Genet. Res., Camb. (1995), **66**, *pp.* 53–69 *With 5 text-figures*

Modifiers of mutation-selection balance: general approach and the evolution of mutation rates

ALEXEY S. KONDRASHOV
Section of Ecology and Systematics, Corson Hall, Cornell University, Ithaca, NY 14853, USA, Phone: (607) 255-4617, FAX: (607) 255-8088
E-mail: alex_kondrashov@qmrelay.mail.cornell.edu

(*Received 3 February 1994 and in revised form 5 December 1994*)

Summary

A general approach is developed to estimate secondary selection at a modifier locus that influences some feature of a population under mutation-selection balance. The approach is based on the assumption that the properties of all available genotypes at this locus are similar. Then mutation-selection balance and weak associations between genotype distributions at selectable loci and the modifier locus are established rapidly. In contrast, changes of frequencies of the modifier genotypes are slow, and lead to only slow and small changes of the other features of the population. Thus, while these changes occur, the population remains in a state of quasi-equilibrium, where the mutation-selection balance and the associations between the selectable loci and the modifier locus are almost invariant. Selection at the modifier locus can be estimated by calculating quasi-equilibrium values of these associations. This approach is developed for the situation where distributions of the number of mutations per genome within the individuals with a given modifier genotype are close to Gaussian. The results are used to study the evolution of the mutation rate. Because beneficial mutations are ignored, secondary selection at the modifier locus always diminishes the mutation rate. The coefficient of selection against an allele which increases the mutation rate by υ is approximately $\upsilon\delta^2/[U(2-\rho)] = \upsilon\hat{s}$, where U is the genomic deleterious mutation rate, δ is the selection differential of the number of mutations per individual in units of the standard deviation of the distribution of this number in the population, ρ is the ratio of variances of the number of mutations after and before selection, and $\hat{s}$ is the selection coefficient against a mutant allele in the quasiequilibrium population. However, the decline of the mutation rate can be counterbalanced by the cost of fidelity, which can lead to an evolutionary equilibrium mutation rate.

1. Introduction

Kondrashov (1995), referred to below as K-95, studied an equilibrium between two forces, unidirectional unconditionally deleterious mutations and directional soft selection against them, in a population with an invariant system of reproduction. Here I will consider the evolution of reproduction under the balance of these two forces. A comparison of the absolute mean fitnesses of separate clones with different modes of reproduction is enough to deal with the evolution of obligate apomixis (Maynard Smith, 1978). In contrast, the evolution of various features of amphimixis, such as mutation, recombination, outcrossing, mate choice, etc., depends on individual selection within the population.

This selection can be studied using explicit genetic models, where, in addition to a locus (loci) subject to direct selection, a 'modifier' locus (loci) is introduced, which may have no direct effect on fitness, but influences some feature of reproduction or migration (Karlin & MacGrenor, 1974; Taylor & Williams, 1982; Liberman & Feldman, 1986; Charlesworth, 1990, 1993). Associations (non-independence, linkage disequilibrium) between the distributions of the alleles at selectable and modifier loci can develop, causing secondary selection at the modifier locus. Sometimes the direction of evolution at such loci can be predicted from the direction of change of the mean fitness or the genetic load (Karlin & McGregor, 1974), but this is not always true if modifiers can recombine with the consequences of their action (e.g. Feldman *et al.* 1980;

Kondrashov, 1984; Altenberg & Feldman, 1987; Charlesworth, 1990, 1993; Wiener & Feldman, 1993).

In several studies (e.g. Taylor & Williams, 1982; Charlesworth, 1990) the properties of all available genotypes at a modifier locus (loci) were assumed similar ('basic assumption'). This leads to profound simplification of the models due to separation of fast and slow variables (see Mishchenko & Rozov, 1980). Surprisingly, this was, to my knowledge, never discussed explicitly, although such separation was used in models where all selection is weak (Passekov & Singh, 1991; Nagylaki, 1993). I consider, at the heuristic level, the general consequences of the basic assumption for models with modifiers, and then apply it to modifiers of mutation-selection balance, where the selectable loci are also subject to mutation. Then I study the evolution of the mutation rate.

Soon after the intraspecific variability of mutation rates (Muller, 1928) and their genetic determination (Dubovskij, 1935; Demerec, 1937; see Ives, 1950 and Dobzhansky, 1951, pp. 58–65) were discovered, the idea that mutability itself evolves was proposed. According to Dubinin *et al.* (1936, p. 973), 'The mutation rate... is the expression of an equilibrium existing between the greater or lesser harmfulness of the mutations (as the presence of some aberrations in the species) and the necessity of them for the evolutionary plasticity of the species (without which the species would have disappeared)'. However, Sturtevant (1937) offered another solution to the question '... why does the mutation rate not become reduced to zero? No answer seems possible at present, other than the surmise that the nature of genes does not permit such a reduction. In short, mutations are accidents, and accidents will happen' (p. 466).

Thus, two alternative hypotheses, according to which the mutation rate is either optimal in some evolutionary sense or a minimal possible one, were formulated (Shapiro, 1938; Shapiro & Ignatiev, 1945). Depending on the mode of selection, either hypothesis can be relevant (see Kimura, 1967; Leigh, 1970, 1973; Gillespie, 1981*a*, *b*; Liberman & Feldman, 1986; Ishii *et al.* 1989; Sasaki, 1994). Because here the mutations will be unconditionally deleterious, the minimal hypothesis will be applicable.

First, I will introduce six new characteristics of selection, in addition to those described in K-95 (Section 2). Then, a general approach to the analysis of the models with the modifier loci under the basic assumption will be discussed, with the emphasis on the modifiers of the mutation-selection balance (Section 3). A modifier of the mutation rate will be studied in Section 4, ignoring its direct effect on fitness, while in Section 5 the cost of fidelity will be explicitly taken into account. Section 6 presents consideration of the impact of the evolution of the mutation rate on the population.

2. Characteristics of selection

Here we need to describe the influence of small changes of the distribution of a quantitative trait $x, p(x)$, before selection on the results of selection (see K-95) for additional details and notation). If $p(x)$ is Gaussian, only the effects of changes of M and V have to be considered. Three results of selection are important for us: $\bar{W}$ and the mean $\tilde{M}$ and variance $\tilde{V}$ of the distribution after selection $\tilde{p}(x)$.

Formally, the characteristics we seek are partial derivatives of a result of selection with respect to a feature of the population before it. I will present these in terms J_k. The expressions that depend only on J_k, denoted by Ks below, remain invariant under constant $W(X)$ (the numbering of characteristics is a continuation of that in K-95).

7. $$\frac{\partial \bar{W}}{\partial M} = \frac{1}{\sigma} J_1 = \frac{1}{\sigma} K_1^0; \qquad (1a)$$

8. $$\frac{\partial \bar{W}}{\partial V} = \frac{1}{2\sigma^2}(J_2 - J_0) = \frac{1}{\sigma^2} K_2^0; \qquad (1b)$$

9. $$\frac{\partial \tilde{M}}{\partial M} = \frac{1}{J_0^2}(J_2 J_0 - J_1^2) = K_1^1; \qquad (1c)$$

10. $$\frac{\partial \tilde{M}}{\partial V} = \frac{1}{2\sigma J_0^2}(J_3 J_0 - J_2 J_1) = \frac{1}{\sigma} K_2^1; \qquad (1d)$$

11. $$\frac{\partial \tilde{V}}{\partial M} = \frac{\sigma}{J_0^3}(J_3 J_0^2 - 3J_2 J_1 J_0 + 2J_1^3) = \sigma K_1^2; \qquad (1e)$$

12. $$\frac{\partial \tilde{V}}{\partial V} = \frac{1}{2J_0^3}(J_4 J_0^2 - 2J_3 J_1 J_0 - J_2^2 J_0 + 2J_2 J_1^2) = K_2^2. \qquad (1f)$$

The derivations are straightforward and involve partial differentiation of the expressions for $\bar{W}$, $\tilde{M}$, or $\tilde{V}$ with respect to M or V (Appendix 1). Numerical values of these characteristics are presented in Fig. 1 (characteristics 7 and 8 were calculated under $\bar{W} = 1$, while the rest are invariant to the multiplication of $W(X)$ by a positive constant). With simultaneous small changes of both M and V their influences must be added.

Because $K_1^0 = \delta$, the value of $\bar{W}$ always decreases with M, provided that $W(X)$ is a decreasing function of X. In contrast, K_2^0 may be both positive or negative under such $W(X)$, so that $\bar{W}$ may change in either direction with increased V. for example, under truncation selection an increase in variance increases the mean fitness if more than half of the population is truncated, and decreases it otherwise (see Kondrashov, 1984, fig. 1). The value of $\tilde{M}$ always increases with M, because $K_1^1 = \rho > 0$ (K-95). With the $W(X)$ used here, K_2^1 and K_1^2 are always negative, while K_2^2 is always positive (Fig. 1).

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

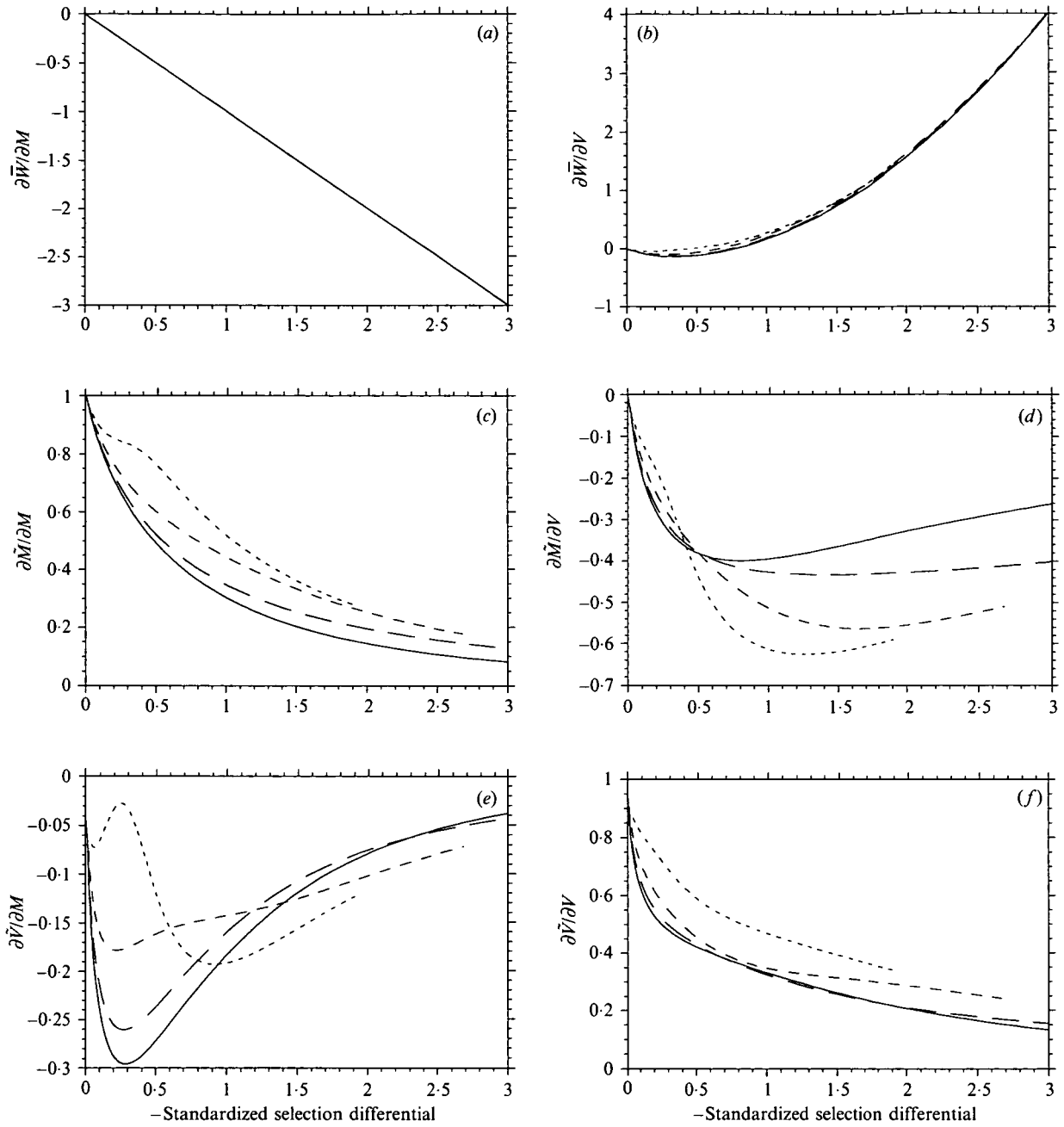

Fig. 1. Six characteristics of selection (7–12, see eqn (1)) as functions of the negative standardized selection differential $-\delta$. The fitness function described by (12) from K-95 was used with $w_w = 0$ (solid lines), 1 (long dashes), 2 (medium dashes), and 4 (short dashes).

3. Selection at the modifier locus

(i) *General approach: quasi-equilibrium and associations between the selectable and the modifier loci*

Models that include both selectable and modifier loci can be, as any multilocus model, very complicated. If all the loci have comparable effects, the dynamics of all genotype frequencies must be considered simultaneously, often leading to analytically untractable equations. Fortunately, the situation is radically simplified under the basic assumption that the properties of all available genotypes at the modifier locus (loci) are similar (e.g. they all cause similar mutation rates or have similar effects on fitness, if selection also acts directly on this locus), while selection (and, perhaps, other processes, e.g. recombination and mutation) at the selectable loci, as well as recombination between selectable and modifier loci, is strong (i.e. the differences in fitness and the rates of recombination, mutation, etc. are not small). In order to take advantage of this simplification, it is convenient, instead of considering the frequencies of all genotypes, to describe the populaton by variables of three types: (1) genotype frequencies at the selectable loci; (2) magnitudes of associations between

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

the distributions of alleles at selectable and modifier loci, and (3) genotype frequencies at the modifier loci (Slatkin, 1972).

The suitable measure of these associations must not depend directly on the genotype frequencies at the modifier loci (see below). Such a measure was introduced by Nei & Li (1980) to describe an association between inversions and electromorphs. It is based on comparison of the frequencies of electromorphs (in our case, genotypes at the selectable loci) among the components of the population having different inversions (genotypes at the modifier loci), and, thus, does not depend directly on the frequencies of the latter. In contrast, the usual coefficient of linkage disequilibrium, as well as the related measure Z, treat both the loci involved in the same way (Crow & Kimura, 1970, pp. 196–197), and are thus unsuitable for our purpose. The basic assumption leads to the following hierarchy of processes:

1. Changes of genotype frequencies at selectable loci, being influenced by fast factors, are rapid, and, in the first approximation, do not depend on what happens at the modifier loci. Ignoring the possibility of other attractors, these frequencies can approach equilibria (to be more precise, quasi-equilibria, see below) with arbitrary properties.

2. Changes of associations between the selectable and the modifier loci are rapid, because one of the factors that influences them, recombination between these loci, is fast. However, the equilibrium (quasi-equilibrium) values of these associations are small, because the only factor that causes them to deviate from zero, the difference between the properties of the modifier genotypes, is small, while a rapid recombination pushes them to zero. At equilibrium the Nei-Li-type measures of these associations do not depend, as a first approximation, on the changes of the genotype frequencies at the modifier loci.

3. Changes of the genotype frequencies at the modifier loci can be initially either slow or rapid, depending on the initial values of the associations betwen selectable and modifier loci. However, after these associations become small (which happens soon), these changes become slow, because both secondary selection, which appears as the result of these associations, and direct selection (if any) at the modifier locus are weak. Still, the results of the changes of the genotype frequencies at the modifier loci can be arbitrary, including the replacement of one modifier allele with another one. Even this, however, will lead to only small changes of the genotype frequencies of selectable loci (Karlin & McGregor, 1974, p. 96), as well as of the magnitude of associations.

Thus, the genotype frequencies at the selectable loci, as well as the magnitudes of associations between these loci and the modifier locus (loci), rapidly reach an invariant state, as a first approximation, or quasi-equilibrium (Crow & Kimura, 1970, p. 197). Then, the changes of the frequencies of the modifier genotypes proceed on the quasi-equilibrium background that is not significantly influenced by them. Actually, a locus with a small difference among the properties of all its genotypes is not necessarily a modifier locus, but may be just one of the selectable loci. These simplifications depend only on the basic assumption, and do not require that all but one genotype at the modifier locus are rare. Here I will use this approach to separate the rapid establishment of mutation-selection balance from slow evolution of the mode of reproduction.

(ii) *Description of the population with a modifier of the mutation-selection balance*

Here a population consists of several components, each containing all individuals with a particular modifier genotype, and a component is characterized by its distribution of x (see K-95). The simplest Nei-Li-type measure of the association between selectable and modifier loci is then the difference between the mean numbers of mutations among the individuals with different modifier genotypes. Besides, I will consider the differences among variances of the numbers of mutations within the components, which is the Nei-Li-type analogue to the higher-order linkage disequilibria.

Because of the basic assumption, the quasi-equilibrium differences between the components will be small. In addition, the distributions of x within the components will be assumed Gaussian (see K-95). Then, the differences between them can be completely described by the differencs between their means and variances (Taylor & Williams, 1982, considered only the differences between the mean values, while both differences were used by Charlesworth, 1990), while the distribution in the whole population is also approximately Gaussian (Appendix 2). All this allows us to use (1) in order to find secondary selection at the modifier loci.

I will consider a single modifier locus m with two alleles, m_1 and m_2. If during syngamy gametes unite without 'memory' of the previous diplophase, and there is no diploid apomixis, frequencies of genotypes at the beginning of the haplophase (in meiospores) are dynamically sufficient (see Lewontin, 1974), i.e. allow us to predict the future. The distributions of x among the meiospores with genotypes m_1 and m_2, $p_1(x)$ and $p_2(x)$, are Gaussian with means M_1 and M_2 and variances V_1 and V_2, respectively, and the frequencies of m_1 and m_2 in all the meiospores are a and 1-a, respectively. To describe secondary selection at m, we have to find quasi-equilibrium values of $\mu = M_2 - M_1$ and $\nu = V_2 - V_1$. Thus, we have to derive equations that connect the values of μ and ν in the successive generations, which requires describing their transformations during each process in the life cycle. If only the diplophase is dynamically sufficient, the differences between the distributions of x in three

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

diploid genotypes must be used as dynamic variables. This situation is not considered here.

(iii) *Modifier of mutation-selection balance: diploid phase unimportant*

Let us start from the simplest case where everything happens in the haplophase, i.e. the life cycle is meiosis–selection–mutation–syngamy. Because of (1 *c–f*), if the modifier is not under direct selection, the values of μ and ν after and before selection are connected by (here and below the higher-order terms are ignored):

$$\left.\begin{aligned}\mu' &= K_1^1\mu + \hat{\sigma}^{-1}K_2^1\nu,\\ \nu' &= \hat{\sigma}K_1^2\mu + K_2^2\nu,\end{aligned}\right\} \quad (2)$$

where $\hat{\sigma} = \sigma[\hat{p}(x)] = -U/\delta$ (eqn 15, K-95) is the standard deviation of the quasiequilibrium distribution of x. During mutation a dummy modifier does not change the values of μ and ν: $\mu'' = \mu'$ and $\nu'' = \nu'$. Reproduction (here, the succession of syngamy and meiosis), leads, with obligate amphimixis and free recombination, to the following changes (capital letters denote the values in the next generation):

$$\left.\begin{aligned}M &= \mu''/2,\\ N &= (\mu'' + \nu'')/4\end{aligned}\right\} \quad (3)$$

(Appendix 3). Therefore, the quasi-equilibrium values of μ and ν, $\hat{\mu}$ and $\hat{\nu}$, can be found from the equations

$$\left.\begin{aligned}\mu &= (\tfrac{1}{2})K_2^1\mu + (\tfrac{1}{2})\hat{\sigma}^{-1}K_2^1\nu,\\ \nu &= (\tfrac{1}{4})(K_1^1 + \hat{\sigma}K_1^2)\mu + (\tfrac{1}{4})(\hat{\sigma}^{-1}K_2^1\nu + K_2^2)\nu.\end{aligned}\right\} \quad (4)$$

The only solution of this linear system is $\hat{\mu} = \hat{\nu} = 0$. From Fig. 1 we can see that the absolute values of K_1^1, K_1^2, K_2^1, and K_2^2 are all below one. Thus, if $\hat{\sigma}$ is not too small ($\hat{\sigma} > 1$ is sufficient), this solution is always stable, at least for the modes of selection studied here. This is not surprising, as the modifier does nothing, and the associations between it and the selectable loci should tend to zero. Too small $\hat{\sigma}$ are inconsistent with Gaussian $p(x)$. The 'pathological' case of $K_1^1 = \rho > 2$, which leads to instability, was discussed in K-95.

The intensity of secondary selection at the modifier locus can be conveniently characterized by a coefficient of selection against a modifier genotype, i.e. by the difference between the mean fitnesses of those who do not carry this genotype and those who carry it, relative to the mean population fitness. This coefficient is the total differential of relative fitness and, if before selection the mean and the variance of the distribution of x in those that carry a particular modifier genotype differ from their values in the rest of the population by $\hat{\mu}$ and $\hat{\nu}$, respectively, equals:

$$\omega = -\hat{\sigma}^{-1}K_1^0\bar{W}^{-1}\hat{\mu} - \hat{\sigma}^{-2}K_2^0\bar{W}^{-1}\hat{\nu} \quad (5)$$

(see (1 *a*, *b*)). This formula will be used below, where 'real' modifiers will lead to non-zero $\hat{\mu}$ and $\hat{\nu}$. If the influence of a modifier locus on some process is small, the secondary selection at it can be described by the derivative of ω by the effect this allele on the relevant process (Charlesworth, 1990).

(iv) *Modifier of the mutation-selection balance: diploid phase important*

The above analysis is applicable if nothing in the diplophase depends on the genotype at the m locus, either directly (e.g. because of differences in recombination or mutation rates among its diploid genotypes) or due to correlations (e.g. because of diploid selection and different distributions of x in different genotypes). Then (3) describes the transformation of μ and ν during the diplophase.

Otherwise, we have to consider the processes in the diplophase explicitly. In order to do this, we have to investigate the changes of $\mu_1 = M_{12} - M_{11}$, $\mu_2 = M_{22} - M_{11}$, $\nu_1 = V_{12} - V_{11}$, and $\nu_2 = V_{22} - V_{11}$, where M_{11}, M_{12}, and M_{22} and V_{11}, V_{12}, and V_{22}, are the means and the variances of the distributions of x in diploid individuals with genotypes m_1m_1, m_1m_2, and m_2m_2, respectively (Appendix 3).

Transformations of μ_1 and ν_1 (and of μ_2 and ν_2) caused by selection in the diplophase are identical to those of μ and ν caused by selection in the haplophase and are described by equations analogous to (2). Note that we cannot consider selection operating in both phases (K-95). After mutation analogously to the haploid case, $\mu_1'' = \mu_1'$, $\mu_2'' = \mu_2'$, $\nu_1'' = \nu_1'$ and $\nu_2'' = \nu_2'$, as long as m does not influence the mutation rate. Otherwise, μ_1'' and μ_2'' (together with ν_1', and ν_2'', if we consider Poisson mutation), will change, and the relative magnitude of these changes depends on the dominance at m locus, which can be arbitrary. This, together with possible diploid selection, can make the values of μ_1'', μ_2'', ν_1'', and ν_2'' before meiosis arbitrary small numbers, although they still can be determined from μ'' and ν''.

If the frequency of the allele m_1 before meiosis is a, the values of M and N in the meiospores after meiosis depend on μ_1'', ν_1'', μ_2'', and ν_2'' in the following way (Appendix 3):

$$\begin{aligned}M &= (\mu_2'' - \mu_1'')/2 + a(\mu_1'' - \mu_2''/2),\\ N &= (\mu_2'' - \mu_1'' + \nu_2'' - \nu_1'')/4 + a(\mu_1'' - \mu_2''/2 + \nu_1'' - \nu_2''/2)/2.\end{aligned} \quad (6)$$

Thus, in contrast to (3), (6) depends on the frequencies of the modifier genotypes. Particularly, intermediate equilibrium frequencies of these genotypes become possible. Stability of such equilibria can be analysed by standard methods.

This situation is analogous to a classical model of selection at one locus under panmixia, where diploid selection leads to frequency-dependent fitnesses of alleles, although zygote fitnesses are invariant. However, there is no reason to limit our analysis to the case

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

when only one modifier allele is frequent (Charlesworth, 1990), because, if we have already made the basic assumption, further restrictions do not simplify the analysis substantially.

4. Selection for lower mutation rate

(i) *The problem*

We are now equipped to write equations on the changes of μ and ν for a real modifier. Here I will do it for a modifier which influences the mutation rate. Mutations at the modifier locus will be ignored, and recombination will be assumed to be free. Mutation at selectable loci can occur either in haploid or in diploid phase. In the first case, genotypes m_1 and m_2 have the mutation rates U and $U+\upsilon$, respectively. In the second case, genotypes m_1m_1, m_1m_2, and m_2m_2 have mutation rates U, $U+\upsilon_1$, and $U+\upsilon_2$, respectively. First, I will consider only the differences in the mean numbers of mutations among the components of the population, ignoring the differences in variances. After this the complete model will be analysed.

(ii) *Differences in the means only*

Let us start with the case when both mutation and selection occur in the haplophase, and selection precedes mutation. Selection, mutation, and reproduction cause the following transformations of μ (see above, compare with eqn 14 from K-95):

$$\mu \xrightarrow{\text{selection}} K_1^1\mu \xrightarrow{\text{mutation}} K_1^1\mu+\upsilon \xrightarrow{\text{reproduction}} (K_1^1\mu+\upsilon)/2. \tag{7}$$

Thus, the quasiequilibrium value of μ before selection is:

$$\hat{\mu} = \upsilon/(2-K_1^1) \tag{8}$$

and, because of (5), the coefficient of selection against m_2 is

$$\omega = -\upsilon\hat{\sigma}^{-1}K_1^0\,\overline{W}^{-1}/(2-K_1^1). \tag{9}$$

Because $\hat{\sigma} = -U/\delta$ (eqn (15) from K-95), $K_1^0\,\overline{W}^{-1} = \delta$ ((1 *a*) and eqn (7) from K-95), and $K_1^1 = \rho$ ((1 *c*) and eqn (8) from K-95), this implies

$$\omega = \frac{\upsilon\delta^2}{U(2-\rho)} = \upsilon\hat{s}, \tag{10}$$

where $\hat{s} = \delta^2[U(2-\rho)]^{-1}$ is the coefficient of selection against a mutant allele in a quasiequilibrium population (eqn (20) from K-95). Thus, if a modifier allele with the effect υ causes the mutation rate $U+\upsilon$ (additive modifier), the derivative of the coefficient of selection against such allele by its effect is

$$\frac{d\omega}{d\upsilon} = \frac{\delta^2}{U(2-\rho)}. \tag{11}$$

Alternatively, if a modifier allele with the effect ϵ causes the mutation rate $U(1+\epsilon)$ (multiplicative modifier), this derivative is

$$\frac{d\omega}{d\epsilon} = \frac{\delta^2}{(2-\rho)} \tag{12}$$

and thus does not depend on U (Fig. 2).

If mutation occurs before selection (which is probably unrealistic, because new mutations are manifested only in the next generation), we have the following transformations of μ:

$$\mu \xrightarrow{\text{mutation}} \mu+\nu \xrightarrow{\text{selection}} K_1^1(\mu+\upsilon) \xrightarrow{\text{reproduction}} K_1^1(\mu+\upsilon)/2 \tag{13}$$

and the quasiequilibrium μ before mutation is

$$\hat{\mu} = \upsilon K_1^1/(2-K_1^1). \tag{14}$$

However, here selection acts after mutation, so that before selection the difference in the mean number of mutations in individuals with alleles m_2 and m_1 is $\hat{\mu}+\upsilon$. Thus, according to (5),

$$\begin{aligned}\omega &= -\upsilon\hat{\sigma}^{-1}K_1^0\,\overline{W}^{-1}[K_1^1/(2-K_1^1)+1] \\ &= -2\upsilon\hat{\sigma}^{-1}K_1^0\,\overline{W}^{-1}/(2-K_1^1)\end{aligned} \tag{15}$$

and here ω is two times larger than in (9), so that the life cycle (13) leads to the twofold increase of the intensity of selection for a lower mutation rate compared to (7).

Let us now consider selection and mutation in the diplophase. Because haploid selection is absent, at the beginning of the diplophase $\mu_1 = \mu$ and $\mu_2 = 2\mu$. Then μ_1 and μ_2 are transformed in the following way:

$$\left.\begin{aligned}\mu_1 &\xrightarrow{\text{selection}} K_1^1\mu_1 \xrightarrow{\text{mutation}} K_1^1\mu_1+\upsilon_1, \\ \mu_2 &\xrightarrow{\phantom{\text{selection}}} K_1^1\mu_2 \xrightarrow{\phantom{\text{mutation}}} K_1^1\mu_2+\upsilon_2.\end{aligned}\right\} \tag{16}$$

After this, the value of μ in the next generation (after meiosis) can be found from (6). If the modifier is semidominant, i.e. $\upsilon_1 = \upsilon$ and $\upsilon_2 = 2\upsilon$, (6) reduces to (3) (see A 3–8), and $\hat{\mu}$ can be found from (8). Then the rest of the analysis also coincides with the case of haploid selection, so that secondary selection at m does not depend on in which phase selection and mutation occur.

Semidominance at m seems plausible, because for genes with small effects heterozygotes usually have properties intermediate between those of the homozygotes. However, in principle we can consider any pattern of mutation rates in three diploid genotypes at m. The frequencies of m_1 in meiospores of successive generations are connected by ((6) and (16)):

$$M = (K_1^1\mu+\upsilon_2-\upsilon_1)/2+a(\upsilon_1-\upsilon_2/2) \tag{17}$$

which yields

$$\hat{\mu} = [\upsilon_2-\upsilon_1+a(2\upsilon_1-\upsilon_2)]/(2-K_1^1). \tag{18}$$

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

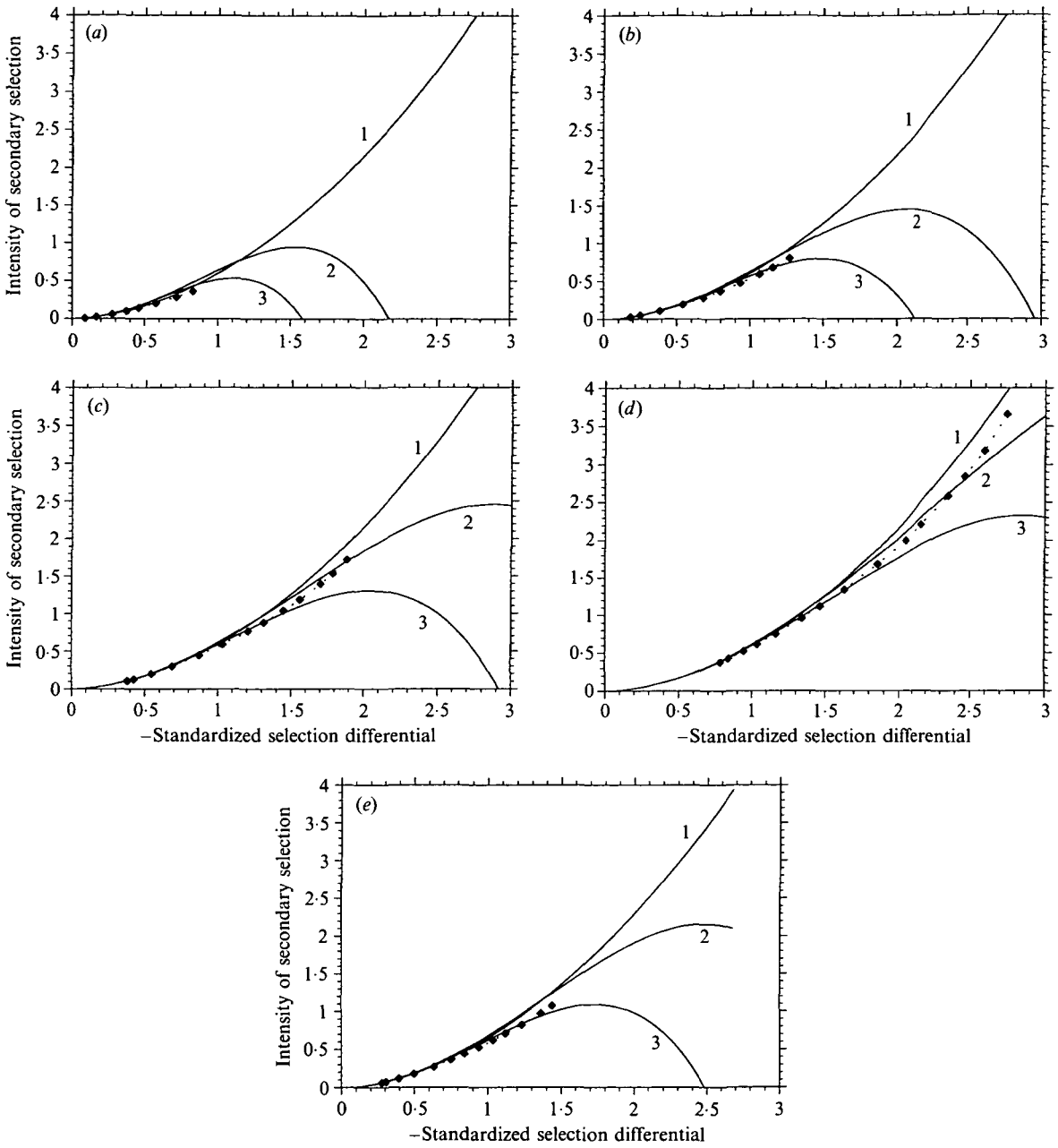

Fig. 2. Intensity of the secondary selection for lower mutation rate. The derivative of the coefficient of selection against a multiplicative modifier with respect to its effect is presented as a function of the negative standardized selection differential $-\delta$. Strict truncation (i.e. $W_w = 0$) (*a*)–(*d*) $W_w = 2{\cdot}0$ (*e*). (*a*) $U = 1{\cdot}0$; (*b*) $U = 2{\cdot}0$; (*c*) $U = 4{\cdot}0$; (*d*) $U = 8{\cdot}0$; (*e*) $U = 3{\cdot}0$. Only changes of the mean are taken into account (1), complete model with shift mutation (2); (*c*) complete model with Poisson mutation (3); (*d*) numerical results (diamonds).

Of course, (18) reduces to (8) if $v_2 = 2v_1$. Otherwise, we can see that $\hat{\mu} = 0$ if

$$a = \hat{a} = (v_2 - v_1)/(v_2 - 2v_1). \tag{19}$$

Thus, if the frequency of allele m_1 after (and before) meiosis equals $\hat{a}$, the distributions of x in all three modifier genotypes have the same means. Therefore, as long as we ignore the differences in their variances, $\hat{a}$ is an equilibrium frequency, and under $a = \hat{a}$ the population is at equilibrium (not just quasi-equilibrium), because no changes occur in it.

It is easy to see that this is possible (i.e. $0 < \hat{a} < 1$) in two cases: (1) $v_1 < 0$ and $v_1 < v_2$ and (2) $v_1 > 0$ and $v_1 > v_2$. Obviously, in case (1) a heterozygote has the mutation rate lower than both homozygotes, while in case (2) it has the highest mutation rate. In case (1) the equilibrium is stable, because μ is positive when $a < \hat{a}$, so that allele m_2 have lower fitness and a grows, while when $a > \hat{a}$, μ is negative and a decreases. In this case the locus m is, in effect, overdominant (Karlin & McGregor, 1974, pp. 68, 78, 97). In contrast, in case (2) the locus m is underdominant and $\hat{a}$ is unstable, so that either m_1 or m_2 is fixed, depending on the initial conditions.

If mutation in the diplophase occurs before selection, selection at m gets stronger, similarly to the case

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

of haploid selection (15). The unrealistic situations when mutation occurs in haplophase and selection – in diplophase, or vice versa, can also be considered.

(iii) *Differences in both means and variances*

Let us now take into account both μ and ν. Assume that selection precedes mutation, that they both occur in the haplophase, and that during mutation difference in mutation rates influences only μ, and not ν (shift mutation, see K-95). This leads to the following transformations of μ and ν:

$$\left.\begin{aligned} &\mu \xrightarrow{\text{selection}} K_1^1\mu+\hat{\sigma}^{-1}K_2^1\nu \xrightarrow{\text{mutation}} K_1^1\mu+\hat{\sigma}^{-1}K_2^1\nu \\ &\quad +\upsilon \xrightarrow{\text{reproduction}} (K_1^1\mu+\hat{\sigma}^{-1}K_2^1\nu+\upsilon)/2, \\ &\nu \longrightarrow \hat{\sigma}K_1^2\mu+K_2^2\nu \longrightarrow \hat{\sigma}K_1^2\mu+K_2^2 \\ &\quad \longrightarrow (K_1^1\mu+\hat{\sigma}^{-1}K_2^1\nu+\upsilon+\hat{\sigma}K_1^2\mu+K_2^2\nu)/4. \end{aligned}\right\} \quad (20)$$

Thus, the quasiequilibrium values of μ and ν can be found from the system:

$$\left.\begin{aligned} &(K_1^1-2)\mu+\hat{\sigma}^{-1}K_2^1\nu=-\upsilon \\ &(K_1^1+\hat{\sigma}K_1^2)\mu+(\hat{\sigma}^{-1}K_2^1+K_2^2-4)\nu=-\upsilon. \end{aligned}\right\} \quad (21)$$

These values are:

$$\left.\begin{aligned} \hat{\mu} &= \frac{\upsilon(4-K_2^2)}{(2-K_1^1)(4-K_2^2)-K_2^1(K_1^2+2\sigma^{-1})}, \\ \hat{\nu} &= \frac{u(2+\sigma K_1^2)}{(2-K_1^1)(4-K_2^2)-K_2^1(K_1^2+2\sigma^{-1})}. \end{aligned}\right\} \quad (22)$$

Using the same reasoning as in deriving (9)–(12), we can see that for a multiplicative modifier the derivative of the coefficient of selection against it by its effect is

$$\frac{d\omega}{d\epsilon}=\frac{\delta^2[4-K_2^2-K_2^0(2/U-\delta^{-1}K_1^2)]}{(2-\rho)(4-K_2^2)-K_2^1(K_1^2-2\delta/U)}. \quad (23)$$

In the case of Poisson mutation, where both μ and ν are incremented by υ after mutation (K-95), the system of equations analogous to (21) differs from (21) only by having -2υ, instead of υ, in the right-hand part of the second equation. This leads to

$$\frac{d\omega}{d\epsilon}=\frac{\delta^2[4-K_2^2-(\delta/U)K_2^1-K_2^0(4/U-\rho/U-\delta^{-1}K_1^2)]}{(2-\rho)(4-K_2^2)-K_2^1(K_1^2-2\delta/U)}. \quad (24)$$

Data from Fig. 2 show that the difference of (23) or (24) from (2) is smal, as long as δ is consistent with the Gaussian approximation (i.e. when the quasi-equilibrium mean number of mutations is not too small). Obviously, under a given U this requires $-\delta \ll U$ (eqn 19 from K-95). Note, that with Poisson mutation after mutation $\sigma=\sqrt{U}$ even when all the individuals with at least one mutation are eliminated by selection, so that $-\delta$ cannot be smaller than $\sqrt{U}$. Thus, the influence of the differences in variances of the distributions of x among the genotypes of the modifier locus on secondary selection in it is small.

The alternative succession of events (mutation–selection–reproduction) and selection and/or mutation in the diplophase can all be analysed analogously. As before, taking the variance of x into account does not lead to large changes of secondary selection at m (the results are not reported).

(iv) *Very low numbers of mutations per genome*

When $\hat{M}[p]$ is too small to use the Gaussian approximation, we still can easily consider the evolution of the mutation rate if $\hat{M}[p] \ll 1$ (see K-95). In this case, which occurs if U is small while selection against mutations is not too weak, no individual carries more than one mutant allele, and M_1 and M_2 are the frequencies and meiospores carrying a mutation among those with alleles m_1 and m_2, respectively. The association between a modifier locus and the selectable loci can be completely described by μ. Because individuals with a mutation are rare, selection reduces their frequency by the factor of $1-s$, where $s>0$ is the coefficient of selection against such individuals. Then, with haploid selection and mutation following selection, we have, analogously to (7):

$$\mu \xrightarrow{\text{selection}} \mu(1-s) \xrightarrow{\text{mutation}} \mu(1-s)+\upsilon \xrightarrow{\text{reproduction}} (\mu(1-s)+\upsilon)/2. \quad (25)$$

Thus, the quasi-equilibrium value of μ before selection is:

$$\hat{\mu}=\upsilon/(1+s) \quad (26)$$

and the selection coefficient against m_2 is

$$\omega=\upsilon s/(1+s). \quad (27)$$

In the case of haploid selection and mutation preceding selection, ω is two times bigger, analogously to (15). Diploid selection can also be analysed.

(v) *Numerical results*

A numerical model which incorporates an infinite number of selectable loci subject to mutation and a modifier of the mutation rate was developed in Kondrashov (1984) (the current implementation of this model is written in Think *C* and is available on request). Fig. 2 presents the data from investigation of this model. Two alleles of m were considered. Allele m_1 caused the mutation rate U, and the allele m_2 to the mutation rate $U(1+\epsilon)$, where $\epsilon=10^{-7},\ldots,10^{-2}$ (the results were almost identical in these cases, while with higher ϵ the population state became too far from quasi-equilibrium). We can see that the agreement with the analytical estimates is good, as long as the Gaussian approximation is valid.

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

5. The cost of fidelity and the equilibrium mutation rate

(i) *The problem*

So far we have ignored selection that can act directly on the modifier of the mutation rate. Thus, the decrease of this rate was always favoured due to secondary selection, because deleterious mutations, although not affecting the individuals in which they first appear, reduce the fitness of their progeny. However, if the mutation rate approaches zero, the cost of fidelity of DNA handling in terms of both time and energy grows unlimitedly (see Kirkwood *et al.* 1986). The balance between secondary selection (due to genetics) for lower mutation rate and direct selection (due to physiology) for higher mutation rate may lead to an equilibrium rate.

In order to find this equilibrium rate we need to know w_{phys}, the fitness of an individual with a given mutation rate in it, relative to that of individuals with very high rate that suffer no cost of fidelity. In this section, I will distinguish between u, the mutation rate in individuals of a given genotype, and U, the average mutation rate in the population. Because w_{phys} depends on intraorganismal processes, it is a function of u alone.

Although no direct data are available, it seems very plausible that w_{phys} monotonously increases with u, being zero if $u = 0$ and one if $u \to \infty$. I will consider a family of such functions:

$$w_{phys}(u) = u^n/(C^n + u^n). \tag{28}$$

The fitness equals one-half of maximal (i.e. 0·5) when $u = C$. When n grows, w_{phys} approaches both 0 and 1 faster, so that the cost of fidelity rapidly changes from negligible with $U > C$ to prohibitive with $U < C$ (Fig. 3). Perhaps, in reality, $n \geqslant 1$.

The total fitness of a genotype at the modifier locus is $w_{tot} = w_{gen}\, w_{phys}$, where w_{gen} is the fitness component caused by the secondary selection considered in Section 4. This component has a different structure in amphimictic and apomictic populations, which therefore will be studied separately.

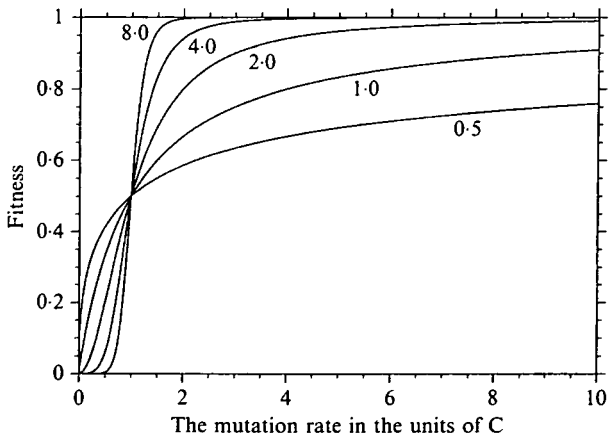

Fig. 3. The fitness of an organism as the function of u/C under various values of n (marked in the figure), according to (28).

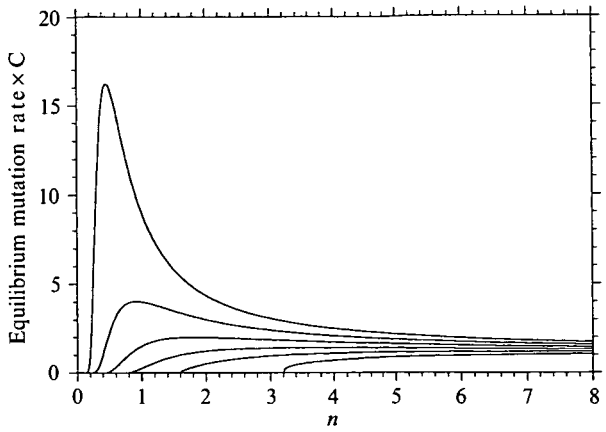

Fig. 4. The equilibrium mutation rate (multiplied by C) as the function of n under $\alpha = 0{\cdot}1$, $0{\cdot}2$, $0{\cdot}4$, $0{\cdot}8$, $1{\cdot}6$, and $3{\cdot}2$ (the corresponding lines are of the decreasing heights), according to (32).

(ii) *Equilibrium mutation rate in amphimictic populations*

In a quasi-equilibrium amphimictic population where the Gaussian approximation is applicable and the average mutation rate is U, a genotype with a similar mutation rate u has the relative fitness, due to individual secondary selection,

$$w_{gen}(u, U) = 1 - \alpha(u - U)/U, \tag{29}$$

where $\alpha = \delta^2/(2-\rho) > 0$ is a parameter of the mode of selection (I assume that selection occurs before mutation and use the simplest formula (10), because taking into account the variance does not make much difference). Thus, here w_{gen} depends on both u and U.

With amphimixis, the problem of the equilibrium mutation rate must be addressed in terms of the average mutation rate in the population. We need to find an evolutionary stable value of U, U_{eq}, i.e. such a U that will not change as long as all individuals have similar values of u. Consider a function

$$f(u, U) = \frac{\partial w_{tot}(u, U)}{\partial u} = \frac{\partial [w_{gen}(u, U)\, w_{phys}(u)]}{\partial u}. \tag{30}$$

We can find U_{eq} from the condition $f(U_{eq}, U_{eq}) = 0$. Because

$$f(u = U, U) = \frac{U^{n-1}}{(C^n - U^n)^2}[-\alpha U^n - \alpha C^n + nC^n] \tag{31}$$

this implies that either $U_{eq} = 0$ or

$$U_{eq} = C\left(\frac{n}{\alpha} - 1\right)^{1/n} \tag{32}$$

(Fig. 4). It is clear from (31) that equilibrium (32) is always globally stable as long as it yields a positive

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

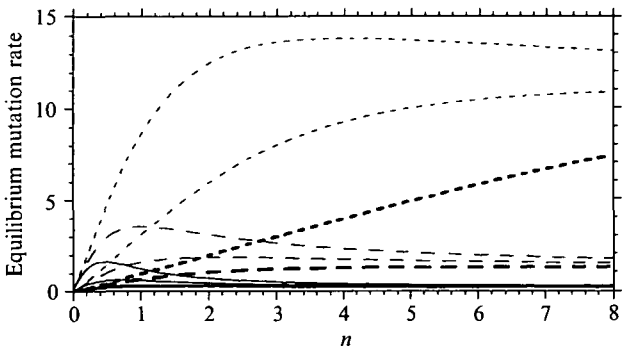

Fig. 5. The equilibrium mutation rate as the function of n under $C = 0{\cdot}1$ (solid lines), $1{\cdot}0$ (medium dashed), and $10{\cdot}0$ (short dashes) and $S = 1{\cdot}0$, $0{\cdot}25$, and $0{\cdot}0625$ (the corresponding lines of the same pattern are of the increasing heights, the lines corresponding to $S = 1{\cdot}0$ are thick), according to (35).

U_{eq}, while otherwise $U_{eq} = 0$ is globally stable. Because α cannot be larger than approximately 5 ($\delta < 2(3)$ and $\rho \leqslant 1$, see K-95), $U_{eq} \approx C$ under large n. If $\alpha < n/2$, i.e. when direct selection against mutations and secondary selection to reduce their rate is weak, $U_{eq} > C$, while otherwise $U_{eq} < C$, although this is important only when n is not high.

If $\alpha > n$, the mutation rate decreases without limit, which would eventually lead to extinction, because $w_{tot} \to 0$ when $u \to 0$. In principle, this may be possible, because we consider individual selection. However, in the course of such a decline of U, $\hat{M}[p(x)]$ would tend to zero (K-95, eqn (19)), which makes the assumption that $p(x)$ is Gaussian inapplicable. Thus, (32) cannot be used unless it predicts $U_{eqn} \gg \alpha^{\frac{1}{2}}$, because only then $\hat{M}[p] \gg 1$.

When (32) yields a smaller U_{eq}, we may consider the asymptote for small $\hat{M}[p]$ (26). Letting $S = s/(1+s)$ (mutation occurs after selection), or $S = 2s/(1+s)$ (mutation before selection), we have

$$w_{gen}(u, U) = 1 - S(u - U) \tag{33}$$

and

$$f(u = U, U) = \frac{U^{n-1}}{(C^n + U^n)^2}[-SU^{n+1} - SC^n U + nC^n]. \tag{34}$$

This implies that either $U_{eq} = 0$ or it can be found from an equation

$$-U^{n+1} - C^n U + \frac{n}{S} C^n = 0. \tag{35}$$

This equation always has exactly one positive root $U_{eq} > 0$ (Fig. 5), which is globally stable. Obviously, U_{eq} grows with C, i.e. when the cost of fidelity increases, and declines if S grows, i.e. when selection against those that carries a mutant allele gets stronger. We can use (34) only if $U_{eqn}/s \ll 1$, because otherwise $\hat{M}[p]$ is too large for (26) to be applicable. When $C \to \infty$, U_{eq} approaches n/S, which implies $\hat{M}[p] \approx n/s^2$. Thus, with high C (35) cannot be used unless $n \ll 1$, which is probably unrealistic.

With $C \ll 1$, we have approximately

$$U_{eq} \approx \sqrt[n+1]{(nC^n/S)} \tag{36}$$

and, as in the case of (31), under high n $U_{eq} \approx C$. With $n = 1$, $U_{eq} = \sqrt{(C/S)} > C$, and then U_{eq} decreases and approaches C when n grows. With high n, $C \ll s$ is sufficient for $U_{eq}/s \ll 1$, while with low n, C must be even smaller. Of course, if (32) predicts a too low U_{eq}, while (35) predicts a too high U_{eq}, the real equilibrium mutation rate does not allow either asymptotic to be applied, and its evolution have to be studied numerically.

(iii) *Equilibrium mutation rate in apomictic populations*

In Section 4 I did not study selection for lower mutation rates in apomictic populations, because it has been done before. The fitness of a clone with a mutation rate u is, relative to that of a clone with zero mutation rate, e^{-u} (Kimura & Maruyama, 1966). In contrast to amphimixis, here we do not have to consider U, because an apomictic population is just a set of genetically independent clones. Thus, w_{gen}, as well as w_{phys}, depends only on u, and the problem of the evolution of the mutation rate must be addressed in terms of group selection. A stable equilibrium rate u_{eq} simply maximizes the total fitness of a clone with $u = u_{eq}$. Therefore, $g(u_{eq}) = 0$, where

$$g(u) = \frac{\partial[w_{gen}(u)\, w_{phys}(u)]}{\partial u}$$

$$= \frac{e^{-u}u^{n-1}}{(C^n + u^n)^2}(-u^{n+1} - C^n u + nC^n) \tag{37}$$

and a stable equilibrium $u_{eq} = 0$ can be found from an equation

$$-u^{n+1} - C^n u + nC^n = 0. \tag{38}$$

Obviously, (38) is a special case of (35) with $S = 1$. Thus, with apomixis the equilibrium mutation rate is no lower than with amphimixis under low $\hat{M}[p]$. The equality is reached only when mutations are lethal and occur before selection, which corresponds to $S = 1$. In this case recombination has no effect, while otherwise it makes selection for reducing the mutation rate less efficient. Even with a very high C, $u_{eq} < n$, so that with apomixis the equilibrium mutation rate cannot be very high, although, of course, there is no limitations on using (38), in contrast to (35).

6. Impact of changes of the mutation rate on the population

Even if all the modifier genotypes available at a given moment are similar, in the long run the trait that is controlled by many modifier loci can change sub-

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

stantially. Thus, even a slow evolution, such that in its course a population always remains in quasi-equilibrium, can eventually lead to large changes.

Let us consider the results of such evolution of the mutation rate. When U is invariant, the mutation-selection equilibrium under soft selection, in an amphimictic population depends on U and on two parameters of such selection, δ and ρ (K-95, eqn 19). If, in contrast, we allow U rate to evolve, its equilibrium value U_{eqn} is determined by the cost of fidelity (which in the case of (28) depends on two parameters, C and n), together with δ and ρ (in the form of α) (32). In terms of these 'ultimate' parameters, the mean number of mutations in the genome at the mutation-selection equilibrium is

$$\hat{M}_{eq} = C^2[(2-\rho)/\delta^2][\sqrt[n]{(n(2-\rho)/\delta^2-1)}]^2 \qquad (39)$$

while the variance $\hat{V}_{eq} = \hat{M}_{eq}(2-\rho)^{-1}$ (K-95, eqn 19). Analogously to (32), (39) can be applicable only when it predicts $\hat{M}_{eq} \gg 1$.

Discussion

(i) *General approach to the evolution of modifiers*

The basic assumption that all the genotypes at the modifier loci that simultaneously have non-negligible frequencies are similar seems plausible, because evolution of the reproductive traits is thought to be usually slow. However, there are some exceptions (Nothel, 1987; Brooks, 1988; Korol *et al.* 1990; Morgan & Barrett, 1990). Some reproductive traits (e.g. the rates of mutation or recombination) can vary continuously, although individual alleles may still influence them drastically. In contrast, the mode of reproduction (amphimixis *v.* apomixis), as well as some forms of inbreeding (outcrossing *v.* selfing) are inherently discrete, and a single allele substitution can naturally have a large effect on them (Hebert *et al.* 1989). However, even here two modifier alleles can be similar, if they set similar probabilities of different outcomes or similar allocation of resources to them.

Besides creating a hierarchy of the processes leading to quasi-equilibrium (Section 3), the basic assumption causes some simplifications specific to the model of mutation-selection balance which I have considered. In particular, it allows us to assume that all the distributions involved are Gaussian and to use the formulae (1). Modifiers of various processes (Taylor & Williams, 1982; Charlesworth, 1990; Iwasa *et al.* 1991; Uyenoyama & Waller, 1991) can be investigated similarly, as long as the quasi-equilibrium differences between the distributions of the number of mutations within different modifier genotypes can be found. For modifiers of meiosis (recombination) this can probably be achieved by applying Bulmer's (1985) approach, as was done by Charlesworth (1990) for Gaussian selection, to any mode of soft selection. The models become more complex when, as in the case of facultative apomixis or selfing of diploids, the haplophase becomes dynamically insufficient (Charlesworth *et al.* 1991).

(ii) *Minimal v. optimal mutation rates*

If, as it is assumed here, the mutation process always converts 'normal' alleles into unconditionally deleterious ones, a zero mutation rate is favoured. Actually, the same is true under a wider range of assumptions. A zero rate is favoured under any invariant mode of selection, even if a genotype with the highest fitness is heterozygous, so that some mutations can increase the fitness on some genetic backgrounds, as long as the population size is infinite and mating is random (Karlin & McGregor, 1974; Liberman & Feldman, 1986; Twomey & Feldman, 1990).

Under other conditions, however, selection can favour a positive mutation rate. Initially, the optimal hypothesis was based on group selection (Berg, 1941, 1942, 1948) and related genetic load (Kimura, 1960, 1967; Levins, 1967) arguments, but later the analysis based on individual selection, which is more appropriate for amphimictic populations, was developed (Leigh, 1970, 1973). Under invariant direct selection positive mutation rates can be favoured only when an equilibrium population is polymorphic and either the population size is finite (Gillespie, 1981*a*), mating is non-random (Holsinger & Feldman, 1983), or selection acts on fecundity (Holsinger *et al.* 1986; Twomey & Feldman, 1990). Changing selection generally favours positive mutation rates (Gillespie, 1981*b*; Semenov & Terkel, 1985; Ishii *et al.* 1989), although zero mutation rates can also be established under some values of the parameters (Leigh, 1970, 1973; Gillespie, 1981*b*).

Although we do not have a general criterion which tells us when selection favours zero or a positive mutation rate, it seems that in any model in which the vast majority of new non-neutral mutant alleles are unconditionally deleterious a zero-rate is always favoured. Even if such mutant alleles are only as frequent as those which can be utilized for adaptation under some conditions, secondary individual selection favours a zero rate (Leigh, 1973, pp. 3 and 13). Thus, the minimal hypothesis seems applicable to amphimictic populations.

The minimal hypothesis is supported by the data (see Lindahl, 1993) on the enormous rate of 'spontaneous' DNA damage *in vivo* (although under some other condition DNA might be more stable, Poinar, 1993) which suggests that the cost of fidelity of molecular processes can be substantial (Kirkwood *et al.* 1986), thus providing the explanation of why the actual mutation rates may be far from zero (see Mohrenweiser, 1994).

In contrast, for apomictic populations, where group selection and genetic load arguments are relevant, the optimal hypothesis may be applicable, at least

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

theoretically (Leigh, 1973; Eshel, 1973; Painter, 1975). However, selection for mutability reduction caused by deleterious mutations is also much stronger with apomixis than with amphimixis. In nature the genomic deleterious mutation rate U is probably only about 0·003 per cell division in unicellular forms, that are either obligately or facultatively apomictic (Drake, 1991). This argues against any significant role of beneficial mutations in the evolution of mutation rate even with obligate apomixis, but more data, particularly on obligately apomictic multicellular forms and their amphimictic relatives, are needed.

(iii) *Selection for reduction of the mutation rate*

The following qualitative analysis may help to understand the main result (10). Suppose that before selection the distribution of x in the genotype which increases the mutation rate by υ is simply shifted by υ to the right. Because of (1*a*), (5), and eqn (7) from K-95 this leads to the decrease of relative fitness by $\upsilon\delta/\hat{\sigma}$, which, because $\hat{\sigma} = U/\delta$ (eqn 16 from K-95) is equal to $\upsilon\delta^2/U$. Because $\rho \approx 1$ (K-95), this is similar to (10), which was obtained for mutation occurring after selection, and is only about two times smaller than (15), where mutation occurs before selection. Thus, ignoring the action of the modifier allele in the previous generations leads to a discrepancy close to a factor of two. Consideration of the difference in variances leads to only small refinements (Fig. 2).

As long as the modifier locus m is assumed to be semidominant, secondary selection for reduction of the mutation rate does not depend on whether selection and mutation occur in the haplophase or in the diplophase. Otherwise, selection in the diplophase can lead to new phenomena, including stable polymorphism at m. However, substantial deviations from semidominance at a locus with small allele effects do not seem plausible.

The intensity of selection against an allele which increases the mutation rate by a factor $1+\epsilon$ (a 'multiplicative' modifier), which seems realistic, does not depend on U and is proportional to the square of the standardized selection differential, δ^2 (K-95). Thus, when the negative selection differential $-\delta$, which is a parameter of soft selection, gets smaller, the intensity of secondary selection for a lower mutation rate rapidly declines. This may imply that U (and, therefore, $\hat{M}[p]$) should be much higher in species with a low intensity of selection (e.g. low fecundity, high random mortality, etc.), although the analysis of evolutionary equilibrium mutation rates is required for such conclusions (see below).

The above analysis can be applied only if $\hat{M}[p] \gg 1$, which requires $U \gg -\delta$. Thus, although (10) predicts that selection for reduction mutation rate tends to infinity when $U \to 0$, in fact (27) must be used in this case, instead of (10). Because according to (10) $\omega = \upsilon\hat{s}$, comparison of (10) and (27) suggests that intensity of secondary selection for a lower mutation rate is similar under any $\hat{M}[p]$.

Thus, when direct selection acts against unidirectional unconditionally deleterious mutations, the intensity of secondary selection for a lower mutation rate is, analogously to that of apparent stabilizing selection on a quantitative trait (Kondrashov & Turelli, 1992, eqn 23) approximately proportional to the coefficient of selection against a mutant allele. It is unclear why this coefficient is so important. However, the similarity between these two cases is not surprising. Actually, the only difference between 'secondary' and 'apparent' selection is that the former is directional and thus leads to the evolution at the modifier locus, while the latter is stabilizing and the distribution of the quantitative trait remains constant.

I have ignored the differences of mutation rates between sexes (see Redfield, 1994). This situation requires special analysis and the results will probably depend on whether the mutation rates in two sexes can evolve independently. Only 'ideal' amphimixis was considered, with no linkage, random mating, or facultative apomixis. Perhaps these factors can lead to a stronger secondary selection for lower mutation rates, because they increase associations of a modifier allele with the consequences of its action (Sturtevant, 1937).

Under apomixis, the fitness of a modifier allele is simply the mean fitness of a clone which carries it, and the fitness of a genotype with the mutation rate U is e^{-U} (Kimura & Maruyama, 1966, see above). Thus, the coefficient of selection against an allele which increases the mutation rate by υ is simply υ. Comparing this with (10) or (27) we can conclude that a secondary selection for a lower mutation rate is always stronger with apomixis than with amphimixis (only when lethal mutations occur before selection these two cases are identical). Thus, although amphimixis can increase the efficiency of direct selection against mutations (see Kondrashov, 1988, 1993), it leads to a weaker selection for a lower mutation rate.

(iv) *Equilibrium mutation rate*

Even if secondary selection always favours a zero mutation rate, such a rate is impossible because of the cost of fidelity (Kirkwood *et al.* 1986) and a non-zero evolutionary stable equilibrium mutation rates U_{eq} will be established. With a given C, the absolute fitness of a population with $U = U_{eq}$ grows with n, i.e. when the decline of fitness of an individual with the decrease of the mutation rate in it gets more drastic. Of course, under a given $w_{phys}(u)$ the U_{eq} and the equilibrium absolute fitness decline as the intensity of secondary selection against mutations increases, while with $-\delta \to 0$ $U_{eq} \to \infty$. Thus, a high intensity of soft selection (e.g. in the species with high fecundity) may

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

lead to a low absolute fitness, because of low U_{eq}. However, if n is high, U_{eq} is mostly determined by C.

Formula (32) can be used only if it predicts U_{eq} large enough so that $\hat{M}_{eq} \gg 1$ (39). Alternatively, if $\hat{M}_{eq} \ll 1$, (35) must be used, while the intermediate cases can be studied numerically. Apparently, under any $\hat{M}_{eq}$ amphimixis leads to a higher U_{eq} than apomixis. The difference can be especially large when $C \gg 1$. In this case, the transition from amphimixis to apomixis (and, perhaps, from outcrossing to selfing) could lead to a profound, although perhaps slow, decline of the mutation rate. This is consistent with the fact that selfing usually leads to lower molecular variability (Barrett & Husband, 1990; Njiokou *et al.* 1993).

This might have interesting implications for the evolution of amphimixis and apomixis. Although transition from apomixis to amphimixis may reduce the mutation load in the short run, subsequent slow evolution can lead to a much higher mutation rate and, perhaps, even to a higher load. However, this may only complicate the backward transition to apomixis, because the new clones will suffer more because of the higher mutation rate.

Thus, amphimictic and apomictic populations may differ mostly not by the mutation loads, but by the mutation rates. We can expect apomixis in the forms where $C \ll 1$, so that the U_{eq} is low. If, in contrast, $C \gg 1$, amphimixis can evolve, leading to the high U_{eq}. Transition to amphimixis might increase C, e.g. because of the high number of cell divisions in the male germ line (see Redfield, 1994). Low mutation rates in unicellulars, which are at least facultatively apomictic (Drake, 1991), as well as a correlation between amphimixis and a high genome size, are consistent with this conclusion.

(v) *Consequences of evolution of the mutation rates*

If the rate of evolution of mutation rate is slow, the state of the population, particularly the mean and the variance of the number of mutations per genome, can follow it practically without delay. Thus, the population remains in quasi-equilibrium and there is no coevolution of the mutation rate and the population state, because these two processes have different characteristic time scales.

If the mutation rate reaches an evolutionary stable equilibrium U_{eq}, the mean number of mutations per genome reaches $\hat{M}_{eq}$, which is extremely sensitive to $-\delta$. According to our model, the parameters of soft selection, $-\delta$ and ρ, determine (together with C and n) U_{eq}, as well as $\hat{M}$ and $\hat{V}$. If w_{pop} has a plateau with small x, the increase of U_{eq} does not cause a higher load (K-95). In contrast, under a given hard selection the load is determined by U_{eq}.

(vi) *The necessary experimental data*

The importance of the processes considered in this paper depends on several parameters which must be measured experimentally. First, we must know the fraction that unconditionally deleterious mutations contribute to new mutations. Conditionally beneficial mutations can be ignored, as was done here, only if this fraction is high. Then, we must know the genomic deleterious mutation rate, and the mode of selection against such mutations. Very little is known about this so far (Crow & Simmons, 1983; Houle *et al.* 1992). Finally, we need to know the dependence of the fitness of an organism on the mutation rate, which has never been measured directly, although some indirect data are available (Nothel, 1987). The potential importance of slightly, although unconditionally, deleterious mutations makes the relevant experimental work desirable.

This and the paper K-95 were started seven years ago, but various external factors did not allow me to finish them. I am grateful to the Section of Ecology and Systematics of Cornell University which finally gave me a chance to do so. Brian Charlesworth and an anonymous reviewer made many useful comments on the first drafts.

Appendix 1

Partial derivatives of a Gaussian density

$$p_{M,V}(x) = \frac{1}{\sqrt{(2\pi V)}} e^{-(x-M)^2/2V} \qquad (A\,1a)$$

with respect to its mean and variance are:

$$\frac{\partial p}{\partial M} = \frac{x-M}{V} p(x), \qquad (A\,1b)$$

$$\frac{\partial p}{\partial V}\left(-\frac{1}{2V} + \frac{(x-M)^2}{2V^2}\right) p(x). \qquad (A\,1c)$$

From the following general formula (all integrals with omitted limits are taken from $-\infty$ to ∞)

$$\frac{\partial I}{\partial \alpha} = \int \frac{\partial f(x,\alpha)}{\partial \alpha} dx, \quad \text{where} \quad I = I(x,\alpha) = \int f(x,\alpha)\, dx \qquad (A\,1d)$$

derivatives of the I_ks (eqn 2 from K-95) are given by:

$$\frac{\partial I_k}{\partial M} = \int \frac{x-M}{V} x^k p(x)\, w(x)\, dx = \frac{1}{V}(I_{k+1} - MI_k), \qquad (A\,1e)$$

$$\frac{\partial I_k}{\partial V} = \int \left(-\frac{1}{2V} + \frac{(x-M)^2}{2V^2}\right) x^k p(x)\, w(x)\, dx$$

$$= -\frac{1}{2V} I + \frac{1}{2V^2}(I_{k+2} - 2\hat{M} I_{k+1} + \hat{M}^2 I_k). \qquad (A\,1f)$$

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

Let us prove, for example that

$$\frac{\partial \tilde{V}}{\partial M} = -\frac{1}{V} I_0^{-3}(I_3 I_0^2 - 3I_2 I_1 I_0 + 2I_1^3). \qquad (A\,1g)$$

Using (A 1*e*) and (A 1*f*), we have:

$$\frac{\partial \tilde{V}}{\partial M} = \frac{\partial}{\partial M}\left(\frac{I_2}{I_0} - \frac{I_1^2}{I_0^2}\right) = I_0^{-2}\left(\frac{\partial I_2}{\partial M} I_0 - \frac{\partial I_0}{\partial M} I_2\right)$$

$$-2\frac{I_1}{I_0} I_0^{-2}\left(\frac{\partial I_1}{\partial M} I_0 - \frac{\partial I_0}{\partial M} I_1\right)$$

$$= \frac{1}{V} I_0^{-2}\left[I_3 I_0 - M I_2 I_0 - I_2 I_1 + M I_2 I_0 - \frac{2I_1}{I_0}(I_2 I_0 - M I_1 I_0 - I_1^2 + M I_1 I_0)\right]$$

$$= \frac{1}{V} I_0^{-3}(I_3 I_0^2 - 3I_2 I_1 I_0 + 2I_1^3). \qquad (A\,1h)$$

Equations (1) are written in terms of J_ks, instead of I_ks, because expressions which depend only on J_ks remain invariant, as long as $W(X)$ remains invariant. To obtain an expression in terms of J_ks from that in terms of I_ks, we have to use the formulae

$$\left.\begin{aligned} J_0 &= I_0, \\ J_1 &= (I_1 - M I_0)/\sigma, \\ J_2 &= (I_2 - 2M I_1 + M^2 I_0)/\sigma^2, \\ &\cdots \end{aligned}\right\} \qquad (A\,1i)$$

which follow from eqn 3 (K-95).

Appendix 2

Consider a population which consists of two components with frequencies a and $1-a$. The distribution of a quantitative trait x in this population, $p(x)$, is

$$p(x) = a p_1(x) + (1-a) p_2(x), \qquad (A\,2a)$$

where $p_1(x)$ and $p_2(x)$ are the distributions of x within the first and second components, respectively. If the mean and the variance of p_1 and p_2 are M_1 and V_1, and M_2 and V_2, respectively, and $\mu = M_2 - M_1$ and $\nu = V_2 - V_1$, the mean and the variance of p, M and V, are:

$$M = \int x[a p_1(x) + (1-a) p_2(x)]\,dx = a M_1 + (1-a) M_2$$

$$= M_1 + (1-a)\mu,$$

$$V = \int x^2[a p_1(x) + (1-a) p_2(x)]\,dx - M^2$$

$$= a V_1 + (1-a) V_2 + a M_1^2 + (1-a) M_2^2 - M^2$$

$$= V_1 + (1-a)\nu + a(1-a)\mu^2. \qquad (A\,2b)$$

Assume now that both p_1 and p_2 are Gaussian. Then,

$$p(x) = \frac{1}{\sqrt{(2\pi)}}\left(\frac{a}{\sqrt{V_1}} e^{-(x-M_1)^2/2V} + \frac{1-a}{\sqrt{(V_1+\nu)}} e^{-(x-M_1-\mu)^2/2(V+\nu)}\right). \qquad (A\,2c)$$

Let us compare $p(x)$ with the Gaussian distribution $p_G(x)$ having the same mean and variance:

$$P_G(x) = \frac{1}{\sqrt{(2\pi V)}} e^{-(x-M)^2/2V}. \qquad (A\,2d)$$

Now, using Taylor expansions $e^x = 1 + x + \ldots$ and $(1+x)^A = 1 + Ax + \ldots$, it is easy to show that, if we take into account formulae (A 2*b*) for M and V and ignore all but linear terms in μ and ν, $p(x)$ and $p_G(x)$ can be reduced to the same form

$$\frac{1}{\sqrt{(2\pi V_1)}}\left(e^{\frac{-(x-M_1)^2 + 2(1-a)\mu(x-M_1) + (1-a)\nu(x-M_1)^2/V_1}{2V_1}} - \frac{1-a}{2V_1}\nu e^{-(x-M_1)^2/2V_1}\right) \qquad (A\,2e)$$

Appendix 3

Before syngamy, the distributions of x among the gametes with genotypes m_1 and m_2, $p_1(x)$ and $p_2(x)$, are Gaussian with means M_1'' and $M_2'' = M_1'' + \mu''$ and variances V_1'' and $V_2'' = V_{1+\nu''}''$, respectively, and the frequencies of m_1 and m_2 in all the gametes are a and $1-a$, respectively. After syngamy and at the beginning of the diplophase the distributions of x in the components of the population with genotypes $m_1 m_1$, $m_1 m_2$, and $m_2 m_2$, $p_{11}(x)$, $p_{12}(x)$, and $p_{22}(x)$, are assumed to be Gaussian and similar, with the means M_{11}, $M_{12} = M_{11} + \mu_1$, and $M_{22} = M_{11} + \mu_2$ and variances V_1'' and $V_2'' = V_1'' + \nu''$, respectively, and the respectively. Particularly, in the case of panmixia of gametes, $M_{11} = 2M_1''$, $M_{12} = 2M_1'' + \mu''$, and $M_{22} = 2M_1'' + 2\mu''$ and $V_{11} = 2V_1''$, $V_{12} = 2V_1'' + \nu''$, and $V_{22} = 2V_1'' + 2\nu''$, respectively, while the frequencies of $m_1 m_1$, $m_1 m_2$, and $m_2 m_2$ are $f_{11} = a^2$, $f_{12} = 2a(1-a)$, and $f_{22} = (1-a)^2$, respectively. Below panmixia will be always assumed, although the same approach can be used without it, as long as the haplophase remains dynamically sufficient, so that μ_1, μ_2, ν_1, and ν_2 can be determined from μ'' and ν'', while the frequencies of diploid genotypes can be determined from a.

Let us now consider the end of the diplophase, just before meiosis. At this point the means and the variances of x in individuals with genotypes $m_1 m_1$, $m_1 m_2$, and $m_2 m_2$ are M_{11}'', $M_{12}'' = M_{11}'' + \mu_1''$, and $M_{22}'' = M_{11}'' + \mu_2''$ and V_{11}'', $V_{12}'' = V_{11}'' + \nu_1''$, and $V_{22}'' = V_{11}'' + \nu_2''$. The distributions of x in the meiospores produced by different zygotes have the following means and variances (Bulmer, 1985)

zygote genotype	mean	variance
$m_1 m_1$	$M_{11}''/2$	$(M_{11}'' + V_{11}'')/4$
$m_1 m_2$	$(M_{11}'' + \mu_1'')/2$	$(M_{11}'' + \mu_1'' + V_{11}'' + \nu_1'')/4$
$m_2 m_2$	$(M_{11}'' + \mu_2'')/2$	$(M_{11}'' + \mu_2'' + V_{11}'' + \nu_2'')/4.$

(A 3*a*)

Let us assume first that the frequencies of genotypes

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

m_1m_1, m_1m_2, and m_2m_2 remained unchanged during the diploid phase, which implies that selection in this phase is absent. Then among all meiospores with the genotype m_1, the fractions a and $1-a$ are produced by individuals m_1m_1 and m_1m_2, respectively. Thus the mean value of x among meiospores m_1 is

$$M_1 = [aM''_{11} + (1-a)(M''_{11} + \mu''_1)]/2. \quad \text{(A 3}b\text{)}$$

Similarly, the fractions a and $(1-a)$ of all m_2 meiospores are produced by individuals m_1m_2, and m_2m_2, respectively, and the mean value of x among meiospores m_2 is

$$M_2 = [a(M''_{11} + \mu''_1) + (1-a)(M''_{11} + \mu''_2)]/2. \quad \text{(A 3}c\text{)}$$

The variance of x in all meiospores m_1, V_1, equals to the weighted average of the variances in those produced by zygotes m_1m_1 and m_1m_2, ignoring the term containing μ^2 (A 2b). Thus,

$$V_1 = [a(M''_{11} + V''_{11}) + (1-a)(M''_{11} + \mu''_1 + V''_{11} + \nu''_1)]/4. \quad \text{(A 3}d\text{)}$$

Similarly,

$$V_2[a(M''_{11} + \mu''_1 + V''_{11} + \nu''_1) + (1-a)(M''_{11} + \mu''_2 + V''_{11} + \nu''_2)]/4. \quad \text{(A 3}e\text{)}$$

Therefore, subtracting (A 3c) from (A 3b) and (A 3e) from (A 3d) we obtain (6). Let us now consider 2 cases.

(1) Nothing happens during the diplophase, so that population before meiosis is the same as after syngamy. Then, $\mu''_2 = 2\mu''_1 = 2\mu''$ and $\nu''_2 = 2\nu''_1 = 2\nu''$, so that (6) reduces to (3) and selection at the modifier locus does not depend on the allele frequencies at it.

(2) The modifier acts during the diplophase. Then, in general, μ''_1, μ''_2, ν''_1, and ν''_2 can be arbitrary small numbers, which can be determined from μ'' and ν'', and the equations (6) must be used. Thus, M and N and, consequently, selection at the modifier locus, depends on a. However, in a special case of semi-dominant modifier, where the properties of the genotype m_1m_2 are exactly in between of those of the genotypes m_1m_1 and m_2m_2, the equalities $\mu''_2 = 2\mu''_1$ and $\nu''_2 = 2\nu''_1$ are still valid, although $\mu''_1 \neq \mu''$ and $\nu''_1 \neq \nu''$. Then (6) reduces to (3).

Let us now consider selection in the diplophase. Let the average fitnesses of individuals with genotypes m_1m_1, m_1m_2, and m_2m_2 be $1+s_{11}$, $1+s_{12}$, and $1+s_{22}$, respectively. Mutation does not change the modifier genotype frequencies, so that before meiosis they are

$$\left.\begin{aligned} f''_{11} &= \frac{(1+s_{11})f_{11}}{1+s_{11}f_{11}+s_{12}f_{12}+s_{22}f_{22}} \approx f_{11}(1+\epsilon_{11}),\\ f''_{12} &= \frac{(1+s_{12})f_{12}}{1+s_{11}f_{11}+s_{12}f_{12}+s_{22}f_{22}} \approx f_{12}(1+\epsilon_{12}),\\ f''_{22} &= \frac{(1+s_{22})f_{22}}{1+s_{11}f_{11}+s_{12}f_{12}+s_{22}f_{22}} \approx f_{22}(1+\epsilon_{22}), \end{aligned}\right\} \quad \text{(A 3}f\text{)}$$

where

$$\left.\begin{aligned} \epsilon_{11} &= s_{11} - s_{11}f_{11} - s_{12}f_{12} - s_{22}f_{22},\\ \epsilon_{12} &= s_{12} - s_{11}f_{11} - s_{12}f_{12} - s_{22}f_{22},\\ \epsilon_{22} &= s_{22} - s_{11}f_{11} - s_{12}f_{12} - s_{22}f_{22}. \end{aligned}\right\} \quad \text{(A 3}g\text{)}$$

The approximate equalities are true because selection is weak (distributions of x in all three genotypes differ only slightly) and thus s_{11}, s_{12}, and s_{22} (and ϵ_{11}, ϵ_{12} and ϵ_{22}) are small numbers, of the order of μ and ν.

Now we can assess the influence of changes of f_{11}, f_{12} and f_{22} due to diploid selection on M and N. For example, instead of (A 3b) we have

$$\begin{aligned} M_1 &= \frac{(a^2+\epsilon_{11})M''_{11} + \tfrac{1}{2}[2a(1-a)+\epsilon_{12}](M''_{11}+\mu''_1)}{2(a^2+\epsilon_{11}+[2a(1-a)+\epsilon_{12}]/2)}\\ &= \frac{M''_{11}(a+\epsilon_{11}+\epsilon_{12}/2) + \tfrac{1}{2}\mu''_1 a(1-a)}{2a+2\epsilon_{11}+\epsilon_{12}}\\ &= [M''_1 + \mu''_1(1-a)]/2 + O(\mu)O(\epsilon) + O(\epsilon^2). \quad \text{(A 3}h\text{)} \end{aligned}$$

Thus, (A 3h) differs from (A 3b) only by the terms that are quadratic in the differences between distributions of x within different modifier genotypes. The same is true for M_2, V_1, and V_2, so that the changes of modifier genotype frequencies due to selection can be ignored when M, N, and selection at the modifier locus are estimated.

The above derivations assumed that recombination is free. In this case the total number of mutant alleles in a diploid genotype is enough to predict the distribution of x in the meiospores it produces (we ignore homozygotes by mutant alleles, because these alleles are rare). Thus meiosis alone is sufficient to make Bulmer's (1985) results applicable. If, in contrast, there is linkage, formulae (A 3a) for variance become more difficult and depend on whether selection acts in haploid or diploid phase, because in the latter case it creates correlations between distributions of mutations in maternal and paternal genomes of diploid organisms. Perhaps in this case the approach of Charlesworth (1990) should be used.

References

Altenberg, L. & Feldman, M. W. (1987). Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. *Genetics* **117**, 559–572.

Barrett, S. C. H. & Husband, B. C. (1990). Variation in outcrossing rates in *Eichhornia paniculata*: the role of demographic and reproductive factors. *Plant Species Biology* **5**, 41–55.

Berg, R. L. (1941). Role of small mutations in the evolution of mutability. *Doklady Akademii Nauk SSSR* **32**, 71–74.

Berg, R. L. (1942). Mutability as dependent on the degree of isolation of populations of *Drosophila melanogaster*. *Doklady Akademii Nauk SSSR* **36**, 71–75.

Berg, R. L. (1948). On relationship between mutability and selection in natural populations of *Drosophila melanogaster*. *Zhurnal Obshchej Biologii* **9**, 299–313 (in Russian).

Bodmer, W. F. & Felsenstein, J. (1967). Linkage and

https://doi.org/10.1017/S001667230003439X Published online by Cambridge University Press

selection: theoretical analysis of the deterministic two locus random mating model. *Genetics* **57**, 237–265.
Brooks, L. D. (1988). The evolution of recombination rates. In *The Evolution of Sex: an Examination of Current Ideas* (ed. R. E. Michod and B. R. Levin), pp. 87–105.
Charlesworth, B. (1990). Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genetical Research* **55**, 199–221.
Charlesworth, B. (1993). Directional selection and the evolution of sex and recombination. *Genetical Research* **61**, 205–224.
Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1991). Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genetical Research* **57**, 177–194.
Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Theory*. Harper and Row, New York.
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 Jr), pp. 1–35. Academic Press, London.
Demerec, M. (1937). Frequency of spontaneous mutations in certain stocks of *Drosophila melanogaster*. *Genetics* **22**, 469–478.
Dobzhansky, T. (1951). *Genetics and the Origin of Species*. New York, Columbia University Press (3rd edn).
Drake, J. W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proceedings of the National Academy of Sciences USA* **88**, 7160–7164.
Dubinin, N. P., Geptner, M. A., Demidova, Z. A. & Dyachkova, L. I. (1936). Genetic constitution and gene dynamics of wild populations of *Drosophila melanogaster*. *Biologicheskij Zhurnal* **5**, 939–976.
Dubovskij, N. V. (1935). On the question of the comparative mutability of stocks of *Drosophila melanogaster* of different origin. *Comptes Rendus (Doklady) le l'Academie des Sciences de l'URSS* **4**, 95–97.
Eshel, I. (1973). Clone selection and the evolution of modifying features. *Theoretical Population Biology* **4**, 196–208.
Feldman, M. W., Christiansen, F. B. & Brooks, L. D. (1980). Evolution of recombination in constant environment. *Proceedings of the National Academy of Sciences USA* **77**, 4838–4841.
Gillespie, J. H. (1981). Mutation modification in a random environment. *Evolution* **35**, 468–476.
Gillespie, J. H. (1981). Evolution of the mutation rate at a heterotic locus. *Proceedings of the National Academy of Sciences USA* **78**, 2452–2454.
Hebert, P. D. N., Beaton, M. J., Schwartz, S. S. & Stanton, D. J. (1989). Polyphiletic origins of asexuality in *Daphnia pulex*. I. Breeding-system variation and levels of clonal diversity. *Evolution* **43**, 1004–1015.
Holsinger, K. E. & Feldman, M. W. (1983). Modifiers of mutation rate: evolutionary optimum with complete selfing. *Proceedings of the National Academy of Sciences USA* **80**, 6732–6734.
Holsinger, K. E., Feldman, M. W. & Altenberg, L. (1986). Selection for increased mutation rates with fertility differences between matings. *Genetics* **112**, 909–922.
Houle, D., Hoffmaster, D. K., Assimacopoulos, S. & Charlesworth, B. (1992). The genomic mutation rate for fitness in *Drosophila*. *Nature* **359**, 58–60.
Ishii, K., Matsuda, H., Iwasa, Y. & Sasaki, A. (1989). Evolutionary stable mutation rate in a periodically changing environment. *Genetics* **121**, 163–174.
Ives, P. T. (1950). The importance of mutation rate genes in evolution. *Evolution* **4**, 236–252.
Iwasa, Y., Pomiankowski, A. & Nee, S. (1991). The evolution of costly mate preferences. II. The "handicap" principle. *Evolution* **45**, 1431–1442.
Karlin, S. & McGregor, J. (1974). Towards a theory of the evolution of modifier genes. *Theoretical Population Biology* **5**, 59–103.
Kimura, M. (1960). Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load. *Journal of Genetics* **57**, 21–34.
Kimura, M. (1967). On the evolutionary adjustment of spontaneous mutation rates. *Genetical Research* **9**, 23–34.
Kimura, M. & Maruyama, T. (1966). The mutation load with epistatic gene interactions in fitness. *Genetics* **54**, 1337–1351.
Kirkwood, T. B. L., Rosenberger, R. F. & Galas, D. J. (1986). *Accuracy in Molecular Processes*. Chapman and Hall, London.
Kondrashov, A. S. (1984). Deleterious mutations as an evolutionary factor. I. The advantage of recombination. *Genetical Research* **44**, 199–214.
Kondrashov, 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. *J. of Heredity* **84**, 372–387.
K-95 = Kondrashov, A. S. (1995). Dynamics of unconditionally deleterious mutations: Gaussian approximation and soft selection. *Genetical Research* **65**, 113–121.
Kondrashov, A. S. & Turelli, M. (1992). Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics* **132**, 603–618.
Korol, A. B., Preigel, I. A. & Preigel, S. I. (1990). Variability of crossing-over in higher organisms. Kishinev, Shtiinca (in Russian).
Leigh, E. G. (1970). Natural selection and mutability. *American Naturalist* **104**, 301–305.
Leigh, E. G. (1973). The evolution of mutation rates. *Genetics* (Supplement) **73**, 1–18.
Levins, R. (1967). Theory of fitness in a heterogeneous environment. VI. The adaptive significance of mutation. *Genetics* **56**, 163–178.
Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. Harvard University Press, Cambridge.
Liberman, U. & Feldman, M. W. (1986). Modifiers of mutation rate: a general reduction principle. *Theoretical Population Biology* **30**, 125–142.
Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715.
Mishchenko, E. F. & Rozov, N. Kh. (1980). *Differential Equations with Small Parameters and Relaxation Oscillations*. Plenum Press, New York.
Mohrenweiser, H. (1994). Impact of the molecular spectrum of mutational lesions on estimates of germinal gene-mutation rates. *Mutation Research* **304**, 119–137.
Morgan, M. T. & Barrett, S. C. H. (1990). Outcrossing rates and correlated mating within a population of *Eichhornia paniculata* (Pontederiaceae). *Heredity* **64**, 271–280.
Muller, H. J. (1928). The measurement of gene mutation rate in *Drosophila*, its high variability, and its dependence upon temperature. *Genetics* **13**, 279–357.
Nagylaki, T. (1993). The evolution of multilocus systems under weak selection. *Genetics* **134**, 627–647.
Nei, M. & Li, W.-H. (1980). Non-random association between electromorphs and inversion chromosomes in finite populations. *Genetical Research* **35**, 65–83.
Njiokou, F., Bellec, C., Berrebi, P., Delay, B. & Jarne, P. (1993). Do self-fertilization and genetic drift promote a very low genetic variability in allotetraploid *Bulinus truncatus* (Gastropoda: Planorbidae) populations? *Genetical Research* **62**, 89–100.
Nothel, H. (1987). Adaptation of *Drosophila melanogaster* populations to high mutation pressure: evolutionary

adjustment of mutation rates. *Proceedings of the National Academy of Sciences USA* **84**, 1045–1048.
Painter, P. R. (1975). Clone selection and the mutation rate. *Theoretical Population Biology* **8**, 74–80.
Passekov, V. P. & Singh, D. K. (1991). Weak selection in a logistic model of density-dependent genotype fitnesses. *Zhurnal Obshchej Biologii* **54**, 509–520 (in Russian).
Poinar, G. O. (1993). Recovery of antediluvian DNA. *Nature* **365**, 700.
Redfield, R. J. (1994). Male mutation rates and the cost of sex for females. *Nature* **369**, 145–147.
Sasaki, A. (1994). Evolution of antigen drift/switching: continuously evading pathogens. *Journal of Theoretical Biology* **168**, 291–308.
Semenov, M. A. & Terkel, D. A. (1985). On the evolution of the mechanisms of genetic variability by indirect action of selection. *Zhurnal Obshchej Biologii* **46**, 271–277 (in Russian).
Shapiro, N. I. (1938). The mutation process as an adaptive character of a species. *Zoologicheskij Zhurnal* **17**, 592–601 (in Russian).
Shapiro, N. I. & Ignatiev, M. V. (1945). Evolution of mutability. *Uspekhi Sovremennoj Biologii* **20**, 325–344 (in Russian).
Slatkin, M. (1972). On treating the chromosome as the unit of selection. *Genetics* **72**, 157–168.
Sturtevant, A. H. (1937). Essays on evolution. I. On the effects of selection on mutation rate. *Quarterly Review of Biology* **12**, 464–467.
Twomey, M. J. & Feldman, M. W. (1990). Mutation modification with multiplicative fertility selection. *Theoretical Population Biology* **37**, 320–342.
Uyenoyama, M. K. & Waller, D. M. (1991). Coevolution of self-fertilization and inbreeding depression. III. Homozygous lethal mutations at multiple loci. *Theoretical Population Biology* **40**, 173–210.
Wiener, P. & Feldman, M. W. (1993). The effects of the mating systems on the evolution of migration in a spatially heterogeneous population. *Evolutionary Ecology* **7**, 251–269.
Woodruff, R. C., Thompson, J. N., Seeger, M. A. & Spivey, W. E. (1984). Variation in spontaneous mutation and repair in natural population lines of *Drosophila melanogaster*. *Heredity* **53**, 223–234.