Journal Compilation. © 2008 Cambridge University Press. *Genet. Res., Camb.* (2007), **89**, *pp.* 373–387. 373
doi:10.1017/S0016672308009555 Printed in the United Kingdom

Genet. Res., Camb. (1976), **26**, *pp.* 221–235

The maintenance of genetic variability by mutation in a polygenic character with linked loci

By RUSSELL LANDE

Museum of Comparative Zoology, Harvard University, Cambridge, Mass. 02138, U.S.A.

(*Received* 17 *February* 1975)

SUMMARY

It is assumed that a character under stabilizing selection is determined genetically by n linked, mutable loci with additive effects and a range of many possible allelic effects at each locus. A general qualitative feature of such systems is that the genetic variance for the character is independent of the linkage map of the loci, provided linkage is not very tight. A particular detailed model shows that certain aspects of the genetic system are moulded by stabilizing selection while others are selectively neutral. With reference to experimental data on characters of *Drosophila* flies, maize, and mice, it is concluded that large amounts of genetic variation can be maintained by mutation in polygenic characters even when there is strong stabilizing selection. The properties of the model are compared with those of heterotic models with linked loci.

1. INTRODUCTION

It is now generally accepted that the evolution of species usually proceeds by a continuous transformation of their characters. The study of extant populations shows that these characters often display considerable heritable variation and typically are influenced by many genes of small effect (Falconer, 1960). The question of how heritable variation for such characters is maintained is still largely unresolved.

It has been argued that most polygenic or quantitative characters are under natural selection for an intermediate optimum phenotype, and such stabilizing selection has been demonstrated for many characters (Rendel, 1943; Lerner, 1954; Allard & Jain, 1962; Lack, 1966). Fisher (1930) showed theoretically that stabilizing selection depletes genetic variability in polygenic characters and this has been confirmed by other workers both theoretically (Robertson, 1956; Lewontin, 1964*b*) and experimentally (e.g. Waddington, 1960; Scharloo, 1964; Gibson & Bradley, 1974). Lewontin (1964*b*) found that with directional dominance many genes can be kept segregating by stabilizing selection, but that only a limited amount of genetic variance can be maintained (less than with a single locus). Gale & Kearsey (1968) and Kearsey & Gale (1968) gave examples where two and three linked loci can be kept segregating if they have very unequal effects; this requirement implies that even with more loci the effective number of loci must be small and that only a

limited amount of genetic variance could be maintained. Such schemes cannot account for the genetic variability in typical polygenic characters which have large effective numbers of loci (Falconer, 1960).

It is usually thought that mutation cannot by itself offset the loss of genetic variability caused by stabilizing selection to maintain the levels of heritable variation observed in natural artificial populations. Selective mechanisms for maintaining genetic variability are often invoked, usually heterozygote advantage and heterogeneous environments. But even in nearly constant, uniform environments sexual populations are known to maintain considerable genetic variation for quantitative characters. An example is the Kaduna population of *Drosophila melanogaster* which has been kept in such conditions for hundreds of generations (Lopez-Fanjul & Hill, 1973). Heterozygote advantage or 'heterosis' is a commonly observed phenomenon but there are two possible explanations for it (East & Jones, 1919). The first is heterozygote superiority at many individual loci and the second is that different inbred lines are homozygous for different partially recessive deleterious genes. It is particularly important to distinguish between these two possible causes of hybrid vigour when considering the maintenance of genetic variability as the first can contribute to the amount of variability maintained while the second cannot. It is well established that many deleterious recessive and partly recessive genes exist in normal outbred individuals. Furthermore, inbred lines have fitnesses as great as wild type under some conditions (East & Jones, 1919; Dobzhansky, Holtz & Spassky, 1942; Dobzhansky & Spassky, 1944). Many species of plants preferentially inbreed to various degrees when there is the possibility of increased outbreeding. These observations seem to argue against any general heterosis at the level of single loci; and very few cases of single locus heterozygote advantage are known (Lewontin, 1974).

Because of computational difficulties, all analytic models and simulations of polygenic systems have ignored either linkage or mutation. All theoretical knowledge of the behaviour of linked genes has thus been derived from models which have only selective mechanisms for the maintenance of genetic variability. This has diverted attention from the possibility that mutation may be a powerful force maintaining genetic variation in polygenic characters and fostered the notion that heterozygote advantage at the level of single loci is necessary to explain observed amounts of genetic variation in characters of adaptive significance.

Here a model of a phenotypic character under stabilizing selection influenced by many linked, mutable genes is investigated. Because stabilizing selection acts not to preserve but to destroy genetic variation, this provides a framework for evaluating the power of mutation to maintain genetic variability in polygenic characters. This evaluation is made in light of the available experimental evidence on mutation in quantitative characters. The effect of linkage on the correlations of alleles at different loci is also derived and a comparison is made with previous models of linkage that incorporate selective mechanisms for maintaining genetic variability.

2. ASSUMPTIONS

In modelling a quantitative character it is certainly more realistic, at least in large populations, to consider that at each locus there are many alleles with a range of effects rather than simply two effects, + and −, as is usually assumed. This is evident from the molecular structure of proteins which allows a potentially large number of amino acid substitutions which would have some effect on the function of the molecule and therefore on the phenotype as well. The first assumption concerns the mutation process.

(1) At the ith locus, each allele mutates with probability μ_i each generation, and the distribution of mutant effects is the same for all alleles. The variance of the mutational changes at the ith locus is denoted as m_i^2. At each locus, the changes produced by mutation have a mean of zero so that there is no mutation pressure influencing the population mean.

To simplify matters it is further assumed that:

(2) The population mates randomly with respect to the character.

(3) The effective population size is large enough to ignore genetic drift.

(4) There is no genotype-environment interaction and the allelic effects are additive within and between loci.

3. QUALITATIVE ANALYSIS

Consider a metrical character affected by n diploid loci. The effect of an allele at locus i is denoted x_i. The covariance of allelic effects between locus i and locus j in the gametes in generation t will be written as $C_{ij}(t)$. From the assumptions it is easily shown that mutation changes the $C_{ij}(t)$ by an amount $\delta_{ij}\mu_i m_i^2$ per generation, where $\delta_{ij} = 1$ if $i = j$ and is zero otherwise. Thus mutation increases the variances but does not alter the covariances. This is not intended as a universal description of the mutation process. In reality, there must be a limit to the range of allelic effects and hence the genetic variance at each locus. If the genetic variance at a locus were at its maximum or saturation level, mutation could not increase the genetic variance at that locus and would decrease the magnitude of the covariances between that locus and the other loci. Such extreme saturation must be rare because mutagenic agents generally increase the genetic variance of quantitative characters. There is some experimental evidence that the amount of new genetic variance produced by mutation each generation is roughly constant independent of the background level of genetic variance. The data of Scossiroli & Scossiroli (1959), Yamada & Kitagawa (1961) and Clayton & Robertson (1964) show that the genetic variance produced in quantitative characters of *Drosophila* by a given dose of X-rays is approximately the same in lines with low and high heterozygosity. No other evidence bearing on this point was found. The present description of mutation may thus be of some relevance to natural populations.

The above property of mutation will now be used to derive a qualitative feature of systems of linked mutating loci under stabilizing selection. The assumption of

additive effects allows each gamete to be assigned a value of $\sum_{i=1}^{n} x_i$. Stabilizing selection, in favouring those gametes with values closest to the optimum, produces negative correlations in allelic effects at closely linked loci, so that positive and negative deviations from the optimum tend to cancel (Mather, 1941; Lewontin, 1964*b*). This creates a pool of hidden genetic variation which is stored in linked combinations.

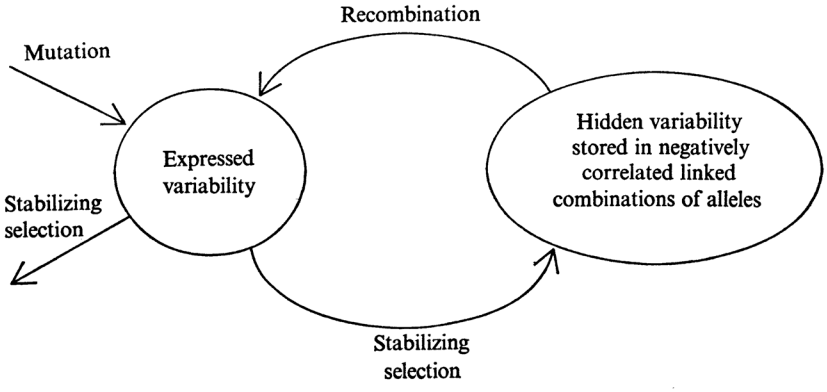

Fig. 1. Flow diagram of genetic variability. See text for explanation.

The influence of stabilizing selection, mutation and recombination on the expressed and hidden genetic variance can be summarized as follows: stabilizing selection converts expressed variance into hidden variance and depletes expressed variation. Mutation does not alter the hidden variation but contributes directly to the expressed genetic variance. Recombination decreases the correlations between loci, converting hidden variation into expressed variation. These relations are depicted in a flow diagram of genetic variability in Fig. 1. When this system has reached equilibrium, the flow out of expressed variation into hidden variation must equal the flow into expressed variation from hidden variation, regardless of the rate of recombination of the loci. The amount of expressed variability is then determined only by the input from mutation and the output on the left due to stabilizing selection. Therefore if there are no saturation effects, the linkage structure of the loci will have no influence on the amount of expressed genetic variance maintained by mutation.

4. A DETAILED MODEL

Kimura (1965) analyzed a one locus, multiallelic model of mutation and stabilizing selection. He used continuous time and a quadratic deviations fitness function. With the previous assumptions and the additional assumption that the changes in allelic effect produced by mutation are small, his model shows that the equilibrium distribution of allelic effects is approximately normal with the average effect at the

optimum. Latter (1970) investigated a discrete generation model of a single locus with mutation and a Gaussian fitness function. Using Kimura's result, he derived a simple formula for the amount of genetic variance maintained. The approach of Latter is extended here to describe many linked loci.

A population with discrete generations is considered in which the cyclic order of events is selection, recombination, mutation and random mating. A primed letter is used to denote loci from one gamete, say the paternal gamete, the unprimed letters denote loci from the maternal gamete. A subscript w indicates that selection has operated.

In zygotes formed by random mating in generation t, there may be a covariance between the effects of alleles at distinct loci i and j in the same gamete (in terms of covariances, $C_{ij}(t) = C_{i'j'}(t) \neq 0$), but there is no covariance between the effects of alleles from different gametes ($C_{ij'}(t) = C_{i'j}(t) = 0$). Selection creates covariances between the effects of alleles from different gametes, which are denoted as

$$b_{ij}(t) = (C_{ij'}(t))_w = (C_{i'j}(t))_w \neq 0.$$

If locus i and locus j recombine, in the new gametes i will be linked to j' and i' will be linked to j, and the recombined fraction will have covariance $b_{ij}(t)$ while the non-recombined fraction will have covariance $(C_{ij}(t))_w$. Letting the recombination fraction between locus i and locus j be r_{ij}, the effect of selection, recombination and mutation on the variances and covariances is summarized as

$$C_{ij}(t+1) = (1-r_{ij})\,(C_{ij}(t))_w + r_{ij}\,b_{ij}(t) + \delta_{ij}\mu_i m_i^2. \tag{1}$$

A specific model of selection will be adopted to express the variables $(C_{ij}(t))_w$ and $b_{ij}(t)$ in terms of the variables $C_{ij}(t)$ to solve these equations.

The stabilizing selection which acts on the phenotypes is taken to be a Gaussian function so that fitness decreases with deviation, z, away from the optimum as

$$W(z) = \exp\left\{\frac{-z^2}{2w^2}\right\}. \tag{2}$$

Roughly w gives the width or range of phenotypes around the optimum which have a high fitness. A large w indicates weak stabilizing selection and a small w indicates strong stabilizing selection. In keeping with the assumption of no genotype-environment interaction, the distribution of phenotypes, z, among individuals with a genotype value X, is taken to be normal with a variance of σ_e^2 which is the environmental variance.

$$p(z|X) = \frac{1}{\sqrt{(2\pi\sigma_e^2)}} \exp\left\{\frac{-(z-X)^2}{2\sigma_e^2}\right\}. \tag{3}$$

The relative fitness of individuals with genotype X is

$$\tilde{W}(X) \propto \int W(z)\,p(z|X)\,dz \propto \exp\left\{\frac{-X^2}{2(w^2+\sigma_e^2)}\right\}. \tag{4}$$

The environmental variance, σ_e^2, acts in conjunction with the strength of selection on the phenotypes to limit the strength of selection on the genotypes; even for very

strong selection on the phenotypes (small w) the width of the fitness function acting on the genotypes cannot be less than σ_e.

The joint distribution of allelic effects in maternal gametes is denoted as

$$f(x,t) = f(x_1, \ldots, x_n, t)$$

and in paternal gametes as $f(x',t) = f(x_1', \ldots, x_n', t)$. It is assumed that there is no sexual dimorphism so that $f(x,t) = f(x',t)$ for $x = x'$. With random mating the joint distribution of allelic effects in zygotes is $F(x,x',t) = f(x,t)f(x',t)$. Since the genes are additive $X = \sum_{i=1}^{n}(x_i + x_i')$ and the distribution of alleles in zygotes after selection is

$$F_w(x,x',t) \propto f(x,t)f(x',t)\,\bar{W}\left(\sum_{i=1}^{n}(x_i + x_i')\right). \tag{5}$$

It will be shown later than under certain conditions the distribution of allelic effects in gametes is approximately multivariate normal and so can be written

$$f(x,t) \propto \exp\{-\tfrac{1}{2}(x-\bar{x}(t))\,C^{-1}(t)\,(x-\bar{x}(t))^T\}, \tag{6}$$

where $C^{-1}(t)$ is the inverse of the variance–covariance matrix $C(t)$ and

$$\bar{x}(t) = (\bar{x}_1(t), \ldots, \bar{x}_n(t))$$

is the vector of mean effects of alleles at the various loci. The superscript T denotes transpose.

The distribution of allelic effects in zygotes after selection is then

$$F_w(x,x',t) \propto \exp\left\{-\tfrac{1}{2}\left[(x-\bar{x}(t))\,C^{-1}(t)\,(x-\bar{x}(t))^T + (x'-\bar{x}(t))\,C^{-1}(t)\,(x'-\bar{x}(t))^T + \frac{1}{w^2+\sigma_e^2}\,(x\mathbf{1}x^T + 2x'\mathbf{1}x^T + x'\mathbf{1}x'^T)\right]\right\}, \tag{7}$$

where $\mathbf{1}$ is a matrix every element of which is 1. This expression has the multivariate normal form

$$F_w(x,x',t) \propto \exp\{-\tfrac{1}{2}(x-\bar{x}_w(t), x'-\bar{x}_w(t))\,K^{-1}(t)\,(x-\bar{x}_w(t), x'-\bar{x}_w(t))^T\}, \tag{8}$$

where $K^{-1}(t)$ is the inverse of the variance–covariance matrix of allelic effects within and between gametes in the selected zygotes and $\bar{x}_w(t)$ is the vector of mean effects after selection. Because of the symmetry between maternal and paternal gametes the matrix $K(t)$ must be of the form

$$K(t) = \begin{pmatrix} C_w(t) & b(t) \\ b(t) & C_w(t) \end{pmatrix}, \tag{9}$$

where $C_w(t)$ and $b(t)$ are symmetric. Separating linear and quadratic terms in (7) and (8) and ignoring constant factors it is found that

$$K^{-1}(t) = \begin{pmatrix} C_w(t) & b(t) \\ b(t) & C_w(t) \end{pmatrix}^{-1} = \begin{pmatrix} C^{-1}(t) + \dfrac{\mathbf{1}}{w^2+\sigma_e^2} & \dfrac{\mathbf{1}}{w^2+\sigma_e^2} \\ \dfrac{\mathbf{1}}{w^2+\sigma_e^2} & C^{-1}(t) + \dfrac{\mathbf{1}}{w^2+\sigma_e^2} \end{pmatrix} \tag{10}$$

https://doi.org/10.1017/S0016672308009555 Published online by Cambridge University Press

and

$$(\bar{x}_w(t), \bar{x}_w(t))\, K^{-1}(t) = (\bar{x}(t), \bar{x}(t)) \begin{pmatrix} C^{-1}(t) & 0 \\ 0 & C^{-1}(t) \end{pmatrix}. \tag{11}$$

The dynamics of the mean effects will be analyzed first. Because recombination and mutation do not change the mean effects, $\bar{x}_w(t) = \bar{x}(t+1)$. Equation (11) is then

$$(\bar{x}(t+1), \bar{x}(t+1))\, K^{-1}(t) \begin{pmatrix} C(t) & 0 \\ 0 & C(t) \end{pmatrix} = (\bar{x}(t), \bar{x}(t)). \tag{12}$$

Using (10) this reduces to two identical matrix equations, which yield the relations

$$\bar{x}_i(t+1) + \frac{2R_i(t)}{w^2+\sigma_e^2} \sum_{i=1}^{n} \bar{x}_i(t+1) = \bar{x}_i(t), \tag{13}$$

where $2R_i(t) = 2\sum_{j=1}^{n} C_{ij}(t)$ is the expressed genetic variance attributable to locus i. Since the mean phenotype in generation t is equal to the mean genotype by assumption 4,

$$\bar{z}(t) = 2\sum_{i=1}^{n} \bar{x}_i(t).$$

Denoting the total expressed genetic variability as

$$\sigma_g^2(t) = 2\sum_{i=1}^{n} R_i(t)$$

equations (13) may be summed and rearranged to give

$$\Delta\bar{z}(t) = -\left(\frac{\sigma_g^2(t)}{w^2+\sigma_e^2+\sigma_g^2(t)}\right)\bar{z}(t). \tag{14}$$

The mean phenotype converges to the optimum at a rate which depends on the expressed genetic variance but not on the hidden genetic variance. The change in the mean effect of alleles at locus i can be obtained from (13) and (14).

$$\Delta\bar{x}_i(t) = -\left(\frac{R_i(t)}{w^2+\sigma_e^2+\sigma_g^2(t)}\right)\bar{z}(t). \tag{15}$$

The direction of change in the mean effect of alleles at any locus is determined only by whether the mean phenotype, $\bar{z}(t)$, is above or below the optimum.

We now proceed to the analysis of the genetic variation. Multiplying both sides of equation (10) by $K(t)$ in the form of (9) gives two matrix equations for $C_w(t)$ and $b(t)$ which can be solved to give

$$C_w(t) = \tfrac{1}{2}\left(C^{-1}(t) + \frac{2\mathbf{1}}{w^2+\sigma_e^2}\right)^{-1} + \tfrac{1}{2}C(t), \tag{16a}$$

$$b(t) = C_w(t) - C(t). \tag{16b}$$

Upon performing the matrix inversion (16a) becomes

$$C_w(t) = C(t) - \frac{C(t)\,\mathbf{1}\,C(t)}{w^2+\sigma_e^2+\sigma_g^2(t)}. \tag{17}$$

https://doi.org/10.1017/S0016672308009555 Published online by Cambridge University Press

The ijth term in this equation is

$$(C_{ij}(t))_w = C_{ij}(t) - \frac{R_i(t)\,R_j(t)}{w^2+\sigma_e^2+\sigma_g^2(t)}. \tag{18}$$

(16b) and (18) may be used to convert the basic equations (1) into recurrence relations

$$\Delta C_{ij}(t) = -\frac{R_i(t)\,R_j(t)}{w^2+\sigma_e^2+\sigma_g^2(t)} - r_{ij}C_{ij}(t) + \delta_{ij}\mu_i m_i^2. \tag{19}$$

Since the primary interest is in the maintenance of genetic variability, the equilibrium condition, $\Delta C_{ij}(t) = 0$, will be examined. A hat, $\hat{}$, signifies equilibrium value. The diagonal equations in (19) with $i = j$ are then $\hat{R}_i^2 = \mu_i m_i^2(w^2+\sigma_e^2+\hat{\sigma}_g^2)$. The only admissible solution to this equation is

$$\hat{R}_i = +\sqrt{(\mu_i m_i^2(w^2+\sigma_e^2+\hat{\sigma}_g^2))} \tag{20}$$

because $2R_i(t)$ is the amount of variance expressed by locus i which must be positive. Using (20), the off-diagonal equations in (19) have the solution

$$\hat{C}_{ij} = -\frac{\sqrt{(\mu_i m_i^2 \mu_j m_j^2)}}{r_{ij}} \quad \text{for} \quad i \neq j. \tag{21a}$$

From the definition of $R_i(t)$

$$\hat{C}_{ii} = \sqrt{(\mu_i m_i^2)}\left(\sqrt{(w^2+\sigma_e^2+\hat{\sigma}_g^2)} + \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{\sqrt{(\mu_j m_j^2)}}{r_{ij}}\right). \tag{21b}$$

The total expressed genetic variation is obtained from $\hat{\sigma}_g^2 = 2\sum_{i=1}^{n} \hat{R}_i$.

$$\hat{\sigma}_g^2 = 2\left(\sum_{i=1}^{n}\sqrt{(\mu_i m_i^2)}\right)\sqrt{\left(w^2+\sigma_e^2+\left(\sum_{i=1}^{n}\sqrt{(\mu_i m_i^2)}\right)^2\right)} + 2\left(\sum_{i=1}^{n}\sqrt{(\mu_i m_i^2)}\right)^2, \tag{21c}$$

Equations (21) are the unique equilibrium solution to the dynamic equations (19). The dynamic system (19) does not seem to be of the type that would give limit cycle behaviour. A computer study of two and three locus cases confirmed that this equilibrium solution is always globally stable.

It is evident from (21b) that this solution cannot be valid when linkage is very tight ($r_{ij} \to 0$) because in actuality there must be some upper limit to the $\hat{C}_{ii}$ as explained in the qualitative analysis; this model is intended to apply only for those ranges of parameters where there are no saturation effects. However, linkage could be a cause of saturation only for extremely low recombination rates because $\sqrt{(\mu_i m_i^2)} \ll \sqrt{(w^2+\sigma_e^2+\hat{\sigma}_g^2)}$.

The equilibrium correlations can be written in terms of the solution (21) as

$$\hat{\rho}_{ij} = \frac{\hat{C}_{ij}}{\sqrt{\left(\left(\hat{R}_i + \sum_{\substack{j=1\\ j\neq i}}^{n}|\hat{C}_{ij}|\right)\left(\hat{R}_j + \sum_{\substack{i=1\\ i\neq j}}^{n}|\hat{C}_{ij}|\right)\right)}}. \tag{22}$$

We now return to the question of when the equilibrium distribution of allelic effects in the gametes, $f(x,t)$, is approximately multivariate normal. One way of ascertain-

https://doi.org/10.1017/S0016672308009555 Published online by Cambridge University Press

ing this is to start with a multivariate normal distribution for $f(x,t)$ and then to inquire under what conditions one generation of selection, recombination and mutation will yield an approximately multivariate normal distribution. Each of these will be dealt with in turn.

(i) From equation (7) the distribution of effects in gametes after selection on zygotes is multivariate normal because it is a marginal distribution of $F_w(x,x',t)$ which is multivariate normal.

(ii) Recombination produces a gametic distribution which is a sum of multivariate normal distributions with equal sets of variances, but unequal sets of covariances and correlations (see equation (1)). Such a sum will be approximately multivariate normal if the component distributions have approximately equal correlations,

$$|(C_{ij}(t))_w - b_{ij}(t)|/\sqrt{((C_{ii}(t))_w (C_{jj}(t))_w)} \ll 1 \quad \text{for} \quad i \neq j$$

or, by application of (16*b*)

$$\sqrt{\left(\frac{C_{ii}(t)\,C_{jj}(t)}{(C_{ii}(t))_w\,(C_{jj}(t))_w}\right)}\,|\rho_{ij}(t)| \ll 1 \quad \text{for} \quad i \neq j.$$

Near equilibrium, the first factor is close to one because from (1), $\hat{C}_{ii} = (\hat{C}_{ii})_w + \mu_i m_i^2$ and the solution (21*b*) shows that $\mu_i m_i^2 \ll \hat{C}_{ii}$. The conditions arising from recombination are thus

$$|\hat{\rho}_{ij}| \ll 1 \quad \text{for} \quad i \neq j. \tag{23}$$

(iii) Kimura (1965) demonstrated for a one locus model that, if the changes produced by mutation are small, the equilibrium distribution of allelic effects is approximately normal. It is necessary to know precisely how small the mutational changes must be. If $m_i^2 \ll (C_{ii}(t))_w$ then the effects of the new mutants will be almost normally distributed with variance $m_i^2 + (C_{ii}(t))_w$ and the new mutants will have nearly the same variance as the non-mutants. Since mutation does not alter the covariances, these conditions ensure that the various component distributions, and hence their sum, will remain approximately multivariate normal. As in (ii) it is noted that near equilibrium $(C_{ii}(t))_w \simeq \hat{C}_{ii}$ and these conditions become $m_i^2 \ll \hat{C}_{ii}$, or using the solution (21*b*)

$$m_i \ll \sqrt{(\mu_i)}\left(\sqrt{(w^2 + \sigma_e^2 + \hat{\sigma}_g^2)} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\sqrt{(\mu_j m_j^2)}}{r_{ij}}\right). \tag{24}$$

For a single locus this reduces to $m_i \ll \sqrt{(\mu_i(w^2 + \sigma_e^2))}$ so the standard deviation of mutational changes must be much smaller, possibly by orders of magnitude, than the width of the selection function acting on the genotypes. For more loci these conditions become less restrictive, with linkage permitting larger mutational changes.

5. DISCUSSION

The equilibrium state of this genetic system has a somewhat peculiar type of stability. The equilibrium phenotype distribution is normal with mean at the optimum. Equation (14) is equivalent to the well-known result that the response of the population mean is equal to the heritability times the selection pressure on the mean (Falconer, 1960). The variance–covariance structure of allelic effects at the different loci in the gametes converges to a unique globally stable equilibrium independent of the mean effects at the individual loci. The peculiarity is that, even though the variance–covariance structure and the grand mean effect of alleles at all loci (the mean genotype or phenotype) both have a stable equilibrium, the mean effects at the individual loci do not (equation (15)). The mean effect of alleles at any given locus then depends largely on historical conditions and chance events. Random genetic drift should be particularly important as there are many patterns of change in the mean effects which are selectively neutral because they do not alter the phenotype distribution. With n loci there are n degrees of freedom for changing the mean effects. The single restriction on these changes, that the mean genotype converges to the optimum, leaves $n-1$ degrees of freedom for selectively neutral changes to occur. Therefore most random changes in the mean effects of alleles at the individual loci are selectively neutral. As a consequence, considerable genetic differentiation between populations could take place by random genetic drift without any phenotypic divergence. Such a process may be, in part, responsible for the genetic differences observed between sibling species (Lewontin, 1974).

We now consider the equilibrium structure of the genetic variation and how it is maintained. In the case of one locus, the equilibrium solution to the detailed model (21) is identical to that of Latter (1970) when infinite population size is used in his formulas. Latter has shown his expressions, which employ a Gaussian fitness function, to be in agreement with those of Kimura (1965) when selection is weak, as implicitly assumed by Kimura in his use of the quadratic deviations fitness function. With weak selection the Gaussian fitness function becomes equivalent to a quadratic deviations fitness function. Kimura (1965) applied his one locus calculation to a polygenic character, assuming that the loci were uncorrelated. In general, with weak selection, the equilibrium amount of expressed genetic variability agrees with Kimura's formula. Even though he assumed there were no correlations between the loci, his calculation is correct because the equilibrium expressed genetic variance is not a function of the linkage relation of the loci.

Both the qualitative analysis and the solution to the detailed model demonstrate that the amount of expressed variability maintained by mutation does not depend on the arrangement of the loci in the genome. Thus the mean fitness of individuals in the population is also independent of the structure of the linkage map. This behaviour contrasts sharply with heterotic models where the mean fitness is strongly dependent on the linkage arrangement of the loci (Lewontin, 1964*a*, 1971; Franklin & Lewontin, 1970).

Unlike the rate of recombination between the loci, the number of recombining

loci has a strong influence on the amount of expressed genetic variability maintained This can be seen most easily by considering a character with an effective number, n_E, of equally mutable loci ($\mu_i m_i^2 = \mu_j m_j^2$). Now the total rate of production of new genetic variation per generation by mutation is

$$\sigma_m^2 = 2 \sum_{i=1}^{n} \mu_i m_i^2.$$

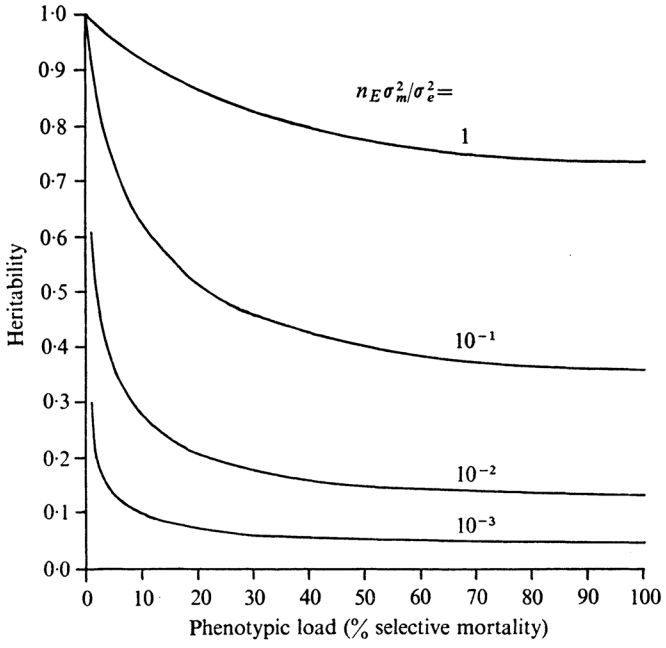

Fig. 2. The heritability maintained by mutation as a function of the strength of stabilizing selection. The heritability is calculated from equation (25) as $h^2 = \hat{\sigma}_g^2/(\sigma_e^2+\hat{\sigma}_g^2)$. The per cent selective mortality is $100(1-\sqrt{(w^2/(w^2+\sigma_e^2+\hat{\sigma}_g^2))})$ which is calculated from the fitness function in equation (2) acting on a normal phenotype distribution with variance $\sigma_e^2+\hat{\sigma}_g^2$ and mean at the optimum. The numbers above each line on the graph are values of the parameter $n_E\sigma_m^2/\sigma_e^2$.

This is a measurable parameter, whereas the mutation rates, μ_i, and the magnitude of mutational changes, m_i, at the individual loci are not. With n_E equally mutable loci, $\mu_i m_i^2 = \sigma_m^2/(2n_E)$ and the expressed genetic variance can be rewritten from (21*c*) as

$$\hat{\sigma}_g^2 = \sqrt{\left(2n_E\sigma_m^2\left(w^2+\sigma_e^2+\frac{n_E\sigma_m^2}{2}\right)\right)}+n_E\sigma_m^2. \tag{25}$$

Thus for a fixed total rate of production of genetic variance by mutation, σ_m^2, and environmental variance, σ_e^2, the higher the effective number of loci the more expressed genetic variability is maintained.

Fig. 2 shows the heritability which can be maintained under various amounts of selective mortality due to stabilizing selection for different values of the

parameter $n_E\sigma_m^2/\sigma_e^2$. It can be seen that when selection is weak, a small increase in the strength of selection can produce a major decrease in the heritability, but when selection is strong any further increase in the strength of selection has little influence on the heritability. This occurs because there is a limit to the strength of stabilizing selection that can be exerted on the genotypes which is determined by the environmental variance of the character (equation (4)). The power of mutation to maintain genetic variation against the force of stabilizing selection is best evaluated from some specific examples.

The first measurements of the production of genetic variation by mutation were performed on *Drosophila melanogaster*. Paxman (1957) estimated σ_m^2 for abdominal and sternopleural bristles and also summarized previous measurements by other workers. Estimates of σ_m^2/σ_e^2 are given in Table 1. It seems reasonable to use 1×10^{-3} as an estimate of σ_m^2/σ_e^2 for both of these characters. The effective number of loci influencing abdominal bristles in *D. melanogaster* has been estimated with low accuracy as about 100 (Falconer, 1960). A rough estimate of the composite parameter $n_E\sigma_m^2/\sigma_e^2$ for this character is therefore about 1×10^{-1}. An estimate of 16 effective loci influencing sternopleural bristles was made by Barnes & Kearsey (1970). A rough estimate of $n_E\sigma_m^2/\sigma_e^2$ for sternopleural bristles is thus about $1{\cdot}6 \times 10^{-2}$.

Table 1. *Rates of production of genetic variance for quantitative characters of* D. melanogaster

	σ_m^2				
Character	Durrant & Mather	Clayton & Robertson	Paxman	σ_e^2	Estimated σ_m^2/σ_e^2
Abdominal bristles	0·006	0·005	0·0035	5·0	1×10^{-3}
Sternopleural bristles	0·002	—	0·001	2·2	1×10^{-3}

σ_e^2 for sternopleural bristles is from Lopez-Fanjul & Hill (1973) and was calculated as $(1-h^2)$ times the phenotypic variance for the Kaduna population. σ_e^2 for abdominal bristles is from Clayton, Morris & Robertson (1957) and is the average variance of males and females from inbred lines of the Kaduna population. Values of σ_m^2 were summarized by Paxman (1957), whose own values have been multiplied by 5/2 because they applied only to chromosome II.

Russell, Sprague & Penny (1963) measured mutation rates for nine quantitative characters of maize in lines that had been selfed at least ten generations. They detected an average of 0·064 mutations per gamete per character. A conservative estimate of σ_m^2/σ_e^2 for these characters can be obtained as $2/\sigma_e^2$ times the rate of mutation detected per gamete times the minimum detectable variance produced by a mutation. The factor of 2 accounts for diploidy. The authors measured 100 offspring from each plant to determine the significance of differences in the line means by the t test and found that almost all the detected differences were significant at the 1 % level. With additive effects, and taking σ_e^2 as the variance within lines, the minimum detected effect is $2m > t_{0\cdot01}\sqrt{(2\sigma_e^2/100)}$ so the minimum variance produced by a detectable mutation is $m^2 > 3{\cdot}3 \times 10^{-2}\sigma_e^2$. A conservative estimate of σ_m^2/σ_e^2 averaged over the nine characters is thus $2(0{\cdot}064)\,(3{\cdot}3 \times 10^{-2}) = 4 \times 10^{-3}$. With

as few as 10 effective loci, most of these characters would have $n_E\sigma_m^2/\sigma_e^2$ in the range of 10^{-2} to 10^{-1} or larger.

Many investigators have measured the divergence between the mean values of quantitative characters in two or more inbred lines derived from a single long-inbred line. Grewal (1962) and Hoi-Sen (1972) have reviewed several such studies. These authors estimated the average rate of increase of variance between lines for 27 and 31 mouse skeletal characters respectively as $3\times10^{-3}\sigma_e^2$ per generation. To convert this into units of σ_m^2/σ_e^2 two correction factors must be taken into account. First, because divergence between lines is due to homozygous mutant alleles, the measured effect is $2m$ rather than m(if there is no dominance) which inflates the variance by a factor of 4. Second, as pointed out by Grewal (1962) this technique detects only a haploid complement of mutations which deflates the variance by a factor of 2. Therefore the average value of σ_m^2/σ_e^2 for these characters is $\frac{2}{4}(3\times10^{-3}) = 1{\cdot}5\times10^{-3}$. With as few as 10 effective loci, most of these characters would have values of $n_E\sigma_m^2/\sigma_e^2$ near 10^{-2}.

These examples indicate that for quantitative characters, $n_E\sigma_m^2/\sigma_e^2$ may typically be near 10^{-2} or 10^{-1}. It can be seen from Fig. 2 that when $n_E\sigma_m^2/\sigma_e^2$ is in this range, high heritabilities can be maintained even when stabilizing selection is very strong. It is concluded that mutation can be a potent force maintaining genetic variation in polygenic characters under stabilizing selection.

Up to this point only the genetic variance which is expressed in the phenotype has been discussed. It is also of some interest to examine the amount of hidden genetic variance which is stored in the negative covariances between effects of alleles at different loci in the gametes. From (21*a*) it can be observed that the covariances, and thus the hidden genetic variance, do not depend on the strength of selection, but are purely a function of the internal properties of the genetic system: the mutability and linkage map of the loci. Large amounts of hidden genetic variation can be accumulated by closely linked loci.

The equilibrium structure of the correlations between alleles at different loci in this model of stabilizing selection differs greatly from that in heterotic models. Consider two loci, i and j. From (22) it can be seen that if other loci are moved closer to i and j on the linkage map the magnitude of correlation between i and j will decrease. This behaviour is the opposite of the imbedding effect found in the heterotic systems (Lewontin, 1964*a*). Furthermore, increasing the number of loci greatly decreases the average magnitude of correlation between the loci, again opposite to the result for the heterotic systems (Franklin & Lewontin, 1970). This can be deduced from the fact that the correlation matrix, $\hat{P}$, must be non-negative definite (Cramér, 1951, pp. 295–6) and in particular

$$(1,\ldots,1)\,\hat{P}(1,\ldots,1)^T = n(n-1)\,\overline{\hat{\rho}_{ij}}+n \geqslant 0,$$

where $\overline{\hat{\rho}_{ij}}$ is the average correlation between loci. As all of the correlations are negative, this relation may be expressed in terms of the average magnitude of correlation,

$$\overline{|\hat{\rho}_{ij}|} \leqslant \frac{1}{n-1}. \tag{26}$$

https://doi.org/10.1017/S0016672308009555 Published online by Cambridge University Press

Thus a polygenic character under stabilizing selection must have a small average magnitude of correlation between loci, regardless of the linkage arrangement of the loci. This result is quite robust and does not depend on the assumptions used in this model, but only on all of the correlations being negative, which is a general feature of stabilizing selection (Mather, 1941; Lewontin, 1964*b*).

Thoday & Gibson (1972) have detected natural and artificial stabilizing selection on sternopleural bristle number in *Drosophila melanogaster* by measuring a negative correlation between the effects of chromosomes II and III. Similar measurements could also be made on parts of chromosomes. Measuring such correlations can yield information about the type of selection operating on the character, but to construe this as a test of the model would be erroneous, since if the correlations are all negative it is impossible for condition (26) to be violated. However, the predictions concerning the amount of genetic variation maintained, and the independence of the covariances and the strength of selection, could be tested by artificial selection experiments. With *Drosophila* there are various experimental approaches for modifying recombination rates to determine whether or not they influence the amount of genetic variation maintained.

Finally it should be mentioned that models with assortative mating and multivariate fitness functions (which allow for pleiotropic gene action) could be investigated by methods similar to those used here.

I thank Jerry Coyne, Marcus Feldman, Donal Hickey, Richard Lewontin, Alan Robertson and Montgomery Slatkin for encouragement and criticism. An anonymous reviewer helped improve the generality of (26) and made other useful suggestions. This work was partially supported by an Atomic Energy Commission grant to R. Lewontin and a National Science Foundation grant to M. Feldman.

REFERENCES

ALLARD, R. W. & JAIN, S. K. (1962). Population studies in predominantly self-pollinated species. II. Analysis of quantitative genetic changes in a bulk-hybrid population of barley. *Evolution* **16**, 90–101.

BARNES, B. W. & KEARSEY, M. J. (1970). Variation for metrical characters in *Drosophila* populations. I. Genetic analysis. *Heredity* **25**, 1–10.

CLAYTON, G. A., MORRIS, J. A. & ROBERTSON, A. (1957). An experimental check on quantitative genetical theory. I. Short-term responses to selection. *Journal of Genetics* **55**, 131–151.

CLAYTON, G. A. & ROBERTSON, A. (1964). The effects of X-rays on quantitative characters. *Genetical Research* **5**, 410–422.

CRAMÉR, H. (1951). *Mathematical Methods of Statistics*. Princeton: Princeton University Press.

DOBZHANSKY, TH., HOLTZ, A. M. & SPASSKY, B. (1942). Genetics of natural populations. VIII. Concealed variability in the second and fourth chromosomes of *Drosophila pseudoobscura* and its bearing on the problem of heterosis. *Genetics* **27**, 463–490.

DOBZHANSKY, TH. & SPASSKY, B. (1944). Genetics of natural populations. XI. Manifestation of genetic variants in *Drosophila pseudoobscura* in different environments. *Genetics* **29**, 270–290.

EAST, E. M. & JONES, D. F. (1919). *Inbreeding and Outbreeding: Their Genetic and Sociological Significance*. Philadelphia: Lippincot.

FALCONER, D. S. (1960). *Introduction to Quantitative Genetics*. New York: Ronald Press.

FISHER, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

FRANKLIN, I. & LEWONTIN, R. C. (1970). Is the gene the unit of selection? *Genetics* **65**, 707–734.

GALE, J. S. & KEARSEY, M. J. (1968). Stable equilibria in the absence of dominance. *Heredity* **23**, 553–561.

GIBSON, J. B. & BRADLEY, B. P. (1974). Stabilizing selection in constant and fluctuating environments. *Heredity* **33**, 293–302.

GREWAL, M. S. (1962). The rate of divergence of sublines in the C57BL strain of mice. *Genetical Research* **3**, 226–237.

HOI-SEN, Y. (1972). Is sub-line differentiation a continuing process in inbred strains of mice? *Genetical Research* **19**, 53–59.

KEARSEY, M. J. & GALE, J. S. (1968). Stabilizing selection in the absence of dominance: an additional note. *Heredity* **23**, 617–620.

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

LACK, D. (1966). *Population Studies of Birds*. Oxford: Clarendon Press.

LATTER, B. H. D. (1970). Selection in finite populations with multiple alleles. II. Centrepetal selection, mutation and isoallelic variation. *Genetics* **66**, 165–186.

LERNER, I. M. (1954). *Genetic Homeostasis*. Edinburgh: Oliver and Boyd.

LEWONTIN, R. C. (1964*a*). The interaction of selection and linkage. I. General considerations. Heterotic models. *Genetics* **49**, 49–67.

LEWONTIN, R. C. (1964*b*). The interaction of selection and linkage. II. Optimum models. *Genetics* **50**, 757–782.

LEWONTIN, R. C. (1971). The effect of genetic linkage on the mean fitness of a population. *Proceedings of the National Academy of Science U.S.A.* **68**, 984–986.

LEWONTIN, R. C. (1974). *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

LOPEZ-FANJUL, C. & HILL, W. G. (1973). Genetic differences between populations of *D. melanogaster* for a quantitative trait. *Genetical Research* **22**, 51–68.

MATHER, K. (1941). Variation and selection of polygenic characters. *Journal of Genetics* **41**, 159–193.

PAXMAN, G. J. (1957). A study of spontaneous mutation in *Drosophila melanogaster*. *Genetica* **29**, 39–57.

RENDEL, J. M. (1943). Variations in the weights of hatched and unhatched ducks' eggs. *Biometrika* **33**, 48–58.

ROBERTSON, A. (1956). The effect of selection against extreme deviants based on deviation or homozygosis. *Journal of Genetics* **54**, 236–248.

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

SCHARLOO, W. (1964). The effect of stabilizing and disruptive selection on the expression of *cubitus interruptus* in *Drosophila*. *Genetics* **50**, 553–562.

SCOSSIROLI, R. E. & SCOSSIROLI, S. (1959). On the relative role of mutation and recombination in responses to selection for polygenic traits in irradiated populations of *D. melanogaster*. *International Journal of Radiation Biology* **1**, 61–69.

THODAY, J. M. & GIBSON, J. B. (1972). A simple test for stabilizing and disruptive selection. *Egyptian Journal of Genetics and Cytology* **1**, 47–50.

WADDINGTON, C. H. (1960). Experiments on canalizing selection. *Genetical Research* **1**, 140–150.

YAMADA, Y. & KITAGAWA, O. (1961). Doubling doses for polygenic mutations in *Drosophila melanogaster*. *Japanese Journal of Genetics* **36**, 76–83.