 1 - \bar{W},$$

when there is no epistasis.

Inbreeding a population at equilibrium under random mating ordinarily increases genetic load. Assuming no interallelic interactions, MORTON *et al.* explicitly defined inbreeding load as

$$L_i = L_r + (L_l - L_r)f,$$

¹ This work was supported in part by Public Health Service grant GM-10476.

where f is WRIGHT's inbreeding coefficient (in the sense of his fixation index [WRIGHT 1965]), L_r is the random mating load and L_i is the expected load when $f = 1$. The regression of load on inbreeding in such systems is linear as given by the above expression.

INBREEDING LOAD IN POPULATIONS INVOLVING EPISTASIS AND GAMETIC BALANCE

Consider an infinite population and a genetic model involving two alleles at each of the loci A and B . Let the fitnesses of individual genotypes be given by w_{ij} (Table 1). If interallelic interactions are taken into account, but the restriction imposed that there is "linkage equilibrium" ($D = g_{11}g_{33} - g_{13}g_{31} = 0$), the expression for mean population fitness in random mating populations takes the form (COCKERHAM 1954)

$$W_r = g_{11}\bar{w}_{AB} + g_{13}\bar{w}_{Ab} + g_{31}\bar{w}_{aB} + g_{33}\bar{w}_{ab},$$

where g_{11} , g_{13} , g_{31} and g_{33} are the gametic frequencies and $\bar{w}_{AB}$, $\bar{w}_{Ab}$, $\bar{w}_{aB}$ and $\bar{w}_{ab}$ are the average gametic fitnesses (computed from relative zygotic fitnesses listed in Table 1). Genetic load for the random mating population thus becomes

$$L_r = 1 - W_r = 1 - (g_{11}\bar{w}_{AB} + g_{13}\bar{w}_{Ab} + g_{31}\bar{w}_{aB} + g_{33}\bar{w}_{ab}).$$

With inbreeding the average fitness for the two locus case can be expressed as

$$W_i = W_r - f [W_r + g_{11}g_{33}(e_{22} + e_{11}) - g_{13}g_{31}(e_{12} + e_{21})]$$

TABLE 1

Gametic and zygotic frequencies and the fitness matrix for the two locus case

Female gametes	Male gametes	AB	Ab	aB	ab	Marginal means
AB	g_{11}	g_{11}	g_{13}	g_{31}	g_{33}	$\bar{w}_{AB}$
		w_{22}	w_{21}	w_{12}	w_{11}	
		(g_{11}^2)	$(g_{11}g_{13})$	$(g_{11}g_{31})$	$(g_{11}g_{33})$	
Ab	g_{13}	w_{21}	w_{20}	w_{11}	w_{10}	$\bar{w}_{Ab}$
		$(g_{11}g_{13})$	(g_{13}^2)	$(g_{13}g_{31})$	$(g_{13}g_{33})$	
		w_{12}	w_{11}	w_{02}	w_{01}	
aB	g_{31}	$(g_{11}g_{31})$	$(g_{13}g_{31})$	(g_{31}^2)	$(g_{31}g_{33})$	$\bar{w}_{aB}$
		w_{11}	w_{10}	w_{01}	w_{00}	
		$(g_{11}g_{33})$	$(g_{13}g_{33})$	$(g_{31}g_{33})$	(g_{33}^2)	
ab	g_{33}					$\bar{w}_{ab}$
Marginal means		$\bar{w}_{AB}$	$\bar{w}_{Ab}$	$\bar{w}_{aB}$	$\bar{w}_{ab}$	

Following COCKERHAM (1954), the epistatic parameters are:

$$\begin{aligned} e_{22} &= w_{22} - w_{12} - w_{21} + w_{11} \\ e_{21} &= w_{21} - w_{20} - w_{11} + w_{10} \\ e_{12} &= w_{12} - w_{11} - w_{02} + w_{01} \\ e_{11} &= w_{11} - w_{10} - w_{01} + w_{00} \end{aligned}$$

$-(g_{11}w_{22} + g_{13}w_{20} + g_{31}w_{02} + g_{33}w_{00})] + F[g_{11}g_{33}(e_{22} + e_{11}) - g_{13}g_{31}(e_{12} + e_{31})]$, where F is HALDANE's probability of joint identity by descent for two loci (HALDANE 1949) and the e_{ij} are as defined in Table 1. By collecting terms and substituting L_i for $1 - W_i$, L_r for $1 - W_r$, etc., this expression can be put in the form

$$L_i = L_r + (L_l - L_r)f + (f - F)\gamma,$$

where

$$\gamma = g_{13}g_{31}(e_{22} - e_{12} - e_{21} + e_{11}).$$

The effect of epistasis on the regression of genetic load on inbreeding clearly depends upon the probability of recombination since the magnitude of F and of the gametic product $g_{13}g_{31}$ are both dependent upon the probability of crossing over between loci. Furthermore the effect of epistasis depends on the sign of γ since $(f - F) \geq 0$. When γ is negative epistasis will decrease inbreeding load, producing a positive (upward) curvilinearity in the regression of load on inbreeding. When γ is positive epistasis will increase inbreeding load, producing a negative (downward) curvilinearity in the regression of load on inbreeding.

DERIVATION OF INBREEDING LOAD FOR POPULATIONS INVOLVING EPISTASIS AND GAMETIC UNBALANCE

To obtain the relationship between inbreeding and load in populations in which there is both epistasis and gametic unbalance it is first necessary to determine how gametic frequencies in the random mating population are related to zygotic frequencies after one generation of inbreeding. Then the appropriate load expression can be derived easily by multiplying zygotic frequencies by their fitness values.

The relationship between gametic frequencies in random mating populations and zygotic frequencies after inbreeding:

Assumptions and notation: (A) an infinitely large diploid population; (B) the probability of recombination between loci A/a and B/b is c and the probability of no recombination is $(1 - c) = c'$; (C) gametic frequencies are as in Table 1 ($D = g_{11}g_{33} - g_{13}g_{31}$); and (D) zygotic frequencies, after random assortment of gametes with selection relaxed, are as given in Table 1.

The probability that a randomly chosen zygote, x , from the population of Table 1 will produce a gamete, g_1 , which is AB is $g_{11} - cD$. Given that the gamete g_1 from x is AB , then the zygote y , related to x by f and F , may produce an AB gamete, g_2 , as a result of any one of the following four events:

1. A and B are derived from one or two common ancestors;
2. A is derived from a common ancestor and B from another source;
3. B is derived from a common ancestor and A from another source; or
4. both A and B are derived from another source.

The sum of the probabilities of these four events gives the expected frequency of the zygote $AABB$, $P(AABB)$, after one generation of inbreeding. The derivation of probabilities of each of the above events follows.

1. *The probability of the first event, that the allelic pair AB is derived from*

one or two common ancestors: The probability that the allelic pair of gamete g_2 is identical by descent to that of g_1 , given g_1 is AB , was shown by HALDANE (1949) to be equal to F_{xy} . However, for the two locus case, gametic frequencies in random mating populations are, after inbreeding, not generally related to the zygotic frequencies by the function, F_{xy} , but they are related by the functions F_{xy}^{AB} , F_{xy}^{Ab} , F_{xy}^{aB} and F_{xy}^{ab} (the probabilities of joint identity by descent for allelic pairs, with sum equal to F_{xy}). HALDANE's relationship holds only for the special case where g_1 and g_2 are statistically independent (i.e., where $D = 0$). This can be seen in the expression

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d) / P(g_1 \text{ is } AB) = g_{11} F_{xy} / g_{11} = F_{xy},$$

where the symbol $\cap$ indicates joint probability and the subscript d indicates identity by descent. The expression holds, as noted above, only where the probabilities $P(g_1 \text{ is } AB)$ and $P(g_2 \text{ is } A_d B_d)$ are statistically independent.

A derivation of the more generally useful functions F_{xy}^{AB} , F_{xy}^{Ab} , F_{xy}^{aB} and F_{xy}^{ab} follows. The case of inbreeding by complete selfing (i.e., the case where F_{xy}^{AB} is F_{xx}^{AB} since $x \equiv y$) will be considered first:

$$\begin{aligned} F_{xy}^{AB} &\equiv F_{xx}^{AB} = P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d) \\ &= P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d \mid x \text{ is } AB/AB) P(x \text{ is } AB/AB) \\ &+ P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d \mid x \text{ is } AB/Ab) P(x \text{ is } AB/Ab) \\ &+ P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d \mid x \text{ is } AB/aB) P(x \text{ is } AB/aB) \\ &+ P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d \mid x \text{ is } AB/ab) P(x \text{ is } AB/ab) \\ &+ P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d \mid x \text{ is } Ab/aB) P(x \text{ is } Ab/aB) \\ &= g_{11}^2 [(\frac{1}{2})c'^2 + (\frac{1}{2})c^2] + g_{11}g_{13} [(\frac{1}{2})c'^2 + (\frac{1}{2})c^2] + \\ &\quad g_{11}g_{31} [(\frac{1}{2})c'^2 + (\frac{1}{2})c^2] + (\frac{1}{2})c'^2 g_{11}g_{33} + (\frac{1}{2})c^2 g_{13}g_{31}. \end{aligned}$$

Substituting $g_{11}g_{33} - D$ for $g_{13}g_{31}$ in this expression we obtain, by collecting terms and simplifying,

$$F_{xy}^{AB} = P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d) = g_{11} [(\frac{1}{2}) - c'c] - (\frac{1}{2})c^2 D.$$

Consequently the probability of event (1), that A and B are derived from one or two common ancestors, given g_1 is AB , is $P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_d) / P(g_1 \text{ is } AB) = \{ g_{11} [(\frac{1}{2}) - c'c] - (\frac{1}{2})c^2 D \} / (g_{11} - cD)$. (1)

The probabilities of joint identity by descent for the other three possible allelic pairs, F_{xy}^{Ab} , F_{xy}^{aB} and F_{xy}^{ab} can be derived following the same probability arguments. They are:

$$\begin{aligned} F_{xy}^{Ab} &= g_{13} [(\frac{1}{2}) - c'c] + (\frac{1}{2})c^2 D \\ F_{xy}^{aB} &= g_{31} [(\frac{1}{2}) - c'c] + (\frac{1}{2})c^2 D \\ F_{xy}^{ab} &= g_{33} [(\frac{1}{2}) - c'c] - (\frac{1}{2})c^2 D \end{aligned}$$

The summation of these four probabilities is identical to the joint inbreeding function for the selfing case, $F_{xx} = (\frac{1}{2}) - c'c$.

For sib mating the probability of joint identity by descent for the allelic pair ij , F_{xy}^{ij} , (where x and y are sibs) is

$$F_{xy}^{ij} = (\frac{1}{8})g_{ij}(2c'^4 + 2c'^2 c^2 + c^2) \pm (\frac{1}{2})c^2 D [(\frac{1}{4}) + (\frac{1}{2})c'^2].$$

The derivation of this expression (see Appendix I) takes a different form from that of the selfing case, which involves uniparental reproduction.

Again $F_{xy} = F_{xy}^{AB} + F_{xy}^{Ab} + F_{xy}^{aB} + F_{xy}^{ab}$, the inbreeding function of HALDANE

(1949) and SCHNELL (1961). It appears that F_{xj}^{ij} for any inbreeding system can be obtained by multiplying the appropriate inbreeding function, F_{xy} , by g_{ij} and adding (or subtracting) the term $K \cdot D(Z - e^2)$. In this term, K is the coefficient of the polynomial appropriate for F_{xy} and Z is the probability of joint identity by descent for the allele pair ij through one parent, when that parent is the double heterozygote. We have not, however, explored this relationship for mating systems other than selfing and sib mating.

2. *The probability of the second event, that A is derived from a common ancestor and B is from another source, given that g_1 is AB can be stated explicitly as*

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_s) / P(g_1 \text{ is } AB).$$

The subscript s indicates identity in state to the homologue of g_1 , but not identity by descent.

$$\begin{aligned} & P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_s) \\ &= P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_s \mid x \text{ is } AB/AB) P(x \text{ is } AB/AB) \\ &+ P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_s \mid x \text{ is } AB/aB) P(x \text{ is } AB/aB) \\ &= [(1/2)c'c + (1/2)c'c] g_{11}^2 + [(1/2)c'c + (1/2)c'c] g_{11}g_{13} = c'cg_{11}(g_{11} + g_{13}). \end{aligned}$$

Thus,

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_d B_s) / P(g_1 \text{ is } AB) = c'cg_{11}(g_{11} + g_{13}) / (g_{11} - cD). \quad (2)$$

3. *The probability of the third event, that B is derived from a common ancestor and A is from another source can be seen, owing to symmetry, to be*

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_s B_d) / P(g_1 \text{ is } AB) = c'cg_{11}(g_{11} + g_{31}) / (g_{11} - cD) \quad (3)$$

4. *The probability of the fourth event, that both A and B are from another source can be stated explicitly as*

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_s B_s) / P(g_1 \text{ is } AB).$$

$$\begin{aligned} & P(g_1 \text{ is } AB \cap g_2 \text{ is } A_s B_s) = P(g_1 \text{ is } AB \cap g_2 \text{ is } A_s B_s \mid x \text{ is } AB/AB) P(x \text{ is } AB/AB) \\ &= [(1/2)c'^2 + (1/2)c^2] g_{11}^2. \end{aligned}$$

Consequently the probability that both A and B are from another source, given g_1 is AB, is

$$P(g_1 \text{ is } AB \cap g_2 \text{ is } A_s B_s) / P(g_1 \text{ is } AB) = g_{11}^2 [(1/2)c'^2 + (1/2)c^2] / (g_{11} - cD). \quad (4)$$

Thus, the total probability that g_1 is AB and g_2 is AB, after one generation of complete selfing, is obtained by summing expressions (1)–(4) and multiplying by $g_{11} - cD$, i.e., $P(AABB) = P(g_1 \text{ is } AB \cap g_2 \text{ is } AB) = P(g_2 \text{ is } AB \mid g_1 \text{ is } AB) P(g_1 \text{ is } AB) = [(1) + (2) + (3) + (4)] [g_{11} - cD] = \{[g_{11}(1/2) - c'c] - (1/2)c^2D\} + [c'cg_{11}(g_{11} + g_{13})] + [c'cg_{11}(g_{11} + g_{31})] + [(1/2) - c'c] g_{11}^2. \quad (5)$

$P(g_1 \text{ is } AB) = [(1) + (2) + (3) + (4)] [g_{11} - cD] = [g_{11}(1/2) - c'c] - (1/2)c^2D$
The relationship between this general expression for $P(AABB)$ and the special case (where $D = 0$) derived by HALDANE can be clarified as follows: Substitute F for $[(1/2) - c'c]$, $(f - F)$ for $c'c$ (since $f = 1/2$) and $(1 - 2f + F)g_{11}^2$ for $[(1/2) - c'c] g_{11}^2$. Expression (5) thus becomes $P(AABB) = (1 - 2f + F)g_{11}^2 + (f - F)g_{11}(g_{11} - g_{33}) + fg_{11} - (1/2)c^2D$, where F is HALDANE's probability of joint identity by descent (F_{xy}) for complete selfing.

The expected frequencies of the other nine genotypes after one generation of selfing are obtained by similar probability arguments. These frequencies, in terms of the coefficient of inbreeding, f , and the probability of joint identity by descent,

TABLE 2

Probabilities with which the 4 classes of gametes from zygote x unite with the 4 classes of gametes from zygote y, if x and y are related by f and F, and the mating system is selfing. $\phi = 1 - 2f + F$

$P(AABB) = \phi g_{11}^2 + (f-F) g_{11} (g_{11} - g_{33}) + fg_{11} - (\frac{1}{2}) c^2 D$
$P(AABb) = 2 \phi g_{11} g_{13} + (f-F) (2g_{11} g_{13} + g_{11} g_{33} + g_{13} g_{31})$
$P(AAbb) = \phi g_{13}^2 + (f-F) g_{13} (g_{13} - g_{31}) + fg_{13} + (\frac{1}{2}) c^2 D$
$P(AaBB) = 2 \phi g_{11} g_{31} + (f-F) (2g_{11} g_{31} + g_{11} g_{33} + g_{13} g_{31})$
$P(AB/ab) = 2 \phi g_{11} g_{33}$
$P(Ab/aB) = 2 \phi g_{13} g_{31}$
$P(Aabb) = 2 \phi g_{13} g_{33} + (f-F) (2g_{13} g_{33} + g_{11} g_{33} + g_{13} g_{31})$
$P(aaBB) = \phi g_{31}^2 + (f-F) g_{31} (g_{31} - g_{13}) + fg_{31} + (\frac{1}{2}) c^2 D$
$P(aaBb) = 2 \phi g_{31} g_{33} + (f-F) (2g_{31} g_{33} + g_{11} g_{33} + g_{13} g_{31})$
$P(aabb) = \phi g_{33}^2 + (f-F) g_{33} (g_{33} - g_{11}) + fg_{33} - (\frac{1}{2}) c^2 D$

F_{xy} , are listed in Table 2. These probabilities replace those derived by NEI (1965) who assumed relaxation of selection on zygotes during one round of random mating but did not account for the subsequent reduction in gametic unbalance in his derivation of genotypic frequencies.

Derivation of genetic load in populations maintaining gametic unbalance: The notation of COCKERHAM (1954) can be followed in assigning fitness values to each genotype (see Table 1). The average fitness of the population is obtained by multiplying genotypic frequencies by appropriate fitness values and taking the sum of the products. Thus, after one generation of selfing,

$$W_i = W_r - f [W_r + g_{11}g_{33}(e_{22}+e_{11}) - g_{13}g_{31}(e_{21}+e_{12}) - W_i] \\ + F [g_{11}g_{33}(e_{22}+e_{11}) - g_{13}g_{31}(e_{12}+e_{21})] - (\frac{1}{2})c^2D(w_{22}-w_{02}-w_{20}+w_{00}).$$

It follows that genetic load, which equals $1 - W_i$, is now

$$L_i = (1-W_r) + f [(1-W_i) - (1-W_r)] + (f-F) [\gamma + D(e_{22} + e_{11})] \\ + (\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}),$$

where

$$\gamma = g_{13}g_{31} (e_{22} - e_{12} - e_{21} + e_{11}).$$

Thus, for populations inbred by selfing, when there is epistasis and $D \neq 0$, genetic load for the 2-locus model will be:

$$L_i = L_r + (L_i - L_r)f + (f-F) [\gamma + D(e_{22} + e_{11})] + (\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}) \quad (6)$$

Note that the final term of expression (6), which will be designated ω , measures the reduction in gametic unbalance resulting from one generation of relaxed selection. The effect of relaxing selection depends on the kind of inbreeding but not on the amount; e.g. for the sib mating case ω is $(\frac{1}{2})c^2D[(\frac{1}{4}) + (\frac{1}{2})c'^2] \times (w_{22} - w_{02} - w_{20} + w_{00})$.

The effect of gametic unbalance on inbreeding load: The effect of gametic unbalance on inbreeding load can be visualized most easily by considering two kinds of epistatic systems as follows. For epistatic models in which homozygotes are of equal fitness the load expression reduces to

$$L_i = L_r + (L_i - L_r)f + (f-F) [\gamma + D(e_{11} + e_{22})],$$

since the final term, $(\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}) = 0$. Gametic unbalance in such systems produces upward curvilinearity in the regression of load on inbreeding when D and $(e_{11} + e_{22})$ are of opposite sign and downward curvilinearity when they are of the same sign.

When $\omega \neq 0$ the load expression becomes

$$L_i = L_r + (L_i - L_r)f + (f - F)[\gamma + D(e_{11} + e_{22})] + (\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}).$$

As mentioned above, the quantity ω is independent of f and F . Its effect is to increase inbreeding load by a constant amount, i.e., when load is regressed on f the effect is to displace the regression line upward along the ordinate. Neither the slope nor the curvilinearity of the regression is affected. However, since the magnitude of ω depends on the mating system, a spurious upward curvilinearity is imposed on the regression of load on inbreeding if load is regressed on the values of f obtained from more than one mating system. Note also that this term depends on the number of generations of inbreeding since, with selection relaxed, the amount of gametic unbalance diminishes each generation. Consequently if the f 's are obtained from more than one generation of inbreeding, there will also be spurious upward curvilinearity should load be regressed on inbreeding.

Some specific cases: We have chosen six epistatic models (Table 3) to illustrate the effect of epistasis and gametic unbalance on inbreeding load. These models were chosen primarily on the basis that they have all been previously proposed in explanation of experimental observations (see JAIN and ALLARD 1966). The pertinent features of these models can be summarized as follows:

TABLE 3

Some specific examples of 2-locus epistasis and the values of e_{ij}

(1) Heterotic model	(2) Cumulative heterotic model
.6 .8 .6	.4 .7 .3
.8 1.0 .8	.7 1.0 .6
.6 .8 .6	.5 .8 .4
$(e_{22}=e_{21}=e_{12}=e_{11}=0)$	$(e_{22}=e_{21}=e_{12}=e_{11}=0)$
(3) Optimum model	(4) Mixed optimum, heterotic model
.6 .9 1.0	.50 .80 .75
.9 1.0 .9	.80 1.0 .80
1.0 .9 .6	.75 .80 .50
$(e_{22}=e_{21}=e_{12}=e_{11}=-.2)$	$(e_{22}=e_{11}=-.10, e_{21}=e_{12}=-.15)$
(5) Diminutive heterotic model	(6) Mixed over- and underdominance
.5 .9 .5	.9 .5 .9
.9 1.0 .9	.5 1.0 .5
.5 .9 .5	.9 .5 .9
$(e_{22}=e_{11}=-.3, e_{21}=e_{12}=.3)$	$(e_{22}=e_{11}=.9, e_{21}=e_{12}=-.9)$
(7) Mixed over- and underdominance model	
.8 .5 .8	
.6 1.0 .5	
.7 .6 .6	
$(e_{22}=.7, e_{21}=-.5, e_{12}=-.8, e_{11}=.5)$	

- (1) A heterotic model (nonepistatic): $\gamma = 0$, $D = 0$, $e_{11} + e_{22} = 0$ and $w_{22} - w_{02} - w_{20} + w_{00} = 0$. Inbreeding load for this model can be expressed as

$$L_i = L_r + (L_l - L_r)f.$$
- (2) A cumulative heterotic model: $\gamma = 0$, $D \geq$ or ≤ 0 as a function of c and initial zygotic frequencies, $e_{11} + e_{22} = 0$ and $w_{22} - w_{02} - w_{20} + w_{00} = 0$. Inbreeding load for this model can be expressed as

$$L_i = L_r + (L_l - L_r)f.$$
- (3) An optimum model: $\gamma = 0$, $D < 0$, $e_{22} + e_{11} < 0$ and $w_{22} - w_{02} - w_{20} + w_{00} < 0$. Inbreeding load for this model for the selfing case is given by

$$L_i = L_r + (L_l - L_r)f + (f - F)D(e_{22} + e_{11}) + (\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}).$$
- (4) A mixed optimum heterotic model: $\gamma > 0$, $D \geq$ or ≤ 0 as a function of c and initial zygotic frequencies, $e_{22} + e_{11} = 0$, and $w_{22} - w_{02} - w_{20} + w_{00} < 0$. Inbreeding load for this model can be expressed (for the selfing case) as

$$L_i = L_r + (L_l - L_r)f + (f - F)\gamma + (\frac{1}{2})c^2D(w_{22} - w_{02} - w_{20} + w_{00}).$$
- (5) A diminutive heterotic model: $\gamma < 0$, $D \geq$ or ≤ 0 as a function of c and initial zygotic frequencies, $e_{22} + e_{11} < 0$, and $w_{22} - w_{02} - w_{20} + w_{00} < 0$. Inbreeding load can be expressed as

$$L_i = L_r + (L_l - L_r)f + (f - F)[\gamma + D(e_{22} + e_{11})].$$
- (6-7) Models of mixed over-and under dominance: $\gamma > 0$, $D \geq$ or ≤ 0 as a function of c and initial zygotic frequencies, $e_{22} + e_{11} > 0$ and $w_{22} - w_{02} - w_{20} + w_{00} = 0$. Inbreeding load for this model is

$$L_i = L_r + (L_l - L_r)f + (f - F)[\gamma + D(e_{22} + e_{11})],$$
 when linkage is sufficient to maintain gametic unbalance; or

$$L_i = L_r + (L_l - L_r)f + (f - F)\gamma,$$
 when linkage is insufficient to maintain gametic unbalance.

The first step in investigating genetic load generated by the above epistatic models was to establish equilibrium gametic and zygotic frequencies. This was done by computer simulation following procedures discussed by JAIN and ALLARD (1966). The parameters D , f and L_r were then determined for each of the six models (Table 4). Two breeding schemes were followed to establish inbreeding load. With scheme A the sibship families were established directly from the equilibrium random mating populations. With scheme B there was one generation of random mating with selection relaxed (all $w_{ij} = 1$) to eliminate any correlation among uniting gametes (with no selection, fixation index also equals zero) before establishing the sibship families. Note that scheme A does not result in fixation index, f , equal to $\frac{1}{8}$ and $\frac{1}{4}$, respectively, for half-sib and full-sib matings since, at equilibrium, the value of f is not zero for the cases studied (Table 4). In order to obtain values of f higher than those produced by a single generation of half- or full-sib matings, a mixed selfing and random mating series with the proportion of selfing, $s = 0, .25, .50, .75$ and 1.0 (giving $f = (\frac{1}{2})s$ under scheme B in the next generation) was also investigated.

KEMPTHORNE (1957) has shown that f and F_{xy} appropriate to such mixed mating systems are simple linear functions of s . Loads for breeding scheme B were computed from expression (6) and checked by computer simulation. However, loads for breeding scheme A were obtained by simulation alone because the

TABLE 4

Some parameters of populations at equilibrium under random mating, and under various kinds and levels of inbreeding

Model	c	D	$\hat{f}$	L_r ($f=0$)	Sibling system†	Half sib	Full sib	Percent selfing (s)			Isogen. ($f=1$)	Regression coefficients	
								.25	.50	1.0		β_1	β_2
(1) Heterotic (no epistasis)	.01, .50	0	—, .1250	.2000	(a)	.2219	.2438	.2250	.2500	.3000	.4000	.2000	0
(2) Optimum	.01	—, .2262	—, .0005	.0095	(b)	.2250	.2500						
					(a)	.0114	.0125						
(3) Mixed optimum- heterotic	.50	—, .0248	—, .0446	.0901	(b)	.0120	.0136	.0109	.0123	.0151	.0190	.0141	—, .0039
					(a)	.1045	.1140						
					(b)	.1075	.1200	.1026	.1151	.1401	.1802	.1132	—, .0207
					(a)	.1494	.1638						
(4) Heterotic (diminutive)	.50	—, .0158	—, .1113	.1898	(b)	.1518	.1685	.1510	.1676	.2008	.2659	.1348	—, .0029
					(a)	.2116	.2312						
					(b)	.2152	.2383	.2129	.2359	.2820	.3671	.1925	—, .0140
					(a)	.2105	.2443						
(5) Heterotic (cumulative)	.50	.0151	—, .1475	.1724	(b)	.2181	.2592	.2161	.2556	.3345	.4924	.3162	—, .0002
					(a)	.1996	.2280						
					(b)	.2050	.2393	.2078	.2433	.3142	.4917	.2528	.0641
					(a)	.3201	.3488						
(6) Mixed under-, overdominance	.01	—, .2444	—, .0520	.0600	(b)	.3279	.3643	.3279	.3643	.4371	.5829	.2913	0
					(a)	.0695	.0748						
					(b)	.0714	.0785	.0661	.0722	.0843	.1000	.0575	—, .0164
					(a)	.2766	.2718						
(7) Mixed under-, overdominance	.01	—, .2329*	—, .1333	.1296	(b)	.2777	.2734	.2672	.2594	.2437	.1000	.0370	—, .2037
					(a)	.1449	.1571						
					(b)	.1480	.1634	.1442	.1589	.1882	.2407	.1235	—, .0116
					(a)	.2983	.3049						
	.50	.0046	—, .0787	.2879	(b)	.3001	.3076	.2933	.2987	.3096	.2610	.1057	—, .1273

* One of the two stable equilibria reached from initially all Ab/aB .† (a) and (b) refer to the two schemes (a) without and (b) with one cycle of random mating prior to inbreeding (see text). Under scheme (a), half-sib and full-sib matings give $f_{hs} = f + (1/6)(1-f)$ and $f_{fs} = f + (1/4)(1-f)$ respectively. Scheme (b) gives $f_{hs} = 1/6$, $f_{fs} = 1/4$, and under mixed mating $f_m = (1/2)s$.

theoretical load expressions require that $f = 0$ prior to the initiation of inbreeding, or that genotypic frequencies in the equilibrium population be known.

The results illustrate three main features of the effect of inbreeding on genetic loads in populations where epistasis maintains gametic unbalance. First, it can be seen from Table 4 that, although loads are generally greater under inbreeding than under random mating, load ratios (L_r/L_i) greater than unity are possible. For example, with breeding scheme B and loose linkage the load ratio is greater than unity for the mixed under- and overdominance models. Consequently it is expected that inbreeding should improve mean population fitness with certain combinations of selective values when underdominance is one of the features of the genetic model, or whenever γ is large and positive.

Second, the results suggest that the curvilinearity imposed upon the regression of load on inbreeding by epistasis is, in most instances, probably insufficient to be detected experimentally by regression analysis. In the equation $L_i = L_r + \beta_1 f + \beta_2 f^2$, β_1 and β_2 are, respectively, the linear and quadratic regression coefficients. The nonepistatic models, as expected on the basis of $e_{ij} = 0$ give $\beta_2 = 0$. For the epistatic models investigated, β_1 is generally large. The value of β_2 is often small except for models of mixed under- overdominance in case of $c = .50$ but at lower levels of inbreeding this nonlinearity is not likely to be detected experimentally. Thus, while certain models can be developed that involve significant nonlinear

FIGURES 1 and 2.—The regression of genetic load (L_i) on inbreeding (f). Numbers on right margin refer to genetic models given in the text. Recombination value, $c = .01$ (Figure 1); $c = .50$ (Figure 2). Inbreeding was imposed by varying proportions of selfing following one generation of random mating (Scheme B). Note the linear and quadratic components of regression as also given by the values of β_1 , β_2 (Table 4).

regression (e.g. KING 1966), for the kinds of models studied, load tends to increase approximately linearly with the level of inbreeding. This relationship is depicted graphically in Figures 1 and 2.

It is apparent from these figures that the regression coefficient, β_1 , is strongly correlated with the degree of heterosis and that the primary measure it provides is of the level of heterozygote advantage. This merely confirms what has been known for a long time, that inbreeding is a powerful technique for detecting heterosis.

Third, in systems where epistasis and gametic unbalance are important, mating systems which generate equivalent values of the fixation index (f), may not produce equivalent genetic load. This results from the fact that mating systems which generate an equivalent f generally do not generate an equivalent F_{xy} . This can result in different amounts of load generated by sib mating and the mixed selfing and random mating system with $s = .50$ (Table 4), both of which give $f = .25$. The relationship between f and F varies from one mating system to another because of the zygotic associations of the type discussed by HALDANE (1949) and BENNETT and BINET (1956). Obviously this effect must be taken into account should load comparisons be made between different inbreeding systems.

Finally, it seems worthwhile to emphasize that for load analyses in experimental populations the realized values of f will always be less than expected from various systems of inbreeding when epistasis and/or overdominance maintain an excess of heterozygotes unless one round of random mating without selection can be performed prior to inbreeding (see Table 3). The effect of this bias is to produce a *spurious* upward curvilinearity in the regression of load on inbreeding similar to that which can be seen in the data of DOBZHANSKY, SPASSKY and TIDWELL (1963), MALOGOLOWKIN-COHEN *et al.* (1964) and SPASSKY *et al.* (1965) in *Drosophila pseudoobscura* and *D. willistoni*. One might imply from this sort of data that in these species there exist epistatic systems in which γ (or an analogous parameter for n -locus epistatic systems) is negative and large, and D and $(e_{22} + e_{11})$ are of opposite sign. This may well be the case. However, such curvilinearity could easily be an anomaly resulting from regressing load on inappropriate values of f owing to overdominance in the base population; it could also result from the effect of differences in F_{xy} associated with the mating systems utilized to obtain different degrees of inbreeding.

The authors wish to thank Drs. J. F. CROW and P. L. WORKMAN for reading the manuscript critically, and the Computer Center at the University of California, Davis, for the use of its facilities.

SUMMARY

Genetic load was examined in populations in which epistasis maintains gametic unbalance. Both analytical and Monte Carlo simulation methods were utilized.—An explicit genetic load expression was obtained for 2-locus cases by utilizing the concept of joint identity by descent for allelic pairs, F_{xy}^{ij} . This expression shows

that the regression of load on inbreeding can be curvilinear upward or downward depending on the type of epistasis. The extent of curvilinearity varies with selective values and decreases with tighter linkage. In general, the curvilinearity is negligible, particularly in the lower range of values of WRIGHT's inbreeding coefficient, f . This result suggests that in practice analysis of regression of load on inbreeding is a poor discriminator of epistasis.—The probability of identity by descent for allelic pairs, F_{xy}^{ij} , depends on the type as well as the amount of inbreeding. The regression of load on inbreeding in epistatic systems consequently also depends on the mating system and when this is not taken into account, a *spurious* upward curvilinearity may be imposed on the regression of load on inbreeding. A spurious upward curvilinearity can also be imposed on the regression in populations where the maintenance of heterozygotes results in negative values of the fixation index (i.e., where $f < 0$). Technical problems associated with regressing genetic load on f reinforce the conclusion that regression analysis in systems similar to those studied would often fail to detect the presence of the epistatic component of fitness (or load).

LITERATURE CITED

- BENNETT, J. H., and F. E. BINET, 1956 Association between Mendelian factors with mixed selfing and random mating. *Heredity* **10**: 51–55.
- 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.
- CROW, J. F., 1958 Some possibilities for measuring selection intensities in man. *Human Biol.* **30**: 1–13.
- DOBZHANSKY, TH., B. SPASSKY, and T. TIDWELL, 1963 Genetics of natural populations. XXXII. Inbreeding and the mutational and balanced genetic loads in natural populations of *Drosophila pseudoobscura*. *Genetics* **48**: 361–373.
- HALDANE, J. B. S., 1937 The effect of variation on fitness. *Am. Naturalist* **71**: 337–349. ——— 1949 The association of characters as a result of inbreeding and linkage. *Ann. Eugenics.* **15**: 15–23.
- JAIN, S. K., and R. W. ALLARD, 1966 Effects of linkage and epistasis on population changes under inbreeding. *Genetics* **53**: 633–659.
- KEMPTHORNE, O., 1957 *An Introduction to Genetic Statistics*. Wiley, New York.
- KING, J. L., 1966 The gene interaction component of the genetic load. *Genetics* **53**: 403–413.
- LEVENE, H., 1963 Inbred genetic loads and the determination of population structure. *Proc. Natl. Acad. Sci. U.S.* **50**: 587–592.
- MALOGOLOWKIN-COHEN, CH., H. LEVENE, N. P. DOBZHANSKY, and A. S. SIMMONS, 1964 Inbreeding and the mutational and balanced loads in natural populations of *Drosophila willistoni*. *Genetics* **50**: 1299–1311.
- 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. Natl. Acad. Sci. U.S.* **42**: 855–863.
- NEI, M., 1965 Effect of linkage on the genetic load manifested under inbreeding. *Genetics* **51**: 679–688.

- SCHNELL, F. W., 1961 Some general formulations of linkage effects in inbreeding. *Genetics* **46**: 947-957.
- SPASSKY, B., TH. 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**: 653-654.
- WRIGHT, S., 1965 The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**: 395-420.

APPENDIX I: Derivation of F_{xy}^{AB} when x and y are full sibs.

The probability of identity by descent between two gametes g_x and g_y derived from the full sibs x and y , given that g_x is AB , can be expressed as $P(g_y \text{ is } A_d B_d | g_x \text{ is } AB) = F_{xy}$. Since g_x can arise as a nonrecombinant, g_x^{NR} , or recombinant gamete, g_x^R ,

$$P(g_y \text{ is } A_d B_d | g_x \text{ is } AB) = P(g_x^{NR} \text{ is } AB \cap g_y \text{ is } A_d B_d) / P(g_x \text{ is } AB) + P(g_x^R \text{ is } AB \cap g_y \text{ is } A_d B_d) / P(g_x \text{ is } AB). \quad (1)$$

The first term on the right side of this equation gives the probability that g_y is AB , given that alleles A and B in g_x were derived from the same parent of x , i.e., g_x^{NR} signifies a nonrecombinant gamete of x while g_x^R indicates a recombinant gamete of x . Taking appropriate parental genotypes into account, we have

$$\begin{aligned} &P(g_x^{NR} \text{ is } AB \cap g_y \text{ is } A_d B_d) \\ &= 2P(g_x^{NR} \text{ is } AB \cap g_y \text{ is } A_d B_d | \text{one parent of } x \text{ or } y \text{ is } AABB)P(AABB) \\ &\quad + 2P(g_x^{NR} \text{ is } AB \cap g_y \text{ is } A_d B_d | \text{one parent is } AABb)P(AABb) + \dots \text{etc.} \\ &= 2[c'[(\frac{1}{2})c']^3 g_{11}^{21} + c^2[(\frac{1}{2})c']^2 g_{11}^{21}] + 2\{[(\frac{1}{2})c']^4 (2g_{11}g_{13}) \\ &\quad + [(\frac{1}{2})c]^2 [(\frac{1}{2})c']^2 (2g_{11}g_{13})\} + \dots \text{etc.} \\ &= [(\frac{1}{8})(2c'^4 + 2c^2 c'^2)(g_{11}^{21} + g_{11}g_{13} + g_{11}g_{31}) + (\frac{1}{4})g_{11}g_{33}c'^4 + (\frac{1}{4})g_{11}g_{31}c^2 c'^2] \\ &= [(\frac{1}{8})g_{11}(2c'^4 + 2c^2 c'^2) - (\frac{1}{4})c^2 c'^2 D] \end{aligned} \quad (2)$$

Note that for $D = 0$, expression (2) reduces to that of HALDANE (1949).

The second term on the right side of expression (1) is the probability that g_y is AB , given that alleles A and B of g_x were derived from recombination between parental genotypes of x , that is,

$$\begin{aligned} &P(g_x^R \text{ is } AB \cap g_y \text{ is } A_d B_d) \\ &= P(g_y \text{ is } A_d B_d | g_x \text{ is recombinant} \cap g_x \text{ is } AB)P(g_x \text{ is recombinant} \cap g_x \text{ is } AB). \end{aligned} \quad (3)$$

Now, since $P(g_x \text{ is } AB) = g_{11} - cD$ and

$P(g_x \text{ is nonrecombinant} \cap g_x \text{ is } AB) = c'g_{11}^{21} + (\frac{1}{2})c'(2g_{11}g_{13}) + (\frac{1}{2})c'(2g_{11}g_{31}) + (\frac{1}{2})c'(2g_{11}g_{33}) = g_{11}c'$, we have, $P(g_x \text{ is recombinant} \cap g_x \text{ is } AB) = g_{11} - cD - g_{11}c' = c(g_{11} - D)$. The probability that y has alleles A and B derived from two parents carrying these alleles respectively is $\frac{1}{4}$, so that

$$\begin{aligned} &P(g_y \text{ is } AB | y \text{ has alleles } A \text{ and } B \text{ derived from two parents}) \\ &= [(\frac{1}{2})c](\frac{1}{4}) \end{aligned}$$

and therefore, (3) becomes

$$P(g_x^R \text{ is } AB \cap g_y \text{ is } A_d B_d) = (\frac{1}{8})c^2 (g_{11} - D). \quad (4)$$

From (2) and (4), we finally have

$$P(g_y \text{ is } A_d B_d \mid g_x \text{ is } AB) = [(1/8)g_{11}(2c'^4 + 2c^2c'^2 + c^2) - (1/2)c^2D \times [(1/4) + (1/2)c'^2]] / (g_{11} - cD)$$

and

$$F_{xy}^{AB} = P(g_y \text{ is } A_d B_d \mid g_x \text{ is } AB)P(g_x \text{ is } AB) = \{(1/8)g_{11}(2c'^4 + 2c^2c'^2 + c^2) - (1/2)c^2D [(1/4) \pm (1/2)c'^2]\}. \quad (5)$$

Expression (5) reduces to HALDANE's expression $(1/8)(2c'^4 + 2c^2c'^2 + c^2)$ for the case $D = 0$ where $F_{xy} = g_{11}F_{xy}^{AB} + g_{13}F_{xy}^{Ab} + g_{31}F_{xy}^{aB} + g_{33}F_{xy}^{ab}$. For $D \neq 0$, analogous expressions for F_{xy}^{Ab} , F_{xy}^{aB} , F_{xy}^{ab} need also be derived in the same way as F_{xy}^{AB} given here.