Dynamics of unconditionally deleterious mutations: Gaussian approximation and soft selection

ALEXEY S. KONDRASHOV
Section of Ecology and Systematics, Corson Hall, Cornell University, Ithaca, NY 14853, USA

Summary

This paper studies the influence of two opposite forces, unidirectional unconditionally deleterious mutations and directional selection against them, on an amphimictic population. Mutant alleles are assumed to be equally deleterious and rare, so that homozygous mutations can be ignored. Thus, a genotype is completely described by its value with respect to a quantitative trait x, the number of mutations it carries, while a population is described by its distribution $p(x)$ with mean $M[p]$ and variance $V[p] = \sigma^2[p]$. When mutations are only slightly deleterious, so that $M \gg 1$, before selection $p(x)$ is close to Gaussian with any mode of selection. I assume that selection is soft in the sense that the fitness of a genotype depends on the difference between its value of x and M, in units of σ. This leads to a simple system of equations connecting the values of M and V in successive generations. This system has a unique and stable equilibrium, $\hat{M} = (U/\delta)^2(2-\rho)$ and $\hat{V} = (U/\delta)^2$, where U is the genomic deleterious mutation rate, δ is the selection differential for x in units of σ, and ρ is the ratio of variances of $p(x)$ after and before selection. Both δ and ρ are parameters of the mode of soft selection, and do not depend on M or V. In an equilibrium population, the selection coefficient against a mutant allele is $\hat{s} = \delta^2[U(2-\rho)]^{-1}$. The mutation load can be tolerable only if the genome degradation rate $v = U/\sigma$ is below 2. Other features of mutation-selection equilibrium are also discussed.

1. Introduction

Measurement of the total genomic deleterious mutation rate U is difficult, because mutations with slight effects cannot be individually detected at the phenotypic level. Still, the slowly accumulating data suggest that in multicellular eukaryotes U may be of the order of 1 or higher (Mukai, 1964; Mukai *et al.* 1972; Crow, 1979; Crow & Simmons, 1983; Charlesworth *et al.* 1990; Agren & Schemske, 1993). If so, selection against mutations can be important for the evolution of reproduction (Kondrashov, 1988, 1993) and aging (Charlesworth, 1993; Partridge & Barton, 1993), maintenance of phenotypic (Kondrashov & Turelli, 1992) and molecular (Charlesworth *et al.* 1993) variability, and some other phenomena.

In this paper I consider the balance between a unidirectional mutation process, that creates unconditionally deleterious alleles, and directional selection against such alleles, that removes them. The difference between models that assume irreversible *v.* reversible mutation, or between those that assume unconditionally deleterious *v.* conditionally beneficial mutation, was discussed by Kondrashov & Turelli (1992). In the

case considered here the properties of the mutation-selection equilibrium at a single locus are well understood (Haldane, 1927, 1937; Wright, 1929; see also Hofbauer, 1985). Such analysis can be extended to the case of two loci (Karlin & McGregor, 1972; Allendorf, 1979). However, if we consider the whole genome simultaneously, which is necessary to describe what actually happens in nature, a similar approach holds only when the mutation rate is low and the selection against mutations is very strong, so that an individual cannot carry more than one or very few mutations.

Here I consider the mutation process for the whole genome, ignoring random drift and assuming that all mutations are equally deleterious. This implies that selection acts on the total number x of mutant alleles in the genome (the 'genome contamination'), and the distribution of x, $p(x)$, is sufficient to describe a population. Then it is possible, making some additional reasonable assumptions, to derive equations which describe changes of $p(x)$ during selection, mutation, and reproduction, and to study them numerically (Kimura & Maruyama, 1966; Kondrashov, 1982).

Unfortunately, this approach lacks generality, which makes analytical results desirable. Charlesworth (1990) obtained such results by assuming that the function relating the number of mutations to fitness, $w(x)$, is Gaussian (when mean value of $w(x)$ is negative, Gaussian selection is effectively directional because $x \geqslant 0$). Then, a Gaussian $p(x)$ remains such after selection. Assuming that $p(x)$ is always Gaussian, Charlesworth described the population by its mean $M[p]$ and variance $V[p] = \sigma^2[p]$, instead of $p(x)$ itself. Even here, however, the algebraic equations for the equilibrium values of $M[p]$ and $V[p]$ must be solved numerically.

In all these studies selection was assumed hard, i.e. $w(x)$ remained the same under any $p(x)$. Here I show that if selection is soft, in the sense that the fitness of a genotype is dependent on the population state, investigation of the mutation-selection balance can be much simpler. I will retain the assumption that $p(x)$ before selection is Gaussian, because numerical data show that in an amphimictic population with free recombination and high $M[p]$ this is a good approximation under any mode of selection. Thus, soft selection will not be necessarily Gaussian. Here an invariant mode of reproduction is considered, while its evolution is investigated in the accompanying paper (Kondrashov, 1995). First, I review some characteristics of selection acting on a quantitative trait.

2. Characteristics of selection

(i) *The problem*

Consider a quantitative trait x, which can take any value from $-\infty$ to ∞, with the population distribution $p(x)$. Selection $w(x)$ causes substitution of $p(x)$ by $\tilde{p}(x)$:

$$\tilde{p}(x) = w(x)p(x)/\overline{W}, \quad \overline{W} = \int w(x)p(x)\,dx, \qquad (1)$$

where $\overline{W}$ is the mean population fitness. Here and in the following integrals are taken from $-\infty$ to ∞. To
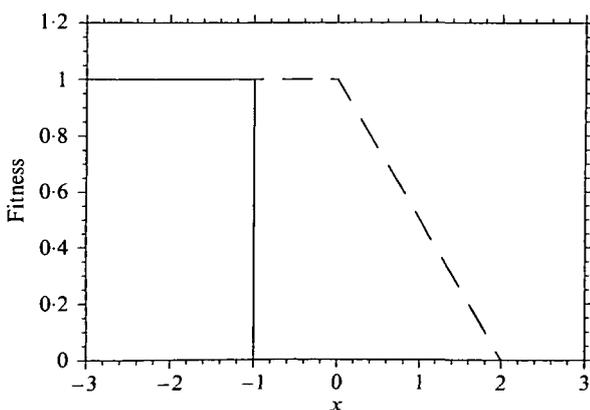
Fig. 1. Fitness functions according to eqn (11): $w_m = -1$ and $w_w = 0$ (solid line), and $w_m = 1$ and $w_w = 2$ (dashed line).
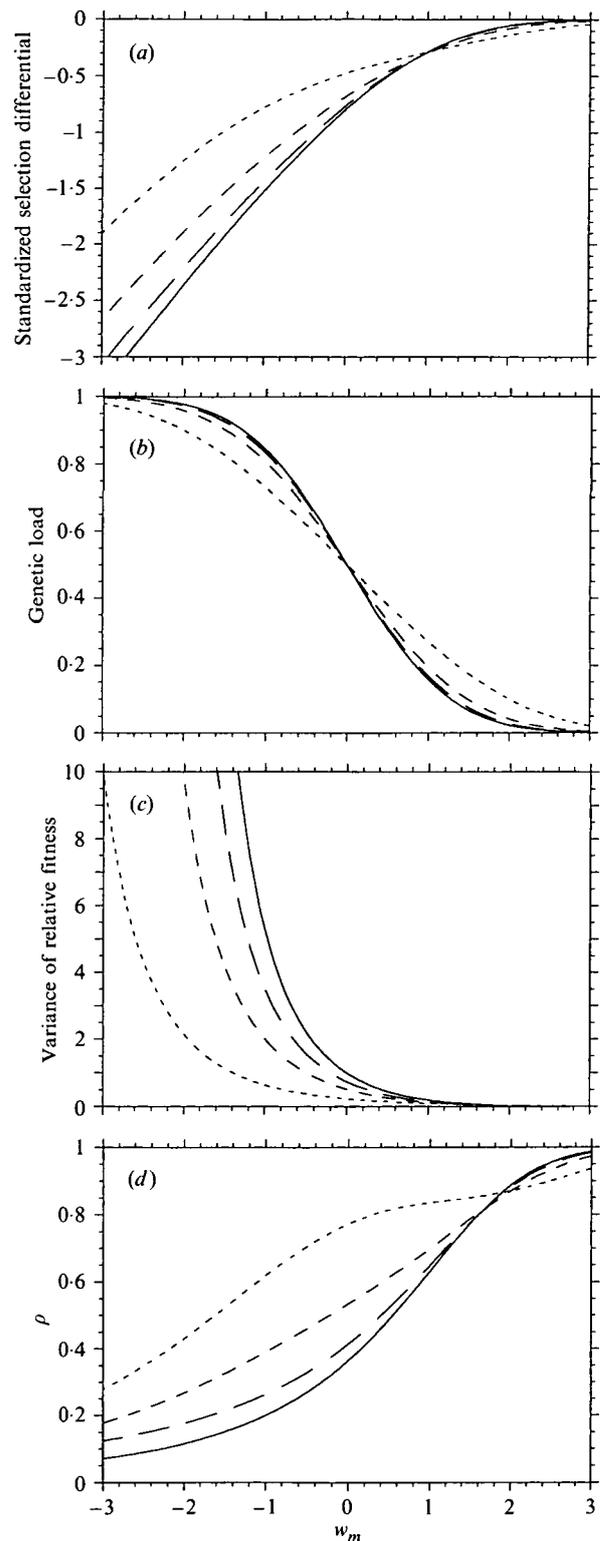
Fig. 2. Characteristics of selection (eqn (11)) as functions of w_m with $w_w = 0$ (solid lines), 1 (long dashes), 2 (medium dashes), and 4 (short dashes). Standardized selection differential (*a*), genetic load (*b*), variance of relative fitness (*c*), ratio of variances of x after and before selection (*d*).

describe various aspects of selection, several characteristics (functionals), which depend on $w(x)$ and $p(x)$, will be used. Most of them describe intrapopulation processes and thus depend only on relative fitness and
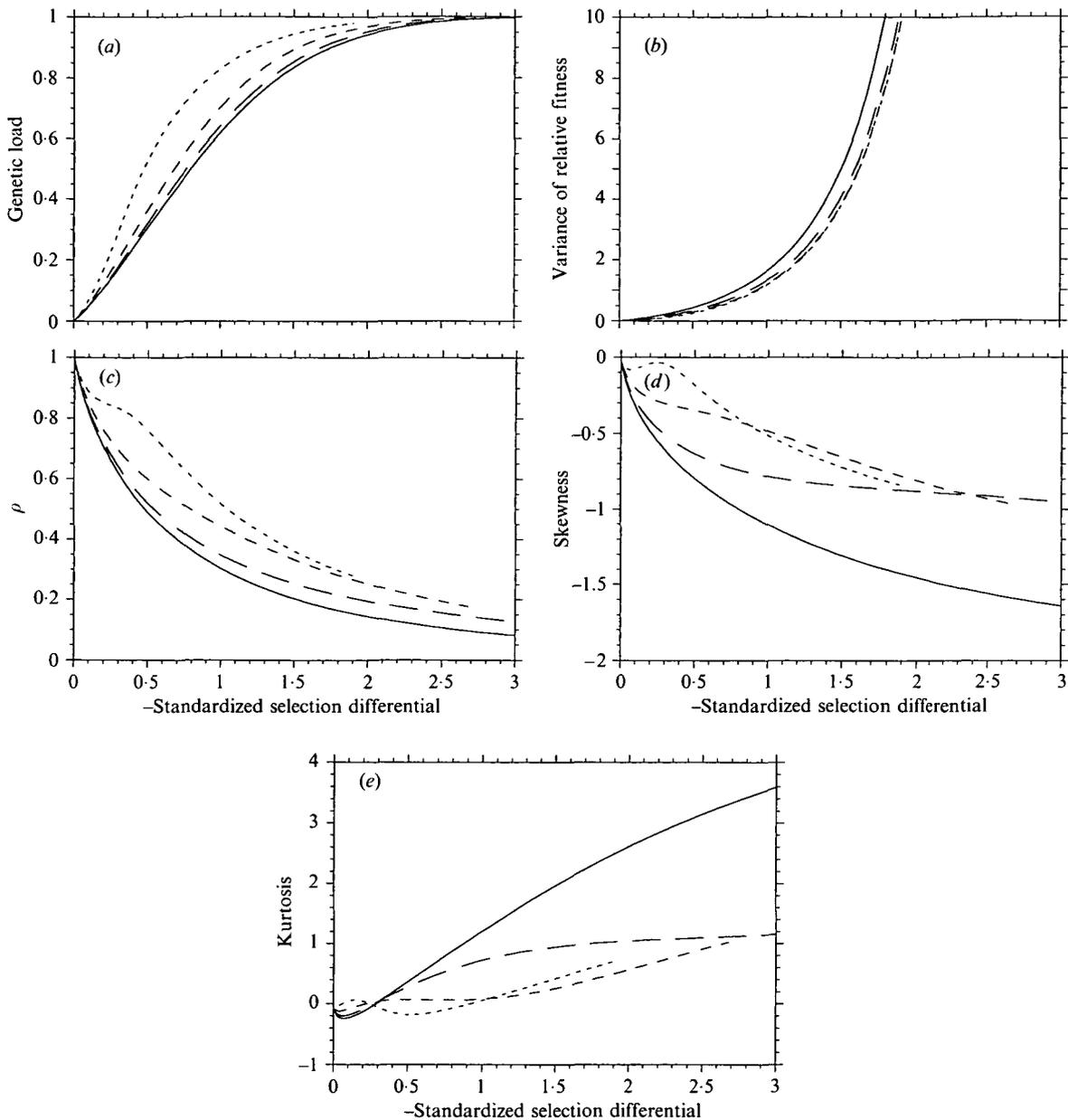
Fig. 3. The same as Fig. 2, but the characteristics of selection are plotted as functions of the negative standardized selection differential. Genetic load (*a*), variance of relative fitnesses (*b*), ratio of variances of *x* after and before selection (*c*), skewness after selection (*d*), kurtosis after selection (*e*).

are invariant to multiplication of $w(x)$ by any positive constant.

It is often convenient to use the standardized trait value $X = (x - M[p])/\sigma[p]$ as an independent variable, instead of x. The population distribution of X is $P(X) = \sigma[p]p(M[p] + \sigma[p]X)$ with $M[P] = 0$ and $V[P] = \sigma[P] = 1$. Fitness as a function of X is defined by $W(X) = w(M[p] + \sigma[p]X)$.

The kth non-central moment of the function $w(x)p(x)$ is:

$$I_k = \int x^k w(x)p(x)\,dx. \tag{2}$$

Analogously, J_k stands for the kth moment of

$W(X)P(X)$. Connections between J_k and I_k can be easily established using the formulae:

$$I_k = \int (M[p] + \sigma X)^k W(X) P(X)\,dX. \tag{3}$$

Obviously, $I_0 = J_0 = \overline{W}$. For some quantities introduced below expressions in terms of both I_k and J_k are useful.

(ii) *Characteristics of selection*

I consider twelve characteristics of selection, which can be divided into three groups. The first group (characteristics 1 and 2) describes the action of selection itself, the second group (characteristics 3–6)

describes the changes which selection causes in the distribution of the trait, and the third group (characteristics 7–12, which will be considered in the accompanying paper) describes the effects of small changes of the population on the results of selection. I denote characteristics of $\tilde{p}(x)$ by respective symbols with the sign '~', e.g. $\tilde{M}$ instead of $M[\tilde{p}(x)]$.

(1) Genetic load (Crow, 1970), the reduction of the mean population fitness relative to the maximal possible fitness:

$$L = (W_{max} - \overline{W})/W_{max} = 1 - \int w(x)\,p(x)\,\mathrm{d}x/W_{max}. \quad (4)$$

(2) Populational variance of relative fitness:

$$D[p] = \int ((w(x) - \overline{W})/\overline{W})^2 p(x)\,\mathrm{d}x$$

$$= I_0^{-2} \int w(x)^2 p(x)\,\mathrm{d}x - 1. \quad (5)$$

$D[p]$ determines the relative increase of the mean fitness due to selection (Fisher, 1930) and may be called opportunity for selection (Crow, 1958) or evolvability of fitness (Houle, 1992).

(3) Selection differential (Fisher, 1930), the difference between the mean values of the trait after and before selection:

$$\Delta = \tilde{M} - M = I_1/I_0 - M. \quad (6)$$

Sometimes the standardized selection differential (or effectiveness of selection, Crow & Kimura, 1970, p. 227) δ should be used:

$$\delta = \Delta/\sigma[p] = (I_1/I_0 - M[p])/\sigma[p] = J_1/J_0. \quad (7)$$

Of course, $\delta < 0$ if $w(x)$ is a decreasing function of x. In the quantitative genetic literature δ is often called 'selection intensity', although this term is more appropriate for L.

(4) Ratio of variances of the trait after and before selection

$$\rho = \tilde{V}/V = (I_2/I_0 - (I_1/I_0)^2)/I_2^2 = J_2/J_0 - (J_1/J_0)^2. \quad (8)$$

Clearly, $V[\tilde{p}]$ and $D[p]$ are different quantities. If $p(x)$ is Gaussian, under exponential selection $\rho = 1$; under synergistic epistasis (i.e. if $(\ln w(x))'' < 0$ for all x) $\rho < 1$; under diminishing returns epistasis (i.e. if $(\ln w(x))'' > 0$ for all x) $\rho > 1$ (Shnol & Kondrashov, 1993).

(5)–(6) Even if $p(x)$ is Gaussian, $\tilde{p}(x)$ remains Gaussian only if $w(x)$ is Gaussian (see Charlesworth, 1990, eqn A1). Otherwise, $\tilde{p}(x)$ is non-Gaussian. Its third and fourth central moments are:

$$\tilde{M}_3 = I_3/I_0 - 3I_1 I_2/I_0^2 + 2I_1^3/I_0^3 \quad (9)$$

$$\tilde{M}_4 = I_4/I_0 - 4I_1 I_3/I_0^2 + 6I_1^2 I_2/I_0^3 - 3I_1^4/I_0^4. \quad (10)$$

The standardized third and fourth cumulants, $\tilde{M}_3/\tilde{\sigma}^3$ and $\tilde{M}_4/\sigma^4 - 3$, describe skewness and kurtosis and characterize the deviation of $\tilde{p}(x)$ from Gaussian (Kendall & Stuart, 1977, pp. 58 and 88).

(iii) *Numerical results*

Consider a particular class of fitness functions $w(x)$ defined by:

$$w(x) = \begin{cases} A, & x \leqslant w_m - w_w/2 \\ A\left(\dfrac{w_w}{w_m} - \dfrac{x}{w_w} + \dfrac{1}{2}\right), & w_m - w_w/2 < x \leqslant w_m + w_w/2 \\ 0, & x > w_m + w_w/2. \end{cases} \quad (11)$$

Thus, when x grows from $w_m - w_w/2$ to $w_m + W_w/2$, $w(x)$ declines linearly from A to 0 (Fig. 1). Four families of $w(x)$, with $w_w = 0$, 1, 2, and 4, and $-3 \leqslant w_m \leqslant 3$, were used and characteristics of such selection acting on a population with Gaussian $p(x)$ with $M = 0$ and $V = 1$ were calculated (Figs. 2 and 3). Obviously, these results can also be interpreted as the characteristics of $W(X)$. The MacFORTRAN program is available on request.

3. Mutation–selection equilibrium

(i) *Soft selection*

If the absolute fitness of a genotype depends not only on its trait value x, but also on $p(x)$, selection is called soft (Wallace, 1975; Maynard Smith, 1978). I will consider such selection, assuming that

$$W(X) = W((x - M[p])/\sigma[p]) = w(M[p] + \sigma[p]X) \quad (12)$$

remains invariant under any $p(x)$. Thus, while $w(x)$ changes with $p(x)$, the fitness of a genotype with a given standardized contamination X remains the same, and the data from Figs. 2 and 3 are relevant to any state of the population with a given $W(X)$. However, the values of M and V, under which $W(X)$ equilibrates with the mutation rate U, remain to be found.

(ii) *Complete model*

Consider an amphimictic population with discrete generations and selection in only one phase of the life cycle. If selection acts on haploids, individuals are haploid, and U is the deleterious mutation rate per haploid genome. If selection acts on diploids, individuals are diploid and U is the mutation rate per diploid genome. With free recombination both cases lead to the same equations.

Both possible life cycles, selection–mutation–reproduction and selection–reproduction–mutation, will be studied, although the first succession is more realistic (see below). Reproduction is a combined process; syngamy followed by meiosis if selection acts on haploids, or meiosis followed by syngamy if selection acts on diploids. The genome contamination, x, will be treated as a continuous variable. Before selection, $p(x)$ will be assumed Gaussian. Thus, a population is completely described by $M[p]$ and $V[p]$.

Three processes of the life cycle cause the following changes in them.

Changes in $M[p]$ and $V[p]$ after selection are described by eqns (7) and (8). I will consider two modes of mutation. With the shift mode, each genome acquires exactly U new mutations, so that $M[p]$ is incremented by U, while $V[p]$ remains invariant (the reasons why it may be convenient to ignore the increment of $V[p]$ due to fresh mutations are described below). With the Poisson mode mutations occur independently, and the number of new mutations per genome has a Poisson distribution with mean U, so that both $M[p]$ and $V[p]$ are incremented by U. Because mutations contribute to x additively, reproduction does not change $M[p]$, while with free recombination the variance of $p(x)$ after reproduction, $V'[p]$, is connected with $M[p]$ and $V[p]$ by:

$$V'[p] = 0 \cdot 5(M[p] + V[p]) \qquad (13)$$

(Bulmer, 1985, eqn 9.16; note that here, if rare mutant alleles are distributed independently, $V[p] = M[p]$). Thus, in every generation reproduction leads to a two-fold reduction of pairwise associations between mutations and of the deviation of the variance from its value without associations $M[p]$ (Bulmer, 1985, eqn 9.17).

Thus, the three processes of the life cycle cause the following changes to M and V:

selection	mutation	reproduction	
$M \to M + \sigma\delta \to M + \sigma\delta + U$		$\to M + \sigma\delta + U$	$(14a)$
$V \to \rho V$	$\to \rho V$	$\to 0 \cdot 5(M + \sigma\delta + U)$	
		$+ 0 \cdot 5(\rho V)$	
selection	mutation	reproduction	
$M \to M + \sigma\delta \to M + \sigma\delta + U$		$\to M + \sigma\delta + U$	$(14b)$
$V \to \rho V$	$\to \rho V + U$	$\to 0 \cdot 5(M + \sigma\delta + U)$	
		$+ 0 \cdot 5(\rho V + U)$	
selection	reproduction	mutation	
$M \to M + \sigma\delta \to M + \sigma\delta + U$		$\to M + \sigma\delta + U$	$(14c)$
$V \to \rho V$	$\to -(M + \sigma\delta)$		
	$+ -(\rho V)$	$\to 0 \cdot 5(M + \sigma\delta)$	
		$+ 0 \cdot 5(\rho V)$	
selection	reproduction	mutation	
$M \to M + \sigma\delta \to M + \sigma\delta$		$\to M + \sigma\delta + U$	$(14d)$
$V \to \rho V$	$\to -(M + \sigma\delta)$		
	$+ -(\rho V)$	$\to 0 \cdot 5(M + \sigma\delta)$	
		$+ 0 \cdot 5(\rho V) + U$	

Cases $(14a)$ and $(14c)$ correspond to shift mutation, and cases $(14b)$ and $(14d)$ to Poisson mutation. The final values of M and V after the chain of transformations $(14a$–$d)$ start the next generation. They are the same in $(14b)$ and $(14d)$. Thus, with Poisson mutation the overall changes of M and V do not depend on the order of the processes.

(iii) *Mutation-selection equilibrium*

Equilibrium values of M and V, $\hat{M}$ and $\hat{V}$, can be found from eqn (14) by equating the final and the initial values of M and V. The equilibrium is always unique and for eqn $(14a)$ it is

$$\hat{M} = (U/\delta)^2(2-\rho)$$
$$\hat{V} = (U/\delta)^2, \qquad (15)$$

where δ (7) and ρ (8) are determined by $W(X)$ and are invariant with respect to M and V. Actually, the positive equilibrium exists only if $\rho < 2$. A 'pathological' case $\rho > 2$ means that diminishing returns epistasis (see Shnol & Kondrashov, 1993) is so strong that selection increases the variance more that two-fold. Under selection regimes (11), in contrast, epistasis is synergistic and $\rho < 1$. The assumption that $p(x)$ is Gaussian and, thus, is defined over the whole x axis (while, in fact, $x \geqslant 0$) can be used only if $\hat{M}[p] \gg 0$ when $\hat{M}[p] \gg \hat{\sigma}[p]$ and the frequencies of genotypes with small x are negligible.

Some forms of $W(X)$ yield analytical expressions for δ and ρ. For example, with exponential soft selection $W(X) = e^{-sX}$, $\delta = -s$ and $\rho = 1$. Usually, however, the relevant integrals J_0, J_1, and J_2 must be calculated numerically. However, this need be done only once for a given $W(X)$ after which the solution (15) can be obtained.

At mutation-selection equilibrium, the increase of $M[p]$ due to the mutation, U, is, of course, equal to its decrease due to selection, $-\Delta = -\delta\sigma[p]$ (see Wright, 1977, v. 3, p. 475). Therefore, at equilibrium $-\delta = v$, where v is the genome degradation rate (Kondrashov, 1984, 1988), the genome deleterious mutation rate in units of $\sigma[p]$:

$$v = U/\sigma[p]. \qquad (16)$$

Thus, the data from Fig. 3 may be interpreted as the dependencies, under various $W(X)$, of the characteristics of selection in equilibrium population on the genome degradation rate.

Because Δ, the number of mutations removed with the given $W(X)$ and δ, is determined by $\sigma[p]$, the equilibrium is always reached when $\hat{\sigma}[p] = -U/\delta$, which counterbalances a given U. The value $\hat{M}$ which corresponds to such $\hat{\sigma}[p]$ depends on ρ. Note, that ρ alone determines the ratio $\hat{V}/\hat{M} = (2-\rho)^{-1}$, while a conventional coefficient of variation $\hat{\sigma}/\hat{M}$ depends on both ρ and δ. Thus, $\hat{V} = \hat{M}$ only in the absence of epistasis (exponential selection). With synergistic epistasis $\hat{V} < \hat{M}$, while with diminishing returns epistasis $\hat{V} > \hat{M}$. Obviously, the first case means repulsion, and the second coupling associations between distributions of different mutations.

Although the Poisson mode of mutation is appropriate if x is treated as a discrete variable, with continuous x the shift mutation model should probably be used for the following reason. Consider the action of non-epistatic selection on a population
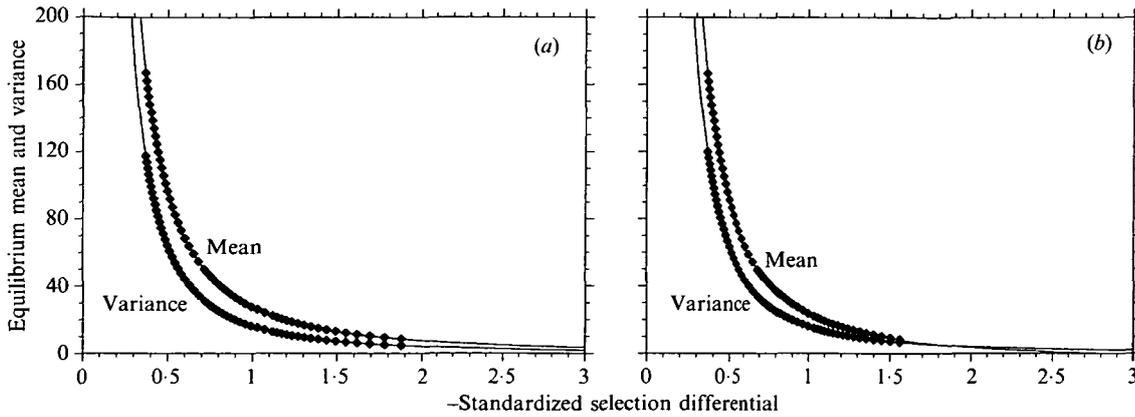
Fig. 4. The mean and the variance of the number of mutations per genome in the equilibrium population before selection with the life cycle selection–mutation–reproduction as a function of negative standardized selection differential $-\delta$ under $U = 4.0$ and $w_w = 0$ (strict truncation). The numerical results ($\blacklozenge$) and analytical estimates obtained from eqn (15) using the values of δ and ρ calculated for a trait with Gaussian distribution (lines) are presented. (*a*) Shift mutation, (*b*) Poisson mutation.

in linkage equilibrium. Such selection implies $w(x) = (1-s)^x$ with discrete x and $w(x) = e^{-sx}$ with continuous x, while linkage equilibrium among rare alleles implies Poisson $p(x)$ with discrete x and Gaussian $p(x)$ with $M[p] = V[p]$ with continuous x. Multiplicative selection decreases both M and V of Poisson $p(x)$ by the same quantity (Maynard Smith, 1978, chapter 3). In contrast, exponential selection decreases only M of Gaussian $p(x)$ leaving V unchanged. Thus, with continuous x exponential selection followed by shift mutation produces the same overall changes of M and V as multiplicative selection and Poisson mutation in the discrete case. Thus, the continuous approximation of an inherently discrete x, which requires large M, implies that the variance introduced by the mutation itself can be ignored.

Therefore, we will mostly consider the case (14*a*), although the difference between all four cases is not large. The value of $\hat{V}$ is the same everywhere. In cases (14*b*) and (14*d*) $\hat{M}$ is smaller by U, while in case (14*c*) it is larger by U, than in case (14*a*). Thus, in contrast to (14*a*), in the other three cases $\hat{V}/\hat{M}$ depends on U, but the relative differences between the values of $\hat{M}$ in different cases decreases with the U and/or δ.

(iv) *Stability of the equilibrium*

The Jacobian of the recurrent eqns (14) in a neighborhood of the equilibrium is

$$\begin{pmatrix} 1 & -\dfrac{\delta^2}{2U} \\ \dfrac{1}{2} & \dfrac{1}{2}\left(-\dfrac{\delta^2}{2U}+\rho\right) \end{pmatrix}. \tag{17}$$

Investigation of its eigenvalues shows that the equilibrium is stable unless $U < \delta^2/(8+4\rho)$. As far as with realistic selection $\delta < 2$ (or, maybe, 3) due to mutation load arguments (see below), stability is guaranteed if $U > 0.5$ (or 1). The instability of equilibrium with small U and very strong selection indicates that the

assumption that $p(x)$ is Gaussian becomes too inaccurate when $M[p]$ is small. If $U \to \infty$, the largest eigenvalue is less than 1 only by a quantity of the order of $1/U$, so that stability is 'weak' and the population approaches the equilibrium slowly. Although the global stability of the equilibrium is not proven, there is little doubt that the population would approach it starting from any initial state.

(v) *Numerical results*

The above analysis is based on the assumption that $\hat{p}(x)$ is Gaussian before selection. I have checked the accuracy of this approximation numerically. (The THINK C program used is available on request.) In the case of Poisson mutation and hard selection the recursion formulae are presented in Kimura & Maruyama (1966, eqn 3.1). For shift mutation with non-integer U, I assumed that the number of mutations acquired by an organism can be equal to either the largest integer smaller than U or the smallest integer larger than U with probabilities such that the expected increase is U. With soft selection, where $W(X)$ is invariant, $w(x) = W((x-M[p])/\sigma[p])$ is recalculated every generation using the current $M[p]$ and $\sigma[p]$.

The data presented in Fig. 4 demonstrate an excellent agreement between the numerically calculated and analytically predicted values of $\hat{M}$ and $\hat{V}$. The same agreement holds under other modes of selection and mutation rates, as long as $\hat{M}$ is not too small (data not reported). Actually, $\hat{M} = 10$ and even less is sufficient for good agreement. As expected, the equilibrium under soft selection seems unique and globally stable and the rate of convergence to it is rather slow when $\hat{M}$ is high (data not reported).

This agreement occurs because the true equilibrium $p(x)$ before selection is very close to Gaussian, although after strong truncation-like selection $p(x)$ is very different from Gaussian (Fig. 5). Thus, it is not surprising that the values of all characteristics of selection at equilibrium are very similar to those
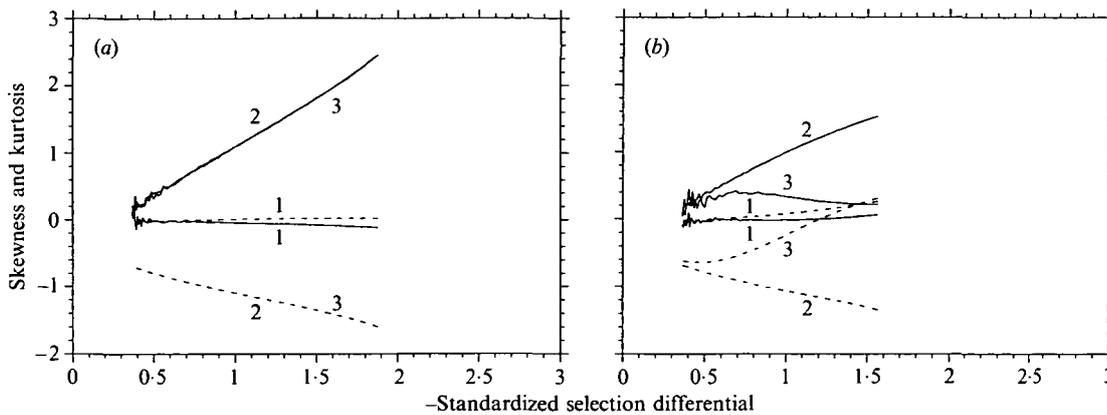
Fig. 5. Skewness (broken lines) and kurtosis (solid lines) before selection (1), mutation (2), and reproduction (3) under the same parameters as in Fig. 4 with shift (a) and Poisson (b) mutation.

predicted analytically for a given $W(X)$ and Gaussian $p(x)$ (data not reported). The main factor which returns $p(x)$ to Gaussian is reproduction (Bulmer, 1985, eqn 9.18). Shift mutation does not alter the shape of $p(x)$ (that is why the lines 2 and 3 almost coincide in Fig. 5a), while Poisson mutation introduces negative skewness and decreases kurtosis (Fig. 5b). Actually, if reproduction acts alone, $p(x)$ tends to Poisson, and not Gaussian distribution. However, the skewness and kurtosis of the Poisson distribution are $M^{-1/2}$ and M^{-1}, respectively (Kendall & Stuart, 1977, Ch. 3 and 5), and the difference is small when $\hat{M} \gg 1$.

(vi) *Selection and equilibria at individual loci*

Let us relate the properties of the mutation-selection equilibrium at the level of the total genome contamination to the events at individual loci. Consider a locus with normal allele B and rare mutant allele b. Mutations occur only from B to b with the frequency μ. The relative fitnesses of individuals with genotypes B and b (or BB and Bb, if individuals are diploid) are 1 and $1-s$, respectively. Thus, the frequency of a mutant allele, q, increases by μ because of mutation and decreases by $1-s$ due to selection. Therefore, at equilibrium:

$$\hat{q} = \mu/s. \tag{18}$$

This formula, which is a limiting case of a more general result (Wright, 1929; Haldane, 1937), is valid when μ is small, $s \gg \mu$, and, if selection acts on diploids, when homozygotes bb can be ignored.

Analogously, for many loci with rare and equally deleterious mutations, where $-\Delta = sM[p]$, at equilibrium $-\Delta = U$ and (Crow, 1970):

$$\hat{M}[p] = U/\hat{s}. \tag{19}$$

Here we must write $\hat{s}$, instead of s, because selection coefficient against a mutant allele may now depend on the population state. Substituting the expression for $\hat{M}$ (15) into (19), we obtain;

$$\hat{s} = \frac{\delta^2}{U(2-\rho)}. \tag{20}$$

Note, that if we assume that the distribution of x in individuals carrying the allele b differs from $p(x)$ just by a shift of 1 to the right, the relative fitness of this allele is (Kimura & Crow, 1979):

$$1-s = \int w(x)p(x-1)\,dx/I_0, \tag{21}$$

and with Gaussian $p(x)$, because $p(x-1) = p(x) + p'(x) + \ldots$ (this Taylor expansion can be used if $\sigma[p] \gg 1$), selection coefficient against b is

$$s = (1/(V[p]\,I_0))(MI_0 - I_1) = -(1/V[p])\,\Delta. \tag{22}$$

At equilibrium this gives $\hat{s} = U/\hat{V}[p]$, instead of the correct $\hat{s} = U/\hat{M}[p]$. Thus, eqn (22) overestimates $\hat{s}$ with synergistic epistasis, where $\hat{V} < \hat{M}$, and underestimates $\hat{s}$ with diminishing returns epistasis.

Obviously, the assumption that the distribution of genome contamination in individuals carrying a mutant allele at some locus differs from $p(x)$ just by a shift of 1 is true only if this allele is distributed independently from all others. If this is true for all loci, $M = V$. With $V < M$ mutations are distributed more uniformly than randomly, and individuals with the allele b at some locus differ from the whole population less than by shift by 1 to the right. In contrast, if $V > M$ the presence of b indicates a larger difference. However, under realistic ρ the discrepancy between eqns (20) and (22) is not large.

4. Discussion

(i) *Applicability of the approach*

My analytical results are based on the assumption that in equilibrium populations the genome contamination before selection can be treated as a quantitative trait with a Gaussian distribution $p(x)$. This is a good approximation as long as $\hat{M}[p] \gg 1$ (Fig. 5), and the analytical predictions about $\hat{M}$ and $\hat{V}$ are quite accurate (Fig. 4). When $\hat{M}[p]$ is too small for the Gaussian approximation to be valid, another asymptotic approach can be used if $\hat{M}[p] \ll 1$, so that most individuals are mutation-free but some carry a single mutation. This situation is similar to the case of

a single locus. As usual, the intermediate situation, where neither asymptotic result is valid, is most difficult to study analytically.

Above I have considered a life cycle which included 'reproduction', i.e. the succession syngamy-meiosis (if selection occurs in the haplophase) or meiosis-syngamy (if selection occurs in the diplophase). If, however, selection occurs in both phases the approach based on Gaussian approximation cannot be generally used, because it will be invalid at the beginning of the diplophase, due to selection in the preceding haplophase. The Gaussian approximation requires that all episodes of selection are separated by meiosis.

Only the 'ideal' amphimixis was studied, with free recombination and no facultative apomixis. Perhaps, the Gaussian approximation remains applicable if these assumptions are relaxed to some extent. Then $\hat{M}$ depends on how eqn (13) is modified by the presence of linkage or facultative apomixis, while $\hat{V}$ always remain the same.

(ii) *Possible values of δ and ρ*

These two parameters, the standardized selection differential (7) and the ratio of variances of $p(x)$ after and before selection (8), are determined by the soft selection function $W(X)$. Any values $\delta \leqslant 0$ and $\rho \geqslant 0$ are possible under non-increasing $W(X)$, but not all of them are realistic. Only when $-\delta < 2$ (3), the mutation load can be not too close to one (Fig. 3). With such δ values, $\rho \geqslant 0.2$. Both L and ρ are minimal, with a given δ, under truncation selection, but under other forms of selection considered here their values are only slightly higher, if δ remains the same (Shnol & Kondrashov, 1994, fig. 1). Thus, we can conclude that $0 \leqslant -\delta < 2$ and $0.2 < \rho \leqslant 1$ ($\rho > 1$ requires diminishing returns epistasis, which is probably unrealistic).

(iii) *Mutation–selection equilibrium*

At equilibrium the genome degradation rate $v = -\delta$ and thus at equilibrium $v < 2$ (3) (Kondrashov, 1988). This implies that $\hat{V} > U^2/4$ (15). The corresponding $\hat{M}$ is the same under exponential selection, or slightly larger under synergistic epistasis. However, it cannot be larger than $2\hat{V}$ even if $\rho = 0$, although, in fact, $\rho < 0.2$ is impossible (see above). According to eqn (20), $\hat{M} > U^2/4$ implies $\hat{s} < 4/U$. Thus, the larger U is, the weaker selection at separate loci must be. Knowledge of v, together with $W(X)$, allows us to find all other characteristics of selection in the equilibrium population (Fig. 3).

Some of the current conclusions should probably be changed if, as is certainly the case in nature, different mutations are deleterious to different degrees. Particularly, the conclusions about $\hat{s}$ may hold for the average coefficient of selection against new mutations, while the average coefficient of selection against persisting alleles must be much lower, because such alleles are enriched with very slightly deleterious ones. However, this requires a separate study.

(iv) *Comparison with hard selection*

Once a stable equilibrium is reached, $p(x)$ remains invariant and there is no difference between hard and soft selection. Thus, our analysis of equilibrium can also be applied to hard selection, provided that the dependence of fitness on X at equilibrium is $W(X)$. However, in this case the deviation of the population from equilibrium leads to changes of $W(X)$, so that δ and ρ depend on M and V and it is generally impossible to determine $\hat{M}$ and $\hat{V}$ from eqn (15) and even to prove the existence or uniqueness of equilibrium.

Of course, the problem can always be studied numerically and, if the equilibrium is reached, $\hat{W}(X)$ can be calculated and used to determine its properties (Figs. 2 and 3). For example, with exponential hard selection $w(x) = \exp(-sx)$, $\hat{M} = \hat{V} = U/s$ (see Maynard Smith, 1978, Ch. 3). Because in this case $\delta = -\sigma s$ and $\rho = 1$, eqn (15) gives the correct $\hat{M}$ and $\hat{V}$, while eqn (20) yields $\hat{s} = (\hat{\sigma}s)^2/U = s$, as expected.

It is intuitively clear (and was repeatedly shown numerically) that with synergistic epistasis equilibrium is 'even more' stable under hard selection than under soft selection, because the growth of $M[p]$ causes a more rapid increase of the strength of selection. In contrast, with diminishing returns epistasis (e.g. if $w(x) = 1/x$) equilibrium under hard selection may be absent. In this case $M[p]$ increases indefinitely (Kimura & Maruyama, 1966; Charlesworth, 1990), because with the growth of $M[p]$ selection becomes less efficient.

If $w(x)$ has alternating regions of synergistic and diminishing returns epistasis, several stable equilibria (probably separated by unstable ones), each of which corresponds to a zone of rapid decline of $w(x)$, are possible. This might happen if the environment contains several resources (or several negative factors), and they have different critical values of contamination to utilize (or to avoid) them. No data, however, confirm the existence of such selection in nature.

(v) *Experimental data and implications*

Predictions from the above analysis depend on the values of U, δ, and ρ. If $U \ll 1$, the current approach can be used with moderate exponential selection against deleterious alleles, so long as $-\delta$ is small, so that $\hat{M}[p] \gg 1$. For example, if $U = 0.1$ and $W(X) = e^{-0.02X}$, $\hat{M} = \hat{V} = 25$ and $L = 1 - e^{-U}$ is approximately 0.1, if the fitness of the mutation-free genotype is regarded as W_{max}.

In contrast, under $U > 1$ our approach can always be used, but the mutation load is tolerable only with the synergistic epistasis between slightly deleterious mutations. Thus, data on a high U value provide indirect evidence for synergistic epistasis. Some data suggest that such epistasis is possible (Mukai, 1969; Malmberg, 1977), but much more information is needed. Unfortunately, there are no direct data on the natural mutation load, and it is not clear how it could

be measured. Thus, now we can only say that fecundities of most species seem to be consistent with high load, provided that selection against mutations is an important cause of natural mortality.

Evolvability of fitness is probably easier to measure than the load, because it depends on the mean, instead of the maximal, fitness. Some data show that evolvability of traits associated with fitness, usually its components, can be rather high (Mukai *et al.* 1974; Houle, 1992). The most promising approach to measuring *D* in nature seems to be the measurement of the correlation of fitnesses between relatives in species with low fecundity and presumably low random mortality. This, together with more data on *U*, can decide the importance of the phenomena studied in this paper.

References

Agren, J. & Schemske, D. (1993). Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution* **47**, 125–135.

Allendorf, F. W. (1979). Rapid loss of duplicate gene expression by natural selection. *Heredity* **43**, 247–258.

Bulmer, M. G. (1985). *The Mathematical Theory of Quantitative Genetics*, 2nd edition. Oxford: Oxford University Press.

Charlesworth, B. (1990). Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genetical Research* **55**, 199–221.

Charlesworth, B. (1993). Evolutionary mechanisms of senescence. *Genetica* **91**, 11–19.

Charlesworth, B., Charlesworth, D. & Morgan, M. T. (1990). Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature* **347**, 380–382.

Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology* **30**, 1–13.

Crow, J. F. (1970). Genetic loads and the cost of natural selection. In *Mathematical Topics in Population Genetics* (ed. K. Kojima), pp. 128–177. Heidelberg: Springer.

Crow, J. F. (1979). Minor viability mutants in Drosophila. *Genetics* **92** (Suppl.), s165–s172.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper and Row.

Crow, J. F. & Kimura, M. (1979). Efficiency of truncation selection. *Proceedings of the National Academy of Sciences, USA* **76**, 396–399.

Crow, J. F. & Simmons, M. J. (1983). The mutation load in Drosophila. In *The Genetics and Biology of Drosophila*, Vol. 3c (ed. M. Ashburner, H. L. Carson and J. N. Thompson), pp. 2–35. New York: Academic Press.

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

Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection. Part V. Selection and mutation. *Proceedings of the Cambridge Philosophical Society* **23**, 838–844.

Haldane, J. B. S. (1937). The effect of variation on fitness. *American Naturalist* **71**, 337–349.

Hofbauer, J. (1985). The selection mutation equation. *Journal of Mathematical Biology* **23**, 41–53.

Houle, D. (1992). Comparing evolvability and heritability of quantitative traits. *Genetics* **130**, 195–204.

Houle, D., Hoffmaster, D. K., Assimacopoulos S. & Charlesworth, B. (1992). The genomic mutation rate for fitness in Drosophila. *Nature* **359**, 58–60.

Karlin, S. & McGregor, J. (1971). On mutation-selection balance for two-locus haploid and diploid populations. *Theoretical Population Biology* **2**, 60–70.

Kendall, M. G. & Stuart, A. (1977). *The Advanced Theory of Statistics*, 4th edition, vol. 1. London: Charles Griffin and Co.

Kimura, M. & Crow, J. F. (1978). Effect of overall phenotypic selection on genetic change at individual loci. *Proceedings of the National Academy of Sciences, USA* **75**, 6168–6171.

Kimura, M. & Maruyama, T. (1966). The mutation load with epistatic gene interactions in fitness. *Genetics* **54**, 1337–1351.

Kondrashov, A. S. (1982). Selection against harmful mutations in large sexual and asexual populations. *Genetical Research* **40**, 325–332.

Kondrashov, A. S. (1984). Deleterious mutations as an evolutionary factor. I. The advantage of recombination. *Genetical Research* **44**, 199–217.

Kondreashov, A. S. (1988). Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440.

Kondrashov, A. S. (1993). Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity* **84**, 372–387.

Kondrashov, A. S. (1995). Modifiers of mutation-selection balance: general approach and the evolution of mutation rates. *Genetical Research*, in press.

Kondrashov, A. S. & Turelli, M. (1992). Deleterious mutations, apparent stabilizing selection and maintenance of quantitative variation. *Genetics* **132**, 603–618.

Malmberg, R. L. (1977). The evolution of epistasis and the advantage of recombination in populations of bacteriophage T4. *Genetics* **86**, 607–621.

Maynard Smith, J. (1978). *The Evolution of Sex*. Cambridge: Cambridge University Press.

Mukai, T. (1964). The genetic structure of natural populations of Drosophila melanogaster. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* **50**, 1–19.

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

Muller, H. J. (1950). Our load of mutations. *American Journal of Human Genetics* **2**, 111–176.

Partridge, L. & Barton, N. H. (1993). Optimality, mutation and the evolution of ageing. *Nature* **362**, 305–311.

Shnol, E. E. & Kondrashov, A. S. (1993). Effect of selection on the phenotypic variance. *Genetics* **134**, 995–996.

Shnol, E. E. & Kondrashov, A. S. (1994). Some relations between different characteristics of selection. *Journal of Mathematical Biology*, **32**, 835–840.

Wallace, B. (1975). Hard and soft selection revisited. *Evolution* **29**, 465–473.

Wright, S. (1929). Fisher's theory of dominance. *American Naturalist* **63**, 274–279.

Wright, S. (1977). *Evolution and the Genetics of Populations*, vol. 3. Chicago: University of Chicago Press.