

# Effects of identity disequilibrium and linkage on quantitative variation in finite populations

HIDENORI TACHIDA AND C. CLARK COCKERHAM

*Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203, USA*

*(Received 17 March 1988 and in revised form 15 July 1988)*

## Summary

Identity disequilibrium, ID, is the difference between joint identity by descent and the product of the separate probabilities of identity by descent for two loci. The effects of ID on the additive by additive ( $a * a$ ) epistatic variance and joint dominance component between populations and in the additive, dominance and  $a * a$  variance within populations, including the effects on covariances of relatives within populations, were studied for finite monoecious populations. The effects are formulated in terms of three additive partitions,  $\eta_b$ ,  $\eta_a$  and  $\eta_d$ , of the total ID, each of which increases from zero to a maximum at some generation dependent upon linkage and population size and decreases thereafter.  $\eta_d$  is about four times the magnitude of the other two but none is of any consequence except for tight linkage and very small populations. For single-generation bottleneck populations only  $\eta_d$  is not zero. With random mating of expanded populations  $\eta_b$  remains constant and  $\eta_a$  and  $\eta_d$  go to zero at a rate dependent upon linkage, very fast with free recombination. The contributions of joint dominance to the genetic components of variance within and between populations are entirely a function of the  $\eta$ 's while those of  $a * a$  variance to the components are functions mainly of the coancestry coefficient and only modified by the  $\eta$ 's. The contributions of both to the covariances of half-sibs, full-sibs and parent-offspring follow the pattern expected from their contributions to the genetic components of variance within populations except for minor terms which most likely are of little importance.

## 1. Introduction

Random mating finite populations, populations that have been through a bottleneck and experimental populations initiated with a few parents or inbred lines are generally treated as distinct entities. For many characterizations including quantitative genetic variation they all fit into the same framework, differing only in details. In studying the quantitative genetic variation within and between finite populations it was found that a joint dominance contribution for a pair of loci was determined entirely by identity disequilibrium (Cockerham, 1984*a*). The question arises as to the effects of identity disequilibrium on other joint effects of loci, epistatic variance.

Identity disequilibrium (ID) is the difference between the joint probability of identity by descent and the product of the separate probabilities of identity by descent for two loci. It is never negative and occurs for pedigree systems of mating for linked loci, increasing as recombination decreases (Weir & Cockerham, 1969*a*). It occurs in finite random mating populations, even for unlinked loci when it is a measure of the

variation in inbreeding coefficients of members of the population with varying pedigrees, and is further enhanced by linkage (Weir & Cockerham, 1969*b*).

In random mating finite populations, initially in linkage equilibrium, ID is the principal agent for the variation in actual inbreeding or heterozygosity (Weir, Avery & Hill, 1980) and for the variation in linkage disequilibrium (Weir & Hill, 1980).

We wish to extend the analysis to include the effects of ID on additive by additive ( $a * a$ ) variance in finite populations. The  $a * a$  variance with self-fertilization was studied in some detail with an emphasis on the effects of ID on the covariances of relatives (Cockerham, 1984*b*). In this case ID was entirely a function of linkage and the relative effect of ID on the covariances of relatives decreased as the relatedness of the relatives increased.

To extend the analysis to finite populations we consider random mating monoecious populations as in Cockerham (1984*a*). The initial population is in linkage and Hardy–Weinberg equilibrium and gives rise to replicate finite populations of size  $N$  each

generation. At some time the replicate populations are expanded,  $N = \infty$ , and random mating is continued in each. The purpose of the expanded populations is to diagnose the effects of finite population history on quantitative genetic variation and of continued random mating to determine the permanency or transient behaviour of the effects. Obviously, populations that have been through a bottleneck are included in this framework. For the expanded populations we partition the  $a * a$  variance into portions within and between populations and further subdivide the portion within populations into portions that behave as additive and  $a * a$  variances. The total ID is also partitioned and the effects of the partitions on the various portions of  $a * a$  variance are formulated for the expanded population and for modifications that occur with continued random mating. A parallel partitioning is also given for joint dominance effects. The contributions of the joint dominance and  $a * a$  variance to the covariance of half-sibs, full-sibs and parent-offspring within populations are formulated.

**2. ID measures**

The analysis is facilitated by measures of the probability of identity by descent. For two loci the double identity by descent measures are for a pair of genes at one locus and a pair of genes at the other locus. For random mating monoecious populations only three are required (Cockerham, 1984a):  $\tilde{\theta}$  for genes on two gametes,  $\tilde{\gamma}$  for genes on three gametes, and  $\tilde{\Delta}$  for genes on four gametes. The measures are the same for all configurations within and between random individuals. Consequently,  $\tilde{\theta}$  is the joint inbreeding coefficient  $\tilde{F}$ , and so on. Transition equations from which these two-locus descent measures can be obtained are given in Appendix A in Cockerham (1984a).

One single-locus measure, the coancestry coefficient  $\theta$ , which is also the inbreeding coefficient  $F$ , is needed, and  $\theta_t = 1 - (1 - 1/2N)^t$ .

It was found useful (Cockerham, 1984a) in partitioning the joint dominance contributions for two loci for finite populations to partition the total ID,  $\eta = \tilde{\theta} - \theta^2$ , into three parts associated with different variance components.

$$\begin{aligned} \eta_b &= \tilde{\Delta} - \theta^2 && : \text{between population variance} \\ \eta_a &= 2(\tilde{\gamma} - \tilde{\Delta}) && : \text{additive variance} \\ \eta_d &= \tilde{\theta} - 2\tilde{\gamma} + \tilde{\Delta} && : \text{dominance variance} \end{aligned}$$

These sum to the total  $\eta_b + \eta_a + \eta_d = \eta$ .

The descent measures for the expanded population will be determined by  $t$ , the number of generations before expansion, and  $N$ , the number of individuals in the replicate populations. With further random mating of the expanded populations the descent measures  $\tilde{\theta}$  and  $\tilde{\gamma}$  will change in time measured now by  $\tau$  with  $\tau = 0$  for the initial expanded populations.

To facilitate the presentation, we speak of gametic,

$\mathcal{G}$ , and non-gametic,  $\mathcal{N}$ , pairs of genes. The status of  $\mathcal{G}$  changes over time. In terms of  $\alpha = (1 + \lambda)/2$  where  $(1 - \lambda)/2 = r$ , the recombination fraction,  $\mathcal{G}_{\tau+1} = \alpha \mathcal{G}_\tau + (1 - \alpha) \mathcal{N}_\tau$ , i.e. gametic genes were gametic in the previous generation with probability  $\alpha$  and non-gametic in the previous generation with probability  $1 - \alpha$ . In contrast in an infinite population  $\mathcal{N}_{\tau+1} = \mathcal{N}_\tau$  since genes on two gametes trace back to genes on one gamete with probability zero. From the definitions of  $\tilde{\theta}$ ,  $\tilde{\gamma}$  and  $\tilde{\Delta}$ ,  $\tilde{\theta}_{\tau+1} = \alpha^2 \tilde{\theta}_\tau + 2\alpha(1 - \alpha) \tilde{\gamma}_\tau + (1 - \alpha)^2 \tilde{\Delta}_\tau$ ;  $\tilde{\gamma}_{\tau+1} = \alpha \tilde{\gamma}_\tau + (1 - \alpha) \tilde{\Delta}_\tau$ ; and  $\tilde{\Delta}_{\tau+1} = \tilde{\Delta}_\tau$ . In terms of initial values  $\tilde{\theta}_\tau = \alpha^{2\tau} \tilde{\theta}_0 + 2\alpha^\tau(1 - \alpha^\tau) \tilde{\gamma}_0 + (1 - \alpha^\tau)^2 \tilde{\Delta}_0$ ;  $\tilde{\gamma}_\tau = \alpha^\tau \tilde{\gamma}_0 + (1 - \alpha^\tau) \tilde{\Delta}_0$ ; and  $\tilde{\Delta}_\tau = \tilde{\Delta}_0$ . Thus, in time, all two-locus descent measures are the same,  $\tilde{\theta}_\infty = \tilde{\gamma}_\infty = \tilde{\Delta}_\infty = \tilde{\Delta}_0$ .

Except for  $\eta_b = \tilde{\Delta}_0 - \theta^2$ , the identity disequilibria go to zero in time:  $\eta_{a\tau} = \alpha \eta_{a\tau-1} = \alpha^\tau \eta_{a0}$  and  $\eta_{d\tau} = \alpha^2 \eta_{d\tau-1} = \alpha^{2\tau} \eta_{d0}$ . The rate of approach to zero per generation for each is a constant,  $1 - \alpha = r$  for  $\eta_a$  and  $1 - \alpha^2 = r(2 - r)$  for  $\eta_d$ . With free recombination the rates are fast, 1/2 for  $\eta_a$  and 3/4 for  $\eta_d$ .

To obtain some notion of the magnitude of  $\eta$ 's in relation to  $N$ ,  $t$  and  $\lambda$  some values are plotted against  $t$  in Fig. 1 for  $N = 4, 20, 100$  and  $\lambda = 0, 0.5, 0.9$ . Each line in the figure that can be distinguished from zero is labelled according to  $\lambda, N$  values.

Each of the  $\eta$ 's is zero initially, builds up to a maximum and eventually goes to zero at fixation. There is considerable similarity in the shape of the curves among the  $\eta$ 's.  $\eta_d$  reaches its maximum quickest and  $\eta_a$  next. An increase in  $\lambda$  increases the time required to reach the maximum. Note the difference in scale for  $\eta_d$  which is about four times that for the other two. While linkage increases the  $\eta$ 's  $\lambda$  must be near 0.9 for the increase to be of much consequence. With  $\lambda = 0$  the  $\eta$ 's are due to the variation among inbreeding coefficients (Weir & Cockerham, 1969b) and of little consequence except for  $\eta_d$  in extremely small populations. In fact, with  $N \geq 20$  it would appear safe to ignore  $\eta_b$  and  $\eta_a$ , particularly when viewed over many pairs of loci where  $\lambda$  on the average is probably very small, and in this case even  $\eta_d$  will be very small.

**3. Partitions of the  $a * a$  variance and joint dominance effects**

For a genotype formed from the union of gametes  $A_k B_k$  and  $A_l B_l$  the  $a * a$  effects in the model for the genotypic value may be symbolized as  $(a_k b_k) + (a_l b_l) + (\overline{a_k b_l}) + (\overline{a_l b_k})$  where overbars indicate non-gametic pairs of genes. In the initial equilibrium population for which these effects are defined each has mean zero and  $\mathcal{E}(ab)^2 = \sigma_{aa}^2$  where  $\mathcal{E}$  denotes expectation. The effects are uncorrelated so that the total  $a * a$  variance among individuals in the population is  $4\sigma_{aa}^2 = \sigma_{\overline{aa}}^2$  for a pair of loci. With inbreeding the effects become correlated and it was shown in Cockerham (1984b) that the  $a * a$  variance in the total variance  $\sigma_p^2$  among

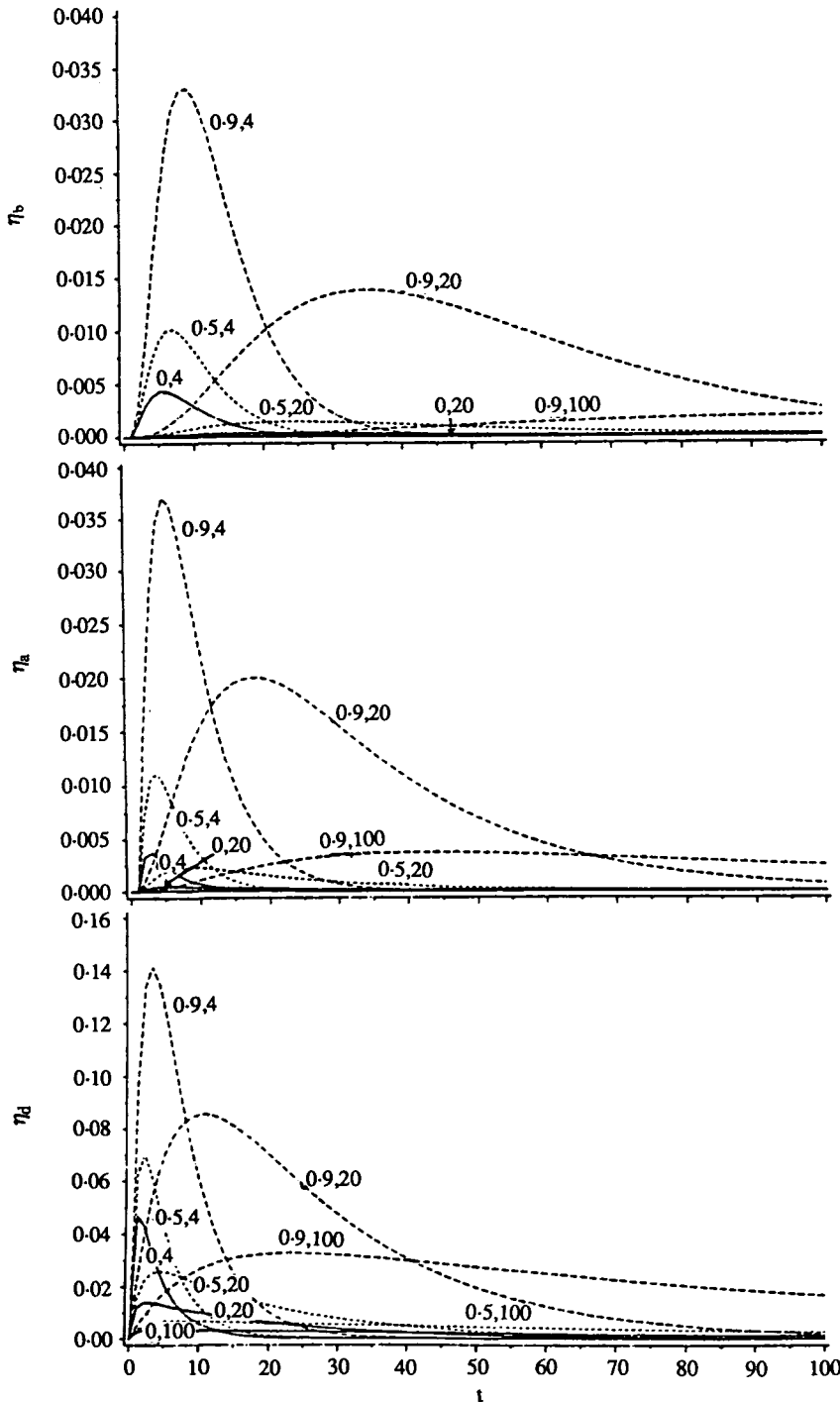


Fig. 1. Plots of  $\eta_b$ ,  $\eta_a$  and  $\eta_d$  against  $t$  for  $N = 4, 20, 100$  and  $\lambda = 0, 0.5, 0.9$ . Plots that can be distinguished from

zero are labelled with values of  $\lambda, N$ .

unrelated but inbred individuals is  $(1 + 2\theta + \tilde{\theta})\sigma_{aa}^2$  since  $\theta = F$  and  $\tilde{\theta} = \bar{F}$ .

Procedures were also established in Cockerham (1984b) for evaluating the  $a * a$  variance in the covariance,  $\mathcal{C}_{XY}$ , between individuals X and Y, including relatives. This variance is  $(\tilde{\theta}_{XY} + \tilde{\gamma}_{X\bar{Y}} + \tilde{\gamma}_{\bar{X}Y} + \tilde{\Delta}_{\bar{X}\bar{Y}})\sigma_{aa}^2$  where an overbar indicates non-gametic genes. Thus, it is necessary to evaluate the double identity measures for the configurations of pairs of genes for X with those for Y.

If X and Y are random members within our finite populations the  $a * a$  variance in  $\mathcal{C}_{XY}$  is  $(\tilde{\theta} + 2\tilde{\gamma} + \tilde{\Delta})\sigma_{aa}^2$ . We partition the total variance into that between populations,  $\sigma_B^2 = \mathcal{C}_{XY}$ , and that within populations  $\sigma_W^2 = \sigma_F^2 - \mathcal{C}_{XY}$ . These partitions for  $\sigma_{aa}^2$  are given in Table 1.

When additive effects within populations are defined in the usual least-squares manner, the variance of these effects constitutes the additive variance,  $\sigma_{A^*}^2$ , within populations as an average value over replicate

Table 1. Coefficients of a \* a variance,  $\sigma_{aa}^2$ , and joint dominance effect,  $2hh'$ , in the components of variance

	$\sigma_B^2$	$\sigma_W^2$	$\sigma_{A^*}^2$	$\sigma_{D^*}^2$	$\sigma_{aa^*}^2$
$\sigma_{aa}^2$	$\tilde{\theta} + 2\tilde{\gamma} + \tilde{\Delta}$	$1 + 2\theta - 2\tilde{\gamma} - \tilde{\Delta}$	$4\theta - \tilde{\theta} - 2\tilde{\gamma} - \tilde{\Delta}$	—	$1 - 2\theta + \tilde{\theta}$
$\sigma_{aa}^2$	$4\tilde{\Delta} + 2\eta_a + \eta_d$	$1 + 2\theta - 3\tilde{\Delta} - \eta_a$	$4(\theta - \tilde{\Delta}) - 2\eta_a - \eta_d$	—	$1 - 2\theta + \tilde{\Delta} + \eta_a + \eta_d$
$2hh'$	$\eta_b$	$\eta_a + \eta_d$	$\eta_a$	$\eta_d$	

populations. Jiang and Cockerham (unpublished) have shown that  $(4\theta - \tilde{\theta} - 2\tilde{\gamma} - \tilde{\Delta})\sigma_{aa}^2$  becomes a part of  $\sigma_{A^*}^2$  and the remainder,  $\sigma_{aa^*}^2$ , we designate as a \* a variance within populations (Table 1). Previously, Cockerham & Tachida (1988) showed that  $4(\theta - \tilde{\Delta})\sigma_{aa}^2$  is a permanent part of  $\sigma_{A^*}^2$ , that is conserved between ancestors and descendants. In the second row of Table 1 the coefficients are formulated utilizing  $\eta_a$  and  $\eta_d$  which are transient parts of the coefficients.

With complete identity equilibrium the coefficient of  $\sigma_{aa}^2$  in  $\sigma_{A^*}^2$  would be  $4(\theta - \theta^2)$ . The percentage reduction for the coefficient  $4(\theta - \tilde{\Delta})$  due to permanent identity disequilibrium is given by  $100 \eta_b / \theta(1 - \theta)$ . These values are plotted against  $t$  for combinations of  $N$  and  $\lambda$  in Figure 2. There is a maximum of about 18% reduction for  $\lambda = 0.9$  and  $N = 4$  with a drastic decrease in the reduction as  $\lambda$  decreases and  $N$  increases. For  $\lambda = 0.9$  and  $N = 100$  the maximum reduction is less than 1%.

The joint dominance effect which involves the product of the inbreeding depressions,  $h$  and  $h'$ , for the two loci are also included in Table 1. It was found in Cockerham (1984a) that the joint dominance contributions belonged in part in  $\sigma_{A^*}^2$  which was substantiated in Cockerham & Tachida (1988) and in part in the dominance variance,  $\sigma_{D^*}^2$ , within populations. The purpose of the inclusion in Table 1 is to demonstrate the behaviour of the coefficients in an expanded random mating population over time. The

coefficients are given in terms of  $\eta$ 's. Only the portion between populations, reflected by  $\eta_b$ , remains constant and the portions within populations go to zero.

4. Joint dominance and a \* a variance in the covariances of relatives

The covariance of relatives is based on the pairs of these relatives being independent. For the covariance of half-sibs within populations, for example, we must correct for the covariance,  $\mathcal{C}_{UU'}$ , between random members of the population which are unrelated other than the relationship brought on by the previous restricted population size. The adjusted covariance is  $\mathcal{C}_{HS}^* = \mathcal{C}_{HS} - \mathcal{C}_{UU'}$ . The contributions of a \* a variance and  $2hh'$  to the covariances of half-sibs and full-sibs,  $\mathcal{C}_{FS}^*$ , are given in the Appendix and to the covariance of parent and offspring,  $\mathcal{C}_{PO}^*$ , in Cockerham & Tachida (1988). To summarize the results, let  $G_a = 4(\theta - \tilde{\Delta})\sigma_{aa}^2 + \eta_a 2hh'$  and  $G_{aa} = (1 - 2\theta + \tilde{\Delta})\sigma_{aa}^2$ . Then ( $\supset$  means contains)

$$\mathcal{C}_{HS}^* \supset \frac{1}{4}G_a + \frac{(1 + \lambda^2)}{16}G_{aa} + \frac{1}{16}[(1 - 6\lambda + \lambda^2)\eta_a - 2\lambda\eta_d]\sigma_{aa}^2$$

$$\mathcal{C}_{FS}^* \supset \frac{1}{2}G_a + \frac{(2 + \lambda^2)}{8}G_{aa} + \frac{1}{8}[(1 - 6\lambda + \lambda^2)\eta_a + (1 - 2\lambda)\eta_d]\sigma_{aa}^2 + \frac{1}{4}\eta_d 2hh'$$

$$\mathcal{C}_{PO}^* \supset \frac{1}{2}G_a + \frac{1}{4}G_{aa} - \frac{1}{4}[(1 + \lambda)\eta_a + \lambda\eta_d]\sigma_{aa}^2$$

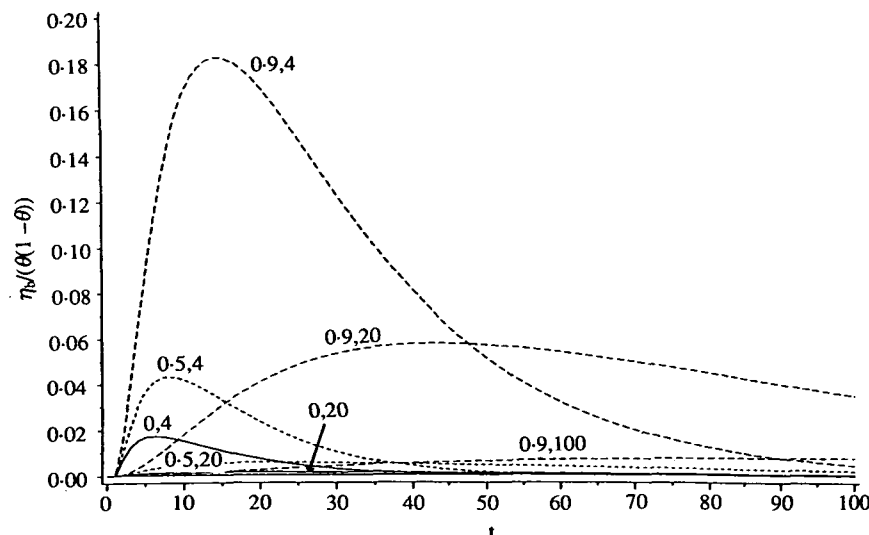


Fig. 2. Percentage reduction in the coefficient  $4(\theta - \tilde{\Delta})$  due to identity disequilibrium. Plots are labelled with values of  $\lambda, N$ .

Table 2. Identity by descent (ibd) measures for one and two loci†

One locus		Two loci			
ibd*	Gene arrangement‡	ibd**	Gene arrangement	ibd**	Gene arrangement
F	(a   a')	$\hat{\Theta}_1$	(ab   a'b')	$\tilde{\Delta}_1$	(a   b)(a'   b')
$\theta$	(a)(a')	$\hat{\Theta}_2$	(ab)(a'b')	$\tilde{\Delta}_2$	(a   a')(b   b')
		$\tilde{\Gamma}_1$	(ab)(a'   b')	$\tilde{\Delta}_3$	(a   b)(a')(b)
		$\tilde{\Gamma}_2$	(ab   a')(b')	$\tilde{\Delta}_4$	(a   a')(b)(b')
		$\tilde{\Gamma}_3$	(ab)(a')(b')	$\tilde{\Delta}_5$	(a)(a')(b)(b')

† Genes are represented by a and by a' for two distinct genes at the A-locus and b and b' for two distinct genes at the B-locus.

‡ Genes separated by | are on different gametes in the same individual and those separated by () are in different individuals.

\* Probability that the two genes a and a' are identical by descent. Symmetrical arrangements are omitted.

\*\* Probability that the two gene pairs, a and a' and b and b', are jointly identical by descent. Symmetrical arrangements are omitted.

The descent measures are for the parental generation. Note that the coefficients of  $G_a$  are the same as for additive variance, of  $G_{aa}$  are the same as for a\*a variance and of  $\eta_a 2hh'$  are the same as for the dominance variance for these relatives in an infinite equilibrium population. There are additional terms involving  $\sigma_{aa}^2$  and identity disequilibrium which are always negative for  $\mathcal{C}_{PO}^*$  but negative for  $\mathcal{C}_{HS}^*$  and  $\mathcal{C}_{FS}^*$  only for large values of  $\lambda$  and depending on  $\eta_a$  and  $\eta_d$ . As noted previously  $\eta_a$  and  $\eta_d$  disappear rapidly for small  $\lambda$  in which case these additional terms disappear. Also, they are small except for very small  $N$  and large  $\lambda$ .

5. Other mating structures

We have considered only random mating monoecious populations including selfing and there is a question as to the extent to which the results are modified by other mating structures such as separate sexes or monogamy. Weir, Avery & Hill (1980) studied the effects of mating structure on variation in inbreeding, and Weir & Hill (1980) studied the effects of mating structure on variation in linkage disequilibrium. While they did not utilize the term 'identity disequilibrium', their results can be expressed in terms of various functions of identity disequilibrium.

In Weir, Avery & Hill (1980) and Weir & Hill (1980) all measures are of non-identity by descent. For a single locus  $P = 1 - F$  and  $\Pi = 1 - \theta$ . The 10 two-locus measures, required to accommodate the avoidance of self-fertilization and separate sexes with hierarchical or monogamous types of matings, are for double non-identity by descent. We place a ~ on each to denote double identity by descent and our descent measures are listed in Table 2.

To accommodate these other mating structures as well as monoecy, the coefficients, Table 1, of  $2hh'$  are  $\tilde{\Delta}_2 - F^2$  in  $\sigma_B^2$ ,  $\tilde{\theta}_1 - \tilde{\Delta}_2$  in  $\sigma_W^2$ ,  $2(\tilde{\Gamma}_2 - \tilde{\Delta}_4)$  in  $\sigma_A^2$ , and  $\theta_1 - 2\tilde{\Gamma}_2 - \tilde{\Delta}_2 + 2\tilde{\Delta}_4$  in  $\sigma_{D^*}^2$ . The coefficients of  $\sigma_{aa}^2$  are  $\tilde{\theta}_2 + 2\tilde{\Gamma}_1 + \tilde{\Delta}_1$  in  $\sigma_B^2$ ,  $1 + 2F + \tilde{\theta}_1 - \tilde{\theta}_2 - 2\tilde{\Gamma}_1 - \tilde{\Delta}_1$  in

$\sigma_W^2$ ,  $4(\theta - \tilde{\Delta}_3 + \tilde{\Gamma}_2 - \tilde{\Gamma}_3)$  in  $\sigma_A^2$ , and the remaining part of the coefficient for  $\sigma_W^2$  in  $\sigma_{aa}^2$ . The coefficients in  $\sigma_A^2$ , were found by the same method of covariance between ancestor and descendant utilized in Cockerham & Tachida (1988). With monoecy all descent measures in a category are equal, e.g.  $\tilde{\Gamma}_1 = \tilde{\Gamma}_2 = \tilde{\Gamma}_3 = \tilde{\gamma}$ .

In the initial expanded population these coefficients are determined by the mating structure and the finite history. With one generation of random mating, either monoecious or dioecious, we find  $F_1 = \theta_1 = \theta_0$  where  $\theta_0$  is the ancestral population value, and

$$\tilde{\Delta}_{1,1} = \tilde{\Delta}_{2,1} = \tilde{\Delta}_{3,1} = \tilde{\Delta}_{4,1} = \tilde{\Delta}_{5,1} = \tilde{\Delta}_{5,0}$$

$$\tilde{\Gamma}_{1,1} = \tilde{\Gamma}_{2,1} = \tilde{\Gamma}_{3,1} = \alpha \tilde{\Gamma}_{3,0} + (1 - \alpha) \tilde{\Delta}_{5,0}$$

$$\tilde{\Theta}_{1,1} = \tilde{\Theta}_{2,1} = \alpha^2 \tilde{\Theta}_{2,0} + 2\alpha(1 - \alpha) \tilde{\Gamma}_{3,0} + (1 - \alpha)^2 \tilde{\Delta}_{5,0},$$

where  $X_{1,\tau}$  is for the  $\tau$ th generation. After that we need only to substitute  $\theta$ ,  $\tilde{\Theta}_2$ ,  $\tilde{\Gamma}_3$  and  $\tilde{\Delta}_5$  for the initial values of  $\theta$ ,  $\tilde{\theta}$ ,  $\tilde{\gamma}$  and  $\tilde{\Delta}$  respectively to accommodate the finite histories of other breeding structures. The identity disequilibriums are  $\eta_b = \tilde{\Delta}_5 - \theta^2$ ,  $\eta_a = 2(\tilde{\Gamma}_3 - \tilde{\Delta}_5)$  and  $\eta_d = \tilde{\Theta}_2 - 2\tilde{\Gamma}_3 + \tilde{\Delta}_5$  where  $\eta_b$  remains constant and  $\eta_a$  and  $\eta_d$  go to zero at the same rate as before. Thus, the formulations in Table 1 need only to be adjusted for the initial values appropriate for the mating structure.

While the results of Weir & Hill (1980) and Weir, Avery & Hill (1980) are functions of ID their formulations and standardizations preclude exact comparisons with ours. The main differences among mating structures in Weir, Avery & Hill were related to the degree of avoidance of mating relatives and were proportionately larger as linkage decreased.

To inquire directly into differences due to mating structure  $\eta_b$ ,  $\eta_a$  and  $\eta_d$  are plotted against  $t$  for monoecy,  $M$ , monoecy with the exclusion of selfing,  $M_e$ , and monogamy,  $M_o$ , all with  $N = 4$ , in Fig. 3. We chose  $N = 4$  to exaggerate the differences in mating structure. Even so the curves are very similar. There is a slight shift to the right for  $M_e$  and  $M_o$ , both of which avoid self-fertilization. Near the maximum  $M_e$  is



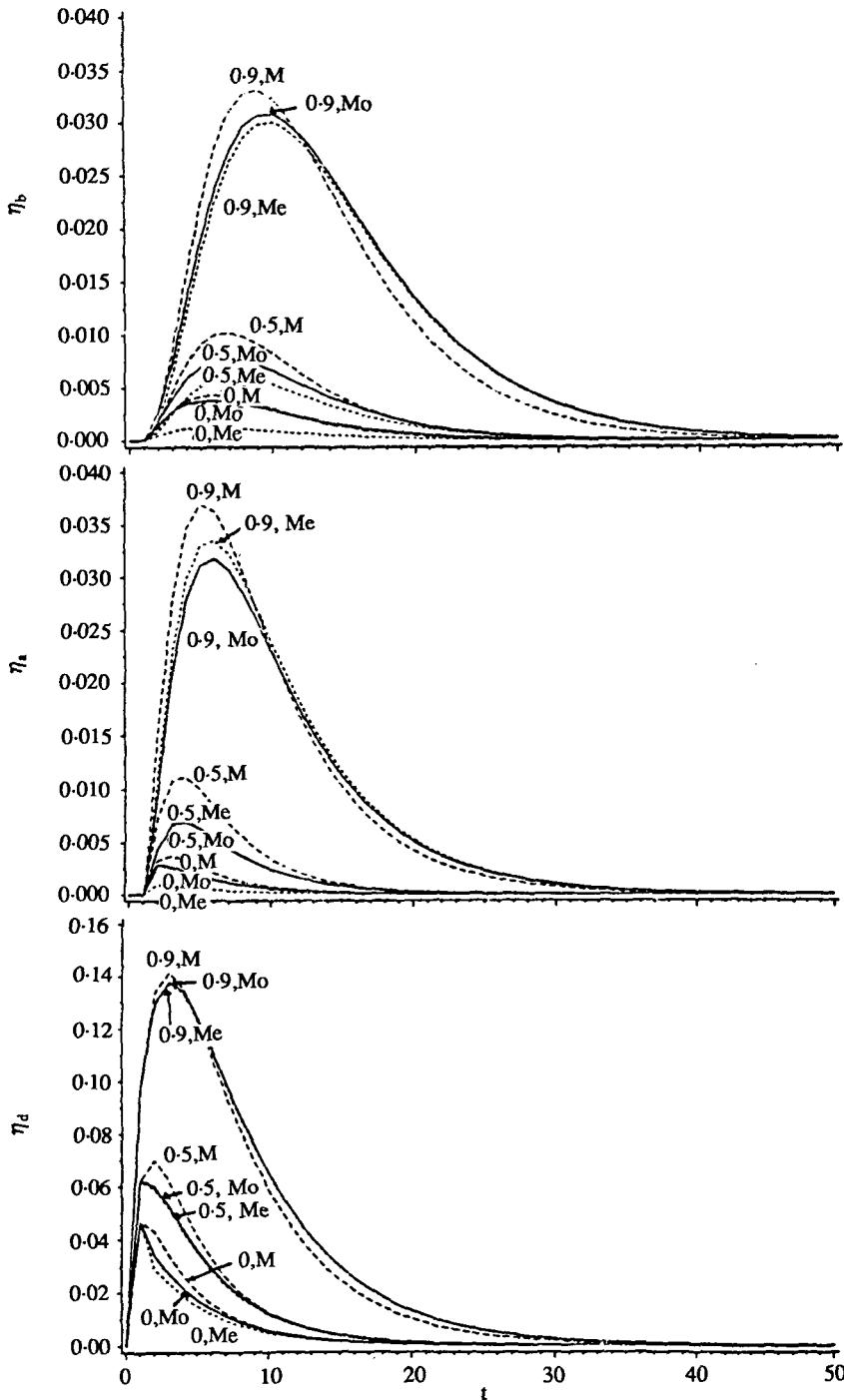


Fig. 3. Plots of  $\eta_b$ ,  $\eta_a$  and  $\eta_d$  against  $t$  for monoecy,  $M$ , monoecy with selfing excluded,  $M_e$ , and monogamy,  $M_o$ ,

all with  $N = 4$ . Plots are labelled with values of  $\lambda$ ,  $X(X = M, M_e, M_o)$ .

intermediate to  $M$  and  $M_o$  reflecting an increased ID due to a higher frequency of full-sib mating than for  $M_o$  but a decreased ID due to no self-fertilization in comparison to  $M$ . The greatest percentage differences are for small  $\lambda$  but then the  $\eta$ 's are of little consequence.

### 6. Discussion

We have utilized expanded populations, which was done in Cockerham & Tachida (1988) and implied in

Cockerham (1984a), as a means of diagnosing the effects of finite population history on quantitative genetic variation. Continued random mating of the expanded populations was utilized to determine the permanency or transient behaviour of the effects. The effects of ID on  $2hh'$  are considerably different from those on  $a * a$  variance.

The coefficients of  $2hh'$  in the components in Table 1 are entirely a function of the  $\eta$ 's. Those,  $\eta_a$  and  $\eta_d$ , for components within populations go to zero with

continued random mating while the variation  $\eta_b 2hh'$ , between populations remains constant. The portion  $\eta_a 2hh'$  in  $\sigma_{A^*}^2$  contributes to permanent response in descendants due to selection (Cockerham & Tachida, 1988) but is reduced each generation of random mating before selection is practised.

The coefficients of  $\sigma_{\bar{a}\bar{a}}^2$  in Table 1 are functions mainly of  $\theta$  and  $\bar{\Delta}$ . The main component of interest is  $\sigma_{A^*}^2$  and  $4(\theta - \bar{\Delta})\sigma_{\bar{a}\bar{a}}^2$  does not change with random mating of the expanded population and contributes the same permanent response in descendants from selection in any generation. The coefficient  $4(\theta - \bar{\Delta})$  is reduced very little by ID, Fig. 2. There is a slight change of  $\sigma_{\bar{a}\bar{a}}^2$  within populations, and of the coefficient of  $\sigma_{\bar{a}\bar{a}}^2$  in  $\sigma_{A^*}^2$  with continued random mating as  $\eta_a$  and  $\eta_d$  go to zero, Table 1.

The differences due to mating structure are primarily in terms of the initial descent measures in the expanded population in Fig. 3 and these differences tend to disappear when adjustment is made for effective population size.

The foregoing results pertain to a population that continues to have a bottleneck for some time. There is considerable simplification for a single-generation bottleneck. For a single-generation bottleneck or an experimental population initiated from  $N$  random outbred parents,  $\eta_b = \eta_a = 0$  and  $\eta_d = (1 + \lambda^2 - 1/N)/4N$  (by simple probability arguments or by letting  $t = 1$  in the expressions of Appendix A of Cockerham, 1984a). This ID dissipates at the rate  $r(2-r)$  with random mating. Also, for a population initiated with  $n$  random inbred lines or homozygous parents the only ID is  $\eta_d = (n-1)/n^2$  which soon dissipates except for tight linkage. There are corresponding simplifications in the covariances of relatives. Bryant *et al.* (1986) let their experimental single-generation bottleneck populations flush to normal size in about five generations before estimating the covariance between mid-parent and offspring which fairly well insured that the estimates are affected little by identity disequilibrium.

Finite population history can modify the genetic components  $\sigma_{A^*}^2$ ,  $\sigma_{D^*}^2$  and  $\sigma_{\bar{a}\bar{a}}^2$  within populations considerably, but linkage and concomitant identity disequilibrium play very minor roles in the modifications, particularly after a few generations of random mating of the expanded populations.

Drs B. S. Weir and C. H. Langley provided helpful comments. Paper no. 11502 of the Journal Series of the North Carolina Agricultural Research Service, Raleigh, N.C. 27695-7601. This investigation was supported in part by Research Grant GM 11546 from the National Institute of General Medical Sciences.

References

Bryant, E. H., McCommas, S. A. & Combs, L. M. (1986). The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. *Genetics* **114**, 1191-1211.

Cockerham, C. C. (1984a). Covariances of relatives for quantitative characters with drift. In *Human Population Genetics: The Pittsburgh Symposium* (ed. A. Chakravarti), pp. 195-208. New York: Van Nostrand Reinhold.  
 Cockerham, C. C. (1984b). Additive by additive variance with inbreeding and linkage. *Genetics* **108**, 487-500.  
 Cockerham, C. C. & Tachida, H. (1988). Permanency of response to selection for quantitative characters in finite populations. *Proceedings of the National Academy of Sciences U.S.A.* **85**, 1563-1565.  
 Weir, B. S., Avery, P. J. & Hill, W. G. (1980). Effect of mating structure on variation in inbreeding. *Theoretical Population Biology* **18**, 396-429.  
 Weir, B. S. & Cockerham, C. C. (1969a). Pedigree mating with two linked loci. *Genetics* **61**, 923-940.  
 Weir, B. S. & Cockerham, C. C. (1969b). Group inbreeding with two linked loci. *Genetics* **63**, 711-742.  
 Weir, B. S. & Hill, W. G. (1980). Effect of mating structure on variation in linkage disequilibrium. *Genetics* **95**, 477-488.

Appendix

Here, we formulate the  $a * a$  variance in the covariance of half-sibs,  $\mathcal{C}_{HS}^* = \mathcal{C}_{HS} - \mathcal{C}_{UU'}$ , and in the covariance of full-sibs,  $\mathcal{C}_{FS}^* = \mathcal{C}_{FS} - \mathcal{C}_{UU'}$ , within populations where H and F are a half-sib and full-sib of an individual S, respectively, and U and U' are unrelated individuals in the same generation, unrelated except for the finite population history.

We will express the covariance in terms of the descent measures,  $\hat{\theta}$ ,  $\tilde{\gamma}$  and  $\bar{\Delta}$  for the parental generation. Consequently,  $\mathcal{C}_{UU'} \supset (\hat{\theta}_{UU'} + \tilde{\gamma}_{\bar{U}\bar{U}'} + \tilde{\gamma}_{U\bar{U}'} + \bar{\Delta}_{\bar{U}\bar{U}'})\sigma_{\bar{a}\bar{a}}^2 = [\alpha^2\hat{\theta} + 2\alpha(2-\alpha)\tilde{\gamma} + (2-\alpha)^2\bar{\Delta}]\sigma_{\bar{a}\bar{a}}^2$  when translated to the parent generation, where  $\supset$  means contains. For half-sibs let P be the common parent and Q and R be the other parents of H and S respectively. We expand the descent measures back to those in the parent. A parental gamete in S from P is parental or recombinant in P with probability  $\alpha$  and  $1-\alpha$  respectively. The same holds for R. Then, for a random parental gamete for S,  $\hat{\theta}_{HS} = [\alpha(\hat{\theta}_{HP} + \hat{\theta}_{HR}) + (1-\alpha)(\tilde{\gamma}_{HP} + \tilde{\gamma}_{HR})]2^{-1}$ . We further expand these measures to include the parent Q

$$\begin{aligned} \hat{\theta}_{HS} &= [\alpha^2(\hat{\theta}_{PP} + \hat{\theta}_{QP} + \hat{\theta}_{QR} + \hat{\theta}_{PR}) \\ &+ \alpha(1-\alpha)(2\tilde{\gamma}_{P\bar{P}} + \tilde{\gamma}_{Q\bar{P}} + \tilde{\gamma}_{P\bar{R}} \\ &+ \tilde{\gamma}_{Q\bar{R}} + \tilde{\gamma}_{\bar{Q}\bar{P}} + \tilde{\gamma}_{\bar{Q}\bar{R}} + \tilde{\gamma}_{\bar{P}\bar{R}}) \\ &+ (1-\alpha)^2(\bar{\Delta}_{P\bar{P}} + \bar{\Delta}_{Q\bar{P}} + \bar{\Delta}_{P\bar{R}} + \bar{\Delta}_{Q\bar{R}})]2^{-2} \\ &= [\alpha^2(\hat{\theta}_{PP} + 3\hat{\theta}) + 2\alpha(1-\alpha)(\tilde{\gamma}_{P\bar{P}} + 3\tilde{\gamma}) \\ &+ (1-\alpha)^2(\bar{\Delta}_{P\bar{P}} + 3\bar{\Delta})]2^{-2}. \end{aligned}$$

We made use of the fact that distinct parents are random members of that generation. Descent measures involving the same individual require simple probabilistic arguments  $\hat{\theta}_{PP} = (1 + \hat{\theta})2^{-1}$ ,  $\tilde{\gamma}_{P\bar{P}} = \theta$  and  $\bar{\Delta}_{P\bar{P}} = (1 + \hat{\theta})2^{-1}$ . For  $\tilde{\gamma}_{HS}$ , note that one gene of the recombinant gamete in S comes from P and the other gene from R. Then, by an expansion,  $\tilde{\gamma}_{HS} = \tilde{\gamma}_{H\bar{P}\bar{R}}$ , where the subscripts  $\bar{P}$  and  $\bar{R}$  denote that one gene is taken from P and R, respectively. Using an expansion

similar to the one used for a parental gamete in the calculation of  $\tilde{\theta}_{HS}$ , this can be further expanded,

$$\begin{aligned} \tilde{\gamma}_{HS} &= [\alpha\tilde{\gamma}_{QP\dot{P}\dot{R}} + (1-\alpha)\tilde{\Delta}_{\dot{Q}\dot{P}\dot{R}} + \alpha\tilde{\gamma}_{PP\dot{P}\dot{R}} + (1-\alpha)\tilde{\Delta}_{\dot{P}\dot{P}\dot{R}}]2^{-1} \\ &= [\alpha(\tilde{\gamma} + \tilde{\gamma}_{PP\dot{P}\dot{R}}) + (1-\alpha)(\tilde{\Delta} + \tilde{\Delta}_{\dot{P}\dot{P}\dot{R}})]2^{-1}. \end{aligned}$$

Again we made use of the fact that distinct parents are random members of that generation. With simple probabilistic arguments, descent measures involving the same individual are calculated to be  $\tilde{\gamma}_{PP\dot{P}\dot{R}} = \tilde{\Delta}_{\dot{P}\dot{P}\dot{R}} = (\theta + \tilde{\gamma})2^{-1}$ . Because of the symmetry,  $\tilde{\gamma}_{HS}$  is the same as  $\tilde{\gamma}_{HS}$ . Finally, applying the expansion on each recombinant gamete in H and S, we obtain  $\tilde{\Delta}_{\dot{H}\dot{S}} = \tilde{\Delta}_{\dot{Q}\dot{P}\dot{P}\dot{R}}$ . Two genes from P are at the same locus or different loci with probability 1/2 and independently they come from the same gamete or different gametes of P with probability 1/2. Therefore, making use of the fact that distinct individuals are random members of the parent generation,  $\tilde{\Delta}_{\dot{H}\dot{S}} = (\theta + \tilde{\gamma} + 2\tilde{\Delta})2^{-2}$ . Combining these and correcting for the contribution to  $\mathcal{C}_{UU}$ , already calculated, we obtain the a \* a variance in the covariance between half-sibs within populations,

$$\begin{aligned} \mathcal{C}_{HS}^* &\supset [(\theta - \tilde{\Delta}) + (1 + \lambda^2)(1 - 2\theta + \tilde{\Delta} + \eta_a)/16 \\ &\quad - \lambda(3\eta_a + \eta_d)/8] \sigma_{aa}^2. \end{aligned}$$

For full-sibs let the two parents be P and Q. Then, utilizing the fact that any measure involving the same parent twice is the same for P or Q,

$$\begin{aligned} \tilde{\theta}_{FS} &= [\alpha^2(\tilde{\theta}_{PP} + \tilde{\theta}) + 2\alpha(1-\alpha)(\tilde{\gamma}_{PP} + \tilde{\gamma}) \\ &\quad + (1-\alpha)^2(\tilde{\Delta}_{PP} + \tilde{\Delta})]2^{-1} \\ \tilde{\gamma}_{FS} &= \tilde{\gamma}_{FS} = \alpha\tilde{\gamma}_{PPQ} + (1-\alpha)\tilde{\Delta}_{PPQ} \\ \tilde{\Delta}_{FS} &= \tilde{\Delta}_{PPQ\dot{Q}} = (1 + 2\theta + \tilde{\theta} + 2\tilde{\gamma} + 2\tilde{\Delta})2^{-3}. \end{aligned}$$

With appropriate substitutions the a \* a variance in the covariance of full-sibs is obtained,

$$\begin{aligned} \mathcal{C}_{FS}^* &\supset [2(\theta - \tilde{\Delta}) + (2 + \lambda^2)(1 - 2\theta + \tilde{\Delta} + \eta_a)/8 \\ &\quad + (\eta_d - \eta_a)/8 - \lambda(3\eta_a + \eta_d)/4] \sigma_{aa}^2. \end{aligned}$$

The contributions of joint dominance effects to  $\mathcal{C}_{HS}^*$  and  $\mathcal{C}_{FS}^*$  were given in Cockerham (1984a)

$$\mathcal{C}_{HS}^* \supset \frac{1}{4}\eta_a 2hh', \quad \mathcal{C}_{FS}^* \supset \frac{1}{2}\eta_a 2hh' + \frac{1}{4}\eta_d 2hh'.$$