Design and Analysis of Experiments on Random Drift and Inbreeding Depression

Michael Lynch

Department of Ecology, Ethology, and Evolution, University of Illinois, Champaign, Illinois 61820

Manuscript received November 16, 1987
Revised copy accepted July 7, 1988

ABSTRACT

While the genetic consequences of inbreeding and small population size are of fundamental importance in many areas of biology, empirical research on these phenomena has proceeded in the absence of a well-developed statistical methodology. The usual approach is to compare observed means and variances with the expectations of Wright's neutral, additive genetic model for quantitative characters. If the observations deviate from the expectations more than can be accounted for by sampling variance of the parameter estimates, the null hypothesis is routinely rejected in favor of alternatives invoking evolutionary forces such as selection or nonadditive gene action. This is a biased procedure because it treats sequential samples from the same populations as independent, and because it ignores the fact that the expectations of the neutral additive genetic model will rarely be realized when only a finite number of lines are studied. Even when genes are perfectly additive and neutral, the variation among the properties of founder populations, the random development of linkage disequilibrium within lines, and the variance in inbreeding between lines reduce the likelihood that Wright's expectations will be realized in any particular set of lines. Under most experimental designs, these sources of variation are much too large to be ignored. Formulas are presented for the variance-covariance structure of the realized within- and between-line variance under the neutral additive genetic model. These results are then used to develop statistical tests for detecting the operation of selection and/or inbreeding depression in small populations. A number of recommendations are made for the optimal design of experiments on drift and inbreeding, and a method is suggested for the correction of data for general environmental effects. In general, it appears that we can best understand the response of populations to inbreeding and finite population size by studying a very large number (>100) of self-fertilizing or full-sib mated lines in parallel with one or more stable control populations.

THE genetic consequences of inbreeding and small population size are of importance in many areas of biology. Population bottlenecks, through their influence on random genetic drift within loci and linkage disequilibrium between loci, are thought to play a major role in the speciation process (MAYR 1963; TEMPLETON 1980; but see BARTON and CHARLESWORTH 1984). The deleterious consequences of inbreeding are believed to be involved in the evolution of the mating systems of many plants and animals (CHARLESWORTH and CHARLESWORTH 1987). Inbreeding depression is also a serious concern in selective breeding programs (FALCONER 1981), in the maintenance of endangered species (SOULÉ 1986), and in the protection of human welfare (CAVALLI-SFORZA and BODMER 1971). Finally, the genetic stability of control lines is an implicit assumption in many experiments in population genetics (HILL 1972a–d), and uniformity of genetic stock is an essential requirement for many areas of biomedical research (FESTING 1979).

There is therefore a need for a statistical theory for the analysis of the dynamics of quantitative characters in finite populations. The expectations are already well-understood for the case of additive gene action (WRIGHT 1951). Under these circumstances, and in the absence of selection, the expected within- and between-population genetic variances, $\sigma^2_{gw}(0)(1 - F)$ and $2\sigma^2_{gw}(0)F$, are proportional to the genetic variation in the base population, $\sigma^2_{gw}(0)$, and to the expected degree of inbreeding, F. Although the within-population variance declines to zero and the betweenpopulation variance builds up to $2\sigma^2_{gw}(0)$, the expected mean phenotype over all populations remains stable in time.

ROBERTSON (1952) showed that the situation is more complicated with dominance. The presence of rare recessive alleles can cause an initial inflation of the genetic variation within populations and can alter the rate of divergence of population means. The limiting values of the within- and between-population genetic variances due to variation in the base population are the same as in the case of additivity, but there is an overall change in the mean genotypic value caused by inbreeding depression. Similar complications arise if there are epistatic interactions between loci (HILL 1982; GOODNIGHT 1987).

In contrast to the considerable attention that has

been given to the expected dynamics of neutral quantitative characters, studies on the variation around the expectations are relatively rare (BULMER 1976, 1980; AVERY and HILL 1977, 1979; WEIR, AVERY and HILL 1980; COCKERHAM and WEIR 1983; LYNCH and HILL 1986). The theory developed in these papers is essential for evaluating the consistency of observational data with expected patterns, but has had remarkably little influence on the design and analysis of empirical studies. Observed increases in the genetic variance within inbred populations are generally attributed to the dominance effect described by ROBERTSON (RASMUSON 1952; BRYANT, MCCOMMAS and COMBS 1986) although such changes are also possible with additive genes in populations with linked loci and/or variable pedigree structure. Many published studies on inbreeding depression exist in which there was inadequate or no control and no attention given to the nonindependence of sequential samples or the importance of genetic drift.

In this paper, methods are developed that allow tests of the null hypothesis that observed genetic changes in small populations are consistent with a neutral additive gene system. The general approach will be to assume that L independent replicate lines, each with expected effective size N_e, are isolated from a base population with additive genetic variance $\sigma^2_{gw}(0)$. The mean phenotypes, and the additive genetic variance within and between lines, are then monitored over $t = 0,k$ generations. These have expected values of $\tilde{\mu}_g(t)$, $\tilde{\sigma}^2_{gw}(t)$, and $\tilde{\sigma}^2_{\bar{g}}(t)$ respectively, but when only a finite number of lines is observed, the realized values $\mu(t)$, $\sigma^2_{gw}(t)$, and $\sigma^2_{\bar{g}}(t)$ will vary around the expectations from experiment to experiment. Due to imperfections in the estimation procedure, the observations $\bar{z}(t)$, $V_{gw}(t)$, and $V_{\bar{g}}(t)$ will also deviate from the true realized values somewhat. The main focus here is on variation in the realization of the process of random genetic drift rather than on the variance of parameter estimates caused by sampling error on the part of the investigator. The first source of variation (realization variance) is a function of population genetic structure and, for a fixed system of mating, is largely beyond the control of the investigator, while the second (sampling variance) can at least be minimized by the use of large sample sizes. Expressions for the sampling variance of population parameters are readily available in textbooks of quantitative genetics, and the two sources of error can be treated as independent and additive. To simplify the presentation, a balanced experimental design will be assumed throughout.

CORRECTING FOR ENVIRONMENTAL TRENDS

A potential source of error in the analysis of unselected lines is the presence of general environmental effects that cause the genotypic mean and/or variance to shift between generations. Even in the most carefully designed laboratory experiments, there are many uncontrolled sources of variation, including unconscious shifts in the behavior of the investigator, and these may obscure the genetic interpretation of phenotypic observations in many different ways. For example, a directional environmental trend that influences all individuals in the same manner can lead to the erroneous conclusion that directional selection or inbreeding depression is operating. If genotype × environment interaction is present, the mean phenotypes of different lines will vary in response to general environmental effects, and in extreme situations, the direction of response may differ between lines. Finally, if the sensitivity to environmental effects increases with inbreeding (LERNER 1954; FALCONER 1981), genotype × environment interaction can inflate the apparent rate of divergence of mean phenotypes.

While many empirical studies on random drift and inbreeding lack controls, those that have employed them often suggest parallel trends between controls and experimental lines. WRIGHT's (1977) inbreeding experiments with guinea pigs are especially dramatic in this respect. Clear evidence for the development of genotype × environment interaction with inbreeding has arisen in experiments with corn (OBILANA and HALLAUER 1974; BARTUAL and HALLAUER 1976) and with *Tribolium* (BRAY, BELL and KING 1962). Thus, the need for controls and a technique for utilizing the information they provide is very real.

In the case of selection experiments, the usual approach to removing general environmental effects has been to subtract the mean of a contemporaneous control from the mean of the selected population. An implicit assumption of this treatment is that both the control and selected populations respond in the same manner to general environmental effects; *i.e.*, there is no genotype × environment interaction. Moreover, as HILL (1972b) has pointed out, this procedure actually can obscure the genetic response of the selected population if the control mean is subject to substantial sampling error. An alternative approach, suggested by MUIR (1986a, b), is to treat the control line(s) as a covariate. This has the advantage of allowing for genotype × environment interaction of an arbitrary level.

There are two important considerations in the choice and employment of controls. First, it is essential that the control line is maintained in such a way that phenotypic changes between generations are caused solely by general environmental effects. This can be accomplished by using clonally propagated genotypes, highly inbred strains, or a large outbred population. In the latter case, precautions have to be made to prevent evolution of the mean phenotype by natural selection. Second, the control should provide a strong

signal of the general environmental effect, *i.e.*, explain a maximum amount of the variance of the mean phenotypes of the experimental line. Since highly inbred lines sometimes have enhanced environmental sensitivity, they might fulfill this criterion provided their response to the environment is highly correlated with that of the experimental lines (MUIR 1986b).

Let the observed mean of experimental line i in generation t be

$$\bar{z}(t,i) = \tilde{\mu}_g(0) + \Delta\bar{g}(t,i) + e_g(t,i) + e_s(t,i) \quad (1)$$

where $\Delta\bar{g}(t,i)$ = the cumulative change in the mean genotypic value up to generation t, and $e_g(t,i)$ and $e_s(t,i)$ refer to deviations of the observed phenotypic mean from $[\tilde{\mu}_g(0) + \Delta\bar{g}(t,i)]$ caused by general and special environmental effects, the latter including measurement error. For the control line, the mean genotypic value is assumed to remain constant during the experiment, so that

$$\bar{z}(t,c) = \bar{z}(c) + e_g(t,c) + e_s(t,c) \quad (2)$$

where $\bar{z}(c)$ is the mean phenotype of the control averaged over the entire experiment. If the control is replicated, then the elements of this equation refer to averages over all replicates.

Since the general environmental effects are the only correlated components of the control and experimental line means, a partial regression of $\bar{z}(t,i)$ on $\bar{z}(t,c)$ provides a way of factoring out the general environmental effects,

$$\bar{z}(t,i) = a(i) + b'(i)\bar{z}(t,c) + d(i)t + e(t,i). \quad (3)$$

Because of the sampling error of the control means, $b'(i)$ provides a slightly biased estimate of the parameter $\beta(i) = \sigma[e_g(i),e_g(c)/\sigma^2[e_g(c)]$. An unbiased estimator is

$$b(i) = b'(i)\left[1 + \frac{\sigma^2[e_s(c)]\sigma^2(t)}{\sigma^2(t)\sigma^2[e_g(c)] - \sigma^2[t,e_g(c)]}\right] \quad (4)$$

where $\sigma^2[e_s(c)]$ is the sampling variance of the control mean within generations, $\sigma^2[t,e_g(c)]$ is the squared covariance of $\bar{z}(t,c)$ and t, and $\sigma^2[e_g(c)]$ is the variance of control means between generations in excess of the sampling variance. This improved estimator may be employed by substituting the observed variance and covariance components. The corrected line means are then estimated by

$$\bar{z}^*(t,i) = \bar{z}(t,i) - b(i)[\bar{z}(t,c) - \bar{z}(c)]. \quad (5)$$

$d(i)$ in Equation 3 is an estimate of the rate of evolution of the corrected mean phenotypes.

As a consequence of inbreeding and drift, the relative response of an experimental line to general environmental effects may change throughout the course of an experiment, in which case the correction factor would need to be a function of time. The occurrence of such change could be examined with the model

$$\bar{z}(t,i) = a(i) + [b'(i) + g(i)t]\bar{z}(t,c) + d(i)t + e(t,i), \quad (6)$$

$g(i)$ providing an estimate of the development of genotype × environment interaction with time. The use of several genetically unique controls will increase the information on general environmental effects for experimental lines and can be implemented by the addition of the appropriate terms to the prevoius formulas. Whatever the approach, it is important to realize that each experimental line may develop its own unique response to general environmental effects and should be corrected independently of all other lines.

As an example of the application of the technique outlined above, the results of an inbreeding experiment with *Drosophila melanogaster* will be considered. Starting from a large base population, KIDWELL and KIDWELL (1966) extracted 20 lines and maintained each by single full-sib matings through 20 generations. The base population was also maintained throughout the experiment, and the progeny of four single pair matings from it served as a control for each generation. An undisclosed number of individuals were assayed for abdominal bristle number and body weight at irregular intervals. There is a general upward trend in bristle number in both the control and experimental lines throughout the experiment, and the dynamics of mean body weight are also roughly concordant between groups (Figure 1). Based on this visual comparison, the authors concluded that the lines were strongly influenced by general environmental effects but not by inbreeding depression.

A more explicit test of this hypothesis can be made by regressing the means of inbred lines on the control means and the inbreeding coefficient. (For small populations, inbreeding depression is expected to scale linearly with F, not t.) For bristle number, the interaction term of Equation 6 was not significant, but the fitted coefficients for Equation 3 are $b' = 0.50 \pm 0.25$, and $d = 2.34 \pm 1.40$ ($r^2 = 0.85$). Thus, the experimental lines were influenced by the same environmental effects as the control, but on average, were only 50% as responsive. Due to insufficient information, b' could not be corrected for sampling bias, but this is not expected to be large. The corrected values of the experimental lines, obtained from Equation 5, are given in Figure 1. There is an expected genetic gain of approximately two bristles under complete inbreeding. The same analysis applied to body weight again indicated that the inbred lines were slightly less responsive to general environmental effects than the control ($b' = 0.77 \pm 0.14$). The partial regression on F ($d = -9.2 \pm 34.8$) was not significant, consistent with the conclusion that there was no inbreeding depression for body weight. Eighty-seven percent of

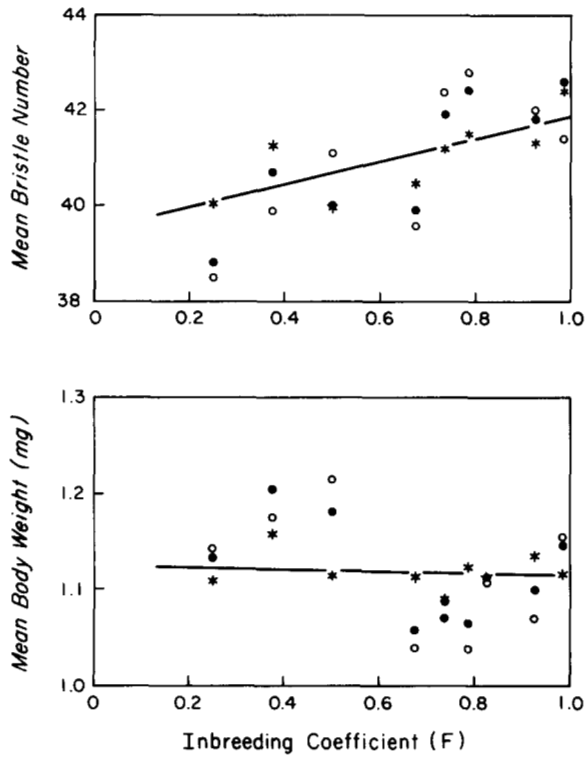

FIGURE 1.—Observed phenotypic means in inbred (closed circles) and control (open circles) lines of *D. melanogaster*, and corrected values for the inbred lines (stars) obtained by use of Equation 5. The partial regressions are represented by the solid lines. Data are from KIDWELL and KIDWELL (1966).

the temporal variance of mean body weight in the experimental lines was accounted for by variation in the controls.

In addition to their influence on mean phenotypes, general environmental effects may also cause spurious, temporal fluctuations in the components of variance. In principle, the procedures outlined above can be extended to the correction of variance components. However, since the sampling error of variances tends to be very large, reliable corrections of this sort will require large sample sizes.

At least in the case of plants, there may be a way of avoiding all of the above statistical procedures. Provided they are kept cool and dry, seeds can usually be stored for many years. This allows one to grow members of all generations in a randomized design simultaneously (RUSSELL, SPRAGUE and PENNY 1963; HALLAUER and SEARS 1973; CORNELIUS and DUDLEY 1974). Even in this case, however, special precautions need to be taken to ensure that phenotypic expression is not influenced significantly by the duration of seed storage or by properties of the seed that may be conditioned by general environmental effects experienced by the maternal plant.

DIVERGENCE OF MEAN PHENOTYPES

As outlined in the introduction, a simple prediction of population genetic theory is that the divergence of mean genotypic values among populations is proportional to the degree of inbreeding within populations. This result is expected only for neutral quantitative traits with a purely additive genetic basis, and even then, it begins to break down with the accumulation of new mutations (CHAKRABORTY and NEI 1982; LYNCH and HILL 1986). There are more subtle assumptions embedded in the theory as well, including the condition that the mode of gene action remains stable with a change in genetic background. Even in the ideal case, the realized between-line variance will be distributed around the expectation $2\sigma^2_{gw}(0)F(t)$. Therefore, in order to evaluate the consistency of observations with the neutral, additive gene theory, a statistical description of the between-line variance is required.

The usual protocol in genetic drift experiments is to maintain several independent lines under the same experimental conditions and with the same mating system. Let each line be initiated simultaneously with N_m males and N_f females randomly extracted from the same base population. Each generation and within each line, N_m males are mated randomly to N_f/N_m females, and n offspring are measured from each full-sib family. The phenotype of the kth offspring of the mating between male i and female j may be represented as

$$z_{ijk} = \frac{g_{mi} + \Delta g_{mik} + g_{fj} + \Delta g_{fjk}}{2} + c_j + e_{ijk} \qquad (7)$$

where g_{mi} and g_{fj} are the additive genetic values of the parents, $(\Delta g_{mik} + \Delta g_{fjk})/2$ is the deviation of the offspring from the midparent additive genetic value caused by segregation, c_j is the common environmental effect of female j, and e_{ijk} is the special environmental effect. (Common environmental effects above those caused by maternal environment occur sometimes, but here they are assumed to be unimportant.)

Because of sampling error of the founder phenotypes, the initial variance of line mean phenotypes has the expectation

$$\tilde{\sigma}^2_z(0) = \frac{\tilde{\sigma}^2_{gw}(0) + \sigma^2_c + \sigma^2_e}{N_m + N_f}, \qquad (8)$$

assuming that none of the cofounders are full-sibs, where σ^2_c and σ^2_e represent the variance of common and special environmental effects. In the following generations, the means are based on nN_f measures, and account must be taken of the genetic and common-environment covariance between full-sibs as well as of the segregational variance within full-sib families (HILL 1972c). Letting $N_e = 4N_mN_f/(N_m + N_f)$, $F(t) = 1 - [1 - (1/2N_e)]^t$, and $\tilde{\sigma}^2_g(0) = \tilde{\sigma}^2_{gw}(0)/(N_m + N_f)$, the

expected variance of mean phenotypes for $t \geq 1$ is

$$\tilde{\sigma}_{\bar{z}}^2(t) = [\tilde{\sigma}_{\bar{g}}^2(0) + 2\tilde{\sigma}_{gw}^2(0)F(t)] + \frac{1}{nN_f}\left[\frac{\tilde{\sigma}_{gw}^2(0)[1 - F(t-1)]}{2} + n\sigma_c^2 + \sigma_e^2\right]. \quad (9)$$

The first term in this formula, which represents the true dispersion of line means, is cumulative over generations, while the remaining terms refer to the variance due solely to finite sample size within lines. Because of the genetic continuity of populations in time, there is an expected covariance between the mean phenotypes in the same line in subsequent generations,

$$\tilde{\sigma}_{\bar{z}}(0,t) = \tilde{\sigma}_{\bar{g}}^2(0) \qquad \text{for } t > 0, \quad (10)$$

$$\tilde{\sigma}_{\bar{z}}(t,t') = \tilde{\sigma}_{\bar{g}}^2(0) + 2\tilde{\sigma}_{gw}^2(0)F(t) \qquad \text{for } 1 < t < t'. \quad (11)$$

It can be seen from the preceding formulas that the contribution of the segregational and special environmental effects variance to the variance of mean phenotypes is inversely proportional to the total sample size (nN_f), whereas the contribution of common environmental effects is inversely proportional to the number of full-sib families (N_f). In genetic experiments, it is desirable to remove these sources of variation in order to obtain an estimate of the variance of mean genotypic values unbiased by sampling error. Normally, this can be accomplished by manipulating the mean squares of a nested analysis of variance. The within- and between-family components of variance can be isolated from the genetic variance between lines by letting $V_{\tilde{g}}(t) = [MS_{\text{lines}} - MS_{\text{fam(line)}}]/nN_f$ be the estimate of $\sigma_{\bar{g}}^2(t)$. A slight problem arises if the lines consist of single families, as in the case of selfing and full-sib mating, since the common environmental effects variance cannot be partitioned from the variance of genotypic means. This problem can be eliminated by temporarily expanding each line into S families prior to analysis and substituting S for N_f in the expression for $V_{\tilde{g}}(t)$ (LYNCH 1984).

Since the number of lines, L, employed in experiments is usually rather small, it is of practical importance to have expressions for the sampling variances and covariances of the realized variances of line means. Starting from the same base population, suppose that an infinite number of divergence experiments, identical in all respects except the realization of the drift process, could be run. At time t, each experiment will have developed a level of between-line genetic variance, $\sigma_{\bar{g}}^2(t)$. Variation will arise among the $\sigma_{\bar{g}}^2(t)$ because of variation in the within-population genetic variance among the founder populations, variance in inbreeding that develops among the lines, and the observation of a finite number of lines. Moreover, since the line means are a function of their past history, the between-line variances for any particular experiment will be correlated in time. With finite sample sizes, it is also necessary to account for the genetic and environmental variance within and between families and the covariance between the cumulative drift variance and the segregation variance.

Expressions for the variance and covariance of between-line variance can be obtained by assuming that measurements have been taken on a scale on which the genetic and environmental effects are normally and independently distributed. In that case, the expected variance of a variance component is twice the expected variance squared divided by the sample size minus one, and approximate expectations for $\sigma^2[\sigma_{\bar{z}}^2(t)]$ and $\sigma[\sigma_{\bar{z}}^2(t), \sigma_{\bar{z}}^2(t')]$ can be obtained by Taylor expansion (APPENDIX).

Equations A1–A3 have been written so that the variance and covariance due to finite sample size are described by the terms within the large brackets. For $t \geq 1$, these terms decline to zero as the sample size within lines (nN_f) increases. The earlier terms in the formulas describe the variance and covariance of the true between-line genetic variance, and for a given effective population size, can be reduced only by increasing the number of lines. Thus, ignoring the variance of inbreeding, the squared coefficient of variation of $\sigma_{\bar{g}}^2(t)$ can be seen to be on the order of $2/(L-1)$, and depending on the variation in inbreeding and environmental effects, it can be considerably greater. This implies that studies of phenotypic divergence need to be very large to be statistically reliable. For example, if it is desirable to reduce the standard error of the between-line variance to 10% of the expectation, approximately 200 lines would need to be sampled (400 and 300 in the case of selfing and full-sib lines).

Of additional concern is the variation among estimates of $\tilde{\sigma}_{\bar{g}}^2(t)$ caused by variation in inbreeding (WEIR, AVERY and HILL 1980; COCKERHAM and WEIR 1983). Under most mating schemes, some individuals mate by chance with closer relatives than do others. This results in variation in F among members of the same population, and because of sampling, accumulates as between-population variance in average inbreeding. Variance in inbreeding is of special interest because it cannot usually be estimated from empirical data. Detailed pedigree records may be possible under some experimental protocols, but uncertainties regarding paternity are common. Moreover, the linkage relationships of constituent loci, which influence σ_F^2, are unknown for virtually all quantitative characters.

Ignoring the variance among founder means, the squared coefficient of variation of $\sigma_{\bar{g}}^2(t)$ is $\{2[1 + (N_m + N_f)^{-1}] + [\sigma_F^2(t)/F^2(t)]\}/(L-1)$. Thus, roughly speaking, the variation in inbreeding between lines is of potential concern if $\sigma_F^2(t)/F^2(t)$ is of the order of 0.2

TABLE 1

Values of the squared coefficient of variation of mean line inbreeding, $\sigma_F^2(t)/F^2(t)$

		Free recombination					0.1 Morgan chromosome				
	Generation	2	4	8	16	32	2	4	8	16	32
$N = 1$	MS	0.00	0.00	0.00	0.00	0.00	0.10	0.03	0.00	0.00	0.00
$N = 2$	MS	0.03	0.02	0.00	0.00	0.00	0.08	0.09	0.03	0.00	0.00
	ME, DR, DH	0.00	0.00	0.00	0.00	0.00	0.45	0.28	0.18	0.14	0.09
$N = 4$	MS	0.02	0.02	0.01	0.00	0.00	0.05	0.10	0.09	0.03	0.00
	ME	0.06	0.02	0.00	0.00	0.00	0.22	0.19	0.14	0.05	0.01
	DR	0.14	0.03	0.01	0.00	0.00	0.38	0.22	0.14	0.05	0.01
	DH	0.54	0.08	0.02	0.00	0.00	1.02	0.34	0.17	0.06	0.01
$N = 16$	MS	0.01	0.01	0.01	0.00	0.00	0.01	0.04	0.07	0.09	0.06
	ME		0.01	0.01	0.00	0.00	0.11	0.07	0.10	0.10	0.07
	DR	0.11	0.02	0.01	0.00	0.00	0.11	0.09	0.11	0.10	0.07
	DH	1.00	0.14	0.04	0.01	0.00	1.44	0.36	0.19	0.13	0.08
$N = 64$	MS	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.04	0.05
	ME			0.00	0.00	0.00			0.04	0.05	0.06
	DR			0.00	0.00	0.00			0.04	0.05	0.06
	DH	0.60	0.20	0.04	0.01	0.00	1.00	0.50	0.16	0.09	0.07

Obtained from data in Table III of WEIR, AVERY and HILL (1980). Mating is random in all four mating schemes: MS = ideal monoecious population including random selfing, ME = monoecy with selfing excluded, DR = dioecy with each offspring produced by a random pairing of male and female gametes, sex ratio assumed to be 1:1, DH = monogamous, dioecious population. In the case of a 0.1 Morgan chromosome, the loci are assumed to be randomly distributed. Slight errors due to rounding may be present. Values under MS were obtained by use of Equation 27 of WEIR and COCKERHAM (1969).

or more. Since the variance in inbreeding is a function of several high-order identity-by-descent measures (WEIR and COCKERHAM 1969), its computation is not a simple matter. Fortunately, WEIR, AVERY and HILL (1980) have published values of $\sigma_F^2(t)$ for a range of mating systems and population sizes. These are converted to estimates of $\sigma_F^2(t)/F^2(t)$ in Table 1.

For freely recombining loci, the variance in inbreeding is zero under those mating schemes in which the pedigree structure is fixed [obligate self-fertilization, full-sib mating, the special systems of mating of WRIGHT (1921), and the circular systems of mating of KIMURA and CROW (1963)] and is of minor importance when there is random pairing of gametes. However, if the sexes are separate and matings are monogamous, $\sigma_F^2(t)/F^2(t)$ can be large enough to be of concern in the first 2–4 generations. Linkage inflates the variance in inbreeding under all systems of mating by causing positive correlations in identity by descent at loci in the same individual. But even if most pairs of loci are very tightly linked, $\sigma_F^2(t)$ can be considered to be of negligible significance after 6 or so generations have passed. If lines are maintained by self-fertilization or full-sib mating there is little reason for concern with $\sigma_F^2(t)$ in any generation.

The preceding theory leads to several recommendations for the design and analysis of experiments on the consequences of small population size for the between-population variance. First, if at all possible, one or more contemporaneous control lines should be maintained so that the estimated mean phenotypes of the experimental lines can be adjusted for general environmental effects. Second, effort should be made to remove the contribution of common (maternal) environmental effects and other sources of within-population variance from the estimates of $\sigma_{\bar{g}}^2(t)$. Even when such corrections can be made, a great deal of confidence should not be placed on the results of the first couple of generations of inbreeding. Thereafter, the sampling variance of the between-line variance under the assumption of the neutral additive gene model may be taken to be approximately $2V_{gw}^2(0)\{4F^2(t)[1 + (N_m + N_f)^{-1}] + (N_m + N_f)^{-2}\}$ plus the variance due to the estimation procedure.

For fixed resources that allow the monitoring of N_eL individuals/generation, the efficiency of estimation of $\tilde{\sigma}_{\bar{g}}^2(t)$ is maximized by making the lines as small as possible—selfing in the case of self-compatible species, full-sib mating in the case of dioecy. Both extreme forms of mating have additional advantages. First, any desired amount of inbreeding is attained in a minimum amount of time. Second, except in the case of extremely strong linkage, the variance in inbreeding among lines can be ignored in all generations. If it is desirable to study the effects of different population sizes, the maximum avoidance of inbreeding schemes of WRIGHT (1921) or the circular designs of KIMURA and CROW (1963) are recommended since they eliminate most of the variation in inbreeding. These, however, have the side effect of at least doubling the effective population size relative to the actual population size and of postponing the generation in which inbreeding begins.

LANDE's test for the selective divergence of mean phenotypes: As a test for natural selection, LANDE (1977) suggested the use of the statistic $\theta = V_{\bar{g}}(t)/[t\bar{V}_{gw}/$

N_e] where $\bar{V}_{gw}$ is the average observed additive genetic variance within lines over t time units of isolation. The numerator and denominator of θ are estimates of the observed and expected between-line variance under the hypothesis of neutral, additive genes (assuming $t < N_e$). LANDE argued that, for a normal sampling distribution of population means, θ will be F-distributed under the null hypothesis of random genetic drift. In terms of quantities described above, application of Fisher's F test to this statistic assumes that the numerator is χ^2-distributed with expectation $2\tilde{\sigma}^2_{gw}(0)F(t)$ and variance $2[2\tilde{\sigma}^2_{gw}(0)F(t)]^2/(L-1)$. Equation A2 indicates that this is asymptotically true for large populations provided the between-line variance in inbreeding is negligible. However, for very small populations, which are often employed in genetic drift experiments, the variance of $V_{\bar{g}}(t)$ is greater than that expected under a χ^2 distribution—at least twice as great in the case of self-fertilization, and at least 1.5 times as great with full-sib mating. Thus, when $(N_m + N_f)$ is very small or when substantial variation in inbreeding is likely to have occurred, the treatment of θ as Fisher's F may cause a substantial probability of inadvertantly rejecting the null hypothesis of neutral, additive genes. BRYANT, COMBS and MCCOMMAS (1986) relied on LANDE's test to reject this hypothesis after putting populations of houseflies through single-generation bottlenecks of 2 to 16 pairs. For the above reasons, and because of nonindependence and possible nonnormality of the characters they studied, the confidence level of their rejection has perhaps been overstated.

REGRESSIONS INVOLVING PARAMETER ESTIMATES

It is common procedure to regress the phenotypic means of populations on time to test for the operation of stabilizing or directional selection (CHARLESWORTH 1984; MANLY 1985). Even when environmental trends can be ruled out, such a treatment of data raises certain difficulties in small populations since drift can give rise to directional changes in mean phenotypes within individual lines. Standard statistical tests for the significance of a regression coefficient are inappropriate for two reasons: the mean phenotypes estimated from the same line at different times are not independent, and the sampling variances of the means decline in time as a consequence of the loss of genetic variation within finite populations. HILL (1972a, b) has dealt with the first difficulty in the context of regressions of selection response on selection differential, but as pointed out by FELSENSTEIN (1985), the problem of nonindependence is almost always ignored in evolutionary analyses.

Suppose the mean phenotype pooled over L lines has been evaluated over $(k + 1)$ consecutive generations. The standard least-squares expression for the regression coefficient is

$$b_{\bar{z}t} = \sum_{t=0}^{k} [\bar{z}(t) - \bar{z}](t - \bar{t}) / \sum_{t=0}^{k} (t - \bar{t})^2 \qquad (12)$$

where $\bar{z}$ is the mean phenotype averaged over all lines and generations. The expected value of $b_{\bar{z}t}$ is zero under the assumption of neutral additive genes. The expected sampling variance of $b_{\bar{z}t}$ may be written as

$$\sigma^2(b_{\bar{z}t}) = \left[\frac{12}{k(k+1)(k+2)}\right]^2 \sum_{t=0}^{k} \sum_{t'=0}^{k} (t - \bar{t})(t' - \bar{t}) \tilde{\sigma}_{\bar{z}}(t,t')/L \qquad (13)$$

under the assumption that the sampling variance of the grand mean ($\bar{z}$) is negligible. Substitution of Equations 9–11 shows that the sampling variance of $b_{\bar{z}t}$ is attributable to four causes: the genetic variance among initial line means, the variance and covariance of genotypic means resulting from drift, the sampling variance of means due to segregational variance, and the sampling variance of means due to environmental effects. While Equation 13 applies to the special case in which means are available for k + 1 consecutive generations, the entire approach can be generalized to situations in which means are missing for some generations. This requires only that the proper variance and covariance expressions be substituted for the $\tilde{\sigma}^2_{\bar{z}}(t,t')$ in Equation 13.

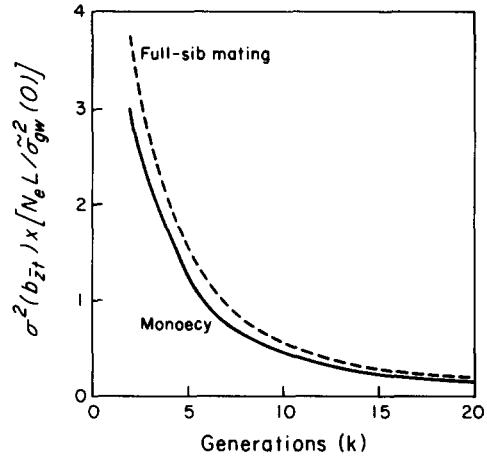

FIGURE 2.—Minimum sampling variance for the regression of phenotypic means on generation number for a neutral quantitative character in a finite population, shown for increasing numbers of consecutive generations. Note that regression analysis requires that $k \geq 2$. Measurement error of the means is assumed to be negligible. The actual, expected sampling variance is obtained by multiplying the plotted values by $\tilde{\sigma}^2_{gw}(0)/N_e L$. Results are given for ideal monoecious populations and for full-sib mating.

Figure 2 illustrates the relationship of $\sigma^2(b_{\bar{z}t})$ to the experimental duration (k) for ideal monoecious populations as well as for full-sib mating ($N_e \simeq 2.5$) computed with Equation 13. The terms describing measurement error are assumed to be negligible, and the first generation in the regression is taken to be the

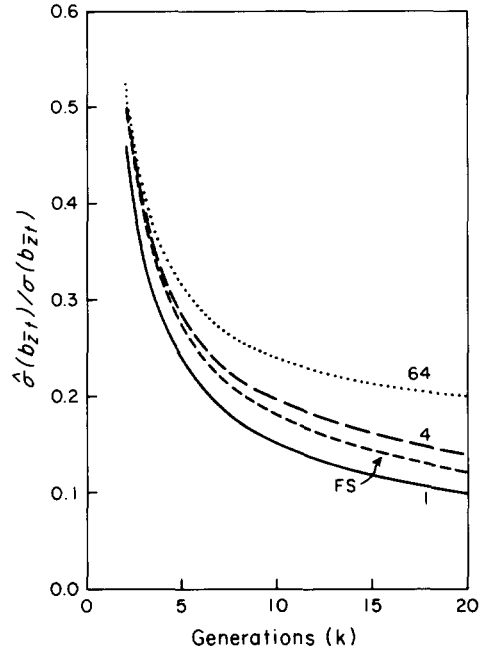

FIGURE 3.—Standard error for the regression coefficient $b_{\bar{z}t}$ under the assumption of bivariate normality and independent sampling relative to the true expectation. Results are given for ideal monoecious populations of three effective sizes and for full-sib mating.

offspring of the founders so that the large sampling variance of the founder means can be avoided. In that case, $\sigma^2(b_{\bar{z}t})$ is directly proportional to the genetic variation in the base population and inversely proportional to N_eL. When means are available for only the first three generations ($k = 2$), $\sigma^2(b_{\bar{z}t})$ is not less than $3\tilde{\sigma}^2_{gw}(0)/N_eL$. With increasing numbers of generations, it declines exponentially. By dividing the plotted values by N_eL and taking the square root, it is possible to gain some appreciation of the magnitude of regression coefficients that are compatible with random genetic drift. For example, when data are available for 10 generations for a single full-sib mated line, $\sigma^2(b_{\bar{z}t})/\sigma^2_{gw}(0) \simeq 0.67/2.5 = 0.27$. The standard error of $b_{\bar{z}t}$ in units of initial genetic standard deviations is therefore at least 0.5. Since the presence of environmental variation and finite sample size can cause further error, in this case a regression coefficient within the range $\pm\ \tilde{\sigma}_{gw}(0)$ certainly would be compatible with a neutral hypothesis.

It is instructive to examine the bias in the sampling variance that would arise if one were to rely on the standard expression, $\hat{\sigma}^2(b_{\bar{z}t}) = \tilde{\sigma}^2_{\bar{z}}(1 - \rho^2)/(k + 1)\sigma^2_t$ which is obtained under the assumption of independent sampling and bivariate normality. Under the neutral hypothesis, the expected correlation ρ between $\bar{z}(t)$ and t is zero, and for $k + 1$ consecutive samples, $(k + 1)\sigma^2_t = k(k + 1)(k + 2)/12$. The sampling variance of the means over the experiment, $\tilde{\sigma}^2_{\bar{z}}$, can be expressed in terms of Equations 9–11.

Ratios of the standard errors, $\hat{\sigma}(b_{\bar{z}t})/\sigma(b_{\bar{z}t})$, for the case in which measurement error is negligible and consecutive means are available starting with the offspring of the founders, are given in Figure 3. The bias in the traditional estimator of the standard error of a regression coefficient is clearly too large to be ignored. Initially, $\sigma(b_{\bar{z}t})$ is on the order of twice $\hat{\sigma}(b_{\bar{z}t})$, and this factor increases severalfold with increasing numbers of generations.

Tests for inbreeding depression: While the procedures outlined above provide a means of evaluating whether a temporal sequence of mean phenotypes is consistent with the neutral additive gene hypothesis, a rejection of the null hypothesis should not be misconstrued as acceptance of the hypothesis of natural selection. When nonadditive interactions exist within and/or between loci, inbreeding can cause a shift in mean phenotypes in the absence of counterbalancing selection. The most common experimental design employed in the detection of inbreeding depression is to subject a series of isolated lines to a regular program of inbreeding. The consecutive line means are then regressed on the expected inbreeding coefficient,

$$b_{\bar{z}F} = \sum_{t=0}^{k} [\bar{z}(t) - \bar{z}][F(t) - \bar{F}] / \sum_{t=0}^{k} [F(t) - \bar{F}]^2 \quad (14)$$

where $\bar{F}$ is the mean expected inbreeding coefficient over the experiment. Such an analysis suffers from the same difficulties noted previously. The mean phenotypes obtained from consecutive samples of the same line are not independent. Moreover, the distribution of $F(t)$ is highly skewed, eventually piling up at values very close to one.

The sampling variance of $b_{\bar{z}F}$ under the neutral additive model can be computed by use of the procedures outlined above. Figure 4 provides the results for the special case in which environmental sources of variance are of negligible importance and the analysis begins with the founder ($F = 0$) generation. When viewed as a function of the expected inbreeding in the final experimental generation, $\sigma^2(b_{\bar{z}F})$ depends very little on the effective population size. However, with larger N_e, it takes longer to reach a given degree of inbreeding, and hence in early generations the results from selfed or full-sib mated lines are much more reliable than those from larger lines. The sampling variance of $b_{\bar{z}F}$ is very high if the cumulative inbreeding is less than 0.25, and diminishes to a minimum of approximately $2.3\tilde{\sigma}^2_{gw}(0)/L$ once inbreeding has proceeded beyond $F = 0.9$. Thus, relatively large departures from the expectation $b_{\bar{z}F} = 0$ can arise in inbreeding experiments even in the absence of dominance. Suppose, for example, that full-sib mating is performed on 10 lines for 10 generations ($F = 0.86$). The sampling variance of $b_{\bar{z}F}$ is then approximately $\tilde{\sigma}^2_{gw}(0)/4$. An observed $|b_{\bar{z}F}| \leqq \tilde{\sigma}_{gw}(0)$ clearly would be consistent with an additive gene model.

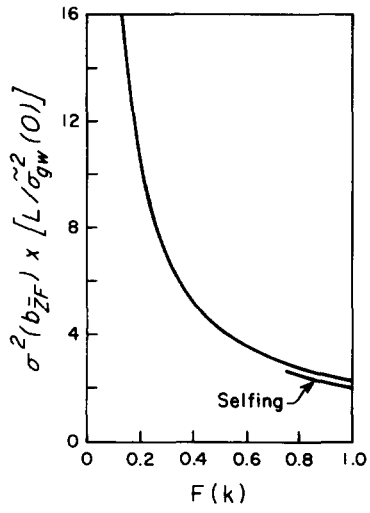

FIGURE 4.—Sampling variance of the regression of mean phenotype on the expected inbreeding coefficient under the assumption of neutral, additive gene action. The major curve applies to full-sib mating and all cases of monoecy except selfing. The plotted values are minimum estimates as they do not include variance from environmental variation or finite sample size. $F(k)$ is the expected level of inbreeding in the final generation of an experiment. The actual, expected sampling variance is obtained by multiplying the plotted values by $\tilde{\sigma}^2_{gw}(0)/L$.

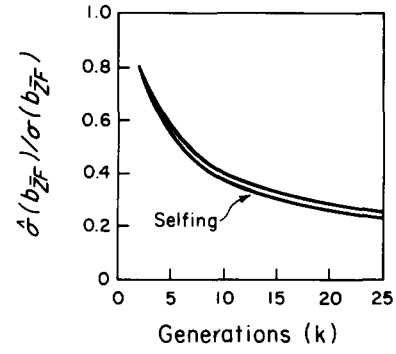

FIGURE 5.—Ratio of the standard error of $b_{\bar{z}F}$ based on normal regression theory relative to the expectation under random genetic drift. The upper curve applies to dioecious and monoecious populations with $N_e > 1$.

It is again useful to consider the bias that is incurred by using the standard expression for the variance of a regression coefficient as has been done in existing studies. Using the approach outlined above and focusing on the special case in which environmental sources of variance can safely be ignored, it is seen in Figure 5 that the bias depends primarily on the duration of the experiment and very little on population size. The usual standard error, $\hat{\sigma}(b_{\bar{z}F})$, always underestimates the preferred measure $\sigma_{b_{\bar{z}F}}$. The bias increases with the experimental duration, asymptoting at $\hat{\sigma}_{b_{\bar{z}F}}/\sigma_{b_{\bar{z}F}} \simeq 0.25$ beyond 20 generations.

As a rough check of the validity of conclusions on inbreeding depression derived from regression analysis, the statistic $J = |b_{\bar{z}F}|/2[\phi V_{gw}(0)/L]^{1/2}$, where ϕ represents the plotted values in Figure 4, may be useful. If J exceeds one, the observed regression coefficient deviates from zero by more than two estimated standard errors, and one is justified in suspecting the presence of inbreeding depression. Of course, the true standard error of $b_{\bar{z}F}$ cannot be known with certainty since the additive genetic variance in the base population is an estimate. The expectation of $[\phi V_{gw}(0)/L]^{1/2}$ will also be less than the true standard error of $b_{\bar{z}F}$ since measurement error has been ignored, but the bias should be small if the number of families and offspring within families assayed is large.

Estimates of $b_{\bar{z}F}/V^{1/2}_{gw}(0)$ are given for several species and characters in Table 2. Not all of the reported experiments were designed like the scenario presented previously. However, almost all of the values of $b_{\bar{z}F}/V^{1/2}_{gw}(0)$ are in excess of one and three exceed five. Most of the data sets to which the J statistic may be applied are in strong agreement with the inbreeding depression hypothesis. For example, for yield in corn, $L = 248$, $\phi = 2.1$, and $b_{\bar{z}F}/V^{1/2}_{gw}(0) = -3.27$, yielding $J = 17.8$. However, it is questionable whether there is inbreeding depression for thorax length in *Drosophila* ($J = 0.9$), offspring number in *Tribolium* ($J = 1.2$), and internode length in barley ($J = 0.2$).

The magnitude of inbreeding depression suggested by Table 2 is fully compatible with a Mendelian model without epistatic interactions. For single loci, $|b_{\bar{z}F}|/\tilde{\sigma}_{gw}(0) > 1$ will arise with complete dominance when recessive alleles have frequency <0.3, and with overdominance, this condition is met over a broad range of gene frequencies (Figure 6). In theory, there is no upper bound to $|b_{\bar{z}F}|/\tilde{\sigma}_{gw}(0)$ since in the case of overdominance there is always a gene frequency at which there is no additive genetic variation, all of it appearing in the dominance component.

In closing this section, another popular method of testing for inbreeding depression should be mentioned. Frequently, the mean phenotypes from a single generation of a control and contemporaneous inbred population are compared by use of a standard t test or analysis of variance. This procedure seems indefensible since the expected variance of an inbred population exceeds that of the control. Moreover, the differences that can arise between control and inbred line means as a consequence of random drift as opposed to inbreeding depression are ignored.

A simple modification of the t test takes these problems into account. Suppose that n offspring are monitored from each of L independent random-mated mothers and from L independent consanguineously mated mothers, all derived from the same base population. Under the null model of neutral additive gene action, the difference between the mean phenotypes of the two types of progeny ($\Delta\bar{z}$) has zero expectation. The observed difference must be evaluated against the standard error of the difference caused by sam-

TABLE 2

Results from various inbreeding experiments involving the regression, $b_{\bar{z}F}$, of mean phenotype on expected inbreeding coefficient

Species	Character	Ref.	L	N_e	$F(k)$	$b_{\bar{z}F}$	$b_{\bar{z}F}/V_{gw}^{1/2}(0)$
Mouse	Litter size	1, 2	10–6	2.5	0.63	−5.10	−3.64
		3	20	2.5–32	0.95	−3.86	−2.74
	3-week weight (g)	3	20	2.5–32	0.95	−2.54	−3.25
	8-week weight (g)	3	20	2.5–32	0.95	−4.94	−2.63
	Postweaning gain (g)	3	20	2.5–32	0.95	−2.18	−1.77
Sheep	Clean fleece weight (lb)	4	–	–	0.28	−4.4	−6.07
	Staple length (cm)	4	–	–	0.28	−1.2	−1.80
	Body weight (lb)	4	–	–	0.28	−29.1	−6.29
D. melanogaster	Wing length (μm)	5	20	2.5–12	0.98	−34.8	−1.77
		6	10	2.5–4	0.75	−52.0	−1.82
	Thorax length (μm)	6	10	2.5–4	0.75	−16.8	−1.02
	Abdominal bristle number	7	20–17	2.5	0.99	+1.82	+1.08
T. castaneum	Offspring number	8	48	10–100	0.64	−2.56	−0.66
Barley	Internode length	9	7	1	0.99	−2.36	−0.29
Corn	Plant height (cm)	10	248	1	0.99	−48.0	−3.73
		11, 12	60	1–2.5	0.94	−55.4	−3.21
	Ear height (cm)	10	248	1	0.99	−30.0	−2.83
		11, 12	60	1–2.5	0.94	−26.9	−1.69
	Ear-leaf width (cm)	10	248	1	0.99	−1.34	−1.38
	Ear length (cm)	10	248	1	0.99	−4.40	−3.69
	Ear diameter (cm)	10	248	1	0.99	−10.08	−5.10
	Kernel depth (mm)	10	248	1	0.99	−6.47	−4.69
	Yield (g/ha)	10	248	1	0.99	−44.9	−3.27
	Days to silking	10	248	1	0.99	+4.6	+1.19

L = number of lines, N_e = effective population size (in some cases several treatments were utilized), $F(k)$ = maximum level of inbreeding at the end of the experiment, $V_{gw}^{1/2}(0)$ = additive genetic standard deviation in the base population. The work with sheep did not involve discrete lines, but utilized members of a large population at various levels of inbreeding.

References: 1) BOWMAN and FALCONER (1960), 2) ROBERTS (1960), 3) EISEN and HANRAHAN (1974), 4) MORLEY (1954), 5) TANTAWY and REEVE (1956), 6) TANTAWY (1957), 7) KIDWELL and KIDWELL (1966), 8) RICH *et al.* (1984), 9) BATEMAN and MATHER (1951), 10) HALLAUER and SEARS (1973), 11) CORNELIUS and DUDLEY (1974), 12) CORNELIUS (1972).

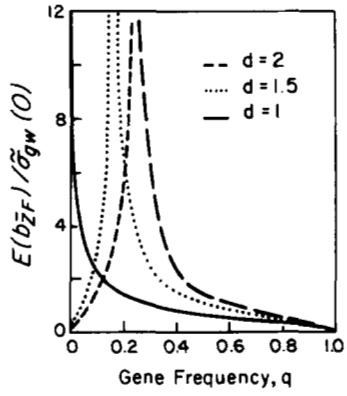

FIGURE 6.—Expected standardized inbreeding depression caused by a single diallelic locus with various dominance coefficients (d) and gene frequencies (q). Following standard theory (FALCONER 1981) and letting the three genotypic values at a locus be $2a$, $(1 + d)a$, and 0, it can be shown that $E(b_{\bar{z}F}) = 2pqad$ and $\tilde{\sigma}_{gw}^2(0) = 2pqa^2[1 + d(q - p)]^2$.

pling. Using Equation 9, it can be shown that the sampling variance is

$$\sigma_{\Delta\bar{z}}^2 = \frac{1}{L}\left[2\tilde{\sigma}_{gw}^2(0) + 2\sigma_c^2 + \frac{(3/4)\tilde{\sigma}_{gw}^2(0) + 2\sigma_e^2}{n}\right] \quad (15a)$$

when the inbred progeny are acquired by self-fertilization, and

$$\sigma_{\Delta\bar{z}}^2 = \frac{1}{L}\left[\frac{5\tilde{\sigma}_{gw}^2(0)}{4} + 2\sigma_c^2 + \frac{(7/8)\tilde{\sigma}_{gw}^2(0) + 2\sigma_e^2}{n}\right] \quad (15b)$$

when they are acquired by full-sib mating.

These formulas are difficult to implement unless one has information on the components of variance in the base population. If, however, a single offspring is monitored from each family, then $n = 1$, and the preceding expressions become

$$\sigma_{\Delta\bar{z}}^2 = \frac{1}{L}[2\tilde{\sigma}_z^2(0) + (3/4)\tilde{\sigma}_{gw}^2(0)] \quad (15c)$$

$$\sigma_{\Delta\bar{z}}^2 = \frac{1}{L}[2\tilde{\sigma}_z^2(0) + (1/8)\tilde{\sigma}_{gw}^2(0)] \quad (15d)$$

where $\tilde{\sigma}_z^2(0)$ is the phenotypic variance in the random-bred population. Since the additive genetic variance is less than the phenotypic variance, these two quantities can be no greater than $(11/4)\tilde{\sigma}_z^2(0)/L$ and $(17/8)\tilde{\sigma}_z^2(0)/L$, respectively.

Thus, a conservative test for inbreeding depression based on a single generation of consanguineous mating can be performed as follows. Subject L females to random mating and L different females to consangui-

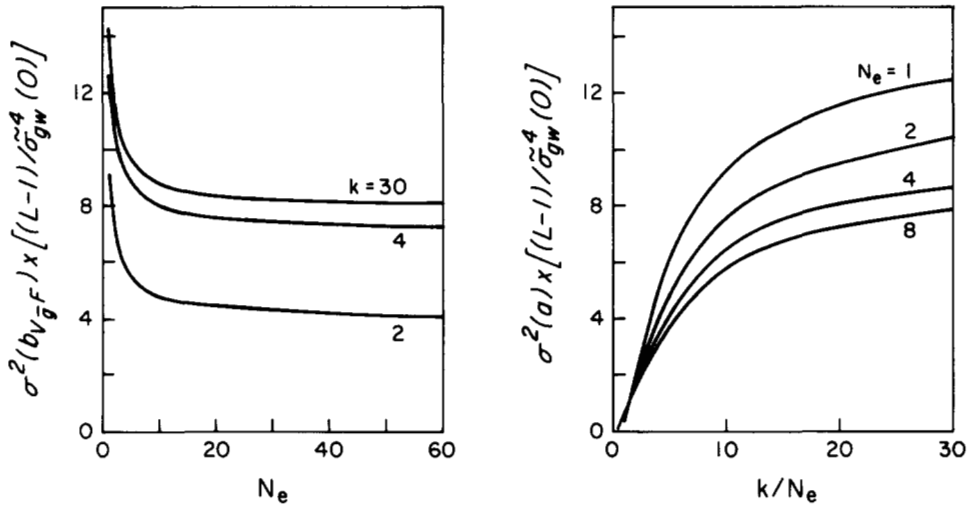

FIGURE 7.—The sampling variance of the regression of $V_{\bar{g}}(t)$ on $F(t)$ as a function of the effective population size (left), and the sampling variance of the intercept as a function of k/N_e (right), where k is the final generation in the regression. The sampling variances under a particular experimental setting are obtained by multiplying the plotted values by $\sigma^4_{gw}(0)/(L-1)$. The results are for ideal monecious populations.

neous mating, and from each of these families measure a single random offspring. Then compute the statistic

$$t = \frac{\Delta\bar{z}}{[\phi V_z(0)/L]^{1/2}} \tag{16}$$

where $V_z(0)$ is the observed phenotypic variance of random-bred progeny and $\phi = 11/4$ with self-fertilization and 17/8 with full-sib mating. Provided the character is approximately normally distributed, t may be treated as t–distributed with $L - 1$ degrees of freedom. Suppose, for example, that the experiment consisted of 10 inbred and 10 random-bred families. Rejection of the null hypothesis of no inbreeding depresson at the 95% level then requires that $t > 2.26$. For self-fertilization and full-sib mating, $\Delta\bar{z}$ would have to exceed 1.2 and 1.0 phenotypic standard deviations respectively for this criterion to be met.

In the case of self-compatible plants that produce multiple flowers, the construction of diallels can further increase the power of a short-term test of inbreeding depression. Pairs of parent plants (A and B) can be used to produce two reciprocal outcrosses ($A \times B$ and $B \times A$) and two inbreds ($A \times A$ and $B \times B$). In this case, $\Delta\bar{z}$ is the difference between the mean phenotypes of offspring from the two types of mating. An advantage of this approach is that it eliminates the contribution of maternal effects and parent sampling to $\sigma^2_{\Delta\bar{z}}$. If n of each progeny type are obtained from each of L pairs of parents,

$$\sigma^2_{\bar{z}} = \frac{1}{L}\left[\frac{\tilde{\sigma}^2_{gw}(0)}{4} + \frac{(3/8)\tilde{\sigma}^2_{gw}(0) + \sigma^2_e}{n}\right].$$

Even with $n = 1$, this quantity is less than $\tilde{\sigma}^2_z(0)/L$. Thus, a very conservative test for inbreeding depression using diallels is to maximize L, measuring one of each of the four types of progeny per parent pair, and then to employ Equation 16 with $\phi = 1$.

Analysis of the between-line variance: Regression analysis can be applied profitably to temporal data on variances as well as on means. The regression of the between-population variance on the cumulative inbreeding ($b_{V_{\bar{g}}F}$), for example, is a useful test of drift theory since its expectation is $2\tilde{\sigma}^2_{gw}(0)$ for additive genes. The intercept of such a regression (a) is also of interest since, in the case of neutrality, it provides a pooled estimate of the between-line variance attributable to factors other than drift (measurement error). Following the procedures outlined above,

$$\begin{aligned}\sigma^2(b_{V_{\bar{g}}F}) = \Big\{\sum_{t=0}^{k}[F(t) - \bar{F}]^2\Big\}^{-2} \cdot \sum_{t=0}^{k}\sum_{t'=0}^{k}\{E[F(t) \\ - \bar{F}]E[F(t') - \bar{F}]\sigma[\sigma^2_{\bar{g}}(t),\sigma^2_{\bar{g}}(t')] + E[F(t) \\ - \bar{F}]E[V_{\bar{g}}(t') - \bar{V}_{\bar{g}}]\sigma[F(t),\sigma^2_{\bar{g}}(t')] + E[V_{\bar{g}}(t) \\ - \bar{V}_{\bar{g}}]E[V_{\bar{g}}(t') - \bar{V}_{\bar{g}}]\sigma[F(t),F(t')]\},\end{aligned} \tag{17}$$

$$\begin{aligned}\sigma^2(a) = \frac{1}{k+1}\sum_{t=0}^{k}\sigma^2[\sigma^2_{\bar{g}}(t)] \\ - \frac{2}{k(k+1)}\sum_{t=0}^{k}\sum_{t'=t+1}^{k}\sigma[\sigma^2_{\bar{g}}(t),\sigma^2_{\bar{g}}(t')] + \bar{F}^2\sigma^2(b_{V_{\bar{g}}F}),\end{aligned} \tag{18}$$

where $\bar{V}_{\bar{g}}$ is the mean between-line variance over an experiment of $k + 1$ consecutive generations, $\sigma[F(t), \sigma^2_{\bar{g}}(t')] = 2\sigma^2_{gw}(0)\lambda^{t'-t}\sigma^2_F(t)/(L-1)$, and $\sigma[F(t),F(t')] = \lambda^{t'-t}\sigma^2_F(t)/(L-1)$ with $t \le t'$. The variance of $b_{V_{\bar{g}}F}$ has been derived under the assumption that the variance of $\bar{V}_{\bar{g}}$ is of negligible significance and ignores the variance of $V_{\bar{g}}(t)$ due to measurement error. Therefore, Equations 17 and 18 give lower bounds on the sampling variances of the regresson parameters under the assumption of neutrality.

The solution of Equation 17 indicates that $\sigma^2(b_{V_{\bar{g}}F})$ increases approximately twofold with the duration of the experiment, although it is essentially stable for $k \ge 4$ provided most pairs of loci are unlinked (Figure 7). For large k, the standard error of $b_{V_{\bar{g}}F}$ is no less than $3.8\tilde{\sigma}^2_{gw}(0)/(L-1)^{1/2}$ for self-fertilizing lines and $2\sqrt{2}\tilde{\sigma}^2_{gw}(0)/(L-1)^{1/2}$ for large N_e. For small N_e and $k/N_e < 2$, the standard error of the intercept is no less

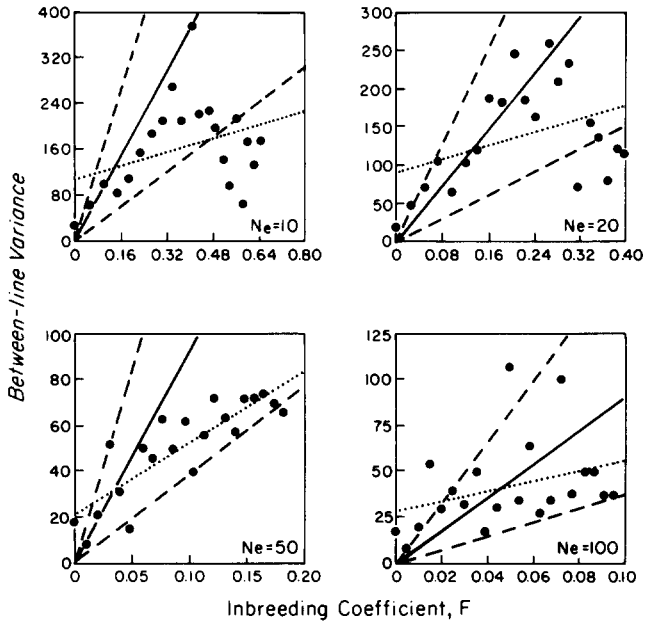

FIGURE 8.—Observed levels of between-line variance for pupal weight for *Tribolium* populations at four effective sizes (RICH *et al.* 1984) as a function of the expected level of inbreeding (solid points). Solid lines are the expected regressions under the neutral, additive genetic hypothesis; dashed lines are conservative 90% confidence limits; and dotted lines are the least-squares regressions.

than $0.9\tilde{\sigma}^2_{gw}(0)\cdot[k/N_e(L-1)]^{1/2}$. With increasing k/N_e, $\sigma^2(a)$ gradually approaches a value on the order of $\sigma^2(b_{V_g F})$.

As an example of the application of the preceding formulas, the results of a large drift experiment with laboratory cultures of *Tribolium castaneum* (RICH *et al.* 1984) will be examined. The authors followed 12 replicate populations at four population sizes (1:1 sex ratio, random mating) over 20 consecutive generations. Each generation, the mean pupal weight (μg) of each population was obtained from a bulk sample of 100 random individuals. The additive genetic variance was estimated to be 460 in the base population. The observed $V_{\bar{g}}(t)$ are plotted as a function of $F(t)$ in Figure 8, along with the expected divergence $920F(t)$ (solid lines). Any interpretation of the results of this study is weakened by the lack of a control. The authors argued that the downward trend in $V_{\bar{g}}(t)$ in the last few generations of three of the four treatments was due to the suppression of random drift and the operation of stabilizing selection. However, the same result could have arisen as a response to a shift in the laboratory environment that influenced the expression of variation.

The dashed lines in Figure 8 give the limits of the between-line variance beyond which there is a less than 5% chance for the realization of the drift process in either direction. Since these bounds are based on a χ^2 distribution, which underestimates the dispersion somewhat, and also ignore measurement error, they may be regarded as conservative confidence limits. Nevertheless, almost all of the observations, with the exception of the late generations at $N_e = 10$ and 20, lie within these limits. There are substantially more observations below (54) than above (26) the expectation, possibly because the additive genetic variance in the base population was overestimated somewhat.

TABLE 3

Least-squares estimates of the regression coefficients and intercepts for the data in Figure 8

N_e	$b_{V_g F}$	$\sigma(b_{V_g F})$	a	$\sigma(a)$
10	148	412	109	154
20	222	402	91	87
50	312	396	21	38
100	292	394	28	19

The standard errors were obtained from Equations 17 and 18 under the assumption of unlinked loci.

The least-squares regressions of the data are given by the dotted lines in the figure. The slope of each regression is less than the expected 920, but all are within 2 SE of the expectation (Table 3). The intercepts are all above the expectation of zero, perhaps due to measurement error, but are well within 2 SE of it. Thus, this fairly conservative analysis indicates that the observed patterns, even in the absence of a control, are consistent with a hypothesis of random drift of neutral, additive genes. There is a significant probability that the observed delines in $V_{\bar{g}}(t)$ late in the experiment arose by chance. In the case of the two smallest effective population sizes the chances of $V_{\bar{g}}(t)$ returning toward the expectation late in the experiment were small since the lines must have already become fixed at many loci.

STATISTICAL PROPERTIES OF THE WITHIN-POPULATION GENETIC VARIANCE

As in the case of the between-line variance, there are several reasons why the realized dynamics of the within-line variance may depart substantially from the expectation even when the assumptions of neutrality and additivity are met: variation in the genetic variance among founder populations, variance in inbreeding, deviations from Hardy-Weinberg equilibrium, and linkage disequilibrium. Although substantial theoretical progress on these matters has been made (AVERY and HILL 1977, 1979; BULMER 1976, 1980), the existing work relies on several simplifying assumptions in the interest of analytical tractability: a base population in linkage equilibrium, no variance in inbreeding between populations, and $t < 2N_e$. Since the latter two assumptions will often be violated in empirical studies, it is necessary to relax them. The following analysis will focus on unlinked loci, since for most organisms the majority of pairs of loci are expected to be on different chromosomes.

As first pointed out by BULMER (1976), the dominant source of variance of $\sigma^2_{gw}(t)$ is the random devel-

opment of linkage disequilibrium that inevitably develops in finite populations, even for unlinked loci. From AVERY and HILL (1977), with two alleles/locus,

$$\sigma^2[\sigma^2_{gw}(t)] = 4 \sum_{i=1}^{n} \sum_{j=1}^{n} a_i^2 a_j^2 [q_{i0}(1 - q_{i0})q_{j0}(1 - q_{j0})]\{\phi_{ijt} + 2\theta_{ijt} - [1 - F(t)]^2\}/L \quad (19)$$

where q_{i0} is the initial gene frequency at the ith locus, $\phi_{ijt} = E[q_{it}(1 - q_{it})q_{jt}(1 - q_{jt})]/[q_{i0}(1 - q_{i0})q_{j0}(1 - q_{j0})]$, $\theta_{ijt} = E(D^2_{ijt})/[q_{i0}(1 - q_{i0})q_{j0}(1 - q_{j0})]$, and D^2_{ijt} is the squared linkage disequilibrium between loci i and j. An analytical expression is available for ϕ_{iit} in CROW and KIMURA (1970), whereas $\theta_{iit} = 0$. The two-locus expectations, ϕ_{ijt} and θ_{ijt}, are independent of gene frequencies and can be evaluated by use of the moment-generating matrix of HILL and ROBERTSON (1968), which applies to systems of invariant inbreeding. To accommodate variance in inbreeding, a Taylor expansion was performed on the elements of the matrix, letting $\sigma^2(N_e) \simeq \sigma^2_F(1)/4F(1)$.

Obviously, Equation 19 is a rather complicated function, but great simplification can be gained following the logic of AVERY and HILL (1977). First, note that the squared expectation of the within-population variance is

$$\tilde{\sigma}^4_{gw}(t) = 4 \sum_{i=1}^{n} \sum_{j=1}^{n} a_i^2 a_j^2 [q_{i0}(1 - q_{i0})q_{j0}(1 - q_{j0})][1 - F(t)]^2. \quad (20)$$

Second, note that the fraction of terms in Equations 19 and 20 attributable to pairs of loci is $(n - 1)/n$, so that with large numbers of loci the contribution of single-locus terms becomes diminishingly small. Thus, the squared coefficient of variation is

$$\Gamma = \sigma^2[\sigma^2_{gw}(t)]/\tilde{\sigma}^4_{gw}(t) \simeq \frac{1}{L}\left[\frac{\phi_{ijt} + 2\theta_{ijt}}{[1 - F(t)]^2} - 1\right]. \quad (21)$$

This function is plotted for ideal monoecious populations and for full-sib mating in Figure 9. The predicted variance is that which is expected within a large progeny group.

For populations of effective sizes of 4 or greater, $\Gamma \simeq 4/3N_eL$ from the very outset of an experiment. The same conclusion was reached by AVERY and HILL (1977), showing that for unlinked loci, their results hold very well even for large t/N_e and are influenced only negligibly by variance in inbreeding. The results of BULMER (1980), obtained in a different manner, suggest Γ may be closer to $5/3N_eL$, but this discrepancy has little bearing on the following conclusions. For smaller populations, the variance in $\sigma^2_{gw}(t)$ caused by linkage is somewhat larger. With full-sib mating Γ rises from $0.5/L$ to $0.8/L$ by 10 generations of inbreeding, at which point it would be very difficult to acquire accurate estimates of $\sigma^2_{gw}(t)$ since inbreeding

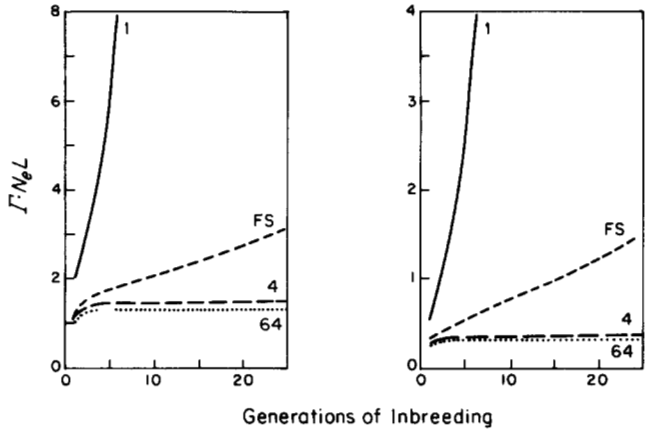

FIGURE 9.—The squared coefficient of variation of the within-population genetic variance multiplied by N_eL, considering only the variation caused by linkage disequilibrium. The panel on the left refers to a large progeny group after the denoted generations of inbreeding; the panel on the right refers to the situation after the progeny group has been mated randomly for a single generation.

has proceeded to 90%. With self-fertilization, Γ increases from $2.0/L$ to $6.3/L$ at five generations of inbreeding.

Since linkage disequilibrium is a transient phenomenon, the sampling variance of $\sigma^2_{gw}(t)$ can be reduced substantially by expanding and randomly mating within each line prior to analysis. The improvement in accuracy can be determined by recomputing ϕ_{ijt} and θ_{ijt} from the moment-generating matrix after allowing for an additional generation with $(1/2N_e) \simeq 0$. The results of a single generation of such treatment are shown in Figure 9, where it can be seem that Γ is reduced to between 50% and 25% of its previous value if the loci are unlinked.

Several other sources of variation of $\sigma^2_{gw}(t)$ exist. First, there is the variance in the initial within-line variance caused by a finite number of founders. If the lines are established with independent members of the base population, then $\sigma^2[\sigma^2_{gw}(0)] = 2\tilde{\sigma}^4_{gw}(0)/L(N_m + N_f)$ assuming normally distributed breeding values. This initial variation is propagated through all subsequent generations as $2[1 - F(t)]^2\tilde{\sigma}^4_{gw}(0)/L(N_m + N_f)$, giving a squared coefficient of variation of $2/LN_e$ for monoecious populations.

BULMER (1980) has pointed out that deviations from Hardy-Weinberg equilibria within and between loci are an additional source of variation of $\sigma^2_{gw}(t)$. However, if a substantial number of individuals in a progeny group are evaluated, this source of variation can safely be ignored. BULMER (1980) also notes that variation in the "true" genetic variance is caused by the propagation of random variation in heterozygosity of individual loci. This is the variance attributable to the single locus terms that were ignored in the derivation of Equation 21. A simple statement in terms of observable parameters is not possible here since the variance of heterozygosity is a function of gene frequencies. If, however, gene effects and frequencies

are assumed to be uncorrelated, it follows that the variance of the "true" genetic variance is inversely related to the number of loci. Thus, if there are a large number of independent loci, this additional source of variation is likely to be small relative to the sources described above.

Finally, it should be noted that the previous derivations have been performed under the assumption of unlinked loci and a base population in linkage equilibrium. Linkage will cause greater variation than that noted above, but for most chromosomal systems the inflation is expected to be fairly small, most likely less than 50% (AVERY and HILL 1977). Linkage disequilibrium in the base population causes additional problems not only for the variance of $\sigma^2_{gw}(t)$ but also for its expectation. $\tilde{\sigma}^2_{gw}(t)$ will no longer change in proportion to the inbreeding coefficient and may actually increase if there is substantial negative disequilibrium initially. AVERY and HILL (1977) have pointed out that with unlinked loci, the values in Figure 9 are reached in 3–4 generations regardless of the initial state of the population.

Summing up all of the sources of variation of $\sigma^2_{gw}(t)$, it can be seen that only a rough statement can be made as to the efficiency of an experimental design. For lines with effective sizes of four or greater, the squared coefficient of variation of $\sigma^2_{gw}(t)$ is at least $(4/3N_eL) + (2/N_eL)$. The coefficient of variation is therefore on the order of $2/\sqrt{N_eL}$. Therefore, if it is desirable to keep the standard error at a level of 10% of the expectation, the design must be such that $N_eL = 400$; *e.g.*, 100 lines of $N_e = 4$, or 25 of $N_e = 16$. With full-sib mating, the coefficient of variation of $\sigma^2_{gw}(t)$ is expected to be of the order $\sqrt{2/L}$ for the first 10 generations. Therefore, for the same level of accuracy, 200 lines would need to be evaluated. Finally, for self-fertilizing lines, the coefficient of variation is expected to average more than $\sqrt{7/L}$ over the first five generations of inbreeding, so at least 700 lines would need to be monitored. These guidelines are on the conservative side as they do not include the additional, and usually substantial, variation due to the deviation of the measurement $V_{gw}(t)$ from the population parameter $\sigma^2_{gw}(t)$.

In evaluating the overall dynamics of observational estimates of $\sigma^2_{gw}(t)$ account also needs to be taken of the fact that subsequent observations are not independent of each other. Much, but not all, of the covariance between observations in the same population is caused by the slow decay of linkage disequilibrium. Both AVERY and HILL (1977) and BULMER (1980) have considered this problem, and from the former it can be deduced that for unlinked loci, the covariance between $\sigma^2_{gw}(t)$ and $\sigma^2_{gw}(t')$ is close to $2^{2-t'+t}[1 - F(t')]\sigma^2_{gw}(0)/3N_eL$. Such covariance would need to be incorporated into tests of the significance and linearity of the decline of $\sigma^2_{gw}(t)$ with $F(t)$. However, in light of the rough nature of the expressions for the variances and covariances of the $\sigma^2_{gw}(t)$, no further attempt to develop the regression theory will be made here.

The preceding results emphasize that the variance of the realization of the drift process [the dispersion of $\sigma^2_{gw}(t)$ around $\tilde{\sigma}^2_{gw}(t)$] should not be ignored in studies of the dynamics of additive genetic variance within small populations. Up to now, however, sampling variance [the dispersion of $V_{gw}(t)$ around $\sigma^2_{gw}(t)$] has been treated as the sole error in empirical studies, a procedure that can only falsely encourage the rejection of the neutral, additive gene model. Based on this approach, LINTS and BURGOIS (1984) and BRYANT, MCCOMMAS and COMBS (1986) have fostered the idea that bottlenecks cause an inflation of the additive genetic variance within populations. The conclusion of LINTS and BURGOIS (1984) is based on a single line of *D. melanogaster* of somewhat uncertain N_e. They compared the realized additive genetic variance, obtained from selection experiments, for sternopleural bristle number in the bottlenecked population and "control" lines. No statistical comparison of the lines was actually performed, but even if there had been one, the results would have been questionable, since each line was evaluated in a different environment.

More credible, but still difficult to evaluate, are the results of BRYANT, MCCOMMAS and COMBS (1986). They extracted houseflies from a large base population and subjected four replicates to single-generation bottlenecks of 2, 8 and 32 individuals. The lines were allowed to expand for 5 generations to approximately 2000 flies prior to analysis. This is fortuitous since, as noted above, such treatment reduces the variance of $\sigma^2_{gw}(t)$ caused by linkage disequilibrium. The authors measured eight genetically correlated characters and concluded that several of them exhibited significant increases in additive genetic variance relative to the base population. These comparisons, however, are based entirely on the variance of $V_{gw}(t)$ due to estimation procedure. Even if all of the variation due to linkage disequilibrium were eliminated by the experimental protocol, the variation in $\sigma^2_{gw}(t)$ caused by the variation in the initial genetic variance among the founder populations is too substantial to ignore. The coefficient of variation caused by this source of variation alone, $\sqrt{2/N_eL}$, is 0.45, 0.25 and 0.12 for the three bottleneck treatments.

There are ways of reducing the problem of variation in the additive genetic variance among founder populations. For example, if N_e families in the base population each contribute single individuals to each founder population, the variance in initial additive genetic variance among lines would be reduced by 50% relative to the situation in which the founders are drawn randomly from the base. If several highly inbred lines are available, each founder population can be established with identical genetic properties.

The need for inbred lines can be avoided if individuals can be multiplied temporarily by vegetative means.

DISCUSSION

It has been shown that for characters with an additive genetic basis, aside from the presence of general environmental effects and possible operation of evolutionary forces such as selection and mutation, there are two causes for the deviations of quantitative genetic estimates from the simple expectations of finite population theory: variation in the realization of the drift/inbreeding process and sampling variance resulting from the estimation procedure. Virtually all existing empirical studies have considered only the latter source of variation, and hence are biased in the direction of falsifying the null hypothesis of neutral additive genes. Under most experimental protocols, this bias is much too large to be ignored.

The major sources of the "realization variance" are the evaluation of a finite number of lines, variance in the genetic properties of the founder populations, linkage disequilibrium in the base and study populations, and variance in inbreeding between lines. Some protocols have been suggested to minimize the realization variance, but in no case can it be eliminated entirely.

Although the random development of linkage disequilibrium and variance in inbreeding are difficult to observe with quantitative characters, they are very real problems in the analysis of very small populations. Fortunately, if most pairs of loci are unlinked the expected variance of the realized changes in genotypic means and variances can still be described in terms of observable quantities (additive genetic variance in the base population, effective population size, and numbers of generations and lines). In cases where a few genes with major effects are linked, the true realization variances will be even higher than those defined above. Other factors will insure that this is so. First, linkage disequilibrium in the base population is always a problem, but because it cannot be easily quantified, there is no simple way to incorporate it into statistical procedures. The best way to guard against this problem is to expand and randomly mate the base population for several generations prior to the extraction of experimental lines. Second, the theory in this paper only considers the genetic variation preexisting in the base population, and therefore assumes that new mutations are not a significant source of variance. Such an assumption is reasonable for experiments of short duration (say, <6 generations of inbreeding), but since many uncertainties still exist over the rate of polygenic mutation (TURELLI 1984; LYNCH 1988), the potential noise caused by mutation in longer experiments should not be ignored. Expressions for the variance of the realized within- and between-line variance caused by mutation are presented in LYNCH and HILL (1986).

Keeping in mind that the expressions for the realization variances presented in this paper are on the conservative side and do not include the sampling variance due to the estimation procedure, their application should improve the utility of the neutral additive genetic model as a null hypothesis in testing for various genetic properties and evolutionary consequences of small population size. As in all statistical analyses, since the form of the distributions of realized means, variances, and regression coefficients could vary substantially from case to case, most of the suggested tests should be considered to be approximations and interpretative guides. Since most of the realized variances are expressed in terms of the additive genetic variance in the base population, a major priority in any drift or inbreeding experiment should be to accurately estimate this parameter at the outset. The only way to apply the formulas for the realization variances is to substitute the observation $V_{gw}(0)$ for its expectation $\tilde{\sigma}^2_{gw}(0)$, so if $V_{gw}(0)$ is an underestimate, the standard errors of parameter estimates will be also.

In order to ease the presentation of some rather complicated formulas, the theory in this paper has been developed under the assumption of a balanced experimental design. This, of course, will amost never be true in large experiments of long duration. Procedures for dealing with unbalanced data sets are complicated, but are outlined in most statistics texts. Use of the relationship matrix (SORENSEN and KENNEDY 1983) although computationally demanding, should allow a generalization of many of the procedures outlined above to arbitrary distributions of family size and structure.

Of greater concern here is the problem as to whether lost individuals and/or lines are a biased sample of the whole population. In the mouse, for example, pronounced inbreeding depression usually follows the first few generations of full-sib mating, and then the overall population appears to recover (BOWMAN and FALCONER 1960; LYNCH 1977; CONNOR and BELLUCCI 1979). This is due to the differential extinction and survival of lines that do and do not exhibit the deleterious effects of inbreeding. A large part of the problem with the mouse is its relatively low reproductive capacity. Corn, which clearly suffers from inbreeding depression but has high fecundity, generally has survivorship of selfed lines in excess of 99%. On the other hand, while experiments with high fecundity organisms may have minimal problems with selection between lines, by allowing the maintenance of "spares," they inflate the problem of selection within lines.

The utility of the neutral, additive genetic model is its simplicity (nondependence on the gene frequency distribution) and unambiguous predictions. A confi-

dent rejection of this model is tantamount to accepting that alternative modes of gene action (dominance or epistasis) or evolutionary forces (selection or migration) are of significant importance. An explicit test of any alternative hypothesis would require the development of statistical methods comparable to those described in this paper but tailored to the specific evolutionary scenario.

Given the fundamental significance of drift and inbreeding to so many issues in biology and the widespread application of the existing theory to practical problems, the lack of large, reliable data sets on the subject is striking. Aside from the excellent data on corn, almost all of the existing information is from laboratory populations of *Drosophila*, *Tribolium*, and mice, and from human surveys, and these studies are less than ideal in many respects. The major conclusion of this paper is that if a confident understanding of the genetic consequences of finite population size and inbreeding depression is desirable, experimental analyses are going to have to involve a large number of lines (ideally, over 100), monitored over several generations with parallel controls. In order to evaluate how natural selection and random genetic drift interact, it will be necessary to perform experiments simultaneously at different population sizes to allow the relative magnitude of the two forces to vary. This calls for an effort on a scale somewhat larger than many population geneticists are used to. To put it in perspective, however, the necessary investment for such work would be a pittance compared to the cost of sequencing a human genome or building a supercollidor. It is not clear that the benefits would be so dwarfed.

This work was supported by National Science Foundation grants BSR 83-06072 and BSR 86-00487. Many thanks to W. G. HILL for helpful comments.

LITERATURE CITED

AVERY, P. J., and W. G. HILL, 1977 Variability in genetic parameters among small populations. Genet. Res. **29:** 193–213.

AVERY, P. J., and W. G. HILL, 1979 Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. Genetics **91:** 817–844.

BARTON, N. H., and B. CHARLESWORTH, 1984 Genetic revolutions, founder effects, and speciation. Annu. Rev. Ecol. Syst. **15:** 133–164.

BARTUAL, R., and A. R. HALLAUER, 1976 Variability among unselected maize inbred lines developed by full-sibbing. Maydica **21:** 49–60.

BATEMAN, A. J., and K. MATHER, 1951 The progress of inbreeding in barley. Heredity **5:** 321–348.

BOWMAN, J. C., and D. S. FALCONER, 1960 Inbreeding depression and heterosis of litter size in mice. Genet. Res. **1:** 262–274.

BRAY, D. F., A. E. BELL and S. C. KING, 1962 The importance of genotype by environment interaction with reference to control populations. Genet. Res. **3:** 282–302.

BRYANT, E. H., L. M. COMBS and S. A. MCCOMAS, 1986 Morphometric differentiation among experimental lines of the housefly in relation to a bottleneck. Genetics **114:** 1213–1223.

BRYANT, E. H., S. A. MCCOMAS and L. M. COMBS, 1986 The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics **114:** 1191–1211.

BULMER, M. G., 1976 The effect of selection on genetic variability: a simulation study. Genet. Res. **28:** 101–117.

BULMER, M. G., 1980 *The Mathematical Theory of Quantitative Genetics.* Oxford University Press, Oxford, U.K.

CAVALLI-SFORZA, L. L., and W. F. BODMER, 1971 *The Genetics of Human Populations.* W. H. Freeman, New York.

CHAKRABORTY, R., and M. NEI, 1982 Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. Genet. Res. **39:** 303–314.

CHARLESWORTH, B., 1984 Some quantitative methods for studying evolutionary patterns in single characters. Paleobiology **10:** 308–318.

CHARLESWORTH, D., and B. CHARLESWORTH, 1987 Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. **18:** 237–268.

COCKERHAM, C. C., and B. S. WEIR, 1983 Variance of actual inbreeding. Theor. Popul. Biol. **23:** 85–109.

CONNOR, J. L., and M. J. BELLUCCI, 1979 Natural selection resisting inbreeding depression in captive wild housemice (*Mus musculus*). Evolution **33:** 929–940.

CORNELIUS, P. L., 1972 Quantitative genetic evaluation of full sib-mating and selfing as methods of inbreeding in *Zea mays* L. Ph.D. thesis, University of Illinois.

CORNELIUS, P. L., and J. W. DUDLEY, 1974 Effects of inbreeding by selfing and full-sib mating in a maize population. Crop Sci. **14:** 815–819.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Burgess, Minneapolis, Minn.

EISEN, E. J., and J. P. HANRAHAN, 1974 Genetic drift and inbreeding depression measured from control populations of mice. Can. J. Genet. Cytol. **16:** 91–104.

FALCONER, D. S., 1981 *Introduction to Quantitative Genetics.* Longman, New York.

FELSENSTEIN, J., 1985 Phylogenies and the comparative method. Am. Nat. **125:** 1–15.

FESTING, M., 1979 *Inbred Strains in Biomedical Research.* Oxford University Press, New York.

GOODNIGHT, C. J., 1987 On the effect of founder events on epistatic genetic variance. Evolution **41:** 80–91.

HALLAUER, A. R., and J. H. SEARS, 1973 Changes in quantitative traits associated with inbreeding in a synthetic variety of maize. Crop Sci. **13:** 327–330.

HILL, W. G., 1972a Estimation of realised heritabilities from selection experiments. I. Divergent selection. Biometrics **28:** 747–765.

HILL, W. G., 1972b Estimation of realised heritabilities from selection experiments. II. Selection in one direction. Biometrics **28:** 767–780.

HILL, W. G., 1972c Estimation of genetic change. I. General theory and design of control populations. Anim. Breed. Abstr. **40:** 1–15.

HILL, W. G., 1972d Estimation of genetic change. II. Experimental evaluation of control populations. Anim. Breed. Abstr. **40:** 193–213.

HILL, W. G., 1982 Dominance and epistasis as components of heterosis. Z. Tierz. Zuechtungsbiol. **99:** 161–168.

HILL, W. G., and A. ROBERTSON, 1968 Linkage disequilibrium in finite populations. Theor. Appl. Genet. **38:** 226–231.

KIDWELL, J. F., and M. M. KIDWELL, 1966 The effects of inbreeding on body weight and abdominal chaeta number in *Drosophila melanogaster.* Can. J. Genet. Cytol. **8:** 207–215.

KIMURA, M., and J. F. CROW, 1963 On the maximum avoidance of inbreeding. Genet. Res. **4:** 399–415.

LANDE, R., 1977 Statistical tests for natural selection on quantitative characters. Evolution **31:** 442–446.

LERNER, I. M., 1954 *Genetic Homeostasis.* Oliver & Boyd, Edinburgh.

LINTS, F. A., and M. BOURGOIS, 1984 Population crash, population flush and genetic variability in cage populations of *Drosophila melanogaster*. Genet. Sel. Evol. **16:** 45–56.

LYNCH, C. B., 1977 Inbreeding effects upon animals derived from a wild population of *Mus musculus*. Evolution **31:** 526–537.

LYNCH, M., 1984 Spontaneous mutations for life-history characters in an obligate parthenogen. Evolution **39:** 804–818.

LYNCH, M., 1988 The rate of polygenic mutation. Genet. Res. **51:** 137–148.

LYNCH, M., and W. G. HILL, 1986 Phenotypic evolution by neutral mutation. Evolution **40:** 915–935.

MANLY, B. F. J., 1985 *The Statistics of Natural Selection*. Chapman & Hall, New York.

MAYR, E., 1963 *Animal Species and Evolution*. Belknap Press, Cambridge, Mass.

MORLEY, F. H. W., 1954 Selection for economic characters in Australian Merino sheep. IV. The effect of inbreeding. Aust. J. Agric. Res. **5:** 305–316.

MUIR, W. M., 1986a Estimation of response to selection and utilization of control populations for additional information and accuracy. Biometrics **42:** 381–391.

MUIR, W. M., 1986b Efficient design and analysis of selection experiments. pp. 269–282. In: *Proceedings 3rd World Congress on Genetics Applied to Livestock Production* (Agricultural Communications, Edited by E. DICKERSON and R. K. JOHNSON). University of Nebraska, Lincoln.

OBILANA, A. T., and A. R. HALLAUER, 1974 Estimation of variability of quantitative traits in BSSS by using unselected maize inbred lines. Crop. Sci. **14:** 99–103.

RASMUSON, M., 1952 Variation in bristle number of *Drosophila melanogaster*. Acta Zool. **33:** 277–307.

RICH, S. S., A. E. BELL, D. A. MILES and S. P. WILSON, 1984 An experimental study of genetic drift for two quantitative traits in *Tribolium*. J. Hered. **75:** 191–195.

ROBERTS, R. C., 1960 The effects on litter size of crossing lines of mice inbred without selection. Genet. Res. **1:** 239–252.

ROBERTSON, A., 1952 The effect of inbreeding on the variation due to recessive genes. Genetics **37:** 189–207.

RUSSELL, W. A., G. F. SPRAGUE and H. L. PENNY, 1963 Mutations affecting quantitative characters in long-time inbred lines of maize. Crop. Sci. **3:** 175–178.

SORENSEN, D. A., and B. W. KENNEDY, 1983 The use of the relationship matrix to account for genetic drift variance in the analysis of genetic experiments. Theor. Appl. Genet. **66:** 217–220.

SOULÉ, M. (Editor), 1986 *Conservation Biology*. Sinauer Assoc., Sunderland, Mass.

TANTAWY, A. O., 1957 Genetic variance of random-inbred lines of *Drosophila melanogaster* in relation to coefficients of inbreeding. Genetics **42:** 121–136.

TANTAWY, A. O., and E. C. R. REEVE, 1956 Studies on quantitative inheritance. IX. The effects of inbreeding at different rates in *Drosophila melanogaster*. Z. Indukt. Abstammungs. Vererbls. **87:** 648–667.

TEMPLETON, A. R., 1980 The theory of speciation via the founder principle. Genetics **94:** 1011–1038.

TURELLI, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. **25:** 138–193.

WEIR, B. S., and C. C. COCKERHAM, 1969 Group inbreeding with two linked loci. Genetics **63:** 711–742.

WEIR, B. S., P. J. AVERY and W. G. HILL, 1980 Effect of mating structure on variation in inbreeding. Theor. Popul. Biol. **18:** 396–429.

WRIGHT, S., 1921 Systems of mating. II. The effects of inbreeding on the genetic composition of a population. Genetics **6:** 124–143.

WRIGHT, S., 1951 The genetical structure of populations. Ann. Eugen. **15:** 323–354.

WRIGHT, S., 1977 *Evolution and the Genetics of Populations. Vol. 3, Experimental Results and Evolutionary Deductions*. University of Chicago Press, Chicago.

Communicating editor: M. TURELLI

APPENDIX

Approximate expectations for the sampling variance and covariance of the between-line variance

For $t = 0$,

$$\sigma^2[\sigma_z^2(0)] = \frac{2\tilde{\sigma}_{gw}^4(0)}{(L-1)(N_m + N_f)^2} + \left[\frac{2(\sigma_c^4 + \alpha_e^4)}{(L-1)(N_m + N_f)^2}\right]. \quad \text{(A1)}$$

For $nN_f \gg 1$ and $t > 0$,

$$\alpha^2[\sigma_z^2(t)] = \sigma^2[\sigma_g^2(0)] + \frac{4\tilde{\sigma}_{gw}(0)}{L-1}\left[2\left(2(1 + \frac{1}{N_m + N_f})F^2(t) + \sigma_F^2(t)\right]\right.$$
$$+ \left[\frac{2}{N_f(L-1)}\left\{\frac{\tilde{\sigma}_{gw}^4(0)}{n}\left[2\frac{F(t)[1 - F(t-1)]}{N_e}\right.\right.\right.$$
$$\left.\left.\left. - \lambda\sigma_F^2(t-1)\right] + \frac{1}{N_f}\left[\sigma_c^4 + \frac{\sigma_e^4}{n^2}\right]\right\}\right]. \quad \text{(A2)}$$

For $t' > t$,

$$\sigma[\sigma_z^2(t), \sigma_z^2(t')] = \sigma^2[\sigma_g^2(0)] + \frac{4\tilde{\sigma}_{gw}^4(0)}{L-1}\left\{2\left[1 + \frac{1}{N_m + N_f}\right]F(t)F(t') + \lambda^{t'-t}\sigma_F^2(t)\right\}$$
$$+ \left[\frac{2\tilde{\sigma}_{gw}^4(0)}{nN_F(N_m + N_f)(L-1)}\left[F(t)[1 - F(t'-1)]\right.\right.$$
$$+ F(t')[1 - F(t-1)]$$
$$+ \frac{[1 - F(t-1)][1 - F(t'-1)]}{4nN_f}$$
$$- \frac{\lambda^{t'-t}\tilde{\sigma}_{gw}^4(0)}{nN_f(L-1)}\left[\lambda^{-1}\sigma_F^2(t) + \lambda\sigma_F^2(t-1)\right.$$
$$\left.\left.\left. - \frac{\sigma_F^2(t-1)}{4nN_f}\right]\right] \quad \text{(A3)}$$

where $\lambda = 1 - (1/2N_e)$, $\sigma^2[\sigma_g^2(0)]$ is the first term of equation (A1), and $\sigma_F^2(t)$ is the expected variance in the cumulative inbreeding among replicate lines at time t.