Genet. Res., Camb. (1988) **51**, *pp.* 47–53 *Printed in Great Britain*

Variance components of fitness under stabilizing selection

HIDENORI TACHIDA AND C. CLARK COCKERHAM
Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203, USA

(*Received 10 February 1987 and in revised form 26 June 1987*)

Summary

Variance components of fitness under the stabilizing selection scheme of Wright (1935) for metric characters are calculated, extending his original analysis to the case with any number of alleles and multiple characters assuming additivity of gene effects. They are calculated in terms of the moments of the effects of alleles at individual loci for the metric characters. From these formulas, the variance components of fitness are evaluated at the mutation–selection equilibria predicted by the 'Gaussian' approximation (Lande, 1976), which is applicable if the per locus mutation rate is high, and the 'House of Cards' approximation (Turelli, 1984), which is applicable if the per locus mutation rate is low. It is found that the additive variance of fitness is small compared to non-additive variance in the 'Gaussian' case, whereas the additive variance is larger than non-additive variance in the 'House of Cards' case if the number of loci per character and the number of characters affected by each locus are not too large. With the assumption that a significant portion of fitness is due to this type of stabilizing selection, it is suggested that the real parameters are in the range where the 'House of Cards' approximation is applicable, since available data on variance components of fitness components in Drosophila show that the additive variance is far larger than the non-additive variance. It is noted that the present method does not discriminate the two approximations if the average values of the metric characters deviate from the optimum values. Other limitations of the present method are also discussed.

1. Introduction

Natural populations have considerable genetic variation with respect to most metric characters. One of the mechanisms proposed to account for the maintenance of their variation is the balance between stabilizing selection and mutation (Latter, 1960). Stabilizing selection acts against deviants from an optimal value and mutation supplies new deviants, thus creating the equilibrium state. Among the varieties of models, one which is simple and appeals to our biological intuition is the one originally proposed by Kimura (1965), and extended by Lande (1976) and Fleming (1979). It assumes that gene effects are additive for the quantitative character, that reduction of the fitness is proportional to the square of the phenotype deviation from the optimum (Wright, 1935) and that the effects of segregating alleles are normally distributed.

Since the model equation cannot be solved exactly, two types of approximation, 'Gaussian' and 'House of Cards', have been proposed. The former is applicable when per locus mutation rates are high, while the latter is applicable when per locus mutation rates are not too high (Turelli, 1984, 1986). These two approximations give different predictions as to the dependence of genetic variance on mutation and selection parameters and the number of loci. For example, once the gametic mutation rate is fixed for the quantitative character, the 'Gaussian' approximation predicts that the equilibrium variance is proportional to the square root of the number of contributing loci and to the average magnitude of mutation, while the 'House of Cards' approximation predicts that it is independent of both the number of loci and the magnitude of mutation (Turelli, 1984). However, it is difficult to find methods which will distinguish between these and other models.

One possible method is to utilize genetic covariances of relatives to make inferences about the magnitudes of additive, dominance and epistatic components of variance (Cockerham, 1963, 1980). Although the metric character is additive in the present context, fitness, when considered as a metric character, does have non-additive components. Although estimates of variance components of fitness are not available at present, there have been several experimental studies which carried out this genetic variance decomposition

in fitness components, viability and fecundity, in Drosophila (Mukai *et al.* 1974; Mukai & Nagano, 1983; Tachida *et al.* 1983; Kusakabe & Mukai, 1984; Rose & Charlesworth, 1981), and those studies have almost invariably shown that dominance and epistatic variances are small in comparison with the additive variance.

In the present paper we calculate additive, dominance and epistatic variances of fitness in terms of the moments of the original metric characters at individual loci in Wright's stabilizing selection model. Wright (1935) calculated the variances due to a single metric character in terms of effects and frequencies of genes assuming two alleles at each locus, but for our purposes a general setting which allows any number of alleles at each locus and any number of metric characters with pleiotropy is necessary. Although Lande (1976), Fleming (1979) and Turelli (1984) assumed the normal function, $\exp(-cY^2)$, where Y is the deviation from the optimum of the metric character, instead of the quadratic function, $1-cY^2$, to express stabilizing selection, the results are almost the same at equilibrium because the equilibrium phenotypic variance is small compared to $(1/c)$, and in this case the normal function is well approximated by the square function.

2. Model, effects and components of variance

Let $\boldsymbol{X}=(x_1, x_2, \ldots, x_r)^T$ be the vector of phenotypic values of r metric characters of an individual affected by n loci in such a way that gene effects and environmental effects are all additive,

$$\boldsymbol{X}=\boldsymbol{M}+\sum_{i=1}^{n}(\boldsymbol{a}_{i1}+\boldsymbol{a}_{i2})+\boldsymbol{e}. \tag{1}$$

$\boldsymbol{M}$, $\boldsymbol{a}$ and $\boldsymbol{e}$ are the vectors of the means, additive effects of genes and the environmental effects, respectively. The subscripts 1 and 2 distinguish the effects due to the maternally and paternally derived genes. Some loci may not contribute to every metric character, in which case the $\boldsymbol{a}$'s take zero values. The environmental effects are assumed to be multivariate normally distributed, $N(0, V_e)$ and independent of the genotypic values, $\boldsymbol{M}+\sum_{i=1}^{n}(\boldsymbol{a}_{i1}+\boldsymbol{a}_{i2})$. We assume linkage equilibrium and no inbreeding, so that all $\boldsymbol{a}$ vectors are mutually independent. With these assumptions, the variance–covariance matrix, V_X, of $\boldsymbol{X}$ is

$$V_X=2\sum V_{ai}+V_e,$$

where V_{ai} is the variance–covariance matrix of the genic effects at the ith locus. The total genetic variance–covariance matrix for the metric trait is $V_A=2\sum V_{ai}$.

We consider a fitness function $W(\boldsymbol{Y})$ based on $\boldsymbol{Y}=\boldsymbol{X}-\theta$, where θ is the vector of the optimum values, such that

$$W(\boldsymbol{Y})=1-\sum_{k=1}^{r}c_k y_k^2=1-\sum_{k=1}^{r}c_k\left(m_k+\sum_{i=1}^{n}A_{ik}+e_k\right)^2 \tag{2}$$

where m_k, A_{ik} and e_{ik} are the kth elements of the vectors $\boldsymbol{M}-\theta$, $\boldsymbol{A}_i=\boldsymbol{a}_{i1}+\boldsymbol{a}_{i2}$ and $\boldsymbol{e}$, respectively. Although the general form of the fitness function is

$$W(Y)=1-(\tfrac{1}{2})\,\boldsymbol{Y}W^{-1}\boldsymbol{Y},$$

where W is a symmetric positive definite matrix (see Lande, 1980 for the normal version of this type of fitness function), there is an orthogonal transformation T of $\boldsymbol{Y}$ which transforms this expression into the form of (2) without violation of any assumptions already made or any consequent loss of generality. If we view this as an approximation to the normal function used by Lande (1976) and Turelli (1984) in the single character case, c's correspond to their $(2w^2)^{-1}$.

In the following, we use subscripts i, j and h to index loci and k and l to index characters. The summations are from 1 to n for the former and from 1 to r for the latter.

Next we decompose the fitness value into effects and their variances using conditional expectations. To distinguish fitness and components of fitness from the primary metric characters, we use Greek letters to denote the former. The genotypic value, Γ, for the fitness of genotype I is

$$\Gamma=E[W(\boldsymbol{Y})\,|\,I]=1-\sum_k c_k\{(m_k+\sum_i A_{ik})^2+E[e_k^2]\},$$

where $E[\,]$ and $E[\,|\,I]$ denote expectations without conditioning and conditioned on the genotype I, respectively. The mean fitness, μ, for the population is

$$\mu=E[W(\boldsymbol{Y})]=1-\sum_k c_k(m_k^2+V_{Ak}+V_{ek}), \tag{3}$$

where V_{Ak} and V_{ek} are the kth diagonals of the variance–covariance matrices V_A and V_e, respectively. The environmental component, ϵ, of fitness for genotype I is

$$\epsilon=W(\boldsymbol{Y})-E[W(\boldsymbol{Y})\,|\,I]=-\sum_k c_k[2e_k(m_k+\sum_i A_{ik})+e_k^2-V_{ek}],$$

with a variance of

$$E[\epsilon^2\,|\,I]=\sum_k\sum_l c_k c_l\{4(m_k+\sum_i A_{ik})\times(m_l+\sum_i A_{il})E[e_k e_l]+E[e_k^2 e_l^2]-V_{ek}V_{el}\}.$$

This leads to an environmental variance of fitness in the population of

$$\sigma_e^2=\sum_k\sum_l 2c_k c_l\{V_{ekl}^2+2m_k m_l V_{ekl}+2V_{ekl}(\sum_i E[A_{ik}A_{il}])\} \tag{4}$$

where V_{ekl} is the klth element of V_e. The last term in the parentheses may be viewed as genotype by environment interaction variance (Wright, 1935).

https://doi.org/10.1017/S0016672300023934 Published online by Cambridge University Press

The additive fitness effect, α_i, of the ith gene is defined as

$$\alpha_i = E[W(Y)\,|\,i] - \mu = -\sum_k c_k(a_{ik}^2 - V_{aik} + 2m_k a_{ik})$$

where V_{aik} is the kth diagonal element of V_{ai}. The variance of α_i is

$$E[\alpha_i^2] = \sum_k \sum_l c_k c_l\{E[a_{ik}^2 a_{il}^2] - V_{aik} V_{ail} + 4m_k E[a_{ik} a_{il}^2] + 4m_k m_l E[a_{ik} a_{il}]\}.$$

The total additive variance, σ_A^2, for fitness is the sum of these terms over $2n$ genes,

$$\sigma_A^2 = 2\sum_k \sum_l c_k c_l \sum_i \{E[a_{ik}^2\ a_{il}^2] - V_{aik} V_{ail} + 4m_k E[a_{ik} a_{il}^2] + 4m_k m_l E[a_{ik} a_{il}]\}. \quad (5)$$

The dominance and additive-by-additive effects and variances are due to interaction of distinct pairs of genes and can be handled in the same way. Let α_{ij} be the interaction effect for genes i and j. Since

$$E[W(Y)\,|\,i,j] = 1 - \sum_k c_k^2[m_k^2 + 2\sum_h V_{ahk} + V_{ek} + (a_{ik}^2 - V_{aik}) + (a_{jk}^2 - V_{ajk}) + 2a_{ik} a_{jk} + 2m_k(a_{ik} + a_{jk})],$$

then

$$\alpha_{ij} = E[W(Y)\,|\,i,j] - \alpha_i - \alpha_j - \mu = 2\sum_k c_k a_{ik} a_{jk}$$

with variance

$$E[\alpha_{ij}^2] = 4\sum_k \sum_l c_k c_l V_{aik} V_{ajl}.$$

For each pair of loci there are four of these variances, each with the same value, so that the total additive-by-additive variance is

$$\sigma_{AA}^2 = 16\sum_k\sum_l c_k c_l \sum_{i<j}\sum V_{aik} V_{ajl}. \quad (6)$$

There is one dominance variance for each locus, which leads to a total dominance variance of

$$\sigma_D^2 = 4\sum_k \sum_l c_k c_l \sum_i V_{aik} V_{ail}. \quad (7)$$

Since σ_{AA}^2 is the sum of $2n(n-1)$ terms over all characters and σ_D^2 is the sum of n similar terms over all characters, σ_{AA}^2 becomes large compared to σ_D^2 as n increases, as was noted by Wright (1935) for the case of a single character with two alleles per locus.

Interaction effects for three and more genes can be shown to be zero, or alternatively, one can show that the total variance in fitness is composed of

$$\sigma_A^2 + \sigma_D^2 + \sigma_{AA}^2 + \sigma_\epsilon^2.$$

The higher-order interaction variances do not appear because the fitness function $W(Y)$ is quadratic. If the fitness function has higher-order terms, corresponding higher interaction variances of fitness will be non-zero.

With appropriate data on fitness or some of its components, the genetic variance components of fitness can be estimated utilizing covariances of relatives (Cockerham, 1963, 1980).

3. Equilibrium structures for the 'Gaussian' and the 'House of cards' approximation

To evaluate the genetic variance components of fitness, and their equilibrium values, we need the moments for genotypic value of the metric characters at each locus. These moments are calculated from the equilibrium functions obtained by the 'Gaussian' and the 'House of Cards' distribution approximations (Lande, 1980; Turelli, 1984, 1985). For both approximations, the equilibrium means, m_k, and the third moments, $E[a_{ik}^2 a_{il}]$, are zero if the effects of mutation are symmetric as originally assumed, but the two approximate distributions are distinguished by their kurtosis. Let Φ_{ik} be the measure of kurtosis for the kth character at the ith locus defined as

$$\Phi_{ik} = E[a_{ik}^4]/3\,V_{aik}^2.$$

For the Gaussian model this is one. For the House of Cards model, an analytical form is not available for the general r-character case, but for the single-character case, corresponding to the case of no pleiotropy, from (3·15) of Turelli (1984) an approximation is obtained as

$$\Phi_{ik} = c_k V_{mik}/3u_i(1 + 2c_k V_{ek}), \quad (8)$$

where u_i is the mutation rate at the ith locus and V_{mik} is the variance of the effect of a mutation on the kth character at that locus. Where the House of Cards model is applicable, $V_{mik} \gg V_{aik} \simeq u_i(1 + 2c_k V_{ek})/c_k$ and Φ_i is large compared to one. For the case of two characters, a reviewer pointed out that, from equation (2.5) in the appendix of Turelli (1985), the kurtosis (Φ_{i1}) of the first character is expressed as

$$\Phi_{i1} = (2\beta_i + 1)\,c_1 V_{mi1}/3u_i(1 + 2c_1 V_{e1})$$

where $\beta_i = (c_2 V_{mi2}/c_1 V_{mi1})^{\frac{1}{2}} > 0$, if the effects of both mutation and environment on the two characters are uncorrelated. Thus the kurtosis becomes even larger compared to the single-character case with the same parameters. Although the general formula for Φ_{ik} is unknown, a large kurtosis can be expected to be a general feature for the House of Cards approximation.

Utilizing the measure of kurtosis and the other equilibrium conditions, the genetic variance components of fitness at equilibrium are

$$\sigma_A^2 = 2\sum_k \sum_l c_k c_l \sum_i \pi_{ikl}[(3\Phi_{ik} - 1)(3\Phi_{il} - 1)]^{\frac{1}{2}} V_{aik} V_{ail}$$

$$\sigma_D^2 = 4\sum_k \sum_l c_k c_l \sum_i V_{aik} V_{ail}$$

$$\sigma_{AA}^2 = 16\sum_k \sum_l c_k c_l \sum_{i<j}\sum V_{aik} V_{ajl}$$

where $\pi_{ikl}\,(<1)$ is the correlation coefficient of a_{ik}^2 and a_{il}^2. From these expressions, we see that the magnitude

https://doi.org/10.1017/S0016672300023934 Published online by Cambridge University Press

of σ_A^2 relative to σ_D^2 decreases as the number of characters increases unless all π's are one.

For the Gaussian approximation ($\Phi = 1$), σ_A^2 is not greater than σ_D^2. In the evaluation of σ_{AA}^2, note that there are probably a lot of zeros among V_{Ai}'s if we consider all the loci contributing to fitness through metric characters. To assess the magnitude of σ_{AA}^2 in comparison with σ_A^2 and σ_D^2, let us consider an idealized situation where the loci and quantitative characters contributing to fitness are divided into subsets, each subset of loci controlling one of the subsets of characters. In other words, these subsets of loci have no pleiotropic connexions or fitness interactions with each other. There may be many such subsets. Because of the assumption of linkage equilibrium, the variance components of fitness measurable are sums of the variance components for respective subsets of loci and the relative magnitude of the total σ_{AA}^2 to the total σ_A^2 or σ_D^2 is the ratio of the sums over these subsets and in this sense an average over them. Thus, we consider one of these subsets and use the same notations n and r to denote the numbers of loci and characters of this limited subset of loci, respectively. If $V_{aik} = V_{ak}$ and $\pi_{ikl} = \pi_{kl}$ for all loci

$$\sigma_A^2 = 4n \sum_k \sum_l c_k c_l \pi_{kl} V_{ak} V_{al} < 4n \sum_k \sum_l c_k c_l V_{ak} V_{al} = \sigma_D^2,$$

$$\sigma_{AA}^2 = 8n(n-1) \sum_k \sum_l c_k c_l V_{ak} V_{al} = 2(n-1)\sigma_D^2 \geqslant 2(n-1)\sigma_A^2.$$

For the House of Cards approximation, on the other hand, when there are equal variances, covariances and kurtosis for all loci,

$$\sigma_A^2 = 2n \sum_k \sum_l c_k c_l \pi_{kl} \{(3\Phi_k - 1) \times (3\Phi_l - 1)\}^{\frac{1}{2}} V_{Ak} V_{al},$$

and the relative magnitude of the additive variance to other components therefore depends on kurtosis. If the additional assumption is made that all characters are symmetric, so that $c_k = c$, $V_{ak} = V_a$, $\pi_{kl} = \pi\,(k \neq l)$ and $\Phi_k = \Phi$, then the variance components of fitness are

$$\sigma_A^2 = 2n(3\Phi - 1)[1 + (r-1)\pi] r c^2 V_a^2$$
$$\sigma_D^2 = 4n c^2 r^2 V_a^2$$
$$\sigma_{AA}^2 = 8n(n-1) c^2 r^2 V_a^2.$$

Thus, σ_A^2 is larger than σ_{AA}^2 if

$$\Phi > \frac{4(n-1)r + (r-1)\pi + 1}{3[1 + (r-1)\pi]}.$$

For example, in the single-character case ($r = 1$), if $u = 10^{-5}$, $V_e = 1$, $c = 0{\cdot}026$, $V_m = 0{\cdot}05$, which are within the plausible range according to Turelli (1984), $\Phi = 41$ from the approximate formula (8) in comparison to the exact value of 47 obtained by numerical methods. In any case, Φ is large and σ_A^2 is larger than σ_{AA}^2 for $n < 31$. The number of loci under which σ_A^2 is larger than σ_{AA}^2 will decrease as the number, r, of characters with pleiotropy increases unless the correlation coefficient, π, is one.

In summary, in the parameter range where the 'Gaussian' approximation applies, the additive variance is equal to or smaller than the dominance variance, and both are much smaller than the additive-by-additive variance. In contrast, in the parameter range where the 'House of Cards' approximation applies, the additive variance is much larger than the dominance variance and larger than the additive-by-additive variance, unless the number of loci and the number of pleiotropic characters becomes too large.

4. Discussion

In order to make inferences about the maintenance mechanism from experimental results of fitness components such as viability and fecundity, not fitness itself, we must make two assumptions. The first one is that a significant part of the fitness component is due to the type of stabilizing selection acting on metric characters described in the present paper. In a mutation accumulation experiment, Mukai (1969) showed that the decrease of average viability of the lines had a significant quadratic component with respect to the number of generations of accumulation, although later experiments with shorter accumulation periods did not show this trend (Mukai *et al.* 1972; Ohnishi, 1977). Secondly, it is well established that the average viability of homozygous lines is lower than that of heterozygous lines (see Dobzhzansky, 1970). In the present model, if the inbreeding coefficient, F, is not zero, then (3) becomes

$$\mu(F) = 1 - \sum_k c_k [m_k^2 + (1+F) V_{Ak} + V_e]. \qquad (9)$$

Thus we expect inbreeding depression, as noted by Lande & Schemske (1985). Finally, genotype–environment interaction expected from (4) has been reported in viability (Dobzhansky & Levene, 1955; Tachida & Mukai, 1985). Although these characteristics are also explained by other models, they are compatible with the present model.

The second assumption is that fitness components for those loci controlling these metric characters are proportional to their total fitness. This is a somewhat controversial issue. Simmons, Preston & Engels (1980) and Yamazaki & Hirose (1984) reported that the correlation coefficient between the total fitness and the viability is not significantly different from zero, whereas Mackay (1986) reported otherwise in homozygous lines. If the correlation is not substantial, the mean of the metric character may be different from the optimum. We will return to this point later.

In a series of experiments conducted by Mukai's group, the additive and dominance variances of viability at the chromosomal level, denoted by ${}_M\sigma_A^2$ and ${}_M\sigma_D^2$, were estimated using *Drosophila mel-*

https://doi.org/10.1017/S0016672300023934 Published online by Cambridge University Press

anogaster captured from several natural populations (Mukai *et al.* 1974; Tachida *et al.* 1983; Mukai & Nagano, 1983; Kusakabe & Mukai, 1984). These two variances are related to the variance components of the present model in the following way:

$$\left.\begin{aligned} {}_M\sigma_A^2 &= \sigma_A^2 + \sigma_{AA}^2/2 \\ {}_M\sigma_D^2 &= \sigma_D^2 + \sigma_{AA}^2/2 \end{aligned}\right\} \quad (10)$$

From table 1 of Mukai (1985), which summarizes the results of these experiments with other unpublished results, we see that ${}_M\sigma_A^2$ is always greater than ${}_M\sigma_D^2$ (ten times greater except for one population). This implies $\sigma_A^2 \gg \sigma_D^2, \sigma_{AA}^2$.

Rose & Charlesworth (1981) estimated the additive and dominance variances of fecundity and longevity of *Drosophila melanogaster* using sib analysis. Since longevity is clearly not proportional to fitness, we will discuss their results on fecundity. Their additive and dominance variances, ${}_R\sigma_A^2$ and ${}_R\sigma_D^2$, are related to the variance components as follows:

$$\left.\begin{aligned} {}_R\sigma_A^2 &= \sigma_A^2 + \overline{[1+(1-2R)^2]}\sigma_{AA}^2/4 \\ {}_R\sigma_D^2 &= \sigma_D^2 + \sigma_{AA}^2/2 \end{aligned}\right\}, \quad (11)$$

where R is the recombination rate between a pair of loci, and the overbar in the first equation is used to indicate an average over all pairs of loci. Since $(1-2R)^2$ is less than or equal to one, their result shows that ${}_R\sigma_A^2 \gg {}_R\sigma_D^2$, which again implies $\sigma_A^2 \gg \sigma_D^2, \sigma_{AA}^2$.

From the previous section, the parameter range where the 'Gaussian' approximation applies predicts $\sigma_A^2 \ll \sigma_{AA}^2$ for fitness, whereas that where the 'House of Cards' approximation applies predicts various relationships depending on the relative magnitude of the number of loci to kurtosis. The two experiments mentioned above indicate that the 'Gaussian' approximation is not applicable if the two assumptions made at the beginning of this section hold. They are compatible with the 'House of Cards' approximation if the numbers of loci and pleiotropic characters are not too large compared to the kurtosis of the distribution of the genetic value of the metric characters at each locus. Here the number of loci means the number of loci of a subset of all loci contributing to fitness, explained in the previous section. Of course, all loci may be interrelated and it may not be possible to divide them into unrelated subsets. However, we expect many elements of V_{ai}'s are zero and therefore the number of loci of each effectively unrelated subset is relatively small compared to the total number of loci contributing to fitness. The number of loci in this sense is not well known at present. Some experiments show that total genomic mutation rates for metric characters are of the order of 0·01 in maize and mouse (Russel, Sprague & Penny, 1963; Hoi-Sen, 1972). If we assume that the per locus mutation rate is less than 10^{-4}, the number of loci becomes greater than 100, in which case the 'House of Cards' approximation may also have larger additive-by-additive variance and is incompatible with these experiments. However, there are some uncertainties in interpreting these data for high mutation rates (see the discussion on pp. 175–178 of Turelli, 1984).

In the present analysis we assumed linkage equilibrium among loci. Since a slight linkage disequilibrium is known to occur in these models (Lande, 1976), it is necessary to evaluate the effect of it on our conclusion. This can be done for the 'Gaussian' approximation for one character using the covariances between genic effects calculated by Lande (1976). The effect is found to be negligible unless linkage among loci is very tight. Because the calculation is very technical, it is given in the Appendix. We may expect that this is also true for the 'House of Cards' approximation.

If the means of metric characters are different from the optima, inflation of the additive variance through the fourth term on the right-hand side of Equation (5) occurs. This may happen if the mean is at another equilibrium (the existence and stability of these other equilibria were shown for the two-allele model by Barton, 1986), or if the mean effect of mutation is different from zero (Lande, 1980). Generally speaking, if the magnitude of the square of the deviation is of the order of the genetic variance due to one locus, the additional term is of the same order as the first term in (5) in the Gaussian case and the conclusion does not change. However, if the deviation is large, the additive variance for fitness may be larger than the additive-by-additive variance also in the Gaussian case. Then the comparison of the additive and additive-by-additive genetic variances does not have the power to discriminate between these two approximations.

One of the major limitations of this study is that we used data on one of the components of fitness rather than the total fitness for which the present theory was developed, because data on total fitness are not available. Thus, our conclusion with regard to the appropriateness of the two approximations depends critically on the assumption that the fitness component measured is proportional to total fitness. This is very difficult to test because of the polygenic nature of the genes contributing to fitness.

Appendix

Effects of linkage disequilibrium on the estimation of variance components when the Gaussian approximation is appropriate

Here we demonstrate that linkage disequilibrium does not affect the conclusion significantly in the 'Gaussian' approximation unless linkage among loci is very tight. Our strategy is to show that the covariances of relatives used in the experiments are not affected much by linkage disequilibrium. We consider Mukai's

https://doi.org/10.1017/S0016672300023934 Published online by Cambridge University Press

experiments. In his experiments, ${}_M\sigma_A^2$ and ${}_M\sigma_D^2$ are estimated by the variance $\{\mathrm{var}[W(X)]\}$ and the covariance $\{\mathrm{cov}[W(X), W(Y)]\}$ between two individuals, X and Y, who share one second chromosome using the following formulas:

$$\begin{aligned} {}_M\sigma_A^2 &= 2\,\mathrm{cov}[W(X), W(Y)] \\ {}_M\sigma_D^2 &= \mathrm{var}[W(X)] - 2\,\mathrm{cov}[W(X), W(Y)]. \end{aligned} \quad \text{(A 1)}$$

Under linkage equilibrium, these variances and covariances are expressed in terms of the additive, dominance and additive-by-additive variances and we obtain relationship (10). Here we calculate $\mathrm{var}[W(X)]$, taking linkage disequilibrium into account under the 'Gaussian' approximation, and show that it does not differ much from that under linkage equilibrium. The same is true for $\mathrm{cov}[W(X), W(Y)]$.

For simplicity we consider one characer and also ignore the effect of the environment. Since the mean of the character is zero in the 'Gaussian' approximation, the fitness function is

$$W(Y) = 1 - c\left[\sum_{i=1}^{n} (a_{i1} + a_{i2})\right]. \quad \text{(A 2)}$$

In order to make the distinction clear, we use $\epsilon[\,]$ to denote an expectation under linkage disequilibrium. The underlying distribution is a multivariate normal. Then, after straightforward though tedious calculations, we obtain

$$\begin{aligned} \mathrm{var}[W(Y)] = c^2\Big(&2\sum_i \{\epsilon[a_{i1}^4] - (\epsilon[a_{i1}^2])^2\} + 4\sum_i \epsilon[a_{i1}^2]^2 \\ &+ 8\sum_{i\neq j}\sum \epsilon[a_{i1}^2]\,\epsilon[a_{j1}^2] \\ &+ \sum_{i\neq j}\sum \{8\epsilon[a_{i1}^3 a_{j1}] \\ &+ 6(\epsilon[a_{i1}^2 a_{j1}^2] - \epsilon[a_{i1}^2]\epsilon[a_{j1}^2]) \\ &+ 8\epsilon[a_{i1}^2]\,\epsilon[a_{i1}\,a_{j1}] + 4(\epsilon[a_{i1}\,a_{j1}])^2\} \\ &+ 4\sum_{i\neq j\neq k}\sum\sum \{3\epsilon[a_{i1}^2\,a_{j1}\,a_{k1}] \\ &+ \epsilon[a_{i1}^2]\,\epsilon[a_{j1}\,a_{k1}] \\ &+ 2\epsilon[a_{i1}\,a_{j1}]\,\epsilon[a_{i1}\,a_{k1}]\} \\ &+ \sum_{i\neq j\neq k\neq l}\sum\sum\sum \{\epsilon[a_{i1}\,a_{j1}\,a_{k1}\,a_{l1}] \\ &+ \epsilon[a_{i1}\,a_{j1}]\,\epsilon[a_{kl}\,a_{l1}]\}\Big). \end{aligned} \quad \text{(A 3)}$$

Note that the first three terms are the additive, dominance and additive-by-additive variances, respectively, calculated using the one-locus marginal distribution [see (5), (6) and (7) of the text]. The remaining terms are due to linkage disequilibrium. There are $n(n-1)$, $n(n-1)(n-2)$ or $n(n-1)(n-2)(n-3)$ terms for the summation with two, three or four $\sum$s.

To evaluate these, we calculate moments of a's using the multivariate normal distribution obtained by Lande (1976). For simplicity, we assume that all loci are equivalent. Then, from (21 *a*) and (21 *b*) of Lande (1976), the correlation coefficient, ρ, of two a's at different loci is

$$\rho = -\frac{(V_m)^{\frac{1}{2}}}{AR} \quad \text{and} \quad A = \left(\frac{1}{2c}\right) + V_e + V_g + (n-1)\frac{(V_m)^{\frac{1}{2}}}{R}, \quad \text{(A 4)}$$

where R and V_g are the recombination rate between two loci and the genetic variance of the quantitative character, respectively, and other parameters are the same as in the text. For the 'Gaussian' approximation to be appropriate, one of the necessary conditions listed in Lande (1976) is

$$(V_m)^{\frac{1}{2}} \ll AU/n$$

where $U = \sum u_i$ is the gametic mutation rate [see (24) of Lande, 1976]. Thus

$$|\rho| \ll U/(nR). \quad \text{(A 5)}$$

Using the correlation coefficient, ρ, and $V_a = \epsilon[a_{i1}^2]$, the second-last term in (A 3) is

$$\begin{aligned} c^2 \sum_{i\neq j\neq k\neq l}\sum\sum\sum \epsilon[a_{i1}\,a_{j1}\,a_{k1}\,a_{l1}] &= 3c^2 n(n-1)\times(n-2) \\ \times(n-3)\,\rho^2 V_a^2 &\ll 3c^2 n(n-1)\,(U/R)^2\,V_a^2. \end{aligned} \quad \text{(A 6)}$$

The magnitude of U is thought to be of the order 10^{-2} (see Turelli, 1984) and the additive genetic variance is $8c^2 n(n-1)\,V_a^2$. Therefore, unless the gametic mutation rate is much greater than the recombination rate, the contribution of this term to $\mathrm{var}[W(Y)]$ is negligible compared to that of the additive-by-additive genetic variance of fitness. All other terms due to linkage disequilibrium are shown to be negligible in a similar way.

We thank Drs R. Lande, M. Turelli, B. S. Weir, T. Wright, Z. Zeng and two anonymous reviewers for their constructive comments on the manuscript. Paper no. 10907 of the journal series of the North Carolina Agricultural Research Service, Raleigh, NC 27695-7601. This investigation was supported in part by NIH Research Grant no. 11546 from the National Institute of General Medical Sciences.

References

Barton, N. H. (1986). The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genetical Research* **47**, 209–216.

Cockerham, C. Clark (1963). Estimation of genetic variances. In *Statistical Genetics and Plant Breeding* (ed. W. D. Hanson and H. F. Robinson), pp. 53–94. National Academy of Sciences and National Research Council Publication No. 982, Washington, D.C., USA.

Cockerham, C. Clark (1980). Random and fixed effects in plant genetics. *Theoretical and Applied Genetics* **56**, 119–131.

Dobzhansky, Th. (1970). *Genetics of the Evolutionary Process*. New York: Columbia University Press.

Dobzhansky, Th. & Levene, H. (1955). Genetics of natural populations. XXIV. Developmental homeostasis in natural populations of *Drosophila pseudobscura*. *Genetics* **40**, 779–808.

Fleming, W. H. (1979). Equilibrium distributions of continuous polygenic traits. *SIAM Journal of Applied Mathematics* **36**, 148–168.

Hoi-Sen, Y. (1972). Is subline differentiation a continuing process in inbred strains of mice? *Genetical Research* **19**, 53–59.

Kimura, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences, U.S.A.* **54**, 731–736.

Kusakabe, S. & Mukai, T. (1984). The genetic structure of natural populations of *Drosophila melanogaster*. XVII. A population carrying genetic variability explicable by the classical hypothesis. *Genetics* **108**, 393–408.

Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research* **26**, 221–235.

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutation. *Genetics* **94**, 203–215.

Lande, R. & Schemske, D. W. (1985). The evolution of self-fertilization and inbreeding depression in plants, I. Genetic models. *Evolution* **39**, 29–40.

Latter, B. D. H. (1960). Natural selection for an intermediate optimum. *Australian Journal of Biological Sciences* **13**, 30–35.

Mackay, T. F. C. (1986). A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster*. *Genetical Research* **47**, 59–70.

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

Mukai, T. (1985). Experimental verification of the neutral theory. In *Population Genetics and Molecular Evolution* (ed. T. Ohta & K. Aoki). Berlin: Springer.

Mukai, T., Cardellino, R. A., Watanabe, T. K. & Crow, J. F. (1974). The genetic variance for viability and its components in a local population of *Drosophila melanogaster*. *Genetics* **78**, 1195–1208.

Mukai, T., Chigusa, S. I., Mettler, L. E. & Crow, J. F. (1972). Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 335–355.

Mukai, T. & Nagano, S. (1983). The genetic structure of natural populations of *Drosophila melanogaster*. XVI. Excess of additive genetic variance of viability. *Genetics* **105**, 115–134.

Ohnishi, O. (1977). Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effect of polygenic mutations. *Genetics* **87**, 529–545.

Rose, M. R. & Charlesworth, B. (1981). Genetics of life history in *Drosophila melanogaster*. I. Sib analysis of adult females. *Genetics* **97**, 173–186.

Russel, W. A., Sprague, G. F. & Penny, L. H. (1963). Mutations affecting quantitative characters in long-time inbred lines of maize. *Crop Science* **3**, 175–178.

Simmons, M. J., Preston, C. R. & Engels, W. R. (1980). Pleiotropic effects on fitness of mutations affecting viability in *Drosophila melanogaster*. *Genetics* **94**, 467–475.

Tachida, H., Matsuda, M., Kusakabe, S. & Mukai, T. (1983). Variance component analysis for viability in an isolated population of *Drosophila melanogaster*. *Genetical Research* **42**, 207–217.

Tachida, H. & Mukai, T. (1985). The genetic structure of natural populations of *Drosophila melanogaster*. XIX. Genotype–environment interaction in viability. *Genetics* **111**, 43–55.

Turelli, M. (1984). Heritable genetic variation via mutation–selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology* **25**, 138–193.

Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation–selection balance for polygenic traits. *Genetics* **111**, 165–195.

Turelli, M. (1986). Gaussian versus non-gaussian genetic analysis of polygenic mutation–selection balance. In *Evolutionary Process and Theory* (ed. S. Karlin and E. Nevo), pp. 607–628. New York: Academic Press.

Wright, S. (1935). The analyses of variance and the correlation between relatives with respect to deviations from an optimum. *Journal of Genetics* **30**, 243–256.

Yamazaki, T. & Hirose, Y. (1984). Genetic analysis of natural populations of *Drosophila melanogaster*. II. The measurement of fitness and fitness components in homozygous lines. *Genetics* **108**, 213–221.